

Methods in
Molecular Biology 1961

Springer Protocols

An illustration on a dark blue background depicting CRISPR gene editing. It features several DNA double helix structures in shades of red and orange. A pair of hands, one holding a pair of scissors, is shown cutting through a DNA strand. Other elements include a white mouse, a pink pig, and a blue virus-like particle.

Yonglun Luo *Editor*

CRISPR Gene Editing

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>

CRISPR Gene Editing

Methods and Protocols

Edited by

Yonglun Luo

Department of Biomedicine, Aarhus University, Aarhus, Denmark

BGI-Shenzhen, Shenzhen, China

 **Humana Press**

Editor

Yonglun Luo
Department of Biomedicine
Aarhus University
Aarhus, Denmark

BGI-Shenzhen
Shenzhen, China

ISSN 1064-3745 ISSN 1940-6029 (electronic)
Methods in Molecular Biology
ISBN 978-1-4939-9169-3 ISBN 978-1-4939-9170-9 (eBook)
<https://doi.org/10.1007/978-1-4939-9170-9>

Library of Congress Control Number: 2019933103

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Open Access Chapter 3 is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>). For further details see license information in the chapter.

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Humana Press imprint is published by the registered company Springer Science+Business Media, LLC, part of Springer Nature.

The registered company address is: 233 Spring Street, New York, NY 10013, U.S.A.

Preface

When the series editor, Dr. John M. Walker, asked me approximately 1 year ago whether I was interested in editing a book on CRISPR technologies, I immediately took his generous offer and started my first and challenging editorial journey. Firstly, as a junior researcher, it has always been a great honor and encouraging experience to be capable of contributing with knowledge and outreach to our scientific community. Secondly, the particular research field and technology that we will be focusing on in this book is one of the most fast growing and important breakthroughs during the past decade. And most importantly, as a researcher working on gene editing for over 10 years, I deeply realize the importance of having a good serial of protocols, experimental tips, and notes to increase the success rate and outcomes of scientific projects.

When I first started the “Pig and Health” PhD project back in 2008, my objective was to recapitulate the pathogenesis of human diseases, e.g., breast cancer and diabetes, using genetically tailored pig models. One major milestone of my PhD study was disrupting the BRCA1 gene in primary porcine fibroblasts by homologous recombination and subsequently generating a BRCA1 knockout pig by somatic cell nuclear transfer. Generating a gene knockout (KO) animal during the pre-TALENs and pre-CRISPR era was rather technically challenging and time-consuming. Although I spent almost 3 years and finally got my BRCA1 KO pig, I wish my project was conducted now rather than 10 years ago. Zinc finger nucleases (ZFNs), known as the second generation of programmable DNA nucleases, were already available during that period. But this technology was not very broadly adopted by the scientific community. One main reason is that the ZFN technology is relatively difficult for the design and generation, and there is a lack of user-friendly protocols and methods instructing the generation and functional validation of ZFNs.

When the transcription activator-like effector (TALE) protein was engineered as a programmable DNA endonuclease (TALEN) for gene editing, as compared to ZFNs, TALEN-based gene editing was more rapidly applied by the whole scientific community. One important driving force of the TALEN technology is conventional web tools and protocols developed for TALEN vector design and construction. Using one the most popular TALEN assembly methods developed by Daniel F. Voytas’s group, although one has to select the different modular plasmids from a large stock of premade ones, it is much easier when compared to ZFNs to generate a couple of TALEN constructions.

Since the first proof-of-principle study of harnessing the clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein 9 (Cas9) for gene editing in 2012, led by Jennifer Doudna and Emmanuelle Charpentier, the CRISPR-Cas9-based gene editing technology rapidly took over ZFNs and TALENs and was successfully applied for genome editing in almost all cell types and organisms. To date, many Cas9 orthologs, different CRISPR systems from bacteria and archaea have been repurposed for gene editing. And many more genetic manipulation tools have been built on top of the CRISPR-Cas9 system, such as CRISPR-based gene activation, gene interference, base editing, DNA methylation, and histone acetylation. Among all the already harnessed CRISPR-Cas systems utilized for gene editing, the CRISPR-Cas9 system is still the most extensively developed and broadly used one.

The CRISPR-Cas9 technology is commonly known and described as simple, efficient, and cost-effective. However, a successful CRISPR gene editing experiment/project requires strategic planning and user-friendly guidelines to select the most suitable CRISPR-Cas system and target sites with high activity and specificity. Also, how to quantify the CRISPR gene editing activity (indel), how to efficiently deliver the CRISPR-Cas9 components into target cells or tissues, and how to enrich and isolate gene-edited cells with desired genetic modifications: these are among the most frequently asked questions and experiments when conducting a CRISPR gene editing study. This book is intended to assist undergraduates, graduates, and researchers with detailed guidelines and methods for the CRISPR gene editing field.

I would like to thank all the contributing authors at the front lines of developing CRISPR technology and applications. Methods covering CRISPR gRNA design, CRISPR delivery, CRISPR activity quantification (indel quantification), and examples of applying CRISPR gene editing in human pluripotent stem cells, primary cells, gene therapy, and genetic screening are included in this book. Without their contributions, this book could never have been written.

My grateful thanks also to my colleagues from the Lars Bolund Institute of Regenerative Medicine and the DREAM team for their assistance in editing the book.

Aarhus, Denmark

Yonglun Luo

Contents

<i>Preface</i>	<i>v</i>
<i>Contributors</i>	<i>ix</i>
PART I METHODS FOR CRISPR-gRNA DESIGN AND QUANTIFICATION OF ACTIVITY	
1 CRISPR-gRNA Design	3
<i>Maria Pallarès Masmijtjà, Nastassia Knödlseder, and Marc Güell</i>	
2 Tracking CRISPR's Footprints	13
<i>Lin Lin and Yonglun Luo</i>	
3 Rapid Quantitative Evaluation of CRISPR Genome Editing by TIDE and TIDER	29
<i>Eva Karina Brinkman and Bas van Steensel</i>	
4 Fast and Quantitative Identification of Ex Vivo Precise Genome Targeting-Induced Indel Events by IDAA	45
<i>Saskia König, Zhang Yang, Hans Heugh Wandall, Claudio Mussolino, and Eric Paul Bennett</i>	
5 Functional Evaluation of CRISPR Activity by the Dual-Fluorescent Surrogate System: C-Check	67
<i>Lin Lin and Yonglun Luo</i>	
PART II METHODS FOR CRISPR DELIVERY	
6 CRISPR-Cas9 Delivery by Artificial Virus (RRPHC)	81
<i>Suleixin Yang, Qinjie Wu, Yuquan Wei, and Changyang Gong</i>	
7 Production and Validation of Lentiviral Vectors for CRISPR/Cas9 Delivery ...	93
<i>Laura Barrett Ryo, Emil Aagaard Thomsen, and Jacob Giehm Mikkelsen</i>	
8 Rapid and Simple Screening of CRISPR Guide RNAs (gRNAs) in Cultured Cells Using Adeno-Associated Viral (AAV) Vectors.....	111
<i>Julia Fakhiri, Manuela Nickl, and Dirk Grimm</i>	
9 Electroporation-Based CRISPR/Cas9 Gene Editing Using Cas9 Protein and Chemically Modified sgRNAs	127
<i>Anders Laustsen and Rasmus O. Bak</i>	
PART III CRISPR GENE EDITING IN HUMAN iPSCs	
10 Efficient Gene Editing of Human Induced Pluripotent Stem Cells Using CRISPR/Cas9	137
<i>Saniye Yumlu, Sanum Bashir, Jürgen Stumm, and Ralf Kühn</i>	

11	Editing the Genome of Human Induced Pluripotent Stem Cells Using CRISPR/Cas9 Ribonucleoprotein Complexes.....	153
	<i>Michaela Bruntraeger, Meg Byrne, Kathleen Long, and Andrew R. Bassett</i>	
12	Conditional Gene Knockout in Human Cells with Inducible CRISPR/Cas9	185
	<i>Kirsten E. Snijders, James D. Cooper, Ludovic Vallier, and Alessandro Bertero</i>	
PART IV CRISPR GENE EDITING IN OTHER CELL TYPES		
13	CRISPR/Cas9 as a Genome Editing Tool for Targeted Gene Integration in CHO Cells	213
	<i>Daria Sergeeva, Jose Manuel Camacho-Zaragoza, Jae Seong Lee, and Helene Faustrup Kildegaard</i>	
14	Rapid and Efficient Gene Deletion by CRISPR/Cas9	233
	<i>Signe Neldeborg, Lin Lin, Magnus Stougaard, and Yonglun Luo</i>	
15	Genome Editing in Mice	249
	<i>Lisbeth Ahm Hansen and Ernst-Martin Fuchtbauer</i>	
16	CRISPR/Cas9-Mediated Gene Tagging: A Step-by-Step Protocol.....	255
	<i>Xi Xiang, Conghui Li, Xi Chen, Hongwei Dou, Yong Li, Xiuqing Zhang, and Yonglun Luo</i>	
17	Gene Editing in Primary Cells of Cattle and Pig	271
	<i>Petra Vochozkova, Kilian Simmet, Eva-Maria Jemiller, Annegret Wünsch, and Nikolai Klymiuk</i>	
PART V CRISPR GENE THERAPY AND SCREENING		
18	Toward In Vivo Gene Therapy Using CRISPR.....	293
	<i>Kristian Alsbjerg Skipper and Jacob Giehm Mikkelsen</i>	
19	CRISPR Gene Therapy of the Eye: Targeted Knockout of <i>Vegfa</i> in Mouse Retina by Lentiviral Delivery	307
	<i>Andreas Holmgaard, Sidsel Alsing, Anne Louise Askou, and Thomas J. Corydon</i>	
20	In Vivo Editing of the Adult Mouse Liver Using CRISPR/Cas9 and Hydrodynamic Tail Vein Injection	329
	<i>Francesco Niola, Frederik Dagnæs-Hansen, and Morten Frödin</i>	
21	CRISPR-Based Lentiviral Knockout Libraries for Functional Genomic Screening and Identification of Phenotype-Related Genes	343
	<i>Emil Aagaard Thomsen and Jacob Giehm Mikkelsen</i>	
	<i>Index</i>	359

Contributors

- SIDSEL ALSING • *Department of Biomedicine, Aarhus University, Aarhus, Denmark*
- ANNE LOUISE ASKOU • *Department of Biomedicine, Aarhus University, Aarhus, Denmark*
- RASMUS O. BAK • *Department of Biomedicine, Aarhus University, Aarhus, Denmark;*
Aarhus Institute of Advanced Studies (AIAS), Aarhus University, Aarhus, Denmark
- SANUM BASHIR • *Max-Delbrück-Centrum für Molekulare Medizin, Berlin, Germany; Berlin*
Institute of Health, Berlin, Germany
- ANDREW R. BASSETT • *Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton,*
Cambridge, UK
- ERIC PAUL BENNETT • *Faculty of Health Sciences, Department of Odontology, Copenhagen*
Center for Glycomics (CCG), University of Copenhagen, Copenhagen, Denmark
- ALESSANDRO BERTERO • *Wellcome Trust-MRC Stem Cell Institute, Anne McLaren*
Laboratory, University of Cambridge, Cambridge, UK; Department of Surgery, University
of Cambridge, Cambridge, UK; Department of Pathology, University of Washington,
Seattle, WA, USA
- EVA KARINA BRINKMAN • *Division of Gene Regulation and Oncode Institute, Netherlands*
Cancer Institute, Amsterdam, The Netherlands
- MICHAELA BRUNTRAEGER • *Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton,*
Cambridge, UK
- MEG BYRNE • *Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge,*
UK
- JOSE MANUEL CAMACHO-ZARAGOZA • *The Novo Nordisk Foundation Center for*
Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark
- XI CHEN • *BGI-Shenzhen, Shenzhen, China*
- JAMES D. COOPER • *Wellcome Trust-MRC Stem Cell Institute, Anne McLaren Laboratory,*
University of Cambridge, Cambridge, UK; Division of Cardiovascular Medicine,
University of Cambridge, Cambridge, UK
- THOMAS J. CORYDON • *Department of Biomedicine, Aarhus University, Aarhus, Denmark;*
Department of Ophthalmology, Aarhus University Hospital, Aarhus, Denmark
- FREDERIK DAGNÆS-HANSEN • *Department of Biomedicine, Aarhus University, Aarhus,*
Denmark
- HONGWEI DOU • *BGI-Qingdao, Qingdao, China*
- JULIA FAKHIRI • *Department of Infectious Diseases/Virology, Heidelberg University Hospital,*
Cluster of Excellence Cell Networks, Heidelberg, Germany; BioQuant Center, University of
Heidelberg, Heidelberg, Germany
- MORTEN FRÖDIN • *Biotech Research and Innovation Centre (BRIC), University of*
Copenhagen, Copenhagen, Denmark
- ERNST-MARTIN FÜCHTBAUER • *Department of Molecular Biology and Genetics, Aarhus,*
Denmark
- CHANGYANG GONG • *State Key Laboratory of Biotherapy and Cancer Center, West China*
Hospital, Sichuan University, Chengdu, China

- DIRK GRIMM • *Department of Infectious Diseases/Virology, Heidelberg University Hospital, Cluster of Excellence Cell Networks, Heidelberg, Germany; BioQuant Center, University of Heidelberg, Heidelberg, Germany; German Center for Infection Research (DZIF), Partner Site Heidelberg, Heidelberg, Germany*
- MARC GÜELL • *Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Barcelona, Spain*
- LISBETH AHM HANSEN • *Department of Biomedicine, Aarhus University, Aarhus, Denmark*
- ANDREAS HOLMGAARD • *Department of Biomedicine, Aarhus University, Aarhus, Denmark*
- EVA-MARIA JEMILLER • *Institute for Molecular Animal Breeding and Biotechnology, LMU Munich, Munich, Germany*
- HELENE FAUSTRUP KILDEGAARD • *The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark*
- NIKOLAI KLYMIUK • *Institute for Molecular Animal Breeding and Biotechnology, LMU Munich, Munich, Germany*
- NASTASSIA KNÖDLSIEDER • *Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Barcelona, Spain*
- SASKIA KÖNIG • *Medical Center—University of Freiburg, Institute for Transfusion Medicine and Gene Therapy and Center for Chronic Immunodeficiency at Center for Translational Cell Research (ZTZ), Freiburg, Germany*
- RALF KÜHN • *Max-Delbrück-Centrum für Molekulare Medizin, Berlin, Germany; Berlin Institute of Health, Berlin, Germany*
- ANDERS LAUSTSEN • *Department of Biomedicine, Aarhus University, Aarhus, Denmark*
- JAE SEONG LEE • *Department of Molecular Science and Technology, Ajou University, Suwon, Republic of Korea*
- CONGHUI LI • *BGI-Qingdao, Qingdao, China*
- YONG LI • *BGI-Shenzhen, Shenzhen, China*
- LIN LIN • *Department of Biomedicine, Aarhus University, Aarhus, Denmark*
- KATHLEEN LONG • *Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK*
- YONGLUN LUO • *Department of Biomedicine, Aarhus University, Aarhus, Denmark; BGI-Shenzhen, Shenzhen, China; Guangdong Provincial Key Laboratory of Genome Read and Write, Shenzhen, China; BGI-Qingdao, Qingdao, China*
- JACOB GIEHM MIKKELSEN • *Department of Biomedicine, Aarhus University, Aarhus, Denmark*
- CLAUDIO MUSSOLINO • *Medical Center—University of Freiburg, Institute for Transfusion Medicine and Gene Therapy and Center for Chronic Immunodeficiency at Center for Translational Cell Research (ZTZ), Freiburg, Germany*
- SIGNE NELDEBERG • *Department of Clinical Medicine, Aarhus University, Aarhus, Denmark*
- MANUELA NICKL • *Department of Infectious Diseases/Virology, Heidelberg University Hospital, Cluster of Excellence Cell Networks, Heidelberg, Germany; BioQuant Center, University of Heidelberg, Heidelberg, Germany; German Center for Infection Research (DZIF), Partner Site Heidelberg, Heidelberg, Germany*
- FRANCESCO NIOLA • *Biotech Research and Innovation Centre (BRIC), University of Copenhagen, Copenhagen, Denmark*
- MARIA PALLARÈS MASMITJÀ • *Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Barcelona, Spain*
- LAURA BARRETT RYØ • *Department of Biomedicine, Aarhus University, Aarhus, Denmark*

- DARIA SERGEEVA • *The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark*
- KILIAN SIMMET • *Institute for Molecular Animal Breeding and Biotechnology, LMU Munich, Munich, Germany*
- KRISTIAN ALSBJERG SKIPPER • *Department of Biomedicine, Aarhus University, Aarhus, Denmark*
- KIRSTEN E. SNIJDERS • *Wellcome Trust-MRC Stem Cell Institute, Anne McLaren Laboratory, University of Cambridge, Cambridge, UK; Department of Surgery, University of Cambridge, Cambridge, UK*
- MAGNUS STOUGAARD • *Department of Clinical Medicine, Aarhus University, Aarhus, Denmark*
- JÜRGEN STUMM • *Max-Delbrück-Centrum für Molekulare Medizin, Berlin, Germany; Berlin Institute of Health, Berlin, Germany*
- EMIL AAGAARD THOMSEN • *Department of Biomedicine, Aarhus University, Aarhus, Denmark*
- LUDOVIC VALLIER • *Wellcome Trust-MRC Stem Cell Institute, Anne McLaren Laboratory, University of Cambridge, Cambridge, UK; Department of Surgery, University of Cambridge, Cambridge, UK; Wellcome Trust Sanger Institute, Hinxton, UK*
- BAS VAN STEENSEL • *Division of Gene Regulation and Oncode Institute, Netherlands Cancer Institute, Amsterdam, The Netherlands*
- PETRA VOCHOZKOVA • *Institute for Molecular Animal Breeding and Biotechnology, LMU Munich, Munich, Germany*
- HANS HEUGH WANDALL • *Faculty of Health Sciences, Department of Cellular and Molecular Medicine, Copenhagen Center for Glycomics (CCG), University of Copenhagen, Copenhagen, Denmark*
- YUQUAN WEI • *State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, Sichuan University, Chengdu, China*
- QINJIE WU • *State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, Sichuan University, Chengdu, China*
- ANNEGRET WÜNSCH • *Institute for Molecular Animal Breeding and Biotechnology, LMU Munich, Munich, Germany*
- XI XIANG • *BGI Education Center, University of Chinese Academy of Sciences, Shenzhen, China; BGI-Shenzhen, Shenzhen, China*
- SULEIXIN YANG • *State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, Sichuan University, Chengdu, China*
- ZHANG YANG • *Faculty of Health Sciences, Department of Cellular and Molecular Medicine, Copenhagen Center for Glycomics (CCG), University of Copenhagen, Copenhagen, Denmark*
- SANIYE YUMLU • *Max-Delbrück-Centrum für Molekulare Medizin, Berlin, Germany; Berlin Institute of Health, Berlin, Germany*
- XIUQING ZHANG • *BGI-Shenzhen, Shenzhen, China*

Part I

Methods for CRISPR-gRNA Design and Quantification of Activity



Chapter 1

CRISPR-gRNA Design

Maria Pallarès Masmitjà, Nastassia Knödlseeder, and Marc Güell

Abstract

Gene editing has great therapeutic impact, being of interest for many scientists worldwide. Clustered regularly interspaced short palindromic repeats (CRISPR) technology has been adapted for gene editing to serve as an efficient, rapid, and cost-effective tool. To fulfill CRISPR experiment's goals, two components are important: an endonuclease and a gRNA. The most commonly used endonucleases are Cpf1 and Cas9 and are described in depth in this chapter. The gRNA targets the genome site to be edited, giving great importance to its design to obtain increased efficiency and decreased off-target events. In this chapter, we describe different tools to design suitable gRNAs for a variety of experimental purposes.

Key words CRISPR, gRNA design, Genome editing, Cas9, Cpf1

1 Introduction

Technologies for genome engineering are on high demand. The adaptation of CRISPR, a new tool for gene editing, has revolutionized the field of genome engineering [1, 2]. Nowadays, CRISPR is the most used technology in gene editing, with Cas9 being the most used CRISPR system [3].

Basically, this new technology consists of two components: A single- or double-strand endonuclease, such as Cas9 or Cpf1, and a synthetic single-guide RNA (gRNA) which will direct the CRISPR system to a target region in the genome. The Cas9-gRNA complex scans the protospacer adjacent motif (PAM) region and forms Watson-Crick base pairing to a target DNA fragment of 20 nucleotides [3]. This conformation enables the endonuclease to cleave site specifically the target DNA. Finally, the endogenous DNA repair machinery of the cells repairs the double-stranded break and introduces the desired changes.

Minimizing to a two-component system enabled genome engineering to be easier than ever before. The gRNA can be changed to any target DNA sequence of interest.

This chapter provides a methodology for gRNA design and includes all necessary information and options when designing an adequate gRNA for your CRISPR experiment.

2 Nuclease Types

Choosing the right nuclease is key for our experiment success. There are many different nucleases to use; the CRISPR-Cas system is divided into 2 classes, which are divided into 6 types and 19 subtypes [4]. However, the most used endonucleases are Cas9 and Cpf1, both from class 2, which both work as a single unit for recognition of the target and cleavage (Table 1).

Cas9 is a dual RNA-guided DNA endonuclease of type II. It cleaves double-stranded DNA.

It has preference for PAM (protospacer-adjacent motif) sequence 5'NGG'3 following the 20 bp gRNA target [5], and it generates blunt ends three nucleotides upstream of the PAM site.

Multiple Cas9 variants have been developed including a nickase Cas9 and a dead Cas9. Nickase Cas9 (nCas9) has either of the two nuclease domains (HNH and RuvC) [6] inactivated whereas if both nuclease domains are inactivated it becomes Dead cas9 (dCas9). nCas9 generates single-strand breaks (SSBs), a pair of nCas9 can be used to generate paired nicks instead of DSBs, and this would reduce off-target cleavage [7]. dCas9 will work as a site-specific DNA-binding vehicle able to be combined with other effectors instead of a genome editing tool [8].

On the other hand, Cpf1 is a CRISPR type V endonuclease which recognizes PAMs 5'-TTN which is more frequent in the human genome than Cas9 PAMs requirements, and also improves the editing of TA-rich parts of the genome [9]. This endonuclease cleaves in a staggered manner creating five nucleotide 5' overhang 18–23 bases away from PAM. It is also smaller than Cas9 nuclease enabling a better packaging in viruses if needed; however, there are two relevant types of Cas9: Cas9 from *Streptococcus pyogenes*

Table 1
Features of Cas9 and Cpf1

Type	Cas9 II	Cpf1 V
Cleavage	Double-stranded DNA	Double-stranded DNA
Protospacer-adjacent motif (PAM)	NGG'3	5'-TTN
Ends	Blunt end 3 nt upstream PAM	5 nucleotide 5' overhang 18–23 nt from PAM

(SpCas9) and Cas9 from *Staphylococcus aureus* (SaCas9) which is smaller and better for adeno-associated viruses packaging and transduction [10].

3 gRNA Efficiency

gRNAs will guide the CRISPR system molecules to the specific site intended to edit in the genome. It has a scaffold that will bind the endonuclease and a spacer of 20 nucleotides which target specific sequence in the genome.

The efficiency of the gRNA can vary from a sequence to another, and it is crucial for the genome editing experiment. Some factors are known to correlate with high efficiency gRNAs such as higher stability of the duplex between gRNA and target DNA [11]. Scores tend to correlate with gRNA efficiency, but their accuracy depends on the experimental system considered [12]. For instance, Moreno-Mateos score is useful for zebrafish injection experiments [13] (where gRNAs need to be very stable after injections including G-rich sequence) but inaccurate in cultured mammalian cells [14]. This score is also reliable for assays based on delivery of gRNAs produced synthetically or by T7 in vitro transcription [15]. The most common scores are based on datasets where gRNA is continuously expressed from a U6 promoter so they do not need to be very stable. One of the most used for mammalian cells is the score by Doench et al. [16]; also the Wang score is useful when designing second generation libraries of human gRNAs in some cancer cell lines [17]. It is recommended to use scoring systems trained with the same expression system planned to be used. The web tool <http://crispor.org> provides multiple scores such as Moreno-Mateos and Doench simultaneously to help select efficient gRNA for different applications.

Recently, approaches based on machine learning such as CRISTA have been developed, which determines the propensity of a genomic site to be cleaved by a gRNA [18].

4 gRNA Off Target

CRISPR RNA-guided endonucleases have been tested for its off-target activity in human cells. Cpf1 is highly specific in human and does not show much measurable off target [19]. There is an essential need of identifying Cas9-induced off-target mutations, to prevent unwanted chromosomal translocations or unintended mutations, in human research to ensure safe use of the CRISPR nucleases in genome engineering and treatment of diseases [20, 21]. Off targets are generally induced by the targeted site and never represent a random change [20].

Off-target cleavage represents a major challenge for CRISPR/Cas9-mediated genome editing, as Cas9 is able to cleave even if there is no complete complementarity between genomic site and gRNA sequence [22]. It was believed that Cas9 specificity was strictly controlled by PAM and the 20 nt guide sequence, but recently it was shown that it was controlled by a 5 bp sequence 5' of the PAM, termed "Seed Sequence" [23]. Through the recent findings of the importance of the sequence in proximity to the PAM, some rules for the design were proposed:

1. U-rich seeds: WHY? Decrease of sgRNA abundance and increase of specificity.
2. Mis-matches: WHERE? Tolerated better at the 5' end of the sgRNA.
3. GC content: HOW? Exceptionally high or low GC content in sgRNA leads to an increase in activity. Guanine is preferred as first base after PAM and cytosine as the base for the fifth position after PAM [23].

Researchers try to find different solutions for minimizing and eliminating off-target mutations in CRISPR experiments. One way is to minimize the concentration of used gRNA or/and Cas9 in human cells. Another way is to change the gRNA sequence by either truncating the 3' ends within the tracrRNA-derived sequence or by addition of two guanines on the 5' ends prior to the complementary region. Paired nickases represent an additional approach, where two gRNAs and a cas9 nickase are generating nicks at the target site [7].

It was shown that 5' truncated gRNAs at their complementary region, which have at least 17-18 nt of complementary function, are as efficient as their full-length construct, but with a lower off-target cleavage [24]. While all the mentioned approaches decrease off-target cleaves, also on-target efficiency gets reduced and through truncated guides, indel formation can occur and lead to less overall target sites in the genome [22]. Slaymaker et al. engineered the cas9 nuclease in a structure-guided manner where they neutralize positive charges in the nt-groove. Hereby, off-target indel formation could be decreased and on-target cleavage stabilized [22].

The off-target activity of the CRISPR system can be carefully minimized by designing a precise gRNA. A study to minimize the off-target binding of the gRNAs was performed using the earlier study on the on-target activity as a model (Rule Set 1) and setting up computational design rules to create better gRNA libraries [16]. It is important to mention the CFD score (cutting frequency determination), which predicts off-target scores combined with Doench et al. designed gRNA designer 2016 which can predict activity and specificity of guides [16].

In a similar way, in order to facilitate Cpf1-based genome editing, a study of >11,000 different target sequences and guide RNAs was performed [9] and a tool to predict indel sequences for given target sequences was created (<http://big.hanyang.ac.kr/cindel>).

5 gRNA Design

To design a satisfactory CRISPR gene editing experiment, here are some important steps:

1. **Selection of a suitable endonuclease.** We described cas9 and cpf1 before, the most used endonucleases in CRISPR experiments.
2. Target selection of the gRNA upstream 5'-NGG for cas9 or downstream of TTN for cpf1.
3. Selecting an efficient gRNA with reduced off-target. Different candidates can be found using different tools described in Table 2.
4. Design ssODN template (optional for homology-directed repair (HDR)).

As mentioned before, the specificity of the nuclease depends on the 20 nt sequence on the gRNA that will base pair with our target sequence. The two critical points to design a suitable gRNA are the efficiency to target our gene of interest avoiding large off-target events. In the previous section about gRNAs, we mentioned some useful scores to design it such as Wang score and CFD score or tools to calculate off target like sgRNA scorer for cas9 or hanyang for cpf1 gRNAs. Select an efficiency score that has been trained on the same gRNA production system planned in your experiment (U6 expression, synthetic/T7 in vitro transcribed).

6 gRNA Expression, Construction, and Delivery

There are a wide variety of approaches to deliver the CRISPR machinery (gRNAs and endonucleases) to the experimental system of election; here we will mention some of them:

1. By PCR amplification: enclose the gRNA onto the reverse primer used to amplify a U6 promoter template. The amplicon can be co-transfected into the cells with cas9 expression plasmid. This is a good method for quick screening of different gRNAs candidates and for generating libraries.

Table 2
Selected websites for gRNA design

Websites for design		
Link	Name	Description
http://tools.genome-engineering.org	CRISPRDesign	To take a large input sequence, identify and rank good target sites and predict their off-target equivalent sites and also specify the necessary oligos and primers; you also can select 20 bp nucleotide sequences and score them and design the primers manually.
http://crispor.org	CRISPOR	To design, evaluate and clone guide sequences for CRISPR experiments. It provides multiple scores to design a more accurate gRNA depending on the system used. It also works with cpfl gRNAs.
https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design	sgRNA designer 2016	To rank gRNA candidates maximizing the on-target candidates and minimizing the off-target using Rule Set 2 from Doench et al.
https://genome.ucsc.edu	UCSC genome browser	To look for CRISPR gRNA. It has already all the reliable gRNA sequences designed and simply choose by looking at the target gene region
http://big.hanyang.ac.kr/cindel	CINDEL	To predict indel sequences for given target sequences for Cpfl gRNAs. The sequence must be 27 bp maximum.

2. By plasmid-based procedure: oligo pairs encoding the 20 nt guide sequence which will be annealed and ligated into a plasmid (which can be the cas9 expression plasmid).
3. Topocloning the gRNAs onto an episomal vector and co-transfecting it with the endonuclease plasmid.
4. Endonucleases and gRNAs can also be transfected into the cells as mRNA and RNA, or RNA-protein complexes.

7 Design of Repair Template (Only for HDR Applications)

Previously, there was the need to use plasmid-based donor repair templates containing homology arms flanking the site of alteration. Nowadays, ssODNs [25] can be used for short modifications in a defined locus without cloning. For high HDR efficiency, ssODNs contain flanking sequences around 40 bp on each side homologous to the target regions.

It has been shown that by adding ssDNA donors complementary to the strand that cas9 dissociates from first, the HDR rate in

human cells can be increased using either cas9 or nickase variants [26].

8 Functional Validation of gRNA

1. Cell culture and transfection with endonucleases and gRNAs.
2. Isolation of the clonal cell lines (optional). There are many techniques to isolate the successfully edited clonal cell lines: to isolate cells with specific modifications you can use FACS enrichment or serial dilutions followed by an expansion period to establish a new clonal cell line.

9 On-Target Efficiency

There are different methodologies to calculate the success of a CRISPR experiment.

1. *Gene editing efficiency can be estimated using SURVEYOR nuclease assay* [27].
2. Sanger sequencing can be used to assess the mutation repertoire generated.
3. Fragment analysis based to estimate population of indels.
4. *Deep sequencing* can simultaneously measure efficiency and mutations repertoire [28, 29].

10 Off-Target Analysis

Many different high-throughput sequencing techniques have evolved lately to analyze on-target and off-target efficiency. Two highly used systems are GUIDE-seq and CIRCLE-seq [30, 31].

11 Conclusion

CRISPR is a relatively new tool for gene editing. Appropriate experimental design may impact significantly the quality of the results.

References

1. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155:733–740
2. Zhang F, Wen Y, Guo X (2014) CRISPR/Cas9 for genome editing: progress, implications and challenges. *Hum Mol Genet* 23:R40–R46
3. 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* 337:816–821
4. Nishimasu H, Nureki O (2017) Structures and mechanisms of CRISPR RNA-guided effector nucleases. *Curr Opin Struct Biol* 43:68–78
 5. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE et al (2013) RNA-guided human genome engineering via Cas9. *Science* 339:823–826
 6. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV (2006) A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct* 1:7
 7. Ran FA, Hsu PD, Lin C-Y, Gootenberg JS, Konermann S, Trevino AE et al (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154:1380–1389
 8. Wen W-S, Yuan Z-M, Ma S-J, Xu J, Yuan D-T (2016) CRISPR-Cas9 systems: versatile cancer modelling platforms and promising therapeutic strategies. *Int J Cancer* 138:1328–1336
 9. Kim HK, Song M, Lee J, Vipin Menon A, Jung S, Kang Y-M et al (2016) In vivo high-throughput profiling of CRISPR–Cpf1 activity. *Nat Methods* 14:153–159
 10. Friedland AE, Baral R, Singhal P, Loveluck K, Shen S, Sanchez M et al (2015) Characterization of *Staphylococcus aureus* Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications. *Genome Biol* 16:257. <https://doi.org/10.1186/s13059-015-0817-8>
 11. Wong N, Liu W, Wang X (2015) WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system. *Genome Biol* 16:218
 12. Haeussler M, Concordet J-P (2016) Genome editing with CRISPR-Cas9: can it get any better? *J Genet Genomics* 43:239–250
 13. Moreno-Mateos MA, Vejnar CE, Beaudoin J-D, Fernandez JP, Mis EK, Khokha MK et al (2015) CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods* 12:982–988
 14. Malina A, Cameron CJF, Robert F, Blanchette M, Dostie J, Pelletier J (2015) PAM multiplicity marks genomic target sites as inhibitory to CRISPR-Cas9 editing. *Nat Commun* 6:10124
 15. Haeussler M, Schönig K, Eckert H, Eschstruth A, Mianné J, Renaud J-B et al (2016) Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol* 17:148
 16. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF et al (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* 34:184–191
 17. Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ et al (2015) Identification and characterization of essential genes in the human genome. *Science* 350:1096–1101
 18. Abadi S, Yan WX, Amar D, Mayrose I (2017) A machine learning approach for predicting CRISPR-Cas9 cleavage efficiencies and patterns underlying its mechanism of action. *PLoS Comput Biol*. Public Library of Science 13:e1005807
 19. Kleinstiver BP, Tsai SQ, Prew MS, Nguyen NT, Welch MM, Lopez JM et al (2016) Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat Biotechnol*. NIH Public Access 34:869
 20. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Keith Joung J et al (2013) High frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol*. NIH Public Access 31:822
 21. Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S et al (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*. Cold Spring Harbor Laboratory Press 24:132
 22. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science* 351:84–88
 23. Zhang XH, Tee LY, Wang XG, Huang QS, Yang SH (2015) Off-target effects in CRISPR/Cas9-mediated genome engineering. *Mol Ther Nucleic Acids*. Cell Press 4:e264. <https://doi.org/10.1038/mtna.2015.37>
 24. Jeffry D, Sander JKJ (2014) CRISPR-Cas systems for genome editing, regulation and targeting. *Nat Biotechnol*. NIH Public Access 32:347
 25. Yang L, Guell M, Byrne S, Yang JL, De Los Angeles A, Mali P et al (2013) Optimization of scarless human stem cell genome editing. *Nucleic Acids Res* 41:9049–9061
 26. Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE (2016) Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol* 34:339–344

27. Shrock E, Güell M (2017) CRISPR in animals and animal models. *Prog Mol Biol Transl Sci.* Academic Press 152:95–114
28. Guell M, Yang L, Church GM (2014) Genome editing assessment using CRISPR Genome Analyzer (CRISPR-GA). *Bioinformatics* 30:2968–2970
29. Pinello L, Canver MC, Hoban MD, Orkin SH, Kohn DB, Bauer DE et al (2016) Analyzing CRISPR genome-editing experiments with CRISPResso. *Nat Biotechnol* 34:695–697
30. Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V et al (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* 33:187–197
31. Tsai SQ, Nguyen NT, Malagon-Lopez J, Topkar VV, Aryee MJ, Joung JK (2017) CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat Methods* 14:607–614



Tracking CRISPR's Footprints

Lin Lin and Yonglun Luo

Abstract

The programmable clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated 9 (Cas9) and CRISPR-Cas9-derived gene editing and manipulation tools have revolutionized biomedical research over the past few years. One important category of assisting technologies in CRISPR gene editing is methods used for detecting and quantifying indels (deletions or insertions). These indels are caused by the repair of CRISPR-Cas9-introduced DNA double-stranded breaks (DSBs), known as CRISPR's DNA cleavage footprints. In addition, CRISPR-Cas9 can also leave footprints to the DNA without introducing DSBs, known as CRISPR's DNA-binding footprints. The indel tracking methods have contributed greatly to the improvement of CRISPR-Cas9 activity and specificity. Here, we review and discuss strategies developed over that past few years to track the CRISPR's footprints, their advantages, and limitations.

Key words CRISPR, Cas9, Indels, DSB, Indel frequency, Off-target

1 The Important Role of Endogenous DSB Repair Machineries in CRISPR Gene Editing

Living organisms have developed an intrinsic mechanism to repair lesions introduced to their genetic materials. Some of these genetic lesions can be lethal, such as DNA double-stranded breaks (DSBs). Unrepaired DSBs will lead to chromosomal abnormality and cell death. In mammalian cells, several DSB repair mechanisms have evolved to cope with these lethal DNA damages. Two major DSB repair pathways have been broadly accepted to be the main DNA repair mechanisms in mammalian cells: non-homologous end joining (NHEJ) and homology-directed repair (HDR) [1]. More and more evidences have uncovered that microhomology-mediated end joining (MMEJ) and single strand annealing (SSA) are two DNA repair mechanisms that mammalian cells frequently used to repair DSBs (Fig. 1) [2, 3].

The DNA repair machinery is a very complicated but systematically organized system that cells have developed to maintain their

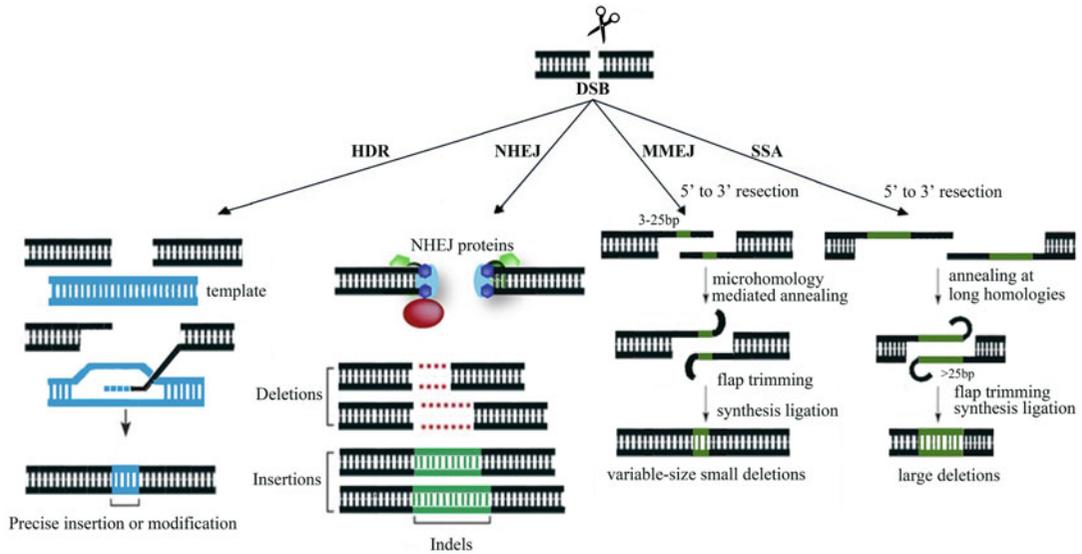


Fig. 1 Schematic illustration of DSB repair by homology-directed repair (HDR), non-homologous end joining (NHEJ), microhomology-mediated end joining (MMEJ), and single strand annealing in mammalian cells. HDR leads to precise and intended gene insertion, replacement, or modification. Indels (deletions or insertions) are typically introduced by NHEJ. Variable-size small and large deletions are introduced by MMEJ and SSA, respectively, after the repair of DSBs

genomic stability. Errors introduced to the DSBs repair pathways are the major cause of oncogenesis in humans. In 2015, the Nobel Prize in Chemistry was awarded to Tomas Lindahl, Paul Modrich, and Aziz Sancar, to acknowledge their work on deciphering the molecular mechanisms of DNA repair processes [4–6]. Although great progresses have been achieved in understanding the mechanism of DNA repair machinery in mammalian cells, and many genes such as *LIG4*, *KU70*, *KU80*, *DNA-PKcs*, *MRN*, *ARTEMIS*, *RAD51*, *TP53*, *BRCAL*, and *NSBI* are found to be involved in the DSB repair pathways, our knowledge of the cellular DNA repair machinery is still expanding [7].

DSBs are not always a bad thing. Traditional gene knockout by homologous recombination or HDR has greatly contributed to the generation of cell and animal models for biomedical research [8]. However, this approach is hampered by its extremely low efficiency. It has long been observed that introduction of DSBs can lead to disruption of gene function, or knockout, and can dramatically enhance the efficiency of HDR. Thus, there has been a long period of hunting for programmable DNA nucleases, which can be harnessed to introduce DSBs to a specific gene locus. The discovery of meganucleases (MGNs) and zinc finger nuclease (ZFNs) are two keystones for the foundation of gene editing, which is conceptionally defined as intentional modification of genes using a molecular scissor [9–11]. These two generations of

gene editing tools are not very broadly adopted mostly due to the technical difficulty and the workload of designing and generating a working MGNs or ZFNs. The Transcription Activator-Like Effector Nucleases (TALENs) technology is the second revolution in the gene editing history. Compared to MGNs and ZFNs, TALENs hold a much more stringent protein-to-DNA code for targeting. The DNA-binding domain of a TALEN is defined by the number and order of four types of modular repeated domains. These highly conserved 33–34 amino acids repeated domains contain two divergent 12th and 13th amino acids, also known as repeat variable diresidue (RVD): NG, HD, NI, NN/NH/NK that target T, C, A, G nucleotide respectively. Nevertheless, it is still quite time-consuming to generate a pair of TALEN vectors, although the whole procedure of generating TALENs has been greatly simplified compared to MGNs and ZFNs. The third generation of revolution in gene editing history is the RNA-guided endonucleases, also known as the clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein (Cas) gene editing technologies [12–14].

The CRISPR-Cas gene editing technology has overcome almost all the disadvantages of previous gene editing tools. Unlike MGNs, ZFNs, and TALENs, the CRISPR gene editing tool is based on RNA-to-DNA code for targeting. Thus, the sole Cas protein, with the *Streptococcus Pyogenes* Cas9 protein (SpCas9) being the most broadly used one, can be used for gene editing in almost all cells and organisms [15]. Designing and generating a new CRISPR targeting component is now simply achieved through changing the guide sequences presented in a small RNA molecule, known as guide RNA (gRNA) or small guide RNA (sgRNA). In the SpCas9-based CRISPR gene editing system, the gRNA is a chimeric RNA (crRNA:tracrRNA), which interacts with SpCas9 and guides SpCas9 to the target site through a Watson-Crick base pairing between the guide sequences of crRNA and the target site. As the guide or spacer sequences are the same as the target site (protospacer), the bacterial Cas protein has evolved an extremely smart system to distinguish self and non-self sequences, which is the of protospacer adjacent motif (PAM)-dependent DNA cleavage system [12]. Once all these features are met, the endonuclease activity of Cas9 protein is activated and a DSB is introduced in the target sites (Fig. 2).

One essentially common feature and outcome of all these gene editing tools is the introduction of a DSB to the target site. These DSBs are repaired by endogenous DNA repair machineries, leading to DNA alterations such as small deletions or insertions (called indels) at the DSB sites. These indels are just like footprints left by CRISPR to the DNA. Thus, quantification of the indel frequency has been used to quantify the overall activity and specificity (off-target) of these gene editing tools. Several methods have been

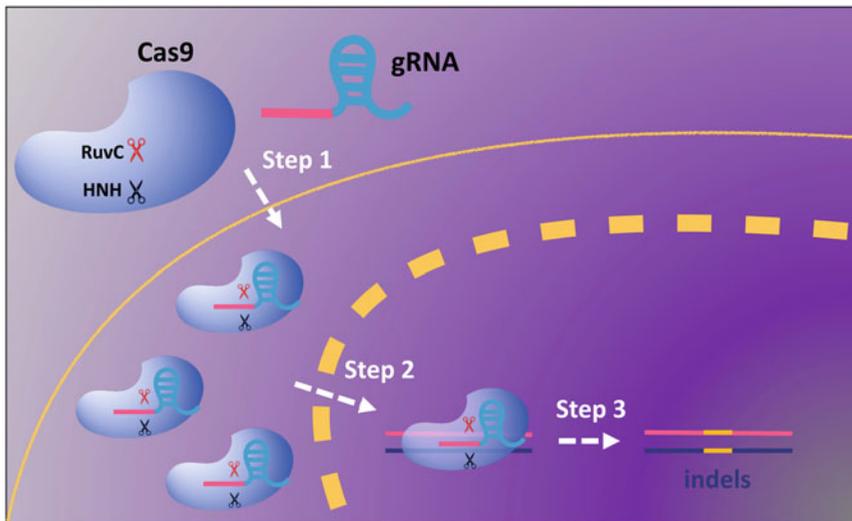


Fig. 2 Schematic illustration of CRISPR-Cas9-mediated gene targeting and the core components. The SpCas9 protein has two nuclease domains: RuvC and HNH. The first step of CRISPR gene editing is delivery of Cas9 and gRNA into cells, which can form Cas9:gRNA complex. The next step is the nucleus entry of Cas9:gRNA complex, mediated by the nucleus localization signal (NLS) peptides fused Cas9. The last step of CRISPR gene editing is the repair of DSBs introduced by Cas9:gRNA with DSBs repair machineries

developed over the past decade for quantifying indel frequency and specificity (Fig. 3). To give readers a systematic overview and help with choosing the right method for their CRISPR gene editing experiment, this chapter reviews the advantage and disadvantage of different methods.

2 CRISPR's Footprint Detection by Mismatch-Specific Endonuclease Assay (MSE)

The mismatch-specific endonucleases have been the gold standard of measuring overall activities of gene editing tools [16]. This assay is based on the mismatch-specific endonucleases such as Surveyor nuclease and T7 endonuclease 1 (T7E1) that are capable of cleaving heteroduplex DNAs [17]. The whole procedure of quantifying CRISPR activity by Surveyor assay or T7E1 is relative simple. First, genomic DNA or cell lysate is made from cells or tissues that are typically a few days after being treated with Cas9-gRNA components. The delivery of Cas9-gRNA can be in various forms, such as transfections in plasmid DNA, RNP, mRNA, or transduction through viral vectors [18–21]. Second, the targeted region is amplified by PCR using a specific primer and proofreading DNA polymerase. As the repairing of DSBs introduced by CRISPR/Cas9 is mainly mediated by NHEJ and MMEJ in mammalian cells, the targeted PCR product is a mixture of WT amplicons and amplicons with various indels. Third, the PCR product is purified and

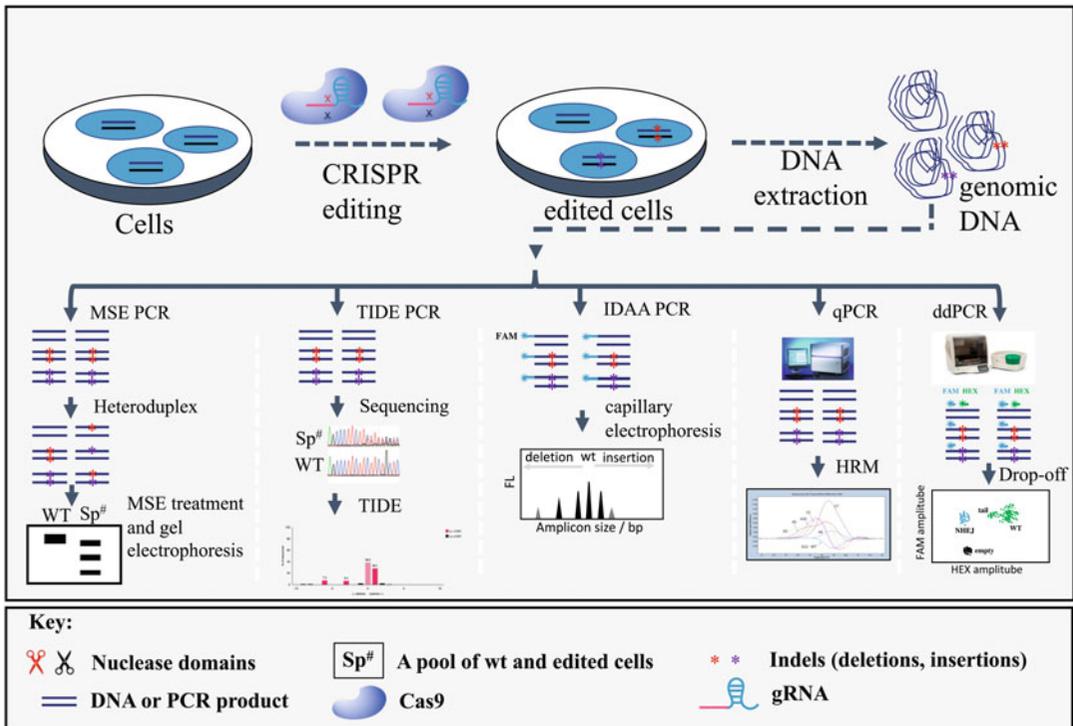


Fig. 3 Schematic illustration of the procedures and methods used for CRISPR indel profiling. Mismatch-specific enzyme assay, TIDE, IDAA, HRM, and Drop-Off methods are presented

subjected to the generation of heteroduplex DNA through denaturing and annealing. And finally, the heteroduplex DNA is digested with mismatch-specific endonuclease and indel frequency is quantified by measuring the electrophoresis band intensity using software such as Image J.

The assay is in principle very straightforward, as the mismatch-specific endonucleases can recognize any mismatches. However, to achieve a successful mismatch-specific endonuclease assay, two steps must be fully optimized: (1) Amplification of target region by PCR. This PCR should be very specific. (2) The generation of heteroduplex PCR product. The purity of PCR product, buffers, denaturing and annealing programs should be well optimized to generate authenticable heteroduplex PCR products. In addition to the practical steps, a few studies have highlighted that mismatch-specific endonuclease assay could underestimate the actual gene editing efficiency [22, 23]. Increasing evidence proves that indels generated by CRISPR for a specific site are not fully random [24]. For example, a given CRISPR target site has an activity as high as 90%, and the majority of indels is one base pair deletion. Most of the indel-containing PCR product will not form heteroduplex DNA. Thus, the mismatch-specific endonuclease assay can only be used to estimate the activity of gene editing tools.

3 CRISPR's Footprint Detection by TIDE and TIDER Assays

TIDE (Tracking of Indels by DEcomposition) and TIDER assays are developed by Bas van Steensel's group from the Netherlands Cancer Institute [25, 26]. According to our experience [27, 28], TIDE and TIDER assays are by far the most convenient, cost-effective, and fastest methods for quantifying indel frequency and genotyping. The assay is based two PCRs, two standard capillary sequencing (or Sanger sequencing), and sequencing trace decomposition algorithm. Unlike enzymatic-based assays which can only estimate the overall indel frequency, TIDE can provide quite accurate and sensitive calculation of indel frequency for each class of indels. Very detailed instructions and a friendly web-based TIDE analysis tool are provided by Bas van Steensel's group (*see* Chapter 3 of this book), which really benefit the whole gene editing community (<https://tide.deskgen.com>). Similarly, another web tool, CRISP-ID, has been developed for genotyping up to three alleles from a single Sanger sequencing trace [29]. The CRISP-ID serves as a very good method for genotyping gene-edited single clones.

4 CRISPR's Footprint Detection by IDAA Assay

IDAA (Indel Detection by Amplicon Analysis) is another method developed by Eric P. Bennett and colleagues from Copenhagen University [23, 30]. This assay is based on tri-primer amplicon labeling and DNA fragment size analysis using capillary electrophoresis, also see another similar method reported by Ramlee, M.K. and colleagues [31], also *see* Chapter 4 of this book. One conventional and smart design of the IDAA system is the utilization of a unique FAM-label primer, which can be used for almost any target site. This strategy greatly reduces the cost of using directly fluorescent-labeled target primers. IDAA has much more advantages over enzymatic-based indel quantification methods. Compared to the TIDE method, the IDAA is similar in its sensitivity, reproductively, and cost for indel quantification. However, IDAA can only give the quantification of each indel types such as one base pair insertion, while TIDE can provide more detailed information of the indels such as the constitution of this one base pair insertion. Nevertheless, IDAA provides a very good, conventional and cost-effective method for quantifying indel frequency.

5 CRISPR's Footprint Detection by High Resolution Melting Curve

High resolution melting (HRM) curve assay is based on the size- and base-dependent melting temperature of PCR amplicons

[32, 33]. The HRM method has been used for epigenetic genotyping, high-throughput genotyping, and SNP genotyping. Compared to enzymatic-based methods and TIDE, the HRM method can provide a much more cost-effective and high-throughput option for screening a large number of gene-edited clonal cells (Fig. 3). For HRM, the PCR reaction and melting curve determination can be performed in the same reaction tube consecutively, thus avoiding many practical steps such as purification of PCR product, concentration measuring, enzymatic-reaction and gel electrophoresis for enzymatic-based method, or preparation of Sanger sequencing for TIDE. However, the HRM method is not sensitive enough to measure overall indel frequency as compared to the aforementioned methods. In addition, the HRM method cannot quantify the frequency of different types of indels.

6 CRISPR's Footprint Detection by Drop-Off Assay

The Drop-Off assay is based on droplet digital PCR (ddPCR) for quantification of indel frequency and genotyping, developed by Lynne-Marie Postovit and colleagues from the University of Alberta and Bio-Rad Laboratorie [34]. This approach uses two fluorescent-labeled duplexed primer probes-based ddPCR for indel detection and quantification [35]. One reference probe binds to a distal none-target site of the amplicon, and another NHEJ/drop-off probe binds the nuclease target site. The Drop-Off assay has a great advantage for distinguishing mono- and bi-allelic modified cell clones, as well as having a relatively high sensitivity (up to 0.6% tested). However, since the method is based on probe hybridization, the Drop-Off method might not be able to distinguish 1 bp indels from wt. We and many other groups have found that, for most gRNA, 1 bp indels (1 bp deletion or insertion) are the most frequent indel types. The Drop-Off will underestimate overall indel frequency when used to evaluate gRNA efficiencies.

7 CRISPR's Footprint Detection by Surrogate Reporter Systems

Another approach of assaying the Cas9 gRNA activity is using surrogate reporter vectors. The finding that plasmids and extra-chromosomal DNAs can also utilize the endogenous DNA repair machineries leads the development of several surrogate reporter vector to sensing DBS and repair in cells [36]. Surrogate vectors have been made based on HDR, NHEJ, SSA, and fluorescent- or antibiotic-based reporters [37–41]. The common feature of all these surrogate reporter vectors is that a surrogate target site, which is identical to the actual target site in the genome, is introduced to the inactivated fluorescence or antibiotic gene of the

reporter vector. Previously, we have developed a system based on Split EGFP and SSA-mediated DSB repair, called C-Check [38], also *see* Chapter 5 of this book. Upon introduction of a DSB to the surrogate target site of C-Check vector, SSA-mediated DSB repair will generate a functional EGFP. We have validated that the C-Check-based CRISPR activity correlates well with TIDE-based quantification [28]. Although surrogate vectors cannot distinguish the different indel profiles, C-Check and other fluorescent surrogate reporter vectors serve as an extremely useful tool for enrichment of gene-edited cells [42].

8 CRISPR's Footprint Detection by Targeted Amplicon Next-Generation Sequencing (TA-NGS)

Targeted amplicon next-generation sequencing (TA-NGS) is a very straightforward method for both quantifying the activity and indel profile of CRISPR-Cas9 gene in cells. Similar to enzymatic-based assay, TIDE, or IDAA, TA-NGS is based on PCR amplification of the target region from cells treated with CRISPR-Cas9 components. The PCR product is then subjected to next-generation sequencing using any of the commercially available NGS platforms. Currently, many sequencing NGS platforms have been used for such purposes including Illumina miSeq system, BGISEQ500, Ion Torrent, etc. The TA-NGS method has several unique advances that other methods lack. First, TA-NGS has the highest sensitivity and R^2 coefficient when compared with TIDE, T7E1, and IDAA [23]. Second, TA-NGS is the only method that can give most spectrum of indels in the cells, including what kind of indels and their percentage. However, the TA-NGS method also has its disadvantages. First, the TA-NGS method is more expensive and takes a relatively longer period (approximately 1 week) as compared to other methods. Secondly, handling of the TA-NGS data could be impractical for most labs that are not familiar with NGS data. However, several web tools have been developed to facilitate the application of TA-NGS in gene editing experiments, such as the AGEseq [43], CRISPResso [44], BATCH-GE [45], and CRISPR-DAV [43] for analysis of CRISPR-Cas9 genome editing outcomes from deep sequencing data.

9 Genome-Scale Detection of CRISPR's Footprints

In addition to activity, specificity is another major concern of CRISPR-Cas9 or CRISPR-Cas9-derived gene editing and manipulation tools. Currently, by engineering the catalytic domain of Cas9 protein and using protein fusion strategy, the CRISPR-Cas9 system

has been developed for nicking single strand, gene interference, gene activation, DNA methylation, histone acetylation, base editing, etc. [46–51]. When discussing the specificity of CRISPR-Cas9, we have to look at it from two different angles: (1) cleaving nontarget sites and (2) binding to nontarget sites; all together we refer to these as CRISPR footprints.

All aforementioned methods are focusing on detecting and/or measuring the CRISPR-introduced cleavages and indels. In order to unbiasedly detect the global CRISPR cleavage footprint, a few *in vitro* and *in vivo* genome-scale CRISPR cleavage footprint methods have been developed. The commonly used *in vitro* methods for detecting CRISPR cleavage footprint are Digenome-seq [52] and CIRCLE-seq [53], which are based on *in vitro* nuclease-digested genomes (Digenome-seq), circularization of fragmented genomic DNA followed by nuclease digestion (CIRCLE-seq), and selective enrichment and identification of adapter-tagged DNA ends by sequencing (SITE-Seq) [54]. The disadvantage of *in vitro* methods is that they cannot recapitulate the effects of chromatin structure on CRISPR-Cas9 cleavage footprints. A few *in vivo* methods have been developed to detect the genome-wide CRISPR cleavages in cells, including high-throughput, genome-wide translocation sequencing (HTGTS) [55], capturing of double-stranded oligodeoxynucleotides into DSBs (GUIDE-seq) [56], capturing of integrase-defective lentiviral vectors (IDLVs) into DSBs (IDLV-seq) [57], and directly labeling DSBs *in situ* with biotinylated linker and streptavidin enrichment (BLESS) [58]. All these *in vivo* methods commonly rely on the DSBs introduced by CRISPR/Cas9, and capture DSBs chemically or through inserting a DNA tag. In addition, whole-genome sequencing (WGS) has been broadly used to identify CRISPR cleavage footprints [59, 60]. However, the WGS method is only recommended to genotype whether off-targets have been introduced to CRISPR-edited cell clones or gene-edited animals. Due to the high frequency of somatic mutations and polymorphism, cautions should be carefully taken when defining single-nucleotide polymorphisms (SNP), indels (deletions or insertions) resulted from CRISPR cleavage or simply from parental inheritance or somatically accumulated mutations.

Another major type of CRISPR's footprint is binding to nontarget sites. This kind of footprints is more prominent than the cleavage footprint. The binding of CRISPR-Cas9 to DNA is depending on both the Cas9 protein's DNA-binding ability and the Watson-Crick base pairing between the gRNA spacer and the target site. Through neutralizing Cas9 protein's DNA-binding domain, high-fidelity Cas9 proteins have been generated [61, 62]. Unlike DNA cleavage, CRISPR-Cas9 is much more loosely depending on the matches between the gRNA spacer and the target site. Three mismatches between the gRNA spacer and the target site can almost completely prevent DNA cleavage by

CRISPR-Cas9, whereas CRISPR-Cas9 can still firmly bind to sites with 9 mismatches [63]. The different characteristics between CRISPR DNA cleavage and binding footprints highlight the importance of developing and using the right method to validate the specificity of CRISPR-Cas9-based gene engineering tools. For example, the catalytically inactivated Cas9 (dCas9) has been harnessed for epigenome editing such as histone acetylation, DNA demethylation, DNA methylation, etc. Previously, we showed that the fusion of dCas9 to the DNA methylation catalytic domain of DNMA3A or DNMT3B can achieve efficient methylation of the CpG dinucleotides adjacent to the target site [64]. Using targeted bisulfite sequencing for a few selected potential off-target sites selected based on sequence similarity, we cannot detect off-target methylation. However, using unbiased whole-genome bisulfite sequencing, we found that the dCas9-DNMT3A and dCas9-DNMT-3B can cause quite a large number of unspecific methylations, which is enriched in open chromatin regions, promoter, 5'UTR regions, and CpG island. Similarly, when defining the specificity of Cas9-derived effectors that rely on DNA binding of dCas9-gRNA, unbiased CRISPR footprint detection methods should be chosen.

10 CRISPR Indels Are Not Random

It is commonly thought that the DSBs introduced by gene editing tools, e.g., CRISPR/Cas9 in mammalian cells are predominantly repaired by NHEJ and MMEJ pathway, which is error prone and random. The formation of indels has thus been considered as random and unpredictable in cells. As our knowledge of the indel profiles increases from CRISPR-edited studies, it becomes clear that these indel events are not random. Megan van Overbeek and colleagues analyze indel profiles of 223 target sites in human cells and found that classical NHEJ and MMEJ are the major DNA repair machineries in mediated DSB repair, and the protospacer sequences (target site) determine the indel outcomes [24]. Consistent with that, we also observed that for each gRNA, the indel profiles are consistent across experimental repeats (Fig. 4) and different cell types [28] (Fig. 5).

Another frequently observed characteristics of gene editing indels is the highly frequent insertion of 1 bp at the break site [65], also see Figs. 4 and 5. It has been found that small indels occur earlier during CRISPR gene editing. Consistent with that, a study from Eric Bennett's group systematically evaluates the effect of different delivery formats (lentivirus transduction, plasmid lipofection, and ribonuclear protein electroporation) and time on CRISPR cleavage indel profiles [65], and observes similar indel profile characteristics. Large indels and deletions, which are

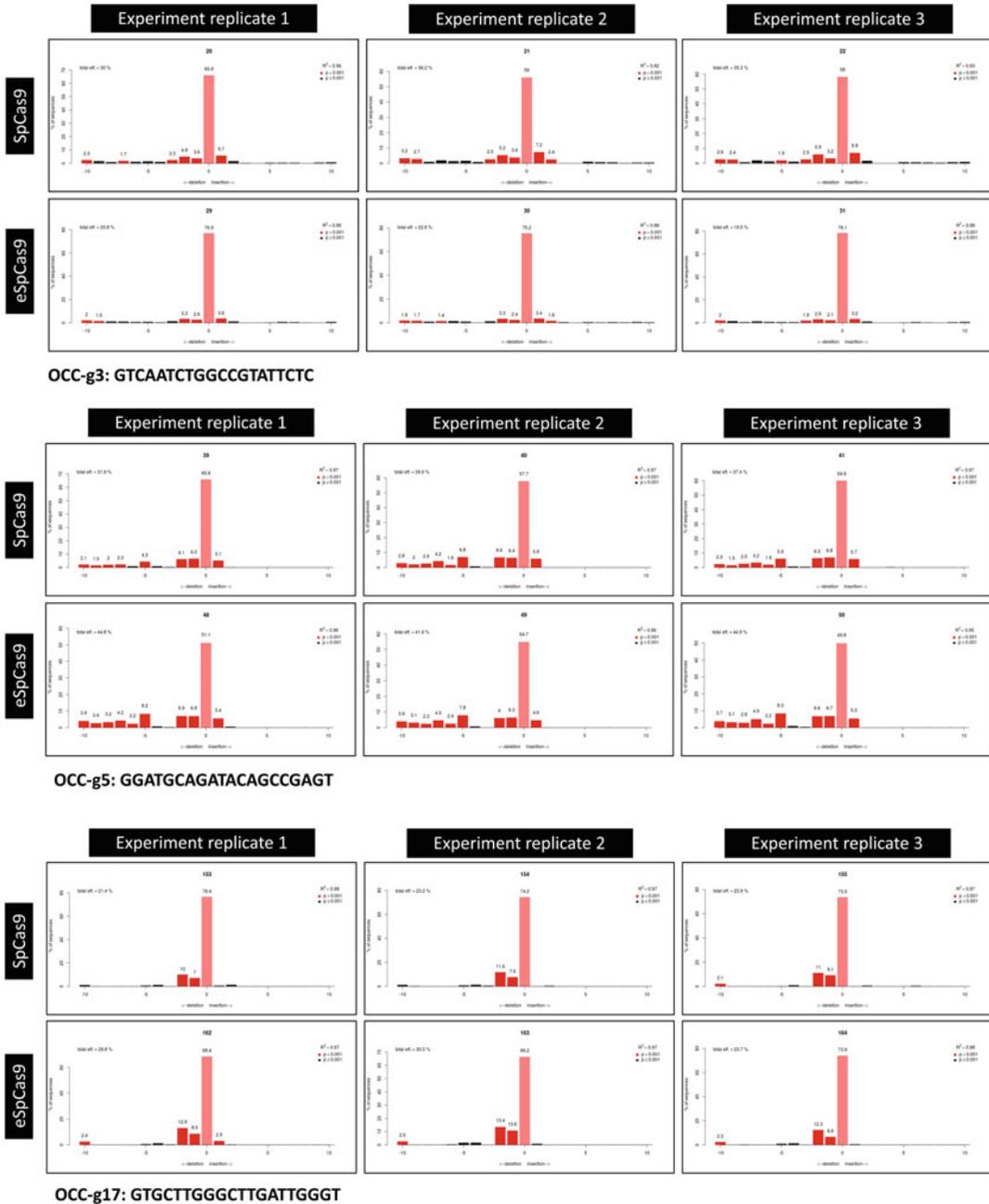


Fig. 4 Profiling of CRISPR indel patterns of three gRNAs by TIDE in HEK293T cells. Experiments have been conducted in triplicates and tested with wild-type SpCas9 and the enhanced SpCas9 (eSpCas9)

mediated by MMEJ or SSA, appear much later [24]. These large indels are more frequently found in clonal gene-edited cells, which have been prorogated from a single cell to a cell clone. It is thus not surprising to find large deletions and complex rearrangements after repair of DSBs introduced by CRISPR Cas9 in a report recently [66].

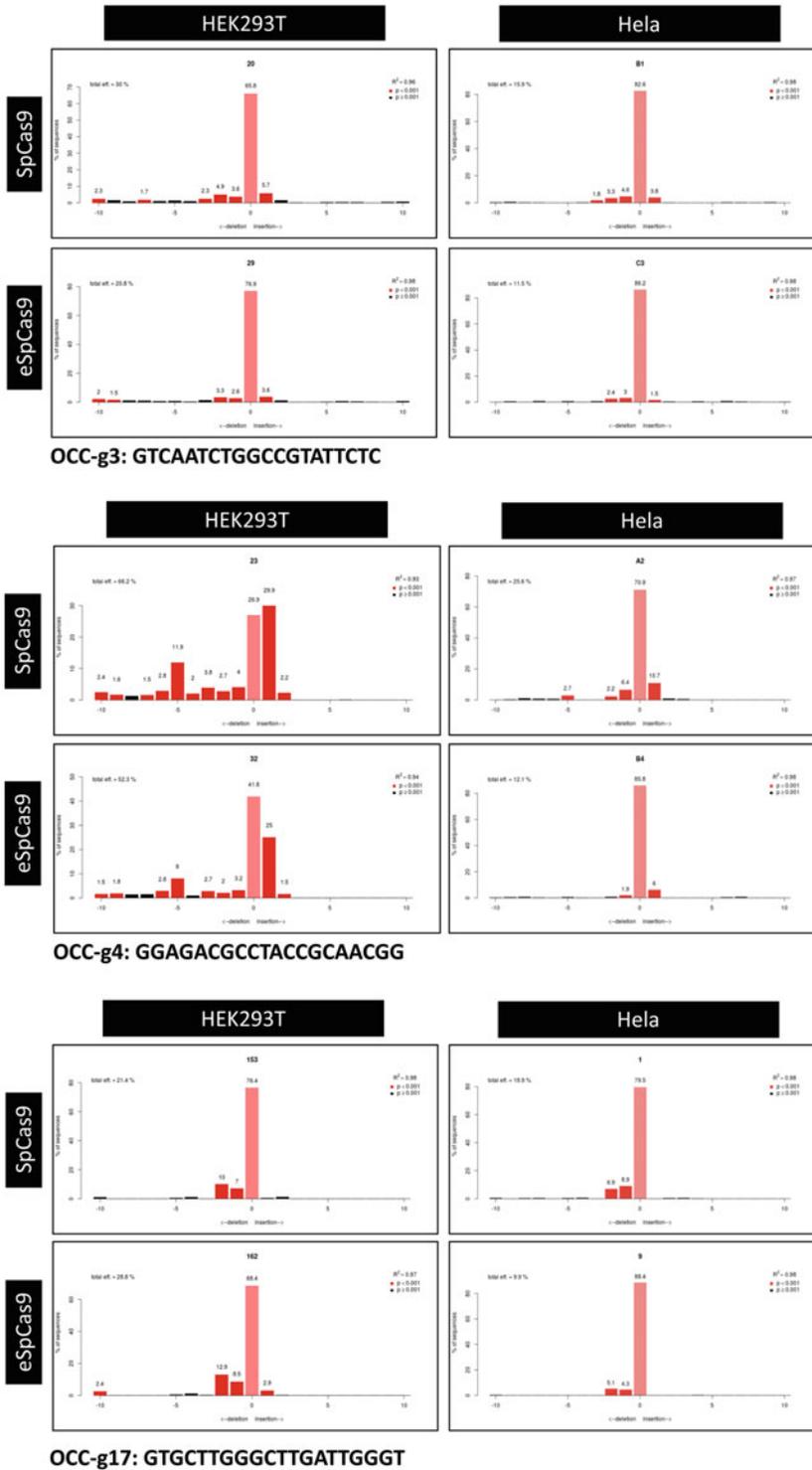


Fig. 5 Profiling of CRISPR indel patterns of three gRNAs with wild-type SpCas9 or the enhanced SpCas9 (eSpCas9) by TIDE in HEK293T and HeLa cells

One advantage of DSBs not being repaired randomly is to predict the indel outcomes for each CRISPR target site. Based on MMEJ-mediated DSB repair, the CRISPor web tool incorporates the function for predicting indels [67]. And recently, the first achievement of predicting CRISPR chromosomal editing events was reported [68], which is achieved by specific PAM configurations. However, since the overhang cleavage profiles also affect indel outcome, a full prediction of indel outcome for a single gRNA is still required and will be extremely useful for increasing the accuracy of *in silico* predicting CRISPR activity in the future.

11 Concluding Remarks

The CRISPR gene editing technology has now been broadly adopted by almost all gene-related applications and research groups. Many assisting methods developed for CRISPR gene editing have contributed greatly to the technology's evolution, such as *in silico* web tools for CRISPR design, software for off-target prediction and data analysis, methods to enhance delivery, modifications to enhance targeting editing efficiency, methods for tracking the CRISPR's footprints, etc. Currently, most methods for detecting CRISPR indels or activity are focusing on small indels. Although compared to these small indels, large indels are much rarer and occur much later after CRISPR gene editing. Large indels are more frequently found in cells if the CRISPR gene editing vectors are stably integrated. Efficient, cost-effective and high-throughput methods for tracking CRISPR-induced large indels are required in the future.

Since there are many methods for tracking CRISPR's footprints, the next question is which one we should choose for our experiment, also see outstanding questions. This is quite depending on the expertise and accessibility to the required facilities of each research group. In our experience, a combination of these methods is recommended to achieve a successful CRISPR gene editing outcome. For each CRISPR gene editing project, first, we carefully designed a few gRNAs, TIDE assay primers, and C-Check vectors. Second, we measure each gRNA's activity by TIDE and C-Check in the actually targeting cells and surrogate cells (HEK293). Finally, the best gRNA will be selected based on TIDE and C-Check results. C-Check is further used to enrich cells carrying the desired mutations by fluorescence-activated cell sorting.

The CRISPR technology is still fast growing and more improvements are needed to bring the technology to successful clinical applications. Tracking the CRISPR's footprints is of great importance. We have recently discovered that when a piece of DNA/gene is deleted by CRISPR, this piece of DNA, instead of just simply degrades, forms a functional extrachromosomal circular

DNA in cells [69]. This discovery expands the genome engineering applications of CRISPR. A better understanding of how DBSs are repaired in cells, the outcome of editing loci, and even the faith of the deleted DNA or genes by CRISPR will continue to contribute to evolutions of the CRISPR gene editing.

Acknowledgments

L.L. is supported by grants from the Lundbeck Foundation. Y.L. is supported by BGI-Shenzhen, BGI-Qingdao, and grants from the Shenzhen Sanming Medical Project. We thank the whole team of Lars Bolund Institute of Regenerative Medicine (LBI), BGI, for their work and assistance on the CRISPR technologies, and especially Jun Wang from LBI for assistance with preparing Fig. 1. Y.L. is also supported by the Guangdong Provincial Key Laboratory of Genome Read and Write (No. 2017B030301011).

Disclaimer Statement: The views expressed in this article are the personal views of the author and may not be understood or quoted as being made on behalf of or reflecting the position of the Lars Bolund Institute of Regenerative Medicine, BGI, or one of its working parties.

References

1. Chang HHY et al (2017) Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol* 18(8):495–506
2. McVey M, Lee SE (2008) MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet* 24(11):529–538
3. Ceccaldi R et al (2016) Repair pathway choices and consequences at the double-strand break. *Trends Cell Biol* 26(1):52–64
4. Lahue RS et al (1989) DNA mismatch correction in a defined system. *Science* 245(4914):160–164
5. Lindahl T (1974) An N-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues. *Proc Natl Acad Sci U S A* 71(9):3649–3653
6. Sancar A, Rupp WD (1983) A novel repair enzyme: UVRABC excision nuclease of *Escherichia coli* cuts a DNA strand on both sides of the damaged region. *Cell* 33(1):249–260
7. Mehta A, Haber JE (2014) Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harb Perspect Biol* 6(9):a016428
8. Capecchi MR (2005) Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat Rev Genet* 6(6):507–512
9. Epinat JC et al (2003) A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Res* 31(11):2952–2962
10. Arnould S et al (2007) Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells. *J Mol Biol* 371(1):49–65
11. Boch J et al (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326(5959):1509–1512
12. Jinek M et al (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096):816–821
13. Cong L et al (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121):819–823

14. Mali P et al (2013) RNA-guided human genome engineering via Cas9. *Science* 339 (6121):823–826
15. Barrangou R, Doudna JA (2016) Applications of CRISPR technologies in research and beyond. *Nat Biotechnol* 34(9):933–941
16. Yeung AT et al (2005) Enzymatic mutation detection technologies. *BioTechniques* 38 (5):749–758
17. Zhu X et al (2014) An efficient genotyping method for genome-modified animals and human cells generated with CRISPR/Cas9 system. *Sci Rep* 4:6420
18. Liu C et al (2017) Delivery strategies of the CRISPR-Cas9 gene-editing system for therapeutic applications. *J Control Release* 266:17–26
19. Li L et al (2018) Non-viral delivery systems for CRISPR/Cas9-based genome editing: challenges and opportunities. *Biomaterials* 171:207–218
20. Senis E et al (2014) CRISPR/Cas9-mediated genome engineering: an adeno-associated viral (AAV) vector toolbox. *Biotechnol J* 9 (11):1402–1412
21. Schmidt F, Grimm D (2015) CRISPR genome engineering and viral gene delivery: a case of mutual attraction. *Biotechnol J* 10 (2):258–272
22. Ehrke-Schulz E et al (2016) Quantification of designer nuclease induced mutation rates: a direct comparison of different methods. *Mol Ther Methods Clin Dev* 3:16047
23. Sentmanat MF et al (2018) A survey of validation strategies for CRISPR-Cas9 editing. *Sci Rep* 8(1):888
24. van Overbeek M et al (2016) DNA repair profiling reveals nonrandom outcomes at Cas9-mediated breaks. *Mol Cell* 63 (4):633–646
25. Brinkman EK et al (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42 (22):e168
26. Brinkman EK et al (2018) Easy quantification of template-directed CRISPR/Cas9 editing. *Nucleic Acids Res* 46(10):e58
27. Lin L et al (2017) Fusion of SpCas9 to E. coli Rec A protein enhances CRISPR-Cas9 mediated gene knockout in mammalian cells. *J Biotechnol* 247:42–49
28. Jensen KT et al (2017) Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. *FEBS Lett* 591(13):1892–1901
29. Dehairs J et al (2016) CRISP-ID: decoding CRISPR mediated indels by Sanger sequencing. *Sci Rep* 6:28973
30. Yang Z et al (2015) Fast and sensitive detection of indels induced by precise gene targeting. *Nucleic Acids Res* 43(9):e59
31. Ramlee MK et al (2015) High-throughput genotyping of CRISPR/Cas9-mediated mutants using fluorescent PCR-capillary gel electrophoresis. *Sci Rep* 5:15587
32. D'Agostino Y et al (2016) A rapid and cheap methodology for CRISPR/Cas9 zebrafish mutant screening. *Mol Biotechnol* 58 (1):73–78
33. Samarut E et al (2016) A simplified method for identifying early CRISPR-induced indels in zebrafish embryos using High Resolution Melting analysis. *BMC Genomics* 17:547
34. Findlay SD et al (2016) A digital PCR-based method for efficient and highly specific screening of genome edited cells. *PLoS One* 11(4): e0153901
35. Pinheiro LB et al (2012) Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal Chem* 84(2):1003–1011
36. Kim H et al (2011) Surrogate reporters for enrichment of cells with nuclease-induced mutations. *Nat Methods* 8(11):941–943
37. Wen Y et al (2017) A stable but reversible integrated surrogate reporter for assaying CRISPR/Cas9-stimulated homology-directed repair. *J Biol Chem* 292(15):6148–6162
38. Zhou Y et al (2016) Enhanced genome editing in mammalian cells with a modified dual-fluorescent surrogate system. *Cell Mol Life Sci* 73(13):2543–2563
39. Fu L et al (2016) A simple and efficient method to visualize and quantify the efficiency of chromosomal mutations from genome editing. *Sci Rep* 6:35488
40. Yang Y et al (2016) Highly efficient and rapid detection of the cleavage activity of Cas9/grRNA via a fluorescent reporter. *Appl Biochem Biotechnol* 180(4):655–667
41. Ramakrishna S et al (2014) Surrogate reporter-based enrichment of cells containing RNA-guided Cas9 nuclease-induced mutations. *Nat Commun* 5:3378
42. Hussmann D et al (2017) IGF1R depletion facilitates MET-amplification as mechanism of acquired resistance to erlotinib in HCC827 NSCLC cells. *Oncotarget* 8 (20):33300–33315

43. Xue LJ, Tsai CJ (2015) AGEseq: analysis of genome editing by sequencing. *Mol Plant* 8 (9):1428–1430
44. Pinello L et al (2016) Analyzing CRISPR genome-editing experiments with CRISPResso. *Nat Biotechnol* 34(7):695–697
45. Boel A et al (2016) BATCH-GE: batch analysis of next-generation sequencing data for genome editing assessment. *Sci Rep* 6:30330
46. Qi LS et al (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152 (5):1173–1183
47. Ran FA et al (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154 (6):1380–1389
48. Hilton IB et al (2015) Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat Biotechnol* 33(5):510–517
49. Vad-Nielsen J et al (2018) Simple method for assembly of CRISPR synergistic activation mediator gRNA expression array. *J Biotechnol* 274:54–57
50. Xiong K et al (2017) RNA-guided activation of pluripotency genes in human fibroblasts. *Cell Reprogram* 19(3):189–198
51. Gaudelli NM et al (2017) Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551 (7681):464–471
52. Kim D et al (2015) Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods* 12 (3):237–243. 1 p following 243
53. Tsai SQ et al (2017) CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat Methods* 14(6):607–614
54. Cameron P et al (2017) Mapping the genomic landscape of CRISPR-Cas9 cleavage. *Nat Methods* 14(6):600–606
55. Frock RL et al (2015) Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat Biotechnol* 33 (2):179–186
56. Tsai SQ et al (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* 33 (2):187–197
57. Wang X et al (2015) Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat Biotechnol* 33(2):175–178
58. Crosetto N et al (2013) Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat Methods* 10 (4):361–365
59. Smith C et al (2014) Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. *Cell Stem Cell* 15(1):12–13
60. Veres A et al (2014) Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. *Cell Stem Cell* 15(1):27–30
61. Slaymaker IM et al (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science* 351(6268):84–88
62. Kleinstiver BP et al (2016) High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529 (7587):490–495
63. Kucsu C et al (2014) Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol* 32 (7):677–683
64. Lin L et al (2018) Genome-wide determination of on-target and off-target characteristics for RNA-guided DNA methylation by dCas9 methyltransferases. *Gigascience* 7(3):1–19
65. Kosicki M et al (2017) Dynamics of indel profiles induced by various CRISPR/Cas9 delivery methods. *Prog Mol Biol Transl Sci* 152:49–67
66. Kosicki M et al (2018) Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol* 36:765–771
67. Haeussler M et al (2016) Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol* 17(1):148
68. Shou J et al (2018) Precise and predictable CRISPR chromosomal rearrangements reveal principles of Cas9-mediated nucleotide insertion. *Mol Cell* 71(4):498–509.e4
69. Henrik Devitt Møller LL, Xi X, Petersen TS, Huang J, Yang L, Kjeldsen E, Jensen UB, Zhang X, Liu X, Xun X, Wang J, Yang H, Church GM, Bolund L, Regenberg B, Luo Y (2018) CRISPR-C: circularization of genes and chromosome by CRISPR in human cells. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gky767>



Rapid Quantitative Evaluation of CRISPR Genome Editing by TIDE and TIDER

Eva Karina Brinkman and Bas van Steensel

Abstract

Current genome editing tools enable targeted mutagenesis of selected DNA sequences in many species. However, the efficiency and the type of introduced mutations by the genome editing method are largely dependent on the target site. As a consequence, the outcome of the editing operation is difficult to predict. Therefore, a quick assay to quantify the frequency of mutations is vital for a proper assessment of genome editing actions. We developed two methods that are rapid, cost-effective, and readily applicable: (1) TIDE, which can accurately identify and quantify insertions and deletions (indels) that arise after introduction of double strand breaks (DSBs); (2) TIDER, which is suited for template-mediated editing events including point mutations. Both methods only require a set of PCR reactions and standard Sanger sequencing runs. The sequence traces are analyzed by the TIDE or TIDER algorithm (available at <https://tide.nki.nl> or <https://deskgen.com>). The routine is easy, fast, and provides much more detailed information than current enzyme-based assays. TIDE and TIDER accelerate testing and designing of DSB-based genome editing strategies.

Key words CRISPR-Cas, Genome editing, Indel mutation, Mutagenesis, DNA mutational analysis/methods, Sanger sequencing, Web tool, Algorithm

1 Introduction

CRISPR-based systems are popular and widely used for genome editing in the field of molecular biology. CRISPR endonuclease Cas9 introduces a DSB into the genomic DNA with high precision. Due to the error-prone repair mechanisms of the cell, this often results in insertions or deletions at the targeted site [1]. This is exploited to make functional knock-outs of specific genes and regulatory elements [2–4]. Alternatively, to gain more control over the nature of the mutations, strategies have been developed that introduce small nucleotide changes around a precisely targeted site by using a donor template [5, 6]. In the latter approach the genomic DNA around the DSB break is replaced by the DNA of the donor template through homology-directed repair (HDR),

resulting in the introduction of a designed mutation with high accuracy [7, 8]. This precise editing creates the possibility to generate and study specific disease-causing nucleotide variants [6, 9]. Typically, one starts with a homogeneous cell line and ends up with a pool of cells with a complex mix of indels and/or designer mutations [10–12]. To study a mutation of interest, clonal mutant lines need to be isolated from the cell pool. Because this is a very labor-intensive process it is important to know *a priori* the efficiency in which the desired mutation(s) have been introduced. However, a complicating factor is that the efficacy of the programmable nucleases can vary dramatically depending on the sequence that is targeted. In addition, different cell types have a varying performance in transfection capability. These factors make the efficacy of CRISPR experiment difficult to predict. For this reason it is usually necessary to test several guide RNAs (gRNAs) that lead the endonuclease to the site of interest. This is even more critical when a template-directed strategy is applied, which often has a low efficiency because HDR repair pathways are generally less active than error-prone non-templated repair [10, 12]. Hence, a quick and easy assay to estimate the frequencies of the diverse introduced mutations in the cell pool is of key importance.

We developed two methods that can accurately quantify the efficiency of either indels or template-directed mutations in a pool of cells. Both methods are rapid and cost-effective. The method TIDE (Tracking of Indels by DEcomposition) identifies and quantifies indels. It requires only a pair of standard Sanger sequence traces of two PCR products [13]. The sequence traces are then analyzed using an easy-to-use web tool. Note that TIDE can only detect overall indel frequencies, but not nucleotide substitutions or specifically designed indels. For the latter purpose we developed TIDER (Tracking of Insertions, DEletions, and Recombination events) [14]. This method can estimate the incorporation frequency of template-directed mutations, including point mutations, and distinguish them from a background of additional indels that originate from competing erroneous repair pathways. Although TIDER can also quantify indels alone, TIDE is slightly simpler to implement and therefore more suited for the assessment of non-templated editing experiments. The corresponding web tools for both TIDE and TIDER are freely accessible at <http://tide.nki.nl>.

2 Materials

2.1 Guide RNA Design

TIDE and TIDER are suitable for any species in which genomic editing experiments can be performed. CRISPR guide RNAs can be designed using various online design web tools (e.g., <http://crispr.mit.edu/>, <https://chopchop.rc.fas.harvard.edu/>, <https://www.deskgen.com/>).

2.2 DNA Purification Buffers and Solutions

Usually, 1–3 days after transfection genomic DNA is isolated. Genomic DNA of a minimum of 1000 cells should be isolated to get a comprehensive sampling of the complexity of the mutations that are introduced by the repair of the CRISPR-Cas9 double strand break. A standard genomic DNA isolation Kit (e.g., BioLine ISOLATE II Genomic DNA Kit) can be used according to the manufacturer’s protocol. DNA can also be isolated with the protocol for isolation of high-molecular-weight DNA from mammalian cells using proteinase K and phenol/chloroform extraction [15].

2.3 PCR Amplification of Control and Experimental Sample DNA

PCR reactions are carried out with primers surrounding the expected break site. We advise to amplify and sequence a stretch of DNA 500–1500 bp enclosing the designed editing site. The projected break site should be located preferably ~200 bp downstream from the sequencing start site.

1. Genomic DNA.
2. PCR primers (*see Note 1*).
3. PCR master mix (example makes 50 μ L):

21- \times μ L	H ₂ O
2 μ L	Primer a (10 μ M stock)
2 μ L	Primer b (10 μ M stock)
\times μ L	Genomic DNA (~50 ng)
25 μ L	2 \times pre-mix of buffer, Taq polymerase, and dNTPs (e.g., BioLine MyTaq)

4. PCR program:

Step	Temperature	Time (min:s)	Number of cycles
Initial denaturation	95 °C	1:00	1
Denaturation	95 °C	0:15	25–30 \times
Annealing	55–58 °C	0:15	
Extension	72 °C	0:10	
Final extension	72 °C	1:00	
	4 °C	Hold	

5. Check an aliquot of the PCR product on 1–2% agarose gel. A sharp single band should be visible.
6. Purify the PCR product using a kit according to manufacturer’s instructions (e.g., BioLine ISOLATE II PCR and Gel Kit).

**2.4 Two-Step PCR
Amplification
of Reference DNA
(TIDER Only)**

1. Genomic DNA.
2. PCR primers (*see* **Notes 2** and **3**).
3. PCR master mix (example makes 50 μL):

	PCR mix1	PCR mix2
21- \times μL	H ₂ O	H ₂ O
2 μL	Primer a (10 μM stock)	Primer d (10 μM stock)
2 μL	Primer c (10 μM stock)	Primer b (10 μM stock)
\times μL	Genomic DNA (~50 ng)	Genomic DNA (~50 ng)
25 μL	2 \times pre-mix of buffer, Taq polymerase, and dNTPs (e.g., BioLine MyTaq)	2 \times pre-mix of buffer

4. PCR program:

Step	Temperature	Time (min:s)	Number of cycles
Initial denaturation	95 °C	1:00	1
Denaturation	95 °C	0:15	25 \times
Annealing	55–58 °C	0:15	
Extension	72 °C	0:10	
Final extension	72 °C	1:00	
	4 °C	Hold	

5. Purify PCR product using kit and manufacture instructions (e.g., BioLine ISOLATE II PCR and Gel Kit).
6. Anneal the following two PCR products for 1 min 95 °C, cool down to 20 °C (0.1 degrees/s).

48 μL	Annealing buffer (=10 mM Tris, 50 mM NaCl, 1 mM EDTA)
1 μL	PCR mix1
1 μL	PCR mix2

7. Extend the annealed products and amplify the joined product.

18 μL	H ₂ O
2 μL	Primer a (10 μM stock)
2 μL	Primer b (10 μM stock)
3 μL	Annealed oligo mix
25 μL	2 \times pre-mix of buffer, Taq polymerase, and dNTPs (e.g., BioLine MyTaq)

8. PCR program:

Step	Temperature	Time (min:s)	Number of cycles
Initial denaturation	95 °C	1:00	1
Denaturation	95 °C	0:15	25×
Annealing	55–58 °C	0:15	
Extension	72 °C	0:10	
Final extension	72 °C	1:00	
	4 °C	Hold	

9. Check the PCR product on 1–2% agarose gel. A sharp single band should be visible.
10. Purify the PCR product using a kit and manufacturer's instructions (e.g., BioLine ISOLATE II PCR and Gel Kit).

2.5 Sanger Sequencing

We strongly recommend that all PCR products (control, experimental sample(s), and for TIDER also the reference) are sequenced in parallel. Purified PCR samples are prepared for Sanger sequencing with the following protocol or can be send for commercial Sanger sequencing.

1. Purified PCR samples (100 ng).
2. PCR primers. Similar primers as in Subheading 2.3 can be used (*see* **Notes 1** and **3**).
3. PCR master mix (example makes 20 µL):

15.5-× µL	H ₂ O
0.5 µL	Primer a or primer b (10 µM stock)
× µL	Purified PCR samples (100 ng)
4 µL	BigDye (e.g., BigDye [®] terminator v3.1 of Applied Biosystems)

PCR program:

Step	Temperature	Time (min:s)	Number of cycles
Initial denaturation	96 °C	1:00	1
Denaturation	96 °C	0:30	30×
Annealing	50 °C	0:15	
Extension	60 °C	4:00	
	4 °C	Hold	

4. Samples are analyzed by a Sanger sequence instrument (e.g., Applied Biosystems 3730 × 1 DNA Analyzer). Sequence trace files must be saved in .ab1 or .scf format.

2.6 Equipment

1. Cell counter.
2. Microcentrifuge.
3. PCR cycler.
4. Nanodrop.

2.7 Software

The TIDE and TIDER web tools are both available at <https://tide.nki.nl> or <https://deskgen.com>.

3 Methods

3.1 Control and Experimental Sample Generation

For both methods genomic DNA is isolated from the cell pool that was transfected with the nuclease or guide RNA alone (control) and from cells exposed to both Cas9 and guide RNA (experimental sample). For TIDER the experimental sample is also co-transfected with the donor template. Then a region of about 500–1500 base pairs around the target site is amplified by PCR from DNA of the control and experimental sample (Fig. 1a, b). Next, the PCR amplicons are subjected to conventional Sanger sequencing. In the PCR product of the experimental sample, the sequence trace may consist of a combination of multiple sequences derived from unmodified DNA and DNA that has acquired a mutation (Fig. 2a).

3.2 Reference Sample Generation (TIDER Only)

TIDER is required for genome editing experiments in the presence of a donor template. In addition to the control and experimental sample trace (*see* Subheading 3.1), TIDER requires one extra Sanger sequencing trace called “reference.” The reference is similar to the control sequence, except that it carries the desired base pair changes as designed in the donor template (Fig. 2e). There are two paths to obtain the reference sequence as described below.

The reference sequence can be easily created in a 2-step PCR protocol based on site-directed mutagenesis [16]. Here, two additional primers are required that overlap and carry the desired mutation(s) (mutated primers c, d, which are reverse complement of each other) (Fig. 1c). These primers are used in combination with the primers used for the amplification of the control and experimental sample (control primers a, b). The control forward primer a is combined with the reverse mutated primer c and the forward mutated primer d with the control reverse primer b, resulting in two PCR amplicons that incorporate the designed mutations. Then the two amplicons are denatured and hybridized at the complementary ends in an annealing reaction. The second PCR uses the

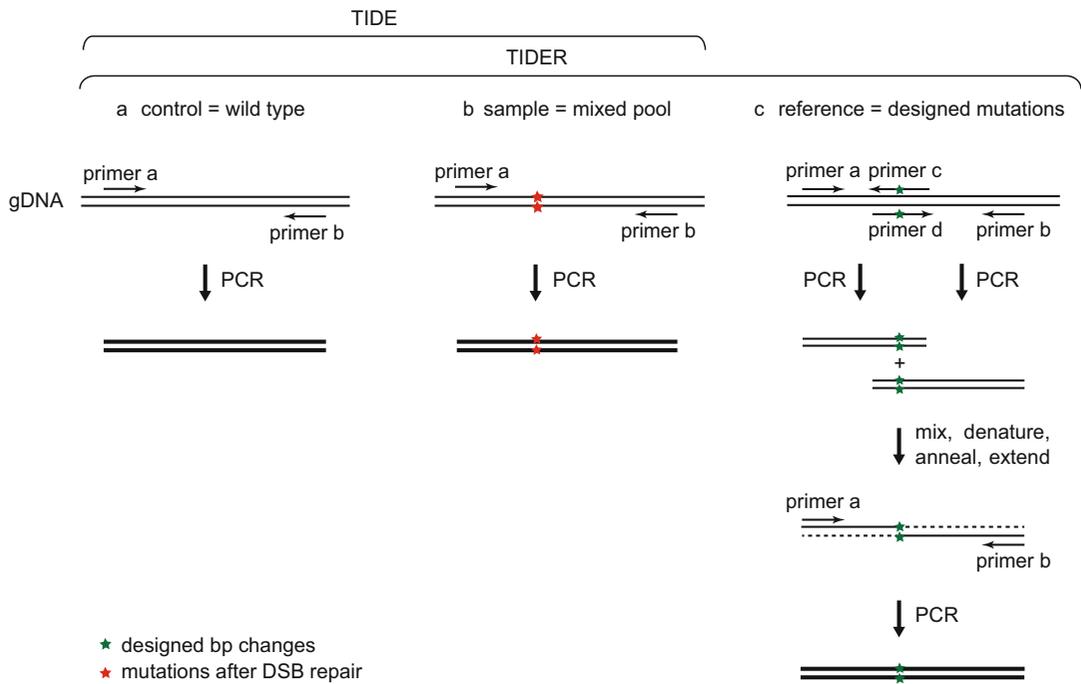


Fig. 1 Method to generate the required input samples for TIDE and TIDER. Control and test samples can be obtained by PCR using primers spanning the CRISPR target site (primers a, b). The reference sequence (TIDER only) can be created in a similar way as site-directed mutagenesis [16] (see Subheading 3.2 for detailed explanation)

annealing mixture as a template and the control forward and reverse (primers a and b) as primers. This PCR starts with an extension step followed by exponential amplification. This results in a PCR product carrying the designed mutations (see **Notes 2** and **3**).

Alternatively, the reference DNA can be ordered as synthesized DNA. The design should include a similar DNA code as the PCR product of the control sample, except that it should carry the designed mutation(s) as in the donor template. The annealing sequences for the forward and reverse primers (a, b) should also be present in the synthesized fragment. Similar to the control and test sample, the reference can be amplified with primer a, b (see **Note 3**).

3.3 Web Tool

The PCR products of the control, optional reference, and experimental sample are processed by conventional Sanger sequencing. The resulting sequence trace files (.ab1 or .scf format) are then uploaded into the TIDE or TIDER web tool (both available at <http://tide.nki.nl> and <https://deskgen.com>). In addition, a character string representing the guide RNA sequence (20 nt) is required as input (see **Notes 4** and **5**). Then, the software will perform several calculations. First, the guide RNA sequence is

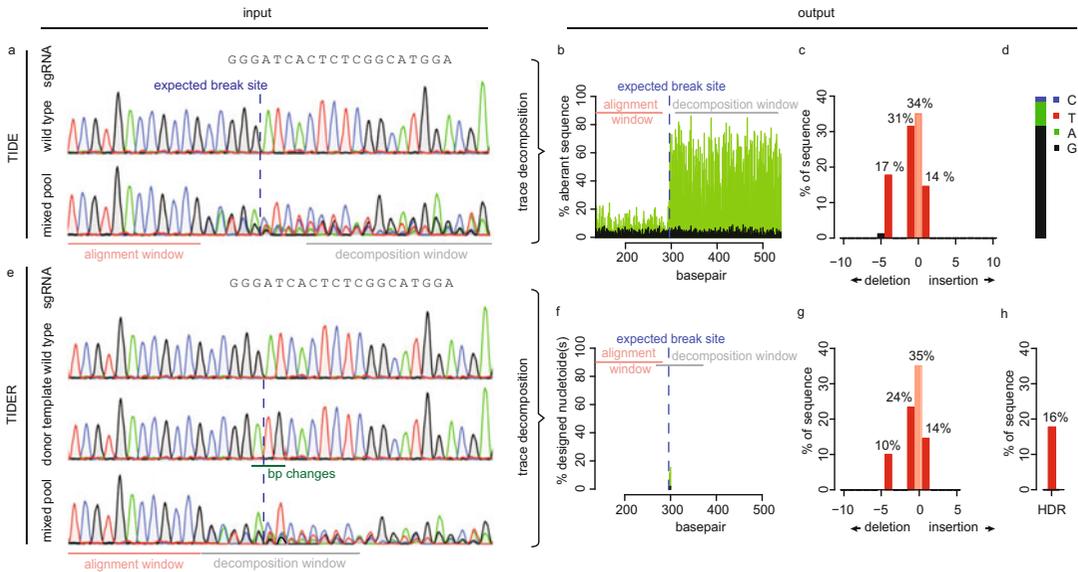


Fig. 2 Overview of TIDE and TIDER algorithm. Due to imperfect repair (and repair by homology-directed repair with a donor template) after cutting by a targeted nuclease, the DNA in the cell pool consists of a mixture of indels (and designed mutations). The various introduced mutations in the pool are disentangled by TIDE or TIDER. **(a)** TIDE requires as input a guide RNA sequence string and two sequences are required: (1) wild-type control, (2) composite test sample. **(b)** For quality control the aberrant sequence signal is visualized in control (black) and treated sample (green), the expected break site (vertical dotted line), region used for alignment (pink bar), and the region used for decomposition (gray bar). A constant composite sequence signal is yielded after the break site. **(c)** Trace decomposition yields the spectrum of indels with their frequencies. **(d)** In presence of +1 insertions, the base composition is estimated. **(e)** Input files for TIDER are identical to TIDE and one additional sequence file with designed mutations in the used donor template. **(f)** Quality plot showing only the proportion of desired mutated nucleotide(s) as designed in donor template that is/are present in the control (black) and treated sample (green). The region for alignment (pink bar) and decomposition (gray bar) as used in TIDER are represented. **(g and h)** Decomposition gives the spectrum of indels **(g)** and the HDR events **(h)** with their frequencies

aligned to the control sequence in order to determine the position of the expected Cas9 break site. Next, in all Sanger sequence traces an alignment window is automatically selected that runs from 100 to 15 bp upstream of the break site. The sequence segment in this window of the experimental sample (and the optional reference) is aligned to that of the control in order to determine any offset between the sequence reads. Users may change the default settings for these calculations, which is necessary when alignment problems occur with these settings (*see* **Notes 6** and **7**). Subsequently, two output plots are generated: one plot that can help with quality control and one that displays the indel/HDR spectrum.

3.4 Quality Control

For generation of the quality control plot the signals of all nucleotides: A, G, T, C at each position in the sequence file are used. In general, each position in the sequence trace is represented

by one predominant nucleotide signal indicative of the actual nucleotide. The minor signals from the other three nucleotides are normally considered as background. In TIDE(R) the percentage of these aberrant nucleotides is plotted along the sequence trace of the control and the experimental sample. Thus, a value of 0% at a position indicates that the detected nucleotide does not differ from the control sequence while a value of 100% indicates that the expected nucleotide was not detected at all (and instead only one or more of the other three nucleotides) (Fig. 2b). The percentages of aberrant nucleotides in the control should be low along the whole sequence trace. However, the experimental sample consists of a mixture of multiple sequences due to the presence of indels and possible point mutations. Around the break site the sequences start to deviate from the control, which is visible with consistently elevated signal of the aberrant sequence signal. Note that there is a 25% chance that an identical nucleotide in a mutated sequence is found as is present in the control sequence at the same position, because there are only 4 different nucleotides available. This plot allows the user to visually inspect the sequence deviation caused by the targeted nuclease and enables to verify the alignments and quality of the data. It is important to confirm that (1) the break site is located as expected, (2) the aberrant signal is only increasing around the break site and (3) remains elevated downstream of the break site. The sequence trace downstream of the break site is decomposed into its individual sequence components. The region used for this purpose is marked as the decomposition window. All parameters in TIDE(R) have default settings but can be adjusted if necessary. The user can interactively change the alignment and decomposition windows. Choosing a different decomposition window is often a remedy to circumvent locally poor sequence traces, which should be avoided (*see* **Notes 8–10**).

For TIDER two additional quality plots are generated. In one, the aberrant signal of the reference trace compared to the control trace is plotted. This can be used to verify whether the designed mutation(s) is/are present at the expected location. In the second one, the percentage of the designed mutation(s) present in the experimental sample is plotted, representing the relative incorporation of the donor template (Fig. 2f).

3.5 Mutation Detection by Decomposition

For the detection of individual mutations with the corresponding frequencies, the TIDE and TIDER software perform a decomposition of the mixed sequence signal in the experimental sample. This composite sequence trace is a linear combination of the wild type (control) and the mutated sequences. For TIDE, the decomposition is performed on a sequence segment downstream of the break site. As a rule of thumb, the larger the decomposition window

is chosen, the more robust the estimation of mutations is (*see Note 9*). To perform the decomposition, a set of sequence trace models are generated that contain all possible indels of size $\{0..n\}$ (n is by default set to 10). The models are derived from the control trace and contain all nucleotide peak signals of the decomposition window shifted by the appropriate number of positions to the left or right. A wild-type trace (shift 0) is also added as a model. Then, using non-negative linear modeling the combination of trace models that can best explain the composite sequence trace in the experimental sample is determined (Fig. 2c) (*see Note 11*). An R^2 value is calculated as a measure of the goodness of fit (*see Notes 10 and 12*), and the statistical significance of the detection of each indel is calculated.

For TIDER the mutation detection is more complex. It is mandatory that the decomposition window in TIDER covers the location of the designed mutation(s) in the donor template (*see Notes 9 and 13*). In contrast to TIDE, the decomposition window of TIDER spans by default only 100 bp. In case only few base pair changes are introduced, the sequence with the designed mutation will be very similar to the wild-type sequence. The smaller decomposition window of TIDER emphasizes the difference between the control and reference better. Simulations of all possible insertions and deletions are generated from the control file and placed in a decomposition matrix together with the control and reference. Subsequently, decomposition of the experimental sample is performed thereby choosing the best combination of the models in the decomposition matrix. This results in an estimation of the incorporation frequency of template-directed mutation(s) and distinguishes these from the background of indels that are introduced by error-prone repair (*see Note 14*).

The reliability of TIDE and TIDER depends on the quality of the input samples (*see Note 15*). For an accurate TIDE (R) estimation it is recommended that (1) $R^2 > 0.9$ and (2) aberrant signals upstream of the break site are below 10% in the quality plot. This applies to all files: control, reference, and experimental sample. To verify the results the samples can be sequenced from the opposite strand (*see Note 13*).

3.6 Sequence Determination of the +1 Insertion (TIDE Only)

During repair of CRISPR-Cas9 a single base pair is frequently inserted at one of the DNA ends of the break [13, 17, 18]. TIDE provides an estimate of the base composition of this insertion. This may be of interest if one wishes to obtain a particular sequence variant (Fig. 2d). For longer insertions this base calling is computationally complicated and currently not implemented.

4 Notes

1. *Primer design recommendations for control and experimental sample.* Primers a, b need to cover the CRISPR target site. The length of the PCR product can vary, but there should be at least >50 bp up- and downstream of the break site for the alignment (*see* **Notes 6** and **7**) and decomposition windows respectively (*see* **Note 9**).
2. *Primer design recommendations for reference sample.* Primer c, d should carry the designed mutation(s) as present in the donor template (*see* Subheading **3.2**, **Note 3**). It is advised to include at least 10 complementary nucleotides on the 3' side of the mutation(s).
3. *Donor plasmid contamination in isolated genomic DNA.* Potentially, a donor template that was transfected into the cells could co-purify with genomic DNA and be co-amplified in the PCR if it contains the primer sequences. This could result in an overestimation of the HDR events. This is generally not a problem with short ssODN donors, but with plasmid templates with long homology arms the primers a, b should be chosen outside of these homology arms. Alternatively, the donor plasmid may be cleared from the cells by a few passages of culturing.
4. *Nuclease type.* TIDE(R) is currently designed for regular Cas9. But it can be used to analyze data from another nuclease, by entering in the web tool the DNA sequence around the expected cut site. The TIDE(R) web tool assumes that the DSB is induced between nucleotides 17 and 18 of the guide RNA sequence string (Fig. **3f**). Note that if the exact breakpoint is unknown, TIDE will estimate the amount of the indels correctly, but the nucleotide composition of the +1 insertion will not be reliable. TIDER will only work when the exact cutting position is known and when the nuclease is a blunt cutter.
5. *No guide RNA match.* Sometimes a mismatch occurs in the control sequence at the location of the sgrRNA. This will stop the TIDE(R) analysis. In this case, edit the base annotation in the chromatogram file into IUPAC nucleotides of the expected control sequence (Fig. **3g**). The peak signals in the chromatogram should not be altered. Viewing and editing of chromatogram files can be performed with Snapgene or Chromas software.
6. *Alignment cannot be performed.* By default, the alignment window begins at nucleotide number 100, because the first part of the sequence read tends to be of low quality. The end of the alignment window is set automatically at 15 bp upstream

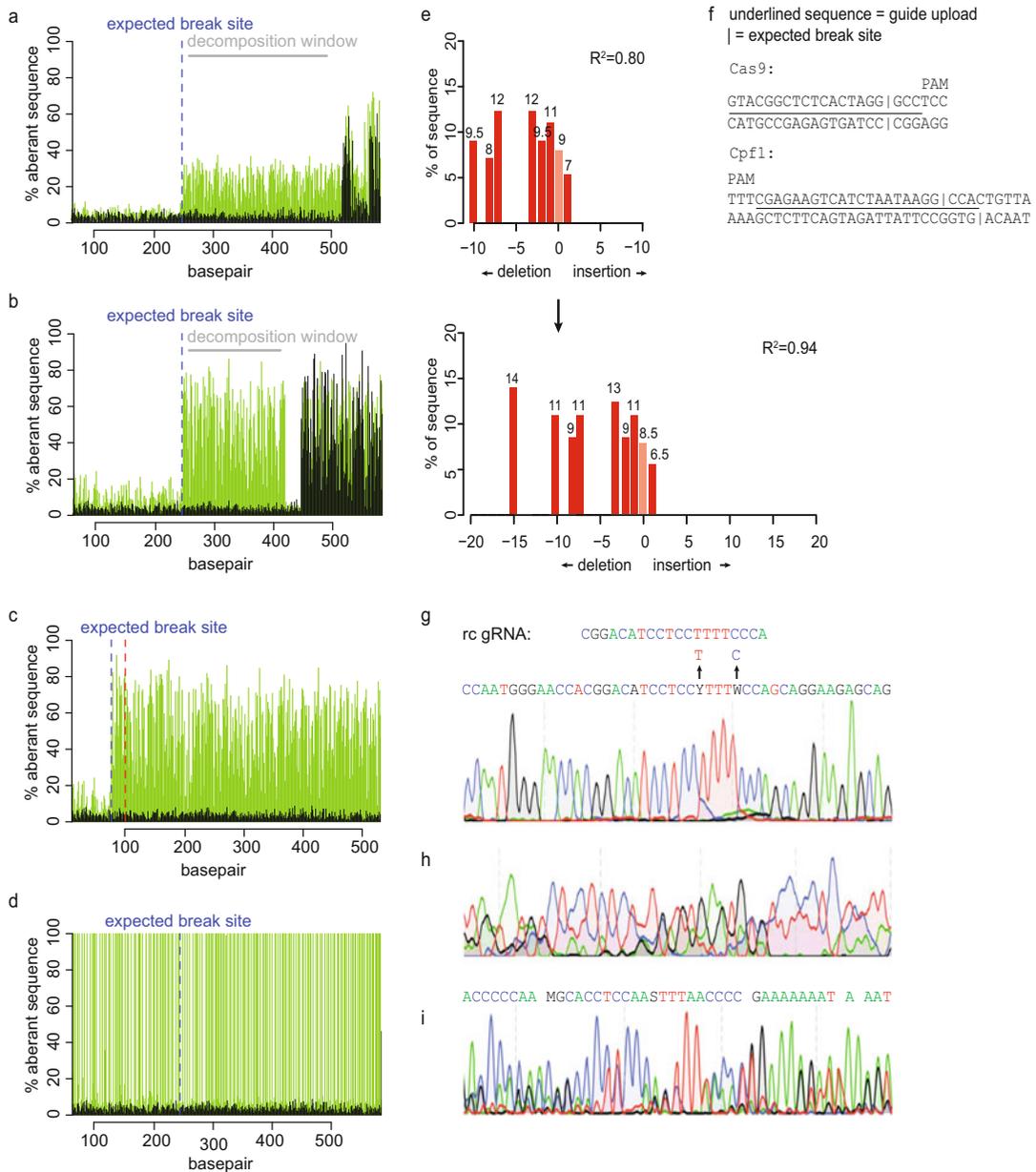


Fig. 3 Troubleshooting with TIDE and TIDER. All parameters in TIDE(R) have default settings but can be adjusted if necessary. Different settings are often a remedy to solve error messages. **(a–i)** Examples of most common error messages with the recommended setting changes. **(a, b)** Avoid the decomposition window to overlap with high aberrant signal in the control. This occurs often near the ends of the sequence traces **(a)** or in a stretch of repetitive sequences **(b)**. **(c, d)** Alignment problems can occur when the alignment window is too small (default is from 100 until 15 bp upstream break site) **(c)** or when the wrong nucleotides in the files are aligned **(d)**. The alignment window can be changed closer or further to 1 in the sequence trace. **(e)** The presence of indels larger than the default of 10 is not included in decomposition and can result in low R^2 score. Indel size can be changed. **(f)** The use of other nucleases than Cas9 in TIDE(R) works when the guide RNA string is mimicked to the 20 nt Cas9 guide RNA that cuts between nucleotide 17 and 18. **(g)** A mismatch in the IUPAC nucleotide annotation that prevents the recognition of guide RNA in the control sequence can be solved by editing the chromatogram file to the expected nucleotides. **(h, i)** Poor sequence quality will not give reliable results in TIDE(R)

of the break site. When this window is too small or when the break site is located upstream of nucleotide 100, the alignment cannot be performed correctly. Then the start of the alignment window should be set manually closer to nucleotide number 1 (Fig. 3c).

7. *Incorrect alignment.* When the beginning of the sequence trace is of poor quality, the alignment function can make a mistake. This results in a quality plot with a high aberrant sequence signal along the whole length of the sequence trace (Fig. 3d). The aberrant sequence signal should only increase around the expected cut site (blue dashed line). In case of poor alignment, the start of the alignment window needs to be adjusted until a proper alignment is achieved (default of 100).
8. *Quality plot recommendations.* In the experimental sample, around the break site a consistently elevated signal is expected, which is due to indels introduced at the break site. The starting position of this elevated signal may be used to verify that breaks were induced at the expected location. The control trace should have a low and equally distributed aberrant sequence signal along the whole trace. The reference trace in the case of TIDER should only have high scores at the positions of the altered nucleotides. Fluctuations in the control and reference signal reflect local variation in the quality of the sequence trace. Near the end of the sequencing traces the aberrant signal is often high, typically due to the lower quality of the trace toward the end (Fig. 3a). When a sequence stretch of poor local quality is present in the decomposition window the calculations of TIDE(R) are compromised. The boundaries of the decomposition window can be manually adjusted to remove the region that is of low quality; this will improve the estimations. Another area to avoid in the decomposition window is a stretch of repetitive sequences. These regions can be recognized in the quality plot as a sudden stretch without aberrant nucleotides (Fig. 3b). Such region might confound the decomposition of the sequence trace.
9. *Decomposition window recommendations.* For TIDE, the default decomposition window spans the entire sequence trace from the break site until the end of the sequence minus the size of the maximum indel. When the boundaries of the decomposition window cannot fulfill this constraint, the software will report that the boundaries are not acceptable. For example, this can occur when the break site is too close to the end of sequence trace. To address this, the decomposition window boundaries should be set further apart or a smaller indel size should be chosen. Alternatively, new primers have to be designed according to **Note 1**. For TIDER the decomposition window is by default 20 bp upstream of the break to 80 bp

downstream from the break. This smaller window compared to TIDE has more discriminatory power for subtle designed base pair changes.

10. *Goodness of fit.* R^2 is a measure for the reliability of the estimated values. For example, if the R^2 value is 0.95, it means that 95% of the variance can be explained by the model; the remainder 5% consists of random noise, very large indels, non-templated point mutations, and possibly more complex mutations. Decomposition results with a low R^2 must be interpreted with caution. A low R^2 can be caused when the settings are not optimal or when the sequence quality is not good (*see Note 15*). A low R^2 value can also arise when a sequence stretch with a poor local quality is present in the decomposition window (*see Note 8*). Furthermore, the presence of indels larger than the maximum indel size that is considered can affect the R^2 (default of 10). By default these are not modeled, which may result in a low R^2 score. The size range of indels that are modeled can be manually changed to larger number to test if this improves the fit (Fig. 3e).
11. *Allele-specific indels.* The different bars in the plot represent the insertions, deletions, and/or template-directed mutations in the cell population. These mutations are not specific of an allele. To determine allele-specific information a cell clone needs to be isolated and analyzed again by TIDE(R). A diploid cell gives a percentage of ~50% per mutation.
12. *Overall efficiency.* The overall efficiency refers to the estimated total fraction of DNA with mutations around the break site. It is calculated as $R^2 \times 100\%$ wild type.
13. *Distal designed mutations.* It has been reported that the incorporation of donor template sequence is less efficient when the designed point mutations are further away from the break site [19]. This often leads to a variation in incorporation frequently of the distal and proximal designed mutations as can be observed in the quality plots. Such a situation may confound TIDER estimates. The decomposition window can be restricted to either the proximal or the distal mutations to resolve the individual efficiencies.
14. *Natural versus designed mutations.* In general, TIDER is able to discriminate “naturally” occurring deletions and insertions from templated “designed” indels. Only in the presence of a small designed deletion (-1, -2) near the expected break site the designed mutation may be underestimated [14]. In case the designed mutation consists of an insertion larger than +1, TIDER does not consider natural insertions of the same size, because the decomposition becomes less robust. This is generally acceptable, because natural insertions larger than +1 are rarely observed [13, 17].

15. *Poor sequence quality.* When the sequence has poor quality overall, TIDE(R) will yield poor results with a low R^2 value (see **Note 10**) since too much noise is present in the data. The quality plot will show an overall high aberrant sequence signal in the control (the reference) and the experimental sample, before and after the break site (see **Note 8**). It is recommended to check the chromatograms of the samples (Fig. 3h) for poor sequencing quality. If so, these samples cannot be analyzed reliably by TIDE(R). Note that sometimes the peak signals in the chromatogram appear normal, but the file can contain wrongly unannotated or additional annotated nucleotides (Fig. 3i). TIDE(R) gives a warning when the spacing between the nucleotides in the chromatogram of the sequence trace is not consistent, which is often an indication for wrongly unannotated or additional annotated nucleotides. In case of this warning, the chromatograms should be carefully investigated (use Snapgene or Chromas software).

Acknowledgments

We thank Marcel de Haas, Stefano Manzo, and Ruben Schep for critical reading of the manuscript. This work was supported by the Netherlands Organization for Scientific Research ZonMW-TOP grant 91211061, and European Research Council Advanced Grant 694466.

Competing Interests: EKB and BvS declare competing financial interests: As inventors of TIDE and TIDER software, they receive licensing payments under their employer's rewards-to-inventors scheme.

References

1. Brinkman EK, Chen T, de Haas M, Holland HA, Akhtar W, van Steensel B (2018) Kinetics and fidelity of the repair of Cas9-induced double-strand DNA breaks. *Mol Cell* 70:801–813
2. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA et al (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823
3. 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* 337:816–821
4. Mali P, Esvelt KM, Church GM (2013) Cas9 as a versatile tool for engineering biology. *Nat Methods* 10:957–963
5. Yoshimi K, Kunihiro Y, Kaneko T, Nagahora H, Voigt B, Mashimo T (2016) ssODN-mediated knock-in with CRISPR-Cas for large genomic regions in zygotes. *Nat Commun* 7:10431
6. Inui M, Miyado M, Igarashi M, Tamano M, Kubo A, Yamashita S, Asahara H, Fukami M, Takada S (2014) Rapid generation of mouse models with defined point mutations by the CRISPR/Cas9 system. *Sci Rep* 4:5396
7. Storici F, Snipe JR, Chan GK, Gordenin DA, Resnick MA (2006) Conservative repair of a chromosomal double-strand break by single-

- strand DNA through two steps of annealing. *Mol Cell Biol* 26:7645–7657
8. Wang Z, Zhou ZJ, Liu DP, Huang JD (2008) Double-stranded break can be repaired by single-stranded oligonucleotides via the ATM/ATR pathway in mammalian cells. *Oligonucleotides* 18:21–32
 9. Ma H, Marti-Gutierrez N, Park SW, Wu J, Lee Y, Suzuki K, Koski A, Ji D, Hayama T, Ahmed R et al (2017) Correction of a pathogenic gene mutation in human embryos. *Nature* 548(7668):413–419
 10. Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE (2016) Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol* 34:339–344
 11. Horlbeck MA, Witkowski LB, Guglielmi B, Replogle JM, Gilbert LA, Villalta JE, Torigoe SE, Tjian R, Weissman JS (2016) Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *Elife* 5:e12677
 12. Yang L, Guell M, Byrne S, Yang JL, De Los Angeles A, Mali P, Aach J, Kim-Kiselak C, Briggs AW, Rios X et al (2013) Optimization of scarless human stem cell genome editing. *Nucleic Acids Res* 41:9049–9061
 13. Brinkman EK, Chen T, Amendola M, van Steensel B (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42:e168
 14. Brinkman EK, Kousholt AN, Harmsen T, Leemans C, Chen T, Jonkers J, van Steensel B (2018) Easy quantification of template-directed CRISPR/Cas9 editing. *Nucleic Acids Res* 46(10):e58
 15. Green MR, Sambrook J (2012) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
 16. Kunkel TA (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc Natl Acad Sci U S A* 82:488–492
 17. van Overbeek M, Capurso D, Carter MM, Thompson MS, Frias E, Russ C, Reece-Hoyes JS, Nye C, Gradia S, Vidal B et al (2016) DNA repair profiling reveals nonrandom outcomes at Cas9-mediated breaks. *Mol Cell* 63:633–646
 18. Dorsett Y, Zhou Y, Tubbs AT, Chen BR, Purman C, Lee BS, George R, Bredemeyer AL, Zhao JY, Sodergren E et al (2014) HCoDES reveals chromosomal DNA end structures with single-nucleotide resolution. *Mol Cell* 56:808–818
 19. Beumer KJ, Trautman JK, Mukherjee K, Carroll D (2013) Donor DNA utilization during gene targeting with zinc-finger nucleases. *G3* 3:657–664

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Fast and Quantitative Identification of Ex Vivo Precise Genome Targeting-Induced Indel Events by IDAA

Saskia König, Zhang Yang, Hans Heugh Wandall, Claudio Mussolino, and Eric Paul Bennett

Abstract

Recent developments in gene targeting methodologies such as ZFNs, TALENs, and CRISPR/Cas9 have revolutionized approaches for gene modifications in cells, tissues, and whole animals showing great promise for translational applications. With regard to CRISPR/Cas9, a variety of repurposed systems have been developed to achieve gene knock-out, base editing, targeted knock-in, gene activation/repression, epigenetic modulation, and locus-specific labeling. A functional commonality of all CRISPR/Cas9 applications is the gRNA-dependent targeting specificity of the Cas9/gRNA complex that, for gene knock-out (KO) purposes, has been shown to dictate the indel formation potential. Therefore, the objective of a CRISPR/Cas9 KO set up is to identify gRNA designs that enable maximum out-of-frame insertion and/or deletion (indel) formation and thus, gRNA design becomes a proxy for optimal functionality of CRISPR/Cas9 KO and repurposed systems. To this end, validation of gRNA functionality depends on efficient, accurate, and sensitive identification of indels induced by a given gRNA design. For in vitro indel profiling the most commonly used methods are based on amplicon size discrimination or sequencing. Indel detection by amplicon analysis (IDAA™) is an alternative sensitive, fast, and cost-efficient approach ideally suited for profiling of indels induced by Cas9/gRNA with similar sensitivity, specificity, and resolution, down to single base discrimination, as the preferred next-generation sequencing-based indel profiling methodologies. Here we provide a protocol that is based on complexed Cas9/gRNA RNPs delivered to primary peripheral blood mononuclear cells (PBMCs) isolated from healthy individuals followed by quantitative IDAA indel profiling. Importantly, the protocol described benefits from a short “sample-to-data” turnaround time of less than 5 h. Thus, this protocol describes a methodology that provides a suitable and effective solution to validate and quantify the extent of ex vivo CRISPR/Cas9 targeting in primary cells.

Key words Indel detection by amplicon analysis (IDAA™), NGS, Ex vivo precise genome targeting, PBMCs, Indel “finger print”, Primary cells, CD34+, CRISPR/Cas9, RNP, Synthetic gRNA, ProfileIt™

1 Introduction

Precise genome targeting events mediated by designer nucleases (DNs) such as CRISPR/Cas9, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), or

meganucleases (MNs) [1] induce formation of double-strand DNA breaks (DSBs) that are mostly repaired by the cellular error-prone non-homologous end joining (NHEJ) pathway [2] leaving small insertions or deletions (indels) or a combination of both at the repaired DSB site. The simplicity and efficiency of gene editing methodologies make them ideally suited to be employed for inactivation of gene function in any cellular context. By definition, indels are less than 1 kb in length while larger sized indels are referred to as copy number variants [3]. In the context of DN-induced indels, their sizes are commonly smaller than 30 bp. However, indel formation is a complex process guided not only by the nature of the induced DSB (blunt or staggered) but also by the primary targeted sequence that may contain microhomology regions driving the repair process [4]. Notwithstanding, the use of DN based on the widely used *S. pyogenes* Cas9 (in the following referred to as Cas9) generally results in predominant single base indels [5–9], which imposes important requirements to the indel detection methodology used [10, 11]. The ability to detect indels is relatively new, and “gold standards” for indel detection are not yet established [3]. Up to now, indel detection has primarily been accomplished by polymerase chain reaction (PCR) followed by direct size discrimination by electrophoretic or sequence-based methodologies [10]. Regarding the latter, indel detection by next-generation sequencing (NGS) approaches has been widely adopted. However, the choice of sequencing platform can have a profound effect on indel detection accuracy and heavily depends on technically challenging bioinformatic alignment analysis of the massive amount of output sequence data generated [3, 12, 13]. In this context, a critical issue is how to evaluate the sensitivity and specificity of indel detection by NGS. While each of the different NGS platforms, including the various bioinformatic algorithms used, has strengths and limitations, a common issue is the lack of “gold standards” for detection and annotation. In short of such reference standards, it is difficult to assess the indel detection sensitivity and specificity of a given NGS assays. Manual review of NGS indel calls is laborious and time consuming. With increasing application of NGS methods for evaluation of editing events in the context of precise genome targeting, rigorous and standardized methods for indel detection that do not require extensive manual review will be required. To this end the recently described TIDE [14] and ICE-D methodologies [15] have provided an alternative sequencing solution to NGS-based indel detection. These methods are however based on Sanger sequencing which limits the indel detection sensitivity to 5–10% [16, 17] and their sensitivity and resolution for complex cell pool indel profiling still remains to be determined. IDAA methodology [11] provides a fast and accurate alternative to sequence-based indel detection (Fig. 1). IDAA depends neither on laborious amplicon library preparation followed by NGS nor on

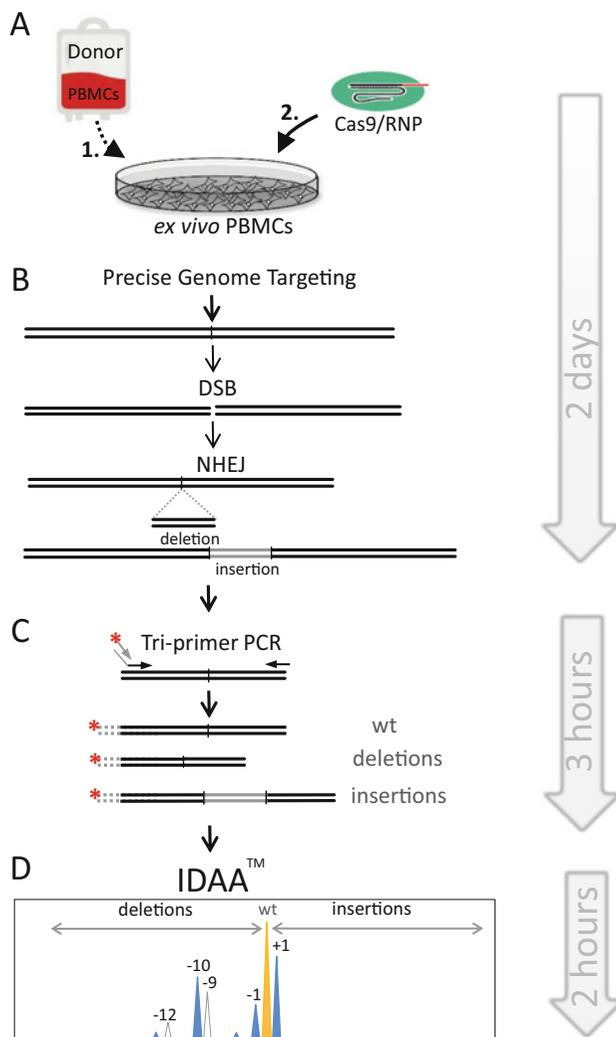


Fig. 1 Schematic outline of the ex vivo IDAA indel profiling workflow. (a) Human healthy donor PBMCs are isolated, cultured *ex vivo* (1) followed by CRISPR/Cas9 RNP delivery (2). (b) Designer nucleases precisely targeted to a defined genomic locus introduce DNA DSBs that are repaired by cellular repair mechanisms and result in indel formation at the targeted site. (c) Tri-primer amplification across the targeted site by use of the universal fluorophore labeled FAMfor primer enables uniform labeling of amplicons. (d) Fluorophore labeled amplicons are analyzed by capillary electrophoretic fragment analysis, followed by quantitative indel identification by ProfileIt™. The illustration depicts the indel “finger print,” with WT/unmodified amplicon peak shown in yellow and indels out-of-frame or in-frame shown in blue and white respectively

extensive bioinformatic data deconvolution, but is based on absolute unbiased and direct identification of fragments amplified from individual indel events in a given sample. The simplicity of IDAA enables cost-effective indel analysis, superior throughput, and

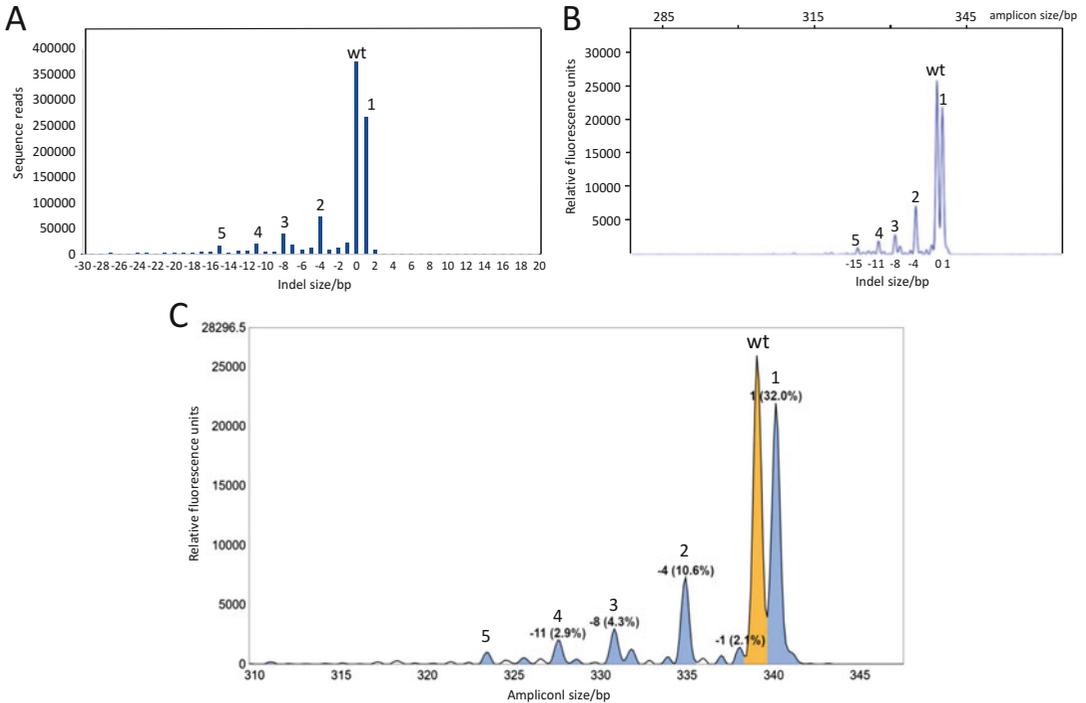


Fig. 2 Comparative profiles for indel detection of a sample by NGS and IDAA™/ProfileIt™ quantitative analysis. **(a)** MiSeq NGS profile of amplicons derived from a cell lysate 2 days post *COSMC* Cas9 nucleofection as previously described [11]. The five most dominant indel events are labeled 1–5 with indication of unmodified wild type peak (WT). **(b)** IDAA profile derived from analysis of the same cell lysate as displayed in **(a)**, analyzed by GeneMapper software (ThermoScientific, USA), and with labeling of the same dominant indel events. **(c)** Depicts the indel “finger print” of the same IDAA sample as shown in panel A and B, but analyzed by ProfileIt™, with quantification of the five most dominant indels shown as % of total indels including unmodified/WT product (shown in yellow) and with all out-of-frame indels shown in blue. The quantitative results for the profile shown are presented in Table 1

importantly with quantitative detection and sensitivity similar to NGS [8, 10] (Fig. 2 and Table 1). The core of IDAA is based on two simple principles: a tri-primer amplicon labeling step that allows for uniform fluorophore amplicon labeling (Fig. 1) and standard capillary electrophoretic analysis of labeled amplicons using commonly available standard sequenator instrumentation. The complete turnaround time for IDAA is 4–5 h from targeted cell sample preparation to quantitative indel profiling. IDAAs high-throughput amenability has been described for in vitro cell line engineering and Cas9-induced germline indel transmission rates estimation in whole animals [7–9, 18]. This protocol focuses on the use of IDAA for indel profiling of primary human cells manipulated ex vivo, with particular emphasis on its applicability as an indel profiling step in ex vivo gene editing applications [19] (Fig. 3). In addition, by comparing the indel profiles of the same gRNA in a

Table 1
ProfileIt quantification of indels shown in Figs. 2c and 3

Figure	WT/bp	Total/%	Out-of-frame/%	First indel/%	Second indel/%	Third indel/%	Fourth indel/%	Fifth indel/%
2c	339	62.2	59.6	+1/32.0	-4/10.6	-8/4.3	-11/2.9	-1/2.1
3a (lower)	431	38.3	34.6	+1/29.3	-13/2.5	-279/2.1	-21/1.6	-16/1.4
3b (upper)	431	76.5	67.2	+1/44.5	-1/5.0	+2/4.3	-279/3.9	-16/3.8
3b (lower)	431	66.7	54.5	+1/25.1	-1/7.6	-16/6.7	-32/6.1	-21/5.5
3c (upper)	431	66.3	34.5	-15/13.1	-5/6.9	1/5.2	-1/5.0	-12/4.2
3c (lower)	431	89.0	42.3	-5/27.9	-15/22.0	-3/7.0	-6/5.4	-12/5.1

NB! 5 most frequent indel events are given as % relative to total indels, including unmodified allele

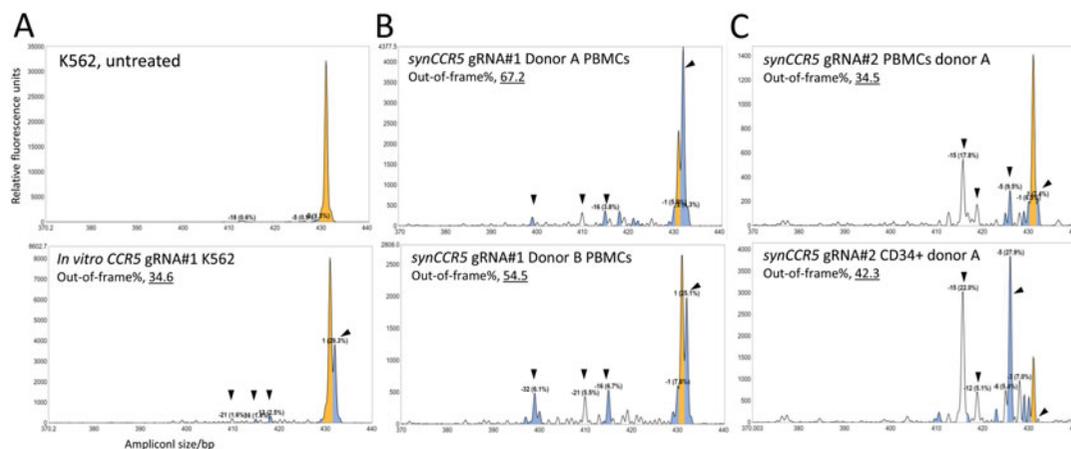


Fig. 3 IDAA™/ProfileIt™ analysis of ex vivo delivered Cas9 RNP to human cells. An in vitro transcribed gRNA (gRNA #1) and two different synthetic gRNA designs (*synCCR5* gRNA #1 and #2), targeting the human *CCR5* between the corresponding amino acids 80 and 120, were complexed with Cas9 and delivered to K562 cell line, isolated donor PBMCs, or purified CD34+ cells respectively. (a) Depicts the indel “finger print” from in vitro transcribed gRNA#1, as described in the protocol, complexed with Cas9 delivered to K562 cells (lower panel). Out-of-frame indel formation monitored three days post nucleofection as compared to untreated K562 cells (upper panel) of 34.6% was achieved. (b) Depicts the indel “finger print” derived from RNP complexed with *synCCR5* gRNA#1 delivered to PBMCs isolated from different human healthy donors (A or B). Similar indel profiles (“indel finger prints”) are observed three days post nucleofection in PBMCs from two different donors, marked by black arrow heads. Relative difference in WT allele presence is attributed to differences in nucleofection efficiency. (c) Depicts the indel “finger print” from RNP complexed with *synCCR5* gRNA#2 delivered either to PBMCs or CD34+ cells isolated from the same human healthy donor A. Similar indel “finger prints” are observed 3 days post nucleofection in PBMCs and CD34+ cells. Relative difference in WT allele presence is attributed to differences in nucleofection efficiency. In the three panels, the quantification of the five most dominant indels is shown as % of total indels including unmodified/WT product (shown in yellow) and out-of-frame indel % is indicated and the respective indels shown in blue. Indel quantification for all profiles shown can be found in Table 1

human cell line and primary cells, we show that the indel profile serves as an “finger print” that is strictly dependent on gRNA design and on the target sequence.

2 Materials

2.1 Primary Human PBMC Isolation and Culture

1. Leukocyte reduction system (LRS) chamber containing the blood sample for further processing.
2. 50 mL Falcon Conical Centrifuge Tubes.
3. PBS.
4. T75 Tissue Culture Flasks.
5. Biocoll Separating Solution (Biochrom, Berlin, Germany).
6. RPMI 1640 with GlutaMAX (Thermo Fisher Scientific, Waltham, USA) + 10% FCS (PAA, Pasching, Austria).
7. Human interleukin 2 improved sequence (IL-2 IS), premium grade (Miltenyi Biotec, Bergisch Gladbach, Germany).
8. Tissue culture plates.
9. NucleoCounter NC250 (ChemoMetec, Allerød, Denmark) or similar device.
10. Humidified incubator at 37 °C and 5% CO₂.

2.2 Thawing and Activation of Primary Human PBMCs

1. Primary human peripheral blood mononuclear cells (PBMCs; *see Note 1*).
2. RPMI 1640 with GlutaMAX (Thermo Fisher Scientific, Waltham, USA) + 10% FCS.
3. Human T cell activation kit (Miltenyi Biotec, Bergisch Gladbach, Germany).
4. Human interleukin 2 improved sequence (IL-2 IS), premium grade (Miltenyi Biotec, Bergisch Gladbach, Germany).
5. Anti-CD25 antibody, human (Miltenyi Biotec, Bergisch Gladbach, Germany).
6. Tissue culture plates.
7. NucleoCounter NC250 (ChemoMetec, Allerød, Denmark) or similar device.
8. DynaMag15 (Invitrogen, Carlsbad, USA) or similar.
9. Humidified incubator at 37 °C and 5% CO₂.

2.3 PCR-Template Preparation for In Vitro Transcription of the gRNA Targeting the Human CCR5 Gene

1. Eukaryotic expression vectors containing a U6-driven gRNA targeting the gene of interest (e.g., gRNA#1 or gRNA#2 used in this study, targeting the human *CCR5* coding sequence between the corresponding amino acids 80 and 120). As an example use the MLM3636 Addgene, cat. no. 43860 for cloning the desired gRNA.

2. Forward and reverse oligonucleotides (100 μ M; *see* **Note 2**).
3. Q5[®] Hot Start High-Fidelity DNA Polymerase and 5 \times Q5 Reaction Buffer (New England Biolabs, Ipswich, USA).
4. 10 mM dNTP Solution Mix (New England Biolabs, Ipswich, USA).
5. QIAquick PCR Purification Kit (Qiagen, Hilden, Germany).
6. RNaseZAP (Sigma, St. Louis, USA).
7. Nuclease-free water (Ambion, Austin, USA).
8. 10 mg/mL ethidium bromide (Roth, Karlsruhe, Germany).
9. 6 \times Orange Loading dye: 100 mg Orange G (Roth, Karlsruhe, Germany), 10 mL Glycerol, 40 mL dH₂O. Aliquot and store at -20 °C.
10. Agarose.
11. TAE gel running buffer: 40 mM Tris-HCl pH 8.0, 20 mM acetic acid, 1 mM EDTA.
12. RNase free 1.5 mL safe-lock tubes (Eppendorf, Hamburg, Germany).
13. PCR Cycler.
14. Nanodrop1000 (PEQLAB, Erlangen, Germany) or other DNA quantification device.
15. Agarose gel electrophoresis apparatus.
16. Fusion FX (Vilber, Eberhardzell, Germany) or other UV imager.
17. Benchtop centrifuge.

2.4 In Vitro Transcription and Purification

1. HiScribe T7 High Yield RNA Synthesis Kit (New England Biolabs, Ipswich, USA).
2. DNase I 2000 units/mL (Rnase free; New England Biolabs, Ipswich, USA).
3. MEGAclean Transcription Clean-Up Kit (Ambion, Austin, USA).
4. DEPC-treated water.
5. 10 \times TBE buffer: 890 mM Tris, 890 mM boric acid, 20 mM EDTA (*see* **Note 3**).
6. Urea.
7. Acrylamide 30% (29:1) (Biorad, Hercules, USA).
8. TEMED (Carl Roth, Karlsruhe, Germany).
9. APS (Carl Roth, Karlsruhe, Germany).
10. Low range ssRNA ladder (New England Biolabs, Ipswich, USA).

11. 2× RNA loading dye (New England Biolabs, Ipswich, USA).
12. Mini-PROTEAN Tetra Vertical Electrophoresis Cell (Biorad, Hercules, USA).

2.5 CRISPR-Cas9 Components Delivery into PBMCs

1. Recombinant *Streptococcus pyogenes* Cas9 (spCas9) protein, high concentration (PNA Bio, Thousand Oaks, USA).
2. In vitro transcribed gRNA from Subheading 2.4 or in alternative synthetic gRNA chemically modified to include 2'-O-methyl analogs and 3' phosphorothioate internucleotide linkages at the first three 5' and 3' terminal RNA residues (e.g., Synthego CRISPREvolution sgRNA EZ Kit).
3. P3 Primary Cell 4D Nucleofector × kit (Lonza, Basel, Switzerland).
4. 4D Nucleofector X device (Lonza, Basel, Switzerland).
5. Flow cytometry device such as the Accuri C6 (BD Biosciences, Allschwil, Switzerland).

2.6 Cell Lysis and DNA Extraction

1. CoboExtract DNA extraction solution (Cobo Technologies).

2.7 Tri-primer IDAA PCR

1. FAMFOR, Fwd and Rev IDAA primers can be ordered in a kit format from TAG Copenhagen (<http://www.TAGC.com>) or custom-synthesized on a 10-nmol scale, desalted, or reverse-phase column purified by standard oligo service providers (*see Note 4*).
2. Design IDAA Fwd and Rev primers to amplify a 200–550 bp product spanning the nuclease cut site using Primer3 or similar software. Add the following 5' extension to the IDAA Fwd primer: 5'-AGCTGACCGGCAGCAAATTG-3' (*see Note 5*).
3. PCR plates, 96 wells.
4. Axygen 8-strip PCR tubes (Fischer Scientific).
5. Thermocycler with programmable temperature stepping functionality, 96 wells, for tri-primer fluorophore labeling of amplicons a Veriti Thermocycler is recommended or Eppendorf Mastercycler with vapo.protect.
6. Benchtop microcentrifuge.
7. Eppendorf Thermomixer R, dry block heating and cooling shaker.
8. Agarose gel electrophoresis system.
9. Novex XCell SureLock Mini-Cell.
10. Blue-light transilluminator and orange filter goggles.
11. TEMPase Hot Start DNA Polymerase with 10× ammonium buffer and MgCl₂ (Ampliqon).

12. dNTP 100 mM (Ampliqon).
13. UltraPure DNase/RNase-free dH₂O.
14. TAE buffer, 10×.
15. UltraPure agarose.
16. MassRuler Low Range DNA Ladder.
17. TrackIt CyanOrange loading buffer (Life Technologies).
18. SYBR Safe DNA Gel Stain (Thermo Fisher Scientific).

2.8 Amplicon Analysis by Capillary Electrophoresis

1. A capillary electrophoresis instrument that supports denaturing capillary electrophoresis (*see Note 6*).
2. Hi-Di Formamide (Thermo Fisher Scientific).
3. Prepare aliquots and store them at $-20\text{ }^{\circ}\text{C}$ for up to 3 months (*see Note 7*).
4. GS500LIZ or GS600LIZ size standard (Thermo Fisher Scientific).

2.9 ProfileIT Indel Profiling and Quantification

1. ProfileIt™ software will work on any modern PC or mobile device supporting one of the following web browsers: Google Chrome 32+, Firefox 45+, Safari 10+, Microsoft Edge 12+.
2. License to Viking ProfileIt™ indel profiling software analysis program (<https://viking.sdu.dk/pages/software/profileit/>).

3 Methods

3.1 Primary Human PBMC Isolation and Culture

Estimated time required: 5 h.

1. Transfer the blood sample from the leukocyte reduction system (LRS) chamber to a 50 mL Falcon tube and note the volume (*see Note 8*).
2. Wash the chamber with 30 mL PBS (PAN Biotech, Aidenbach, Germany) by adding it directly to the cells in the 50 mL Falcon and transfer the solution into a T75 culture flask.
3. Add 29 mL PBS for each 6 mL of cell suspension (*see Note 9*).
4. Prepare 50 mL Falcon tubes with 15 mL of Biocoll (Biochrom, Berlin, Germany) for the density gradient separation (*see Note 10*).
5. Carefully overlay Biocoll with 35 mL cell suspension and centrifuge at $400 \times g$ for 30 min at $20\text{ }^{\circ}\text{C}$ (*see Note 11*).
6. Discard the supernatant and transfer the cellular layer into a new 50 mL Falcon tube.
7. For washing, fill up to 50 mL with PBS and centrifuge at $300 \times g$ for 10 min at $20\text{ }^{\circ}\text{C}$.

8. Discard the supernatant and repeat the washing step once.
9. Resuspend the pellet of freshly isolated human PBMC in 40 mL PBS and determine the cells concentration using a Nucleocounter or a similar device.
10. Use the cells immediately for further experiments or freeze them down in FCS + 10% DMSO and store cells in liquid nitrogen in appropriate aliquots.
11. Prepare a suitable amount of culture medium (RPMI 1640 with GlutaMAX, 10% FCS, 50 U/mL IL-2).
12. Culture PBMCs in the freshly prepared medium from Subheading 3.1, step 11 and keep the cells at a concentration not lower than 2×10^6 cells/mL and 1.3×10^6 cells/cm² at 37 °C with 5% CO₂.

3.2 Thawing and Activation of Primary Human PBMCs

Estimated time required: 1 h.

1. To metabolically activate PBMCs prepare CD2, CD3, and CD28 coupled magnetic beads according to the manufacturer's instructions and store at 4 °C. The final beads concentration is 1×10^8 beads/mL (*see Notes 12 and 13*).
2. If using freshly isolated PBMCs, proceed directly to **step 4**. If using frozen PBMCs, before thawing them, prepare 9 mL prewarmed medium indicated in Subheading 3.1, **step 11** and thaw the cells in a 37 °C water bath.
3. Immediately add the cell solution into the prewarmed medium. Subsequently, centrifuge for 5 min at $300 \times g$ and resuspend the cells in an appropriate volume of medium, depending on the estimated cell amount (*see Note 14*).
4. Determine the cells concentration using a cell counter or equivalent device.
5. Take the required amount of beads based on the number of PBMCs to be activated (cells:beads ratio of 2:1), and wash the beads by gently resuspending them in 100–200 µL of medium.
6. Centrifuge at $300 \times g$ for 5 min, discard the supernatant, and resuspend the beads in a volume of medium corresponding to the initial volume of the beads suspension. Add the beads to the cells and seed them in the cell concentration mentioned above (Subheading 3.1, **step 12**).
7. Incubate the cells at 37 °C with 5% CO₂ for 3 days.
8. Monitor PBMCs activation via flow cytometry, using a device such as Accuri C6, upon staining the cells with an anti-human CD25 antibody (diluted 1:50) following manufacturer's instructions (*see Note 15*).

9. After three days, remove beads using a commercially available magnet system as the DynaMag15 (Invitrogen, *see Note 16*) following manufacturer's instructions. Determine the cell concentration and culture as suggested in Subheading 3.1, **step 12** to recover the cells before nucleofection.

3.3 PCR-Template Preparation for In Vitro Transcription of the gRNA Targeting the Human CCR5 Gene

Estimated time required: 1 h.

1. Mix 1 μL of 10 μM forward and reverse primers (*see Note 2*) with 1 μL of 10 mM dNTPs, 10 μL 5 \times Q5 Reaction Buffer, 0.5 μL Q5 Hot Start High-Fidelity DNA Polymerase, 1 ng of plasmid template, and fill up to 50 μL with nuclease-free water. Cycling conditions: 30 s—98 $^{\circ}\text{C}$, 30 \times (5 s—98 $^{\circ}\text{C}$, 30 s—69 $^{\circ}\text{C}$, 20 s—72 $^{\circ}\text{C}$), 2 min—72 $^{\circ}\text{C}$ (*see Notes 17 and 18*).
2. Check the PCR reaction on a 1% agarose gel in 1 \times TAE buffer with ethidium bromide by mixing 3 μL of PCR reaction, 9 μL water, and 2 μL 6 \times orange loading dye. Let the gel run for 30 min at 100 V and visualize with a UV light device.
3. If the reaction is successful, purify the rest of the PCR reaction using QIAquick PCR Purification Kit and follow the manufacturer's instructions (*see Note 19*).
4. Determine the DNA concentration using a Nanodrop or an equivalent device.
5. The purified PCR amplicons can be used immediately for the in vitro transcription or stored at -20°C (*see Note 21*).

3.4 In Vitro Transcription and Purification

Estimated time required: 1 day.

1. Perform the in vitro transcription using the HiScribe T7 High Yield RNA Synthesis Kit following the manufacturer's protocol for short transcripts (*see Notes 18 and 20*). Briefly, mix 1.5 μL Reaction Buffer (10 \times), 1.5 μL ATP (100 mM), 1.5 μL UTP (100 mM), 1.5 μL CTP (100 mM), 1.5 μL GTP (100 mM), 1.5 μL T7 RNA Polymerase Mix, 1 μg of PCR template from Subheading 3.3, **step 5** and fill up to 20 μL with nuclease-free water. Incubate for at least 16 h overnight at 37 $^{\circ}\text{C}$.
2. To remove the DNA amplicons add 70 μL H₂O, 10 μL DNase Buffer I, and 2 μL DNase I. Incubate for 15 min at 37 $^{\circ}\text{C}$.
3. For purifying the gRNA use the MEGAclean Transcription Clean-Up Kit and follow the manufacturer's protocol (*see Notes 22–24*).
4. Measure the concentration and check the quality of RNA with a Nanodrop or equivalent device (*see Note 25*) and aliquot the gRNA in suitable amounts for later analysis. Store at -80°C .

5. As additional quality control the gRNAs can be visualized on a denaturing PAGE. Therefore prepare a 10% PAGE gel with 8 M Urea. For 10 mL mix 1 mL of 10× TBE, 4.8 g Urea, 3.35 mL Acrylamide 30%, 29:1 and fill up to 10 mL with ddH₂O. Mix at 37 °C until everything is dissolved.
6. Add 10 μL TEMED and 100 μL of 10% APS, pour immediately and insert a suitable comb.
7. After polymerization is complete (generally about 30 min) remove the comb and any bottom spacers from the gel. Fill the lower reservoir of the electrophoresis tank with 1× TBE buffer and place the gel into the lower tank (*see Note 26*).
8. For sample preparation, mix 500 ng of gRNA in 5 μL H₂O and 5 μL loading dye (2×). To properly evaluate the length of the RNA, use 2 μL low range ssRNA ladder mixed with 8 μL loading dye (2×). Incubate the samples and the ladder at 70 °C for 10 min, to denature the RNA and avoid the formation of secondary structures, and put immediately on ice. Wash the wells with 1× TBE buffer to remove residual urea and gel pieces and load the samples and the ladder. Run the gel at 220 V for 30–60 min to 1 h until the lower dye front reaches the end of the gel. Afterwards stain the gel with EtBr by mixing 20 mL H₂O with 7.5 μL EtBr and incubating for about 20 min. Afterwards wash 2–3 times with water and visualize the bands using a UV light device.

3.5 CRISPR-Cas9 Components Delivery into PBMCs

Estimated time required: 2 h.

1. Prepare a 48-wells plate with 500 μL/well of the PBMCs culture medium and equilibrate the plate in the incubator at 37 °C with 5% CO₂ until nucleofection.
2. Prepare a master mix of nucleofection solution for all reactions by mixing the solution and supplement (*see Note 27*).
3. Count the cells using a device such as NucleoCounter NC250 (ChemoMetec, Allerød, Denmark) or equivalent and aliquot 1×10^6 cells for each planned reaction (*see Note 28*). Place the reaction tubes into an incubator at 37 °C with 5% CO₂ until further procedure.
4. Prepare the RNPs by mixing the gRNA from Subheading 3.4 (or a synthetic gRNA as indicated in Subheading 2.5, item 2) and the spCas9 protein in a 5:1 ratio (e.g., for 1×10^6 cells, use 100 pmol gRNA and 20 pmol spCas9 protein). Mix the two components by gentle pipetting and incubate the reaction tube for 10 min at 37 °C in the incubator (*see Notes 28–32*).

5. While precomplexing the RNPs, centrifuge the aliquoted cells for 5 min at $300 \times g$ and resuspend them in 20 μL nucleofection solution.
6. Add the 20 μL of cell solution to the RNP mix and gently pipet the whole mixture into the nucleofection cuvette without creating any bubble. Perform the nucleofection using the EO-115 program (*see Note 33*).
7. Immediately after nucleofection, transfer the solution to one well of the 48-wells plate from Subheading 3.5, **step 1** and use 100 μL of this medium to recover all remaining cells from the nucleofection cuvette.
8. Incubate the cells at 37 °C with 5% CO₂ until further analysis.
9. The viability of the cells can be assessed one day after nucleofection by flow cytometric analysis and staining with a viability marker like 7AAD, by using a small aliquot of cells (e.g., 50 μL). If needed cells can be further cultured as explained in Subheading 3.1, **step 12** (*see Note 34*).
10. Indel profiling can be performed three days post nucleofection for an optimal result. Due to the enhanced dynamics in indel formation for RNP Cas9 delivery indel profiling can effectively be done 1 day post RNP delivery [9]. To this end, transfer the cells to 1.5 mL tubes and pellet them by centrifugation at $300 \times g$ for 5 min at room temperature. Proceed to IDAA profiling or store the cells at -20 °C for later analysis (*see Note 35*).

3.6 Cell Lysis and DNA Extraction

Estimated time required: 30 min.

1. Add CoboExtract solution to the cell pellets isolated in Subheading 3.5, **step 10**, to yield a lysate of 20–10,000 cells μL^{-1} (*see Note 36*). Transfer lysates to Eppendorf or PCR tubes and incubate in heat block or a thermocycler for 20 min at 65 °C followed by 10 min at 98 °C.
2. CoboExtracted cell lysates are now ready for IDAA tri-primer amplification or can be stored at -20 °C (*see Note 37*).

3.7 Tri-primer IDAA PCR

Estimated time required: 3 h.

1. Resuspend the IDAA Fwd primer to a final concentration of 2.5 μM and the IDAA Rev and FAMFOR primers to 25 μM in 10 mM Tris (pH 8.0) (*see Note 4*).
2. *Setting up the IDAA PCR.* Set up the following IDAA reaction for each sample to be analyzed. Include a nontransfected control sample.

Component	Amount (μL)	Final concentration
Ammonium PCR buffer (Ampliqon), 10× (Important Note 39)	1.25	1×
dNTP, 100 mM (25 mM each)	0.125	1 mM
IDAA Fwd primer, 2.5 μM	0.125	0.025 μM
IDAA Rev primer, 25 μM	0.125	0.25 μM
FAMFOR primer, 25 μM (<i>see Note 38</i>)	0.125	0.25 μM
DNA polymerase (Ampliqon) (<i>see Note 39, important</i>)	0.12	0.5 U
MgCl ₂ , 25 mM	0.7	2.5 mM
DMSO (<i>see Note 39, important</i>)	0.625	5%
CoboExtract cell lysate ^a from Subheading 3.6 , step 2 (<i>see Note 37</i>)	1	20–20,000 cells
ddH ₂ O	8.33	
Total	12.5	

^aAlternatively, 1–100 ng purified genomic DNA may be used

3. Perform PCR using the following “touch down”^a cycling conditions:

Cycle number	Denature	Anneal	Extend
1	95 °C, 15 min		
2–16 ^a touch down	95 °C, 30 s	^a 72 °C, 30 s	72 °C, 30 s
17–41	95 °C, 30 s	58 °C, 30 s	72 °C, 30 s
42			72 °C, 30 min

^aDecrease annealing temperature by 1 °C in each subsequent touch down cycles

4. Run 2.5 μL of the PCR product on a 1.5% (w/v) agarose gel to check amplification. There should be only one clear band and no smear (*see Notes 40 and 41*).
5. Proceed with amplicon analysis of tri-primer PCR product by capillary electrophoresis in the following section. Alternatively, tri-primer PCR products can be shipped to certified IDAA analysis custom service providers (*see Note 42*).

3.8 Amplicon Analysis by Capillary Electrophoresis

Estimated time required: 2 h.

1. Create a master mix of the size standard and formamide as detailed below and aliquot 10.3 μL into separate wells of a 96-well plate. Add 0.5 μL of each appropriately diluted IDAA

PCR from Subheading 3.7, **step 5** to each separate well of the 96 plate (*see Note 43*).

Component	Amount (μL)
Hi-Di Formamide (ABI)	10
GS500LIZ size standard (ABI)	0.3

- Denature for 5 min at 98 °C.
- Mix by vortexing and spin down briefly.
- Load plate(s) onto sequenator instrument and run fragment analysis according to manufacturer's instructions (*see Note 44*).

3.9 ProfileIT Indel Profiling and Quantification

Estimated time required: 30 min.

- After capillary electrophoretic run completion, all run files (.fsa files), containing the individual raw sample data, are uploaded into profileIt™.
- Log on to ProfileIt with username/password provided.
- Go to *Projects* in the header.
- Select *new*.
- Add "Project Name."
- Select Project and click *create*.
- Select *new project* from the list.
- Click the button *new task*.
- Select *ProfileIt*.
- Add a "Task name" and click *next*.
- Upload .fsa files.
- Any of the uploaded files may be marked as the "control sample" and "negative control" may be specified.
- Click *run task*.
- View Results.
- Select a normalization/WT peak by clicking on the desired peak in the control sample profile and select *set as normalization peak*. Normalization peak will be shown in yellow as shown in Figs. 1, 2c, and 3. Select marking of *out-of-frame INDELS*, which will be shown in blue as in Figs. 1, 2c, and 3.
- Click the button *export data*. Save or show *overall statistics* (download in XLXS format). All quantitative information is presented as shown in Table 1.
- Display individual sample profiles to see detailed info of profiles, zoom-in on peak areas, and to download images.

4 Notes

1. PBMCs can be used fresh or aliquoted as required and frozen in FCS containing 10% DMSO in liquid nitrogen.
2. In vitro transcription of the gRNA is performed with a commercially available T7 RNA polymerase-based kit. Thereby, the PCR template used for transcription has to contain a T7-driven gRNA. One can use the following general primers to generate gRNAs targeting the sequence of choice by replacing the “x” with the target site: T7_TargetX_fw: 5'ttaatacgaactcactata**GG**xxxxxxxxxxxxxxx; T7_gRNA_rv: 5'aaaagcaccgactcgggtgccac. Importantly, for the proper T7-driven transcription, the target site should begin with a 5'GG as highlighted in bold capital letter.
3. Weighing and mixing of boric acid should be performed under a fume hood.
4. The 6-FAM fluorophore is light-sensitive and should be stored in the dark at $-20\text{ }^{\circ}\text{C}$, where it will be stable for at least 6 months.
5. Test new IDAA primers in the tri-primer setup on control samples (non-nuclease treated cells) with roughly same cell concentration as the samples to be genotyped. There should be only one clear band and no smear, when analyzed by 1.5% (wt/vol) agarose gel electrophoresis.
6. The resolution of IDAATM enables the discrimination of DNA fragments down to single base pairs. This performance is met by instruments such as ABI's 310, 3100, 3130, 3500, or 3730 Genetic Analyzers (Applied Biosystems/Thermo Fisher Scientific). The examples shown in this protocol have all been generated using the ABI3500×L instrument, but similar results can be obtained using older models such as an ABI3130 instrument [11, 18].
7. Only freeze–thaw twice.
8. Generally the blood sample is collected in a leukocyte reduction system (LRS) chamber that is used to separate the white blood cells from blood products such as platelets and red blood cells.
9. Include the 30 mL PBS used for washing the chamber into the calculation.
10. To further separate PBMCs from, e.g., Erythrocytes and Granulocytes a density gradient is applied using a hydrophilic polymer.
11. Don't use the break for the centrifugation to avoid mixing of the cellular layers.
12. The beads can be stored up to 4 months at $4\text{ }^{\circ}\text{C}$.

13. Generally we recommend not to vortex the loaded beads and only mix by flicking the tube.
14. The cell concentration should not be lower than 2×10^6 cells/mL.
15. By looking at the forward and sideward scatter, activated cells increase in size and granularity. Expression levels of the CD25 marker on the cell surface can be analyzed as its higher expression correlates with a higher activation status.
16. Use a magnet suitable for the volume of the cell suspension.
17. Optimize PCR conditions for each primer pair to avoid any unspecific band.
18. The input material for in vitro transcription should be 1 μ g of PCR template with a final concentration of at least 90 ng/ μ L. Thereby, more than one PCR reaction might be needed.
19. PCR reactions from the same plasmid template can be pooled and eluted from one column to increase the concentration.
20. Change gloves regularly to avoid contaminations. All steps should be performed under RNase-free conditions by using RNase-free material (e.g., safe-lock tubes, pipette tips) and the working area and equipment (e.g., pipettes, tip boxes, racks) should be cleaned with RNaseZAP prior to start working. However, use RNaseZAP sparingly as it can lead to the degradation of the in vitro transcribed RNA. For reproducible results, make aliquots of kit components.
21. Prolonged storage or repeated freeze and thaw might reduce the quality of the PCR amplicon. This will subsequently impact on the yield and quality of the resulting gRNA.
22. Perform all centrifugation steps at maximum speed in bench-top centrifuge.
23. For elution use option 1 and incubate the filter cartridge in the elution tube at 70 °C for 10 min before centrifugation.
24. Immediately after elution, store the RNA on ice until further processing.
25. For pure RNA the expected ratios are: about 2 for 260 nm/280 nm and about 2.0–2.2 for 260/230 nm.
26. Try to remove bubbles by using a syringe and fill the upper reservoir with $1 \times$ TBE to cover the wells. Pre-run and warm the gel for at least 30 min at 200 V.
27. For each sample, mix 16.4 μ L of nucleofection solution with 3.6 μ L of supplement solution. Prepare a master mix according to the total number of samples to transfect. The nucleofection solution should be kept at room temperature prior to use.

28. Perform the procedure under RNase-free conditions and use RNase-free material.
29. The gRNAs and Cas9 protein should be stored in suitable aliquots to avoid repeated freeze and thaw cycles which might lead to a decreased activity of the nucleases.
30. All reactions should contain the same volume and the volume of the nucleofected reagents should not exceed 10% (means a maximum of 2 μL for a 20 μL reaction) of the total volume. RNase-free water can be used to fill up the volume.
31. Cells should not be longer than 5–10 min in the nucleofection solution as this may impact on cell viability. When having many samples, prepare the cells and RNPs in several steps.
32. Transfection efficiency can be monitored by adding a sample transfected with an mRNA encoding for a green fluorescent protein. Lower cell numbers might lead to increased toxicity.
33. Do not perform more than two reactions at once as RNPs might degrade and viability of the cells impaired when keeping them for too long in the nucleofection solution.
34. If the cells should be maintained and expanded for a longer time, repeat the activation every seven days and remove the beads after 3 days of activation. Depending on the application, cells can be harvested for analysis, such as qPCR, T7 Endonuclease 1 (T7E1) assay, or IDAA assay, every 7 days after reactivation.
35. Cell pellets are stable at $-20\text{ }^{\circ}\text{C}$ for extended periods of time, more than several months.
36. The cell number required for indel profiling by IDAA™ has a large dynamic range going from thousands of cells down to 10 cells or less [9]. The CoboExtract cell lysate template input volume into the tri-primer PCR reaction should be limited to 1–2 μL , since excessive content of lysate affects the performance and fidelity of the PCR reaction. However, depending on the objective, a broad range of cell numbers can be used in the tri-primer PCR reaction as described in this protocol.
37. CoboExtracted cell lysates are stable for extended periods of time, more than several months.
38. Importantly, observe the 10:1:10 molar ratio of FAMFOR primer:IDAA Fwd primer:IDAA Rev primer for optimal amplicon labeling and yields.
39. DNA polymerases, including Taq, possess a template-independent 3' nucleotide extension activity [20–22]. Most commonly a single base (preferentially adenine) is added to the 3' end of the amplicon. This activity should be maximized to completion as illustrated in Fig. 4. As recently noted [23],

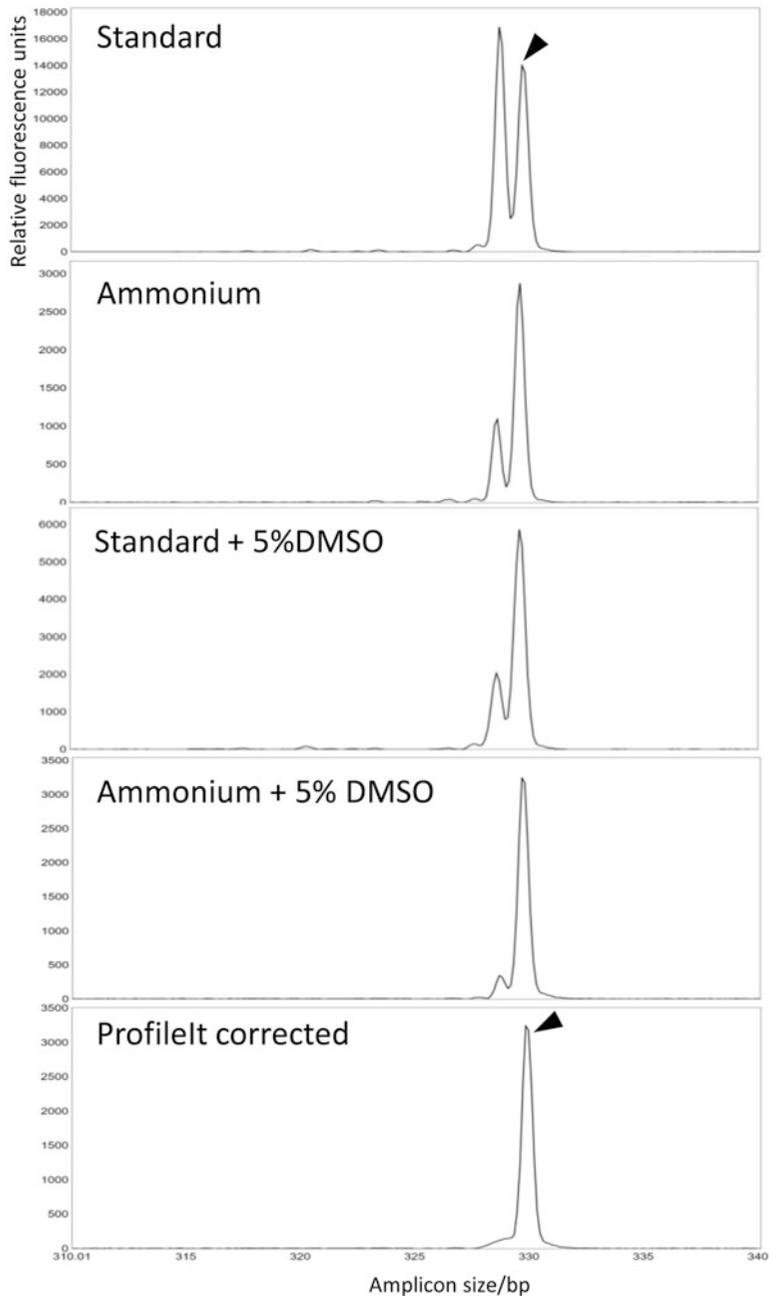


Fig. 4 Illustration of the 3'-template independent Taq DNA polymerase activity. The 3' template independent activity of several DNA polymerases is well documented [20–22]. Indel profiles shown are derived from *ST6GALNAC1* locus amplified from human genomic DNA [7]. From top to bottom the panels show how optimization of PCR buffer, from standard KCl buffer, ammonium buffer, standard KCl buffer plus 5% DMSO, ammonium buffer plus 5% DMSO, and in the latter case followed by ProfileIT™ correction can improve the 3' template independent activity from split “stutter” peak formation to a single peak. The fully 3' elongated fragment/peak is indicated with an arrow head. Importantly, careful inspection of WT peak profile and optimal 3' extension of amplicons by use of optimal 5'- “end base” design [24] is required for reliable IDAA quantification

decreased level of amplicon extension may for some targets be observed when strict ammonium buffer, addition of 5% DMSO (as previously recommended [8] and shown herein) and/or optimal reverse primer 5'-end base design, preferable G [24], are not used. Thus, inclusion of a nontargeted control sample is highly recommended, in order to assess the level of potential residual un-extended amplicon, which ultimately can be resolved using ProfileIT™ as shown in Fig. 4.

40. PCR product may be stored in the dark at 4 °C for up to 2 weeks or at –20 °C for up to 6 months.
41. The tri-primer product should be a single clear product when analyzed by standard agarose gel; however, minor amounts of unspecific product in most cases do not interfere with the results.
42. Tri-primer PCR products are stable at room temperature and may directly be shipped to custom service providers certified for IDAA analysis, such as Cobo Technologies (<https://coboscientific.com/genome-editing/indel-detection-by-amplicon-analysis-IDAA/>). The turnaround time for Quantitative IDAA analysis will depend on service provider and may range from 1 to 14 days.
43. In order not to exceed the upper threshold limit of the instrument, the amount of product generated in the IDAA tri-primer PCR, amount of diluted tri-primer PCR product may need to be diluted up to 1:150 fold. Normally PCR product may be stored in the dark at 4 °C for up to 2 weeks or at –20 °C for up to 6 months. The optimal amount of IDAA PCR to analyze may vary from 0.1 to 1 µL. If analyzing other than 200–450 bp IDAA PCR amplicons, choose the size standard accordingly, but optimal resolution is achieved for tri-primer amplicons smaller than 600 bp.
44. The 6-FAM fluorophore is unstable in formamide and IDAA PCRs should therefore be analyzed within 1–2 days after preparation for the analytical run.

Acknowledgments

We thank Vasili Korol and Ilia A. Solov'yov from the University of Southern Denmark, Department of Physics, Chemistry and Pharmacy, for development of ProfileIt™ and Camilla Andersen from Copenhagen Center for Glycomics, Department of Odontology, University of Copenhagen, for excellent technical assistance. This work was supported by the Witten/Herdecke University internal research promotion Grant No. IFF2017-12, the German Duchenne Foundation “Aktion Benni & Co.” starting grant to E.E.-S.,

the Danish National Research Foundation [DNRF107], the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 765269, and the German Federal Ministry of Education and Research (BMBF). Z.Y. received support from the Lundbeck Foundation and H.H.W. received support from ERC-2017-COG Type of action: ERC-COG; 772735; GlycoSkin.

Conflict of Interest Statement: E.P.B. declares that a patent application covering the IDAA™ method is pending, and acts as scientific advisor for Cobo Technologies Aps.

References

- Chandrasegaran S, Carroll D (2016) Origins of programmable nucleases for genome engineering. *J Mol Biol* 428:963–989
- Sfeir A, Symington LS (2015) Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway *Trends Biochem Sci* 40:701–714
- Sehn JK (2014) Insertions and deletions (Indels). In: *Clinical genomics*. Elsevier, Amsterdam, pp 129–150
- Deriano L, Roth DB (2013) Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu Rev Genet* 47:433–455
- van Overbeek M, Capurso D, Carter MM, Frias E, Russ C, Reece-Hoyes JS, Nye C, Vidal B, Zheng J, Hoffman GR et al (2016) DNA repair profiling reveals nonrandom outcomes at. *Mol Cell* 63:633–646
- Paquet D, Kwart D, Chen A, Sproul A, Jacob S, Teo S, Olsen KM, Gregg A, Noggle S, Tessier-Lavigne M (2016) Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 533:125–129
- Narimatsu Y, Joshi H, Zhang Y, Gomes C, Yen H, Lorenzetti F, Furukawa S, Schjoldager K, Hansen L, Clausen H, Bennett EP et al (2018) A validated gRNA library for CRISPR/Cas9 targeting of the human glycosyltransferase genome. *Glycobiology* 28 (5):295–305
- Lonowski LA, Narimatsu Y, Riaz A, Delay CE, Yang Z, Niola F, Duda K, Ober EA, Clausen H, Wandall HH et al (2017) Genome editing using FACS enrichment of nuclease-expressing cells and indel detection by amplicon analysis. *Nat Protoc* 12:581–603
- Kosicki M, Rajan SS, Lorenzetti FC, Wandall HH, Narimatsu Y, Metzakopian E, Bennett EP (2017) Dynamics of indel profiles induced by various CRISPR/Cas9 delivery methods. *Prog Mol Biol Transl Sci*. Elsevier Inc 152:49–67
- Bennett EP, Jacobi AM, Garrett RR, Behlke MA (2018) Detection of insertion/deletion (indel) events after genome targeting: pro's and con's of the available methods. In: Appasani K (ed) *GENOME EDITING and ENGINEERING: from Talens, ZFNs and CRISPRs to molecular surgery*. Cambridge University Press, Cambridge, pp 181–194
- Yang Z, Steentoft C, Hauge C, Hansen L, Thomsen AL, Niola F, Vester-Christensen MB, Frodin M, Clausen H, Wandall HH et al (2015) Fast and sensitive detection of indels induced by precise gene targeting. *Nucleic Acids Res* 43:e59. 1–8
- Sandmann S, De Graaf AO, Karimi M, Van Der Reijden BA, Hellström-Lindberg E, Jansen JH, Dugas M (2017) Evaluating variant calling tools for non-matched next-generation sequencing data. *Sci Rep* 7:1–12
- Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ (2012) Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol* 30:434–439
- Brinkman EK, Chen T, Amendola M, van Steensel B (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42:e168–e168
- Hsiao T, Maures T, Waite K, Yang J, Kelso R, Holden K, Stoner R (2018) Inference of CRISPR edits from sanger trace data. *bioRxiv* 1101:10. <https://doi.org/10.1101/251082>
- Jamuar S, Lam A-TN, Kircher M, D'Gama A, Wang J, Barry BJ, Zhang X, Hill RS, Partlow JN, Topcu M et al (2014) Somatic mutations in cerebral cortical malformations. *N Engl J Med* 371:733–743
- Mohamed S, Penaranda G, Gonzalez D, Camus C, Khiri H, Boulmé R, Sayada C,

- Philibert P, Olive D, Halfon P (2014) Comparison of ultra-deep versus Sanger sequencing detection of minority mutations on the HIV-1 drug resistance interpretations after virological failure. *AIDS* 28:1315–1324
18. Yang Z, Wang S, Halim A, Schulz MA, Frodin M, Rahman SH, Vester-Christensen MB, Behrens C, Kristensen C, Vakhrushev SY et al (2015) Engineered CHO cells for production of diverse, homogeneous glycoproteins. *Nat Biotechnol* 33:2014–2017
 19. Cornu TI, Mussolino C, Cathomen T (2017) Refining strategies to translate genome editing to the clinic. *Nat Med* 23:415–423
 20. Clark JM (1988) Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res* 16:9677–9686
 21. Hu G (1993) DNA polymerase-catalyzed addition of nontemplated extra nucleotides to the 3' end of a DNA fragment. *DNA Cell Biol* 12:763–770
 22. Smith JR, Carpten JD, Brownstein MJ, Ghosh S, Magnuson VL, Gilbert DA, Trent JM, Collins FS (1995) Approach to genotyping errors caused by nontemplated nucleotide addition by Taq DNA polymerase. *Genome Res* 5:312–317
 23. Sentmanat MF, Peters ST, Florian CP, Connelly JP, Pruett-Miller SM (2018) A survey of validation strategies for CRISPR-Cas9 editing. *Sci Rep* 8(1):888
 24. Magnuson VL, Ally DS, Nylund SJ, Karanjwala ZE, Rayman JB, Knapp JI, Lowe AL, Ghosh S, Collins FS (1996) Substrate nucleotide-determined non-templated addition of adenine by Taq DNA polymerase: implications for PCR-based genotyping and cloning. *Biotechniques* 21(4):700–9



Functional Evaluation of CRISPR Activity by the Dual-Fluorescent Surrogate System: C-Check

Lin Lin and Yonglun Luo

Abstract

Rapid evaluation of the CRISPR gRNA activity is an essential step of employing the technology in editing genes. Through machine learning strategy, the rule sets for in silico designing gRNAs with high activity has greatly improved. However, there are still discrepancies between different prediction rule sets, and between the predicted and actual gRNA activities. Thus, experimentally validating gRNA activity is still the gold standard in defining the best gRNAs for gene editing experiments. One such approach for experimentally selecting gRNAs with high activity is fluorescent surrogate reporter vectors. We had previously developed a dual-fluorescent surrogate system, called C-Check, which based on single-strand annealing repair of the DNA double-strand breaks introduced by CRISPR-Cas9 to generate a functional EGFP. The system offers a tool for rapid functional evaluation of CRISPR gRNA activity, as well as for enrichment of gene edited cells. In this chapter, we will give a step-by-step instruction on the design, generation, and application of the C-Check system for quantifying gRNA activities.

Key words CRISPR, Cas9, Reporter vector, Surrogate vector, C-Check, Gene editing

1 Introduction

The discovery of programmable DNA endonucleases (ZFNs, TALENs, and CRISPR-Cas9) has greatly accelerated both academic and industrial applications that involves the modification of the genetic codes [1–3]. The ease of vector design, generation and efficiency undoubtedly and rapidly made CRISPR-Cas9 as the most popular selection in the tool box of gene editing enzymes. Just through changing the typically 20 nt guide sequences of the small guide RNA (gRNA), the sole Cas9 protein can be redirected to the targeted site which is complementary to the 20 nt guide sequences and comprises a proximity DNA motif, known as protospacer adjacent motif (PAM). The PAM for Cas9 protein from *Streptococcus pyogenes* (SpCas9), the most broadly used CRISPR-Cas9 system, is 5'-NGG-3' where "N" represents any nucleotide followed by two guanine (G). The evolutionary requirement of

PAM gives CRISPR-Cas9 a unique and powerful feature of discriminating self and non-self target sites. This PAM-dependent feature contributes greatly to the specificity of CRISPR-Cas9 [4].

One common question that every user of the CRISPR-Cas9 technology encounters is that what is the activity of the selected gRNA. During the last few years, several methods have been developed to streamline the quantification of CRISPR-Cas9 activity [5–8]. When an active Cas9-gRNA component is delivered to the cells, a DNA double-strand break (DSB) will be introduced to the target site. Mammalian cells have evolved several DSB repair pathways to repair these lethal DNA lesions, such as non-homologous end joining (NHEJ), homology-directed repair (HDR), microhomology-mediated end joining (MMEJ), and Single-Strand Annealing (SSA). One way of quantifying the CRISPR-Cas9 activity is through harnessing the cellular DSB repair pathway to generate a functional surrogate vector, which contains the same locus as the intended target site in the genome. To simplify the cloning as well as quantifying CRISPR-Cas9 activity, we have generated a dual-fluorescent reporter vector, called C-Check [7]. This vector employs the SSA-mediated repair of two truncated EGFP genes, which contains 500 bp homologous sequences. The C-Check vector contains several unique features (Fig. 1): (1) Golden-Gate Assembly technology was employed to simplify the cloning procedure and increase efficiency; (2) A Lac-Z expression cassette was included in the cloning site to enable blue-and-white selection of

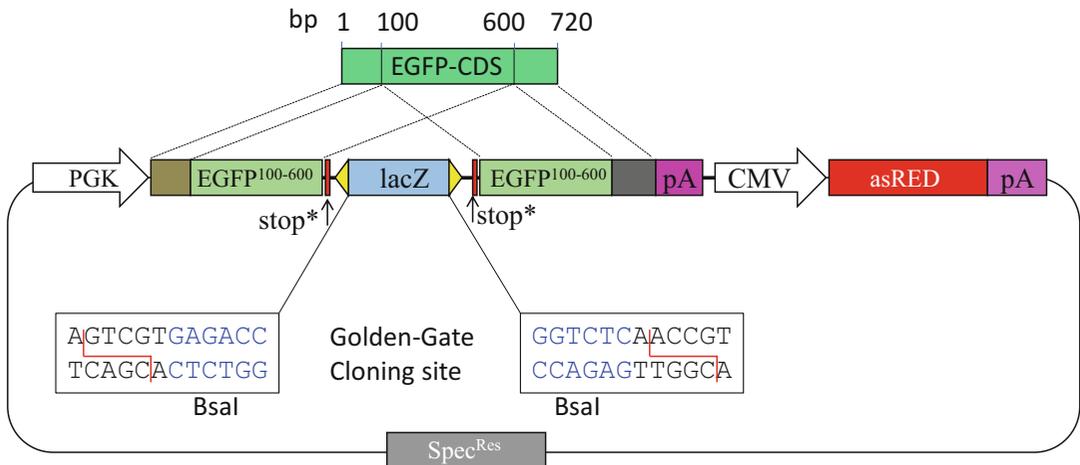


Fig. 1 Schematic illustration of the C-Check vector. The C-Check vector contains two expression cassettes. First, a consecutive AsRED expression cassette is used as a reference control for normalization. Second, a truncated EGFP expression cassette, of which the EGFP genes is splitted into two parts with 500 nt in homology. A *Bsal*-based golden-gate cloning site is used for insertion of any surrogate DNA for C-Check-based CRISPR activity assay. Upon introduction of double-stranded DNA breaks (DSB) at the surrogate DNA by, e.g., CRISPR, DSB repaired by single-strand annealing with the 500 nt homologous sequences will generate a functional EGFP expression cassette

positive clones; (3) A constitutive AsRed expression cassette was included for normalization of transfection and expression normalizations (Fig. 1). We demonstrated that the CRISPR-Cas9 activity quantified by C-Check correlates well with the functional readouts of indel frequency in cells [9–11].

Although the C-Check vector described here is for quantifying CRISPR-Cas9 activity, the system can be used to quantify the activity of any programmable DNA nucleases that can introduce DSBs. In this method chapter, we provide a detail instruction and guideline of how to design, generate, and quantify the C-Check-based CRISPR-Cas9 activity.

2 Materials

2.1 Generation of C-Check Vector

1. P119_C-Check (Addgene Plasmid #66817) (*see Note 1*).
2. Restriction enzymes: *BsaI* (*Eco31I*), *BsmBI*, *BbsI*, *BamHI*, *KpnI* (*see Note 2*).
3. T4 DNA ligase.
4. 10× T4 DNA ligase buffer.
5. LB medium.
6. LB agar plates with 50 µg/mL spectinomycin.
7. 0.5 M IPTG.
8. 100 mg/µL X-gal.
9. Chemically competent *E. coli* cells (*see Note 3*).
10. 5× KCM buffer.
11. DNA oligonucleotides.
12. C-Check screening primer (Forward): 5'-TGGTGAGCAA GGGCGAGGAGCTG
13. DreamTaq DNA polymerase.
14. 10 mM dNTP.
15. ddH₂O.
16. 10× NEBuffer2.
17. NucleoSpin Gel and PCR Clean-up kit.
18. Plasmid miniprep kit.
19. Plasmid midiprep kit.
20. 1% Agarose gel.
21. Heating block.
22. Thermal cycler.
23. Gel electrophoresis system.

2.2 C-Check Transfection

1. Human embryonic kidney 239T (HEK293T) cells.
2. D10 medium: DMEM (high glucose), 10% FBS, 1× GlutaMAX, and 1× P/S.
3. Fetal bovine serum (FBS).
4. 0.05% Trypsin-EDTA.
5. PBS without calcium and magnesium.
6. Tissue culture tested plasticwares: 6-well, 24-well.
7. X-tremeGENE 9 transfection reagent.

2.3 Evaluate C-Check by Flow Cytometry

1. Fetal bovine serum (FBS).
2. 0.05% Trypsin-EDTA (phenol red free).
3. 4% formaldehyde.
4. PBS without calcium and magnesium.
5. 96-well round well plates.
6. Flow cytometry.
7. Software for analyzing FCS data file.

3 Methods

Carry out all procedures at room temperature in a classified laboratory unless otherwise specified.

3.1 The Principle of Choosing the Optimal C-Check Design Approach

The C-Check vectors can be generated through (1) assembly of synthesized oligonucleotides (synthetic approach) or (2) PCR amplification of the targeted region (PCR approach). The principal of choosing synthetic or PCR approach depends on the number of gRNAs to be evaluated and how the gRNA target sites are distributed in the genome. Please follow these principles to choose the optimal approach:

1. Number of gRNAs (1–3), use the synthetic approach.
2. Number of gRNA (over 3), if all gRNAs target the same genomic locus within a very close region (less than 500 bp), use PCR approach.
3. Number of gRNA (over 3), if all gRNAs are located in a broad genomic region (over 500 bp), use a combination of synthetic and PCR approach.
4. Design one C-Check vector for each gene/locus.

3.2 C-Check Vector Generation—Synthetic Approach

1. Design C-Check oligonucleotides according to Table 1 (*see Note 4*).
2. Synthesize C-Check oligonucleotides, desalt purification. For oligos longer than 70 nt, use HPLC purification.

Table 1
Guide for designing C-Check oligonucleotides by synthetics

Oligo name	5' linker	Target sequences ^a	3' linker	Oligo to be order
GOI-CC-For	GTCGGAt	(target strand:NNN)	ataGGT	GTCGGAt(target strand:NNN) ataGGT
GOI-CC-Rev	CGGTACCtat	(reverse compliment strand:NNN)	aTC	CGGTACCtat(reverse compliment strand:NNN)aTC

^aEach target site should contain the 20 nt protospacer sequences and the PAM (NGG for SpCas9). Multiple sites can be placed in a tandem manner, so one C-Check vector can be used to evaluate the activity of several gRNAs

3. Dilute oligos to 100 μ M with TE buffer.
4. Prepare oligo annealing mixture in a 200 μ L PCR tube: 1 μ L each of 100 μ L C-Check oligos (sense and antisense), 2 μ L 10 \times NEBuffer 2, 16 μ L ddH₂O.
5. Mix the oligos thoroughly and quick spin using a bench top centrifuge.
6. Place the tube in a thermal cycler.
7. Denature the oligos at 95 $^{\circ}$ C for 5 mins, then decrease the temperature to 25 $^{\circ}$ C by about 2 degrees per minute using the ramping rate of delta C (*see Note 5*).
8. Save the annealed oligos at 4 $^{\circ}$ C before use. For long-term storage, save annealed oligos at -20° C.
9. Digest 1 μ g of the P119 C-Check plasmid with *BsaI* restriction enzyme.
10. Separate the digested P119 C-Check vector with 1% agarose Gel electrophoresis.
11. Purify the C-Check plasmid backbone band (6693 bp) using a gel clean-up kit.
12. Prepare an ice box and a 200 μ L PCR tube.
13. Place the PCR tube on ice, and prepare the C-Check ligation reaction containing: 50 ng of the C-Check plasmid backbone, 1 μ L of the anneal C-Check oligos, 1 μ L T4 DNA ligase, 2 μ L 10 \times T4 ligase buffer, and add ddH₂O to a total volume of 20 μ L.
14. Place the ligation mixture in a thermal cycler, set to the right temperature and time for the ligase used.
15. Once the ligation is completed, save the ligation product at -20° C until use.

3.3 C-Check Vector Generation—PCR Approach

Most steps of the PCR approach are the same as the synthetic approach. Here we highlight those steps unique for the PCR approach.

Table 2
Primer designing scheme for PCR-based generation of C-Check vectors

For <i>BsaI</i> -based cloning of C-Check vector	Notes
C-Check-F: ATAAGGTCTCAGTCGGAt---SS--- C-Check-R: ATAAGGTCTCACGGTACctat---AS---	Target sites should not contain <i>BsaI</i> recognition site.
For <i>BsmBI</i> -based cloning of C-Check vector	
C-Check-F: ATAACGTCTCAGTCGGAt---SS--- C-Check-R: ATAACGTCTCACGGTACctat---AS---	Target sites should not contain <i>BsmBI</i> recognition site.
For <i>BbsI</i> -base cloning of C-Check vector	
SSA-F: ATAAGAAGACATGTCGGAt---SS--- SSA-R: ATAAGAAGACATCGGTACctat---AS---	Target sites should not contain <i>BbsI</i> recognition site.

1. Design PCR primers using the primer designing scheme provided in Table 2.
2. Synthesize DNA oligos using any certificated supplier.
3. Perform PCR using isogenic genomic DNA template from cells which will be used for subsequent gene editing (*see Note 6*).
4. Analyze the PCR specificity with a 1.5% agarose gel electrophoresis. The PCR should give a unique and intense band of expected size. If not, pre-optimization of the PCR should be carried out.
5. Purify the PCR product using a PCR clean-up kit.
6. Digest 200–500 ng of PCR product using the corresponding restriction enzyme (*see Table 2*).
7. Purified digested PCR product using a 1.5% agarose gel.
8. Ligate the PCR product to the purified P119 C-Check backbone from Subheading 3.2 and carry out the ligation steps accordingly.

3.4 Transformation and PCR Screening of C-Check Vector

1. Transform chemically competent *E. coli* cells using 1–2 μL of the C-Check ligation product.
2. Spread transformed cells on a LB agar plate containing 50 $\mu\text{g}/\text{mL}$ spectinomycin, IPTG and X-Gal (*see Note 7*).
3. For PCR-based screening of the C-Check vector, prepare two sets of 200 μL PCR tubes (lysate set and LB set). The positive rate of C-Check cloning is over 95%, pick up to 3 maximum colonies (white clones) for PCR-based screening.
4. Add 50 μL LB medium (50 $\mu\text{g}/\text{mL}$ spectinomycin) to each well of the LB set tubes, mark the tubes as LB1, LB2, LB3.
5. Add 30 μL ddH₂O to each well of the lysate set, mark the tube as LS1, LS2, LS3.

6. Use a sterile 100 μ L pipette tip, pick one white clone, firstly dip into LB1, gently swirl for 1 s.
7. Then place the tip in LS1 of the lysate tube.
8. Repeat pick the rest of colonies as **steps 6 and 7**.
9. Cap the lysate set and LB set tubes.
10. Place the LB set tubes at 37 °C incubator.
11. Place the lysate set tubes in a thermal cycler, lyse the cells at 98 °C for 10 min.
12. Save the lysate at 4 °C, and use 1 μ L lysate as template for PCR screening.
13. If using the synthetic approach for generating C-Check vector, PCR screening is carried out using the C-Check screening primer (Forward) and the antisense C-Check oligo (*see Note 8*).
14. If using the PCR approach for generating C-Check vector, PCR screening can be carried out using the same condition as amplifying the fragment as in Subheading 3.3, **step 3**.
15. Analyze PCR with a 1.0% agarose gel. For synthetic approach, the length of PCR product should be 600 bp + the length of C-Check antisense oligo. For PCR approach, the length of PCR product will be the same as the cloning fragment.
16. Set up LB culture of the PCR positive clones, 5 mL LB + 50 μ g/mL spectinomycin.
17. Purify C-Check plasmid using a miniprep kit.
18. Validate the C-Check plasmids with restriction enzymes (RE) digestion: BamHI and KpnI. Positive clones will result in two bands: one of 6.6 kb and the other of 600 bp+C-Check oligo/fragment.
19. Set up LB culture of one RE positive clone for midi-prep.
20. Purify plasmid using a commercial midi-prep kit.
21. Save plasmid at 4 °C until used.

3.5 Transfection of HEK293T Cells with C-Check and CRISPR

1. Seed 50,000 HEK293T cells per well to a 24-well plate, prepare triplicate for each transfection and enough wells of cells (*see Note 9*).
2. Transfect HEK293T cells with X-tremeGENE 9 or other transfection reagents. For each transfection, use a molar ratio of Cas9:gRNA:C-Check = 1:1:1.
3. Inspect cells using a fluorescent microscope 48 h post transfection (Fig. 2).

3.6 Analysis of Transfected Cells by Flow Cytometry

1. Prepare 5% FBS-PBS.
2. Wash cells with PBS twice.

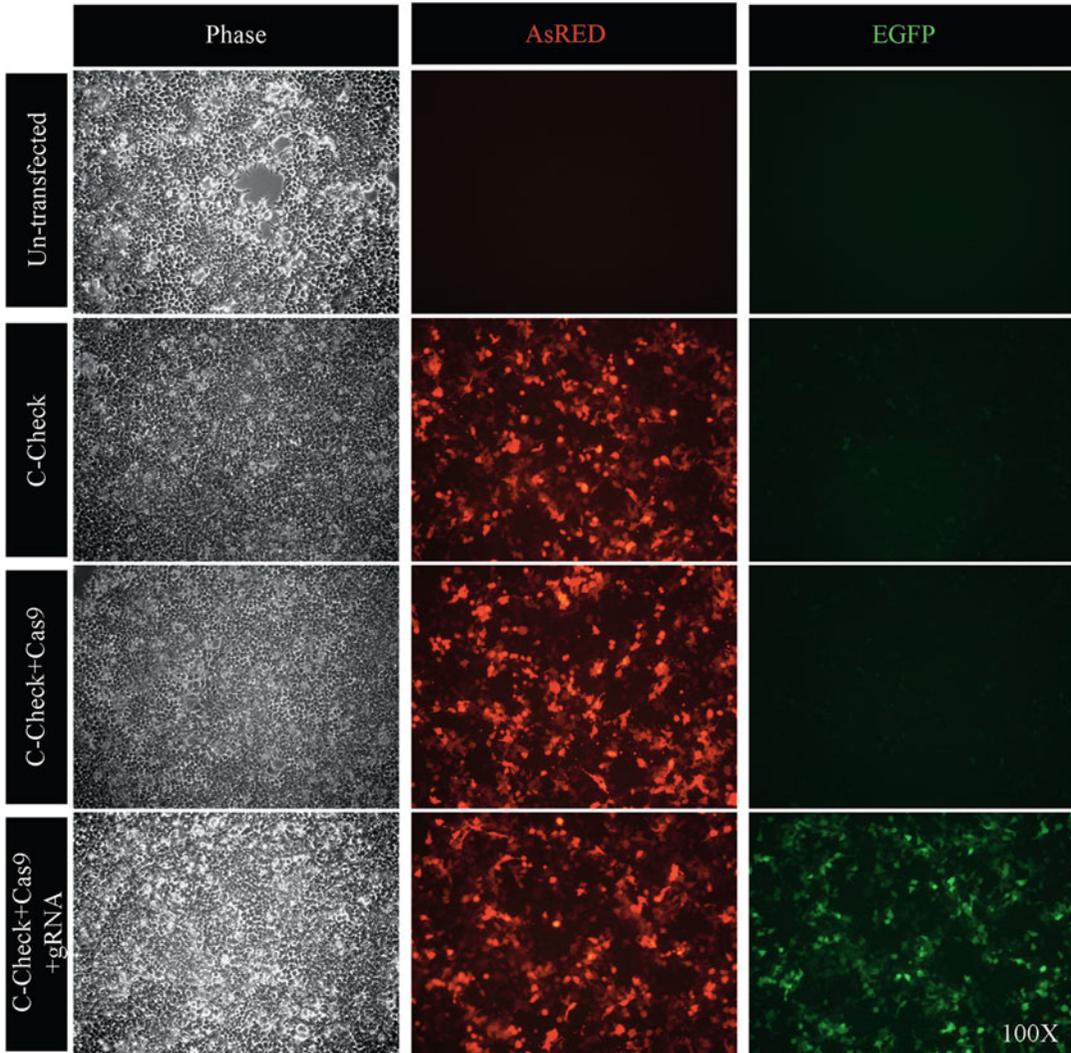


Fig. 2 Fluorescent microscopy pictures of C-Check-based measurement of CRISPR activity. Un-transfected HEK293T cells, HEK293T cells transfected with C-Check only, C-Check + Cas9, and C-Check+Cas9+gRNA targeting the surrogate region. Magnification, 100×

3. Add 100 μ L 0.05% Trypsin-EDTA (phenol red free) to each well (*see Note 10*).
4. Incubate at 37 °C for 4–5 mins.
5. Add 500 μ L 5% FBS-PBS to each well to stop trypsin.
6. Pipette up and down a few times and transfer the cells to tubes or 96-well round-well plate that is compatible with flow cytometry analysis, kept on ice until analysis.
7. If the cells are not analyzed immediately, fixed the cells with 4% formaldehyde for 10 min.
8. Wash cells twice with PBS.

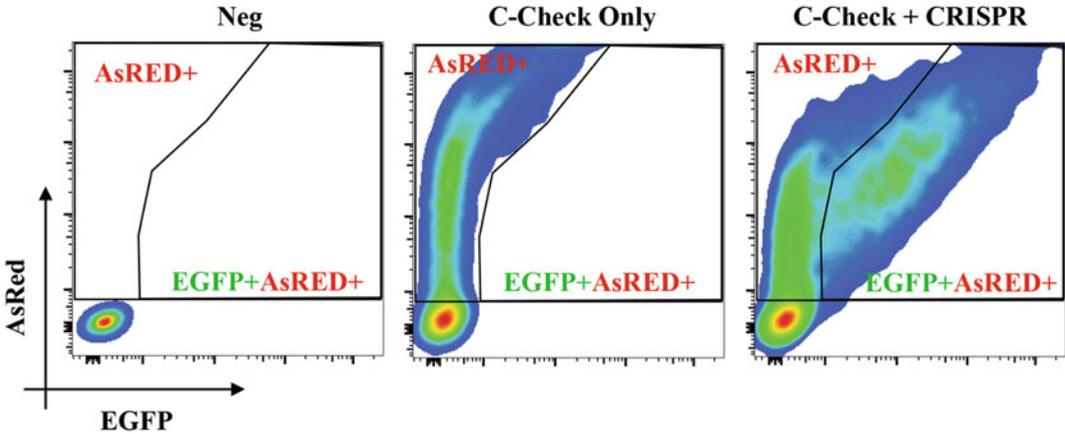


Fig. 3 FACS analysis of C-Check-based measurement of CRISPR activity. Gatings for transfected cells (AsRED+) and for cells that underwent SSA-mediated repair (EGFP+AsRED+) are presented

9. Resuspend the cells in 500 μ L PBS and save at 4 $^{\circ}$ C until flow cytometry analysis. The cells can be kept at 4 $^{\circ}$ C for one week.
10. To set up the flow cytometry, use the un-transfected cells, cells transfected with C-Check plasmid only, and cells transfected with C-Check, Cas9 and gRNA (*see Note 11*).
11. The C-Check fluorescence intensity will give a hook effect; analyze the data using the gating principles shown in Fig. 3.

4 Notes

1. The C-check plasmid can be ordered from Addgene or can be requested directly from us.
2. We used FastDigest restriction enzymes from [Thermo Scientific](#) to shorten the waiting time of digestion. Restriction enzymes from other suppliers will also work but require longer digestion time.
3. The chemically competent cells we used for our study is made by ourselves. If using commercially made ones, it is important to use competent cells that are recombination deficient. The C-Check vector contains internal homologous sequences. If the competent cells are proficient in recombination, this will greatly decrease the success of generating the C-Check vector.
4. Artificial stop codons were added as linker to the C-Check oligos to prevent both the read through and the introduction of alternative start codon.
5. The denaturing step should be controlled to minimum length, as the triphosphate stability is not great at 95 $^{\circ}$ C. If not using a thermal cycler, the anneal oligos can be prepared using a

heating block, but prepare the annealing mixture in a 1.5 mL EP tube instead. First denature at 95 °C for 5 min, then turn off the heating block and let it cool down naturally.

6. It's crucial to use isogenic genomic DNA from the cells (not the HEK293T cells used for C-Check assay) which will be used for the subsequent targeting. SNPs are highly frequent in cells. The use of isogenic genomic DNA ensures that the gRNAs activity measured exactly reflect that in the targeted cells. Proofreading DNA polymerase should be used to avoid the introduction of mutations by PCR.
7. The addition of IPTG and X-Gal to the LB plate is to exclude negative clones. We observed that approximately 1–2 blue clones (negative) can be observed for a good digestion.
8. Instead of using the antisense C-Check oligo as the reverse primers for PCR, one of the CRISPR oligos can also be used as reverse primer.
9. For each C-Check transfection, the following control groups should be prepared: (1) Un-transfected cells; (2) C-Check plasmid only; (3) C-Check + Cas9 expression plasmid.
10. The reason of using phenol red free trypsin is to avoid the steps of washing the cells after trypsinization.
11. Use the 488 nm laser and 530/30 filter for EGFP, 561 nm laser and 586/15 filter for AsRED. The voltages for 488 nm laser and 561 nm laser have to be optimized for according to the instrument used.

Acknowledgments

This work is partially supported by the Lundbeck Foundation (R219–2016-1375, R173–2014-1105), the Danish Research Council for Independent Research (DFF-1337–00128), the Sapere Aude Young Research Talent Prize (DFF-1335–00763A), the Innovation Fund Denmark (BrainStem), and Aarhus University Strategic Grant (AU-iCRISPR). Y.L is also supported by the Guangdong Provincial Key Laboratory of Genome Read and Write (No. 2017B030301011).

References

1. Mali P et al (2013) RNA-guided human genome engineering via Cas9. *Science* 339 (6121):823–826
2. Boch J et al (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326(5959):1509–1512
3. Henrik Devitt Møller LL, Xi X, Petersen TS, Huang J, Yang L, Kjeldsen E, Jensen UB, Zhang X, Liu X, Xun X, Wang J, Yang H, Church GM, Bolund L, Regenberg B, Luo Y (2018) CRISPR-C: circularization of genes and chromosome by CRISPR in human cells.

- Nucleic Acids Res. <https://doi.org/10.1093/nar/gky767>
4. Leenay RT, Beisel CL (2017) Deciphering, communicating, and engineering the CRISPR PAM. *J Mol Biol* 429(2):177–191
 5. Brinkman EK et al (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42(22):e168
 6. Yang Z et al (2015) Fast and sensitive detection of indels induced by precise gene targeting. *Nucleic Acids Res* 43(9):e59
 7. Zhou Y et al (2016) Enhanced genome editing in mammalian cells with a modified dual-fluorescent surrogate system. *Cell Mol Life Sci* 73(13):2543–2563
 8. Yang Y et al (2016) Highly efficient and rapid detection of the cleavage activity of Cas9/gRNA via a fluorescent reporter. *Appl Biochem Biotechnol* 180(4):655–667
 9. Jensen KT et al (2017) Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. *FEBS Lett* 591(13):1892–1901
 10. Liu B et al (2018) STAT3 associates with vacuolar H(+)-ATPase and regulates cytosolic and lysosomal pH. *Cell Res* 28(10):996–1012
 11. Lin L et al (2017) Fusion of SpCas9 to *E. coli* Rec A protein enhances CRISPR-Cas9 mediated gene knockout in mammalian cells. *J Biotechnol* 247:42–49

Part II

Methods for CRISPR Delivery



CRISPR-Cas9 Delivery by Artificial Virus (RRPHC)

Suleixin Yang, Qinjie Wu, Yuquan Wei, and Changyang Gong

Abstract

Since its first harnessing in gene editing in 2012 and successful application in mammalian gene editing in 2013, the CRISPR-Cas9 system exerted magnificent power in all gene-editing-related applications, indicating a sharp thrive of this novel technology. However, there are still some critical drawbacks of the CRISPR-Cas9 system that hampered its broad application in gene editing. Efficient delivery of the Cas9 protein and its partner small guide RNA (sgRNA) to the target cells or tissue is one of the technical bottlenecks. CRISPR-Cas9 delivery via DNA plasmids still plays the big role in gene editing methods. With regard to the disadvantages of CRISPR-Cas9 plasmids, the most acute barrier lies in its large size (>10 kb) and the subsequent low transfection efficiency by conventional transfection method. In this chapter, what we present is an easy method by fabricating CRISPR-Cas9 plasmids into nanoparticle system and efficiently delivered into target cells to achieve gene editing.

Key words CRISPR-Cas9, Transfection, Artificial virus, Branched polyethylene imine, Heptafluorobutyric anhydride

1 Introduction

Technical revolutions always lead to an outbreak of research and applications in both academic and industrial communities. Gene editing has attracted massive attentions due to its broad applications in science, biology, and health. Over the last few decades, there has been a great shortage of efficient gene editing tools. But with the discovery of TALENs and especially the third-generation gene editing tool Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas9 [1–4], gene editing technology was universally utilized in laboratories around the world and play a valuable and indispensable role in molecular botany[5–7], genetics [8, 9], and pharmacology [10]. With the booming of scientific research applying these gene editing tools, most researchers noticed some tough merits that hampered the CRISPR-Cas9's performance to be fully brought out, such as off-target effect, gene-correction efficiency, and difficulty of delivery [11]. The large size of CRISPR-Cas9 plasmids were broadly discussed [12, 13] which

affects packaging efficiency. Different kinds of methods have been tried to circumvent the inherent weakness that led to low efficiency of transfection. Virus-mediated CRISPR-Cas9 delivery, such as adeno-associated virus, lentivirus, and adeno virus [14–18], were mostly adopted, while the innate immune responses or DNA sensing mechanism can often affect the long-term performance. Nevertheless, none of them sharply improved the condition [19].

In this chapter, we describe a simple method of fabricating a kind of powerful “artificial virus” nanoparticle [20] with high-transfection efficiency to deliver CRISPR-Cas9 plasmids [21], known as RRPHC. The nanoparticle is based on the amphiphobicity of fluorinated compounds and phase-separation [22], which augmented affinity to lipid membranes like cell lipid bilayer membranes and lysosomal/endosomal membranes. The fluorinated branched polyethylene imine (bPEI) was used as the core material to condense the CRISPR-Cas9 plasmids, followed by furtherly wrapped with the cell penetrating peptide R8-RGD modified hyaluronan shell.

2 Materials

Prepare all solutions using ultrapure water and analytical grade reagents.

1. Anhydrous methanol.
2. Triethylamine.
3. Ethylene glycol-monomethyl ether.
4. *N*-(3-Dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDCI).
5. *N*-hydroxysuccinimide (NHS).
6. Ninhydrin.
7. Hydrindantin.
8. Branched polyethylene imine (bPEI, molecular weight = 1800).
9. Heptafluorobutyric anhydride.
10. Hyaluronic acid (HA, molecular weight = 35 KD).
11. [c(RGDfK)-RRRRRRRR-Cys] (RGD-R8) was synthesized by Chinapeptides (Suzhou, China).
12. Maleimide-PEG₂₀₀₀-NH₂.
13. Phosphate-buffered saline (PBS) solution: 8 g/L sodium chloride, 0.2 g/L potassium chloride, 1.42 g/L disodium phosphate, 0.24 g/L monopotassium phosphate, pH 7.4.

14. 2-(*N*-morpholino) ethanesulfonic acid (MES) buffer: 0.1 M 2-(*N*-morpholino) ethanesulfonic acid, pH 6.5.
15. Sodium acetate buffer: 0.2 M sodium acetate, pH 5.4.
16. Rotavapor.
17. Magnetic Stirrer.
18. Lyophilizer.
19. Microplate Reader.
20. Inverted fluorescence microscopy.
21. Flow cytometer.
22. Dialysis membrane (1000 Da).

3 Methods

3.1 Synthesis of Fluorinated PEI (PFs) and RGD-R8-PEG-HA (RRPH)

The “artificial virus” was constituted by a plasmid-loading core PFs and RRPHC shell. Here in the convenience of distinguishing, we denoted the plasmid-loading nanosphere as PFs complex (PFC). PFC is subsequently coated by RRPHC, named as RRPHC, which resembled virus on a larger scale. Core materials PFs (*s* means grafting ratio) laid the foundation of high transfection efficiency. PFs can be easily synthesized through a one-step method (Fig. 1):

1. Replace air with nitrogen to keep an absolute water-free environment.
2. Dissolve 200 mg bPEI in 5 mL anhydrous methanol in a round-shaped flask in an ice bath with magnetic stirring.
3. After 30 mins, when the sticky bPEI was totally dissolved, add 200 μ L of anhydrous triethylamine into the bPEI solution.
4. Assemble the constant pressure drop funnel on top of the aforementioned round-shaped flask.
5. Dissolve 220 μ L heptafluorobutyric anhydride in 10 mL anhydrous methanol in the constant pressure drop funnel with magnetic stirring.
6. Add the heptafluorobutyric anhydride solution from **step 5** into the bPEI solvent system from **step 4** at a speed of 3 drops per second.
7. Displace the ice bath until last drop of heptafluorobutyric anhydride solution.
8. 48 hours later, the liquid in the flask was collected.
9. Dialyze the liquid from **step 8** against double-distilled (dd) water using a dialysis membrane for 3 days with 3 times renewal per day.

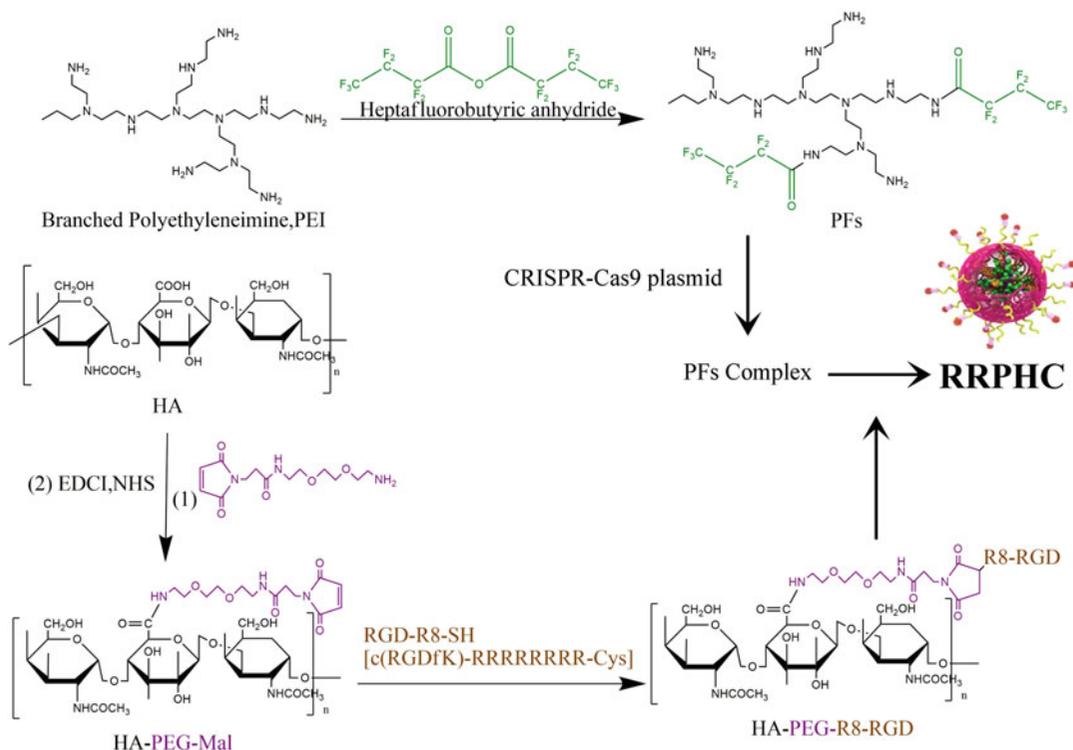


Fig. 1 Synthesis of PFs and RRPHC

10. Lyophilize the liquid from **step 9** for 2 days, followed by its measurement of weight and store at -20°C .
11. Dissolve 42.3 mg HA in 20 mL 25°C MES buffer and the carboxylic groups of HA were activated by EDCI and NHS for 2 h.
12. Add 200 mg Mal-PEG-NH₂ to the reaction and proceed for 24 h.
13. The liquid from **step 12** was dialyzed against double-distilled water for 3 days in PBS in a dialysis membrane (MWCO, 3500 Da).
14. Add 15 mg RGD-R8 into the liquid collected from **step 13** and react for another 48 h.
15. Dialyze the final product in PBS for 2 days and 1 day in ddH₂O and lyophilize for 2 days.

3.2 Characterization of PFs and RRPHC

To characterize the syntheses of PFs and RRPHC. Three different chemical shifts of ^{19}F -NMR spectrum indicated the grafts of fluorines (Fig. 2a). FTIR image showed the existence of carbonyl group (1686 cm^{-1}) and C-F bond (1225 cm^{-1}) compared to bPEI (Fig. 2b). And for RRPHC, the characteristic absorption peaks: carboxylic group (3442 cm^{-1}), ester group (1162 and

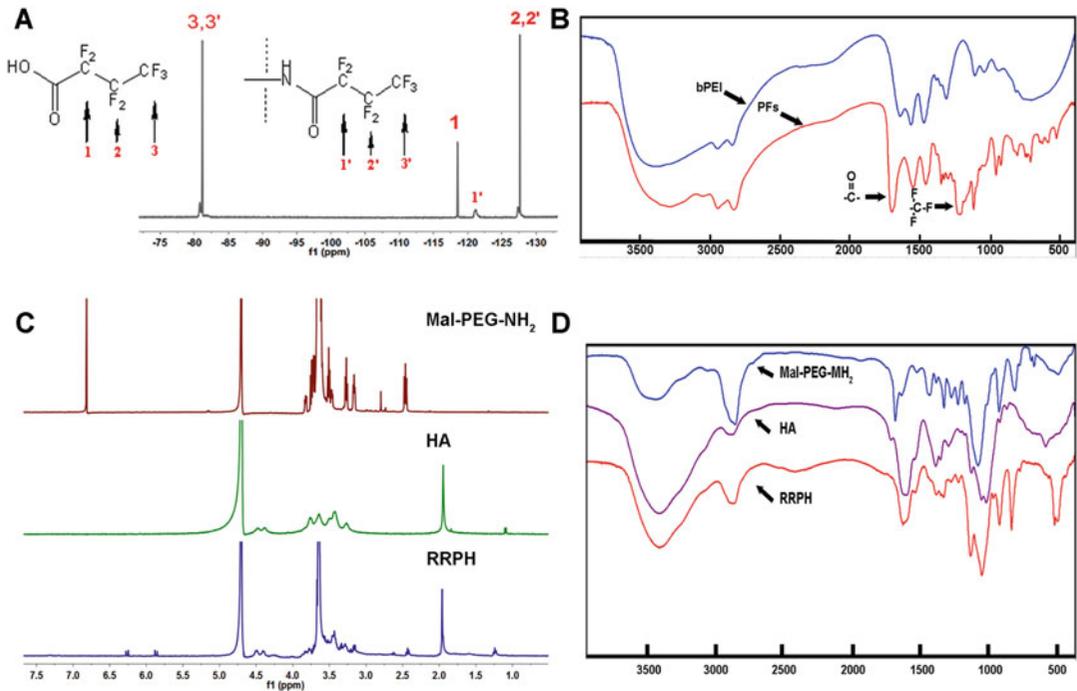


Fig. 2 NMR and FTIR spectra of PFs (a, b) and RRPHC (c, d)

1075 cm^{-1}), guanidine group (1636 cm^{-1}) (Fig. 2d) testified the successful modifications of PEG and R8-RGD on HA.

Grafting ratio of fluorophores on bPEI was measured by an improved ninhydrin method [23]:

1. Dissolve 170 mg ninhydrin and 30 mg hydrindantin in 20 mL ethylene glycol-monomethyl ether.
2. Mix 100 μL ethylene glycol-monomethyl ether solution from **step 1** with 100 μL sodium acetate buffer.
3. Add 100 μL bPEI 1.8 K or PFs to the liquid from **step 2** and heat in boiling water for 10 mins.
4. Cool down the mixture for several minutes, and dilute it with 300 μL ethanol/water (v/v ~60:40) solution.
5. Measure the absorption values at 570 nm at different concentrations of primary amine using a UV-Vis spectroscopy.
6. Draw the calibration curve (Fig. 3) basing on data like table below (Table 1), which shows the absorption value (X) vs. molarity of primary amines (Y).
7. Retrodict (Eq. 1) samples' Abs and deduce the molarity (X').
8. Obtain the equation (Eq. 2), which numbers of grafted fluoro groups (dubbed as n) could be solved.

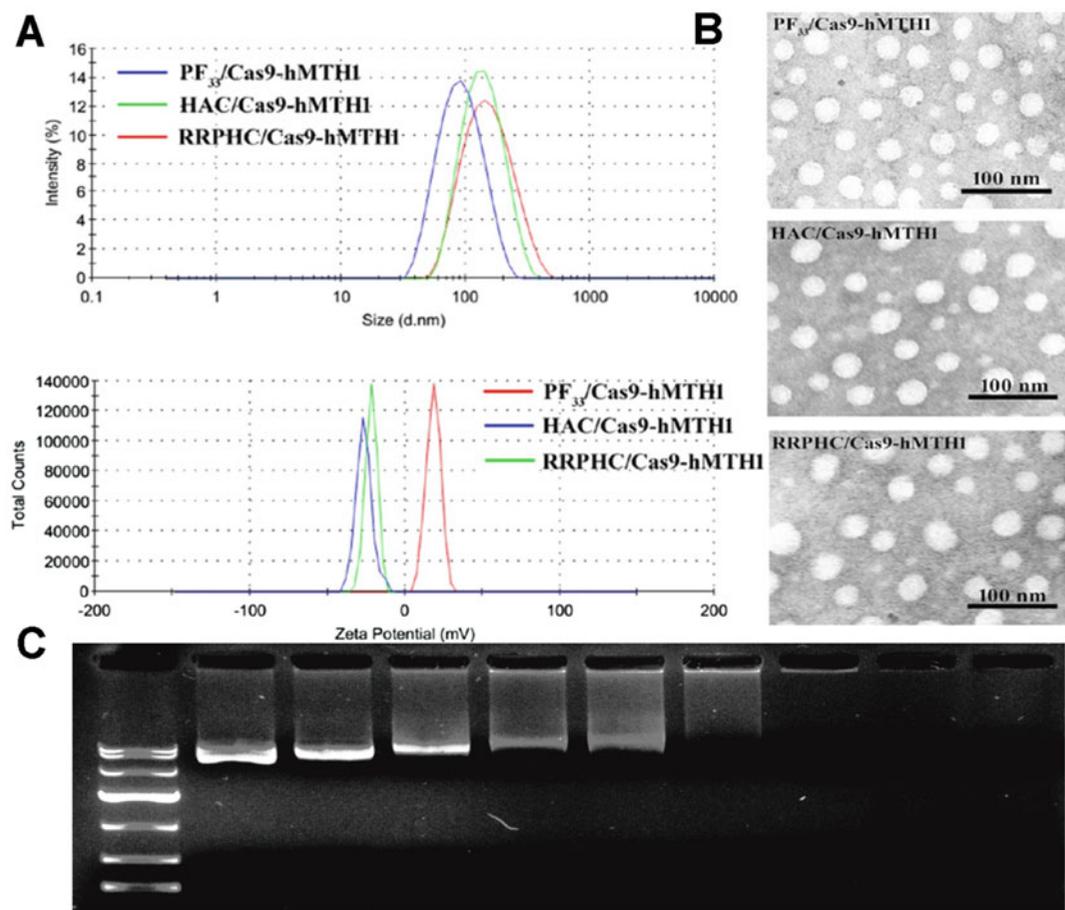


Fig. 3 Evaluation of size and zeta potential (a), TEM (b), and condensation ability (c) of RRPHC. Reprinted from Ref. [19] with permission (a and b)

Table 1
Evaluation of grafting ratio of the PFs

Theoretical PFs	Measured PFs	<i>n</i>	Grafting ratio (%)
PF ₂₀	PF ₁₅	2.42	15.1
PF ₃₀	PF ₂₃	3.68	23.3
PF ₄₀	PF ₃₀	4.80	30.1
PF ₅₀	PF ₃₃	5.28	33.2

n stands for the average number of heptafluorobutyric acid groups grafted on every sample

- Divide *n* by 16 (numbers of primary amines which we reckon each bPEI 1.8 K molecule possess after calculation) equals the grafting ratio.

From the characterization of Table 1, we can see that the tested grafting ratio was lower than theory.

$$\Upsilon = 406.7X - 6.572 \quad (1)$$

Equation 1, Υ stands for Abs while X for concentration of primary amine

$$\frac{C}{(196n + 1800)} \times (16 - n) = X' \quad (2)$$

Equation 2, C stands for mass concentration and X' for retrodict molarity.

3.3 Preparation of Ternary Complexes RRPHC

In this protocol, we chose PF₃₃ (as showed previously, 33 represents the grafting ratio which reached 33%) in our experiments and the NPs are assembled by simple layer by layer mixing.

1. Dilute 5 μg plasmids in 50 μL PBS in a 1.5 mL EP tube.
2. Disperse 50 μg PF₃₃ in 50 μL serum-free media in a 1.5 mL EP tube.
3. Mix the plasmids and PF₃₃ together and incubate at room temperature for 25 mins.
4. Add 30 μL RRPHC stock solution (5 $\mu\text{g}/\mu\text{L}$, RRPHC dissolved in PBS) into the tube and incubate for 15 mins to form ultimate nanoparticles. A schematic procedure is shown in Fig. 4.

3.4 Evaluation of Size, Zeta Potential, and Condensation Ability

The size and zeta potentials of RRPHC could greatly impact the efficiency, stability, and safety in vitro and in vivo. Here we used dynamic light scattering to measure the size and zeta data of the nanoparticle. Before coated by RRPHC, the diameter was 77.5 ± 3.6 nm, and has a zeta potential of $+21.3 \pm 2.3$ mV. Furthermore, as showed in Fig. 5a, after coated by RRPHC, the diameter escalated to 131.3 ± 4.2 nm, and zeta potential met an overhaul to -21.8 ± 1.8 mV. Also, the size change was observed via transmission electron microscopy (Fig. 5b). These phenomena together indicated the successful wrapping of the negative electrical RRPHC. Usually the nanoparticles possess a diameter of 100–200 nm, and the zeta potential can reach approximately -20 mV.

In order to test the condensation ability of the nanoparticles, gel electrophoresis assays were conducted (Fig. 3c). 500 ng CRISPR-Cas9 GFP plasmids are respectively incubated with different amounts of PFs and subsequent RRPHC, and the “disappeared” band would show at how low a ratio can our materials tightly wrap the plasmids, thus protecting it from degradation.

3.5 Choosing Best Transfection Ratio

The best transfection ratio is chosen based on testing lots of ratios. To test that, we typically seed 293T cells (1×10^5 cells/well) to several 12-well plates one day prior to transfection. 24 hours later, the media was replaced by fresh serum-free media to synchronize

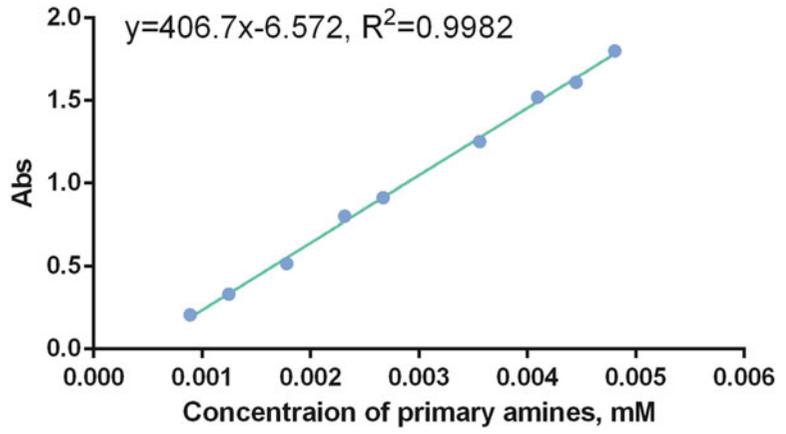


Fig. 4 Line that reflect the relationship of concentration of primary amines and Abs

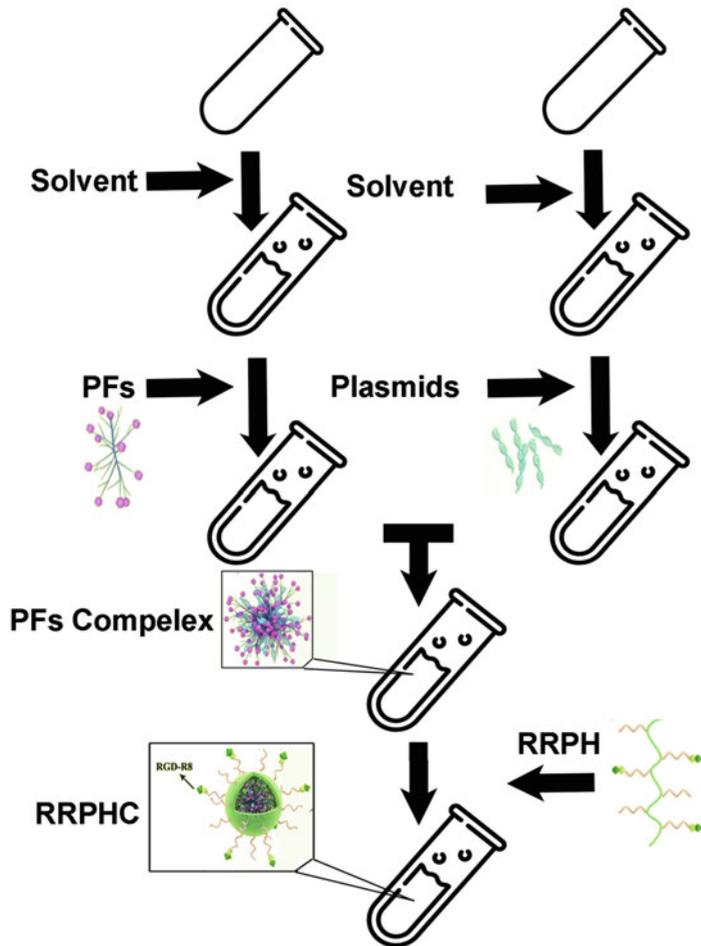


Fig. 5 Concise procedures of preparing RRPHC nanoparticle

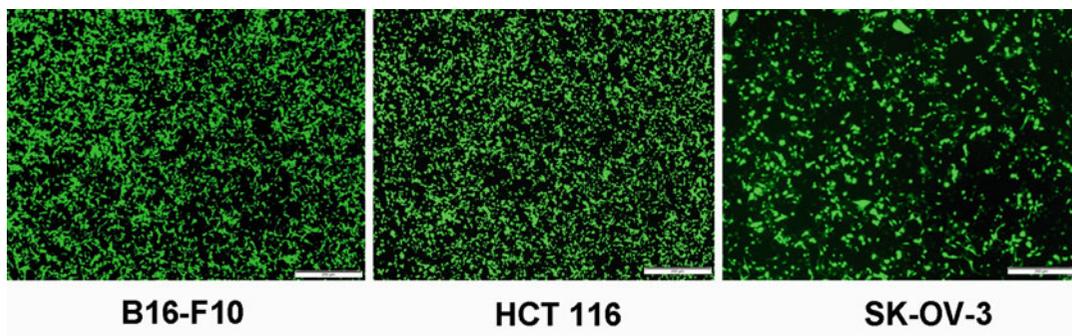


Fig. 6 Transfection results using RRPHC on different types of malignant tumor cells

the cells by serum starvation. Then 1 μg CRISPR-Cas9-GFP plasmids were mixed with different amount of PFs to get binary complexes with ratios ranging from 1:1 to 15:1, and incubate for 25 mins. RRPHC was subsequently added into former tubes to incubate for another 15 mins to generate nanoparticles coating plasmids. The nanoparticles with different ratios were added to each well and after 6 h, the serum-free medium was replaced with normal cell culture medium with 10% fetal bovine serum (FBS). To evaluate the efficiency of transfection, fluorescent images were taken 24 h later after transfection (Fig. 6) via a fluorescent microscopy. Transfection efficiency was also determined by flow cytometer (FCM).

4 Future Prospects

To efficiently deliver the CRISPR-Cas9 plasmids into cells for achieving genome editing, efficient delivery methods such as our approach are in huge demand. By overcoming the inherent drawbacks in relation to the large plasmid size for transfection, the RRPHC platform provides several advantages: tightly enough to encapsulate the plasmids and readily enough to deliver the plasmids into nucleus. In our previous work, we found RRPHC possess staggering potency for its high transfection efficiency in cancer cells such as B16-F10, HCT116, and SK-OV-3 (Fig. 6) and at the same time, low cytotoxicity, readiness of fabrication, low-cost, and great tumor-targeting ability *in vivo* [20] implying the prosperous applications of the artificial virus (RRPHC). For further improvement, the RGD-R8 moiety may be replaced by other functional ligands and peptides if necessary. In addition, projects on further exploration of PFs with higher grafting ratios, different forms (linear, branched, and super-branched), and molecular weight (1,200, 10,000, 25,000, etc.) of PEI are being carried out. We expect to furtherly develop the RRPHC method to achieve better CRISPR gene editing efficiencies and translation into clinical trials.

References

1. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (2013) Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8 (11):2281–2308. <https://doi.org/10.1038/nprot.2013.143>
2. Wang HY, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153 (4):910–918. <https://doi.org/10.1016/j.cell.2013.04.025>
3. Hsu PD, Lander ES, Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157(6):1262–1278. <https://doi.org/10.1016/j.cell.2014.05.010>
4. Yang L, Mali P, Kim-Kiselak C, Church G (2014) CRISPR-Cas-mediated targeted genome editing in human cells. *Methods Mol Biol* 1114:245–267. https://doi.org/10.1007/978-1-62703-761-7_16
5. Weeks DP, Jiang WZ (2014) Genome editing in plants using CRISPR/Cas9/sgRNA technologies. *In Vitro Cell Dev-An* 50:S21–S21
6. Xing HL, Dong L, Wang ZP, Zhang HY, Han CY, Liu B, Wang XC, Chen QJ (2014) A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol* 14:327. <https://doi.org/10.1186/S12870-014-0327-Y>
7. Belhaj K, Chaparro-Garcia A, Kamoun S, Patron NJ, Nekrasov V (2015) Editing plant genomes with CRISPR/Cas9. *Curr Opin Biotechnol* 32:76–84. <https://doi.org/10.1016/j.copbio.2014.11.007>
8. Butler M, van der Meer L, Yu J, Beeby T, Kuiper R, van Leeuwen F (2017) CRISPR/CAS9 based reverse genetics identifies therapeutic targets to improve therapy response in paediatric acute lymphoblastic leukemia. *Pediatr Blood Cancer* 64:S5–S6
9. Gonzalez F (2016) CRISPR/Cas9 genome editing in human pluripotent stem cells: harnessing human genetics in a dish. *Dev Dyn* 245 (7):788–806. <https://doi.org/10.1002/Dvdy.24414>
10. Naylor J, Suckow AT, Seth A, Baker DJ, Sermadiras I, Ravn P, Howes R, Li JL, Snaith MR, Coghlan MP, Hornigold DC (2016) Use of CRISPR/Cas9-engineered INS-1 pancreatic beta cells to define the pharmacology of dual GIPR/GLP-1R agonists. *Biochem J* 473:2881–2891. <https://doi.org/10.1042/Bcj20160476>
11. Cox DBT, Platt RJ, Zhang F (2015) Therapeutic genome editing: prospects and challenges. *Nat Med* 21(2):121–131
12. Li L, He ZY, Wei XW, Gao GP, Wei YQ (2015) Challenges in CRISPR/CAS9 delivery: potential roles of nonviral vectors. *Hum Gene Ther* 26(7):452–462. <https://doi.org/10.1089/hum.2015.069>
13. Doudna JA, Charpentier E (2014) The new frontier of genome engineering with CRISPR-Cas9. *Science* 346(6213):1258096. <https://doi.org/10.1126/Science.1258096>
14. Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, Dahlman JE, Parnas O, Eisenhaure TM, Jovanovic M, Graham DB, Jhunjhunwala S, Heidenreich M, Xavier RJ, Langer R, Anderson DG, Hacohen N, Regev A, Feng G, Sharp PA, Zhang F (2014) CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* 159 (2):440–455. <https://doi.org/10.1016/j.cell.2014.09.014>
15. Senis E, Fatouros C, Grosse S, Wiedtke E, Niopek D, Mueller AK, Borner K, Grimm D (2014) CRISPR/Cas9-mediated genome engineering: an adeno-associated viral (AAV) vector toolbox. *Biotechnol J* 9 (11):1402–1412. <https://doi.org/10.1002/biot.201400046>
16. Kennedy EM, Cullen BR (2015) Bacterial CRISPR/Cas DNA endonucleases: a revolutionary technology that could dramatically impact viral research and treatment. *Virology* 479–480:213–220. <https://doi.org/10.1016/j.virol.2015.02.024>
17. Cheong TC, Compagno M, Chiarle R (2016) Editing of mouse and human immunoglobulin genes by CRISPR-Cas9 system. *Nat Commun* 7:10934. <https://doi.org/10.1038/ncomms10934>
18. Fricano-Kugler CJ, Williams MR, Salinaro JR, Li M, Luikart B (2016) Designing, packaging, and delivery of high titer CRISPR retro and lentiviruses via stereotaxic injection. *J Vis Exp* 111. <https://doi.org/10.3791/53783>
19. Yin H, Kanasty RL, Eltoukhy AA, Vegas AJ, Dorkin JR, Anderson DG (2014) Non-viral vectors for gene-based therapy. *Nat Rev Genet* 15(8):541–555. <https://doi.org/10.1038/nrg3763>
20. Li L, Song LJ, Liu XW, Yang X, Li X, He T, Wang N, Yang SLX, Yu C, Yin T, Wen YZ, He ZY, Wei XW, Su WJ, Wu QJ, Yao SH, Gong CY, Wei YQ (2017) Artificial virus delivers CRISPR-Cas9 system for genome editing of

- cells in mice. *ACS Nano* 11(1):95–111. <https://doi.org/10.1021/acsnano.6b04261>
21. Li L, Song LJ, Yang X, Li X, Wu YZ, He T, Wang N, Yang SLX, Zeng Y, Yang L, Wu QJ, Wei YQ, Gong CY (2016) Multifunctional “core-shell” nanoparticles-based gene delivery for treatment of aggressive melanoma. *Biomaterials* 111:124–137. <https://doi.org/10.1016/j.biomaterials.2016.09.019>
 22. Horváth IT, Rábai J (1994) Facile catalyst separation without water: fluorous biphasic hydroformylation of olefins. *Science* 266(5182):72
 23. Wang MM, Liu HM, Li L, Cheng YY (2014) A fluorinated dendrimer achieves excellent gene transfection efficacy at extremely low nitrogen to phosphorus ratios. *Nat Commun* 5:3053. <https://doi.org/10.1038/Ncomms4053>



Production and Validation of Lentiviral Vectors for CRISPR/Cas9 Delivery

Laura Barrett Ryø, Emil Aagaard Thomsen, and Jacob Giehm Mikkelsen

Abstract

Genetic information transferred by HIV-1-based lentiviral vectors as single-stranded RNA is converted to double-stranded DNA by reverse transcription and subsequently inserted into the genome of recipient cells. Integration into the genome allows stable, long-term expression of genes-of-interest driven by promoter sequences contained within the vector. This technology can be used as a standard method for production of cells stably expressing Cas9 protein and single guide RNA (sgRNA), the key components of the CRISPR genome editing system. Here, we provide a protocol for production and validation of VSV-G-pseudotyped lentiviral vectors for delivery of the CRISPR system and generation of knockout cell lines.

Key words CRISPR/Cas9, Lentiviral, sgRNA, Genome-wide screening, Forward genetics

1 Introduction

Human immunodeficiency virus type 1 (HIV-1)-derived lentiviral vectors, originally developed by Naldini and coworkers [1, 2], remain powerful carriers of genetic information. Several recent reports describe the effective use of lentiviral vectors for therapeutic gene transfer, including gene delivery to hematopoietic stem cells for treatment of immunodeficiencies and leukodystrophies [3–6], but for many researchers, lentiviral vector systems are key tools in everyday studies of molecular biology and genetics. Also for genome editing and delivery of the components of the CRISPR/Cas9 system, lentiviral vectors have become cherished vehicles.

Already in the 1990s, shortly after the invention of lentivirus-based vectors, initial biosafety concerns induced the development of safer lentiviral vector systems, resulting in the third-generation lentiviral vector system [7], which is essentially the system currently used in many laboratories. This system features a chimeric long terminal repeat (LTR) containing a constitutive promoter upstream of the vector transcript. In a later version of this vector, the U3 region in the downstream LTR was partially deleted, facilitating

formation of proviral DNA devoid of HIV-1-derived promoter elements [8, 9]. These vectors, referred to as “self-inactivating” (SIN) lentiviral vectors, allowed integration of vector-derived DNA with expression of the transgene driven strictly by an internal promoter of choice. Early on, cell lines for production of lentiviral vectors were developed [10], but state-of-the-art lentiviral vector production today is based on transfection of four plasmids (three packaging constructs and a transfer vector) into easy-to-transfect HEK293T cells. Here, we provide a protocol for production and validation of lentiviral vectors transferring *Streptococcus pyogenes* SpCas9 and single guide RNA (sgRNA) expression cassettes.

2 Materials

2.1 Production of Lentiviral Vectors

1. Lentiviral packing and transfer plasmids (Addgene IDs: 12253, 12251, 12259)
2. 2.5 M CaCl₂.
3. 2× HEPES Buffered Saline (2×HBS), pH 7.2.
4. dH₂O.
5. HEK293T cells.
6. 15 mL falcon tubes.
7. p15 dishes.
8. 0.45 μM filter.
9. Centrifuge tubes.
10. General reagents and appliances for cell culturing.
11. Ultracentrifuge (We run the samples in a SW28 rotor using the Beckman coulter centrifuge tubes, reference number: 358126).
12. PBS.
13. Parafilm.

2.2 Titration of Lentiviral Particles

1. General reagents and appliances for cell culturing.
2. qPCR master mix.
3. qPCR primer and probe sets for amplification of the Woodchuck Hepatitis Virus post-regulatory element (WPRE) and segments of the albumin gene.
4. pAlbumin available from Addgene (www.addgene.org; Addgene ID: 22037).
5. 96-well optical plate.
6. Optical film.
7. Lightcycler Real-Time PCR system.

8. Antibiotic (depending on the resistance gene in the transfer plasmid).
9. 0.6% methylene blue solution.
10. Regular H₂O.
11. Optional: polybrene.

2.3 Generation of Knockout Cell Lines Using CRISPR/Cas9

1. pLentiCas9-blast [11], pLentiGuide-Puro [11], and pLenti-CRISPR v2 [11] available from Addgene (www.addgene.org; Addgene IDs: 52962, 52963, 52961).
2. Blasticidin (final concentration suitable for your cell line).
3. Puromycin (final concentration suitable for your cell line).
4. General reagents and appliances for cell culturing.

2.4 Assessment of CRISPR Gene Disruption Efficiency by TIDE Analysis

1. Reagents for PCR.
2. Primer set to amplify a fragment encompassing the sgRNA target region.
3. 1% agarose gel.
4. General equipment to run an agarose gel.
5. Gel extraction kit.
6. Barcodes to send samples for sequencing, or in-house Sanger-sequencing equipment.

3 Methods

3.1 Production of Lentiviral Vectors

Initially you will need HEK293T cells and an estimate of the yield of lentiviral vectors required for a certain application. You will also have to consider whether or not you will need to further purify and concentrate viral particles harvested in growth medium from transfected HEK293T cells (we refer to this harvest as a “crude” preparation) by ultracentrifugation.

The production of lentiviral vectors takes 4–6 days depending on the number of harvests from transfected producer cells and whether you will need to up-concentrate the vector preparation by ultracentrifugation.

This protocol below is based on production of lentiviral vectors in p15 dishes. You may want to scale it down and use p10 dishes (*see Note 1*).

1. Day 1, seed 1×10^7 HEK293T cells per p15 dish in 25 mL medium.
2. Day 2, change medium on the HEK293T cells (25 mL DMEM) 1 h prior to transfection.
3. Prepare the transfection mix according to Table 1 (*see Note 2*).

Table 1
Amount of each plasmid used for preparing the transfection mix

Plasmid	per p15 dish (μg)
pRSV-REV	7.26
pMD.2G	9.07
pMDIg/p-RRE	31.46
Transfer plasmid	31.46

Co-transfect the HEK293T cells with the lentiviral packaging and transfer plasmids. We use the third generation lentiviral packaging plasmids available at Addgene (www.Addgene.org). Names of the plasmids that we use and plasmid amounts are provided in Table 1.

4. Add dH₂O to the pDNA mix to a final volume of 1089 μL and mix.
 5. Add 121 μL 2.5 M CaCl₂ and mix by pipetting up and down.
 6. Add 1210 μL 2 \times HBS buffer to the pDNA-CaCl₂ solution.
 7. Vortex the solution thoroughly (10–15 s) immediately after adding the 2 \times HBS.
 8. Incubate the transfection mix at room temperature for 15 min (*see Note 3*). Add the 2.42 mL transfection mix dropwise to the cells and gently swirl the dish to distribute the transfection mix.
 9. Day 3, remove the medium from the p15 dish and add 19 mL fresh medium (*see Note 4*).
 10. Check the transfection efficiency (optional): on day 3 you can evaluate the transfection efficiency by fluorescence microscopy. If you estimate the transfection efficiency to be below 75%, consider starting over. Examples of GFP expression 24 and 48 h after successful transfection with pCCL/PGK-eGFP as the transfer plasmid are provided in Fig. 1.
 11. Day 4, pool the supernatant from the p15 dishes if you have more than one dish transfected with the same pDNA mix.
 12. If you are planning on doing a second harvest on day 5, add 19 mL fresh medium and incubate for another 24 h (*see Note 5*).
 13. Filter the pooled supernatant through a 0.45 μm filter (*see Note 6*).
1. Divide the supernatant equally between an appropriate number of centrifuge tubes (aliquote approximately 24 mL per tube leaving space for the sucrose solution).

3.2 Concentration by Ultracentrifugation

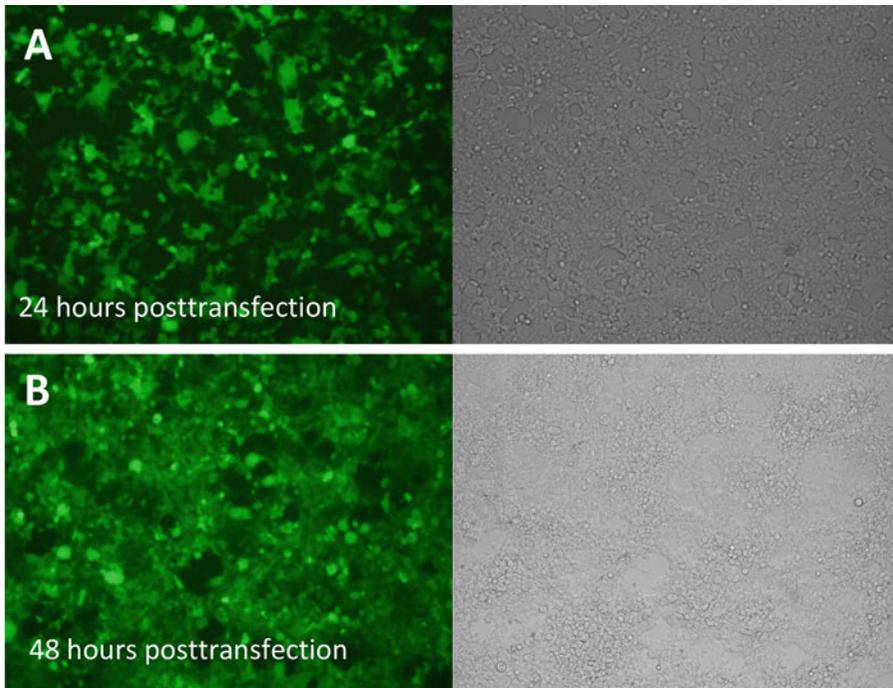


Fig. 1 Expected level of GFP expression 24 and 48 h after DNA transfection. HEK293T cells were seeded in a p15 dish and transfected with the plasmids indicated in Table 1 including pCCL/PGK-eGFP as the transfer plasmid. GFP expression, in this case driven by the phosphoglycerate kinase promoter (PGK), was assessed 24 (a) and 48 (b) hours after transfection by fluorescence microscopy. Right panels show the same area using light microscopy

2. Slowly add 4 mL sucrose solution to the bottom of the centrifuge tubes using a serological pipette (*see Note 7*).
3. Weigh the centrifuge tubes and adjust the volumes to ensure balance in the centrifuge.
4. Ultracentrifuge for 2 h at $113,000 \times g$ at 4°C .
5. After centrifugation, discard the supernatant and let the tubes stand upside down for 2 mins.
6. Remove any remaining supernatant from the tube.
7. Add the volume of $\text{PBS}^{-/-}$ corresponding to the desired fold up-concentration (*see Note 8*).
8. Cover the centrifuge tubes with parafilm and let the virus pellet resuspend overnight at 4°C .
9. Repeat the protocol for supernatant from day 5 and discard the cells.
10. Pool the virus particles from the first and second harvest, make aliquots in appropriate sizes to avoid repeated thawing and freezing, and store the aliquots at -70 to -80°C until use.

3.3 Titration of Lentiviral Particles by qPCR

A viral titer is a numerical expression of the quantity of viral units in a given volume and is usually either a physical titer or a transductional (functional) titer. Physical titers can be estimated by p24 ELISA or qPCR on viral RNA. These assays are generally faster to perform than assays used to determine a transductional titer. However, the assays used to determine a physical titer tend to overestimate the titer, as they measure both functional and defective viral particles. Transductional titers are based on serial transductions of the target cells and indicate the amount of infectious virus particles in a given volume. These assays are in general more laborious than the ones used to determine physical titer. However, they are considered to be more accurate and are specific for the cell type in which the assay is performed. In this protocol, we present two ways to determine the transductional titer of a preparation of lentiviral vectors.

Assessing transductional titer by qPCR is a viable option in both suspension and adherent cells. The protocol that we use is adapted from a protocol developed by Didier Trono's lab (<https://tronolab.epfl.ch/>) and utilizes qPCR to determine the titer of a lentiviral vector batch. This titer assay is based on qPCR measurement of the number of lentiviral DNA copies integrated into the genome of the target cells plus potential episomal DNA intermediates. The number of integration-competent particles in the virus batch can be calculated by normalizing to an endogenous control gene with two copies per cell.

In this set up, the Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE), present in most third-generation lentiviral vectors, is used to determine the number of integrated lentiviral DNA copies, and the gene encoding albumin is used as the endogenous control gene.

1. Day 1, seed 1×10^5 target cells per well in a 6-well dish or in another culture format. Prepare one well for each dilution of the vector batch and one extra well for a non-transduced control (*see Note 9* if the target cells are cultured in suspension).
2. Day 2, transduce the target cells with appropriate dilutions of the vector batch (*see Note 10*). If the target cells can tolerate polybrene, it is possible to supplement with $8 \mu\text{g}/\text{mL}$ polybrene (final concentration) to increase the transduction efficiency.
3. Count the number of cells in the non-transduced well and re-seed them. The number of cells is used to calculate the titer after the qPCR has been performed; re-seeded cells serve as a control in the qPCR reaction.
4. Day 3, change medium on the transduced cells and on control cells.

Table 2
Sequences of primers and probes used for qPCR

Target sequence	Primer/probe	Primer/probe sequence
WPRE	WPRE_forward	5'-GGCACTGACAATTCCGTGGT-3'
	WPRE_reverse	5'-AGGGACGTAGCAGAAGGACG-3'
	WPRE_probe	5'-FAM-ACGTCCTTTCCATGGCTGCTCGC-BHQ-3'
Albumin	Albumin_forward	5'-GCTGTCATCTCTTGTGGGCTGT-3'
	Albumin_reverse	5'-ACTCATGGGAGCTGCTGGTTC-3'
	Albumin_probe	5'-FAM-CCTGTCATGCCACACAAATCTCTCC-BHQ-3'

5. Day 6, extract genomic DNA (gDNA) from each individual well using your preferred protocol.
6. Elute gDNA with dH₂O.
7. Store gDNA at -20 °C until use.

3.4 Quantification of WPRE in Proviral DNA

The primer and probe sets for the qPCR are presented in Table 2. The WPRE primer/probe set is used for quantification of the WPRE sequence, which is contained in the vector and thus inserted into the genome of the target cells after lentiviral transduction. The albumin gene primer/probe set is used for quantification of the endogenous albumin gene, which is used to normalize for the amount of gDNA present in the qPCR reaction. The albumin primer/probe set is specific for the human albumin gene, so if your target cells have nonhuman origin, an alternative primer/probe set should be used. The standard curves for WPRE and the albumin gene are made using plasmids containing WPRE and albumin gene sequences (*see Note 11*).

1. Adjust the DNA concentrations of the two plasmids containing the WPRE and the albumin gene sequence to 1 µg/µL.
2. The first point of the standard curves should be 1×10^8 molecules in 8 µL corresponding to 1.25×10^7 molecules per µL. Adjust the concentration of the two plasmids for the standard curves accordingly (*see Note 12*).
3. Generate the standard curves by making tenfold dilutions until 10 molecules in 8 µL (8 serial dilutions in total).
4. Prepare the gDNA samples. The gDNA sample amount in each qPCR reaction should be between 50 and 100 ng in 8 µL (6.25–12.5 ng/µL).
5. Prepare two master mixes (Table 3), one for quantification of WPRE and one for quantification of the albumin gene according to the following protocol. All samples and standards should be run in technical duplicates as a minimum.

Table 3
Table guide for preparing Maxima qPCR master mix

8.5 μL Maxima qPCR master mix or equivalent
0.17 μL forward primer (10 μM)
0.17 μL reverse primer (10 μM)
0.17 μL probe (10 μM)

6. Place 9 μL of the master mix into the appropriate number of wells in a 96-well optical plate.
7. Add 8 μL DNA sample to each of the appropriate wells.
8. Seal the plate with optical film.
9. Centrifuge the plate at $200 \times g$ for 1 min at room temperature to ensure that all liquid is at the bottom of the wells.
10. Place the 96-well plate in the real-time PCR machine and run the appropriate program depending on the fluorochromes (FAM) and quencher (BHQ) used in the TaqMan probes. Run the amplification reaction as follows: Preincubation for 10 mins at 95 °C followed by 45 cycles: 15 s at 95 °C, 1 min at 60 °C.
11. Analyze the amplification using the standard curve method.
12. Calculate for each sample the WPRE copy number per cell using the following formula:

$$\text{Vector copy number (per cell)} = \frac{\text{Quantity mean of WPRE sequence}}{\text{Quantity mean of albumin gene}} \times 2$$

(see **Note 13**).

13. Calculate the transduction titer using the formula below.

$$\text{Titer} \left(\frac{\text{Integrating particles}}{\text{mL}} \right) = \frac{\text{Number of recipient cells (counted at day 2)} \times \text{Vector copy number per cell}}{\text{Volume of virus in mL}}$$

3.5 Titer Determination of Lentiviral Vectors by Quantification of Colony-Forming Units (CFU)

Another way to determine the number of infectious particles in a lentiviral vector batch is to do a colony-forming units (CFU) assay. Obviously, this method is not an option in suspension cells, but provides a less work-intensive method of assessing the transductional titer on adherent cells. This assay requires that the transfer plasmid contains a selection cassette, which is inserted into the genome of the target cells after lentiviral transduction. If your lentiviral transfer plasmid does not contain a selection cassette, measurement of titer should be performed with qPCR.

1. Day 1, seed 1×10^5 target cells per well in a 6-well dish or in another culture format. Seed one well for each dilution of the vector batch and include an extra well for a non-transduced control for the drug selection.

2. Day 2, transduce the target cells with serial tenfold dilutions of the lentiviral vector batch. If the target cells can tolerate polybrene, it is possible to supplement with 8 $\mu\text{g}/\text{mL}$ polybrene (final concentration) to increase the transduction efficiency. Leave one well un-transduced.
3. Day 3, change medium on the target cells.
4. Day 4, start selection for drug-resistant cells at a concentration suitable for the target cells (*see Note 14*).
5. When colonies have appeared (typically after 4–6 days depending on the cell type), remove the medium from the wells.
6. Wash the wells with PBS.
7. Add an appropriate volume of 0.6% methylene blue solution to each well.
8. Incubate at room temperature for 15 min.
9. Wash the wells with H_2O and leave to air dry.
10. Count the number of colonies manually and calculate the CFU per mL using the following formula:

$$\frac{\text{CFU}}{\text{mL}} = \frac{\text{Number of formed colonies}}{\text{Volume of virus – containing supernatant added in mL.}}$$

3.6 Generation of Knockout Cell Lines Using CRISPR/Cas9

The lab of Feng Zhang has developed two lentiviral systems for easy and efficient delivery of the CRISPR/Cas9 system to mammalian cells: a one-vector system plentiCRISPRv2 encoding all elements necessary for genome editing in one vector, and a two-vector system, by which SpCas9 and the sgRNA are delivered on two separate lentiviral vector constructs (plasmids referred to as plentiCas9-Blast and plentiGuide-Puro, respectively). Both systems allow you to clone-in your 20-nucleotide sgRNA sequence of interest in the context of a lentiviral vector plasmid, allowing incorporation of derived vector RNA into lentiviral particles. These vectors are not the only available vectors for lentiviral delivery of the CRISPR/Cas9 system, but the following protocols are based on our experience with these vectors.

The protocols below describe the generation of knockout cell lines based on 2 to 3 weeks of drug selection for cells carrying one or more integrated copies of the vector. However, for some cell types, which can easily be transduced (e.g., HEK293 cells), high indel frequencies can be achieved by transferring vector-containing supernatant to the recipient cells. Hence, depending on the cell type, indel frequencies above 90% can be achieved 5 days post transduction without selection (Fig. 2a). In the shown example, indel rates of 50% were achieved in normal human dermal fibroblasts (NHDF) after 5 days (Fig. 2b). In our hands, we rarely detect indels within the first day after transduction, but observe the largest

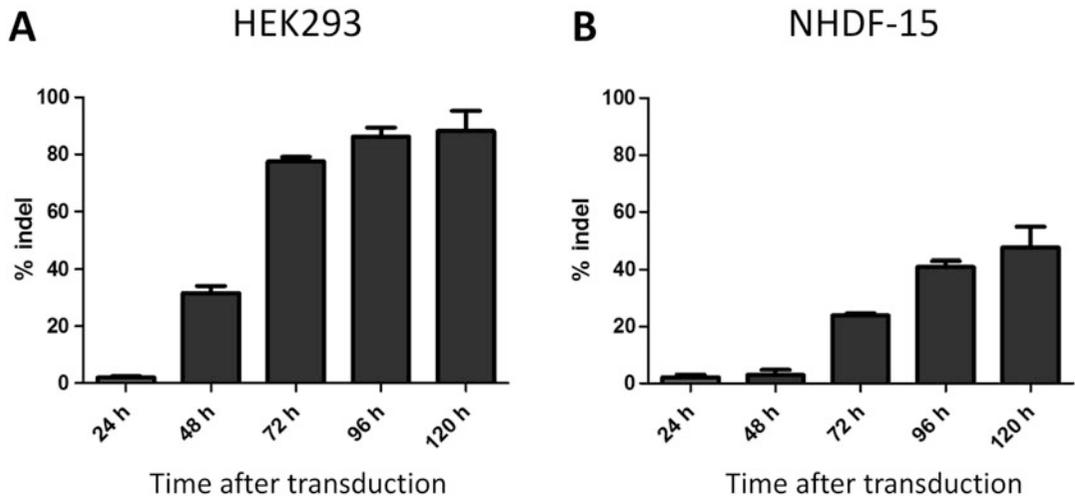


Fig. 2 Gene disruption rates in HEK293 cells and normal human dermal fibroblasts over a period of 5 days. HEK293 (a) and NHDF-15 (b) cells were transduced with LV/lentiCRISPRv2-sgRNA encoding a sgRNA targeting the *SERPING1* gene. 24 h after transduction the virus-containing medium was removed and the cells were harvested in 24-h intervals, and the indel frequencies were determined by TIDE analysis. Data are depicted as mean + SEM

increase in indel frequency between 48 and 72 h after lentiviral delivery of the CRISPR/Cas9 system (Fig. 2a, b). Thus, under conditions allowing high levels of transduction, the majority of the SpCas9-mediated cleavage occurs within this time frame. Although the preferred methodology may obviously depend on the application, we recommend carrying out drug selection for the time indicated in the protocols. This ensures depletion of non-transduced cells and allows use of lower amounts of vector, leading to reduced risks of integrating more than one lentiviral vector in transduced cells.

3.7 Generation of Knockout Cell Lines Using the One-Vector CRISPR/Cas9 System

For the generation of knockout cell lines using lentiviral delivery of the one-vector CRISPR/Cas9 system, clone the sgRNA sequence of interest into the plentiCRISPRv2 [11] plasmid (resulting in plentiCRISPRv2-sgRNA) and prepare a batch of lentiviral vectors (designated LV/lentiCRISPRv2-sgRNA) according to the guidelines in Subheading 3.1 using plentiCRISPRv2-sgRNA as the transfer plasmid.

1. Day 1, seed $1-2 \times 10^5$ target cells per well in a 6-well dish or in another suitable culture format. Seed four wells in total (*see Note 15*).
2. Day 2, transduce cells in three of the wells with the LV/lentiCRISPRv2-sgRNA batch. If the target cells can tolerate polybrene, it is possible to supplement with $8 \mu\text{g}/\text{mL}$ polybrene (final concentration) to increase the transduction

efficiency. The fourth well serves as a control for the puromycin selection applied on day 4.

3. Day 3, change medium on the target cells.
4. Day 4, start selection with puromycin in all four wells at a concentration suitable for the target cells.
5. Renew puromycin-containing medium every second or third day, and maintain puromycin selection for 14 days. Passage the cells if needed. At the first passage after transduction, pool the three wells containing transduced cells. Expect the cells in the non-transduced well to die within the first 4–7 days.
6. Day 19, freeze down aliquots of the cell population for later studies and harvest gDNA for validation of CRISPR/Cas9 activity leading to gene disruption.

3.8 Generation of Knockout Cell Lines Using the Two-Vector CRISPR/Cas9 System

The generation of stable SpCas9-expressing cell lines can be accomplished with the plentiCas9-blast vector. Due to the nature of lentiviral infection, it is recommended to transduce at a low multiplicity of infection (MOI) of 0.1–0.3. The low MOI helps control the number of inserted copies of the SpCas9-encoding cassette per cell. This will produce a heterogeneous population of cells, ideally each with a single insertion, reducing the overall heterogeneity of the population in terms of SpCas9 expression levels.

Prior to creation of SpCas9-expressing cells, produce a batch of LV/lentiCas9-blast vectors and assess the titer (*see* Subheadings 3.1–3.5).

Proceed following the general steps described below.

1. Day 1, seed $2\text{--}5 \times 10^5$ cells per well in a 6-well dish or in another suitable culture format. Seed cells in 4 wells in total.
2. Day 2, transduce 3 wells with LV/lentiCas9-Blast using a viral particle load ensuring low MOI. The needed volume of the vector preparation is determined based on the specific titer in the relevant cell line.
3. Day 3, replace medium with medium containing blasticidin; the fourth well serves as a control for blasticidin selection.
4. From Day 4 to 17, once each well becomes confluent, pool cells from the different wells and proceed by culturing cells in a larger appropriate culture volume. Maintain cells under blasticidin selection and passage as required (*see* **Note 16**).
5. Day 18, freeze down several aliquots of the population and proceed to test the functionality of SpCas9 in the blasticidin-resistant population.
6. (Optional) To establish and characterize individual SpCas9-expressing clones, seed cells individually to start generating expanded cell clones.

3.9 Validation of SpCas9-Expressing Cell Lines and gRNA Transduction

After 2 weeks of blasticidin selection most cell lines should be completely depleted of blasticidin-sensitive cells. Since the blasticidin resistance gene is expressed by the same promoter as SpCas9 and the two genes are fused by a P2A-peptide sequence, any blasticidin-resistant cell should also contain SpCas9 protein. Assessing the functionality and efficiency of SpCas9 in the cells can be done by delivering a sgRNA to the cells (*see Note 17*). We recommend doing this by transducing the cells with LV/LentiGuide-Puro, a vector that contains a sgRNA scaffold as well as a selection marker. A high-activity sgRNA (here designated sgRNA^X) should be cloned into plentiGuide-Puro for this step.

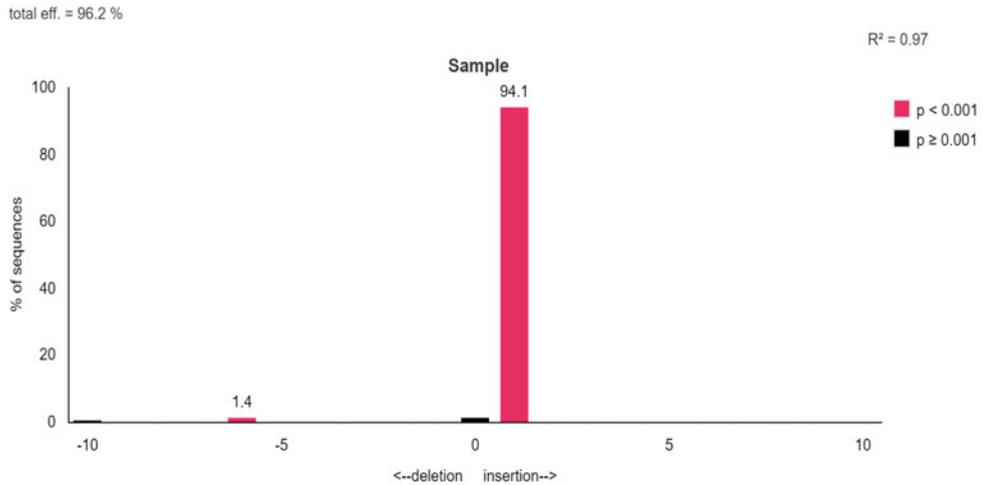
1. Day 1, seed $2\text{--}5 \times 10^5$ cells per well in a 6-well dish or in another suitable culture format. Seed 4 wells in total.
2. Day 2, transduce 3 wells with LV/LentiGuide-Puro-sgRNA^X.
3. Day 3, apply puromycin to all wells; the fourth well will serve as a control for puromycin selection.
4. Day 4–11, maintain puromycin selection and passage cells as usual (*see Note 18*).

Day 11. Harvest gDNA for PCR, sequencing and TIDE analysis, as described in Subheading 3.9 (*see Note 19*).

3.10 TIDE-Based Validation of CRISPR/Cas9-Mediated Gene Knockout

One way to validate CRISPR/Cas9-mediated editing in the cell lines created in Subheadings 3.6 and 3.7 is by tracking of indels by decomposition (TIDE) [12]. The TIDE software requires two chromatograms of the region targeted by the Cas9/sgRNA complex, (1) one from the cell lines generated in Subheadings 3.6 or 3.7 and (2) one from an untreated control population of the same cell line (or alternatively a population of cells transduced with a control vector). The TIDE software analyzes the two supplied chromatograms by using the height of each peak in the chromatogram to determine the relative abundance of aberrant nucleotides. The output is an indel frequency, which represents the fraction of alleles cleaved by Cas9 and subsequently repaired by NHEJ in a cell population.

1. Extract gDNA from the knockout population created in Subheadings 3.6 or 3.7 and from an untreated population of the same cell line using your preferred protocol.
2. PCR-amplify the sgRNA target region using the gDNA as template (*see Note 20*).
3. Run the PCR on a 1% agarose gel and purify the PCR band.
4. Sequence the PCR products; if you follow the recommended primer design, this can be done with your forward primer.

A**B**

Quality control - Aberrant sequence signal

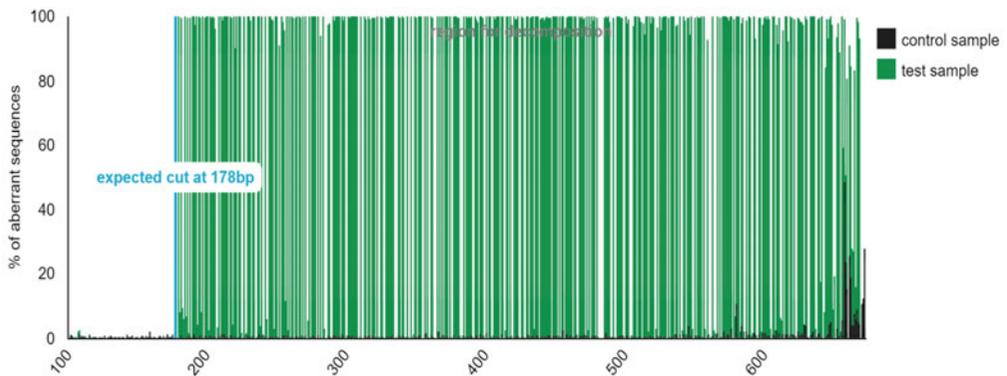


Fig. 3 TIDE analysis performed on primary fibroblasts after lentiviral delivery of the one-vector CRISPR/Cas9 system with a sgRNA targeting the *SERPING1* gene. An example of the output from a TIDE analysis showing an indel frequency of 96.2% with a one base pair insertion as the most common indel (a). The TIDE software also returns a quality control of the chromatograms uploaded depicted as % aberrant sequence with an illustration of the expected cut site (b)

5. Perform TIDE analysis using the TIDE software (<https://tide.nki.nl/>) or alternatively ICE analysis (<https://ice.synthego.com>).

An example of a TIDE analysis is shown in Fig. 3. This TIDE analysis is performed on primary fibroblasts after lentiviral delivery of the one-vector CRISPR/Cas9 system with a sgRNA targeting the *SERPING1* gene.

4 Notes

1. Production specifications are displayed in Table 4 below.
2. Transfection control: The transfection efficiency is absolutely crucial for production of high-titer vector preparations. If the transfer plasmid does not contain a fluorescent marker gene, it is recommended to include a transfection control. You can either transfect cells in a separate dish with a transfer plasmid expressing a fluorescent marker or include 1 µg of a standard expression plasmid (non-lentiviral) encoding a fluorescent marker in the transfection.
3. When carrying out a large-scale production, it is recommended to split the transfections into several rounds. Do not prepare more than 10 transfection mixtures at once, since incubation of the transfection mixture for prolonged periods of time will reduce the potency of the transfection.
4. If you are not ultracentrifuging the vector-containing supernatant, and the recipient cells grow in another medium than DMEM, change to that medium now. The cells are very confluent at this point, so be gentle when adding the fresh medium.

Table 4
List of production specifications used in this method

Reagent	Amount/volume
Day 1	
HEK293T cells per p10 dish	4×10^6
DMEM	10 mL
Day 2	
DMEM 1 h before transfection	10 mL
pRSV-REV	3 µg
pMD.2G	3.75 µg
pMDIg/p-RRE	13 µg
Transfer plasmid	13 µg
dH ₂ O in transfection mix	To a final volume of 450 µL
2 M CaCl ₂	50 µL
2×HBS	500 µL
Day 3	
DMEM	7–10 mL

5. If you are doing a second harvest on day 5, you can store the first harvest at 4 °C, pool it with the second harvest, and then make the aliquots or do the ultracentrifugation at day 5.
6. If you are not up-concentrating the vector preparation by ultracentrifugation, the crude vector-containing supernatant is now ready to use. To avoid repeated thawing and freezing, make aliquots of the crude preparation in appropriate sizes and store the aliquots at -70 to -80 °C until use.
7. When preparing the sucrose cushion aspirate 7 mL of the sucrose solution in a 5 mL serological pipette. Slowly release the sucrose solution into the tube and stop when you reach 3 mL. By doing this, you will have a 4 mL solution in each tube allowing you to keep track of the volume added and avoid disturbing the sucrose cushion by accidentally blowing air into it.
8. We typically add 85 µL to each tube resulting normally in a vector batch with an approximately 300-fold higher concentration of vectors relative to the crude prep.
9. If the recipient cells are suspension cells, you can combine day 1 and day 2 and do seeding and transduction in one step.
10. We normally go for 1000 µL, 200 µL, and 50 µL crude virus (vector-containing supernatant) or the equivalent volume of ultracentrifuged virus in a total volume suitable for your culture.
11. We use the pCCL-PGK-Puro-H1-MCS plasmid [13] for the WPRE standard curve, but any vector containing the WPRE element (like pLentiGuide-Puro or pLentiCRISPRv2) can be used. pAlbumin (acquired from Addgene) is used for the Albumin standard curve.
12. The conversion from grams to number of molecules of pAlbumin is carried out as depicted below. As the size of pAlbumin is 7539 bp plasmid and the estimated molecular weight of a base pair is 660 g/mol, the calculation is as follows:

$$\text{Molecular weight of pAlbumin} = 7539 \text{ bp} \times \frac{660 \frac{\text{g}}{\text{mol}}}{\text{bp}} = 5 \times 10^6 \frac{\text{g}}{\text{mol}}$$

If the concentration of the plasmid stock is 1 µg/µL then

$$\text{Molar concentration} = \frac{1 \times 10^{-6} \frac{\text{g}}{\mu\text{L}}}{5 \times 10^6 \frac{\text{g}}{\text{mol}}} = 2 \times 10^{-13} \frac{\text{mol}}{\mu\text{L}}$$

$$\text{and the number of molecules per } \mu\text{L} = 2 \times 10^{-13} \frac{\text{mol}}{\mu\text{L}} \times 6.02 \times 10^{23} \frac{\text{Molecules}}{\text{mol}} = 1.2 \times 10^{11} \frac{\text{molecules}}{\mu\text{L}}$$

Since $\frac{1.2 \times 10^{11}}{1.25 \times 10^7} = 9.6 \times 10^3$, the first sample of the standard curve for Albumin should be a 9.6×10^3 times dilution of the 1 µg/µL pAlbumin plasmid.

13. As two alleles of the albumin gene are assumed to be present in the genome, the relative quantity mean of WPRE and the albumin gene needs to be multiplied by 2 (indicated by “ $\times 2$ ” in the equation above).
14. Keep the cells in selection medium throughout the experiment. When the cells in the non-transduced control well have died off, and the colonies in transduced wells have reached a size visible with the naked eye, the colonies can now be stained and counted.
15. If the target cells are suspension cells, you can combine day 1 and day 2 and do seeding and transduction in one step.
16. Consider expanding your cells throughout the selection period in order to achieve a large cell population for freezing and storage. This will provide aliquots with minimum passage time ready for downstream applications.
17. Alternatively, Cas9 protein levels can be assayed by Western blot using anti-Flag antibodies. Obviously, this only confirms the expression and does not provide an assessment of functionality.
18. By using the two-vector system, the presence of Cas9 protein at a steady state in the cell allows for a more rapid generation of saturated knockout levels after sgRNAs delivery. Thus, one week of puromycin selection will suffice with the two-vector system.
19. With a typical Cas9-expressing population and delivery of a high-activity sgRNA species, indel formation should be near maximum after 1 week of puromycin selection. Expect an indel rate determined by TIDE above 75%.
20. Primer design: When designing primers for the PCR reaction, place the forward primer minimum 200–300 bp upstream of the expected Cas9 cut site to allow enough sequence for alignment with the control sequence. Aim for a PCR amplicon of 700–1000 bp and make sure that the primer set is not able to amplify another region in the genome of the same size.

Acknowledgments

The authors are grateful to Feng Zhang and colleagues for making lentiviral vector plasmids available through Addgene. JGM and members of the JGM lab are grateful for all support provided to projects involving lentiviral vector technologies, including grants from the Danish Council for Independent Research | Medical Sciences (grant DFF-4004-00220), The Lundbeck Foundation (grant R126-2012-12456), the Novo Nordisk Foundation, Aase og Ejnar Danielsens Fond, Agnes og Poul Friis Fond, Carl og Ellen

Hertz' legat til læge- og naturvidenskaben, Harboefonden, Arvid Nilssons Fond, Fonden af 1870, Else og Mogens Wedell-Wedellsborgs Fond, Einar Willumsens Mindelegat, Emil C. Hertz og hustru Inger Hertz's Fond, Holger Hjortenbergs og Hustru Dagmar Hjortenbergs Fond, Krista og Viggo Petersens Fond, and Andersen-Isted Fonden.

References

1. Naldini L, Blomer U, Gage FH, Trono D, Verma IM (1996) Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc Natl Acad Sci U S A* 93:11382–11388
2. Naldini L, Blomer U, Gallay P, Ory D, Mulligan R, Gage FH, Verma IM, Trono D (1996) In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* 272:263–267
3. Aiuti A, Biasco L, Scaramuzza S, Ferrua F, Cicalese MP, Baricordi C, Dionisio F, Calabria A, Giannelli S, Castiello MC et al (2013) Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science* 341:1233151
4. Biffi A, Montini E, Lorioli L, Cesani M, Fumagalli F, Plati T, Baldoli C, Martino S, Calabria A, Canale S et al (2013) Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science* 341:1233158
5. Cartier N, Hacein-Bey-Abina S, Bartholomae CC, Veres G, Schmidt M, Kutschera I, Vidaud M, Abel U, Dal-Cortivo L, Caccavelli L et al (2009) Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science* 326:818–823
6. De Ravin SS, Wu X, Moir S, Anaya-O'Brien S, Kwatema N, Littel P, Theobald N, Choi U, Su L, Marquesen M et al (2016) Lentiviral hematopoietic stem cell gene therapy for X-linked severe combined immunodeficiency. *Sci Transl Med* 8:335ra357
7. Dull T, Zufferey R, Kelly M, Mandel RJ, Nguyen M, Trono D, Naldini L (1998) A third-generation lentivirus vector with a conditional packaging system. *J Virol* 72:8463–8471
8. Miyoshi H, Blomer U, Takahashi M, Gage FH, Verma IM (1998) Development of a self-inactivating lentivirus vector. *J Virol* 72:8150–8157
9. Zufferey R, Dull T, Mandel RJ, Bukovsky A, Quiroz D, Naldini L, Trono D (1998) Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J Virol* 72:9873–9880
10. Xu K, Ma H, McCown TJ, Verma IM, Kafri T (2001) Generation of a stable cell line producing high-titer self-inactivating lentiviral vectors. *Mol Ther* 3:97–104
11. Sanjana NE, Shalem O, Zhang F (2014) Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 11:783–784
12. Brinkman EK, Chen T, Amendola M, van Steensel B (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42:e168
13. Bak RO, Stenderup K, Rosada C, Petersen LB, Moldt B, Dagnaes-Hansen F, Jakobsen M, Kamp S, Jensen TG, Dam TN et al (2011) Targeting of human interleukin-12B by small hairpin RNAs in xenografted psoriatic skin. *BMC Dermatol* 11:5



Rapid and Simple Screening of CRISPR Guide RNAs (gRNAs) in Cultured Cells Using Adeno-Associated Viral (AAV) Vectors

Julia Fakhiri, Manuela Nickl, and Dirk Grimm

Abstract

Genome editing reagents including the recently introduced CRISPR/Cas9 system have become established and widely used molecular tools to answer fundamental biological questions and to target and treat genetic diseases. The CRISPR system, originally derived from bacteria and archaea, can be delivered into cells using different techniques, comprising (1) transfection of mRNA or plasmid DNA, (2) electroporation of plasmid DNA or the Cas9 protein in a complex with a g(uide)RNA, or (3) use of nonviral or viral vectors. Among the latter, Adeno-associated viruses (AAVs) are particularly attractive owing to many favorable traits: (1) their apathogenicity and episomal persistence, (2) the ease of virus production and purification, (3) the safe handling under lowest biosafety level 1 conditions, and (4) the availability of numerous natural serotypes and synthetic capsid variants with distinct cell specificities. Here, we describe a fast and simple protocol for small-scale packaging of CRISPR/Cas9 components into AAV vectors. To showcase its potential, we employ this method for screening of gRNAs targeting the murine miR-122 locus in Hepa1–6 cells (using AAV serotype 6, AAV6) or the 5′LTR of the human immunodeficiency virus (HIV) in HeLaP4-NLtr cells (using a synthetic AAV9 variant). We furthermore provide a detailed protocol for large-scale production of purified AAV/CRISPR vector stocks that permit higher cleavage efficiencies *in vitro* and are suitable for direct *in vivo* applications.

Key words AAV, Adeno-associated virus, CRISPR, gRNA, GuideRNA, T7E1 assay

1 Introduction

In 2012–2013, a series of reports on the adaptation of the prokaryotic CRISPR (clustered regularly interspaced short palindromic repeats) system to mammalian cells have launched a fundamental revolution in biological and medical research [1–4]. Compared to other genome editing tools, such as zinc-finger nucleases (ZFN) or transcription activator-like effector nucleases (TALENs), CRISPR/Cas9 is very easy to customize as this system only requires two

Julia Fakhiri and Manuela Nickl contributed equally to this work.

components, the endonuclease Cas9 and a guide RNA (gRNA), and as one can redirect Cas9 by simply adapting the targeting sequence present in the gRNA [5]. Moreover, one can concomitantly target different genomic loci by multiplexing gRNAs, which facilitates efforts to decipher cellular networks and to answer complex biological questions [6–8]. Last but not least, by fusing a catalytically inactive Cas9 mutant to effector protein domains or other moieties, it is possible to modulate the epigenetic status of specific loci or to tag target DNAs with fluorescent markers, among many other applications [9–14].

Owing to this breadth and versatility, CRISPR has become the most popular tool for genome editing and is already widely used by academic and industrial laboratories around the world. In part, this success is due to the fact that an initial bottleneck in the application of CRISPR technology, i.e., a need of efficient, specific, and safe means for Cas9/gRNA delivery into cells, has recently been overcome by others and us. Most notable and in the center of this article is that the CRISPR machinery can be packaged into, and delivered by, recombinant Adeno-associated viruses (AAV), which permits robust genome editing in numerous different cell types and organs in small and large animals [15–18]. To this end, one can use either two separate AAV vectors (one for each component, Cas9 and gRNA) or all-in-one vectors that carry both parts on a single vector genome [19–27]. The choice of one over the other is highly dependent on the cell type, efficiency of transduction (i.e., gene delivery by viral vectors, which in the case of AAV is primarily determined by the capsid), and the size of the Cas9 variant used (e.g., *SpCas* is 4.1 kilobases [kb] long, versus *SaCas* with only 3.1 kb) [1, 2, 28]. The latter is critical because AAV vectors have an inherent limit for packaging of foreign DNA of around 4.8 kb, which—in the case of *SpCas*—leaves little space for regulatory sequences such as promoters [19]. Of note, the small size of gRNAs and associated promoters in a range of a few hundred base pairs enables their delivery by so-called “self-complementary” AAV vectors, which express more quickly than conventional single-stranded AAV vectors (but can only accommodate up to 2.4 kb of foreign DNA, which is enough for gRNA expression cassettes but not for Cas9) [29–31].

Here, we provide a protocol for rapid screening of gRNAs using a workflow combining small-scale AAV production (Fig. 1) with a “quick & dirty” T7 endonuclease I assay (Fig. 2). Using this two-step method, we routinely prescreen gRNAs in a variety of cell lines and primary cells, as exemplified here with two cell lines, HeLaP4-NLtr and Hepal–6 (Fig. 2). As shown below and as is typical for this method, we achieve potent target DNA cleavage of up to about 50%, which is already sufficient for many downstream applications.

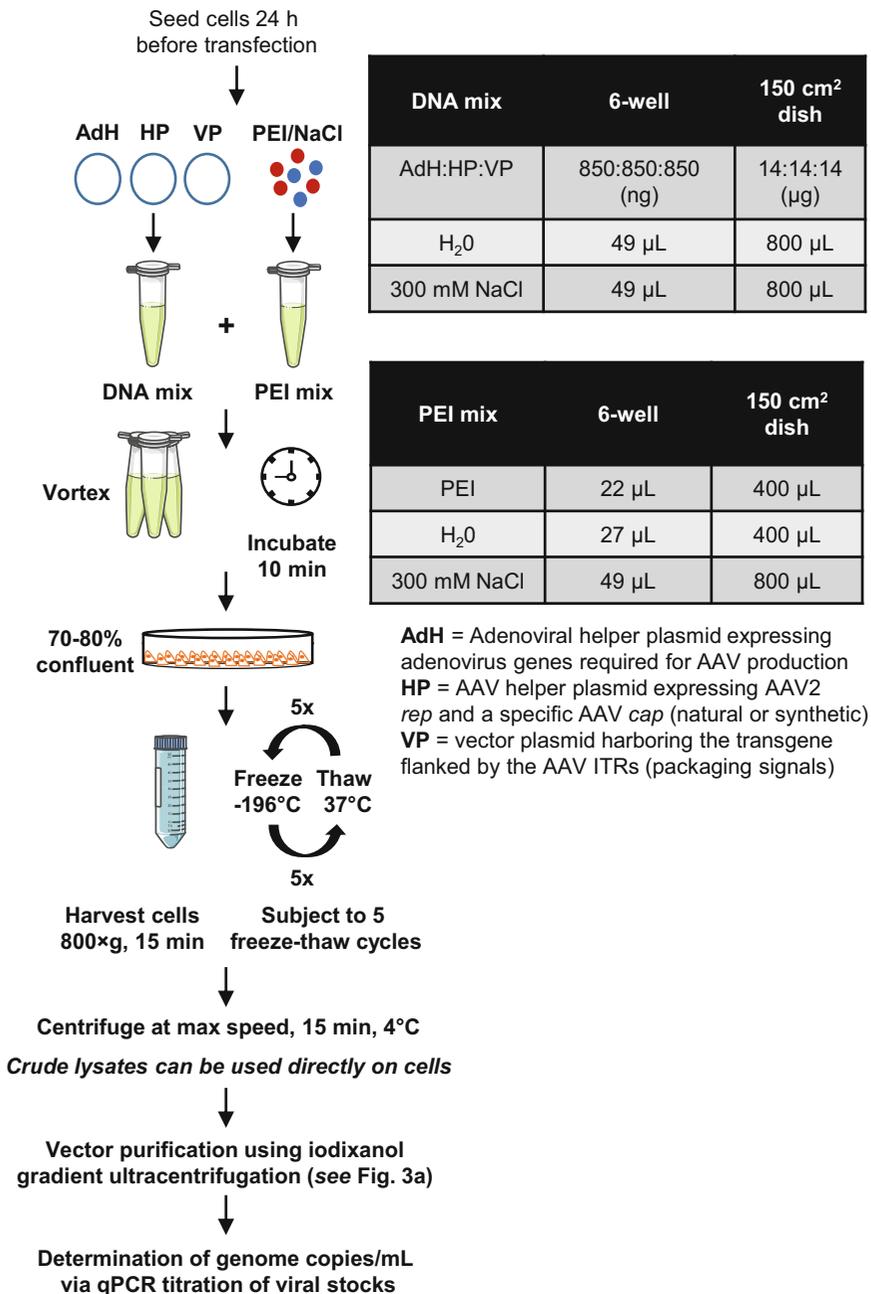


Fig. 1 Small- and large-scale PEI transfection scheme for the production of recombinant AAV vectors. *ITRs* inverted terminal repeats. This figure contains clipart from Servier Medical Art (<https://smart.servier.com/>)

To further improve AAV transduction and thereby enhance DNA editing efficiency, we moreover enclose a protocol for production of highly concentrated and purified viral stocks. To clean up viral “crude lysates,” ultracentrifugation is typically used for separation of full viral particles (i.e., capsids containing an AAV

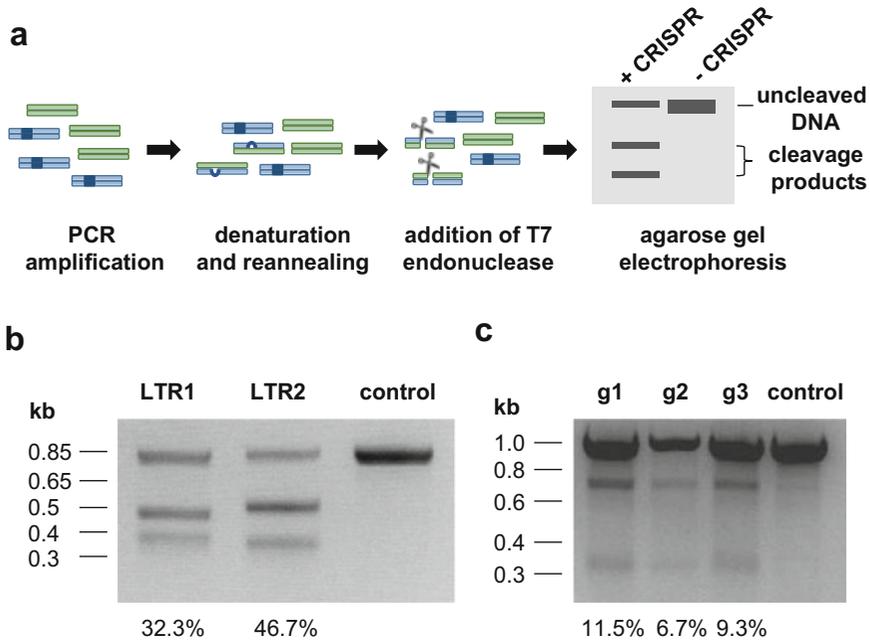


Fig. 2 T7 endonuclease I (T7E1) assay from cells transduced with crude lysates of AAV-CRISPR vectors. (a) Scheme showing the principle of the T7E1 assay. First, the region flanking the gRNA target site (depicted in dark blue) is PCR-amplified, which results in wild-type (blue) or mutated (green) amplicons. Then, the sample is heated and slowly cooled down to allow the formation of heteroduplexes. Finally, T7 endonuclease I is added to the sample which cuts the DNA strands at mismatches (depicted as bulges). The resulting DNA fragments are separated by agarose gel electrophoresis (a typical result is shown schematically). (b) T7E1 assay using *SpCas9* and gRNAs targeting the integrated HIV-1 5' LTR (human immunodeficiency virus 5' long terminal repeat) in HeLaP4-NLtr cells. The first two lanes show samples from cells treated with crude AAV lysates encoding Cas9 and two different gRNAs (LTR1, LTR2). The expected cleavage bands are visible (LTR1: 446, 332 bp; LTR2: 467, 311 bp). (c) T7E1 assay with gRNAs targeting the endogenous miR-122 locus in Hepa1–6 cells. The first three lanes show samples from cells treated with crude AAVs encoding *SaCas9* and three gRNAs that target the same region but differ in length (20, 21, and 22 nucleotides for g1, g2, and g3, respectively). The expected cleavage bands are visible (697, 317 bp). Control = sample from cells only transduced with crude AAV lysates encoding Cas9. Percentages of inserts and deletions (indels), calculated based on band intensities, are indicated below

vector genome) from undesired empty capsids based on their distinct densities. Two commonly used substances for preparing these gradients are iodixanol and cesium chloride. Here, we focus on iodixanol gradient centrifugation, a method that has gained wide popularity since its publication in 1999 [32] due to its (1) applicability to every AAV serotype, (2) speed as compared to cesium chloride purification, and (3) safety, as iodixanol is nontoxic in mammalian cells. Below, we exemplify its use to purify different commonly used AAV serotypes encoding distinct gRNAs, and we provide a representative overview of the titers that can be obtained (Fig. 3). The compatibility with many AAV capsid variants is important as these differ in the efficiency and specificity with which they

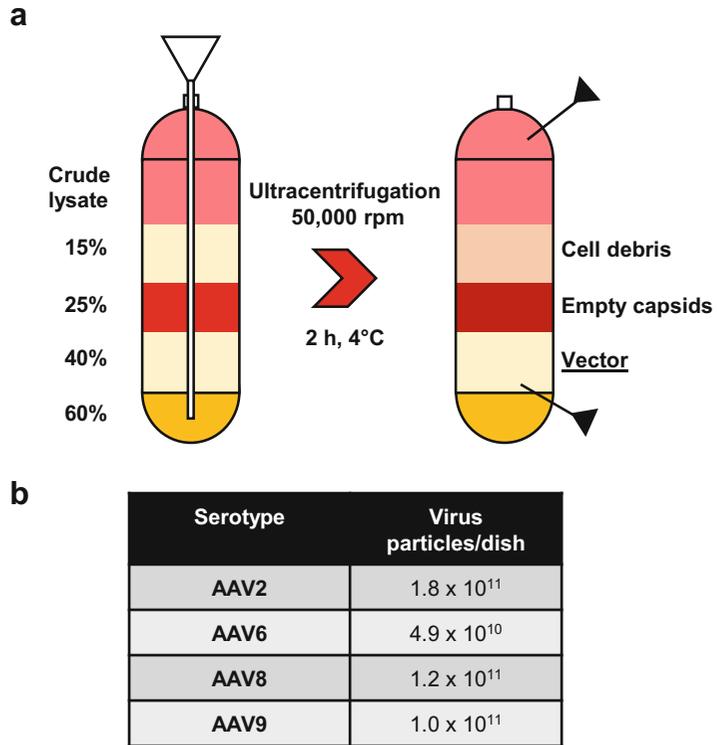


Fig. 3 Large-scale iodixanol purification of AAV vectors. **(a)** Schematic representation of an iodixanol gradient before and after ultracentrifugation. The purified vector particles containing a full-length genome accumulate in the 40% phase. **(b)** Packaging of gRNAs into different, commonly used AAV serotypes and the amount of viral particles per 150 cm² dish as estimated by qPCR analysis

transduce cells *in vitro* and *in vivo* [15, 33–36]. Therefore, a pre-screening of a particular cell line or cell type of interest for the best-performing AAV serotype is highly recommended (*see Note 1*).

Collectively, the detailed protocols presented in this chapter should provide readers with all the essential skills that are needed for the scalable production of basic AAV-CRISPR vectors and for their application in mammalian cells.

2 Materials

All solutions should be prepared with double-distilled water (ddH₂O) and autoclaved or filter-sterilized by passing through a 0.22 μm filter. Solutions should be stored at room temperature, unless indicated otherwise.

2.1 Transfection and Cell Harvest

1. Cell counting device or Neubauer chamber.
2. Beckman Avanti[®] centrifuge or equivalent.

3. Polyethylenimine (PEI) solution: weigh 0.323 g PEI (Polysciences) and add water to a final volume of 1 L. Stir for 1 h at 60 °C until the PEI completely dissolves. Transfer to 50 mL Falcon tubes and perform three freeze/thaw cycles at –80 °C/37 °C. Store PEI aliquots (e.g., in 50 mL aliquots) at –80 °C.
4. 300 mM NaCl: weigh 17.52 g NaCl and add water to a final volume of 1 L.
5. 1 M MgCl₂: weigh 203.3 g MgCl₂·6H₂O and dissolve in 800 mL water. Adjust the volume to 1 L with water.
6. Benzonase buffer: weigh 8.8 g NaCl, then add 2 mL 1 M MgCl₂, 50 mL 1 M Tris–HCl, and 900 mL water. Adjust pH to 8.5 with 1 M NaOH and fill up to 1 L with water.
7. Benzonase nuclease (Merck) 100,000 U.
8. Cell scraper.
9. 50 mL Falcon tubes or 500 mL Corning centrifugation tubes (for small- or large-scale AAV vector preparations, respectively).
10. 1 × 0.25% trypsin/EDTA.
11. Dulbecco's modified eagle's medium (DMEM).
12. Fetal bovine serum (FBS).
13. Penicillin/Streptomycin (Pen/Strep).

2.2 Iodixanol Gradient Centrifugation

1. Ultracentrifuge plus Beckman 70 Ti/70.1 Ti rotors or equivalent (*see Note 2*).
2. 3 or 5 mL syringes and 19- or 21-gauge (G) needles.
3. Beckman tube sealer.
4. 60% iodixanol stock solution (Opti-Prep, Progen Biotechnik GmbH). Keep sterile and protected from light.
5. Sterile Pasteur glass pipettes.
6. Beckman Quick-Seal centrifugation tubes: 14 × 89 mm or 25 × 89 mm.
7. 1 M PBS-MK-NaCl buffer: weigh 58.4 g NaCl, 263 mg MgCl₂, and 149.1 mg KCl. Add 1 × PBS to a final volume of 1 L.
8. 1 × PBS-MK buffer: weigh 263 mg MgCl₂ and 149.1 mg KCl. Add 1 × PBS to a final volume of 1 L.
9. Iodixanol working solutions:
 - 15% iodixanol solution: mix 12 mL 60% iodixanol and 36 mL 1 M PBS-MK-NaCl buffer.
 - 25% iodixanol solution: mix 20 mL 60% iodixanol, 28 mL 1 × PBS-MK buffer, and 120 μL phenol red.

40% iodixanol solution: mix 30 mL 60% iodixanol and 15 mL 1× PBS-MK buffer.

60% iodixanol solution: 60% iodixanol plus 2.5 μL/mL phenol red.

2.3 T7 Endonuclease I (T7E1) Assay

1. PCR thermocycler.
2. Hybridization oven with a rotating platform.
3. Dimethylsulfoxide (DMSO).
4. dNTP mix, 10 mM each.
5. Biozym LE Agarose.
6. Proteinase K.
7. High-fidelity polymerase (Phusion Hot Start Flex).
8. DirectPCR Lysis reagent (Cell) (Viagen Biotech).
9. T7 endonuclease (10,000 U/mL).
10. 50× Tris-Acetate-EDTA (TAE) stock solution: weigh 363 g Tris base and dissolve in 500 mL water. Carefully add 85.65 mL glacial acid and 750 mL 0.1 M EDTA. Adjust pH to 8.5 with 1 M NaOH and fill up to 1.5 L with water.
11. 2% agarose: weigh 14 g agarose and add 700 mL 1× TAE. Boil until completely dissolved. Cool down to 50 °C and add ethidium bromide to a final concentration of 0.5 μg/mL.

3 Methods

3.1 Small-Scale Production of AAV Vectors (Crude Lysates)

The following protocol describes the production of recombinant AAV vectors in a 6-well plate, yielding “crude lysates” that are ideally suited for rapid and parallel gRNA screening.

1. Seed 5×10^5 HEK293T cells per well in DMEM containing 10% FBS and 1% Pen/Strep one day before transfection.
2. Once the cells have reached a density of 70–80% (typically 24 h after plating), prepare the PEI-DNA transfection mix as depicted in Fig. 1 (*see Note 3*).
3. Add the transfection mix dropwise to the cells, gently shake the plate, and place at 37 °C. There is no need to change the growth medium after transfection.
4. After 72 h, remove the medium from the cells and add 1 mL PBS. Harvest the cells by pipetting up and down using a P1000 pipette.
5. Centrifuge at $800 \times g$ for 10 min at 4 °C.
6. Discard PBS and suspend the cells in 300–500 μL PBS.

7. Subject the cell suspension to five freeze-thaw cycles by alternating between liquid nitrogen (or a dry ice/ethanol bath) and 37 °C, to obtain a crude cell lysate.
8. Spin down at $16,100 \times g$ and 4 °C for 10 min to remove cell debris.
9. Transfer vector-containing supernatant to a 1.5 mL microcentrifuge tube and use directly in cells or freeze at -80 °C for long-term storage.

3.2 Large-Scale AAV Vector Production

This protocol and the following (Subheading 3.3) describe the steps for production of highly concentrated and pure AAV vector stocks suitable for direct in vitro application in cells as well as in vivo use.

1. Grow HEK293T cells in a sufficient number of 175 cm² flasks (plate 7×10^6 cells per flask) for subsequent plating of 150 cm² dishes for vector production. For example, five to six 175 cm² flasks will typically yield enough cells for 50 dishes. *See Note 4* for a calculation of the total amount of virus needed and then refer to Fig. 3b to estimate the number of dishes to be seeded.
2. Cell harvest: after 48 h, remove the medium completely and wash the cells in each flask with 15 mL PBS. Add 2 mL 0.25% trypsin/EDTA solution and incubate for 2 min at 37 °C. Stop the reaction by adding at least 8 mL medium and thoroughly suspend the cells by pipetting to obtain a single-cell suspension (clumps can decrease transfection efficiency). Collect all the harvested cells in one flask and count them using a Neubauer chamber or a similar device.
3. Seed 4×10^6 cells per 150 cm² plate in 22 mL medium. Pile ten plates and shake in multiple horizontal directions to achieve a homogenous cell distribution. Incubate the cells for 48 h at 37 °C (*see Note 5*).
4. Prepare PEI-DNA transfection mixes as depicted in Fig. 1. Add the mix dropwise to the cells and incubate for 72 h at 37 °C.
5. Harvest the cells using a cell scraper and transfer into 50 mL Falcon tubes or 500 mL centrifuge buckets (the latter are used for larger volumes).
6. In groups of 10, wash the plates once with 15 mL PBS (transfer the PBS from one plate to the other) and add to the harvested cells.
7. Centrifuge in a Beckman Avanti[®] centrifuge at $800 \times g$ for 15 min.
8. Discard medium completely and suspend the cells in 15 mL PBS (pool multiple cell pellets if necessary).

9. Repeat centrifugation at $800 \times g$ for 15 min and then discard supernatant.
10. Resuspend the cell pellet in 5 mL benzonase buffer for small gradients (up to 10 plates of cells) or in 20 mL for larger gradients. The cell suspension can be stored at $-80\text{ }^{\circ}\text{C}$.
11. Vortex well and subject to five freeze-thaw cycles by alternating between liquid nitrogen (or dry ice/ethanol bath) and $37\text{ }^{\circ}\text{C}$, to obtain a crude cell lysate.
12. To remove plasmid DNA, add benzonase (50 U/mL) and incubate at $37\text{ }^{\circ}\text{C}$ for 1 h, while inverting the tubes every 15 min.
13. Centrifuge the resulting “crude lysate” twice for 15 min at $4000 \times g$ and $4\text{ }^{\circ}\text{C}$, and always save the supernatant. Progress directly to iodixanol gradient centrifugation or freeze at $-20\text{ }^{\circ}\text{C}$.

3.3 Iodixanol Gradient Purification

1. Arrange Beckman ultracentrifugation tubes (14×89 mm for small gradients/ 25×89 mm for large gradients) in a holder and put a sterile Pasteur pipette into each.
2. Transfer the vector cell lysate (5 or 20 mL for small or large gradients, respectively) through the Pasteur pipette into the tube, using a 5 or 10 mL pipette (*see* Fig. 3a).
3. Add each of the different iodixanol solutions in the following order using a 2 mL plastic pipette extended with a P1000 tip, taking care to avoid air bubbles at any stage (*see* Note 6).
 - (a) 1.5 mL (small) or 7 mL (large) 15% iodixanol solution.
 - (b) 1.5 mL (small) or 5 mL (large) 25% iodixanol solution.
 - (c) 1.5 mL (small) or 4 mL (large) 40% iodixanol solution.
 - (d) 1.5 mL (small) or 4 mL (large) 60% iodixanol solution.
4. Carefully add benzonase buffer using a 3 mL syringe and 19G needle to top off the tube.
5. Close the tubes with metal caps and seal using a heating device to seal the tubes.
6. Centrifuge the tubes at $4\text{ }^{\circ}\text{C}$ for 2 h at $230,000 \times g$ in a Beckman 70.1 Ti rotor or at $257,000 \times g$ in a Beckman 70 Ti rotor. Set the acceleration to “max” and deceleration to “slow.”
7. Puncture the tube carefully below the 60–40% interface with a 21G needle attached to a 5 mL syringe (*see* Note 7).
8. Collect up to 1 mL per tube, taking care not to touch the 40–25% interface.
9. Repeat for each tube. As iodixanol is nontoxic, the vectors can be directly added to cells. Optionally, they can be further

concentrated and the iodixanol can be exchanged to PBS [37, 38]. For long-term storage, keep the vectors at $-80\text{ }^{\circ}\text{C}$ in aliquots that are sufficiently small to avoid repeated freezing and thawing.

3.4 Assess Vector Titer

This protocol describes how to assess vector titers using quantitative (q)PCR analysis (SYBR green or probe-based) with appropriate primers binding in the transgene or AAV ITR region.

1. Take a $10\text{ }\mu\text{L}$ sample of the vector (or H_2O as negative control) and add $10\text{ }\mu\text{L}$ TE buffer and $20\text{ }\mu\text{L}$ 2 M NaOH, vortex well, and incubate at $55\text{ }^{\circ}\text{C}$ for 30 min. This step denatures the viral particles and releases the encapsidated DNA.
2. Neutralize the solution by adding $38\text{ }\mu\text{L}$ 1 M HCl and vortex well. At this step, the sample can be stored at $-80\text{ }^{\circ}\text{C}$.
3. Directly before qPCR analysis, bring the volume to 1 mL by adding $922\text{ }\mu\text{L}$ water. Prepare a 1:10 working dilution and measure the samples in triplicates (around $1.5\text{--}5\text{ }\mu\text{L}$ of sample and $10\text{--}15\text{ }\mu\text{L}$ of Master Mix per qPCR reaction).
4. To estimate the viral titer (genome copies/mL), a standard curve with known amounts of plasmid DNA is required. For this, prepare a stock containing 3.5×10^{11} molecules/mL of the AAV vector plasmid used for transfection during production. Use the following formula to determine the number of molecules per mL of your plasmid preparation:

$$\frac{\text{DNA concentration } \left(\frac{\text{g}}{\text{mL}}\right)}{660 \times \text{size of plasmid in bp}} \times 6.02 \times 10^{23}$$

660 = average molecular weight for one bp of DNA (in Da).

6.02×10^{23} = Avogadro's number (the number of molecules in one mole of plasmid DNA).

Prepare 1:10 dilutions ranging from 3.5×10^{11} down to 3.5×10^6 molecules/mL and measure the standards and samples in triplicates.

5. Run an appropriate cycling program on your qPCR instrument according to manufacturer's instructions for the used polymerase. To create the standard curve, plot the Ct values of the standard samples on the y-axis and the logarithm of the number of molecules per ml on the x-axis. Then use the linear equation of the standard curve and the Ct values of your AAV samples to determine the x-values. As your samples have been diluted 1:1000 with water after neutralization, multiply the determined x-values with 1000 to obtain the number of viral genomes per mL. If you titrated a single-stranded AAV construct, this value has to be multiplied by two, because the plasmid used to create the standard curve is double-stranded.

3.5 Transduction of Cells and Cell Lysis

This protocol describes the transduction of cells with two AAVs, the first one encoding the Cas9 endonuclease and the second one a single or multiple gRNA(s) against a desired target sequence.

1. Seed 2500 HeLaP4-NLtr cells or 8000 Hepa1-6 cells per well in 100 μL DMEM containing 10% FBS and 1% Pen/Strep in a 96-well plate (*see Note 8*).
2. Transduce the cells by adding 10 μL of each crude lysate or 2–20 μL of each purified vector (encoding Cas9 or gRNA) (*see Note 4*).
3. After 72 h, remove medium and wash cells with 150 μL 1 \times PBS per well. Dilute DirectPCR (Cell) reagent 1:2 with water and add 11 U/mL Proteinase K. Add 140 μL of this lysis buffer per well in the 96-well plate. Seal the plate with aluminum foil and shake in a hybridization oven with a rotating platform at 150 rpm for 12–16 h at 55 $^{\circ}\text{C}$.
4. Incubate the samples in a PCR cycler for 45 min at 85 $^{\circ}\text{C}$ to inactivate the Proteinase K. Proceed to PCR amplification (*see next chapter*) or freeze at -20°C .

3.6 PCR and “Quick & Dirty” T7E1 Assay

This protocol describes how to determine the target DNA cleavage efficiency after treatment with Cas9/gRNA-encoding AAV vectors using the T7 endonuclease I (T7E1) assay (Fig. 2a).

1. Prepare a 25 μL PCR mix containing 7.5 pmol of each primer, 5 pmol dNTPs, 3% DMSO, 1 \times HF buffer (for Phusion Hot Start polymerase), 0.5 μL Phusion Hot Start Flex, and 5 μL template DNA. Add water to a final volume of 25 μL (*see Note 9*).
2. Run the reaction as follows: 98 $^{\circ}\text{C}$ 10 min; 40 cycles of (1) denaturation (98 $^{\circ}\text{C}$, 10 s), (2) annealing (55–72 $^{\circ}\text{C}$, 15 s), (3) extension (72 $^{\circ}\text{C}$, 30 s), followed by a final extension at 72 $^{\circ}\text{C}$ for 10 min (*see Note 10*).
3. After the PCR, heat the sample for 5 min at 95 $^{\circ}\text{C}$ and let it cool down slowly to 25 $^{\circ}\text{C}$ in increments of -5°C and 5 min per step.
4. Add 5U T7 endonuclease I per sample and incubate for 20 min at 37 $^{\circ}\text{C}$.
5. Stop the reaction by placing the tubes on ice and run the samples immediately on a 2% agarose gel.
6. Visualize cleavage pattern (as shown in Fig. 2b, c) and quantify cleavage using the following formula [39, 40]:

$$\% \text{ DNA cleavage} = 100 \times \left(1 - \sqrt{1 - \text{fraction cleaved}} \right)$$

4 Notes

1. Complete AAV expression systems including the adenoviral helper, an AAV vector genome encoding a fluorescent reporter and one AAV helper plasmid (for production of one AAV serotype) are commercially available as kits (e.g., Stratagene). Further AAV helper plasmids encoding different *cap* genes including synthetic variants can be purchased in addition (e.g., Cell Biolabs), with no need to order one separate kit for each. Ready-to-package AAV-CRISPR plasmids are moreover available from the authors or through Addgene (for example, Addgene plasmid #61591 is an all-in-one AAV-CRISPR construct harboring *SaCas9* and a site for cloning of a gRNA). For fast screening of AAV variants on your cell type of interest, package a fluorescent reporter gene such as GFP (for example, pscAAV-GFP, Addgene #32396) into the different AAV serotypes, using the protocol provided in Subheading 3.1 (see also Fig. 1 for a transfection scheme). Incubate the cells with 10 μ L of each of the different AAV_(x)-GFP vectors (*x* denotes the capsid variant) and estimate the percentage of GFP-positive cells after 48–72 h by fluorescence microscopy or FACS analysis.
2. By changing the run length, other rotors can be used. To determine the optimal run length that results in comparable sedimentation of particles, check the *K*-factor of the rotor and apply the following formula [41]:

$$\frac{t1}{k1} = \frac{t2}{k2}$$

*t*1, *t*2 = duration of the run for rotor 1 and 2, respectively

*k*1, *k*2 = *K*-factor for rotor 1 and 2, respectively

3. It is highly recommended to package a reporter gene (such as pscAAV-GFP, Addgene #32396) in parallel as a positive control, especially for researchers new to the field, to obtain a feeling about the efficiency of the AAV vector system in a given setting.
4. For purified vectors, a MOI of 1×10^5 of each of the two vectors (Cas9/gRNA) is usually sufficient to obtain good editing efficiencies. MOI = multiplicity of infection, i.e., the amount of viral particles per cell. For example, to achieve a MOI of 1×10^5 in 1×10^4 cells, a total of 1×10^9 viral particles is needed. Note that these numbers are estimates, and that optimal MOIs are dependent on a variety of parameters including the efficiency of a selected AAV capsid on a given cell type. Hence, these parameters need to be fine-tuned experimentally in order to maximize subsequent gene editing efficiencies.

5. To save time, the transfection can be performed already after 24 h. Therefore, to obtain an optimal cell density, $6\text{--}7 \times 10^6$ cells per 150 cm^2 dish should be seeded.
6. Iodixanol is a viscous material, especially the 60% stock solution. Therefore, the use of a 2 mL plastic pipette helps to form a continuous, smooth flow of iodixanol and avoids clogging of the Pasteur pipette with air. Air bubbles are tricky to remove and may destroy the gradient. If this problem occurs, trapped air can be removed by carefully tapping the Pasteur pipette from the top. In case this does not help, applying more pressure by sticking and tapping a second Pasteur pipette in the first should solve the problem. At the end, close the opening of the pipette with a thumb and carefully pull it out of the gradient. Remove residual air bubbles by tapping the tube at the table. This last step is important because too much entrapped air can deform and destroy the tubes during ultracentrifugation.
7. To puncture the ultracentrifugation tubes, mount them first in an appropriate ring stand, avoiding too much pressure. Then, puncture the upper part with a 21G needle to permit air inflow. Take a second 21G needle and insert it slightly below the 40–60% interface in an approximately 45° angle (as shown in Fig. 3a). Gently apply pressure on the needle that slowly increases, turn the bevel up and collect around 1 to 1.2 mL. When reaching the 25–40% interface, it is advisable to turn the bevel down again, to minimize the risk of collecting any empty capsids or contaminants.
8. The optimal number of cells seeded per well is dependent on the cell size and time-frame of experiment and should therefore be determined experimentally. As an optional step, the transduction can be performed twice to increase editing efficiency. For this, change medium two days after the first transduction and transduce a second time. After 48 h, continue with **step 2**.
9. Primers used in the T7E1 assay should ideally amplify a product of around 700 to 2500 bp. If the PCR product is very small, the even smaller T7E1 cleavage products will be hard to visualize on the gel. Furthermore, the gRNA target site must not be positioned exactly in the middle between the primers, as the resulting cleavage products would have the same size and therefore could not be separated on the agarose gel. For sufficient separation, cleavage products should differ by at least 100 bp in length.
10. The “quick & dirty” T7E1 assay is only applicable to validated PCRs that yield a single, defined product. If side products are observed, separate the bands through agarose gel electrophoresis and extract the correct DNA band from the gel. Also notable is that the rapid T7E1 assay assumes comparable

amplification in all reaction tubes, which may not always be the case. Thus, for a more quantitative estimation and comparison of gRNAs, the same amounts (200–500 ng) of input DNA should be used. For this purpose, purification of PCR products is necessary using standard protocols such as phenol/chloroform extraction, agarose gel electrophoresis, or spin columns.

Acknowledgments

J.F. and D.G. are grateful for funding from the Cystic Fibrosis Foundation Therapeutics (CFFT, grant number GRIMM15XX0) as well as from the Heidelberg Biosciences International Graduate School HBIGS at Heidelberg University. D.G. acknowledges funding by the Collaborative Research Center SFB1129 (German Research Foundation [DFG], project TP2), the Cluster of Excellence CellNetworks (DFG; EXC81), and the Transregional DFG Collaborative Research Center TRR179 (project TP18, Projekt-nummer 272983813). M.N. and D.G. are grateful for additional funding from the German Center for Infection Research (DZIF; TTU-HIV 04.803). Moreover, M.N. appreciates funding from the Heidelberg Innovation fund FRONTIER (project “A CRISPR way to eradicate viral pathogens”).

References

1. Cong L, Ran FA, Cox D et al (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823. <https://doi.org/10.1126/science.1231143>
2. Mali P, Yang L, Esvelt KM et al (2013) RNA-guided human genome engineering via Cas9. *Science* 339:823–826. <https://doi.org/10.1126/science.1232033>
3. Barrangou R, Doudna JA (2016) Applications of CRISPR technologies in research and beyond. *Nat Biotechnol* 34:933–941. <https://doi.org/10.1038/nbt.3659>
4. 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* 337:816–821. <https://doi.org/10.1126/science.1225829>
5. Boettcher M, McManus MT (2015) Choosing the right tool for the job: RNAi, TALEN, or CRISPR. *Mol Cell* 58:575–585. <https://doi.org/10.1016/j.molcel.2015.04.028>
6. Doudna JA, Charpentier E (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346:1258096. <https://doi.org/10.1126/science.1258096>
7. Hsu PD, Lander ES, Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157:1262–1278. <https://doi.org/10.1016/j.cell.2014.05.010>
8. Porteus M (2016) Genome editing: a new approach to human therapeutics. *Annu Rev Pharmacol Toxicol* 56:163–190. <https://doi.org/10.1146/annurev-pharmtox-010814-124454>
9. Dominguez AA, Lim WA, Qi LS (2016) Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Biol* 17:5–15. <https://doi.org/10.1038/nrm.2015.2>
10. Gilbert LA, Horlbeck MA, Adamson B et al (2014) Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* 159:647–661. <https://doi.org/10.1016/j.cell.2014.09.029>
11. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152:1173–1183. <https://doi.org/10.1016/j.cell.2013.02.022>

12. Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD (2014) A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* 159:635–646. <https://doi.org/10.1016/j.cell.2014.09.039>
13. Kearns NA, Pham H, Tabak B, Genga RM, Silverstein NJ, Garber M, Maehr R (2015) Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat Methods* 12:401–403. <https://doi.org/10.1038/nmeth.3325>
14. Thakore PI, D'Ippolito AM, Song L et al (2015) Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Methods* 12:1143–1149. <https://doi.org/10.1038/nmeth.3630>
15. Grimm D, Zolotukhin S (2015) E Pluribus Unum: 50 years of research, millions of viruses, and one goal--tailored acceleration of AAV evolution. *Mol Ther* 23:1819–1831. <https://doi.org/10.1038/mt.2015.173>
16. Schmidt F, Grimm D (2015) CRISPR genome engineering and viral gene delivery: a case of mutual attraction. *Biotechnol J* 10:258–272. <https://doi.org/10.1002/biot.201400529>
17. Epstein BE, Schaffer DV (2017) Combining engineered nucleases with adeno-associated viral vectors for therapeutic gene editing. *Adv Exp Med Biol* 1016:29–42. https://doi.org/10.1007/978-3-319-63904-8_2
18. Schmelas C, Grimm D (2018) Split Cas9, Not hairs - advancing the therapeutic index of CRISPR Technology. *Biotechnol J* 13(9): e1700432. <https://doi.org/10.1002/biot.201700432>
19. Senis E, Fatouros C, Grosse S et al (2014) CRISPR/Cas9-mediated genome engineering: an adeno-associated viral (AAV) vector toolbox. *Biotechnol J* 9:1402–1412. <https://doi.org/10.1002/biot.201400046>
20. Swiech L, Heidenreich M, Banerjee A et al (2015) In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat Biotechnol* 33:102–106. <https://doi.org/10.1038/nbt.3055>
21. Chew WL, Tabebordbar M, Cheng JK et al (2016) A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat Methods* 13:868–874. <https://doi.org/10.1038/nmeth.3993>
22. Long C, Amoasii L, Mireault AA et al (2016) Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* 351:400–403. <https://doi.org/10.1126/science.aad5725>
23. Nelson CE, Hakim CH, Ousterout DG et al (2016) In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* 351:403–407. <https://doi.org/10.1126/science.aad5143>
24. Tabebordbar M, Zhu K, Cheng JKW et al (2016) In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* 351:407–411. <https://doi.org/10.1126/science.aad5177>
25. Chow RD, Guzman CD, Wang G et al (2017) AAV-mediated direct in vivo CRISPR screen identifies functional suppressors in glioblastoma. *Nat Neurosci* 20:1329–1341. <https://doi.org/10.1038/nn.4620>
26. Wang G, Chow RD, Ye L et al (2018) Mapping a functional cancer genome atlas of tumor suppressors in mouse liver using AAV-CRISPR-mediated direct in vivo screening. *Sci Adv* 4: eaao5508. <https://doi.org/10.1126/sciadv.aao5508>
27. Platt RJ, Chen S, Zhou Y et al (2014) CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* 159:440–455. <https://doi.org/10.1016/j.cell.2014.09.014>
28. Ran FA, Cong L, Yan WX et al (2015) In vivo genome editing using Staphylococcus aureus Cas9. *Nature* 520:186–191. <https://doi.org/10.1038/nature14299>
29. Grimm D, Streetz KL, Jopling CL et al (2006) Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* 441:537–541. <https://doi.org/10.1038/nature04791>
30. McCarty DM, Fu H, Monahan PE, Toulson CE, Naik P, Samulski RJ (2003) Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction in vivo. *Gene Ther* 10:2112–2118. <https://doi.org/10.1038/sj.gt.3302134>
31. McCarty DM, Monahan PE, Samulski RJ (2001) Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Ther* 8:1248–1254. <https://doi.org/10.1038/sj.gt.3301514>
32. Zolotukhin S, Byrne BJ, Mason E et al (1999) Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield. *Gene Ther* 6:973–985. <https://doi.org/10.1038/sj.gt.3300938>
33. Kotterman MA, Schaffer DV (2014) Engineering adeno-associated viruses for clinical gene therapy. *Nat Rev Genet* 15:445–451. <https://doi.org/10.1038/nrg3742>

34. Zincarelli C, Soltys S, Rengo G, Rabinowitz JE (2008) Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. *Mol Ther* 16:1073–1080. <https://doi.org/10.1038/mt.2008.76>
35. Weinmann J, Grimm D (2017) Next-generation AAV vectors for clinical use: an ever-accelerating race. *Virus Genes* 53 (5):707–713. <https://doi.org/10.1007/s11262-017-1502-7>
36. Ellis BL, Hirsch ML, Barker JC, Connelly JP, Steininger RJ 3rd, Porteus MH (2013) A survey of ex vivo/in vitro transduction efficiency of mammalian primary cells and cell lines with Nine natural adeno-associated virus (AAV1-9) and one engineered adeno-associated virus serotype. *Virology* 453:107–117. <https://doi.org/10.1016/j.virus.2013.08.021>
37. Lock M, Alvira M, Vandenberghe LH, Samanta A, Toelen J, Debyser Z, Wilson JM (2010) Rapid, simple, and versatile manufacturing of recombinant adeno-associated viral vectors at scale. *Hum Gene Ther* 21:1259–1271. <https://doi.org/10.1089/hum.2010.055>
38. Strobel B, Miller FD, Rist W, Lamla T (2015) Comparative analysis of cesium chloride- and iodixanol-based purification of recombinant adeno-associated viral vectors for preclinical applications. *Hum Gene Ther Methods* 26:147–157. <https://doi.org/10.1089/hgtb.2015.051>
39. 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. https://doi.org/10.1007/978-1-60761-753-2_15
40. Ran FA, Hsu PD, Lin CY et al (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154:1380–1389. <https://doi.org/10.1016/j.cell.2013.08.021>
41. BeckmanCoulter (2014) Using K-factor to compare rotor efficiency. Technical guide CENT-66APP0. www.beckmancoulter.com



Electroporation-Based CRISPR/Cas9 Gene Editing Using Cas9 Protein and Chemically Modified sgRNAs

Anders Laustsen and Rasmus O. Bak

Abstract

CRISPR/Cas9 is an effective and easy-to-use tool for editing the genome of many human cancer cell lines. However, in some hard-to-transfect cell lines and primary cells, gene editing is more challenging. This protocol details an electroporation-based protocol for the delivery of Cas9 protein from *Streptococcus pyogenes* complexed with chemically modified sgRNAs. We have found this protocol to work very efficiently in numerous cell lines and primary cells that are difficult to transfect by conventional chemical-based transfection methods.

Key words CRISPR/Cas9, Gene editing, Chemically modified sgRNAs, Electroporation, Nucleoporation, Gene knockout

1 Introduction

Upon the discovery of the CRISPR/Cas9 gene editing technology, numerous publications highlighted the great therapeutic potential of the technique [1]. However, most of these initial publications employed plasmid delivery in standard easy-to-transfect cancer cell lines, and it quickly became evident that gene editing in more “exotic” cell lines and primary cells was more challenging [2]. Many cancer cell lines have been adapted for laboratory use, including selection for high transfection rates and high transgene expression levels. Consequently, these cell lines may lack parts of the innate immunological machinery for responding to foreign nucleic acids [3, 4]. In other cell lines or in primary cells where this machinery is intact, plasmid DNA may lead to a widespread interferon response and consequent low transgene expression and/or widespread toxicity. Similar to other gene editing technologies, it also quickly became evident for the CRISPR/Cas9 system, that prolonged expression from DNA-based vectors led to higher off-target activity [5, 6]. Luckily, Cas9 has proven much easier to produce as recombinant protein than, for example, zinc finger

nucleases and TALENs, which has addressed the need for a hit-and-run delivery approach where the nuclease acts only within a short time frame [6]. Another limiting factor of the system is the stability of the 100 nt sgRNAs, which are not protected from nuclease-mediated degradation. However, we and others have shown that the use of synthetic and chemically modified sgRNAs can markedly enhance gene editing [5, 7, 8]. These sgRNAs carry modified nucleotides that render the sgRNAs resistant to nucleases. Synthetic and modified sgRNAs can be complexed with Cas9 protein to form a ribonucleoprotein (RNP) prior to cellular delivery. Combined with electroporation, where a short electrical pulse permeabilizes the cell membrane and enables very efficient uptake of macromolecules, Cas9 RNP constitutes a highly effective methodology for gene editing in hard-to-transfect cells [5, 9].

Here we outline a protocol for electroporation of *Streptococcus pyogenes* Cas9 RNP. We describe steps to optimize this protocol using GFP mRNA delivery as a proxy for Cas9 RNP delivery to identify optimal electroporation conditions. Using this protocol, we generally observe >50% gene editing efficiencies in a wide array of cell types.

2 Materials

2.1 CRISPR/Cas9 Reagents and GFP mRNA

1. Cas9 protein. *Streptococcus pyogenes* Cas9 protein can be purchased from numerous vendors, e.g., IDT, Thermo Fisher Scientific, or NEB. For decreased off-target activity, we recommend the HiFi Cas9 variant from IDT. Store at -20°C .
2. sgRNA. Synthetic chemically modified single guide RNAs can be purchased from Synthego, Thermo Fisher Scientific, TriLink, or Dharmacon. We recommend ordering with the following format with the target-specific nucleotides designated with N, and three nucleotides at both termini modified with 2'-O-methyl 3'phosphorothioate (underlined):
NNNNNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUA-
GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAU-
CAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU-
U. Store at -80°C .
3. GFP mRNA. GFP mRNA is available from numerous commercial sources, e.g., TriLink Biotechnologies, OZ Biosciences, Miltenyi Biotec, and SBI. Alternatively, GFP mRNA can be produced by standard in vitro transcription methods. Store at -80°C .

2.2 Electroporation Devices and Reagents

1. Electroporation device, e.g., the 4D-Nucleofector System (Lonza), the Nucleofector 2b Device (Lonza), or the Neon Transfection System (Thermo Fisher Scientific).

2. Electroporation kits. Cell type-specific electroporation kits can be purchased for the two nucleofection devices (Lonza) or a single universal kit for all cell types can be purchased for the Neon Transfection System. These include electroporation cuvettes/tips and electroporation buffers. For the Lonza devices, homemade electroporation buffers can be used (*see step 3*) and can be combined with repeated use of the nucleofection cuvettes/strips after careful rinsing and sterilization. If using the Nucleofector Iib, generic cuvettes are available (VWR or Thermo Fisher Scientific).
3. Electroporation buffers. Electroporation buffers are included in the electroporation kits mentioned above. For the Lonza nucleofectors, we have found for some cell types that different solutions/buffers work as well as the suggested buffers. These include Opti-MEM, Solution 1 M (as described in [10]), and K562 electroporation solution (as described in [11]). As described in the methods section, we suggest trying several buffers to optimize the electroporation procedure.

3 Methods

We here outline electroporation procedures for the Lonza 4D-Nucleofector System using the Nucleocuvette strip format. For other formats or electroporation devices, please follow the manufacturer's suggestions. Before identifying the optimal electroporation protocol and conditions, we recommend doing a prior screening of a range of sgRNAs for your target genetic locus. This can be performed by plasmid-based expression of Cas9 and the different sgRNAs (e.g., using pX330) as detailed by Ran et al. [12]. Targeted cleavage can subsequently be validated in an easy-to-transfect cell line such as HEK293. Alternatively, the sgRNAs to be screened can be purchased as "ready-to-transfect" synthetic sgRNAs from commercial vendors and screened directly in the relevant cell type. As a positive control sgRNA we recommend using a CCR5-targeting sgRNA, which we have found to be highly potent (CCR5 sgRNA target site: 5'-GCAGCATAGTGAGCCC AGAA-3') [11].

3.1 Identification of Optimal Electroporation Settings Using GFP mRNA

1. Prepare an electroporation parameter matrix according to Fig. 1a. Consult the electroporation device manual, company website, technical specialists, and previous publications for suggestions on which electroporation kit, buffer, and program to use. Optionally, include the electroporation buffers mentioned above.
2. Prepare a cell population with high viability (>90%) and in logarithmic growth phase (*see Note 1*).

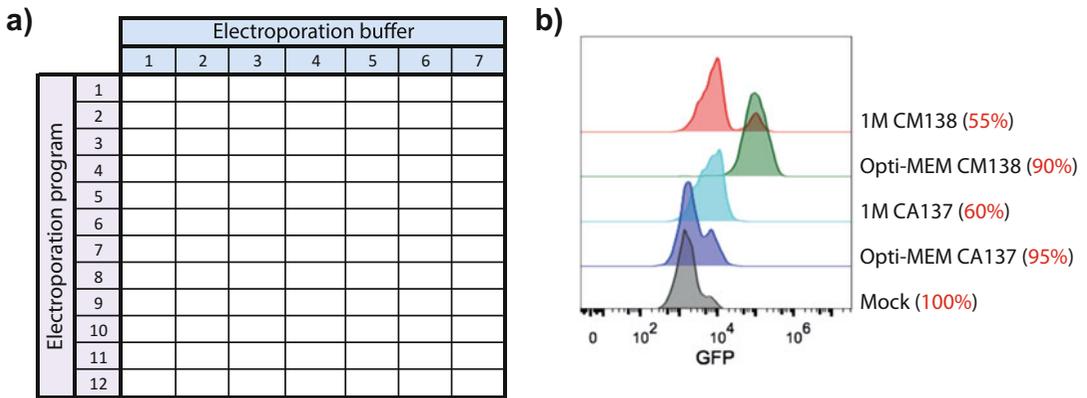


Fig. 1 Optimization of electroporation parameters. **(a)** Schematic representation of an electroporation parameter matrix used to evaluate different combinations of electroporation programs and buffers using GFP mRNA delivery that serves as a proxy for Cas9 RNP delivery efficiencies. **(b)** Example of electroporation optimization results in a hard-to-transfect cell line (PMDC05) using a 2×2 electroporation parameter matrix testing the combination of two different buffers and two different electroporation programs. Offset histograms show expression of GFP and red numbers indicate percentage of viable cells as measured 24 h post-electroporation

- Count the cells and distribute the desired cell number into one Eppendorf tube per electroporation buffer to be tested (*see Note 2*).
- Pellet cells by centrifugation at room temperature at $300 \times g$ for 5 min.
- Aspirate the supernatant and wash the cells by resuspending them in PBS. Pellet the cells by centrifugation at room temperature at $300 \times g$ for 5 min.
- Aspirate the supernatant and gently resuspend the cells in electroporation buffer by pipetting up and down (*see Note 3*). Use 20 μL of buffer per electroporation, i.e., 20 μL per 200,000 cells. The sample that will serve as an un-electroporated sample is then transferred directly into prewarmed culture media in an appropriate culture vessel. To the remaining cell suspension, add 1 μg of GFP mRNA per sample to be electroporated, i.e., 1 μg per 20 μL electroporation buffer (*see Note 4*). Mix by pipetting gently up and down.
- Distribute the cell + GFP mRNA suspension into wells of the Nucleocuvette strip (21 μL /well). Make sure the samples cover the bottom of the wells and avoid air bubbles (*see Note 5*). Attach the lid to the Nucleocuvette strip.
- Program the nucleofection device according to the electroporation parameter matrix designed in **step 1** and electroporate the samples.

9. Immediately after electroporation, add prewarmed culture media to the cells (180 μL per well of the Nucleocuvette strip). Using a pipette with an appropriately small-sized tip, gently pipette up and down and then transfer the cells into a prepared culture vessel with prewarmed media.
10. Culture the cells overnight and analyze GFP expression and cell viability by flow cytometry (*see* **Note 6**).
11. Optimal electroporation settings are chosen based on percent GFP positive cells, GFP intensity (mean fluorescence intensity; MFI), and percent viable cells. An example of results is shown in Fig. 1b. There may be a trade-off between cell viability and GFP expression, and the best parameter will depend on the requirements for the specific application. It may be necessary to proceed with testing a few different settings for Cas9 RNP delivery (Subheading 3.2).

3.2 Electroporation of Cas9 RNP

1. Based on the results from Subheading 3.1, one or more electroporation conditions are evaluated by Cas9 RNP electroporation for the efficiency of INDEL formation.
2. Prepare a cell population with high viability (>90%) and in logarithmic growth phase (*see* **Note 1**).
3. Immediately prior to preparing the cells for electroporation, complex the Cas9 protein and the sgRNA by mixing 6 μg Cas9 protein with 3.2 μg sgRNA in a PCR tube per electroporation sample (*see* **Note 4**). Mix by pipetting and incubate at room temperature for at least 10 min and maximum 1 h.
4. Count the cells and distribute the desired cell number into one Eppendorf tube per electroporation condition (*see* **Note 2**). Include cells for one extra sample that will not be electroporated. This will serve as a reference control during INDEL quantification.
5. Pellet cells by centrifugation at room temperature at $300 \times g$ for 5 min.
6. Aspirate the supernatant and resuspend the cells in PBS.
7. Pellet the cells by centrifugation at room temperature at $300 \times g$ for 5 min.
8. Aspirate the supernatant and gently resuspend the cells in electroporation buffer (20 μL per sample) by pipetting up and down (*see* **Note 3**).
9. Transfer 20 μL of the cell suspension to each of the PCR tubes containing Cas9 RNP and mix by pipetting gently up and down. 20 μL of cell suspension should be left that will serve as an un-electroporated sample and is transferred directly into a prepared culture vessel with prewarmed media.

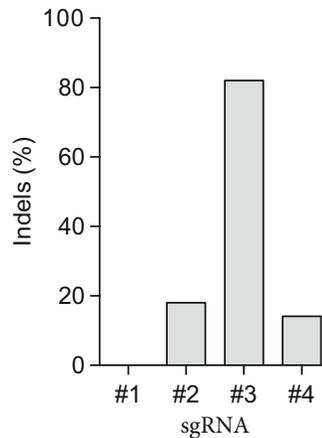


Fig. 2 Gene editing efficiencies after optimization of electroporation conditions. Bar graph shows the results of a TIDE analysis quantifying the Indels created in the PMDC05 cell line after electroporation with four different sgRNAs targeting the same gene. The buffer (Opti-MEM) and electroporation program (CM138) was chosen based on the results obtained in Fig. 1b

10. Distribute the cell + RNP suspensions into different wells of the Nucleocuvette strip. Make sure the samples cover the bottom of the wells and avoid air bubbles (*see Note 5*). Attach the lid to the Nucleocuvette strip.
11. Program the nucleofection device according to the desired electroporation program(s) identified in Subheading 3.1 and electroporate the samples.
12. Immediately after electroporation, add prewarmed culture media to the cells (180 μ L per well of the Nucleocuvette strip). Using a pipette with an appropriately small-sized tip, gently pipette up and down and then transfer the cells into a prepared culture vessel of appropriate size containing prewarmed media (*see Note 7*).
13. Culture the cells for at least four days and then harvest genomic DNA. For INDEL quantification, we routinely use Sanger sequencing of a PCR product spanning the sgRNA target site and analysis with either ICE (ice.synthego.com) or TIDE (*see Note 8*) [13]. *See Fig. 2* for an example of INDEL frequency results obtained in the PMDC05 cell line using four different sgRNAs and optimized electroporation parameters.

4 Notes

1. Cells should be passaged 2–3 days before electroporation to ensure a logarithmic growth phase. For adherent cell types

(e.g., fibroblasts) we recommend electroporating cells when they reach 70–85% confluency as higher cell confluency may cause lower electroporation efficiencies. Trypsin can be used to detach cells without affecting the electroporation procedure.

2. We routinely use 200,000 cells per electroporation for the Nucleocuvette strips and include an extra sample for each electroporation buffer that will serve as an un-electroporated reference for viability and GFP expression during flow cytometry. As an example, if testing four different electroporation buffers and five different electroporation programs, distribute 1,200,000 cells into each of four different Eppendorf tubes (one tube per electroporation buffer, $5 \times 200,000$ cells for electroporation programs and $1 \times 200,000$ cells for an un-electroporated reference).
3. Avoid keeping the cells in the electroporation buffer for longer periods as this may reduce cell viability and gene transfer efficiency.
4. Cas9 protein, sgRNA, and GFP mRNA should be thawed and kept on ice during the procedure to limit loss of efficiency due to degradation of protein/RNA. Follow general guidelines for handling RNA to avoid RNase-mediated degradation.
5. To avoid bubbles, add the cell-suspension carefully to the bottom of the nucleocuvette strip well using an appropriate pipette tip that is thin enough to reach the center of the well without hitting the metal sides (we find p100 pipette tips or the plastic pipettes provided by Lonza appropriate for this). After adding the cell suspension, gently tap the strip to ensure that no bubbles are trapped beneath the cell suspension. If bubbles occur, use a p10 pipette and carefully remove these.
6. Depending on the cell type, GFP expression may be observed a few hours after electroporation. However, we recommend analyzing the cells after ~16–36 h to get a better assessment of the effect of the electroporation protocol/buffer on cell viability.
7. For precise gene editing (gene knock-in, gene correction, or gene changes), the methodology may be combined with AAV-mediated donor delivery directly after electroporation for use of the homology-directed repair pathway [14].
8. Depending on the desired application, we recommend performing additional assays to confirm the action of the sgRNA, such as phenotypic or functional assays (e.g., protein expression).

Acknowledgments

R.O.B. gratefully acknowledges the support from a Lundbeck Foundation Fellowship (R238-2016-3349), an AIAS-COFUND

(Marie Curie) fellowship from Aarhus Institute of Advanced Studies (AIAS) co-funded by Aarhus University's Research Foundation and the European Union's seventh Framework Program under grant agreement no 609033, the Novo Nordisk Foundation (NNF17OC0028894), Innovation Fund Denmark (8056-00010B), the Carlsberg Foundation (CF17-0129), Slagtermester Max Wørzner og Hustru Inger Wørzners Mindelegat, The AP Møller Foundation, and the Riisfort Foundation.

References

1. Barrangou R, May AP (2015) Unraveling the potential of CRISPR-Cas9 for gene therapy. *Expert Opin Biol Ther* 15(3):311–314. <https://doi.org/10.1517/14712598.2015.994501>
2. Mandal PK, Ferreira LM, Collins R, Meissner TB, Boutwell CL, Friesen M, Vrbanac V, Garrison BS, Stortchevoi A, Bryder D, Musunuru K, Brand H, Tager AM, Allen TM, Talkowski ME, Rossi DJ, Cowan CA (2014) Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell Stem Cell* 15(5):643–652. <https://doi.org/10.1016/j.stem.2014.10.004>
3. Burdette DL, Monroe KM, Sotelo-Troha K, Iwig JS, Eckert B, Hyodo M, Hayakawa Y, Vance RE (2011) STING is a direct innate immune sensor of cyclic di-GMP. *Nature* 478(7370):515–518. <https://doi.org/10.1038/nature10429>
4. Hornung V, Rothenfusser S, Britsch S, Krug A, Jahrsdorfer B, Giese T, Endres S, Hartmann G (2002) Quantitative expression of toll-like receptor 1–10 mRNA in cellular subsets of human peripheral blood mononuclear cells and sensitivity to CpG oligodeoxynucleotides. *J Immunol* 168(9):4531–4537
5. Hendel A, Bak RO, Clark JT, Kennedy AB, Ryan DE, Roy S, Steinfeld I, Lunstad BD, Kaiser RJ, Wilkens AB, Bacchetta R, Tsalenko A, Dellinger D, Bruhn L, Porteus MH (2015) Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat Biotechnol* 33(9):985–989. <https://doi.org/10.1038/nbt.3290>
6. Kim S, Kim D, Cho SW, Kim J, Kim JS (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res* 24(6):1012–1019. <https://doi.org/10.1101/gr.171322.113>
7. Yin H, Song CQ, Suresh S, Wu Q, Walsh S, Rhym LH, Mintzer E, Bolukbasi MF, Zhu LJ, Kauffman K, Mou H, Oberholzer A, Ding J, Kwan SY, Bogorad RL, Zatspein T, Koteliansky V, Wolfe SA, Xue W, Langer R, Anderson DG (2017) Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing. *Nat Biotechnol* 35(12):1179–1187. <https://doi.org/10.1038/nbt.4005>
8. Lee K, Mackley VA, Rao A, Chong AT, Dewitt MA, Corn JE, Murthy N (2017) Synthetically modified guide RNA and donor DNA are a versatile platform for CRISPR-Cas9 engineering. *elife* 6:e25312. <https://doi.org/10.7554/eLife.25312>
9. Wong TK, Neumann E (1982) Electric field mediated gene transfer. *Biochem Biophys Res Commun* 107(2):584–587
10. Chicaybam L, Sodre AL, Curzio BA, Bonamino MH (2013) An efficient low cost method for gene transfer to T lymphocytes. *PLoS One* 8(3):e60298. <https://doi.org/10.1371/journal.pone.0060298>
11. Bak RO, Porteus MH (2017) CRISPR-mediated integration of large gene cassettes using AAV donor vectors. *Cell Rep* 20(3):750–756. <https://doi.org/10.1016/j.celrep.2017.06.064>
12. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (2013) Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8(11):2281–2308. <https://doi.org/10.1038/nprot.2013.143>
13. Brinkman EK, Chen T, Amendola M, van Steensel B (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42(22):e168. <https://doi.org/10.1093/nar/gku936>
14. Bak RO, Dever DP, Porteus MH (2018) CRISPR/Cas9 genome editing in human hematopoietic stem cells. *Nat Protoc* 13(2):358–376. <https://doi.org/10.1038/nprot.2017.143>

Part III

CRISPR Gene Editing in Human iPSCs



Chapter 10

Efficient Gene Editing of Human Induced Pluripotent Stem Cells Using CRISPR/Cas9

Saniye Yumlu, Sanum Bashir, Jürgen Stumm, and Ralf Kühn

Abstract

The generation of targeted mutants is a crucial step toward studying the biomedical effect of genes of interest. The generation of such mutants in human induced pluripotent stem cells (iPSCs) is of an utmost importance as these cells carry the potential to be differentiated into any cell lineage. Using the CRISPR/Cas9 nuclease system for induction of targeted double-strand breaks, gene editing of target loci in iPSCs can be achieved with high efficiency. This chapter covers protocols for the preparation of reagents to target loci of interest, the transfection, and for the genotyping of single cell-derived iPSC clones. Furthermore, we provide a protocol for the convenient generation of plasmids enabling multiplex gene targeting.

Key words Pluripotent stem cells, Gene editing, CRISPR, Cas9, Knockout, Knockin, i53, Trex2

1 Introduction

Editing of human induced pluripotent stem cells (iPSCs) is of great importance in order to study the function of targeted genes with the goal of disease modeling and gene correction. CRISPR/Cas9 has led to a new era in the field of genome engineering [1]. Since the first report of its applicability for genome engineering, a myriad of publications describe utilization of CRISPR/Cas9 for an ever-increasing number of genome editing experiments [2, 3]. Gene targeting of desired genomic loci to remodel the genome in a way which can be useful in understanding disease, correction of mutation(s), and introduction of reporter gene cassettes into the genome are just a few examples [4–6]. These goals can be achieved by initiation of site-specific DNA double-strand breaks (DSBs) for introduction of precise mutations, gene correction, or gene inactivation. The targeted DSBs are repaired by one of two predominant DNA repair pathways, namely non-homologous end joining (NHEJ) and homologous recombination (HR). Repair via the NHEJ pathway often results in a variety of sequence modifications known as insertion-deletions (indels) and gives rise to

knockout (KO) mutations. On the contrary, the HR pathway leads to precise repair and is utilized to introduce sequences of interest either via single-stranded oligodeoxynucleotides (ssODNs) or plasmid-based knockin (KI) donor vectors [7–10].

In a typical gene targeting experiment the basic components of the CRISPR/Cas9 system, a Cas9 expression plasmid and a single guide (sg) RNA expression vector are introduced into host cells with or without an ssODN or KI vector. Prior to reagent delivery, the best working transfection method for a particular cell line should be established. If the goal of experiment is to target multiple genes at once, new target sites can be easily addressed by expressing multiple sgRNAs from a single plasmid while the Cas9 expression vector remains invariant [11]. The gradual steps which are required for representative multiplex KO and KI targeting experiments in human iPSCs are explained in more detail in the successive methods section. Nevertheless, the important considerations that should be given to a CRISPR-based targeting experiment are (1) determining the goal of an editing experiment, (2) selection of a target sites and design of sgRNAs, (3) choice of appropriate reagents delivery method (e.g.; plasmid vs. peptide-based), and (4) selection and genotyping of edited cells. As a rule of thumb, selection of guide RNA(s) should be made according to the goals of a specific targeting experiment such as generation of knockout alleles, insertion or correction of few nucleotides, larger knockins or defined deletions of a certain genomic sequence. An important consideration for designing a knockout experiment is to select a guide RNA that targets an exon so that chance frameshift mutations resulting from NHEJ-based repair result in knockout phenotype. However, for introduction of a precise mutation by HR the guide RNA should be designed in a way that it is located in the vicinity of the desired mutation location, in particular when the targeting is performed using ssODNs [12–14].

The protocol detailed in this chapter is the standard practice we routinely use for mutant iPSC cell line generation in our laboratory (Fig. 1). These methods are based on lipofection of human iPSCs using Cas9 and sgRNA expression plasmids for either transient expression (Fig. 2a), or using a stably transfected cell line where Cas9 is expressed from a strong doxycycline inducible promoter (Fig. 2b). Where transient transfections can result in moderate editing efficiency at desired loci, stable Cas9 expressing lines are remarkably efficient in targeting multiple loci in a single experiment. Here, a single plasmid expressing multiple sgRNAs as well as a selection marker for antibiotic or fluorescence-based selection can be transfected transiently with high efficiency and ultimately results in high gene editing. Using this approach, we routinely obtain mutant clones harboring indels at a frequency of 40–85% and homology directed repair (HDR) alleles at a frequency of 30–70%.

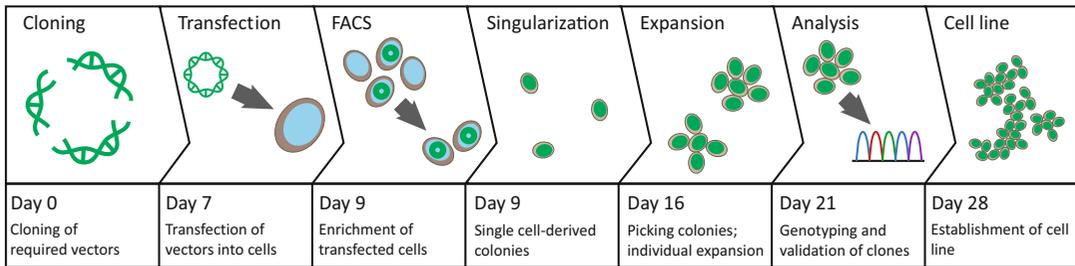


Fig. 1 Workflow for genome editing using plasmids. Having chosen a targeting site of interest and designed an appropriate guide RNA in silico, respective expression vectors are cloned. These vectors are transfected into hiPSCs, transfected cells subsequently enriched utilizing flow cytometry, and disseminated sparsely to ensure arising colonies being single cell-derived. These colonies are isolated, expanded, and successful genome editing verified by sequencing before establishing a cell line

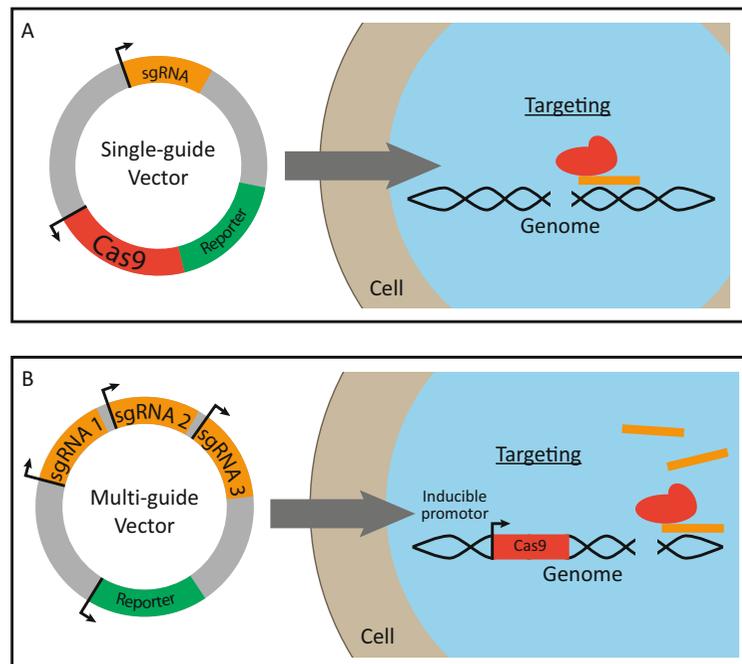


Fig. 2 Overview on both CRISPR/Cas9-mediated targeting strategies elaborated in this methods chapter. **(a)** An expression vector carrying the Cas9 gene and a guide RNA cassette transfected into hiPSCs enables effective targeting of a single locus. **(b)** Simultaneous targeting of several loci can be achieved by combining multiple guide RNA expression cassettes in one vector. Utilizing a transgenic cell line that provides inducible expression of Cas9 further enhances efficacy. Reporter genes included in both plasmids allow for flow cytometry-based enrichment of successfully transfected cells

2 Materials

2.1 Cloning of sgRNAs

1. sgRNA cloning vector (e.g., Addgene #86985; www.addgene.com).
2. A pair of target specific sgRNA oligonucleotides with BbsI overhangs.
3. BbsI restriction enzyme.
4. *AscI* restriction enzyme.
5. T4 Ligase.
6. Herculanase II Fusion DNA Polymerase Kit.
7. dNTP Set.
8. Plasmid Miniprep Kit.
9. Chemically competent *E. coli* bacteria.
10. Carbenicillin.
11. LB medium and LB agar plates.
12. Thermocycler Agarose gel electrophoresis (setup and consumables).
13. DNA gel extraction kit.

2.2 General hiPS Cell Culture

1. hiPS cell lines (we use the human pluripotent stem cell registry (<https://hpscereg.eu/>) lines BIHi001A or BIHi043-A (without inducible Cas9) and BIHi001-A-1 or BIHi043A1 as lines with Doxycycline inducible Cas9).
2. hiPS cell culture media (Essential-8).
3. Y-27632.
4. DPBS $-/-$ ($-Ca^{2+}$, $-Mg^{2+}$).
5. Vitronectin (VTN-N) recombinant Human.
6. Cell culture plates.
7. Accutase solution.
8. Bambanker Serum Free Cell Freezing Media.
9. DMSO.
10. Doxycycline Hydrochloride.
11. Dissociation buffer: 0.18% NaCl, 0.5 mM EDTA in DPBS (to 491.92 mL DPBS add 5 ml of 0.5 M EDTA pH 8.0 and 3.8 mL of 5 M NaCl).

2.3 Lipofection of hiPS Cells

1. Lipofectamine 3000 transfection kit.
2. Opti-MEM I reduced serum medium.

2.4 FACS Enrichment of Transfected hiPS Cells

1. 0.45 μ m PTFE membrane filter.
2. FACS tubes with 35 μ m strainer cap.

3. Penicillin/Streptomycin (10,000 U/mL).
4. Gentamycin (10 mg/mL).
5. RevitaCell Supplement (100×).
6. Flow cytometry cell sorter.

3 Methods

Preparation of CRISPR Vectors for Gene Targeting.

3.1 Cloning of a Single sgRNA

For the selection of sgRNAs we analyze genomic sequences using the CRISPOR webtool (<http://crispor.tefor.net/>). sgRNA oligonucleotides can be cloned into short guidance RNA expression vectors (e.g., Addgene 86985 www.addgene.org) by *BbsI* restriction enzyme overhangs, with N₁–N₂₀ as the selected Cas9 guiding target sequence (*see* **Note 1**).

3.2 Digestion of sgRNA Expression Vector

1. In a total volume of 100 μ L, digest 5 μ g sgRNA expression vector with 50 units of the restriction enzyme *BbsI* and 10 μ L of 10× corresponding restriction buffer. Fill up the reaction to 100 μ L with nuclease-free water and mix carefully. Incubate for 1–2 h at 37 °C.
2. Inactivate the restriction enzyme for 20 min at 65 °C.
3. Load and run the digested vector on a 0.9% agarose gel.
4. Extract the linearized vector using a DNA gel extraction kit (*see* **Note 2**).

3.3 Annealing of sgRNA Oligonucleotides

1. Solve lyophilized sgRNA oligonucleotides at a concentration of 1 μ g/ μ L in 1× TE-buffer.
2. Dilute oligonucleotides to a concentration of 10 ng/ μ L by combining 1 μ L of each oligo and 98 μ L 1× TE-buffer.
3. Denature for 5 min at 98 °C in a heat block (*see* **Note 3**).
4. Cool down the mixture slowly to RT by switching off the heat block.
5. Keep annealed oligonucleotides on ice for immediate use or store at –20 °C for later use (*see* **Note 4**).

3.4 Ligation of sgRNA Oligonucleotides and Transformation

1. Set up a ligation reaction with 100 ng *BbsI* linearized vector, 1.5 μ L annealed oligonucleotides (10 ng/ μ L each), 1 μ L T4 ligase, and 1.5 μ L 10× T4 ligase buffer. Fill up to 15 μ L with nuclease-free water and incubate at 16 °C overnight.
2. Thaw chemically competent *E. coli* cells slowly on ice.
3. Add 5 μ L of the ice-cold ligation reaction to 50 μ L chemically competent *E. coli*. Carefully triturate twice using a pipette to

mix the competent cells with the ligation mix and incubate for 30 min on ice.

4. Perform heat shock in a water bath at 42 °C for 60 s.
5. Incubate for 3 min on ice.
6. Add 1 mL LB medium w/o antibiotics and incubate for 30 min at 37 °C at 200 rpm.
7. Plate the transformation on LB-Agar plates containing appropriate antibiotics, e.g., 50 µg/mL carbenicillin, and incubate O/N at 37 °C.
8. At the following day pick up to five colonies, inoculate 5 mL LB medium containing appropriate antibiotics, and incubate at 37 °C at 200 rpm overnight.
9. Perform plasmid mini preparations and verify the correct sequence by Sanger sequencing using an appropriate primer (e.g., hU6-fwd) (*see* **Notes 5 and 6**).

3.5 Cloning of Multiple sgRNAs via Gibson Assembly

Combining multiple sgRNA expression cassettes via Gibson assembly requires that the individual sgRNAs have to be cloned separately beforehand as described in the previous section. The following protocol provides instructions for assembling three sgRNA expression cassettes in one vector. If it is intended to combine more than three sgRNAs, primer pairs have to be designed with appropriate assembly overhangs, according to manufacturer guidelines (*see* **Note 7**).

3.6 Amplification of Guide RNA Expression Cassettes

1. Perform PCR amplification of the guide RNA cassettes for every sgRNA you want to combine in a vector. Set up the PCR reaction in 50 µL reaction volume with 5–20 ng DNA of the single sgRNA cassette containing plasmids, 0.25 µM of each primer, 250 µM dNTPs, and a proofreading DNA polymerase, e.g., 0.5 µL Herculase II. Fill up with nuclease-free water.
2. Carry out the PCR reaction with following conditions: 98 °C, 2 min; [98 °C, 30 s; 55 °C, 30 s; 72 °C, 30 s] × 30 cycles; 72 °C 2 min.
3. Load the PCR products on a 1% agarose gel and purify the 412 bp fragment using a DNA gel extraction kit (*see* **Note 8**).

3.7 Digestion of sgRNA Expression Vector

1. We suggest to use Addgene plasmid #86985 as backbone for this cloning purpose but any other plasmid with matching sequences and appropriate restriction sites can be used. Digest 5 µg of plasmid with 50 U *AscI* enzyme and 10 µL of 10× corresponding restriction buffer. Fill up the reaction to 100 µL with nuclease-free water and mix carefully. Incubate for 1–2 h at 37 °C.

2. Inactivate the enzyme for 20 min at 65 °C.
3. Load the digested vector on a 0.9% agarose gel and extract the linearized vector backbone using a DNA gel extraction kit. Expected fragment sizes are: 5586 bp for the vector backbone and 383 bp for the empty sgRNA cassette.

3.8 Gibson Assembly Reaction

1. To assemble three sgRNA expression cassettes, use 30 ng of each purified PCR product from Subheading 3.6, 200 ng of AscI digested plasmid backbone, and 10 µL Gibson assembly master mix (2×). Fill up to 20 µL with nuclease-free water and incubate for 1 h at 50 °C.
2. Perform transformation as described in Subheading 4.1.4.
3. When using Addgene plasmid #86985 perform a test digestion with the restriction enzyme NdeI on five to ten plasmid mini preparations. The expected fragment sizes are for three sgRNA expression cassettes 5515 bp; 436 bp; 391 bp and 390 bp when using plasmid #86985 as backbone. Subject two plasmids with the correct digestion pattern for Sanger sequencing using appropriate primers (*see* **Notes 9** and **10**).

3.8.1 Targeting Procedure

Efficient transfection of a Cas9 nuclease and the sgRNA expression vector is indispensable for successful genome editing. The choice of a certain transfection method depends upon its efficacy in a hiPS cell line of interest as well as on the viability of the cell line, when subjected to the transfection method. However, where some cell lines can be efficiently transfected by reagent or chemical-based transfection methods such as Lipofectamine 3000, others respond well to electroporation. To facilitate generation of gene-edited clones we recommend using either a fluorescent marker if FACS sorting is feasible or selection markers like puromycin or hygromycin on the plasmid encoding for Cas9 and sgRNA to enrich transfected cells.

Co-expression of factors facilitating either NHEJ-mediated DNA repair pathway for the introduction of Indel mutations or the HDR pathway to generate precise knockin further improves gene targeting efficiencies. We recommend to co-transfect Trex2 (Fig. 3a), an exonuclease to promote Indel formation, or i53, an inhibitor of 53BP1 that is a key regulator of NHEJ and favors DSB repair by HDR (Fig. 3b).

3.9 Transfection of CRISPR/Cas9 Constructs into hiPS Cells

1. This protocol is based on hiPS cells grown in feeder-free conditions in Essential 8 culture media in a 6-well cell culture plate. Prepare plasmid DNA at a concentration of 0.5–1 µg/µL in deionized water or TE buffer. The DNA used for transfection should be of high quality as poor quality of DNA might decrease the efficacy of transfection.

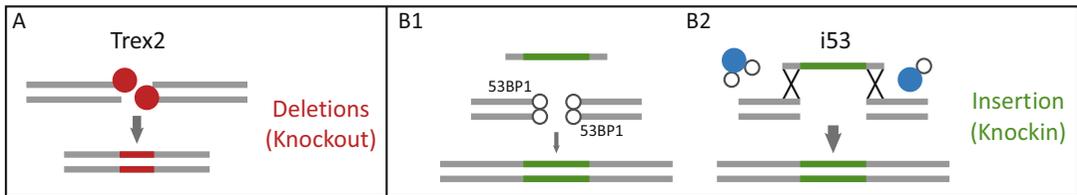


Fig. 3 Trex2 and i53 increase frequencies of indel and HDR formation, respectively. **(a)** In case of a DSB exonuclease Trex2 (red circle) trims nucleotides from 3' termini, resulting in deletions. **(b)** Following DSBs transcription factor 53 binding protein 1 (53BP1) impedes HDR mechanisms. i53 (blue circle) inhibits this effect, therefore increasing the frequency of HDR events

2. 24 h before transfection of the cells with the components for CRISPR/Cas9-mediated gene editing, iPSCs should be passaged. We recommend to use for every targeting experiment three wells of a 6-well plate to obtain sufficient cell numbers after the entire procedure (*see* **Notes 11** and **12**).

Remove medium and wash the cells once with DPBS. Add 0.5 mL Accutase per well of a 6-well plate and incubate for 3–5 min at 37 °C. When colonies begin to detach and fall apart, add 2 mL iPSC growth medium. Pipet carefully up and down to break cell patches to single cells and pellet the cells at 300 g for 4 min. Resuspend iPSCs in an appropriate volume of medium (1–2 mL per well of a 6-well plate) including 10 μ M Y-27632 and count the cells. Plate the cells on cell culture plates with appropriate coating at a density of 5×10^4 to 1×10^5 (*see* **Notes 13–15**).

3. Change the medium ahead of transfection.
4. **(A)** When it is intended to generate cells with InDel mutations to produce knockout mutations, dilute plasmids carrying Cas9 and/or sgRNA. Depending on the size, 1 μ g DNA (for plasmids <8 kb (sgRNA and fluorescent reporter)) or 2 μ g DNA (for plasmids >8 kb (Cas9, sgRNA and fluorescent reporter)) together with 0.5 μ g of a Trex2 and BFP encoding plasmid (e.g., Addgene plasmid #111145) in 125 μ L of Opti-MEM reduced serum medium in tube labeled as A.

(B) When planning to create KnockIn alleles using transient Cas9 transfection and short single-stranded oligonucleotides (ssODN) dilute the following in 125 μ L Opti-MEM: 1.5 μ g plasmid DNA (Cas9, sgRNA and fluorescent reporter), 0.5 μ g expression plasmid encoding for i53 and BFP (e.g., Addgene plasmid #111145) and 30 pmol ssODN (*see* **Note 16**).

When utilizing doxycycline inducible Cas9 expressing hiPSC lines, use 0.5 μ g sgRNA and fluorescent reporter containing plasmid DNA, 0.5 μ g expression plasmid encoding for i53 and BFP and 30 pmol of ssODN each (*see* **Note 17**).

5. Add 3.75 μL of p3000 reagent to tube A and mix well. Dilute 5 μL Lipofectamine 3000 reagent in 125 μL of Opti-MEM reduced serum medium in tube B and mix well. Diluted Lipofectamine 3000 should be used within 15 min, longer times can result in decreased transfection efficacy.
6. Add the content of tube A to tube B, mix well by vigorously pipetting or vortexing and incubate at room temperature for 5 min.
7. Add the DNA-lipid complex to one well of a 6-well plate in a dropwise manner and gently rock the plate to ensure its distribution over the whole well (*see Note 18*).
8. Change medium on the following day. When using iPS cells harboring a doxycycline inducible Cas9, feed the cells for 48 h with medium supplemented with 1 $\mu\text{g}/\text{mL}$ doxycycline. 48 h after transfection proceed with FACS sorting (Subheading 4.3) (*see Note 19*).

3.9.1 FACS Enrichment and Clonal Isolation hiPS Cells

Selecting cells that were successfully transfected improves the rate of gene editing cells. In this protocol, we describe the enrichment of transfected cells by FACS. If there is no opportunity to utilize FACS, other selection methods, e.g., using antibiotics like puromycin or hygromycin, can be applied (*see Note 20*).

3.10 Flow Cytometry-Based Enrichment of Transfected hiPSC

1. Prepare post-FACS medium: 50% fresh Essential 8 medium, 50% conditioned Essential 8 medium (medium that was conditioned for 24 h on hiPS cells with 50–80% confluency of the same line and filtered through a 0.45 μm PTFE membrane filter) supplemented with 10 μM Rock inhibitor Y-27632, RevitaCell Supplement (1 \times), Pen/Strep (1 \times), and Gentamicin (1 \times). Coat a 6-well plate with appropriate coating (*see Notes 21–23*).
2. For each FACS sample, prepare a 1.5 mL microcentrifuge tube with 1 mL post-FACS medium and keep on ice.
3. Pre-warm Essential 8 medium containing 10 μM Y-27632, Pen/Strep (1 \times), and Gentamicin (1 \times).
4. Dissociate cells using Accutase (Subheading 4.2.1).
5. Resuspend cells in 200–300 μL medium per well of a 6-well plate at 70–80% confluency.
6. Strain through a strainer cap with a mesh size of 35 μm into a FACS tube.
7. Put cells on ice and proceed with flow cytometry immediately.
8. Depending on the nature of your experiment and the chosen fluorophore, sort for your desired cell population directly into a cooled 1.5 mL microcentrifuge tube containing 1 mL of post-FACS medium.

9. Seed 300–1000 cells per well in 2 mL post-FACS medium in a 6-well format as quickly as possible after sorting. Prepare two to three wells for each sample (*see Note 24*).
10. Make sure single cells are evenly distributed throughout the well to ensure emergence of single cell-derived colonies.
11. Additionally, seed the remaining cells at a higher density on a separate well. This well can be used after expansion for gDNA extraction from the sorted bulk population and subsequent PCR-based genotyping to determine the rough gene targeting efficacy, for example, by RFLP (*see Notes 25 and 26*).

3.11 Clonal Isolation of hiPSCs

1. Culture the sorted cells for 5 days in 50% conditioned and 50% fresh medium until small colonies emerge. Change medium daily (*see Note 27*).
2. When colonies emerge, usually between day 4 and day 6 switch to regular Essential 8 medium and continue with daily feeding.
3. Monitor the growth of the single cell-derived colonies to exclude the possibility that neighboring colonies fuse together. This can be done by circling neighboring colonies on the bottom side of the well with a marker pen.
4. After approximately 7–10 days colonies are large enough for manual picking.
5. Gently scratch the colonies into smaller, checkered patches either using a small needle or using a 10 μ L pipette tip utilizing a stereo microscope in a sterile environment.
6. Transfer the patches of one clone to one well of a pre-coated 12-well plate containing pre-warmed medium supplemented with 10 μ M Y-27632.
7. Gently resuspend the cell patches by pipetting up and down with a 1000 μ L tip. After a few days colonies may need to be dissociated with EDTA within the plate to break apart the large colony and avoid differentiation of the cells. Change medium daily.
8. When the wells reach a confluency of 80%, detach the cells (*see Subheading 4.2.1*) of each well and transfer 20% of the cells of each clone to a well of a pre-coated 12-well plate (back-up plate).
9. Centrifuge the remaining 80% of cells, discard the supernatant, and perform gDNA extraction, followed by PCR-based genotyping to detect CRISPR/Cas9-edited clones.
10. As soon as the back-up reaches a confluency of 60 to 80%, cryopreserve the cells in the plate (*see Note 28*).
11. Aspirate the medium and cover the wells with dissociation buffer. Incubate the plate for 3–5 min at 37 °C. Aspirate the

Table 1
Summary of analysis methods for genome genome modifications

Assay	Advantage	Disadvantage	Application
Sanger/Tide	Quick and simple; High accuracy for single cell-derived clones	Low accuracy for quantification in bulk population	Indel detection in clonal-derived cells
Cloning/ Sanger	High accuracy	Very laborious	For verification of knockins or indels in single cell-derived clones
Fragment analysis	Quick and simple; High accuracy for single cell-derived clones	Requires special equipment/service	Indel detection in clonal-derived cells or bulk populations
RFLP	Quick and simple; High accuracy for single cell-derived clones	Low accuracy for quantification on bulk populations	Detection of knockins or indels in single cell-derived clones or bulk populations
Deep Seq	High accuracy	Requires special equipment/service	Detection of KI or indels in bulk populations and single cell-derived clones
T7E1/Surveyor	Quick and simple	Need thorough optimization	To pre-test sgRNAs in cell lines

dissociation reagent and add 0.4 mL chilled Bambanker cryo-preservation solution per well of a 12-well plate. Gently dissociate the colonies by repeated pipetting. Seal the plate with Parafilm and store the plate immediately at -80°C in a Styro-foam box until further use.

12. Thaw positive genotyped clones using a standard protocol, expand the cells, and cryo-preserve at least 5 vials.

3.11.1 Analysis Techniques of CRISPR/ Cas9 Induced Modifications

Various methods to analyze genome modifications are available. Most of the techniques are based on PCR amplification of the targeted region, and in such cases, optimization of the PCR reactions is required for clear analysis results. In Table 1, we summarized the most common methods for analyzing indels or point mutations.

4 Notes

1. For sgRNA cloning in sgRNA expression vectors (e.g., Addgene #86985 or #86986) oligonucleotides with suitable overhangs are required. Following sequences can be used to design sgRNA oligonucleotides, with N_1-N_{20} as the selected Cas9 target sequence.

sgRNA-oligo-F 5' -CACC(G)_{N₁}NNNNNNNNNNNNNNNN
 NNNNNN₂₀-3'

sgRNA-oligo-R 5' -AAACNNNNNNNNNNNNNNNNNN
 NNNN(C)-3'

2. Purify the linearized plasmid via gel extraction rather than directly via a spin column to avoid contamination of the DNA with undigested plasmid.
3. For denaturation and annealing of oligonucleotides a PCR cycler programmed for 10 min at 98 °C and successive decrease of 1 °C per minute to a temperature of 4 °C can be used alternatively.
4. The abovementioned vectors allow expression of the sgRNA by the human U6-promoter. This promoter requires a “G” base at the transcription start site. Hence, it is recommended using CRISPR/Cas9 target sites starting with a “G.” Otherwise an additional “G” should be added at the start of the sgRNA sequence. It should be noted that wild-type Cas9 is amenable to the inclusion of an extra “G” but other RNA guided nucleases are not (Cas9-HF/eCas9).
5. When using the vectors mentioned in **Note 2**, sequencing can be performed using subsequent primer sequence. hU6-For: GAGGGCCTATTTCCCATG.
6. Re-digestion of the ligated plasmid using the enzyme *BbsI* improves the cloning efficacy.
7. When using the suggested vector (Addgene #86985 or #86986) for cloning multiple sgRNA expression cassettes in one backbone, the latter primers can be used for cloning:

Gibson-pU6-A_F
 CAGGAAACAGCTATGACCATGAGGGCCCCCTTCA
 CCGAGGGCCTATTTTC

Gibson-pU6-A_R
 CCGATGGCCAGGCCGATGCTGTGATCAAAAAAAG
 CACCGACTCGG

Gibson-pU6-B_F
 ACAGCATCGGCCTGGCCATCGGGCCCCCTTCAC
 CGAGGGCCTATTTTC

Gibson-pU6-B_R
 CTTGGCCATCTCGTTGCTGAAGATCAAAAAAAGC
 ACCGACTCGG

Gibson-pU6-C_F
 TTCAGCAACGAGATGGCCAAGGCCCCCTTCACCG
 AGGGCCTATTTTC

Gibson-pU6-C_R
 GTCAATAATCAATGTCGAATCCGGGATCAAAAAA
 GCACCGACTCGG

Use primer combinations A_F/A_R; B_F/B_R and C_F/C_R if three cassettes are desired to be cloned and combinations A_F/A_R and B_F/C_R if it is desired to clone two sgRNA cassettes.

8. Purify PCR product via gel extraction rather than directly via a spin column to avoid contamination of the purified PCR product with plasmid DNA. Alternatively, the PCR product can be digested with DpnI, which digests only methylated DNA and subsequently purified by spin columns.
9. To improve the transformation efficacy perform a cleanup of the assembled reaction using a PCR purification kit, prior to transformation.
10. When using vector Addgene #86985 or #86986 as plasmid backbone, following primer sequences can be used for sequencing.
 pU6-(BbsI)sgRNAseqF: TTGTGTGGAATTGTGAGCGG
 pU6-(BbsI)sgRNAseqR: GGCTATGAACTAATGACCCCG
11. Passaging iPSCs as single cells one day prior transfection ensures that the cells are in the growth phase.
12. For a targeting experiment only use iPSCs of good quality that are undifferentiated and have a normal karyotype.
13. Depending on the iPSC line and the confluency detaching of the cells can take up to 15 min. Monitor the cells with an inverted microscope every 2 min.
14. The optimal cell number needs to be figured out for every cell line. iPSC cell lines differ in their viability upon transfection [11].
15. Use an appropriate cell culture vessel coating for your cell line. We recommend truncated recombinant human vitronectin (VTN-N) when possible. VTN-N minimizes movement of cells on the plates.
16. We have observed that with transient Cas9-plasmid transfection up to two ssODNs with 30 pmol each can be used safely, e.g., to create heterozygous KIs with one repair template for carrying the desired mutation and a second wild-type template that has a modified PAM sequence.
17. When using doxycycline inducible Cas9 expressing hiPSC lines up to three ssODN with 30 pmol each can safely be transfected with only minor cell death after transfection.
18. Some cell lines are sensitive to Lipofectamine 3000. The addition of 5 μ M Rock inhibitor Y-27632 to the culture medium for 24 h increases viability.
19. When lipofection is not convenient for your cell line try electroporation methods.

20. The procedure of FACS enrichment might induce cell stress; therefore, we suggest processing the cells as fast as possible. Dissociated cells should not be kept on ice for longer than 45 min.
21. Cultivating hiPS cells in 50% conditioned medium dramatically increases cell survival of the FACS [15]. We have observed that the usage of conditioned medium increases the viability of single seeded iPS cells in low density and prevents the cells from spontaneous differentiation.
22. To avoid contaminations during FACS we suggest using penicillin and streptomycin as well as gentamicin during sorting and for another 24 h.
23. We suggest VTN-N, this minimizes cell migration and therefore promotes cultivation of single cell-derived clones.
24. The number of seeded cells has to be determined for every hiPS cell line for different cell lines have variable survival when seeded at low density.
25. The analysis of the bulk cell population can be very helpful to determine the number of single clones that need to be analyzed for the specific gene targeting event.
26. In case of low gene targeting efficacy, two rounds of transfection and sorting can be of advantage to enrich for gene-edited single clones. Therefore, seed all positive cells coming out of the first sort. Once the cells have recovered, and have reached an appropriate confluency, transfect them again, followed by a second sorting step.
27. The addition of 5 μM Rock inhibitor Y-27632 for 1–2 days to the medium (50% conditioned and 50% fresh medium) improves viability of single cell-derived clones.
28. It is recommended to use the Bambanker cryo-preservation solution instead of 10% DMSO when freezing iPSCs in a multi-well plate. Bambanker enhances survival after thawing hiPSCs.

References

1. Shui B, Hernandez Matias L, Guo Y, Peng Y (2016) The rise of CRISPR/Cas for genome editing in stem cells. *Stem Cells Int* 2016:8140168. <https://doi.org/10.1155/2016/8140168>
2. Jinek M, Chylinski K, Fonfara I et al (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816. LP-821
3. Barrangou R (2014) Cas9 targeting and the CRISPR revolution. *Science* 344:707. LP-708
4. Hess GT, Tycko J, Yao D, Bassik MC (2017) Methods and applications of CRISPR-mediated base editing in eukaryotic genomes. *Mol Cell* 68:26–43. <https://doi.org/10.1016/j.molcel.2017.09.029>
5. Hsu PD, Lander ES, Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157:1262–1278. <https://doi.org/10.1016/j.cell.2014.05.010>
6. Ma D, Liu F (2015) Genome editing and its applications in model organisms. *Genom*

- Proteom Bioinformat 13:336–344. <https://doi.org/10.1016/j.gpb.2015.12.001>
7. Chu VT, Weber T, Wefers B et al (2015) Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotechnol* 33:543–548. <https://doi.org/10.1038/nbt.3198>
 8. Kim EJ, Kang KH, Ju JH (2017) CRISPR-Cas9: a promising tool for gene editing on induced pluripotent stem cells. *Korean J Intern Med* 32:42–61. <https://doi.org/10.3904/kjim.2016.198>
 9. Yang D, Scavuzzo MA, Chmielowiec J et al (2016) Enrichment of G2/M cell cycle phase in human pluripotent stem cells enhances HDR-mediated gene repair with customizable endonucleases. *Sci Rep* 6:1–15. <https://doi.org/10.1038/srep21264>
 10. Li H, Beckman KA, Pessino V, et al (2017) Design and specificity of long ssDNA donors for CRISPR-based knock-in. <https://doi.org/10.1101/178905>
 11. Yumlu S, Stumm J, Bashir S et al (2017) Gene editing and clonal isolation of human induced pluripotent stem cells using CRISPR/Cas9. *Methods* 121–122:29–44. <https://doi.org/10.1016/j.ymeth.2017.05.009>
 12. Mohr SE, Hu Y, Ewen-Campen B et al (2016) CRISPR guide RNA design for research applications. *FEBS J* 283:3232–3238. <https://doi.org/10.1111/febs.13777>
 13. Merkle FT, Neuhausser WM, Santos D et al (2015) Efficient CRISPR-Cas9-mediated generation of knockin human pluripotent stem cells lacking undesired mutations at the targeted locus. *Cell Rep* 11:875–883. <https://doi.org/10.1016/j.celrep.2015.04.007>
 14. Cui Y, Xu J, Cheng M et al (2018) Review of CRISPR/Cas9 sgRNA design tools. *Interdiscip Sci* 10(2):455–465. <https://doi.org/10.1007/s12539-018-0298-z>
 15. Peters D, Cowan C, Musunuru K (2014) Genome editing in human pluripotent stem cells. In: *StemBook*. Harvard Stem Cell Institute, Cambridge, MA. <https://doi.org/10.3824/stembook.1.94.1>



Chapter 11

Editing the Genome of Human Induced Pluripotent Stem Cells Using CRISPR/Cas9 Ribonucleoprotein Complexes

Michaela Bruntraeger, Meg Byrne, Kathleen Long, and Andrew R. Bassett

Abstract

Genome editing using the CRISPR/Cas9 system has rapidly established itself as an essential tool in the genetic manipulation of many organisms, including human cell lines. Its application to human induced pluripotent stem cells (hiPSCs) allows for the generation of isogenic cell pairs that differ in a single genetic lesion, and therefore the identification and characterization of causal genetic variants. We describe a simple, effective approach to perform delicate manipulations of the genome of hiPSCs through delivery of Cas9 RNPs along with ssDNA oligonucleotide repair templates that can generate mutations in up to 98% of single cell clones and introduce single nucleotide changes at an efficiency of up to 40%. We describe our use of a T7 endonuclease assay to identify active guide RNAs, and a high-throughput sequencing genotyping strategy that allows the identification of correctly edited clones. We also present our experiences of generating single nucleotide changes at 15 sites, which show considerable variability between both guides and target sites in the efficiency at which such changes can be introduced.

Key words CRISPR/Cas9, Genome editing, Human iPSC, RNP, Point mutation, SNP

1 Introduction

Robust gene editing technologies based on the CRISPR/Cas9 system have greatly facilitated genetic manipulation in many different contexts, and enabled genetic analysis of human cell lines at a scale that was hitherto impractical. Similarly, the advent of human induced pluripotent stem cells (hiPSCs) has made it possible to produce a renewable source of untransformed cell lines from hundreds of individuals. These can be differentiated into a multitude of different cell types to model the cellular phenotypes underlying human disease states, many of which (such as neurons) would be impossible to study except in postmortem tissues. The combination of these two technologies provides a powerful means to understand the genetic contribution to particular diseases, to revert

Michaela Bruntraeger, Meg Byrne, and Kathleen Long contributed equally to this work.

Yongjun Luo (ed.), *CRISPR Gene Editing: Methods and Protocols*, Methods in Molecular Biology, vol. 1961, https://doi.org/10.1007/978-1-4939-9170-9_11, © Springer Science+Business Media, LLC, part of Springer Nature 2019

the causal genomic lesions, and provide potential therapeutic opportunities.

Here we describe a simple and effective method to generate single nucleotide changes in the genome of hiPSCs to generate isogenic cell pairs differing in a single genetic lesion. This is a powerful means to understand the importance of a specific genetic change in a disease state by removing the influence of the genetic background, which can often confound or mask phenotypes (especially subtle ones) when comparing iPSCs derived from patients with healthy individuals [1].

The RNA-guided Cas9 endonuclease from *Streptococcus pyogenes* (*SpCas9*) consists of three components—the Cas9 protein, a CRISPR RNA (crRNA), and a trans-activating CRISPR RNA (tracrRNA) [2]. The first 20 nt of the crRNA determines the specificity of the enzyme by base-pairing with its complementary target site in the genome, which is therefore highly predictable. The choice of target sites is only limited by the need for a protospacer adjacent motif (PAM) sequence which is necessary immediately adjacent to this 20 nt sequence in the genome, although is not actually present in the crRNA itself (Fig. 1). In the case of the most widely used *SpCas9* enzyme, the PAM is NGG, which can be present on either DNA strand, and given an even distribution of each base, should occur approximately every 8 bp.

The Cas9 endonuclease is used to introduce a site-specific double strand break (DSB) essentially anywhere in the genome. This lesion is repaired by the endogenous cellular repair pathways of non-homologous end joining (NHEJ) or various forms of homology directed repair (HDR), although in most mammalian cell types (including hiPSCs) the NHEJ-based pathways predominate [3, 4]. Both forms of repair can be exploited to introduce changes in the genome; NHEJ is somewhat error-prone and can result in small insertions and deletions at the cut site, which can be used to delete small genomic elements or introduce frameshifts in protein coding genes. The method that we present here exploits the HDR pathway by supplying a short ~100 nt ssDNA oligonucleotide (ssODN) that is used as a template for repair, and can be used to introduce defined changes such as single nucleotide polymorphisms, or short insertions or deletions of exogenous sequence.

Our lab has found that the use of Cas9 ribonucleoprotein (RNP) complexes with synthetic guide RNAs (either crRNA/tracrRNAs or full length sgRNAs) combined with a single-stranded DNA oligonucleotide (ssODN) as a repair template is a highly effective way to introduce single nucleotide changes. We describe how to design the reagents, assemble and deliver the RNP complex, test for efficiency using a T7 endonuclease assay (T7E1), and generate and analyze single cell clones using a high-throughput sequencing (HTS) genotyping strategy (Fig. 2).

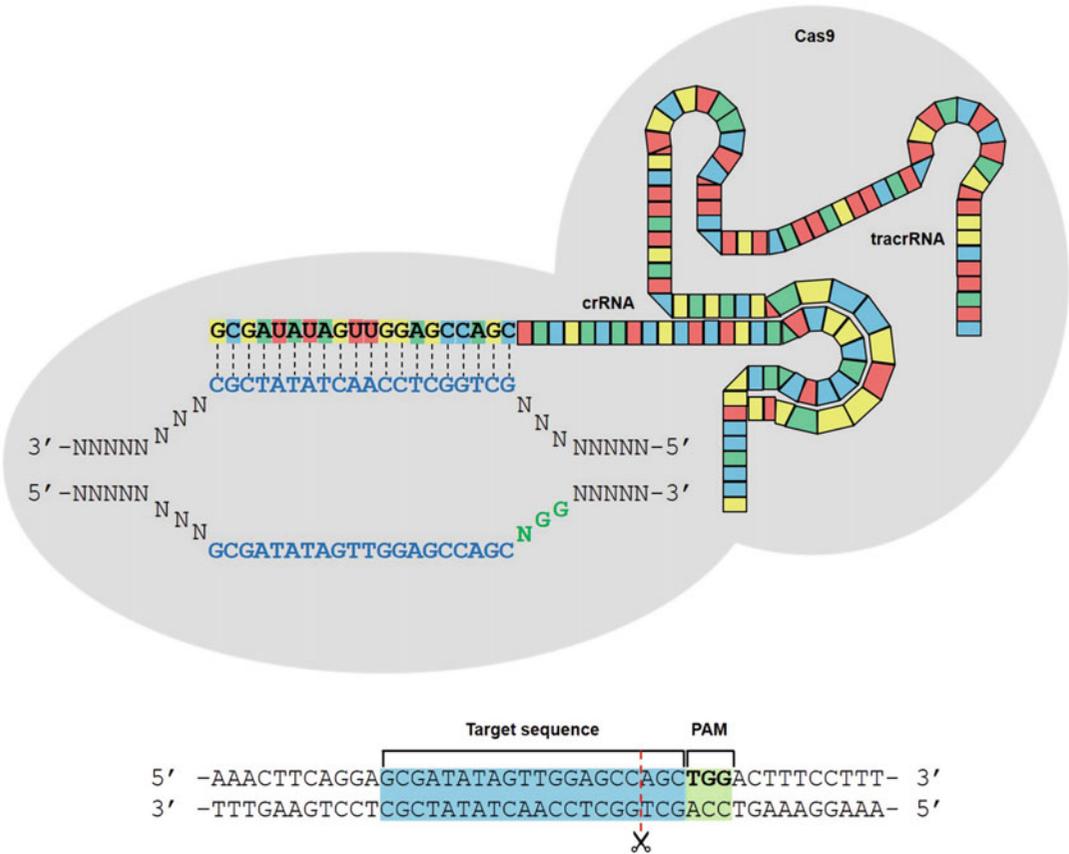


Fig. 1 Schematic of the CRISPR/Cas9 system used for gene editing. The CRISPR RNA (crRNA) contains a unique 20 nt sequence that is complementary to the respective target sequence in the genome. Together with the trans-activating CRISPR RNA (tracrRNA), this complex guides the Cas9 protein to a specific region where it is in turn responsible for generating the double-stranded break three nucleotides away from the protospacer-adjacent motif (PAM)

This method has many advantages over other systems, namely that: (1) The reagents are immediately active and rapidly degraded over 12–24 h, reducing the chances of off-target mutagenesis, minimizing mosaicism, and maximizing HDR rates [5, 6]. (2) There are no cloning steps involved, minimizing the time and effort necessary to generate and validate such constructs. (3) Nonspecific integration of DNA is minimized, since there is no dsDNA being introduced into the cell. (4) Rates of HDR are often increased when using ssDNA templates compared to dsDNA due to use of somewhat different DNA repair mechanisms [7–9]. (5) Toxicity and innate immune responses resulting from transfection of dsDNA plasmids is reduced considerably through use of the RNP complex. (6) Screening by HTS allows high sensitivity to identify mosaic clones and scalability to screen large numbers of single cell colonies. It is also worth noting that a very similar

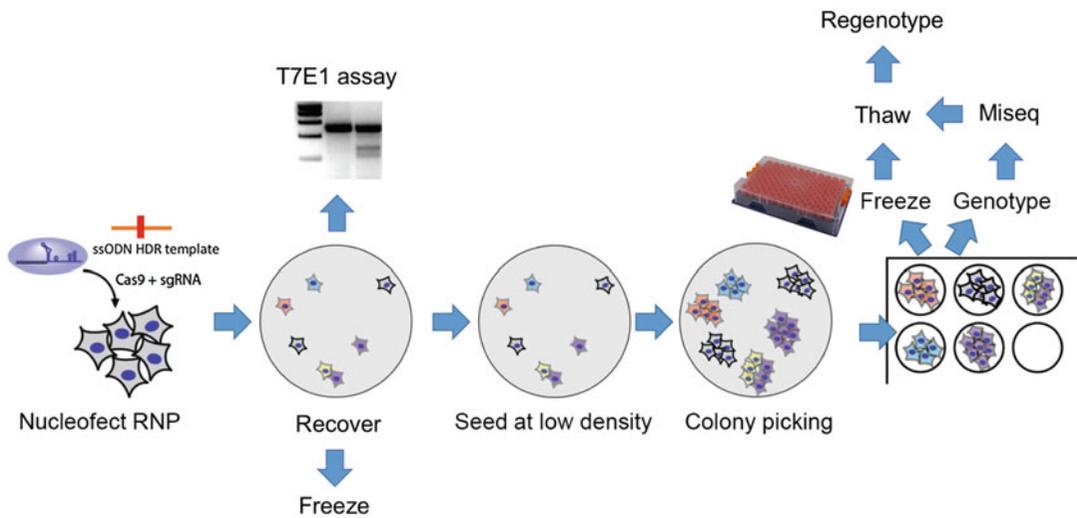


Fig. 2 Workflow of genome editing in hiPSCs. hiPSCs are nucleofected with Cas9 ribonucleoprotein complexes together with a short ~100 nt homology directed repair template. After recovery, cell pools are analyzed for CRISPR efficiency using a T7E1 assay, frozen to maintain a stock and seeded at low density to generate clonal populations. Colonies are picked, and duplicate plates are made for freezing in 96-well matrix vials, and genotyping by high-throughput sequencing (MiSeq). Final clones are grown up and genotypes validated by Sanger sequencing

strategy without a ssODN HDR template can be used to generate knockout mutations through frameshift in protein coding sequences, but in order to maintain effective electroporation, it is important to add a non-targeting ssDNA oligonucleotide.

2 Materials

1. TeSR-E8 (Stem Cell Technologies, 05990).
2. SyntheMAX (Corning, 3535) (*see Note 1*).
3. 10 cm tissue-culture treated dishes.
4. Distilled water.
5. D-PBS.
6. Gentle Cell Dissociation Reagent.
7. 4D Nucleofector System (Lonza, Core unit AAF-1002B and X unit AAF-1002×).
8. CrisprRNAs and tracrRNAs or full length synthetic guide RNAs (sgRNAs) can be purchased from a number of suppliers (e.g., IDT, Sigma, Synthego) (*see Note 2*).
9. Cas9—either expressed and purified from *E. coli*, or purchased (Feldan, A034b-a-0500PMOL or IDT, 1081058).

10. ssODN repair template (IDT, Ultramer DNA Oligo, *see Note 3*).
11. Cas9 storage buffer: 10 mM Tris-HCl, pH 7.4, 300 mM NaCl, 0.1 mM EDTA, 1 mM DTT.
12. P3 Primary Cell 4D-Nucleofector[®] X Kit L (Lonza, V4XP-3012).
13. Accutase.
14. Rock inhibitor Y-27632.
15. 6-well plates.
16. Yolk Sac Lysis Buffer: 10 mM Tris-HCl pH 8.3, 50 mM KCl, 2 mM MgCl₂, 0.45% IGEPAL CA-630, 0.45% Tween 20.
17. Proteinase K.
18. Oligonucleotide primers (*see Note 4*).
19. T7 endonuclease I (T7E1) with buffer.
20. KAPA HiFi HotStart ReadyMix (2×) (KapaBiosystems, KR0370).
21. CloneR (Stem Cell Technologies, 05888).
22. 96-well plates.
23. U-bottom 96-well plate.
24. Matrigel, hESC-qualified (Corning, 354277).
25. Knock-Out Serum Replacement (Life Technologies, 10828028).
26. DMSO.
27. Matrix plates (Thermo, 3725).
28. Mineral oil (Sigma, M8410).
29. E-Gel Low Range Quantitative DNA Ladder.
30. E-Gel 96 Agarose Gels, 2%.
31. E-Base Integrated Power System for running E-Gels: Mother E-base device.
32. SPRI beads.
33. DNA LoBind tubes.
34. KAPA library quantification kit.
35. 24-well plates.
36. PCR Purification Kit.

3 Methods

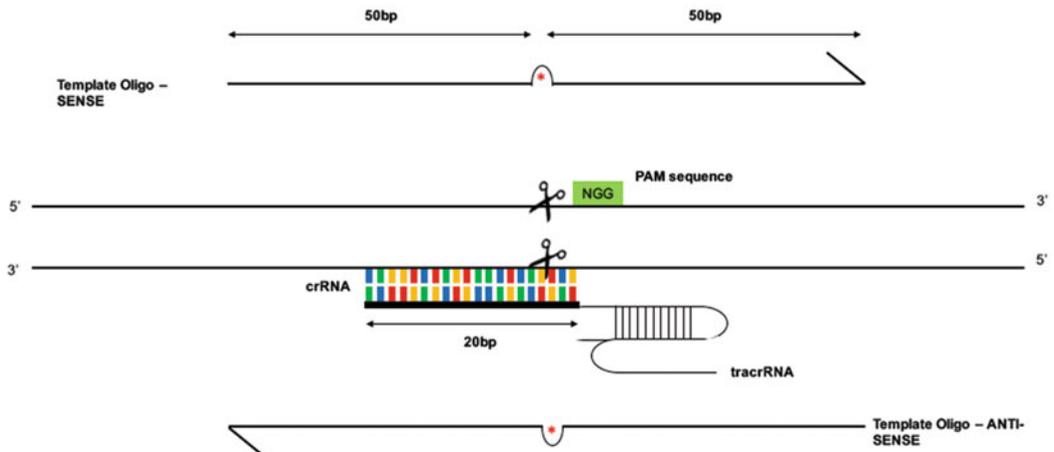
3.1 Gene Targeting Strategy

The choice of CRISPR target site depends on a balance between the position of the cut site in the genome, the specificity of the guide RNA in terms of unwanted or off-target cutting, and the

activity of the Cas9 enzyme when programmed with a particular guide. When introducing single nucleotide changes, the position is the major factor that needs to be considered, since the frequency of introduction of mutations by HDR drops off very rapidly within a few tens of nucleotides either side of the cut site [7, 8]. Ideally, target sites should be chosen such that the cut site is as close as possible to the mutation to be introduced, and preferably no more than 10 nt either side, which normally limits the choice of guides to 1–3 for each desired change. In the case of *SpCas9*, the cut site is 3 nt away from the PAM (i.e., between positions 17 and 18 of the crRNA) (Fig. 3a). Note that in some cases, it is not possible to find an appropriate PAM sequence nearby, in which case it is possible to use NAG PAM sequences, albeit at an approximately 10× lower frequency [10]. Recent protein engineering has resulted in Cas9 enzymes with broadened PAM specificity to include NGA (VQR [11]) and NGN (xCas9 [12]) which could also be employed in the future.

It is well established that *SpCas9* can tolerate mismatches between the guide RNA and target DNA and still maintain its ability to cut DNA [13]. This depends on both the position and number of mismatches, those at the 5′ end of the guide RNA being more tolerated than those at the 3′ (PAM-adjacent) end [13]. In general, we have found that off-targets with 3 or more mismatches are unlikely to be a problem unless the mismatches are clustered at the 5′ end of the guide. Target sites should be screened for potential off-target sites elsewhere in the genome, and if possible chosen to minimize such effects. Given the limitations in choice of gRNA for introducing SNPs, it is sometimes unavoidable to choose those guides with potential off-targets, and in this case, off-targets can be screened for mutations by PCR amplification and sequencing. Many software tools are available to do this, including WGE (<http://www.sanger.ac.uk/htgt/wge/>) [14], which provides pre-computed off-target data for all exonic regions in the genome, and guides should be screened using such software. We also routinely use engineered high-fidelity variants of *SpCas9* (e.g., *eSpCas9_1.1* [15]) to mitigate this problem. Although we have seen very low frequencies of off-target mutation in our system, we cannot exclude the possibility that they exist, so it is important when designing an experiment to control for such effects. This can be achieved for instance by use of two independent guides to introduce the same mutation, which will have completely different off-target sets. Minimally, it is good practice to analyze multiple mutant and control clones, which are highly unlikely to contain the same off-target mutation. It is also sometimes possible to complement the mutation by for instance expression of WT cDNA, or by repair of the mutation in the newly mutated line by a second round of engineering.

A) Guide and Templates:



B) Primers:

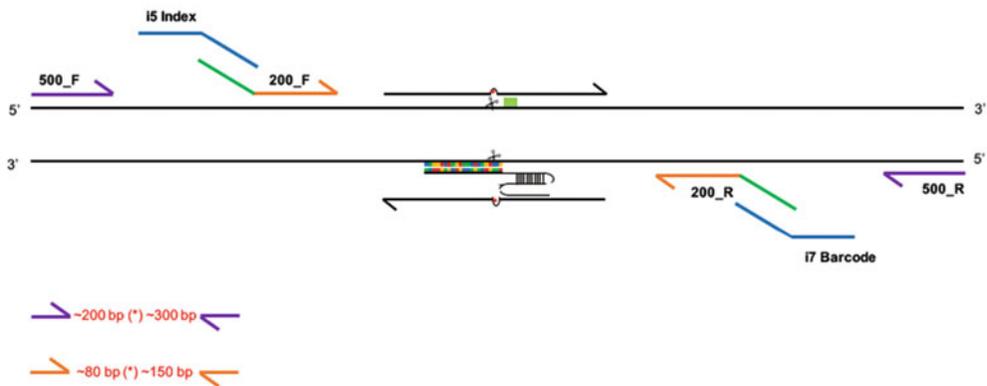


Fig. 3 Schematic illustrating guide RNA, HDR template, and primer design. **(a)** The unique crRNA sequence is the 20 nucleotides directly upstream of the PAM sequence (NGG). HDR templates are single-stranded DNA oligonucleotides ordered in both the sense and antisense direction. They are designed to be ~100 bp in length with the mutation (asterisk) situated in the middle. **(b)** 3 pairs of primers are required for sequence analysis. (1) A set to produce a 500 bp amplicon (500_F and R), (2) a nested pair to produce a 200 bp amplicon with universal appends (200_F and R), and (3) an i5 and i7 primer containing the remainder of the Illumina sequencing adaptors and barcodes to allow for unique identification of DNA fragments

On-target activity is another important consideration when choosing guide RNAs, and there are a number of available pieces of software that can predict the efficacy of a particular guide [16]. In our hands, we have only seen a weak correlation of such predictions to the actual activity of a guide, and find that around 20–25% of guides are not usable due to low activity. Therefore, we routinely screen 2–3 guides per target site by nucleofection individually into cells followed by a T7E1 assay on the pooled

population (Subheading 3.4, Fig. 2). This provides a good measure of the efficiency of each guide, and we normally only continue to single cell cloning with those guides that give a measurable signal in this assay. It also ensures that the nucleofection and Cas9 endonuclease have been effective before substantial time, effort, and money are invested in single cell cloning and genotyping.

In terms of ssODN homology construct design, we have found that the total length of homology should be a minimum of 80 nt, and we normally use ~100 nt oligos with 50 nt of homology either side of the desired change(s). Importantly, when designing the homology construct, it is imperative to block the ability of Cas9 to cut after the desired change has been introduced. If this is not done, Cas9 will introduce further double strand breaks, resulting in small insertions and deletions as well as the intended event. We find that single nucleotide changes in the PAM sequence, or within 5 bases of the 3' end of the guide are sufficient to block re-cutting in most instances, so if the desired change is within this region there is no necessity to introduce further mutations. If this is not the case, additional silent mutations, ideally within the PAM sequence, or in the PAM-adjacent region of the target site can be introduced. Make sure that you do not mutate the PAM from NGG to NAG, since Cas9 still maintains some activity at this sequence, albeit at a lower efficiency in general [13]. We have observed some difference between sense and antisense ssODN templates, and this appears to depend on the position of the cut site relative to the mutation that is introduced (*see Note 5*). We have not observed any difference in efficiency between different purification strategies, so normally use desalted oligonucleotides, although PAGE or HPLC purified ssODNs are equally as effective.

3.2 hiPSC Culture

We routinely use the KOLF2_C1 induced pluripotent stem cell (iPSC) line for editing, which was derived by the Human Induced Pluripotent Stem Cell Initiative (HipSci) consortium [17], and has additionally been derived from a single cell to maximize homogeneity within the population. It is cultured under feeder-free conditions typically on 10 cm tissue culture dishes in TeSR-E8 medium on SyntheMAX substrate at a final coating concentration of 2 $\mu\text{g}/\text{cm}^2$. Cells are routinely passaged 1:5 every 4–5 days. Cells are cultured in a humidified incubator at 37 °C and 5% CO₂.

1. Remove spent medium and wash cell layer once with PBS.
2. Incubated with Gentle Cell Dissociation Reagent (5 mL/10 cm dish) at 37 °C for 3 min.
3. Aspirate Gentle Cell Dissociation Reagent and add 10 mL TeSR-E8 medium.
4. Use a cell scraper to manually detach dissociated colonies from the culture surface (*see Note 6*).

Table 1
Composition of crRNA:tracrRNA complex

Vol (μL)	Component	Amount (pmoles)
1	crRNA	200
1	tracrRNA	200
2.5	IDT Duplex buffer	

5. Triturate cell suspension once with a 10 mL stripette to break up larger clumps of cells into smaller clumps.
6. Transfer 2 mL cell suspension into a fresh 10 cm dish containing an additional 8 mL TeSR-E8 medium.

3.3 Gene Targeting by Nucleofection

1. Resuspend RNA oligos (crRNA and tracrRNA) in IDTE solution to a concentration of 200 μM .
2. Mix crRNA and tracrRNA in equimolar concentrations in a sterile microfuge tube to a final concentration of 45 μM of each using IDT duplex buffer (Table 1).
3. Anneal the gRNA complex by heating to 95 $^{\circ}\text{C}$ for 2 min.
4. Remove from heat source and cool slowly to room temperature.
5. Prior to RNP complex formation, Cas9 protein should be diluted to a final concentration of 4 $\mu\text{g}/\mu\text{L}$ in Cas9 storage buffer (Subheading 2, item 11) (*see* Note 7).
6. Combine 5 μL (20 μg , ~120 pmol) Cas9 with 5 μL (225 pmol) gRNA and incubate at RT for 10–20 min.
7. Remove spent medium from 10 cm dish and wash once with PBS. Ensure that cells are still actively growing, by using cells that have been passaged 3–5 days earlier, and that are approximately 70% confluent (Fig. 4).
8. Add 5 mL of accutase and leave for 6 min at 37 $^{\circ}\text{C}$.
9. Aspirate accutase and scrape the cells in 10 mL TeSR-E8 supplemented with rock inhibitor.
10. Triturate cells to create a single cell suspension.
11. Count the cells and aliquot 1×10^6 cells into a sterile microfuge tube.
12. Centrifuge at $300 \times g$ for 3 min to pellet.
13. Resuspend in 100 μL complete P3 buffer.
14. To use a ssODN repair template (to make base changes by HDR), add 500 pmol of purified ssDNA oligonucleotide (5 μL 100 μM stock) to the RNP complex.

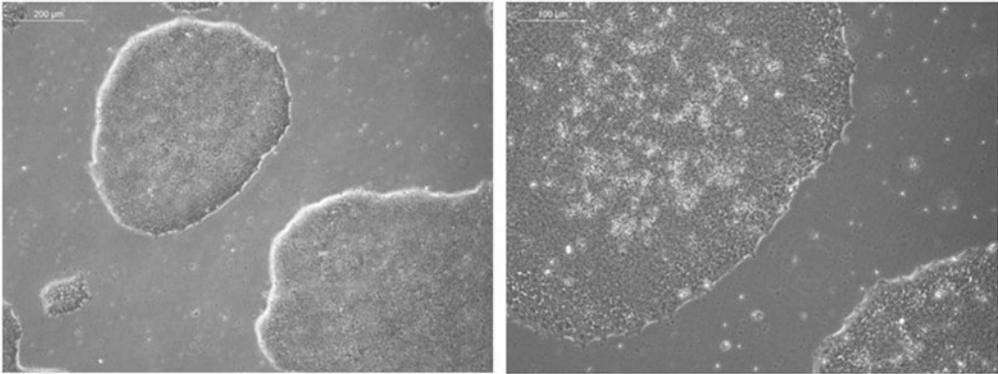
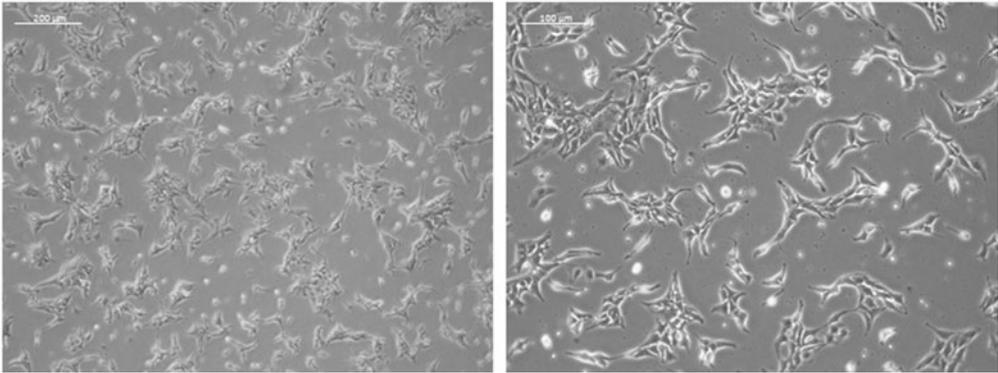
A) Pre-nucleofection:**B) Post-nucleofection:**

Fig. 4 Cell morphology pre-/post-nucleofection. The state of the cells is important in achieving good editing efficiencies. **(a)** KOLF2-C1 cells under a 5× and 10× objective showing ideal morphology of cells prior to nucleofection. Note the defined borders of the colonies and even distribution and compaction of the cells. **(b)** KOLF2-C1 cells under a 5× and 10× objective respectively post-nucleofection. Individual cells will appear less round and more elongated or “spiky.” This is a result of treatment with ROCK inhibitor (Ri) and will subside once Ri is removed and cells begin to form small colonies and compact together

15. Mix the RNP complex with the resuspended cells using a swirling motion with the pipette tip (avoid triturating the cells too much).
16. Transfer cells to a nucleofection cuvette, dispensing carefully between the electrodes to avoid formation of bubbles. (Final amounts per nucleofection: 1×10^6 cells in 100 µL P3 solution, 5 µL (20 µg) Cas9 protein, 5 µL (20 µg) sgRNA).
17. Electroporate with the 4D-Nucleofector on program CA 137.
18. Transfer cells from the cuvette across 2 wells of a 6-well plate (pre-coated with SyntheMAX at 5 µg/cm² containing 2 mL of TeSR-E8 + 10 µM Rock inhibitor) using a disposable Pasteur pipette. Do not pipette up and down, since cells are very fragile at this point (*see* **Note 8**).

19. Incubate for 24 h, then exchange medium for TeSR-E8 without Rock inhibitor (Fig. 4).
20. Feed daily until cells reach approximately 70% confluence for subcloning.

At this stage, it is wise to confirm CRISPR activity using, e.g., T7E1 or TIDE analysis [18] (*see Note 9*). This is an important step before continuing with subcloning, since many guide RNAs (~20–25%) have very low activity, and it is possible to eliminate these at this stage.

3.4 Screening CRISPR Efficacy with a T7E1 Assay

1. For a single, confluent well of a 6-well plate, prepare 1 mL of Yolk Sac Lysis Buffer and supplement it with 20 μ L Proteinase K.
2. Aspirate medium from the well and wash cell layer with PBS, aspirating again.
3. Add the Lysis Buffer mixture to the well and allow to stand for 30–60 s.
4. Collect the lysate in a microfuge tube and incubate at 60 °C for at least 1 h to overnight (*see Note 10*).
5. Inactivate the Proteinase K by incubating the lysate at 95 °C for 10 min.
6. Dilute the lysate 1:10 with nuclease-free water.
7. Use the diluted lysate to prepare a 20 μ L PCR mix according to Table 2, using the set of primers designed to generate a 500 bp product (500_F/R).
8. Run the PCR with the cycling protocol described in Table 3 (*see Note 11*).
9. To 10 μ L of the PCR product, add 1.5 μ L NEBuffer 2 (10 \times) and 1.5 μ L Nuclease-free H₂O. Include a second control reaction which will receive no T7E1.
10. Form the hetero-duplexes by heat-cooling the mixture in a thermal cycler using the protocol described in Table 4.
11. Add 2 μ L of T7E1 (Diluted in NEBuffer 2 (1 \times) to a final concentration of 1 U/ μ L). For control digests, use 2 μ L NEBuffer 2 (1 \times).
12. Incubate at 37 °C for 1 h.
13. Add an appropriate volume of loading dye to the samples.
14. Run the samples on a 2.5% agarose gel at 70 V for 45 min (*see Fig. 5*).

3.5 Subcloning

Edited cells are expanded to ~70% confluency following nucleofection before subcloning. Approximately 1000–2000 cells are subcloned onto 10 cm tissue culture dishes pre-coated with SyntheMAX substrate at a concentration of 5 μ g/cm².

Table 2
Composition of a PCR mix using KAPA HiFi HotStart ReadyMix (2×)

Component	Volume (μL)	Final concentration (μM)
DNA Template	2	
KAPA HiFi HotStart ReadyMix (2×) (KapaBiosystems, KR0370)	5	1×
Forward primer (10 μM)	0.3	0.3
Reverse primer (10 μM)	0.3	0.3
Nuclease-Free Water	Up to 10	

Reaction volumes can be adjusted between 10 and 50 μL by scaling the values stated proportionately

Table 3
PCR cycling program

Step	Temperature (°C)	Duration	Cycle
Initial denaturation	95	3 min	1
Denaturation	98	20 s	35
Annealing	60–75	15 s	
Extension	72	30 s	
Final Extension	72	5 min	1

We recommend optimizing your primers beforehand in order to know accurate annealing temperatures

Table 4
Denaturation—annealing thermocycler program to obtain heteroduplexes

Temperature (°C)	Duration
95	10 min
95–85	Ramp –2 °C/s
85–25	Ramp –0.3 °C/s

1. Remove spent medium and wash mixed pool of edited cells once in PBS.
2. Add 1 mL pre-warmed accutase/well of 6-well plate and incubate at 37 °C for 6 min.
3. Remove the accutase and scrape cells in 10 mL TeSR-E8 supplemented with 1× CloneR.

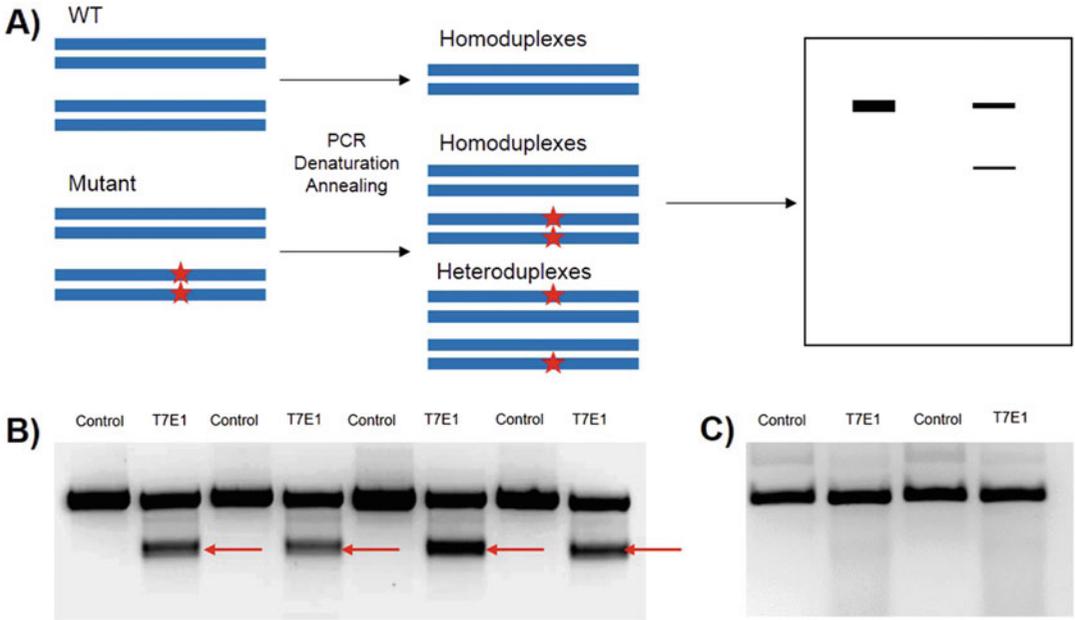


Fig. 5 T7 Endonuclease Assay to assess CRISPR efficiency. (a) The T7E1 assay is based on the formation of homo- and heteroduplexes between PCR amplified genomic WT and mutant DNA fragments. The latter will contain mismatches which are identified and cleaved by the T7E1 enzyme resulting in either cleaved or full-length fragments which can be visualized and distinguished on gel. (b) Example of a positive T7E1 assay, indicating good activity of the guide at this target site. (c) Example of a negative T7E1 assay, indicating poor/nonexistent activity of the guide at this target site

4. Triturate cells to dissociate into single cells for counting (*see Note 12*).
5. Serially dilute cell suspension in order to generate a single cell suspension at 1×10^3 cells/mL.
6. Transfer 1–2 mL of cells at 1×10^3 /mL to a new 10 cm dish with an additional 9 mL TeSR-E8 + CloneR and evenly dispersed across the surface.
7. Transfer to a 37 °C incubator.
8. Cells should be left undisturbed without feeding for 48 h. Afterward, they should undergo daily feeding with 10 mL TeSR-E8 medium for 8–10 days until colonies are approx 1–2 mm in diameter (*see Note 13*).

3.6 Colony Picking

Single cell-derived colonies are manually picked from the culture surface approximately 8–10 days after subcloning. At this point colonies should be 1–2 mm in size, uniform, and round in appearance with bright borders (Fig. 6) (*see Note 14*). Prior to picking, equal numbers of 96-well plates should be prepared with Matrigel or SyntheMAX substrate at a concentration of $5 \mu\text{g}/\text{cm}^2$. Each well should be prefilled with 150 μL TeSR-E8 medium supplemented with rock inhibitor.

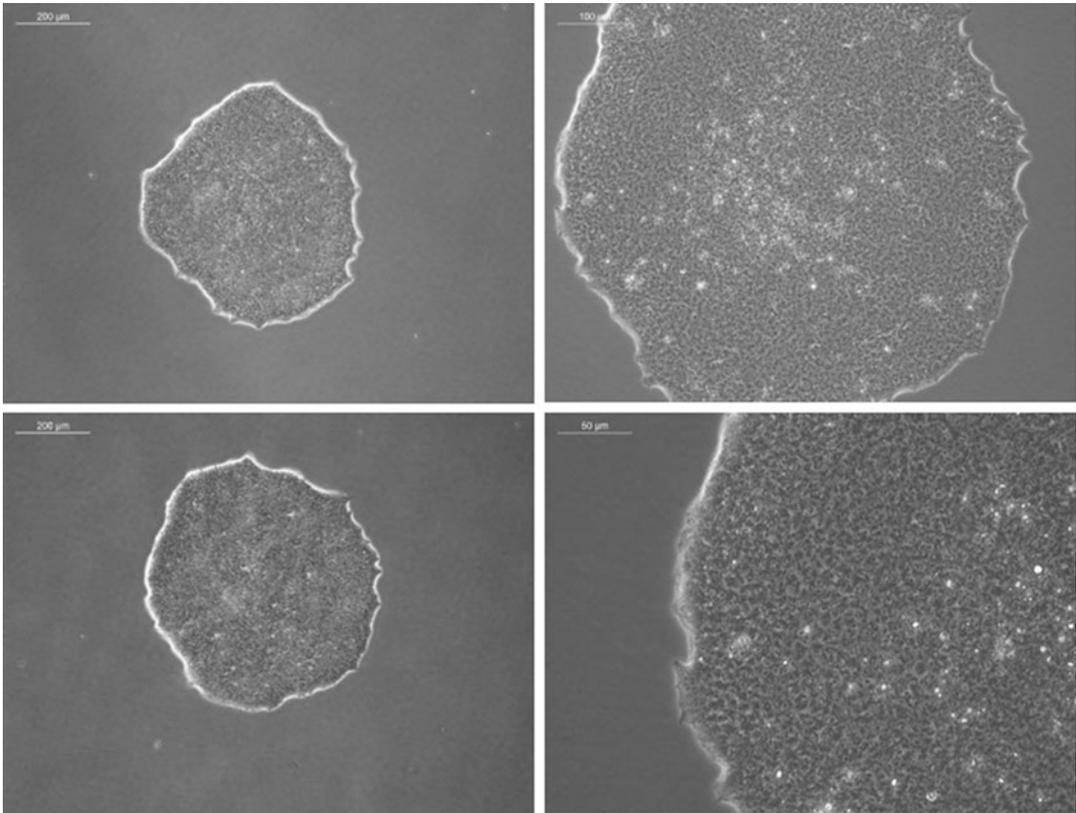


Fig. 6 Ideal morphology of colonies appropriate for picking. Two colonies ~10 days post subcloning under a 5× and 10× objective respectively. It is impossible to be certain that a colony on a plate is derived from a single iPSC, good indicators however, include colony size of 1–2 mm, round in shape with clear, bright borders. Cells should also be even throughout and colonies should not be touching one another

1. Aspirate the medium from the 10 cm dish of colonies.
2. Wash once in PBS.
3. Add 5 mL of gentle cell dissociation buffer and incubate at 37 °C for 3 min (*see Note 15*).
4. Aspirate the Gentle Cell Dissociation Reagent taking care not to disrupt the colonies.
5. Add 10 mL TeSR-E8 medium.
6. Using a dissection microscope and a P20 pipette (set to 10 μL), manually pick individual colonies by scraping from the culture surface and collecting in the pipette.
7. Transfer the cells to a U-bottom 96-well plate containing 100 μL TeSR-E8 medium supplemented with rock inhibitor. Use a fresh tip per colony to prevent cross-contamination.
8. Colonies are usually picked in batches of 96 which can be processed for genotyping as whole plates. Once sufficient numbers of colonies have been picked, triturate the cells 5–6 times to break colonies into smaller fragments (*see Note 16*) and transfer 50 μL into one matrigel and the other 50 μL into

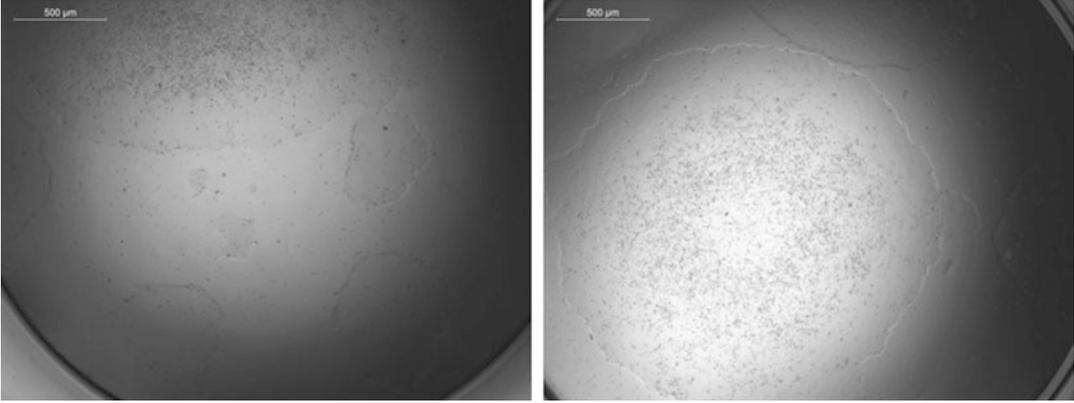


Fig. 7 Freezing from a 96-well plate. Two wells of cells in a 96-well plate that are ready to freeze viewed under a 2.5 \times objective 4–5 days after picking. Typically, confluency levels range between 80 and 100%

one SyntheMAX-coated plate, thereby creating two duplicate 96-well plates of colonies (*see Note 17*).

9. Cells should be cultured for approximately 5 days (or until wells are confluent) with daily 200 μ L TeSR medium changes, before freezing/lysing (*see Note 18*).

3.7 Freezing

Cells are routinely frozen 4–5 days (Fig. 7) after colony picking from the matrigel plates (*see Note 19*).

1. Remove spent medium.
2. Wash colonies once in PBS.
3. Add 25 μ L pre-warmed Accutase/well and incubate for 6 min at 37 $^{\circ}$ C.
4. Quench the reaction with 75 μ L Knock-Out Serum Replacement.
5. Triturate the samples 6–7 times, then add 100 μ L Knock-Out Serum Replacement containing 20% DMSO (final concentration of 10% DMSO).
6. Mix 3–4 times with a pipette, and transfer the entire contents into a 96-well matrix plate (*see Note 20*).
7. Add 100 μ L sterile mineral oil over the top of the freeze medium.
8. Incubate matrix plates at -80° C overnight before transferring to liquid nitrogen for long-term storage.

3.8 Genotyping

When designing a genotyping strategy it is important to consider its ability to identify mixed or mosaic clones (i.e., those colonies deriving from multiple different edited cells), and its scalability to the large number of clones that often must be screened in order to obtain the cell line of interest. Here we present a high-throughput

sequencing based strategy that involves a three-step PCR from crude cell lysates that are easily prepared from 96-well plates, followed by sequencing on an Illumina MiSeq instrument that allows screening of tens of plates on a single run. It is also highly sensitive to detect each individual allele in a mixed population down to <1% of the total alleles, making identification of mixed clones simple and accurate.

Colonies growing on SyntheMAX can be lysed 2–4 days post picking. Lysate plates are used for genotyping of edited clones and are generated in a similar way to the lysates used in the T7E1 assay (Subheading 3.4).

1. Remove spent medium.
2. Wash cells once in PBS.
3. Incubate with 100 μ L Yolk Sac lysis buffer supplemented with 2 μ L Proteinase K/well at 60 °C from 1 h to overnight (*see Note 10*).
4. Inactivate the proteinase K at 95 °C for 10 min.
5. Store lysates at –20 °C or use directly for PCR.
6. Dilute lysates 1:10 in nuclease-free H₂O.
7. Prepare 10 μ L PCR reaction for each well, using the set of primers designed to generate a 500 bp product (500_F/R) (Fig. 3, Table 2).
8. Run PCR thermocycler program (Table 3).
9. Recommended: run 3 μ L of product on a 2% E-gel, made up to 20 μ L in ddH₂O to ensure PCR was successful, with 10 μ L E-gel low range quantitative DNA ladder combined with 10 μ L H₂O.
10. Assemble PCR as in **step 7** (Table 2) using instead the 200_F/R primers, and 0.5 μ L of the round one 500 bp product as template.
11. Run PCR thermocycler program mentioned in **step 8** (Table 3). Recommended annealing temperature of 60 °C.
12. Recommended: run 3 μ L of product on a 2% E-gel, as in **step 9**.

3.8.1 PCR to Generate Barcoded Amplicons

1. Prepare PCR as in Subheading 3.8, **step 7**, using the barcoded i5 and i7 MiSeq primers [19, 20]. Each well of a 96-well plate should have a different i7 primer. If multiple 96-well plates of one gene are being analyzed, each plate should have a different i5 primer. Different amplicons may use the same primer combinations. Use 0.5 μ L of a 1/100 dilution of round 2 product in nuclease-free H₂O for the template.

2. Run PCR thermocycler program mentioned in Subheading 3.8, step 8 (Table 3). Recommended annealing temperature 70 °C.
3. Recommended: run 3 μL of product on a 2% E-gel, as in Subheading 3.8, step 9.
4. Pool 2 μL of each sample and mix well.
5. Add 100 μL of pool to 60 μL of SPRI beads in a LoBind 1.5 mL microfuge tube.
6. Incubate at room temp for 5 min.
7. Place tube on magnet and allow solution to clear. Remove and discard supernatant.
8. Add 200 μL room temperature 80% ethanol, being careful not to disturb beads.
9. Incubate for 30 s, then remove and discard ethanol.
10. Repeat steps 15 and 16 once more.
11. Remove tube from magnet and briefly spin.
12. Replace tube on magnet and remove final traces of ethanol.
13. Air-dry pellet for 2 min, or until all traces of ethanol are gone.
14. Remove tube from magnet and add 52 μL of elution buffer to tube. Resuspend beads.
15. Incubate for 2 min.
16. Place tube on magnet. Once solution is clear, remove 50 μL of purified DNA. Store in DNA LoBind microfuge tube.
17. Quantify DNA concentration using KAPA library quantification kit, according to instructions in technical datasheet (KR0405).
18. Submit DNA for 150 nt paired end sequencing. Aim for a minimum of 1000 reads per sample.
19. Analyze fastq files with CRISPResso (<https://github.com/lucapinello/CRISPResso> [21]). This software can be run from the command line, and a typical desktop computer will be sufficient to analyze hundreds of clones. We have developed a graphical interface for simple identification of clones, but the direct output files will provide all of the information necessary to identify the correct edited clones (Fig. 8).

3.9 Thawing and Expansion

Individual matrix vials containing clones of interest can be recovered from liquid nitrogen storage and reintroduced into culture.

1. Remove vial and thaw rapidly.
2. Transfer contents of vial from under mineral oil into 1 mL pre-warmed TeSR-E8 medium and centrifuge at $300 \times g$ for 3 min.

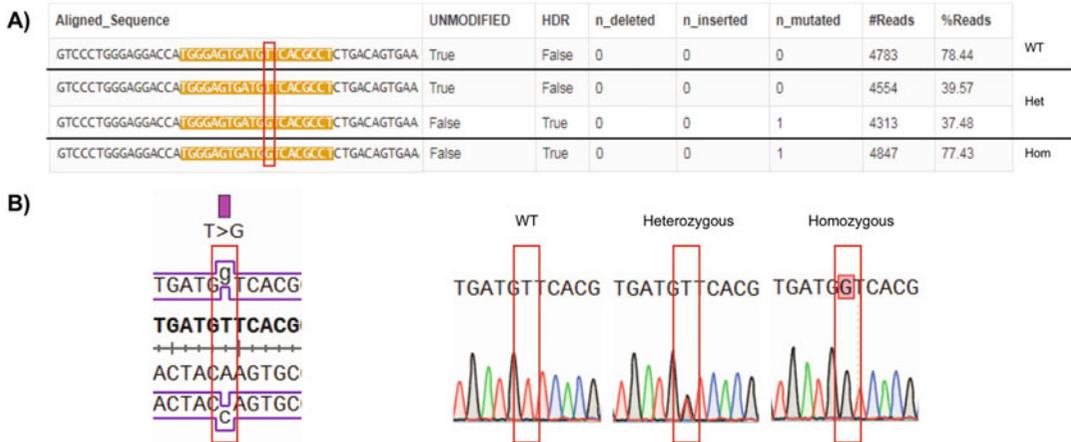


Fig. 8 Primary and secondary genotyping. **(a)** Primary genotyping sequence data of three potential clones processed by MiSeq. CRISPR guide sequences have been highlighted on the aligned sequence and show the number of reads for that sequence and what percentage of the total reads this represents. Due to errors during the PCR and sequencing processes, a single sequence will not represent 100% of the reads, but caution should be taken with sequences present at $>1\%$ of total reads. **(b)** Secondary genotyping of three potential clones by Sanger sequencing. Sequence traces show clones representing WT, heterozygous and homozygous lines and confirm the results from the primary genotyping

- Gently resuspend cell pellet 1 mL TeSR-E8 medium supplemented with CloneR and transfer into a single 24-well (pre-coated with SyntheMAX at $2 \mu\text{g}/\text{cm}^2$).
- Culture overnight at 37°C .
- Perform 1 mL medium change to TeSR-E8 medium and expand as necessary (*see Note 21*).

3.10 Re-genotyping

Secondary sequencing is done on lysates of potential clones of interest that have been thawed and expanded (Fig. 8).

- PCR to generate 500 bp product (Subheading 3.4, step 2).
- Purify the product by PCR purification or gel extraction.
- Sequence the amplicon by Sanger sequencing.
- Analysis of sequence trace can be done on various programs (e.g., SnapGene).

3.11 Expected Results

The efficiency of the editing process depends on a large number of factors, including the efficacy of the chosen guide RNA, the genomic context and sequence of the locus, and the cell line used. Using the KOLF2_C1 hiPSC line, we have observed up to 98% total mutagenic efficiency, counting either indel or HDR based mutations (Fig. 9). This suggests that the delivery of the RNP is virtually 100% efficient, since there is no selection for its uptake, and also that the protein is highly active in order to generate such a high rate of mutations. However, the efficiency of mutagenesis remains

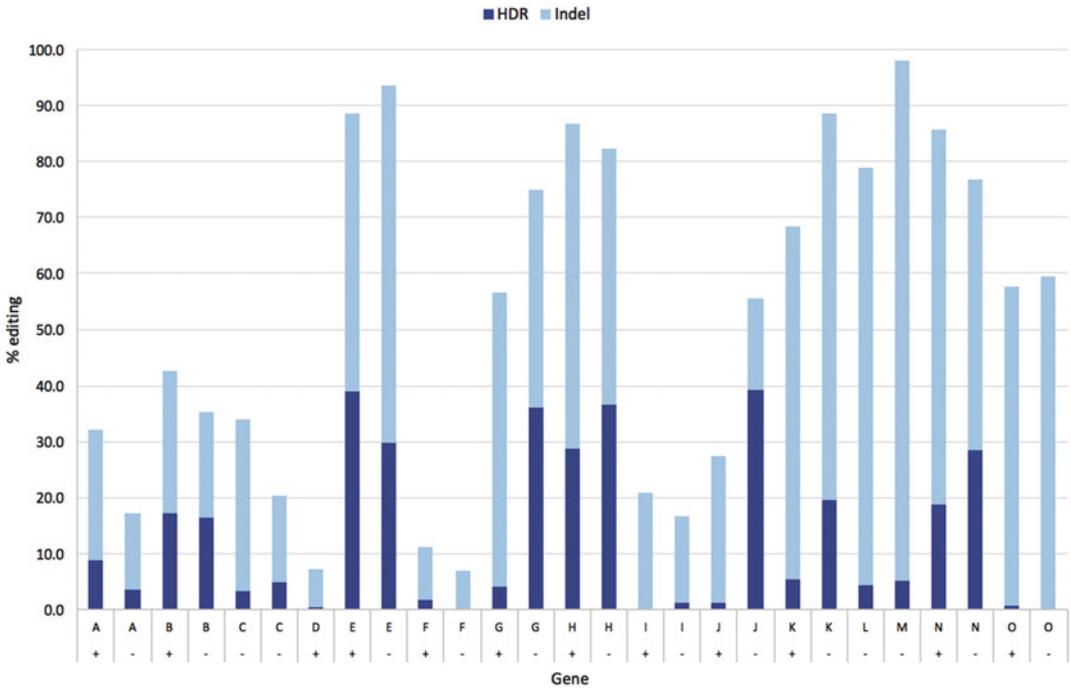


Fig. 9 Analysis of editing efficiency across different genomic loci. Bars show the efficiency of HDR (SNP, dark blue) and indel alleles (light blue) across 15 genomic sites (A–O) highlighting the variation in overall efficiency and proportion of HDR. Strand of HDR template is also indicated (+ or –)

highly variable between sites, with some guide RNAs failing to work altogether (not shown), and some only at very low efficiency. This is likely due to the guide RNA sequence itself, or the chromatin structure at the targeted locus.

In terms of the proportion of mutations that are made up by HDR-driven single nucleotide changes, this is also highly variable between sites, with proportions varying from 0 to 70% of the total mutagenic events. These differences are predominantly due to the different sites which have been targeted, but in some cases, there is also a significant difference between the two strands of ssODN repair template (e.g., sites G or J, Fig. 9), so we recommend performing two experiments per guide, one with the sense ssODN, and a second with the antisense one.

The efficiency of HDR in the best examples is high, with up to 40% of alleles being modified in this way and in some cases more than 25% of single cell clones being homozygous HDR events. Despite this high efficiency at some sites, we still observe a large variability between sites, and very low frequencies of the desired mutation at some loci. Thus, although our system will prove eminently usable in many instances, there is still room for improvement of HDR efficiency. This could be achieved for instance biasing of DNA repair outcomes with small molecules or by manipulation of endogenous cellular factors [22], recruitment of HDR repair

templates to their site of action [23], alterations of cell state and novel methods of selection for the correct event [24, 25]. Equally, natural and engineered CRISPR variants [11, 26] will allow a greater targeting range and enable editing of those sites in the genome currently inaccessible to such mutagenesis due to a lack of appropriate PAM sequences.

The CRISPR/Cas9 system has allowed us for the first time to perform delicate genetic manipulations of human cell lines to understand the functional consequences of genetic mutations. The introduction of single nucleotide variants into iPSCs provides a powerful means to generate and identify causal genetic variants and cellular phenotypes that underlie many genetic diseases, and potentially to manipulate stem cells for therapeutic benefit.

4 Notes

1. Synthemax is prepared by dissolving 10 mg powder in 10 mL sterile distilled water to make a 1 mg/mL stock and stored in aliquots at 4 °C. Further dilution in water 1:40 and 1:100 make a working dilution for coating plates at a concentration of 5 µg/cm² and 2 µg/cm², respectively.
2. We find that full length synthetic, chemically modified guides from Synthego or Sigma are highly effective, and in most cases superior to the crRNA:tracrRNA hybrid. We use the same molar amounts and concentrations as for the crRNA:tracrRNA hybrid, namely an approximately twofold molar excess of guide RNA compared to Cas9 protein.
3. HDR ssDNA oligonucleotides can be ordered from many suppliers, but we typically use Ultramer oligos from IDT, due to the higher proportion of full length products with the correct sequence that are obtained. We find that simple desalted purification is sufficient, but HPLC or PAGE purified oligonucleotides work equally effectively.
4. Three sets of primers are required to allow for analysis of gene editing (Fig. 3b):
 - (a) 500_F/R: A pair of primers producing an amplicon of approximately 500 bp with the mutation of interest situated approximately in the middle. These are used for T7E1 assay and as the first round of a nested set for genotyping.
 - (b) 200_F/R: A pair of primers to generate the amplicon for MiSeq genotyping. They will produce an amplicon of 200–250 bp (max 295 bp to allow assembly of paired end 150 nt reads), with the mutation approximately 80–100 bp from the beginning of the forward primer to maximize sequencing quality. The following appends

consisting of the proximal part of the Illumina sequencing adaptors should be added to the primer

5' ACACTCTTTCCCTACACGACGCTCTTCCGA
TCT-[gene-specific forward] 3'

5' TCGGCATTCTGCTGAACCGCTCTTCCGAT
CT-[gene-specific reverse] 3'

- (c) i5 Index and i7 Barcoded MiSeq primers: These are common to all experiments, and add the remainder of the Illumina sequencing adaptors. Forward primer: i5 index primer, a different primer for each 96-well plate. Reverse primers: i7 index primers, a different index for each well of a 96-well plate (Tables 5 and 6).
5. Recent results suggest that the predominant repair mechanism when using a ssODN is by annealing of the oligo to the resected 5' ends of the DNA that are formed after a double strand break, and its use as a template for DNA polymerases

Table 5
Sequences of the i5 index primers used for MiSeq

Name	Primer sequence	8 base barcode within primer
D501	AATGATACGGCGACCACCGAGATCTAC ACTATAGCCTACACTCTTTCCCTACAC GACGCTCTTCCGAT*C*T	TATAGCCT
D502	AATGATACGGCGACCACCGAGATC TACACATAGAGGCACACTCTTTCCCTACA CGACGCTCTTCCGAT*C*T	ATAGAGGC
D503	AATGATACGGCGACCACCGAGATCTACA CCCTATCCTACACTCTTTCCCTACACGAC GCTCTTCCGAT*C*T	CCTATCCT
D504	AATGATACGGCGACCACCGAGATCTACA CGGCTCTGAACACTCTTTCCCTACACGAC GCTCTTCCGAT*C*T	GGCTCTGA
D505	AATGATACGGCGACCACCGAGATCTACA CAGGCGAAGACACTCTTTCCCTACACGAC GCTCTTCCGAT*C*T	AGGCGAAG
D506	AATGATACGGCGACCACCGAGATCTACA CTAATCTAACACTCTTTCCCTACACGAC GCTCTTCCGAT*C*T	TAATCTTA
D507	AATGATACGGCGACCACCGAGATCTACA CCAGGACGTACACTCTTTCCCTACACGAC GCTCTTCCGAT*C*T	CAGGACGT
D508	AATGATACGGCGACCACCGAGATCTACA CGTACTGACACACTCTTTCCCTACACGAC GCTCTTCCGAT*C*T	GTACTGAC

Table 6
Sequences of the i7 barcode primers used for MiSeq

Name	Primer sequence	8 base barcode within primer
iPCRtag1	CAAGCAGAAGACGGCATAACGATGATGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AACGTGAT
iPCRtag2	CAAGCAGAAGACGGCATAAACATCGGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AAACATCG
iPCRtag3	CAAGCAGAAGACGGCATAATGCCTAAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ATGCCTAA
iPCRtag4	CAAGCAGAAGACGGCATAAGTGGTCAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AGTGGTCA
iPCRtag5	CAAGCAGAAGACGGCATAACACTGTGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ACCACTGT
iPCRtag6	CAAGCAGAAGACGGCATAACATTGGCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ACATTGGC
iPCRtag7	CAAGCAGAAGACGGCATAAGATCTGGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CAGATCTG
iPCRtag8	CAAGCAGAAGACGGCATAACATCAAGTGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CATCAAGT
iPCRtag9	CAAGCAGAAGACGGCATAAGCTGATCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CGCTGATC
iPCRtag10	CAAGCAGAAGACGGCATAACAAGCTAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ACAAGCTA
iPCRtag11	CAAGCAGAAGACGGCATACTGTAGCCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CTGTAGCC
iPCRtag12	CAAGCAGAAGACGGCATAAGTACAAGGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AGTACAAG
iPCRtag13	CAAGCAGAAGACGGCATAACAACCAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AACAACCA
iPCRtag14	CAAGCAGAAGACGGCATAAACCGAGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AACCGAGA
iPCRtag15	CAAGCAGAAGACGGCATAAACGCTTAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AACGCTTA
iPCRtag16	CAAGCAGAAGACGGCATAAAGACGGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AAGACGGA
iPCRtag17	CAAGCAGAAGACGGCATAAAGGTACAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AAGGTACA
iPCRtag18	CAAGCAGAAGACGGCATAACACAGAAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ACACAGAA

(continued)

Table 6
(continued)

Name	Primer sequence	8 base barcode within primer
iPCRtag19	CAAGCAGAAGACGGCATAACGAGATACAGCAGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ACAGCAGA
iPCRtag20	CAAGCAGAAGACGGCATAACGAGATACCTCCAAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ACCTCCAA
iPCRtag21	CAAGCAGAAGACGGCATAACGAGATACGCTCGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ACGCTCGA
iPCRtag22	CAAGCAGAAGACGGCATAACGAGATACGTATCAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ACGTATCA
iPCRtag23	CAAGCAGAAGACGGCATAACGAGATACTATGCAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ACTATGCA
iPCRtag24	CAAGCAGAAGACGGCATAACGAGATAGAGTCAAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AGAGTCAA
iPCRtag25	CAAGCAGAAGACGGCATAACGAGATAGATCGCAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AGATCGCA
iPCRtag26	CAAGCAGAAGACGGCATAACGAGATAGCAGGAAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AGCAGGAA
iPCRtag27	CAAGCAGAAGACGGCATAACGAGATAGTCACTAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AGTCACTA
iPCRtag28	CAAGCAGAAGACGGCATAACGAGATATCCTGTAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ATCCTGTA
iPCRtag29	CAAGCAGAAGACGGCATAACGAGATATTGAGGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ATTGAGGA
iPCRtag30	CAAGCAGAAGACGGCATAACGAGATCAACCACAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CAACCACA
iPCRtag31	CAAGCAGAAGACGGCATAACGAGATCAAGACTAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CAAGACTA
iPCRtag32	CAAGCAGAAGACGGCATAACGAGATCAATGGAAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CAATGGAA
iPCRtag33	CAAGCAGAAGACGGCATAACGAGATCACTTCGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CACTTCGA
iPCRtag34	CAAGCAGAAGACGGCATAACGAGATCAGCGTTAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CAGCGTTA
iPCRtag35	CAAGCAGAAGACGGCATAACGAGATCATACCAAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CATACCAA
iPCRtag36	CAAGCAGAAGACGGCATAACGAGATCCAGTTCAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CCAGTTCA

(continued)

Table 6
(continued)

Name	Primer sequence	8 base barcode within primer
iPCRtag37	CAAGCAGAAGACGGCATAACGAGATCCGAAAGTAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CCGAAAGTA
iPCRtag38	CAAGCAGAAGACGGCATAACGAGATCCGTGAGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CCGTGAGA
iPCRtag39	CAAGCAGAAGACGGCATAACGAGATCCTCCTGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CCTCCTGA
iPCRtag40	CAAGCAGAAGACGGCATAACGAGATCGAACTTAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CGAACTTA
iPCRtag41	CAAGCAGAAGACGGCATAACGAGATCGACTGGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CGACTGGA
iPCRtag42	CAAGCAGAAGACGGCATAACGAGATCGCATAACGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CGCATAACA
iPCRtag43	CAAGCAGAAGACGGCATAACGAGATCTCAATGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CTCAATGA
iPCRtag44	CAAGCAGAAGACGGCATAACGAGATCTGAGCCAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CTGAGCCA
iPCRtag45	CAAGCAGAAGACGGCATAACGAGATCTGGCATAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CTGGCATA
iPCRtag46	CAAGCAGAAGACGGCATAACGAGATGAATCTGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GAATCTGA
iPCRtag47	CAAGCAGAAGACGGCATAACGAGATGACTAGTAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GACTAGTA
iPCRtag48	CAAGCAGAAGACGGCATAACGAGATGAGCTGAAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GAGCTGAA
iPCRtag49	CAAGCAGAAGACGGCATAACGAGATGATAGACAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GATAGACA
iPCRtag50	CAAGCAGAAGACGGCATAACGAGATGCCACATAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GCCACATA
iPCRtag51	CAAGCAGAAGACGGCATAACGAGATGCGAGTAAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GCGAGTAA
iPCRtag52	CAAGCAGAAGACGGCATAACGAGATGCTAACGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GCTAACGA
iPCRtag53	CAAGCAGAAGACGGCATAACGAGATGCTCGGTAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GCTCGGTA
iPCRtag54	CAAGCAGAAGACGGCATAACGAGATGGAGAACAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GGAGAACA

(continued)

Table 6
(continued)

Name	Primer sequence	8 base barcode within primer
iPCRtag55	CAAGCAGAAGACGGCATAACGAGATGGTGCGAAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GGTGCGAA
iPCRtag56	CAAGCAGAAGACGGCATAACGAGATGTACGCAAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GTACGCAA
iPCRtag57	CAAGCAGAAGACGGCATAACGAGATGTCGTAGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GTCGTAGA
iPCRtag58	CAAGCAGAAGACGGCATAACGAGATGTCTGTCAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GTCTGTCA
iPCRtag59	CAAGCAGAAGACGGCATAACGAGATGTGTTCTAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GTGTTCTA
iPCRtag60	CAAGCAGAAGACGGCATAACGAGATTAGGATGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	TAGGATGA
iPCRtag61	CAAGCAGAAGACGGCATAACGAGATTATCAGCAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	TATCAGCA
iPCRtag62	CAAGCAGAAGACGGCATAACGAGATTCCGTCTAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	TCCGTCTA
iPCRtag63	CAAGCAGAAGACGGCATAACGAGATTCTTCACAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	TCTTCACA
iPCRtag64	CAAGCAGAAGACGGCATAACGAGATTGAAGAGAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	TGAAGAGA
iPCRtag65	CAAGCAGAAGACGGCATAACGAGATTGGAACAAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	TGGAACAA
iPCRtag66	CAAGCAGAAGACGGCATAACGAGATTGGCTTCAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	TGGCTTCA
iPCRtag67	CAAGCAGAAGACGGCATAACGAGATTGGTGGTAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	TGGTGGTA
iPCRtag68	CAAGCAGAAGACGGCATAACGAGATTTACGCAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	TTCACGCA
iPCRtag69	CAAGCAGAAGACGGCATAACGAGATAACTCACCAGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AACTCACC
iPCRtag70	CAAGCAGAAGACGGCATAACGAGATAAGAGATCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AAGAGATC
iPCRtag71	CAAGCAGAAGACGGCATAACGAGATAAGGACACGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AAGGACAC
iPCRtag72	CAAGCAGAAGACGGCATAACGAGATAATCCGTCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AATCCGTC

(continued)

Table 6
(continued)

Name	Primer sequence	8 base barcode within primer
iPCRtag73	CAAGCAGAAGACGGCATAACGAGATAATGTTGCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AATGTTGC
iPCRtag74	CAAGCAGAAGACGGCATAACGAGATACACGACCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ACACGACC
iPCRtag75	CAAGCAGAAGACGGCATAACGAGATACAGATTCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ACAGATTC
iPCRtag76	CAAGCAGAAGACGGCATAACGAGATAGATGTACGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AGATGTAC
iPCRtag77	CAAGCAGAAGACGGCATAACGAGATAGCACCTCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AGCACCTC
iPCRtag78	CAAGCAGAAGACGGCATAACGAGATAGCCATGCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AGCCATGC
iPCRtag79	CAAGCAGAAGACGGCATAACGAGATAGGCTAACGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	AGGCTAAC
iPCRtag80	CAAGCAGAAGACGGCATAACGAGATATAGCGACGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ATAGCGAC
iPCRtag81	CAAGCAGAAGACGGCATAACGAGATATCATTCCGAGATCGG TCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ATCATTCC
iPCRtag82	CAAGCAGAAGACGGCATAACGAGATATTGGCTCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	ATTGGCTC
iPCRtag83	CAAGCAGAAGACGGCATAACGAGATCAAGGAGCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CAAGGAGC
iPCRtag84	CAAGCAGAAGACGGCATAACGAGATCACCTTACGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CACCTTAC
iPCRtag85	CAAGCAGAAGACGGCATAACGAGATCCATCCTCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CCATCCTC
iPCRtag86	CAAGCAGAAGACGGCATAACGAGATCCGACAACGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CCGACAAC
iPCRtag87	CAAGCAGAAGACGGCATAACGAGATCCTAATCCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CCTAATCC
iPCRtag88	CAAGCAGAAGACGGCATAACGAGATCCTCTATCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CCTCTATC
iPCRtag89	CAAGCAGAAGACGGCATAACGAGATCGACACACGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CGACACAC
iPCRtag90	CAAGCAGAAGACGGCATAACGAGATCGGATTGCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CGGATTGC

(continued)

Table 6
(continued)

Name	Primer sequence	8 base barcode within primer
iPCRtag91	CAAGCAGAAGACGGCATAACGAGATCTAAGGTCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	CTAAGGTC
iPCRtag92	CAAGCAGAAGACGGCATAACGAGATGAACAGGCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GAACAGGC
iPCRtag93	CAAGCAGAAGACGGCATAACGAGATGACAGTGCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GACAGTGC
iPCRtag94	CAAGCAGAAGACGGCATAACGAGATGAGTTAGCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GAGTTAGC
iPCRtag95	CAAGCAGAAGACGGCATAACGAGATGATGAATCGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GATGAATC
iPCRtag96	CAAGCAGAAGACGGCATAACGAGATGCCAAGACGAGA TCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC	GCCAAGAC

[7, 8]. Consistent with our results, these studies show that for point mutations that are very close (± 5 nt) to the CRISPR cut site, it does not appear to matter which strand of ssODN HDR template is used. However, for mutations that are further away (5–20 nt), there is a significant difference between the two strands of ssODN. A simple rule is that if the mutation to be introduced is downstream of the CRISPR cut site, the antisense template ssODN (i.e., the one that binds to the top strand) should be used. Similarly, if the mutation is upstream of the CRISPR cut site, the sense ssODN should be used. This is irrespective of the strand to which the guide binds.

6. Cells adhere to SyntheMAX very well, so cell scrapers need to be used to manually detach colonies which have not dislodged during incubation with gentle cell dissociation reagent. Do not use excessive trituration to dislodge attached colonies since this will result in cell clusters which are too small for routine passaging, leaving cells vulnerable to poor growth/colony compaction during onward culture. Continual passaging of pluripotent cells as single cells has also been shown to increase accumulation of karyotypic abnormalities [27, 28].
7. When diluting Cas9 protein, DTT should be added fresh from a frozen, 0.1 M stock solution immediately before use.
8. Cells are very delicate following nucleofection; therefore, handling of cells should be minimized. Transfer from cuvette to culture dish using a wide bore Pasteur pipette and disperse over the culture surface using a gentle swirling/back-and-forth movement. Work quickly once mixing the RNP complex,

template, buffer, and cells together, minimizing the length of exposure of the cells to these components. The accutase treatment duration time is somewhat cell line specific and can be reduced to increase viability, but not so much as to compromise obtaining single cells.

9. An alternative to the T7E1 assay is to send PCR products from the 500 bp amplification for sequencing and analyze sequencing traces with TIDE [16] (<https://tide-calculator.nki.nl/>) or ICE (<https://ice.synthego.com/#/>). Note that a WT sequencing read is also required to calibrate such software.
10. Lysis is usually performed for 1 h at 60 °C; however, incubation times can be increased to overnight if necessary.
11. If during the preparation of 500 bp PCR product, PCR is unsuccessful, it is likely to be due to inhibition of the PCR by the lysate, low DNA concentration, or ineffective primers. Primers can be optimized by trying a range of annealing temperatures or adding DMSO to the reaction mix (5% v/v final concentration). Inhibition by components in the lysate can be improved by diluting it, or lysing cells in a larger volume of lysis buffer. If the DNA concentration is too low, try lysing in a smaller volume of lysis buffer or not diluting the DNA before adding to the PCR. Genomic DNA can also alternatively be obtained using a column based system, such as Qiagen DNeasy Blood & Tissue Kit (69504).
12. When counting cells, it should be clear under the microscope if cells are single or in clumps. If there are clumps, try adding an extra mix of the cell suspension or increasing the Accutase treatment time. Cells can also be filtered with a 40 µm cell strainer to help to remove clumps.
13. If there is poor cell survival after subcloning, add an additional centrifugation step to remove residual trace amounts of accutase before counting and diluting the cells. Reducing the accutase treatment duration can help, but not so much as to compromise obtaining single cells.
14. It is important to seed single cells during subcloning to ensure colonies arise from single cells. hiPSCs do not grow well from single cells, and will often survive better as small clumps of cells. It can be difficult to identify colonies generated from single cells but the size and shape of the colonies relative to others on the plate can be a good indicator of success. Try to avoid picking colonies which are growing close to each other or touching (Fig. 10).
15. Incubation in cell dissociation reagent serves simply to aid breaking apart of the colony during trituration and splitting into matrigel/syntheMAX plates; it is not intended to detach and break apart colonies from the plate before picking.

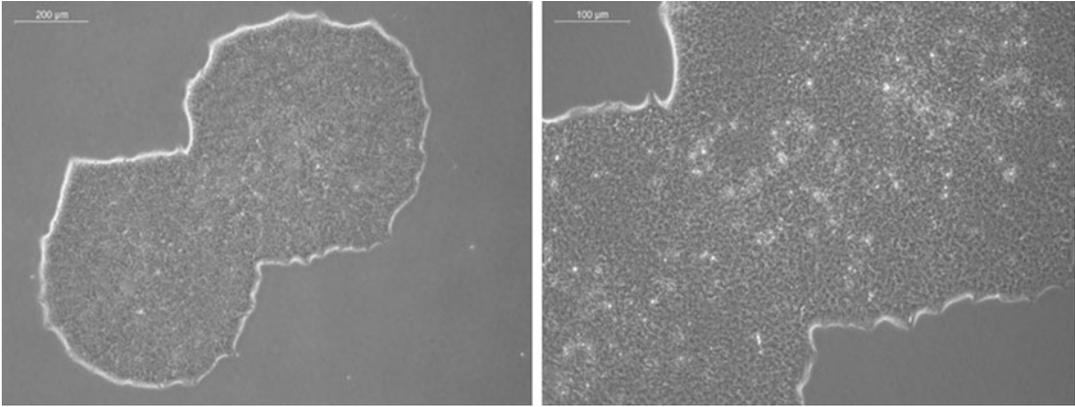


Fig. 10 Inappropriate morphology of colonies for picking. Two colonies ~10 days after subcloning viewed under a 5× and 10× objective. Colonies that do not have a regular round shape are likely to have been derived from two or more cells. Identical genetic sequences can therefore not be ensured and is why such colonies should be avoided when picking

16. If colonies are breaking into very small clumps (~5–10 cells) during colony picking, reduce the incubation time with gentle cell dissociation reagent. If colonies are dissociated too much at this stage there is a high risk of cells floating off into the picking medium which can result in cross contamination of colonies.
17. Matrigel plates serve as the freezing plates in which cell stocks will be bulked up and frozen (Subheading 3.7) for storage until genotyping confirms successful editing. The SyntheMAX plates will be grown up and lysed for genotyping analysis to identify suitable edited clones (Subheading 3.8).
18. If cells are picked out of the tissue culture hood or with the front of the hood open, Penicillin and Streptomycin solution (Thermo 15140122) can be added for 24 h to help prevent contamination.
19. It is impossible to have all 96 clones ready to freeze at the same time, and some wells will have to be frozen when under- or over-confluent.
20. Matrix freeze plates are routinely used to freeze clones in this format before genotyping analysis has been complete. They enable single clones to be removed from LN₂ storage for expansion.
21. Cells are routinely grown in TeSR-E8 medium and can be expanded ~1:5 into a 6-well plate when reaching 80% confluence (Subheading 3.2). Master cell banks are usually prepared from a single 6-well. During preparation of a master cell bank, a small sample of cells (~1/10 of a 6-well) is lysed in yolk sac lysis buffer for secondary genotyping (Subheading 3.10).

References

- Bassett AR (2017) Editing the genome of hiPSC with CRISPR/Cas9: disease models. *Mamm Genome* 28:348–364
- 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* 337:816–821
- Bibikova M, Beumer K, Trautman JK, Carroll D (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science* 300:764
- Shrivastav M, De Haro LP, Nickoloff JA (2008) Regulation of DNA double-strand break repair pathway choice. *Cell Res* 18:134–147
- Kim S, Kim D, Cho SW, Kim J, Kim JS (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res* 24:1012–1019
- Liang X, Potter J, Kumar S, Zou Y, Quintanilla R, Sridharan M, Carte J, Chen W, Roark N, Ranganathan S, Ravinder N, Chesnut JD (2015) Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J Biotechnol* 208:44–53
- Kan Y, Ruis B, Takasugi T, Hendrickson EA (2017) Mechanisms of precise genome editing using oligonucleotide donors. *Genome Res* 27:1099–1111
- Davis L, Maizels N (2016) Two distinct pathways support gene correction by single-stranded donors at DNA nicks. *Cell Rep* 17:1872–1881
- Richardson CD, Kazane KR, Feng SJ, Zelin E, Bray NL, Schäfer AJ, Floor SN, Corn JE (2018) CRISPR–Cas9 genome editing in human cells occurs via the Fanconi anemia pathway. *Nat Genet* 50(8):1132–1139
- Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, Cradick TJ, Marraffini LA, Bao G, Zhang F (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* 31:827–832
- Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales AP, Li Z, Peterson RT, Yeh JR, Aryee MJ, Joung JK (2015) Engineered CRISPR–Cas9 nucleases with altered PAM specificities. *Nature* 523:481–485
- Hu JH, Miller SM, Geurts MH, Tang W, Chen L, Sun N, Zeina CM, Gao X, Rees HA, Lin Z, Liu DR (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556:57–63
- 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* 31:822–826
- Hodgkins A, Farne A, Perera S, Grego T, Parry-Smith DJ, Skarnes WC, Iyer V (2015) WGE: a CRISPR database for genome engineering. *Bioinformatics* 31:3078–3080
- Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science* 351:84–88
- Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z, Wilen C, Orchard R, Virgin HW, Listgarten J, Root DE (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR–Cas9. *Nat Biotechnol* 34:184–191
- Kilpinen H, Goncalves A, Leha A, Afzal V, Alasoo K, Ashford S, Bala S, Bensaddek D, Casale FP, Culley OJ, Danecek P, Faulconbridge A, Harrison PW, Kathuria A, McCarthy D, McCarthy SA, Meleckyte R, Memari Y, Moens N, Soares F, Mann A, Streeter I, Agu CA, Alderton A, Nelson R, Harper S, Patel M, White A, Patel SR, Clarke L, Halai R, Kirton CM, Kolb-Kokocinski A, Beales P, Birney E, Danovi D, Lamond AI, Ouwehand WH, Vallier L, Watt FM, Durbin R, Stegle O, Gaffney DJ (2017) Common genetic variation drives molecular heterogeneity in human iPSCs. *Nature* 546:370–375
- Brinkman EK, Chen T, Amendola M, van Steensel B (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42:e168
- Bronner IF, Quail MA, Turner DJ, Swerdlow H (2014) Improved protocols for illumina sequencing. *Curr Protoc Hum Genet* 80:18.2.1–18.242
- Kozarewa I, Turner DJ (2011) 96-plex molecular barcoding for the illumina genome analyzer. *Methods Mol Biol* 733:279–298
- Pinello L, Canver MC, Hoban MD, Orkin SH, Kohn DB, Bauer DE, Yuan GC (2016) Analyzing CRISPR genome-editing experiments with CRISPResso. *Nat Biotechnol* 34:695–697
- Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, Kuhn R (2015) Increasing the efficiency of homology-directed repair

- for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotechnol* 33:543–548
23. Carlson-Stevermer J, Abdeen AA, Kohlenberg L, Goedland M, Molugu K, Lou M, Saha K (2017) Assembly of CRISPR ribonucleoproteins with biotinylated oligonucleotides via an RNA aptamer for precise gene editing. *Nat Commun* 8:1711
 24. Agudelo D, Durringer A, Bozoyan L, Huard CC, Carter S, Loehr J, Synodinou D, Drouin M, Salsman J, Delleire G, Laganieri J, Doyon Y (2017) Marker-free coselection for CRISPR-driven genome editing in human cells. *Nat Methods* 14:615–620
 25. Zhou Y, Liu Y, Hussmann D, Brogger P, Al-Saaidi RA, Tan S, Lin L, Petersen TS, Zhou GQ, Bross P, Aagaard L, Klein T, Ronn SG, Pedersen HD, Bolund L, Nielsen AL, Sorensen CB, Luo Y (2016) Enhanced genome editing in mammalian cells with a modified dual-fluorescent surrogate system. *Cell Mol Life Sci* 73:2543–2563
 26. Hirano H, Gootenberg JS, Horii T, Abudayyeh OO, Kimura M, Hsu PD, Nakane T, Ishitani R, Hatada I, Zhang F, Nishimasu H, Nureki O (2016) Structure and engineering of francisella novicida Cas9. *Cell* 164:950–961
 27. Ellerstrom C, Strehl R, Noaksson K, Hyllner J, Semb H (2007) Facilitated expansion of human embryonic stem cells by single-cell enzymatic dissociation. *Stem Cells* 25:1690–1696
 28. Thomson A, Wojtacha D, Hewitt Z, Priddle H, Sottile V, Di Domenico A, Fletcher J, Waterfall M, Corrales NL, Ansell R, McWhir J (2008) Human embryonic stem cells passaged using enzymatic methods retain a normal karyotype and express CD30. *Cloning Stem Cells* 10:89–106



Conditional Gene Knockout in Human Cells with Inducible CRISPR/Cas9

Kirsten E. Snijders, James D. Cooper, Ludovic Vallier,
and Alessandro Bertero

Abstract

The advent of the easily programmable and efficient CRISPR/Cas9 nuclease system has revolutionized genetic engineering. While conventional gene knockout experiments using CRISPR/Cas9 are very valuable, these are not well suited to study stage-specific gene function in dynamic situations such as development or disease. Here we describe a CRISPR/Cas9-based OPTimized inducible gene KnockOut method (OPTiKO) for conditional loss-of-function studies in human cells. This approach relies on an improved tetracycline-inducible system for conditional expression of single guide RNAs (sgRNAs) that drive Cas9 activity. In order to ensure homogeneous and stable expression, the necessary transgenes are expressed following rapid and efficient single-step genetic engineering of the *AAVSI* genomic safe harbor. When implemented in human pluripotent stem cells (hPSCs), the approach can be then efficiently applied to virtually any hPSC-derived human cell type at various stages of development or disease.

Key words CRISPR/Cas9, sgRNA, Inducible knockout, Tetracycline, *AAVSI*, Human pluripotent stem cells

1 Introduction

The genomics revolution brought the key challenge of functionally annotating an ever-growing list of both coding and noncoding genetic variants linked to human development, physiology, and disease [1, 2]. As a result, there is a high demand for simple, scalable, rapid, and robust platforms for functional genetic analyses in multiple human cell types, including populations that are transitory during development, rare, or inaccessible. Loss-of-function experiments in human pluripotent stem cells (hPSCs: human embryonic stem cells, hESCs, or human induced pluripotent stem cells, hiPSCs) provide a unique opportunity to address this major challenge, as hPSCs can provide an unlimited source of virtually any somatic cell type of clinical interest [3–5]. However, the study of gene function at specific stages of differentiation requires efficient

methods to conditionally manipulate gene expression in both hPSCs and hPSC-derived cells.

In recent years, our ability to interrogate gene function has been revolutionized by the rise of gene editing methods based on CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) and associated Cas genes, particularly CRISPR/Cas9 [6–8]. The most popular version of this approach is based on a modified prokaryotic type II CRISPR/Cas9 system in which the DNA endonuclease Cas9 protein is selectively targeted to a genomic region by a single guide RNA (sgRNA). In such sgRNA, an approximately 20 base pairs (bp) sequence-specific “target” region (CRISPR RNA, or crRNA) is joined to an invariable “scaffold” sequence (trans-activating crRNA, or tracrRNA). The target region is designed to be perfectly complementary to a specific genomic region that is immediately followed by the essential protospacer adjacent motif (PAM) 5'-NGG-3' site. Upon binding, Cas9 cleaves the DNA to generate double-strand breaks that are predominantly repaired by error-prone non-homologous end joining (NHEJ), resulting in random small insertions or deletions (indels) that can generate loss-of-function alleles [9, 10].

CRISPR/Cas9-based methods for inducible loss-of-function in hPSCs have been reported [11–13]. However, these approaches are either significantly complex and time-consuming, as they involve multiple genome editing steps that must be tailored to each individual gene [12], or are not widely applicable in hPSC-derived cells, as they rely on inducible promoters that are not stably and homogeneously expressed following hPSC differentiation [11, 13–16]. We recently developed an alternative approach that bypasses these limitations by allowing generation of inducible knockout hPSCs with a single gene editing step introducing transgenes that are stably expressed in hPSC-derived cell types [16]. This method, which we named OPTiKO (for OPTimized inducible gene KnockOut), allows streamlined, highly efficient, homogeneous, and rapid induction of loss-of-function mutations in both hPSCs and hPSC-derived lineages [16]. Recently, this has also been used as an *in vitro* platform for the screening of efficient sgRNAs to be used for CRISPR/Cas9 gene editing in human embryos [17].

OPTiKO combines the CRISPR/Cas9 technology with an optimized tetracycline-derepressible (TET-OFF) system: hPSCs are genetically engineered to carry a constitutively expressed Cas9 protein that is only activated following tetracycline-induced expression of the sgRNA (Fig. 1). To its core, this relies on an H1 RNA polymerase III promoter that was modified to include a tet operator (tetO2) sequence after the TATA box [18] (H1-TO; Fig. 2). The tetO2 is strongly bound by a tetracycline-sensitive repressor protein that was codon-optimized to maximize its expression in human cells (OPTtetR) [16]. While this prevents sgRNA expression in normal culture conditions, as a result of tetracycline

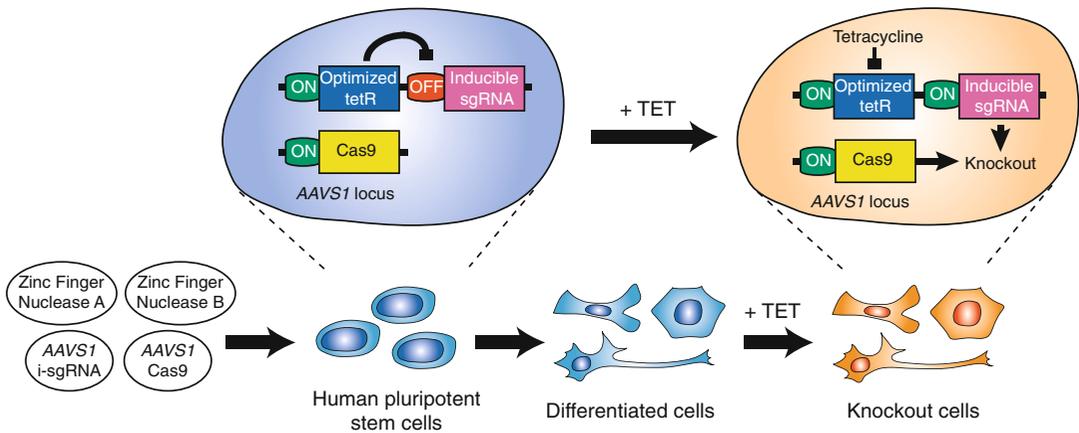


Fig. 1 The optimized inducible knockout (OPTiKO) system. Schematic of the generation of OPTiKO human pluripotent stem cells (hPSCs) and of the resulting genetic circuit which allows tetracycline (TET)-dependent induction of gene knockout both in hPSCs and hPSC-derived cells

treatment the tetracycline-bound OPTtetR undergoes a conformational change that relieves its binding to the HI-TO promoter, which is therefore derepressed. To ensure strong, homogeneous, and stable expression even after hPSC differentiation, all of the necessary transgenes are integrated into the well-characterized *AAVS1* genomic safe harbor [19–22]. This is achieved by means of single-step zinc finger nuclease (ZFN)-facilitated gene editing [23], which was optimized to ensure on-target recombination efficiency greater than 90% [16, 20]. OPTiKO relies on targeting each allele of the *AAVS1* locus with one of two targeting plasmids: the first (pAAV-Neo_CAG-Cas9) carries the constitutively expressed *Streptococcus Pyogenes* Cas9 transgene, while the second (pAAV-Puro_siKO) carries the inducible sgRNA cassette and the constitutive OPTtetR (Figs. 2a and 3a). As these vectors also carry two distinct gene trap-based drug resistance genes, respectively, for neomycin and puromycin, hPSCs that express both transgenes are selected by means of dual drug selection. When the resulting OPTiKO hPSCs carry an inducible sgRNA designed to generate NHEJ-mediated loss-of-function mutations, gene knockout can be induced at any stage of hPSC differentiation by addition of tetracycline (Fig. 1).

The following protocol describes all of the procedures required for the generation and validation of OPTiKO hPSCs. First, the sgRNA specific to the gene of interest is designed and cloned as a double-strand oligonucleotide into pAAV-Puro_siKO (Subheading 3.1; Fig. 2). Then, hPSCs are nucleofected with the *AAVS1* ZFN plasmids and targeting vectors, and dual-gene targeted lines are selected using puromycin and neomycin (Subheading 3.2; Fig. 3). Clonal hPSC lines are isolated and characterized by genomic PCR to identify correctly targeted clones (Subheading 3.3; Fig. 4).

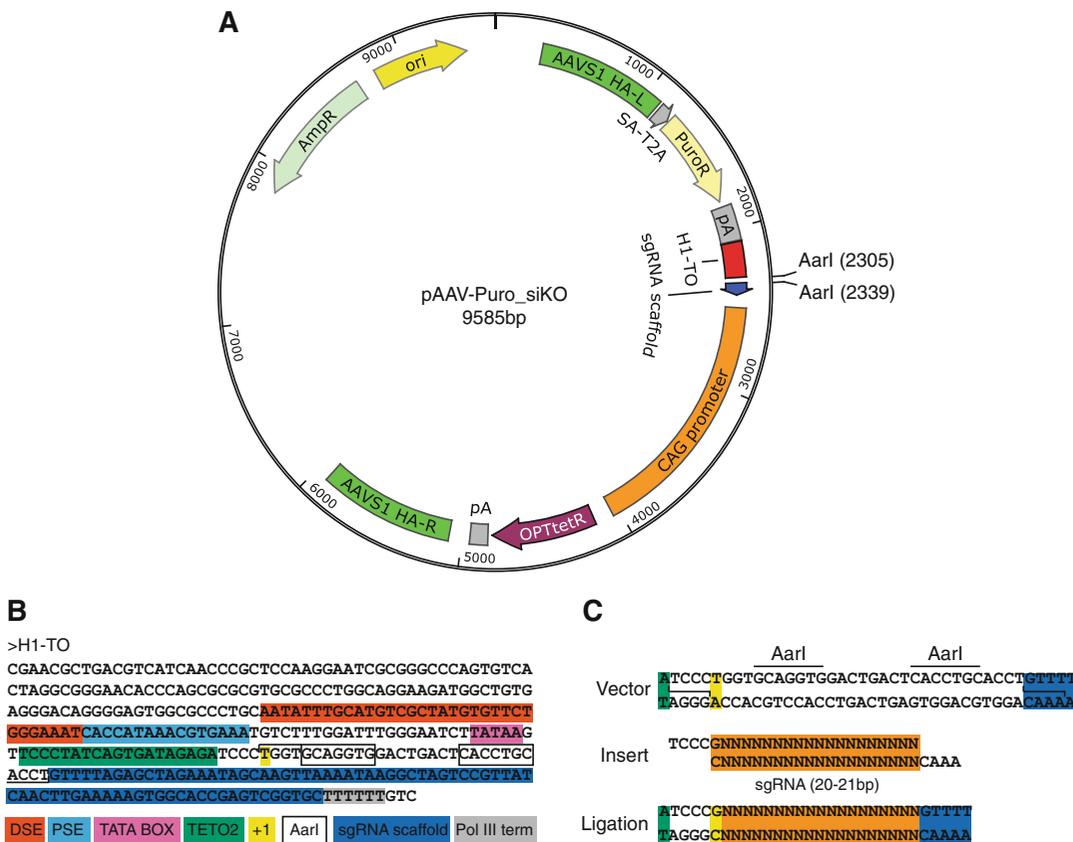


Fig. 2 Cloning of sgRNAs into the OPTiKO plasmid. (a) Map of the *AAVS1* targeting vector carrying the inducible sgRNA: the transgene is integrated following homologous directed repair (HDR), and employs a gene-trap approach to drive the drug resistance gene through the endogenous promoter of the *AAVS1* locus (Fig. 4). *HA-L/R* left/right homology arms, *SA* splice acceptor, *T2A* self-splicing T2A sequence, *PuroR* puromycin resistance cDNA, *pA* polyadenylation sequence, *H1-TO* tetracycline-inducible H1 RNA Polymerase III promoter containing a tetO2 sequence, *CAG* CMV early enhancer, chicken β -actin, and rabbit β -globin hybrid constitutive promoter, *OPTtetR* codon-optimized tetracycline-responsive repressor protein cDNA, *AmpR* ampicillin resistance, *ori* high-copy origin of replication. (b) Nucleotide sequence of the tetracycline-inducible H1-TO RNA Polymerase III promoter. Key features are color coded. The restriction enzyme sites used for sgRNA cloning are shown in boxes (panel c). *DSE* distal sequence element, *PSE* proximal sequence element, *TETO2* tet operator; +1 start position of RNA transcription, *Pol III term* poly(T) sequence inducing transcription termination. (c) Schematics of the sgRNA cloning procedure described in Subheading 3.1

Finally, the resulting OPTiKO hPSCs are validated to induce efficient gene knockout by confirming loss of the protein product upon tetracycline treatment (Subheading 3.4; Fig. 5). The whole procedure can be completed in 4–9 weeks depending on the experimental design and expertise of the investigator, and allows generation of OPTiKO hPSCs with >90% efficiency. The protocol shares similarities with a related method to generate inducible knockdown hPSCs, which can serve as additional reference [16, 24].

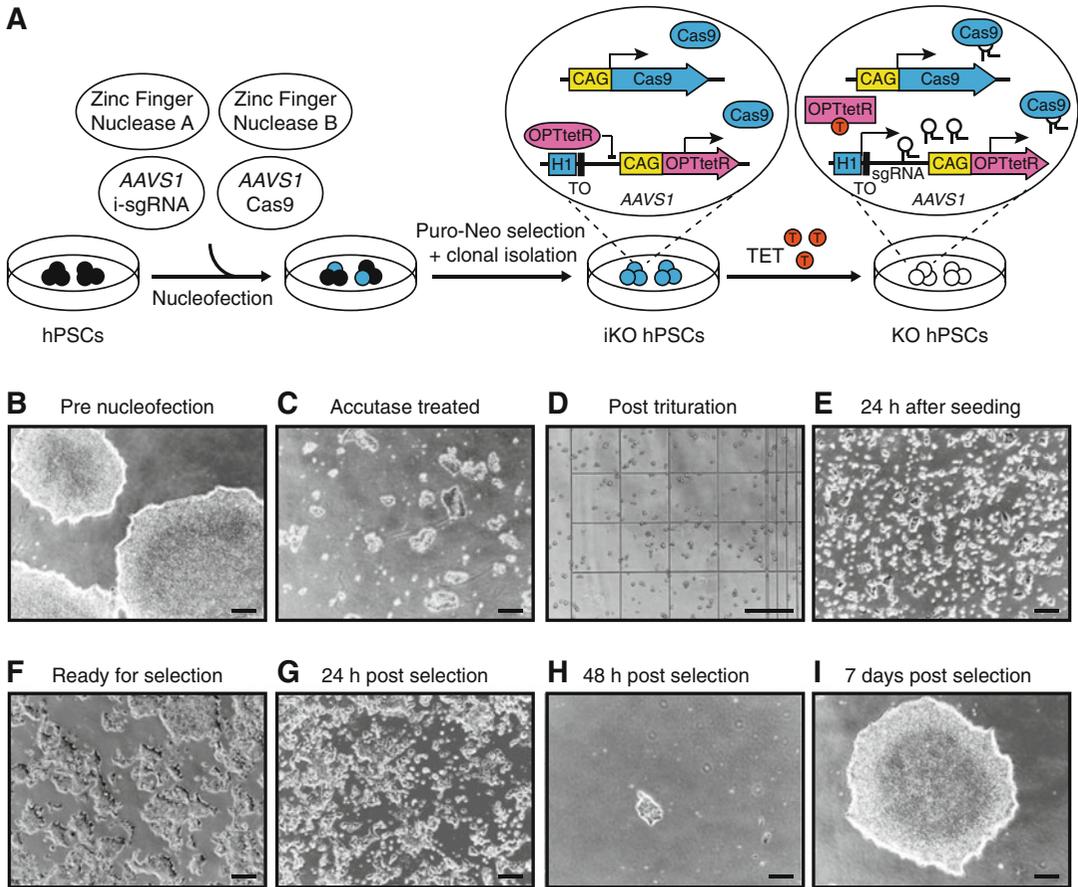


Fig. 3 Generation of OPTiKO hPSCs by gene targeting. **(a)** Schematic of the gene targeting procedure described in Subheading 3.2. The resulting OPTiKO transgenic alleles and their functionality in the presence or absence of tetracycline (TET) are shown. *ZFN* zinc finger nuclease, *i-sgRNA* inducible sgRNA, *iKO* inducible knockout, *KO* knockout. **(b–i)** Representative phase-contrast images of hPSCs at the indicated stages of the gene targeting procedure. Scale bars: 250 μ M

2 Materials

Prepare all solutions using analytical grade chemicals, and DNase- and RNase-free ultrapure deionized water (dH₂O). Store all reagents at room temperature, unless indicated otherwise below or in the supplier's instructions. Hazardous reagents must be handled with care while wearing appropriate personal protective equipment and in accordance with local safety regulations. Waste materials are to be disposed according to the relevant regulations. Standard equipment used in molecular biology and cell culture research is required for this protocol (e.g. micropipettes, a tabletop microcentrifuge, and a serological pipette controller). Certain materials are required at multiple stages of the protocol.

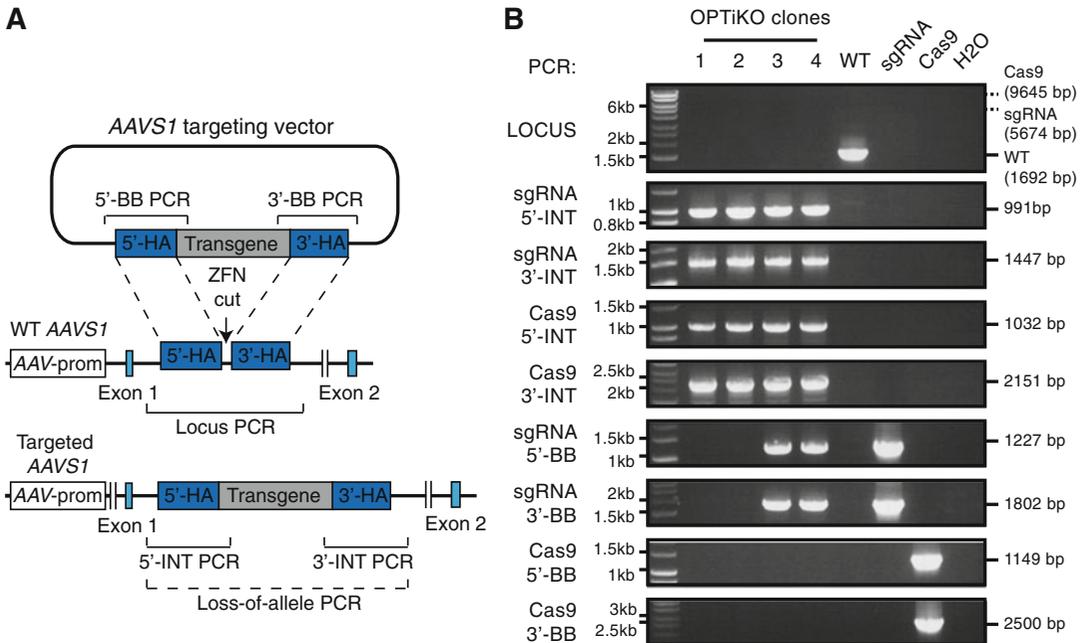


Fig. 4 Genotyping of OPTiKO hPSCs. **(a)** Schematic of the *AAVS1* gene targeting event that generates the OPTiKO transgenic alleles through homologous recombination of the donor plasmids (features are not in scale). AAV prom: endogenous promoter of the *AAVS1* locus (*PPP1R12C* gene), which drives the gene-trap drug resistance genes. The genotyping strategies used to identify correctly targeted hPSCs are shown (Subheading 3.3). Locus PCR PCR product of wild-type *AAVS1* locus (indicating a non-targeted allele); Loss-of-allele: potential PCR amplification that fails onto the targeted allele due to large size and high GC-content; 5'/3'-INT PCR: PCR product of transgene 5'/3'-end integration region (indicative of expected transgene targeting); 5'/3'-BB PCR: PCR product of vector backbone 5'/3'-end (indicative of nonspecific off-target plasmid integration). **(b)** Representative example of genotyping results from 4 OPTiKO clonal sublines. For all clones the *AAVS1* locus is correctly targeted with both transgenes, while clones 3 and 4 also carry randomly integrated copies of the targeting plasmid (Table 2). All clones could be used for further experiments. The predicted size of each PCR amplicon is indicated (Table 1), and a molecular weight control is shown on the left. WT control PCR from wild-type hPSCs, sgRNA control PCR from pAAV-Puro_siKO plasmid, Cas9 control PCR from pAAV-Neo_CAG-Cas9 plasmid, H₂O no template control

2.1 Molecular Cloning

1. pAAV-Puro_siKO plasmid (Addgene #86696).
2. Oligonucleotides for sgRNA target (custom, *see* Subheading 3.1, step 2).
3. AarI restriction enzyme (2 U/ μ L; ThermoFisher Scientific).
4. DNA loading dye (6 \times): 60% glycerol, 10 mM Tris-HCl (pH 7.6), 60 mM EDTA, 0.03% bromophenol blue, 0.03% xylene cyanol.
5. TAE electrophoresis buffer (50 \times): 2 M Tris base, 1 M acetic acid, 50 mM EDTA disodium salt.
6. Agarose I, molecular biology grade.
7. Ethidium bromide: 10 mg/mL stock solution in dH₂O.

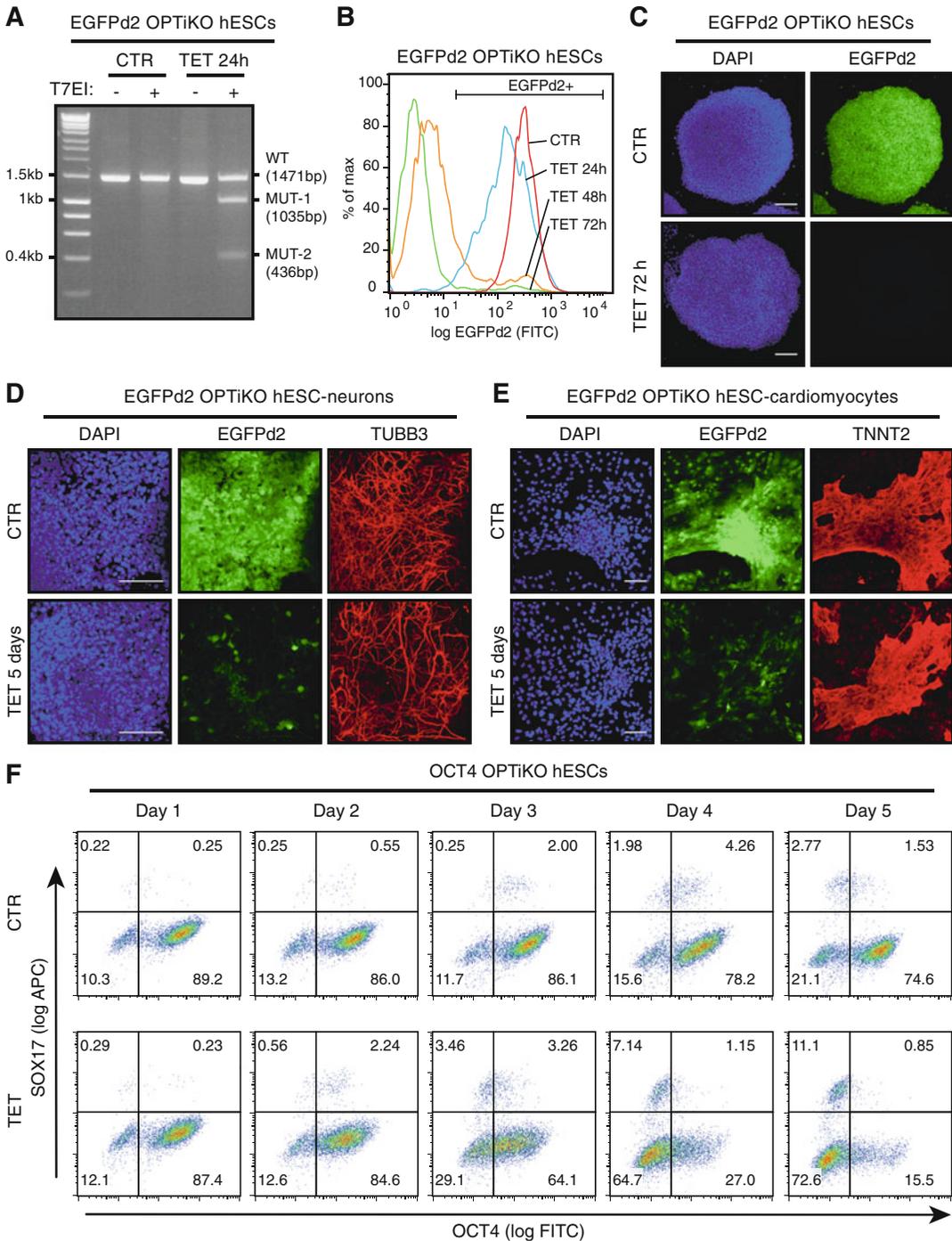


Fig. 5 Performance of OPTiKO in hPSCs and hPSC-derived cells. (a) T7 endonuclease 1 (T7E1) assay in EGFPd2 OPTiKO hESCs [16]. hESCs constitutively expressing a destabilized EGFP transgene (homozygous *ROSA26* CAG-EGFPd2 hESCs) were retargeted with the OPTiKO plasmids to drive an inducible sgRNA against EGFP. Cells were analyzed in the absence (CTR) or presence of tetracycline (TET) for 24 h. A 1.47 Kb fragment around the sgRNA cut site was amplified by PCR and analyzed by T7E1 assay. The presence of indels in

8. Agarose gel in TAE: dissolve the required amount of agarose powder into 100 mL of 1× TAE in a glass bottle. Place a cap on the bottle but leave loose. Incubate at room temperature for 15 min to pre-dissolve. Microwave for about 1–2 min or until all of the powder is fully dissolved, but do not let the solution boil. Allow the solution to cool at room temperature for 5–10 min (the temperature of the solution should not go below 65 °C to prevent premature gelling), and add ethidium bromide to a final concentration of 0.5 µg/mL. After mixing pour the solution into a gel casting tray equipped with the appropriate combs and let the gel set for 30 min before use.
9. DNA electrophoresis apparatus.
10. DNA molecular weight ladder.
11. UV transilluminator.
12. QIAEX II Gel Extraction Kit (QIAGEN).
13. Oligo annealing buffer (10×): 100 mM Tris-HCl (pH 8), 10 mM EDTA, 1 M NaCl.
14. Thermocycler with heated lid.
15. T4 DNA ligase (400 U/µL; New England Biolabs).
16. α-select Gold Efficiency *E. Coli* ($\geq 10^9$ cfu/µg; Bioline),
17. Heated water bath.
18. Humidified bacterial incubator.
19. S.O.C. medium: 2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, and 20 mM glucose. Adjust to pH 7.0 and sterilize by passing through a 0.22 µm filter.
20. Luria Bertani (LB) broth: 1% tryptone, 0.5% yeast extract, 171 mM NaCl. Adjust to pH 7.0 and sterilize by autoclaving.
21. Ampicillin: 100 mg/mL stock solution in dH₂O, store at –20 °C.
22. LB-agar ampicillin bacterial culture petri dishes: dissolve 20 g of agar powder in 1 L of LB broth and mix well. Autoclave on a liquid cycle and let to cool until it is warm enough to touch

Fig. 5 (continued) tetracycline-treated cells led to T7E1-mediated cleavage of the genomic fragment into the indicated products (MUT-1 and MUT-2). **(b)** Quantification of EGFPd2 by flow cytometry in EGFPd2 OPTiKO hESCs. >95% knockout could be achieved following 72 h of induction. **(c–e)** Analysis by immunocytochemistry of EGFPd2 OPTiKO hESCs **(c)**, hESC-derived neurons **(d)**, and hESC-derived cardiomyocytes **(e)**. EGFPd2 fluorescence is shown in green, while for **(d)** and **(e)** lineage-specific markers are shown in red. DAPI: nuclear staining. Scale bars: 100 µM. Widespread loss of EGFPd2 expression could be achieved in all lineages. **(f)** Analysis by flow cytometry of OCT4 OPTiKO hESCs [17]. Cells were analyzed for expression of OCT4 and of the endoderm marker SOX17. The result demonstrates that loss of OCT4 expression led to endoderm differentiation in a subpopulation of hESCs

(approximately 50 °C). Add ampicillin to a concentration of 100 μ M and swirl to mix (do not shake as this will create bubbles). Pour into petri dishes to completely cover the bottom surface. Allow the plates to set for 2 h at room temperature, and store sealed plates at 4 °C for up to 3 months.

23. Bacterial culture orbital shaker.
24. QIAprep Spin Miniprep Kit (QIAGEN).
25. siKO_fw primer: 5'-CGAACGCTGACGTCATCAACC-3'.
26. Glycerol for molecular biology (>99%).
27. QIAfilter Plasmid Midi Kit (QIAGEN).

2.2 Gene Targeting

1. Human pluripotent stem cells (*in house* or from various commercial suppliers).
2. 100 mm tissue culture petri dishes.
3. TeSR-E8 medium (STEMCELL Technologies).
4. 250 μ g/mL Vitronectin XF (STEMCELL Technologies).
5. UltraPure 0.5 M EDTA pH 8.0 (Life Technologies).
6. Humidified tissue culture incubator with CO₂ supply.
7. Biosafety level 2 laminar air flow tissue culture hood.
8. Y-27632 dihydrochloride: 10 mM solution in DMSO. Prepare single-use aliquots and store at -20 °C for up to 6 months.
9. DPBS no calcium and no magnesium.
10. StemPro Accutase Cell Dissociation Reagent (ThermoFisher Scientific).
11. P3 Primary Cell 4D-Nucleofector X Kit L (Lonza).
12. Trypan blue: 0.4% w/vol solution in dH₂O.
13. Hemocytometer.
14. pZFN_AAVS1-R-KKR plasmid (available from the corresponding authors).
15. pZFN_AAVS1-L-ELD plasmid (available from the corresponding authors).
16. pAAV-Neo_CAG-Cas9 plasmid (Addgene #86698).
17. pAAV-Puro_siKO-sgRNA plasmid (custom, *see* Subheading 3.1, step 11).
18. AAVS1-CAGGS-EGFP (Addgene #22212).
19. 4D-Nucleofector Core Unit and X Unit (Lonza).
20. Puromycin dihydrochloride: 10 mg/mL solution in dH₂O. Sterilize by passing through a 0.22 μ m filter, and store at -20 °C for up to 1 year.
21. Geneticin (G418 sulfate): 200 mg/mL solution in dH₂O. Sterilize by passing through a 0.22 μ m filter, and store at -20 °C for up to 1 year.

2.3 Genotyping and Validation

1. 24-well tissue culture petri dishes.
2. 10.000 U/mL penicillin-streptomycin.
3. Wizard SV Genomic DNA Purification Kit (Promega).
4. LongAmp Taq DNA Polymerase (2.5 U/ μ L; New England Biolabs).
5. dNTP mix: 10 mM dATP, 10 mM dCTP, 10 mM dGTP, 10 mM dTTP.
6. Dimethyl sulfoxide (DMSO), PCR grade.
7. Genotyping primers (*see* Table 1 for full list and sequences): 5 μ M stock solutions in dH₂O.
8. Tetracycline hydrochloride (cell-culture grade; Sigma-Aldrich): to prepare a stock 10 mg/mL solution dissolve 50 mg of tetracycline hydrochloride in 5 mL of dH₂O. The resulting solution should have a mild yellow-orange color. Filter-sterilize using a 0.22 μ m filter, and prepare single-use 5 or 10 μ L aliquots. Store at -80 °C for up to 6 months. Prepare and store this reagent protected from direct illumination.

3 Methods

Carry out all procedures at room temperature unless otherwise specified.

3.1 Generation of Inducible sgRNA Targeting Vector

1. Identify a 20 bp sgRNA target sequence against the gene of interest by taking advantage of available sgRNA design tools (*see* Note 1). Testing of the candidate sgRNAs efficiency is optional but is strongly recommended (*see* Note 2). If the sgRNA sequence does not start with a guanine (“G”) or an adenosine (“A”), add a “G” to the 5' end (*see* Note 3).
2. Design the “top” oligonucleotide by adding 5'-TCCC-3' to the 5' end of the sgRNA target sequence (without PAM), and the “bottom” oligonucleotide by adding 5'-AAAC-3' to the 5' end of the reverse complement of the sgRNA target sequence (without PAM). Annealing of these oligonucleotides will create a short double-strand DNA sequence with 5' overhangs suitable for subsequent directional cloning (Fig. 2c). Order both oligonucleotides as desalted purified products from a preferred vendor, and resuspend at a concentration of 200 μ M in dH₂O.
3. Prepare the backbone plasmid by digesting 5 μ g of pAAV-Puro_siKO in a 100 μ L reaction containing 5 μ L (10 U) of AarI, 2 μ L (0.5 μ M) of the provided oligonucleotide, and 10 μ L (1 \times) of the provided 10 \times AarI digestion buffer (remaining volume dH₂O). Incubate at 37 °C overnight (16 h; *see* Note 4). AarI is a type II restriction enzyme that cuts outside of its two recognition sites to create non-complementary 3'

Table 1
Primers for genotyping of OPTiKO clonal lines

PCR type	Primer name	Primer location	Primer sequence (5' → 3')	Amplicon wild-type ^a	Amplicon target ^b	Amplicon plasmid ^c	Temp. ann (°C) ^d	Ext. time ^d
Locus	Locus_fw	Genomic, 5' to 5' HA	CTGTTTCCCC TTCCCAGGCAGGTCC	1692 bp	No band	No band	65	1' 30"
	Locus_rev	Genomic, 3' to 3' HA	TGCAGGGGAACGGGGC TCAGTCTGA					
5' INT sgRNA	Locus_fw	Genomic, 5' to 5' HA	CTGTTTCCCC TTCCCAGGCAGGTCC	No band	991 bp	No band	65	1'
	Puro_rev	Puromycin resistance	TCGTCGGGG TGGCGAGGGCAACG					
5' INT Cas9	Locus_fw	Genomic, 5' to 5' HA	CTGTTTCCCC TTCCCAGGCAGGTCC	No band	1032 bp	No band	60	1'
	Neo_rev	Neomycin resistance	GTGCCCCAGTCATAGCCGAAT					
3'-INT sgRNA	OPTTetR_fw	OPTTetR cDNA	CCACCGAGAAGCAGTACGAG	No band	1447 bp	No band	60	1' 30"
	Locus_rev	Genomic, 3' to 3' HA	TGCAGGGGAACGGGGC TCAGTCTGA					
3' INT Cas9	Cas9_fw	Cas9 cDNA	AAACGGCCGGAAAGAGAATG	No band	2151 bp	No band	60	2'
	Locus_rev	Genomic, 3' to 3' HA	TGCAGGGGAACGGGGC TCAGTCTGA					
5'-BB sgRNA	Backbone_fw	Backbone, 5' to 5' HA	ATGCTTCCGGCTCGTATGTT	No band	No band	1227 bp	60	1' 30"
	Puro_rev2	Puromycin resistance	TGAGGAAGAGTTC TTGCAGTCTC					
5' BB Cas9	Backbone_fw	Backbone, 5' to 5' HA	ATGCTTCCGGCTCGTATGTT	No band	No band	1149 bp	60	1' 30"
	Neo_rev	Neomycin resistance	GTGCCCAAGTCATAGCCGAAT					

(continued)

Table 1
(continued)

PCR type	Primer name	Primer location	Primer sequence (5' → 3')	Amplicon wild-type ^a	Amplicon target ^b	Amplicon plasmid ^c	Temp. ann (°C) ^d	Ext. time ^d
3'-BB sgRNA	OPTTetR_fw Backbone_rev	OPTTetR cDNA Backbone, 3' to 3' HA	CCACCGAGAAGCAGTACGAG ATGCACCCACCGGGTAAAGTT	No band	No band	1802 bp	60	2'
3'/BB Cas9	Cas9_fw Backbone_rev	Cas9 cDNA Backbone, 3' to 3' HA	AAACGGCCGGAAAGAGAATG ATGCACCCACCGGGTAAAGTT	No band	No band	2500 bp	60	2' 30''

^aResult of PCR on wild-type *AAVSI* allele

^bResult of PCR on OPTIKO-targeted *AAVSI* allele

^cResult of PCR on targeting vector (positive control for off-target plasmid integration)

^dVariable parameter in PCR protocol (Subheading 3.3, step 5)

overhangs suitable for seamless directional cloning of the sgRNA target sequence in front of the sgRNA scaffold (Fig. 2a–c).

4. Add 20 μL ($1\times$) of the $6\times$ DNA gel loading dye, and perform standard DNA electrophoresis of the digestion product using a 1% agarose-TAE gel. Include a DNA ladder as molecular weight control. Visualize the DNA using a UV transilluminator, excise the 9551 bp band with a clean scalpel (*see Note 5*), and gel-extract the DNA using QIAEX II Gel Extraction Kit following the manufacturer's instructions. Quantify the linearized plasmid, and adjust the concentration to 50 ng/ μL . Keep on ice for same day use, or store in single-thaw aliquots at $-20\text{ }^{\circ}\text{C}$ for at least 6 months.
5. Prepare the insert by annealing the top and bottom oligonucleotides from Subheading 3.1, **step 2** in a 20 μL reaction containing 5 μL (50 μM) of each oligo and 2 μL ($1\times$) of $10\times$ oligo annealing buffer (remaining volume dH₂O). Incubate in a thermocycler for 5 min at 95 $^{\circ}\text{C}$, followed by slow descent to 4 $^{\circ}\text{C}$ by $-0.1\text{ }^{\circ}\text{C}/\text{s}$, then dilute 1:500 in $1\times$ oligo annealing buffer. Keep on ice for same day use (do not store).
6. Assemble a 10 μL ligation reaction with 4 μL of the diluted annealed oligo, 1 μL (50 ng) of the linearized plasmid from Subheading 3.1, **step 4**, 1 μL (5 U) of T4 DNA Ligase, and 1 μL ($1\times$) of the provided $10\times$ T4 DNA Ligase Buffer (remaining volume dH₂O). Incubate at room temperature for 2 h.
7. Transform 2 μL of the ligation product into 25 μL of α -select *E. Coli* according to manufacturer's instructions using a heated water bath for heat-shock. Recover the transformed bacteria in 250 μL of S.O.C. media for 30 min at 37 $^{\circ}\text{C}$, then plate all of the culture onto a LB-agar petri dish containing 100 $\mu\text{g}/\text{mL}$ of ampicillin. Incubate overnight (16 h) at 37 $^{\circ}\text{C}$ in a humidified incubator.
8. Pick individual bacterial colonies using a sterile tip and inoculate them into 4 mL of LB broth with 100 $\mu\text{g}/\text{mL}$ of ampicillin (*see Note 6*). Incubate in a bacterial culture orbital shaker at 37 $^{\circ}\text{C}$ overnight (16 h) while shaking at 225 rpm.
9. Isolate plasmids from bacterial cultures using the QIAprep Spin Miniprep Kit following manufacturer's instructions. Save 200 μL of each culture and store in the fridge.
10. Confirm the presence of the desired sgRNA and the lack of mutations by performing Sanger DNA sequencing using the siKO_fw primer. Upon successful result, freshly inoculate the bacterial clone by diluting the previously saved liquid culture 1:1000 in 50 mL of LB broth with 100 $\mu\text{g}/\text{mL}$ of ampicillin.

Incubate in a bacterial culture orbital shaker at 37 °C overnight (16 h) while shaking at 225 rpm.

11. Prepare a glycerol stock to be stored at –80 °C for long-term backup by mixing 200 µL of sterile autoclaved 50% glycerol in dH₂O with 200 µL of the bacterial culture. Isolate the plasmid from the remaining culture using the QIAfilter Midiprep Kit according to the manufacturer's instructions. Resuspend the resulting pAAV-Puro_siKO-sgRNA targeting vector in dH₂O at a concentration of 1 µg/mL and store at –20 °C.

3.2 Gene Targeting of Inducible CRISPR/Cas9 Platform in hPSCs

1. Grow hPSCs in mTeSR E8 medium and vitronectin-coated culture dishes according to the manufacturer's instructions in a 37 °C incubator with 5% CO₂ and 95% humidity. Passage cells every 3–5 days using 0.5 mM EDTA to facilitate mechanical dissociation, and seed them as small clumps of 5–20 cells with sub-cultivation ratio of 1:3 to 1:10 (*see Note 7*). hPSCs to be used for nucleofection should be within their exponential growth phase (50–70% confluency; Fig. 3b), and 2×10^6 cells are required for each nucleofection. Positive and negative nucleofection controls are strongly recommended for first-time users or when troubleshooting (*see Note 8*). Volumes are given for cells cultured in 100 mm dishes.
2. 16–24 h before nucleofection feed hPSCs with 10 mL of fresh TeSR-E8 media without antibiotics and supplemented with 10 µM Y-27632 (ROCK inhibitor; *see Note 9*).
3. Before beginning the nucleofection procedure, prepare two vitronectin-coated 100 mm dishes for each nucleofection, aspirate the coating solution, add 10 mL of TeSR-E8 medium supplemented with 10 µM Y-27632, and place in the 37 °C incubator to pre-warm and equilibrate. Thaw all the plasmids and adjust their concentration to 1 µg/mL. Add the supplement to the P3 nucleofection solution from the 4D-Nucleofector kit, and let acclimatize at room temperature.
4. Begin cell collection by aspirating hPSC culture medium, rinsing the cells with 10 mL DPBS, and adding 5 mL of Accutase. Incubate for 3–5 min at 37 °C in the incubator until colonies can be lifted by gently tapping on the side of the dish (Fig. 3c). Add 10 mL of TeSR-E8 medium, and mechanically triturate the colonies into clumps of 3–4 cells using a 5 mL serological pipette (Fig. 3d).
5. Perform a live cell count using a hemocytometer after diluting an aliquot of the cells 1:2 with trypan blue (*see Note 10*). Aliquot 2×10^6 live cells in a separate conical tube for each nucleofection, and pellet the cell suspension by spinning for 3 min at 115 g at room temperature. Remove the supernatant as completely as possible.

6. Prepare the nucleofection mix in a sterile tube by adding 2 μL (2 μg) of pAAV-Puro_siKO-sgRNA, 2 μL (2 μg) of pAAV-Neo_CAG-Cas9, 4 μL (4 μg) of pZFN_AAVS1-R-KKR, 4 μL (4 μg) of pZFN_AAVS1-L-ELD, and finally 100 μL of P3 solution. Using a 1000 μL pipette tip transfer all of the mix to the tube containing the cell pellet, and resuspend the cells very gently by pipetting 3–5 times (*see Note 11*).
7. Rapidly transfer the nucleofection mix to the provided nucleofection cuvette (*see Note 12*), and pulse the cells by operating the program “CA-137” (*see Note 13*). Let the cells recover at room temperature for 5 min.
8. Add 500 μL of TeSR-E8 medium supplemented with 10 μM Y-27632 to the bottom of the cuvette, and using the provided suction pipette very gently transfer half of the cell suspension to each of the two pre-warmed 100 mm dishes from Subheading 3.2, step 3. Distribute drop by drop over the plate surface, and gently shake the plate back and forth then left and right 3–5 times to promote even plating. Incubate overnight in the 37 °C incubator (*see Note 14*).
9. On the following day, visually confirm efficient cell attachment (Fig. 3e), and replace the culture media with fresh TeSR-E8 medium. Subsequently, perform daily media changes. After 3 days from nucleofection, or when cells reach 50–70% confluence (Fig. 3f), begin dual drug selection by adding 0.5 $\mu\text{g}/\text{mL}$ of puromycin and 25 $\mu\text{g}/\text{mL}$ geneticin to the culture media (*see Note 15*). For the first 48 h of selection, further supplement the media with 10 μM Y-27632 (*see Note 9*). Selection should be complete within 48–72 h (Fig. 3g, h; *see Notes 8 and 16*), after which the drugs can be withdrawn. Individual colonies should reach an appropriate size for passaging (1–2 mm; Fig. 3i) by 7–10 days after nucleofection (*see Note 17*).

3.3 Genotyping of OPTIKO hPSCs

1. Mechanically pick an individual colony by using a micropipette equipped with a sterile tip. Use a microscope to facilitate visualization during the procedure. Transfer each colony to a single well of a 24-well tissue culture petri dish, then gently pipette 5–10 times to triturate. Add penicillin-streptomycin to the culture media to minimize the risk of bacterial contamination, and further supplement it with 10 μM Y-27632 to promote hPSC survival. Repeat the procedure for 8–12 colonies (*see Note 18*).
2. Once they reach 50–70% confluency, split each clonal line into two wells of a 24-well plate: one with approximately 1/3 of the cells and the second with the remaining 2/3. Cells in the first well will be grown, while the second well will be used for genotyping.

3. When the cells prepared for genotyping reach >50% confluency, extract genomic DNA using the Wizard Genomic DNA Purification Kit according to the manufacturer's instructions. Adjust the DNA concentration to 25–50 ng/ μ L.
4. Individually assemble the genotyping polymerase chain reactions (PCR) detailed in Table 1 using LongAmp Taq DNA Polymerase (*see* Note 19). For each reaction, prepare a 10 μ L mix containing: 100 ng of genomic DNA, 0.3 μ L (300 μ M) dNTP mix, 0.5 μ L (250 nM) forward primer, 0.5 μ L (250 nM) reverse primer, 0.2 μ L (2%) DMSO, and 0.4 μ L (10 U) LongAmp Taq Polymerase (remaining volume dH₂O). Inclusion of positive and negative controls is strongly recommended (*see* Note 20).
5. Perform the PCR in a thermocycler according to the following program (lid heated at 95 °C): (1) 94 °C for 5 min; (2) 94 °C for 15"; (3) annealing temperature (*see* Table 1) for 30"; (4) 65 °C for extension time (*see* Table 1); (5) repeat steps 2 to 4 for a total of 35 cycles; (6) 65 °C for 3 min; (7) hold at 10 °C.
6. Perform DNA gel electrophoresis for half of each PCR reaction with a 1% agarose-TAE gel, and visualize the results with a UV transilluminator (*see* Subheading 3.1, step 4; Fig. 4b).
7. Determine the genotype of OPTiKO clones by referring to Table 2 (*see* Note 19). Only clones showing dual targeting of both the Cas9 and sgRNA transgenes are OPTiKO hPSCs that should be kept for further experiments.
8. Identify the clones of interest from the second 24-well plate from Subheading 3.3, step 2. Once they are ready to be passaged split them for expansion, banking, and experimental analyses.

3.4 Validation of OPTiKO hPSCs

1. Knockout can be induced in hPSCs or hPSC-derived cells by adding the drug tetracycline to the culture media at a concentration of 1 μ g/mL (*see* Notes 21 and 22). The duration of tetracycline treatment required to induce knockout in the majority of the cells will depend on the efficiency of the sgRNA and on the cell type (Fig. 5), but ought to be in the range of 2–10 days (*see* Note 23).
2. Validation of gene knockout should be performed by means of appropriate techniques described in detail elsewhere (*see* Notes 24 and 25). The quantification of CRISPR/Cas9-induced indels on the genomic locus can be efficiently assessed using the Surveyor or T7 endonuclease 1 assay (Fig. 5a), by DNA Sanger sequencing, and/or by next-generation sequencing [17, 25] (*see* Note 26). Mutations introducing premature stop codons will often result in decreased transcript levels due

Table 2
Inferring the genotype of OPTiKO clonal lines from PCR results

5'- and 3'-INT sgRNA	5'- and 3'-INT Cas9	Locus	5'- and 3'-BB (any)	Possible genotype ^a
Both bands at expected size	Both bands at expected size	No band	No bands	Dual targeting of both sgRNA and Cas9
Both bands of expected size	Both bands of expected size	No band	At least one band at expected size	Dual targeting of both sgRNA and Cas9, with additional sgRNA/Cas9 plasmid copies ^b
Both bands at expected size	One or more band absent or at incorrect size	Band at expected size	Any	Targeting of sgRNA only
One or more band absent or at incorrect size	Both bands at expected size	Band at expected size	Any	Targeting of Cas9 only
Both bands at expected size	Both bands of expected size	Band at expected size	Any	Population is likely not clonal.
No bands	No bands	Band at expected size	At least one band at expected size	Random transgene integration

^aOnly the most common results are described. Refer to **Note 19** and **Fig. 4** to interpret other possible genotyping results

^bAlternatively the targeting plasmid is still present in the cells as an episome

to nonsense-mediated decay, which can be easily measured by quantitative reverse transcription PCR (RT-qPCR) [26]. Finally and most importantly, loss of the protein product can be validated by flow cytometry (**Fig. 5b**), immunocytochemistry (**Fig. 5c–e**), and/or Western blot [16, 27–30] (*see Note 27*).

- Validated OPTiKO hPSCs may now be used to investigate the biological question of interest (**Fig. 5f**; *see Note 28*).

4 Notes

- sgRNA design should consider several aspects including the cut site location (conserved and constitutive exons located toward the start of the coding sequence and/or encoding for crucial protein domains are to be preferred), and the predicted activity and specificity of the sgRNA [31, 32]. Publicly available resources for sgRNA design are numerous and reviewed

elsewhere [33]. We recommend the online tools CRISPR design (<http://crispr.mit.edu>) or GuideScan [34], or the standalone software Protospacer Workbench [35].

2. Even more so than for conventional CRISPR/Cas9 applications, identification of an efficient sgRNA is the most critical factor in the success of OPTiKO experiments [16, 17]. The sgRNA will be expressed from a single transgenic copy, and will therefore be in limiting amounts. As a result, efficient gene editing will be achieved only with a sgRNA highly effective in inducing double-strand DNA breaks (*see* **Note 23**). It is therefore highly recommended to prescreen 3 to 5 sgRNAs using an appropriate method, such as the inexpensive T7 endonuclease I (T7EI) assay [36], and to select the sgRNA showing the highest activity.
3. Distinctly from the U6 promoter (which prefers transcription of sgRNAs starting with guanine), either a guanine or an adenosine may support efficient and precise transcription from the H1 Pol III promoter [37, 38]. Addition of a guanine at the start of sgRNAs not satisfying this criterion does not substantially influence gene editing efficiency.
4. AarI requires binding to two copies of its recognition sequence for cleavage (the second copy is provided by the oligonucleotide to be included in the digestion reaction), and is a slow cutter that requires extended incubation for complete plasmid digestion.
5. Occasional AarI star activity after overnight digestion can lead to the presence nonspecific products smaller in size compared to the expected linearized plasmid. Provided that the linearized plasmid is the predominant digestion product and it can be clearly identified and distinguished after an appropriate electrophoretic run, this will not interfere with the efficiency of subsequent ligation. If this is not the case, however, the digestion time should be shortened to 4–8 h.
6. We recommend screening 4 to 8 bacterial clones. If the procedure was successful, the expected recombinant DNA should be found in more than 90% of the clones. Should the efficiency prove significantly lower, colony PCR can be used as a rapid method to screen several clones [16]. For this, use the siKO_fw primer and the bottom sgRNA oligonucleotide as reverse primer in order to detect an approximately 250 bp product in the clones carrying the sgRNA. Additionally, correct recombinant clones can be screened by diagnostic digestion with AarI, as this will be unable to digest the desired recombinant DNA (provided no AarI site is found within the sgRNA sequence).
7. The culture conditions recommended in this protocol rely on commercially available and commonly used reagents for feeder-

and xeno-free hPSC culture. Nevertheless, OPTiKO hPSCs have been successfully derived from cells maintained in various culture conditions including gelatin-fibronectin coating with an *in house* chemically defined media, matrigel coating with mTeSR1, and laminin-521 coating with E8 [16, 17].

8. A positive control can be performed by using an *AAVSI* targeting vector expressing a constitutive EGFP transgene (AAV-CAGGS-EGFP) instead of the OPTiKO plasmids. Cells can be visually monitored for EGFP expression the day following nucleofection to confirm efficient plasmid delivery, and the nucleofected cells can be subjected to drug selection to confirm efficient gene targeting of the *AAVSI* locus (*see* Subheading 3.2, step 9). Note that this targeting vector confers resistance only to puromycin, so geneticin should be omitted. A negative control nucleofection without any plasmid can be performed to monitor the efficiency of drug selection: the nucleofected cells should be completely killed within 48–72 h of drug selection. Finally, an additional negative control nucleofection only with the OPTiKO plasmids but omitting the *AAVSI* zinc finger nucleases can be performed to confirm the low rate of transgene integration outside of the *AAVSI* locus: no or very few colonies should survive drug selection.
9. The use of antibiotics is optional at all other stages of the procedure but must be avoided the day before and the day after the nucleofection, as they are toxic in cells with increased cell membrane permeability. The addition of ROCK inhibitor before and after nucleofection increases hPSC survival by inhibiting apoptosis [39].
10. Viability should be greater than 90%. If this is significantly lower, increase the duration of Accutase treatment up to 7 min to reduce the amount of mechanical trituration to the minimum required to obtain 3–4 cell clumps.
11. Use of a pipette tip with a large orifice is key to limit shear stress on nucleofected cells. Pipetting should be very gentle and minimized as much as possible. In particular, cells post-nucleofection are incredibly fragile and should be mixed only a couple of times, preferably using a large bore pipette tip.
12. Ensure that no air bubbles are trapped at the bottom of the cuvette. If so, a sharp tap should be sufficient to dislodge them.
13. The combination of the recommended nucleofection solution and nucleofection program has been optimized for H9 hESCs, and proved to work efficiently also in hiPSCs (40–60% nucleofection efficiency). Nevertheless, this might require some degree of optimization for certain hPSC lines. Consult the nucleofector manufacturer's instructions for alternative nucleofection solutions and programs suitable for hPSCs.

14. An alternative to the recommended feeder-free procedure is plating the nucleofected cells onto feeder layers of mitotically inactivated MEF cells [16]. While more expensive, laborious, and not xenofree, this approach may increase gene editing efficiency of certain hPSC lines due to improved survival both after nucleofection and during drug selection. In this case, utilize DR4 MEF (Applied StemCell) that are genetically engineered to contain resistance to puromycin and neomycin, and top up the feeder layer as needed during drug selection (some cell death of the MEF feeder will be observed despite their drug resistance).
15. The recommended drug doses have been optimized for H9 hESCs, and proved to work efficiently also for hiPSCs. Nevertheless, based on the results of the positive and negative controls described in **Note 8**, optimization might be required for certain hPSC lines. In this case, perform a kill curve experiment to identify the minimal dose required to completely select wild-type hPSCs within 48–72 h of drug selection (range: 0.25–2 $\mu\text{g}/\text{mL}$ of puromycin and 12.5–200 $\mu\text{g}/\text{mL}$ geneticin).
16. During the first 48 h of selection it can be beneficial to increase the media volume to 20 mL and to perform media changes every 12 h in order to rapidly remove dead cells, thus reducing the stress they impose on the surviving hPSCs.
17. 5–50 individual hPSC colonies should be obtained in each of the two 100 mm dish. An individual colony is expected to result from clonal expansion of a single cell that experienced the rare gene editing event in which both *AAVS1* alleles have been targeted with the two different transgenes. Therefore, provided that such a pseudo-clonal colony is not in close proximity to any other, this will be considered and referred to as “clonal.” Should a more stringent way to generate clonal lines be desirable, hPSCs should be plated as single cells into 96- or 384-well plates and grown as *bona fide* clones.
18. Given the high efficiency of the gene targeting procedure, the proportion of correctly targeted hPSCs is expected to be greater than 90% [16]. As a matter of fact, while clonal isolation is recommended in order to obtain an isogenic population, this may be entirely bypassed in experimental situations where this is not an important requirement (such as for analyses that will be performed at the single cell level). In such case, simply passage the whole 100 mm dishes into new plates and proceed to validation (Subheading 3.4).
19. The recommended genotyping strategies are illustrated in Fig. 4a. The 5'- and 3'-integration (INT) PCRs are designed to verify site-specific integration of the transgenes. Targeting of

both alleles of the *AAVSI* locus is further confirmed by loss-of-allele PCR for the wild-type locus (this PCR fails following successful transgene integration due to the large size of the amplicon and the high GC content of the CAG promoter). Finally, 5'- and 3'-backbone (BB) PCRs allow to determine if the targeting plasmids have been integrated in a random genomic region. Note that the BB PCRs are optional, as the presence of additional transgenic copies does not interfere with the functionality of the method. Nevertheless, selection of clones with exactly one copy of each transgene can be advantageous in certain experimental settings, for instance when comparing the efficiency of different sgRNAs [17].

20. Recommended controls are: wild-type genomic DNA, 100 ng; pAAV-Neo_CAG-Cas9 (Cas9 targeting plasmid), 1 ng; pAAV-Puro_siKO-sgRNA (sgRNA targeting plasmid), 1 ng; and no template control, dH₂O (Fig. 4b).
21. This dose of tetracycline is not toxic to hPSCs and does not interfere with hPSC differentiation into multiple lineages from all the germ layers [16]. Nevertheless, this might require optimization for a specific sensitive cell type (range: 0.05–2 µg/mL). Tetracycline is unstable in aqueous solutions, and should be added fresh from single-use freshly thawed aliquots (*see* Subheading 2.3). Media changes should be performed at least every other day since the half-life of tetracycline at 37 °C is of approximately 24 h.
22. Given the widespread use of tetracycline as an antibiotic in livestock animals, animal-derived products such as fetal bovine serum (FBS) may be contaminated with tetracycline. When such reagents have to be included as part of the culture media to be used for hPSC maintenance or differentiation, batch testing to confirm lack of detectable tetracycline contamination is strongly recommended. Tetracycline-free FBS is available from commercial suppliers.
23. As introduced in **Note 2**, the speed and efficiency of knockout will be predominantly dictated by the activity of the sgRNA. Should the performance of a given sgRNA be insufficient to appropriately investigate the biological question at hand, an alternative is to build a pAAV-Puro_siKO plasmid containing multiple copies of the sgRNA. This can be achieved by a one-step Gibson assembly reaction using PCR-amplified inducible sgRNA expression cassettes [16]. Of note, this method can also be applied to express multiple distinct sgRNAs against the same gene, or multiple sgRNAs against different genes.
24. For most sgRNAs the level of gene knockout in the absence of the inducer tetracycline should be minimal (less than 5%) even

after prolonged passaging [16, 17]. However, rare exceptionally potent sgRNAs can induce even substantial premature gene knockout. In such case, it is recommended to repeat the procedure using an alternative targeting vector for sgRNA cloning: pAAV-Puro_siKO-2TO (Addgene #86697). In this plasmid the inducible H1 promoter has an additional tetO2 binding site for the tetR before the TATA box, thus reinforcing transcriptional repression in the absence of tetracycline [16, 40]. Note that this modification also results in lower sgRNA expression levels after addition of tetracycline, thus making it a valuable option only in case of exceptionally strong sgRNAs.

25. The kinetics by which acquisition of knockout mutations leads to loss of the protein product is heavily dependent on the stability of both the mRNA and the protein. Furthermore, decrease in protein levels will be more marked in dividing cells (which will serially dilute the protein after each cell cycle) compared to non-proliferative cells. Therefore, the timing required for efficient knockout ought to be determined for each gene and cell type to be studied.
26. The same methods should also be applied to determine the degree of possible CRISPR/Cas9 off-targets effects based on *in silico* predictions. Indeed, while these ought to be minimized by careful design of the sgRNA (*see Note 1*), the method relies on wild-type *Streptococcus Pyogenes* Cas9 protein, which is known to suffer from some degree of off-target activity [41, 42].
27. The OPTiKO method can also be applied to introduce indels in noncoding regions of the genome, which represent an important proportion of disease-associated sequences [1, 2]. In this case, validation will primarily focus on determining the mutations at the DNA level.
28. Analysis of OPTiKO cells maintained in parallel and cultured in the absence of tetracycline must be always implemented to provide a reference for the experiment. Additionally, when investigating a new cell type and/or biological process we recommend to also analyze cells treated with tetracycline but lacking the specific inducible sgRNA, in order to control for potential nonspecific effects due to the drug [43, 44]. For this, OPTiKO cells carrying a scrambled sgRNA are the ideal control, as they also allow to monitor for potential nonspecific effects of CRISPR/Cas9 expression. Alternatively, use wild-type cells, or cells expressing the Cas9 transgene. Since clonal isolation of hPSCs can exacerbate biological variability, it is important to analyze multiple OPTiKO clones for the same sgRNA to ensure that the results are reproducible. Finally,

analysis of OPTiKO cells carrying separate sgRNAs designed to target the same gene will reinforce the conclusion of any experimental analysis by ruling out the possibility that the observed phenotypes are due to unappreciated off-targets of the sgRNA.

Acknowledgments

The authors would like to thank Dr. Mark Kotter, Dr. Sanjay Sinha, and the members of the Vallier, Kotter, and Sinha laboratories for their help with the validation of this methodology. We would also like to thank Dr. Kosuke Yusa for providing the *AAVSI* ZFN plasmids. This work was supported by a British Heart Foundation PhD Studentship (FS/11/77/39327; A.B.); a European Research Council starting grant Relieve IMDs (281335; L.V., A.B.); the Cambridge University Hospitals National Institute for Health Research Biomedical Research Centre (L.V., K.E.S.); the British Heart Foundation Oxbridge Centre of Regenerative Medicine (J.D.C.), and a core support grant from the Wellcome Trust and Medical Research Council to the Wellcome Trust—Medical Research Council Cambridge Stem Cell Institute (PSAG028; L.V.).

References

1. Cooper GM, Shendure J (2011) Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. *Nat Rev Genet* 12:628–640
2. Zhang F, Lupski JR (2015) Non-coding genetic variants in human disease. *Hum Mol Genet* 24:R102–R110
3. Trounson A, DeWitt ND (2016) Pluripotent stem cells progressing to the clinic. *Nat Rev Mol Cell Biol* 17:194–200
4. Avior Y, Sagi I, Benvenisty N (2016) Pluripotent stem cells in disease modelling and drug discovery. *Nat Rev Mol Cell Biol* 17:170–182
5. Murry CE, Keller G (2008) Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* 132:661–680
6. Cong L, Ran FA, Cox D et al (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823
7. Mali P, Yang L, Esvelt KM et al (2013) RNA-guided human genome engineering via Cas9. *Science* 339:823–826
8. Jinek M, Chylinski K, Fonfara I et al (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816–821
9. Wright AV, Nuñez JK, Doudna JA (2016) Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* 164:29–44
10. Hsu PD, Lander ES, Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157:1262–1278
11. González F, Zhu Z, Shi Z-D et al (2014) An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. *Cell Stem Cell* 15:215–226
12. Chen Y, Cao J, Xiong M et al (2015) Engineering human stem cell lines with inducible gene knockout using CRISPR/Cas9. *Cell Stem Cell* 17:233–244
13. Mandegar MA, Huebsch N, Frolov EB et al (2016) CRISPR interference efficiently induces specific and reversible gene silencing in human iPSCs. *Cell Stem Cell* 18:541–553
14. Haenebalcke L, Goossens S, Naessens M et al (2013) Efficient ROSA26-based conditional and/or inducible transgenesis using RMCE-compatible F1 hybrid mouse embryonic stem cells. *Stem Cell Rev* 9:774–785
15. Ordovas L, Boon R, Pistoni M et al (2015) Efficient recombinase-mediated cassette exchange in hPSCs to study the hepatocyte

- lineage reveals AAVS1 locus-mediated transgene inhibition. *Stem Cell Rep* 5:918–931
16. Bertero A, Pawlowski M, Ortmann D et al (2016) Optimized inducible shRNA and CRISPR/Cas9 platforms for in vitro studies of human development using hPSCs. *Development* 143:4405–4418
 17. Fogarty NME, McCarthy A, Snijders KE et al (2017) Genome editing reveals a role for OCT4 in human embryogenesis. *Nature* 550(7674):67–73
 18. Brummelkamp TR, Bernards R, Agami R (2002) A system for stable expression of short interfering RNAs in mammalian cells. *Science* 296:550–553
 19. Smith JR, Maguire S, Davis LA et al (2008) Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. *Stem Cells* 26:496–504
 20. Hockemeyer D, Soldner F, Beard C et al (2009) Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol* 27:851–857
 21. DeKolver RC, Choi VM, Moehle EA et al (2010) Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res* 20:1133–1142
 22. Pawlowski M, Ortmann D, Bertero A et al (2017) Inducible and deterministic forward programming of human pluripotent stem cells into neurons, skeletal myocytes, and oligodendrocytes. *Stem Cell Rep* 8:803–812
 23. Gaj T, Gersbach CA, Barbas CF (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* 31:397–405
 24. Bertero A, Yiangou L, Brown S et al (2018) Conditional manipulation of gene function in human cells with optimized inducible shRNA. *Curr Protoc Stem Cell Biol* 44:5C.4.1–5C.4.49
 25. Santos DP, Kiskinis E, Eggen K et al (2016) Comprehensive protocols for CRISPR/Cas9-based gene editing in human pluripotent stem cells. *Curr Protoc Stem Cell Biol* 38:5B.6.1–5B.6.60
 26. Fan H, Robetorye RS (2010) Real-time quantitative reverse transcriptase polymerase chain reaction. *Methods Mol Biol* 630:199–213
 27. Wang L, Gaigalas AK, Yan M (2011) Quantitative fluorescence measurements with multi-color flow cytometry. *Methods Mol Biol* 699:53–65
 28. Willingham MC (2010) Fluorescence labeling of intracellular antigens of attached or suspended tissue-culture cells. *Methods Mol Biol* 588:153–164
 29. Willingham MC (2010) Fluorescence labeling of surface antigens of attached or suspended tissue-culture cells. *Methods Mol Biol* 588:143–151
 30. Komatsu S (2015) Western blotting using PVDF membranes and its downstream applications. *Methods Mol Biol* 1312:227–236
 31. Tsai SQ, Joung JK (2016) Defining and improving the genome-wide specificities of CRISPR–Cas9 nucleases. *Nat Rev Genet* 17:300–312
 32. Mohr SE, Hu Y, Ewen-Campen B et al (2016) CRISPR guide RNA design for research applications. *FEBS J* 283:3232–3238
 33. Graham DB, Root DE (2015) Resources for the design of CRISPR gene editing experiments. *Genome Biol* 16:260
 34. Perez AR, Pritykin Y, Vidigal JA et al (2017) GuideScan software for improved single and paired CRISPR guide RNA design. *Nat Biotechnol* 35:347–349
 35. MacPherson CR, Scherf A (2015) Flexible guide-RNA design for CRISPR applications using Protospacer Workbench. *Nat Biotechnol* 33:805–806
 36. Bloom K, Ely A, Arbuthnot P (2017) A T7 endonuclease I assay to detect talen-mediated targeted mutation of HBV cccDNA. *Methods Mol Biol* 1540:85–95
 37. Ranganathan V, Wahlin K, Maruotti J, Zack DJ (2014) Expansion of the CRISPR-Cas9 genome targeting space through the use of H1 promoter-expressed guide RNAs. *Nat Commun* 5:4516
 38. Ma H, Wu Y, Dang Y et al (2014) Pol III promoters to express small RNAs: delineation of transcription initiation. *Mol Ther* 3:e161
 39. Watanabe K, Ueno M, Kamiya D et al (2007) A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat Biotechnol* 25:681–686
 40. Henriksen JR, Løkke C, Hammerø M et al (2007) Comparison of RNAi efficiency mediated by tetracycline-responsive H1 and U6 promoter variants in mammalian cell lines. *Nucleic Acids Res* 35:e67
 41. Slaymaker IM, Gao L, Zetsche B et al (2015) Rationally engineered Cas9 nucleases with improved specificity. *Science* 351:84–88
 42. Kleinstiver BP, Pattanayak V, Prew MS et al (2016) High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529:490–495

43. Chatzisyrou IA, Held NM, Mouchiroud L et al (2015) Tetracycline antibiotics impair mitochondrial function and its experimental use confounds research. *Cancer Res* 75:4446–4449
44. Moullan N, Mouchiroud L, Wang X et al (2015) Tetracyclines disturb mitochondrial function across eukaryotic models: a call for caution in biomedical research. *Cell Rep* 10:1681–1691

Part IV

CRISPR Gene Editing in Other Cell Types



CRISPR/Cas9 as a Genome Editing Tool for Targeted Gene Integration in CHO Cells

Daria Sergeeva, Jose Manuel Camacho-Zaragoza, Jae Seong Lee, and Helene Fastrup Kildegaard

Abstract

The emergence of CRISPR/Cas9 system as a precise and affordable method for genome editing has prompted its rapid adoption for the targeted integration of transgenes in Chinese hamster ovary (CHO) cells. Targeted gene integration allows the generation of stable cell lines with a controlled and predictable behavior, which is an important feature for the rational design of cell factories aimed at the large-scale production of recombinant proteins. Here we present the protocol for CRISPR/Cas9-mediated integration of a gene expression cassette into a specific genomic locus in CHO cells using homology-directed DNA repair.

Key words Chinese hamster ovary cells, CRISPR/Cas9, Genome editing, Targeted integration

1 Introduction

CHO cells have a privileged position in biopharma industry as the preferred host for the production of recombinant therapeutic proteins. With a long record of regulatory approvals, therapeutic protein production in CHO cells now represents 60–70% of biopharmaceuticals market [1]. Development of new biopharmaceuticals and biosimilars has put more pressure in biopharma industry for shortening cell line generation time and lowering manufacturing costs. This has prompted a renewed interest in genome engineering of CHO cells to harness the whole biosynthetic capacity for recombinant protein production [1]. Achieving this goal calls for the generation of streamlined cell lines by means of genetic engineering approaches aimed at mitigating the metabolic and cellular bottlenecks that limit production.

Whereas small-scale production of recombinant proteins for screening purposes or preclinical studies can be easily achieved by transient gene expression systems, high productivity is largely

dependent on the long-term retention of the plasmid encoding the gene of interest (GOI), which generally decreases over time due to cell division [2]. Consequently, the large-scale protein production is generally accomplished in stable cell lines where the GOI has been integrated into the genome.

Most methods for gene integration can be grouped into two classes: random integration and targeted integration approaches. Current CHO cell line development technologies are mostly based on random integration and amplification of a recombinant gene together with a selection marker, commonly, dihydrofolate reductase (DHFR) or glutamine synthetase (GS) [3]. Integration of transgene into random genomic sites yields cell lines with a wide range of expression, growth, and stability characteristics and requires extensive clone screening to identify stable high producers. In contrast to random integration, targeted integration allows the selection of a transcriptionally active genomic locus that promotes high and stable transgene expression with minimal perturbation of the genetic context, keeping out confounding factors such as chromosome position effects, copy number variability, and unwanted mutations [4, 5]. Thus, targeted integration approach accelerates development of cell lines with predictable performance and consistent behavior.

Site-specific genome editing tools developed over the last decades rely on engineered nucleases and comprise zinc-finger nucleases (ZFN), transcription activator-like effector nucleases (TALENs), and, the most recent, clustered regularly interspaced short palindromic repeats/CRISPR-associated (CRISPR/Cas) system [6]. These tools can be customized to recognize and cleave a double-stranded DNA molecule at a specific site, which can be subsequently repaired by one of two endogenous cellular pathways mainly: non-homologous end-joining (NHEJ) or homology-directed repair (HDR). NHEJ generates indel mutations at the repair site, which can be harnessed for the knockout generation at the targeted locus. On the other hand, HDR can seamlessly repair DNA break in the presence of a DNA repair template with homologous regions spanning the cleavage site [6]. The DNA template used in HDR can be customized by placing a GOI in between two homology arms and then delivered exogenously in a plasmid vector. Even though HDR occurs at a lower frequency than NHEJ in mammalian cells (particularly in CHO cells) [7], this strategy has been successfully implemented for the targeted integration of a GOI in a genomic locus in several mammalian cells, including CHO cells [4, 5, 8–14].

Compared to ZFNs and TALENs, whose specific site recognition activity relies only on the protein structure of the nuclease by itself, site recognition in the CRISPR/Cas9 system is carried out within a nuclease/RNA complex. Base pairing between a small RNA fragment and the target locus guides the nuclease to the

target site in a sequence-specific manner. The RNA molecule used by this system is called single guide RNA (sgRNA), which is a synthetic fusion of two RNA molecules: a targeting RNA molecule called CRISPR RNA (crRNA) and a scaffold RNA called transactivating crRNA (tracrRNA) [15]. Site recognition occurs by base-pairing of the crRNA portion of the sgRNA and the genomic target sequence when it is immediately next to a small 3 nt sequence called the protospacer adjacent motif (PAM). Upon target recognition, two endonuclease domains within Cas9 (RuvC and HNH) become activated, generating a blunt double-stranded break (DSB) between the third and the fourth nucleotide upstream from the PAM site (Fig. 1a) [15]. The ease, versatility, and speed of creating sgRNAs have fostered the use and evolution of CRISPR/Cas9 as a tool over previous methods, making this system the preferred genome engineering tool currently available.

In this chapter we describe the protocol for CRISPR/Cas9-mediated targeted integration of a gene expression cassette in a predefined genomic site in CHO cells. This method is based on the delivery of three plasmids: Cas9, sgRNA, and a donor plasmid (repair template). The donor plasmid harbors the GOI, a selection marker, and homology arms (HAs) to promote homology-directed

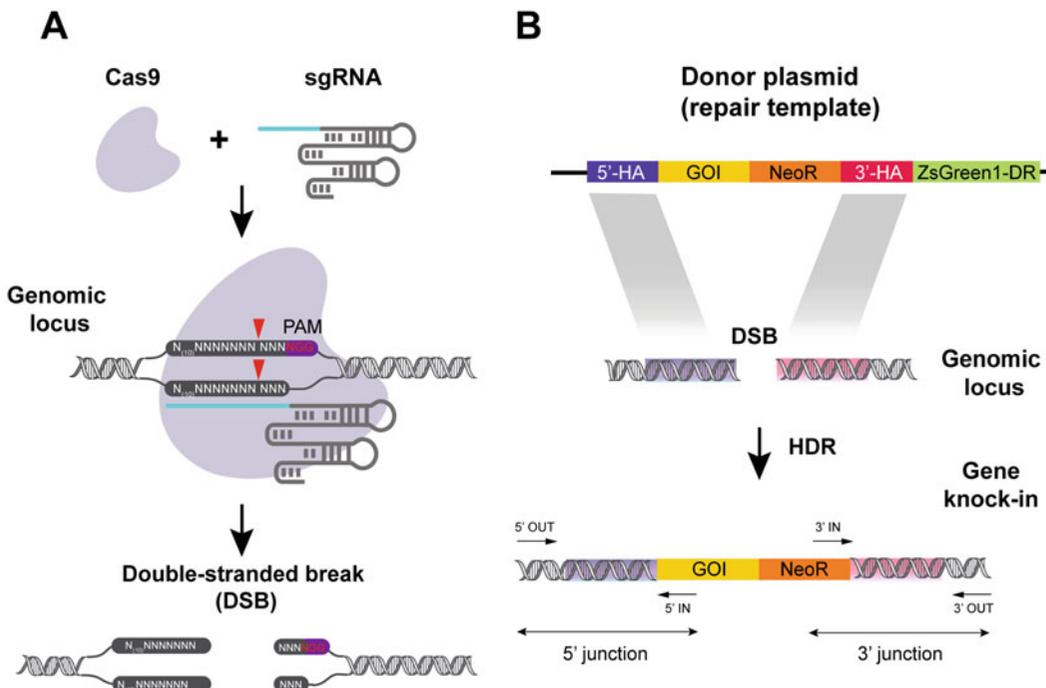


Fig. 1 The use of CRISPR/Cas9 for targeted gene integration. (a) Complex of Cas9 and sgRNA recognizes genomic target sequence in the proximity of PAM site and generates double-stranded break (DSB). (b) DSB can be repaired by homology-directed repair (HDR) when a repair template with homology arms (HAs) is provided. It results in the integration of the GOI in the predefined locus

integration. Upon drug selection and fluorescence-activated cell sorting (FACS), cell lines with homogeneous and stable transgene expression can be generated. This method is easily adaptable to site-specific knock-in of genes in a predefined genomic site in various mammalian cell lines.

2 Materials

2.1 Construction of sgRNA and Donor Plasmid

1. Glycerol stocks of *E. coli* transformed with sgRNA expression plasmid (from Ronda et al. [16]) and donor plasmid (from Lee et al. [5]).
2. 2× YT medium.
3. Kanamycin.
4. Ampicillin.
5. 250–500 mL baffled Erlenmeyer shake flasks.
6. Sterile pipette tips.
7. Incubator with shaker, 37 °C, 250 rpm.
8. Plasmid mini- and midiprep kit.
9. Nuclease-free water.
10. NanoDrop 2000.
11. Oligonucleotides containing sgRNA sequence (for design instructions *see* Subheadings 3.1 and 3.2).
12. PCR primers for amplification of sgRNA backbone and donor plasmid backbone (*see* Subheadings 3.2 and 3.4).
13. PCR primers for amplification of homology arms, drug resistance cassette, and gene expression cassette (*see* Subheadings 3.3 and 3.4).
14. 2× Phusion U Hot Start PCR Master Mix (Thermo Scientific).
15. PCR tubes.
16. Thermocycler.
17. FastDigest DpnI enzyme (Thermo Scientific).
18. 10× FastDigest Green Buffer (Thermo Scientific).
19. 1 kb DNA ladder.
20. 1% agarose gel: 1 g agarose powder dissolved in 100 mL 1× TAE buffer.
21. Gel chamber and power source.
22. PCR and gel purification kit.
23. Heat block.
24. USER enzyme (New England Biolabs).
25. CutSmart buffer (New England Biolabs).

26. 1.5 mL Eppendorf tubes.
27. Mach1 competent *E. coli* cells (Thermo Scientific).
28. Heat block, 37 °C, 300 rpm.
29. Tabletop centrifuge.
30. Sterile spatula.
31. LB-ampicillin and LB-kanamycin agar plates: 15 g/L Agar, 10 g/L Tryptone, 10 g/L NaCl, 5 g/L Yeast Extract, 60 µg/mL ampicillin, or 50 µg/mL kanamycin.

2.2 Preparation of Cas9 Expression Plasmid

1. Glycerol stock with *E. coli* transformed with the codon-optimized Cas9 expression plasmid (from Ronda et al. [16]).
2. Ampicillin.
3. 2× YT medium.
4. 250–500 mL baffled Erlenmeyer shake flask.
5. Sterile pipette tips.
6. Incubator with shaker, 37 °C, 250 rpm.
7. Plasmid midiprep kit.
8. Nuclease-free water.
9. NanoDrop 2000.

2.3 Transfection of CHO-S Cells

1. CHO-S cells (Life Technologies).
2. Cell counter.
3. Growth medium: CD CHO medium (Life Technologies), 8 mM L-glutamine.
4. 15 or 50 mL centrifuge tubes.
5. 6-well plate, flat bottom, non-treated (Corning).
6. Humidified incubator, 37 °C, 5% CO₂, 120 rpm.
7. OptiPro SFM (Life Technologies).
8. FreeStyle MAX reagent (Life Technologies).
9. sgRNA expression plasmid generated in Subheading 3.5.
10. Donor plasmid generated in Subheading 3.5.
11. Cas9 expression plasmid prepared in Subheading 3.6.

2.4 Antibiotic Selection

1. Growth medium: CD CHO medium, 8 mM L-glutamine.
2. Selection drug (hygromycin B or G418).
3. 6-well plate, advanced TC (Greiner bio-one).
4. Humidified incubator, 37 °C, 5% CO₂.
5. TrypLE select (Life Technologies).
6. Phosphate buffered saline (PBS).
7. 6-well plate, flat bottom, non-treated.

2.5 Fluorescence-Activated Cell Sorting (FACS)

1. Fluorescence-activated cell sorter.
2. 384-well plates, flat bottom, tissue culture treated (Corning).
3. FACS sorting medium: CD CHO medium, 8 mM L-glutamine, 1× Antibiotic-Antimycotic (Gibco), and 1.5% HEPES (Life Technologies).
4. FACS tubes.
5. 30 µm cell strainers.
6. Image cytometer or microscope.
7. Humidified incubator, 37 °C, 5% CO₂, no shake.
8. 96-well plates, flat bottom, non-treated (Corning).
9. Clone expansion medium: CD CHO medium, 8 mM L-glutamine, 1× Antibiotic-Antimycotic, and 1 µL/mL Anti-clumping agent (Life Technologies).
10. 96-well plates, V-Shaped (Greiner bio-one).
11. Breathable plastic bag.

2.6 PCR Verification of Clones

1. QuickExtract DNA extraction solution (Epicentre).
2. Primers (for design instructions *see* Subheading 3.10).
3. 2× Phusion PCR Master Mix (Thermo Scientific).
4. Nuclease-free water.
5. PCR tubes or plates.
6. Thermocycler.
7. 1 kb DNA ladder.
8. 1% agarose gel.
9. Gel chamber and power source.

2.7 Expansion of Clones

1. 12-well plates, flat bottom, non-treated (Corning).
2. 6-well plates, flat bottom, non-treated.
3. 125 mL Erlenmeyer shake flasks.
4. Clone expansion medium: CD CHO medium, 8 mM L-glutamine, 1× Antibiotic-Antimycotic, and 1 µL/mL Anti-clumping agent.
5. DMSO.
6. Cryotubes.
7. Humidified incubator, 37 °C, 5% CO₂, 120 rpm.

2.8 Copy Number Analysis

1. GeneJet Genomic DNA purification kit (Thermo Scientific).
2. NanoDrop 2000.
3. 20× TaqMan Gene Expression Assays for GOI and COSMC (*see* Subheading 3.11 for instructions).

4. 2× TaqMan Gene Expression Master Mix (Applied Biosystems).
5. Nuclease-free water.
6. Real-time PCR instrument.
7. Reaction tubes or plates suitable for qRT-PCR.

3 Methods

The following section is a general protocol that we use for site-specific integration in CHO-S cells in our lab (Fig. 2). The protocol includes generation of sgRNA plasmid, Cas9 plasmid, and donor plasmid with GOI by USER cloning [16–18]. After co-delivery of these plasmids and antibiotic selection, cells can be single-cell sorted using FACS. Clones with targeted integration can be verified using PCR analysis of junction sequences and copy number analysis by quantitative real-time PCR (qRT-PCR).

3.1 Design of sgRNA

1. Find a sequence of the genomic locus you want to use for integration in CHO genome (download genomic sequence from NCBI).
2. Copy ~1500–2000 bp of the genomic region and paste it to a CRISPR design tool (e.g., CRISPOR [19] <http://crispor.tefor.net/>).
3. Select appropriate genome and PAM and search for target sequences.
4. In the list of predicted target sequences, select one with the highest specificity score and the lowest off-targets (*see Note 1*).
5. Proceed with sgRNA plasmid construction.

3.2 sgRNA Plasmid Construction by USER Cloning

The sgRNA plasmid consists of two DNA bricks, which can be assembled by USER cloning (Fig. 3): (1) sgRNA backbone, which contains elements for replication in the bacterial host, U6 promoter, sgRNA scaffold and termination signal to generate sgRNA expression cassette. (2) sgRNA annealed oligonucleotides, containing target sequence.

1. Design and order a pair of sgRNA oligos using the following sequences by replacing the N's with your 20 nt target sequence. The first position of the target sequence is a GC pair, shown in bold (if your target sequence starts with another nucleotide, replace it with G for proper U6-driven transcription).

sgRNA_FW_oligo: 5'-GGAAAGGACGAAACACCGNN
NNNNNNNNNNNNNNNNNGTTTTAGAGCTA
GAAAT-3'.

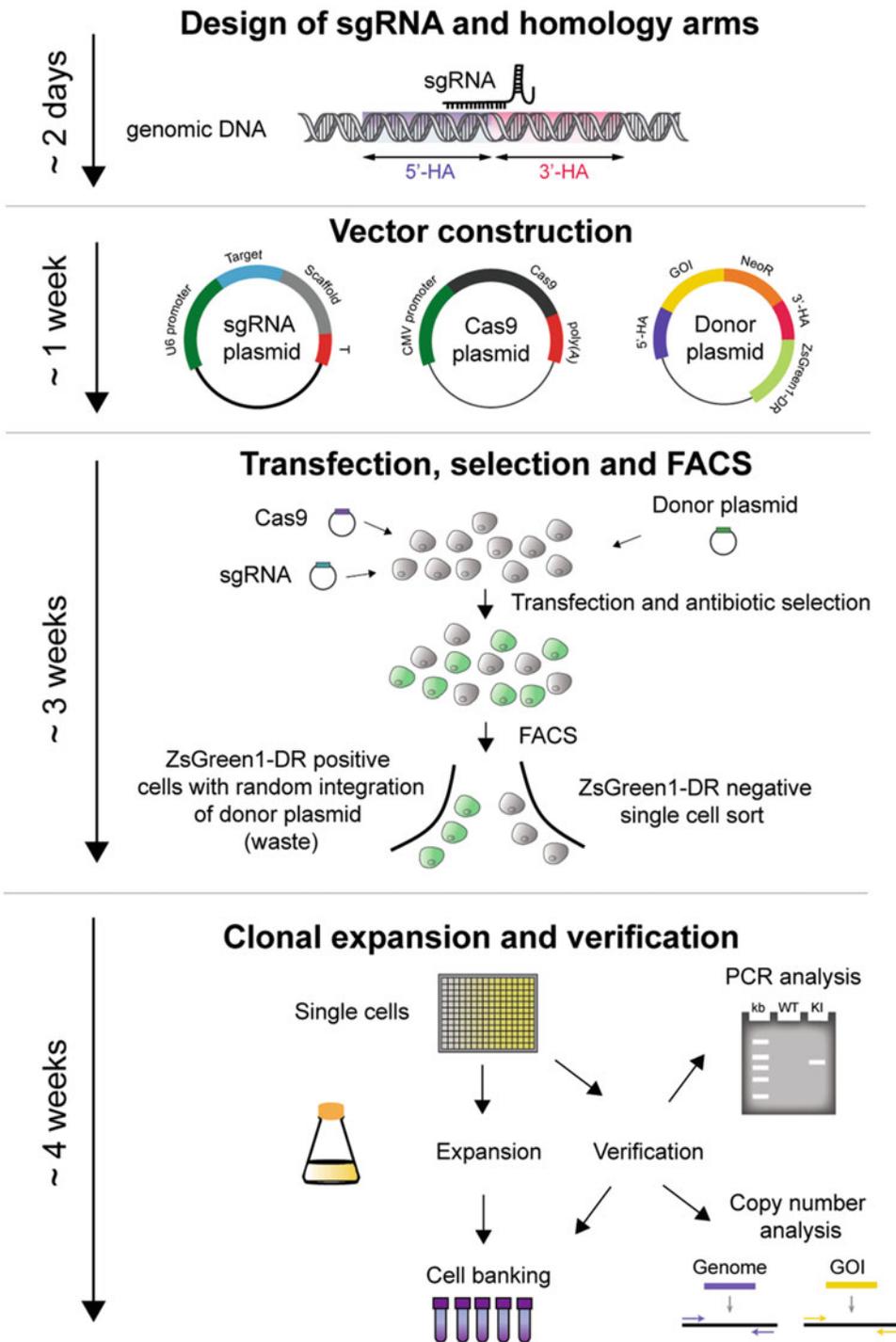


Fig. 2 An overview of the protocol for the generation of CHO cell lines with CRISPR/Cas9-mediated targeted gene integration. The first step of the protocol is to design a sgRNA and homology arms for the specific genomic locus. Second, sgRNA, Cas9, and donor plasmids are constructed. Next, CHO cells are transfected with constructed plasmids and subjected to antibiotic selection and FACS sorting. Clones with targeted integration should be ZsGreen1-DR negative since this fluorescent marker should not be integrated into the genome. Then cells are expanded and verified using PCR and copy number analysis to isolate cell lines with site-specific gene integration

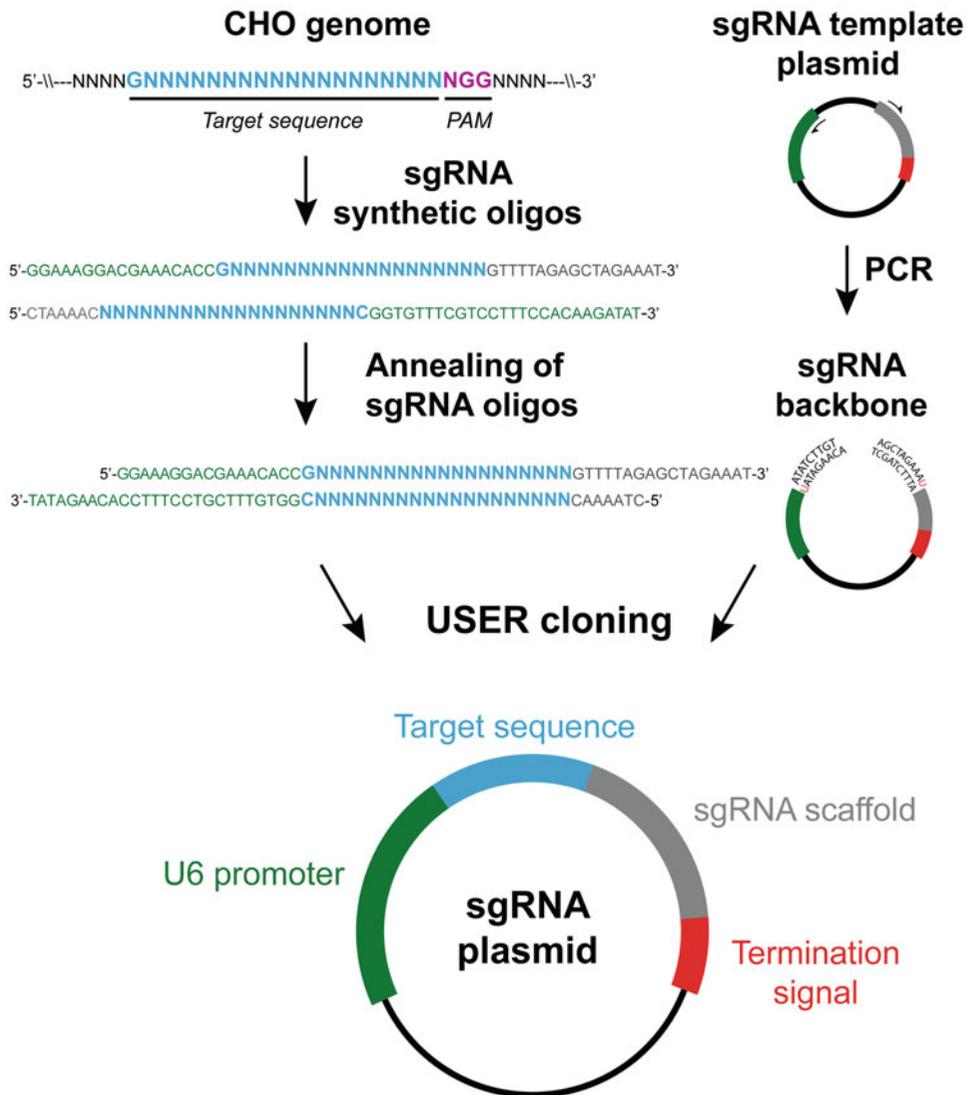


Fig. 3 An overview of the sgRNA plasmid construction. The target sequence with extensions is ordered as two short synthetic oligonucleotides. After annealing these oligos form 3' single-stranded overhangs, which will be used in USER cloning. The universal sgRNA plasmid backbone is amplified using PCR primers containing a single deoxyuracil residue. USER enzyme treatment creates 3'-overhangs on sgRNA plasmid backbone, which allows assembly of the target sequence and backbone, generating sgRNA plasmid

sgRNA_RV_ oligo: 5'-CTAA AACNNNNNNNNNNNNNNNNNNNN
NNNNNNNNCGGTGTTCGTCCTTTCCACAAGA
TAT-3'.

- Anneal sgRNA oligos. Mix the following in an Eppendorf tube: 10 μ L 10 \times NEBuffer 4, 10 μ L sgRNA forward oligo (100 μ M), 10 μ L sgRNA reverse oligo (100 μ M), 70 μ L nuclease-free water.

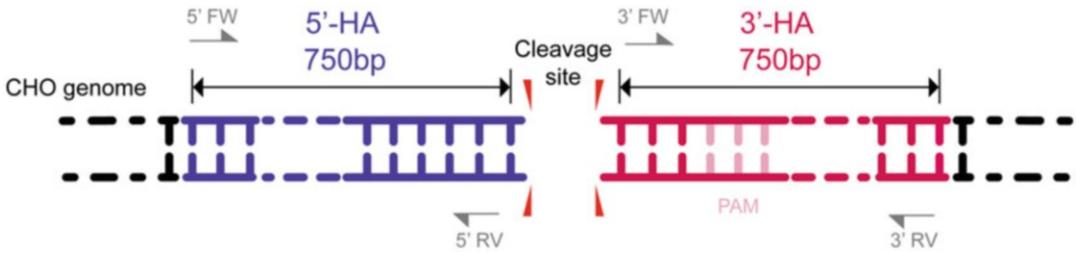


Fig. 4 Design of homology arms. Homology arms are ~750 bp sequences upstream and downstream of CRISPR/Cas9 cleavage site in the genome. The Cas9 cleavage site is located 3 nt upstream of the PAM site

17. Heat shock at 42 °C for 30 s.
18. Place the cells on ice for 1 min and then add 950 μ L of 2 \times YT medium.
19. Incubate at 37 °C for 1 h at 300 rpm.
20. Spin down the cells at 2000 $\times g$ for 5 min.
21. Discard the supernatant and resuspend pellet in 100 μ L of 2 \times YT medium.
22. Plate all the sample (100 μ L) in an agar plate with kanamycin and incubate at 37 °C overnight.

3.3 Design of Homology Arms

1. Open genomic locus sequence from Subheading 3.1 in your preferred sequence viewer software.
2. Identify CRISPR/Cas9 cleavage site as the site between the third and the fourth nucleotide upstream from the PAM sequence within your target site. Your 5' and 3' homology arms (5'-HA and 3'-HA) will correspond to the sequence extending ~750 bp to the left and to the right of the cleavage site, respectively (Fig. 4).
3. Design primers for amplification of homology arms so that they bind approximately 750 bp upstream and downstream of the cleavage site. Using these primers, amplify ~1500 bp of genomic region and use this PCR amplicon as a template for the generation of homology arms.

3.4 Donor Plasmid Construction by USER Cloning

The donor plasmid consists of the following fragments, which can be assembled by USER cloning (Fig. 5): (1) Donor plasmid backbone (amplified from plasmid from Lee et al. [5]). This backbone includes elements for replication in the bacterial host and ZsGreen1-DR expression cassette as a fluorescent reporter to rule out cells with random integration of the donor plasmid using FACS. (2) 5' and 3' homology arms (~750 bp each) to mediate homology-directed targeted integration (amplified from genomic DNA or synthetic template). (3) Drug resistance cassette, harboring neomycin or hygromycin resistance gene together with

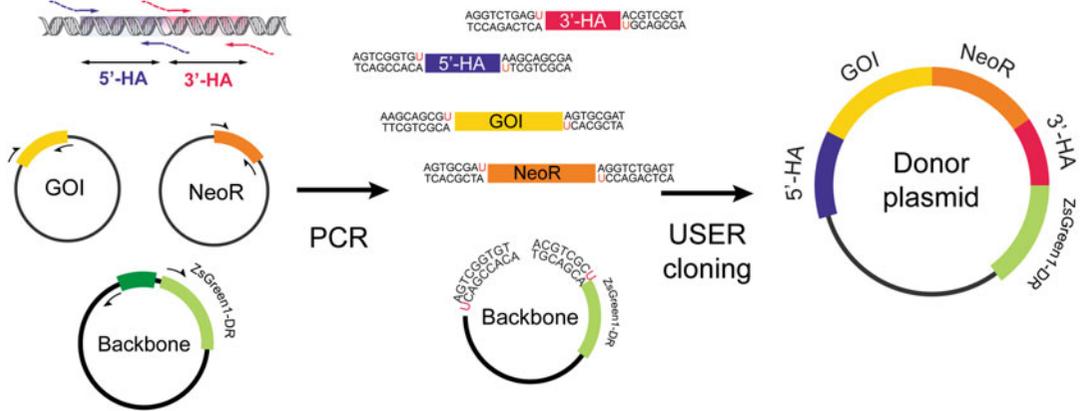


Fig. 5 An overview of the donor plasmid construction. Five DNA bricks (5' and 3' - homology arms (HAs), gene expression cassette (with GOI), drug resistance cassette (e.g., NeoR) and plasmid backbone) are amplified using primers containing USER-linkers. After USER treatment PCR fragments are assembled in the desired order and form the donor plasmid

promoter and terminator sequences (amplified from another plasmid or synthetic template). (4) Gene expression cassette, harboring gene of interest (GOI) with the desired promoter and terminator sequence (amplified from another plasmid or synthetic template).

1. Request the donor backbone plasmid from Lee et al. [5] and make a bacterial glycerol stock.
2. Inoculate bacterial stock in 4 mL of 2 × YT medium in a 14 mL bacterial culture tube with 60 µg/mL ampicillin. Incubate overnight at 37 °C with shaking at 250 rpm.
3. Extract plasmid using a miniprep kit. Resuspend in nuclease-free water and quantify using NanoDrop 2000.
4. Design primers for amplification of homology arms (HAs), gene expression cassette (GOI), and drug resistance (DR) cassette. Primers for HAs should bind to 5' and 3' ends of genomic region and at the cleavage site (Fig. 4). Add following uracil-containing overhangs on 5'-end of your primers for USER cloning (*see Note 2*):
 - 5'-HA_FW primer tail (Linker A): 5'-AGTCGGTGU-3'.
 - 5'-HA_RV primer tail (Linker B): 5'-ACGCTGCTU-3'.
 - GOI_FW primer tail (Linker B): 5'-AAGCAGCGU-3'.
 - GOI_RV primer tail (Linker O2): 5'-ATCGCACU-3'.
 - DR_FW primer tail (Linker O2): 5'-AGTGC GAU-3'.
 - DR_RV primer tail (Linker D): 5'-ACTCAGACCU-3'.
 - 3'-HA_FW primer tail (Linker D): 5'-AGGTCTGAGU-3'.
 - 3'-HA_RV primer tail (Linker O1): 5'-AGCGACGU-3'.

5. Order following uracil-containing primers for amplification of donor plasmid backbone:
 Donor_BB_FW primer (Linker O1): 5'-ACGTCGCU GTTGACATTGATTATTGACT-3'.
 Donor_BB_RV primer (Linker A): 5'-ACACCGACU GAGTCGAATAAGGGCGACACCCCA-3'.
6. Amplify DNA bricks for USER cloning. For homology arms amplification use the DNA of genomic region (25–250 ng) amplified in the Subheading 3.3. For amplification of donor plasmid backbone use plasmid (1–10 ng) prepared in step 3. For amplification of DR and GOI use your source of sequence (another plasmid or synthetic template).
7. Mix the following in a PCR tube: 2.5 μ L primer forward (10 μ M), 2.5 μ L primer reverse (10 μ M), 25 μ L 2 \times Phusion U Hot Start PCR Master Mix, 1 μ L template, 19 μ L nuclease-free water.
8. Place the PCR tubes in the thermocycler and run the following program: 98 $^{\circ}$ C for 30 s; 35 cycles of 98 $^{\circ}$ C for 10 s, T_m for 30 s, 72 $^{\circ}$ C for 15–30 s/kb; 72 $^{\circ}$ C for 10 min.
9. If plasmid was used as a template for PCR reaction, treat the PCR product with DpnI enzyme as was described in the Subheading 3.2. Purify DNA bricks and measure concentration using NanoDrop 2000.
10. Perform USER cloning reaction. Mix the components with an equimolar ratio in a PCR tube as shown in Table 1.
11. Incubate the mix at 37 $^{\circ}$ C for 40 min and 25 $^{\circ}$ C for 30 min.

Table 1
Components for donor plasmid assembly

Component	USER reaction (μ L)	Negative control (μ L)
Donor plasmid backbone	1	1
5'-HA	1	–
3'-HA	1	–
Drug resistance cassette	1	–
Gene expression cassette	1	–
CutSmart buffer	1	1
USER enzyme	1	1
Nuclease-free water	3	7

12. Add 1.5 μL of USER reaction to 15 μL of *E. coli* Mach1 competent cells in an Eppendorf tube and incubate on ice for 30 min.
13. Heat shock at 42 °C for 30 s.
14. Place the cells on ice for 1 min and then add 950 μL of 2 \times YT medium.
15. Incubate at 37 °C for 1 h at 300 rpm.
16. Spin down the cells at 2000 $\times g$ for 5 min.
17. Discard the supernatant and resuspend pellet in 100 μL of 2 \times YT medium.
18. Plate the sample (100 μL) in an agar plate with ampicillin and incubate at 37 °C overnight.

3.5 Analysis and Preparation of sgRNA and Donor Plasmids

1. Pick colonies from each plate of sgRNA and donor plasmid transformation using a sterile pipette tip and inoculate 4 mL of 2 \times YT medium with the corresponding antibiotic in a 10 mL bacterial culture tube.
2. Incubate at 37 °C and 250 rpm overnight.
3. Extract plasmid using a miniprep kit and use this preparation for Sanger sequencing. Use sequencing primers that bind before the U6 promoter of the sgRNA plasmid and sequencing primers covering both 5' and 3' homology arms and sequence in between these elements of the donor plasmid.
4. Analyze sequencing results using a sequencing analysis software to make sure that your plasmids do not contain any mutations.
5. Grow a bacterial culture of correct transformants for plasmid purification. Inoculate 100–200 mL of 2 \times YT medium in baffled shake flasks, supplemented with the corresponding antibiotics, and incubate at 37 °C and 250 rpm overnight.
6. Extract plasmid using a midiprep kit and resuspend in endotoxin-free water. Quantify plasmid yield using NanoDrop 2000 and dilute plasmids to 500–1000 ng/ μL . These preparations will be used for the CHO-S transfection.

3.6 Preparation of Cas9 Expression Plasmid

1. Request Cas9 plasmid from Ronda et al. [16].
2. Inoculate 100–200 mL of 2 \times YT medium with 60 $\mu\text{g}/\text{mL}$ ampicillin in shake flasks with the bacterial stock using a sterile pipette tip and incubate overnight at 37 °C and 250 rpm.
3. Extract plasmid using a midiprep kit and resuspend in endotoxin-free water. Quantify plasmid DNA using NanoDrop 2000, dilute it to 500–1000 ng/ μL , and save this preparation for the transfection.

3.7 Transfection of CHO-S Cells

1. Cultivate CHO-S cells in the shake flask using CD CHO medium supplemented with 8 mM L-glutamine. For transfection use CHO-S cells at low passage and viability above 95%.
2. One day before transfection measure viable cell density (VCD) and calculate the volume needed to seed at 7×10^5 cells/mL in 30 mL of fresh medium.
3. Spin down the required volume $200 \times g$ for 5 min and discard the supernatant.
4. Resuspend cells in 5 mL of prewarmed medium (CD CHO + 8 mM L-glutamine) and transfer to 25 mL of media in a 125 mL shake flask.
5. Incubate cells at 37 °C, 5% CO₂, and shake 120 rpm for 16–24 h.
6. On the day of transfection, measure VCD and dilute cells in prewarmed media (CD CHO + 8 mM L-glutamine) to a final VCD of 1×10^6 cells/mL.
7. Add 3 mL of diluted cells to a 6-well plate and place the plate in the incubator at 37 °C, 5% CO₂, and shake 120 rpm.
8. Dilute plasmids in OptiPRO SFM to a final volume 60 μL. Mix sgRNA, Cas9 and donor plasmids using a total amount of 3.75 μg of DNA in a 1:1:1 (w/w) ratio (Cas9:sgRNA:donor).
9. Invert transfection reagent tube (FreeStyle MAX). Dilute 3.75 μL of transfection reagent with OptiPRO SFM to a total volume 60 μL. Mix gently. Incubate for 5 min.
10. Add diluted transfection reagent to diluted plasmid mix.
11. Incubate lipid-DNA mix for 8 min.
12. Immediately after incubation, transfer the transfection mix (120 μL) to cells in the 6-well plate.
13. Incubate cells at 37 °C, 5% CO₂, and shake 120 rpm for 3 days (*see Note 3*).

3.8 Antibiotic Selection

1. Measure VCD of transfected cells and calculate the volume needed to seed at 3×10^5 cells/mL in 3 mL of medium (CD CHO + 8 mM L-glutamine).
2. Spin down the calculated volume $200 \times g$ for 5 min and discard the supernatant.
3. Resuspend the pellet in 3 mL of fresh growth medium with added selection drug (*see Note 4*) and seed in a 6-well advanced TC plate.
4. Incubate at 37 °C, 5% CO₂, no shake for 3–4 days. Live cells will attach to the plate.
5. After 3–4 days carefully exchange spent medium for fresh medium with selection drug, removing dead cells present in the spent media.

6. Repeat **step 5** every 3–4 days.
7. Detach cells after around 2 weeks of selection. Remove old media, briefly wash cells with 3 mL PBS, add 0.3 mL of TrypLE reagent, and incubate for 3–5 min at 37 °C. Then add 3 mL of prewarmed selection media to the cells.
8. Measure VCD. Cell concentration and viability of selected cells should reach $5\text{--}10 \times 10^5$ cells/mL and 60–80%, respectively.
9. Transfer detached cells to non-treated 6-well plate and cultivate in suspension at 37 °C, 5% CO₂, 120 rpm.
10. When cells have recovered to >90% of viability, transfer stable cell pool to the 125 mL shake flask in 15–20 mL of media at a seeding density of 3×10^5 viable cells/mL. Keep passaging cells in selection media until FACS.

3.9 Fluorescence-Activated Cell Sorting (FACS)

1. Prepare 384-well flat bottom plates with 30 µL of sorting media for single cell sorting.
2. Filter cells through a 30 µm cell strainer into a FACS tube to eliminate clumps and debris.
3. Use wild-type CHO-S and CHO-S transiently transfected with ZsGreen1-DR as gating control for FACS.
4. Single cell sort stable cell pool, selecting for ZsGreen1-DR negative cells (ZsGreen1-DR positive corresponds to cells with a random integration of the donor plasmid), into one or more prewarmed (37 °C) 384-well plates. If FACS is not available, limited dilution can be used instead (*see Note 5*).
5. Spin plates at $200 \times g$ for 5 min to make sure cells reach the medium.
6. Place cells in a breathable plastic bag to limit evaporation and incubate cells at 37 °C, 5% CO₂, no shake.
7. After 10–14 days check for surviving cells using a microscope or image cytometer. Cell count should preferably be >1000 cells in a well or confluency >50%.
8. Transfer selected colonies from 384-well plates to 96-well plates. Carefully pipet cells up and down and move 30 µL to a 96-well flat bottom plate with 180 µL of clone expansion medium.
9. Leave cells for maximum 4 days. When the clones have a confluency >50%, carefully pipet up and down and transfer 50 µL cell suspension to a 96-well V-shaped plate. Refill the flat-shaped plate with 50 µL of fresh media.
10. Spin down the V-shaped 96-well plate at $1000 \times g$ for 5 min. Remove the supernatant and add 20 µL of QuickExtract DNA extraction solution to cell pellets.

11. Resuspend the pellets and move them to PCR tubes or plates. Incubate at 65 °C for 15 min and 95 °C for 5 min. Store at –20 °C. This DNA will be used for PCR verification of clones.
12. Within 4 following days verify clones using PCR analysis.

3.10 PCR Verification of Clones

1. Design primers that flank 5' and 3' donor:genome junctions (Fig. 1b). You should have primers binding outside of homology arms in the genomic locus (5' and 3' OUT primers), a primer specific to 5' end of gene expression cassette and a primer recognizing 3' end of drug resistance cassette (5' and 3' IN primers). Use 5' OUT primer with 5' IN primer and 3' IN primer with 3' OUT primer to amplify 5' and 3' junction sequences, respectively.
2. Mix the following components in a PCR tube for 5' junction PCR (for one reaction): 10 µL 2× Phusion PCR Master Mix, 1 µL primer 5' forward OUT (10 µM), 1 µL primer 5' reverse IN (10 µM), 1 µL DNA template (genomic DNA extract from Subheading 3.9), 7 µL nuclease-free water. Use genomic DNA extracted from wild-type CHO-S as a template in the negative control.
3. Place the PCR tubes in a thermocycler and run the program for touchdown PCR as shown in Table 2.
4. Run PCR products on a 1% agarose gel and select clones with expected amplicon size. No PCR amplicons with expected size should be present in the negative control.
5. If desired, sequence respective junction PCR amplicons by Sanger sequencing to ensure that the junction sequences are correct.
6. Repeat steps 2–5 for 3' junction PCR on selected 5' junction-positive clones using 3' forward IN primer with 3' reverse OUT primer.

Table 2
PCR program for PCR verification of clones

Temperature (°C)	Time (min)	Number of cycles
98	00:30	1
98	00:10	10
T _m + 10	00:30	
–1 per cycle		
72	15–30 s/kb	
98	00:10	30
T _m	00:30	
72	15–30 s/kb	
72	10:00	1
4	∞	1

3.11 Expansion of Clones

1. Select 5' junction- and 3' junction-positive clones. It indicates that your GOI was inserted in the selected genomic locus.
2. Move the selected clones from the 96-well plate when >90% confluent to a 12-well flat bottom plate.
3. Maintain clones in the 12-well plate until confluent, then move 1 mL of cells to a 6-well flat bottom plate with 2 mL media.
4. When cells are confluent in 6-well plate, harvest 1×10^6 cells for copy number analysis (*see* Subheading 3.12) and replenish the harvested volume with fresh media.
5. When confluent, move cells from 6-well plate to a 125 mL shake flask and seed at 3×10^5 cells/mL in 15–20 mL.
6. Freeze the clones that have one-copy integration (*see* Subheading 3.12). Spin 1×10^7 cells $200 \times g$ 5 min, discard media, resuspend pellet in 1 mL medium (CD CHO + 8 mM L-glutamine) with 5–10% DMSO, and transfer cells to cryotube. Freeze in a Styrofoam box at -80°C the first 24 h before moving to permanent storage at -180°C .

3.12 Copy Number Analysis by qRT-PCR

1. Extract DNA from 1×10^6 cells using Genomic DNA purification kit. Measure the concentration using NanoDrop 2000. Dilute genomic DNA to 10 ng/ μL using nuclease-free water.
2. Design TaqMan assay for your GOI using, e.g., PrimerQuest software (<https://www.idtdna.com/PrimerQuest/>). Order the primers and assess their specificity and efficiency (*see* Note 6). Select specific primers with efficiency between 90 and 105%. Order respective FAM dye-labeled TaqMan probe and test efficiency of your TaqMan assay.
3. Order the following TaqMan assay with VIC dye-labeled MGB probe for endogenous one-copy gene *COSMC* and test its efficiency:
 COSMC_FW primer: 5'-ACCCGAACCAGGTAGTAGAA-3'.
 COSMC_RV primer: 5'-ACATGTCCAAAGGCCCTAAG-3'.
 COSMC probe: 5'-AGTGACAGCCATATTGGAACAGCATCC-3'.
4. Calculate the number of reactions that you need (including no template control) to perform copy number analysis of your clones by qRT-PCR. Have at least three replicates of each reaction.
5. Prepare the reaction mix and pipet in PCR tubes or plates. Mix components (for one reaction): 10 μL 2 \times TaqMan Gene Expression Master Mix, 1 μL 20 \times TaqMan Gene Expression Assay (FAM), 1 μL 20 \times TaqMan Gene Expression Assay (VIC), 2 μL DNA template, 6 μL nuclease-free water.
6. Run samples on a qRT-PCR instrument with following conditions: 50°C for 2 min; 95°C for 1 min; 40 cycles of 95°C for 15 s, 60°C for 1 min.

7. Calculate GOI copy number for each clone using the formula:

$$\text{CN} = \frac{(1 + \text{Ef}(\text{COSMC}))^{\text{Ct}_{\text{mean}}(\text{COSMC})}}{(1 + \text{Ef}(\text{GOI}))^{\text{Ct}_{\text{mean}}(\text{GOI})}}$$

where: CN is the number of GOI copies in the genome, Ct is the threshold cycle, Ef(COSMC) is the efficiency of COSMC TaqMan assay, Ef(GOI) is the efficiency of GOI TaqMan assay.

8. Select clones that have one copy of your GOI (*see Note 7*).

4 Notes

1. We recommend to design and test at least two different sgRNA. Select the sgRNA with the highest DNA cleavage efficiency. To validate sgRNA efficiency you can use T7 endonuclease assay (*see* protocol from [18]).
2. USER cloning allows direct assembly of multiple DNA bricks by using unique 3' single-stranded DNA overhangs (USER linkers). Here we provide universal USER linkers used in our laboratory, although other linkers can be designed and used as well (*see* [17]).
3. Other transfection methods can be used for plasmid delivery, e.g., electroporation. Follow instructions provided by the supplier of electroporation system.
4. The drug concentration recommended for antibiotic selection depends on the selection marker used and must be adjusted for the specific cell line by performing a killing curve experiment with different concentrations of the selection drug. For CHO-S cells harboring neomycin resistance cassette, we recommend using 500 µg/mL G418, for hygromycin resistance cassette we recommend using 600 µg/mL hygromycin. Prepare media containing selection drug fresh each time (drug can be added to each well individually).
5. Limiting dilution is an alternative way to isolate single cells. In this case, you will need to screen more clones, as you cannot rule out cells with random integration of the donor plasmid by FACS enrichment.
6. It is important to validate your primers for qRT-PCR and assess amplification efficiency of a qPCR reaction for accurate measurement of GOI copy number. Design several sets of primers and screen them using the standard curve and melt curve analysis.
7. Digital PCR can be used instead of qRT-PCR for copy number analysis using the same TaqMan assays. Digital PCR is less dependent on primer efficiency and provides a linear response to the number of copies present, which can result in more accurate copy number estimation.

References

1. Templeton N, Young JD (2018) Biochemical and metabolic engineering approaches to enhance production of therapeutic proteins in animal cell cultures. *Biochem Eng J* 136:40–50
2. Lee J-H, Park J-H, Park S-H, Kim S-H, Kim JY, Min J-K, Lee GM, Kim Y-G (2018) Co-amplification of EBNA-1 and PyLT through dhfr-mediated gene amplification for improving foreign protein production in transient gene expression in CHO cells. *Appl Microbiol Biotechnol* 102(11):4729–4739
3. Kim JY, Kim Y-G, Lee GM (2012) CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Appl Microbiol Biotechnol* 93:917–930
4. Lombardo A, Cesana D, Genovese P, Di Stefano B, Provasi E, Colombo DF, Neri M, Magnani Z, Cantore A, Lo Riso P, Damo M, Pello OM, Holmes MC, Gregory PD, Gritti A, Broccoli V, Bonini C, Naldini L (2011) Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat Methods* 8:861–869
5. Lee JS, Kallehauge TB, Pedersen LE, Kildegaard HF (2015) Site-specific integration in CHO cells mediated by CRISPR/Cas9 and homology-directed DNA repair pathway. *Sci Rep* 5:8572
6. Carroll D (2014) Genome engineering with targetable nucleases. *Annu Rev Biochem* 83:409–439
7. Cristea S, Freyvert Y, Santiago Y, Holmes MC, Urnov FD, Gregory PD, Cost GJ (2013) In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration. *Biotechnol Bioeng* 110:871–880
8. Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKolver RC, Katibah GE, Amora R, Boydston EA, Zeitler B, Meng X, Miller JC, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R (2009) Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol* 27:851–857
9. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823
10. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-guided human genome engineering via Cas9. *Science* 339:823–826
11. 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
12. Orlando SJ, Santiago Y, DeKolver RC, Freyvert Y, Boydston EA, Moehle EA, Choi VM, Gopalan SM, Lou JF, Li J, Miller JC, Holmes MC, Gregory PD, Urnov FD, Cost GJ (2010) Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Res* 38:e152
13. Lee JS, Grav LM, Pedersen LE, Lee GM, Kildegaard HF (2016) Accelerated homology-directed targeted integration of transgenes in Chinese hamster ovary cells via CRISPR/Cas9 and fluorescent enrichment. *Biotechnol Bioeng* 113:2518–2523
14. Gaidukov L, Wroblewska L, Teague B, Nelson T, Zhang X, Liu Y, Jagtap K, Mamo S, Tseng WA, Lowe A, Das J, Bandara K, Bajjura S, Summers NM, Lu TK, Zhang L, Weiss R (2018) A multi-landing pad DNA integration platform for mammalian cell engineering. *Nucleic Acids Res* 46:4072–4086
15. 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* 337:816–821
16. Ronda C, Pedersen LE, Hansen HG, Kallehauge TB, Betenbaugh MJ, Nielsen AT, Kildegaard HF (2014) Accelerating genome editing in CHO cells using CRISPR Cas9 and CRISPy, a web-based target finding tool. *Biotechnol Bioeng* 111:1604–1616
17. Lund AM, Kildegaard HF, Petersen MBK, Rank J, Hansen BG, Andersen MR, Mortensen UH (2014) A versatile system for USER cloning-based assembly of expression vectors for mammalian cell engineering. *PLoS One* 9: e96693
18. Grav LM, la Cour Karottki KJ, Lee JS, Kildegaard HF (2017) Application of CRISPR/Cas9 genome editing to improve recombinant protein production in CHO Cells. *Methods Mol Biol* 1603:101–118
19. Haeussler M, Schönig K, Eckert H, Eschstruth A, Mianné J, Renaud JB, Schneider-Maunoury S, Shkumatava A, Teboul L, Kent J, Joly JS, Concordet JP (2016) Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol* 17(1):148. <https://doi.org/10.1186/s13059-016-1012-2>



Rapid and Efficient Gene Deletion by CRISPR/Cas9

Signe Neldeborg, Lin Lin, Magnus Stougaard, and Yonglun Luo

Abstract

CRISPR/Cas9 is a powerful genetic engineering technology that enables the introduction of genomic changes such as deletions and insertions of specific bits of DNA in cells with high precision. Compared to other programmable DNA nuclease such as ZFNs and TALENs, the specific binding of the Cas9 nuclease is mediated by a small guide RNA (gRNA), which can easily be designed to target any locus in the genome. The ease of generating novel gRNA vectors and its high efficiency has rapidly made CRISPR-Cas9 the dominant tool in gene editing applications, including gene knockout, knockin, tagging, etc. Here we describe our method for rapid and efficient generation of gene knockout or deletion cells using CRISPR/Cas9 within the time span of one month. The design of gRNAs, plasmid cloning, transfection, cell culturing, positive clone selection, and screening can be obtained from this method.

Key words Gene editing, Genomic engineering, Transfection, Endonuclease, gRNA, Deletion, DNA

1 Introduction

The CRISPR-Cas9 system utilized in genome engineering originates from bacteria where it serves as an adaptive immune system responding to viral infections [1]. In 2012, CRISPR-Cas9 was first harnessed for genome editing [2] and has since then been applied rapidly to edit genes in almost all species including humans [3, 4]. Simple gene editing utilizing CRISPR-Cas9 is based on a sole DNA endonuclease, Cas9, that can be used for any genomic locus in any species. The DNA nuclease domains of the Cas9 protein is not active until it interacts with a small guide RNA (gRNA) and this Cas9-gRNA complex subsequently forms a complementary complex with the targeted locus which harbors a protospacer adjacent motif (PAM) sequence. Once the nuclease activity of Cas9 is activated, a double-strand DNA break (DSB) is introduced to the target site. Un-repaired DSBs are lethal to cells as this can lead to major chromosomal alterations. Therefore, cells have developed two major DSB repair pathways: non-homologous end joining (NHEJ) and homology-directed repairs. Other DNA

repair pathways have recently been described involved in repairing the DSBs introduced by CRISPR-Cas9, such as microhomology-mediated end joining and single-strand annealing. Repair of DSBs introduced by CRISPR-Cas9 is predominantly by the NHEJ pathway, which leads to small insertion or deletions (known as indels). Therefore, the majority of CRISPR-Cas9 applications are gene knockouts caused by the introduction of indels to the target gene. Another major CRISPR-Cas9 application is the deletion of a specific piece of DNA or a whole gene, which is achieved through the co-delivery of Cas9 and pairs of gRNAs targeting the desired loci. We have previously showed that several factors can affect the efficiency of CRISPR-Cas9 gene editing, such as gRNA design, chromatin accessibility, and secondary energy of gRNA guide sequences [5]. Since the majority of indels introduced by a single gRNA is under 10 bp, several methods have been developed during the last few years for indel detection [summarized by Luo, *see* Chapter 2 of this book]. However, these methods are laborious and we have found that it's more straightforward to generate knockout/deletion cells through co-delivery of Cas9 and pairs of gRNAs [6]. One major reason for this strategy is the ease of genotyping. To ensure that Cas9 and the gRNA encoding plasmids are simultaneously delivered into the same cell, we have previously developed an efficient CRISPR gRNA expression array [6]. A detailed protocol for the CRISPR gRNA expression array has been published in Bio-protocols [7]. In this chapter, we will focus on describing the detailed experimental procedure for generating gene knockout by CRISPR, which comprises the design of gRNAs, rapid production and assembly of gRNA expression plasmids, delivery of gRNA and *Cas9* plasmids into human cells, selection of transfection positive cells, design of PCR screening primers, PCR screening of deletion positive clones. Using this method, the generation of gene knockout or deletion cells can be achieved within one month.

2 Materials

2.1 Primer Design

We typically use CLC Main Workbench software (QIAGEN) or other primer design software such as primer 3 (<http://bioinfo.ut.ee/primer3-0.4.0/>). When choosing primer targeting regions, regions containing repeated sequences are excluded using the RepeatMasker web tool (<http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker>). Primer oligos with a GC content of 40–60% and an annealing temperature of 57–63 °C are selected.

2.2 Preparation of PCR Product for Sanger Sequencing

1. Primers (stock 100 µM).
2. Thermal cycler (any brand).
3. 10 mM dNTP.

4. Platinum™ Pfx DNA Polymerase or other equivalent polymerase (*see Note 1*).
5. Agarose gel electrophoresis equipment.
6. Agarose.
7. Ethidium Bromide.
8. UV transilluminator.
9. DNeasy Blood & Tissue Kit or equivalent kit for DNA purification.
10. NucleoSpin Gel and PCR clean up kit or equivalent kit.
11. Mix2Seq kit or equivalent kit for Sanger sequencing.

2.3 *In Silico* gRNA Design

In this book serial, Mac Gruell has summarized different web tools for designing CRISPR gRNAs. One major driving force for the rapid adaptation of the CRISPR-Cas9 engineering tool by the scientific community is owing to all these open web tools, which are used by numerous scientists from all over the world to design their gRNA guide sequences as well as analyzing data. We have typically used the web tool CRISPOR (*see Note 2*).

1. Computer with internet access.
2. Web link to CRISPOR: <http://crispor.tefor.net>.

2.4 *Synthesis of CRISPR gRNA Oligos*

All CRISPR gRNA oligonucleotides were ordered from Sigma Aldrich, desalt purification, 100 μM in TE buffer.

2.5 *gRNA Oligo Annealing*

1. Heating block or thermal cycler.
2. Pipette tips: 10p, 20p.
3. Ice or cool rack for PCR tubes.
4. 200 μL PCR tubes.
5. 1.5 mL Eppendorf tubes.
6. Single-stranded gRNA oligos (100 μM, each).
7. 10× NEBuffer 2.
8. Distilled H₂O.

2.6 *Golden-Gate Assembly of CRISPR gRNA Plasmids*

1. Thermal cycler.
2. Pipette tips: P10, P20.
3. Annealed gRNA oligos from Subheading 2.5.
4. The All-in-One modular gRNA plasmids (available from Addgene, *see Table 1* for corresponding Addgene plasmid IDs).
5. T4 ligase (5 U/μL).
6. Fast digest BpiI(BbsI) restriction Enzyme.
7. 10× T4 ligase buffer.
8. Distilled H₂O.

Table 1
Addgene plasmid IDs

Plasmid name	Addgene ID	Description	Cloning site for inserting the guide sequences	Cloning site for second Golden-Gate assembly
pMA-SpCas9-g1	80784	For clone gRNA expression block, modular 1	BbsI	BsaI
pMA-SpCas9-g2	80785	For clone gRNA expression block, modular 2	BbsI	BsaI
pMA-SpCas9-g3	80786	For clone gRNA expression block, modular 3	BbsI	BsaI
pMA-SpCas9-g4	80787	For clone gRNA expression block, modular 4	BbsI	BsaI
pMA-SpCas9-g5	80788	For clone gRNA expression block, modular 5	BbsI	BsaI
pMA-SpCas9-g6	80789	For clone gRNA expression block, modular 6	BbsI	BsaI
pMA-SpCas9-g7	80790	For clone gRNA expression block, modular 7	BbsI	BsaI
pMA-SpCas9-g8	80791	For clone gRNA expression block, modular 8	BbsI	BsaI
pMA-SpCas9-g9	80792	For clone gRNA expression block, modular 9	BbsI	BsaI
pMA-SpCas9-g10	80793	For clone gRNA expression block, modular 10	BbsI	BsaI
Cas9 plasmids				

2.7 List of Plasmids

The modular gRNA expression plasmids can be requested by Addgene or from Yonglun Luo's lab, and the Cas9 expression plasmids can be ordered from Addgene (Table 1).

2.8 Transformation of *E. coli* Cells/Bacteria Work in General/ Colony Screening/ Plasmid Purification

1. Shaking incubator (37 °C).
2. Thermal Incubator for bacterial (37 °C).
3. Ice.
4. 200 µL PCR tubes.

4. 6-, 24-, and 96-well cell culturing plates.
5. P10 culture dishes.
6. Pipette filter tips: p10, p100, and p1000.
7. 10 mL Serological pipettes.
8. 1 × PBS.
9. Cell culturing medium: DMEM (high glucose), 10% FBS, 1% P/S, 1% Glutamine.

**2.10 Transfection/
Puromycin Selection**

1. Shaking incubator (37 °C).
2. Thermal incubator.
3. Sterile flow bench.
4. 24-well cell culturing plates.
5. Pipette filter tips: p10, p100, and p1000.
6. Opti-MEM™.
7. Xtreme Gene 9.
8. Cell culturing medium: DMEM, 10% FBS, 1%P/S, 1% Glutamine.
9. Plasmids.
10. 10 mg/mL Puromycin.

**2.11 PCR Screening
for Deletion Positive
Cell Clones**

1. 10 mM dNTP.
2. Platinum™ Pfx DNA Polymerase or other equivalent polymerase.
3. Agarose gel electrophoresis equipments.
4. Cell lysis buffer CS: 50 mM KCl, 1.5 mM MgCl₂, 10 mM Tris-Cl, pH 8.5, 0.5% NP40, 0.5% Tween 20, 400 µg/mL Proteinase K.

3 Methods

Carry out all methods at room temperature unless otherwise specified. Working with E. coli and tissue culture should be conducted in classified laboratories.

**3.1 Validate the
Targeted Regions in
Your Cells of Interest
by Sanger Sequencing
(see Note 3)**

1. Define the intended target regions based on Repeatmasker analysis and the aim of your experiment.
2. Once you have selected the targeted regions (typically which exon(s), 500–700 bp), design PCR primers to amplify this region for Sanger sequencing.
3. Purify genomic DNA from desired host cells which will be targeted in later CRISPR experiments (*see Note 4*).

4. Measure DNA concentration using a spectrophotometer.
5. Set up the Pfx-based PCR reaction on ice: 1 μL genomic DNA (25 $\text{ng}/\mu\text{L}$), 0.6 μL target specific forward primer (10 μM), 0.6 μL target specific reverse primer (10 μM), 0.6 μL dNTP (10 mM), 0.3 μL MgSO_4 (50 mM), 0.24 μL Platinum *Pfx* DNA polymerase, 3 μL 10 \times *Pfx* Amplification buffer, 3 μL 10 \times PCR_X enhancer solution, and 6.06 μL H_2O in a 200 μL PCR tube (total 15 μL) (*see Note 5*).
6. A master mix (w/o PCR primers) can be prepared when screening multiple target sites.
7. Place the PCR tubes in a thermal cycler. The following PCR program works for most reactions: 94 $^\circ\text{C}$ for 2 min; 35 cycles at 94 $^\circ\text{C}$ for 20 s, 60 $^\circ\text{C}$ for 30 s, 68 $^\circ\text{C}$ for 45 s; and followed by 68 $^\circ\text{C}$ for 7 min. Save the PCR product at 4 $^\circ\text{C}$ or proceed directly.
8. Purify PCR products using the NucleoSpin Gel and PCR clean up kit according to manufacturer's protocol.
9. Measure DNA concentration using a spectrophotometer.
10. Diluted PCR product to 5 $\text{ng}/\mu\text{L}$.
11. Combine 15 μL diluted PCR product with 2 μL target specific forward primer.
12. Sequence the PCR product using a qualified sequencing platforms.
13. Once the sequences of the targeted cell lines are validated by Sanger sequencing, they are ready to be used for gRNA design (Fig. 1).

3.2 Design CRISPR gRNAs

It is very straightforward to design CRISPR gRNA using the CRISPOR web tool (<http://crispor.tefor.net>). A detailed instruction on designing gRNAs has been provided in this book serial as well. In this section, we just highlight a few important points to increase the success of designing the best possible gRNA. We typically pick gRNAs based on the following criteria:

Off-target, the most similar off-target site should have no less than 3 mismatches to the target site.

Restriction sites, as we generate the CRISPR gRNA vector using Golden-Gate assembly, gRNAs should not contain restriction enzyme recognition sites for BsmBI, BsaI, or BbsI.

Poly T, we exclude target sites containing poly T (over 4), as this will be used as the termination signal for the U6 promoter.

Once the guide sequences are selected, add adapter nucleotides to the gRNA sense and antisense sequences:

The genomic sequence containing the PAM sequence is always the sense strand in the gRNA.

If target sequence starts with a G add adapter nucleotides:

3.5 Transformation of *E. coli* Cells (see Note 9)

1. Prepare an ice box.
2. Place one 1.5 EP tube on ice.
3. Heat LB medium to 37 °C.
4. Add 5 µL of ligation product from Subheading 3.3, 35 µL of H₂O and 10 µL of 5× KCM buffer to the EP tube.
5. Pipette up and down to mix.
6. Incubate on ice for at least 2 min.
7. Add 50 µL of competent *E. coli* cells to the reaction mixture, gently pipette or swirl the tube to mix (!! Keep the reaction on ice all the time).
8. Incubate on ice for 20 min.
9. Remove the tube from ice and incubate at room temperature for 10 min.
10. Add 400 µL of preheated LB medium.
11. Incubate in a shaking incubator for 1 h at 37 °C, with 200–250 rpm.
12. Plate 50–100 µL of transformation mixture onto ampicillin LB-agar plate (50 µg/mL).
13. Seal the plate with parafilm and incubate, bottom up, at 37 °C overnight.

3.6 Colony PCR Screening

This rapid colony screening method was developed for PCR-based screening of transformed *E. coli* clones containing the target plasmid.

1. For each gRNA colony plate, prepare two 8-well 200 µL PCR strips (culture strip and lysate strip).
2. Add 50 µL LB with ampicillin 50 µg/mL to each well in one PCR strip (culture strip).
3. Add 30 µL ddH₂O to each well in the other PCR strip (Lysate strip).
4. Use a 100 µL pipette tip to gently pick one colony.
5. First dip into one well in the culture strip.
6. Then dip into the corresponding well of the lysate strip, leave the pipette tip in the well.
7. Repeat **steps 4–6**, and pick 3–5 colonies from each gRNA colony plate.
8. The culture strip is incubated at 37 °C for 3–5 h and stored at 4 °C.
9. The lysate strip is lysed on a thermal cycler at 98 °C for 10 min.
10. Prepare an ice box.

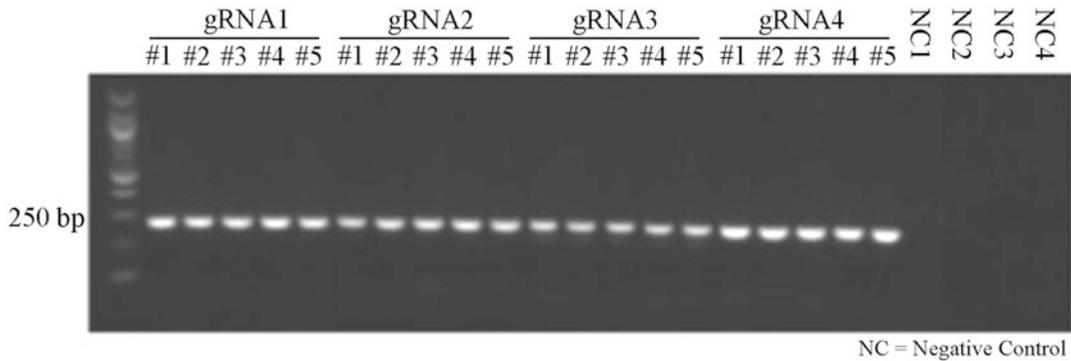


Fig. 2 PCR screening of bacterial clones transformed with gRNAs. Agarose gel 1.5%. Bands visualized with ethidium bromide. M = 1000 bp marker. Run at 80 V in 1 × TAE buffer. PCR product size a 250 bp. Negative controls are water control for each gRNA

11. Set up the following PCR screening reaction on ice: 1 μ L *E. coli* lysate, 0.6 μ L U6 promoter forward primer (5 μ M), 0.6 μ L gRNA antisense oligo (5 μ M), 0.3 μ L dNTP (10 mM), 0.08 μ L DreamTaq polymerase, 1.5 μ L 10× DreamTaq buffer, 10.92 μ L ddH₂O. *A master mix can be prepared when screening multiple colonies.*
12. Place reaction in a thermal cycler.
13. Perform PCR reaction using the following PCR program: 1 hold at 95 ° C for 3 min; 35 cycles at 95 ° C for 20 s, 58 ° C for 20 s, and 72 ° C for 30 s; 1 hold at 72 ° C for 7 min.
14. Save the PCR product at 4 ° C.
15. Check the PCR product in an 1.5% agarose gel. The PCR product should be approximately 250 bp (Fig. 2).

3.7 Plasmid Purification

1. Prepare a 14 mL Falcon tube or similar which can be used for bacterial culture.
2. Add 5 mL LB + ampicillin (50 μ g/mL).
3. Add 10 μ L of PCR-positive *E. coli* culture from the culture strip, step up two preps for each of the gRNA clones.
4. Incubate in a shaking incubator at 37 ° C overnight, 200–250 rpm.
5. Purify plasmid containing *E. coli* culture using a plasmid prep kit (*see Note 10*).
6. Validate the plasmid by Sanger sequencing with the U6-F primer.

3.8 Transfection of Targeted Cells

The transfection protocol used here is conducted using the XtremeGene9 transfection reaction. Other cell transfection reagents should be used according to manufacturer's instruction.

1. Seed an appropriate number of the targeted cells into a 24 well plate 24 h before transfection (*see Note 11*).
2. Prepare a transfection mix by combining 150 ng of plentiCRISPR-v2 plasmid and 50 ng of each gRNA containing plasmid with 25 μL Opti-MEM.
3. Add 0.75 μL Xtreme-Gene9 to the transfection mix, mix well by tapping.
4. Incubate at room temperature for 30 min.
5. Add the transfection mix to the cells in a dropwise manner.
6. Mix gently by swirling the 24-well plate from side to side.
7. Cells are incubated at 37 ° C and 5% CO₂ for 24 h.

3.9 Antibiotic Selection of Targeted Cells

1. 24 h post transfection, change the culture media to fresh media containing puromycin (1 $\mu\text{g}/\text{mL}$) (*see Note 12*).
2. Incubate cells at 37 ° C and 5% CO₂ in puromycin selection medium for 3–4 days, or until all cells in the non-transfected control well are dead.
3. Change the culture media back to normal culture media without puromycin, and culture the cells for 2–3 days more.

3.10 Establishment of Single Cell Clones by Manual Picking or FACS Sorting

1. Trypsinize the cells with Trypsin-EDTA.
2. Resuspend cells in 1 \times PBS with 10% FBS and kept on ice.
3. If using a stereomicroscope to pick single cells, we typically place the stereomicroscope in a sterile flow bench and use either a mouth pipette or P10 pipette to pick cells manually.
4. Or single cells are sorted directly into 96-well plate by flow cytometry if a FACS CORE facility is available.
5. Culturing single cell in a 96-well plate is somehow tricky for some cell types, we recommend using 50% condition medium to increase the success rate of establishing colonies from single cells (*see Note 13*).
6. Culture cells at 37 ° C and 5% CO₂, with medium change every 3 days.
7. Depending on the cell type, colonies normally evolve after approximately 2 weeks.

3.11 Colony PCR Screening

1. For each cell colony to be screened prepare cell lysis buffer CS.
2. Culture media is removed from cells and cells are washed in 1 \times PBS.
3. Prepare one 96-well tissue culture plate, add 60 μL complete cell medium to each well.
4. Prepare a 96-well PCR plate.

5. Mark on the lid of the corresponding clones that are ready for genotyping.
6. Wash the corresponding wells with PBS twice.
7. Add 30 μL Trypsin-EDTA, incubate at 37 ° C for 3–4 min.
8. Add 90 μL complete cell medium with 5 ng/mL.
9. Use a 100 μL tip, set to 60 μL , pipette up and down a few times (*see Note 14*).
10. Transfer 60 μL of cells to the 96-well tissue culture plate.
11. Transfer the remaining cells to the same position of the 96-well PCR plate.
12. Repeat **steps 4–9** until all clones are passaged.
13. For the 96-well plate with cells, place back to the incubator with medium change every 2–3 days.
14. For the PCR plate, proceed to the genotyping.
15. Spin down at 500 $\times g$, 5 min.
16. Pool the supernatant and tap on a piece of filter paper but not too hard.
17. Quick spin the plate, typically there are about 10 μL medium remained (*see Note 15*).
18. Add 50 μL cell lysis buffer CS per well to the cells.
19. Lysis cells at 65 ° C for 30, 95 ° C for 10.
20. Add 50 μL ddH₂O to the lysate per well.
21. Use 1 μL for PCR screening.
22. A PCR is set up on ice by combining 1 μL cell lysate, 0.6 μL target specific forward primer (10 μM), 0.6 μL target specific reverse primer (10 μM), 0.6 μL dNTP (10 mM), 0.3 μL MgSO₄ (50 mM), 0.24 μL Platinum *Pfx* DNA polymerase, 3 μL 10 \times *Pfx* Amplification buffer, 3 μL 10 \times PCR_X enhancer solution and 6.06 μL H₂O in a 200 μL PCR tube. A master mix can be prepared when screening multiple colonies.
23. The tube is placed in a thermal cycler and run at 94 ° C for 2 min, 35 cycles at 94 ° C for 20 s, 60 ° C* for 30 s, 68 ° C for 45 s followed by 68 ° C for 7 min.
24. Check PCR product with 1.5% agarose gel electrophoresis (Fig. 3).

4 Notes

1. The platinum *Pfx* polymerase will be discontinued from Thermal Scientific. Since we performed PCR from crude cell lysate, it's crucial to use a polymerase that is compatible with PCR on

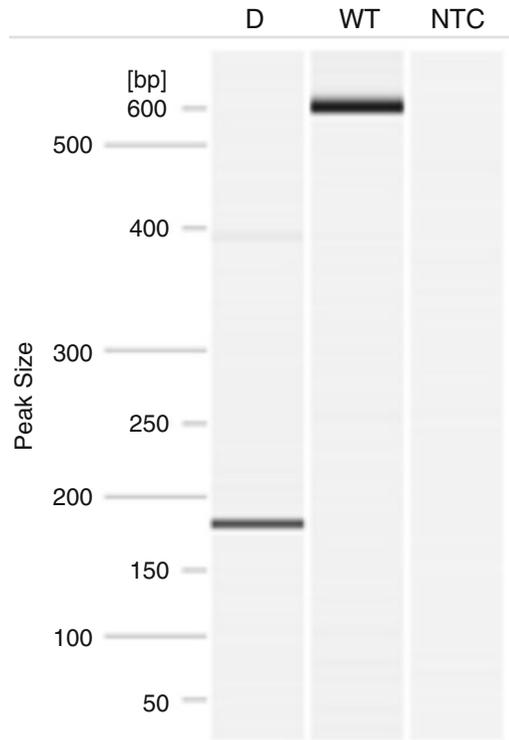


Fig. 3 PCR screening of transfected HEK293T cells. *D* transfected HEK293T cell DNA with 441 bp deletion in BRCA2 exon 20. Deleted PCR product is seen at 183 bp. *WT* non-transfected HEK293T cell DNA. Non-deleted PCR product is seen. *NTC* non-template/no DNA control, H₂O. Capillary electrophoresis

non-purified DNA samples. It's recommend that authors should optimize the PCR on cell lysate before genotyping hundreds of clones.

2. The CRISPOR web tool provides several scores for choosing the best gRNA guide sequences. No current in silico methods can predict the on-target efficiency well and, therefore, we normally do not select gRNA based on that parameter. However, we have previously discovered that the secondary energy of the guide sequences affect gRNA activity [5]. We, therefore, typically analyze the delta G of the 20 nt guide sequences using <http://unafold.rna.albany.edu/?q=mfold/rna-folding-form>.
3. Although the reference sequences for all genes are available from, e.g., NCBI (<https://www.ncbi.nlm.nih.gov/>), Ensembl (<https://www.ensembl.org/index.html>), or genome browser (<https://genome.ucsc.edu/>), there are variations present, e.g., SNPs, in the targeted gene that differs between cell lines. Therefore, we typically sequence the intended target

regions of the cell line of interest before conducting the massive editing experiments.

4. DNA purification can be done using your preferred methods. We typically use the DNeasy Blood & Tissue Kit.
5. Using Pfx, we have found that using 2× reaction buffer and 2× enhancer buffer work better. If using other DNA polymerases, please select proof-reading ones.
6. Some old-version thermal cyclers do not have the delta C option. Alternatively, you can use a heating block to anneal the gRNA oligos: first denature the reaction mixture (in one 1.5 EP tube) at 95 °C for 5 min, then remove the heating block and let it cooling down naturally.
7. A master mix without annealed insert can be prepared without the annealed gRNA oligo if multiple gRNAs are prepared at the same time.
8. This Golden-Gate Assembly reaction program is optimized for the restriction enzymes (RE) and T4 ligase used in this protocol. If using enzymes from different suppliers, the reaction temperature for the RE and T4 enzymes should be modified accordingly. For example: N cycles at XX1 °C (XX1, temperature for the RE) for T1 minutes (T1, incubation time for RE reaction) and XX2°C (XX2, temperature for the T4 ligase) for T2 min (T2, incubation time for T4 reaction). The cycles number N can be extended if the positive colony rate is low. However, we have found that with 10 cycles the positive colony rate is over 95% for all gRNAs tested in our lab.
9. This optimized protocol is used for chemically competent *E. coli* cells.
10. We used the PureLink™ HQ Mini Plasmid DNA Purification Kit as the plasmid can be used for subsequent cell transfection directly. Alternatively, the plasmid can be purified with midi- or maxi-prep kits.
11. 50×10^4 HEK293T and SK-OV-3 cells are seeded out at 50–60% confluency.
12. The concentration of puromycin should be optimized for each cell type. Normally, 1 µg/mL of puromycin works for most cell types.
13. Preparation of 50% condition medium: prepare a flask of wild-type cells (60–70%), change to fresh medium, and culture overnight. Save the overnight culture medium, filter the medium with a 0.4 µM filter. Mix the filtered medium with fresh medium at a ratio of 1:1. Save the condition medium at 4 °C and use within one week.

14. Typically pipetting 5 times are enough. You can check under the microscope if the cells are de-touched from the well. If not, pipette a few more times.
15. Some wells might have a little bit more medium left. For those with too much medium, you can remove it by pipette.

Acknowledgments

The CRISPR projects are partially supported by grants from the Lundbeck Foundation (R219–2016-1375, R173–2014-1105), the Danish Research Council for Independent Research (DFR–1337–00128), the Sapere Aude Young Research Talent Prize (DFR-1335–00763A), the Innovation Fund Denmark (BrainStem), and Aarhus University Strategic Grant (AU-i-CRISPR). Y.L. is also supported by the Guangdong Provincial Key Laboratory of Genome Read and Write (No. 2017B030301011).

References

1. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315:1709–1712
2. 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* 337:816–821
3. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-guided human genome engineering via Cas9. *Science* 339:823–826
4. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823
5. Jensen KT, Floe L, Petersen TS, Huang J, Xu F, Bolund L, Luo Y, Lin L (2017) Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. *FEBS Lett* 591(13):1892–1901
6. Vad-Nielsen J, Lin L, Bolund L, Nielsen AL, Luo Y (2016) Golden Gate Assembly of CRISPR gRNA expression array for simultaneously targeting multiple genes. *Cell Mol Life Sci* 73:4315–4325
7. Vad-Nielsen J, Lin L, Jensen KT, Nielsen AL, Luo Y (2016) A Golden Gate-based protocol for assembly of multiplexed gRNA expression arrays for CRISPR/Cas9. *Bio-protocol* 6:23. <https://doi.org/10.21769/BioProtoc.2059>



Genome Editing in Mice

Lisbeth Ahm Hansen and Ernst-Martin Füchtbauer

Abstract

Programmable nucleases like CRISPR/Cas9 enable to edit the mouse genome directly in the zygote. Several methods have been successfully used for this. Here we describe injection into one of the pronuclei of the zygote and electroporation of zygotes. Alternative methods will be mentioned.

Key words CRISPR/Cas9, Mouse, Genome editing, Pronucleus injection, Electroporation

1 Introduction

Programmable nucleases, in particular the CRISPR/Cas9 system, have created unforeseen possibilities to modify animal genomes by gene editing in the zygote or early embryos. The principals of genome editing and design of experiments are well described in other chapters of this book, so only methods directly related to the application in mice will be described here. In addition to the general advices for the design of gRNA, the genetic background of the mouse strain should be considered, which is, e.g., possible with “Breaking Cas” (<http://bioinfo.gp.cnb.csic.es/tools/breakingcas>) [1]. This chapter will therefore focus on application of CRISPR/Cas9 techniques to modify the genome in the mouse germ line. Genome editing in zygotes differs from genome editing in somatic cells in several ways. Mechanisms for repair of DSBs might differ, transfer of the CRISPR/Cas9 components differs, and, perhaps most importantly, the resulting founder mice can be bred, which opens the possibility to remove off target mutations if there should be any.

It would require an entire book to describe all techniques needed for the generation, isolation, culture, injection, or electroporation, and final transfer of early murine embryos. Readers who are not familiar with these techniques should consult one of the numerous publications describing the generation of genetically modified mice, e.g., [2, 3].

2 Materials

2.1 CRISPR/Cas9

1. RNase free RNA duplex buffer: 100 mM potassium acetate, 30 mM HEPES, pH 7.5.
2. Cas9 dilution buffer: 300 mM NaCl, 10 mM Tris-HCl pH 7.4, 1 mM DTT, 0.1 mM EDTA, 500 µg/mL BSA, 50% Glycerol (*see Note 1*).
3. Custom designed crRNA (Alt-R™ crRNA) and generic tracrRNA dissolved to a concentration of 100 µM in RNA duplex buffer (*see Notes 2 and 3*).
4. Donor DNA (*see Note 4*).

2.2 Handling of Mouse Embryos

1. Glass capillaries made from Pasteur pipettes.
2. Soft silicon tubes (outer diameter 8 mm, inner diameter 5 mm) for mouth pipetting.
3. Mouth piece from glass micropipettes (Paul Marienfeld GmbH, Lauda-Königshofen). Alternatively pipette tips for 200 µL pipettes without filter can be used.
4. Pipette tip for 1000 µL pipettes without filter as adaptor for the Pasteur pipette.
5. Disposable Petri dishes 6 and 10 cm.
6. Sterile dissection tools.
7. Dissection microscope set up for transmitted and reflected light microscopy.
8. CO₂ incubator set at 37 °C, 5% CO₂, and 100% humidity.
9. M2 Medium.
10. Hyaluronidase, bovine testes Type IV-S, embryo tested, dissolved to 10 mg/mL in M2.
11. KSOM (Merck).
12. OptiMEM (Gibco).

2.3 Pronucleus Injection

1. For detailed description of micro-injection setup refer to reference [2, 3].
2. Injection needles. If no needle puller is available, injection needles can be ordered from several companies (e.g., Biomedical-Instruments, Zöllnitz, Germany, or Eppendorf).
3. PN injection buffer: 10 mM Tris-HCl pH 7.4, 0.1 mM EDTA.

2.4 Electroporation

1. BioRad Gene Pulser XCell electroporator (*see Note 5*).
2. GSS-500 cell fusion chamber normally used for electro-fusion of 2 cell embryos (BLS Biological Laboratory Equipment Maintenance and Service Ltd. Hungary) (*see Note 6*).

3 Methods

3.1 Generation of RNA Duplex and Ribo-protein Complexes

1. Mix 5 μL crRNA (100 μM), 5 μL tracrRNA (100 μM), 35 μL RNA duplex buffer and denature for 5 min at 95 $^{\circ}\text{C}$.
2. Put the tube on the bench or turn off the heat block to cool the RNA slowly to room temperature (*see Note 7*).
3. For Cas9-RNA complex formation mix 15 μL RNA duplex (from **step 2**), 12.5 μL RNA duplex buffer, 2.5 μL Cas9 protein (61 μM) (*see Note 8*), incubate for 15 min at RT.
4. For 100 μL injection buffer mix 30 μL RNP complex (from **step 3**), 70 $-x$ μL pronucleus injection buffer and, if needed x μL donor DNA (*see Note 9*).

3.2 Pronucleus Injection

1. Superovulate and mate female mice according to standard procedures.
2. Isolate zygotes from the ampulla according to standard procedures.
3. Add hyaluronidase solution (10 mg/mL) to the drop with the zygotes in a ratio of approx 1:1 and allow the zygotes to incubate for around 5 min, until the cumulus cells will release the zygotes and fall down to the bottom of the dish.
4. Collect all the zygotes and wash them in 3 drops of M2 to wash away cumulus cells and debris.
5. In a Petri dish, make 2 drops of KSOM of ~ 40 μL and 1 drop of ~ 60 μL , covered with sterile mineral oil. Equilibrate overnight at 37 $^{\circ}\text{C}$ with 5% CO_2 .
6. Wash the zygotes in the 2 smaller drops to rinse off M2 and place them in the larger last drop for incubation.
7. Transfer the zygotes into the injection chamber and perform PN injection following standard procedures.
8. In a Petri dish, make 2 drops of KSOM of ~ 40 μL and 1 drop of ~ 60 μL , covered with sterile mineral oil. Equilibrate overnight at 37 $^{\circ}\text{C}$ with 5% CO_2 .
9. Wash the zygotes in the 2 smaller drops to rinse off M2 and place them in the larger last drop for incubation.
10. Incubate zygotes over night to the 2-cell stage and transfer them to a pseudopregnant mouse following standard procedure.

3.3 Electroporation of Zygotes

1. Superovulate and mate female mice according to standard procedures.

2. Mix 5 μL crRNA (100 μM) with 5 μL tracrRNA (100 μM), incubate at 95 $^{\circ}\text{C}$ for 5 min, cool to room temperature slowly, spin briefly (*see Note 10*).
3. Prepare the electroporation buffer, mix 13.6 – x μL OptiMEM, 1.3 μL RNA duplex (50 μM , from **step 2**) in RNA duplex buffer, 1.1 μL Cas9 protein (61 μM), incubate at RT for 15 min, add x μL donor DNA if needed (*see Note 11*).
4. Isolate zygotes from the ampulla according to standard procedures.
5. Add hyaluronidase solution (10 mg/mL) to the drop with the zygotes in a ratio of approx 1:1 and allow the zygotes to incubate for around 5 min, until the cumulus cells will release the zygotes and fall down to the bottom of the dish.
6. Collect all the zygotes and wash them in 3 drops of M2 to wash away cumulus cells and debris (*see Note 12*).
7. In a Petri dish, make a small drop (e.g. 3 μL) of electroporation buffer (from **step 3**) and add the zygotes to the drop with as little OptiMEM as possible.
8. Place 5 μL electroporation buffer between and over the electrodes of the fusion chamber (*see Fig. 1*) and add the zygotes with minimal additional volume.
9. Electroporate with 2 square pulses of 30 V and 3 ms duration with an interval of 100 ms.
10. Carefully collect all the electroporated zygotes from the chamber and wash them in 3 drops of M2 to rinse off the electroporation mix.
11. In a Petri dish, make 2 drops of KSOM of ~40 μL and 1 drop of ~60 μL , covered with sterile mineral oil. Equilibrate overnight at 37 $^{\circ}\text{C}$ with 5% CO_2 .
12. Wash the zygotes in the 2 smaller drops to rinse off M2 and place them in the larger last drop for incubation.
13. Incubate them over night to the 2-cell stage and transfer them to a pseudopregnant mice following standard procedure.

4 Notes

1. Use this to store diluted Cas9 protein when the original concentration is too high. According to IDT, Cas9 can also be diluted in, i.e., RNA duplex buffer and stored at +4 $^{\circ}\text{C}$ or –80 $^{\circ}\text{C}$, but not at –20 $^{\circ}\text{C}$.
2. Typically crRNA is ordered as 2 nmol which should be dissolved in 20 μL ; typically tracrRNA is ordered as 5 nmol, which should be dissolved in 50 μL .

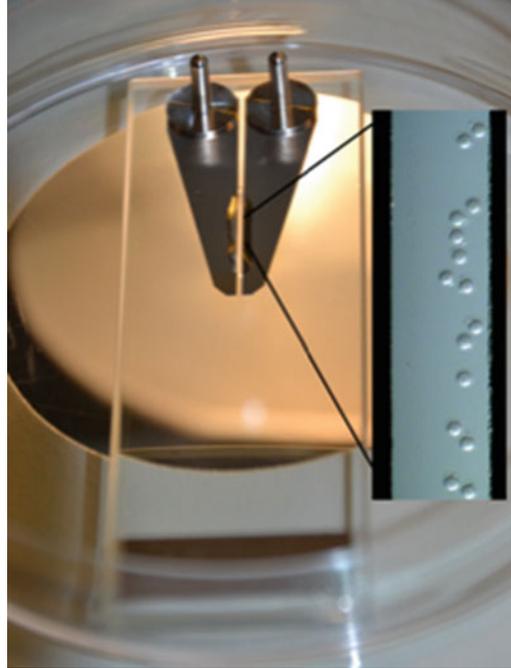


Fig. 1 GSS-500 cell fusion chamber in a 10 cm petri dish on the stage of a dissection microscope. The gap between the two electrodes is overlaid with 5 μ L electroporation buffer. Note that the fusion chamber is not connected with the power supply. The connector can be used to fix the chamber to the rim of the Petri dish (not shown). The insert to the right shows zygotes in the fusion chamber ready to be electroporated. The gap between the electrodes is 500 μ m wide

3. Alternatively a ds DNA oligo coding for the sequence specific part of a gRNA may be cloned into an all-in-one vector, e.g., pX330, *see* Chapter 14.
4. If the experiment includes the insertion of a donor DNA either ssDNA oligos with micro homologies of appr. 50 bases on each side of the CAS9 cut site can be used or, for longer insertions long ssDNA molecules also with micro-homologies can be used [4].
5. Any pulse generator which can deliver double pulses of 30 V, 3 ms duration with a pause of 100 ms between the pulses can be used.
6. BioRad electroporation cuvettes with an electrode distance of 0.1 mm can also be used [5, 6], but retrieval of embryos after electroporation is more difficult using narrow cuvettes.
7. The RNA duplex has now a concentration of 11.1 μ M.

8. The RNP complex has now a concentration of approx. 5 μM . Some protocols use an up to fivefold excess of RNA duplex over Cas9 protein.
9. ssDNA oligos or long ssDNA should have a final concentration of 5–20 ng/ μL , super coiled plasmid DNA of 50 ng/ μL .
10. The RNA duplex has now a concentration of 50 μM .
11. ssDNA oligos should have a final concentration of 0.5–1 $\mu\text{g}/\mu\text{L}$, long ssDNA of 5–20 ng/ μL .
12. It is not necessary to remove or weaken the zona pellucida with acid tyrode or pronase [5].

References

1. Oliveros JC, Franch M, Tabas-Madrid D, San-León D, Montoliu L, Cubas P, Pazos F (2016) Breaking-Cas-interactive design of guide RNAs for CRISPR-Cas experiments for ENSEMBL genomes. *Nucleic Acids Res* 44: W267–W271
2. Behringer R, Gertsenstein M, Nagy KV, Nagy A (2014) *Manipulating the mouse embryo: a laboratory manual*, 4th edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
3. Pease S, Saunders TL (2011) *Advanced protocols for animal transgenesis*, an ISTT manual. Springer, Berlin, Heidelberg
4. Quadros RM, Miura H, Harms DW, Akatsuka H, Sato T, Aida T, Redder R, Richardson GP, Inagaki Y, Sakai D, Buckley SM, Seshacharyulu P, Batra SK, Behlke MA, Zeiner SA, Jacobi AM, Izu Y, Thoreson WB, Urness LD, Mansour SL, Ohtsuka M, Gurumurthy CB (2017) Easi-CRISPR: a robust method for one-step generation of mice carrying conditional and insertion alleles using long ssDNA donors and CRISPR ribonucleoproteins. *Genome Biol* 18:92
5. Tröder SE, Ebert LK, Butt L, Assenmacher S, Schermer B, Zevnik B (2018) An optimized electroporation approach for efficient CRISPR/Cas9 genome editing in murine zygotes. *bioRxiv*. 281402. <https://doi.org/10.1101/281402>
6. Wefers B, Bashir S, Rossius J, Wurst W, Kühn R (2017) Gene editing in mouse zygotes using the CRISPR/Cas9 system. *Methods* 121–122:55–67



CRISPR/Cas9-Mediated Gene Tagging: A Step-by-Step Protocol

Xi Xiang, Conghui Li, Xi Chen, Hongwei Dou, Yong Li, Xiuqing Zhang, and Yonglun Luo

Abstract

CRISPR/Cas9 provides a simple and powerful tool for modifying almost any DNA of interest. One promising application of the CRISPR/Cas9 system is for tagging genes with a fluorescence marker or tag peptides. For such a purpose, FLAG, HIS, and HA tags or fluorescence proteins (EGFP, BFP, RFP, etc.) have been broadly used to tag endogenous genes of interest. The advantages of generating fluorescence tagging proteins are to provide easy tracing of the subcellular locations, real-time monitoring the expression and dynamics of the protein in different conditions, which cannot be achieved using traditional immunostaining or biochemistry assays. However, the generation of such a gene-tagged cell line could be technically challenging. In this chapter, we demonstrate the generation of tagging the porcine *GAPDH* (*pGAPDH*) gene *GFP* by CRISPR/Cas9-based homology-directed repair.

Key words CRISPR/Cas9, Gene tagging, Homologous recombination, *GAPDH*, Reporter gene, *copGFP*

1 Introduction

The homologous recombination (HR) pathway is employed by cells to repair harmful lesions introduced to the genome which is known as DNA double-strand breaks (DSB). Harnessing the HR pathway enables scientists to introduce exogenous DNA fragments into a specific genomic locus, which is known as gene targeting. To date, gene targeting strategy is widely used for gene knockout [1–3], knocking in of functional elements (such as *loxP*, reporter genes) to a safe genomic harbor [4], modifying or replacing endogenous DNA with other species-derived genetic material [5], fuse tag proteins or fluorescence proteins to the target gene, etc.

Before the discovery of programmable DNA nucleases such as ZFNs, TALENs, and CRISPR/Cas, the traditional gene targeting method employs only a donor vector. A donor vector is a DNA molecule that contains exogenous DNA fragments (or gene of

interest) flanked by two homologous arms, which are usually derived from context of target loci. As the success of gene targeting highly depends on cell cycles and the HR pathway, the efficiency of traditional targeting method is known to be disappointingly low in mammalian cells [6].

It has been noticed that introduction of DSBs to the targeted locus can increase HR efficiency. To date, there are three generation of programmable DNA nucleases that are developed for such purpose, including ZFNs, TALENs, and CRISPR/Cas9. The CRISPR/Cas9 is a simple and efficient tool to introduce DSBs into specified loci of genome. In mammalian cells, the DSBs are predominantly repaired by the non-homologous end joining (NHEJ), while only a very small proportion of DSBs are replaced by the homology-directed repair (HDR) pathways [7]. Although the efficiency of HDR in cells are affected by various factors, introduction of DSBs by CRISPR/Cas9 will greatly enhance gene targeting by HDR in specified loci as compared to traditional gene targeting [8].

One application of gene targeting is gene tagging, which allows us to real-time monitor the endogenous activities of target proteins. Unlike immunostaining and biochemistry assays, proteins fused with a fluorescence protein enable live-cell imaging. The gene-tagged cells provide a valuable tool to detect the subcellular location and expression spatiotemporally of the targeted gene [9]. In addition to gene tagging with fluorescence proteins, genes tagged with small molecular tags, like His-tag, Flag-tag, HA-tag, etc., can be easily purified and detected utilizing standardized antibodies, solving the problem of unspecific or unavailable antibodies.

In this chapter, we take pig *pGAPDH* gene as an example, giving a detailed protocol of fusing a *GFP* with *pGAPDH* by CRISPR/Cas9 in PK15 cell line.

2 Materials

2.1 Web Tools

1. Download *pGAPDH* gene sequence from the National Center of Biotechnology Information (NCBI) which the website is freely accessible at <https://www.ncbi.nlm.nih.gov>.
2. Many web tools are available for gRNA design, such as MIT CRISPR (<http://crispr.mit.edu>), CRISPor (<http://crispor.tefor.net>), CHOPCHOP (<http://chopchop.cbu.uib.no>), E-CRISP (<http://www.e-crisp.org/E-CRISP/>).

2.2 Molecular Cloning and Cell Engineering

1. Primers: dissolved in ddH₂O to 10 μM.
2. High fidelity DNA polymerase (e.g., PrimeSTAR premix of Takara).
3. General Taq polymerase.

4. ExTaq premix (Takara).
5. Genome DNA was isolated from PK15 cell line.
6. pCDH-EF1-*copGFP* (Addgene ID #73030).
7. Zero Blunt™ PCR Cloning Kit.
8. Standard gel electrophoresis reagents.
9. DNA markers.
10. LentiCRISPRv2 plasmid (Addgene ID 98290).
11. T4 DNA ligase.
12. 10× T4 ligase buffer.
13. Esp3i (BsmBI).
14. pcDNA3.1(+) plasmid.
15. Ampicillin.
16. Competent cells.
17. LB medium.
18. DNA extraction kit.
19. Endo-free plasmid purification kit.
20. PK15 (ATCC® CCL-33) cell line.
21. 24-well plate, 96-well plate, 10 cm dish for cell culture.
22. Complete cell culture medium: Dulbecco's Modified Eagle's Medium (DMEM), 10% FCS, 100 U/mL penicillin, 100 mg/mL streptomycin in, 1% Glutamine.
23. 0.5% Trypsin-EDTA.
24. PBS.
25. Opti-MEM.
26. Lipofectamin2000.
27. pEGFP-N1-FLAG plasmid.
28. T7 Endonuclease I (NEB).
29. 10× NEBuffer2.
30. 50 mg/mL G418.
31. Stereomicroscope.

3 Methods

Carry out all the experiment steps at room temperature unless otherwise specified.

3.1 Gene Fusion Strategy at pGAPDH Locus

The seamless fusion of *pGAPDH* gene with *copGFP* is conducted by CRISPR/Cas9-mediated HDR (Fig. 1). To achieve that, a Cas9, a *pGAPDH* targeting gRNA, and a donor vector are needed. The

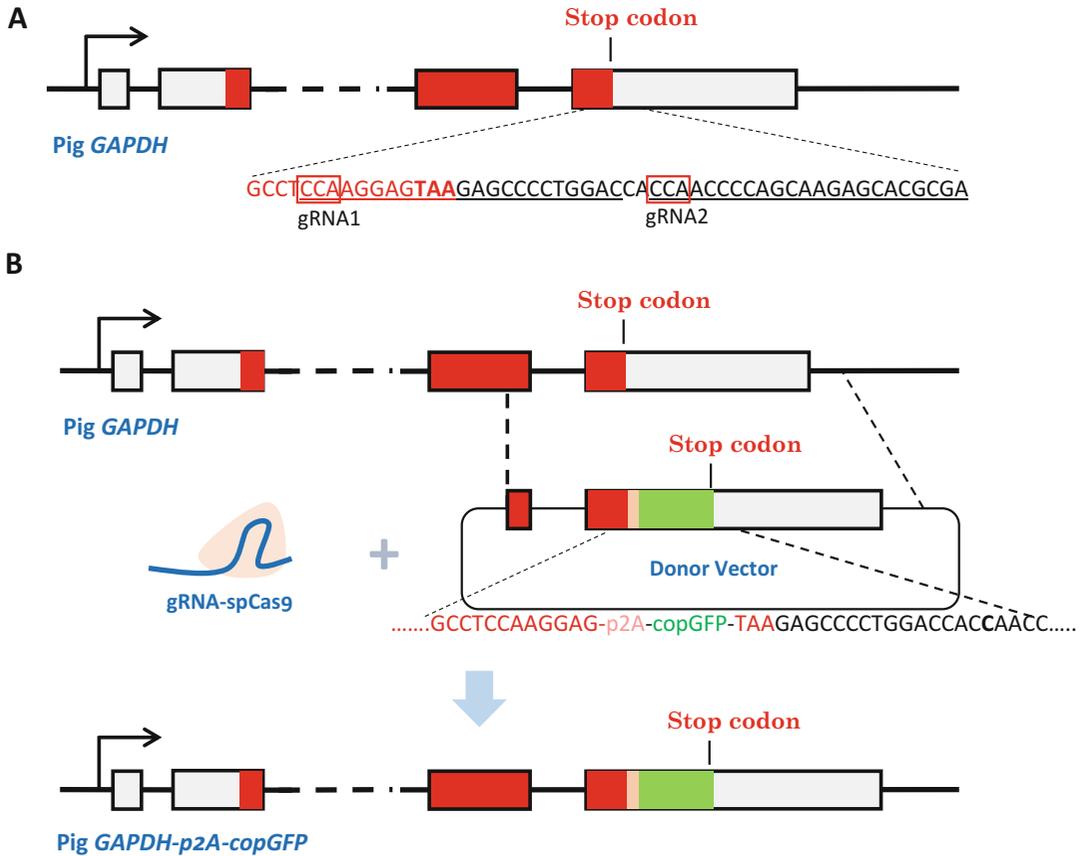


Fig. 1 Gene tagging strategy at *pGAPDH* locus. (a) Two gRNAs nearby the stop codon site of *pGAPDH*. (b) Schematic diagram of CRISPR/cas9-mediated HDR and the structure of donor vector. Red letters are part of the last exon sequence of *GAPDH*; Letters inside the red frames are PAM of gRNAs. The pale green blocks are the noncoding parts of *GAPDH* and solid red blocks are coding sequence of *GAPDH*. Pink blocks represent p2A element

donor plasmid contains the *copGFP* coding sequence which is flanked by left and right homologous arms (HAs). The total HA length for accuracy integration of large fragment is normally 1–2 kb [10, 11]. However, we also find that HAs of 0.5–1 kb can yield fairly high efficiency of homologous recombination [9]. A short and self-cleavage peptide P2A [12] was used to link the target protein to fluorescence reporter protein (*see Note 1*).

3.2 Design *pGAPDH* gRNA and Screening Primers

1. Download target gene sequence from NCBI: <https://www.ncbi.nlm.nih.gov>, select database “gene,” search gene *GAPDH* and *sus scrofa* as the species (Gene ID: 396823).
2. Click “Full Report” and choose the “Gene table” option. *pGAPDH* has 12 exons. Unfold the exon table of mRNA [NM_001206359.1](https://www.ncbi.nlm.nih.gov/nuccore/NM_001206359.1). Then there is a table listing the region and sequence hyperlink of the 12 exons. Click the coding sequence of the last exon (64129952 to 64129883), copy the

FASTA sequence and paste to a word document and named as “*pGAPDH* targeting sequence.”

- Expand and select the region from 64,127,952 to 64,131,883 (include 2000 bp on either side of the stop codon) in current web page. Copy and paste the sequence to the word “*pGAPDH* targeting sequence.” Search the last exon sequence in this region and mark the exon by changing the sequence color to be red. The last three base pairs TAA is the stop codon of *pGAPDH*.
- Design gRNA via the Zhang’s lab web tool: <http://crispr.mit.edu>. Select “target genome” option to be “pig (SusScr3).” Select 100–150 bp sequence include the stop codon TAA in the middle as target sequence and paste into the web tool frame “target sequence” (*see Note 2*), then click “submit.” The system needs approximately 10–30 min to calculate and predict the off-target efficiency of candidate gRNAs. Select gRNAs with high score (high target specificity and low off-target efficiency). It’s better to select gRNAs within or close to the *pGAPDH* stop codon (*see Note 3*).

In this protocol, we selected gRNAs:

pGAPDH-gRNA1: GTCCAGGGGCTCTTACTCCTTGG

pGAPDH-gRNA2: TCGCGTGCTCTTGCTGGGGTGG

(sequence underlined are PAMs)

- Design primers for measuring gRNA activity. In this experiment, we used T7 endonuclease I assay (T7E1) to test the gRNAs activities. To achieve this, the forward and reverse were designed at 200–400 bp proximity flanking the Cas9 target site to allow clear visualization of cleaved bands by gel electrophoresis [13].

The screen primers for *pGAPDH* are (*see Note 4*):

pGAPDH -iden-F: TGTAGGTTTGGGTTGGAACAG

pGAPDH -iden-R: CAGCAAGCTCTAGTTCTCCTT

The amplicon size is 821 bp.

3.3 gRNA Assembly

- Synthesize the gRNA oligos as below:

pGAPDH-gR1-SS: CACCGTCCAGGGGCTCTTACTCCT

pGAPDH-gR1-AS: AAACAGGAGTAAGAGCCCCCTGGAC

pGAPDH-gR2-SS: CACCGTCGCGTGCTCTTGCTGGGGT

pGAPDH-gR2-AS: AAACACCCCAGCAAGAGCACGCGAC

For the backbone plasmid of LentiCRISPRv2, which is designed to be digested by BsmBI, the linker sequence model is: gRNA-SS: CACCN(20); gRNA-AS: AAACN(20). If the first base pair of gRNA was NOT G, then an additional G should be added at the first base pair of 5’end (*see Note 5*). Then the linker sequence model will be: gRNA-SS: CACCGN(20); gRNA-AS: AAACN(20)C.

2. Dissolve the gRNA oligos in ddH₂O to 10 μM.
3. Mix 10 μL gRNA-SS and 10 μL gRNA-AS in a PCR tube.
4. Anneal the oligos in a thermocycler at 95 °C for 5 min, ramp down to 25 °C at 5 °C/min.
5. Prepare the following Golden Gate Assembly reaction in a PCR tube:

Reagents	1×
LentiCRISPRv2	100 ng
Annealed oligos	1 μL
Esp3i (BsmBI)	1 μL
T4 Ligase	1 μL
10 × T4 ligase buffer	2 μL
H ₂ O	To be 20 μL

6. Place the PCR tube in a thermal cycler and apply the following program:

Condition	Cycle number
37 °C for 5 min and 22 °C for 10 min	10
37 °C for 30 min	1
75 °C for 15 min	1
Save at 4 °C	

7. Transformation. Transform 1 μL of the ligation product from **step 6** into 50 μL competent cells (mix gently), incubate on ice for 30 min, and heat at 42 °C for 90 s (or follow the protocols of your own lab). Add 150 μL LB medium into the mixture and incubate in a shaking incubator at 37 °C for 1 h. Plate 50 μL on LB ampR plate and distribute evenly (store the remaining at 4 °C until colonies are observed). Incubate the plates at 37 °C up-side-down for 16 h (±2 h).
8. In the next day, pick up two colonies for each gRNA and conduct PCR screening using the following steps: (1) Pick up one colony using a pipette with 10 μL tip and suspend the colony in 6 μL ampicillin LB media by gently pipetting up and down; (2) take 1 μL of the bacterial liquid from **step 1** for PCR screening.

9. Set up colony PCR screening reaction.

Reagents	1×
Extaq Premix	10 μL
U6-F ^a	0.4 μL
gRNA-AS	0.4 μL
Bacterial	1 μL
H ₂ O	8.2 μL
Total	20 μL

^aU6-F primer sequence used in the experiment: CGAAAGGTCGGGCAGGAAGA

10. Perform PCR with the following program.
Positive amplicon length is 287 bp.

Condition	Cycle number
98 °C 3 min	1
98 °C 10 s 58 °C 10 s 72 °C 30 s	30
72 °C 5 min	1
4 °C hold	

11. Expend the positive colony of each gRNA in 20 mL ampR LB medium at 37 °C in a shaking incubator and overnight. Extract the plasmids using an endo-free plasmid purification kit.

3.4 T7EI Assay of gRNAs

Measure the gRNAs activities by T7EI assay [8] or sanger sequencing (*see Note 5*).

1. Recover the PK15 cells in a 24-well plate (*see Note 6*). Seed 0.05×10^6 cells in a 24-well plate 1 day before transfection.
2. The second day, change medium when the cell confluence is around 50–70%. Then prepare the transfection reagents as below:

Reagents A	1×
Lipofectamine2000	2 μL
Opti MEM	23 μL

Reagents B	1×
LentiCRISPRv2-gRNA	500 ng
Opti MEM	To 25 μL

3. Let the reagent A and B incubating at room temperature separately for 5 min and then add B into A gently in a dropwise manner (pipette up and down to mix). Incubate the A–B mixture at room temperature for 15 min. Add the A–B mixture into cell culture medium in a slowly dropwise manner and shake the cell plate gently.
4. Put the cells in 37 °C incubator with 5% CO₂, change medium 24 h later after transfection.
5. Harvest cells 48 h after transfection and extract genome DNA using general DNA extraction kit (or lysis the cells using lysis buffer contains proteinase K).
6. Prepare PCR reactions as below.

Reagents	1×
Extaq Premix	10 μL
<i>pGAPDH</i> -iden-F	0.4 μL
<i>pGAPDH</i> -iden-R	0.4 μL
Genome DNA	100 ng
H ₂ O	To 20 μL

PCR programme is:

Condition	Cycle number
98 °C 3 min	1
98 °C 10 s 58 °C 10 s 72 °C 30 s	32
72 °C 5 min	1
4 °C hold	1

The amplicon length is 821 bp.

7. T7E1 digestion. Mix 18 μL PCR product with 2 μL NEB buffer2 in a 200 μL PCR tube. Denature and anneal the PCR products at 95 °C for 5 min, ramp down to 25 °C at 5 °C/min. Then add 0.2 μL T7E1 in the tube and mix thoroughly. Keep the reaction in 37 °C for 1 h and then conduct the electrophoresis gel assay (*see Note 7*).

3.5 Generation of *pGAPDH* Donor Vector

For gene targeting by HDR, it's common to use a donor vector with left and right homologous arms flanking the gene of interest for insertion. It's known that generating the targeting vector could be laborious and time-consuming. In this protocol, we describe the

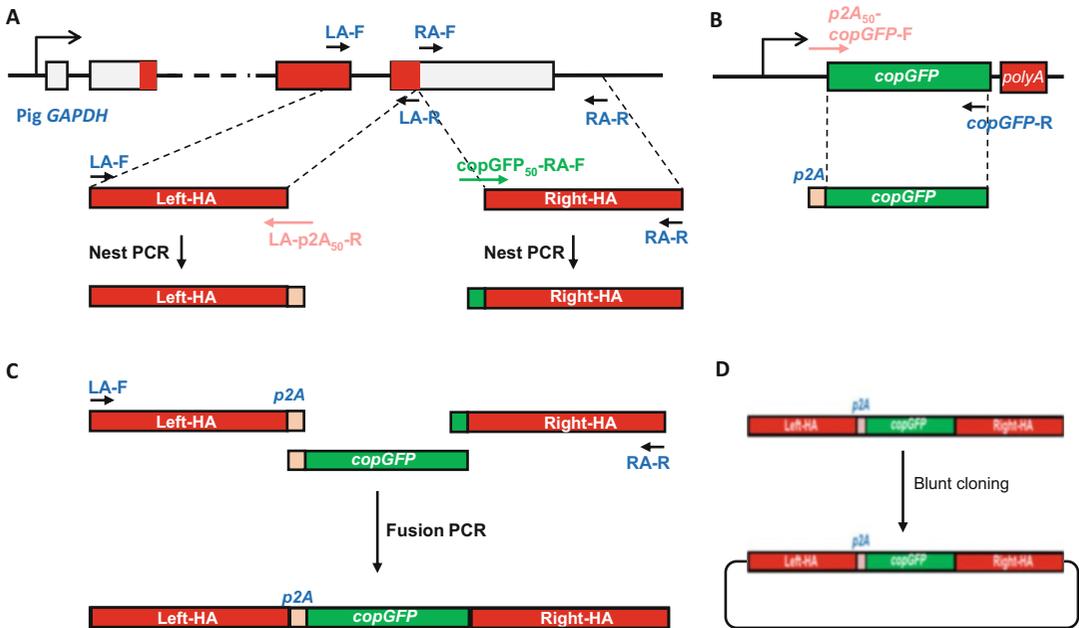


Fig. 2 Donor vector construction. (a) Cloning of left and right HAs by nested-PCR. After the second round PCR, 50 bp of 5'p2A-copGFP were fused to the 3'end of Left-HA and 50 bp of 3'copGFP were fused to the 5'end of Right-HA. (b) Cloning of copGFP. After the PCR, 50 bp of p2A were fused to 5'end of copGFP. (c) The three fragments fused to be one intact fragment by fusion PCR. (d) Cloning of the fused fragment into blunt cloning vector

rapid generation of donor vector by seamless fusion PCR (Fig. 2). For fusion PCR, 50 bp homologous sequence in the junction region of the two DNA fragments is required. Nest PCR was conducted to add 50 bp homologous adaptors to the junction sites of each HA (Fig. 2a). For *copGFP* cloning, the p2A sequence was fused to 5'end of *copGFP* by PCR using a long forward primer (Fig. 2b). Then the three fragments were fused together to be one intact fragment by one round PCR (Fig. 2c). Then the fusion product was gel purified and cloned into a blunt cloning vector (Fig. 2d).

1. HAs and *copGFP* cloning. Two rounds of PCR were conducted to clone the left and right HAs. And one round PCR for cloning of p2A-*copGFP*. Used high fidelity DNA polymerase during the cloning and fusion steps. Primers used in this part of experiment are listed in Table 1.

- (a) First round of PCR for LA and RA amplification:
Prepare PCR reactions as below.

Table 1
Primers for cloning of fusion fragments

Primer name	5'-3' sequence*
LA-F	ACTGAGCCAGAGCACGTTCT
LA-R	TTTCCCTTGGGCTAACCTACT
LA-p2A50-R	TCCATCCCTCCGGAGCCTCCGCCACCTGATCCGCCACCTCCGC TCCACCTCCTTGGAGGCCATGTGGA
RA-F	CTCATTTCCCTGGTAGGGTGCTGG
RA-R	TCTTCAAGCATCAGCCGAAGA
copGFP50-RA-F	ATTCTGCCGTGGACGGCACCCGCCGACCCGGCTCCACCGGAT CTCGCTAAGAGCCCCTGGACCACCAACC
p2A50-copGFP-F	GGTGGGAGCGGAGGTGGCGGATCAGGTGGCGGAGGCTCCGG AGGGATGGAGAGCGACGAGAGCGG
copGFP-R	TTAGCGAGATCCGGTGGAGC

*pink sequence represents p2A and the green sequence represents part of copGFP.

Reagents	1×
PrimeSTAR Premix	5 μL
LA-F (RA-F)	0.2 μL
LA-R (RA-R)	0.2 μL
PK15 genome DNA	100 ng
H ₂ O	To 10 μL

PCR programme:

Condition	Cycle number
98 °C 3 min	1
98 °C 10 s 55 °C 5 s 72 °C 90 s	30
72 °C 5 min	1
4 °C hold	1

(b) Second round of PCR for LA and RA amplification:
Prepare PCR reactions as below.

Reagents	1×
PrimeSTAR Premix	5 µL
LA-F (copGFP50-RA-F)	0.2 µL
LA-p2A50-R (RA-R)	0.2 µL
PCR product of first round	100 ng
H ₂ O	To 10 µL

PCR programme:

Condition	Cycle number
98 °C 3 min	1
98 °C 10 s 55 °C 5 s 72 °C 90 s	30
72 °C 5 min	1
4 °C hold	1

For LA, the target band length is 1097 bp and for RA is 1022 bp.

(c) *copGFP* PCR,
Prepare PCR reactions as below.

Reagents	1×
PrimeSTAR Premix	5 µL
p2A50- <i>copGFP</i> -F	0.2 µL
<i>copGFP</i> -R	0.2 µL
pCDH-EF1- <i>copGFP</i>	10 ng
H ₂ O	To 10 µL

PCR programme:

Condition	Cycle number
98 °C 3 min	1
98 °C 10 s 55 °C 5 s 72 °C 60 s	30
72 °C 5 min	1
4 °C hold	1

The copGFP fragment is 804 bp.

- (d) Three fragments seamless fusion PCR.
Prepare PCR reactions as below.

Reagents	1×
PrimeSTAR Premix	25 µL
LA-F	1 µL
RA-R	1 µL
PCR products of each fragment	1 + 1 + 1 µL
H ₂ O	20 µL

PCR programme:

Condition	Cycle number
98 °C 3 min	1
98 °C 10 s 55 °C 5 s 72 °C 60 s	30
72 °C 5 min	1
4 °C hold	1

The fused fragment is 2923 bp.

- Gel purification of the fused fragment and clone it to blunt cloning vector following the manufacturer's instruction.
- Prepare endo-free plasmid for cell transfection.

3.6 CRISPR Gene Tagging in Porcine Cells

1. Seed 0.05×10^6 PK15 cells in a 24-well plate and conduct transfection when cell density is between 50 and 70%. Transfect cells with Lipofectamine2000 as below:

Reagents	1 ×
Lipofectamine2000	2 μL
<i>pGAPDH</i> -gRNA1	200 ng
Donor vector	200 ng
Linearized pcDNA3.1(+) ^a	100 ng
Opti MEM	To 50 μL

^aWorks as a *Neo* gene provider for G418 screen and the neo plasmid ratio is usually 4:1 in our lab

Conduct the transfection procedure as Subheading 3.4, steps 2–4.

2. Keep transfected cells in the incubator for 48 h without disturbance.
3. 48 h post transfection, trypsinize cells.
4. Seed 1000 cells in a 10 cm petri dish (*see Note 8*), prepare 3–5 dishes would be sufficient for the screening (*see Note 9*).
5. Add G418 (800 μg/mL for PK15 cells) 24 h later and change medium every 2 days.
6. Generally, the cells begin to die 6–8 days after addition of G418 and colonies are ready for picking up after approximately 12–14 days of G418 screening.
7. Mark the GFP-positive colonies under a fluorescence microscope (*see Note 10*) and transfer the green colonies to 96-well plates by ring-cloning (*see Note 11*).
8. Genotype the cell clones by PCR and sanger sequencing.
9. Expand the positive colonies and freeze in liquid nitrogen.

4 Notes

1. After gene fusion, the stop codon of target gene should be deleted and replaced precisely by the coding sequence of p2A-copGFP. If there were frame-shift indels around the junction site, the target gene function would be damaged and the reporter would not work.
2. The input sequence length is required to be 23–250 bp in the website.

3. For selection of gRNAs outputted by the webtool, except concerning about the predicted score, it is better to choose the gRNA which contains the stop codon in the guide sequence. For these gRNAs will not cleave the donor vector itself and also will not disrupt the junction site again in the successful gene-tagged cells. If there were not suitable gRNAs which contain stop codon, it is better to choose gRNAs close and downstream of the stop codon for minimized disruption of target gene function. At the same time, it is required to mutate the selected gRNA PAM sequence on the right HA of donor vector to avoid cleavage of donor vector and re-cut of the successful tagged gene by this group of gRNAs.
4. It is better to design and synthesize 2 pairs of primers in case one of them does not work well.
5. The U6 promoter will recognize the first G and treat it as the start site of transcription.
6. Transfection efficiency will affect the gene editing efficiency of CRISPR/cas9 in different type of cells. For some hard-to-transfect cells, it is hard to test gRNAs activities via either T7E1 or sanger sequencing. In this situation, we recommend to change another easy-to-transfect cell type of the same species or utilizing the c-check system [9] to screen for the best gRNA.
7. For electrophoresis gel test of the T7E1 digestion product, load all the 20.2 μ L sample in one lane to make sure that even the weak bands after T7E1 digestion could be seen in the gel images. The total length of digested bands should be the same with the main band. In this experiment, the digested bands should be 554 bp + 267 bp = 821 bp for gRNA1 and 529 bp + 292 bp = 821 bp for gRNA 2.
8. After recovering, passage the cells at least 1 time before transfection.
9. The number of cells seeded depends on transfection efficiency of the cell type. For easy-to-transfect cell type, seeding 1000 cells in one 10 cm dish would be sufficient for colonies screening.
10. At the beginning it would be difficult to see the green fluorescence positive cells under microscope. It is better to check the cell colonies 8–12 days after G418 screening.
11. The GAPDH protein only expresses in cytoplasm, so you can see a round black area which refer to the position of cell nuclear in *copGFP*-knockin positive cells (Fig. 3).

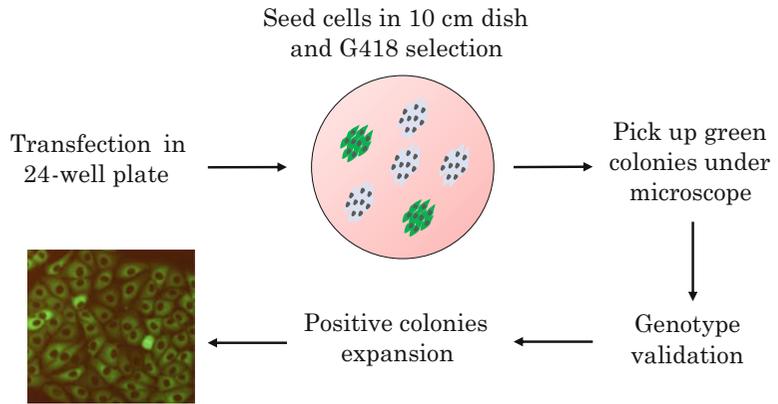


Fig. 3 Schematic illustration of generating gene-tagged cells by CRISPR

Acknowledgment

This work is partially supported by Guangdong Provincial Key Laboratory of Genome Read and Write (No. 2017B030301011).

References

1. Plump AS et al (1992) Severe hypercholesterolemia and atherosclerosis in apolipoprotein E-deficient mice created by homologous recombination in ES cells. *Cell* 71(2):343–353
2. Nandi AK et al (1988) Regulated expression of genes inserted at the human chromosomal beta-globin locus by homologous recombination. *Proc Natl Acad Sci U S A* 85 (11):3845–3849
3. De Lozanne A, Spudich JA (1987) Disruption of the Dictyostelium myosin heavy chain gene by homologous recombination. *Science* 236 (4805):1086–1091
4. Li ZW et al (1996) Generation of mice with a 200-kb amyloid precursor protein gene deletion by Cre recombinase-mediated site-specific recombination in embryonic stem cells. *Proc Natl Acad Sci U S A* 93(12):6158–6162
5. Brightbill HD et al (2010) Antibodies specific for a segment of human membrane IgE deplete IgE-producing B cells in humanized mice. *J Clin Invest* 120(6):2218–2229
6. Kuwayama H (2012) Enhancement of homologous recombination efficiency by homologous oligonucleotides. In: Gowder SJT (ed) *Cell interaction*. IntechOpen, London
7. Hsu PD, Lander ES, Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157(6):1262–1278
8. Cong L et al (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121):819–823
9. Zhou Y et al (2016) Enhanced genome editing in mammalian cells with a modified dual-fluorescent surrogate system. *Cell Mol Life Sci* 73(13):2543–2563
10. Dickinson DJ et al (2013) Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. *Nat Methods* 10(10):1028–1034
11. Yang H et al (2013) One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 154(6):1370–1379
12. Szymczak AL et al (2004) Correction of multi-gene deficiency in vivo using a single ‘self-cleaving’ 2A peptide-based retroviral vector. *Nat Biotechnol* 22(5):589–594
13. Ran FA et al (2013) Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8 (11):2281–2308



Chapter 17

Gene Editing in Primary Cells of Cattle and Pig

Petra Vochozkova, Kilian Simmet, Eva-Maria Jemiller, Annegret Wünsch, and Nikolai Klymiuk

Abstract

Gene Editing by CRISPR/Cas has revolutionized many aspects of biotechnology within a short period of time. This is also true for the genetic manipulation of livestock species, but their specific challenges such as the lack of stem cells, the limited proliferative capacity of primary cells, and the genetic diversity of the pig and cattle populations need consideration when CRISPR/Cas is applied. Here we present guidelines for CRISPRing primary cells in pig and cattle, with a specific focus on testing gRNA in vitro, on generating single cell clones, and on identifying modifications in single cell clones.

Key words CRISPR/Cas, Primary cells, gRNA, Single cell clones

1 Introduction

Although Gene Editing has emerged only recently, it is already evident that this technology resembles one of the most important innovations in biotechnology history. The full power of Gene Editing became obvious when the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas system was shown to efficiently modify mammalian genomes in vitro [1, 2]. The CRISPR technology is derived from a kind of a prokaryotic adaptive immune system that defends bacteria from phage infection. Basically, this system introduces exogenous phage DNA into a defense locus within the bacterial genome [3] and then activates these acquired DNA segments by transcription. These phage-specific so-called crRNA transcripts form a ribonucleoprotein complex with a linker RNA and an endonuclease [4] that eventually destroys re-infecting phage DNA by introducing a double strand break (DSB). The most relevant aspect in this system is that the Cas protein is unspecific *per se* and is directed to its target site solely by

Petra Vochozkova and Kilian Simmet contributed equally to this work.

Yonglun Luo (ed.), *CRISPR Gene Editing: Methods and Protocols*, Methods in Molecular Biology, vol. 1961, https://doi.org/10.1007/978-1-4939-9170-9_17, © Springer Science+Business Media, LLC, part of Springer Nature 2019

the crRNA. The translation of this simple principle to a Gene Editing tool is a charming approach, because theoretically any RNA can be used to guide the Cas enzyme to a desired target site. In biotechnology, simplified systems are used and there the RNA directing Cas is most often called guiding RNA (gRNA). Directing Cas endonucleases by a short RNA segment to a target site is very simple and broadly applicable compared to previous Gene Editing tools such as ZFN or TALEN. Since its introduction, the CRISPR/Cas system revolutionized the entire field of life science in a dramatic and sustained way, allowing breakthroughs that would have been illusive a short time ago [5–8].

For sure, the invention of CRISPR/Cas resembled also a change of paradigm for the genome modification of livestock species, an attempt that has been a toilsome undertaking a couple of years ago and limited its application to a handful of specialized laboratories worldwide. Nowadays, an uncounted number of groups are working on CRISPR/Cas-based Gene Editing in livestock, succeeding in modifying the genome of diverse species [9–12], but to our knowledge there is no sufficient guideline on how to use this technology in cells of these species. Based on our extended experience in generating genetically modified livestock models [13–19] as well as on our expertise in Gene Editing [20, 21] we feel competent to give such a guideline for Gene Editing procedures in primary cells of cattle and pig. Primary cells are of outstanding interest, as in cattle and pig (as well in most other species) no stem cells exist and their manipulation thus represents a fast-forward way to produce genetically modified embryos or animals by somatic cell nuclear transfer (SCNT).

We describe the entire process from designing gRNA to transfection of pig or cattle primary cells and give protocols for evaluating gRNA efficiency as well as for generating and analyzing single cell clones. We will, however, not give instructions for designing recombination vectors as those mostly need very specific considerations regarding the target site or the kind of modifications intended to be introduced. Basic guidelines for such might be found in other publications [22–24]. Numerous CRISPR/Cas systems have been identified in different prokaryotic species and were sufficiently adapted for Gene Editing, but we will exclusively refer to the by far most widely used system that has been adapted from *Streptococcus pyogenes*.

2 Materials

2.1 Designing Appropriate gRNAs

1. Design gRNA for candidate regions with bioinformatic tools.
 - (a) chopchop (<http://chopchop.cbu.uib.no/>, [25]).
 - (b) e-crisp (<http://www.e-crisp.org/E-CRISP/>, [26]).

(c) CRISPR Design (<http://crispr.mit.edu/>, [27]).

2. Synthesize the gRNA by commercial suppliers.

2.2 Isolation and Cultivation of Primary Cells

1. Cell culture medium: DMEM (high glucose with GlutaMAX and pyruvate), supplemented with 0.1 mM β -mercaptoethanol (0.1 mM), 1% non-essential amino acids, 1% HEPES, 10% fetal calf serum (FCS).
2. Sterile culture dishes coated with Collagen R 0.2% (Serva, Heidelberg, Germany) diluted 1:10 in cell culture grade H₂O (Biowest, Nuaille, France), before use.
3. 37 °C incubator with humidified atmosphere of 5% CO₂.
4. Washing buffer: phosphate buffered saline (PBS, 0.2 g KCl, 8.0 g NaCl, 0.2 g KH₂PO₄, 2.14 g Na₂HPO₄ + 7 × H₂O in 1 L cell culture grade H₂O).
5. Passaging/dissociation reagent: 0.4% Difco Trypsin +0.032% EDTA in PBS.
6. Freezing medium: FCS containing 10% DMSO.
7. Equipment for freezing: 0.5 mL or 1.5 mL Cryo Vials, Cool-Cell freezing container, liquid nitrogen storage container.

2.3 Nucleofection and Recovery

1. Amaxa Nucleofector[®] transfection device.
2. Basic Fibroblast Nucleofector[™] kit including nucleofection solution, cuvettes and Pasteur pipets (all Amaxa components by Lonza, Basel, Switzerland).
3. Culture dishes; 100 and 35 mm.

2.4 Evaluation of Mixed Clonal Population

1. Genomic (g) DNA isolation kit (DNA Easy, Invitrogen by ThermoFisherScientific).
2. End-point PCR kit (HotStart Taq and Long-range Taq, both Qiagen, Hilden, Germany; Herculase II, Agilent Technologies, Santa Clara, CA, USA).
3. Commercially synthesized primers.
4. Agarose.
5. DNA extraction/PCR clean-up kit (NucleoSpin Gel and PCR cleanup, Macherey-Nagel, Düren, Germany; Double-Pure Combi Kit, Bio&Sell).

2.5 Generation of Single Cell Clones

1. 96-well half area Costar 3696 culture plates (Corning Costar, New York, NY, USA) or standard 96-well culture plates.
2. Cellavista automated cell culture microscope (Synentec, Elms-horn, Germany).
3. Enriched cell culture medium (composed as described in Sub-heading 2.2 with the exception of 15% FCS).

2.6 Evaluation of Single Cell Clones

1. Proteinase K, Tris (Tris(Hydroxymethyl)-Amino-Methan), DTT (di-Thiothreitol), SDS (Sodium dodecyl-Sulfate), EDTA (Ethylen-di-Amino-tetra-Acetic acid), NaCl, iPrOH (iso-Propanol), EtOH (Ethanol).
2. Resolving buffer (T-buffer (1×): 10 mM Tris/HCl, pH 8.0.
3. PK-buffer (10×): 200 mM Tris/HCl pH 8.0, 1 M NaCl, 40 mM EDTA pH 8.0.
4. End-point PCR kit, agarose and DNA extraction kit (as described in Subheading 2.4).
5. Components for Sanger sequencing: BigDye Terminator and 5× BigDye sequencing buffer (both ThermoFisherScientific).
6. FastStart Essential DNA Green Master (Roche, Basel, Switzerland).
7. LightCycler96 (Roche).
8. DNA standards: DNA isolated by DNA Easy kit (Invitrogen by ThermoFisherScientific) and diluted in elution buffer.
9. Raw data processing by the LightCycler96 software (version 1.1.0.1320, Roche) and Microsoft Excel 2016 (Microsoft, Redmond, WA, USA).

3 Methods

3.1 Designing Appropriate gRNAs

gRNA binding sites are theoretically easy to design, as the only strict requirement of the *Streptococcus pyogenes* CRISPR system is the placement of the gRNA sequence upstream of a so-called NGG protospacer adjacent motif (PAM) that is directly recognized by Cas9 via a protein-DNA interaction [28]. In practice, however, numerous aspects influence the efficacy of a given gRNA to introduce a DSB. A manifold of in silico prediction tools has been developed for defining promising gRNA sequence and many of them are available on open-access platforms. We follow a procedure where we use three different programs (see Subheading 2.1) for evaluating candidate regions independently. Although the predictions diverge significantly most of the times, there is also a certain degree of consistency between the respective suggestions (example given in Fig. 1). Promising gRNAs can be chosen on the basis of these rankings, but decision-making requires also some intuition and experimental evaluation (see Notes 1 and 2).

According to the decision for one or more candidate gRNA, those can be synthesized in a way that fits to the preferred transfection method. Due to historical reasons, we mainly rely on the CRISPR/Cas components described in [2], using distinct plasmid vectors encoding the Cas9 protein and transcribing a gRNA-transcript, as well as an optional recombination vector. Other,

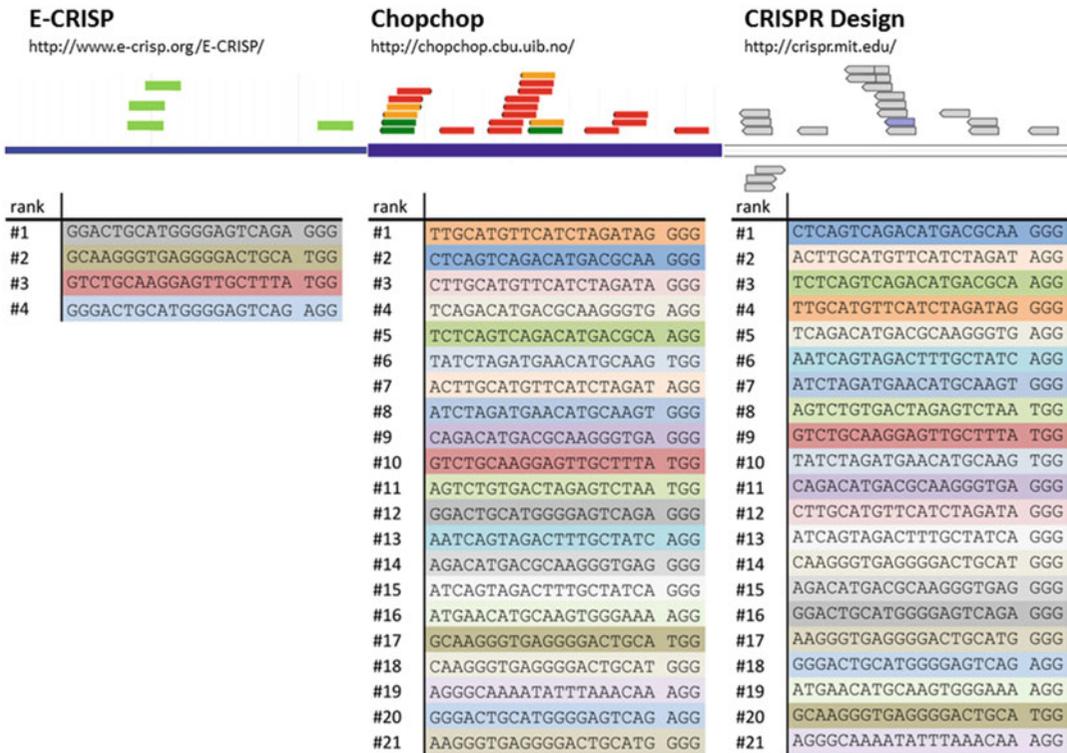


Fig. 1 Designing gRNA sequences. We search for gRNA with E-Crisp, Chopchop, and CRISPR Design and compare the suggested candidates. As a representative example we show the data for a 245 nt target region from the intron4 of the porcine OTC gene. The candidates are shown in their proposed rankings with their approximate positions indicated above on the basis of screen shots from the respective programs. Some gRNA are identified by all programs (marked by identical color) but at different rank. The algorithmic basis for the identification and ranking differs for the three methods and is explained in more detail at the respective web pages. Of note, also the genomic data used for estimating the off-target effects differ for the respective programs, being Sscrofa10.2 for E-Crisp and Chopchop and susScr3 for Crispr Design

more sophisticated vectors might work as well, including approaches that use a single plasmid encoding all Gene Editing components (*see Note 3*).

3.2 Isolation and Cultivation of Primary Cells

Isolation and cultivation of primary cells has been described in numerous publications. We follow the procedure described previously [29], with some minor adaptations:

1. Isolate bovine and porcine primary cells:
 - (a) Isolate bovine cells from ear fibroblasts or kidney tissue.
 - (b) Isolate porcine cells are exclusively from kidneys.
2. After mincing and collagenase-digest of a 0.5 cm × 0.5 cm × 0.5 cm piece of tissue, seed $\frac{1}{8}$ – $\frac{1}{4}$ of the resulting cell pellet onto a 100 mm dish (*see Note 4*).

3. After 24 h of incubation, wash off the non-adherent cells and let the adherent cells grow to 80–90% confluence. Exchange the cell culture medium every second day.
4. For passaging, wash the cells twice with PBS. Add the dissociation reagent and incubate at 37 °C for 3–5 min under visual control (*see Note 5*).
5. Stop trypsin reaction by adding cell culture medium and centrifuge at $150 \times g$ for 5 min.
6. For further cultivation, resuspend the cells in cell culture medium and seed at appropriate density (approx. $1.5\text{--}2.0 \times 10^4$ cells/cm² $\approx 1 \times 10^6$ cells/100 mm dish).
7. For cryo-preservation, resuspend portions of 1×10^6 cells in precooled cryo-conservation medium (8 °C) and transfer to a cryo-vial, which is placed in a CoolCell device and then stored at –80 °C. After cooling for at least 3 h, vials can be transferred to liquid nitrogen.

3.3 Nucleofection and Recovery

The efficacy of introducing exogenous DNA into primary cells of pig or cattle heavily depends on the transfection method. We have gained best experience with the Amaxa Nucleofection technology for this purpose [29], basically following the manufacturer's protocol but with some important aspects to consider:

1. In both, pig and bovine, seed 1×10^6 somatic cells (passage 2–3) on a 100 mm culture dish and culture for 24 h to obtain exponentially growing cells (log phase) containing a high proportion of mitotic cells.
2. Before nucleofection, the following preparatory work should be considered:
 - (a) Pre-warm the nucleofection solution at RT.
 - (b) Mix the DNA to be transfected in a 1.5 mL reaction tube and pre-warm the DNA mix at RT (*see Note 6*).
 - (c) Equilibrate 35 mm dishes containing 2 mL of culture medium in a cell culture incubator.
3. Dissociate cells by trypsin/EDTA. Resuspend the cells in cell culture medium and determine the number of cells per mL.
4. Centrifuge 0.5×10^6 cells and resuspend cell pellet in 100 μ L Basic Nucleofector™ solution at RT.
5. Transfer the cell suspension to the tube containing the prepared DNA mix, and then mix the DNA-cell suspension by pipetting up and down carefully.
6. Transfer the entire suspension into a cuvette and place in the Amaxa/Lonza Nucleofector® transfection device.

7. Use the predefined nucleofection program U12 for both, pig and cattle cells (*see* **Note 7**).
8. Wash out the nucleofected cell suspension of the cuvette with the Pasteur pipet from the Nucleofection kit using 500 μL of the equilibrated culture medium and transfer the cell suspension into 35 mm cell culture plate (*see* **Note 8**).
9. 24 h after nucleofection, exchange the medium to remove dead cells and culture cells for another 24 h. Use the transfected cells for isolation of gDNA from a mixed clonal population (*see* Subheading 3.4) or for the generation of single cell clones (*see* Subheading 3.5).

3.4 Evaluation of a Mixed Clonal Population

For reliably predicting how effective gRNAs are in reality, we evaluate each gRNA for its ability to introduce non-homologous end joining (NHEJ) mutations. This is done by analyzing mutation events in a mixed clonal populations after nucleofection.

1. After recovery from nucleofection, detach and centrifuge the cells as described in Subheading 3.2 and wash the pellet once in 1 mL PBS.
2. Use the cell pellet for DNA isolation or store at $-80\text{ }^{\circ}\text{C}$ for later analysis.
3. Isolate the gDNA with conventional isolation kits according to the manufacturer's protocols (recommended amount is 1×10^6 cells).
4. Amplify the target site by end-point PCR from gDNA (*see* **Note 9**) and analyze PCR product by electrophoresis on a 0.7% agarose gel.
5. Excise the band of the PCR product from the gel and extract DNA with a conventional kit according to the manufacturer's protocols.
6. Use the purified PCR amplicon for evaluation of target sites by distinct methods such as T7 endonuclease approach, next-generation sequencing, TIDE or Sanger sequencing of cloned PCR products (*see* **Notes 10** and **11**).

3.5 Generating Single Cell Clones

Single cell clones are generated for identification of a desired modification in an individual clone prior to further experimental procedures such as somatic cell nuclear transfer.

1. Passage the recovered nucleofected cells from Subheading 3.3 and resuspend the cells in enriched cell culture medium (15% FCS) (*see* **Note 12**):
 - (a) For culture of cells without antibiotic selection, 100–200 cells per 10 mL.
 - (b) For positive selection, $1\text{--}2 \times 10^4$ cells per 10 mL.

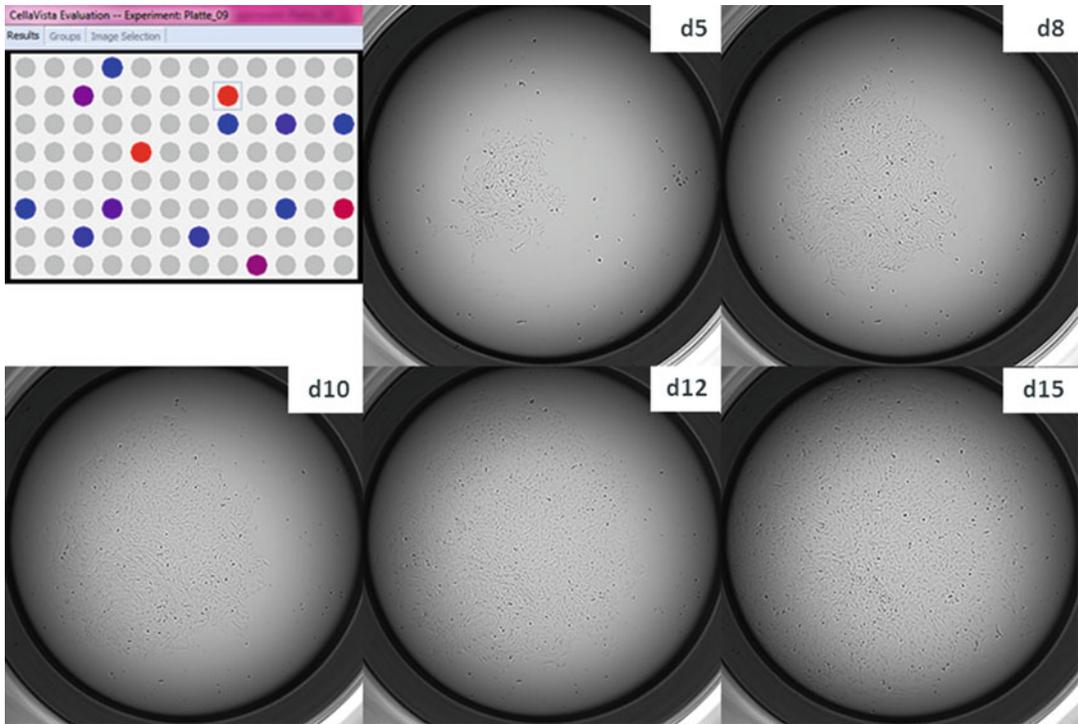


Fig. 2 Screening for single cell clones. The Cellavista automated microscope identifies proliferating cell clones at a 96-well scale (upper left image). The system also facilitates the documentation of proliferation for specific cell clones as shown here for well B8 in the exemplary plate

2. Transfer 100 μL of the cell suspension to each well of a 96-well half area plates (*see Note 13*).
3. Let the cells grow under the conditions described in Subheading 3.2. Exchange culture medium every 48–72 h.
4. Starting at day 5–6 after seeding, scan the plates every second day for proliferating single cell clones with the Cellavista cell culture microscope (Fig. 2) (*see Note 14*).
5. At 50–90% confluence, harvest the culture of each single cell clone with 20 μL of trypsin/EDTA solution and split each culture into 2 wells of a standard 96-well plate (*see Note 15*).
6. After reaching 80–90% confluence, trypsinize both wells of the proliferating clones again:
 - (a) Stop the trypsin reaction in one well by adding cryo-preservation medium and transfer the cells into a cryovial (*see Note 16*).
 - (b) In the other well, stop the trypsin reaction with cell culture medium. Centrifuge the cells and freeze the cell pellet at $-80\text{ }^{\circ}\text{C}$ for DNA isolation (*see Subheading 3.6*).

3.6 Evaluation of Single Cell Clones: gDNA Isolation

Single cell clones derived from primary cells have a limited proliferation capacity compared to stem cells. Thus, the maximum outcome of material from clonal selection is in the range of 10^4 – 10^5 cells overall. Consequently, gDNA isolation needs to be efficient and the quality needs to meet the requirements for identifying mono-allelic modification by qPCR. We have tested a number of gDNA isolation methods, including commercial kits for forensic DNA isolation, but we learned that the high-salt precipitation protocol described here is superior to any column-based method, not only by its low costs and easy procedure, but also regarding yield and quality.

1. Resuspend each cell pellet in 100 μ L 1 \times PK-buffer, then add 10 μ L 10% SDS and 4.4 μ L 1 M DTT, incubate at 60–65 °C for 1 h.
2. Add 2 μ L Proteinase K (20 mg/mL) and incubate for another hour at 60 °C.
3. Add 30 μ L NaCl (4.5 M) and put the reaction tube on ice immediately for 10 min; centrifuge for 20 min at RT, full speed.
4. Transfer supernatant to a new tube, add 0.7 \times volume 100% iPrOH (approx. 110 μ L), centrifuge for 20 min at RT, full speed.
5. Remove supernatant and add 500 μ L 70% EtOH, incubate overnight in fridge (*see* **Notes 17** and **18**).
6. Re-pellet DNA for 2.5 min at full speed.
7. Carefully remove supernatant and air-dry for 6–10 min (*see* **Note 19**).
8. Resuspend pellet in an appropriate volume of T-buffer (30–60 μ L) and incubate for 1 h at 60 °C.

3.7 Evaluation of Single Cell Clones: Endpoint-PCR-Based Evaluation Including Sanger Sequencing

1. Amplify the target region from each single cell clone individually.
2. Purify the PCR product directly from PCR reaction using the PCR clean-up kit and resuspend DNA in 30–40 μ L of T-buffer.
3. Load 10–20 μ L of the purified PCR product on agarose gel for confirming the abundance of the desired amplicon.
4. Use 20–100 ng of the purified PCR product for Sanger sequencing with BigDye Terminator (*see* **Note 20**).
5. Analyze sequencing electropherograms for introduced mutations and polymorphic positions indicating the amplified alleles (*Fig. 3*).

For many HR-based approaches a conventional end-point PCR for screening is difficult or even impossible to establish. This is definitely true when HR vectors such as bacterial artificial

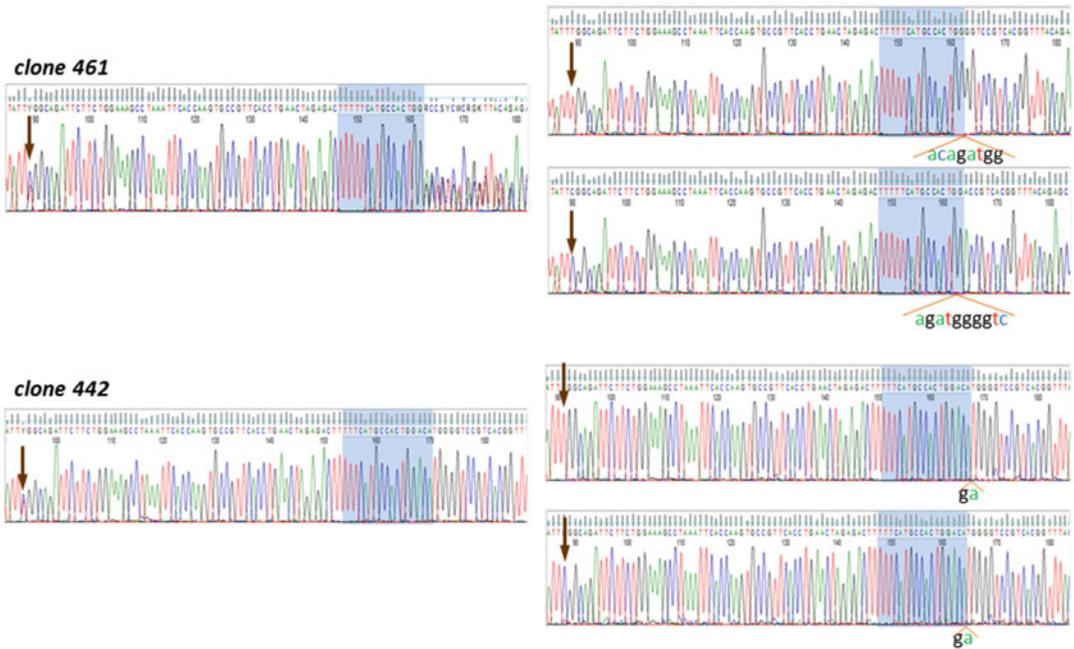


Fig. 3 Sanger Sequencing of single cell clones. After amplification of the target region, the PCR products are sequenced and the electropherograms are examined for mutations and polymorphisms. Clone 461 is a typical example where different mutations were introduced into the two alleles, indicated by the mixed peaks in the right part and by the polymorphic position in the left part of the electropherogram (brown arrow). Cloning of the PCR product revealed the specific modification of the respective alleles (right side). Clone 442 shows a continuous electropherogram throughout the amplicon, but the polymorphic site indicates that the same modification has been introduced on both alleles. Again, this finding was confirmed by sequencing cloned PCR products (right side)

3.8 Evaluation of Single Cell Clones: qPCR-Based Loss-of-Wild-Type-Allele (LOWA) Approach

chromosomes (BACs) are used, because of their extended homologous arms of several 10 kbs [14, 16]. It is also often difficult to establish end-point PCR spanning the homologous arms of smaller targeting vectors, in particular if the low amount of available template DNA from single cell clones is considered. Constructing control templates for the latter case might be an option, but this is very laborious. Thus, we rely on a qPCR-based loss-of-wild-type-allele (LOWA) approach as a standard method to detect the relative copy number of the target site versus reference sites somewhere else in the genome (Fig. 4).

1. Design and optimize one qPCR for the target site and two for different reference genes (see Note 21).
2. Analyze a given set of clones in 96-well plates.
 - (a) Run each qPCR in a 96-well plate under its optimized conditions.
 - (b) Compose each plate identically, comprising a standard curve, mostly ranging from 4000 copies/ μ L to

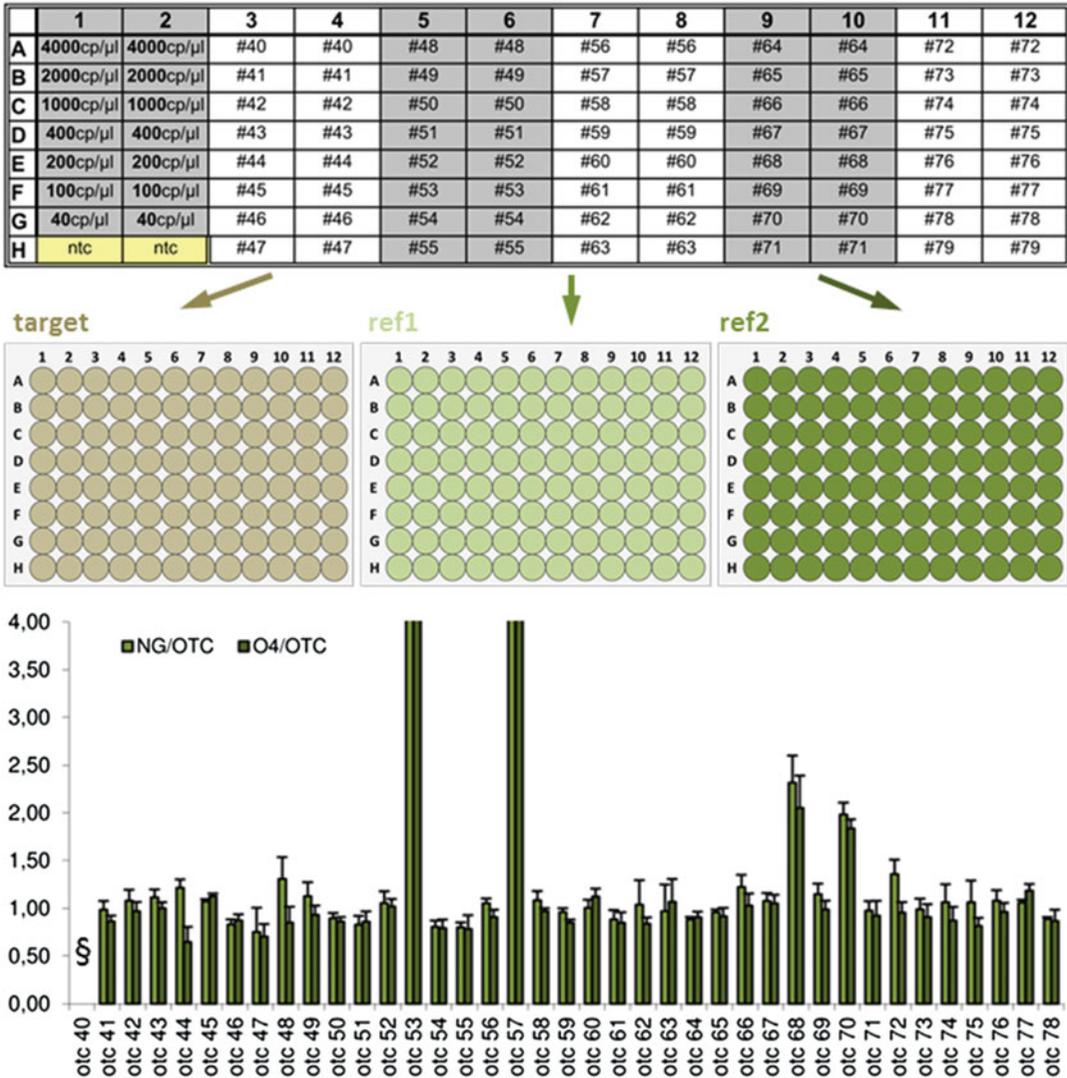


Fig. 4 Loss-of-wild-type-allele (LOWA) approach. A set of clones comprise a standard curve and a non-template control (columns 1 & 2) as well as DNA from 40 single cell clones, with each sample run as a duplicate (upper panel). Three different plates, one with the target qPCR and two others with reference qPCR are run independently at their optimized conditions (middle panel). The copy numbers of the three qPCR is determined for each single cell clones to calculate the relative copy number ratios (ref1: target) and (ref2: target) on a MicroSoft Excel basis (lower panel). Clones with low copy numbers (<40 mean copies for both reference sites) are not examined (indicated by the paragraph sign)

- 40 copies/μL, non-template controls, and 40 single cell clones (Fig. 4).
- (c) Run each standard concentration and each sample on a plate in duplicates.

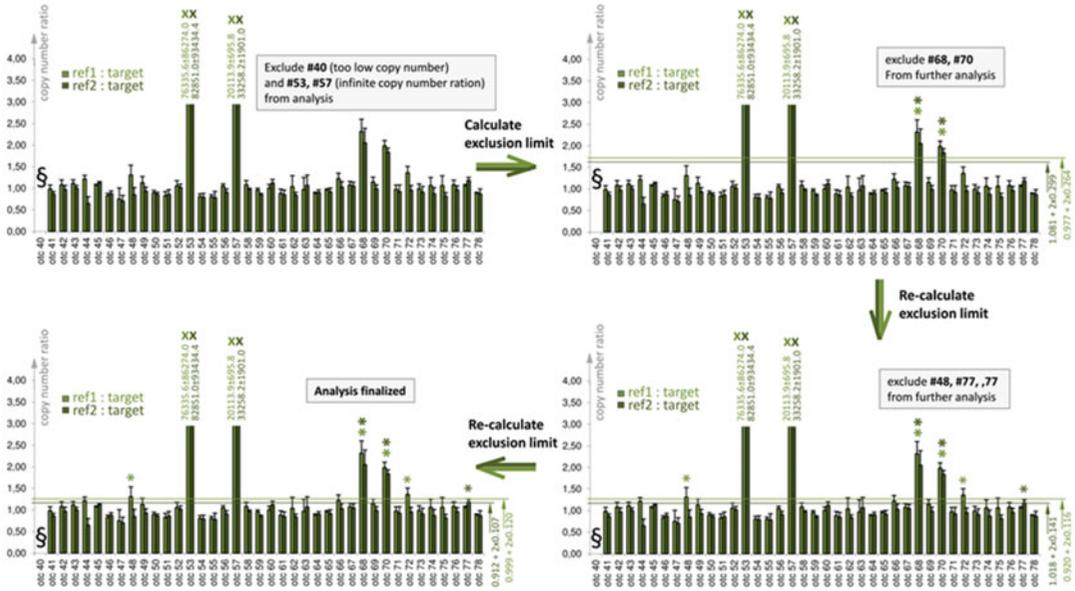


Fig. 5 Iterative calculation of recombination events in single cell clones by LOWA. A specific data set is given as an example. Clones with low copy numbers (<40 mean copies for both reference sites, indicated by the paragraph sign) or with infinite copy number ratios (bi-allelic targeting, indicated by “x”) are excluded from further analysis (upper left). Then exclusion limits for copy number ratios are calculated from mean values and standard deviations of the remaining cell clones. Clones with copy number ratios > than exclusion criteria are further excluded (upper right, indicated by asterisks) and calculation of mean value and standard deviation is repeated. Recalculation is repeated until no more clones appear above the exclusion limit (lower panel)

3. Determine the copy numbers on the basis of the respective standard curves for each site in each sample using the Light-cycler 96 software.
4. Export copy numbers to Excel, determine and plot the relative copy number ratios (reference site: target site) for each set of clones (*see Note 22*).
5. In a statistical manner, perform following steps for identification of modified cell clones in an iterative way (Fig. 5).
 - (a) Cell clones that obviously underwent bi-allelic modification are identified by a copy number ratio that is tending toward infinity (*see Note 23*). Exclude these clones from the analysis.
 - (b) Calculate mean value and standard deviation for the remaining copy number ratios.
 - (c) Define an exclusion limit for copy number ratio (mean value +2× standard deviation).
 - (d) Cell clones with copy number ratios above this limit are supposed to retain only one of the initial wild-type alleles and are, thus, supposed to be modified on one allele.

- (e) In a second run, exclude the clones above the exclusion limit from the analysis. Recalculate the copy number ratio of mean value $+2 \times$ standard deviation.
- (f) Repeat the iterative recalculation until no additional clones are identified above the exclusion limit (*see Note 24*).
- (g) Cell clones that exceed the final exclusion limit for both, (ref1: target) and (ref2: target), are seen as modified on one of the alleles (*see Note 25*).

4 Notes

1. One of the most often discussed drawbacks of Gene Editing are additional nuclease induced modifications somewhere else in the genome, so-called off-target effects. Indeed, such by-products of Gene Editing need to be considered, but on the other hand it is extremely difficult to really estimate the off-targeting prevalence. Any of the described prediction tools consider off-targets on the basis of the available reference genomes, but the quality of these genomes is far from perfect. This is particularly true for the pig, where the Sscrofa10.2 genome was only of limited use for either in silico evaluation or NGS-based full genome sequencing. The newly available Sscrofa11.1 version seems to be a significant improvement, but by now we cannot estimate how reliable the database really is. The quality of the available bovine reference genomes (since a couple of years it is mostly UMD3.1), in contrast, has been classified as highly reliable.
2. Predicting gRNAs on the basis of pig and cattle reference genomes may also be hampered by the significant genetic diversity in the pig and cattle populations. As a consequence, binding of a suggested gRNAs might be impaired or even prevented if single nucleotide polymorphisms appear in the target cell line. Thus, we recommend verifying the region of interest by sequencing the target cell line before gRNAs are designed. Polymorphisms might be helpful, on the other hand, when they appear in vicinity to the gRNA binding site, because they allow discrimination of the two alleles during screening of single cell clones.
3. We, however, refer from such attempts, as we realized that the assembly of such combination vectors for each specific purpose is time-consuming. Rather, for transcribing any single gRNA we have a plasmid commercially synthesized for low cost and at fast supply (normally within 7 working days) that transcribes the desired gRNA under the control of the human U6

promoter. This allows minimal hands-on time and immediate co-transfection of the gRNA plasmids together with the other components upon arrival. The necessity for introducing multiple distinct DNA vectors into a single cell at once is not a major obstacle in our hands, as the nucleofection efficiency of DNA into bovine and pig primary cells is regularly 80–90%.

4. The proliferation capacity of primary cells depends on several factors such as the tissue source, the age of the donor animal and the time between taking the sample and the starting of the isolation process. Thus, the isolation protocols of primary cells might need some adaptation.
5. The passaging process is a volatile process. Some adaptation might be necessary as well, and visual control of the releasing process is important. As a rule of thumb, we recommend the following for passaging primary cells: For dispatching cells, trypsinization solution should sufficiently cover the cells; trypsinization is stopped by cell culture medium of a volume that is normally used for cultivation.
6. The optimal concentration of DNA (gRNA and Cas9 plasmids and, optionally, targeting vector) should be determined for each cell type and for each gRNA. In general, mutation frequency normally increases at high concentrations, but at the cost of cell viability. Therefore, we recommend optimizing the transfection process by nucleofecting several aliquots of 5×10^5 cells with varying concentrations of gRNA and Cas9 plasmids ranging 250–1000 ng/ μ L each. Due to the limited capacity of total exogenous DNA in a transfection reaction, concentrations of gRNA and Cas9 plasmids are reduced (200–300 ng/ μ L) when combined with a targeting vector, which normally ranges between 200–750 ng/ μ L.
7. We evaluated the transfection efficiency of different programs supplied by the manufacturer and found program U12 to be the most efficient for our cell types [29]. Nevertheless, we agree to the manufacturer's recommendations to test a variety of suggested programs for any new cell type to optimize the nucleofection procedure.
8. It is essential to transfer the nucleofected cells carefully into the 35 mm dish. It is important that the cell suspension from the cuvette is added to the culture dish in a dropwise manner, as recommended by manufacturer.
9. Before amplifying the target site from nucleofected mixed clonal populations, the PCR primer pairs should be tested for their specificity to both alleles in the respective target cell lines, because the primer binding sites might be affected by polymorphisms. In that sense the abundance and amplification of

SNPs within a PCR product are helpful indicators that both alleles have been amplified.

10. NGS seems to be the fastest and easiest way to determine gRNA efficacy. However, not all laboratories have immediate access to such high-end methods and commercial NGS analysis is often expensive and time-consuming until results are eventually available. T7 endonuclease assays still represent a standard method, but in our laboratory we found it difficult to establish endonuclease approaches for different target sites. Thus, we regularly clone PCR products into plasmids and sequence them for estimating the gRNA efficacy. Analyzing 100 plasmids gives a good estimation of gRNA activity. Sanger sequencing of PCR products and analysis of the electropherograms by TIDE [30] looks a very promising short-cut of the plasmid cloning approach, but we do not have enough experience to comment on its suitability for estimating gRNA efficacy in primary cells.
11. In our laboratory we have evaluated gRNAs for 15 distinct target sites in primary cells, mainly from the pig, and found that NHEJ mutation frequency varies widely in a range between <1–40%. It is of note that even the efficacy of gRNAs that differ just in position of a few nucleotides or have a different orientation can vary widely.
12. By following this strategy, one cell is statistically seeded into each well. This should be, however, optimized for each cell type, due to varying proliferation capacities of cells in clonal selection.
13. Standard 96-well plates work as well, but we prefer using half area plates, because their structure facilitates screening with the Cellavista automated microscope.
14. For sure, it is possible to scan the plates by hand using a phase contrast microscope, but only an automated system allows effective production of single cell clones due to the much higher throughput.
15. The proliferation capacity of single cell clones varies widely, so the growth of each single cell clone has to be followed tightly and the timepoint of passaging needs to be chosen individually. In general, we follow the proliferation process for 5–6 days before we passage well proliferating clones to standard 96-well plates. If split too early, cell clones often fail to proliferate further.
16. For convenience of work, we maintain these backup cells at -80°C for up to one month until screening of the clones is finalized and appropriately modified clones are identified. These clones are then thawed, propagated in three 96-wells

and re-frozen in liquid nitrogen until used for experimental purposes such as SCNT [31].

17. Removing traces of iPrOH or EtOH is a critical step, so we recommend to carefully remove iPrOH and EtOH with a small (200 μ L) pipet tip. When using a 1 ml pipet, a small pipet tip should be placed at the top of the larger 1 mL pipet tip to facilitate removing the alcohol entirely.
18. The overnight washing in 70% EtOH is strongly recommended to achieve high DNA quality. Incubation for only a few hours dramatically impairs its usage for end-point PCR and qPCR.
19. Remaining EtOH as well as over-drying of the DNA pellet, e.g., in a vacuum concentrator or speed-vac, will impair sensitive analysis such as qPCR. Air drying after precise removal of EtOH gives the best results in our laboratory.
20. For Sanger sequencing of PCR products amplified from the gDNA of single cell clones we slightly modified the manufacturer's protocol:
 - (a) A sequencing reaction of 10 μ L is prepared of 4 μ L 5 \times BigDye sequencing buffer, 1 μ L BigDye Terminator, 5–10 pmol primer and template PCR product.
 - (b) The cycling program is generally done with 1 min at 95 $^{\circ}$ C, 40 \times (5 s at 95 $^{\circ}$ C, 10 s at 54 $^{\circ}$ C, 4 min at 60 $^{\circ}$ C), 10 min at 4 $^{\circ}$ C. The annealing temperature can be lowered to 50 $^{\circ}$ C in case of A-T-rich templates or stepwise increased to 60 $^{\circ}$ C in case of GC-rich templates.
 - (c) The sequencing reaction is mixed with 75 μ L freshly made precipitation solution (H₂O:125 mM EDTA: EtOH = 10:5:60), incubated on ice for 15 min and centrifuged for 30 min at 4 $^{\circ}$ C.
 - (d) The DNA pellet is washed with 150 μ L 70% EtOH, vortexed and centrifuged again for 2.5 min.
 - (e) EtOH is carefully removed and the pellet is air-dried for 6–10 min and resolved in 30 μ L double-distilled water.
 - (f) Analysis of the purified sequencing reactions is done on a 48-capillary sequencer.
21. Designing and running qPCR is challenging, in particular if one is intending to discriminate a 2:1 copy number ratio on the basis of only a few hundred templates per reaction. We found it practicable to design primers for amplicons not longer than 250 nt by conventional primer design programs and run the primers with the FastGreen intercalating dye. For more details, specific guidelines for establishing qPCR might be considered [32, 33]. Once appropriate qPCR have been established, they can serve for any new target site.

22. If qPCR sets run perfectly, mono- and bi-allelically modified cell clones can be clearly identified visually as illustrated in the Fig. 4. In nonoptimal qPCR sets, however, the described statistical method may be helpful to identify the modified clones appropriately.
23. The calculated copy number ratios for bi-allelically modified clones vary widely. This is particularly true if only few copies are in the sample, or if background noise is significant. We found it, thus, practicable to set the exclusion criterion for bi-allelically modified clones to 10 and to exclude samples that contain less than 40 copies/ μL of the reference genes from the analysis.
24. We have based the definition of our exclusion criteria on the assumption that the copy number ratios within a given set of clones distribute according to Gauss. Thus, defining an exclusion criterion of mean value $+2\times$ standard deviation would mean that 2.5% of unmodified clones would lie above this limit, according to statistical deviation. We have not tested this statistical assumption, but our extended experience in using these exclusion criteria correctly identified modified clones for many different target sites and fits for our purposes.
25. For sure, further examination of the clones identified by the LOWA approach is recommended. This might include the screening of the non-recombined allele for NHEJ-based modifications, the integration of the desired modification, etc. The limited amount of DNA from single cell clones, however, prevents the analysis for off-target events by NGS. The latter might be theoretically feasible after whole genome amplification (WGA), but we do not have experience with this method.

References

1. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121):819–823. <https://doi.org/10.1126/science.1231143>
2. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-guided human genome engineering via Cas9. *Science* 339(6121):823–826. <https://doi.org/10.1126/science.1232033>
3. 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. <https://doi.org/10.1126/science.1225829>
4. Bolze A, Mahlaoui N, Byun M, Turner B, Trede N, Ellis SR, Abhyankar A, Itan Y, Patin E, Brebner S, Sackstein P, Puel A, Picard C, Abel L, Quintana-Murci L, Faust SN, Williams AP, Barretto R, Duddridge M, Kini U, Pollard AJ, Gaud C, Frange P, Orbach D, Emile JF, Stephan JL, Sorensen R, Plebani A, Hammarstrom L, Conley ME, Selleri L, Casanova JL (2013) Ribosomal protein SA haploinsufficiency in humans with isolated congenital asplenia. *Science* 340(6135):976–978. <https://doi.org/10.1126/science.1234864>
5. Lupianez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, Santos-Simarro F, Gilbert-Dussardier B, Wittler L, Borschiwer M, Haas SA, Osterwalder M, Franke M, Timmermann B, Hecht J, Spielmann M, Visel A, Mundlos S (2015) Disruptions of topological chromatin domains

- cause pathogenic rewiring of gene-enhancer interactions. *Cell* 161(5):1012–1025. <https://doi.org/10.1016/j.cell.2015.04.004>
6. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153(4):910–918. <https://doi.org/10.1016/j.cell.2013.04.025>
 7. Wu J, Platero-Luengo A, Sakurai M, Sugawara A, Gil MA, Yamauchi T, Suzuki K, Bogliotti YS, Cuello C, Morales Valencia M, Okumura D, Luo J, Vilarino M, Parrilla I, Soto DA, Martinez CA, Hishida T, Sanchez-Bautista S, Martinez-Martinez ML, Wang H, Nohalez A, Aizawa E, Martinez-Redondo P, Ocampo A, Reddy P, Roca J, Maga EA, Esteban CR, Berggren WT, Nunez Delicado E, Lajara J, Guillen I, Guillen P, Campistol JM, Martinez EA, Ross PJ, Izpisua Belmonte JC (2017) Interspecies Chimerism with mammalian pluripotent stem cells. *Cell* 168(3):473–486. e415. <https://doi.org/10.1016/j.cell.2016.12.036>
 8. Zhu Z, Verma N, Gonzalez F, Shi ZD, Huangfu D (2015) A CRISPR/Cas-mediated selection-free knockin strategy in human embryonic stem cells. *Stem Cell Rep* 4(6):1103–1111. <https://doi.org/10.1016/j.stemcr.2015.04.016>
 9. Carlson DF, Lancto CA, Zang B, Kim ES, Walton M, Oldeschulte D, Seabury C, Sonstegard TS, Fahrenkrug SC (2016) Production of hornless dairy cattle from genome-edited cell lines. *Nat Biotechnol* 34(5):479–481. <https://doi.org/10.1038/nbt.3560>
 10. Niu D, Wei HJ, Lin L, George H, Wang T, Lee IH, Zhao HY, Wang Y, Kan Y, Shrock E, Lesha E, Wang G, Luo Y, Qing Y, Jiao D, Zhao H, Zhou X, Wang S, Wei H, Guell M, Church GM, Yang L (2017) Inactivation of porcine endogenous retrovirus in pigs using CRISPR-Cas9. *Science* 357(6357):1303–1307. <https://doi.org/10.1126/science.aan4187>
 11. Nottle MB, Salvaris EJ, Fiscaro N, McIlpatrick S, Vassiliev I, Hawthorne WJ, O’Connell PJ, Brady JL, Lew AM, Cowan PJ (2017) Targeted insertion of an anti-CD2 monoclonal antibody transgene into the GGTA1 locus in pigs using FokI-dCas9. *Sci Rep* 7(1):8383. <https://doi.org/10.1038/s41598-017-09030-6>
 12. Whitworth KM, Rowland RR, Ewen CL, Triple BR, Kerrigan MA, Cino-Ozuna AG, Samuel MS, Lightner JE, McLaren DG, Mileham AJ, Wells KD, Prather RS (2016) Genetically edited pigs are protected from porcine reproductive and respiratory syndrome virus. *Nat Biotechnol* 34(1):20–22. <https://doi.org/10.1038/nbt.3434>
 13. Bahr A, Kaser T, Kemter E, Gerner W, Kurome M, Baars W, Herbach N, Witter K, Wunsch A, Talker SC, Kessler B, Nagashima H, Saalmuller A, Schwinzer R, Wolf E, Klymiuk N (2016) Ubiquitous LEA29Y expression blocks T cell co-stimulation but permits sexual reproduction in genetically modified pigs. *PLoS One* 11(5):e0155676. <https://doi.org/10.1371/journal.pone.0155676>
 14. Klymiuk N, Blutke A, Graf A, Krause S, Burkhardt K, Wuensch A, Krebs S, Kessler B, Zakhartchenko V, Kurome M, Kemter E, Nagashima H, Schoser B, Herbach N, Blum H, Wanke R, Aartsma-Rus A, Thirion C, Lochmuller H, Walter MC, Wolf E (2013) Dystrophin-deficient pigs provide new insights into the hierarchy of physiological derangements of dystrophic muscle. *Hum Mol Genet* 22(21):4368–4382. <https://doi.org/10.1093/hmg/ddt287>
 15. Klymiuk N, Bocker W, Schonitzer V, Bahr A, Radic T, Frohlich T, Wunsch A, Kessler B, Kurome M, Schilling E, Herbach N, Wanke R, Nagashima H, Mutschler W, Arnold GJ, Schwinzer R, Schieker M, Wolf E (2012) First inducible transgene expression in porcine large animal models. *FASEB J* 26(3):1086–1099. <https://doi.org/10.1096/fj.11-185041>
 16. Klymiuk N, Mundhenk L, Kraehe K, Wuensch A, Plog S, Emrich D, Langenmayer MC, Stehr M, Holzinger A, Kroner C, Richter A, Kessler B, Kurome M, Eddicks M, Nagashima H, Heinritzi K, Gruber AD, Wolf E (2012) Sequential targeting of CFTR by BAC vectors generates a novel pig model of cystic fibrosis. *J Mol Med (Berl)* 90(5):597–608. <https://doi.org/10.1007/s00109-011-0839-y>
 17. Klymiuk N, van Buerck L, Bahr A, Offers M, Kessler B, Wuensch A, Kurome M, Thormann M, Lochner K, Nagashima H, Herbach N, Wanke R, Seissler J, Wolf E (2012) Xenografted islet cell clusters from INSLEA29Y transgenic pigs rescue diabetes and prevent immune rejection in humanized mice. *Diabetes* 61(6):1527–1532. <https://doi.org/10.2337/db11-1325>
 18. Simmet K, Reichenbach M, Reichenbach HD, Wolf E (2015) Phytohemagglutinin facilitates the aggregation of blastomere pairs from Day 5 donor embryos with Day 4 host embryos for chimeric bovine embryo multiplication.

- Theriogenology 84(9):1603–1610. <https://doi.org/10.1016/j.theriogenology.2015.08.012>
19. Wuensch A, Baehr A, Bongoni AK, Kemter E, Blutke A, Baars W, Haertle S, Zakhartchenko V, Kurome M, Kessler B, Faber C, Abicht JM, Reichart B, Wanke R, Schwinzer R, Nagashima H, Rieben R, Ayares D, Wolf E, Klymiuk N (2014) Regulatory sequences of the porcine THBD gene facilitate endothelial-specific expression of bioactive human thrombomodulin in single- and multitransgenic pigs. *Transplantation* 97(2):138–147. <https://doi.org/10.1097/TP.0b013e3182a95cbc>
 20. Klymiuk N, Fezert P, Wunsch A, Kurome M, Kessler B, Wolf E (2014) Homologous recombination contributes to the repair of zinc-finger-nuclease induced double strand breaks in pig primary cells and facilitates recombination with exogenous DNA. *J Biotechnol* 177:74–81. <https://doi.org/10.1016/j.jbiotec.2014.01.018>
 21. Simmet K, Zakhartchenko V, Philippou-Massier J, Blum H, Klymiuk N, Wolf E (2018) OCT4/POU5F1 is required for NANOG expression in bovine blastocysts. *Proc Natl Acad Sci U S A* 115(11):2770–2775. <https://doi.org/10.1073/pnas.1718833115>
 22. Bouabe H, Okkenhaug K (2013) A protocol for construction of gene targeting vectors and generation of homologous recombinant embryonic stem cells. *Methods Mol Biol* 1064:337–354. https://doi.org/10.1007/978-1-62703-601-6_24
 23. Bouabe H, Okkenhaug K (2013) Gene targeting in mice: a review. *Methods Mol Biol* 1064:315–336. https://doi.org/10.1007/978-1-62703-601-6_23
 24. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (2013) Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8(11):2281–2308. <https://doi.org/10.1038/nprot.2013.143>
 25. Labun K, Montague TG, Gagnon JA, Thyme SB, Valen E (2016) CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res* 44(W1):W272–W276. <https://doi.org/10.1093/nar/gkw398>
 26. Heigwer F, Kerr G, Boutros M (2014) E-CRISP: fast CRISPR target site identification. *Nat Methods* 11(2):122–123. <https://doi.org/10.1038/nmeth.2812>
 27. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, Cradick TJ, Marraffini LA, Bao G, Zhang F (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* 31(9):827–832. <https://doi.org/10.1038/nbt.2647>
 28. Anders C, Niewoehner O, Duerst A, Jinek M (2014) Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 513(7519):569–573. <https://doi.org/10.1038/nature13579>
 29. Richter A, Kurome M, Kessler B, Zakhartchenko V, Klymiuk N, Nagashima H, Wolf E, Wuensch A (2012) Potential of primary kidney cells for somatic cell nuclear transfer mediated transgenesis in pig. *BMC Biotechnol* 12:84. <https://doi.org/10.1186/1472-6750-12-84>
 30. Brinkman EK, Chen T, Amendola M, van Steensel B (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42(22):e168. <https://doi.org/10.1093/nar/gku936>
 31. Kurome M, Kessler B, Wuensch A, Nagashima H, Wolf E (2015) Nuclear transfer and transgenesis in the pig. *Methods Mol Biol* 1222:37–59. https://doi.org/10.1007/978-1-4939-1594-1_4
 32. Bustin S, Huggett J (2017) qPCR primer design revisited. *Biomol Detect Quantif* 14:19–28. <https://doi.org/10.1016/j.bdq.2017.11.001>
 33. Bustin SA, Benes V, Nolan T, Pfaffl MW (2005) Quantitative real-time RT-PCR—a perspective. *J Mol Endocrinol* 34(3):597–601. <https://doi.org/10.1677/jme.1.01755>

Part V

CRISPR Gene Therapy and Screening



Toward In Vivo Gene Therapy Using CRISPR

Kristian Alsbjerg Skipper and Jacob Giehm Mikkelsen

Abstract

CRISPR, a revolutionizing technology allowing researchers to navigate in and edit the genome, is moving on the fast track toward clinical use for ex vivo correction of disease-causing mutations in stem cells. As we await the first trials utilizing ex vivo CRISPR editing, implementation of CRISPR-based gene editing as an in vivo treatment directly in patients still remains an ultimate challenge. However, quickly accumulating evidence has provided proof-of-concept for efficacious editing in vivo. Attempts to edit genes directly in animals have largely relied on classical vector systems based on virus-based delivery of gene cassettes encoding the Cas9 endonuclease and single guide RNA, the key components of the CRISPR system. However, whereas persistent gene expression has been the primary goal of gene therapy for decades, things may be different in the case of CRISPR delivery. Is short-term presence of the CRISPR components perhaps sufficient for efficacy and ideal for safety?—and are strategies needed for restricting immune recognition of the bacteria-derived editing tool? Here, while answers to these questions still blow in the wind, we review prominent examples of genome editing with focus on targeting of genes with CRISPR in liver, muscles, and eyes of the mouse.

Key words Genome editing, Knockout, AAV, Alpha-1 antitrypsin deficiency, Hypercholesterolemia, Muscular dystrophy, Eye, Cas9 immune responses

1 Introduction

Enough said about CRISPR? Well, a revolution does not fade away overnight, and this one—fueled by the groundbreaking discovery of a genome navigation system—just got started and does not seem to slow down now or in the near future. In the shadows of some of the hype that inevitably followed in the wake of the discovery of CRISPR, genome editing has quickly become a standard technology in laboratories worldwide. Gene knockouts and genome-wide CRISPR screens are moving into everyday research, and recent discoveries demonstrating clinical-grade potency in stem cells [1, 2] promise upcoming clinical trials utilizing CRISPR for ex vivo genetic therapies in stem cells. It is not unlikely, hence, that editing in hematopoietic stem cells for treatment of disease of

the blood, like immunodeficiencies and sickle cell disease, will make it into clinical trials in the near future.

For years, the ultimate goal of gene transfer technologies and gene therapy trials has been life-long expression of one or more therapeutic genes. And it still is. Vectors have been developed allowing persistent gene expression from stably integrated or episomal platforms for long-term treatment of dividing and nondividing cells, respectively. Currently, the gene therapy community leans toward using lentiviral vectors for integration of gene cassettes and vectors based on adeno-associated virus (AAV) for purposes that do not require genomic insertion of a transgene. With the emergence of genome editing systems, initially based on zinc-finger nucleases and TAL-effector nucleases and now primarily on CRISPR, and their use in the next generation of genetic therapies, it has come easily to the community to find inspiration in conventional gene delivery systems for ferrying gene editing tool kits to cells or tissues. In a typical CRISPR setup based on either lentiviral or AAV-based vectors, one gene cassette encodes the Cas9 endonuclease (often SpCas9 derived from *Streptococcus pyogenes*) from an RNA polymerase II promoter, whereas a corresponding single guide RNA (sgRNA) is produced from a separate expression cassette driven by an RNA polymerase III promoter. This is a powerful combination, but the ultimate goal of vector-based delivery has changed: rather than aiming for life-long expression, potent delivery combined with short-term activity of the locus-targeted endonuclease are now major criteria for successful and safe editing of the genome. For ex vivo treatments in stem cells, it appears that electroporation of ribonucleoprotein (RNP) complexes consisting of recombinant Cas9 and synthetic sgRNA fulfills the criteria, facilitating high levels of targeting in a short window of time [2, 3].

For correction of disease-causing mutations in the human genome, the creation of a double-stranded DNA break using CRISPR in a predetermined genomic location is a job half done. For purposes that require exchange of one or more nucleotides, and where classical gene knockout does not suffice, the delivery of the endonuclease needs to be combined with the delivery of a DNA donor for repair typically by homologous recombination. In hematopoietic stem cells, integrase-defective lentiviral vectors (IDLVs) can serve as donors for repair [4, 5], and donors based on AAV serotype 6 have more recently been shown to facilitate high levels of repair [6]. In fact, a combination of RNP-based SpCas9/sgRNA delivery and AAV6-directed delivery of the donor sequence currently seems as the most promising approach for correction of mutations in hematopoietic stem cells [1, 2] and may soon be offered to patients in the first clinical trials of their kind.

As we await the first trials utilizing ex vivo CRISPR editing, implementation of CRISPR-based gene editing as a treatment directly in patients still remains an ultimate goal, which is

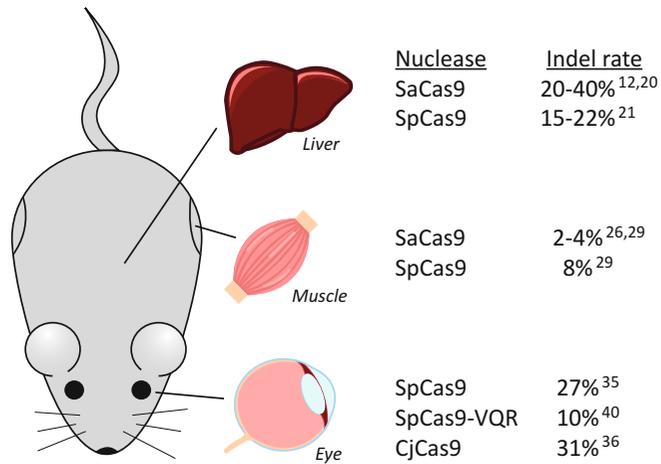


Fig. 1 Graphic overview of in vivo indel rates that have been achieved by AAV-directed CRISPR delivery in liver, muscle, and eye of the mouse. Key references are listed in parentheses. The relevant organs are schematically represented and do not reflect the exact anatomy of mouse organs

challenged by the need of combining potent in vivo delivery of Cas9/sgRNA/donor tool kits with short-lived endonuclease activity and, not least, securing the timing between the formation of double-stranded breaks and DNA donor accessibility. So far, efforts to edit genes directly in animals have largely relied on conventional gene delivery systems with the inherent risk of prolonging exposure of the genome to endonuclease activity. Here, we review the current status of virus-based genome editing in the mouse and with focus on the liver, the muscle, and the eye highlight some of the prime attempts that have been made to establish CRISPR-based genome editing in vivo. Examples of gene disruption rates that have been achieved by in vivo CRISPR delivery to these organs are listed in Fig. 1.

2 CRISPR-Based Gene Knockout and Editing in Mouse Liver

Step by step over the last two decades AAV-based gene augmentation therapies in the liver have emerged as viable treatments for liver-based disorders. For long, treatment of hemophilia has been considered the holy grail of gene therapy, and recent findings indeed document stunning therapeutic effects in patients suffering from hemophilia B treated with AAV8 vectors carrying an optimized factor IX gene expression cassette leading to high levels of persistent factor IX production [7]. Although long-term follow-up of efficacy and safety is still needed, it seems that the path to the holy grail has unearthed, only to fuel hopes for the next generation of genetic treatments targeting the liver.

Evidence of CRISPR-directed treatment in the liver was first provided in studies utilizing hydrodynamic co-delivery of a plasmid encoding SpCas9 and sgRNA and single-stranded DNA oligo (ssDNA) as a donor for repair. In a mouse model of hereditary tyrosinemia (HT), the authors repaired the disease variant gene encoding fumarylacetoacetate hydrolase (FAH) in 0.4% of the hepatocytes. In this disease model, however, selective advantage and expansion of repaired hepatocytes compensated for the limited correction efficacy and resulted in amelioration of the disease phenotype [8]. Taking advantage of potent liver transduction using adenovirus-based vectors, Cheng and coworkers managed to introduce indels in 90% of targeted *Cebpa* alleles in mouse liver [9], whereas Ding et al. introduced knockout mutations in more than half of the *Pcsk9* alleles in mouse livers treated with adenovirus-encoded SpCas9 and sgRNA. Notably, this level of targeting of the *Pcsk9* gene led to reduced levels of secreted PCSK9. Since PCSK9 supports degradation of the low-density lipoprotein (LDL) receptor, *Pcsk9* knockout was accompanied by increased levels of LDL-receptor leading to a reduction in the plasma cholesterol levels [10]. In a third study exploiting adenovirus-based delivery, indel formation was achieved in >20% of the targeted *pten* alleles, resulting in massive liver steatosis and liver phenotypes mimicking phenotypes typical of *pten* deletion [11]. Targeted genomic alterations by this approach were detectable as soon as 3 days after adenoviral vector injection [10] and were maintained, as expected, after liver regeneration induced by partial hepatectomy [9].

Due to the liver toxicity induced by adenovirus-based delivery strategies, adenoviral gene delivery is currently not considered viable for therapeutic treatment, and AAV-based vectors are by far the most popular vehicles for delivery of genetic material to the liver. Although early findings gave an indication of fairly low targeting rates in mouse liver treated with SpCas9/sgRNA-encoding AAV vectors (roughly 0.5% targeting of the miR-122 locus), a rapidly growing number of articles demonstrate efficacy of AAV-delivered CRISPR components in the liver. Along with the launch of the SaCas9 protein derived from *Staphylococcus aureus* and the development of SaCas9 sgRNA design tools, Ann Ran and coworkers showed robust knockout of the *Pcsk9* gene in mouse livers treated with AAV-8 capsids carrying a vector containing both a SaCas9 expression cassette driven by the liver-specific thyroxine-binding globulin promoter and an U6-driven matching sgRNA [12]. Indel formation was observed in >40% of the targeted alleles, leading to a massive drop in serum PCSK9 and an accompanying reduction in blood cholesterol levels. In another approach using AAV8 as a carrier, Thakore and coworkers delivered an expression cassette encoding a catalytically “dead” SaCas9 variant (dSaCas9) fused to the Krüppel-associated box (KRAB) epigenetic repressor motif [13]. The dSaCas9-KRAB repressor was targeted to the *Pcsk9*

gene by a sgRNA species expressed from a co-delivered AAV vector, resulting in transcriptional silencing of the gene and reduction of PCSK9 serum levels to 20% of the levels found in control animals. This resulted in a moderate reduction of serum levels of LDL and cholesterol. Although this approach also resulted in significant transcriptional changes elsewhere in the genome, side effects that will still need to be addressed, CRISPR-based regulators of gene expression may have durable, long-term therapeutic effects in the liver (in this particular case through 10 weeks) even without creating double-stranded breaks. The opposite phenotype leading to hypercholesterolemia and formation of atherosclerotic plaques was observed in mice in which knockout mutations in the gene encoding the LDL-receptor (up to 54% of the alleles) were introduced by transduction of the liver with AAV vectors encoding locus-specific sgRNAs [14].

Patients suffering from alpha-1 antitrypsin deficiency (A1ATD), a hereditary liver disease typically caused by homozygosity for the Z-variant of *SERPINA1* gene, have an increased risk of developing fibrosis, cirrhosis, and hepatocellular carcinoma. Liver toxicity in A1ATD patients is caused by intracellular retention and aggregation of mutant alpha-1 antitrypsin produced from the *SERPINA1* Z-allele, but homozygous patients also suffer from the lack of alpha-1 antitrypsin function, which is needed to control serine protease activity in the lung. A combination of CRISPR-based genetic intervention in the liver combined with AAV-based A1AT replacement gene therapy or delivery of a donor for direct repair of the mutation could serve as a new treatment modality addressing both the lung and liver pathology of A1ATD. Albeit with varying success, evidence of *SERPINA1* gene delivery to the liver has already been established [15–19]. Now, examples of robust CRISPR targeting of the *SERPINA1* Z-variant in the mouse, aiming at knocking out or repairing the Z-variant, are starting to emerge. Using AAV8-delivered SaCas9, Shen and colleagues achieved indel rates around 20% in the livers of PiZ mice carrying the human disease-causing Z-allele [20]. This resulted in markedly reduced levels of mutated *SERPINA1* mRNA and for one group of mice almost complete knockdown of A1AT expression, leading to reduction of aggregate formation in the liver. To achieve targeted correction of the Z-allele, the authors co-delivered two AAV8 vectors, one encoding SaCas9 and one carrying a mutant allele-specific sgRNA expression cassette along with donor template for homology-directed repair. This treatment resulted in production of A1AT RNA transcripts (up to 4.8% of total A1AT transcripts) encoding the repaired M form of A1AT, as measured 56 days after treatment, and reduced formation of A1AT aggregates in the liver. Roughly similar efficacies were obtained by Song et al., who intravenously co-injected two AAV vectors, one vector packaged in AAV9 capsids encoding SpCas9 and one AAV8-based vector

carrying the sgRNA expression cassette and the donor template [21]. By including a Myc tag in the donor vector, the authors could specifically detect corrected M-type A1AT, and found that 14% of the hepatocytes stained positively for corrected A1AT protein in PiZ mice that were treated one week after birth. The indel rate at the targeted locus was roughly 22% in these mice, whereas correction was detected at 2% of the targeted alleles. In adult PiZ mice, the authors were able to introduce indels at 15% of the targeted alleles and induce production of Myc-tagged A1AT in approximately one fifth of the hepatocytes [21].

The liver as a target organ for therapeutic genome editing has innumerable applications, which may collectively benefit from efforts that optimize and balance efficacy safety. Recent reports show potent targeting of the Factor IX gene [22] and CRISPR-directed cDNA insertion as a strategy for treatment of hemophilia B [22] potentially showing a way for next-generation genetic therapies for hemophilia. As of now, AAV-based vehicles are the obvious—and perhaps also easy—choice for efficient delivery of CRISPR components. We will see an increasing number of reports documenting targeted indel formation and repair in the liver, and focus on optimized gene expression and boosted CRISPR activity will probably lead to even higher levels of targeting. The safety of such approaches, based on long-term expression of Cas9 protein and sgRNA, should be looked at with caution, and therapeutic translation may depend on systems allowing strict regulation of gene expression and the ability to shut-down Cas9/sgRNA production. The implementation of self-limiting expression circuits allowing targeted, sgRNA-directed disruption of the Cas9 expression cassette is a feasible option for limiting the duration of Cas9 expression [23–25].

3 CRISPR Treats Muscular Dystrophies: AAV-CRISPR Partnership New Black Also in the Muscle

More than a handful of recent papers have documented robustness of genome editing using CRISPR in muscles leading to treatment of mouse models of muscular dystrophies. Duchenne muscular dystrophy (DMD) is a muscular degenerative disease caused by mutations, often deletions of one or more exons, in the dystrophin gene causing frameshifts or loss of production of functional dystrophin.

Using SaCas9 packaged in AAV8-vectors and a second AAV-vector encoding two sgRNAs targeting introns flanking exon 23 in the dystrophin gene, intramuscular and systemic delivery of the CRISPR system restored muscle function in the mdx model of Duchenne muscular dystrophy [26]. In this case, a fairly

low exon 23 deletion rate (>2%) resulted in a marked increase of exon 23-deleted transcripts supposedly due to protection of these transcripts from nonsense-mediated decay. In a parallel study, corresponding findings were made in the mdx mouse made using a similar saCas9/sgRNA delivery vector approach using AAV9 [27]. The same strategy was successful using a dual AAV9 vector approach by which spCas9 was expressed from a miniCMV promoter [28]. More recently, the delivery approach for treatment of DMD has been further refined by exploiting a strong muscle-specific promoter, CK8, driving expression of spCas9 in injected muscles for treatment of the exon 53 mutation in the mdx^{4cv} mouse model [29] and by delivering only a single sgRNA targeting a splice acceptor site, allowing exon skipping and restoration of the reading frame [30]. In an innovative approach, which may potentially increase safety of the interventional procedure, AAV9 capsids were exploited to co-deliver the spCas9 gene on one vector and expression cassettes expressing a shorter, inactive sgRNA (referred to as a “dead gRNA,” dgRNA) and a transactivation domain on another vector [31]. By exploiting the synergistic activation mediator (SAM) system developed by Konermann and colleagues [32], the MS2-P65-HSF1 (MPH) activation domain was recruited to epigenetically silenced genes in mdx mouse muscles by engineered dgRNAs containing two MS2 loop domain [31]. By this approach, targeted activation of the *klotho* gene was found to improve the strength of AAV-treated muscles, leading to amelioration of the DMD phenotype. Similarly, the *utrophin* gene could be activated in transgenic mdx mice expressing SpCas9 causing partial rescue of DMD in the model [31].

4 Treating Senses: CRISPR Genome Editing in Eye and Ear

The eye represents an easy accessible, immuno-privileged, and highly compartmentalized system making it an ideal target for in vivo genome editing. Indeed, several studies have utilized the CRISPR/Cas9 system to target retinal degenerative diseases. So far, the main focus has been on targeting inherited, monogenic diseases of the eye such as retinitis pigmentosa (RP) and Leber congenital amaurosis 10 (LCA10) [24, 33–38], but more complex diseases, like age-related macular degeneration (AMD) and primary open-angle glaucoma (POAG), have also been targeted [39–43].

In a proof-of-concept study a disease-causing mutation was targeted in the S334ter-3 model of autosomal dominant RP. By subretinal injection of plasmid DNA encoding sgRNA and SpCas9 followed by electroporation, Bakondi and coworkers achieved allele-targeted disruption of the *Rho*^{S334}-allele in approximately 0.32% of retinal cells. It should be noted that the sgRNA used was only allele-specific in the sense that it targeted the transgene,

mouse-derived *Rho*^{S334}-allele, in a rat background. Nevertheless, the researchers demonstrated robust retinal preservation and increased visual acuity in treated eyes [33]. Likewise, another study employing subretinal injections of plasmid DNA, this time targeting the *Rho*^{P23H} variant, also demonstrated *Rho* knockout. As above, the sgRNAs used in this study were not allele-specific and also targeted wild-type alleles [38]. To achieve allele-specific targeting of the *Rho*^{P23H} variant, the engineered SpCas9-VQR nuclease recognizing a 5'NGA PAM motif was exploited by utilizing the AAV9-PHP.B serotype for intravitreal delivery [34]. Remarkably, using this approach, the researchers achieved transduction of the entire retina. Indel rates were however mediocre in a FACS-enriched population with around 10% efficiency. Even so, targeting the *Rho*^{P23H} variant appeared allele-specific, as no detectable cleavage was seen in the wild-type *Rho* allele. Curiously, this contradicts a recent study where the researchers used the same sgRNA as above and demonstrated that only a truncated version enabled the Cas9 protein to distinguish between wild-type and targeted genomic variant [35].

The great genetic diversity in disease-causing mutations in RP has spurred researchers to seek out mutation-independent approaches for genetic treatment of RP. To that end, reprogramming rod to cone photoreceptors by *Nrl* repression has been suggested as a possible therapy [44]. In a recent report, Yu and coworkers cotransduced AAV8-packaged SpCas9 and sgRNAs targeting *Nrl*. Remarkably, deep sequencing showed indel formation in up to 98% of transduced rods in C57bl/6j mice. The researchers moved on and demonstrated that ablation of *Nrl* in three different mouse models of rod-induced retinal degeneration led to significant improvement of the disease phenotype [36]. In addition to generating a complete knockout of *Nrl*, Cas9-mediated transcriptional repression has been explored as a possible option for targeting *Nrl*. Although less efficient than a *Nrl* knockout control, delivery of a catalytically inactive split-SpCas9 (split-dCas9) fused to a KRAB repression domain together with sgRNAs targeting the transcriptional start site of *Nrl* has been shown to mediate robust knockdown leading to improved visual function [37].

An intronic variant (c.2991+1655A > G) in the *CEP290* gene constitutes around 15% of all LCA10 cases. By expressing dual sgRNAs flanking the cryptic splice site expressed from an AAV5 vector and co-delivered together with SpCas9, a study performed in C57BL/6J mice showed that introduction of an intronic deletion did not interfere with normal splicing of *CEP290* [24], thus paving the way for studies in more complex models. Interestingly, without enrichment for cells with successful transduction using FACS, remarkably high indel rates ranging from 7.5 to 26.5% in total retinal cells were obtained.

In contrast to the monogenic diseases described above resulting from single, disease-causing variants, AMD is caused by a dysregulation of angiogenic cytokines such as vascular endothelial growth factor A (VEGFA), leading to choroidal neovascularization (CNV). Treatments have thus focused on attenuating the signal transduction with neutralizing VEGF antibodies. Several studies have also focused on reducing CNV by introducing knockout mutations into either *vegfa*, its receptor *vegfr2*, or the transcription factor *bif1a*. In a pioneering study, researchers packaged expression cassettes encoding the smaller CjCas9 derived from *Campylobacter jejuni* linked to eGFP with a T2A peptide and a sgRNA targeting *vegfa* or *bif1a* in all-in-one AAV9 vectors. At day 42 after intravitreal injection, the indel rates in RPE cells ranged between 22 and 31% for *vegfa* or *bif1a*, respectively, leading to a 20–24% reduction in CNV in a laser-induced CNV model [43]. In addition to the CjCas9, the SpCas9 has been utilized to treat AMD and delivered to the retina using both dual AAVs, lentiviral vectors (LVs) and as RNPs, all of which led to reduced CNV. While dual AAV1-based delivery by intravitreal injections yielded an approximate 2% indel rate in *vegfr2* in total retinal cells [39], subretinal injection of LVs yielded up to 84% in RPE cells expressing Cas9 [41]. Furthermore, by subretinal injection of *vegfa*-targeting SpCas9 RNPs, indel rates around 22% were seen in enriched RPE cells [40].

As with the eye, the ear offers easy access. Interestingly, in contrast to the eye where AAVs have primarily been utilized to deliver DNA encoding the editing machinery, recombinant Cas9 protein has been delivered directly to the ear in complex with sgRNAs [45, 46]. By packaging these Cas9 RNPs into cationic lipids and injecting directly into the cochlea, editing efficiencies up to 25 % were reported [45]. In a recent study, this approach was used to target the disease-causing *Tmc1^{Bth}* allele in the *Tmc1^{Bth/+}* mouse model. Although the on-target editing efficiency was considerably lower than what had previously been demonstrated, the researchers showed a significant increase in hair-cell survival and rescue of progressive hearing loss in neonatal mice [46].

5 Have Patients Been Exposed to Cas9 Before?: Preexisting Immune Responses Against Cas9

As it stands out from the emerging literature, current strategies for efficient in vivo genome editing are based primarily on administration—in some cases systemic—of viral vectors co-expressing substantial amounts of either SpCas9 or SaCas9 and matching sgRNAs. Overall, this strategy results in high levels of targeted DNA cleavage (Fig. 1). In most cases, however, this approach

results in long-term production of the CRISPR components, resulting in prolonged exposure to Cas9 endonucleases. In the mouse, expression of SaCas9 evokes a cellular immune response and induced production of Cas9-responsive T cells as well as production of Cas9-specific antibodies [47], supporting the notion that Cas9 proteins are recognized as antigens.

Clinical translation of novel CRISPR-based genome therapeutics may strongly depend on the immune response to Cas9 in humans. Importantly, previous reports have stated that the majority of, if not all, human adults have been exposed to *S. pyogenes* [48] and *S. aureus* [49, 50] and therefore are positive for antibodies against these bacterial strains. This prompted recent investigations of preexisting immune responses to Cas9 in humans [51]. Among donors, 65% and 79% were seropositive for SpCas9 and SaCas9, respectively, whereas anti-SaCas9 cytotoxic T-cells were detected in half of the donors [51]. Although anti-SpCas9 T-cells could not be detected in all individuals, possibly due to technical reasons, these findings collectively bring to the attention that preexisting humoral and cell-based immune response may act upon Cas9 protein and Cas9-presenting cells upon administration of Cas9-encoding vectors. For ex vivo genome editing purposes, this may not constitute notable problems, but the related complications for in vivo use are currently unclear and supposedly represent an important challenge for direct clinical translation. It is speculated that sustained expression of either SpCas9 or SaCas9 will lead to immune clearance of Cas9-presenting cells and potentially to toxicity in tissues in which a larger percentage of the cells express Cas9 [51]. Such predictions call for further refinements of CRISPR technologies; could designed Cas9 variants potentially escape immune detection and maintain efficacy?—or could natural Cas9 variants, for example from bacterial species that are not present in humans, be more appropriate for in vivo use?

6 Concluding Remarks

Along with focusing on the CRISPR components that are effective and slip under the radar of the immune system, one may focus also on the delivery approach. AAV vectors are cherished for their capacity to support very high levels of transgene expression, but perhaps, in the case of CRISPR delivery, less is in fact better. If Cas9 protein is delivered exclusively near or at the chromatin and with high cell specificity, it may not be necessary to overexpress genome editing tool kits in transduced cells with the risk of boosting Cas9 presentation to T cells. This is obviously easier said than done, but alternative delivery systems may potentially support robust levels of endonuclease activity with minimal cellular exposure to the protein. Recently, we established virus-based delivery of enzymes, including

endonucleases, allowing direct delivery of ready-to-go genome editing tool kits [5, 52–54]. By facilitating direct protein delivery to the nucleus resulting in robust and time-restricted enzymatic activity [55], this approach may potentially allow safe CRISPR-based intervention, but evidence of potency in vivo is still lacking. Also, nonviral delivery approaches may have therapeutic potential. Recently, nonviral delivery of chemically stabilized sgRNA along with Cas9-encoding mRNA co-packaged in lipid nanoparticles was shown to induce knockout of >80% of targeted *Pcsk9* alleles in mouse liver upon a single intravenous injection [56]. Like we have seen for conventional gene therapies, different delivery strategies may suit different tissues and different therapeutic purposes. Although AAV-based delivery systems are currently at the forefront of in vivo genome editing, alternative approaches designed specifically for delivery of genomic tool kits will be out of the shadows.

Acknowledgments

Research focusing on genome engineering technologies in the laboratory of JGM is made possible through support of the Danish Council for Independent Research | Medical Sciences (grant DFF-4004-00220), The Lundbeck Foundation (grant RI26-2012-12456), the Novo Nordisk Foundation, the Hørslev Foundation, Aase og Ejnar Danielsens Fond, Grosserer L. F. Foghts Fond, Agnes og Poul Friis Fond, Oda og Hans Svenningsens Fond, Ellen og Mogens Wedell-Wedellsborgs Fond, Einar Willumsens Mindelegat, Snedkermester Sophus Jacobsen & Hustru Astrid Jacobsens Fond, Familien Hede Nielsens Fond, Carl og Ellen Hertz' legat til læge- og naturvidenskaben, Harboefonden, Arvid Nilssons Fond, Fonden af 1870, Emil C. Hertz og hustru Inger Hertz's Fond, Holger Hjortenberg og Hustru Dagmar Hjortenbergs Fond, Krista og Viggo Petersens Fond, Frode V. Nyegaard og Hustrus Fond, and Andersen-Isted Fonden.

References

1. Bak RO, Dever DP, Reinisch A, Cruz Hernandez D, Majeti R, Porteus MH (2017) Multiplexed genetic engineering of human hematopoietic stem and progenitor cells using CRISPR/Cas9 and AAV6. *Elife* 6:e27873
2. Dever DP, Bak RO, Reinisch A, Camarena J, Washington G, Nicolas CE, Pavel-Dinu M, Saxena N, Wilkens AB, Mantri S et al (2016) CRISPR/Cas9 beta-globin gene targeting in human haematopoietic stem cells. *Nature* 539:384–389
3. DeWitt MA, Magis W, Bray NL, Wang T, Berman JR, Urbinati F, Heo SJ, Mitros T, Munoz DP, Boffelli D et al (2016) Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. *Sci Transl Med* 8:360ra134
4. Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA, Ando D, Urnov FD, Galli C, Gregory PD et al (2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol* 25:1298–1306
5. Cai Y, Laustsen A, Zhou Y, Sun C, Anderson MV, Li S, Ulbjerg N, Luo Y, Jakobsen MR, Mikkelsen JG (2016) Targeted, homology-

- driven gene insertion in stem cells by ZFN-loaded ‘all-in-one’ lentiviral vectors. *Elife* 5:e12213
6. Wang J, Exline CM, DeClercq JJ, Llewellyn GN, Hayward SB, Li PW, Shivak DA, Surosky RT, Gregory PD, Holmes MC et al (2015) Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. *Nat Biotechnol* 33:1256–1263
 7. George LA, Sullivan SK, Giermasz A, Rasko JEJ, Samelson-Jones BJ, Ducore J, Cuker A, Sullivan LM, Majumdar S, Teitel J et al (2017) Hemophilia B gene therapy with a high-specific-activity factor IX variant. *N Engl J Med* 377:2215–2227
 8. Yin H, Xue W, Chen S, Bogorad RL, Benedetti E, Grompe M, Kotliansky V, Sharp PA, Jacks T, Anderson DG (2014) Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat Biotechnol* 32:551–553
 9. Cheng R, Peng J, Yan Y, Cao P, Wang J, Qiu C, Tang L, Liu D, Tang L, Jin J et al (2014) Efficient gene editing in adult mouse livers via adenoviral delivery of CRISPR/Cas9. *FEBS Lett* 588:3954–3958
 10. Ding Q, Strong A, Patel KM, Ng SL, Gosis BS, Regan SN, Cowan CA, Rader DJ, Musunuru K (2014) Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing. *Circ Res* 115:488–492
 11. Wang D, Mou H, Li S, Li Y, Hough S, Tran K, Li J, Yin H, Anderson DG, Sontheimer EJ et al (2015) Adenovirus-mediated somatic genome editing of Pten by CRISPR/Cas9 in mouse liver in spite of Cas9-specific immune responses. *Hum Gene Ther* 26:432–442
 12. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS et al (2015) In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* 520(7546):186–191
 13. Thakore PI, Kwon JB, Nelson CE, Rouse DC, Gemberling MP, Oliver ML, Gersbach CA (2018) RNA-guided transcriptional silencing in vivo with *S. aureus* CRISPR-Cas9 repressors. *Nat Commun* 9:1674
 14. Jarrett KE, Lee CM, Yeh YH, Hsu RH, Gupta R, Zhang M, Rodriguez PJ, Lee CS, Gillard BK, Bissig KD et al (2017) *Sci Rep* 7:44624
 15. Calcedo R, Somanathan S, Qin Q, Betts MR, Rech AJ, Vonderheide RH, Mueller C, Flotte TR, Wilson JM (2017) Class I-restricted T-cell responses to a polymorphic peptide in a gene therapy clinical trial for alpha-1-antitrypsin deficiency. *Proc Natl Acad Sci U S A* 114:1655–1659
 16. Chiuchiolo MJ, Kaminsky SM, Sondhi D, Hackett NR, Rosenberg JB, Frenk EZ, Hwang Y, Van de Graaf BG, Hutt JA, Wang G et al (2013) Intrapleural administration of an AAVrh.10 vector coding for human alpha-1-antitrypsin for the treatment of alpha-1-antitrypsin deficiency. *Hum Gene Ther Clin Dev* 24:161–173
 17. Flotte TR, Trapnell BC, Humphries M, Carey B, Calcedo R, Rouhani F, Campbell-Thompson M, Yachnis AT, Sandhaus RA, McElvaney NG et al (2011) Phase 2 clinical trial of a recombinant adeno-associated viral vector expressing alpha-1-antitrypsin: interim results. *Hum Gene Ther* 22:1239–1247
 18. Gruntman AM, Flotte TR (2017) Therapeutics: gene therapy for alpha-1 antitrypsin deficiency. *Methods Mol Biol* 1639:267–275
 19. Mueller C, Gernoux G, Gruntman AM, Borel F, Reeves EP, Calcedo R, Rouhani FN, Yachnis A, Humphries M, Campbell-Thompson M et al (2017) 5 Year expression and neutrophil defect repair after gene therapy in alpha-1 antitrypsin deficiency. *Mol Ther* 25:1387–1394
 20. Shen S, Sanchez ME, Blomenkamp K, Corcoran EM, Marco E, Yudkoff CJ, Jiang H, Teckman JH, Bumcrot D, Albright CF (2018) Amelioration of alpha-1 antitrypsin deficiency diseases with genome editing in transgenic mice. *Hum Gene Ther* 29(8):861–873
 21. Song CQ, Wang D, Jiang T, O’Connor K, Tang Q, Cai L, Li X, Weng Z, Yin H, Gao G et al (2018) In vivo genome editing partially restores alpha-1-antitrypsin in a murine model of AAT deficiency. *Hum Gene Ther* 29(8):853–860
 22. Singh K, Evens H, Nair N, Rincon MY, Sarcar S, Samara-Kuko E, Chuah MK, Vanden-Driessche T (2018) Efficient in vivo liver-directed gene editing using CRISPR/Cas9. *Mol Ther* 26:1241–1254
 23. Petris G, Casini A, Montagna C, Lorenzin F, Prandi D, Romanel A, Zasso J, Conti L, Demichelis F, Cereseto A (2017) Hit and go CAS9 delivered through a lentiviral based self-limiting circuit. *Nat Commun* 8:15334
 24. Ruan GX, Barry E, Yu D, Lukason M, Cheng SH, Scaria A (2017) CRISPR/Cas9-mediated genome editing as a therapeutic approach for leber congenital amaurosis 10. *Mol Ther* 25:331–341
 25. Merienne N, Vachey G, de Longprez L, Meunier C, Zimmer V, Perriard G,

- Canales M, Mathias A, Herrgott L, Beltrami-nelli T et al (2017) The self-inactivating Kami-Cas9 system for the editing of CNS disease genes. *Cell Rep* 20:2980–2991
26. Nelson CE, Hakim CH, Ousterout DG, Thakore PI, Moreb EA, Castellanos Rivera RM, Madhavan S, Pan X, Ran FA, Yan WX et al (2016) In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* 351:403–407
 27. Tabebordbar M, Zhu K, Cheng JK, Chew WL, Widrick JJ, Yan WX, Maesner C, Wu EY, Xiao R, Ran FA et al (2016) In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* 351:407–411
 28. Long C, Amoasii L, Mireault AA, McAnally JR, Li H, Sanchez-Ortiz E, Bhattacharyya S, Shelton JM, Bassel-Duby R, Olson EN (2016) Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* 351:400–403
 29. Bengtsson NE, Hall JK, Odom GL, Phelps MP, Andrus CR, Hawkins RD, Hauschka SD, Chamberlain JR, Chamberlain JS (2017) Muscle-specific CRISPR/Cas9 dystrophin gene editing ameliorates pathophysiology in a mouse model for Duchenne muscular dystrophy. *Nat Commun* 8:14454
 30. Amoasii L, Long C, Li H, Mireault AA, Shelton JM, Sanchez-Ortiz E, McAnally JR, Bhattacharyya S, Schmidt F, Grimm D et al (2017) Single-cut genome editing restores dystrophin expression in a new mouse model of muscular dystrophy. *Sci Transl Med* 9(418): ean8081
 31. Liao HK, Hatanaka F, Araoka T, Reddy P, Wu MZ, Sui Y, Yamauchi T, Sakurai M, O'Keefe DD, Nunez-Delgado E et al (2017) In vivo target gene activation via CRISPR/Cas9-mediated trans-epigenetic modulation. *Cell* 171:1495–1507. e1415
 32. Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H et al (2015) Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 517:583–588
 33. Bakondi B, Lv W, Lu B, Jones MK, Tsai Y, Kim KJ, Levy R, Akhtar AA, Breunig JJ, Svendsen CN et al (2016) In vivo CRISPR/Cas9 gene editing corrects retinal dystrophy in the S334ter-3 rat model of autosomal dominant retinitis pigmentosa. *Mol Ther* 24:556–563
 34. Giannelli SG, Luoni M, Castoldi V, Massimino L, Cabassi T, Angeloni D, Demontis GC, Leocani L, Andreazzoli M, Broccoli V (2018) Cas9/sgRNA selective targeting of the P23H Rhodopsin mutant allele for treating retinitis pigmentosa by intravitreal AAV9.PHP.B-based delivery. *Hum Mol Genet* 27:761–779
 35. Li P, Kleinstiver BP, Leon MY, Prew MS, Navarro-Gomez D, Greenwald SH, Pierce EA, Keith Joung J, Liu Q (2018) Allele-specific CRISPR-Cas9 genome editing of the single-base P23H mutation for rhodopsin-associated dominant retinitis pigmentosa. *CRISPR J* 1:55–64
 36. Yu W, Mookherjee S, Chaitankar V, Hiriyanna S, Kim JW, Brooks M, Ataeijannati Y, Sun X, Dong L, Li T et al (2017) Nrl knockdown by AAV-delivered CRISPR/Cas9 prevents retinal degeneration in mice. *Nat Commun* 8:14716
 37. Moreno AM, Fu X, Zhu J, Katrekar D, Shih YV, Marlett J, Cabotaje J, Tat J, Naughton J, Lisowski L et al (2018) In situ gene therapy via AAV-CRISPR-Cas9-mediated targeted gene regulation. *Mol Ther* 26(7):1818–1827
 38. Latella MC, Di Salvo MT, Cocchiarella F, Benati D, Grisendi G, Comitato A, Marigo V, Recchia A (2016) In vivo editing of the human mutant rhodopsin gene by electroporation of plasmid-based CRISPR/Cas9 in the mouse retina. *Mol Ther Nucleic Acids* 5:e389
 39. Huang X, Zhou G, Wu W, Duan Y, Ma G, Song J, Xiao R, Vandenberghe L, Zhang F, D'Amore PA et al (2017) Genome editing abrogates angiogenesis in vivo. *Nat Commun* 8:112
 40. Kim K, Park SW, Kim JH, Lee SH, Kim D, Koo T, Kim KE, Kim JH, Kim JS (2017) Genome surgery using Cas9 ribonucleoproteins for the treatment of age-related macular degeneration. *Genome Res* 27:419–426
 41. Holmgaard A, Askou AL, Benckendorff JNE, Thomsen EA, Cai Y, Bek T, Mikkelsen JG, Corydon TJ (2017) In vivo knockout of the Vegfa Gene by lentiviral delivery of CRISPR/Cas9 in mouse retinal pigment epithelium cells. *Mol Ther Nucleic Acids* 9:89–99
 42. Jain A, Zode G, Kasetti RB, Ran FA, Yan W, Sharma TP, Bugge K, Searby CC, Fingert JH, Zhang F et al (2017) CRISPR-Cas9-based treatment of myocilin-associated glaucoma. *Proc Natl Acad Sci U S A* 114:11199–11204
 43. Kim E, Koo T, Park SW, Kim D, Kim K, Cho HY, Song DW, Lee KJ, Jung MH, Kim S et al (2017) In vivo genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni*. *Nat Commun* 8:14500

44. Montana CL, Kolesnikov AV, Shen SQ, Myers CA, Kefalov VJ, Corbo JC (2013) Reprogramming of adult rod photoreceptors prevents retinal degeneration. *Proc Natl Acad Sci U S A* 110:1732–1737
45. Zuris JA, Thompson DB, Shu Y, Guilinger JP, Bessen JL, Hu JH, Maeder ML, Joung JK, Chen ZY, Liu DR (2015) Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat Biotechnol* 33:73–80
46. Gao X, Tao Y, Lamas V, Huang M, Yeh WH, Pan B, Hu YJ, Hu JH, Thompson DB, Shu Y et al (2018) Treatment of autosomal dominant hearing loss by in vivo delivery of genome editing agents. *Nature* 553:217–221
47. Chew WL, Tabebordbar M, Cheng JK, Mali P, Wu EY, Ng AH, Zhu K, Wagers AJ, Church GM (2016) A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat Methods* 13:868–874
48. Mortensen R, Nissen TN, Blauenfeldt T, Christensen JP, Andersen P, Dietrich J (2015) Adaptive immunity against *Streptococcus pyogenes* in adults involves increased IFN-gamma and IgG3 responses compared with children. *J Immunol* 195:1657–1664
49. Colque-Navarro P, Jacobsson G, Andersson R, Flock JI, Mollby R (2010) Levels of antibody against 11 *Staphylococcus aureus* antigens in a healthy population. *Clin Vaccine Immunol* 17:1117–1123
50. Dryla A, Prustomersky S, Gelbmann D, Hanner M, Bettinger E, Kocsis B, Kustos T, Henics T, Meinke A, Nagy E (2005) Comparison of antibody repertoires against *Staphylococcus aureus* in healthy individuals and in acutely infected patients. *Clin Diagn Lab Immunol* 12:387–398
51. Charlesworth CT, Deshpande PS, Dever DP, Dejene B, Gomez-Ospina N, Mantri S, Pavel-Dinu M, Camarena J, Weinberg KI, Porteus MH (2018) Identification of pre-existing adaptive immunity to Cas9 proteins in humans. *BioRxiv*. <https://doi.org/10.1101/243345>
52. Al-Mashhadi RH, Sorensen CB, Kragh PM, Christoffersen C, Mortensen MB, Tolbod LP, Thim T, Du Y, Li J, Liu Y et al (2013) Familial hypercholesterolemia and atherosclerosis in cloned minipigs created by DNA transposition of a human PCSK9 gain-of-function mutant. *Sci Transl Med* 5:166ra161
53. Cai Y, Bak RO, Mikkelsen JG (2014) Targeted genome editing by lentiviral protein transduction of zinc-finger and TAL-effector nucleases. *Elife* 3:e01911
54. Choi JG, Dang Y, Abraham S, Ma H, Zhang J, Guo H, Cai Y, Mikkelsen JG, Wu H, Shankar P et al (2016) Lentivirus pre-packed with Cas9 protein for safer gene editing. *Gene Ther* 23:627–633
55. Skipper KA, Nielsen MG, Andersen S, Ryo LB, Bak RO, Mikkelsen JG (2018) Time-restricted PiggyBac DNA transposition by transposase protein delivery using lentivirus-derived nanoparticles. *Mol Ther Nucleic Acids* 11:253–262
56. Yin H, Song CQ, Suresh S, Wu Q, Walsh S, Rhym LH, Mintzer E, Bolukbasi MF, Zhu LJ, Kauffman K et al (2017) Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing. *Nat Biotechnol* 35:1179–1187



CRISPR Gene Therapy of the Eye: Targeted Knockout of *Vegfa* in Mouse Retina by Lentiviral Delivery

Andreas Holmgaard, Sidsel Alsing, Anne Louise Askou, and Thomas J. Corydon

Abstract

Genome editing and knockout by virus-based delivery of CRISPR/Cas9 may provide a new option to cure inherited and acquired ocular diseases. Here we describe development and application of lentivirus-based delivery vectors enabling knockout of the *Vegfa* gene. We show that *Streptococcus pyogenes* (Sp) Cas9 and single-guide RNAs (sgRNAs) delivered by such vectors selectively can ablate the vascular endothelial growth factor A (*Vegfa*) gene in mouse retina following a single administration. These findings may contribute to the development of a new therapeutic path in the treatment of ocular diseases including exudative age-related macular degeneration (AMD).

Key words DNA cloning, Lentiviral vector, CRISPR/Cas9, miRNA, Antiangiogenesis, AMD, VEGF, Retina, RPE

1 Introduction

Successful interruption of functional genes in living cells depends on reliable tools allowing efficient and targeted knockout (KO) within the genome. CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9) has accomplished in a very short time to become one of the most important scientific breakthroughs. This machinery holds the potential to treat severe diseases through molecular surgery of the genome. Based on our recent findings [1] we here describe in detail the generation and application of virus-based gene therapy by CRISPR/Cas9-mediated gene KO as a new option for treatment of inherited and acquired ocular diseases of the retina.

To overcome the limitations of a “two-vector design,” the generated lentiviral vectors (LVs), simultaneously encode SpCas9 and sgRNAs targeting murine *Vegfa*, as well as the fluorescent eGFP marker protein. Using this approach, we demonstrated robust KO of *Vegfa* and subsequent reduction of VEGFA protein

in transduced cells. The top three of the designed sgRNAs induced *in vitro* indel formation at a high frequency, ranging from 44 to 93%. Notably, the amount of VEGFA protein was significantly reduced in these cells. A single unilateral, subretinal injection of LVs resulted in disruption of *Vegfa* with an *in vivo* indel efficacy of up to 84% in eGFP⁺ retinal pigment epithelial (RPE) cells isolated five weeks post LV-delivery. Sequencing of *Vegfa*-specific amplicons demonstrated the spectrum of indels, ranging from 4-bp deletions to 2-bp insertions. Since dysregulation of the *Vegfa* gene, involved in abnormal angiogenesis, is linked to exudative AMD [2–4], the LV-based approach for delivery of editing and KO tools to retinal cells described in this chapter has potential in amelioration of genes involved in ocular diseases, including exudative AMD.

The LVs were generated by a simple cloning step, where designed sgRNAs were inserted into an RNA scaffold under the control of a U6 promoter in the lentiCRISPRv2 plasmid as depicted in Fig. 1a–c. Simultaneously, the LV constructs encode SpCas9 under the control of the elongation factor 1 α short (EFS) promoter. To ease visualization of CRISPR/Cas9 expressing cells the lentiCRISPRv2 vector harbors the DNA sequence of a FLAG tag fused to the SpCas9 cDNA sequence followed by DNA sequences encoding a P2A self-cleaving peptide and the eGFP marker protein (Fig. 1b) [1]. Broad Institute CRISPRko and MIT CRISPR Design software were used to validate on-target and off-target efficacy, respectively. In addition to the cloning design (Fig. 1a, b) and LV production (Fig. 1c), the figure illustrates the functional validation of the various sgRNAs in transduced cells and in the mouse retina following subretinal injection of LVs (Fig. 1d–o) [1]. In order to identify *in vivo* KO events in mouse retinal cells it was necessary to include a fluorescence-activated cell sorting (FACS) step enabling isolation of eGFP⁺ cells. Efficient KO of the *Vegfa* gene was documented five weeks post injection in eGFP⁺ cells. Representative examples of KO events assessed by Surveyor assay or Tracking of Indels by Decomposition (TIDE) analysis are depicted in Fig. 2a–d, respectively. In conclusion, we have generated LVs for delivery of CRISPR/Cas9 to the retina that may contribute to the development of new KO and editing therapies for amelioration of diseases involving intraocular neovascularization.

2 Materials

Prepare all solutions using ultrapure water.

2.1 Cloning

1. *BsmBI* restriction enzyme.
2. Calf intestinal alkaline phosphatase (CIP).
3. Compatible enzyme reaction buffers.

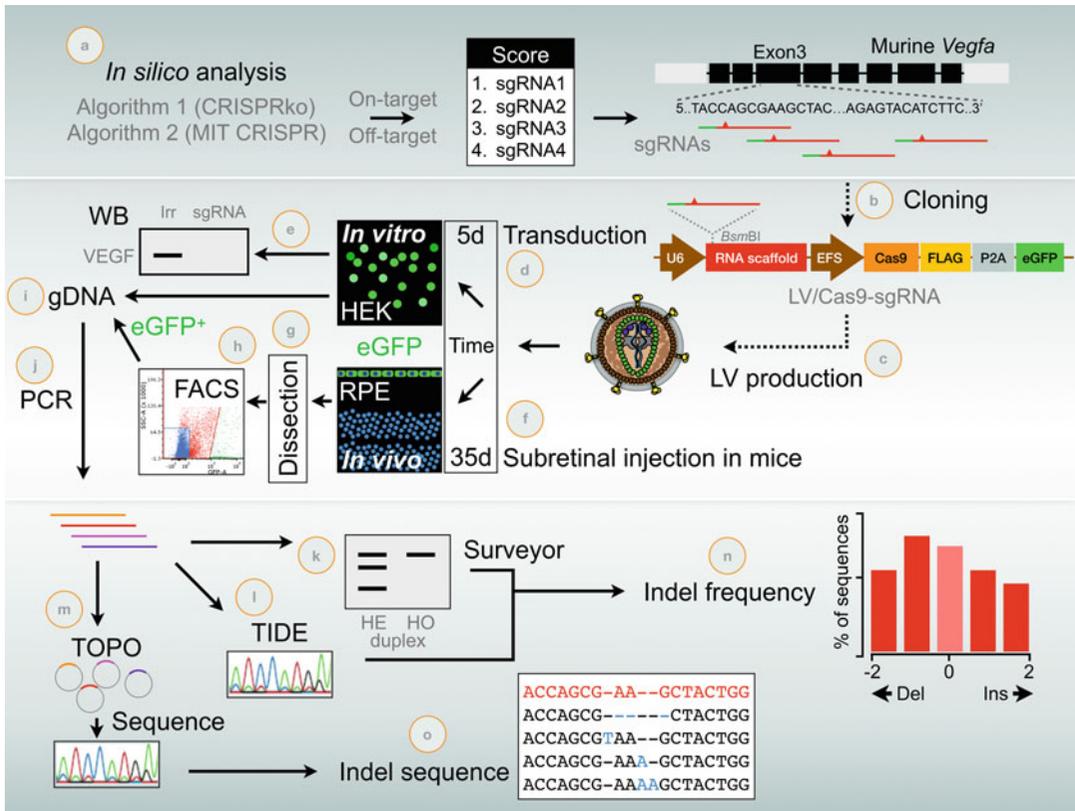


Fig. 1 Schematic depiction of cloning design and functional validation of the various sgRNAs. (a) In silico evaluation using CRISPRko and MIT CRISPR on-line tools for on-target and off-target analysis, respectively. (b) Cloning of annealed oligos representing the various sgRNAs (selected in (a)) into the *BsmBI* restriction site of the LV/Cas9-sgRNA vector. (c) Production of LVs. (d) Transduction of HEK293-Vegfa cells. (e) Assessment of the VEGFA protein level in HEK293-Vegfa cells by means of Western blot analysis. (f) Delivery of LVs to RPE cells of the retina following subretinal injection. (g) Dissection of retinal tissue and isolation of RPE cells in single cell suspension. (h) FACS analysis resulting in the isolation of eGFP⁺ RPE cells harboring the LV/Cas9-sgRNA vector, which thus also express Cas9 and relevant sgRNAs. (i) Isolation of gDNA from HEK293-Vegfa and RPE cells. (j) Amplification of *Vegfa* sequences (exon 3) using relevant primers (see Table 1). (k) Surveyor analysis of *Vegfa* amplicons. A representative fragment pattern is depicted. An illustrative example of obtained *Vegfa* KO results using Surveyor assay is shown in Fig. 2a. (l) Sequencing and TIDE analysis of *Vegfa* amplicons. (m) TOPO cloning of *Vegfa* amplicons followed by DNA sequencing. (n) Calculation of indel frequency following Surveyor analysis using the equation shown in Fig. 2b (see also Note 7). Regarding TIDE, the indel frequency was determined by sequence-based analysis (see also Note 8). Illustrative examples of *Vegfa* KO results obtained with TIDE analysis are shown in Fig. 2c, d. (o) Assessment of indel sequence following TOPO cloning and sequence analysis. *Cas9* CRISPR-associated protein 9, *Del* deletion, *eGFP* enhanced green fluorescent protein, *eGFP*⁺ eGFP positive, *EFS* elongation factor 1 α short promoter (an RNA polymerase II promoter), *FACS* fluorescence-activated cell sorting, *FLAG* octapeptide tag, *gDNA* genomic DNA, *HE* hetero, *HEK* human embryonic kidney cells, *HO* homo, *Ins* insertion, *Irr* irrelevant (non-targeting) sgRNA, *P2A* sequence encoding a self-cleaving 2A peptide, *RPE* retinal pigment epithelial cells, *sgRNAs* single-guide RNAs, *U6* human U6 promoter (an RNA polymerase III promoter), *VEGF* vascular endothelial growth factor, *WB* Western blotting

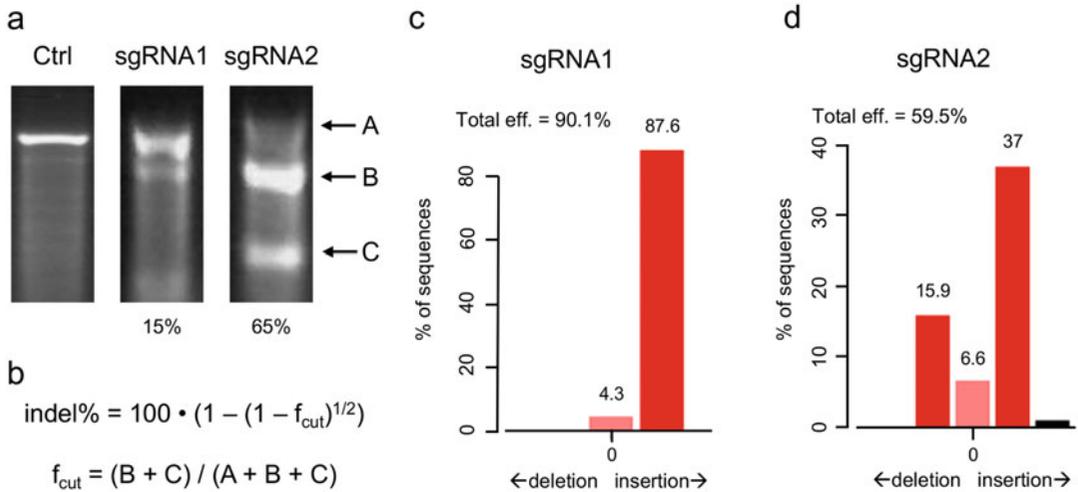


Fig. 2 Assessment of indel formation. **(a)** Representative examples of agarose gel patterns observed following Surveyor analysis of gDNA purified from cells treated with control sgRNA (Ctrl) or sgRNAs targeting *Vegfa* (sgRNA1–2). Based on the Surveyor analysis, sgRNA1 results in 15% indel formation whereas sgRNA2 produces up to 65% indel formation. The DNA band marked with A represents uncut amplicons whereas the B and C bands represent amplicons which have been cut. **(b)** The equation used to calculate indel frequency based on the Surveyor agarose gel shown in **(a)**. See also **Note 7**. **(c, d)** Assessment of indels formed using TIDE analysis following treatment with sgRNA1 or sgRNA2. **(c)** In contrast to the result from the Surveyor analysis sgRNA1 displays high indel formation efficacy following TIDE analysis. See also **Notes 5–7**. **(d)** TIDE analysis showing high indel formation efficacy of sgRNA2 resembling the result from the Surveyor analysis. See also **Notes 5–7**. *Ctrl* control, *eff* efficiency

4. Loading dye for subsequent gel analysis.
5. LentiCRISPRv2-eGFP plasmid.
6. Designed sgRNAs oligos.

2.2 Gel Purification

1. 1× TAE: BIO-RAD 50× TAE buffer, diluted in water.
2. 7 × 10 cm UV-transparent agarose gel trays.
3. Preparative comb for agarose gel trays (containing two preparative and two reference wells).
4. 8- and 15-well combs for agarose gel trays.
5. Preparative agarose gel (0.7%): Prepare a 0.7% agarose gel by adding 0.35 g agarose to 50 mL 1% TAE and heat the solution to boiling. Leave the solution to cool to approx. 60 °C and add 1.5 μL GelRed or other nucleic acid gel stain. Cast the gel within a 7 × 10 cm UV-transparent gel tray and immediately insert a preparative comb (containing two preparative and two reference wells) without introducing air bubbles.
6. Agarose gel (1%): Prepare a 1% agarose gel by adding 0.5 g agarose to 50 mL 1% TAE and heat the solution to boiling. Leave the solution to cool to 60 °C and add 1.5 μL GelRed or

other nucleic acid gel stain. Cast the gel within a 7×10 cm UV-transparent gel tray and immediately insert 1 or 2 (depending on the number of samples to be analyzed) 15-well combs without introducing air bubbles.

7. DNA ladder.
8. QIAquick or other gel DNA extraction kit.
9. Gel loading buffer.
10. Nanodrop spectrophotometer.
11. Scalpel.

2.3 Ligation

1. T4 DNA ligase.
2. Compatible reaction buffer for the T4 DNA ligase.
3. pCR-Blunt-II-TOPO vector.

2.4 Transformation

1. *Escherichia coli* (*E. coli*) XL2-Blue or other ultracompetent cells.
2. β -Mercaptoethanol.
3. LB medium: Mix 10 g tryptone, 5 g yeast extract, and 10 g NaCl with 800 mL water. Adjust pH to 7.4 and bring the final volume up to 1 L. Sterilize by autoclaving.
4. LB ampicillin medium: LB medium containing 0.1 mg/mL ampicillin.
5. SOC-medium: Mix 19 μ L 1 M MgSO_4 , 19 μ L 1 M MgCl_2 , and 38 μ L 20% glucose with 1.8 mL LB medium in a sterile environment. Prepare fresh.
6. 9 cm diameter Petri dish.
7. LB agar ampicillin plates: Add 15 g agar to 800 mL LB medium and bring final volume up to 1 L utilizing LB medium. Autoclave to dissolve agar and sterilize. Let the LB agar solution cool to approx. 45 °C and add 0.1 mg/mL ampicillin. Swirl the flask gently to mix and avoid air bubbles. Pour 10 mL LB agar ampicillin solution into a Petri dish and leave at room temperature (RT) to solidify.
8. Sterile inoculation needles.

2.5 PCR

1. DNA sequences of PCR primers are listed in Table 1. Primer-1 and primer-2 were used for amplification of *Vegfa* cDNA sequence in HEK293-Vegfa cells (*see Note 1*). Primer-3 and primer-4 were used for amplification of *Vegfa* exon 3 from gDNA isolated from RPE. Primer-5 and primer-6 were used for verification of correct insertion of sgRNA in the lentiCRISPRv2-eGFP plasmid.

Table 1
Primers

ID	Name	Sequence (5'-3')	Annealing temp. (°C)	Amplicon size (bp)
1	<i>mVegfa</i> cDNA Fwd.	ATGAACTTTCTGCTCTCTTG	61	567
2	<i>mVegfa</i> cDNA Rev.	CCTTGGCTTGTCACATCT		
3	<i>mVegfa</i> exon 3 Fwd.	AAAGGTCACGAAAGCAGATGG	66	777
		TCAA		
4	<i>mVegfa</i> exon3 Rev.	GTGTATATACATAGCTG		
		TCCCCGG		
5	U6 Fwd.	CCCATGATTTCCTTCATATTTGC	63	349
6	lentiCRISPRv2 Rev.	TTCCCACTCCTTTCAAGACC		

mVegfa murine Vegfa, *Fwd* forward, *Rev* reverse

2. Phusion High-Fidelity DNA Polymerase or another high-fidelity DNA polymerase and compatible reaction buffer.
3. 100% DMSO.
4. PCR tubes.
5. MinElute or other PCR DNA fragment purification kit.
6. Redistilled water.

2.6 Lentivirus Production

1. p15 dish.
2. Plasmids: pRSV-REV, pMD-2G and pMDLg/pRRE for third-generation LV production.
3. CaCl₂.
4. HEPES buffer solution.
5. Beckmann tubes.
6. Ultracentrifuge.
7. p24 ELISA assay or another measure for titer estimation [5].
8. Polybrene for LV transduction of cells.
9. Hank's balanced salt solution (HBSS).

2.7 Cell Cultivation

1. HEK293-T cells.
2. HEK293-Vegfa cells.
3. Dulbecco's Modified Eagle Medium, DMEM.
4. Fetal calf serum, FCS.
5. Glutamine.
6. Penicillin/streptomycin.
7. Trypsin.
8. PBS.

9. PBS-EDTA.
10. Cell cultivation flasks.
11. 6-well plates.
12. RIPA lysis buffer or another cell lysis buffer.
13. cOmplete™ Mini or another protease inhibitor.

2.8 Animal Experiments

2.9 Subretinal Injections

1. C57BL/6J mice.
1. Carprofen.
2. Ketamine (Ketador).
3. Medetomidin (Cepetor).
4. 1% tropicamide solution (Mydricyl®).
5. Metaoxedrin 10%.
6. Carbomer eye gel (Viscotears®).
7. Microscope (e.g., Zeiss OPMI 1 FR pro).
8. Coverslips.
9. 5 µL Hamilton syringe.
10. Dumont style 3C Dumoxel tweezers.
11. 34-gauge (G) needle.
12. 0.5–1 mg/kg Antisedan/Alzane.
13. Heating pad.

2.10 Dissection and Isolation of RPE Cells for FACS

1. Dumont style 3C Dumoxel tweezers.
2. Vannas scissors.
3. Dispensable pipettes.
4. 2.5 mL syringe.
5. 27-G needle.
6. Petri dish.
7. PBS.
8. HBSS.
9. Trypsin/EDTA.
10. Filter papers.

2.11 Immunostaining of Cryosections

1. Optimal cutting temperature (OCT) compound.
2. Cryostat.
3. DAPI (Sigma-Aldrich), 2 µg/µL.
4. PBS.
5. Triton X-100.

6. Normal goat serum.
7. ProLong™ Gold antifade reagent.
8. Paraformaldehyde (PFA).

2.12 Western Blot Analysis

1. Mini-PROTEAN® Tetra Vertical Electrophoresis Cell (BIO-RAD).
2. Tris buffered saline with 0.1% Tween20 (TBS-T).
3. Tris/Glycine/SDS (TGS).
4. Mini-Protean® TGX™ Precast Gels (BIO-RAD) or other gels for Western blotting.
5. Trans-Blot® Turbo™ Mini PVDF Transfer Packs (BIO-RAD) or other membranes for Western blotting.
6. Trans-Blot® Turbo™ Transfer System (BIO-RAD) or another blotting system.
7. Clarity™ Western ECL Substrate (BIO-RAD) or another visualization substrate for Western blotting.
8. ImageQuant LAS 4000 mini or another imaging system.
9. Polyclonal rabbit anti-VEGFA antibody.
10. Polyclonal Goat Anti-Rabbit Immunoglobulins/HRP conjugate.
11. Protein ladder.

2.13 Indel Analysis

1. Surveyor® Mutation Detection Kit (Integrated DNA Technologies) or another mismatch detection kit.
2. TIDE software: <https://tide-calculator.nki.nl/>.
3. Zero Blunt PCR Cloning Kit (Thermo Fischer) or another kit for TOPO cloning.

3 Methods

Carry out all procedures at RT unless otherwise specified.

3.1 Selection of Single-Guide RNAs for spCas9 Targeting of Murine Vegfa

1. Obtain the genomic sequence of the murine *Vegfa* gene, i.e., by searching GenBank (NC_000083).
2. Identify target locus: Determine exon-intron structure and identify which splice variants are expressed in the target tissue. Alternative splicing variants of *Vegfa* share exon 1–5. *Vegfa* exon 3 was chosen as it is an early exon and it is the largest exon.
3. Predict sgRNA sequences using sgRNA prediction software. Broad Institute CRISPRko and MIT CRISPR Design sgRNA were used to survey *Vegfa* exon 3. On this basis 4 sgRNAs designated sgRNA1–4 were designed (*see Note 2*).

4. Estimate on-target efficacy by CRISPRko, as described in [6]. In this case sgRNA2 and sgRNA3 were selected for further analysis on this basis.
5. Evaluate off-targeting by MIT CRISPR Design, as described in [7]. On this basis sgRNA1 and sgRNA4 were selected for further analysis.
6. Order sgRNA sequences, for example, at Eurofins, as oligonucleotides with a 5'-CACC-3' or 5'-AAAC-3' addition to the 5' end of the sense and antisense strand, respectively. These overhangs are compatible with the *Bsm*BI cut-sites in lentiCRISPRv2-eGFP.

**3.2 Cloning
of sgRNAs into
lentiCRISPRv2-egfp
Plasmid**

1. Mix 5 μ g of lentiCRISPRv2-egfp with reaction buffer and 5 units (U) of *Bsm*BI restriction enzyme in a total volume of 100 μ L and place at 37 °C overnight to linearize the vector (*see Note 3*).
2. Add 20 U of CIP to the linearized vector reaction mixture and leave for 15 min at 37 °C.
3. Add loading dye to the mixture and mix gently.
4. Purify the linearized vector using a 0.7% preparative agarose gel. Place the gel in a submerged horizontal electrophoresis cell and cover with 1 \times TAE buffer.
5. Apply the restriction digest mixture to the well and run the gel at low voltage (40–60 V) for 1–2 h to ensure separation of digested plasmid from undigested plasmid. This step also ensures that the restriction enzymes are removed from the DNA.
6. Briefly verify band separation utilizing a UV transilluminator.
7. Cut out the DNA band representing the linearized lentiCRISPRv2-egfp plasmid with a scalpel.
8. Purify the DNA fragment from the agarose gel by applying the QIAquick or another gel extraction kit according to the supplied protocol.
9. Measure DNA concentration, e.g., by applying 2 μ L of the DNA solution on a nanodrop spectrophotometer according to the user's protocol.
10. Mix 400 ng of each pair of oligonucleotide sgRNA strands in a total volume of 40 μ L. Transfer annealed fragments to a 100 °C water bath and allow to cool to RT. Annealing can also be performed using a PCR machine.
11. Make separate ligation mixtures for each sgRNA sequence. Mix 6 μ L of annealed oligonucleotides with 100 ng linearized vector plasmid.

12. Add 1 U T4 DNA ligase and ligase buffer in a total volume of 20 μL .
13. As negative control, replace annealed oligonucleotides with water.
14. Mix gently by pipetting and incubate for 1 h at RT.
15. Thaw ultracompetent cells, e.g., XL2-Blue cells on ice (*see Note 4*).
16. Aliquot 25 μL cells into each distinct tube on ice for each sgRNA.
17. Add 0.5 μL β -mercaptoethanol and mix gently by tabbing the tubes.
18. Incubate the cells for 10 min on ice, gently swirling the tubes every 2 min.
19. Add 1.5 μL of each ligation mixture to the distinct tubes.
20. Gently mix by pipetting and incubate on ice for 30 min.
21. Heat shock the cells by submerging the tubes in a 42 $^{\circ}\text{C}$ water bath for 30 s.
22. Incubate on ice for 2 min.
23. Add 225 μL of freshly made SOC-medium to each tube and incubate at 37 $^{\circ}\text{C}$ in a shaking incubator (225–250 rpm) for 1 h.
24. Plate 50 μL of each transformation mixture on separate LB agar ampicillin plates.
25. Incubate at 37 $^{\circ}\text{C}$ overnight.
26. Compare number of colonies on each sgRNA transformation plate with the negative control plate. Several colonies should be identified on the plates, while no colonies should appear on the negative control as the vector should not be able to reanneal.
27. Select a number of colonies to be screened for insertion of the sgRNA.
28. Use a sterile inoculation needle to transfer cells from a single colony to a small tube containing 2 mL LB ampicillin medium.
29. Incubate the culture overnight (12–16 h) at 37 $^{\circ}\text{C}$ with shaking at 225–250 rpm.
30. Purify plasmid DNA from approximately 1 mL of each of the cultures utilizing the QIAprep Spin or another Miniprep kit according to the manufacturer's protocol.
31. PCR amplification across the sgRNA insert: For each colony to be tested, mix 2 μL of purified plasmid with 5 pmol of U6 fwd and lentiCRISPRv2 rev primers (Table 1, primer-5 and primer-6) with 10 μL Phusion High-Fidelity DNA Polymerase and

Table 2
PCR program for thermocycler used for amplification of gDNA

Cycles	Temperature (°C)	Time
1	98	1 min
35–40 ^a	98	10 s
	61, 63, 66 ^b	25 s
	72	1 min
1	72	7 min

^a35 cycles were used for HEK293-Vegfa gDNA amplification and for confirmation of sgRNA insertion in the lentiCRISPRv2-egfp plasmid. 40 cycles were used for amplification of gDNA from FACS isolated RPE cells

^bAnnealing temperatures of 61 °C were used for HEK293-Vegfa gDNA amplification, 63 °C were used for confirmation of sgRNA insertion and 66 °C were used for amplification of gDNA from FACS isolated RPE cells

0.6 µL 100% DMSO in a PCR tube. Bring total volume to 20 µL with water.

32. Run PCR according to the program presented in Table 2 (63 °C annealing temperature).
33. Mix the amplified PCR product with loading dye and analyze over a 1% agarose gel. Place the gel in a submerged horizontal electrophoresis cell using 1× TAE buffer.
34. Apply the PCR products to the wells and run the gel at medium voltage (70–80 V) for 1 h. Uncut lentiCRISPRv2-egfp plasmid may be used as negative control.
35. Use a UV transilluminator to verify correct band size. The PCR products with correct sgRNA insertion should be 349 bp in length.
36. Select positive clone based on correct band size and sequence PCR products to verify correct sgRNA insertion.
37. The remaining 1 mL of each of the cultures is added to 100 mL LB ampicillin medium and incubated for 12–16 h at 37 °C with shaking at 225–250 rpm before plasmid DNA extraction, i.e., by Qiagen Plasmid Midi Kit or another plasmid purification kit.

3.3 Lentivirus Production

Lentiviral particles are produced according to the procedure described in [1, 5, 8]. A detail method for product of lentivirus can be found in this book serial by Ryø et al. In brief:

1. Day one: 1×10^7 HEK-293 T cells are seeded in a p15 dish. The number of p15 dishes varies with experimental purpose. In vitro use requires only one p15 and ultracentrifugation is optional. In vivo application typically requires >9 p15 dishes and requires ultracentrifugation for concentration of LV. The cells are maintained in DMEM supplemented with FCS,

glutamine, and penicillin/streptomycin (hereafter named DMEM-complete).

2. Day two: Mix 7.26 μg of pRSV-REV, 9.07 μg of pMD-2G, 31.46 μg of pMDI μ /pRRE, and 31.46 μg of the cloned lentiCRISPRv2-egfp plasmid with water to a final volume of 1089 μL (*see Note 5*).
3. Add 121 μL 2.5 M CaCl_2 to the DNA, and pipet the mix into 1210 μL 2 \times HEPES buffer. Incubate for 15 min.
4. Add 2.42 mL dropwise to the p15 dish.
5. Day 3: Apply 17 mL fresh DMEM-complete medium to the cells.
6. Day 4: Harvest medium, filter in 0.22 μm filters, and spin on an ultracentrifuge to concentrate the LV preparation. Dissolve centrifuged LV particles in HBSS. The volume of HBSS depends on the desired concentration for subsequent experiments.
7. Estimate titer by a p24 ELISA assay, according to the manufacturer's protocol.

Following production of LV particles, functional analysis of the sgRNAs can be verified *in vitro* in transduced cells and *in vivo* in RPE cells of mouse retina after subretinal administration of LVs. The following sections describe such analyses. *See also [1]*.

3.4 Transduction of HEK293-Vegfa Cells

1. Day 1: Seed 200,000 HEK293-Vegfa cells in separate 6-well plates for each LV/Cas9-sgRNA LV preparation, including the irrelevant and a non-transduced control sample (Fig. 1d).
2. Day 2: Add fresh DMEM-complete containing 8 $\mu\text{g}/\text{mL}$ polybrene to the cells.
3. Transduce the cells by adding 500 ng of each LV preparation for the sgRNAs to be tested.
4. Day 3: Add fresh DMEM-complete.
5. Day 7: Validate transduction (5 days post transduction) by fluorescence analysis of eGFP signal from the LV preparations (Fig. 1d).
6. Harvest cells. Wash cells in PBS-EDTA and trypsinize by adding a suitable trypsin solution to the cells. Spin for 3 min at $240 \times g$, and discard the supernatant. Resuspend in PBS and repeat spin. Freeze the dry cell pellet at -20°C or proceed with genomic DNA purification.

3.5 Genomic DNA Purification and Amplification of the Murine Vegfa Gene

1. Purify gDNA from the cell pellets using DNeasy Blood and Tissue kit or another gDNA purification kit, following the manufacturer's protocol (Fig. 1i).

2. Measure DNA concentration, e.g., by applying 2 μL of the DNA solution on a nanodrop spectrophotometer according to the user's protocol.
3. Amplify the murine *Vegfa* gene by PCR on gDNA using the *mVegfa* cDNA fwd and rev primers (Table 1, primer-1 and primer-2).
4. Mix 100 ng gDNA with 5 pmol of forward and reverse primers, 10 μL Phusion High-Fidelity DNA Polymerase and 0.6 μL 100% DMSO in a PCR tube. Bring total volume to 20 μL with water.
5. Run PCR according to the PCR program in Table 2 (61 $^{\circ}\text{C}$ annealing temperature, *see* Table 2) (Fig. 1j).
6. Verify correct PCR product length on an agarose gel. The PCR product should be 567 bp in length.
7. Purify the PCR product using the MinElute PCR Purification Kit or another PCR purification kit according to the manufacturer's protocol.

3.6 Assessment of Genomic Indel Formation in HEK293-Vegfa Cells by Surveyor Nuclease Digestion Assay

Surveyor nuclease digestion assay was performed using the Surveyor Mutation Detection Kit (Integrated DNA Technologies) according to the manufacturer's protocol (*see* Note 6 and Fig. 1k).

1. Mix 200 ng of purified PCR product with water in a total of 8 μL in a PCR tube to allow for Heteroduplex formation.
2. Use a thermocycler to denature and reanneal DNA strands at 95 $^{\circ}\text{C}$ for 10 min, ramp down to 25 $^{\circ}\text{C}$ at 5 $^{\circ}\text{C}/\text{min}$ and 25 $^{\circ}\text{C}$ for 1 min.
3. Add 2 μL of Surveyor Nuclease S, 1 μL of Surveyor Enhancer S, and 1 μL MgCl_2 solution to the tube and incubate at 42 $^{\circ}\text{C}$ for 1 hour.
4. Following incubation, stop the Nuclease S by adding 1.4 μL of Surveyor Stop Solution and mix.
5. Add loading buffer to the tube and analyze using an agarose gel running at 70 V for \sim 1 h. *See* also Notes 6 and 7.
6. Analyze the gel using a UV transilluminator. *See* Fig. 1k for an example of a Surveyor assay agarose gel.
7. Quantify the DNA bands using ImageJ (imagej.nih.gov/ij/) [1].
8. Calculate indel frequency [9] using the following equation: $\text{indel}\% = 100 \cdot (1 - (1 - f_{\text{cut}})^{1/2})$, in which f_{cut} is the cleaved fraction of total DNA. *See* also Figs. 1n and 2a, b, and Note 7.

3.7 Assessment of Genomic Indel Formation in HEK293-Vegfa Cells by TIDE Analysis

1. Perform sequence analysis of purified PCR product with *mVegfa* cDNA fwd primer (Table 1, primer-1).
2. Apply the TIDE software (<https://tide.nki.nl/>) to assess indel formation in the sequences (Fig. 1l, n).

3. Insert sequenced control reference and edited sequences in the software tool together with the sgRNA sequence.
4. TIDE outputs an indel score and delivers an overview of the length of indels and an analysis of +1 bp insertions. *See* also Figs. 11, n and 2c, d, and **Note 8**.

3.8 Assessment of Functional Knockout by Western Blotting

1. Day 1: Seed 200,000 HEK293-Vegfa cells in separate 6-well plates for each LV/Cas9-sgRNA LV preparation, including the irrelevant and a non-transduced control sample.
2. Day 2: Add fresh DMEM-complete containing 8 µg/mL polybrene to the cells.
3. Transduce the cells by adding 500 ng of each LV preparation for the sgRNAs to be tested.
4. Day 3: Add fresh DMEM-complete the next day.
5. Day 7: Validate transduction (five days post transduction) by fluorescence analysis of eGFP signal from cells treated with LV preparations.
6. Wash cells 3× in PBS.
7. Apply 150 µL RIPA + cOmplete Mini or another lysis buffer containing protease inhibitors to each well.
8. Incubate on ice on a tilting board for 10–15 min.
9. Transfer the lysate to an Eppendorf tube.
10. Spin at 18,000 × *g* for 15 min at 4 °C.
11. Transfer supernatant to a clean Eppendorf tube.
12. Mix 800 µL water with 200 µL Protein Assay (BIO-RAD) for each sample to be measured (double measurements and standard curve).
13. Add (0, 1, 2.5, 5, 7.5, and 10 µg) BSA for the standard curve or 1–3 µL lysate for each sample and mix.
14. Measure absorbance at 595 nm in a spectrophotometer.
15. Calculate protein concentration according to the standard curve.
16. Mix samples of 15 µg protein, 4 µL 5× loading dye and PBS in a total of 20 µL.
17. Heat for 5 min at 100 °C.
18. Quick spin the samples.
19. Apply the samples (and an appropriate molecular marker) to a SDS-PAGE gel, i.e., 4–15% Mini-PROTEAN® TGX™ Precast Protein gels (BIO-RAD).
20. Run in a vertical electrophoresis cell in tris/glycine/SDS (TGS) running buffer (BIO-RAD) for 30 min at 200 V.

21. Use the Trans-Blot[®] Turbo[™] Transfer System to transfer proteins to a PVDF membrane.
22. Block with 5% w/v skimmed milk powder in TBS-T for 1 hour.
23. Wash 3 × 5 min with TBS-T.
24. Incubate with primary antibody against VEGFA (ab46154, abcam) 1:1000 in TBS-T overnight at 4 °C.
25. Repeat **step 23**.
26. Incubate with secondary goat-anti-rabbit HRP conjugated antibody (DAKO) 1:10,000 in TBS-T for 1 h.
27. Repeat **step 8**.
28. Visualize protein bands using the Clarity[™] Western ECL Substrate (BIO-RAD) for 5 min.
29. Use ImageQuant LAS 4000 mini to take pictures of the membranes. Graphic illustration of a representative Western blot is shown in Fig. 1e.
30. Repeat **steps 8–14** with anti-Histone-H3 antibody 1:10,000.
31. Use ImageJ to quantify and compare VEGFA and H3 protein bands.

3.9 Subretinal Injection of LV/Cas9-sgRNAs in C57BL/6J Mice

Subretinal injections were performed as previously described [1, 5, 8]. In brief:

1. Dilute LV preparations in HBSS to 70 ng/μL and place on ice.
2. Treat mice with NSAID carprofen prior to and following subretinal injections.
3. Anesthetize C57BL/6J mice with a combination of 60–100 mg/kg ketamine/medetomidine hydrochloride (Ketador) and 0.5–1 mg/kg Cepetor. Dilate pupils with a drop of 1% tropicamide solution (Mydracil[®]) and Metaoxedrin 10% prior to subretinal injection. Apply generous amounts of ViscoTears to the eyes and place a round coverslip (10 mm) on each eye.
4. Inject 2 μL subretinally in one eye with a Hamilton syringe 34-G needle while holding the eye with a Dumont style 3C Dumoxel tweezer.
5. Remove Viscotears from the fur and whiskers.
6. Inject mice with 0.5–1 mg/kg Antisedan and place on the stomach on a heating pad for waking.

3.10 Cryosectioning of LV/Cas9-sgRNA Injected Retinas

Analyze mice for eGFP expression in the RPE cells 35 days post injection.

1. Euthanize LV/Cas9-sgRNA treated mice by cervical dislocation.

2. Enucleate treated eyes and transfer them to 4% PFA for 5 min.
3. Wash briefly in PBS and place in a Petri dish with a filter paper soaked in PBS.
4. Remove excess tissue surrounding the eye and remove lens and cornea by dissection and fixate eyes for 2 h in 4% PFA at RT [5].
5. Wash eyes in PBS and incubate in a 30% sucrose solution ON at 4 °C.
6. Orientate and transfer eyes to tubes containing OCT compound and freeze on dry ice.

3.11

Immunostaining of Mouse Retina Sections

1. Thaw slides.
2. Rehydrate slides in PBS for 5 min.
3. Block in 100 μ L of PBS + 10% NGS + 1% triton for 2 h in a humidity chamber. Incubate slides covered with Parafilm.
4. Remove blocking buffer.
5. Add 2 μ L DAPI (2 μ g/ μ L) to 50 mL PBS and place the slides in the DAPI solution for 3 min.
6. Wash 2 \times 5 min in redistilled water protected from light.
7. Dip in 70% Ethanol and let the slides dry upright for 20–30 min protected from light.
8. Add 1–2 drops of ProLong™ Gold antifade reagent to each slide and mount cover glass. Analyze slides using a confocal laser scanning microscope. *See* also Fig. 1f and Note 9.

3.12 *In Vivo* Genome Editing

To ensure a sufficient number of dissected RPE cells in which KO has occurred, we utilized the fluorescent eGFP marker protein simultaneously expressed from the LV/Cas9 vector. eGFP⁺ RPE cells can be isolated by FACS and subsequently analyzed for KO events following purification of gDNA. Dissect murine retina 35 days post injection and isolate RPE for evaluation of indel formation (Fig. 1g).

1. Euthanize LV/Cas9-sgRNA injected mice by cervical dislocation.
2. Enucleate LV-treated eyes and transfer them to a Petri dish containing ice-cold HBSS.
3. Remove excess tissue surrounding the eye with vannas scissors.
4. Open eyes by cutting along ora serrata.
5. Remove the lens and cornea.
6. Remove the neuroretina.
7. Place the eye cup in ice-cold HBSS containing 1% trypsin/EDTA (2 mg/mL) at 37 °C for 1 h.

8. Flush the RPE cells of the eye cup with ice-cold HBSS using a 2.5-mL syringe with a 27-G needle.
9. Stop the trypsin digestion by adding 0.5 mL FCS.
10. Spin down cells at $500 \times g$ at 4 °C for 3 min. Discard supernatant.
11. Resuspend pellet in 1 mL HBSS containing 1% trypsin/EDTA (2 mg/mL) and place at 37 °C for 1 h.
12. Spin down cells at $500 \times g$ at 4 °C for 3 min. Remove supernatant.
13. Resuspend pellet in ice-cold HBSS and vortex gently. The resulting single cell suspension is ready for FACS analysis (*see below*).
14. Pass single cell suspension of murine RPE cells dissected from the retina through a 50 μ m filter to avoid large clumps of cells in the FACS machine.
15. Perform FACS analysis without discarding cells due to aggregation or cell size (Fig. 1h). *See Note 10.*
16. Isolate cells with high eGFP expression for further analysis.
17. Freeze isolated RPE cells at -20 °C for further analysis (*see Note 11*).

3.13 Genomic Indel Analysis of In Vivo Edited Cells

1. Purify gDNA from eGFP⁺ RPE cells using QIAamp DNA Micro Kit according to the manufacturer's protocol (Fig. 1i).
2. Use 10–100 ng of gDNA for PCR amplification (Fig. 1j). Perform PCR with 10 μ L Phusion High-Fidelity DNA Polymerase and 10 pmol of *mVegfa* exon 3 fwd and rev primers (Table 1, primer-3 and primer-4). Run PCR according to Table 2 with 66 °C annealing temperature.
3. Analyze the amplified PCR product in a 1% analytic agarose gel. Place the gel in a submerged horizontal electrophoresis cell and cover with $1 \times$ TAE buffer.
4. Add loading dye to the PCR products and subject the samples to the wells. Run the gel at medium voltage (70–80 V) for 1 h.
5. Verify correct band size using a UV transilluminator. A single uniform 777 bp long PCR band is expected in this setup.
6. Isolate and purify the PCR product from the gel using MinElute PCR Purification Kit or similar PCR purification kit.
7. TIDE analysis (Fig. 1l, n) of eGFP⁺ RPE cells was performed as described in Subheading 3.7, but with *mVegfa* exon 3 fwd primer instead (Table 1, primer-3).

3.14 TOPO Cloning and Analysis of eGFP⁺ RPE Cells

TOPO cloning was performed using the Zero Blunt TOPO PCR Cloning Kit (Fig. 1m).

1. Use 100 ng of PCR product amplified from murine RPE gDNA (Subheading 3.6, step 3 pkt. 6). Mix with 1 μ L of Salt solution (applied in the Zero Blunt TOPO PCR Cloning Kit), 1 μ L of pCR-Blunt-II-TOPO vector and water in a total of 6 μ L.
2. Incubate for 5 min.
3. Transform into *E. coli* as described in Subheading 3.2, step 15.
4. Prepare a PCR reaction with 10 μ L Phusion High-Fidelity DNA Polymerase, 10 pmol of *mVegfa* exon 3 fwd and rev primers (Table 1, primer-3 and primer-4) and water in a total of 20 μ L.
5. Use a sterile inoculation needle to transfer cells from a single colony to the PCR tube.
6. Perform PCR as described in Table 2 with 66 °C annealing temperature.
7. Analyze 2 μ L of PCR product in a 1% analytic agarose gel as described above.
8. Purify PCR product using MinElute PCR Purification Kit or a similar PCR purification kit.
9. Sequence PCR product using *mVegfa* exon 3 fwd primer (Table 1, primer-3).
10. Align sequences for each clone and identify indel types (Fig. 1m–o). In the specific case, we frequently identified indels ranging from 2-bp insertions to 4-bp deletions in *Vegfa*-specific amplicons [1] (see also Note 12).

4 Notes

1. HEK293-vegfa cells stably express high levels of murine *Vegfa* [10].
2. The location of sgRNAs were based on identification of an early coding exon which is common among splice variants and is a large exon, which allows for design of multiple sgRNAs. Exon 3 of the murine *Vegfa* gene was chosen. We designed our sgRNAs using Broad Institute CRISPRko (broadinstitute.org/gpp/public/analysis-tools/sgrna-design) and MIT CRISPR Design (<http://crispr.mit.edu/>). Broad Institute CRISPRko assesses on-target efficiency from 0 to 1, with 1 as the most efficient. In addition, the sgRNAs were ranked according to off-targets risk. MIT CRISPR Design assesses off-target risk on a scale from 1 to 100, where high scores indicate low off-target risk. Initially, we chose sgRNAs both

based on off-target risk and on-target efficiency. However, as on-target efficiency was of the essence in our setup, we used the most efficient on-target cleaving sgRNAs for in vivo studies.

3. As the LV/Cas9-sgRNA plasmid is linearized with a single restriction enzyme addition of CIP to the digested vector is imperative to prevent religation of the linearized plasmid DNA in a subsequent ligation reaction. Ligation mix without fragment was used as negative control for the transformation.
4. Competent cells are heat sensitive. Hence, tubes containing competent cells should be transferred on ice immediately after removing them from the freezer as well as thawed on ice. In addition, the duration of the heat pulse is critical. Following transformation plates should be left upside down. To avoid drying-out plates must be wrapped in Parafilm. The negative control plate should result in very few colonies, preferably fewer than 10. It is preferable to include a negative plate (e.g., linearized and dephosphorylated vector) and compare the number of colonies on this plate to the number on a positive plate. The number of colonies on the latter should be many times greater than the number on the negative control plate. If there are no or very few colonies on the positive plate, or the number of colonies is comparable to the number on the negative plate, the transformation should be repeated. In this case circular plasmid could be included as a positive control for correct transformation conditions. Ligation reactions can be stored at -20°C . If retransformation using the same ligation reaction is unproductive, the ligation should be repeated. Preferentially, redo restriction digest and purification of DNA vector and fragment.
5. The applied LV particles are produced using a third-generation LV-packing system. Relevant packing plasmids (in each case three different) and LV/Cas9-sgRNA plasmids (Fig. 1c) are used to transfect HEK-293 T cells [1, 5, 8]. For cell studies, crude LV preparations can be harvested and for subretinal injections, the LV preparations are typically concentrated by ultracentrifugation. In each case the titers of the LV preparations can be assessed, e.g., by determining the HIV-1 p24 concentration by means of ELISA [11].
6. The key component of the Surveyor Nuclease Digestion assay is the Surveyor Nuclease. Upon denaturation and random reannealing of PCR amplicons, DNA strands with different indels and DNA strands with non-edited strands will randomly pair up. In some cases, this will result in mismatches between the two strands due to the indels formed, i.e., heteroduplexes. Reannealing between matching strands will result in homoduplexes. The Surveyor Nuclease detects mismatches and cleaves the DNA sequence, making the formation of heteroduplexes

essential for the Surveyor analysis. These heteroduplexes are detected as edited DNA sequences, while homoduplexes will be evaluated as non-edited DNA.

For reliable readouts of the Surveyor analysis, a few assumptions are made. Firstly, that the wild-type sequence is the predominant allele in the analysis mixture. Secondly, a variety of indels are formed, as one unique indel will not result in mismatches with reannealing with other edited DNA strands. In some cases, one unique indel is formed in a large proportion of the cells, making this the predominant allele. In this case, reannealing will result in a large quantity of homoduplexes, which will result in false negative results of the Surveyor analysis. An example of this is sgRNA1 (Fig. 2a, c) which yields a score of ~15% by Surveyor analysis, but sequencing and TIDE analysis reveals that a large quantity (87%) of +1 indels contain a unique A insertion. Homoduplexes of this +1 A insertion created a false negative result in the surveyor analysis. For this reason, wild-type PCR amplicons can be mixed with the sample PCR amplicons prior to denaturation and reannealing to account for this phenomenon.

Another pattern of indel formation was obtained using sgRNA2. In this case several different indels were formed, and the results obtained from Surveyor and TIDE analysis were comparable (Fig. 2a, d).

7. Following Surveyor Nuclease Digestion, the samples are run on an agarose gel. The gel is photographed, and three DNA bands may be identified (Fig. 2a). The DNA bands with lowest mobility correspond to the original PCR amplicon, i.e., uncut homoduplexes (Band A, Fig. 2a). The two high mobility bands (B and C, Fig. 2a) correspond to the two fragments created by the Surveyor Nuclease cutting heteroduplexes, and should together resemble the total length of the PCR amplicon. Calculation of the indel frequency is performed using the equation $\text{indel\%} = 100 \cdot (1 - (1 - f_{\text{cut}})^{1/2})$ (see Fig. 2b). f_{cut} is the amount of cleaved DNA out of total DNA: $f_{\text{cut}} = (B + C) / (A + B + C)$, where A, B, and C are the three DNA bands with A being the largest DNA band and C being the smallest cleavage product.
8. TIDE software is available at tide-calculator.nki.nl. TIDE analysis requires three components: the sgRNA sequence, the sequence of a control and the sample to be evaluated. A few factors can be modulated to assess indel frequency: alignment window, decomposition window, and indel range. These can be modulated according to the cut site and the size of indels you want to assess. Based on the analysis, a total efficiency score is provided indicating the fraction of altered sequences.
9. Subretinal injection of LV particles results in RPE-specific expression [11, 12]. In the present case, RPE-specific expression was assessed by immunofluorescence analysis of eGFP

expression in sections of mouse retina 35 days after subretinal delivery of 140 ng p24 LV/Cas9-sgRNA (Fig. 1f).

10. During the FACS procedure no exclusion criteria were applied. The overall aim of the procedure was to isolate all eGFP positive cells. Normally, cells can be excluded due to size or aggregation of cells. As we identified robust eGFP expression limited to the RPE cell layer following cryosectioning, it was not necessary to exclude cells due to size. Aggregation was not an issue as isolation of eGFP positive RPE cells was the aim. In hindsight, the procedure may have been improved by applying RPE65 antibody to initially isolate RPE cells and then evaluate these for eGFP expression.
11. A pool of RPE cells from 5 to 10 dissected retinas resulted in between 2000 and 12,000 eGFP⁺ RPE cells.
12. TOPO cloning was performed to assess the types of indels. Colonies were PCR amplified for exon 3 of murine *Vegfa* and sequencing analysis of the amplicons revealed the types of indels. Sequences were aligned according to the wild-type sequence and indel frequency was estimated based here on. While TIDE scores overall editing efficacy and identifies bases of +1 insertions, TOPO cloning reveals precise deletions, and insertions larger than +1 base pairs. This is a way to further assess the types of indels formed following double strand repair.

Acknowledgments

This work was supported by the Faculty of Health Sciences (PhD scholarship to AH), The Danish Council for Independent Research (Grant No. 4183-00017B), the Gene Therapy Initiative Aarhus (GTI-Aarhus) funded by the Lundbeck Foundation (Grant No. R126-2012-12456), the Lundbeck Foundation (Grant No. R165-2013-15631), The Danish Eye Research Foundation, Aase og Ejnar Danielsen's Foundation, Knud and Edith Eriksen's Foundation, Maskinfabrikant Jochum Jensen og hustru Mette Marie Jensen f. Poulsens Mindelegat, the Riisfort Foundation and Svend Helge Schrøder og hustru Ketty Lydia Larsen Schrøders fond. Cell sorting was performed at the FACS Core Facility, Aarhus University, Denmark.

References

1. Holmgaard A, Askou AL, Benckendorff JNE et al (2017) In vivo knockout of the *Vegfa* gene by lentiviral delivery of CRISPR/Cas9 in mouse retinal pigment epithelium cells. *Mol Ther Nucleic Acids* 9:89–99. <https://doi.org/10.1016/j.omtn.2017.08.016>
2. Askou AL (2014) Development of gene therapy for treatment of age-related macular

- degeneration. *Acta Ophthalmol* 92(Thesis 3):1–38. <https://doi.org/10.1111/aos.12452>
3. Corydon TJ (2015) Antiangiogenic eye gene therapy. *Hum Gene Ther* 26(8):525–537. <https://doi.org/10.1089/hum.2015.064>
 4. Miller JW, Adamis AP, Shima DT et al (1994) Vascular endothelial growth factor/vascular permeability factor is temporally and spatially correlated with ocular angiogenesis in a primate model. *Am J Pathol* 145(3):574–584
 5. Askou AL, Benckendorff JNE, Holmgaard A et al (2017) Suppression of choroidal neovascularization in mice by subretinal delivery of multigenic lentiviral vectors encoding antiangiogenic microRNAs. *Hum Gene Ther Methods* 28(4):222–233. <https://doi.org/10.1089/hgtb.2017.079>
 6. Doench JG, Fusi N, Sullender M et al (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* 34(2):184–191. <https://doi.org/10.1038/nbt.3437>
 7. Hsu PD, Scott DA, Weinstein JA et al (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* 31(9):827–832. <https://doi.org/10.1038/nbt.2647>
 8. Askou AL, Corydon TJ (2018) Development of multigenic lentiviral vectors for cell-specific expression of antiangiogenic miRNAs and protein factors. *Methods Mol Biol* 1715:47–60. https://doi.org/10.1007/978-1-4939-7522-8_4
 9. Guschin DY, Waite AJ, Katibah GE et al (2010) A rapid and general assay for monitoring endogenous gene modification. *Methods Mol Biol* 649:247–256. https://doi.org/10.1007/978-1-60761-753-2_15
 10. Pihlmann M, Askou AL, Aagaard L et al (2012) Adeno-associated virus-delivered polycistronic microRNA-clusters for knockdown of vascular endothelial growth factor in vivo. *J Gene Med* 14(5):328–338. <https://doi.org/10.1002/jgm.2623>
 11. Askou AL, Aagaard L, Kostic C et al (2015) Multigenic lentiviral vectors for combined and tissue-specific expression of miRNA- and protein-based antiangiogenic factors. *Mol Ther Methods Clin Dev* 2:14064. <https://doi.org/10.1038/mtm.2014.64>
 12. Balaggan KS, Ali RR (2012) Ocular gene delivery using lentiviral vectors. *Gene Ther* 19(2):145–153. <https://doi.org/10.1038/gt.2011.153>



In Vivo Editing of the Adult Mouse Liver Using CRISPR/Cas9 and Hydrodynamic Tail Vein Injection

Francesco Niola, Frederik Dagnæs-Hansen, and Morten Frödin

Abstract

CRISPR/Cas9 technology allows facile modification of the genome in virtually any desired way through the use of easily designed plasmid constructs that express a gRNA targeting a genomic site-of-interest and Cas9. Hydrodynamic tail vein injection, on the other hand, is a simple method to deliver “naked” plasmid DNA to 5–40% of the hepatocytes of the liver of adult mice. Here, we describe how these two techniques can be combined to create a workflow for fast, easy, and cost-efficient *in vivo* genome editing of the adult mouse liver. Using this method, large cohorts of mice with genetically modified livers can be established within 3 weeks to generate models for gene function in normal physiology and diseases of the liver.

Key words Hydrodynamic tail vein injection, Liver, CRISPR/Cas9, Indels, Genetic mouse models

1 Introduction

The liver is the second largest organ in the body and the base of numerous genetic, metabolic diseases. Furthermore, primary liver cancer ranks second in mortality among all malignancies worldwide. For these reasons, the mouse liver has been subject to intense genetic engineering for the past many years with the aim of generating genetic models of disease or evaluating gene therapy approaches for liver as well as for extrahepatic diseases.

Hydrodynamic tail vein delivery of plasmid-based genetic engineering reagents to the adult mouse liver has emerged as a powerful method for liver engineering. In this approach, “naked” DNA is injected manually within 4–7 s into the lateral tail vein of the mouse in a volume of physiological saline solution corresponding to ~10% of the body weight—roughly equivalent to the total blood volume of the animal [1, 2]. The bolus injection causes transient heart failure and decreased heart rate, resulting in accumulation of the injected solution in the inferior vena cava and an increased hydrostatic pressure [3]. This forces the solution into the liver via the hepatic veins and sinusoids, *i.e.*, in the reverse direction of normal

blood flow. The mechanism of cellular uptake of plasmid DNA is not fully elucidated but appears to occur via enlargement of sinusoidal fenestrae and generation of transient membrane pores and vesicles in the hepatocytes [3–5]. HTV injection targets the liver tissue relatively uniformly at the gross level [1, 2, 6]. Within the individual liver lobules, however, hepatocytes in the region around the hepatic venule (zones 2 and 3) constitute the vast majority of the transfected cells [5, 7]. The reported total efficiencies vary, but 5–40% of liver cells may become transfected, as estimated by transgene/reporter gene expression. If several plasmids are co-injected, a high degree of co-delivery of plasmids to individual hepatocytes has been demonstrated [8, 9]. Liver damage associated with HTV injection is very limited; only ~5% of hepatocytes become necrotic or apoptotic within the first day post-injection and most of these are cleared after a few days [1, 2, 5]. While the liver is the all-predominant target tissue, other organs like kidney, lung, spleen, and heart are also transfected by HTV injection, albeit at >1000-fold lower levels [2, 4]. Because of these features, a large number of studies have used HTV injection to manipulate the mouse liver through delivery of plasmid DNA, DNA fragments, bacterial artificial chromosomes, siRNAs, oligos, proteins, and more since the method was reported two decades ago [10].

More recently, a series of studies have shown that the mouse liver can be genome edited in almost any desired way through hydrodynamic delivery of CRISPR/Cas9 reagents. Thus, co-delivery of oligo donor and gRNA/Cas9 has been used to knockin correct a point mutation responsible for a metabolic disease [11]. Liver cancer has been evoked by co-delivery of two gRNAs and Cas9 to indel mutate (knock out) two cancer genes [12], by delivery of 10 gRNAs to knock out 10 cancer genes [8] as well as through delivery of genome-wide gRNA libraries for knock-out mutagenesis screens [6]. Finally, we and others have co-delivered two gRNAs and Cas9 to engineer a large (400 kb) chromosomal deletion detected in fibrolamellar hepatocellular carcinoma patients to demonstrate a resultant *Dnajb1-Prkaca* fusion gene as the causative oncogenic driver [13, 14]. In these studies, the vector systems either transiently expressed the CRISPR/Cas9 constructs in the liver cells [11–14] or stably inserted a CRISPR/Cas9 transposon expression cassette by a co-delivered transposase construct [6, 8]. With respect to editing frequencies, one study reported indel formation in 2–4% and knockin mutagenesis in 0.1% of the hepatocytes 2 weeks after injection [12]. However, as the experiment conferred positive selection on edited cells, initial editing rates may have been lower. Another study reported 0.4% knockin editing frequencies 6 days after injection [11].

Here, we detail the procedures for in vivo editing of the mouse liver using CRISPR/Cas9 and HTV injection. First, a vector system is chosen and a series of gRNAs are designed for the desired

genomic target site and cloned into the vector. The best CRISPR/Cas9 construct is identified by transfection into a mouse cell line followed by analysis of editing efficiency. Next, the best CRISPR/Cas9 construct is delivered to the liver of adult mice through HTV injection. Finally, a couple of the injected mice are analyzed at an appropriate time point after the injection to assess whether the desired level and type of editing has been achieved.

2 Materials

2.1 CRISPR/Cas9 Constructs

The choice of which CRISPR/Cas9 vector system to use naturally depends on the intended experiment. Various vectors are detailed in the studies referenced in Introduction, but in principle, any vector system may work, as long as the promoter that drives expression is functional in hepatocytes. Here, we will outline the general steps for generating and testing a CRISPR/Cas9 construct for HTV injection using pX330-U6-chimeric_BB-CBh-hSpCas9 (Addgene plasmid #42230) as an example. This vector co-expresses gRNA driven by the U6 promoter and *Streptococcus pyogenes* Cas9 driven by the chicken β -actin hybrid promoter and has been used in several liver editing studies [11–14]. The control mouse cohort is injected with pX330-U6-chimeric_BB-CBh-hSpCas9 without or with scrambled/irrelevant gRNA sequence. Plasmids should be of high quality and purity with minimal contaminants, in particular of endotoxins. This can be achieved by purifying plasmids with EndoFree Plasmid Kits from Qiagen.

2.2 For the Injections

1. Mice (6–12 weeks old) (*see Note 1*).
2. Heat chamber (box and heat lamp with a 120 watt red-ray bulb). Cover the bottom of the chamber with bedding—an uncovered bottom burns the feet of the mice.
3. Anesthesia: isoflurane, gas chamber, and anesthetics mask.
4. Scale.
5. Timer.
6. Injection working station with mouse tail restrainer.
7. CRISPR/Cas9 plasmid constructs (amounts are discussed below).
8. Luer-Lock syringes (3-mL for mice <30 g; 4-mL for mice >30 g), compatible 23-gauge, 1¼ in. injection needles and polyethylene tubing with an inside diameter of 0.023 in. to make the injection needle setup, as described below.
9. Ringer's solution (147 mM NaCl, 4 mM KCl, 1.13 mM CaCl₂), autoclaved or filter sterilized.
10. 0.20 μ m cellulose acetate filters and 50-mL Luer-Lock syringe.

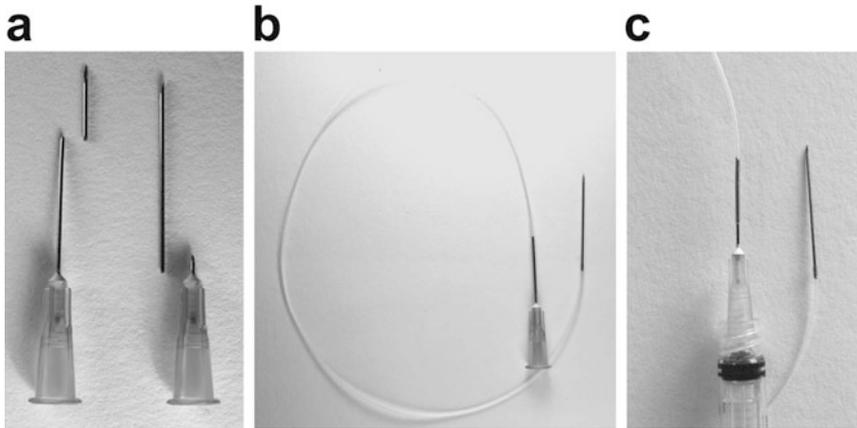


Fig. 1 Assembly of the injection needle setup

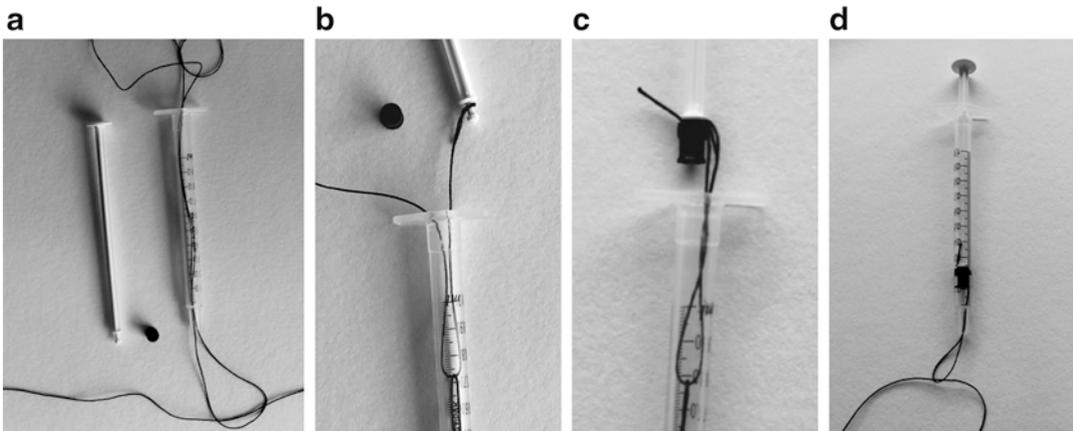


Fig. 2 Assembly of the tourniquet

2.3 Injection Needle Setup

To prepare the injection needle setup, use a pair of tweezers to gently break a 23-gauge, 1¼ in. long Luer-Lock compatible syringe needle close to the hub and another one close to the bevel (Fig. 1a). Connect the injection needle and the needle with the hub with approximately 15 cm polyethylene tubing having an inside diameter of 0.023 in. (Fig. 1b). Mount the hub on a Luer-Lock syringe (Fig. 1c) (*see Note 2*). We change the injection needle in this set after three injections—thus, calculate how many needles to prepare according to the number of mice that will be injected. We change the other parts of the set when changing injection solution—thus, prepare as many of these parts as there are different solutions to inject (*see Note 3*).

To prepare the tourniquet, remove the seal from the plunger of a 1-mL syringe (Fig. 2a). Pass both ends of a nylon or cotton thread through the needle adapter and all through the barrel of the syringe in order to create a loop protruding from the adapter (Fig. 2a). Pass

a spare thread through the loop (Fig. 2a). Tie both ends of the thread that forms the loop to the plunger in such a way that the length of the loop is 2–3 cm shorter than the barrel of the syringe (Fig. 2b). Position the seal back on the top of the plunger (Fig. 2c) and insert the plunger into the barrel. Use the spare thread to pull the loop out of the adapter (Fig. 2d). The plunger can now be used to pull the thread and narrow the loop around the mouse tail to form a tourniquet. We normally have 2–3 tourniquets available for an HTV experiment in case one breaks during use.

2.4 Liver Harvesting

1. Dissection instruments (forceps and scissors of sizes suitable for mice).
2. Disposable scalpels.
3. Phosphate-buffered saline (PBS).
4. Petri dishes (10 cm diameter).
5. A tray with ice.
6. Liquid nitrogen in a suitable container.
7. 4% Formaldehyde.
8. 70% Ethanol.

2.5 Testing Editing Efficiency

1. Nucleic acids extraction kit for tissue.

3 Methods

3.1 Generating and Testing the CRISPR/Cas9 Constructs

1. Design 4–6 gRNAs for the intended genomic target site using your favorite gRNA design tool, as, for example, WTSI Genome Editing (WGE; <https://www.sanger.ac.uk/htgt/wge/>) [15], which is very user friendly. When gRNAs have been selected, sequence the target sites of the gRNAs in the mouse strain to be edited to ensure that the mice do not contain a SNP which would interfere with gRNA binding.
2. Order the gRNA sequences as sense and antisense oligonucleotides, anneal the oligos, and clone them into the *BbsI* cloning site of pX330-U6-chimeric_BB-CBh-hSpCas9, as described [16].
3. Transfect a mouse cell line such as Neuro-2a with the individual CRISPR/Cas9 constructs.
4. After 2 days, extract genomic DNA and determine the frequency and size of indels elicited at the target sites of the various gRNAs. This may be done by Indel Detection by Amplicon Analysis of the PCR amplified target sites, as described in detail [17] and exemplified in Fig. 3a (see Note 4).

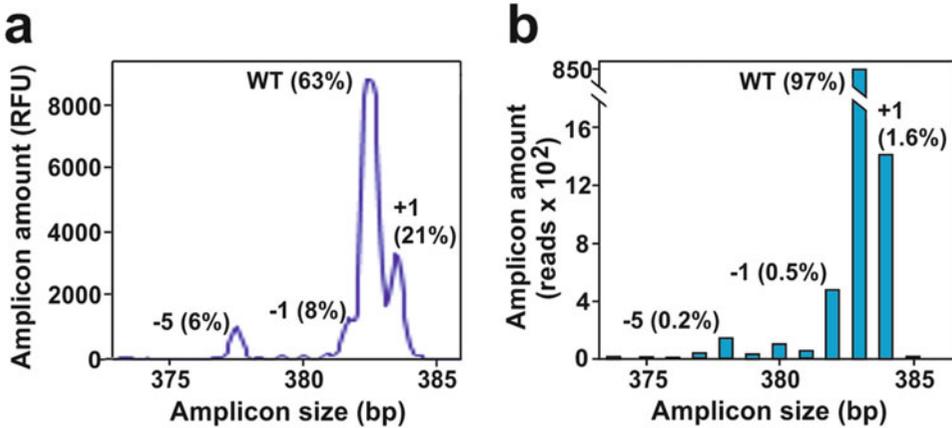


Fig. 3 Evaluation of indel mutagenesis. **(a)** Example of indels elicited in the *trp53* locus in Neuro-2a cells 2 days after transfection with a given gRNA construct, as determined by IDAA. **(b)** Example of indels elicited in the *trp53* locus in the liver 2 months after HTV injection of the same gRNA construct used in **(a)**, as determined by Next-Generation Sequencing. Both assays reveal the size and frequency of indel mutations elicited by the gRNA construct (some are highlighted). Note that the two most frequent indels, a 1 bp deletion and a 1 bp insertion, are the same in Neuro-2a cells and in liver. This illustrates that the indel profile of a given gRNA is often similar across cell types, meaning that testing the gRNAs in a murine cell line is often highly predictive of editing outcome in the liver. Note also that the analyses reveal that the two major indels produce a frameshift and thereby a functional gene knockout

3.2 Hydrodynamic Tail Vein Injection (See Note 5)

3.2.1 DNA Solution Preparation

1. On the day of HTV injection, prepare a volume of Ringer's solution appropriate for the number and size of mice to be injected in a 50 mL Falcon tube or similar. The total volume of solution to be injected per mouse corresponds to 9% of the mouse body weight (*see Note 6*). Thus, if the mouse weighs 25 g, it should be injected with 2.25 mL ($25 \times 0.09 = 2.25$). We use a worksheet providing the volume of solution needed per injection according to mouse weight (worksheet with mouse weights at increments of 0.2 g) to assist a rapid execution of the injection and reduce the risk of mistakes.
2. Add CRISPR/Cas9 plasmid construct to a concentration of 15–60 $\mu\text{g}/\text{mL}$ (*see Note 7*). Filter the solution through a 0.20 μm filter on a 50-mL Luer-Lock syringe into a sterile 50 mL Falcon tube or similar. Make sure that the DNA solution is at room temperature (20–25 °C) at the time of injection. NB: blood contains aggressive DNase activities and therefore the DNA solution must not come in contact with used needles or syringes.

3.2.2 Animal Preparation

1. Weigh the mouse on a scale (*see Note 8*). Take note of its weight and the appropriate volume of DNA solution to be injected in that particular mouse. Take these notes on a worksheet with six columns for: (1) mouse weight before injection, (2) mouse weight after injection, (3) the calculated volume to

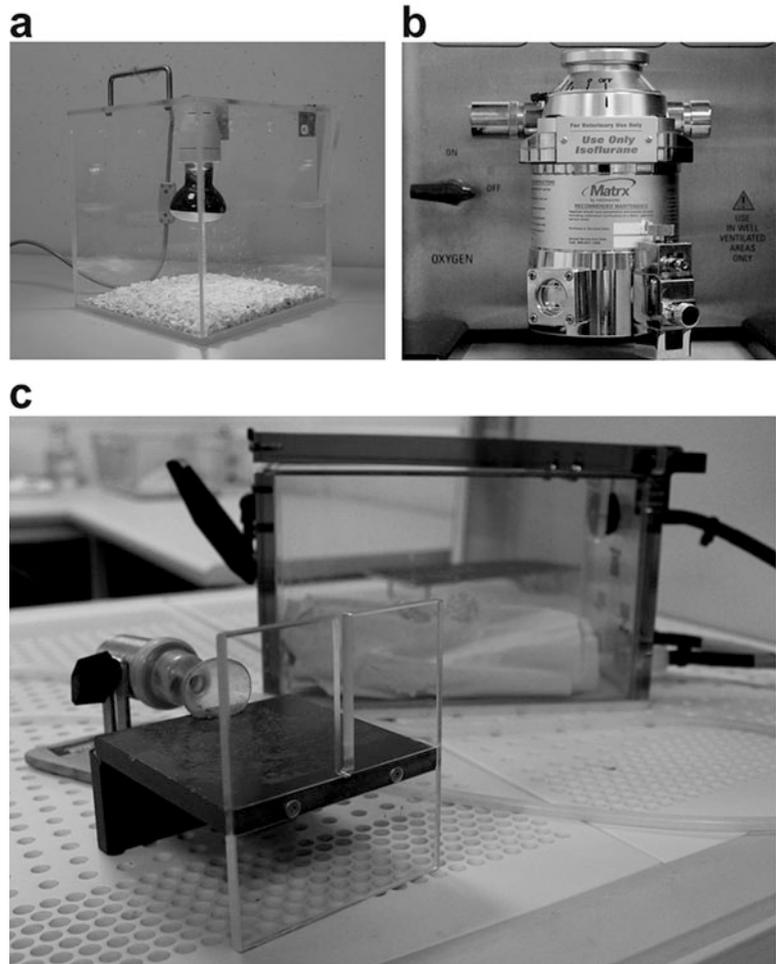


Fig. 4 Examples of a heating chamber (a), a gas vaporizer (b), and a gas chamber and work station setup (c)

inject, (4) the volume injected based on pre- and post-injection weight, (5) the plasmid construct injected, and (6) comments, if any.

2. Load a 3-mL Luer-Lock syringe with a volume of DNA solution appropriate for the mouse plus extra 200–300 μL (*see Note 9*). Remove any small bubbles from the syringe by tapping it or by moving the plunger a little bit up and down. Connect the needle set to the syringe, then push the air out of the needle and tubing and eject the excess DNA solution into a small waste container such that the syringe holds the exact volume appropriate for the mouse to be injected.
3. Place the mouse in a heat chamber (Fig. 4a) and start the timer. Optimal heating time depends mainly on the heat chamber used, and typically is 2–4 min. NB: optimal heating is critical,

as it makes the tail vein swell, thereby becoming more visible and easier to inject. Subsequent steps until insertion of the injection needle in the tail vein (Subheading 3.2.3, **step 1**) must therefore be executed as fast as possible such that the mouse cools down as little as possible (*see Note 10*).

4. Move the mouse to an anesthetic gas chamber (Fig. 4c). Gas chambers are kept in a ventilated hood to minimize the exposure of the operator to anesthetics. Cover the bottom of the gas chamber with paper towel or similar to reduce heat-absorption from the mouse.
5. Once the mouse is anesthetized, transfer it to an injection station and place it such that its muzzle is inside an anesthetic mask in order to keep the mouse lightly anesthetized (Fig. 4c) (*see Note 11*).
6. Lay the mouse on one side, insert the tail of the mouse into the loop of the tourniquet, and then pull the plunger to narrow the loop around the tail and compress the veins (*see Note 12*).

3.2.3 Hydrodynamic Tail Vein Injection

1. Hold and pull the tail with one hand. With the other hand, insert the needle with the bevel facing up into the lateral vein facing up 2–3 cm from the trunk of the mouse (*see Note 13*). The needle should be at a 30–45° angle from the plane of the tail. As the needle enters into the vein, move it to a position almost parallel with the tail and then push it forward to insert most of the needle into the vein (Fig. 5a) (*see Note 14*).
2. Loosen the tourniquet loop from the tail. Some blood should now flow into the tubing, if the needle has been positioned correctly into the vein (Fig. 5b).
3. Hold the syringe stable and start pushing the plunger to discharge all of its content—it is suggested to keep the syringe lying down on the operating table when injecting. Use a steady movement and a constant pressure to inject the entire DNA solution in 4–7 s. NB: it is critical to inject within this short period of time, as delivery to the liver drops steeply with longer injection times. You should feel minimal resistance during the injection, indicating correct positioning of the needle in the vein.
4. Withdraw the needle from the tail. Weigh the mouse on the scale and take note of the post-injection weight on the worksheet. Label the tail of the mouse with a marker to indicate that it has been injected and move the mouse back to its cage. Observe that the mouse recovers fully, which should occur after 2–6 min (*see Note 15*).

3.3 Liver Harvesting and Analysis of Editing Outcome

1. At an appropriate time point post-injection, sacrifice a couple of mice from the injected cohort following national guidelines (*see Note 16*).

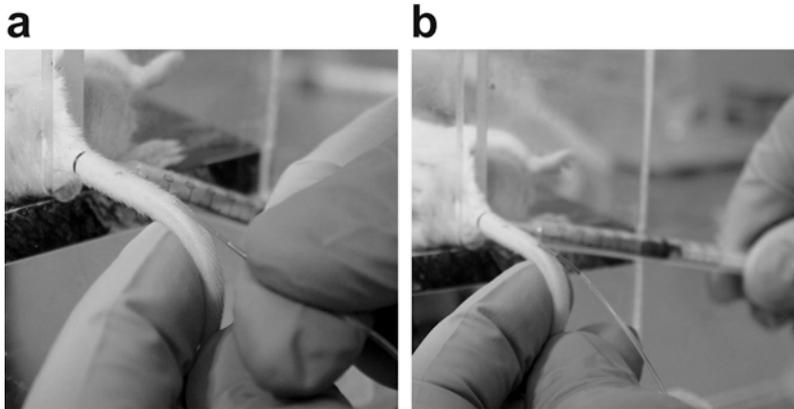


Fig. 5 Steps of the hydrodynamic tail vein injection procedure. (a) Insertion of the injection needle into one of the lateral tail veins. (b) Release of the tourniquet prior to injection. Note some backflow of blood into the tubing, indicating correct positioning of the needle in the vein

2. Collect the liver, place it in a Petri dish on ice and rinse it quickly in ice-cold PBS. Collect a small sample (or several) of the liver with a clean pair of scissors (or disposable scalpels) and place it in a collection tube (*see Note 17*). Flash-freeze the sample in liquid nitrogen. Place the remaining liver in a container with 4% formaldehyde for 24 h at 4 °C and thereafter change to 70% ethanol for fixation and preservation of the tissue for immunohistochemistry, if needed.
3. Extract genomic DNA from one of the frozen tissue samples using an appropriate kit according to the manufacturer's instructions (*see Note 18*).
4. Evaluate whether desirable indel sizes and frequencies have been obtained by performing indel mutation analysis of the gRNA target site. Depending on the extent of editing that has been obtained, this may be performed by IDAA, TIDE or by Next-Generation Sequencing [18]. The latter is exemplified in Fig. 3b.

4 Notes

1. In the CRISPR/Cas9 editing studies reported so far, the mice injected have been 6–10 weeks of age. However, mice of younger or older age can be HTV injected [2]. Any mouse strain can be used, but the tail veins of albino strains are more visible and hence easier to inject. If the mice are from another facility, plan to receive the animals following national guidelines to allow time (typically several days) for the mice to acclimatize to the new environment of your animal facility before performing the experiment.

2. The use Luer-Lock syringes and compatible needles is important, because the relatively high pressure during injection may cause the needle to detach from the syringe if not secured by a Luer-Lock.
3. This flexible needle setup makes it easier to insert the injection needle into the vein as well as to perform a steady injection with the syringe lying on the operating table. With mice <6 weeks of age, the use of a 23-gauge needle may be a challenge, as the veins are smaller. A 27-gauge needle attached to a smaller-diameter tubing is a possibility, but the increased resistance of the smaller-diameter tubing may increase the challenge of injecting the large volume within the 4–7 s required. Alternatively, a 27-gauge, ½ inch needle (Becton Dickinson, cat. no. 305109) mounted directly on a Luer-Lock syringe may be used for mice of any age [1], but may require more practice.
4. Alternatively indels may be characterized using Tracking of Indels by DEcomposition (TIDE) analysis of the PCR amplified target sites [19].
5. In principle, one trained person can perform an HTV injection. However, if injecting larger cohorts (>10 mice), it is highly recommended for speed, accuracy and reproducibility that a small team executes the procedure: one person should focus entirely on the injections. Meanwhile, another one or two persons take care of the mouse weighing, heating, and injection solution calculation. In this way, one mouse can be prepared, while another one is being injected. A team of trained persons can inject a cohort of 40 mice within 2–3 h.
6. Most protocols inject a volume corresponding to 10% of the mouse body weight, but we find that the less stress-full 9% works well. Note, however, that 9% is close to the minimum volume that can be used, as transfection efficiencies drop steeply with lower relative injection volumes [1, 2]. Since the weight of mouse strains vary considerably, weigh 6 mice to estimate the average weight of the cohort to be injected. To determine the needed volume of Ringer's solution, calculate the injection volume for the average mouse and multiply this volume by 1.5 to allow for loss of solution at the various steps of the procedure and for some re-injections (*see Note 14*).
7. The amounts will depend on the CRISPR/Cas9 vectors used: constructs for transient expression are used in the higher range, whereas the transposase-based constructs are used in the lower range—refer to the studies cited in Introduction. We are not aware of dose-response analyses of editing efficiency as a function of injected amount of plasmid, but the ranges given here have been used in all of the published CRISPR/Cas9 studies and, for example, yield indel frequencies in the range of 2–4% at a targeted allele.

8. Optimally, use a scale suitable for weighing un-anesthetized mice. If not available, weigh the mice immediately after they have been anesthetized, but do it fast so that they do not cool down too much after the heat chamber. Note that it is not unusual that the mice decrease a bit in body weight during the heating procedure.
9. Or use a 4-mL Luer-Lock syringe, if injecting mice above 30 g.
10. Black mice heat faster than albino or mice with a lighter coat color. As mice are getting heated, they start getting jumpy. Note that while appropriate heating is critical, excessive heating must be avoided, as the mice may die.
11. A 10-mL syringe (or 5-mL for small mice) can be used as an anesthetic mask; sometimes it is easier to position the muzzle of the mouse such that it receives an appropriate dose of anesthetic in a syringe than in commercial masks.
12. The tourniquet helps keeping the tail vein swollen and thereby easier to inject.
13. Two tail arteries run along the upper and lower side of the tail; the two lateral tail veins run on either side of the tail, midway between the arteries.
14. If correct insertion of the needle fails, it may be re-attempted in the other lateral tail vein. However, if the procedure has taken too long, the mouse may have cooled down and the tail veins have become difficult to see. In that case, the mouse may be transferred back to its cage, allowed to recover and injected at a later time point of the session.
15. We normally achieve 100% survival in a cohort of injected mice. However, occasionally one or two mice in a larger cohort may not recover after the anesthesia and injection and may die or need to be euthanized. The measured weight increase may not exactly correspond to the weight of the injected volume, but it helps evaluate whether the intended volume was approximately injected. It occasionally happens that the deviation is significant and in such cases it may be considered removing the mouse from the experiment.
16. We evaluate editing efficiency at least 3 days post-injection in order to predominantly analyze the cells that have survived the injection (*see* Subheading 1). Editing efficiency may, however, also be assessed at later time points, depending of the nature of the experiment. When modeling liver cancer, we often perform the analysis after e.g., 4 weeks to allow for some expansion of edited cells. Thereby less sensitive, but simpler mutation detection methods like IDAA or TIDE can be used instead of NGS analysis.

17. As HTV injection targets the liver tissue relatively uniformly, in principle, any part of the liver may be sampled. However, we normally sample a part of the large liver lobe.
18. In some cases, analysis of mRNA may be a desired read-out [13] and in such instances RNA should be extracted.

Acknowledgments

This work was supported by the Danish Cancer Society (R146-A9563-B3551 to F.N. and R124-A7632-15-S2 to M.F.), the Danish Council for Independent Research (DFF-Mobilex 4092-00235 to F.N.), and the Novo Nordisk Foundation (NNF17OC0028380 to M.F.)

References

1. Zhang G, Budker V, Wolff JA (1999) High levels of foreign gene expression in hepatocytes after tail vein injections of naked plasmid DNA. *Gene Ther* 5:272–276
2. Liu F, Song Y, Liu D (1999) Hydrodynamics-based transfection in animals by systemic administration of plasmid DNA. *Gene Ther* 6:1258–1266
3. Zhang G, Gao X, Song YK, Vollmer R, Stolz DB, Gasiorowski JZ, Dean DA, Liu D (2004) Hydroporation as the mechanism of hydrodynamic delivery. *Gene Ther* 11:675–682
4. Kobayashi N, Nishikawa M, Hirata K, Takakura Y (2004) Hydrodynamics-based procedure involves transient hyperpermeability in the hepatic cellular membrane: implication of a nonspecific process in efficient intracellular gene delivery. *J Gene Med* 6:584–592
5. Budker VG, Subbotin VM, Budker T, Sebestyén MG, Zhang G, Wolff JA (2006) Mechanism of plasmid delivery by hydrodynamic tail vein injection. II. Morphological studies. *J Gene Med* 8:874–888
6. Xu C, Qi X, Du X, Zou H, Gao F, Feng T, Lu H, Li S, An X, Zhang L, Wu Y, Liu Y, Li N, Capecchi MR, Wu S (2017) piggyBac mediates efficient in vivo CRISPR library screening for tumorigenesis in mice. *Proc Natl Acad Sci U S A* 114:722–727
7. Zhang G, Song YK, Liu D (2000) Long-term expression of human α 1-antitrypsin gene in mouse liver achieved by intravenous administration of plasmid DNA using a hydrodynamics-based procedure. *Gene Ther* 7:1344–1349
8. Weber J, Öllinger R, Friedrich M, Ehmer U, Barenboim M, Steiger K, Heid I, Mueller S, Maresch R, Engleitner T, Gross N, Geumann U, Fu B, Segler A, Yuan D, Lange S, Strong A, de la Rosa J, Esposito I, Liu P, Cadiñanos J, Vassiliou GS, Schmid RM, Schneider G, Unger K, Yang F, Braren R, Heikenwälder M, Varela I, Saur D, Bradley A, Rad R (2015) CRISPR/Cas9 somatic multiplex-mutagenesis for high-throughput functional cancer genomics in mice. *Proc Natl Acad Sci U S A* 112:13982–13987
9. Hu J, Che L, Li L, Pilo MG, Cigliano A, Ribback S, Li X, Latte G, Mela M, Evert M, Dombrowski F, Zheng G, Chen X, Calvisi DF (2016) Co-activation of AKT and c-Met triggers rapid hepatocellular carcinoma development via the mTORC1/FASN pathway in mice. *Sci Rep* 6:20484
10. Bonamassa B, Hai L, Liu D (2011) Hydrodynamic gene delivery and its applications in pharmaceutical research. *Pharm Res* 28:694–701
11. Yin H, Xue W, Chen S, Bogorad RL, Benedetti E, Grompe M, Koteliansky V, Sharp PA, Jacks T, Anderson DG (2014) Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat Biotechnol* 32:551–553
12. Xue W, Chen S, Yin H, Tammela T, Papagiannakopoulos T, Joshi NS, Cai W, Yang G, Bronson R, Crowley DG, Zhang F, Anderson DG, Sharp PA, Jacks T (2014) CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature* 514:380–384

13. Engelholm LH, Riaz A, Serra D, Dagnæs-Hansen F, Johansen JV, Santoni-Rugiu E, Hansen SH, Niola F, Frödin M (2017) CRISPR/Cas9 engineering of adult mouse liver demonstrates that the Dnajb1-Prkaca gene fusion is sufficient to induce tumors resembling fibrolamellar hepatocellular carcinoma. *Gastroenterology* 153:1662–1673
14. Kastenhuber ER, Lalazar G, Houlihan SL, Tschaharganeh DF, Baslan T, Chen CC, Requena D, Tian S, Bosbach B, Wilkinson JE, Simon SM, Lowe SW (2017) DNAJB1-PRKACA fusion kinase interacts with β -catenin and the liver regenerative response to drive fibrolamellar hepatocellular carcinoma. *Proc Natl Acad Sci U S A* 114:13076–13084
15. Hodgkins A, Farne A, Perera S, Grego T, Parry-Smith DJ, Skarnes WC, Iyer V (2015) WGE: a CRISPR database for genome engineering. *Bioinformatics* 31:3078–3080
16. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823
17. Lonowski LA, Narimatsu Y, Riaz A, Delay CE, Yang Z, Niola F, Duda K, Ober EA, Clausen H, Wandall HH, Hansen SH, Bennett EP, Frödin M (2017) Genome editing using FACS enrichment of nuclease-expressing cells and indel detection by amplicon analysis. *Nat Protoc* 12:581–603
18. Güell M, Yang L, Church GM (2014) Genome editing assessment using CRISPR Genome Analyzer (CRISPR-GA). *Bioinformatics* 30:2968–2970
19. Brinkman EK, Chen T, Amendola M, van Steensel B (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42:e168



CRISPR-Based Lentiviral Knockout Libraries for Functional Genomic Screening and Identification of Phenotype-Related Genes

Emil Aagaard Thomsen and Jacob Giehm Mikkelsen

Abstract

Adaptation of the CRISPR system has enabled scientists to probe the genome and interfere with gene function at an unprecedented scale. Adding to the use of CRISPR for generation of individual gene knockout, which is by now conventional, the CRISPR system enables high-throughput functional screening of the genome. By combining the integrative properties of lentiviral vector delivery with the disruptive nature of the CRISPR system, genome-wide CRISPR libraries provide the power to screen among thousands of genes despite the high complexity of the entire genome and identify a list of genes potentially affecting a certain phenotype. Genome-wide CRISPR screening is an advanced technology compiling numerous practical aspects and a series of molecular biology techniques. In this protocol, we describe all steps toward implementing CRISPR knockout screens in your research; we describe the core procedures and key information as well as some tricks and tips needed to successfully perform a CRISPR screen.

Key words CRISPR/Cas9, Lentiviral, sgRNA, Genome-wide screening, Forward genetics

1 Introduction

1.1 General Concepts of CRISPR- Based Screening of the Genome

The discovery of CRISPR, part of the adaptive immune system in bacteria, has revolutionized the vision of editing genomic DNA with high specificity [1]. The CRISPR system, adapted for use in eukaryotic cells, consists of the Cas endonuclease guided to a specific position in the genome by a single guide RNA (sgRNA) molecule. In the most commonly used system derived from *Streptococcus pyogenes*, the SpCas9 endonuclease forms a complex with the sgRNA and binds to genomic DNA through base pairing [1]. The resulting double-stranded break is subsequently predominantly repaired by nonhomologous end-joining (NHEJ), a process known to be highly error-prone, resulting in indels (insertions or deletions) at the cutting site. Efficient targeting of genes leads to knockout mutations that eliminate production of functional

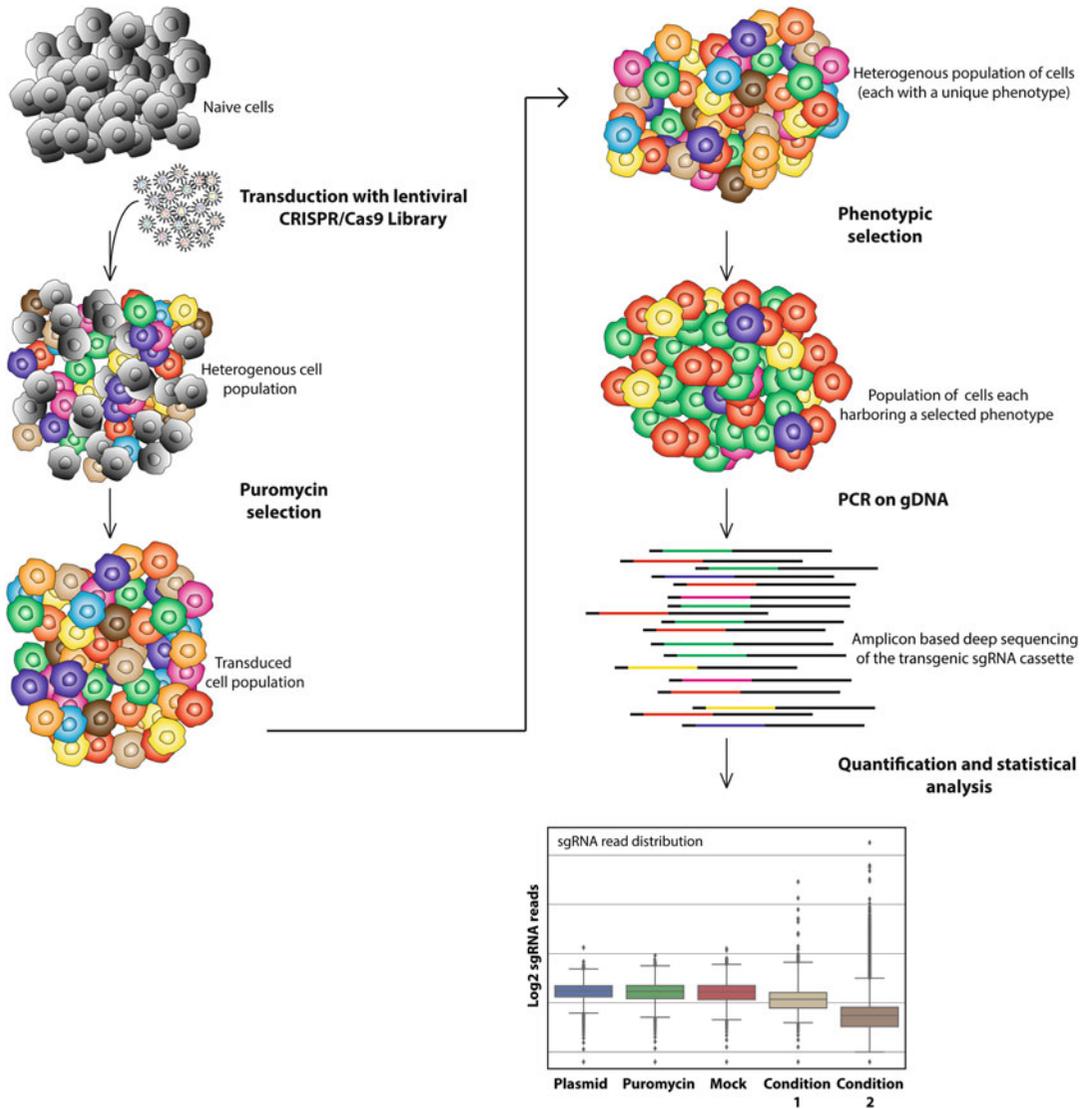


Fig. 1 General CRISPR screen workflow. The conceptual workflow of a general CRISPR screen illustrated in a stepwise schematic manner. The final readout is illustrated by a schematic boxplot representing sgRNA distribution among different samples

protein. By modifying the DNA-binding portion of the sgRNA, it is possible to target and knockout any gene in the genome.

In high-throughput CRISPR library screening approaches (Fig. 1), thousands of different sgRNAs are transferred to a population of cells by lentiviral delivery. Emergence of knockout mutations, determined by the unique sgRNAs expressed in each cell, creates a heterogeneous cell population with knockout mutations in all human genes. Following delivery of a single unique sgRNA to individual cells a unique cellular genotype is generated in each cell. By characterizing cells that are tolerant to a given drug treatment, it

is possible to identify genes that affect the drug response. Other selection modalities can be applied to select for a specific cellular phenotype within the population. Genes affecting the phenotype are identified by next-generation sequencing (NGS) of sgRNA-containing PCR amplicons and downstream bioinformatic analyses. The targeted deep sequencing approach allowing identification of genes that are targeted in CRISPR/Cas9 screens is made possible by the integrative properties of lentiviral vectors. With the controlled delivery by lentiviral vectors, the majority of cells contain one transgenic element, each with a unique sgRNA sequence. With only one sequence per cell, the sgRNA cassette serves as a unique barcode. Identified by NGS, the quantification of these sgRNA sequences, demonstrating either sgRNA depletion or enrichment, will unveil candidate genes associated to the phenotype selected during screening.

1.2 Preserving Representation of the Library

When performing CRISPR-based library screens a large number of different sgRNAs are handled at once and through a series of steps. In order to prevent a bias in the library pool of sgRNAs and maintain the representation of the entire library, different precautions must be taken. The recommended cutoff values, allowing sufficient coverage to be maintained, are based on previous experience with pooled shRNA libraries [2] and CRISPR-based libraries in the original Gecko V2 CRISPR library paper by Shalem and coworkers [3]. Additionally, the technique was developed further with launch of the Brunello libraries in work published by Doench and coworkers [4].

To combat the potential lack of effectivity of specific sgRNAs, pooled CRISPR libraries are in general designed with sgRNA redundancy. In addition, sgRNA redundancy plays a key role in readouts of library screen analyses by consolidating a potential identified gene with hits by multiple sgRNAs supporting this. sgRNA redundancy also contributes to a larger library in total. With unpredictable amplification and lentiviral transfer efficacies for each sgRNA, a theoretical minimum representation of each sgRNA must be kept throughout the entire screening process. A baseline coverage of 500 (e.g., 500 copies of each sgRNA present in the library) will be used in this protocol.

1.3 Knockout Libraries: Some Introductory Remarks

At present, several different libraries suited for different screening modalities and approaches are available on Addgene (www.addgene.org). These include libraries used for knockout, activation, and inactivation screens. Besides genome-wide libraries several smaller customized libraries targeting specific cellular pathways or components exist. The procedures provided in the present protocol describe the essential steps of genome-wide knockout libraries. Most steps, however, will apply to any library in general. The protocols described herein take inspiration from the work and

procedures presented especially by Shalem et al. [3] and Doench et al. [4], who have pioneered these screening methodologies, and in general from instrumental CRISPR work carried out at the Broad Institute Genetic Perturbation Platform. In the protocol presented here, we describe essentials of a knockout screen step by step including some of the tricky aspects that are harder to decipher from current CRISPR screening reports.

2 Materials

2.1 *General Appliances and Reagents*

1. Ethanol 99 and 70%.
2. Isopropanol.
3. MilliQ H₂O or equivalent.
4. Centrifuge suited for 50 mL tubes.

2.2 *Plasmid Amplification*

1. Electrocompetent bacteria; Lucigen Endura (Lucigen) or DH5- α (New England Biolabs) are recommended.
2. Luria Broth (LB) medium.
3. Ampicillin (50 μ g/mL) agar plates.
4. 1.5 mL microcentrifuge tubes.
5. 0.1 cm gap cuvettes (Bio Rad).
6. Bio Rad Gene Pulser (other electroporation systems will also work, but cuvettes should be changed accordingly).
7. Recovery medium, corresponding to bacteria of choice.
8. Loosely capped falcon tubes.
9. Heating cabinet.
10. Maxi/Midi plasmid purification kit.
11. Plasmid from library of choice, several different libraries are readily available from Addgene.
12. Bacterial plate spreader.

2.3 *General Screen*

1. Lentiviral preparation of your library.
2. Puromycin (final concentration suitable for your cell line).
3. Standard reagents for passaging cells.
4. Incubator with settings compatible with the cells used in screen.

2.4 *Genomic DNA Extraction*

1. Water bath or similar.
2. 6 M NaCl.
3. Proteinase K.
4. Lysis Buffer: Tris 10 mM, EDTA 1 mM, NaCl 150 mM, SDS 0.5%, pH 10.5.

2.5 PCR Amplification

1. Herculase II fusion polymerase (Agilent, Santa Clara, CA, USA).
2. Thermocycler.
3. Gel extraction kit.
4. 1.5% agarose gels.
5. Electrophoresis equipment, preferentially with 15–20 well combs 1.5 mm thick in a 15 × 10 cm (W × L) tray.
6. Reverse Phase column purified oligos for primers.

3 Methods

3.1 Plasmid Amplification of Libraries

The following section describes the amplification and reagents used to amplify one genome-wide library (*see Note 1*). Preheat recovery medium (~10 mL) and 21 ampicillin agar plates by placing them in the heating cabinet at 37 °C.

1. Fetch a large amount of ice and place 4 cuvettes (0.1 cm gap) and 4 microcentrifuge tubes (1.5 mL) on ice to prechill.
2. Prepare a second batch of ice and bring up the electrocompetent cells (*see Note 2*) from storage freezer.
3. Let cells thaw slowly on ice (≈8–10 min).
4. Aliquot 25 μL into prechilled microcentrifuge tubes (1.5 mL).
5. Add 2 μL of library plasmid (*see Note 3*) to the 25 μL cells in the prechilled microcentrifuge tubes, mix gently, and avoid bubbles.
6. Transfer the cell and plasmid solution into the narrow gap of the electroporation cuvette and *be sure not to make any bubbles!* (*see Note 4*).
7. As chilling of the tubes will leave traces of water on the outside, wipe the *tubes before inserting* them into your electroporation equipment, such as the Bio-rad gene pulser.
8. Place cuvette in Bio-rad gene pulser, enter the electroporation settings according to your bacteria of choice (*see Note 2*); release the pulse.
9. Immediately following the electroporation, add 973 μL of recovery medium into the cuvette, gently pipet up and down a few times, and transfer then to 14 mL loosely capped falcon tube. Add 1 mL of additional recovery medium to the falcon tube, resulting in a total of 2 mL for each electroporation (*see Note 5*).
10. Repeat previous steps until reaching a total of 4 electroporations for each library.

11. Pool 4 electroporations per library in the loosely capped falcon tube, resulting in a total volume of 8 mL recovery medium.
12. Incubate in shaking incubator for 1 h at 37 °C.
13. Prepare a 40,000× dilution for the control plate by transferring 10 µL from your total on 8 mL to 1 mL recovery media in an Eppendorf tube; be sure to mix it well and then plate 20 µL onto a pre-warmed agar plate.
14. Plate the remaining onto 20 pre-warmed agar plates, 400 µL on each plate, and spread.
15. Incubate at 37 °C for 14 h (*see Note 6*).
16. Count the number of colonies on your control plate, if it is above 100 proceed to harvesting (*see Note 7*).
17. Pipet 3 mL LB medium onto each plate and scrape off the layer of bacteria, repeat this process once. LB medium and bacteria from each plate are pooled (≈ 120 mL total) into a tube for centrifugation. Note the weight of the empty tube.
18. Spin it down and discard supernatant; determine the weight of the tube and calculate the weight of the wet bacterial pellet by subtracting the weight of the empty tube.
19. Distribute into several purification columns or use a larger column. Be sure not to overload the column. Check the capacity of your preferred columns beforehand. Purify according to manufacturer's protocol. Expect between 2 and 3 g of wet bacterial pellet. Using the NucleoBond Xtra Midi kit this is equivalent to approximately six columns (*see Note 8*).

3.2 Production of Lentiviral Vector Libraries

3.2.1 Considerations Related to Viral Vector Production

To estimate the lentiviral transfer capacity (transductional titer) in your cell line of choice, start out by producing a batch of “crude virus” (sterile-filtered virus-containing medium from the producer cells) following the protocol described by Ryø et al. [5], which also describes different techniques for assessing viral titer. It is recommended to perform this assessment with vectors produced using pLentiGuide-puro [6], as this is the vector used in most libraries. We refer to a lentiviral vector carrying vector RNA encoded by the pLentiGuide-puro plasmid as LV/LentiGuide-puro. Assessment of the titer of your prep of LV/LentiGuide-puro in your cell line will help you to decide on the production method for producing the final lentiviral library. If the titer of the lentiviral vector preparation contained in medium from vector-producing cells (“crude virus”) is below 1×10^5 in your cell line, it is recommended to upconcentrate the virus preparation by ultra-centrifugation. The extra workload for producing ultra-centrifuged virus is substantial, but for most cell lines the crude vector preparation in filtered medium will suffice. *See* Subheading 3.3 for guidelines to determine the viral vector requirements for your library setup and calculate the

production scale needed (*see Note 9*). After harvesting the virus make separate small aliquots of your lentiviral vector preparation for titration to avoid any impact of freeze-thaw cycles on your actual library preparation.

3.2.2 Generation of SpCas9-Expressing Cell Line

Most of the commercially available libraries are based on a two-vector approach by which the SpCas9 endonuclease is delivered separately from the sgRNA. Consequently, the sgRNA transfer plasmid can be delivered with higher efficiency. Hence, a cell line with stable SpCas9 expression is required (*see Note 10*), which is the case, for example, when running the Brunello human genome-wide knockout library. For more details on generating a stable SpCas9-expressing cell line and validating SpCas9 function, we refer to an accompanying protocol [5].

3.2.3 Titration of Lentiviral Library Preparation

To ensure sufficient coverage of your library during the screen and to ensure mostly single transduction events per cell it is crucial that your actual batch of library virus is titrated in the specific SpCas9-expressing clone or a cell population intended for the screen. This process of titration can be carried out as previously described.

3.3 General Considerations Related to the Setup

Although genome-wide screening using CRISPR is a powerful technique it is important to emphasize that it will only provide answers to the exact questions that you ask. Consequently, general handling of cells when passaging will influence the final readout. The success of a CRISPR screen relies heavily on the selection scheme. As such, the desired selection scheme should be thoroughly tested in your SpCas9-expressing cells and preferentially also under the exact conditions (culture scale) that you intend to use during the screen. Besides thorough testing it is recommended to include more selection schemes, e.g., different drug concentrations. Here, we illustrate a suggested day-by-day scheme for a standard setup with the Brunello genome-wide knockout library.

For a screen with the Brunello library expand SpCas9-expressing cells to a minimum of 1.6×10^8 cells total, allowing two replicates with $500\times$ coverage (*see Note 11*).

1. Day 1. Seed 8.0×10^7 cells per replicate at minimum density or a density that will allow them to grow unhindered and without passaging until puromycin selection starts.
2. Day 2. Transduce each replicate with a volume of library vector preparation corresponding to 4.0×10^7 functional virus particles (*see Subheading 3.2* on titration of viral particles). This results in an MOI of 0.5, which is expected to result in an acceptable number of cells with multiple transduction events at (*see Note 12*).

3. Day 3 (optional). Refresh medium if polybrene was used during the transduction.
4. Day 4. Refresh medium on the library-transduced cells with medium containing puromycin at a suitable concentration. At this stage, your cells should be near confluent, and the puromycin treatment should quickly thin out the population. However, if cells are confluent, they should be split and slightly diluted. During the puromycin selection cells will have to be passaged or expanded into more flasks (*see Note 13*).
5. Days 4–11. Cells are kept under puromycin selection, adjust passaging of cells so enough cells are available for phenotypic selection after puromycin treatment.
6. Day 11. All cells in each replicate are pooled and counted. Depending on the setup, cells are seeded at 4.0×10^7 per selection scheme, including a mock treatment. Optional: harvest $8.0\text{--}10.0 \times 10^7$ cells per replicate, which may serve as control cells (referred to as “puromycin baseline”) for the downstream analysis of the final library readout, this baseline is recommended.
7. Day 11+. Phenotypic selection of cells is highly depending on the specific drug or other selection modalities. However, a general rule of thumb is not to prolong the screening process for more than 3 weeks. Longer passage times will introduce a growth-related bias in the library-transduced populations.

3.4 Genomic DNA Extraction

Sufficient sgRNA representation, or coverage, must be maintained throughout the screen, during transduction and passaging of cells and after harvesting cells and genomic DNA extraction. Calculate the required number of cells needed for genomic DNA extraction in order to maintain coverage and add 20–30% to account for any loss during the purification or testing of downstream PCR reactions (*see Note 14*). The protocol for DNA extraction is as follows:

1. Harvest cells by pelleting them in a 50-mL falcon tube.
2. Pellet $1.0\text{--}1.5 \times 10^8$ cells in each tube. Usually one tube per sample will do.
3. Resuspend the pellet in 1 mL PBS (*see Note 15*).
4. Add 12 mL of lysis buffer and 60 μ L Proteinase K to each falcon tube.
5. Incubate overnight (14–16 h) at 55 °C in a water bath or similar.
6. The following day, add 4 mL of 6 M NaCl and shake each tube vigorously.
7. Spin down at 4 °C, $3005 \times g$ for 50 min.
8. Transfer supernatant to a new 50-mL falcon tube.

9. Spin down at 4 °C, 3005 × *g* for 50 min.
10. Transfer supernatant to a new 50-mL falcon tube containing 20 mL of ice-cold 99% ethanol and invert roughly 50 times or until a clear precipitate is visible.
11. Spin down at 4 °C, 3005 × *g* for 80 min.
12. Remove supernatant and add 5 mL of 70% ethanol, invert until pellet detach from the tube.
13. Spin down at 4 °C, 3005 × *g* for 40 min.
14. Remove supernatant and let the pellet air dry; resuspend overnight in 2 mL MilliQ water (preferentially on a rocking table).

3.5 PCR Amplification Prior to Next-Generation Sequencing

The preparation of amplicons for subsequent NGS follows a nested PCR approach. The first PCR (PCR #1) must preserve the coverage maintained throughout the screen. Consequently, the genomic DNA that serves as template must correspond to the number of cells that equals the biological coverage. Assuming that each cell contains 6.6 pg of DNA (take major aneuploidy of your cells into account), ~256 μg of genomic DNA should be used as template for PCR #1 in a screen with the Brunello library (77,441 unique sgRNAs) using 500× biological coverage (as determined based on the following calculation: $6.6 \times 10^{-6} \mu\text{g} \times 77,441 \times 500 = 255.55 \mu\text{g}$).

This entails multiple PCR #1 reactions per replicate, which will later be pooled. In this protocol, each PCR #1 reaction is performed with 10 μg genomic DNA as template, resulting in a total of 26 reactions per screen condition.

To help achieve even amplification between samples from all screening conditions, genomic DNA is diluted to ~200 ng/μL (*see Note 16*), as higher concentration of genomic DNA is usually sticky, which can lead to pipetting of uneven amounts of DNA.

3.5.1 PCR #1

For each screen condition prepare a master mix and briefly vortex, before aliquoting 100 μL into separate PCR tubes. PCR #1 with the primers listed here will produce a band of 551 bp (Tables 1 and 2).

The number of cycles should be kept as low as possible (from 16 to 20 cycles) to reduce PCR-introduced bias in the amplicon. PCR #1 does not have to produce a band that is visible by analysis by gel electrophoresis. However, if PCR #2 fails to produce the amplicon resulting in a visible band on the gel, then the number of amplification cycles should be increased. The number of cycles may depend of the purity of the genomic DNA. If increasing the number of cycles does not lead to amplification and production of the amplicon in PCR #2, lowering the amount of genomic DNA in each PCR #1 reaction can help. Obviously, you will then need to perform a higher number of PCR #1 reactions to preserve the coverage.

Table 1
PCR#1 reaction conditions

5 μ L 10 μ M forward primer 5'-TTTCCCATGATTCCTTCATA-3'
5 μ L 10 μ M reverse primer 5'-CGGCGCCAAAGTGGATCTCT-3'
20 μ L Herculase II Buffer X5
2 μ L DNTP mix 10 mM
2 μ L Herculase II polymerase
X μ L DNA (10 μ g)
66 μ L H ₂ O – X

Table 2
Thermocycler conditions

Stage 1: 1 cycle	Stage 2: 16–20 cycles			Stage 3: 1 cycle
95 °C	95 °C	56 °C	72 °C	72 °C
2 min	20 s	30 s	30 s	3 min

3.5.2 PCR #2

All PCR #1 reactions for each sample are pooled. This results in a large volume of PCR #1 product. The number of PCR #2 reactions should be high enough to yield sufficient PCR product for next-generation sequencing while keeping the number of amplification cycles as low as possible with each reaction. Perform one reaction per 1.0×10^4 unique sgRNAs in your library. For the Brunello library, this corresponds to approximately 8 PCR #2 reactions, which should produce enough PCR product for downstream analysis.

Depending on the specific sequencing platform, the primers used in PCR #2 will change accordingly. For a finished and ready-to-sequence pool of amplicons, the primers should contain Adapters (e.g., Illumina-compatible), a stagger region to avoid mono-template issues and a barcode. Depending on the number of samples to be multiplexed, barcodes can be added to the forward primer alone or to both the forward and the reverse primer. Design this type of primers in accordance with your NGS service provider's recommendations (*see Note 17*). As an alternative to ready-to-sequence amplicons, it is possible to produce a pool of amplicons, which is applicable for adapter ligation. This is also an issue that needs to be considered in relation to your NGS provider. With most primers, the reaction conditions listed below should work, but preliminary tests and fine-tuning of the conditions for PCR #2 using your specific primers is recommended. Use the same thermocycler conditions from PCR #1 or adjust the conditions in

Table 3
PCR#2 reaction conditions

5 μ L 10 μ M forward primer
5 μ L 10 μ M reverse primer
20 μ L Herculase II Buffer \times 5
2 μ L DNTP mix 10 mM
1 μ L Herculase II polymerase
5 μ L of pooled PCR#1 reactions
62 μ L H ₂ O

accordance with your specific primers. Keeping the number of cycles low is also crucial during PCR #2 (should be kept between 18 and 24 cycles, Table 3).

3.5.3 Gel Extraction of PCR #2

1. Pool PCR #2 reactions from each sample.
2. Add loading buffer, the total volume amounts to \sim 900 μ L for each sample.
3. Run the gel for at least 50 min at 90 volts to achieve good separation.
4. Purify the amplified PCR band. At this stage, PCR #2 still contains remnants of PCR #1 amplicon as well as potential chimeric products resulting from left-over primers from PCR#1 and primers from PCR#2. As such, good separation on the gel is ideal to certify that the PCR product is as clean as possible for the NGS (*see Note 18*).

A 1.5% agarose gel cast with 15- or 20 well-combs 1.5 mm thick in a gel tray minimum 15 \times 10 cm (W \times L), will produce enough loading capacity to encompass PCR #2 product from one sample. Excise the band precisely to avoid excess agarose and purify it according to your gel extraction kit protocol. With E.Z.N.A Gel Extraction Kit (Omega), one column can bind 25 μ g of DNA. One column per sample is essentially sufficient, but multiple loading rounds are required to load all melted gel material from one sample.

3.6 Analysis of NGS Data

Genome-wide CRISPR screens produce large quantities of data to process and can be analyzed in many ways. It is highly recommended to consult a trained bioinformatician when analyzing library data. A few guidelines and suggestions for data processing are provided in the following section.

1. Initially, check the quality of the data by using software such as FastQC [7]. This will provide a nice and fast indication that the technical part of your sequencing has been successful.

2. Once the quality is secured, the reads will have to be trimmed down to the essential part, namely the sgRNA sequences. Depending on how PCR #2 was performed, the length and start position of your reads may vary, but a tool like Cutadapt [8] will trim down the sequences based on a specific motif. Cutadapt will also permit the user to adjust the promiscuity of the motif and the final length of the trimmed sequence, allowing you to accommodate some of the biases introduced during PCR and sequencing.
3. Once trimmed, the reads will have to be mapped back to an index file containing all the sgRNAs and corresponding gene names within the used library. This is accomplished by using a tool such as Bowtie [9], which is suited for mapping of short sequences. The output of this mapping procedure is a read count file, containing the total number of reads per sgRNA.
4. Quantifying and performing statistical analysis of the reads can be carried out by a tool such as Mageck [10].

4 Notes

1. The procedures described will also apply to smaller libraries like the Human Kinome knockout (Brunello) [4] and the mitochondrial DNA removal library (CRAM) [11]. With these libraries containing fewer sgRNAs, it is possible to reduce the number of electroporations and still obtain 50× coverage of the plasmids. However, following this general procedure will ensure plenty plasmid for downstream production of lentiviral particles carrying the library.
2. Both Lucigen Endura and DH5- α will suffice. However, in our hands Lucigen Endura bacteria have performed best. Lucigen Endura cells require two stock aliquots per amplification, while DH5- α only requires 1. Settings are in accordance with manufacturer's settings.
3. Addgene will ship approximately 10 μ L of library plasmid. Do not waste any of your plasmid by trying to measure the concentration. Quick spin the tube from Addgene and proceed with amplification.
4. Bubbles in the bacterial suspension might cause arching; if this happens consider repeating the electroporation. With a single order of plasmid from Addgene the volume of plasmid may not allow for a second electroporation. If this is the case, proceed with the protocol and evaluate total efficiency based on the control plate.

5. Note that ~20 μL bacterial solution, corresponding to 1/50 of the total volume, is remaining in the gap of the cuvette. For maximum yield, use a p100 pipet to aspirate it.
6. The 14 h of growth is to limit the recombination between the homologous lentiviral LTR regions in the plasmid. Accordingly, the incubation can be done at 32 °C to further reduce this process.
7. It is important to preserve 50 \times coverage during plasmid amplification. For Human Knockout library Brunello [4] with 77,441 sgRNAs, this is obtained with 100 colonies on the control plate. Adjust number of colonies on the control plate to correspond to the specific library of use.
8. To ensure acceptable representation throughout the plasmid pool after amplification, it is recommended to sequence the amplified plasmid pool before proceeding with virus production. However, sequencing of the plasmid can be performed in parallel with sequencing of the samples.
9. With a baseline coverage of 500 using Brunello with 77,441 sgRNAs, 3.87×10^7 cells will have to be transduced; performing the screen in duplicates doubles this number. This corresponds to 7.74×10^7 lentiviral particles needed for a screen. Using a vector with a titer of 1×10^5 in supernatant from vector-producing cells (“crude virus”), the volume of virus needed amounts to 774 mL. Ensure that transferring of this volume to your cells is feasible and adjust the production scale accordingly.
10. It is important to consider whether using a SpCas9-expressing cell clone or a SpCas9-expressing cell population best suits your library setup. By choosing a setup based on a SpCas9 clone, even SpCas9 expression within each cell is ensured during the screen. However, clonal lines will differ from your population.
11. 500 times representation of each sgRNAs provides sufficient coverage for most screens. However, a coverage ranging between 300 and 1000 will work depending on the specific screen setup. In general, a higher coverage is preferable in screens with weaker selection methods or depletion-based selection. On the other hand, some cell types, like primary cells, may be hard-to-transduce and expand, which will automatically result in a lower coverage. The number of transduced cells may be altered in accordance with the specific cell type used in your screen. If your cells grow slowly or you desire to carry out a multitude of different phenotypic selections immediately after puromycin selection, consider seeding and transducing more cells from the beginning.

12. For cells in suspension, transduction can be carried out immediately after seeding on day 1. MOI denotes the multiplicity of infection which can be used as a guideline to control the number of transduction events within each cell. The process of transduction follows the poisson distribution; therefore, if more stringency is needed transduce at an MOI of 0.3. If multiple transduction events per cell is achieved due to a too high MOI, an elevated background and appearance of false positives may be observed in your screen.
13. Passaging of cells at any time during the screen has certain thresholds never to exceed. First and foremost, the combined population of each replicate must never be split below 4.0×10^7 cells, since this represents the baseline of a population with $500\times$ coverage of the library. It should be noted that cells passaged at higher confluence than normal can introduce a condition-based bias, affecting sgRNA distribution in your population.
14. Genomic DNA extraction by salt precipitation does not usually result in a high loss during purification and is a cheap alternative to column-based purification, which is attractive in consideration of the large amount of DNA that must be extracted. However, column-based extraction is also a viable option.
15. The pellet will be large, and once resuspended it will result in a thick white suspension.
16. Re-suspension and dilution of genomic DNA in milliQ H₂O or equivalent grade water is crucial to PCR performance. Although DNA resuspended in water may be less stable than DNA resuspended in TE-buffer, the PCR relies so heavily on high efficiency that genomic DNA must be handled in milliQ H₂O. Dilute enough genomic DNA to perform the necessary PCR reactions and prepare aliquots of the undiluted DNA to be frozen for long-term storage.
17. Design of primers for PCR #2 is flexible, but make sure that you achieve good quality reads spanning the sgRNA sequence region. A standard design of primers compatible with Illumina sequencing platforms is available in the supplementary methods of [3]. Follow the link to www.genome-engineering.org for an excel spreadsheet with primer sequences.
18. Carryover PCR products and primers from PCR #1 is not a major problem as long as good separation on the gel is achieved. Figure 2 provides an example of PCR #2 products separated on a gel. Potential unspecific PCR products will reduce the amount of the actual product. However, by performing additional PCR#2 reactions, sufficient material can be obtained for NGS. It is important, though, to make sure complete separation is achieved, allowing preparation of a pure PCR fragment for NGS.

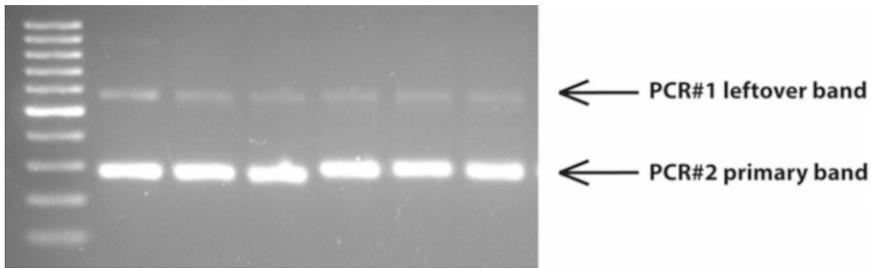


Fig. 2 PCR #2 gel purification. PCR #2 separated on 1.5% agarose gel, clearly displaying a carryover band from PCR #1

Acknowledgments

The authors are grateful to the laboratories of Feng Zhang and John G. Doench for making constructs and libraries available through Addgene. Genome-wide CRISPR screens performed in the laboratory of Jacob Giehm Mikkelsen are made possible through funding by the Danish Council for Independent Research | Medical Sciences (grant DFF-4004-00220), The Lundbeck Foundation (grant R126-2012-12456), the Novo Nordisk Foundation, Else og Mogens Wedell-Wedellsborgs Fond, Einar Willumsens Mindelegat, Emil C. Hertz og hustru Inger Hertz's Fond, Holger Hjortenbergs og Hustru Dagmar Hjortenbergs Fond, Krista og Viggo Petersens Fond, and Andersen-Isted Fonden.

References

1. 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
2. Boutros M, Ahringer J (2008) The art and design of genetic screens: RNA interference. *Nat Rev Genet* 9(7):554–566
3. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS et al (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343(6166):84–87
4. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF et al (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* 34(2):184–191
5. Ryø LB, Thomsen EA, Mikkelsen JG (2019) Production and validation of lentiviral vectors for CRISPR/Cas9 delivery. In: Luo Y (ed) *CRISPR gene editing*, *Methods in molecular biology*, vol 1961. Springer, New York
6. Sanjana NE, Shalem O, Zhang F (2014) Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 11(8):783–784
7. Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
8. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 17:10–12
9. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10(3):R25
10. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F et al (2014) MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* 15(12):554
11. Wu J, Huang B, Chen H, Yin Q, Liu Y, Xiang Y et al (2016) The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* 534(7609):652–657

INDEX

A

- AAV8.....295–298, 300
- AAV9.....297, 299–301
- AAVS1..... 187, 188, 190, 192, 203–205, 207
- Adeno-associated virus (AAV)..... 111–124, 133, 294–299, 301, 302
- Age-related macular degeneration (AMD).....299, 301, 308
- Alpha-1 antitrypsin deficiency (AIATD)..... 297
- Amplicon analysis..... 53, 58, 59, 333, 345
- Annealing..... 17, 31–35, 58, 71, 76, 121, 141, 148, 164, 168, 169, 173, 180, 192, 194, 197, 200, 221, 222, 234, 235, 240, 286, 312, 315, 317, 319, 323, 324
- Antiangiogenesis 307
- Artificial virus81–89

B

- Branched polyethylene imine (bPEI).....82–86
- BRCA1 v, 14

C

- Capillary electrophoresis..... 18, 47, 48, 53, 58, 59, 245
- Cas9 3, 15, 29, 45, 67, 81, 93, 111, 127, 137, 153, 186, 215, 233, 250, 256, 274, 294, 307, 330, 345
- Cas9 protein 15, 20, 21, 67, 108, 127–133, 154, 155, 161, 162, 171, 179, 186, 206, 233, 251, 252, 254, 274, 298, 300–302
- Cas9-responsive T cells 302
- Catalytically inactivated Cas9 22
- Cattle 271–287
- C-Check.....20, 25, 67–76, 268
- CCR5..... 49–51, 55, 129
- Cell lysis52, 57, 121, 238, 243, 244, 313
- Chemically modified sgRNAs..... 127–133
- Chimeric RNA (crRNA)..... 15
- Chinese hamster ovary (CHO) cells 213–231
- Choroidal neovascularization (CNV) 301
- Clonal isolation 30, 144, 146–147, 187, 204, 206

- Cloning..... 8, 50, 68, 72, 73, 122, 140–142, 147, 148, 155, 160, 188, 190–194, 206, 219–226, 231, 256–257, 263, 264, 266, 280, 285, 308, 309, 314–317, 324, 327, 333
- Clustered regularly interspaced short palindromic repeats (CRISPR)..... 4, 15, 29, 45, 67, 81, 93, 111, 127, 137, 153, 186, 214, 234, 249, 256, 274, 293, 307, 330, 345
- Clustered regularly interspaced short palindromic repeats/CRISPR-associated (CRISPR/Cas)v, vi, 4, 15, 214, 255, 271, 272, 274
- Colony-forming units (CFU)..... 100–101
- Colony PCR screening202, 241–244, 261
- Colony picking164–167, 181
- Conditional gene knockout..... 185–207
- Copy number analysis 218–220, 230–231
- CpfI 3–5, 7, 8
- CRISPR RNA (crRNA)5, 154, 155, 158, 159, 161, 171, 186, 215, 250–252, 271
- Cryosectioning 313, 314, 321, 322, 327

D

- Decomposition..... 18, 30, 36–42, 308, 326, 338
- Deletions..... 14, 15, 17, 19, 21–23, 29, 38, 42, 46, 114, 138, 144, 154, 160, 186, 233–247, 296, 298, 300, 308, 309, 324, 327, 330, 334, 343
- Delivery..... vi, 5, 16, 22, 25, 47, 52, 56, 57, 81–89, 93–108, 112, 127, 128, 130, 131, 133, 138, 170, 203, 215, 231, 234, 294–298, 300–303, 307–327, 329, 330, 335, 344
- DNA endonuclease v, 4, 67, 186, 233
- DNA extraction.....52, 57, 218, 228, 257, 262, 273, 274, 311, 317, 346, 350–351, 356
- Donor DNA 39, 250–253, 294
- Donor plasmid construction 216–217, 220–226
- Double-strand breaks (DSBs) 4, 13–16, 20–23, 25, 29, 31, 39, 46, 47, 68, 69, 137, 143, 144, 154, 160, 173, 186, 215, 233, 249, 255, 256, 271, 274
- Droplet digital PCR (ddPCR) 19
- Duchenne muscular dystrophy (DMD) 298, 299

E

Electroporation 22, 127–133, 143, 149, 156, 231, 249–253, 294, 299, 346, 347, 354
Escherichia coli 69, 72, 140, 141, 156, 192, 197, 216, 217, 222, 226, 236–238, 241, 242, 246, 311, 324
 Expansion 146, 169, 170, 181, 200, 204, 218, 228, 230, 296
 Ex vivo 45–64, 293, 294, 302
 Expression 5, 50, 69, 94, 112, 127, 138, 158, 186, 213, 234, 256, 294, 321, 330, 349

F

Fibroblasts v, 101, 102, 105, 133, 273, 275
 Flow cytometry 52, 54, 70, 73, 75, 131, 133, 139, 141, 144–146, 192, 201, 243
 Fluorescence-activated cell sorting (FACS) 25, 75, 122, 140, 143–145, 150, 216, 218–221, 228–229, 231, 243, 300, 308, 309, 313, 317, 322, 323, 327
 Fluorinated PEI 83–84
 Freezing cells 108, 120, 140, 166, 167, 181

G

GAPDH 258, 268
 Gene editing v, vi, 3, 7, 9, 13–18, 20, 22, 25, 26, 46, 48, 67, 72, 81, 89, 122, 127–133, 137–150, 233, 234, 249, 268, 271–287, 294–298
 Gene fusion 257, 267, 330
 Gene targeting 16, 137, 141–143, 146, 150, 157–163, 189, 192, 195–199, 204, 255, 256, 262
 Gene therapy vi, 293–303, 307–327, 329
 Genome editing 4–7, 20, 29–40, 89, 93, 101, 111, 112, 137, 139, 143, 156, 186, 213–231, 233, 249–254, 293–295, 298–303, 322–323, 333
 Genome-wide screening 349
 Genomic DNA (gDNA) 16, 21, 29, 31, 32, 34, 39, 58, 63, 72, 76, 99, 103, 104, 132, 146, 180, 194, 200, 205, 218, 221, 229, 230, 238, 277, 279, 286, 309–311, 317–319, 322–324, 333, 337, 343, 346, 350–351, 356
 Genomic screening 343–356
 Genotyping 18, 19, 21, 60, 138, 146, 154, 156, 160, 166–171, 181, 190, 194, 195, 199–201, 204, 234, 237, 244, 245, 267, 344
 GFP mRNA 128–131, 133
 Gibson assembly 142, 143, 205
 Golden-Gate Assembly 68, 235, 239, 240, 246, 260

Guiding RNA (gRNA) 7, 15, 30, 50, 67, 161, 234, 249, 256, 272, 299, 333
 assembly 234, 235, 239, 240, 259–261
 design vi, 3–9, 50, 234, 235, 239–240, 256, 333
 efficiency 5

H

HeLaP4-NLtr cells 112, 114, 121
 Hemophilia B 295
 Hepa1–6 cells 112, 114, 121
 Hereditary tyrosinemia (HT) 296
 hESCs 157, 185, 191, 192, 203, 204
 High resolution melting (HRM) 17–19
 High-throughput sequencing (HTS) 9, 154, 156, 166
 Homologous recombination (HR) v, 137, 138, 190, 255, 256, 258, 279, 294
 Homology arms (HAs) 8, 39, 188, 214–216, 220, 221, 223–225, 229, 258, 262, 263
 Homology-directed repair (HDR) 7, 8, 13, 14, 19, 29, 30, 36, 39, 68, 133, 138, 143, 144, 154–156, 158, 159, 161, 170, 171, 179, 188, 214, 215, 233, 256–258, 262, 297
 Human embryonic kidney 239T (HEK293T) 23, 24, 70, 73, 74, 76, 94–97, 106, 117, 118, 245, 246
 Human immunodeficiency virus type 1 (HIV-1) 93, 94, 114, 325
 Human induced pluripotent stem cells (hiPSCs) 137–150, 153–181, 185, 203, 204
 Human induced pluripotent stem cells (hiPSCs) culture 158–161
 Hydrodynamic tail vein injection 329–340

I

i53 143, 144
 Immune responses 82, 155, 301–302
 Immunostaining 256, 313, 314, 322
 Indel detection by amplicon analysis (IDAA) 17, 18, 20, 45–64, 334, 337
 Induced pluripotent stem cell (iPSC) 144, 149, 150, 154, 158, 166, 172
 Inducible knockout 186, 187, 189
 Insertion-deletions (indels) 6, 15, 30, 46, 69, 101, 114, 130, 137, 170, 186, 214, 234, 267, 296, 308, 333, 343
 frequency 15–19, 69, 102, 104, 105, 132, 309, 310, 319, 326, 327
 profiling 17, 46–48, 53, 57, 59, 62
 Insertions 14, 15, 18, 19, 21, 29, 36, 38, 39, 42, 46, 68, 103, 105, 138,

154, 160, 186, 234, 253, 262, 294, 298, 308,
 309, 311, 316, 317, 320, 324, 326, 327, 334,
 336, 337, 339, 343
 In vitro transcription 5, 7, 49–52,
 55, 56, 60, 61, 128
 In vivo genome editing 295, 299,
 301, 303, 322–323
 Iodixanol gradient centrifugation 114, 116, 117, 119

K

Knockin (KI) 138, 143, 144,
 147, 255, 268, 330
 Knockout (KO) v, 14, 95, 101–105,
 137, 138, 144, 156, 185–207, 214, 234, 255,
 293–298, 300, 301, 303, 307–327, 330, 334,
 343–356
 Knockout libraries 343–356

L

Lentiviral 93–108, 294, 301,
 307–327, 343–356
 Lentiviral vectors (LVs) 93–108, 294,
 301, 307, 345, 348–349
 Lentivirus 22, 82, 93, 312, 316
 Lentivirus production 312, 316–318
 Ligation 71, 72, 141, 197, 202,
 240, 241, 260, 311, 315, 316, 325, 352
 Lipofection 22, 138, 140, 149
 Liver toxicity 296, 297
 Long terminal repeat (LTR) 93, 114, 355
Loxp 255

M

Meganucleases (MGNs/MNs) 14, 15, 46
 Mice 249–254, 297–301,
 313, 321, 322, 331–335, 337–339
 Microhomology-mediated end joining (MMEJ) .. 13, 14,
 16, 22, 23, 25, 68, 234
 Mouse embryos 70, 186, 250
 Mouse livers 295–298, 303, 329–340
 Muscular dystrophies 298–299
 Mutagenesis 34, 35, 155, 170, 172, 330, 334

N

Next-generation sequencing (NGS) 20, 46, 48,
 200, 277, 280, 285, 322, 334, 337, 345,
 351–354, 356
 Nickases 4, 6, 9
 Non-homologous end joining (NHEJ) 13, 14,
 16, 19, 22, 46, 68, 104, 137, 138, 143, 154, 186,
 214, 233, 234, 256, 277, 285, 287, 343

O

Off-targets 4–9, 15, 21, 22,
 25, 81, 127, 128, 155, 158, 190, 206, 207, 219,
 239, 249, 259, 275, 280, 287, 308, 309, 315, 324
 On-target 6, 8, 9, 159, 187,
 245, 299, 301, 308, 309, 315, 324
 Optimized inducible gene Knockout
 (OPTiKO) 186–191, 195, 199–203, 206

P

PBMC isolation 50, 53, 54
 PCR 7, 16, 30, 46, 69, 94,
 117, 130, 142, 157, 187, 216, 234, 257, 273,
 311, 333, 345
 PCR amplification 17, 20, 31–33,
 70, 142, 147, 158, 190, 216, 263, 265, 316, 317,
 323, 347, 351–353
 Peripheral blood mononuclear cells (PBMCs) 47, 49,
 50, 52, 54, 56, 57, 60
 Pigs v, 256, 259, 271–287
 Plasmid 7, 16, 39, 55, 69,
 81, 94, 127, 138, 155, 187, 214, 234, 254, 257,
 274, 296, 308, 330, 346
 Plasmid amplification 346–348, 355
 Polyethylenimine (PEI) 83–84, 89,
 113, 116
 Primary cells vi, 50, 52, 112,
 127, 157, 192, 271–287, 355
 ProfileIt 47–49, 53, 59, 63, 64
 Pronuclei 250
 Pronucleus injection 250, 251
 Protospacer-adjacent motif (PAM) 3, 4, 6, 25,
 67, 71, 154, 155, 158–160, 172, 186, 194, 215,
 219, 221, 223, 233, 239, 258, 259, 268, 274, 300

R

Reference DNA 32–33, 35
 Repair templates 8, 149, 154, 156,
 157, 161, 171, 214, 215
 Reporter vectors 19, 20, 68
 Retina 299–301, 307–327
 Retinal pigment epithelial (RPE) cells 301, 308,
 309, 313, 317, 318, 321–324, 327
 Retinitis pigmentosa (RP) 299, 300
 RGD-R8-PEG-HA (RRPH) 83–87, 89
 Ribonucleoproteins (RNPs) 16, 47, 49,
 56, 57, 62, 128, 130–132, 153–181, 251, 254,
 271, 294, 301
 RNA duplex buffer 250–252
 RRRPHC 81–89

S

Sanger sequencing..... 9, 18, 19, 30, 33–36, 46, 132, 142, 143, 156, 170, 200, 224, 229, 234, 235, 238–239, 242, 261, 267, 268, 274, 277, 279, 280, 285, 286

Serotypes..... 114, 115, 122, 294, 300

Single cell clones 154, 171, 243, 272–274, 277–283, 285–287

Single guide RNAs (sgRNAs) 6, 39, 52, 94, 128, 138, 154, 186, 215, 297, 307, 344

Single-strand annealing (SSA) 13, 14, 19, 20, 23, 68, 72, 75, 234

Single-strand breaks (SSBs) 4

Single-stranded oligodeoxynucleotides (ssODNs) 7, 8, 39, 138, 144, 149, 154, 156, 157, 160, 161, 171, 173, 179

Somatic cell nuclear transfer (SCNT) v, 272, 277, 286

SpCas9-expressing cell lines 104, 349

Staphylococcus aureus Cas9 (SaCas9)..... 5, 122, 296–299, 301, 302

Streptococcus pyogenes Cas9 (SpCas9) 4, 15, 16, 23, 24, 52, 56, 67, 71, 94, 101–104, 114, 154, 158, 272, 274, 294, 296, 297, 299–302, 307, 308, 314–315, 331, 343, 355

Subcloning..... 162–166, 180, 181

Subretinal injections..... 299, 301, 308, 309, 313, 321, 325, 326

Surrogate reporter..... 19, 20

Surveyor nuclease assay..... 9

Synergistic activation mediator (SAM) 299

Synthetic gRNA 49, 52, 56

T

T7 endonuclease 1 (T7E1) 16, 20, 62, 112, 114, 117, 121, 123, 154, 156, 157, 159, 162, 163, 165, 168, 171, 180, 191, 200, 202, 231, 257, 259, 262, 268

TeSR-E8 156, 158, 161, 162, 164–166, 169, 170, 181, 195, 199

Tetracycline (TET)..... 186–189, 191, 194, 200, 205, 206

Titers..... 98–101, 103, 106, 114, 120, 312, 318, 325, 348, 355

Titration..... 94, 95, 98, 99, 349

Tracking of indels by decomposition (TIDE)..... 17–20, 23–25, 29–40, 46, 95, 102, 104, 105, 108, 132,

163, 180, 277, 285, 308–310, 314, 319, 323, 326, 327, 337, 338

Tracking of insertions, deletions, and recombination events (TIDER)..... 18, 29–40

Trans-activating crRNA (tracrRNA)..... 154–156, 161, 186, 215, 250–252

Transcription activator-like effector nucleases (TALENs) v, 15, 45, 67, 81, 111, 128, 214, 255, 256, 272

Transduction 5, 16, 22, 98–104, 107, 108, 112, 113, 121, 123, 296, 297, 300, 301, 309, 312, 318, 320, 349, 350, 356

Transfection 9, 16, 30, 31, 62, 69, 70, 73, 76, 82, 83, 85, 89, 94–97, 106, 113, 115–118, 120, 122, 123, 127, 128, 138, 140, 143–145, 149, 150, 155, 217, 224, 227, 231, 234, 238, 242, 243, 246, 261, 262, 266–268, 272–274, 276, 284, 331, 334, 338

Transformation 72–73, 141–143, 149, 224, 236–237, 241, 260, 311, 316, 325

Trex2 143, 144

U

Ultracentrifugation 95–97, 107, 113, 115, 119, 123, 316, 325

USER cloning 219–226, 231

V

Vascular endothelial growth factor A (VEGFA) 301, 307–327

Vector generation 70–72

W

Web tools..... v, 5, 18, 20, 25, 30, 34–36, 39, 234, 235, 239, 245, 256, 259

Western blotting..... 108, 201, 309, 314, 320–321

Whole-genome sequencing (WGS) 21

Woodchuck hepatitis virus posttranscriptional regulatory element (WPRE) 94, 98–100, 107, 108

Z

Zeta potential 85, 86

Zinc finger nucleases (ZFNs) v, 14, 15, 45, 67, 111, 127, 187, 189, 203, 207, 214, 255, 256, 272, 294

Zygotes 249, 251–253