RESEARCH ARTICLE

A versatile toolbox for knock-in gene targeting based on the Multisite Gateway technology

Sho Yoshimatsu[1,2☯], **Takefumi Sone**[1☯], **Mayutaka Nakajima**[1], **Tsukika Sato**[1], **Ryotaro Okochi**[1], **Mitsuru Ishikawa**[1], **Mari Nakamura**[1], **Erika Sasaki**[3,4], **Seiji Shiozawa**[1]*, **Hideyuki Okano**[1,4]*

1 Department of Physiology, School of Medicine, Keio University, Shinjuku-ku, Tokyo, Japan, **2** Laboratory for Proteolytic Neuroscience, RIKEN Center for Brain Science, Wako City, Saitama, Japan, **3** Central Institute for Experimental Animals, Kawasaki City, Kanagawa, Japan, **4** Laboratory for Marmoset Neural Architecture, RIKEN Center for Brain Science, Wako City, Saitama, Japan

☯ These authors contributed equally to this work.
* hidokano@keio.jp (HO); shiozawa@keio.jp (SS)

Check for updates

Abstract

Knock-in (KI) gene targeting can be employed for a wide range of applications in stem cell research. However, vectors for KI require multiple complicated processes for construction, including multiple times of digestion/ligation steps and extensive restriction mapping, which has imposed limitations for the robust applicability of KI gene targeting. To circumvent this issue, here we introduce versatile and systematic methods for generating KI vectors by molecular cloning. In this approach, we employed the Multisite Gateway technology, an efficient *in vitro* DNA recombination system using proprietary sequences and enzymes. KI vector construction exploiting these methods requires only efficient steps, such as PCR and recombination, enabling robust KI gene targeting. We show that combinatorial usage of the KI vectors generated using this method and site-specific nucleases enabled the precise integration of fluorescent protein genes in multiple loci of human and common marmoset (marmoset; *Callithrix jacchus*) pluripotent stem cells. The methods described here will facilitate the usage of KI technology and ultimately help to accelerate stem cell research.

OPEN ACCESS

Citation: Yoshimatsu S, Sone T, Nakajima M, Sato T, Okochi R, Ishikawa M, et al. (2019) A versatile toolbox for knock-in gene targeting based on the Multisite Gateway technology. PLoS ONE 14(8): e0221164. https://doi.org/10.1371/journal.pone.0221164

Editor: Zongbin Cui, Chinese Academy of Sciences, CHINA

Received: March 11, 2019

Accepted: July 31, 2019

Published: August 27, 2019

Copyright: © 2019 Yoshimatsu et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: Data presented herein were generated during the "Construction of System for Spread of Primate Model Animals" project, performed under the Strategic Research Program for Brain Sciences and Brain Mapping by Integrated Neurotechnologies for Disease Studies (Brain/MINDS) of the MEXT and the AMED (to H.O. and S. S.); and Scientific Research in Innovative Areas, a

Introduction

Knock-in (KI) gene targeting, a gene engineering method mediated by homologous recombination (HR), can be employed for a wide range of applications in stem cell research. Conventional KI methods intended for use in mammalian pluripotent stem cells (PSCs) required long homology arms of over 5–10 kb [1, 2]. Recently, to improve KI efficiency, site specific-endonucleases (SNs), such as zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats/CRISPR associated protein 9 (CRISPR/Cas9) have been utilized and validated in mammalian PSCs [3–5]. Although only short homology arms of 0.5–1 kb are required for KI with SNs, multiple complicated processes, including multiple digestion/ligation steps and extensive restriction mapping, must be taken to generate KI vectors harboring two homology arms of the 5' side

MEXT Grant-in-Aid project FY2014-2018 "Brain Protein Aging and Dementia Control" (to H.O.). This work was financially supported by Grant Number 15K06924 from Japan Society for the Promotion of Science (JSPS) to T.S. Also, S.Y. was supported by RIKEN Junior Research Associate Program.

Competing interests: T.S. is currently a paid employee of Takara Bio Inc., but all the experiments in this study were performed while he was in Keio University, where he had no conflicts of interest with the company. H.O. serves as a paid scientific advisor at SanBio Co. Ltd. and K Pharma Inc. The other authors declare neither financial nor non-financial competing interests. This does not alter the authors' adherence to PLOS ONE policies on sharing data and materials.

and 3' side, and a drug-resistance cassette flanked by the arms to concentrate the homologous recombinants [3, 4, 6]. So far, these time-consuming and cumbersome steps have imposed limitations for the robust applicability of KI gene targeting.

To address this issue, here we introduce versatile and systematic methods for the generation of KI vectors using the Multisite Gateway technology. The Gateway technology, originating from a DNA recombination system of the Enterobacteria phage λ, utilizes proprietary sequences (*att* sequences) and recombination enzymes (BP clonase and LR clonase) for specific recombination [7]. Distinct from commonly used restriction enzymes that recognize 6–8 bp DNA sequences, the recombination enzymes recognize specific sequences (a core 21 bp sequence in each *att* site), which renders it unnecessary to perform extensive restriction mapping for generating long or complicated vectors. This system involves two of the following reactions for specific recombination between *att* sites: (1) the BP reaction (*att*B × *att*P → *att*L + *att*R), which is mediated by the integrase (Int) and integration host factor (IHF) proteins (so called BP clonase), and (2) the LR reaction (*att*L × *att*R → *att*B + *att*P), which is mediated by the Int, IHF and excisionase (Xis) enzymes (so called LR clonase). In addition, because there are six mutant sequences available for each *att* sequence (*att*B1–6, *att*P1–6, *att*L1–6 and *att*R1–6; each of these *att*Bx or *att*Lx (x: 1–6) is specifically recombined with the corresponding *att*Px or *att*Rx by BP or LR reaction, respectively), multiple recombination cloning can be performed in a specific manner using this system (termed Multisite Gateway technology), the utility of which has been validated in mammalian and plant cells [8, 9].

In the current study, we employ our methods to generate KI vectors based on the Multisite Gateway technology. Although similar attempts have been conducted for mammalian cells previously [6, 10], we developed novel methods for generating KI vectors for (1) the introduction and correction of mutations, and (2) the integration of fluorescent/luminescent reporter gene(s) in specific loci, leaving only minimal residual sequence in the genome following KI and cassette excision. Furthermore, we demonstrate that the latter can be used in both human and common marmoset (marmoset; *Callithrix jacchus*) PSCs.

Materials and methods

Cell culture and transfection

A human induced pluripotent stem cell (hiPSC) line 201B7 was used in the current study, which was previously established [11] and kindly provided by Shinya Yamanaka (Kyoto University). The hiPSCs were cultured on Laminin-511 (Nippi) in StemFit AK02N medium (AK02N; Ajinomoto). For passaging, the hiPSCs were pre-treated with 10 μM ROCK inhibitor, Y-27632 (Merck) in AK02N at 37˚C for one hour. The cells were next dissociated into single cells by TrypLE Select (Thermo Fisher), and were plated to Laminin-511-coated plates in AK02N supplemented with 10 μM Y-27632. Transfection and positive/negative selection of hiPSCs were performed as described previously [12, 13]. In brief, hiPSCs were dissociated into single cells by TrypLE Select, and 10 μg of each KI vector with 5 μg of each Cas9-gRNA vector (for KI) or 10 μg of pCAGS-PBx (for cassette excision) was transfected into 1 × 10^6 hiPSCs by electroporation using a NEPA21 electroporator (NEPA Gene) at an optimized poring pulse condition of 125 mV for 5 ms. The cells were immediately plated onto new Laminin-511-coated plates in AK02N supplemented with 10 μM Y-27632 (Day 0). On day 1, the medium was changed to AK02N without Y-27632. From day 3, cells were treated with 1 μg/ml puromycin (Sigma) or 10 μM ganciclovir (Nacalai Tesque) with 10 μM Y-27632 for positive or negative selection, respectively. From day 5, Y-27632 was removed from the medium.

A marmoset embryonic stem cell (cjESC) line, CMES40 was used in the current study, which was previously established [14]. The cjESCs were cultured on 30 Gy-irradiated mouse embryonic fibroblasts (MEFs) in ES medium (ESM). ESM consisted of 1× Knockout Dulbecco's modified Eagle's medium (Thermo Fisher) supplemented with 20% Knockout Serum Replacement (KSR; Thermo Fisher), 0.1 mM MEM Non-Essential Amino Acids Solution (NEAA; Sigma), 1 mM L-glutamine (L-glu; Thermo Fisher), 0.2 mM β-mercaptoethanol (2ME; Thermo Fisher) and 10 ng/ml human recombinant basic fibroblast growth factor (Thermo Fisher). For passaging, the cjESCs were pre-treated with 10 μM Y-27632 in ESM at 37˚C for one hour. The cells were then dissociated by 0.25% trypsin-ethylenediaminetetraacetic acid (Nacalai Tesque), and mechanically separated from MEFs. After removing MEFs, cjESCs were plated onto new MEFs. Irradiated MEFs were prepared on gelatin-coated 100-mm dishes or 6-well plates, at a density of 2.5×10^6 cells / 100-mm dish or 6-well plate prior to seeding cjESCs. Transfection and positive/negative selection of cjESCs were performed as described previously [15, 16]. In brief, Lipofectamine LTX with PLUS reagent (Thermo Fisher) was used for transfection according to the manufacturer's instructions. For selection, 25 μg/ml hygromycin (Sigma), 1 μg/ml puromycin, 50 μg/ml G418 or 10 μM ganciclovir was used.

Differentiation of PSCs

Spontaneous differentiation of cjESCs was performed by embryoid body (EB) formation, as previously described [14, 16].

Hippocampal differentiation of hiPSCs was performed as previously described [17] with slight modifications. In brief, sub-confluent hiPSCs were dissociated into single cells using TrypLE Select, and transferred into conical V-bottom-wells of low-cell-adhesion 96-well plates (Sumitomo Bakelite) at a density of 10,000 cells per well in 100 μl differentiation medium (Day 0). The differentiation medium consisted of Glasgow's MEM (Thermo Fisher) supplemented with 20% KSR, 0.1 mM NEAA, 1 mM L-glu, 1 mM Sodium pyruvate (Sigma), 0.1 mM 2ME, 3 μM IWR1e (Calbiochem), 5 μM SB431542 (Tocris) and 20 μM Y-27632. The medium was changed every few days until day 18. On day 18, the EBs were collected, and transferred to wells of Ultra-low attachment surface 6-well plates (Corning) in induction medium. The induction medium consisted of DMEM/F-12 (Sigma), 1% N2 supplement (Thermo Fisher), 0.1 mM NEAA, 1 mM L-glu supplemented with 3 μM CHIR99021 (Cayman), 8 ng/ml human recombinant BMP4 (R&D Systems) and 10% FBS. On day 30, the EBs were briefly dissociated by pipetting, and plated onto each well of 8-well chamber slides (Iwaki) which were coated with poly-l-ornithine/fibronectin (Sigma) in neuronal medium consisting of 1× Neurobasal medium (Thermo Fisher) supplemented with 10% FBS, 1 mM L-glu, 1% N2 supplement and 10 ng/ml recombinant human BDNF (Thermo Fisher).

FACS analysis

Before FACS analysis, cells were incubated in Accutase (Nacalai Tesque) at 37˚C for an hour for single-cell dissociation. Following dissociation, the cells were suspended in FACS buffer consisting of 1% inactivated fetal bovine serum (FBS), 5 mM EDTA (Sigma) and 10 μM Y-27632 in 1× PBS buffer, supplemented with 0.1% Propidium iodide staining solution (Sigma) for removing dead cells. FACS analysis was performed using the FACSAria III cell sorter (BD) with the FACSuite software (BD) following the manufacturer's instructions.

Immunochemistry

For immunochemical analysis, cells were fixed with 4% paraformaldehyde or 100% ethanol for 15–30 min at room temperature. After incubating with blocking buffer (PBS containing 0.05% Tween 20 (Sigma) and 10% goat or donkey serum) for 30–60 min at room temperature, the cells were incubated with primary antibodies at 4˚C overnight. After incubation with primary antibodies, the cells were washed with PBS three times, and incubated with Alexa 488, 555 and 647-conjugated secondary antibodies (Thermo Fisher) and 10 mg/ml Hoechst 33258 (Sigma) for one hour at room temperature. Primary antibodies used in the current study are as follows: anti-NANOG (1:500; 1E6C4; Cell Signaling), SSEA4 (1:500; MAB4304; Merck), TRA-1-60 (1:500; MAB4360; Merck), βIII-tubulin (1:400; 2G10; Abcam), αSMA (1:1000; 1A4; Sigma), SOX17 (1:200; AF1924; R&D Systems), HNF3β/FOXA2 (1:500; D5606; Cell Signaling), RFP (1:500, ab62341, Abcam) and VGLUT2 (1:500, ab79157, Abcam).

Vector construction

Plasmid vectors used and described in the current study were listed in the Table A in S1 Supporting Information, and schematic drawings are shown in Figs 1–5, 6A, 7A, 8A and 9. All the vectors will be provided by the corresponding authors upon request. We attached entire DNA sequences of the vectors in the S1 sequence. Recombinant DNA experiments were performed in accordance with institutional guidelines on Recombinant DNA experiments at Keio University, and approved by the Recombinant DNA Experiment Safety Committee of Keio University (approval number: 27–023 and 27–034).

Gateway vectors

pUC-DEST-R3R4(R) (Fig 2B), pDONR-P3P1r (Fig 1B, left), pDONR-P2rP4 (Fig 1B, right), pENTR-L1-PGK-PuroTK-L2 (Fig 2A), pENTR2-L3-PBL-R1 (Fig 3B, left) and pENTR2-R2-PBR-L4 (Fig 3B, right) were generated and described previously [8, 12, 18–20].

pENTR-L1-PGK-HygTK-L2 (Fig 2A), pENTR-L1-PGK-Neo-L2 (Fig 2A), pENTR-L1-PGK-TKNeo-L2 (Fig 2A), pENTR-L1-PGK-TKBsd-L2 (Fig 2A), pUC-DEST-DTA-DEST-R3R4(R) (Fig 2B) were constructed from the corresponding vectors, such as pENTR-L1-PGK-PuroTK-L2 (Fig 2A) or pUC-DEST-R3R4(R) (Fig 2B).

pATD (Fig 2B) was constructed based on pUC-DEST-R3R4(R) and pSINTK.

pENTR2-L3-SfoI-EGFP-PBL-R1 (Fig 5, left), pENTR2-L3-SfoI-Venus-PBL-R1 (Fig 5, left), pENTR2-L3-SfoI-Cerulean-PBL-R1 (Fig 5, left), pENTR2-L3-SfoI-tdTomato-PBL-R1 (Fig 5, left), pENTR2-L3-SfoI-mCherry-PBL-R1 (Fig 5, left), pENTR2-L3-SfoI-iRFP713-PBL-R1 (Fig 5, left), pENTR2-L3-SfoI-IRES-EGFP-PBL-R1 (Fig 5, right) and pENTR2-L3-SfoI-EGFP-sec-NanoLuc-PBL-R1 (Fig 5, right) were constructed by inserting PCR-amplified fluorescent/luminescent protein genes into an *Hpa*I-digested pENTR2-L3-SfoI-R1 by seamless cloning with the GeneArt Seamless Cloning and Assembly Enzyme Mix (Thermo Fisher). To construct pENTR2-L3-SfoI-R1, a bacterial backbone of PX330 (a gift from Feng Zhang (Addgene plasmid # 42230)) was replaced with that of pENTR2-L3-PBL-R1 (Fig 3B, left), and the resultant vector was digested by *Age*I (NEB) and *Eco*RI (NEB), followed by ligation with an *Age*I and *Eco*RI-digested custom DNA fragment (IDT) which harbors the *att*L3 sequence, a *Sfo*I site, a *P2A* self-cleaving peptide sequence, an *Hpa*I site, a *Piggybac* 5'ITR and the *att*R1 sequence (reverse), in order from upstream to downstream.

To construct pDEST-spCas9-tdTomato, the T2A-tdTomato (silent mutations were introduced in the *tdTomato* gene for deleting two *Bbs*I sites) sequence (IDT custom oligo) was amplified and inserted into an *Eco*RI-digested PX330 by seamless cloning. The resultant vector was digested by *Not*I and *Sfo*I, and the *att*R3-CmR-ccdB-*att*R4 (reverse) fragment, amplified from pUC-DEST-R3R4(R), was inserted into the vector by seamless cloning.

Cas9-gRNA vectors

For the Cas9-gRNA vector, we used pSpCas9(BB)-2A-Puro (PX459) (a gift from Feng Zhang (Addgene plasmid # 48139)).

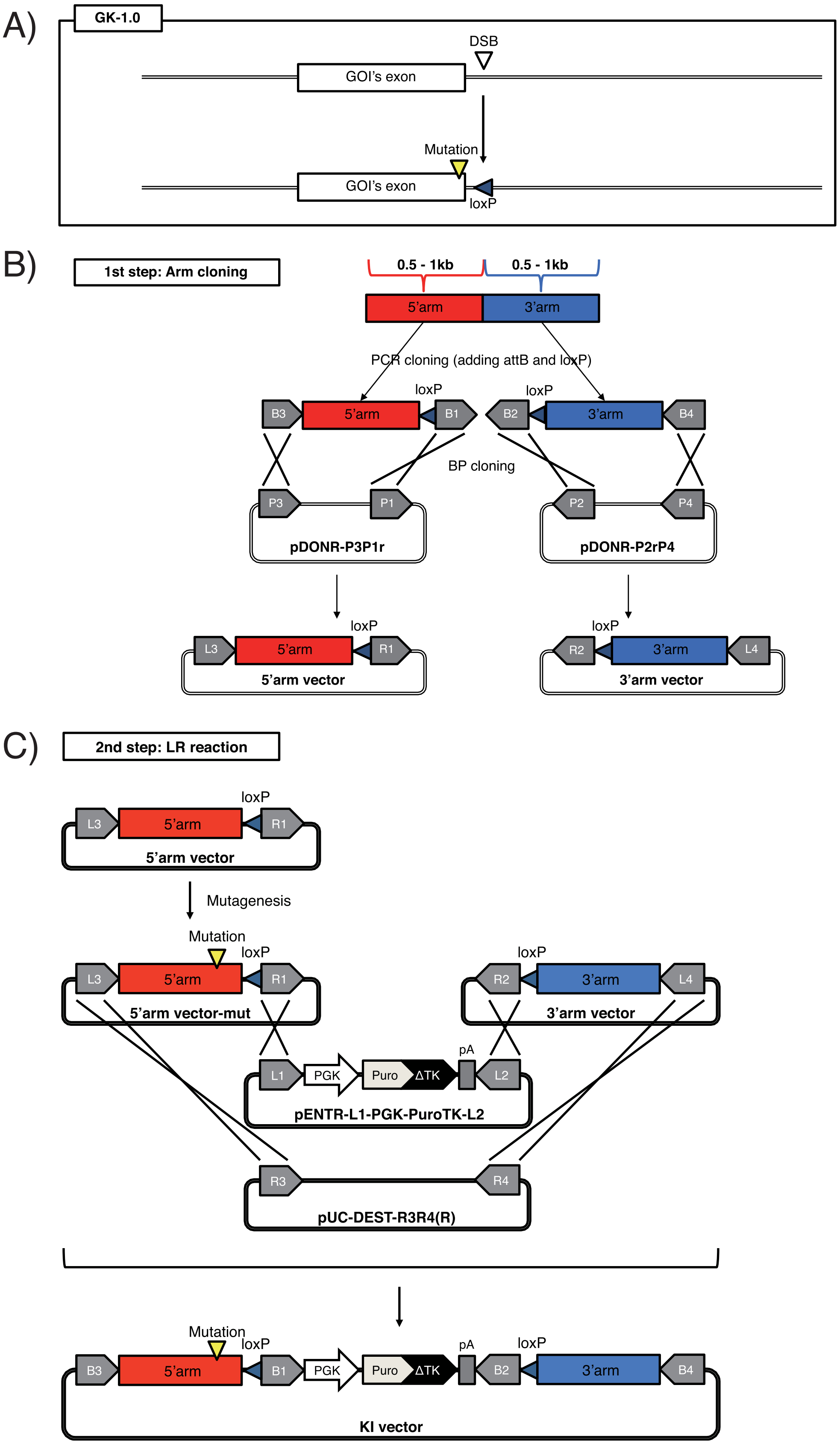
A)
GK-1.0
DSB
GOI's exon
Mutation
GOI's exon
loxP
B)
1st step: Arm cloning
0.5 - 1kb
0.5 - 1kb
5'arm
3'arm
PCR cloning (adding attB and loxP)
loxP
loxP
B3
5'arm
B1
B2
3'arm
B4
BP cloning
P3
P1
P2
P4
pDONR-P3P1r
pDONR-P2rP4
loxP
loxP
L3
5'arm
R1
R2
3'arm
L4
5'arm vector
3'arm vector
C)
2nd step: LR reaction
loxP
L3
5'arm
R1
5'arm vector
Mutagenesis
Mutation
loxP
loxP
L3
5'arm
R1
5'arm vector-mut
R2
3'arm
L4
3'arm vector
pA
L1
PGK
Puro
ΔTK
L2
pENTR-L1-PGK-PuroTK-L2
R3
R4
pUC-DEST-R3R4(R)
Mutation
loxP
pA
loxP
B3
5'arm
B1
PGK
Puro
ΔTK
B2
3'arm
B4
KI vector

Fig 1. Schematic of the GKI-1.0 method. (A) Schematic of the GKI-1.0 method. DSB, double strand break; GOI, gene of interest. (B) The 1st step of GKI-1.0. A 5'-side homology arm (red box) and 3'-side homology arm (blue box) are subcloned into pDONR-P3P1r and pDONR-P2rP4 by PCR and BP cloning, respectively. X-marks indicate recombination that occurs between each corresponding *att*B and *att*P during the BP reaction. (C) The 2nd step of GKI-1.0. X-marks indicate recombination that occurs between each corresponding *att*L and *att*R during the LR reaction. *Chloramphenicol resistance gene (CmR)* and *ccdB* gene are located between *att*R3 and *att*R4 in pUC-DEST-R3R4(R) (not described) for the positive/negative selection of transformants harboring the vector itself or the desired KI vector. PGK, murine *phosphoglycerate Kinase 1 (Pgk-1)* promoter; PuroTK, a bifunctional fusion protein gene between *puromycin-N-acetyltransferase gene (Puro)* and a N-terminal shortened version of herpes simplex virus type 1 *thymidine kinase (ΔTK)* [25]; pA, polyadenylation signal sequence. The entire sequences of pDONR-P3P1r, pDONR-P2rP4 and pUC-DEST-R3R4(R) are attached in the S1 sequence.

https://doi.org/10.1371/journal.pone.0221164.g001

For designing gRNA sequences, we used CRISPRdirect [21]. The gRNA target sequences (not including the PAM sequences) used in the current study are as follows: `CCCCCCTCCA GTTCGCTGTC` (for human *SOX2*), `TGAAGAGTAGCAGTCCCCTT` (for human *PROX1*), `GGAA`

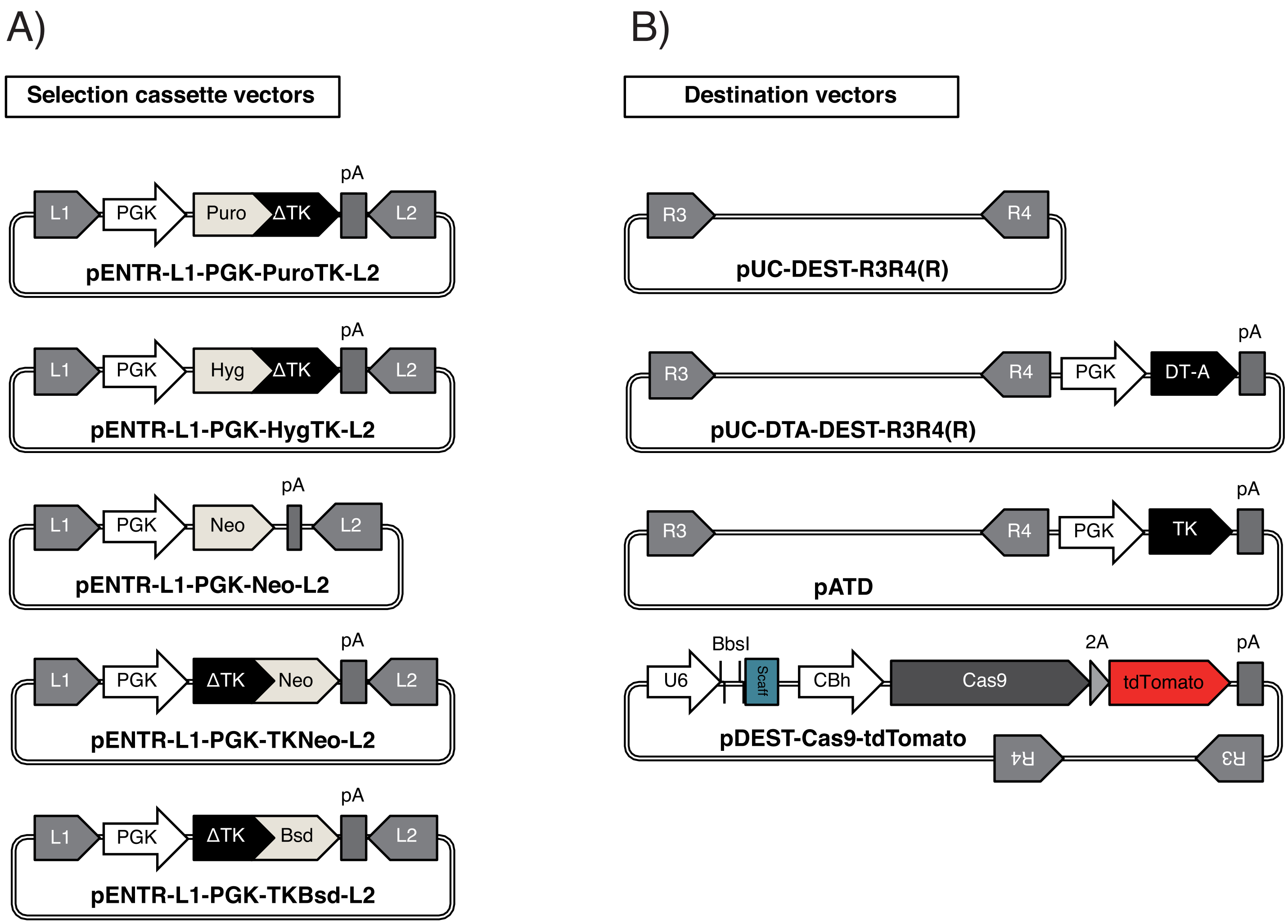
A)
Selection cassette vectors
L1
PGK
Puro
ΔTK
pA
L2
pENTR-L1-PGK-PuroTK-L2
L1
PGK
Hyg
ΔTK
pA
L2
pENTR-L1-PGK-HygTK-L2
L1
PGK
Neo
pA
L2
pENTR-L1-PGK-Neo-L2
L1
PGK
ΔTK
Neo
pA
L2
pENTR-L1-PGK-TKNeo-L2
L1
PGK
ΔTK
Bsd
pA
L2
pENTR-L1-PGK-TKBsd-L2
B)
Destination vectors
R3
R4
pUC-DEST-R3R4(R)
R3
R4
PGK
DT-A
pA
pUC-DTA-DEST-R3R4(R)
R3
R4
PGK
TK
pA
pATD
U6
BbsI
Scaff
CBh
Cas9
2A
tdTomato
pA
pDEST-Cas9-tdTomato
R4
R3

Fig 2. Module vectors for GKI methods. (A) Selection cassette vectors intended for use at the 2nd step of the GKI methods. (B) Destination vectors intended for use at the 2nd step of the GKI methods. The *CmR* and *ccdB* gene are located between *att*R3 and *att*R4 in each destination vector (not described) for the positive/negative selection of transformants harboring the vector itself or the desired KI vector. DT-A, *diphtheria toxin A*; TK, Herpes Simplex Virus-1 *Thymidine Kinase*; U6, mammalian U6 type III RNA polymerase III promoter; BbsI, a sgRNA cloning site (mediated by *Bbs*I digestion); Scaff, scaffold sequence of sgRNA followed by TTTTTT (poly-T terminal signal of RNA polymerase III); CBh, cytomegalovirus enhancer fused to the chicken beta-actin promoter and hybrid intron; Cas9, *spCas9*; 2A, *Thosea asigna* virus 2A peptide sequence (T2A). Whole sequences for all vectors described are provided in the S1 sequence.

https://doi.org/10.1371/journal.pone.0221164.g002

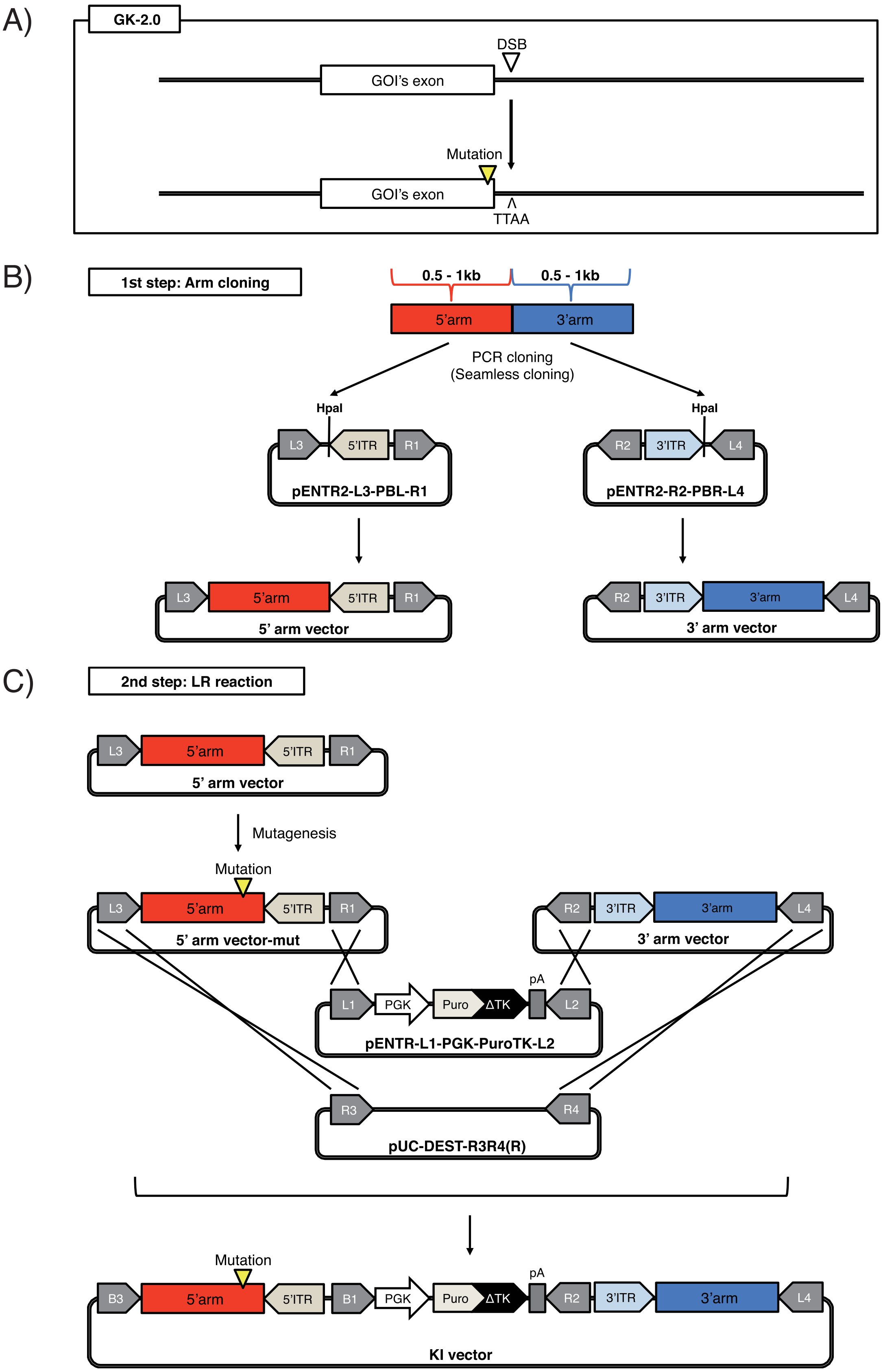
A)
GK-2.0
DSB
GOI's exon
Mutation
GOI's exon
TTAA
B)
1st step: Arm cloning
0.5 - 1kb
0.5 - 1kb
5'arm
3'arm
PCR cloning
(Seamless cloning)
HpaI
HpaI
L3
5'ITR
R1
pENTR2-L3-PBL-R1
R2
3'ITR
L4
pENTR2-R2-PBR-L4
5' arm vector
3' arm vector
C)
2nd step: LR reaction
5' arm vector
Mutagenesis
Mutation
5' arm vector-mut
3' arm vector
L1
PGK
Puro
ΔTK
pA
L2
pENTR-L1-PGK-PuroTK-L2
R3
R4
pUC-DEST-R3R4(R)
B3
B1
KI vector

Fig 3. Schematic of the GKI-2.0 method. (A) Schematic of the GKI-2.0 method. (B) The 1st step of GKI-2.0. A 5'-side homology arm (red box) and 3'-side homology (blue box) arm are subcloned into pENTR2-L3-PBL-R1 and pENTR2-R2-PBR-L4 by PCR and seamless cloning, respectively. HpaI, *Hpa*I restriction enzyme site; 5'ITR and 3'ITR, 5'-side and 3'-side of *Piggybac* inverted terminal repeats. (C) The 2nd step of GKI-2.0. X-marks indicate recombination that occurs between each corresponding *att*L and *att*R during the LR reaction. The entire sequences of pENTR2-L3-PBL-R1 and pENTR2-R2-PBR-L4 are attached in the S1 sequence.

https://doi.org/10.1371/journal.pone.0221164.g003

AATCTTAAGGATCCAT (for marmoset *BLIMP1*) and AAAGCTACTCTCTGCTCATA (for marmoset *STELLA*).

KI vectors

To generate the human SOX2-EGFP KI vector, a 713 bp fragment directly upstream of the termination codon of the human *SOX2* gene was amplified from the human genomic DNA by PCR, and inserted into a *Sfo*I-digested pENTR2-L3-SfoI-EGFP-PBL-R1 (Fig 4B, left) by seamless cloning. Similarly, a 616 bp fragment directly downstream of the termination codon of the human *SOX2* gene was amplified and inserted into an *Hpa*I-digested pENTR2-R2-PBR-L4 by seamless cloning (Fig 4B, right). The resultant vectors were named pENTR2-hSOX2-5'arm-EGFP and pENTR2-hSOX2-3'arm, respectively. In addition, to make the latter unrecognizable by its corresponding gRNA, we performed mutagenesis by PCR using the following primers: TGAGGGTTGGACAGCGAACTGGAGGG and CTGTCCAACCCTCATTAACCCTAGAA. The resultant vector was named pENTR2-hSOX2-3'arm-mut. Then, pUC-DEST-R3R4(R), pENTR-L1-PGK-PuroTK-L2, pENTR2-hSOX2-5'arm-EGFP and pENTR2-hSOX2-3'arm-mut were mixed together and incubated with LR clonase II (Thermo Fisher) at room temperature overnight for LR reaction, and transformed into Stbl3 chemically competent *E. coli* (Stbl3; Thermo Fisher). The resultant vector was named pUC-DEST-hSOX2-EGFP-PuroTK (Fig 4C).

To generate a human PROX1-tdTomato KI vector, a 707 bp fragment directly upstream of the termination codon of the human *PROX1* gene was amplified from the human genomic DNA by PCR and inserted into a *Sfo*I-digested pENTR2-L3-SfoI-tdTomato-PBL-R1 by seamless cloning (Fig 4B, left). Similarly, a 736 bp fragment directly downstream of the termination codon of the human *PROX1* gene was amplified and inserted into an *Hpa*I-digested pENTR2-R2-PBR-L4 by seamless cloning (Fig 4B, right). The resultant vectors were named pENTR2-hPROX1-5'arm-tdTomato and pENTR2-hPROX1-3'arm, respectively. In addition, to make the latter unrecognizable by its corresponding gRNA, we performed mutagenesis by PCR using the following primers: CCCTTTAAATGTCCAAGTTATAT and GGACATTTAAAGGGGACTGCTAC. The resultant vector was named pENTR2-hPROX1-3'arm-mut. Then, pUC-DEST-R3R4(R), pENTR-L1-PGK-PuroTK-L2, pENTR2-hPROX1-5'arm-tdTomato and pENTR2-hPROX1-3'arm-mut were mixed together and incubated with LR clonase II at room temperature overnight for LR reaction and transformed into Stbl3. The resultant vector was named pUC-DEST-hPROX1-tdTomato-PuroTK (Fig 4C).

To generate a marmoset BLIMP1-Venus KI vector, a 611 bp fragment directly upstream of the termination codon of marmoset *BLIMP1* gene was amplified from the marmoset genomic DNA by PCR, and inserted into a *Sfo*I-digested pENTR2-L3-SfoI-Venus-PBL-R1 by seamless cloning (Fig 4B, left). Similarly, a 532 bp fragment directly downstream of the termination codon of the marmoset *BLIMP1* gene was amplified and inserted into an *Hpa*I-digested pENTR2-R2-PBR-L4 by seamless cloning (Fig 4B, right). The resultant vectors were named pENTR2-cjBLIMP1-5'arm-Venus and pENTR2-cjBLIMP1-3'arm, respectively. Then, pUC-DEST-R3R4(R), pENTR-L1-PGK-HygTK-L2, pENTR2-cjBLIMP1-5'arm-Venus and pENTR2-cjBLIMP1-3'arm were mixed together and incubated with LR clonase II at room

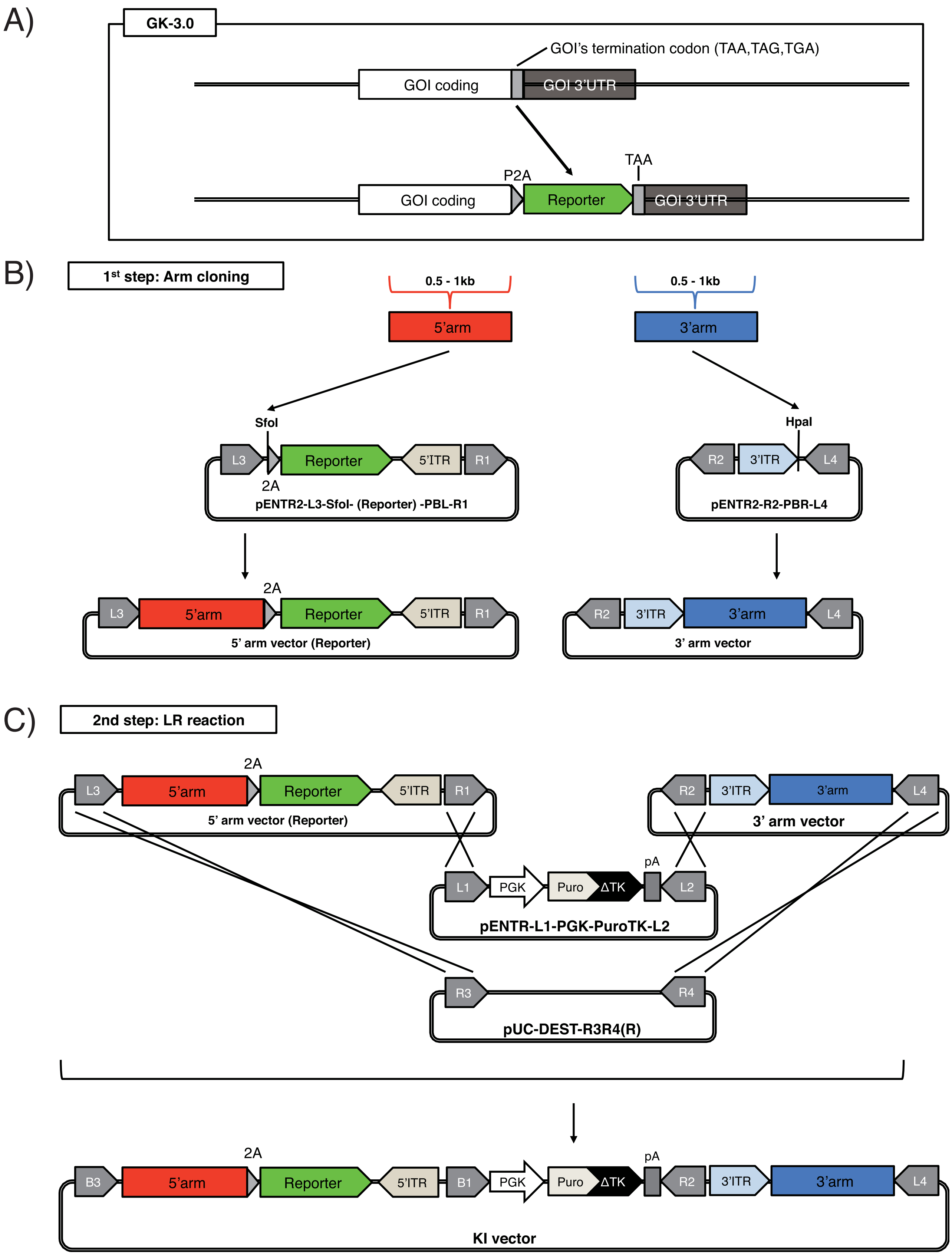
A)
GK-3.0
GOI's termination codon (TAA,TAG,TGA)
GOI coding
GOI 3'UTR
P2A
TAA
GOI coding
Reporter
GOI 3'UTR
B)
1st step: Arm cloning
0.5 - 1kb
5'arm
0.5 - 1kb
3'arm
SfoI
L3
Reporter
5'ITR
R1
2A
pENTR2-L3-SfoI- (Reporter) -PBL-R1
HpaI
R2
3'ITR
L4
pENTR2-R2-PBR-L4
2A
L3
5'arm
Reporter
5'ITR
R1
5' arm vector (Reporter)
R2
3'ITR
3'arm
L4
3' arm vector
C)
2nd step: LR reaction
2A
L3
5'arm
Reporter
5'ITR
R1
5' arm vector (Reporter)
R2
3'ITR
3'arm
L4
3' arm vector
pA
L1
PGK
Puro
ΔTK
L2
pENTR-L1-PGK-PuroTK-L2
R3
R4
pUC-DEST-R3R4(R)
2A
pA
B3
5'arm
Reporter
5'ITR
B1
PGK
Puro
ΔTK
R2
3'ITR
3'arm
L4
KI vector

Fig 4. Schematic of the GKI-3.0 method. (A) Schematic of the GKI-3.0 method. P2A, porcine teschovirus-1 2A peptide sequence. (B) The 1st step of GKI-3.0. A 5'-side homology arm (red box) and 3'-side homology arm (blue box) are subcloned into each pENTR2-L3-SfoI-PBL-R1 reporter vector (described in Fig 5) and pENTR2-R2-PBR-L4 by PCR and seamless cloning, respectively. SfoI, *Sfo*I restriction enzyme site; HpaI, *Hpa*I restriction enzyme site. (C) The 2nd step of GKI-3.0. X-marks indicate recombination that occurs between each corresponding *att*L and *att*R during the LR reaction.

https://doi.org/10.1371/journal.pone.0221164.g004

temperature overnight for LR reaction, and transformed into Stbl3. The resultant vector was named pUC-DEST-cjBLIMP1-Venus-HygTK (Fig 4C).

To generate a marmoset STELLA-Cerulean KI vector, a 500 bp fragment directly upstream of the termination codon of the marmoset *STELLA* gene was amplified from the marmoset

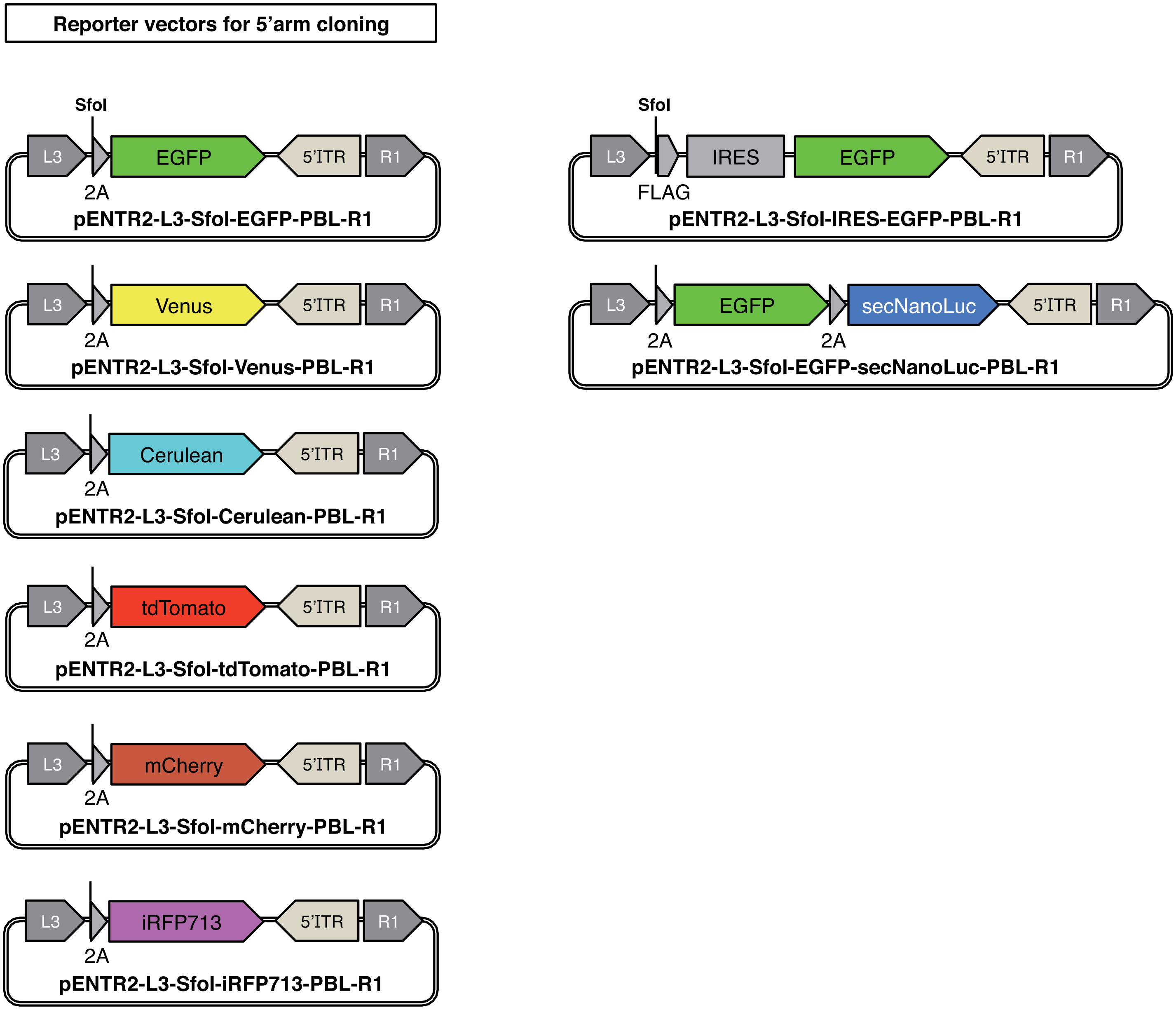
Reporter vectors for 5'arm cloning
SfoI
L3
2A
EGFP
5'ITR
R1
pENTR2-L3-SfoI-EGFP-PBL-R1
Venus
pENTR2-L3-SfoI-Venus-PBL-R1
Cerulean
pENTR2-L3-SfoI-Cerulean-PBL-R1
tdTomato
pENTR2-L3-SfoI-tdTomato-PBL-R1
mCherry
pENTR2-L3-SfoI-mCherry-PBL-R1
iRFP713
pENTR2-L3-SfoI-iRFP713-PBL-R1
SfoI
FLAG
IRES
EGFP
pENTR2-L3-SfoI-IRES-EGFP-PBL-R1
2A
EGFP
2A
secNanoLuc
pENTR2-L3-SfoI-EGFP-secNanoLuc-PBL-R1

Fig 5. Reporter vectors for 5'arm cloning in the GKI-3.0 method. Reporter vectors generated for use in 5' arm cloning with the GK-3.0 method. 2A, porcine teschovirus-1 2A peptide sequence (*P2A*); FLAG, FLAG epitope tag; IRES, internal ribosomal entry site. The entire sequences of the described vectors are attached in the S1 sequence.

https://doi.org/10.1371/journal.pone.0221164.g005

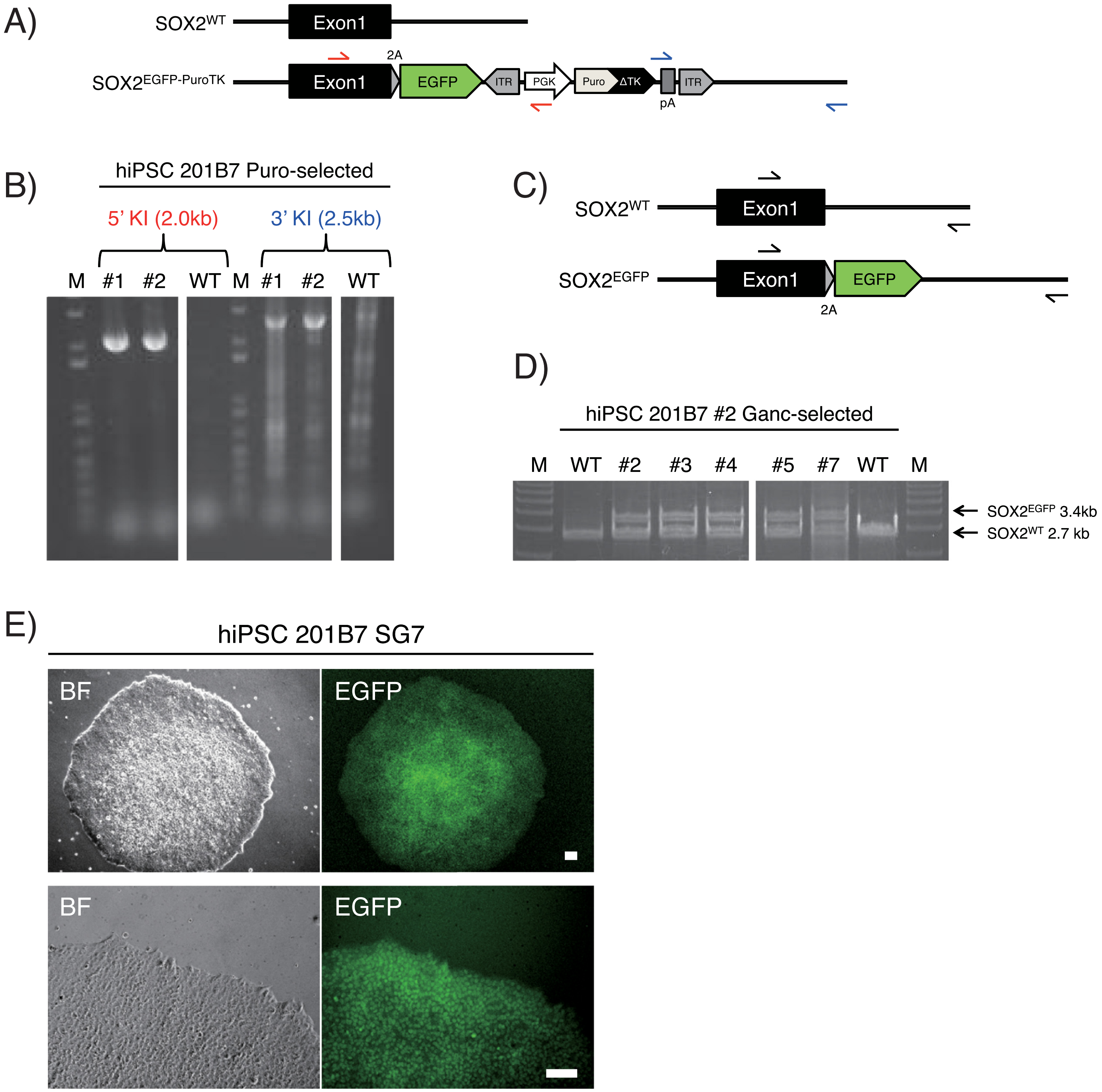
A)
SOX2WT
Exon1
SOX2EGFP-PuroTK
2A
Exon1
EGFP
ITR
PGK
Puro
ΔTK
pA
ITR
B)
hiPSC 201B7 Puro-selected
5’ KI (2.0kb)
3’ KI (2.5kb)
M #1 #2 WT M #1 #2 WT
C)
SOX2WT
Exon1
SOX2EGFP
Exon1
EGFP
2A
D)
hiPSC 201B7 #2 Ganc-selected
M WT #2 #3 #4 #5 #7 WT M
SOX2EGFP 3.4kb
SOX2WT 2.7 kb
E)
hiPSC 201B7 SG7
BF
EGFP
BF
EGFP

Fig 6. Generation of SOX2-EGFP hiPSCs. (A–B) Genotyping PCR analysis of SOX2-EGFP KI hiPSCs. Red arrows indicate the 5’-side genotyping PCR primers recognizing the 5’ external region (hSOX2-forward) and the inner selection cassette (PGK-reverse). Blue arrows indicate the 3’-side genotyping PCR primers recognizing the inner selection cassette (pA-forward) and the 3’ external region (hSOX2-reverse). By using these primer sets, specific PCR bands (2.0 kb and 2.5 kb) appeared in KI hiPSCs (#1 and #2) following transfection of the SOX2-EGFP KI vector and puromycin (Puro) selection. Separated images were cropped from the same gel. M, DNA marker; WT, wild-type 201B7 hiPSCs. (C–D) Genotyping PCR analysis of *PuroTK* cassette excision. Black arrows indicate 5’ and 3’ external primers (hSOX2-forward and hSOX2-reverse) for the analysis. Following transfection of *PBx* and ganciclovir (Ganc) selection, all of the five analyzed clones (#2, 3, 4, 5 and 7) had the cassette excised. Separated images were cropped from the same gel. (E) Representative images of hiPSCs harboring *SOX2*$^{+/EGFP}$ *alleles* (SG7 clone) in bright field (BF) and green fluorescence (EGFP) at low (upper) and high (lower) magnification. Scale bars, 100 μm.

https://doi.org/10.1371/journal.pone.0221164.g006

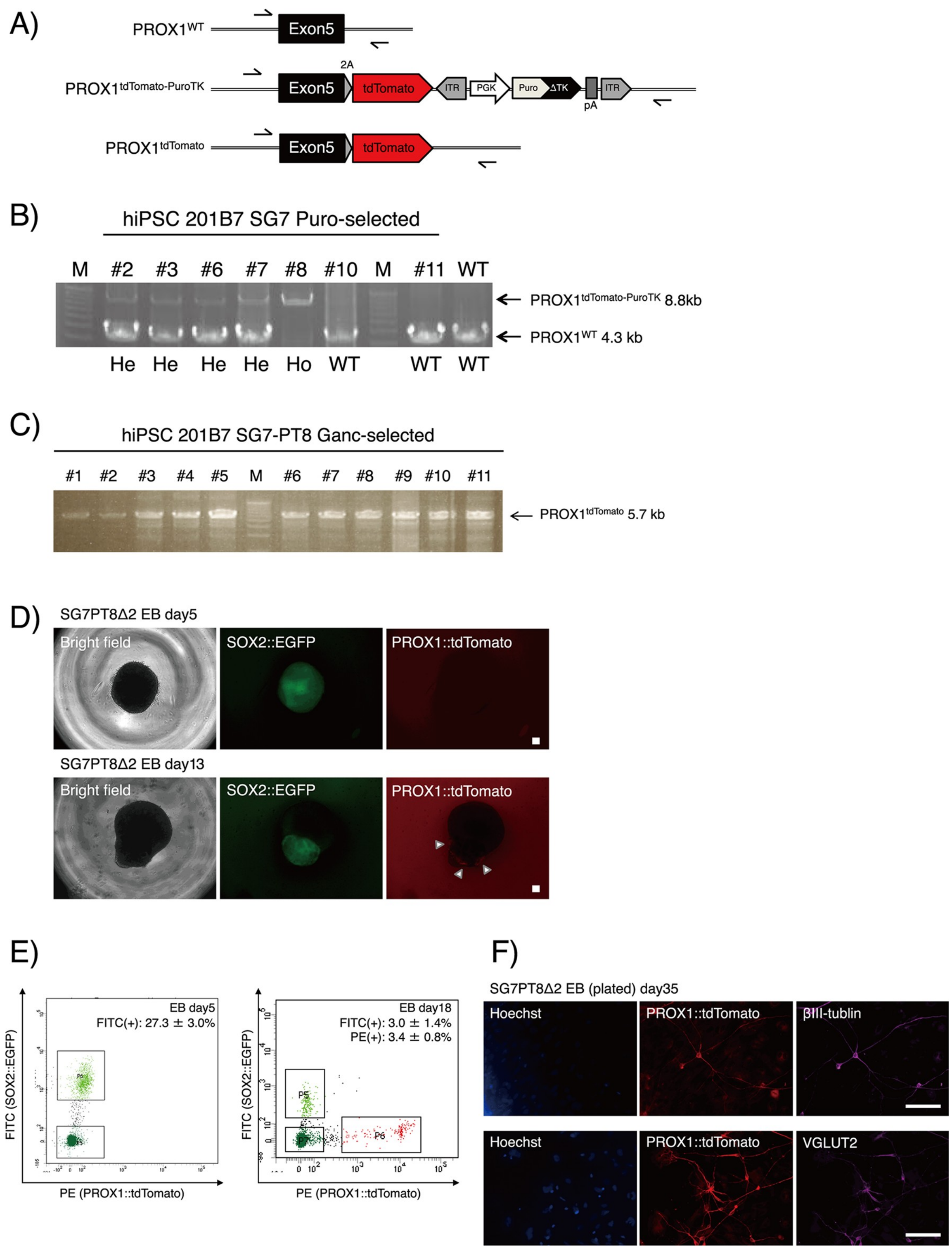

Fig 7. Generation and validation of SOX2-EGFP/PROX1-tdTomato hiPSCs. (A) Graphical schematic of genotyping PCR analysis of PROX1-tdTomato KI hiPSCs. Black arrows indicate the 5' and 3' external primers (hPROX1-forward and hPROX1-reverse) used for the analysis. (B) Genotyping PCR of PROX1-tdtomato KI. In total, we obtained four heterozygous-KI clones (#2, 3, 6 and 7) and one homozygous-KI clone (#8) among seven analyzed clones following puromycin selection. He, heterozygous-KI; Ho, homozygous-KI. (C) Genotyping PCR for the evaluation of *PuroTK* cassette excision. Following the transfection of the *PBx* vector into SG7-PT8 hiPSCs and ganciclovir selection, all of the eleven analyzed clones had the cassette excised homozygously. (D) Representative images of live EBs derived from SG7-PT8-Δ2 hiPSCs during hippocampal differentiation. White arrowheads indicate tdTomato-positive cells located at

the surface of the protrusion. Scale bars, 100 μm. (E) FACS analysis of dissociated EBs on day 5 and day 18 during hippocampal differentiation. Doublets and dead cells were removed from the analysis with gate P2-P4 (not shown). This analysis was biologically and technically triplicated. (F) Representative images of immunochemical analysis of the differentiated SG7-PT8-Δ2 hiPSCs (day 35). Scale bars, 100 μm.

https://doi.org/10.1371/journal.pone.0221164.g007

genomic DNA by PCR, and inserted into a *Sfo*I-digested pENTR2-L3-SfoI-Cerulean-PBL-R1 by seamless cloning (Fig 4B, left). Similarly, a 539 bp fragment directly downstream of the termination codon of the marmoset *STELLA* gene was amplified and inserted into an *Hpa*I-digested pENTR2-R2-PBR-L4 by seamless cloning (Fig 4B, right). The resultant vectors were named pENTR2-cjSTELLA-5'arm-Cerulean and pENTR2-cjSTELLA3-3'arm, respectively. In addition, to make the latter unrecognizable by its corresponding gRNA, we performed mutagenesis by PCR using the following primers: `CTCATATAAATTATATATCAATAATTTTGA` and `ATATATAATTTATATGAGCAGAGAGTAGCTTTC`. The resultant vector was named pENTR2-cjSTELLA-3'arm-mut. Then, pUC-DEST-R3R4(R), pENTR-L1-PGK-PuroTK-L2, pENTR2-cjSTELLA-5'arm-Cerulean and pENTR2-cjSTELLA-3'arm-mut were mixed together and incubated with LR clonase II at room temperature overnight for LR reaction, and transformed into Stbl3. The resultant vector was named pUC-DEST-cjSTELLA-mCerulean-PuroTK (Fig 4C).

Piggybac vectors

For *Piggybac* transposase expression, we used pCMV-HyPBase (kindly provided by Kosuke Yusa and Allan Bradley (Sanger institute)) or pCAGS-PBx (System Bioscience).

To construct a *Piggybac*-based *VASA* reporter vector, a 3011 bp fragment directly upstream of the initiation codon of human *DDX4* (*VASA*) gene was amplified by PCR and inserted into pPB-tdTomato-pNeo. The resultant vector was named pPB-hVASA-tdTomato-pNeo.

Primers for arm cloning, BP reaction and LR reaction in GKI methods

Abbreviations for the vector construction methods in the current study (GKI-1.0, 2.0 and 3.0) are described in the Results section. For GKI-1.0, to amplify the 0.5–1 kb 5'arm sequence containing *att*B1 in the 5' terminus, as well as *att*B3 and *lox*P in the 3' terminus (Fig 1B, left), we performed PCR using three primers as follows: `GGGGCAACTTTGTATAATAAAGTTGNNNNNNNNNNNNNNNNNN` (forward), `TTCGTATAATGTATGCTATACGAAGTTATTNNNNNNNNNNNNNNNNNN` (reverse) and `GGGCTGCTTTTTTGTACAAACTTGATAACTTCGTATAATGTATGCTATAC`, followed by BP reaction with pDONR-P3P1r. N (18 nt) indicates any nucleotide base.

Similarly, to amplify a 0.5–1 kb 3'arm sequence containing *att*B2 and *lox*P in the 5' terminus, as well as *att*B4 in the 3' terminus (Fig 1B, right), we performed PCR using three primers as follows: `TTCGTATAGCATACATTATACGAAGTTATNNNNNNNNNNNNNNNNNN` (forward), `GGGGCAACTTTGTATAGAAAAGTTGNNNNNNNNNNNNNNNNNN` (reverse) and `GGGCAGCTTTCTTGTACAAAGTGGATAACTTCGTATAGCATACATTATAC`, followed by BP reaction with pDONR-P2rP4.

For GKI-2.0, to amplify a 0.5–1 kb 5'arm sequence into an *Hpa*I-digested pENTR2-L3-PBL-R1 (Fig 3B, left), we performed PCR using `CTTTGTATAATAAAGTTGTTNNNNNNNNNNNNNNNNNN` (forward) and `ATGATTATCTTTCTAGGGTTAANNNNNNNNNNNNNNNNNN` (reverse), which was followed by seamless cloning. Similarly, to amplify a 0.5–1 kb 3'arm sequence into an *Hpa*I-digested pENTR2-R2-PBR-L4 (Fig 3B, right), we performed PCR using `CAGACTATCTTTCTAGGGTTAANNNNNNNNNNNNNNNNNN` (forward) and `CAACTTTGTATAGAAAAGTTGTTNNNNNNNNNNNNNNNNNN` (reverse), which was followed by seamless cloning.

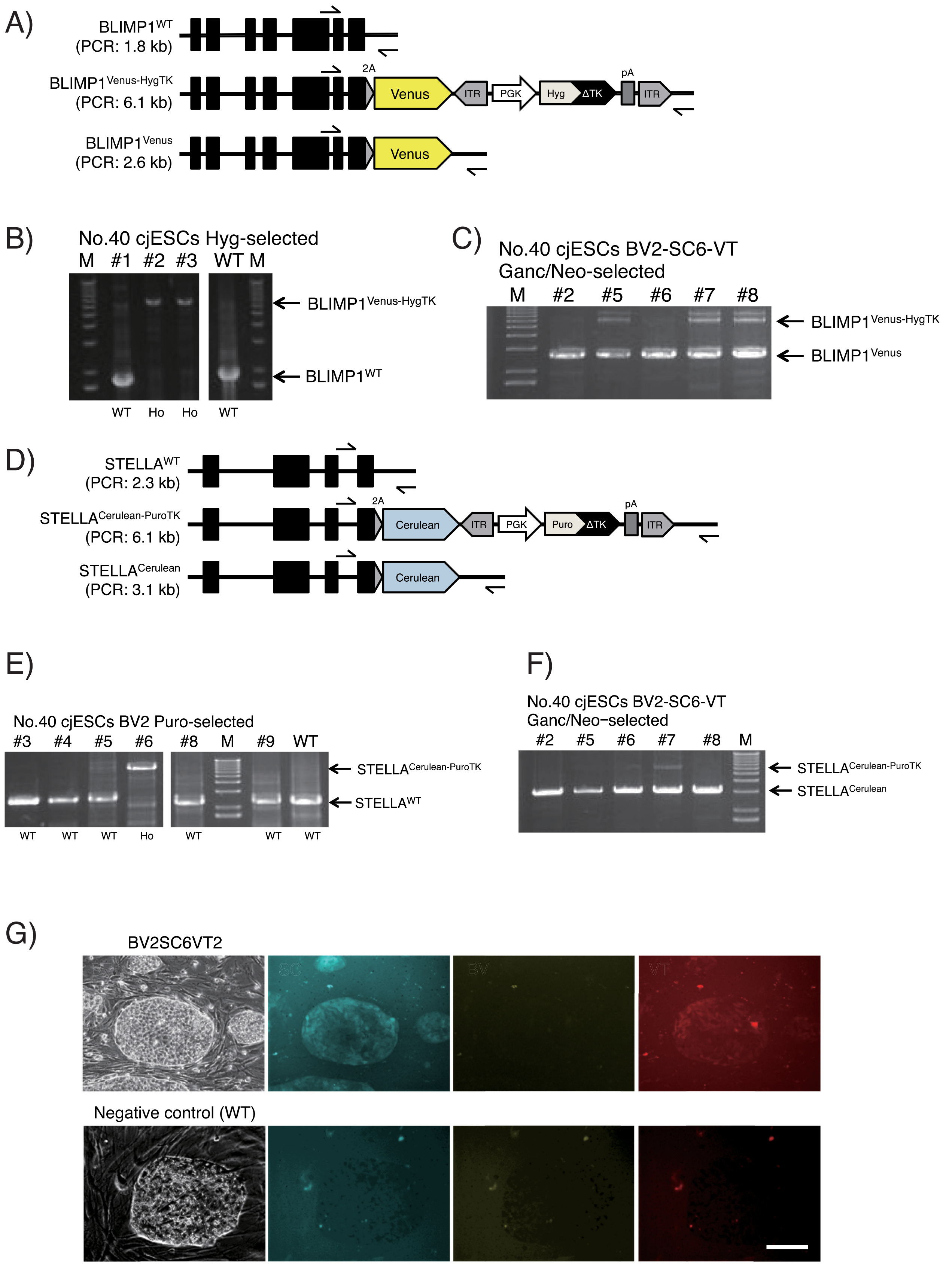
A)
BLIMP1^{WT}
(PCR: 1.8 kb)
BLIMP1$^{Venus-HygTK}$
(PCR: 6.1 kb)
2A
Venus
ITR
PGK
Hyg
ΔTK
pA
ITR
BLIMP1^{Venus}
(PCR: 2.6 kb)
Venus
B)
No.40 cjESCs Hyg-selected
M #1 #2 #3 WT M
BLIMP1$^{Venus-HygTK}$
BLIMP1^{WT}
WT Ho Ho WT
C)
No.40 cjESCs BV2-SC6-VT
Ganc/Neo-selected
M #2 #5 #6 #7 #8
BLIMP1$^{Venus-HygTK}$
BLIMP1^{Venus}
D)
STELLA^{WT}
(PCR: 2.3 kb)
STELLA$^{Cerulean-PuroTK}$
(PCR: 6.1 kb)
2A
Cerulean
ITR
PGK
Puro
ΔTK
pA
ITR
STELLA^{Cerulean}
(PCR: 3.1 kb)
Cerulean
E)
No.40 cjESCs BV2 Puro-selected
#3 #4 #5 #6 #8 M #9 WT
STELLA$^{Cerulean-PuroTK}$
STELLA^{WT}
WT WT WT Ho WT WT WT
F)
No.40 cjESCs BV2-SC6-VT
Ganc/Neo-selected
#2 #5 #6 #7 #8 M
STELLA$^{Cerulean-PuroTK}$
STELLA^{Cerulean}
G)
BV2SC6VT2
SC
BV
VT
Negative control (WT)

Fig 8. Generation of cjESCs harboring PGC-specific reporters. (A–C) Genotyping PCR analysis of the *BLIMP1* locus in cjESCs. Black arrows indicate the 5' and 3' external primers (cjBLIMP1-forward and cjBLIMP1-reverse) used for the analysis. Separated images were cropped from the same gel. (D–F) Genotyping PCR analysis of the *STELLA* locus in cjESCs. Black arrows indicate the 5' and 3' external primers (cjSTELLA-forward and cjSTELLA-reverse) for the analysis. Separated images were cropped from the same gel. (G) Representative images of undifferentiated BV2-SC6-VT2 cjESCs (upper) and WT cjESCs (lower) in bright field (BF), cyan fluorescence (SC), yellow fluorescence (BV) and red fluorescence (VT). Scale bars, 100 μm.

https://doi.org/10.1371/journal.pone.0221164.g008

For GKI-3.0, to amplify a 0.5–1 kb 5'arm sequence into each *Sfo*I-digested pCBh-SfoI reporter vector (Fig 4B, left), we performed PCR using AAGTTGTTGCCACCATGGGCNNNNNNNNNNNNNNNNNNN (forward) and AACAGAGAGAAGTTCGTGGCNNNNNNNNNNNNNNNNNNNN (reverse), followed by seamless cloning. Amplification and seamless cloning of the 0.5–1 kb 3'arm sequence in GKI-3.0 was performed similarly as in GKI-2.0.

Genotyping

For genotyping PCR analysis, cells were lysed overnight at 55˚C in cell lysis buffer consisting of Tris-HCl (0.2 M), EDTA (10 mM), SDS (0.2%) and NaCl (0.2 M) in nuclease-free water with proteinase K (10 μg/ml). Genomic DNA was purified using a standard method with phenol-chloroform and ethanol. PrimeSTAR Max DNA polymerase (Takara Bio) was used for genotyping PCR, according to the manufacturer's instructions. PCR was performed as follows: 30 s at 94˚C; 35 cycles of 10 s at 98˚C and 10 min at 68˚C; then 10 min at 68˚C; and a final incubation at 4˚C. Following PCR reaction, gel electrophoresis (100 V, 30 min) was performed using 1% agarose gels and an 1 Kb Plus DNA Ladder (Thermo Fisher) as the DNA size marker. Primer sequences for genotyping PCR used in the current study are as follows: hSOX2-forward AAGATGCACAACTCGGAGATCAG, PGK-reverse TGCTGTCCATCTGCACGAGACTAGTGAG, pA-forward CGTCCGAGGGCAAAGGAATAGTCGAGA, hSOX2-reverse GGCTACAGGCTTGGACCCGC, hPROX1-forward TGGGCAGTTGCTCCGCTTCA, hPROX1-reverse CCCAGTGGCACACTGACACGA, cjBLIMP1-forward AGCACCATTCTGAAGACACGAAG, cjBLIMP1-reverse GGCAATACCCAAACCTCATTTTC, cjSTELLA-forward GCCACCTCTACCAATCAACATCA, cjSTELLA-reverse AGGATGCTTCCACAAAAACACAA.

DNA sequencing

DNA sequencing analysis was performed using the BigDye Terminator v1.1 cycle sequencing kit (Thermo Fisher) with the 3130xl Genetic Analyzer (Applied Biosystems). The sequencing data presented in the figures were illustrated using the Snap Gene software (GSL Biotech). Primer sequences for DNA sequencing are as follows: cjBLIMP1-seq GAGGACGTGGAGGATGAC, cjSTELLA-seq TGTGTTCCCTGTTCCAACAT.

Results

Construction of KI vectors for the introduction or correction of mutations

Using the Multisite Gateway technology, we first established a method for the construction of vectors aimed at introducing or correcting mutations in specific loci. For this purpose, we built upon our previous methods for constructing protein expression vectors using the Multisite Gateway technology [8] to establish a new method for the generation of KI vectors harboring a bipotential selection marker cassette PGK-Puro-ΔTK (*PuroTK)* flanked by two *lox*P sequences and homology arms. This method was named the GKI-1.0 method (Multisite Gateway-mediated KI vector construction-version 1.0). Although we have previously reported successful KI using vectors generated by GKI-1.0 [12, 13, 16, 18], we have never described the

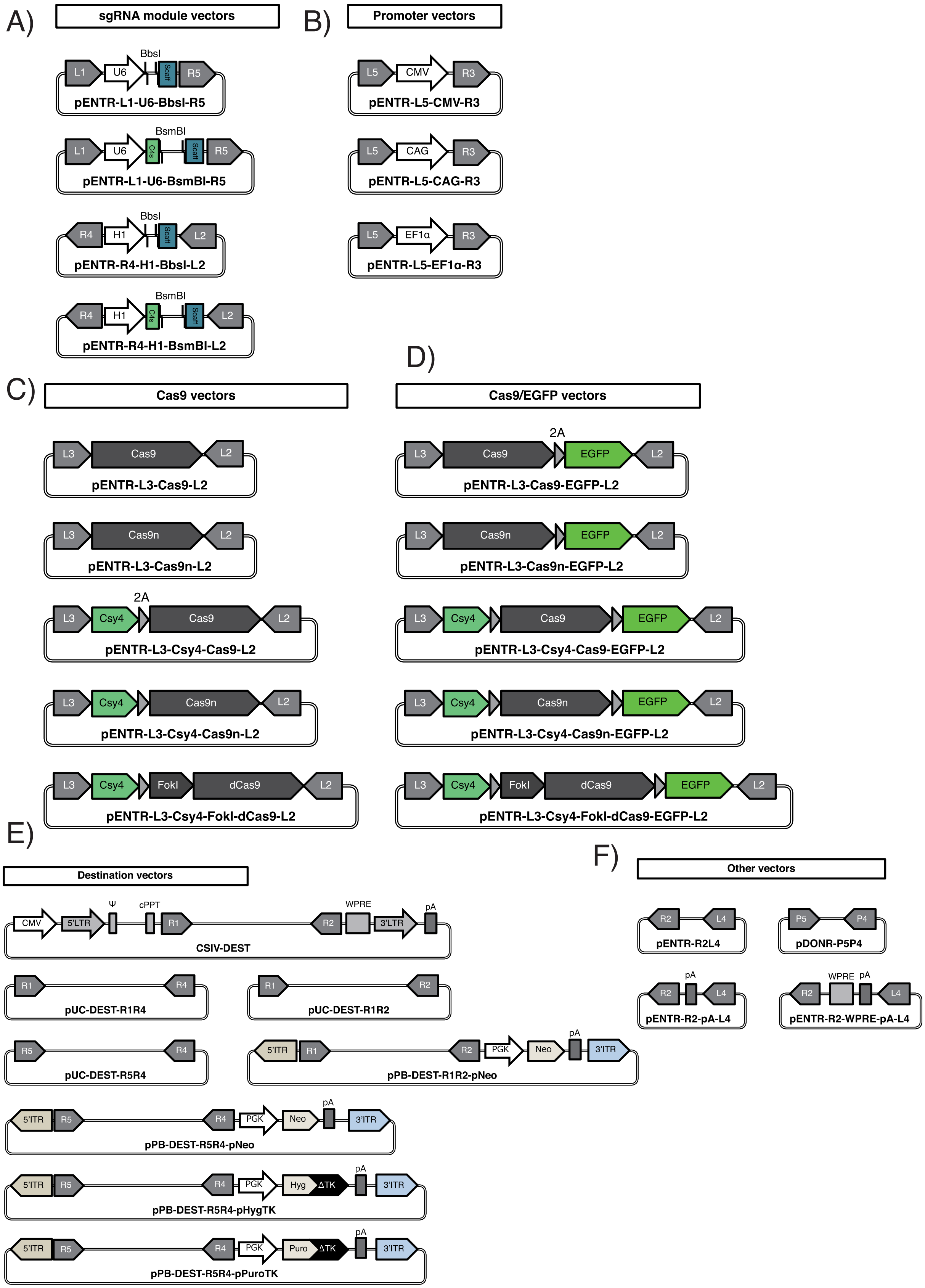
A)
sgRNA module vectors
L1
U6
BbsI
Scaff
R5
pENTR-L1-U6-BbsI-R5
BsmBI
C4s
pENTR-L1-U6-BsmBI-R5
R4
H1
L2
pENTR-R4-H1-BbsI-L2
pENTR-R4-H1-BsmBI-L2
B)
Promoter vectors
L5
CMV
R3
pENTR-L5-CMV-R3
CAG
pENTR-L5-CAG-R3
EF1α
pENTR-L5-EF1α-R3
C)
Cas9 vectors
L3
Cas9
pENTR-L3-Cas9-L2
Cas9n
pENTR-L3-Cas9n-L2
2A
Csy4
pENTR-L3-Csy4-Cas9-L2
pENTR-L3-Csy4-Cas9n-L2
FokI
dCas9
pENTR-L3-Csy4-FokI-dCas9-L2
D)
Cas9/EGFP vectors
EGFP
pENTR-L3-Cas9-EGFP-L2
pENTR-L3-Cas9n-EGFP-L2
pENTR-L3-Csy4-Cas9-EGFP-L2
pENTR-L3-Csy4-Cas9n-EGFP-L2
pENTR-L3-Csy4-FokI-dCas9-EGFP-L2
E)
Destination vectors
CMV
5'LTR
Ψ
cPPT
R1
R2
WPRE
3'LTR
pA
CSIV-DEST
pUC-DEST-R1R4
pUC-DEST-R1R2
pUC-DEST-R5R4
5'ITR
PGK
Neo
3'ITR
pPB-DEST-R1R2-pNeo
pPB-DEST-R5R4-pNeo
Hyg
ΔTK
pPB-DEST-R5R4-pHygTK
Puro
pPB-DEST-R5R4-pPuroTK
F)
Other vectors
L4
pENTR-R2L4
P5
P4
pDONR-P5P4
pENTR-R2-pA-L4
pENTR-R2-WPRE-pA-L4

Fig 9. Multisite Gateway-mediated Cas9/gRNA expression system (GCas method). (A) sgRNA module vectors harboring the *att* sites. U6, mammalian U6 type III RNA polymerase III promoter; H1, H1 RNA polymerase III promoter; BbsI, a single sgRNA cloning site mediated by *Bbs*I digestion; BsmBI, a dual sgRNA cloning site mediated by *BsmB*I digestion; Scaff, scaffold sequence of sgRNA followed by TTTTTT (poly-T terminal signal of RNA polymerase III); C4s, a Csy4-mediated RNA cleavage site. See Figure E in S1 Supporting Information and the corresponding legend for the cloning of sgRNA(s) into the sgRNA module vectors. (B) Promoter vectors harboring *att*L5 and *att*R3. CMV, cytomegalovirus promoter; CAG, cytomegalovirus enhancer fused to the chicken *beta-actin* promoter and introns of chicken *beta-actin* and rabbit *beta-globin*; EF1α, *Elongation factor 1-alpha* promoter. (C–D) Vectors harboring Cas9 alone or Cas9/EGFP harboring *att*L3 and *att*L2. 2As are illustrated as right-pointing triangles. Cas9, *spCas9*; 2A, *Thosea asigna* virus 2A peptide sequence (*T2A*); Csy4, *Csy4* RNase of *Pseudomonas aeruginosa*; FokI, dimerization-dependent wild-type *Fok*I nuclease-domain; Cas9n, mutated *Cas9* nickase (D10A); dCas9, catalytically inactive *Cas9*. (E) Destination vectors for the GCas method. The *CmR* and *ccdB* gene are located between *att*R1 and *att*R2 in each destination vector (not described) for the positive/negative selection of transformants harboring the vector itself or the desired Cas9/gRNA expression vector. Gray arrows and boxes in CSIV-DEST indicate regulatory sequences for lentivirus production such as 5'LTR (5'-side long terminal repeat), Ψ (Packaging signal), cPPT (central polypurine tract), WPRE (woodchuck hepatitis virus posttranscriptional regulatory element) and 3'LTR (3'-side long terminal repeat). 5'ITR and 3'ITR, 5'-side and 3'-side of *Piggybac* inverted terminal repeats. The entire sequences of the vectors described are attached in the S1 sequence. (F) Other module vectors for the GCas method. pA, simian virus 40 polyadenylation signal sequence.

https://doi.org/10.1371/journal.pone.0221164.g009

detailed methodology and the multiple types of module vectors intended for the method, which we will do below.

The schematic for the method is shown in Fig 1A. To introduce or correct mutations in a gene of interest (GOI), SNs are targeted to induce double strand break (DSB) at the vicinity of the mutation site. In addition, to ensure that proper gene expression and splicing are maintained, the selection cassette is excised following homologous recombination, which results in one residual *lox*P sequence (34 bp) in the genome (Fig 1A, bottom).

As shown in Fig 1B and 1C, the GKI-1.0 method consists of two steps: (1) Arm cloning and (2) LR reaction. In the 1st step, 0.5–1 kb homology arms that are either directly upstream or downstream of the DSB site (the 5' arm and 3' arm, respectively) are amplified from the genome by PCR, and inserted into the pDONR vectors (pDONR-P3P1r and pDONR-P2rP4) by BP reaction (Fig 1B). For the PCR process, the following sequences are included at the 5' terminus of the primers to add the respective sequences to the corresponding arms (see Materials and methods): *att*B3 and *att*B1 with *lox*P for the 5' arm, and *att*B2 with *lox*P and *att*B4 for the 3' arm (Fig 1B, top). Following the BP reaction, as a result, the 5'arm vector harbors *att*L3 and *att*R1, and the 3' arm vector harbors *att*R2 and *att*L4 (Fig 1B, bottom).

Prior to the 2nd step, the 5' arm or 3'arm vector may need to undergo mutagenesis PCR depending on the desired construct. In the 2nd step, the 5' arm and 3'arm vectors, one of which harbors the mutation, a pENTR vector harboring the *PuroTK* cassette (pENTR-L1-PGK-PuroTK-L2), and a destination vector (pUC-DEST-R3R4(R)) undergo LR reaction (Fig 1C), after which the desired KI vector is finally obtained.

Furthermore, to enhance the utility of the method for further KI experiments, we generated multiple module vectors for use at the 2nd step of GKI-1.0 (Fig 2A and 2B). Using these module vectors, multiple different types of positive or negative selection markers can now be incorporated into the KI vectors, such as *Hygromycin resistance gene (Hyg)* fused to *ΔTK* (*HygTK*), *Neomycin resistance gene* (*Neo*), *Neo* fused to *ΔTK* (*TKNeo*) and *Blasticidin resistance gene (Bsd)* fused to *ΔTK* (*TKBsd*), which were previously validated for the selection of mammalian cells [22–24] (Fig 2A). In addition, we constructed destination vectors harboring negative selection markers for mammalian cells, such as *diphtheria toxin A (DT-A)* and Herpes Simplex Virus-1 *Thymidine Kinase (TK)*, which can be helpful for concentrating the homologous recombinants (Fig 2B). Furthermore, we constructed a destination vector whose backbone harbors expression cassettes of *Streptococcus pyogenes* Cas9 (spCas9) with tdTomato and a single-guide RNA (sgRNA) cloning site under the U6 promoter, for the construction of an all-in-one KI vector (Fig 2B, bottom). In Figs 1, 2 and 4, pUC-L1-PGK-PuroTK-L2 and pUC-DEST-R3R4(R) are just drawn as a representative selection cassette vector and a representative destination vector. They could be replaced with any other selection cassette vector and destination vector shown in Fig 2A and 2B.

Minimization of residual sequences following KI and cassette excision

Although the generation of KI vectors using GKI-1.0 with multiple module vectors may facilitate KI experiments, the residual *lox*P sequence (34 bp length) in the genome may be problematic in some cases. For the introduction or correction of mutations in a GOI without leaving the *lox*P sequence in the genome, we slightly modified the method by replacing the *lox*P system with *piggyBac* transposon-based cassette excision system. The modified method was named GKI-2.0 (Fig 3A–3C). In the *piggyBac* system, the *piggyBac* transposase recognizes the transposon-specific inverted terminal repeats (ITRs) that are inserted on both ends of the target contents, which allows the transposase to move the contents from the TTAA site from its original position to different TTAA sites on other chromosomes or vectors [26]. As a result, no residual sequences remain and the TTAA sequence is left intact in the original position. Additionally, in an experimental setting, exploiting an excision-competent, integration-deficient *piggyBac* transposase (PBx) [27] enables the complete excision of the selection marker cassettes from the genome following homologous recombination, leaving only the short TTAA sequence behind in the genome (Fig 3A). Again, although we have previously reported successful KI using vectors generated by GKI-2.0 [19, 20], we have not described the detailed methodology, which we do below.

In a similar manner to GKI-1.0, the GKI-2.0 method consists of two steps: (1) Arm cloning and (2) LR reaction (Fig 3B and 3C). In the 1st step, 0.5−1 kb homology arms directly upstream or downstream of the DSB site (the 5'arm and 3'arm, respectively) are amplified from the genome by PCR, and inserted into *Hpa*I-digested pENTR2 vectors (pENTR2-L3-PBL-R1 and pENTR2-R2-PBR-L4) by seamless cloning (Fig 3B). Primer sequences for seamless cloning are described in Materials and Methods.

Prior to the 2nd step, the 5' arm or 3' arm vector may need to undergo mutagenesis PCR, depending on the desired construct. In the 2nd step, the 5'arm and 3'arm vectors, one of which harbors the mutation, a pENTR vector harboring the *PuroTK* cassette (pENTR-L1-PGK-PuroTK-L2), and a destination vector (pUC-DEST-R3R4(R)) undergo LR reaction (Fig 3C), after which the desired KI vector is obtained.

Construction of KI vectors for the integration of fluorescent/luminescent protein genes

Precise integration of fluorescent/luminescent protein genes in specific loci enables adequate and faithful marking of specific cells. When exploiting the bi-cistronic expression system of the self-cleaving 2A peptide or the internal ribosomal entry site (*IRES*) [28], it is possible that the integration of reporter genes does not disrupt the expression of the targeted endogenous genes. Therefore, we established a new method based on GKI-2.0 to generate KI vectors for the precise integration of reporter genes using the 2A peptide sequence, with minimal effects on targeted endogenous genes. We named this method GKI-3.0.

To integrate reporter genes into a GOI locus without disrupting the expression of the gene, SNs are targeted to induce DSB at the vicinity of the termination codon of the gene (Fig 4A). Additionally, reporter genes are fused to the last amino acid-coding codon of the GOI via a 2A peptide sequence. Selection cassettes for concentrating homologous recombinants are excised by the *piggyBac* system following homologous recombination, similarly as in GKI-2.0.

The GKI-3.0 method consists of two steps: (1) Arm cloning and (2) LR reaction (Fig 4B and 4C). In the 1st step, a 0.5−1 kb 5'-side homology arm which is directly upstream of the termination codon of the GOI is amplified from the genome by PCR, and inserted into each *Sfo*I-digested pENTR2-L3-SfoI-PBL-R1 reporter vector by seamless cloning (Fig 4B, left). As shown in Fig 5, we constructed multiple reporter vectors containing fluorescent/luminescent reporter genes intended for use with the GKI-3.0 method, such as *Enhanced green fluorescent protein* (*EGFP*), variant of *Yellow fluorescent protein Venus*, *Cyan fluorescent protein Cerulean*, tandem dimer of *dTomato* (*tdTomato*), *monomeric Cherry fluorescent protein* (*mCherry*), *near-infrared fluorescent protein iRFP713* and *NanoLuc* fused to a secretion signal (*secNanoLuc*) [29].

On the other hand, a 0.5–1 kb 3' side homology arm, which contains the termination codon of the GOI and the direct downstream sequence of the termination codon, is amplified from the genome by PCR, and inserted into an *Hpa*I-digested pENTR2-R2-PBR-L4 (Fig 4B, right), similarly as in GKI-2.0 (Fig 3B, right).

In the 2nd step, the 5' arm vector harboring the reporter genes and the 3' arm vector, a pENTR vector-harboring a selection cassette, and a destination vector that undergoes LR reaction (Fig 4C), after which the desired KI vector is obtained.

Application of the KI vectors generated with GKI-3.0 in human PSCs

To demonstrate the application of the GKI-3.0 method for stem cell research, we constructed two KI vectors, a SOX2-EGFP vector and a PROX1-tdTomato vector, with this method for transfection into human PSCs. In addition, we constructed Cas9-gRNA vectors (PX459 backbone [30]) containing guide RNA sequences targeting the vicinity of the termination codon of each target gene.

Sex determining region Y-box 2 (SOX2) is an indispensable factor for the maintenance of pluripotency and neuroectodermal differentiation from PSCs [31]. A previous study has shown that the depletion of *SOX2* expression in human PSCs leads to altered stem cell marker expression and trophectodermal differentiation [32]. *SOX2* is thus considered to be an important indicator for the pluripotency and differentiation states of PSCs, so we decided that the *SOX2* gene is a suitable target for the application of GKI-3.0.

After transfecting the hSOX2-EGFP KI vector with a corresponding Cas9-gRNA vector into 201B7 hiPSCs [11], we were able to obtain two heterozygous-KI clones (#1 and #2) following puromycin selection. KI was confirmed by genotyping PCR analysis (Fig 6A and 6B). Additionally, transfection of an expression vector of *PBx* into the #2 hiPSCs and ganciclovir selection resulted in the excision of the *PuroTK* cassette in all of the five analyzed clones (#2, 3, 4, 5 and 7; Fig 6C and 6D). Representative images of the #7 hiPSC clone (SG7) harboring the $SOX2^{+/EGFP}$ alleles are shown in Fig 6E, showing a uniform intensity of EGFP fluorescence within the hiPSC colony.

We then attempted to perform an additional KI in the SG7 hiPSCs. *Prospero homeobox protein 1* (*PROX1*) has a crucial role in the development of lymphatic vessels and the dentate gyrus (DG) of the hippocampus [33, 34]. Hippocampal neurogenesis in the DG region, which gives rise to PROX1-positive excitatory granule neurons from SOX2-positive neural stem cells constantly throughout a lifetime [35], has an important role in learning and memory formation, and is known to be affected in neurodegenerative diseases such as Alzheimer's disease [36, 37]. In this context, hiPSCs harboring the double reporters of *SOX2* and *PROX1* could be used for the study of the hippocampal neurogenesis.

After transfecting the hPROX1-tdTomato KI vector with a corresponding Cas9-gRNA vector, we were able to obtain four heterozygous-KI clones (#2, 3, 6 and 7) and one homozygous-KI clone (#8) following puromycin selection. KI was confirmed by genotyping PCR analysis (Fig 7A and 7B). In addition, transfecting the PBx vector into the #8 hiPSCs and ganciclovir selection resulted in the excision of the *PuroTK* cassette in all of the eleven analyzed clones (Fig 7C).

Using one clone obtained following ganciclovir selection (#2 in Fig 7C, named SG7-PT8-Δ2), we differentiated the hiPSCs into hippocampal cells using the SFEBq method [17]. After differentiation for five days, the aggregated EBs showed strong, uniform EGFP fluorescence, while the tdTomato fluorescence could not be detected (Fig 7D, top). We were able to isolate EGFP-positive cells by FACS, which revealed that 27.3 ± 3.0% cells were EGFP-positive (Fig 7E, left), while there were few (less than 0.1%) tdTomato-positive cells. On day 13, strong EGFP fluorescence became restricted to the protrusive region of the EBs, while tdTomato-positive cells began to emerge at the surface of the protrusion (Fig 7D, bottom). On day 18, we quantified EGFP- and tdTomato-positive cells by FACS, which revealed that 3.0 ± 1.4% or 3.4 ± 0.8% cells were EGFP or tdTomato-positive (Fig 7E, right), respectively, with less than 0.1% EGFP/tdTomato-double-positive cells (Fig 7E, right). Our data reveal that during the initial days of differentiation, the percentage of tdTomato-positive cells increased, while that of EGFP-positive cells decreased, which suggests that the cells were indeed differentiating from its pluripotent/neuroectodermal state into hippocampal neurons.

Next, the EBs were collected, and transferred to a suspension culture for further differentiation, followed by a 5-day adherent culture. On day 35, we performed immunochemical analysis of the plated EBs, and observed that tdTomato-positive cells expressed a pan-neuronal marker βIII-tubulin and an excitatory neuronal marker VGLUT2 (Fig 7F), confirming that the cells are hippocampal neurons. Taken together, these results show that the vectors describe here can be used for tracking the differentiation of PSCs into hippocampal neurons, indicating the utility of this double-reporter system using the GKI-3.0 method in human PSCs.

Application of KI vectors generated with GKI-3.0 in marmoset PSCs

In addition to the human PSC system, we performed KI in marmoset PSCs using vectors generated with GKI-3.0. *B-lymphocyte-induced maturation protein-1* (*BLIMP1*) is a critical factor for the specification of primordial germ cells (PGCs) in rodents and primates [38, 39]. In mice, derivation of BLIMP1/STELLA double-positive cells (termed PGC-like cells) from the PSCs resulted in the induction of functional gametes [40–42]. However, the function and expression pattern of *STELLA* in primates is yet to be elucidated [43]. Therefore, monitoring *BLIMP1* and *STELLA* expression is important for the verification of successful induction of germ cells. For this purpose, we constructed two KI vectors harboring BLIMP1-Venus (BV) and STELLA-Cerulean (SC).

After transfecting the BV-KI vector with the corresponding Cas9-gRNA vector into CMES40 cjESCs [14], we were able to obtain two homozygous-KI clones (#2 and 3; BV2 and BV3, respectively) among the three clones analyzed, following hygromycin selection (Fig 8A and 8B). DNA sequencing analysis revealed precise homozygous KI (Figure A1 in S1 Supporting Information).

We further performed an additional KI in the BV2 cjESCs with the SC-KI vector and the corresponding Cas9-gRNA vector. Following transfection and puromycin selection, we were able to obtain one homozygous-KI clone (#6, named BV2-SC6) among the six clones analyzed (Fig 8D and 8E). DNA sequencing analysis revealed precise homozygous KI (Figure B1 in S1 Supporting Information). We were also able to successfully generate a SC-KI cjESC clone named SC1 after transfecting the SC-KI vector into wild-type cjESCs (Figure A2 in S1 Supporting Information).

However, the efficiency of the excision of selection cassettes by PBx expression was low in the cjESC system. Upon excising selection cassettes from the SC1 cjESCs, four PBx transfections and ganciclovir selections were required to obtain cjESCs harboring homozygous *STELLA*^{Cerulean} alleles (SC1-Δ3, 5, 6 and 8; Figure B2 in S1 Supporting Information). Moreover, we did not obtain cjESCs harboring homozygous *BLIMP1*^{Venus} alleles from the BV2 and BV2-SC6 cjESCs within the four PBx transfections and ganciclovir selections (Figure C in S1 Supporting Information).

In order to obtain cassette-excised cjESCs more efficiently, we decided to use an expression vector of hyperactive integration-competent *piggyBac* transposase (*HyPBase*) [44] instead of *PBx*, along with an additional *piggyBac*-based vector to enable positive selection of the transposase-expressing cells. For this additional vector, we used pPB-hVASA-tdTomato-pNeo (VT), which contains a coding sequence of *tdTomato* under the control of the 3-kb human *VASA* promoter and a selection cassette of neomycin flanked by *piggyBac* ITRs (Figure A4 in S1 Supporting Information). Since *VASA* expression significantly increases from the late migrating stage of PGCs following specification and colonization of PGCs [45], the *VASA* reporter construct enables the monitoring of the maturation of PGCs following specification.

After one-time transfection of the VT reporter vector into the BV2-SC6 cjESCs, we isolated five colonies showing a dim tdTomato fluorescence following ganciclovir and G418 selection, since VASA is normally very weakly expressed in primate PSCs [46]. We demonstrated efficient excision of the selection cassettes in both the *BLIMP1* and *STELLA* loci in these isolated clones (2/5 and 3/5 analyzed clones; Fig 8C and 8F). As a result, we were able to obtain a cjESC line harboring homozygous *BLIMP1*^{Venus} and *STELLA*^{Cerulean} alleles, along with the VT transgene. This line was named BV2-SC6-VT2 (#2 in Fig 8C and 8F). The cjESCs harboring the *STELLA*^{Cerulean} alleles showed a weak fluorescence of Cerulean (Fig 8G and Figure C2 in S1 Supporting Information), although we failed to detect the Venus fluorescence from *BLIMP1*^{Venus} in the same cells (Fig 8G). Immunochemical analysis of the BV2-SC6-VT2 cjESCs using pluripotency markers (NANOG, SSEA4 and TRA-1-60) showed that the cells retained an undifferentiated state even after the integration of the triple (BV, SC and VT) reporters (Figure B4 in S1 Supporting Information), as well as a differentiation potential into all the three germ layers following EB formation (Figure C4 in S1 Supporting Information). Thus, the dual (positive and negative) selection strategy enables efficient cassette excision in marmoset PSCs, which facilitates the utility of the GKI-3.0 method in the system.

Multisite Gateway-mediated Cas9/gRNA expression system

Up to this point, we have utilized the PX459-based Cas9/gRNA expression vectors [30] for validating the GKI-3.0 method in human and marmoset PSCs. However, an all-in-one Cas9/gRNA expression vector system with multiple modules would be more useful for KI experiments in mammalian cells. Therefore, we established a new Multisite Gateway-based construction method named GCas, and constructed novel vectors for this method (Fig 9). In this method, one or two sgRNAs are incorporated into the corresponding cloning sites (*Bbs*I for single sgRNA [30] and *Bsm*BI for dual sgRNAs [47]) of the sgRNA module vectors, *att*L1–*att*R5 or *att*R4–*att*L2, which are under the control of either the U6 or H1 promoter, respectively (Fig 9A and Figure E in S1 Supporting Information). Also, in addition to the *att*L3-*att*L2 module vectors containing the wildtype *spCas9* nuclease, we constructed module vectors carrying the nickase mutant (*Cas9n*), a catalytically dead mutant fused to the *Fok*I endonuclease (*FokI-dCas9*) [30, 47] (Fig 9C and 9D), and high-fidelity *spCas9* nucleases such as SpCas9-HF1, SpCas9-HF2, SpCas9-HF4, eSpCas9(1.1) and eSpCas9n(1.1) [48, 49] (Table A in S1 Supporting Information).

To construct a standard Cas9/gRNA expression vector, each *att*L1–*att*R5 sgRNA module vector, *att*L5–*att*R3 promoter vector (Fig 9B), a polyA vector pENTR-R2-pA-L4 (Fig 9F) and a destination vector pUC-DEST-R1R4 (Fig 9E), along with either *att*L3−*att*L2 Cas9 alone (Fig 9C) or Cas9/EGFP vector (Fig 9D), undergo LR reaction, after which the desired Cas9/gRNA vector is obtained. See Figure F in S1 Supporting Information for further applications of the GCas method.

Discussion

In the present study, we introduced new methods, GKI-1.0, 2.0 and 3.0, for generating KI vectors, which are all based on the Multisite Gateway technology. Although SNs greatly improve the KI efficiency in various cells and animals [50], it is still necessary to perform positive selection for concentrating cells in which precise-KI occurred. In this context, since PSCs harbor a prominent potential of unlimited proliferation and differentiation into a plenty of cell types including those from all the three germ layers, it is important to validate the KI vectors generated by the GKI methods in PSCs. Thus, since the application of the KI vectors has been validated in human and marmoset PSCs in the present and previous studies [12, 13, 16, 18–20], we conclude that the methods can be robustly used in cell research.

Two previous studies, which described a KI vector construction method using the Multisite Gateway technology for murine ESCs or a human cell line [6, 10], exploited the Cre-*lox*P system for the excision of the positive selection cassette, which ends up with a 34 bp *lox*P sequence left behind in the genome. However, in our GKI-2.0 and 3.0 methods, we used an alternative approach of cassette excision using the *piggyBac* transposase system [26], which resulted in only a 4 bp (TTAA) residual sequence in the genome following excision. In this approach, when temporary integration of the selection cassette is targeted to a native TTAA sequence in the genome, it enables “footprint-free” KI following cassette excision. Since eukaryotic introns that have not been extensively characterized may in fact harbor functional regions [51, 52], our approach is ideal for reverse genetics.

Additionally, our module vectors for the GKI methods (Fig 2), in particular, the variety of positive/negative selection cassettes in hand (Fig 2A), would be useful for generating PSCs harboring homozygous or multiple KI alleles.

Mammalian adult neurogenesis continues throughout life only in two brain regions, such as DG in the hippocampus and subventricular zone in the cerebral cortex [53]. The spatial restriction of the phenomenon is due to the restricted stem cell niche in the adult brain, which is supported by unique microenvironments of the regions or unique cellular properties of the stem cells located in the regions [37]. Our double reporter system of *SOX2* and *PROX1* in human PSCs, which we generated and validated in the current study, would facilitate the *in vitro* study of this interesting phenomenon.

Although functional gametes have previously been derived from murine PSCs by transplantation into gonads or an *in vitro* long-term culture in reconstituted ovaries [40–42, 54], this has never been achieved in any other mammalian species. Two recent studies reported the successful induction of PGC-like cells from human PSCs using faithful KI reporters of PGC-specific genes such as *BLIMP1*, *TFAP2C* or *NANOS3* [39, 46]. However, because of the ethical difficulties, the human PSC system cannot be used for the functional evaluation of matured post-meiotic gametes, which may be obtained following the emergence of VASA-positive pre-meiotic gonocytes [55, 56]. Even in the murine systems [40–42, 54], embryonic gonadal niches or embryonic gonadal somatic cells are required for obtaining mature germ cells from PSCs, which is most likely accomplished through unidentified cellular or humoral signals [43]. In this context, because the use of human embryonic tissues is limited by ethical and technical issues, generating PSCs harboring faithful PGC reporters from non-human primate species is useful for studying primate gametogenesis. Furthermore, since marmosets have many advantages in reproduction research, including short gestation and sex maturation periods [57], the cjESCs harboring PGC-specific reporters may facilitate germ cell research.

Supporting information

S1 Supporting information. Including supporting methods, figures (Figs A-F), figure legends, a table, references.
(PDF)

S1 sequence. Entire vector sequences in the current study.
(DOCX)

Acknowledgments

We greatly thank Dr. Toshiaki Noce (RIKEN), Dr. Ryusuke Nakajima (RIKEN) and Dr. Hiroyuki Miyoshi (Keio University) for kindly providing DNA resources, Kent Imaizumi (Keio University) for assisting vector construction. We are indebted to Dr. Douglas Sipp (RIKEN) for critical reading of the manuscript. We also thank all the laboratory members of H.O. for their encouragement and generous supports.

This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas "Fluorescence Live imaging" of The Ministry of Education, Culture, Sports, Science, and Technology, Japan.

Author Contributions

Conceptualization: Sho Yoshimatsu, Takefumi Sone, Hideyuki Okano.

Data curation: Sho Yoshimatsu, Takefumi Sone, Mayutaka Nakajima, Tsukika Sato.

Formal analysis: Sho Yoshimatsu.

Funding acquisition: Takefumi Sone, Seiji Shiozawa, Hideyuki Okano.

Investigation: Sho Yoshimatsu, Takefumi Sone.

Methodology: Sho Yoshimatsu, Takefumi Sone, Ryotaro Okochi.

Project administration: Takefumi Sone, Hideyuki Okano.

Resources: Takefumi Sone, Tsukika Sato, Erika Sasaki, Hideyuki Okano.

Supervision: Mitsuru Ishikawa, Erika Sasaki, Seiji Shiozawa, Hideyuki Okano.

Validation: Sho Yoshimatsu, Mayutaka Nakajima, Tsukika Sato, Ryotaro Okochi, Mitsuru Ishikawa.

Visualization: Sho Yoshimatsu, Mayutaka Nakajima, Tsukika Sato.

Writing – original draft: Sho Yoshimatsu.

Writing – review & editing: Sho Yoshimatsu, Takefumi Sone, Mari Nakamura, Hideyuki Okano.

References

1. Thomas KR, Capecchi MR. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell. 1987; 51(3):503–12. Epub 1987/11/06. https://doi.org/10.1016/0092-8674(87)90646-5 PMID: 2822260

2. Zwaka TP, Thomson JA. Homologous recombination in human embryonic stem cells. Nat Biotechnol. 2003; 21(3):319–21. Epub 2003/02/11. https://doi.org/10.1038/nbt788 PMID: 12577066
3. Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKelver RC, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol. 2009; 27(9):851–7. Epub 2009/08/15. https://doi.org/10.1038/nbt.1562 PMID: 19680244
4. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, et al. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol. 2011; 29(8):731–4. Epub 2011/07/09. https://doi.org/10.1038/nbt.1927 PMID: 21738127
5. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014; 157(6):1262–78. Epub 2014/06/07. https://doi.org/10.1016/j.cell.2014.05.010 PMID: 24906146
6. Iiizumi S, Nomura Y, So S, Uegaki K, Aoki K, Shibahara K, et al. Simple one-week method to construct gene-targeting vectors: application to production of human knockout cell lines. Biotechniques. 2006; 41 (3):311–6. Epub 2006/09/23. https://doi.org/10.2144/000112233 PMID: 16989091
7. Hartley JL, Temple GF, Brasch MA. DNA cloning using in vitro site-specific recombination. Genome Res. 2000; 10(11):1788–95. Epub 2000/11/15. https://doi.org/10.1101/gr.143000 PMID: 11076863
8. Sone T, Imamoto F. Methods for constructing clones for protein expression in mammalian cells. Methods Mol Biol. 2012; 801:227–50. Epub 2011/10/12. https://doi.org/10.1007/978-1-61779-352-3_15 PMID: 21987257
9. Mano S, Nishihama R, Ishida S, Hikino K, Kondo M, Nishimura M, et al. Novel gateway binary vectors for rapid tripartite DNA assembly and promoter analysis with various reporters and tags in the liverwort Marchantia polymorpha. PLoS One. 2018; 13(10):e0204964. Epub 2018/10/05. https://doi.org/10.1371/journal.pone.0204964 PMID: 30286137
10. Nyabi O, Naessens M, Haigh K, Gembarska A, Goossens S, Maetens M, et al. Efficient mouse transgenesis using Gateway-compatible ROSA26 locus targeting vectors and F1 hybrid ES cells. Nucleic Acids Res. 2009; 37(7):e55. Epub 2009/03/13. https://doi.org/10.1093/nar/gkp112 PMID: 19279185
11. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 2007; 131(5):861–72. Epub 2007/11/24. https://doi.org/10.1016/j.cell.2007.11.019 PMID: 18035408
12. Ichiyanagi N, Fujimori K, Yano M, Ishihara-Fujisaki C, Sone T, Akiyama T, et al. Establishment of In Vitro FUS-Associated Familial Amyotrophic Lateral Sclerosis Model Using Human Induced Pluripotent Stem Cells. Stem Cell Reports. 2016; 6(4):496–510. Epub 2016/03/22. https://doi.org/10.1016/j.stemcr.2016.02.011 PMID: 26997647
13. Tanaka Y, Sone T, Higurashi N, Sakuma T, Suzuki S, Ishikawa M, et al. Generation of D1-1 TALEN isogenic control cell line from Dravet syndrome patient iPSCs using TALEN-mediated editing of the SCN1A gene. Stem Cell Res. 2018; 28:100–4. Epub 2018/02/18. https://doi.org/10.1016/j.scr.2018.01.036 PMID: 29453127
14. Sasaki E, Hanazawa K, Kurita R, Akatsuka A, Yoshizaki T, Ishii H, et al. Establishment of novel embryonic stem cell lines derived from the common marmoset (Callithrix jacchus). Stem Cells. 2005; 23 (9):1304–13. Epub 2005/08/20. https://doi.org/10.1634/stemcells.2004-0366 PMID: 16109758
15. Shiozawa S, Kawai K, Okada Y, Tomioka I, Maeda T, Kanda A, et al. Gene targeting and subsequent site-specific transgenesis at the beta-actin (ACTB) locus in common marmoset embryonic stem cells. Stem Cells Dev. 2011; 20(9):1587–99. Epub 2010/12/04. https://doi.org/10.1089/scd.2010.0351 PMID: 21126169
16. Yoshimatsu S, Okahara J, Sone T, Takeda Y, Nakamura M, Sasaki E, et al. Robust and efficient knock-in in embryonic stem cells and early-stage embryos of the common marmoset using the CRISPR-Cas9 system. Sci Rep. 2019; 9(1):1528. Epub 2019/02/08. https://doi.org/10.1038/s41598-018-37990-w PMID: 30728412
17. Sakaguchi H, Kadoshima T, Soen M, Narii N, Ishida Y, Ohgushi M, et al. Generation of functional hippocampal neurons from self-organizing human embryonic stem cell-derived dorsomedial telencephalic tissue. Nat Commun. 2015; 6:8896. Epub 2015/11/18. https://doi.org/10.1038/ncomms9896 PMID: 26573335
18. Hosoya M, Fujioka M, Sone T, Okamoto S, Akamatsu W, Ukai H, et al. Cochlear Cell Modeling Using Disease-Specific iPSCs Unveils a Degenerative Phenotype and Suggests Treatments for Congenital Progressive Hearing Loss. Cell Rep. 2017; 18(1):68–81. Epub 2017/01/05. https://doi.org/10.1016/j.celrep.2016.12.020 PMID: 28052261
19. Suda Y, Kuzumaki N, Sone T, Narita M, Tanaka K, Hamada Y, et al. Down-regulation of ghrelin receptors on dopaminergic neurons in the substantia nigra contributes to Parkinson's disease-like motor dysfunction. Mol Brain. 2018; 11(1):6. Epub 2018/02/21. https://doi.org/10.1186/s13041-018-0349-8 PMID: 29458391

20. Nakamoto FK, Okamoto S, Mitsui J, Sone T, Ishikawa M, Yamamoto Y, et al. The pathogenesis linked to coenzyme Q10 insufficiency in iPSC-derived neurons from patients with multiple-system atrophy. Sci Rep. 2018; 8(1):14215. Epub 2018/09/23. https://doi.org/10.1038/s41598-018-32573-1 PMID: 30242188

21. Naito Y, Hino K, Bono H, Ui-Tei K. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. Bioinformatics. 2015; 31(7):1120–3. Epub 2014/11/22. https://doi.org/10.1093/bioinformatics/btu743 PMID: 25414360

22. Seibler J, Schubeler D, Fiering S, Groudine M, Bode J. DNA cassette exchange in ES cells mediated by Flp recombinase: an efficient strategy for repeated modification of tagged loci by marker-free constructs. Biochemistry. 1998; 37(18):6229–34. Epub 1998/06/13. https://doi.org/10.1021/bi980288t PMID: 9572836

23. Karreman C. New positive/negative selectable markers for mammalian cells on the basis of Blasticidin deaminase-thymidine kinase fusions. Nucleic Acids Res. 1998; 26(10):2508–10. Epub 1998/06/10. https://doi.org/10.1093/nar/26.10.2508 PMID: 9580709

24. Schwartz F, Maeda N, Smithies O, Hickey R, Edelmann W, Skoultchi A, et al. A dominant positive and negative selectable gene for use in mammalian cells. Proc Natl Acad Sci U S A. 1991; 88(23):10416–20. Epub 1991/12/01. https://doi.org/10.1073/pnas.88.23.10416 PMID: 1720540

25. Chen YT, Bradley A. A new positive/negative selectable marker, puDeltatk, for use in embryonic stem cells. Genesis. 2000; 28(1):31–5. Epub 2000/10/06. PMID: 11020714

26. Zhao S, Jiang E, Chen S, Gu Y, Shangguan AJ, Lv T, et al. PiggyBac transposon vectors: the tools of the human gene encoding. Transl Lung Cancer Res. 2016; 5(1):120–5. Epub 2016/03/10. https://doi.org/10.3978/j.issn.2218-6751.2016.01.05 PMID: 26958506

27. Li X, Burnight ER, Cooney AL, Malani N, Brady T, Sander JD, et al. piggyBac transposase tools for genome engineering. Proc Natl Acad Sci U S A. 2013; 110(25):E2279–87. Epub 2013/06/01. https://doi.org/10.1073/pnas.1305987110 PMID: 23723351

28. He X, Tan C, Wang F, Wang Y, Zhou R, Cui D, et al. Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. Nucleic Acids Res. 2016; 44(9):e85. Epub 2016/02/07. https://doi.org/10.1093/nar/gkw064 PMID: 26850641

29. Heise K, Oppermann H, Meixensberger J, Gebhardt R, Gaunitz F. Dual luciferase assay for secreted luciferases based on Gaussia and NanoLuc. Assay Drug Dev Technol. 2013; 11(4):244–52. Epub 2013/05/18. https://doi.org/10.1089/adt.2013.509 PMID: 23679848

30. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013; 8(11):2281–308. Epub 2013/10/26. https://doi.org/10.1038/nprot.2013.143 PMID: 24157548

31. Zhang S, Cui W. Sox2, a key factor in the regulation of pluripotency and neural differentiation. World J Stem Cells. 2014; 6(3):305–11. Epub 2014/08/16. https://doi.org/10.4252/wjsc.v6.i3.305 PMID: 25126380

32. Fong H, Hohenstein KA, Donovan PJ. Regulation of self-renewal and pluripotency by Sox2 in human embryonic stem cells. Stem Cells. 2008; 26(8):1931–8. Epub 2008/04/05. https://doi.org/10.1634/stemcells.2007-1002 PMID: 18388306

33. Hong YK, Harvey N, Noh YH, Schacht V, Hirakawa S, Detmar M, et al. Prox1 is a master control gene in the program specifying lymphatic endothelial cell fate. Dev Dyn. 2002; 225(3):351–7. Epub 2002/11/02. https://doi.org/10.1002/dvdy.10163 PMID: 12412020

34. Iwano T, Masuda A, Kiyonari H, Enomoto H, Matsuzaki F. Prox1 postmitotically defines dentate gyrus cells by specifying granule cell identity over CA3 pyramidal cell fate in the hippocampus. Development. 2012; 139(16):3051–62. Epub 2012/07/14. https://doi.org/10.1242/dev.080002 PMID: 22791897

35. Suh H, Consiglio A, Ray J, Sawai T, D'Amour KA, Gage FH. In vivo fate analysis reveals the multipotent and self-renewal capacities of Sox2+ neural stem cells in the adult hippocampus. Cell Stem Cell. 2007; 1(5):515–28. Epub 2008/03/29. https://doi.org/10.1016/j.stem.2007.09.002 PMID: 18371391

36. Mu Y, Gage FH. Adult hippocampal neurogenesis and its role in Alzheimer's disease. Mol Neurodegener. 2011; 6:85. Epub 2011/12/24. https://doi.org/10.1186/1750-1326-6-85 PMID: 22192775

37. Goncalves JT, Schafer ST, Gage FH. Adult Neurogenesis in the Hippocampus: From Stem Cells to Behavior. Cell. 2016; 167(4):897–914. Epub 2016/11/05. https://doi.org/10.1016/j.cell.2016.10.021 PMID: 27814520

38. Ohinata Y, Payer B, O'Carroll D, Ancelin K, Ono Y, Sano M, et al. Blimp1 is a critical determinant of the germ cell lineage in mice. Nature. 2005; 436(7048):207–13. Epub 2005/06/07. https://doi.org/10.1038/nature03813 PMID: 15937476

39. Irie N, Weinberger L, Tang WW, Kobayashi T, Viukov S, Manor YS, et al. SOX17 is a critical specifier of human primordial germ cell fate. Cell. 2015; 160(1–2):253–68. Epub 2014/12/30. https://doi.org/10.1016/j.cell.2014.12.013 PMID: 25543152

40. Hayashi K, Ohta H, Kurimoto K, Aramaki S, Saitou M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell. 2011; 146(4):519–32. Epub 2011/08/09. https://doi.org/10.1016/j.cell.2011.06.052 PMID: 21820164

41. Hayashi K, Ogushi S, Kurimoto K, Shimamoto S, Ohta H, Saitou M. Offspring from oocytes derived from in vitro primordial germ cell-like cells in mice. Science. 2012; 338(6109):971–5. Epub 2012/10/09. https://doi.org/10.1126/science.1226889 PMID: 23042295

42. Hikabe O, Hamazaki N, Nagamatsu G, Obata Y, Hirao Y, Hamada N, et al. Reconstitution in vitro of the entire cycle of the mouse female germ line. Nature. 2016; 539(7628):299–303. Epub 2016/11/04. https://doi.org/10.1038/nature20104 PMID: 27750280

43. Saitou M, Miyauchi H. Gametogenesis from Pluripotent Stem Cells. Cell Stem Cell. 2016; 18(6):721–35. Epub 2016/06/04. https://doi.org/10.1016/j.stem.2016.05.001 PMID: 27257761

44. Yusa K, Zhou L, Li MA, Bradley A, Craig NL. A hyperactive piggyBac transposase for mammalian applications. Proc Natl Acad Sci U S A. 2011; 108(4):1531–6. Epub 2011/01/06. https://doi.org/10.1073/pnas.1008322108 PMID: 21205896

45. Noce T, Okamoto-Ito S, Tsunekawa N. Vasa homolog genes in mammalian germ cell development. Cell Struct Funct. 2001; 26(3):131–6. Epub 2001/09/22. PMID: 11565805

46. Sasaki K, Yokobayashi S, Nakamura T, Okamoto I, Yabuta Y, Kurimoto K, et al. Robust In Vitro Induction of Human Germ Cell Fate from Pluripotent Stem Cells. Cell Stem Cell. 2015; 17(2):178–94. Epub 2015/07/21. https://doi.org/10.1016/j.stem.2015.06.014 PMID: 26189426

47. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nature Biotechnology. 2014; 32(6):569–76. https://doi.org/10.1038/nbt.2908 PMID: 24770325

48. Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature. 2016; 529(7587):490–5. Epub 2016/01/07. https://doi.org/10.1038/nature16526 PMID: 26735016

49. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. Science. 2016; 351(6268):84–8. Epub 2015/12/03. https://doi.org/10.1126/science.aad5227 PMID: 26628643

50. Zhang F, Wen Y, Guo X. CRISPR/Cas9 for genome editing: progress, implications and challenges. Hum Mol Genet. 2014; 23(R1):R40–6. Epub 2014/03/22. https://doi.org/10.1093/hmg/ddu125 PMID: 24651067

51. Hube F, Francastel C. Mammalian introns: when the junk generates molecular diversity. Int J Mol Sci. 2015; 16(3):4429–52. Epub 2015/02/25. https://doi.org/10.3390/ijms16034429 PMID: 25710723

52. Chorev M, Joseph Bekker A, Goldberger J, Carmel L. Identification of introns harboring functional sequence elements through positional conservation. Sci Rep. 2017; 7(1):4201. Epub 2017/06/25. https://doi.org/10.1038/s41598-017-04476-0 PMID: 28646210

53. Ming GL, Song H. Adult neurogenesis in the mammalian brain: significant answers and significant questions. Neuron. 2011; 70(4):687–702. Epub 2011/05/26. https://doi.org/10.1016/j.neuron.2011.05.001 PMID: 21609825

54. Zhou Q, Wang M, Yuan Y, Wang X, Fu R, Wan H, et al. Complete Meiosis from Embryonic Stem Cell-Derived Germ Cells In Vitro. Cell Stem Cell. 2016; 18(3):330–40. Epub 2016/03/01. https://doi.org/10.1016/j.stem.2016.01.017 PMID: 26923202

55. Fang F, Angulo B, Xia N, Sukhwani M, Wang Z, Carey CC, et al. A PAX5-OCT4-PRDM1 developmental switch specifies human primordial germ cells. Nat Cell Biol. 2018; 20(6):655–65. Epub 2018/05/02. https://doi.org/10.1038/s41556-018-0094-3 PMID: 29713018

56. Yamashiro C, Sasaki K, Yabuta Y, Kojima Y, Nakamura T, Okamoto I, et al. Generation of human oogonia from induced pluripotent stem cells in vitro. Science. 2018; 362(6412):356–60. Epub 2018/09/22. https://doi.org/10.1126/science.aat1674 PMID: 30237246

57. Okano H, Hikishima K, Iriki A, Sasaki E. The common marmoset as a novel animal model system for biomedical and neuroscience research applications. Semin Fetal Neonatal Med. 2012; 17(6):336–40. Epub 2012/08/09. https://doi.org/10.1016/j.siny.2012.07.002 PMID: 22871417