

Award Number: W81XWH-15-1-0261

TITLE: Derivation of Parathyroid Gland Cells and Their Progenitors from Induced Pluripotent Stem Cells (iPSCs) for Personalized Therapy

PRINCIPAL INVESTIGATOR: Roy Geoffrey Sargent

CONTRACTING ORGANIZATION: University of California, San Francisco
San Francisco, CA 94103

REPORT DATE: OCTOBER 2017

TYPE OF REPORT: Final Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE OCTOBER 2017		2. REPORT TYPE Final		3. DATES COVERED 1 August 2015 - 31 July 2017	
4. TITLE AND SUBTITLE Derivation of Parathyroid Gland Cells and Their Progenitors from Induced Pluripotent Stem Cells (iPSCs) for Personalized Therapy				5a. CONTRACT NUMBER W81XWH-15-1-0261	
				5b. GRANT NUMBER PR141950 ^(SEP)	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Roy Geoffrey Sargent E-Mail: geoffrey.sargent@ucsf.edu current geoffreysargent@gmail.com				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) UNIVERSITY OF CALIFORNIA, SAN FRANCISCO 1855 FOLSOM ST STE 425 SAN FRANCISCO CA 94103-4249				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This research addresses metabolic pathologies resulting from hypoparathyroidism, parathyroid disease, and injury. The major goals of the project were to establish genetically modified primary parathyroid cell lines, develop protocols to differentiate pluripotent stem cell lines into parathyroid tissues using a genetically marked pluripotent stem cell line, and demonstrate efficient transplantation of in vitro differentiated parathyroid tissues in a hypoparathyroid mouse model. This report discusses the successful creation of an induced pluripotent stem cell line with a red fluorescent protein marker inserted into the GCM2 gene. GCM2 is a unique marker expressed in parathyroid tissues and serves as an easily detected reporter for cells that have differentiated into parathyroid tissues. A novel recombination product was identified in these experiment that has not been reported for iPS cell lines or for gene modifications using CRISPR/Cas9. Several growth conditions and substrates for growth of parathyroid primary cells have been tested that have supported short-term growth, but not apparent long term survival, of parathyroid primary cell lines in culture. Three different induced pluripotent stem cell lines have been used to test in vitro differentiation protocols to generate parathyroid tissue. These experiments have met with mixed success as demonstrated by the transient expression of parathyroid hormone and GCM2, both markers of parathyroid tissues.					
15. SUBJECT TERMS Induced pluripotent stem cells, ips cells, parathyroid, Crispr/cas9, TALENS, pluripotent stem cells, hypoparathyroidism, 2 human homolog (Gcm2/GCMB), parathyroid hormone (PTH) and calcium sensing receptor (CaSR)					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (include area code)
Unclassified	Unclassified	Unclassified	Unclassified	65	USAMRMC

Table of Contents

	<u>Page</u>
1. Introduction.....	1
2. Keywords.....	2
3. Accomplishments.....	2-8
4. Impact.....	8-9
5. Changes/Problems.....	9-10
6. Products.....	10
7. Participants & Other Collaborating Organizations.....	11-12
8. Special Reporting Requirements.....	12
9. Appendices.....	12-65

1. Introduction

This research addresses the metabolic pathologies resulting from hypoparathyroidism. Parathyroid endocrine function can be compromised or lost through injury, therapeutic intervention, e.g., surgery or radiation or as a result of disease (genetic or autoimmune), and can lead to parathyroid gland (PTG) dysfunction and a disruption of calcium homeostasis and metabolism. Hypoparathyroidism is particularly devastating and is becoming more and more prevalent because of the increase in thyroidectomies and thyroid radiotherapy. Due to the multi-organ pathological metabolic manifestations of hypoparathyroidism and the limitations of the present pharmacological therapies, it would be advantageous for patients if there were a cellular therapy that could mimic the native PTG responses to serum calcium fluctuations and compensate for lost PTG function. Induced pluripotent stem cells (iPSCs) have created a new opportunity to develop cellular therapies to repair tissues damaged by injury, therapeutic intervention or disease. This proposal will take a patient's somatic cells, reprogram them into iPSCs, and then differentiate the resultant iPSCs into parathyroid progenitors as an autologous/personalized cellular therapy that bypasses immune-rejection. This is an innovative strategy for restoring parathyroid function, since it bypasses the need for pharmacological supplementation and does not depend on rescuing remnant parathyroid glands for transplantation. The directed differentiation of iPSCs into parathyroid progenitors and mature PTG cells relies on exposure to conditions that mimic, to a degree, embryonic parathyroid organogenesis. The iPSCs are being modified with CRISPR or TALEN technology for sequence specific insertion of a GFP reporter into the parathyroid-specific glial cell missing 2 human homolog (Gcm2/GCMB), parathyroid hormone (PTH) and calcium sensing receptor (CaSR) regulatory sequences to monitor the differentiation of the iPSCs into parathyroid progenitors and mature PTG cells. The studies test the hypothesis that iPSCs can be directed to differentiate along an endodermal lineage pathway and progress to cells that have characteristics consistent with parathyroid cell function, e.g., secrete PTH in low Ca²⁺, GCMCB, express CaSR, express vitamin D receptor (VDR), recapitulate PTG function/morphology in an ex vivo bioactive matrix and a GATA3 heterozygote (+/-) mouse model of hypoparathyroidism. Functional efficacy and reversal of hypoparathyroidism in the GATA3 (+/-) mouse is critical to helping define the parameters for the generation and selection of the parathyroid progenitors for transplantation. This research described in this project lays the groundwork for developing a patient-specific cell-based therapy for hypoparathyroidism and provides the preliminary information to support a more in-depth proposal that could lead to future clinical trials.

2. Keywords

Induced pluripotent stem cells, ips cells, parathyroid, Crispr/cas9, TALENS, pluripotent stem cells, hypoparathyroidism, 2 human homolog (Gcm2/GCMB), parathyroid hormone (PTH) and calcium sensing receptor (CaSR)

3. Accomplishments

What were the major goals of the project?

<p align="center">Specific Aim 1:</p> <p>AIM 1 will involve (a) the establishment and characterization of primary cultures of parathyroid gland epithelial cells, and (b) the generation of a genomic GCMB/GCM2 promoter-GFP reporter iPSC line that will be directed to differentiate along an endodermal lineage specific pathway and that will be assessed of differentiation stage-specific gene expression features.</p>	<p align="center">Timeline</p>	<p align="center">Status</p>
<p>Major Task 1: The establishment and characterization of primary cultures of parathyroid gland epithelial cells.</p>	<p align="center">Months</p>	
<p>Subtask 1: Obtain tissues from patient specimens following surgery.</p>	<p align="center">1-18</p>	<p align="center">limited/unpredictable availability</p>
<p>Subtask 2: Optimize culture conditions for long-term culture of primary cells and characterize cells for parathyroid features.</p>	<p align="center">1-6</p>	<p align="center">unsuccessful due to limited availability of tissues</p>
<p>Subtask 3: Establish GFP-labeled primary PTG cells with high efficiency gene delivery systems</p>	<p align="center">3-9</p>	<p align="center">not completed due to limited availability of tissues</p>
<p>Local IRB/IACUC Approval IRB and GESCR approvals are in place. IACUC approval pending.</p>	<p align="center">1-18</p>	
<p>Major Task 2: Generation of a genomic GCMB/GCM2 promoter-GFP reporter iPSC line that will be directed to differentiate along an endodermal lineage specific pathway and that will be assessed of differentiation stage-specific gene expression features.</p>		
<p>Subtask 1: Generate CRISPR/Cas targeting construct for inserting GFP into GCMB promoter region of a Sendai virus reprogrammed iPSC cell line. Isolate a GFP containing clone by PCR screening or secondary drug selection.</p>	<p align="center">2-3</p>	<p align="center">completed</p>
<p>Subtask 2: Develop differentiation protocol for the progression of iPSCs along a lineage pathway</p>	<p align="center">3-12</p>	<p align="center">unsuccessful</p>

that will lead to parathyroid progenitors and their more mature descendants.		
Subtask 3: Monitor and characterize the differentiation of the cells as function of the expression of various stage-specific transcription factors and through expression of GCMB-GFP	3-12	not completed
Specific Aim 2: AIM 2 will be directed at developing a 3-D bioactive matrix that will promote functional PTG “organogenesis” <i>ex vivo</i> and evaluating a mouse model of hypoparathyroidism for increased parathyroid function by measuring changes in serum calcium and phosphate levels upon transplantation with iPSC-derived parathyroid progenitor cells.		
Major Task 3: Develop a 3D bioactive matrix that will promote functional <i>ex vivo</i> PTG organogenesis		
Subtask 1: Generate bioactive matrices to test for primary parathyroid cell growth and maintenance of function.	6-12	tested not completed due to limited availability of tissues
Subtask 2: Establish parathyroid spheroid culture.	5	not completed due to limited availability of tissues
Subtask 3: Generate a <i>GATA3^(+/-)Rag1^(-/-)</i> mouse model for organoid and iPSC-derived parathyroid progenitor transplantation. It is expected that 10 mice will be needed to make this animal with 2 crosses.	3-9	not completed
Subtask 4: Transplant and assess functional efficacy of the transplanted PTG gland tissue/cells. There will be 6 animals/ experiment and it is anticipated that a there will be 2 conditions (primary cells and iPSC-derived), with 10 experiments overall. Thus it is expected that there will be 120 animals required for the planned assessments	9-18	not completed

What was accomplished under these goals?

Specific Aim 1; Major Task 1: *The establishment and characterization of primary cultures of parathyroid gland epithelial cells.*

Subtask 1: Obtain tissues from patient specimens following surgery.

1 August 2016-31 July 2017 A constant supply of parathyroid samples from patients continued to be an ongoing problem that impacted our experiments in all Specific Aims and prevented progress on Specific Aim 2.. In addition to requiring parathyroid primary cell lines for

Specific Aims 1 and 2, fresh parathyroid tissue was needed for positive controls for immunostaining and quantitative PCR.

Subtask 2: Optimize culture conditions for long-term culture of primary cells and characterize cells for parathyroid features. Our initial experiments to optimize culture conditions focused on testing combinations of growth substrates and media and monitoring cell growth. This work was being performed by Dr Sargent and Cristina Barilla, a graduate student in the laboratory.

Subtask 3: Establish GFP-labeled primary PTG cells with high efficiency gene delivery systems Construction of the GFP-reporter vector for establishing GFP-labeled primary PTG cells was completed. The GFP targeting vector construction is similar to that of the RFP vector used in experiments discussed below.

Specific Aim 1; Major Task 2: *Generation of a genomic GCMB/GCM2 promoter-GFP reporter iPSC line that will be directed to differentiate along an endodermal lineage specific pathway and that will be assessed of differentiation stage-specific gene expression features.*

Subtask 1: Generate CRISPR/Cas targeting construct for inserting GFP into GCMB promoter region of a Sendai virus reprogrammed iPSC cell line. Isolate a GFP containing clone by PCR screening or secondary drug selection.

Forty-two puromycin resistant CF3iPS2 clones were analyzed by PCR to identify GCM2 homologous recombinants of which 21.4% (9 out of the 42 clones) were identified as homologous replacement events (Figure 1). However, the initial PCR analysis was consistent with 11.9% (5 out of the 42 clones) having a different homologous recombination product associated with insertion events (Figure 1) where the donor fragment and plasmid backbone are homologously inserted. Further PCR analysis confirmed that these 5 clones are homologous insertion products (Figure 2).

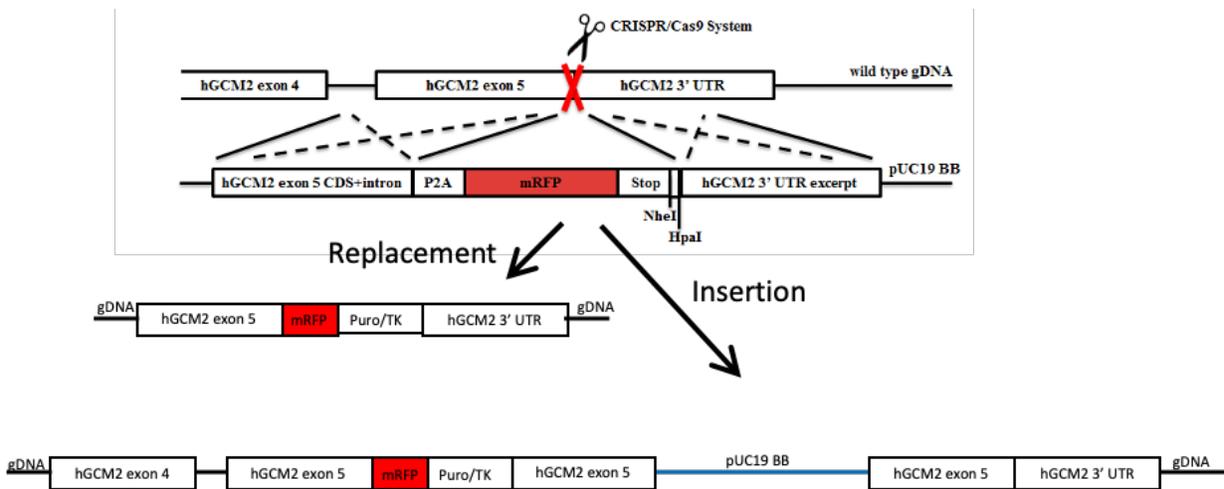


Fig. 1: GCM2 Donor DNA and gDNA Homologous Replacement Enhanced by CRISPR/Cas9 A scheme of homologous recombination between the donor DNA and wild type hGCM2 gDNA. The 5' homology arm is a 967 bp long sequence derived from amplifying the last exon of the wild type GCM2 gene and part of the intron immediately upstream of the coding sequence. The GCM2 stop codon was not included as part of the PCR amplicon. The 3' homology arm is a 925 bp long sequence derived from amplifying part of the 3' GCM2 untranslated region. The CRISPR/Cas9 system enhances homologous recombination through inducing DSB (double strand DNA break) around the stop codon. Replacement results in target genomic DNA replacement with donor fragment whereas insertion results in homologous insertion of the entire donor plasmid with a partial duplication of exon 5.

Clone: 5 25 15 21 4 6 11 12 31 32



Fig. 2: PCR Analysis of potential homologous recombinants for insertion products An inside-out PCR strategy was designed to amplify plasmid backbone linked to genomic DNA absent from the donor DNA vector. A PCR product will be amplified only when plasmid backbone and donor DNA is homologously inserted into the GCM2 target sequence at exon 5. With this strategy, there will be no PCR products amplified from homologous replacement events or random integration events. In the above gel, clones 5 and 25 are presumptive replacement products, clones 15 and 21 are presumptive random integrants, and clones 4, 6, 11, 12, 31, and 32 were identified in the initial screen as potential insertion products. Only clones 4, 6, 12, 31, and 32 have PCR products that positively identify as homologous insertion products. Clone 31 is a likely mixed colony resulting in a reduced amount of amplified product.

Subtask 2: Develop differentiation protocol for the progression of iPSCs along a lineage pathway that will lead to parathyroid progenitors and their more mature descendants.

See Subtask 3

Subtask 3: Monitor and characterize the differentiation of the cells as function of the expression of various stage-specific transcription factors and through expression of GCMB-GFP.

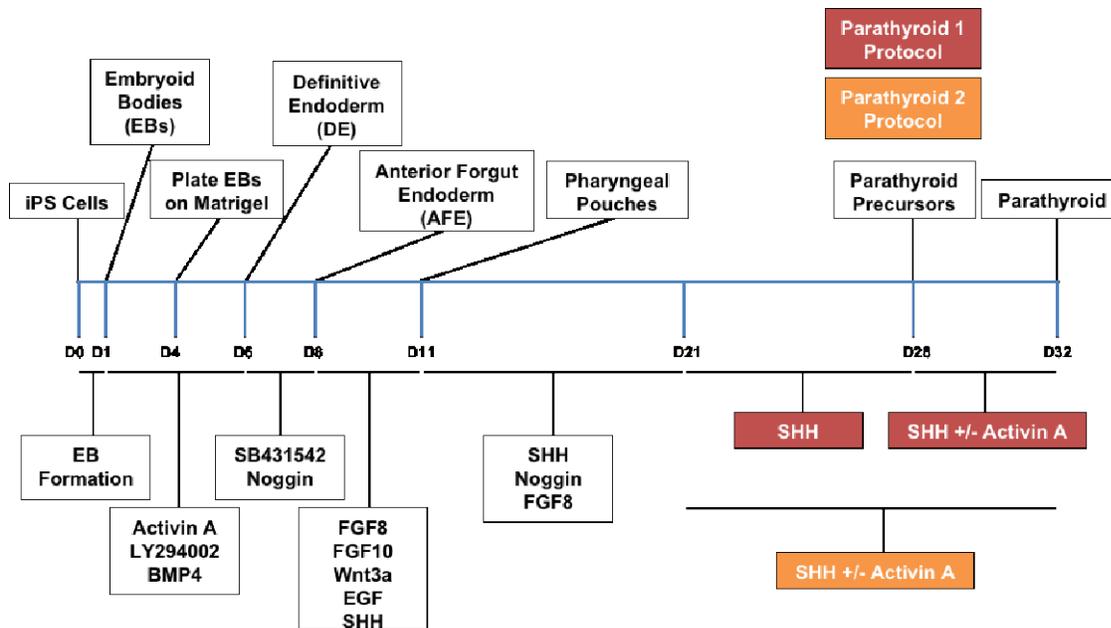


Figure 3: Differentiation Timeline

Subtasks 2 and 3 are closely related and summarized here.

We retested both differentiation protocols outlined in Figure 3 and measured gene expression levels of HoxA3, PBX1, Parathyroid hormone (PTH) and GCM2. HoxA3 and PBX1 are transcription factors expressed during development of the pharyngeal arches in definitive endoderm with HoxA3 expression detected in endodermal tissues of the 3rd and 4th pharyngeal arches about E28 and PBX1 expressed in parathyroid and thyroid progenitor cells at about E32. PTH and GCM2 are genes expressed in parathyroid tissues.

As previously described in our 2016 progress report, we continued to compare the quality of differentiation of three different human iPSC cell lines using embryoid bodies, as indicated in the differentiation timeline (Fig 3) without controlled aggregation (EB-aggregation) or by controlling the number of iPSC cells used to form EB using Aggrewell plates. Regardless of the method used for EB formation, addition of cytokines to the medium to initiate differentiation into parathyroid tissues resulted in cell death. This appears to be influenced by the iPSC cell line used as well as the method used for EB formation. Direct plating of iPSC cells onto tissue culture plates, however, showed the best long-term cell survival and was used for preliminary differentiation experiments into parathyroid lineages.

Gene expression analysis by quantitative PCR of HOXA3, PBX1, PTH, and GCM2 was used as markers for differentiation. Expression of HOXA3 and PBX1 are found in the pharyngeal pouch and surface ectoderm whereas GCM2 and PTH are characteristically expressed in parathyroid tissues. Both HOXA3 and PBX1 are transiently expressed at D11 in cells differentiated by a protocol leading to airway epithelium, whereas expression of these two transcription factors is weak or nonexistent in cells differentiated towards pharyngeal pouches and ultimately parathyroid tissues. Expression of PTH and GCM2 was observed at D8-D11, however expression of these two genes, characteristic of parathyroid tissues, diminished after D11 suggesting the CF3iPS2eGFPb5.1 cell line or the current differentiation protocol is not generating parathyroid tissue. The b5.1 cell line does not have a GFP reporter integrated in the GCM2 gene and we will use the GCM2-GFP modified cell line in future experiments.

Quantitative PCR was also reperformed on RNA samples from previous parathyroid differentiation experiments to quantify expression levels of PTH, HOX3A, and PBX1 expression. Expression of HOXA3 in protocol 1 only occurs in the presence activin A treatment whereas in the protocol 2 experiment. HOXA3 expression did not depend on activin A treatment. There is expression of GCM2 in cells differentiated by either protocol and at approximately the correct time suggesting protocols are proficient for differentiation cells into parathyroid tissue. Transient expression of parathyroid hormone was detected in a control protocol for differentiation of cells into airway epithelium suggesting that it might be possible to modify the airway epithelium differentiation protocol to generate parathyroid tissues.

These results were reproduced during the Fall of 2016. Subsequent research by Dr Suzuki has led to the suggestion that the critical step in the differentiation process is endodermal induction, which was not directly developed in the research funded by this grant. Furthermore, Dr Suzuki has determined the iPS cell lines used for the differentiation is critical. It is possible that our experiments were not successful in part from the iPS lines used here and due to inefficient endoderm induction.

Specific Aim 2; Major Task 3: Develop a 3D bioactive matrix that will promote functional ex vivo PTG organogenesis

Subtask 1: Generate bioactive matrices to test for primary parathyroid cell growth and maintenance of function.

The research for subtask 1 is incorporated into subtasks associated with Specific Aim 1. We have tested several bioactive matrices to establish long-term culture of parathyroid primary cells and will be assessing the influence of these matrices on differentiation of iPS cells into definitive endoderm and parathyroid cells. Due to the unpredictable and limited availability of human parathyroid tissue samples we were not able to make continued progress on testing bioactive matrices.

Subtask 2: Establish parathyroid spheroid culture.

Nothing to report. These experiments depended on a predictable and regular source of parathyroid tissues and development of differentiation protocols. We were unable to find a predictable source of surgical samples which impacted spheroid generation and as described above we were unable to develop reproducible differentiation of iPSC cell lines into parathyroid tissues.

Subtask 3: Generate a GATA3(+/-)Rag1(-/-) mouse model for organoid and iPSC-derived parathyroid progenitor transplantation. It is expected that 10 mice will be needed to make this animal with 2 crosses.

See subtask 4

Subtask 4: Transplant and assess functional efficacy of the transplanted PTG gland tissue/cells. There will be 6 animals/ experiment and it is anticipated that there will be 2 conditions (primary cells and iPSC-derived), with 10 experiments overall. Thus it is expected that there will be 120 animals required for the planned assessments

Not completed due to inadequate supply of human parathyroid tissue samples and inability to repeat initial differentiation results of iPSC cells into PTH-like cells. Budgetary and time limitations prevented hiring of new personnel when Yao, Chosa, Suzuki, and Barilla left for new positions or new programs.

What opportunities for training and professional development has the project provided?

Drs Sargent and Suzuki were responsible for mentoring students (Barilla, Chosa, and Yao) for research activities associated with the grant.

How were the results disseminated to communities of interest?

See attached manuscript to be submitted to academic journal. TBD

What do you plan to do during the next reporting period to accomplish the goals?

Not applicable/nothing to report.

4. Impact

What was the impact on the development of the principal discipline(s) of the project?

Nothing to report.

What was the impact on other disciplines?

As part of our gene modification introducing the GFP reporter into the GCM2 promoter, we discovered that many of the gene targeting products were not gene replacement events but rather were gene insertion events. These latter gene insertion events have been observed in

experiments in our laboratory at in the CFTR and beta-hemoglobin genes, but have not been previously described in iPS cells or for CRISPR/Cas9 treated cells. This is a novel DNA repair product that potentially allows generation of seamlessly modified cell lines on removal of the vector backbone sequences in a second step treatment with CRISPR/Cas9 directed to the vector backbone/genomic DNA sequence junction as described in the attached manuscript in the appendix.

What was the impact on technology transfer?

Nothing to report. No patent applications were filed.

What was the impact on society beyond science and technology?

Nothing to report

5. Changes/Problems

Changes in approach and reasons for change

None to report.

Actual or anticipated problems or delays and actions or plans to resolve them

During the transition period, we did not have access to human parathyroid surgical samples to complete studies on establishing primary human parathyroid cell cultures and identifying bioactive matrices to support parathyroid cell growth. Supply of human parathyroid samples continued to be unpredictable and infrequent which prevented our ability to perform the spheroid studies, construct GFP labelled primary human parathyroid cell lines, and construct the mouse models. The infrequent supply of human tissues not only prevented progress on the Specific Aims, it complicated analysis of differentiation experiments due to missing positive controls for quantitative PCR and immunostaining.

Furthermore, my research staff left for new positions well before the end of the grant. Cristina Barilla and Keisuke Chosa returned to their home institutions to complete their PhD programs, Michael Yao completed his academic year and entered a bachelor's program, and Dr Shingo Suzuki took a new postdoctoral position in Italy.

Changes that had a significant impact on expenditures

Nothing to report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Significant changes in use or care of human subjects

Nothing to report.

Significant changes in use or care of vertebrate animals.

Mouse model was not created due to limited human tissue samples required for spheroid development and inefficient differentiation of iPS cell lines to parathyroid cells.

Significant changes in use of biohazards and/or select agents

Nothing to report

6. Products

Publications: See attached manuscript entitled:

Seamless gene correction in the human cystic fibrosis transmembrane conductance regulator locus by vector replacement and vector insertion events, Shingo Suzuki, Keisuke Chosa, Cristina Barillà, Michael Yao, Orsetta Zuffardi, Hirofumi Kai, Tsuyoshi Shuto, Mary Ann Suico, Yuet W Kan, R Geoffrey Sargent, and Dieter C Gruenert

Patent Applications

No patent applications were generated from this work.

7. Participants & Other Collaborating Organizations

What individuals have worked on the project?

Name	Dieter C Gruenert (deceased)
Project Role	Ex-PI
Research Identifier	NA
Nearest Person Month Worked	0
Contribution to Project	Dr Gruenert was the Principle Investigator and died April 10, 2016
Funding Support	National Institutes of Health-NIDDK 1P01DK08876001A1 and National Institutes of Health R01DK104681

Name	Roy Geoffrey Sargent
Project Role	PI
Research Identifier	NA
Nearest Person Month Worked	12
Contribution to Project	Dr Sargent is the current PI and assumed responsibility for the grant in June, 2016.
Funding Support	1P01DK08876001A1, National Institutes of Health-NIDDK, Development of iPS Cells for Treatment of Hemoglobinopathies, Kan (PI), No Cost Extension.

Name	Shingo Suzuki
Project Role	Post Doctoral Fellow
Research Identifier	NA
Nearest Person Month Worked	6
Contribution to Project	GCMB/GCM2 reporter-gene knockin and mentoring Keisuke Chosa, Cristina Barilla, and Michael Yao
Funding Support	Postdoctoral fellowship from Japan

Name	Cristina Barilla
Project Role	Graduate Student
Research Identifier	NA
Nearest Person Month Worked	6

Contribution to Project	Defining growth conditions for human primary parathyroid tissues and differentiation of parathyroid tissues from pluripotent stem cells.
Funding Support	Graduate student fellowship from University of Padua

Name	Keisuke Chosa
Project Role	Graduate Student
Research Identifier	NA
Nearest Person Month Worked	9
Contribution to Project	
Funding Support	Graduate student fellowship

Name	Michael Yao
Project Role	Volunteer/Intern
Research Identifier	NA
Nearest Person Month Worked	4
Contribution to Project	GCMB/GCM2 reporter-gene vector construction and design of CRISPR/Cas9 oligonucleotides
Funding Support	Laboratory volunteer

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Dr Sargent is no longer employed by UCSF and has no current grant support.

What other organizations were involved as partners?

Nothing to report

8. Special Reporting Requirements

Nothing to report

9. Appendices

Manuscript attached.

**Seamless gene correction in the human cystic fibrosis transmembrane conductance regulator locus
by vector replacement and vector insertion events**

Shingo Suzuki^{1,2,3 #, †}, Keisuke Chosa^{1,4, #}, Cristina Barilla^{1,5}, Michael Yao¹, Orsetta Zuffardi⁵, Hirofumi Kai⁴, Tsuyoshi Shuto⁴, Mary Ann Suico⁴, Yuet W Kan^{2,6,7}, R Geoffrey Sargent^{1,8,10 †}, and Dieter C Gruenert^{1,6,8,9, ††}

1. Department of Otolaryngology–Head and Neck Surgery, University of California–San Francisco, San Francisco, California, 94115, USA;

2. Department of Medicine, University of California–San Francisco, San Francisco, California, 94143, USA;

3. Center for Stem Cell and Regenerative Medicine, Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, TX 77030;

4. Department of Molecular Medicine, Graduate School of Pharmaceutical Sciences, Kumamoto University, Kumamoto, 862-0973, Japan;

5. Department of Molecular Medicine, University of Pavia, Pavia, 27100, Italy;

6. Institutes for Human Genetics, University of California, San Francisco, California, 94143, USA;

7. Department of Laboratory Medicine, University of California, San Francisco, California, 94143, USA;

8. California Pacific Medical Center Research Institute, San Francisco, California, 94115, USA;

9. Department of Pediatrics, University of Vermont College of Medicine, Burlington, Vermont, 05405, USA.

10. Onconetics Pharmaceuticals, 1477 Via Manzanitas, San Lorenzo, CA 94580

#: The authors equally contributed to this works.

†: Corresponding authors

††: This manuscript is in honor of our colleague, friend, and mentor Dieter C Gruenert.

Email address

(S Suzuki) shingo.suzuki@uth.tmc

(K Chosa) drug8street4@gmail.com

(C Barilla) crisrina.barilla01@universitadipavia.it

(M Yao) michaelyao2017@gmail.com

(O Zuffardi) orsetta.zuffardi@unipv.it

(H Kai) hirokai@gpo.kumamoto-u.ac.jp

(T Shuto) tshuto@gpo.kumamoto-u.ac.jp

(Mary Ann Suico) mann@gpo.kumamoto-u.ac.jp

(YW Kan) yw.kan@ucsf.edu

(R Geoffrey Sargent) geoffreysargent@gmail.com

(DC Gruenert) deceased

†: Corresponding authors' contact information

(R Geoffrey Sargent) Tel/Fax, +1-510-207-2046; E-mail, geoffreysargent@gmail.com; Address, Onconetics Pharmaceuticals, Inc., 1477 Via Manzanos, San Lorenzo, CA 94580

(S Suzuki) Tel/Fax, +1-713-500-3438; E-mail, shingo.suzuki@uth.tmc; Address, Center for Stem Cell and Regenerative Medicine, Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, TX 77030

Abstract

Background: Gene correction via homologous recombination (HR) in patient-derived induced pluripotent stem cell (iPSCs) for regenerative medicine are becoming a more realistic approach to develop personalized and mutation-specific therapeutic strategies due to current developments in gene editing and iPSC technology. Cystic fibrosis (CF) is the most common inherited disease in the Caucasian population, caused by mutations in the CF transmembrane conductance regulator (CFTR) gene. Since CF causes significant multi-organ damage and with over 2,000 reported CFTR mutations, CF patients could be one prominent population benefiting from gene and cell therapies. When considering gene-editing techniques for clinical applications, seamless gene corrections of the responsible mutations would be the most desirable approach.

Result: The studies reported here describe the generation of iPSCs from a CF patient homozygous for the W1282X Class I CFTR mutation and the seamless correction of the W1282X CFTR mutation using CRISPR/Cas9 nickase (Cas9n). In addition to the expected HR vector replacement product, we discovered another class of HR products resulting in vector insertions with a partial duplication of the CFTR exon23 sequence. This vector insertion was furthermore removed via intrachromosomal homologous recombination (IHR) enhanced by double nicking with CRISPR/Cas9n which resulted in the seamless correction of CFTR exon23 in CF-iPSCs.

Conclusion: We show here that removal of the drug resistance cassette and generation of seamless gene corrected cell lines by two independent processes: by treatment with the PiggyBac (PB) transposase in replacement product or by IHR between the tandemly duplicated CFTR gene sequences.

Keywords

Cystic fibrosis, iPS cells, Seamless gene correction, Homologous recombination, Vector replacement event, Vector insertion event, Intrachromosomal homologous recombination

Background

Cystic fibrosis (CF) is the most common inherited disease in the Caucasian population, caused by mutations in the CF transmembrane conductance regulator (*CFTR*) gene [1]; over 2,000 disease-causing *CFTR* mutations have been reported (The Clinical and Functional TRanslation of *CFTR* (*CFTR2*) database; available at <http://cftr2.org>). [2] CF patients typically exhibit mucus accumulation, which causes multi-organ damages in many of tubular organs such as lungs, pancreas, liver, kidneys, and intestine. Especially the airway and lung in CF have severe pathologies caused by abnormal mucus accumulation and hyper inflammation accompanied by bacterial infections [3, 4]. Therefore, the treatment for CF patients would require a comprehensive strategy that both correct the underlying genetic defect and repair/regenerate damaged tissues. With this concept, we and others have been developing and have achieved the *CFTR* mutation correction in iPSCs derived from a patient carrying the homozygote *F508del CFTR* mutation found in around 70% of CF allele. Also, these corrected CF-iPSCs were directed into airway epithelium and showed the restoration of *CFTR* function, which is Cl⁻ ion transportation [5-7].

W1282X CFTR is a nonsense mutation of the codon encoding the 1282th *CFTR* amino acid, Tryptophan (Trp, W), at exon23, and is caused by G>A transition to create a stop codon. Thus, *W1282X CFTR* synthesizes a truncated *CFTR* protein, classified as Class I mutation, which exhibits little or no function. The *W1282X* mutation is one of the common *CFTR* mutations following the *F508del* mutation, and approximately 2% of patients with CF have at least 1 copy of the *W1282* mutation (*CFTR2* database). The *W1282X* mutation is found at a high frequency among Ashkenazi-Jewish and Middle Eastern populations, and is observed in approximately 20% of CF patients from those populations [8-10]

Gene editing technologies have been developed very rapidly especially after developments of DNA sequence-specific nucleases such as ZFNs [11], transcription activator-like effector nucleases (TALENs) [12, 13], and the RNA-guided CRISPR-Cas9 nuclease [14, 15]. These nucleases introduce double-strand breaks (DSB) in a targeting sequence, which enhances genome modification by utilizing cellular DNA repair processes such as homologous recombination (HR) and non-homologous end joining (NHEJ). When considering clinical applications, seamless gene corrections of the responsible mutations would be the most desirable approach. Recently, two approaches have been utilized to achieve seamless correction in iPSCs. One approach utilizes HR for gene targeting and the PiggyBac (PB) transposon system to remove drug selectable markers. In this approach, an HR-targeted allele is modified by the desired donor sequence and a selectable marker is inserted that can be removed seamlessly upon the expression of PB transposase (PB_{ase}). This two-step system with the replacement and excision is very common and is widely used to target a variety of genes and cell types [16, 17]. Another is a one-step strategy via short homologous oligo, both single- and double-strand DNA, without a selectable marker as donor DNA, followed by the clonal isolation of modified clone with the limiting dilution technique [18, 19]

In this study, we generated iPSCs from a CF patient homozygous for the *W1282X*, Class I *CFTR* mutation, and performed seamless gene correction in it using CRISPR/Cas9n minimalizing off-target events [2], and PB transposon system to prove the concept that *CFTR* mutations other than *F508del* in patient-derived iPSCs can be targeted. We also developed an alternative strategy to carry out seamless gene correction via an insertion of whole plasmid donor DNA. This novel approach would expand the flexibility to generate seamless HR targeted cell lines in both basic research and gene targeting therapy.

Results

Generation of iPSCs from a CF patient homozygous for the *W1282X* mutation (CF2-iPSCs)

Primary nasal polyp fibroblasts (CFNPF7) from a CF patient homozygous for the *W1282X* mutation were reprogrammed by transduction with four individual retroviruses, each containing one canonical

transcription factor (*OCT4*, *SOX2*, *KLF4*, or *c-MYC*) and grown for 3-4 weeks on mitomycin C-inactivated mouse embryo fibroblast feeders until candidate CF-iPSCs colonies appeared [7]. Ten candidate CF-iPSCs colonies were isolated and two clonal cell lines, clone 3 and clone 9, were selected for further analysis and experimentation. The *W1282X/W1282X CFTR* genotype in *CFTR* exon23 in each clone was confirmed by DNA sequence analysis (Additional file 1: Fig S1A). Cytogenetic analysis of clone 3 and clone 9 between P8.9-P8.12 (where passage number PX.Y.etc = X passages before transduction/reprogramming. Y passages since candidate colony isolation) showed a normal diploid male karyotype (46, XY) (Additional file 1: Fig S1B). Immunocytochemical analysis showed that clone 3 and clone 9 expressed the pluripotent markers NANOG, SSEA4, TRA1-60, and TRA1-81 (Additional file 1: Fig S1C). Pluripotency was further demonstrated in vitro by formation of embryoid bodies (EBs) and immunocytochemical detection of the 3 germ layer markers α -fetoprotein (AFP, endoderm), α -smooth muscle actin (SMA, mesoderm), and β -tubulin 3 (TUJ1, ectoderm). (Additional file 1: Fig S1D). Thus, we designated clones 3 and 9, as CF2-iPS3 and CF2-iPS9 cells, respectively.

***W1282X CFTR* gene targeting in CF2-iPSCs via HR enhanced by CRISPR/Cas9n double nicking**

Homologous recombination (HR) between gene targeting vectors and their chromosomal targets can result in at least two classes of recombination events, vector replacement events or vector insertion events, depending on the site of vector linearization [20, 21]. Recently, targeted vector insertions at homologous chromosomal targets by NHEJ pathways have been reported, including insertion of whole targeting plasmid DNA, (PITCh (Microhomology-mediated end-joining, MMEJ) [22], ObLiGaRe (NHEJ) [23]), homology-independent targeted integration (HITI) [24] and insertion in ESCs (NHEJ [25]). The studies described here demonstrate that gene targeting using non-linearized vectors can result in both vector replacement and vector insertion events by homologous recombination.

Homologous recombinant vector replacement events

Two gene-targeting vectors were constructed with different lengths of homology to the genomic target. Both vectors contain a CAG-Puro Δ TK drug selection cassette flanked by DNA sequences homologous to the *CFTR* exon23 region and both vectors contain the TTAA repeat found in intron 23 for excision of the CAG-Puro Δ TK cassette by PiggyBac (PB) transposase (Fig 1A). The CF2A targeting vector contains a total of 1,366-bp of *CFTR* homologous sequence with a 5' arm of 804-bp homologous *CFTR* sequence and a 3' arm of 566-bp homology to the *CFTR* exon23 region. The CF2B targeting vector contains 2,355-bp of *CFTR* homologous sequence with a 5' arm of 1,207-bp and 3' arm of 1,148-bp of homology, respectively (Fig 1A).

Four pairs of guide RNAs (gRNAs) in CRISPR/Cas9n expression vectors (CRISPR/Cas9n-gRNAs) were designed to cleave at specific genomic DNA sequences adjacent to the *W1282X* mutation. All four pairs of gRNAs differ only in the DNA sequence recognized by the 3' gRNA. In order to introduce DSB on only mutant *CFTR* locus, we designed the 3' gRNA for pair 2 and pair 3 to overlap with the *W1282X* G>A mutation whereas for pair 1 3' gRNA can recognize *wild type* (*wt*) or *W1282X* mutant sequences and pair 2w 3' gRNA is complementary to the *wt* *W1282X* DNA sequence (Additional file 1: Fig S2A). These nickases were assayed for optimal targeting efficiency and mutation specificity using the T7 endonuclease I (T7E1) assay on genomic DNA from CFBE41o- cells (*wt* for the *W1282X* mutation) transfected with each nickase pair. Whereas pair 1 showed 21.2% NHEJ induction, pair 2 and pair 3 gRNA (*W1282X* mutation specific) induced only 1.3% and 9.9% NHEJ induction, respectively. Because the *CFTR* exon23 sequence in CFBE41o- cells is *wt* for the *W1282X* mutations, pair 2 and pair 3 gRNA were expected to have little or no on-target cutting. In order to confirm the specificity of the pair 2 gRNA molecules, we generated pair 2w gRNA that is complementary to the *wt* *W1282X* mutation sequence. The T7E1 assay showed 11.1% of NHEJ induction by pair 2w gRNA, which was about 10 times more effective than pair 2 gRNA, and suggests that pair 2 gRNA targeting could achieve high allele specific discrimination between *W1282X* and *wtCFTR* (Additional file 1: Fig S2B). The pair 2 gRNA was used for targeting experiments in CF2-iPSCs. With similar concept as above to cleave mutant genomic DNA only, but not donor DNA and modified

allele, two CRISPR/Cas9n pairs (pair A/B and C/B) were designed to target adjacent to TTAA sequence where CAG-Puro Δ TK cassette will be inserted (Additional file 1: Fig S2A). Both A/B and C/B pair did not show clear digestion on the targeting locus by T7E1 assay (Data not shown).

In order to target the *W1282X* mutation in *CFTR* exon23, pair 2 gRNA CRISPR/Cas9n expression vectors (CRISPR/Cas9n-pair 2) were co-transfected with the targeting vector CF2A or the targeting vector CF2B into CF2-iPS3 cells. Candidate clones targeted by donor DNA were selected with Puromycin treatment for 2 weeks, and then assayed by an inside-out PCR strategy with one PCR primer complimentary to the PB cassette (AP2 for 5' end or AP3 for 3' end of the PB Cassette) and the other PCR primer complementary to genomic DNA sequence outside of the homology arm in donor DNA (AP1 for the upstream of 5'- or AP4 for 3'-arm). The AP1/2 and AP3/4 primer pairs amplify a 1,412-bp or 1,602-bp PCR fragment, respectively from HR-targeted alleles (Fig 1A, B, and Additional file 1: Fig S3).

Donor DNA alone did not catalyze site-directed HR in clones exhibiting Puromycin resistance regardless of homology as determined by the lack of PCR products by the AP1/2 or AP3/4 primers but positive PCR for the Puro Δ TK cassette (Additional file 1: Fig S3; Expts H and K). However, co-transfection of the CRISPR/Cas9n-pair 2 with the CF2A targeting vector containing shorter homology arms (Expt I) or with the CF2B targeting vector containing longer homology arms (Expt L) resulted in a targeting efficiencies of 5.7% (2/35) or 28.6% (12/42), respectively (Fig 1B, and Additional file 1: Fig S3) as determined by PCR products with both AP1/2 and AP3/4 primers. On the other hands, co-transfection of the CRISPR/Cas9n-pair C/B did not show any targeting on *CFTR* exon23 locus (Additional file 1: Fig S3; Expts J and M), corresponding with the result of T7E1 assay. One candidate, Expt I clone 8 (Ic8), was used for further characterization of HR and for PB cassette excision. To demonstrate that clone Ic8 resulted from vector replacement between genomic DNA and donor DNA, and not due to NHEJ, the junctions between vector DNA and the genomic target for the targeted Ic8 allele was genotyped by sequencing the AP1/2 and AP3/4 PCR products. Gene correction of *W1282X* *CFTR* mutation (TGA>TGG) and the insertion of PB cassette

flanked by TTAA site were likewise confirmed by DNA sequencing (Fig 1C). These analyses demonstrate that the targeted allele in the Ic8 clone has a gene structure consistent with a vector replacement event.

Excision of drug selectable marker with PB system in CF2-iPSCs

Excision of the Puro Δ TK drug selection cassette was achieved by transient expression of the PB transposase, followed by negative selection with Fialuridine (FIAU) or Ganciclovir (GCV). Seven independent GCV-resistant clones and 24 FIAU clones were isolated, and removal of the Puro Δ TK cassette was confirmed using the P8/P4 and P7f/P6r PCR primer pairs (Fig 1A, 1D). Clones e11 from FIAU treatment and e2 from GCV treatment had PCR results consistent with loss of the drug selectable marker (Fig 1D) and its seamless excision of the Puro Δ TK cassette was confirmed by DNA sequencing (Fig 1E, Ic8_{GCVe2} clone has corrected allele and uncorrected allele with 11-bp deletion at the CRISPR/Cas9n-pair2 targeting site, while Ic8_{FIAUe11} clone has corrected allele and uncorrected allele.). The gene corrected CF2-iPS3 cells maintained their pluripotency and normal karyotype. Clones Ic8_{GCVe2} and Ic8_{FIAUe11} expressed the pluripotent markers NANOG, SSEA3, SSEA4, TRA1-60, and TRA1-81 and upon differentiation expressed the three germ layer markers AFP, SMA, and TUJ1 (Fig 2A). In addition, both modified CF2-iPS3 cells clones had a normal male karyotype (Fig 2B).

Homologous recombinant vector insertion events

Vector insertion gene targeting products have been observed for gene targeting using ends-in vectors, linearized inside vector sequences homologous to genomic targets, for immortalized cell lines and mouse embryonic stem cell lines [20, 21]. Based on previous gene targeting experiments in our laboratory, suggesting that HR vector insertion events were an overlooked class of homologous recombinant products in TALEN and CRISPR/Cas9 stimulated gene targeting (K Chosa, S Suzuki, RG Sargent, unpublished data), we analyzed HR clones identified by the AP1/AP2 inside-out PCR strategy for vector insertion events. While the AP1/AP2 PCR product is diagnostic for all HR events, it does not distinguish between vector replacement and vector insertion gene targeting products. For vector insertion events, the AP3/AP4

PCR products were predicted to be 5,848-bp for the CF2A targeting vector and 6,833-bp for the CF2B targeting vector (Fig 3A), whereas vector replacement events would yield an AP3/AP4 PCR product of 1,602-bp (Fig 1A). A high molecular weight AP3/AP4 PCR product was observed for several HR clones (Fig 1B, and Additional file 1: Fig S3) consistent with the predicted PCR product diagnostic for insertion events.

To verify the molecular structure of the putative insertion events, we used a nested PCR strategy with gel-purified AP3/AP4 PCR products as template (Additional file 1: Fig S4A). The predicted sizes of AP3/AP4 PCR products were confirmed, and nested PCR products using primers targeting to the plasmid backbone and *CFTR* DNA sequences confirmed the presence of the vector backbone and homology arms (Additional file 1: Fig S4B, S4C). Restriction enzyme digestions of the AP3/CF44 PCR products also confirmed same molecular structure of insertion event (Additional file 1: Fig S5).

DNA sequencing across the CRISPR target sequence at 5' and 3' junction of genomic and plasmid DNA further demonstrated that the insertion events were due to HR. DNA sequencing across the CRISPR/Cas9n-gRNAs target sequence for both copies of the tandemly duplicated *CFTR* exon23 for 7 clones showed that only one copy of that in clone Lc35.11 was not a perfect homologous recombinant product and contained a duplication of TAA DNA sequence (Fig 3B). In 6/7 clones, both of the tandemly duplicated exon23 alleles contained the A>G corrected "wild type" DNA sequence whereas clone Lc25 retained the *W1282X* mutant allele in the upstream duplicated exon23 gene sequence and a corrected allele in the downstream exon23 gene duplication consistent with the predicted vector insertion event gene structure via HR.

Reanalysis of Puro^R clones for vector insertion events revealed that the majority of HR clones resulted from vector insertion events with the CF2A targeting vector (Fig 3C; 57.1% insertion vs. 5.7% replacement) whereas vector replacement events were more frequent for the CF2B targeting vector (Fig 3C; 28.6% replacement vs. 19.0% insertion). The gene targeting frequency relative to Puro^R colonies screened for the CF2A targeting vector was 62.8% HR compared to 47.6% for the CF2B targeting vector.

Seamless excision of whole plasmid insertion from genomic DNA

The tandem duplication of exon23 found in the vector insertion events allows for several possible strategies for removal of the plasmid backbone and drug selectable markers: by NHEJ, by HR with additional donor DNA, and by Intrachromosomal homologous recombination (IHR) without additional donor DNA (Figs 4 and 5). Clone Ic14 was used for experiments to test for seamless removal of DNA sequences by these three processes. Experiments using the CRISPR/cas9n-pair 2w were performed with (treatment 1.2) and without (treatment 1.1) co-transfection of the targeting small DNA fragments (SDFs) (1,769-bp *wtCFTR*) to test the efficiency of removal by HR with additional donor and by NHEJ, respectively (Fig 4). Since introduction of DSB has previously been shown to stimulate HR between chromosomal tandem duplications [26, 27], Ic14 were transfected with CRISPR/Cas9n-gRNAs targeting the plasmid backbone sequence (M13/T7 gRNA), or with CRISPR/Cas9n-gRNAs targeting the junction between plasmid and *CFTR* genomic sequence (M13/5'CF2A gRNA) (Fig 5).

Transfection with the CRISPR/Cas9n-pair 2w could introduce DSB in either or both copies of the duplicated exon23 sequences in Ic14. The subsequent loss of intervening plasmid, Puro Δ TK, and duplicated *CFTR* genomic sequences by HR with donor SDF or NHEJ would result in GCV-resistant colonies (Fig 4A). After negative selection with GCV, some clones from each treatment (1.1 and 1.2) were tested for the Puro Δ TK cassette excision (PB⁻) using PCR (P8/P4 and P7f/P6r primers) (Fig 4B, 4C) and are thus presumptive recombinant excision clones. The Ic14 cell line transfected with CRISPR/Cas9n-pair 2w showed a frequency of GCV resistant colonies about 11 times higher than the frequency of excision in Ic14 cells co-transfected with CRISPR/Cas9n-pair 2w and donor SDF (Fig 4B, 8th columns). In the presumptive recombinant excision clones, the DNA sequence surrounding exon23 was analyzed by PCR using the CF46/CF47 PCR primer pair to identify clones that potentially arose by NHEJ. Indels created by NHEJ near the CRISPR/Cas9n-pair 2w cut site would result in CF46/CF47 PCR products with amplicon sizes different from the unmodified allele and HR products. Single CF46/CF47 PCR amplicons diagnostic for

seamlessly stitched clones were observed from presumptive recombinant excision clones isolated from CRISPR/Cas9n-pair 2w transfections (Treatment 1.1; Fig 4B, 9th column; Additional file 1: Fig S6, clones 3c and 3g) and were also observed for CRISPR/Cas9n-pair 2w with SDF co-transfection (Treatment 1.2; Fig 4B, 9th column; Additional file 1: Fig S6, clones 1.2-1c, 1.2-1e, 1.2-1g, 1.2-1h, and 1.2-2b). Presumptive seamlessly stitched clones were sequenced to confirm the genotype. The AP1/AP4 amplicons were used for sequence analysis because under the PCR conditions used, the AP1/AP4 primer pair amplifies non-targeted allele (N) (G>A mutant) and PB⁻ recombinant alleles, but not the tandemly duplicated and vector-inserted allele (I) (Fig 4A). Only one clone from treatment 1.1 (Fig 4D, clone 3c) and two clones from condition 1.2 (Fig 4D, clones 1c and 1h) had a DNA sequence demonstrating a seamlessly corrected *CFTR* exon23 sequence.

IHR between tandemly duplicated regions without additional donor DNA was tested using two pairs of CRISPR/Cas9n-gRNAs, targeting near (M13/T7 pair) or at (M13T7/5'CF2A pair) the 3' junction of plasmid DNA backbone and genomic DNA (Fig 5A). First, these two CRISPR/Cas9n pairs were tested for their targeting efficiency with the T7E1 assay in the Ic14 clone (Fig 4C). While the M13T7/5'CF2A pair showed a faint band in the T7E1 digestion, the M13/T7 pair did not demonstrate detectable targeting (Additional file 1: Fig S7). Consistent with the T7E1 assay, we obtained 13 GCV-resistant clones with the M13T7/5'CF2A pair transfection, but only 1 clone with M13/T7 pair transfection (Fig 5B, 5th column). All GCV-resistant colonies from pair 2w (a), M13/T7 (b) and M13T7/5'CF2 (c) were examined for PB cassette excision by PCR (P8/P4 and P6r/P7f pairs), which confirmed the successful excision by pair 2w (a) and M13T7/5'CF2 (c) transfections. In parallel testing of these three conditions, PBase transfection in Ic14 (d) was performed following the same procedure, which revealed a similar efficiency of excision among the three different excision strategies (%Excision: pair 2w (a); 0.80×10^{-3} , M13T7/5'CF2 (c); 1.30×10^{-3} , PBase (d); 4.11×10^{-3}) (Fig 5B, 7th column). Importantly, when we carried out CF46/47 PCR in all PB cassette-excised clones from M13T7/5'CF2 (c), all clones showed only single band, while some clones from pair 2w (a) showed multiplex bands (Fig 5C). Furthermore, representative PB cassette-

excised clones as well as candidates of seamless corrected clones were tested for the presence of the plasmid backbone using the insertion specific PCR, and successful excision was confirmed (Fig 5D, and E). On the other hand, as we expected, the PBase overexpression (d) successfully excised only the PB cassette from Ic14 genomic DNA, leaving the plasmid DNA backbone intact (Fig 5D, and E). Finally to confirm seamless stitching through IHR, we sequenced all 12 recombinant clones and confirmed the seamless correction in the exon 23 allele (Fig 5F, and 5G, representative sequences from clones 34, 35 and 36). Similar results were obtained in an independent trial followed by negative selection with FIAU (Additional file 1: Fig S8). These data suggest that successful excision via IHR can provide seamless corrected clones. We confirmed normal pluripotent stem cell characteristics in clones Ic14_{GCV}1.2-c, Ic14_{FIAUC}-b10.9 and Ic14_{GCV}-e36, that had successfully undergone HR with additional donor and IHR without it. All three clones with excised insertions by both HR and IHR expressed the pluripotent markers in iPSCs culture condition and three germ layer markers in the randomly differentiated cultures (Fig 6A). Also, all the modified CF2-iPS3 cells had normal male karyotypes (Fig 6B). Thus, we demonstrate a novel and effective strategy to obtain seamless gene correction by IHR of vector insert products.

Discussion

Here we show seamless gene corrections in CF-iPSCs carrying the *W1282X* mutation using two techniques. One approach uses established vector replacement HR products in conjunction with the PiggyBac (PB) transposase system to remove drug selectable markers. We also demonstrate a novel strategy using site-specific, homologous recombinant, vector insertion events followed by excision of drug selectable markers and duplicated DNA sequences via intrachromosomal HR (IHR).

The relative frequency of CRISPR/Cas9 stimulated vector replacement events ranged from 5.7% to 28.6% homologous recombinants per puromycin resistant colony. This frequency appears to be dependent on the amount of homology in the vector arms with the more extensive homology correlated with the higher relative recombination frequency. Whereas increasing the total amount of vector homology

approximately 1.7-fold, from 1,370 bp (vector CF2A, Fig 1A) to 2,355 bp (vector CF2B, Fig 1A), we observed an approximate 5-fold increase in vector replacement products (Fig 1B). This result is in contrast to similar observations in mouse embryonic stem cells without CRISPR/Cas9 stimulation where increasing the total homology present in targeting vectors from 1.3 kbp to 6.8 kbp resulted in approximately a 250-fold increase in homologous recombination [28].

Unlike previous gene targeting experiments in mouse ES cells using vectors linearized inside homologous vector arms, or at the border of arm homology and vector sequences (Fig 7) [20, 21], we transfected the CF2-iPS3 cells with circular targeting vectors. In addition, the CRISPR/Cas9 nickases targeting only or favoring the *W1282X* mutant sequence was used to maintain the circular donor vector as opposed to previous work inducing the whole plasmid DNA insertion via NHEJ by cutting both genomic and donor DNA (Genome research 2013, NAR2016). If HR is initiated between a chromosomal break and a circular targeting vector, then the double-strand-break-repair (DSBR) and Holliday models for homologous recombination predict two possible HR products: vector replacement events and vector insertion events [29-31]. In support of this prediction, we observed vector insertion events for both the CF2A and CF2B targeting vectors (Fig 3).

The majority of vector insertion products are homologous recombinants with only one product out of seven sequenced clones (one out of 14 junctions) showed evidence of an accompanying sequence duplication (Lc35.11; Fig 3B), unlike the previous strategy of the whole plasmid DNA insertion resulting in a high rate of indel inductions within junctions via NHEJ (Genome research 2013, NAR2016).

Interestingly, the majority of clones with vector insertion products also had *W1282X* gene conversion events resulting in reversion of A-to-G such that both of the duplicated exon 23 sequences contained the *WT* CFTR gene sequence. The TAA sequence duplication in the Lc35.11 clone could be explained by mis-templating of DNA synthesis on the vector DNA, accompanying DSBR, that would also be responsible for incorporating the observed A to G reversion.

Whereas we observed a relative HR vector insertion frequency of 57.1% for the CF2A vector, there was only a 19% frequency of HR vector insertion using the CF2B vector. For the CF2A vector, this represents an approximate 10-fold higher frequency of insertion products to replacement products within the same puromycin resistant cell population. The ratio of HR vector insertion events as compared to replacement events appears to be inversely dependent on the length of homology present in vector arms since the ratio of insertion to replacement products decreases from 10:1 to 1:1.5 (Fig 3).

Excision of the puromycin Δ TK drug resistance cassette from the lc8 cell line, containing a vector replacement product, was accomplished by transfecting cells with a PB transposase expression vector and selecting for GCV or FIAU resistant colonies. Both drug selection protocols yielded approximately the same relative frequency of cell lines with the puromycin Δ TK drug resistance cassette removed (1/7 versus 1/6) as determined by PCR analysis. DNA sequencing confirmed the CF2-iPS3_{FIAU}e11 clone contained a seamless excision of the puromycin Δ TK drug resistance cassette whereas the CF2-iPS3_{GCV}e2 clone also contained an 11 basepair deletion in the untargeted CFTR exon 23, adjacent to CRISPR/Cas9n pair2 targeting site and to the same TAA sequence found duplicated in the insertion product Lc35.11 clone. Both clonal lines had normal diploid karyotypes, showed expression for pluripotent markers, and could be differentiated into cells from the 3 germ layers (Fig 2).

Two strategies were used to stimulate recombination between the exon 23 duplication in the lc14 cell line containing a vector insertion product. The first approach used the CRISPR/Cas9 2w nickases that could cut in either or both copies of exon 23 in the duplicated locus. If both copies of exon 23 received double strand breaks then the resulting intermediate would be repaired by NHEJ and if only one copy is cut the resulting product could be repaired by HR, or possibly by NHEJ. To reduce the possibility of NHEJ, we performed experiments where a short DNA fragment (SDF) homologous to the CFTR exon 23 was cotransfected to provide an exogenous substrate for repair of the double strand breaks induced by the

CRISPR/Cas9 2w nickases. Cotransfection of the CFTR SDF did not appear to dramatically improve recovery of recombinant cell lines and may have interfered with recovery of GCV resistant colonies (Fig 4B). However, in experiments with the CRISPR/Cas9 2w nickase cotransfected with the CFTR SDF, 1/5 GCV resistant clones were homologous recombinants versus 1/12 (Fig 5B) or 0/8 homologous recombinants without the SDF (Fig 4B). We are conducting experiments to further quantify the influence of an exogenous template for CRISPR/Cas9 stimulated DSB involving insertion products with tandem gene duplications.

The second approach to stimulate IHR between the exon 23 duplications tested the ability of double strand breaks near the CFTR homology, but inside the vector backbone, or at the edge of the vector/CFTR sequences to induce recombination. Transfection of the lc14 cell line with the CRISPR/Cas9n M13T7/5'CF2 stimulated appearance of GCV resistant colonies at a frequency similar to transfection with the CRISPR/Cas9n 2w or with transfection of the PBase transpose expression vector. However, for cells treated with the CRISPR/Cas9n M13T7/5'CF2, 12/12 gancyclovir recombinants were confirmed as seamless homologous products that reconstructed a gene corrected *W1282X* sequence (Fig 5G, clones c-e34, e35 and e36 are shown as representatively). With cells treated with the CRISPR/Cas9 2w nickase, 3/3 gancyclovir resistant clones were due to NHEJ and deleted all or part of exon 23 (Fig 5G, clones a-e57, e61 and e97).

Conclusion

Generation of homologous recombinant cell lines using vector insertion products has the potential to create several genetically different cell lines from one gene targeting event. This study demonstrates the potential to create seamless gene corrections for cell lines in the same number of steps as commonly used strategies involving the PiggyBac transpose. By creating double strand breaks at the border of vector sequences and homologous genomic sequence we were able to stimulate HR between tandemly duplicated CFTR exon 23 and propose that the CRISPR/Cas9 M13T7 nickase could serve as an almost

universal nickase to be paired with a genomic DNA nickase to stimulate IHR in vector insertion products. We also suggest that one could engineer cell lines with vector insertion products using vectors containing multiple gene sequence changes. Recombination in the intervals between the sequence changes, induced by a CRISPR/Cas9 targeted to the interval, would then result in cell lines containing the desired mutation(s).

Methods

Cells and Culture Conditions

All studies that involve human tissues were approved by the UCSF Committee on Human Research (CHR) and California Pacific Medical Center (CPMC) Institutional Review Board (IRB). Primary CF nasal polyp fibroblasts (CFNPF7), homozygous for the *W1282X CFTR* mutation, were isolated from tissue sample in Dr. Gruenert's laboratory. CFNPF7, used to generate CF patient-derived iPSCs, and human embryonic kidney 293T (HEK293T) cells, used to generate retrovirus, were grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% FBS, 2 mM L-Glutamine (UCSF Cell Culture Facility (CCF), San Francisco, CA), 1 x Non-essential amino acid (UCSF CCF), and 1 x PenStrep (UCSF CCF) and were routinely subcultured by trypsinization with 0.05% trypsin-versene (UCSF CCF). CFNPF7 and HEK293T cells were grown on tissue culture plastic coated with 0.1% gelatin (Sigma-Aldrich, St Louis, MO). Mouse embryo fibroblasts (MEFs) (Millipore, Billerica, MA) were cultured, inactivated with mitomycin C (MMC) (Roche Diagnostics, Mannheim, Germany) as described previously [7], and used as feeder cultures. Immortalized CFBE41o- cells, are homozygous for the *F508del CFTR* mutation [32, 33], were routinely grown in supplemented Eagle's Minimal Essential Medium and subcultured with PET (0.02 % trypsin-versene (UCSF CCF), 1% polyvinylpyrrolidone (Sigma-Aldrich), and 0.2% EGTA (Sigma-Aldrich)) as described previously [7] and used for the T7 endonuclease I assay (T7E1 assay). CF-iPSCs were routinely cultured on Matrigel (BD Biosciences, San Jose, CA) in mTeSR1 medium (StemCells Inc., Vancouver, BC, Canada) and subcultured by non-enzymatic dissociation with ReLeSR (StemCell Inc.).

Cell passage number is denoted as $P_{n_1.n_2.n_3. \dots n_z}$, where n_1 = number of passages as primary cells before reprogramming, n_2 = number of passages since reprogramming, n_3 = number of passages after transfection with a donor DNA, etc where each period delineates the onset of a specific protocol or treatment that alters the character of the cells [34].

Production of recombinant retrovirus and cell reprogramming

CF-iPSCs were generated by retroviral transduction with the canonical transcription factors *OCT4*, *KLF4*, *SOX2*, and *c-MYC* as described previously [7] and according to guidelines developed by the Stem Cell Research Oversight (SCRO) Committee at the CPMC Research Institute (CPMCRI) and the UCSF Gamete and Embryonic Stem Cell Research (GESCR) Committee. Details are described in Additional file 1: Supplemental Materials and Methods.

Generation of embryoid bodies (EBs)

CF-iPSCs lines were cultured in mTeSR1 medium on Matrigel, and then harvested by treatment with Dispase, followed by centrifugation (5 min at 200 g). Cell pellets were re-suspended in 2 ml of mTeSR1 supplemented with 10 μ M Y27632 (Selleckchem, Houston, TX), re-plated into 1 well of a 6-well low attachment tissue culture plate (Corning-Costar) and cultured at 37 °C under 5% CO₂. Within 24 hrs, floating 3 dimensional spherical cell clumps indicative of EBs were visible. EBs were cultured for additional 7 days in mTeSR1 with feeding every other day and then transferred to a multi-well tissue culture plate coated with 0.1% gelatin for attachment. The cells were then grown for an additional 7 days in DMEM supplemented with 10% FBS, 2 mM L-Glutamine, 1 x Non-essential amino acid (UCSF CCF), and 1 x PenStrep before immunostaining.

Immunocytochemical analysis

Cells were grown in multi-well plates on Matrigel (iPSCs) or on gelatin-coated surfaces (non-iPSCs) for immunostaining using primary antibodies and fluorescently labeled secondary antibodies (Additional file

1: Table S1). Briefly, cells washed with PBS were fixed in 4% PFA for 30 min at room temperature (RT) and washed three times, for 5 min each, with ice-cold PBS. Samples were then permeabilized with 0.25% Triton-X for 10 min at RT and incubated for 45 min with 5% serum and 1% BSA in PBS containing 0.1% Tween20 (PBST) at RT with gentle agitation, followed by an overnight incubation at 4 °C with primary antibody in 1% BSA in PBST. Samples were then subjected to 3 x 5 min washes with PBS by gentle agitation. After the third wash, the cells were incubated with secondary antibody in 1% BSA, PBST for 1 hr at RT in the dark, and again washed 3 times as indicated above. Samples were sealed with Dapi Fluoromount-G (SouthernBiotech, Birmingham, AL) and coverslips and examined by fluorescence microscopy.

Cytogenetic analysis

CF-iPSCs were treated with 10 ng/ml of Colcemid (Invitrogen) at 37 °C for 1 hr. The cells were harvested and G-banded according to standard cytogenetic protocols [35]. Metaphase cells were analyzed and karyotyped with the CytoVision system (Leica Microsystems, Wetzlar, Germany) by the Department of Laboratory Medicine Cytogenetics Core, UCSF. Chromosomes were also stained using Quinacrine (Sigma Aldrich) for 30min incubation and karyotype analysis was made through Genikon software (Nikon, Tokyo, Japan) by Department of Molecular Medicine, University of Pavia.

PCR

Genomic DNA was isolated with GeneJET Genomic DNA Purification Kit (Thermo Fisher scientific, Waltham, MA) and used for PCR reactions. PCR were performed using 2X MyTaq HS Mix (BIOLINE, Tuanton, MA) or PrimeSTAR GXL DNA Polymerase (Takara Bio USA Inc., Mountain View, CA) according to the manufacturer's instructions. The amplification products were separated by 0.8 - 2% agarose electrophoresis and imaged with UV light on Geldoc 2000 imaging instrument (Bio-Rad, Hercules, CA). PCR primers are indicated in Additional file 1: Table S2.

Donor DNA vector construction

The PB transposon vector pCAG-Puro Δ TK was derived from pCAG-Puro Δ TK.Neo [17] by removing the Neomycin gene. Two *CFTR* exon23-targeting vectors (CFTRexon23-pCAG-Puro Δ TK.Neo; CF2A or CF2B) were constructed using recombinant PCR. Briefly, CF2A consists of 804-bp of the 5' homology CFTR arm containing the intron sequences upstream of *CFTR* exon23 (CF2A-2 fw, 5'-TTGCAGGTCTCTGCTTCTGG -3') through the first TTAA site from exon23 in the intron downstream of exon23 (TGTTTTTTTAA). The 3' homology arm consists of 566-bp from the intron TTAA site above (TTAACAGCTC) to GAGCACCTCC (CF2A rv, 5'-GGAGGTGCTCCTGGCATTTTA -3') in the same intron. CF2B consists of 1207-bp of the 5' homology CFTR arm containing AGAACACAGA (CF2B fw, 5'-AGAACACAGAGTTGGGGCTC -3') through the same TTAA site as CF2A construct. The 3' homology arm consists of 1148-bp from the intron TTAA site to GGCCAGAGTT (CF2B rv, 5'-AACTCTGGCCCACTTGGTTTT -3') in the same intron. These homology arms were amplified by PCR and joined with the pCAG-Puro Δ TK cassette by recombinant PCR to create the final targeting construct. All primers used for donor DNA construction are shown in Table S3.

Generation of CRISPR/Cas9n-gRNAs

Guide RNA targeting sequences were designed and selected using Web-based software, Optimized CRISPR design, developed by Zhang Lab, MIT; <http://crispr.mit.edu>. pSpCas9n(BB)-2A-Puro (PX462), a gift from Dr. Feng Zhang (Addgene plasmid # 48141), was used as the both gRNA and SP-Cas9n expression vector. The oligos listed in Additional file 1: Table S4 were used to assemble gRNA targeting specific sequence by following the established protocol [36].

Correction of *W1282X* mutation in CF2-iPS3 cells

CF2-iPS3 cells were co-transfected with the donor DNA plasmid (CF2A or CF2B) in the absence or presence of CRISPR/Cas9n-gRNAs. Briefly, CF2-iPS3 cells were harvested as single cell suspension with Accutase (StemCell Inc.). Then, 1.5×10^6 cells were nucleofected with 2.5 μ g of donor DNA plasmid with

or without 2.5 μg of each pairs of CRISPR/Cas9n-gRNAs using the 4D Nucleofector X (Lonza) with the P3 Primary Cell solution and Program CA137. Transfected CF2-iPS3 cells were plated in mTeSR1 supplemented with 10 μM Y27632 in Matrigel-coated plate for 24 hr post transfection. Two to three days after transfection, the culture medium was switched to the selection medium, mTeSR1 containing 0.5 - 1.0 $\mu\text{g}/\text{ml}$ of Puromycin (Sigma Aldrich), and the cells were continuously cultured under Puromycin selection up to 14 days. During the selection, 7 - 10 days post-transfection, all colonies were manually picked up and individually transferred to 24-well plates coated with Matrigel. Genomic DNA were isolated from individual clones and amplified by PCR with primers AP1/AP2, AP3/AP4, and P6r/P7f to screen successful vector replacement and insertion event, and each PCR product was separated on a 1 - 2% agarose gel containing ethidium bromide and visualized under UV light.

Excision of PB cassette

In order to remove the PB cassette containing the drug selection marker from modified CF2-iPS3, PB transposase (PBase) were transfected into these cells, followed by negative selection with Fialuridine (FIAU) (Santa Cruz Biotechnology Inc., Dallas, TX) or Ganciclovir (GCV) (Sigma-Aldrich) [37]. The modified PBase (R372A/K375A/D450N) [38] expression vector was kindly provided by Drs. YW Kan and Lin Ye at UCSF. The PBase expression vector was nucleofected in as described above. After the nucleofection, the cells were passaged twice as single cells every 2 - 3 days with Y27632, and plated into 60mm petri dishes, as single cells, at 10^6 cells/dish. Negative selection for loss of the TK gene was with mTeSR1 medium containing 0.25 μM FIAU, or 0.2 or 2 μM GCV. After negative selection, some colonies were clonally isolated and cultured individually in 24-well plates. PCR with primers P8/P4, P6r/P7f were performed on genomic DNA harvested from each clones to screen a clone with successfully excised PB cassette (map with PCR primer locations). The percentage of excision was calculated from the number of PB Cassette negative clones (P8/P4⁻ and P6r/P7f⁻) out of the cell numbers right before the selection in 60mm dish (10^6 cells), according to the following formulas:

$$f_{excision} = \frac{\text{Excised clones}}{\text{tested clones}}$$

$$\%Excision = 100 * \frac{\text{All appeared colonies} * f_{excision}}{10^6}$$

Excision of whole plasmid insertion

CRISPR/Cas9n targeting *CFTR* exon23 or the junction of vector backbone and *CFTR* gene were nucleofected in whole vector-inserted CF2-iPSCs (Ic14) to introduce Intrachromosomal HR. After nucleofection, excised clones were screened with 0.2 or 2 μ M GCV and PCR as described above. The percentage of excision was calculated also as described above. Following Fig 4C flow, seamlessly excised clones were screened by PCR with CF46/CF47 primer pair and confirmed by Sanger sequencing on AP1/4 PCR amplicon.

Declarations

Ethics approval and consent to participate

All studies that involve human tissues were approved by the UCSF Committee on Human Research (CHR) and California Pacific Medical Center (CPMC) Institutional Review Board (IRB).

Patient-derived iPSC was generated by according to guidelines developed by the Stem Cell Research Oversight (SCRO) Committee at the CPMC Research Institute (CPMCRI) and the UCSF Gamete and Embryonic Stem Cell Research (GESCR) Committee.

Consent for publication

Not applicable

Availability of data and materials

Datasets supporting the conclusions of article are including within the article and its additional file.

Competing interests

The authors declare there are no competing interests.

Funding

These studies were supported by NIH Program Project Grant (PPG) [DK088760], the US Department of Defense [Award W81XWH-15-1-0261] and the Japan Society for the Promotion Science (JSPS) program on Strategic Young Researcher Overseas Visits Program for Accelerating Brain Circulation [Grant Number R2803].

Author Contributions

Conceptualization: S.S., R.G.S., and D.C.G.

Data curation: S.S., K.C., M.Y. and C.B.

Formal analysis: S.S., K.C. and C.B.

Investigation: S.S., K.C., C.B., R.G.S., and D.C.G.

Project administration: R.G.S., and D.C.G.

Resources: O.Z., Y.W.K., D.C.G., and R.G.S.

Supervision: O.Z., T.S., H.K., Y.W.K., D.C.G., and R.G.S.

Original draft: S.S., K.C., M.S., and R.G.S.

Review/editing the draft: S.S., K.C., M.S., and R.G.S.

Acknowledgements

We acknowledge supports for S.S. and K.C. from Japan Society for the Promotion of Science (JSPS) Research Fellowship for Young Scientists and the support of S.S. by JSPS Postdoctoral Fellowship for Research Abroad.

References

1. Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N, Chou JL, et al.: **Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA.** *Science* 1989, **245**:1066-1073.
2. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F: **Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity.** *Cell* 2013, **154**:1380-1389.
3. Hartl D, Gaggar A, Bruscia E, Hector A, Marcos V, Jung A, Greene C, McElvaney G, Mall M, Doring G: **Innate immunity in cystic fibrosis lung disease.** *J Cyst Fibros* 2012, **11**:363-382.
4. O'Sullivan BP, Freedman SD: **Cystic fibrosis.** *Lancet* 2009, **373**:1891-1904.
5. Crane AM, Kramer P, Bui JH, Chung WJ, Li XS, Gonzalez-Garay ML, Hawkins F, Liao W, Mora D, Choi S, et al: **Targeted correction and restored function of the CFTR gene in cystic fibrosis induced pluripotent stem cells.** *Stem Cell Reports* 2015, **4**:569-577.
6. Firth AL, Menon T, Parker GS, Qualls SJ, Lewis BM, Ke E, Dargitz CT, Wright R, Khanna A, Gage FH, Verma IM: **Functional Gene Correction for Cystic Fibrosis in Lung Epithelial Cells Generated from Patient iPSCs.** *Cell Rep* 2015, **12**:1385-1390.
7. Suzuki S, Sargent RG, Illek B, Fischer H, Esmaili-Shandiz A, Yezzi MJ, Lee A, Yang Y, Kim S, Renz P, et al: **TALENs Facilitate Single-step Seamless SDF Correction of F508del CFTR in Airway Epithelial Submucosal Gland Cell-derived CF-iPSCs.** *Mol Ther Nucleic Acids* 2016, **5**:e273.

8. Shoshani T, Augarten A, Gazit E, Bashan N, Yahav Y, Rivlin Y, Tal A, Seret H, Yaar L, Kerem E, et al.: **Association of a nonsense mutation (W1282X), the most common mutation in the Ashkenazi Jewish cystic fibrosis patients in Israel, with presentation of severe disease.** *Am J Hum Genet* 1992, **50**:222-228.
9. Castellani C, Cuppens H, Macek M, Jr., Cassiman JJ, Kerem E, Durie P, Tullis E, Assael BM, Bombieri C, Brown A, et al: **Consensus on the use and interpretation of cystic fibrosis mutation analysis in clinical practice.** *J Cyst Fibros* 2008, **7**:179-196.
10. Bobadilla JL, Macek M, Jr., Fine JP, Farrell PM: **Cystic fibrosis: a worldwide analysis of CFTR mutations--correlation with incidence data and application to screening.** *Hum Mutat* 2002, **19**:575-606.
11. Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, Rupniewski I, Beausejour CM, Waite AJ, Wang NS, Kim KA, et al: **An improved zinc-finger nuclease architecture for highly specific genome editing.** *Nat Biotechnol* 2007, **25**:778-785.
12. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF: **Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting.** *Nucleic Acids Res* 2011, **39**:e82.
13. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, et al: **A TALE nuclease architecture for efficient genome editing.** *Nat Biotechnol* 2011, **29**:143-148.
14. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F: **Multiplex genome engineering using CRISPR/Cas systems.** *Science* 2013, **339**:819-823.
15. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM: **RNA-guided human genome engineering via Cas9.** *Science* 2013, **339**:823-826.

16. Yusa K, Rashid ST, Strick-Marchand H, Varela I, Liu PQ, Paschon DE, Miranda E, Ordonez A, Hannan NR, Rouhani FJ, et al: **Targeted gene correction of alpha1-antitrypsin deficiency in induced pluripotent stem cells.** *Nature* 2011, **478**:391-394.
17. Ye L, Wang J, Beyer AI, Teque F, Cradick TJ, Qi Z, Chang JC, Bao G, Muench MO, Yu J, et al: **Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5Delta32 mutation confers resistance to HIV infection.** *Proc Natl Acad Sci U S A* 2014, **111**:9591-9596.
18. Sargent RG, Suzuki S, Gruenert DC: **Nuclease-mediated double-strand break (DSB) enhancement of small fragment homologous recombination (SFHR) gene modification in human-induced pluripotent stem cells (hiPSCs).** *Methods Mol Biol* 2014, **1114**:279-290.
19. Miyaoka Y, Chan AH, Judge LM, Yoo J, Huang M, Nguyen TD, Lizarraga PP, So PL, Conklin BR: **Isolation of single-base genome-edited human iPS cells without antibiotic selection.** *Nat Methods* 2014, **11**:291-293.
20. Hasty P, Crist M, Grompe M, Bradley A: **Efficiency of insertion versus replacement vector targeting varies at different chromosomal loci.** *Mol Cell Biol* 1994, **14**:8385-8390.
21. Valancius V, Smithies O: **Testing an "in-out" targeting procedure for making subtle genomic modifications in mouse embryonic stem cells.** *Mol Cell Biol* 1991, **11**:1402-1408.
22. Nakade S, Tsubota T, Sakane Y, Kume S, Sakamoto N, Obara M, Daimon T, Sezutsu H, Yamamoto T, Sakuma T, Suzuki KT: **Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9.** *Nat Commun* 2014, **5**:5560.
23. Maresca M, Lin VG, Guo N, Yang Y: **Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining.** *Genome Res* 2013, **23**:539-546.
24. Suzuki K, Tsunekawa Y, Hernandez-Benitez R, Wu J, Zhu J, Kim EJ, Hatanaka F, Yamamoto M, Araoka T, Li Z, et al: **In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration.** *Nature* 2016, **540**:144-149.

25. He X, Tan C, Wang F, Wang Y, Zhou R, Cui D, You W, Zhao H, Ren J, Feng B: **Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair.** *Nucleic Acids Res* 2016, **44**:e85.
26. Sargent RG, Meservy JL, Perkins BD, Kilburn AE, Intody Z, Adair GM: **Role of the nucleotide excision repair gene ERCC1 in formation of recombination-dependent rearrangements in mammalian cells.** 2000, **28**:3771-3778.
27. Jasin M: **Genetic manipulation of genomes with rare-cutting endonucleases.** *Trends Genet* 1996, **12**:224-228.
28. Hasty P, Rivera-Pérez J, Bradley A: **The length of homology required for gene targeting in embryonic stem cells.** *Mol Cell Biol* 1991, **11**:5586-5591.
29. Holliday R: **A mechanism for gene conversion in fungi.** *Genet Res Camb* 1964, **5**:282-304.
30. Holliday R: **Genetic recombination in fungi. In Replication and Recombination of Genetic Material.** (ed. W.J. Peacock and R.D. Brook). *Canberra: Australian Academy of Science* 1968:157-174.
31. Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW: **The double-strand-break repair model for recombination.** *Cell* 1983, **33**:25-35.
32. Gruenert DC, Willems M, Cassiman JJ, Frizzell RA: **Established cell lines used in cystic fibrosis research.** *J Cyst Fibros* 2004, **3 Suppl 2**:191-196.
33. Illek B, Maurisse R, Wahler L, Kunzelmann K, Fischer H, Gruenert DC: **Cl transport in complemented CF bronchial epithelial cells correlates with CFTR mRNA expression levels.** *Cell Physiol Biochem* 2008, **22**:57-68.
34. Gruenert DC, Basbaum CB, Welsh MJ, Li M, Finkbeiner WE, Nadel JA: **Characterization of human tracheal epithelial cells transformed by an origin-defective simian virus 40.** *Proc Natl Acad Sci U S A* 1988, **85**:5951-5955.
35. Barch MJ, Knutsen T, Spurbeck JL, Association of Genetic T: *The AGT cytogenetics laboratory manual.* Philadelphia: Lippincott-Raven Publishers; 1997.

36. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F: **Genome engineering using the CRISPR-Cas9 system.** *Nat Protoc* 2013, **8**:2281-2308.
37. Chen YT, Bradley A: **A new positive/negative selectable marker, puDeltatk, for use in embryonic stem cells.** *Genesis* 2000, **28**:31-35.
38. Li X, Burnight ER, Cooney AL, Malani N, Brady T, Sander JD, Staber J, Wheelan SJ, Joung JK, McCray PB, Jr., et al: **piggyBac transposase tools for genome engineering.** *Proc Natl Acad Sci U S A* 2013, **110**:E2279-2287.

Figure legend

Fig 1 Seamless gene correction with PiggyBac (PB) and CRISPR/Cas9n-gRNAs in Class I CF-iPSCs

(CF2-iPS3 cells). (A) Illustration of targeting strategy via HR with donor DNA carrying PB cassette, described in methods. PCR primers (AP1/AP2, AP3/AP4, and P6r/P7f) used to confirm successful HR are illustrated next to the annealing site. AP1/AP2; intron at upstream of *CFTR* exon23 through the 5' end of PB cassette. AP3/AP4; 3' end of PB cassette through intron at downstream of exon23. P6r/P7f; Puromycin resistant gene at PB cassette. Black lines in both genomic DNA and donor DNA represent homology regions. Grey line in genomic DNA is non-homology region. Bold black line is plasmid donor DNA backbone. (B) HR efficiency assessed by PCR shown in Additional file 1: Fig S3, and gel pictures of genotyping for represented colonies. M; 1kb or 100bp DNA ladder plus, g; CF2-iPS3 genomic DNA, g+p; the mixture of CF2-iPS3 genomic DNA and donor DNA. High molecular weight (HMW) fragments (> 3kbp) observed in AP3/AP4 PCR were further investigated in Fig 3, and Additional file 1: Fig S4, S5. (C) Sequence chromatogram showing the 5' and 3' junction of PB cassette (shown in red box) and *W1282X* mutation site on *CFTR* exon23 (shown in blue box) in Ic8 clone. Sanger sequencing was performed on AP1/AP2 and AP3/AP4 amplicons to present only targeted allele sequence. (D) The summary of the PB cassette excision. The excision was identified PB cassette negative clones (P8/P4- and P7f/P6r-) according to the PCR results as shown in gel pictures. M; 100bp DNA ladder plus. (E) Sequence of excised clones screened from PBase-transfected Ic8 clones selected with 0.2 μ M of GCV (Ic8_{GCVe2}, top clone) or with FIAU (Ic8_{FIAUe11}, bottom clone). Adenine (A) and Guanine (G) capitalized and highlighted in red representing *W1282X* mutation (G>A) and the correction (A>G), respectively. Dot represents deleted nucleotide.

Fig 2 Characterization of corrected Class I CF mutation in CF2-iPS3 cells. (A, B) Characterization of

Ic8GCVe2 and Ic8FIAUe11 clones by (A) immunocytochemical analysis for ES markers (NANOG, SSEA3, SSEA4, TRA1-60, TRA1-81), and three-germ layers (Endoderm: alpha-fetoprotein; AFP, Mesoderm: alpha-smooth muscle actin; SMA, Ectoderm: beta-tubulin 3; TUJ1), and by (B) karyotype.

Fig 3 Identification of plasmid DNA insertion event. (A) Schematic illustration of site-specific insertion event with donor DNA backbone. AP3/AP4 PCR amplifies 5,848-bp product from CF2A donor transfection and 6,833-bp product from CF2B donor transfection, which are confirmed in Fig 1B, and Additional file 1: Fig S3. (B) Representative sequence alignments at 5' and 3' junctions as the consequence of each donor DNA insertion, CF2A (I) or CF2B (L) donor. Lc35 clone initially isolated was a mixed population of Lc35.7 and Lc35.11 cells. Targeting sequences of pair 2 gRNA are shown in red and blue. Adenine (A) and Guanine (G) capitalized and representing W1282X mutation (G>A) and the correction (A>G), respectively. (C) Comparison of frequency between vector replacement and site-specific insertion event. The frequency of each event was calculated with the number of positive clones out of the number of picked and analyzed PuroR clones.

Fig 4 Seamless excision of inserted plasmid DNA backbone with CRISPR/Cas9n-pair 2w and donor SDF. (A) Strategy of excising inserted backbone in order to obtain seamless corrected CF2-iPS3 cells using CRISPR/Cas9n-pair 2w. Two methods were performed by using CRISPR/Cas9n-pair 2w, the pair targeting site of correction experiment in Fig 1, but targeting wt exon23, in the presence (treatment 1.2) or absence (treatment 1.1) of donor SDF. The gRNA targeting site is represented by black diamond and connected by the line for a pair. Indel is indicated by */. I; Inserted allele, N; unmodified allele. (B) Comparison of the seamless excision efficacy in Ic14 clone between treatment 1.2 and 1.1. (C) Flow of seamless corrected clone isolation. (D) Sequence histogram of AP1/4 amplicon shows single peak of A in W1282X site in Ic14 (Pre-excision), and mix peak of A and G in Ic14 1.1-3c, 1.2-1c and 1.2-1h (Post-excision). Red arrow points W1282X mutation site.

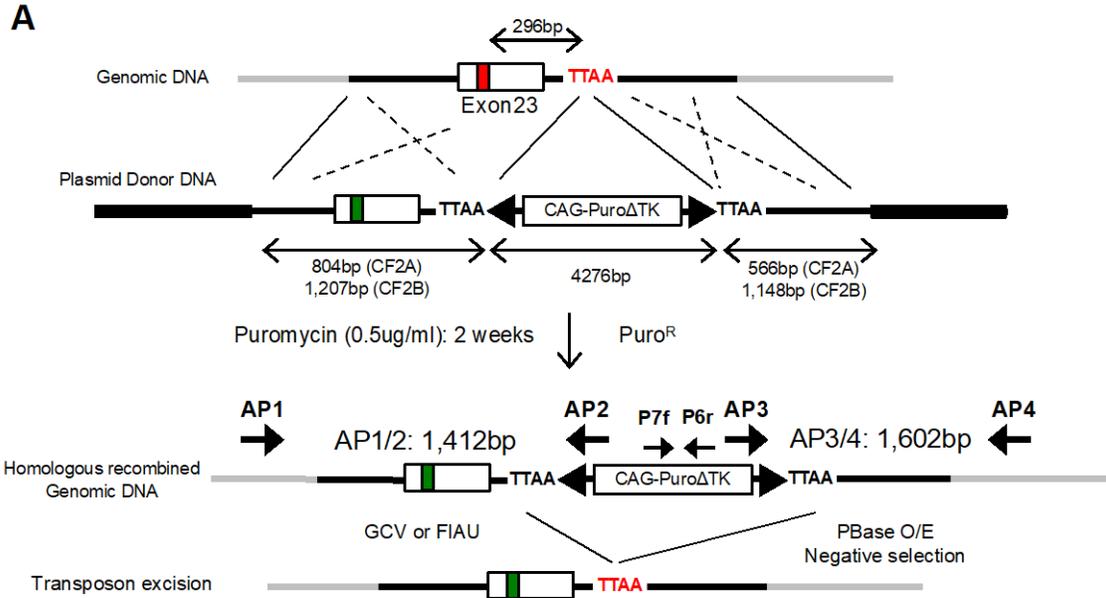
Fig 5. Seamless excision of inserted plasmid DNA backbone with IHR. (A) Strategy of excising the inserted plasmid DNA backbone using the CRISPR/Cas9n-gRNAs targeting the junction of plasmid donor backbone and genomic DNA. Newly generated gRNA targeting two regions were used, one of which cut

the backbone of plasmid DNA (M13/T7 gRNA, treatment b) and another targeted the junction between backbone and 5'-homology arm of CFTR on donor DNA (M13T7/5'CF2 gRNA, treatment c). For comparison, pair 2w gRNA alone (treatment a) and PBase alone (treatment d) were also tested. (B) Summary of the excision efficacy for each strategy. Ic14 (P8.13.14) was transfected with four treatments and followed the flow for the isolation of seamless corrected clones in Fig 4C (C) Selection of possible seamless excised clones with CF46/47 PCR. PB cassette negative clones from treatment a, and c were tested in a multiplicity of PCR amplicon. Clones a-e26, a-e29, a-e33 (circled in red), and all 12 clones from treatment c had single amplicon. (D) Schematic illustration of predicted outcomes from each treatment with primer targeting sites used to test their gene structure. (E) Confirmations of the removal of inserted plasmid DNA backbone using the insert specific primer pairs in representative clones in treatment a, c, and d. (F) Sequence histogram of AP1/4 amplicon from Ic14 and Ic14 c-e36. Red arrow points to W1282X mutation site. (G) Sequence of representative clones screened from Ic14 clones with excised inserted backbone by treatment a, and c. Red capital letter shows W1282X mutation site. All three Ic14GCVa clones have indel in corrected allele at upstream of W1282X mutation site shown as dot (deletion) or blue letter (insertion). All three Ic14GCVc clones have corrected and uncorrected allele without indels.

Fig 6 Characterization of corrected CF2-iPS3 cells with removed insertion. (A, B) Characterization of representative corrected clones, Ic14GCV1.2-1c, Ic14FIAUc-b10.9, and Ic14GCVc-e36, by (A) immunocytochemical analysis for three germ layers and ES markers and (B) karyotyping.

Fig 7. Gene targeting experiments in mouse ES cells using linearized vector. i) The linearization of plasmid donor DNA by digesting at at the border of homology and vector sequences results in vector replacement event and ii) The linearization by cutting within sequences homologous to genomic targets results in vector insertion event. Plasmid DNA back bone; Black, Homology sequences in plasmid Donor DNA; Red and in genomic DNA; Blue, Non homology region in genomic DNA; Gray.

Figure 1: Seamless gene correction with PiggyBac and CRISPR/Cas9 nickases in Class I mutant CF-iPSCs



B

Experiment	cell number	donor DNA (ug)	CRISPR (ug each)	Puro (ug/ml)	Puro ^R	HDR
H	1.5 x 10 ⁶	CF2A (2.5)	-	0.5	28	0
I	1.5 x 10 ⁶	CF2A (2.5)	pair2 (2.5)	0.5	35	2 (5.7%)
J	1.5 x 10 ⁶	CF2A (2.5)	C/B (2.5)	0.5	36	0
K	1.5 x 10 ⁶	CF2B (2.5)	-	0.5	31	0
L	1.5 x 10 ⁶	CF2B (2.5)	pair2 (2.5)	0.5	42	12 (28.6%)
M	1.5 x 10 ⁶	CF2B (2.5)	C/B (2.5)	0.5	32	0

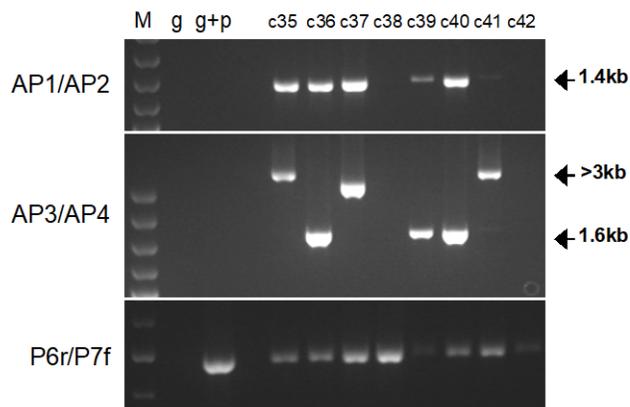
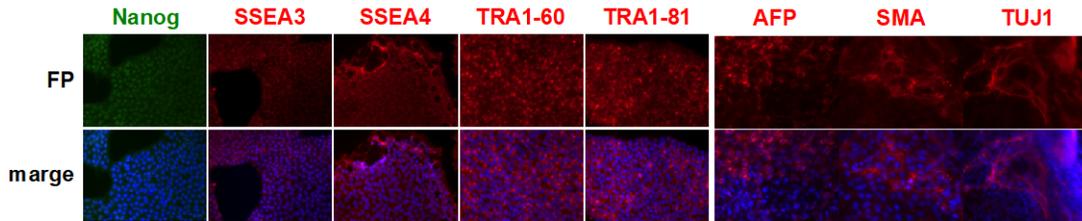


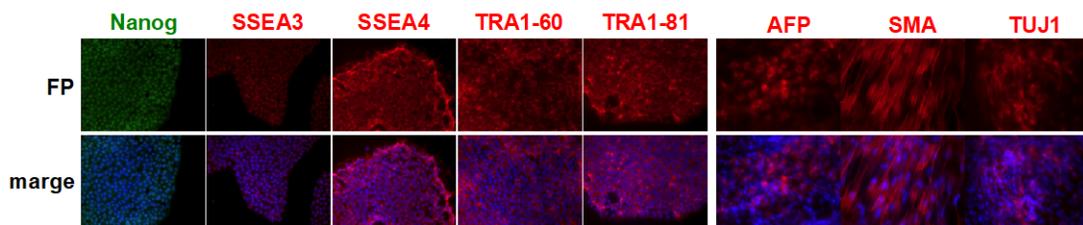
Fig 2. Characterization of Corrected Type I CF mutation in CF-iPSCs

A

I.c8_{GCV}e2



I.c8_{FIAU}e11



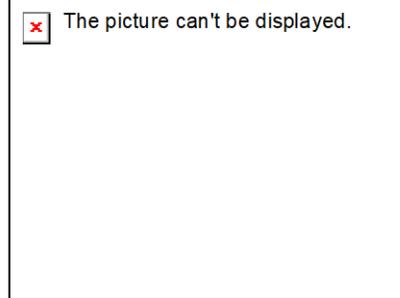
B

I.c8_{GCV}e2



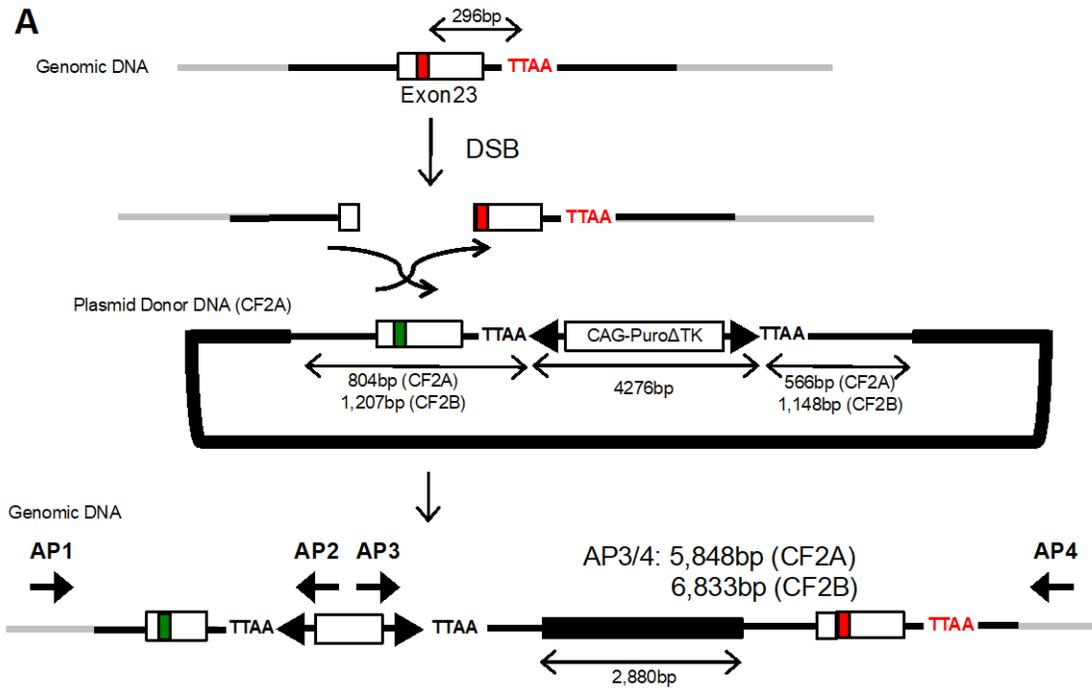
**Normal male (46, XY)
(4 tests)**

I.c8_{FIAU}e11



**Normal male (46, XY)
(20 tests)**

Figure 3: Identification of plasmid DNA insertion event



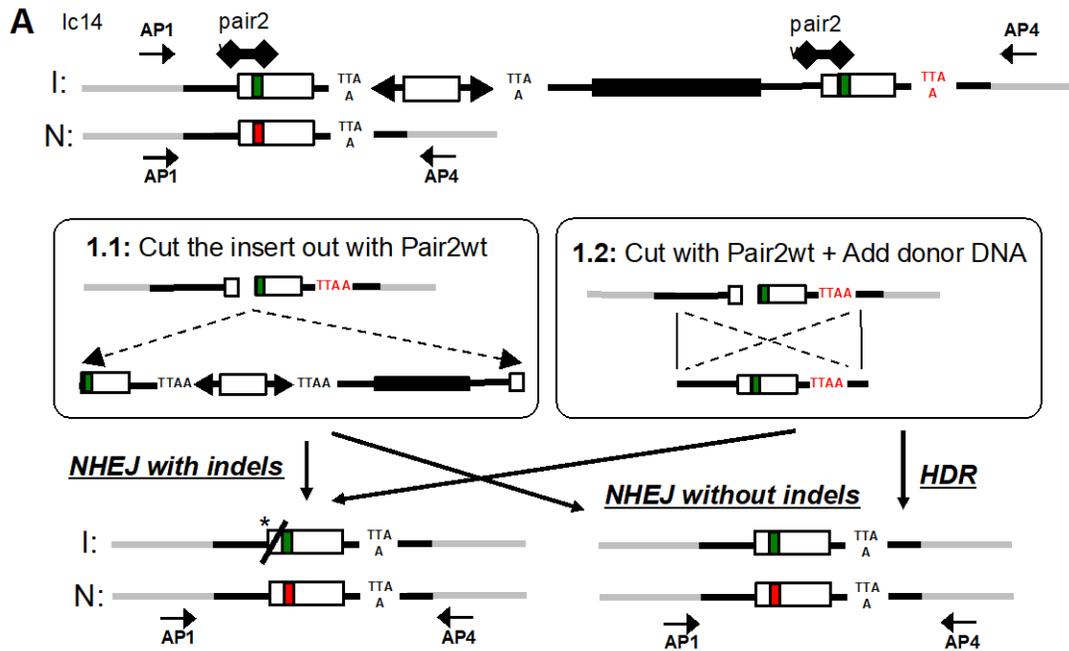
B

Experiment	cell number	donor DNA (ug)	CRISPR (ug each)	Puro ^R	HDR	Site-specific Insertion
H	1.5 x10 ⁶	CF2A (2.5)	-	28	0	0
I	1.5 x10 ⁶	CF2A (2.5)	pair2 (2.5)	35	2 (5.7%)	20 (57.1%)
K	1.5 x10 ⁶	CF2B (2.5)	-	31	0	0
L	1.5 x10 ⁶	CF2B (2.5)	pair2 (2.5)	42	12 (28.6%)	8 (19%)

C

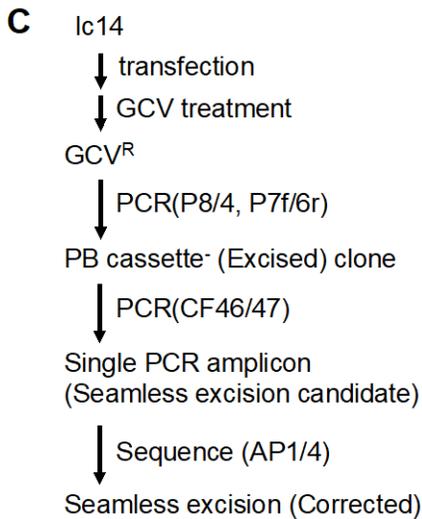
	5' junction	3' junction
I.c2	cca gatcgatggtgtgtccttggg attcaataactttgcaacagtg Gagg	cca gatcgatggtgtgtccttggg attcaataactttgcaacagtg Gagg
I.c14	cca gatcgatggtgtgtccttggg attcaataactttgcaacagtg Gagg	cca gatcgatggtgtgtccttggg attcaataactttgcaacagtg Gagg
I.c20	cca gatcgatggtgtgtccttggg attcaataactttgcaacagtg Gagg	cca gatcgatggtgtgtccttggg attcaataactttgcaacagtg Gagg
L.c1	cca gatcgatggtgtgtccttggg attcaataactttgcaacagtg Gagg	cca gatcgatggtgtgtccttggg attcaataactttgcaacagtg Gagg
L.c25	cca gatcgatggtgtgtccttggg attcaataactttgcaacagtg Aagg	cca gatcgatggtgtgtccttggg attcaataactttgcaacagtg Gagg
L.c35.7	cca gatcaataatgatcttga attcaataactttgcaacagtg Gaaa	cca gatcgatggtgtgtccttggg attcaataactttgcaacagtg Gagg

Fig 4. Seamless excision of inserted plasmid DNA backbone



B Excision Treatments

	cell numbers	vector (ug)	dsDNA	GCV (uM)	GCV ^R	PB ⁻ (- /test)	%excision	CF46/47 (single/test)	corr /seq	corr /excision
1.1	?	pair2 (?ug each)	-	0.2	165	12/18	11.00 x10 ⁻³	2/12	1/2	1/12
1.2	?	pair2 (?ug each)	?ug	0.2	17	10/17	1.00 x10 ⁻³	5/10	2/5	1/10



D Sequence of AP1/4

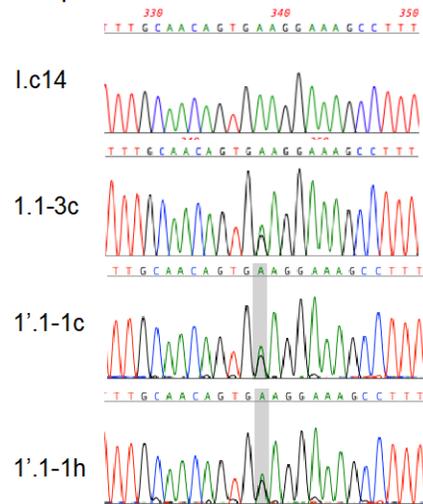
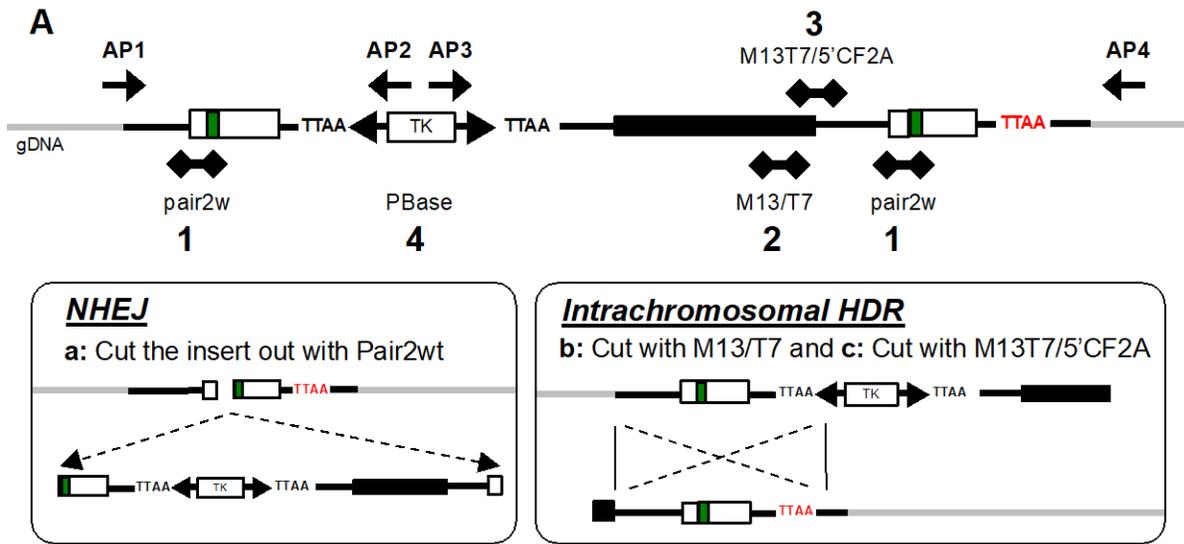


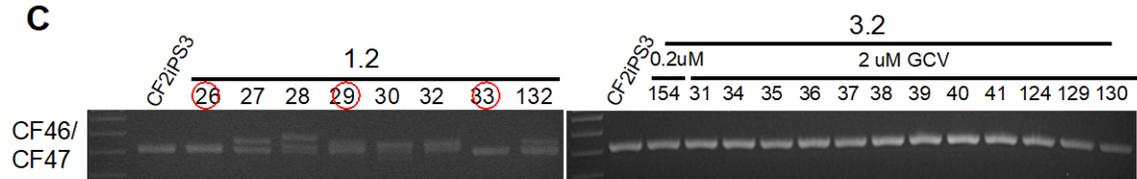
Fig 5. Seamless excision of inserted plasmid DNA backbone

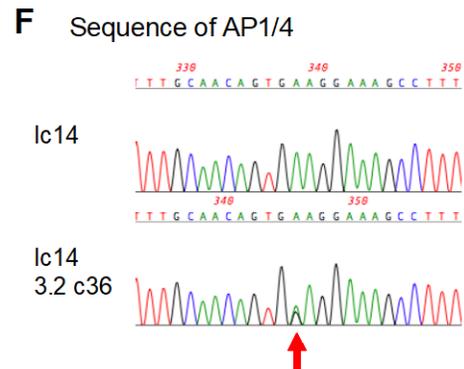
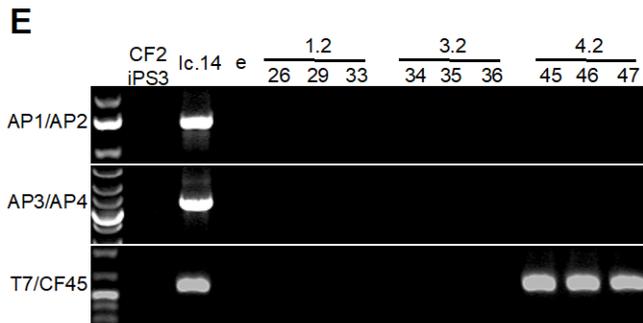
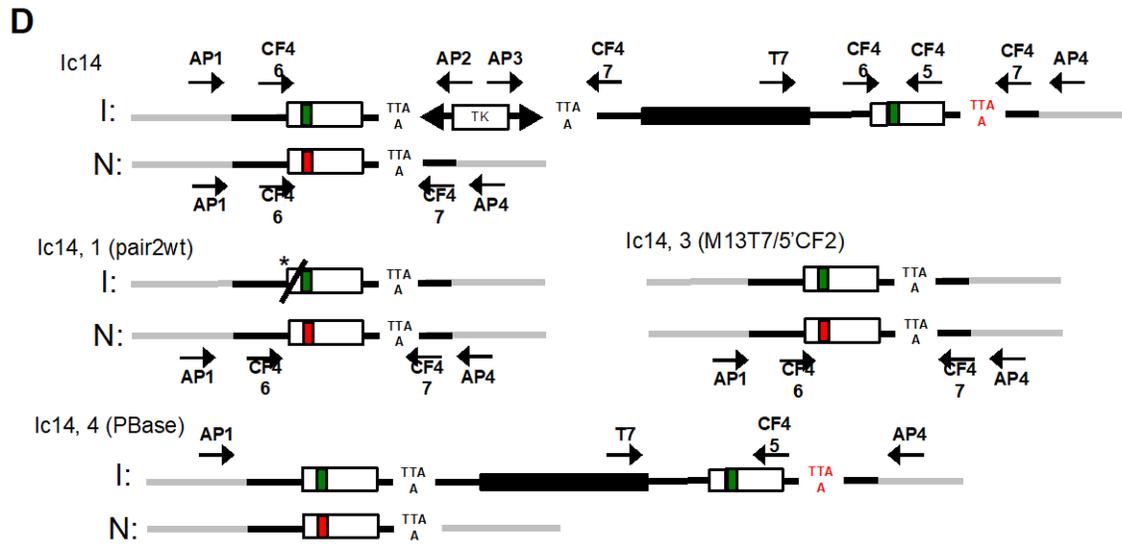


B I.c14 P8.13.14

	cell numbers	vector (ug)	GCV (uM)	GCV ^R	PB Cassette (-/test)	%excision	CF46/47 (single/test)	corr /seq	corr /excision
a	2.0 x 10 ⁶	pair2wt (3.33ug each)	2.0	10	8/10	0.80 x 10 ⁻³	3/8	0/3	0/8
b	2.0 x 10 ⁶	M13/T7 (3.33ug each)	2.0	1	0/1	-	-	-	-
c	2.0 x 10 ⁶	M13T7/5'CF2 (3.33ug each)	2.0	13	12/12	1.30 x 10 ⁻³	12/12	3/3	12/12
d	2.0 x 10 ⁶	PBase (6.0ug)	2.0	48	6/7	4.11 x 10 ⁻³	NE	NE	NE

C





G

I.c14_{GCV}1.2-e57: W1282? (G>?) + 26bp deletion/ W1282X (G>A)

I: ccagatcgatggtgtgtccttggg aaagcctttggagtgateaccacag

N: ccagatcgatggtgtgtccttgggattcaataactttgcaacagtg **A**aggaaagcctttggagtgateaccacag

I.c14_{GCV}1.2-e61: W1282? (G>?) + 23bp deletion/ W1282A (G>A) + 10bp insertion

I: ccagatcgatggtgtgtccttggga ggaagcctttggagtgateaccacag

N: ccagatcgatggtgtgtccttgggattcaata **GGATTC AATA**actttgcaacagtg **A**aggaaagcctttggagtgateaccacag

I.c14_{GCV}1.2-e97: W1282W (G>G) + 31bp insertion/ W1282X (G>A)

I: ccagatcgatggtgtgtccttgggattcaataactttgcaa **INSC**agtg**G**aggaaagcctttggagtgateaccacag

N: ccagatcgatggtgtgtccttgggattcaataactttgcaacagtg **A**aggaaagcctttggagtgateaccacag

I.c14_{GCV}3.2-e34, -e35, -e36: W1282W (G>G) / W1282X (G>A)

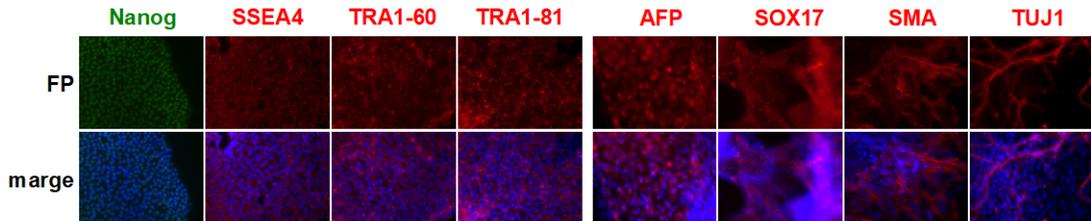
I: ccagatcgatggtgtgtccttgggattcaataactttgcaacagtg **G**aggaaagcctttggagtgateaccacag

N: ccagatcgatggtgtgtccttgggattcaataactttgcaacagtg **A**aggaaagcctttggagtgateaccacag

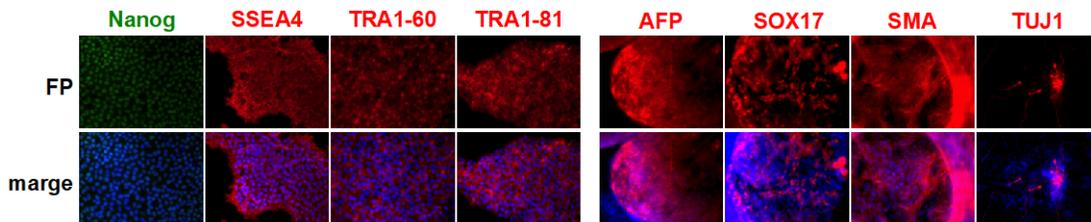
Fig 6. Characterization of corrected CF-iPSCs removed insertion

A

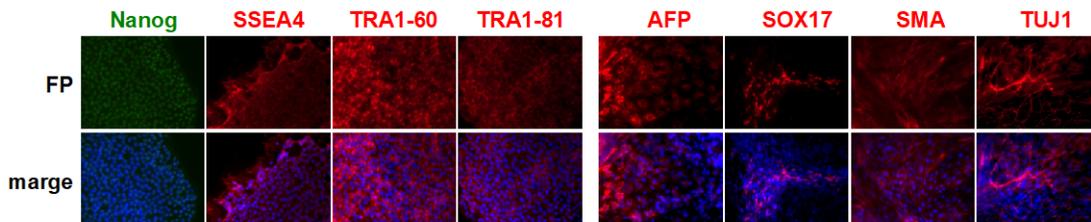
I.c14_{GCV}1'.1-1c



I.c14_{FIAU}3.1-b10.9



I.c14_{GCV}3.2-e36



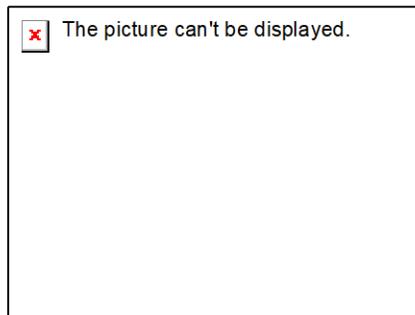
B

I.c14_{GCV}1'.1

I.c14_{FIAU}3.1-b10.9

I.c14_{GCV}3.2-e36

On going



On going

**Normal male (46, XY)
(20 tests)**

Supplemental Figure

Figure S1: Generation of W1282X CF-iPSCs

Figure S2: Assessment of CRISPR/Cas9n for W1282X targeting

Figure S3: All PCR picture for 1st screening on Puro^R clones

Figure S4: Identification of HMW by PCR

Figure S5: Identification of HMW by RE

Figure S6: Screening for seamless excision

Figure S7: T7E1 assay for BB-targeting CRISPR/Cas9n

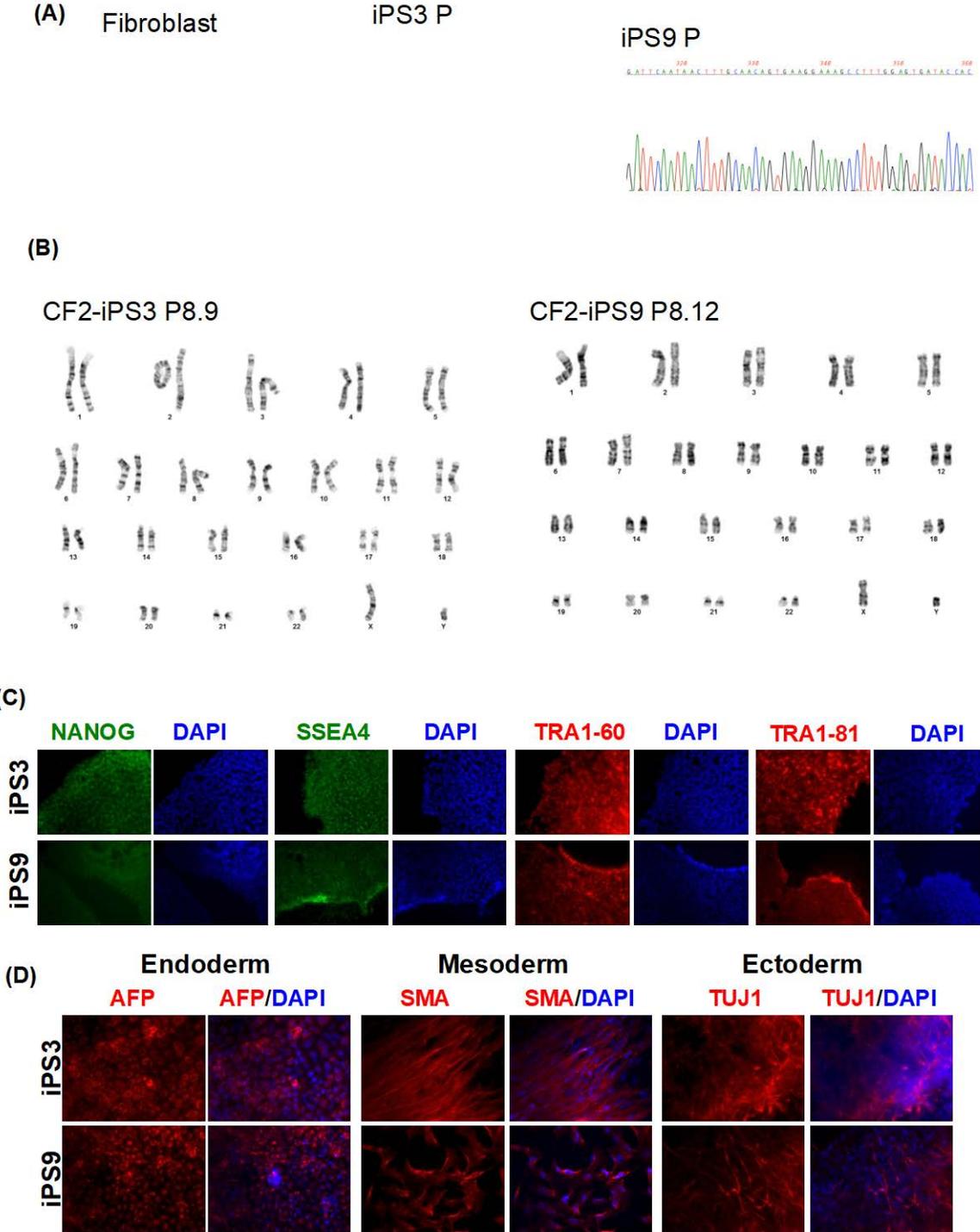
Figure S8: Excision of insertion and Negative selection with FIAU

Table S1: PCR primer

Table S2: Oligo for guide RNA synthesis

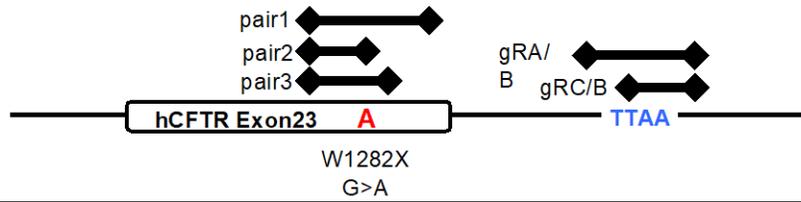
Table S3: Antibody list

Figure S1: Generation of W1282X CF-iPSCs



Assessment of CRISPR/Cas9n for W1282X targeting

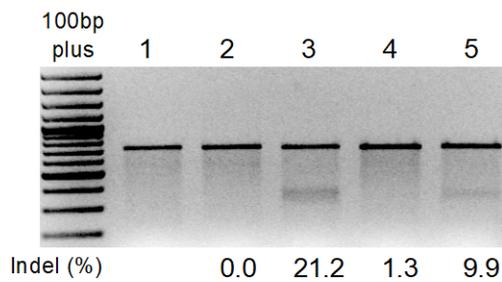
A



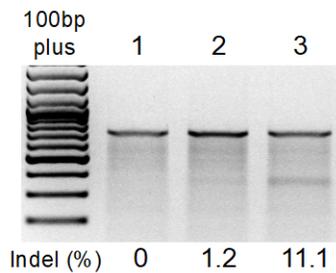
CRISPR Cas9n	left gRNA (5'>3')	Target mutation	right gRNA (5'>3')	Target mutation	gRNA offset
Pair 1	CCCAAGACACACCATCGATC tgg	No	AGCCTTTGGAGTGATACCAC agg	No	28 bp
Pair 2			CAATAACTTTGCAACAGTGA agg	Yes	3 bp
Pair 3			CAACAGTGAAGGAAAGCCTT tgg	Yes	14 bp
Pair w2			CAATAACTTTGCAACAGTGG agg	No	3 bp

gR	left gRNA (5'>3')	gR	right gRNA (5'>3')	offset
A	CAAGGCTCCCACTGTAAATT tgg	B	CTTTTTTGTTCCTGCTTCAG tgg	21 bp
C	TGTTAAAAAACAAGCCCA agg			3 bp

B



lane	Nuclease pair	Fragment size (bp)	
1	GFP	709	
2	gR left	709	
3	gR pair1	357	352
4	gR pair2	345	364
5	gR pair3	350	359

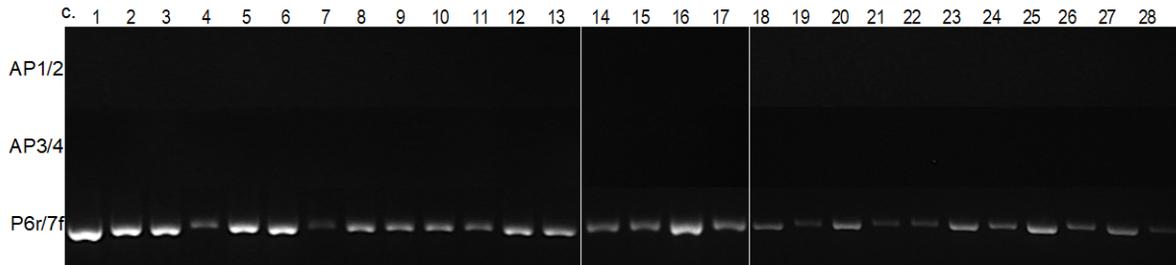


lane	Nuclease pair	Fragment size (bp)	
1	GFP	709	
2	gR pair 2	345	364
4	gR pair w2	345	364

Figure S3: All PCR pic for 1st screening on PuroR

Experiment	cell number	donor DNA (ug)	CRISPR (ug each)	Puro (ug/ml)	Puro ^R	HDR	Insertion
H	1.5 x10 ⁶	CF2A (2.5)	-	0.5	28	0	0
I	1.5 x10 ⁶	CF2A (2.5)	pair2 (2.5)	0.5	35	2 (5.7%)	20 (57.1%)
J	1.5 x10 ⁶	CF2A (2.5)	C/B (2.5)	0.5	36	0	0
K	1.5 x10 ⁶	CF2B (2.5)	-	0.5	31	0	0
L	1.5 x10 ⁶	CF2B (2.5)	pair2 (2.5)	0.5	42	12 (28.6%)	8 (19%)
M	1.5 x10 ⁶	CF2B (2.5)	C/B (2.5)	0.5	32	0	0

Experiment H



Experiment I

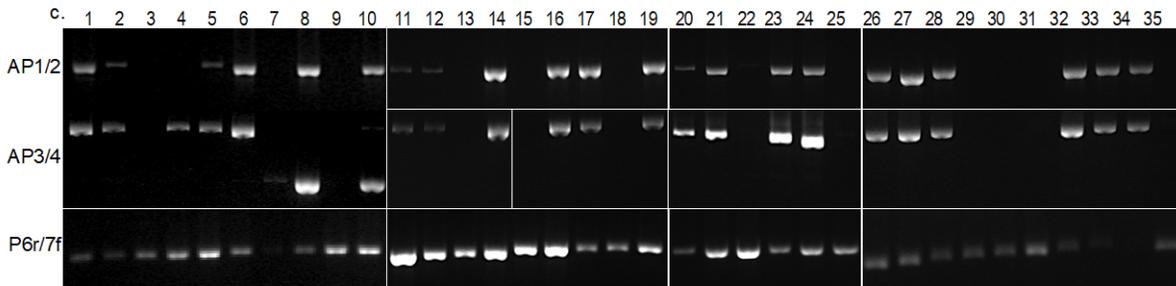
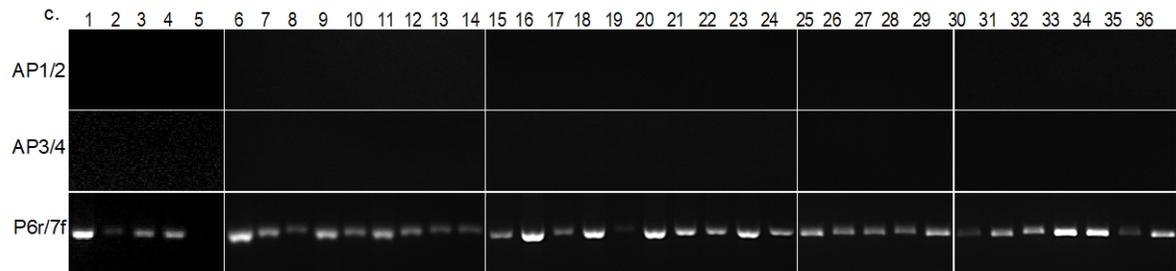
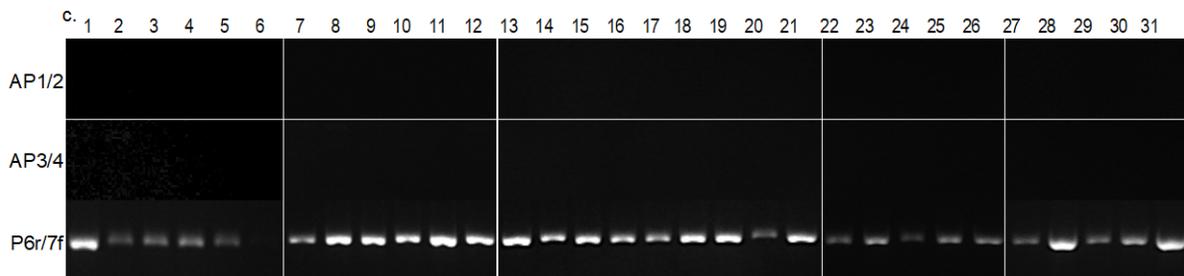


Figure S3: All PCR picture for 1st screening on Puro^R clones

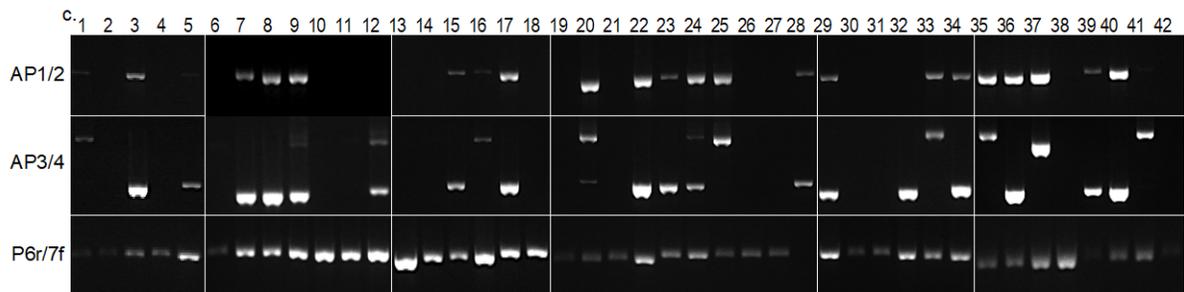
Experiment J



Experiment K



Experiment L



label M

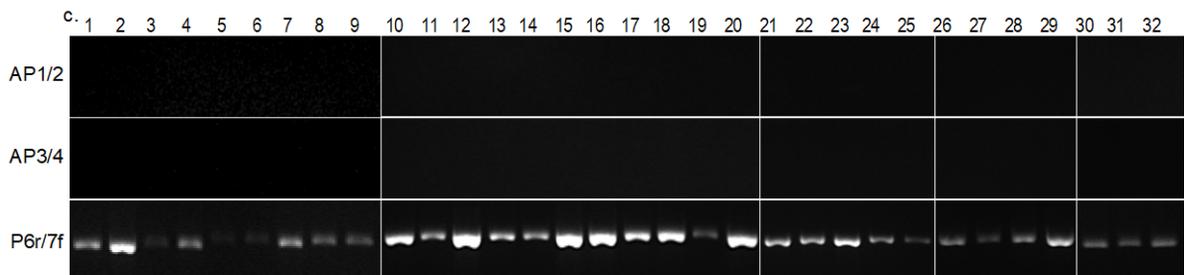
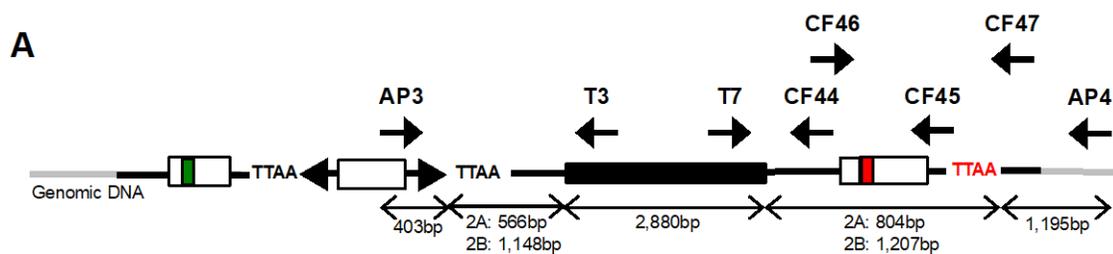


Figure S4: Identification of HMW by PCR



B

Donor DNA (5': bp/ 3': bp)	Primer pair (F/R) and product size (bp)				
	AP3/AP4	AP3/CF44	AP3/T3	CF46/CF47	T7/CF45
CF2A (804/566)	5,848	4,003	1,018	709	583
CF2B (1207/1148)	6,833	4,988	1,600	709	986

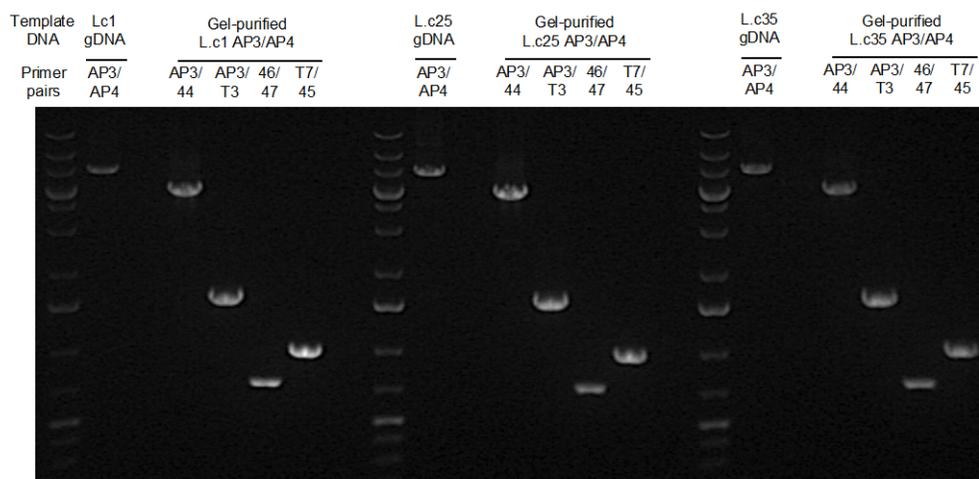
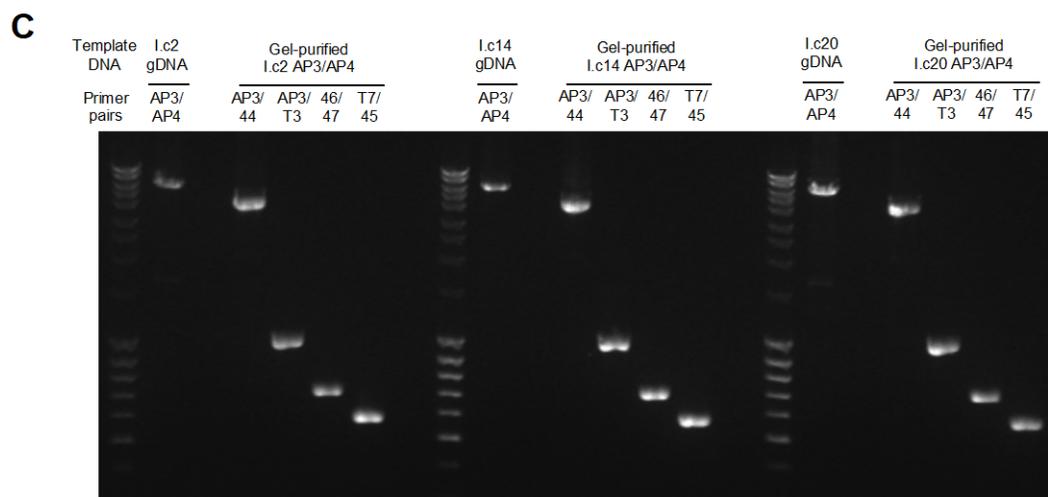
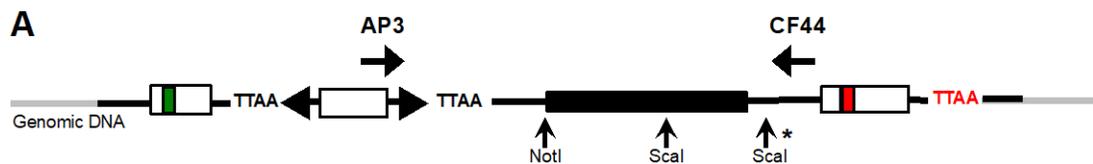


Figure S5: Identification of HMW by RE



B

Donor DNA (5': bp/ 3': bp)	Restriction Enzyme and Fragment size (bp)			
	-	NotI	Scal	NotI/Scal
CF2A (804/566)	4,003	3,030 973	2,753 1,250	1,780 1,250 973
CF2B (1207/1148)	4,988	3,433 1,555	3,335 1,308 345	1,780 1,555 1,308 345

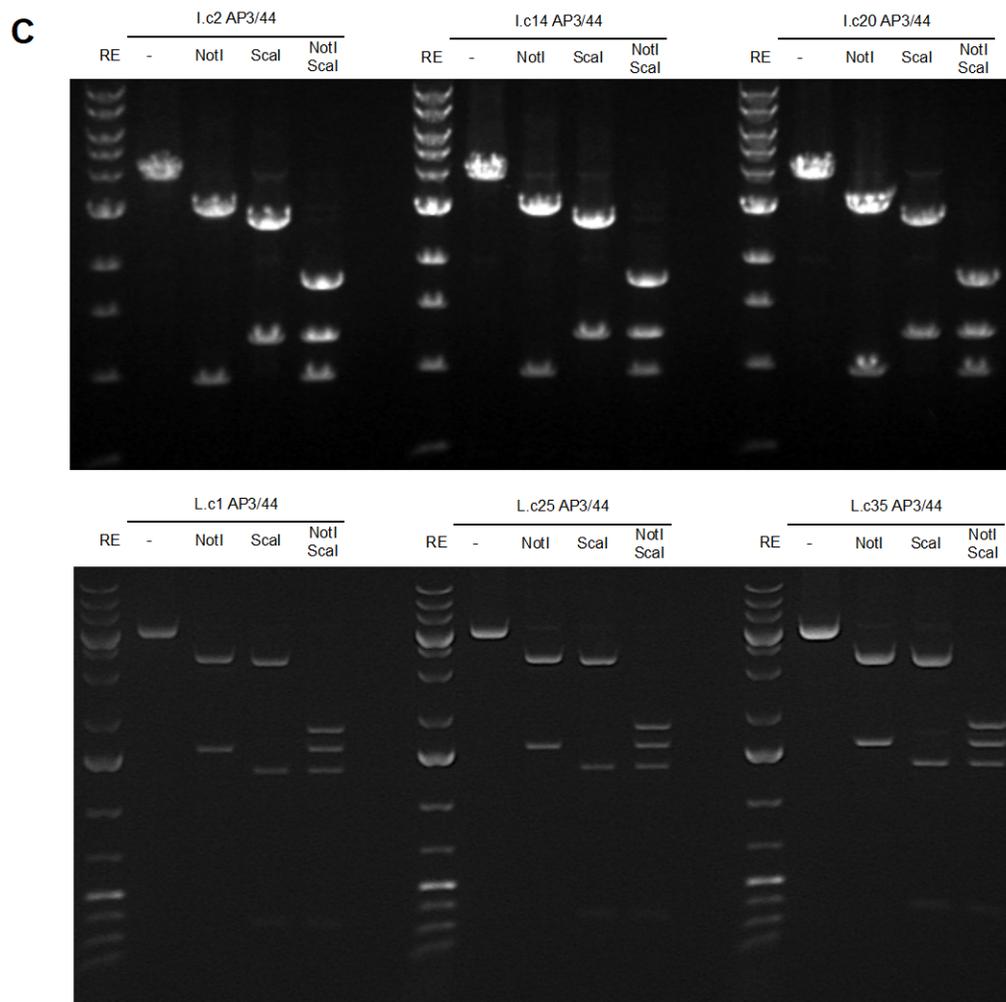


Figure S6: Screening for seamless excision

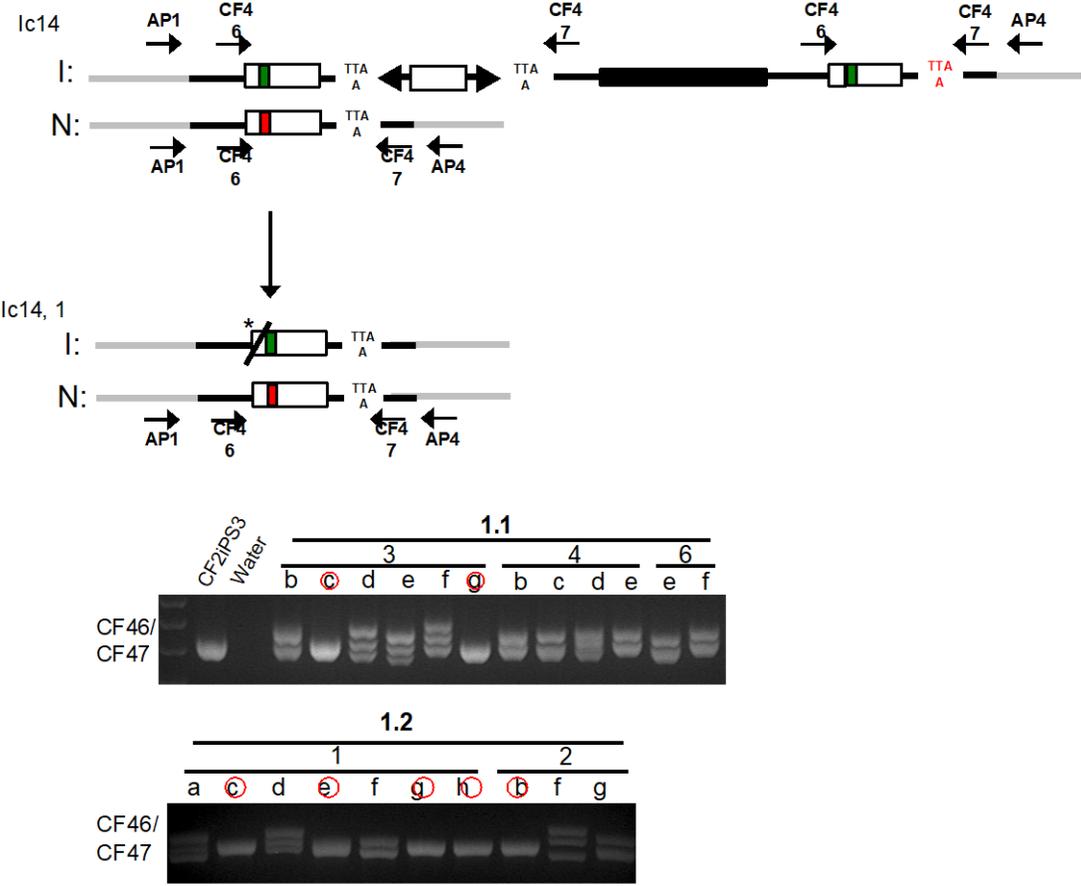
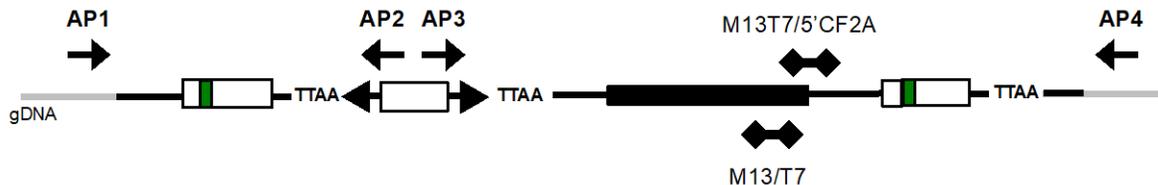


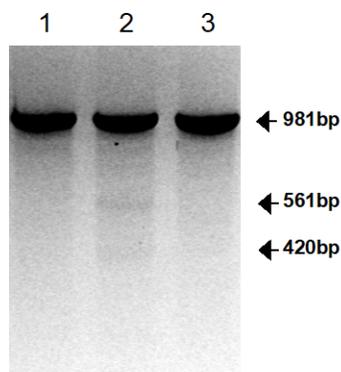
Figure S7: T7E1 assay for plasmid DNA backbone-targeting CRISPR/Cas9n

A



CRISPR Cas9n	left gRNA (5'>3')		right gRNA (5'>3')	gRNA offset
M13	TCGTTTTACAACGTCGTGAC tgg	T7	CGACTCACTATAGGGCGAAT tgg	22 bp
M13T7	GTCGTATTACGCGCGCTCAC tgg	5'CF2A	GCCTTGCAGGTCTCTGCTTC tgg	24 bp

B



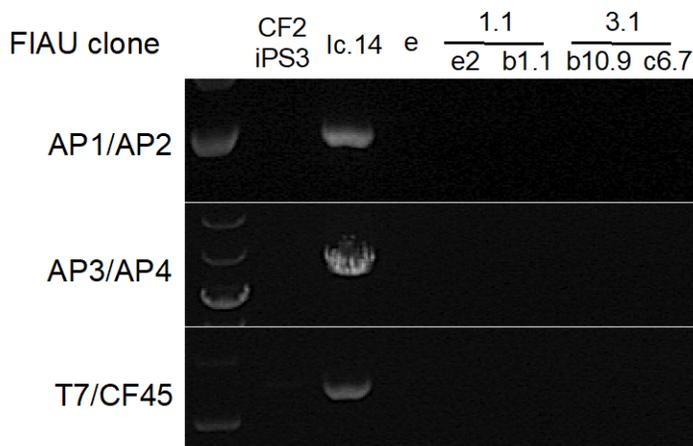
lane	Nuclease pair	Fragment size (bp)	
1	M13/T7	393	588
2	M13T7/5'CF2A	420	561
3	PBase	981	

Fig S8. Excision of insertion and Negative selection with FIAU

A I.c14 P8.13.11

Experiment	cell numbers	vector (ug each)	FIAU (uM)	Picked	PB ⁻ (-/test)	corr/seq
a	1.5 x10 ⁶	pair2 (2.5)	0.25	96	2/96	0/2
b	1.5 x10 ⁶	M13/T7(2.5)	0.25	96	0/96	NA
c	1.5 x10 ⁶	M13T7/5' CF2(2.5)	0.25	96	2/96	2/2
d	1.5 x10 ⁶	PBase (4.5)	0.25	96	1/96	NA

B



C

• Sequence of excised clone

CF2-iPS3 I.c14_{FIAU}1.1-e2: W1282W (G>G) + 8bp deletion/ W1282A (G>A)

| : ccagatcgatggtgtgtcttgggattcaataac agtgAggaaagcctttggagtgataccacag

N: ccagatcgatggtgtgtcttgggattcaataactttgcaacagtgAaggaaagcctttggagtgataccacag

CF2-iPS3 I.c14_{FIAU}1.1-b1: W1282A (G>A) + 6bp insertion/ W1282? (G>?) + ?bp insertion

| : ccagatcgatggtgtgtcttgggattcaataactttgcaa INS

N: ccagatcgatggtgtgtcttgggattcaataactttgcaacag caacagtgAaggaaagcctttggagtgataccacag

CF2-iPS3 I.c14_{FIAU}3.1-b10: W1282W (G>G)/ W1282X (G>A)

CF2-iPS3 I.c14_{FIAU}3.1-c6: W1282W (G>G)/ W1282X (G>A)

| : ccagatcgatggtgtgtcttgggattcaataactttgcaacagtg gaggaaagcctttggagtgataccacag

N: ccagatcgatggtgtgtcttgggattcaataactttgcaacagtgAaggaaagcctttggagtgataccacag