Stem Cell Reports
Report

ISSCR

OPEN ACCESS

Universal Correction of Blood Coagulation Factor VIII in Patient-Derived Induced Pluripotent Stem Cells Using CRISPR/Cas9

Chul-Yong Park,[1,2] Jin Jea Sung,[1,3] Sung-Rae Cho,[4] Jongwan Kim,[5] and Dong-Wook Kim[1,2,3,*]
[1]Department of Physiology, Yonsei University College of Medicine, Seoul 03722, Korea
[2]Severance Biomedical Research Institute, Yonsei University College of Medicine, Seoul 03722, Korea
[3]Brain Korea 21 Plus Project for Medical Science, Yonsei University College of Medicine, Seoul 03722, Korea
[4]Department and Research Institute of Rehabilitation Medicine, Yonsei University College of Medicine, Seoul 03722, Korea
[5]S.Biomedics Co., Ltd, 28 Seongsui-ro, 26-gil, Seongdong-gu, Seoul 04797, Korea
*Correspondence: dwkim2@yuhs.ac
https://doi.org/10.1016/j.stemcr.2019.04.016

SUMMARY

Hemophilia A (HA) is caused by genetic mutations in the blood coagulation *factor VIII* (*FVIII*) gene. Genome-editing approaches can be used to target the mutated site itself in patient-derived induced pluripotent stem cells (iPSCs). However, these approaches can be hampered by difficulty in preparing thousands of editing platforms for each corresponding variant found in HA patients. Here, we report a universal approach to correct the various mutations in HA patient iPSCs by the targeted insertion of the *FVIII* gene into the human *H11* site via CRISPR/Cas9. We derived corrected clones from two types of patient iPSCs with frequencies of up to 64% and 66%, respectively, without detectable unwanted off-target mutations. Moreover, we demonstrated that endothelial cells differentiated from the corrected iPSCs successfully secreted functional protein. This strategy may provide a universal therapeutic method for correcting all genetic variants found in HA patients.

INTRODUCTION

Hemophilia A (HA) is one of the most common inherited bleeding disorders, with an incidence of 1 in 5,000 males worldwide (Berntorp and Shapiro, 2012). HA is caused by various genetic mutations (2,015 unique variants) within the X-linked coagulation *factor VIII* (*FVIII*) gene, including large deletions, insertions, inversions, and point mutations (FVIII variant database; www.factorviii-db.org/). At present, the intravenous infusion of recombinant FVIII protein is an available treatment option; however, this therapy is not curative and is associated with high costs, lifelong treatment, and the formation of FVIII-inactivating antibodies. Thus, the development of a fundamental method for treating HA is required.

Human induced pluripotent stem cells (iPSCs) are a versatile cell source for transplanting autologous cells with restored genes to compensate for mutated genes, for understanding cellular and molecular disease mechanisms and disease modeling (Cherry and Daley, 2013), and for therapeutic applications including drug discovery (Shi et al., 2017). Indeed, autologous retinal pigment epithelial cells differentiated from iPSCs were transplanted into a patient with neovascular age-related macular degeneration (Mandai et al., 2017). In addition, patient iPSCs were used for genome editing to correct genetic mutations (Li et al., 2015; Park et al., 2015a; Xu et al., 2017).

The type II clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) system is a versatile tool for genome editing (Jinek et al., 2012). Recently, nuclease-mediated genome editing was performed in HA patient iPSCs to correct the endogenous *FVIII* locus to the normal gene orientation without using ectopic donor DNA or with the targeted insertion of donor DNA harboring partial *FVIII* exons (Park et al., 2015b; Wu et al., 2016), strategies that are typically applied to inversion mutations. Interestingly, more than 60% of HA patients exhibit other genetic variations including point mutations, large deletions, insertions, or duplications (Graw et al., 2005), which suggests that the application of a universal correction strategy is required for all types of genetic variations that occur in HA. In addition, protocols to differentiate iPSCs into liver sinusoidal endothelial cells (LSECs), which are the primary producers of the FVIII protein, have also not been available to date (Park et al., 2016a). Thus, to overcome these limitations we have developed a universal correction approach by accessing a genomic safe harbor site and expressing functional FVIII without any restrictions in the type of variation or the cell type used for transplantation.

In this study, we chose the human *H11* locus as a safe harbor site and inserted the functional *FVIII* gene into this site using the CRISPR/Cas9 system in a targeted manner in both deleted- and inverted-patient iPSCs. Importantly, we found that the mRNA expression induced FVIII activity in the cultured supernatants. Moreover, we demonstrated that no off-target mutations were found in the corrected clones. To our knowledge, this is the first report to demonstrate a targeted *FVIII* insertion in a safe harbor locus that resulted in the phenotypic correction of the *FVIII* deficiency in HA patient iPSCs. This approach may provide

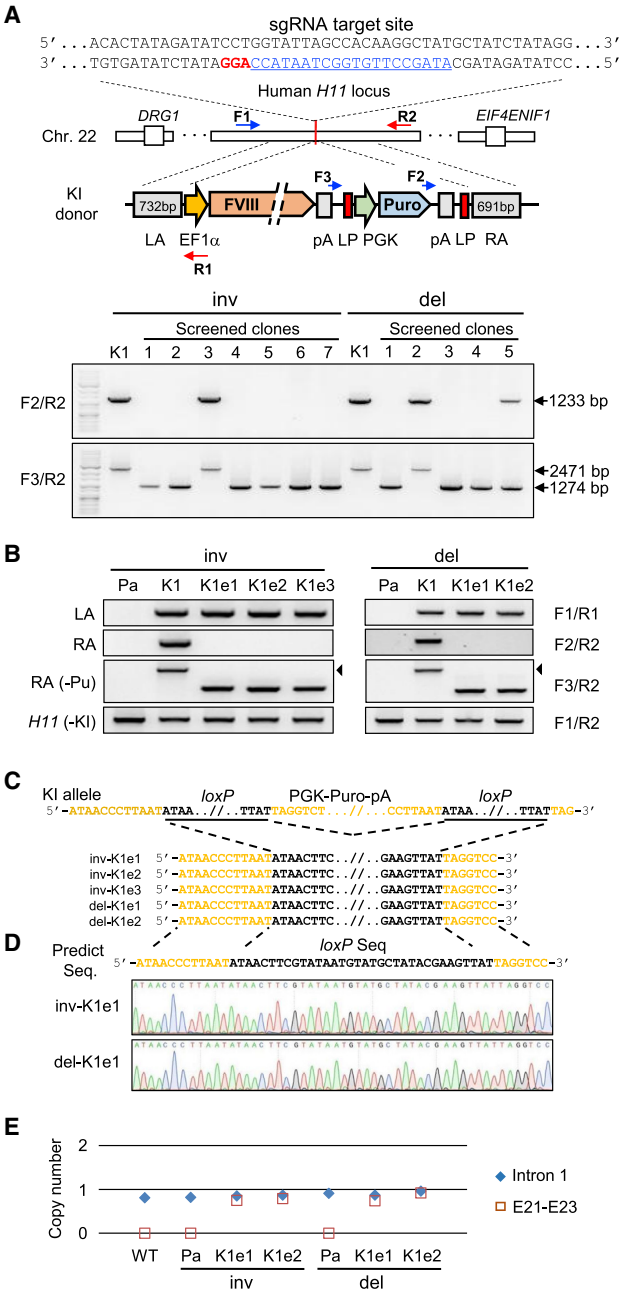

Figure 1. Targeted Insertions of the *FVIII* Gene into the Human *H11* Locus of HA Patient-Derived iPSCs

(A) Schematic overview depicting the *FVIII* gene knockin strategy into the *H11* locus located in human chromosome 22. The bases underlined in blue indicate the sgRNA target site. The protospacer adjacent motif sequence is shown in red. The five specific primers used for genotyping are shown. Lower panel shows PCR-based screening for excision of the selection cassette.

(B) PCR-based genotype analysis to confirm the removal of the puromycin expression cassette in the targeted iPSC clones. Black arrowheads indicate the DNA bands containing the puromycin cassette in the inv-K1 and del-K1 clones.

a universal correction method for application to all types of genetic variations found in HA patients.

RESULTS

Generation of *FVIII* Deleted Patient iPSCs

First, we derived *FVIII* deleted iPSC clones from adipose tissue-derived mesenchymal stem cells obtained from a patient diagnosed with several exon deletions (exons 8–22) using episomal vectors. We selected a total of nine embryonic stem cell-like colonies (termed Epi1 to Epi9), which were maintained onto a feeder layer followed by adaptation in feeder-free culture conditions (Figure S1A). After seven passages, we confirmed the absence of the episomal vectors in all the clones except one by PCR (Figure S1B). For the remaining experiments we chose the Epi6 line, which does not contain the *EBNA-1* sequence encoded in the vectors and expresses pluripotency markers including SSEA4, TRA-1-60, OCT4, NANOG (Figure S1C), *SOX2*, and *Lin28* (Figure S1D). The ability to differentiate *in vitro* was further confirmed in the Epi6 line (Figure S1E), which also exhibited a normal karyotype (Figure S1F).

Synthetic Single Guide RNA Design and Validation of Nuclease Activity

To insert *FVIII* into the *H11* locus using a targeted approach, we designed the synthetic single guide RNA (sgRNA) that recognized the target site using web-based *in silico* tools (crispr.mit.edu) (Figure 1A). We co-transfected either HEK-293T cells or each iPSC clone with both Cas9 and sgRNA vectors. To test the nuclease activity, we performed a T7 endonuclease I (T7E1) assay. The nuclease activity was relatively high, inducing small insertions and deletions (indels) mutations with a frequency of 15% at the *H11* locus based on the T7E1 assay results (Figure S2A). However, the nuclease activity in each iPSC clone was relatively low due to low transfection efficiency, inducing indels with a frequency of 4% and 4.3%, respectively (Figure S2A). In addition, deep sequencing analyses revealed that the frequency of indels at the target site was 27.9% (Figure S2B), with various indels found at the target

(C) Post-excision DNA sequence analysis between the two underlined *loxP* sites.

(D) Chromatograms showing the targeted excision of the puromycin cassette from the knockin allele.

(E) Droplet digital PCR (ddPCR) analysis to determine the copy number of the inserted fragment in each clone. The *RPP30* gene served as a reference. The primer sequences for the intron 1 locus (Intron 1) of the endogenous *FVIII* gene and for exon 21 to exon 23 (E21–E23) of the knockin donor template are described in the ddPCR analysis section of Experimental Procedures.

See also Figures S2 and S3; Tables S1 and S2.

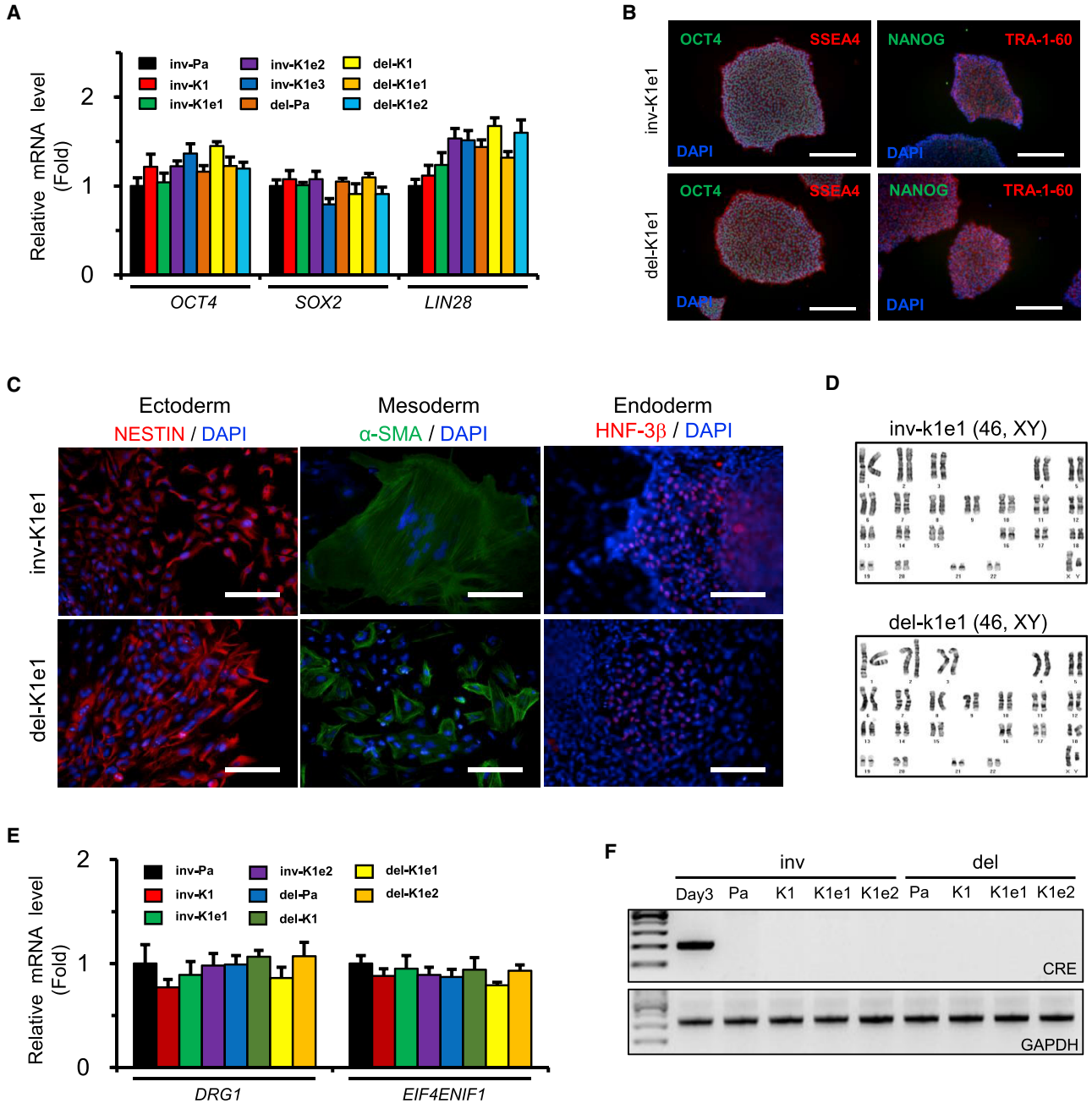

Figure 2. Analyses of Pluripotency from the Corrected iPSC Clones

(A) Expression of endogenous *OCT4*, *SOX2*, and *LIN28* in parental patient and corrected iPSC clones. The expression level of each gene was normalized to that of *GAPDH*. Data are means ± SEM of three independent experiments.

(B) Expression of the pluripotency markers OCT4, NANOG, SSEA-4, and TRA-1-60 detected by immunocytochemistry. The DAPI signal indicates the total cell content in the image. Scale bars, 200 μm.

(C) Expression of marker proteins representing the ectoderm (NESTIN), mesoderm (α-smooth muscle actin [α-SMA]), and endoderm (hepatocyte nuclear factor-3β [HNF-3β]). The DAPI signal indicates the total cell content in the image. Scale bars, 100 μm.

(D) Karyotype analyses were performed in the corrected iPSC clones.

(E) Expression of *DRG1* and *EIF4ENIF1* in parental and corrected patient iPSC clones. The expression level of each gene was normalized to that of *GAPDH*. Data are means ± SEM of three independent experiments.

(legend continued on next page)

site (Figure S2C). Moreover, no off-target mutations were detectably induced at eight homologous sites that differed from the on-target site by up to four nucleotides (Figure S2B).

Targeted Knockin of the *FVIII* Gene into the Human *H11* Locus

In parallel, we cloned the *FVIII* expression cassette under the control of the human elongation factor 1α (EF1α) promoter and a puromycin expression cassette into the backbone vector to construct donor DNA, followed by cloning flanked by both homology arms (Figures 1A and S3A). Thereafter, we electroporated Cas9 ribonucleoproteins (RNPs) and donor DNA into both *FVIII* inverted iPSCs (inv-Pa) and *FVIII* deleted iPSCs (del-Pa). Following additional culturing and puromycin selection, we screened drug-resistant colonies by PCR-based genotyping to find colonies harboring the targeted knockin using the specific primer sets listed in Table S1 (Figure S3A). Nine out of 14 colonies (64.3%, in case of inv-Pa) and 18 out of 27 colonies (66.7%) exhibited positive PCR bands for the inv-Pa and del-Pa knockin junctions on an agarose gel, respectively (Table S2). Following three passages, we derived two clones (inv-K1 and inv-K2) from inv-Pa iPSCs and three clones (del-K1 to del-K3) from del-Pa iPSCs, and demonstrated that all of the clones were single-allele knockin clones (Figure S3B). Targeted knockins from all the clones were further verified by Sanger sequencing of PCR amplicons (Figures S3C and S3D). Next, to excise the puromycin expression cassette from the clones, we chose the inv-K1 and del-K1 clones because they had no indels, even in the untargeted allele (−KI allele) (Figure S3C). After the transient expression of *Cre* recombinase in these clones, we screened seven colonies by PCR-based genotyping using the F3/R2 primers (Figure 1A). The successful excision of the selection cassette was confirmed by PCR-based genotyping using specific primers (Figure 1B and Table S1). Following the verifications of recombination between the two *loxP* sites by Sanger sequencing (Figures 1C and 1D), we derived the three puromycin-excised clones (inv-K1e1 to inv-K1e3) from the inv-K1 clone, and the two clones (del-K1e1 and del-K1e2) from the del-K1 clone (Figure 1B). Next, we confirmed that four knockin clones (inv-K1e1, inv-K1e2, del-K1e1, and del-K1e2) contained only one copy of the *FVIII* knockin fragment using droplet digital PCR analysis (Figure 1E). These results revealed targeted integration of the *FVIII* gene at the *H11* locus.

Analyses of Pluripotency and Off-Target Mutations in Genetically Corrected Clones

Next, we investigated whether the corrected clones maintained their pluripotent characteristics compared with the parental clones. Indeed, all the corrected clones actively transcribed pluripotency genes, including *OCT4*, *SOX2*, and *LIN28* (Figure 2A), and maintained similar levels of pluripotency marker proteins such as SSEA4, TRA-1-60, OCT4, and NANOG compared with their parental iPSC clones (Figure 2B). In addition, the corrected clones successfully differentiated into the three germ layers, as shown by positive immunostaining for NESTIN (ectoderm), α-smooth muscle actin (mesoderm), and hepatocyte nuclear factor-3β (endoderm) (Figure 2C). Furthermore, the clones exhibited a normal 46, XY karyotype by G-banding (Figure 2D). We then confirmed by qPCR analysis that no significant differences were found in the expression levels of *DRG1* and *EIF4ENIF1* genes between the parental and knockin iPSC clones (Figure 2E). No PCR bands corresponding to the *Cre* sequence were amplified from each knockin clone (Figure 2F). We then investigated whether off-target mutations were induced in the corrected clones by the nuclease used in this study. Ten potential off-target sites were examined by targeted deep sequencing in the four corrected clones and the two types of parental clones. We verified that no significant off-target mutations were induced at the sites listed in Table S3 (Figure S4). To further verify off-target mutations, we performed whole-genome sequencing analysis for three iPSC clones (del-Pa, del-K1e1, and del-K1e2) using Illumina NovaSeq6000 and yielding 40× coverage. We demonstrated that no off-target mutations were found (Table S4). These results revealed that the nuclease used in this study did not induce off-target mutations in two of the knockin clones, which is in line with a recent study demonstrating the high specificity of Cas9-mediated nuclease in the clonal populations of pluripotent stem cells (Park et al., 2015b; Veres et al., 2014).

Phenotypic Correction of the FVIII Deficiency *In Vitro*

After the successful targeted insertion of the *FVIII* gene in the *H11* locus, we examined the phenotype of the corrected clones using an *in vitro* culture system. Using semi-quantitative RT-PCR, the *FVIII* mRNA expression levels in the *H11* locus were measured at the undifferentiated stage. As expected, no PCR bands corresponding to *FVIII* exons 21 and 23 were amplified from the two types of patient iPSCs due to either incorrect splicing (inv-Pa) or deletion of the corresponding region (del-Pa); however, corresponding

(F) Detection of the *Cre* expression vector sequence remaining in each clone by PCR. The *GAPDH* gene was used as a quality control for the total isolated DNA. Total DNA isolated from the cells 3 days after electroporation was used as the positive control for the *Cre* expression vector.

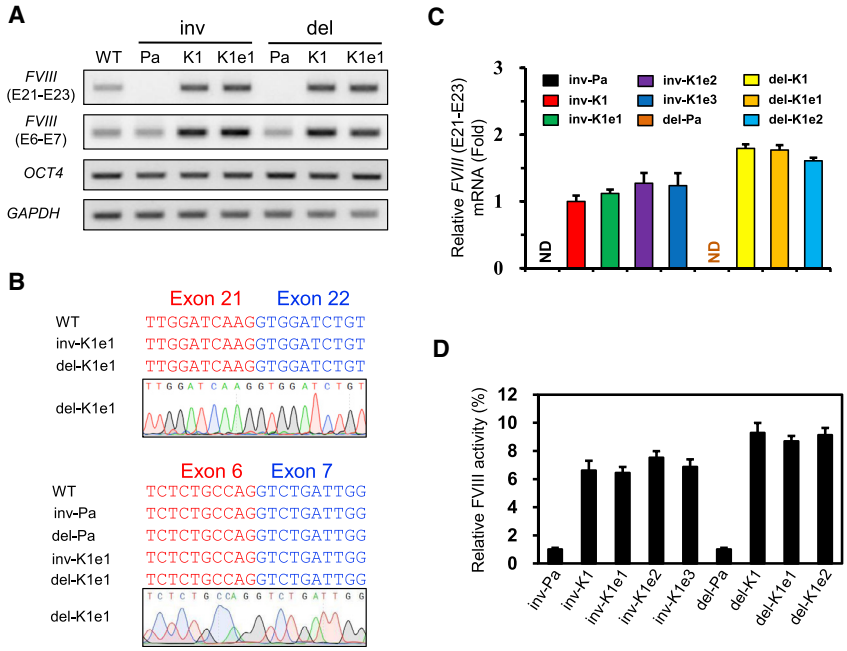

Figure 3. Phenotypic Rescue of the Expression of *FVIII* Gene from the Corrected iPSC Clones

(A) RT-PCR analysis to detect the expression of *FVIII* and *OCT4* in undifferentiated, corrected iPSC clones. *GAPDH* expression was used for normalization.

(B) Chromatograms showing the sequences of amplified DNA bands from each clone (related to A).

(C) Results of the qPCR analysis showing the *FVIII* expression levels in undifferentiated patient iPSC clones and corrected clones. *GAPDH* expression was used for normalization. Data are means ± SEM of three independent experiments. ND, not detected.

(D) The FVIII activity was determined after 20-fold concentration in supernatants obtained from either patient or corrected clones. Data indicate activity detected per 1 × 10^6 iPSCs. Data are means ± SEM of three independent experiments.

bands to *FVIII* exons 6 and 7 were amplified from both inv-Pa and del-Pa iPSC clones (Figure 3A). In contrast, PCR bands corresponding to exons 21 and 23 were amplified in the corrected iPSCs, and were detected regardless of the presence or absence of the selection cassette (Figure 3A). We also performed Sanger sequencing to verify the sequences of the amplified DNA bands (Figure 3B). In additional qPCR analyses, we demonstrated that there were no differences in the expression of the *FVIII* gene between each post-excision clone (inv-K1e1 to inv-K1e3, or del-K1e1 and del-K1e2) (Figure 3C). We then investigated whether the induction of *FVIII* expression corresponded to increased FVIII activity in the corrected clones. As shown in Figure 3D, all of the corrected clones, regardless of whether they had the selection cassette, exhibited highly elevated FVIII activity levels compared with levels in the parental patient iPSCs.

Next, the four corrected iPSC clones, including the two types of patient iPSCs, were differentiated into endothelial cells, a source of FVIII production (Shahani et al., 2010), as previously reported (Harding et al., 2017). We did not detect any impairment in the differentiation of the clones into endothelial cells, which exhibited cobblestone-like morphologies at the time of differentiation on day 4 (Figure 4A) and positive staining for endothelial cell markers such as CD31 and von Willebrand factor (vWF) at the end of differentiation (Figure 4B). We further evaluated the levels of *FVIII* mRNA using RT-PCR and qPCR. No PCR bands corresponding to *FVIII* exons 21 and 23 were detected in the patient endothelial cells that differentiated from inv-Pa and del-Pa clones (Figure 4C). However, as expected, the *FVIII* mRNA was detected in endothelial cells differentiated from the corrected iPSC clones (Figure 4C). In addition, all cells that differentiated from the corrected clones showed significantly elevated FVIII activity levels compared with the patient iPSCs (Figure 4D). These results indicated that the phenotype of the patient iPSCs could be corrected by the expression of the *FVIII* gene via targeted knockin at the *H11* locus.

DISCUSSION

In typical genome editing, correcting the endogenous locus itself is an ideal strategy (Park et al., 2016b). However, there are some current limitations to targeting the endogenous site in HA: (1) there are no protocols for differentiating iPSCs into LSECs, which are the primary producers of the FVIII protein; (2) microvascular endothelial cells cannot fully correct the phenotype because they produce only a small amount of FVIII protein; and (3) it is not feasible to prepare sgRNA, including donor DNA, for the 2,015 unique mutations found in HA. To overcome these limitations, we attempted the targeted knockin of the functional *FVIII* gene into a safe harbor site for a universal correction regardless of mutation types. Although there was an attempt to target *FVIII* gene at ribosomal DNA locus in HA patient iPSCs using TALENickases, this attempt has not demonstrated successful functional restoration in an *in vivo* system (Pang et al., 2016).

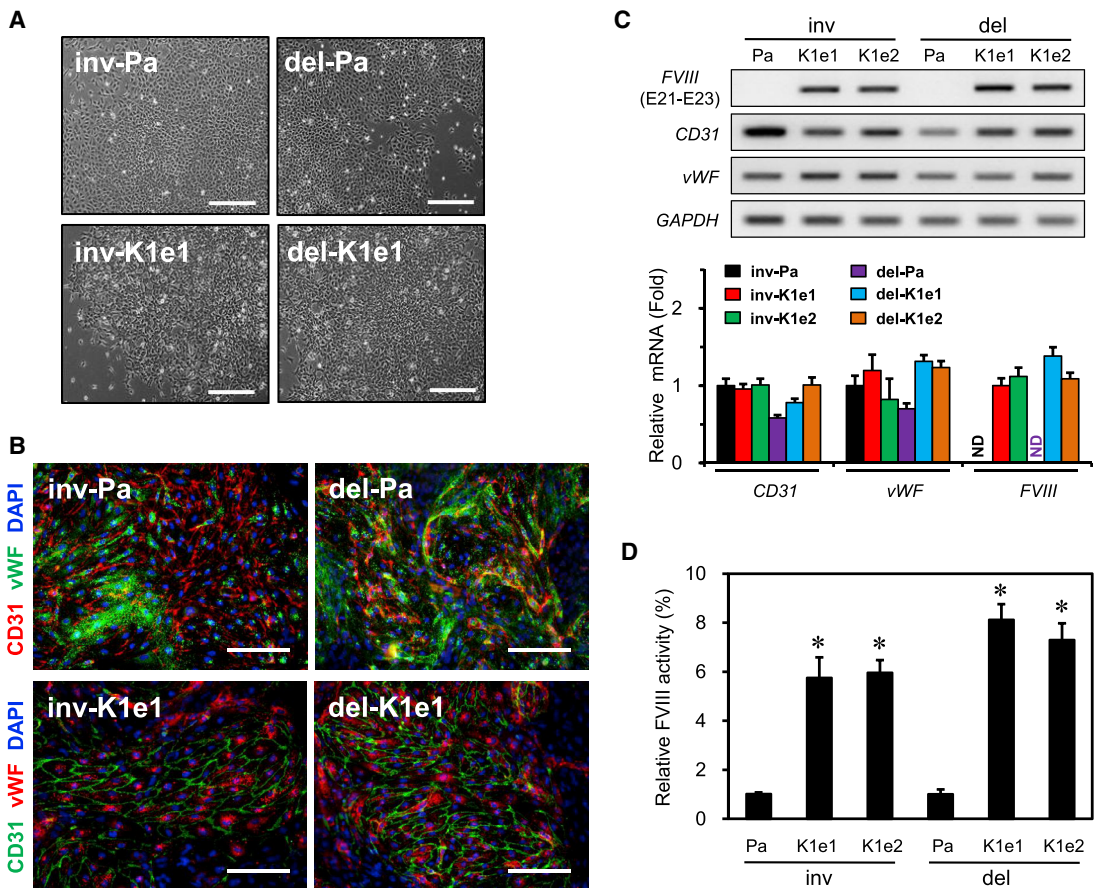

Figure 4. Functional Correction of FVIII Deficiency in Endothelial Cells Differentiated from Corrected iPSC Clones

(A) Phase image of cobblestone-like morphologies at the time of differentiation on day 4 in the indicated clones. Scale bars, 200 μm.

(B) Expression of marker proteins (CD31 and vWF) representing endothelial cells (ECs) derived from parental patient and corrected clones. The DAPI signal indicates the total cell content in the image. Scale bars, 100 μm.

(C) Expression of *FVIII* mRNA including EC marker genes were verified in cells differentiated from patients and corrected clones using RT-PCR (upper panel) and qPCR (lower panel). *GAPDH* expression was used for normalization. Data are means ± SEM of three independent experiments. ND, not detected.

(D) The FVIII activity was determined after 20-fold concentration in supernatants obtained from patient or corrected clones after differentiation into ECs. Data indicate activity detected per 1×10^6 ECs. Data are means ± SEM of three independent experiments. *$p < 0.001$ compared with the parental cells (Student's t test).

The *H11* locus was identified as a safe harbor site and used to express transgene in iPSCs (Turan et al., 2016; Zhu et al., 2014) or animal model (Ruan et al., 2015). Interestingly, this locus is an intergenic sequence; thus, knockin of a transgene into the site does not induce the gene disruption observed with other safe harbor sites such as AAVS1 and the albumin locus. Using this advantage, we chose the *H11* locus to develop our universal correction platform via targeted integration of the *FVIII* gene. In addition, we used two types of patient iPSCs, FVIII inverted and large deleted patient iPSCs, for a proof of principle. In our experiments, we generated corrected clones with high efficiency in the two different types of iPSCs and demonstrated the functional recovery of mRNA expression as well as the secretion of the FVIII protein from the corrected cells *in vitro*. Nevertheless, the functional effects following transplantation in an *in vivo* system still need to be confirmed.

Off-target mutations are a concern regarding the therapeutic use of engineered nucleases. To avoid off-target effects, we chose unique target sequences that differed from any other site in the human genome by at least three nucleotides. We also used 5′-GGX_{20} sgRNAs transcribed *in vitro*, which reduced off-target mutations without reducing on-target activity as previously reported (Cho et al., 2014), and electroporated Cas9 RNPs into the patient iPSCs (Kim et al., 2014). In addition, targeted deep sequencing was performed to validate the absence of unwanted mutations in the corrected clones following

the removal of the puromycin expression cassette, resulting in no off-target mutations. These results are consistent with previous reports of the reduced induction of off-target mutations in individual clones of edited cells (Park et al., 2015b; Veres et al., 2014).

In summary, we used the *H11* locus to develop a universal correction platform, and precisely targeted the *FVIII* gene via error-free knockin with high efficiency. Finally, we verified that endothelial cells (ECs) successfully differentiated from the corrected iPSCs containing the *FVIII* gene and secreted functional protein *in vitro* system. This approach is not only simple but may also provide a universal platform to correct for the various mutations found in HA using the same sgRNA and donor DNA for one target site, the human *H11* locus.

EXPERIMENTAL PROCEDURES

Ethical Statement

The generation and analyses of iPSCs from the HA patients were approved by Yonsei University Institutional Review Board (IRB #4-2012-0028). All volunteers who participated in this study signed written informed consent forms before donating cells for the generation of iPSCs.

Preparations of Donor Plasmid and Guide RNA for SpCas9

To construct the donor plasmid, we used the pCDNA4/BDD-FVIII plasmid (www.addgene.org, no. 41035) (Peters et al., 2013) as a backbone. See Supplemental Experimental Procedures for details of the protocol.

Off-Target Analysis and Targeted Deep Sequencing

Ten potential off-target sites differing by up to four nucleotides from the on-target site were searched using a web-based *in silico* tool (crispr.mit.edu) (Ran et al., 2013). For deep sequencing analysis, PCR amplicons for each off-target site and the on-target site were prepared from genomic DNA using high-fidelity PrimeSTAR Max DNA polymerase (Takara Bio) and the specific primer sets listed in Table S3. Following PCR purification of the resulting PCR products using an *AccuPrep* PCR Purification Kit (Bioneer, Korea), the PCR amplicons were subjected to paired-end sequencing using a MiSeq system (Illumina) at LAS (Korea).

CRISPR/Cas9-Mediated Correction by *FVIII* Knockin

Cas9 RNPs and *FVIII* knockin donor DNA were electroporated into the patient iPSCs as previously described (Park et al., 2015b) with slight modifications. See Supplemental Experimental Procedures for details of the protocol.

Excision of the Puromycin Selection Cassette

To excise the puromycin expression cassette from the *FVIII* knockin clones, we electroporated 2×10^5 knockin iPSCs with 1 μg of the *Cre* expression plasmid (pCAG-*Cre*:GFP; no. 13776, www.addgene.org). The isolation of clones was performed by single cell passaging as previously described (Park et al., 2016c). The excision of the puromycin cassette was confirmed by PCR-based genotyping using the F3 and R2 primers. After removal of the puromycin expression cassette, we confirmed the absence of the *Cre* expression vector in each clone by PCR using the CRE-F and CRE-R primers listed in Table S1.

Differentiation into Endothelial Cells

To induce the differentiation of the iPSCs into ECs, we used a previously described protocol (Harding et al., 2017) with slight modifications. See Supplemental Experimental Procedures for details of the protocol.

Measurement of FVIII Activity

To measure FVIII activity, we concentrated culture supernatants 20-fold using an Amicon Ultra-4 centrifugal filter (Millipore). The FVIII activities were measured in the culture supernatants using the commercially available Coamatic Factor VIII chromogenic assay kit (Instrumentation Laboratory) according to the manufacturer's instructions. See Supplemental Experimental Procedures for details of the protocol.

Statistical Analysis

All data are expressed as means ± SEM of at least three independent experiments. Statistically significant differences were estimated using a Student's t test. A resulting p value of <0.05 was considered statistically significant.

ACCESSION NUMBERS

The deep sequencing data files reported in this study have been deposited in the GEO database repository (www.ncbi.nlm.nih.gov/geo/) under the accession number GEO: GSE124663. The whole-genome sequencing data files have also been deposited in the SRA database (www.ncbi.nlm.gov/sra/) under the accession number PRJNA515982.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.stemcr.2019.04.016.

AUTHOR CONTRIBUTIONS

C.-Y.P. designed and carried out the experiments. S.-R.C. and J.J.S. helped with iPSC clone derivation and validation. C.-Y.P. and D.-W.K. interpreted the results. C.-Y.P., J.K., and D.-W.K. were in charge of critical revision. C.-Y.P. and D.-W.K. wrote the manuscript.

ACKNOWLEDGMENTS

C.-Y.P. was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (2016R1C1B1008742). D.-W.K. was supported by the Bio & Medical Technology Development Program of the NRF (2017M3A9B4042580), the Korea Health Technology R&D Project from the Ministry of Health and Welfare (HI18C0829), and the

Faculty Research Grant of Yonsei University College of Medicine (6-2017-0190).

Received: August 15, 2018
Revised: April 17, 2019
Accepted: April 17, 2019
Published: May 16, 2019

REFERENCES

Berntorp, E., and Shapiro, A.D. (2012). Modern haemophilia care. Lancet *379*, 1447–1456.

Cherry, A.B., and Daley, G.Q. (2013). Reprogrammed cells for disease modeling and regenerative medicine. Annu. Rev. Med. *64*, 277–290.

Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., and Kim, J.S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res. *24*, 132–141.

Graw, J., Brackmann, H.H., Oldenburg, J., Schneppenheim, R., Spannagl, M., and Schwaab, R. (2005). Haemophilia A: from mutation analysis to new therapies. Nat. Rev. Genet. *6*, 488–501.

Harding, A., Cortez-Toledo, E., Magner, N.L., Beegle, J.R., Coleal-Bergum, D.P., Hao, D., Wang, A., Nolta, J.A., and Zhou, P. (2017). Highly efficient differentiation of endothelial cells from pluripotent stem cells requires the MAPK and the PI3K pathways. Stem Cells *35*, 909–919.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Kim, S., Kim, D., Cho, S.W., Kim, J., and Kim, J.S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res. *24*, 1012–1019.

Li, H.L., Fujimoto, N., Sasakawa, N., Shirai, S., Ohkame, T., Sakuma, T., Tanaka, M., Amano, N., Watanabe, A., Sakurai, H., et al. (2015). Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. Stem Cell Reports *4*, 143–154.

Mandai, M., Watanabe, A., Kurimoto, Y., Hirami, Y., Morinaga, C., Daimon, T., Fujihara, M., Akimaru, H., Sakai, N., Shibata, Y., et al. (2017). Autologous induced stem-cell-derived retinal cells for macular degeneration. N. Engl. J. Med. *376*, 1038–1046.

Pang, J., Wu, Y., Li, Z., Hu, Z., Wang, X., Hu, X., Liu, X., Zhou, M., Liu, B., Wang, Y., et al. (2016). Targeting of the human F8 at the multicopy rDNA locus in Hemophilia A patient-derived iPSCs using TALENickases. Biochem. Biophys. Res. Commun. *472*, 144–149.

Park, C.Y., Halevy, T., Lee, D.R., Sung, J.J., Lee, J.S., Yanuka, O., Benvenisty, N., and Kim, D.W. (2015a). Reversion of FMR1 methylation and silencing by editing the triplet repeats in fragile X iPSC-derived neurons. Cell Rep. *13*, 234–241.

Park, C.Y., Kim, D.H., Son, J.S., Sung, J.J., Lee, J., Bae, S., Kim, J.H., Kim, D.W., and Kim, J.S. (2015b). Functional correction of large factor VIII gene chromosomal inversions in hemophilia a patient-derived iPSCs using CRISPR-Cas9. Cell Stem Cell *17*, 213–220.

Park, C.Y., Lee, D.R., Sung, J.J., and Kim, D.W. (2016a). Genome-editing technologies for gene correction of hemophilia. Hum. Genet. *135*, 977–981.

Park, C.Y., Sung, J.J., and Kim, D.W. (2016b). Genome editing of structural variations: modeling and gene correction. Trends Biotechnol. *34*, 548–561.

Park, C.Y., Sung, J.J., Choi, S.H., Lee, D.R., Park, I.H., and Kim, D.W. (2016c). Modeling and correction of structural variations in patient-derived iPSCs using CRISPR/Cas9. Nat. Protoc. *11*, 2154–2169.

Peters, R.T., Toby, G., Lu, Q., Liu, T., Kulman, J.D., Low, S.C., Bitonti, A.J., and Pierce, G.F. (2013). Biochemical and functional characterization of a recombinant monomeric factor VIII-Fc fusion protein. J. Thromb. Haemost. *11*, 132–141.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308.

Ruan, J., Li, H., Xu, K., Wu, T., Wei, J., Zhou, R., Liu, Z., Mu, Y., Yang, S., Ouyang, H., et al. (2015). Highly efficient CRISPR/Cas9-mediated transgene knockin at the H11 locus in pigs. Sci. Rep. *5*, 14253.

Shahani, T., Lavend'homme, R., Luttun, A., Saint-Remy, J.M., Peerlinck, K., and Jacquemin, M. (2010). Activation of human endothelial cells from specific vascular beds induces the release of a FVIII storage pool. Blood *115*, 4902–4909.

Shi, Y., Inoue, H., Wu, J.C., and Yamanaka, S. (2017). Induced pluripotent stem cell technology: a decade of progress. Nat. Rev. Drug Discov. *16*, 115–130.

Turan, S., Farruggio, A.P., Srifa, W., Day, J.W., and Calos, M.P. (2016). Precise correction of disease mutations in induced pluripotent stem cells derived from patients with limb girdle muscular dystrophy. Mol. Ther. *24*, 685–696.

Veres, A., Gosis, B.S., Ding, Q., Collins, R., Ragavendran, A., Brand, H., Erdin, S., Talkowski, M.E., and Musunuru, K. (2014). Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. Cell Stem Cell *15*, 27–30.

Wu, Y., Hu, Z., Li, Z., Pang, J., Feng, M., Hu, X., Wang, X., Lin-Peng, S., Liu, B., Chen, F., et al. (2016). In situ genetic correction of F8 intron 22 inversion in hemophilia A patient-specific iPSCs. Sci. Rep. *6*, 18865.

Xu, X., Tay, Y., Sim, B., Yoon, S.I., Huang, Y., Ooi, J., Utami, K.H., Ziaei, A., Ng, B., Radulescu, C., et al. (2017). Reversal of phenotypic abnormalities by CRISPR/Cas9-mediated gene correction in Huntington disease patient-derived induced pluripotent stem cells. Stem Cell Reports *8*, 619–633.

Zhu, F., Gamboa, M., Farruggio, A.P., Hippenmeyer, S., Tasic, B., Schule, B., Chen-Tsai, Y., and Calos, M.P. (2014). DICE, an efficient system for iterative genomic editing in human pluripotent stem cells. Nucleic Acids Res. *42*, e34.

Stem Cell Reports, Volume *12*

Supplemental Information

Universal Correction of Blood Coagulation Factor VIII in Patient-Derived Induced Pluripotent Stem Cells Using CRISPR/Cas9

Chul-Yong Park, Jin Jea Sung, Sung-Rae Cho, Jongwan Kim, and Dong-Wook Kim

Supplemental Information

Universal correction of blood coagulation factor VIII in patient-derived induced pluripotent stem cells using CRISPR/Cas9

Chul-Yong Park, Jin Jea Sung, Sung-Rae Cho, Jongwan Kim and Dong-Wook Kim

Supplemental Figure 1

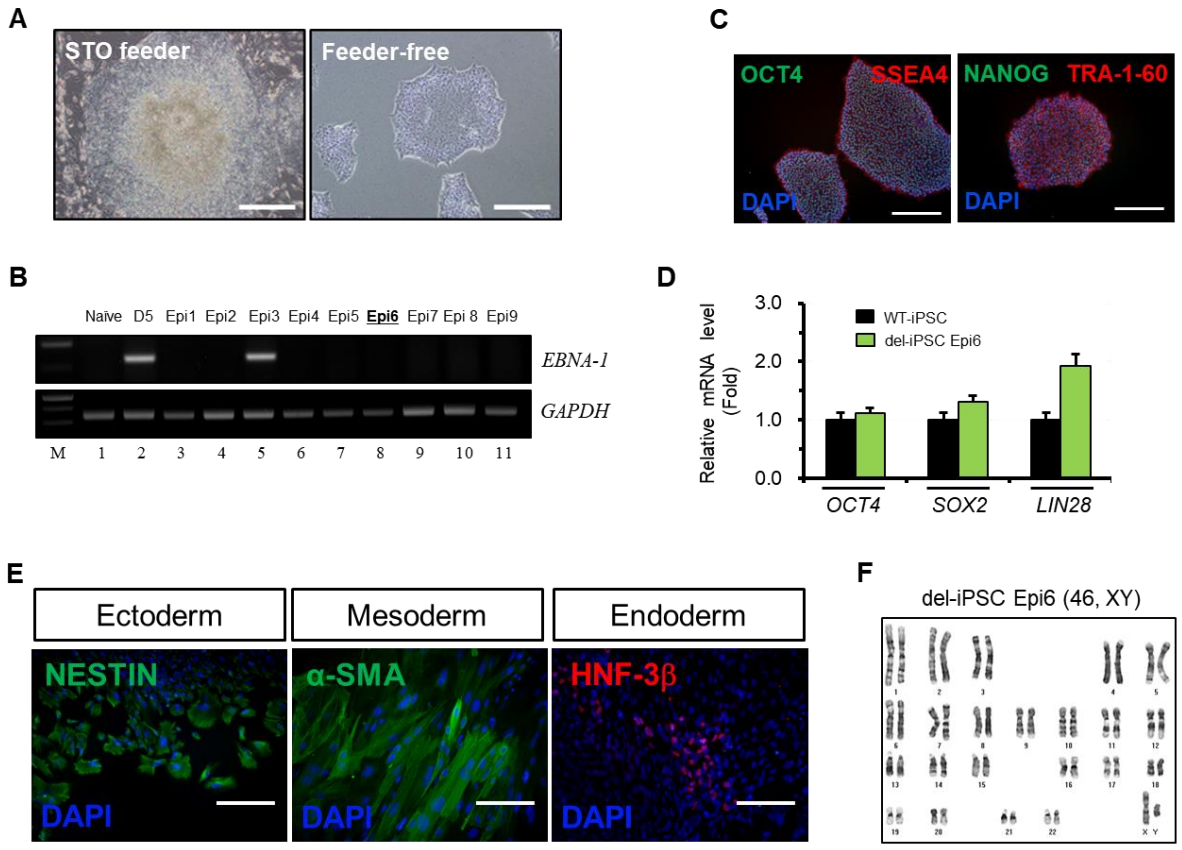

Figure S1. Generation of *FVIII* Deleted Patient-iPSCs.

(A) Human ESC-like morphology of expanded patient-iPSCs (Epi6) onto a STO feeder layer or in a feeder-free culture conditions. Scale bars, 500 μm (feeder layer) or 200 μm (feeder-free culture).

(B) Detection of episomal vector sequence (*EBNA-1*) that remained in established iPSC clones (Epi1 to Epi9). The *GAPDH* gene is used as a quality control for isolated total DNA. Total DNA isolated from the cells before (naïve) and after electroporation (day 5, D5) was used as negative and positive controls for episomal vector DNA, respectively.

(C) The expression of the pluripotency markers OCT4, NANOG, SSEA-4, and TRA-1-60 detected by immunocytochemistry. The DAPI signal indicates the total cell content in the image. Scale bar, 200 μm.

(D) The mRNA quantification of endogenous *OCT4*, *SOX2*, and *LIN28* from the indicated clones were measured by qPCR and normalized to *GAPDH* expression. Data are means ± SEM of three independent experiments.

(E) The expression of marker proteins representing ectoderm (NESTIN), mesoderm (α-

smooth muscle actin, α-SMA), and endoderm (hepatocyte nuclear factor-3β, HNF-3β). Scale bar, 100 μm. The DAPI signal indicates the total cell content in the image.
(F) G-banding analyses were performed in the Epi6 clones. Related to Figure 1.

Supplemental Figure 2

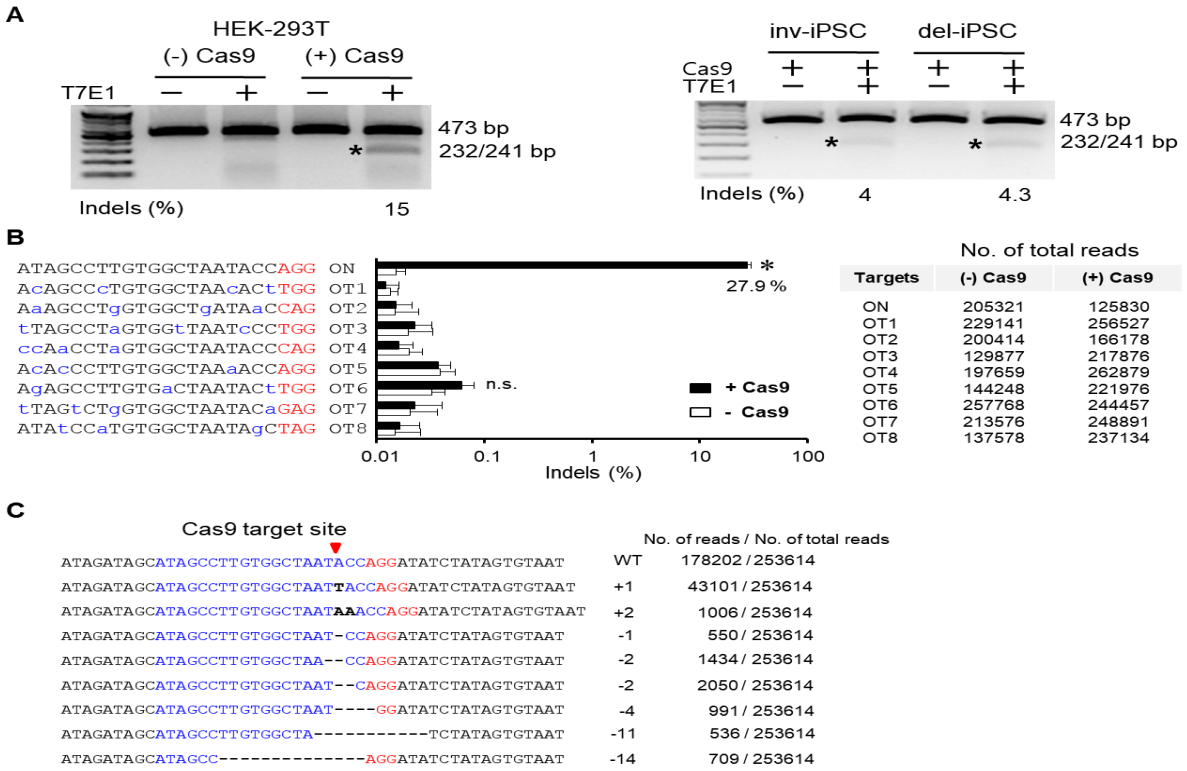

Figure S2. sgRNA design for the human *H11* locus and validation of nuclease activity at on- and off-target sites.

(A) Indels frequencies at the on-target site were analyzed in HEK-293T cells (left) and each iPSC clone (right) using the T7E1 assay. The asterisk indicates the predicted position of DNA bands cleaved by T7E1.

(B) Indels in the on- and off-target site were analyzed by targeted deep sequencing. Eight potential off-target sites in the human genome that differed from the on-target site by up to four nucleotides were examined in HEK-293T cells. Mismatched nucleotides and PAM sequences (5'-NRG-3', R = A or G) are shown in blue and in red, respectively. Data are means ± SEM of three independent experiments. The total number of reads is provided as an indicator of deep sequencing sensitivity. n.s., not significant compared with absence of Cas9; *, $p < 0.001$ compared with absence of Cas9 (Student's t-test).

(C) Data indicates the partial DNA sequences of indels that are analyzed from deep sequencing. The number of inserted or deleted bases is indicated in the right column. Red arrow indicate cleavage position in the target sequences. The number of reads with WT and each edited sequence is indicated. Related to Figure 1.

Supplemental Figure 3

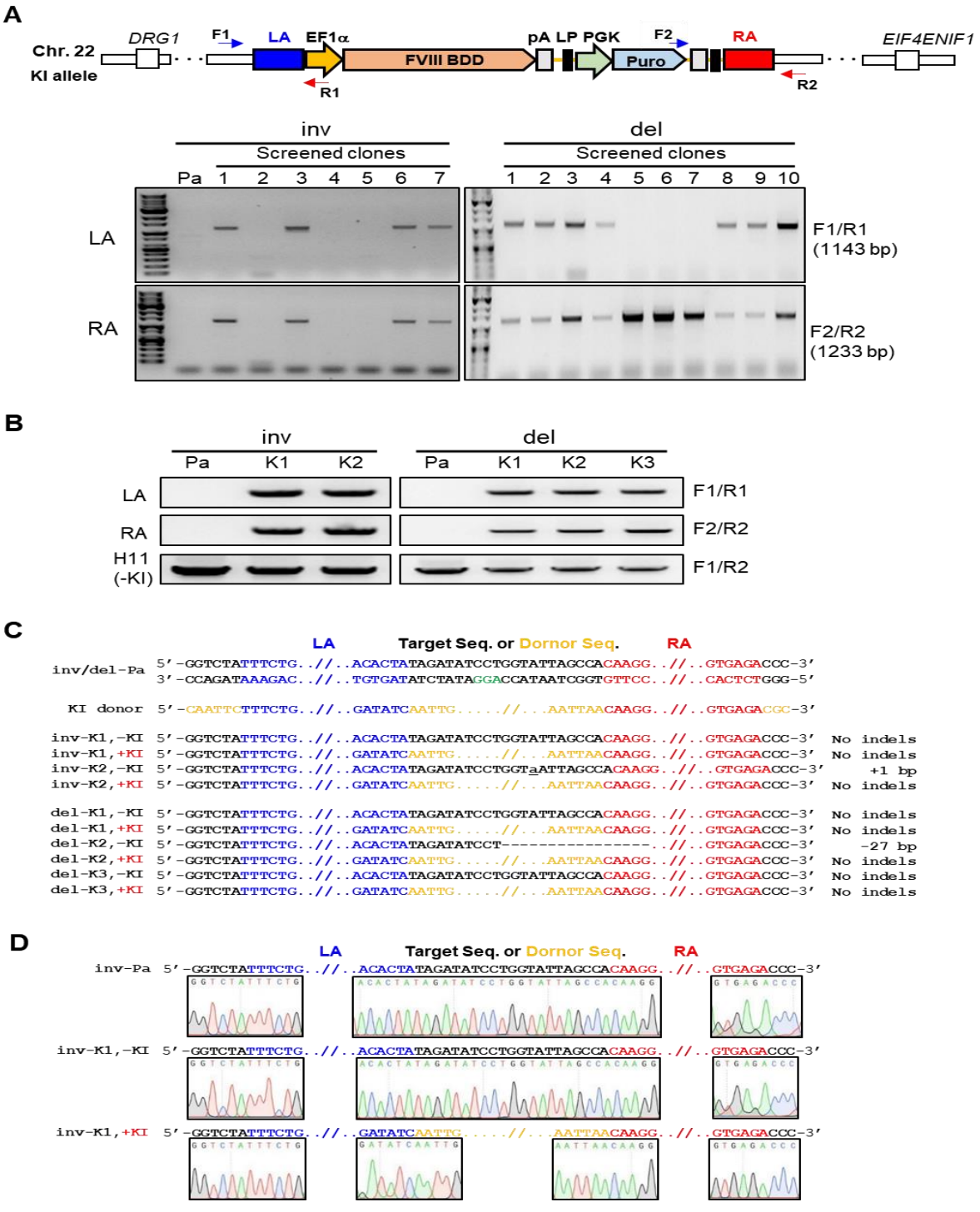

Figure S3. Targeted insertions of the *FVIII* donor DNA.

(A) Genome structure of *FVIII* inserted allele in the target site. The four specific primers used for genotyping are shown. PCR-based screening for targeted knock-in using the four specific primers listed in Table S1 (lower).

(B) PCR-based genotype analysis to confirm the targeted insertion of donor DNA in the

indicated iPSCs. The four specific primers used for genotyping are shown.

(C) The DNA sequences of knock-in junctions in the indicated iPSC clones including the template donor DNA. Underlined lowercase letter indicates inserted base. Dashes indicate deleted bases. The donor sequences are shown in deep yellow. The original genome sequences are shown in black. The left and right arm sequences are shown in blue and red, respectively. The PAM sequence is shown in green.

(D) Chromatograms showing the targeted insertion of donor DNA in the corrected clones. Related to Figure 1.

Supplemental Figure 4

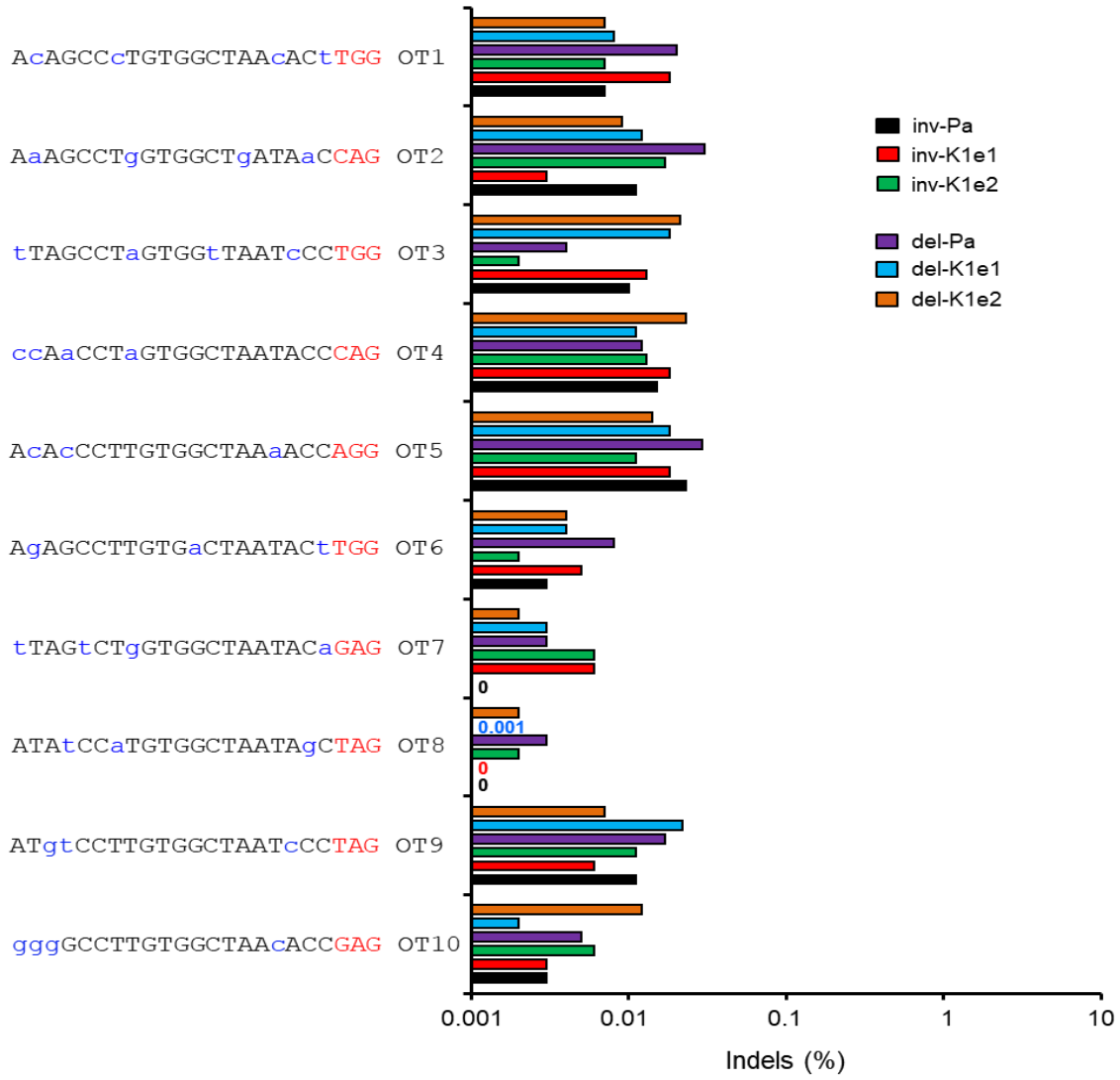

Figure S4. Off-target analyses in the corrected iPSCs by targeted deep sequencing.
Ten potential off-target sites differing by up to four nucleotides from the on-target site were examined in the corrected clones by targeted deep sequencing. Mismatched nucleotides and PAM sequences (5'-NRG-3', R = A or G) are shown in blue and in red, respectively. Related to Figures 1 and 2.

Table S1. Sequences of primers used in this study. Related to Figures 1, 2, 3, and 4.

Primer[a]	Sequence (5' to 3')	Used for the experiment of
F1	CAACCCAGTCCTCCTTACCTTGC	genotype PCR
R1	CCGTTGCGAAAAAGAACGTTCAC	genotype PCR
F2	GCCTGAGGGGATCAATTCTCTAG	genotype PCR
R2	GTTAGGCCTGTGTCAACAGTTTGG	genotype PCR
F3	GGCCCTTTCGTCTTCAAGAATTCCG	genotype PCR
H11-F	GTGAGCTAAGGAATGTGATACAG	T7E1 assay
H11-R	AGGCTGAGACAGAAGAATCGCCT	T7E1 assay
Oct4-F	CCTCACTTCACTGCACTGTA	qPCR
Oct4-R	CAGGTTTTCTTTCCCTAGCT	qPCR
Sox2-F	CCCAGCAGACTTCACATGT	qPCR
Sox2-R	CCTCCCATTTCCCTCGTTTT	qPCR
Lin28-F	AGCCAAGCCACTACATTC	qPCR
Lin28-R	AGATACGTCATTCGCACA	qPCR
DRG1-F	ACGATGAGCAGCACCTTAGC	qPCR
DRG1-R	CGACGAAGCTTAGCAAGACG	qPCR
EIF4ENIF1-F	GACCTGAAGAAGCCTCCTGCCTC	qPCR
EIF4ENIF1-R	TTCTCAGGGTCCCAGACACCATC	qPCR
GAPDH-F	TGCACCACCACCTGCTTAGC	qPCR and RT-PCR
GAPDH-R	GGCATGGACTGTGGTCATGAG	qPCR and RT-PCR
F8-exon21-F	CCGGATCAATCAATGCCTGGAG	qPCR and RT-PCR
F8-exon23-R	ATGAGTTGGGTGCAAACGGATG	qPCR and RT-PCR
F8-exon6-F	GCCTGGCCTAAAATGCACACAGTC	RT-PCR
F8-exon7-R	ATTGAGTGCACTTCAGGAGTGGTG	RT-PCR
CD31-F	GGTCAGCAGCATCGTGGTCAACATAAC	qPCR
CD31-R	TGGAGCAGGACAGGTTCAGTCTTTCA	qPCR
vWF-F	TCGGGCTTCACTTACGTTCT	qPCR
vWF-R	CCTTCACTCGGACACACTCA	qPCR
CRE-F	CTACCTGTTTTGCCGGGTCAGAAA	PCR
CRE-R	AAACTCCAGCGCGGGCCATATCTC	PCR
GAPDH-F	GAACATCATCCCTGCCTCTACTG	PCR
GAPDH-R	CAGGAAATGAGCTTGACAAAGTGG	PCR
EBNA-1-F	ATGGACGAGGACGGGGAAGA	PCR
EBNA-1-R	GCCAATGCAACTTGGACGTT	PCR

[a] Primer direction: F, forward; R, reverse

Table S2. Summary of targeting efficiency by PCR screening and Sanger sequencing. Related to Figure 1.

	iPSCs	No. of puromycine resistant colonies/ $5x10^5$ iPSCs	No. of PCR (+) colonies/ No. of screened colonies	No. of targeted KI clones by sequencing/ No. of screened clones	No. of targeted KI clones confirmed by FVIII activity/ No. of screened clones
Pre-excision of puro cassette	Inv	25	9/14 (64.3%)	2/2	2/2
	Del	37	18/27 (66.7%)	3/3	3/3
Post-excision of puro cassette	Inv	-	6/7 (85.7%)	3/3	3/3
	Del	-	5/7 (71.4%)	2/2	2/2

Table S3. Sequences of each target site and primers used in on- and off-target amplification. Related to Figures 1 and 2.

Target site	Target Seq. (5' to 3')	Forward primer Seq. (5' to 3')	Reverse primer Seq. (5' to 3')
On-target (On, Chr. 22)	ATAGCCTTGTGGCTAATACC	GATTTGTTTGAGAGAAACTACC	CCTTGAGCTTTAAAGACCCC
Off-target 1 (OT1, Chr. 17)	ACAGCCCTGTGGCTAACACT	ATGGTTAGCAGTCAACCGTG	CCACAAACAAGAGGTCATCC
Off-target 2 (OT2, Chr. 9)	AAAGCCTGGTGGCTGATAAC	CTTCACTACCATCCCCCACT	ACTTTTGACATCCATCGCCG
Off-target 3 (OT3, Chr. 6)	TTAGCCTAGTGGTTAATCCC	TTAGCACTAAGATGGACCAC	TACCTGCTTGGCACCATGTC
Off-target 4 (OT4, Chr. 20)	CCAACCTAGTGGCTAATACC	GTAGTTTTAAAATAGGAGTCAGGG	GGTTCAATAACCATTGTTCCC
Off-target 5 (OT5, Chr. 5)	ACACCCTTGTGGCTAAAACC	AGTGGACAGAATTCCCTCTG	TACACTAGGGGTCATTGAGC
Off-target 6 (OT6, Chr. 14)	AGAGCCTTGTGACTAATACT	AGGAAAGCGTCTACTGTTAG	GTGTTAAATCTAGTGTGTTGC
Off-target 7 (OT7, Chr. X)	TTAGTCTGGTGGCTAATACA	CCTTCAGAGTGGAATGCTAT	ATTAGAAGCGTCCTGGTAGG
Off-target 8 (OT8, Chr. 4)	ATATCCATGTGGCTAATAGC	CAGGCCAGACTATAGAAGTT	AAGTTCAAATGCCCAATGGG
Off-target 9 (OT9, Chr. 7)	ATGTCCTTGTGGCTAATCCC	TCTCTGGGGCTGAAACCCAA	GGAGGTCTTTGTGTCTTAGC
Off-target 10 (OT10, Chr. X)	GGGGCCTTGTGGCTAACACC	AGACCTATCAGAGAGCCTAG	TAAGTTCCCCACAGCATCTC

Table S4. Analysis of potential off-target sites with whole genome sequencing. Related to Figures 1 and 2.

Variants		del-Pa	del-K1e1	del-K1e2
All variants	Single nucleotide polymorphisms (SNPs)	3,489,220	3,565,265	3,536,560
	Small insertions and deletions (INDELs)	539,134	588,046	565,157
	Structural variations (SVs)	5,170	5,556	5,546
	Copy number variations (CNVs)	703	659	730
Filtered Variants (variants not present in the common variant database; 1000 genomes variants)	SNPs	373,168	387,401	365,403
	INDELs	333,263	375,037	303,124
Clone specific variants	SNPs	22,946	28,240	20,739
	INDELs	41,695	67,897	31,086
Candidate variants at the 2,325 potential off-target sites (by up to six nucleotides mismatches)	SNPs and INDELs	Not applicable	0	0

Supplementary experimental procedures

Cell Cultures

HEK-293T (ATCC) cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% antibiotics. Adipose tissue-derived mesenchymal stem cells (MSCs) isolated from an *FVIII* deleted patient were cultured in DMEM (low glucose) supplemented with 10% FBS, 0.0145 g/L ascorbic acid, and 1% antibiotics on collagen type I coated plates. Human wild-type iPSC (WT-iPSCs Epi3 line) (Park et al., 2014), *FVIII* inverted patient-derived iPSCs (Pa2, intron 22 inverted iPSCs Epi5 line) (Park et al., 2015b), *FVIII* deleted patient-derived iPSCs (Epi6 line), and *FVIII* knock-in clones generated from the *FVIII* inverted or deleted iPSC clones were maintained in iPSC culture medium [DMEM/F12 medium containing 4 ng/mL basic fibroblast growth factor (bFGF; PeproTech), 20% knockout serum replacement (Invitrogen), 1% nonessential amino acids (Invitrogen), and 0.1 mM 2-mercaptoethanol (Sigma)]. In some experiments, iPSC clones were cultured in StemMACS™ iPS-Brew XF medium (Miltenyi Biotec) according to the manufacturer's instructions for feeder-free culturing.

Preparations of Donor Plasmid and Guide RNA for SpCas9

To construct the donor plasmid, we used the pCDNA4/BDD-FVIII plasmid (www.addgene.org, no. 41035) (Peters et al., 2013) as a backbone. The *MfeI* and *MauBI* restriction sites in the backbone were used to insert 5'-homology arm (Left arm, LA) and 3'-homology arm (Right arm, RA) by In-Fusion cloning, respectively. Next, the human EF1α promoter sequence was cloned between the left arm and the *FVIII* open reading frame. Afterwards, the puromycin expression cassette flanked by *loxP* sites was inserted between the *FVIII* gene cassette and the right arm by In-Fusion cloning. After the construction of the donor plasmid, we verified the sequences of the cloned DNA by Sanger sequencing at Solgent, Inc. (Korea). The recombinant Cas9 protein derived from *Streptococcus pyogenes* (SpCas9) was purchased from ToolGen, Inc. (Korea). A 5'-GGX$_{20}$ sgRNA (5'-ATAGCCTTGTGGCTAATACC-3') for SpCas9 was transcribed *in vitro* under the control of the T7 promoter using the MEGAshortscript™ kit (Ambion) according to the manufacturer's protocol followed by purification as previously described (Kim et al., 2014).

RNA Isolation, RT-PCR, and Quantitative PCR (qPCR) Analysis

Total RNAs were purified from cells using the Easy-Spin™ Total RNA Extraction Kit (iNtRON Biotech) according to the manufacturer's instructions. The cDNAs were synthesized from 1 μg of purified total RNA using the PrimeScript™ RT Master Mix (Takara Bio). For quantification of

the mRNA levels, qPCR was performed using the SYBR® Premix Ex-Taq (Takara Bio) and a CFX96 Real-Time System (Bio-Rad). Ct values for each gene were normalized to the Ct values for *GAPDH*. Semiquantitative RT-PCR was performed using the EmeraldAmp® GT PCR Master Mix (Takara Bio). For the amplification of *FVIII* mRNA from the knock-in clones, a forward primer located in exon 21 was used in combination with a reverse primer located in exon 23. To confirm the expression of *vWF* and *CD31* mRNA, we used previously reported primer sets (Wang et al., 2007; Morita et al., 2015). The specific primer sequences are shown in Table S1.

T7E1 Assay

To validate the activity of the CRISPR/Cas9 designed for the target site, a T7E1 assay was performed as previously described (Guschin et al., 2010, Park et al., 2015a). In brief, HEK-293T cells were co-transfected with 1 μg SpCas9 plasmid and 2 μg sgRNA plasmid purchased from ToolGen, Inc. (Korea) using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. Four days after co-transfection, genomic DNA segments encompassing the nuclease target sequence were amplified using high fidelity PrimeSTAR® Max DNA polymerase (Takara Bio) and the specific primer set listed in Table S1. The PCR amplicon was treated with T7E1 to cut the mismatched DNA. Following agarose gel electrophoresis, the band intensities of the uncleaved and cleaved fragments were determined.

Generation of iPSCs

The *FVIII* deleted patient-iPSCs were generated from MSCs obtained from a patient with severe HA who was clinically confirmed by the Korea Hemophilia Foundation Clinic using episomal reprogramming vectors (pCXLE-hOCT3/4-shp53-F, pCXLE-hSK, and pCXLE-hUL; no. 27077, 27078, and 27080 respectively; www.addgene.org) as previously reported (Okita et al., 2011). Briefly, expanded MSCs were electroporated with reprogramming vectors (1 μg each, total 3 μg) using the Neon Transfection System (Life Technologies). After being pulsed three times at a voltage of 1,650 for 10 ms, the cells were grown further in low glucose DMEM supplemented with 10% FBS on collagen type I coated plates. A week after electroporation, the cells were transferred onto mouse SIM Thioguanine/Ouabain-resistant mouse fibroblast cell line (STO) fibroblasts (ATCC) as a feeder layer. Human ESC-like iPSCs colonies were picked mechanically and further cultured for characterization. Afterwards, iPSCs were adapted in a feeder-free culture conditions for use in the *FVIII* knock-in experiments.

Three Germ Layer Differentiation

The *In vitro* differentiation of the iPSCs into three germ layers was performed as previously

described (Xu et al., 2017). The iPSCs were partially dissociated into clumps using collagenase type IV (Invitrogen). The clumps were then transferred onto low-attachment tissue plates to form embryonic bodies (EBs) and cultured in 5% FBS containing iPSC culturing medium without bFGF for a week. Next, the EBs were attached on Matrigel-coated coverslips and further cultured for 10 days to allow spontaneous differentiation of EBs into cells representing the three germ layers.

CRISPR/Cas9-Mediated Correction by *FVIII* Knock-in

Cas9 ribonucleoproteins (RNPs) and *FVIII* Knock-in donor DNA were electroporated into the patient-iPSCs as previously described (Park et al., 2015b) with slight modifications. Briefly, the SpCas9 protein (15 μg) was mixed with 20 μg of *in vitro* transcribed sgRNA and incubated for 10 min at room temperature to allow the formation of Cas9 RNPs. The patient iPSCs were pretreated with a 10 μM ROCK inhibitor (Y-27632, Sigma) for at least 2 hr before electroporation. After washing with PBS, the cells were treated with 1 × Versene solution (Life Technologies) for 3 min. Next, the cells were scraped and dissociated into approximately single cells. After centrifugation, 5×10^5 cells were electroporated with Cas9 RNPs and 3 μg donor DNA using the Neon Transfection System (Life Technologies). After being pulsed one time at a voltage of 950 for 10 ms, the cells were cultured in StemMACS™ iPS-Brew XF medium (Miltenyi Biotec) with 10 μM Y-27632 for 1 day. Four days post-transfection, the cells were selected using 0.5 μg/mL puromycin. To isolate clonal populations of the targeted cells, we performed three rounds of passaging.

PCR Analysis and DNA Sequencing of the Knock-in Junction

Genomic DNA was isolated from the cells using the DNeasy Blood & Tissue kit (Qiagen) according to the manufacturer's instructions. To confirm the targeted knock-in of donor DNA into the *H11* site, the DNA fragments for each knock-in junction were amplified using the EmeraldAmp® GT PCR Master Mix (Takara Bio) and the specific primer sets listed in Table S1. Following PCR purification, the sequence of each DNA amplicon was verified by Sanger sequencing at Solgent, Inc. (Korea).

Droplet Digital PCR (ddPCR) Analysis

Genomic DNA purified from each iPSC clone was subjected to ddPCR analysis to determine the copy number of the knock-in fragment using ddPCR SuperMix (without UTP; Bio-Rad) according to the manufacturer's protocol. The ddPCR reaction mixtures were converted into approximately 20,000 droplets using the QX200 Droplet Generator (Bio-Rad) and then were transferred into a 96-well plate for thermal cycling. Following PCR, the droplets were read

using the QX200 Droplet Reader (Bio-Rad) and were analyzed using Quantasoft software (Bio-Rad). The copy numbers were determined by calculating the ratio of the target molecule concentration to that of the reference molecule using the following equation: copy number = (*A*/*B*) N_B (A: concentration of target species, B: concentration of reference species, N_B: number of copies of reference loci in the genome). The following primer sets were used for ddPCR: Intron1-F: 5'-CGGGTTAGGATGGTTGTGATG-3' and Intron1-R: 5'-ATGACGAAGAGAAGCAATGGAC-3'; F8-Exon21-F: 5'-CGGATCAATCAATGCCTGGAG-3' and F8-Exon23-R: 5'-GAAGAGTGCTGCGAATGCT-3'. The primers for human *RPP30* (Bio-Rad; dHsaCP1000485) were used as a reference.

Whole-Genome Sequencing and Variant Calling

Genomic DNA was extracted from each clone using the G-DEX™ IIc genomic DNA extraction kit (iNtRON Biotech) according to the manufacturer's protocols. To construct the sequencing library, purified DNA was randomly fragmented, followed by 5' and 3' adaptor ligation. Adaptor-ligated libraries were then PCR amplified and gel purified. The libraries were subjected to sequencing using the Illumina NovaSeq6000 at Macrogen (South Korea). The sequencing depth was 40X. The sequencing data were converted into raw data using illumine package bcl2fastq (ver. 2.20.0). Adapters are not trimmed away from the reads. The paired-end sequences generated by the NovaSeq instrument were mapped to the human genome using Isaac Aligner (ver. 01.15.02.08) and the UCSC assembly hg19 (original GRCh37 from NCBI, Feb. 2009) was used as the reference sequence. Variants were called using the Isaac Variant Caller (ver. 2.0.13) for single-nucleotide variants and small indels, Control-FREEC (ver. 6.4) for copy-number variants, and Manta (ver. 0.20.2) for structural variants. We applied bioinformatic filters to more than three million variants to discard common variations listed in a public database (1000 Genomes Project, ftp://ftp.1000genomes.ebi.ac.uk). We then focused on identification of small indels including single nucleotide variations. We also identified unique variants induced in each clone by removing variations that were found simultaneously in all clones including the patient iPSC clone, because those were not induced by the nuclease. In parallel, we searched for 2,325 potential off-target sites that differed up to six nucleotides from the on-target site using Cas-OFFinder (www.rgenome.net) (Bae et al., 2014). To identify candidate variants as off-target mutations, we then compared 2,325 potential off-target sites with specific variants found in the each corrected clone.

Differentiation into Endothelial Cells

To induce the differentiation of the iPSCs into endothelial cells (ECs), we used a previously described protocol (Harding et al., 2017) with slight modifications. Briefly, the iPSCs were

pretreated with a 10 μM ROCK inhibitor (Y-27632, Sigma) for at least 2 hr before passaging. After washing with PBS, the cells were treated with 1× Versene solution (Life Technologies) for 3 min. Next, the cells were scraped and dissociated into small clumps. After centrifugation, the clumps were seeded onto Matrigel-coated culture dishes and cultured further in StemMACS™ iPS-Brew XF medium (Miltenyi Biotec) with 10 μM Y-27632 for 1 day. After 2 days, the culture medium was changed to STEMdiff™ APEL™ Medium (STEMCELL Technologies) with 6 μM CHIR99021 (Tocris Bioscience) for 2 days to induce the mesoderm. Afterwards, the cells were cultured in STEMdiff™ APEL™ Medium with 25 ng/mL bone morphogenetic protein 4 (BMP4, R&D Systems), 10 ng/mL bFGF (PeproTech), and 50 ng/mL vascular endothelial growth factor (VEGF, R&D Systems) for 2 days to induce vascular progenitors. Finally, the cells were cultured in EC Growth Medium MV2 (ECGM-MV2, PromoCell) with 50 ng/mL VEGF to generate endothelial progenitors, followed by media changes every 2 days.

Immunofluorescence Staining and Karyotype Analysis

Immunofluorescence staining was performed as previously described (Park et al., 2015a). Briefly, cells were fixed in a 4% paraformaldehyde solution containing 0.2% Triton X-100 for 10 min at room temperature (RT), washed three times with PBS, and incubated in blocking solution containing 5% normal goat serum and 2% bovine serum albumin for 1 hr at RT. Next, the cells were incubated with the primary antibodies for 2 hr at RT. The following primary antibodies were used: anti-OCT4 (Santa Cruz, SC9081), anti-NANOG (Abcam, AB21624), anti-SSEA-4 (Millipore, MAB4304), anti-TRA-1-60 (Millipore, MAB4360), anti-NESTIN (Millipore, MAB5326), anti-α-SMA (Sigma, A5228), anti-HNF-3β (Santa Cruz, SC6554), anti-CD31 (BD Bioscience, 555444), and anti-vWF (Millipore, AB7356). Afterwards, the cells were washed twice with PBS and incubated with fluorescence-conjugated secondary antibodies (Alexa Fluor® 488 or 594; Invitrogen) for 1 hr at RT. The cells were then washed again with a PBS containing 0.1% Tween 20 and mounted onto coverslips using a mounting medium. For the visualization of nuclei, 4′, 6-Diamidino-2-Phenylindole (DAPI; Vector Laboratories) was used. The images were captured and analyzed using an Olympus IX71 fluorescence microscope. For karyotype analysis, the chromosomes isolated from each iPSC clone were stained with Giemsa for G-banding analysis and analyzed using the Chromosome Image Processing System at GenDix Inc. (Korea).

Measurement of FVIII Activity

To measure FVIII activity, culture supernatants were concentrated 20-fold using an Amicon® Ultra-4 centrifugal filter (Millipore). The FVIII activities were measured in the culture

supernatants using the commercially available Coamatic® Factor VIII chromogenic assay kit (Instrumentation Laboratory) according to the manufacturer's instructions. The FVIII activity measurements were performed in a 96-well microplate and the absorbance at 405 nm was determined by an endpoint reading using a microplate reader (Molecular Devices). A standard curve was prepared by diluting the human calibration plasma (Instrumentation Laboratory).

Supplemental references

Bae, S., Park, J., and Kim, J.-S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 30, 1473–1475.

Guschin, D.Y., Waite, A.J., Katibah, G.E., Miller, J.C., Holmes, M.C., and Rebar, E.J. (2010). A rapid and general assay for monitoring endogenous gene modification. Methods Mol. Biol. *649*, 247-256.

Morita, R., Suzuki, M., Kasahara, H., Shimizu, N., Shichita, T., Sekiya, T., Kimura, A., Sasaki, K., Yasukawa, H., and Yoshimura, A. (2015). ETS transcription factor ETV2 directly converts human fibroblasts into functional endothelial cells. Proc. Natl. Acad. Sci. USA *112*, 160-165.

Okita, K., Matsumura, Y., Sato, Y., Okada, A., Morizane, A., Okamoto, S., Hong, H., Nakagawa, M., Tanabe, K., Tezuka, K.*, et al.* (2011). A more efficient method to generate integration-free human iPS cells. Nat. Methods *8*, 409-412.

Park, C.Y., Kim, J., Kweon, J., Son, J.S., Lee, J.S., Yoo, J.E., Cho, S.R., Kim, J.H., Kim, J.S., and Kim, D.W. (2014). Targeted inversion and reversion of the blood coagulation factor 8 gene in human iPS cells using TALENs. Proc. Natl. Acad. Sci. USA *111*, 9253-9258.

Wang, Z.Z., Au, P., Chen, T., Shao, Y., Daheron, L.M., Bai, H., Arzigian, M., Fukumura, D., Jain, R.K., and Scadden, D.T. (2007). Endothelial cells derived from human embryonic stem cells form durable blood vessels in vivo. Nat. Biotechnol. *25*, 317-318.