ORIGINAL ARTICLE

Advanced cell-based modeling of the royal disease: characterization of the mutated *F9* mRNA

L. MARTORELL,*†‡ E. LUCE,§¶** J.L. VAZQUEZ,†† Y. RICHAUD-PATIN,††
S. JIMENEZ-DELGADO,†† I. CORRALES,† N. BORRAS,† S. CASACUBERTA-SERRA,* A. WEBER,§¶**
R. PARRA,‡‡‡ C. ALTISENT,‡‡ A. FOLLENZI,§§ A. DUBART-KUPPERSCHMITT,§¶**
A. RAYA,††¶¶*** F. VIDAL†‡††† and J. BARQUINERO* (iD)

*Gene and Cell Therapy Laboratory, Vall d'Hebron Research Institute, Universitat Autònoma de Barcelona; †Congenital Coagulopathies Laboratory, Blood and Tissue Bank (BST); ‡Molecular Diagnosis and Therapy Unit, VHIR-UAB, Barcelona, Spain; §INSERM Unité Mixte de Recherche (UMR_S) 1193; ¶Université Paris-Sud; **Département Hospitalo-Universitaire Hepatinov, Paul Brousse Hospital, Villejuif, France; ††Center of Regenerative Medicine in Barcelona (CMRB); ‡‡Hemophilia Unit, Vall d'Hebron University Hospital, Barcelona, Spain; §§University of Piemonte Orientale, Novara, Italy; ¶¶Catalan Institution for Research and Advanced Studies (ICREA), Barcelona; ***Biomedical Research Networking Center on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain; and †††Biomedical Research Networking Center on Cardiovascular Diseases, Madrid, Spain

Essentials

- The Royal disease (RD) is a form of hemophilia B predicted to be caused by a splicing mutation.
- We generated an iPSC-based model of the disease allowing mechanistic studies at the RNA level.
- *F9* mRNA analysis in iPSC-derived hepatocyte-like cells showed the predicted abnormal splicing.
- Mutated *F9* mRNA level was very low but we also found traces of wild type transcripts.

Summary. *Background:* The royal disease is a form of hemophilia B (HB) that affected many descendants of Queen Victoria in the 19th and 20th centuries. It was found to be caused by the mutation F9 c.278-3A>G. *Objective:* To generate a physiological cell model of the disease and to study *F9* expression at the RNA level. *Methods:* Using fibroblasts from skin biopsies of a previously identified hemophilic patient bearing the F9 c.278-3A>G mutation and his mother, we generated induced pluripotent stem cells (iPSCs). Both the patient's and mother's iPSCs were differentiated into hepatocyte-like cells (HLCs) and their *F9* mRNA was analyzed using next-generation sequencing (NGS). *Results and Conclusion:* We demonstrated the previously predicted aberrant splicing of the *F9* transcript as a result of an intronic nucleotide substitution leading to a frameshift and the generation of a premature termination codon (PTC). The *F9* mRNA level in the patient's HLCs was significantly reduced compared with that of his mother, suggesting that mutated transcripts undergo nonsense-mediated decay (NMD), a cellular mechanism that degrades PTC-containing mRNAs. We also detected small proportions of correctly spliced transcripts in the patient's HLCs, which, combined with genetic variability in splicing and NMD machineries, could partially explain some clinical variability among affected members of the European royal families who had lifespans above the average. This work allowed the demonstration of the pathologic consequences of an intronic mutation in the F9 gene and represents the first *bona fide* cellular model of HB allowing the study of rare mutations at the RNA level.

Keywords: factor IX; hemophilia B; high-throughput nucleotide sequencing; induced pluripotent stem cells; RNA splicing.

Correspondence: Jordi Barquinero, Gene & Cell Therapy Laboratory, Vall d'Hebron Research Institute (VHIR), Pssg. Vall d'Hebron 119-129, 08035 Barcelona, Spain
Tel.: +34 932746726
E-mail: jordi.barquinero@vhir.org

Introduction

The royal disease (RD) is a form of hemophilia B most famously transmitted from Queen Victoria (1819–1901) to the European royal families during the 19th and 20th centuries. Because of its historical impact, it has attracted the

attention of historians and biomedical researchers for decades. The mutation (F9 c.278-3A>G) was identified in DNA from bone remains of members of the Russian royal family. It was predicted to create an abnormal splicing site two nucleotides ahead of the wild-type (wt) site in the *F9* transcript, resulting in the inclusion of two additional nucleotides at the exon 3/exon 4 boundary, a frameshift and a premature termination codon (PTC) in the mature mRNA [1]. This was functionally validated in a model based on COS7 cells transfected with a minigene construct (truncated intron 3 – exon 4 – truncated intron 4). However, this approach did not allow the capture of the complexity of splicing of the full-length *F9* transcript in hepatocytes, thus limiting the impact of this study. Because of the strict tissue specificity of the F9 promoter, a more precise analysis of the molecular mechanism leading to the bleeding phenotype would require studying *F9* mRNA in the hepatocytes of patients carrying this mutation.

We previously identified an 18-year-old Spanish patient with severe hemophilia B (HB), unrelated to the royal family, carrying this mutation [2]. The basal levels of coagulant factor (F) IX were < 1% in the patient and in the range of 30–40% in the carrier mother. A detailed description of the clinical history of this patient was previously reported [2]. The carrier has a history of hematomas, menorrhagia and bleeding episodes after minor surgeries, for which she has received replacement therapy with recombinant FIX. In the patient, a more sensitive quantification of his endogenous FIX plasma levels was not possible because he has been on prophylactic treatment since the age of 3 years and a liver biopsy to study *F9* mRNA was precluded for ethical reasons. To circumvent these limitations, we took advantage of the induced pluripotent stem cell (iPSC) technology to create a cellular model that captured the patient's genomic complexity and to study a more physiological *F9* expression upon *in vitro* differentiation into hepatocyte-like cells (HLCs).

Materials and methods

Human subjects

Diagnosis of the patient's HB and of the mother's carrier status was clinically and genetically confirmed at the Hemophilia Unit of Vall d'Hebron University Hospital [2]. Both mother and son agreed to undergo a skin biopsy for this study and signed a written informed consent previously approved by our institutional Ethics Committee and Review Board, which also approved the study.

Generation and maintenance of HB-patient-specific iPSCs (HB-iPSCs)

A 3-mm full-thickness skin punch biopsy was obtained from the volar surface of the forearm. The biopsy specimen was cut into 30–40 ~0.5-mm pieces and plated in HDF medium (DMEM containing 10% fetal bovine serum [FBS], 2 mM GlutaMAX, 50 U mL^{-1} penicillin and 50 mg mL^{-1} streptomycin [all from Gibco/Thermo Fisher Scientific, Waltham, MA, USA]). Primary cultures of fibroblasts were reprogrammed to pluripotency by retroviral transduction with a 1 : 1 : 1 : 1 mixture of retroviruses encoding FLAG-tagged Oct4, Sox2, Klf4 and c-Myc, as previously described [3]. Transduced cells were trypsinized and seeded onto irradiated CCD1112Sk (ATCC CRL2429) human foreskin fibroblasts (irrHFF) 2–3 days after infection. One day later, the medium was replaced with hESC medium (KO-DMEM supplemented with 20% KO-serum replacement, 2 mM Glutamax, 50 mM 2-mercaptoethanol, non-essential amino acids [all from Gibco/Thermo Fisher Scientific] and 10 ng mL^{-1} of human FGF-basic [Peprotech, Rocky Hill, NJ, USA]). Cells were maintained at 37 °C in 5% CO_2 and the medium was changed daily until iPSC colonies appeared. These patient-specific iPSC lines were maintained and passed by mechanical dissociation of the colonies, at a 1 : 3 ratio, and cultured on a feeder layer of irrHFF with hESC medium. Alternatively, the colonies were subjected to limited trypsin digestion and passaging onto Matrigel-coated plates with hESC medium pre-conditioned by irradiated mouse embryonic fibroblasts (MEFs).

Characterization of HB-iPSC lines

The patient-specific iPSC lines were characterized as previously described [3]. Briefly, the iPSC lines were selected based on colony morphology, growth characteristics, and the capacity to sustain long-term passaging (> 20 passages) while maintaining a normal karyotype. Alkaline phosphatase (AP) activity was measured using an AP blue/red membrane substrate solution kit (Sigma, St Louis, MO, USA) according to the manufacturer's guidelines. The expression of pluripotency-associated nuclear factors (OCT4, SOX2 and NANOG) and surface markers (SSEA3, SSEA4, TRA1-60 and TRA1-81), and the *in vitro* pluripotent differentiation ability of the cells to endoderm, mesoderm and neuroectoderm were analyzed by immunofluorescence as previously described [4] (Fig. 1). Briefly, the cells were fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS) at room temperature (RT) for 10 min, permeabilized for 15 min in 0.5% Triton in PBS, and then blocked for 2 h in Triton-X100 with 3% donkey serum (Millipore, Darmstadt, Germany). The primary antibodies used in all immunophenotypic studies are listed in Table S1. Secondary antibodies were of the Alexa Fluor series from Jackson ImmunoResearch (West Grove, PA, USA), and used at a 1 : 250–1 : 500 dilution. Images were taken using a Leica SP5 confocal microscope. Cell nuclei were counterstained with 0.5 µg mL^{-1} DAPI (4′,6-diamidino-2-phenylindole).
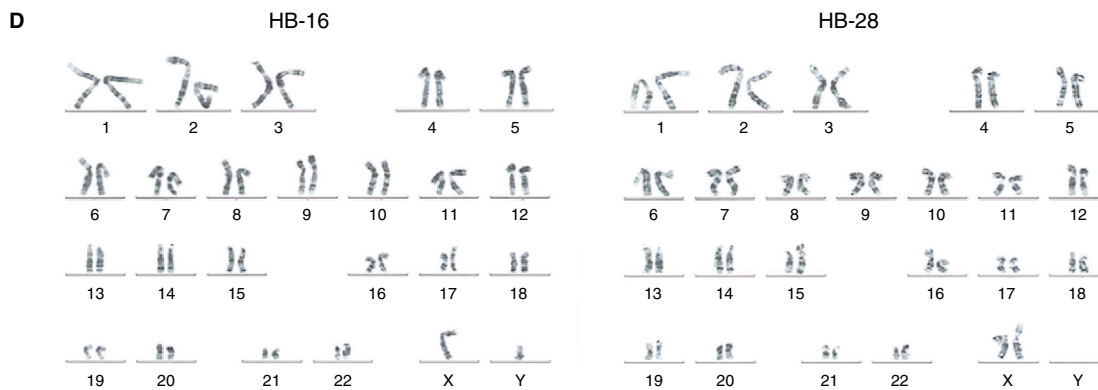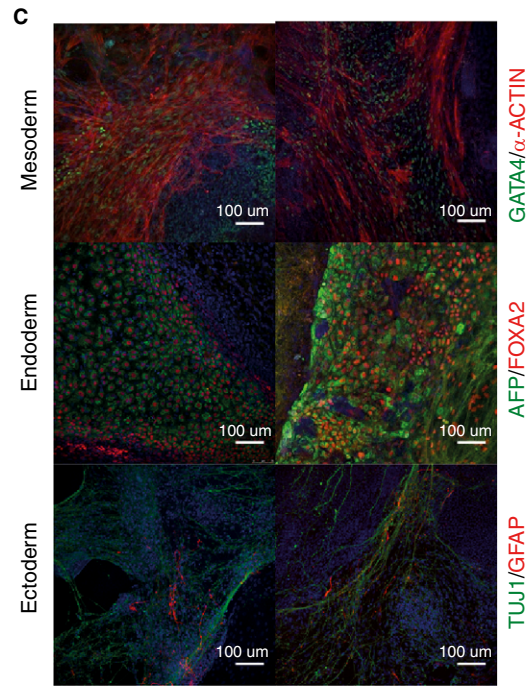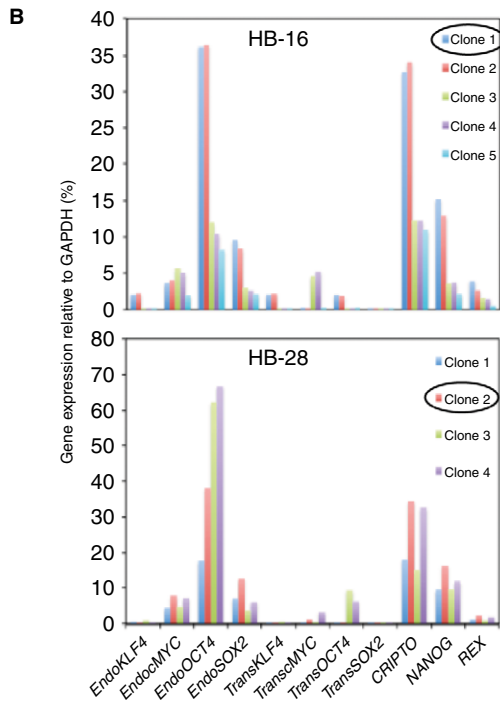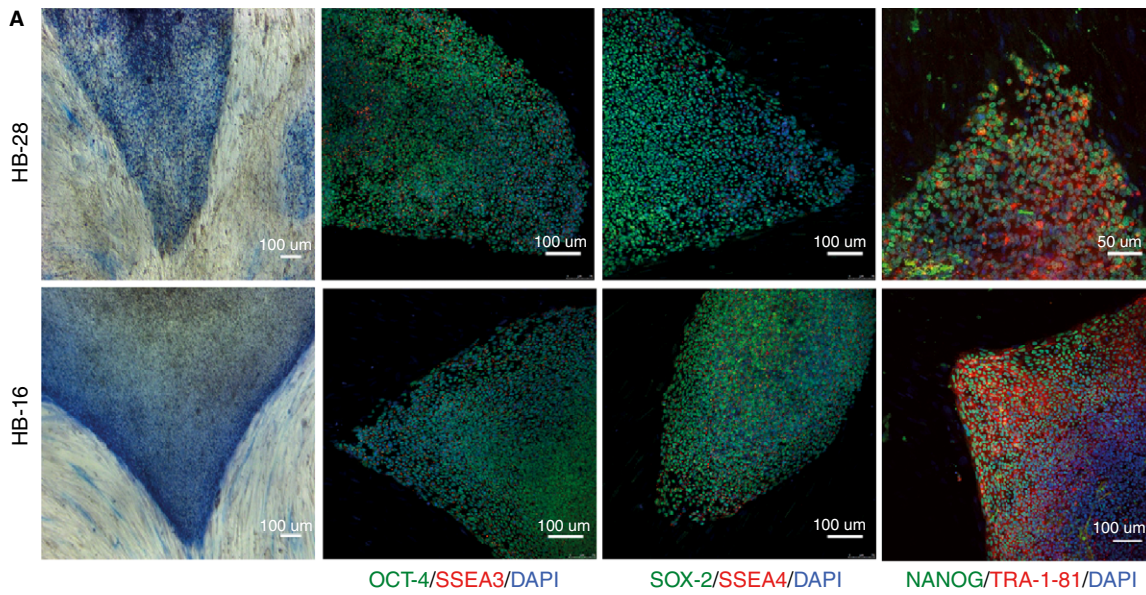
Fig. 1. Characterization of royal disease (RD) patient's and mother's induced pluripotent stem cells (iPSCs). (A) Alkaline phosphatase (AP) staining of representative iPSC colonies hemophilia B (HB)-28 (mother) and HB-16 (patient) (left), and immunofluorescence images showing expression of the pluripotency markers OCT4, SOX2, SSEA3 and 4, NANOG and TRA-1-81. (B) Relative expression of endogenous and exogenous pluripotency genes detected by RT-qPCR using specific primers (Table S3). iPSC clones were selected based on the activation of endogenous (endo) genes associated with pluripotency and the silencing of (exogenous) transgenes. (C) iPSCs could be differentiated into the three germ layers, as confirmed by the expression of specific markers detected using immunofluorescence (endoderm, AFP and FOXA2; mesoderm, smooth muscle α-actin (SMA) and GATA4; ectoderm, TUJ1 and GFAP) (see Table S1 for the antibodies used). (D) Normal karyotype of HB-16 (46XY, left) and HB-28 iPSCs (46XX, right), analyzed at passages 15 and 7, respectively. [Color figure can be viewed at wileyonlinelibrary.com]

Silencing of the reprogramming transgenes (Oct4, Sox2, Klf4 and c-Myc) and activation of the endogenous pluripotency-associated genes (OCT4, NANOG, CRIPTO and REX1) in HB-iPSCs were confirmed as follows. Trizol-extracted total RNA and the SuperScript III Reverse Transcriptase kit (Life Technologies/Thermo Fisher Scientific, Carlsbad, CA, USA) were used for first-strand cDNA synthesis. A qPCR was then performed using SYBR Green PCR Mastermix (Life Technologies/Thermo Fisher Scientific) and the following thermocycling parameters: initial denaturation for 10 min at 95 °C and then 40 cycles of 95 °C for 15 s and 60 °C for 1 min, ending with a dissociation curve stage. Non-transduced primary cells served as the negative control, and GAPDH as the endogenous control. The regulatory sequences of interest were amplified by two subsequent PCRs using the primer pairs indicated in Table S2. The resulting amplified products were cloned using the Dual Promoter TA Cloning kit (Life Technologies/Thermo Fisher Scientific), amplified in DH5α cells, purified and sequenced. A minimum of two iPSC clones from each of the patient and his mother were selected that fulfilled all these criteria. Frozen stocks of early (< 10) and late (> 20) passage cells were stored in liquid N_2.

Short tandem repeat (STR) analysis

To confirm that the iPSCs originated from the patient and his mother, DNA from PBMCs, fibroblasts and iPSC samples was subjected to STR analysis using the AmpFLSTR Identifiler PCR amplification kit (Applied Biosystems, Foster, CA, USA) as previously described [2].

Hepatic differentiation of HB-iPSCs

Hepatic differentiation was induced using a protocol that mimics liver development *in vivo* [5], with a few modifications (Fig. 2A). Briefly, iPSC clones were cultured for 2 days in chemically defined medium supplemented with polyvinyl alcohol (CDM-PVA) (50% F12 GlutaMAX [Gibco] + 50% IMDM [Gibco] supplemented with 1% concentrated lipid + 450 μM thioglycerol + 7 μg mL^{-1} insulin + 15 μg mL^{-1} transferrin [Roche, Basel, Switzerland] + 1% Pen/Strep + polyvinyl alcohol [Sigma, Saint Louis, MO, USA] + 100 ng mL^{-1} Activin A [Peprotech] + 80 ng mL^{-1} fibroblast growth factor [FGF] 2 [R&D Systems, Minneapolis, MN, USA]), 2 days in CDM-PVA medium + 100 ng mL^{-1} Activin A + 80 ng mL^{-1} FGF2 + 10 ng mL^{-1} bone morphogenetic protein (BMP) 4 (R&D Systems) + 10 μM Ly294002 (Calbiochem, Merck Millipore, Billerica, MA, USA) + 3 μM Chir99021 (Stemgent, Miltenyi Biotec, Bergisch Gladbach, Germany) and 1 day without Chir99021. Definitive endoderm differentiation was then specified by incubation with RPMI1640/B27 (Life Technologies) + 100 ng mL^{-1} Activin A + 80 ng mL^{-1} FGF2 for 1 day and RPMI/B27 + 50 ng mL^{-1} Activin A for 2 additional days. Finally, the differentiation of hepatoblasts was driven by incubation in RPMI1640/B27 + 20 ng mL^{-1} BMP4 + 10 ng mL^{-1} FGF10. Cells were passaged on collagen I-coated plates on day 11 and cultured in a hepatocyte culture medium (Lonza, Levallois, France) + 20 ng mL^{-1} hepatocyte growth factor (HGF) (Peprotech) + 20 ng mL^{-1} Oncostatin M (Peprotech) + 0.1 mM dexamethasone (Sigma) until day 27.

Immunofluorescence of iPSC-derived HLCs

iPSC-derived cells were washed twice with PBS and fixed with 4% paraformaldehyde (Euromedex, Souffelweyersheim, France) for 15 min at room temperature (RT) followed by three washes with PBS. The fixed cells were permeabilized with 0.1% Triton X-100 for 15 min and blocked for 1 h with 3% bovine serum albumin (BSA) (Sigma) in PBS. The cells were then incubated with the primary antibodies (Table S1) overnight at 4 °C in the presence of 1% BSA in PBS. After three washes with PBS, they were incubated with the secondary antibodies (1 : 1000) in the dark with agitation for 1 h at RT. Cell nuclei were visualized by DAPI staining (0.5 μg mL^{-1}) (Sigma), mounting the cells on glass slides with Faramount Aqueous mounting medium (Dako/Agilent Technologies, Santa Clara, CA, USA). Images were obtained using the EVOS FL Auto microscope and its software (Thermo Fisher Scientific).

RNA isolation, RT-PCR and qPCR

Total RNA was isolated using the RNeasy Mini kit (Qiagen, Hilden, Germany) or Trizol (Invitrogen, Rockville, MD, USA) according to the manufacturers' recommendations. HB-iPSCs and iPSC-derived HLCs were also tested by RT-PCR to analyze the expression of specific markers during hepatic differentiation (days 0, 5, 11 and 21,
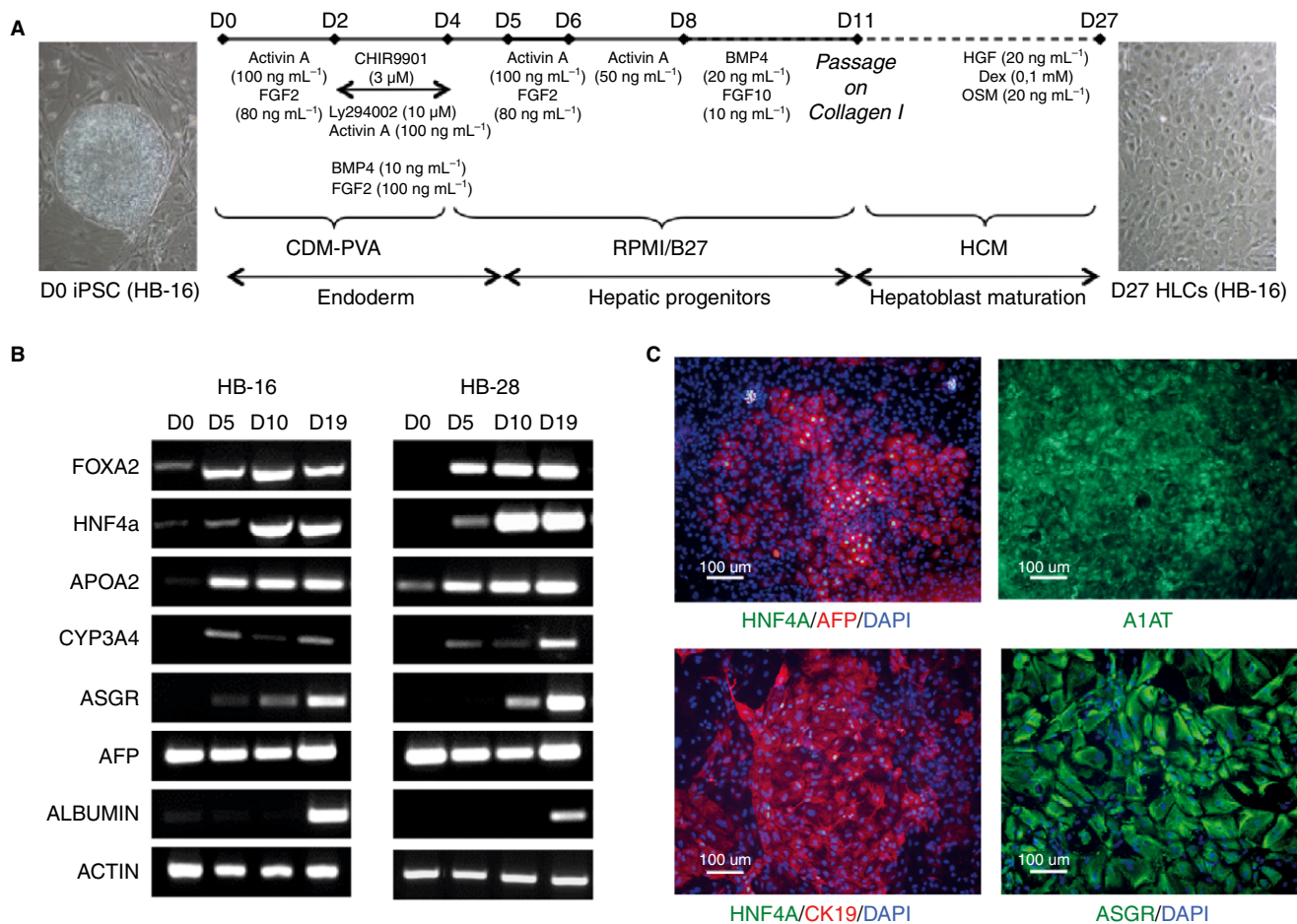
Fig. 2. Characterization of hepatocyte-like cells (HLCs) derived from hemophilia B (HB)-16 and HB-28 induced pluripotent stem cell (iPSC) clones. (A) Protocol used for iPSC differentiation into HLCs. Phase-contrast microscopy shows the morphology of the cells at days 0 (left) and 27 (right) of differentiation. (B) RT-PCR analyses showing the expression pattern during hepatic differentiation of both HB-16 and HB-28 iPSC clones. (C) Expression of hepatocyte markers analyzed by immunofluorescence in patient's (HB-16-derived) HLCs. The same hepatocyte markers were observed with HB-28 HLCs (data not shown). AFP, α-fetoprotein; APOA2, apolipoprotein A2; ASGR, asialoglycoprotein receptor; A1AT, α1-anti-trypsin; CK19, cytokeratin 19; CYP3A4, cytochrome P450 3A4; FOXA2 and HNF4a, hepatocyte transcription factors. [Color figure can be viewed at wileyonlinelibrary.com]

corresponding to the pluripotent state, endoderm, hepatic progenitors and mature hepatocytes, respectively). A total of 500 ng of total RNA extracted with the RNeasy Mini Kit (Qiagen) in a final volume of 25 μL were retrotranscribed using the High-Capacity cDNA reverse transcription kit (Thermo Fisher Scientific), according to the manufacturer's recommendations. An equal amount of cDNA was used to establish the expression of OCT4, NANOG, GATA4, SOX17, FOXA2, AFP, HNF4, ALB, CK18, CK19, CYP34A and F9 (primers listed in Table S2).

For quantitative analyses of factor *F9* mRNA in HLCs, RT-qPCR was performed using TaqMan Universal Master Mix II, with uracil n-glycosylase (Thermo Fisher Scientific) according to the manufacturer's instructions. GAPDH (4333764F) served as the endogenous control and *F9* mRNA was detected in a Taqman assay (Hs_01592597_m1).

F9 mRNA sequencing

Total RNA was retrotranscribed using the High Capacity cDNA reverse transcription kit (Thermo Fisher Scientific) as described above. The primers for massive sequencing were designed using the *F9* cDNA (GenBank no. NM_000133.3) to amplify the affected region of *F9* mRNA, resulting in a 340-bp amplicon (A region) (Fig. 3A and Table S2). Cells from the human hepatocellular carcinoma cell line Huh-7 or adult primary hepatocytes (obtained from liver biopsies) served as positive controls. After cDNA synthesis, the three amplicons described in Fig. 3A were amplified using the FastStart High Fidelity PCR system (Roche Diagnostics, Basel, Switzerland) with an initial denaturation at 94 °C for 3 min, 30 cycles at 94 °C for 30 s, 65 °C for 30 s, and 72 °C for 1 min, followed by a final extension at 72 °C for 3 min. The PCR products were specifically indexed,

separated on a 1% agarose gel, and visualized by Syber® Safe staining (Thermo Fisher Scientific). They were then quantified with the Qubit® dsDNA BR assay kit in a Qubit fluorometer, normalizing the amounts using the SequalPrep™ normalization plate kit (Thermo Fisher Scientific). The normalized products were sequenced in a Miseq sequencer using a Miseq reagent Nano kit v2 (Illumina, San Diego, CA, USA). The obtained sequences were assembled and aligned with the consensus wt F9 cDNA sequence using the software CLC Genomics Workbench (Qiagen).

X chromosome inactivation analysis

The pattern of X inactivation was determined using the HUMARA assay with some modifications [6]. Briefly, genomic DNA was digested overnight with the methylation-sensitive Hpa II endonuclease, coupled with PCR analysis using FAM-fluorescently labeled primers. The amplified products were separated using an ABI 3130 genetic analyzer (Applied Biosystems) as previously described [7]. The fragment size was estimated by comparison with the internal size standard Gene Scan ROX (Applied Biosystems).

Results

Patients, iPSC generation and characterization

Skin biopsies were taken from the patient and his mother and the subsequently isolated fibroblasts were reprogrammed using γ-retroviral vectors encoding the four Yamanaka factors. Among all the colonies that were generated and isolated, representative colonies from the patient and the mother were selected, amplified and fully characterized as iPSC clones [3,4].

The stemness of the iPSCs was confirmed by (i) positive staining for AP (Fig. 1A); (ii) immunofluorescence-based detection of the expression of stem cell markers (Fig. 1A, Table S1); (iii) silencing of the transduced genes and induction of endogenous pluripotency genes (Fig. 1B); and (iv) the presence of early derivatives of the three germ layers after spontaneous differentiation *in vitro* (Fig. 1C). HB-iPSCs maintained a normal karyotype (Fig. 1D) and short tandem repeat (STR) analysis confirmed the origin of the iPSC clones (Table S3). One iPSC clone for the patient and one for the mother, named HB-16 and HB-28, respectively, were selected for further studies.

iPSC differentiation into HLCs

Hepatocytes are the only cells of the body that express full-length *F9* mRNA and produce functional FIX. HB-16 and HB-28 iPSCs were thus differentiated into HLCs by adapting a protocol mimicking normal liver development, as previously described [5], (Fig. 2A). During this process, cells were first differentiated into definitive endoderm cells expressing FOXA2 at day 5 (Fig. 2B); then they were committed into bipotent hepatic progenitors co-expressing hepatocytic and biliary markers (HNF4a, AFP, APOA2 and CK19) at day 10 (Fig 2B, C, left panels). Finally, at day 27, cell morphology closely resembled that of hepatocytes (Fig. 2A, right picture) and HLCs expressed mature hepatocyte markers such as CYP3A4, ASGR, albumin and A1AT (Fig 2B, C, right panels). We thus amplified exons 2–8 of *F9* RNA, a region that can only be amplified using hepatocyte- or HLC-derived RNA that covers the patient's mutation, from both HB-16 and HB-28-HLC RNA.

Molecular analyses of F9 mRNA

RT-qPCR of *F9* mRNA from HLCs and comparisons with primary hepatocyte *F9* mRNA levels showed a relative expression level of 36.9% ± 2.7 in the mother's and 2.1% ± 0.7 in the patient's HLCs (Fig. 3B). The incomplete differentiation/maturation into hepatocytes may account for the low level of *F9* mRNA expression in the HB-28 HLCs compared with primary hepatocytes, but it does not explain its much lower expression in the HB-16 HLCs. Indeed, HLCs from both the patient and the mother exhibited the same pattern of differentiation markers with no striking differences between them. However, nonsense-mediated decay (NMD), a cellular mechanism that degrades PTC-containing mRNAs, could explain the very low level of *F9* mRNA observed in the HB-16 HLCs that express only the mutated transcript. Nevertheless, the interpretation of the level of *F9* mRNA expression in HB-28 HLCs must take into account from which X chromosome the *F9* transcripts are issued. Indeed, the X-chromosome inactivation pattern (XCIP) could have been modified during reprogramming of the mother's skin fibroblasts into HB-28 iPSCs [8]. Thus, depending on the actual XCIP pattern, either or both of the normal and mutated *F9* transcripts could be expressed in the HB-28 HLCs. To elucidate this point and to deeply analyze the sequence of the mutated transcript, we performed NGS analysis. We used three different sets of primers to amplify *F9* mRNA by RT-PCR in the HLCs (Fig. 3A, C).

Amplicon sequencing showed the creation of an abnormal splicing site in the vast majority of the transcripts from HB-16 HLCs that resulted in the inclusion of two additional nucleotides to *F9* mRNA in the exon 3/exon 4 boundary, leading to a frameshift and the generation of a PTC (Fig. 3A), as previously predicted [1]. However, the remaining 2.9% of the reads displayed the wt sequence (Table 1, Fig. 3D), suggesting some normal splicing of the mutated transcript and the existence of a residual FIX expression, although very limited if we take into account that the *F9* mRNA expression was very low in
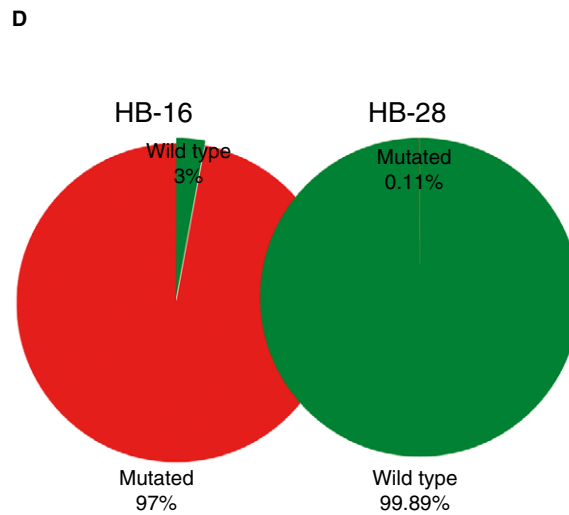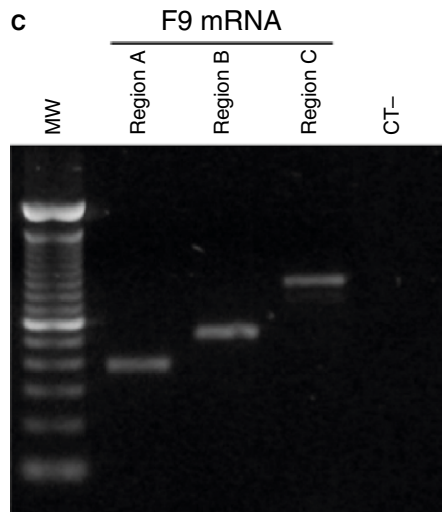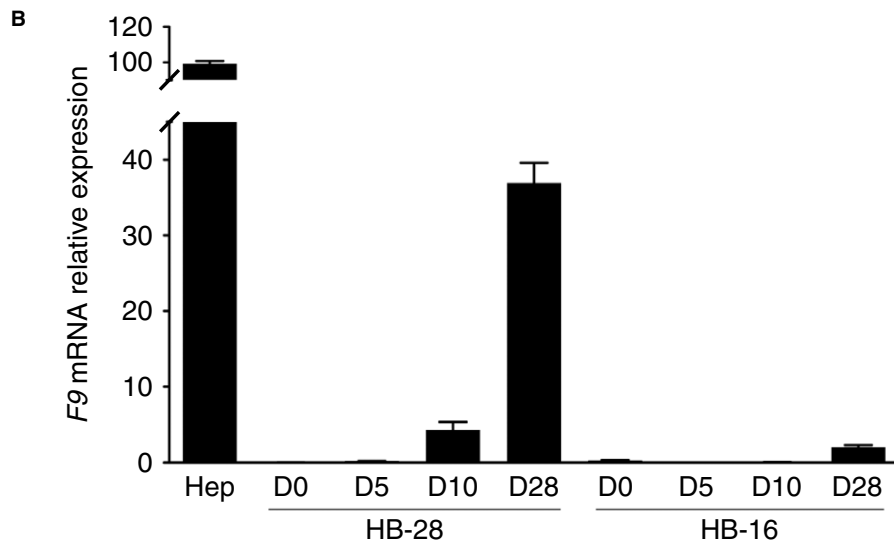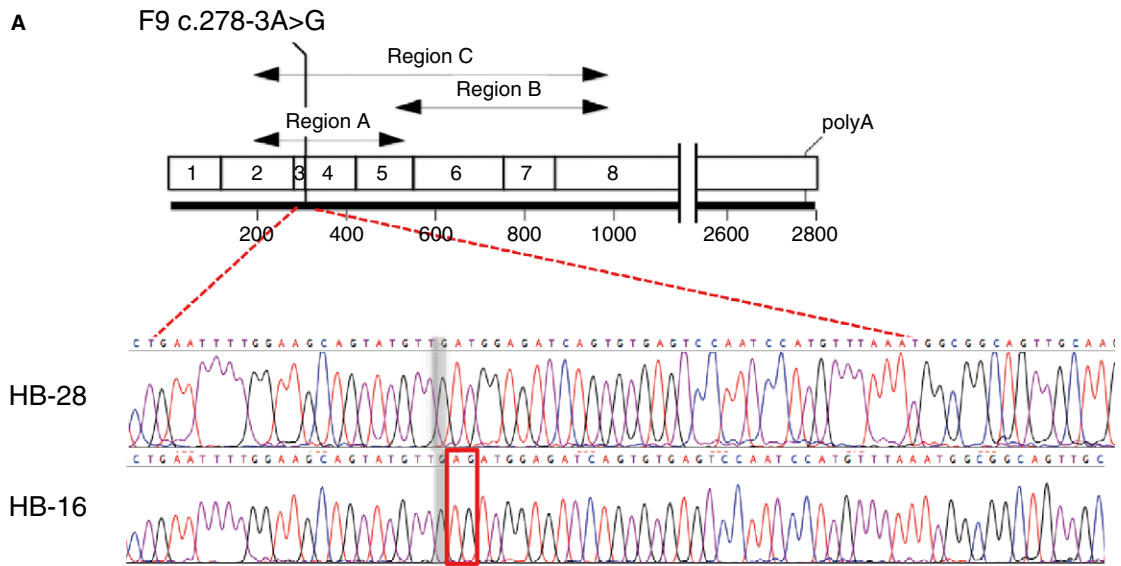
A F9 c.278-3A>G

Region C

Region B

Region A

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | polyA |

200 400 600 800 1000 2600 2800

HB-28

CTGAATTTTGGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAG

HB-16

CTGAATTTTGGAAGCAGTATGTTGAGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGC

B

F9 mRNA relative expression

Hep | D0 | D5 | D10 | D28 | D0 | D5 | D10 | D28

HB-28 | HB-16

C F9 mRNA

MW | Region A | Region B | Region C | CT−

D

HB-16
Wild type 3%
Mutated 97%

HB-28
Mutated 0.11%
Wild type 99.89%

Fig. 3. *F9* mRNA expression in induced pluripotent stem cell (iPSC)-derived hepatocyte-like cells (HLCs). (A) Location of the mutation (F9 c278-3A>G) and the regions amplified by RT-PCR (region A, from exons 2–5: 340 bp; region B, from exons 5–8: 478 bp; region C, exons 2–8: 797 bp), and Sanger sequencing showing inclusion of the AG dinucleotide (red box) at the exon 3/exon 4 boundary, as a result of an abnormal splicing. (B) Relative *F9* mRNA levels were measured by RT-qPCR during cell differentiation using the Taqman assay (Hs_01592597_m1) and GAPDH (4333764F) as the endogenous control. The results are expressed as the mean percentages of the control levels, measured in primary human hepatocytes (*n* = 3). *F9* mRNA expression level in hemophilia B (HB)-16 HLCs was 2.05% ± 0.65 (*n* = 5) on day 27 of differentiation, whereas in HB-28 HLCs expression was first detectable on day 10 and reached 36.9% ± 2.71 (*n* = 5) on day 27. Error bars represent standard deviations. (C) Representative *F9* mRNA amplification in HB-16 HLCs on day 27. (D) *F9* mRNA massive sequencing. Illumina GA analysis of sequences spanning the affected exons of the *F9* mRNA provided relatively uniform sequence coverage, with a sequencing depth for each nucleotide position in the range of ~6000 to 30000 reads. The abnormal dinucleotide inclusion in the F9 mRNA resulting from the F9 c.278-3A>G mutation was the only unambiguous genetic variant identified in the analysis. The reads containing the exon 3 and 4 joining were extracted and aligned with the *F9* reference sequence. Pie charts show the percentages of mutated (red) and wt (green) amplicons from *F9* mRNA transcripts in HB-16 and HB-28 HLCs. Hep, primary human hepatocytes; MW, molecular weight, 100 bp ladder; CT-, negative control. [Color figure can be viewed at wileyonlinelibrary.com]

HB-16 HLCs. *F9* mRNA analysis of HB-28 HLCs indicated an extremely low percentage (0.1%) of mutated transcripts (Table 1), which led us to investigate the XCIP in the HB-28 iPSCs and the derived HLCs. Analyses using the HUMARA assay showed that the X-chromosome carrying the mutation was inactivated in the HB-28-iPSC clone and in all derived HLCs (Figure S1). This result explains the nearly exclusive presence of wt *F9* mRNA in HB-28 HLCs. Moreover, this result supports the hypothesis of the degradation of the mutated transcript by NMD in the HB-16 HLCs to explain the reduced level of *F9* mRNA in those cells compared with HB-28 HLCs. Finally, by combining the proportion of *F9* mRNA detected in HB-16 HLCs with respect to the levels observed in primary hepatocytes (about 2%) and the percentage of this mRNA that we found correctly spliced (about 3%), despite the limitations of this cellular model in mimicking the RD, we can roughly estimate that the total wt *F9* mRNA level in these RD HLCs is around 0.06% (3% of correctly spliced mRNA out of the 2% of the total *F9* mRNA level with respect to that of normal hepatocytes; see Table 1), which is consistent with the patient's severe phenotype.

Discussion

The present work describes the creation of a cellular model of HB, based on the iPSC paradigm. In this case, it allowed us to generate patient-specific HLCs carrying the RD mutation, to isolate hepatocyte-specific *F9* mRNA and to fully characterize the splicing mutation by NGS in a more physiological context. It also represents the first definitive demonstration, in the absence of

currently living descendants, that the F9 c.278-3A>G mutation is truly responsible for the hemophilic syndrome in the affected members of the European royal families. In addition, these results represent advances in two fields: first, it is the first time that HB has been modeled using patient-specific iPSCs, and second, it allows a more physiological (and thus more relevant) characterization of this mutation in comparison with that of a previous report that could only predict some of the consequences of this splicing mutation at the RNA level, as it was based on an artificial minigene-based system. The present work, which relies on the use of HLCs expressing *F9* mRNA, confirms the previously predicted consequences of the mutation at the RNA level, and provides novel information that could not be anticipated with the previously used approach. Indeed, in the case of intronic splicing mutations such as this one, *F9* transcripts must be analyzed, and this requires either primary cells of the relevant tissue (e.g. hepatocytes) or HLCs generated from the patient's iPSCs. However, in our case, because primary hepatocytes could not be obtained, the only way to get a real and meaningful picture of what occurs in the patient's hepatocytes was to differentiate the patient's iPSCs into HLCs. It is not possible to draw definitive conclusions on the reality of the abnormal splicing unless the entire mutated F9 gene is expressed in hepatocytes or HLCs.

These results also provide new knowledge that may help in understanding the long-discussed issue of the unexpectedly high longevity (given the severity of the RD and the absence of existing replacement therapies at that time) of certain affected members of the European royal families. Indeed, one of the most intriguing aspects of the disease is that the average life expectancy of some

Table 1 NGS analysis of *F9* mRNA amplicons

	% of *F9* mRNA (with respect to HPHs)	Total # of reads	# of mutated reads	% of mutated reads	# of wild-type reads	% of wild-type reads	% of wild-type *F9* mRNA/HPHs (column 2 × column 7)
HB-16 (patient)	2.1 ± 0.7	434225	421546	97.1	12679	2.9	0.06
HB-28 (mother)	36.9 ± 2.7	1049194	1049	0.1	1038702	99.9	36.86

HPHs, human primary hepatocytes.

affected European royalty members was clearly longer than that of contemporary or even more recently documented untreated patients with severe HB, estimated to be 11-25 years [9,10]. Indeed, although some affected members died at a very early age, others lived for several decades, such as Prince Leopold (1853–1884), Prince Waldemar of Prussia (1889–1945) or Alfonso and Gonzalo de Borbon (1907–1933 and 1914–1934, respectively, who were killed in car accidents).

These observations of historical records, along with previous reports showing a correlation between residual expression of *F8* mRNA and hemophilia A (HA) severity [11], suggest the possibility that some affected individuals might have suffered a less severe form of the disease. However, considering (i) the FIX activity (< 1%) at the time of diagnosis in the HB-16 patient and (ii) the fact that the F9 c.278-3A>G mutation was also reported by others as severe [12–14], it seems unlikely that the very low residual level of wt *F9* mRNA that we detect could have resulted in a moderate phenotype in the affected members of the royal families, unless there are other sources of variability. The confirmation of the predicted splicing mutation and the finding that mutated *F9* mRNA could be targeted by NMD, represent novel variables supporting this notion, as it has been shown that variability in the efficiencies of splicing [15,16] and NMD [17] machineries (e.g. in the expression levels or the existence of more or less active variants of the involved proteins) can result in huge inter-individual phenotypic differences. Indeed, a study carried out on 10 patients with cystic fibrosis and the 3849 + 10 kb C>T mutation showed a high percentage of variability (0–28%) in the levels of aberrantly spliced RNA transcribed in their nasal epithelia [15]. Thus, it is tempting to speculate, although this is merely theoretical, that these and/or other additional genetic traits might have attenuated the bleeding phenotype in some affected individuals of the royal families, although other socio-economic considerations, including special care received by these privileged individuals, may also have played an important role.

In summary, we report the first iPSC-based cellular model of HB. This technology allowed us to establish a *bona fide* HB model to characterize rare mutations, such as the one causing the RD, and we propose its use as a powerful tool to elucidate mutational mechanisms at the RNA level and to investigate correction strategies by gene therapy, genome editing and/or drugs targeting specific mutations.

Addendum

L. Martorell performed differentiation experiments and molecular analyses and co-wrote the manuscript. J. L. Vazquez, Y. Richaud-Patin, S. Jimenez-Delgado and A. Raya generated and characterized iPSCs. E. Luce, A. Weber and A. Dubart-Kupperschmitt developed the HLCs differentiation protocols, generated and characterized the HLCs and contributed to the writing of the manuscript. I. Corrales contributed to the NGS experiments. N. Borras contributed to the lyonization experiments. S. Casacuberta-Serra contributed to the differentiation experiments. R. Parra and C. Altisent recruited the patients and contributed to data analysis. A. Follenzi participated in the design of the study and in the discussion of the results. A. Dubart-Kupperschmitt, A. Raya, F. Vidal and J. Barquinero equally contributed to the design, obtaining of funding and supervision of the different parts of the study. J. Barquinero co-wrote the manuscript and coordinated the study.

All authors read and approved the final manuscript.

Disclosure of Conflict of Interests

The authors state that they have no conflict of interest.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. List of primary antibodies used for iPSC and differentiated cell characterization.
Table S2. List of primers used for amplification of pluripotency-associated mRNAs, liver-specific mRNAs, *F9* mRNA, and DNA methylation analyses.
Table S3. STR analysis.
Fig. S1. X-chromosome inactivation (XCI) analysis.

References

1 Rogaev EI, Grigorenko AP, Faskhutdinova G, Kittler EL, Moliaka YK. Genotype analysis identifies the cause of the "royal disease". *Science* 2009; **326**: 817.

2 Ramirez L, Altisent C, Parra R, Vidal F. The 'royal disease' mutation in a Spanish patient. *J Thromb Haemost* 2010; **8**: 2316–7.

3 Raya A, Rodriguez-Piza I, Navarro S, Richaud-Patin Y, Guenechea G, Sanchez-Danes A, Consiglio A, Bueren J, Izpisua Belmonte JC. A protocol describing the genetic correction of somatic human cells and subsequent generation of iPS cells. *Nat Protoc* 2010; **5**: 647–60.

4 Aasen T, Raya A, Barrero MJ, Garreta E, Consiglio A, Gonzalez F, Vassena R, Bilic J, Pekarik V, Tiscornia G, Edel M, Boue S, Izpisua Belmonte JC. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat Biotechnol* 2008; **26**: 1276–84.

5 Yang G, Si-Tayeb K, Corbineau S, Vernet R, Gayon R, Dianat N, Martinet C, Clay D, Goulinet-Mainot S, Tachdjian G, Tachdjian G, Burks D, Vallier L, Bouille P, Dubart-Kupperschmitt A, Weber A. Integration-deficient lentivectors: an effective strategy to purify and differentiate human embryonic stem cell-derived hepatic progenitors. *BMC Biol* 2013; **11**: 86.

6 Allen RC, Zoghbi HY, Moseley AB, Rosenblatt HM, Belmont JW. Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am J Hum Genet* 1992; **51**: 1229–39.

7 Cram DS, Ma K, Bhasin S, Arias J, Pandjaitan M, Chu B, Audrins MS, Saunders D, Quinn F, deKretser D, McLachlan R. Y chromosome analysis of infertile men and their sons conceived through intracytoplasmic sperm injection: vertical transmission of deletions and rarity of de novo deletions. *Fertil Steril* 2000; **74**: 909–15.

8 Pasque V, Plath K. X chromosome reactivation in reprogramming and in development. *Curr Opin Cell Biol* 2015; **37**: 75–83.

9 Konkle BA, Huston H, Nakaya Fletcher S. Hemophilia B. *GeneReviews® [Internet]* In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJH, Bird TD, Ledbetter N, Mefford HC, Smith RJH, Stephens K, eds. Seattle, WA: University of Washington, 1993–2017.

10 Miguelino MG, Powell JS. Clinical utility and patient perspectives on the use of extended half-life rFIXFc in the management of hemophilia B. *Patient Prefer Adherence* 2014; **8**: 1073–83.

11 Liang Q, Xiang M, Lu Y, Ruan Y, Ding Q, Wang X, Xi X, Wang H. Characterisation and quantification of F8 transcripts of ten putative splice site mutations. *Thromb Haemost* 2015; **113**: 585–92.

12 Ketterling RP, Drost JB, Scaringe WA, Liao DZ, Liu JZ, Kasper CK, Sommer SS. Reported in vivo splice-site mutations in the factor IX gene: severity of splicing defects and a hypothesis for predicting deleterious splice donor mutations. *Hum Mutat* 1999; **13**: 221–31.

13 Kwon MJ, Yoo KY, Kim HJ, Kim SH. Identification of mutations in the F9 gene including exon deletion by multiplex ligation-dependent probe amplification in 33 unrelated Korean patients with haemophilia B. *Haemophilia* 2008; **14**: 1069–75.

14 Miller CH, Benson J, Ellingsen D, Driggers J, Payne A, Kelly FM, Soucie JM, Craig Hooper W, Hemophilia Inhibitor Research Study I. F8 and F9 mutations in US haemophilia patients: correlation with history of inhibitor and race/ethnicity. *Haemophilia* 2012; **18**: 375–82.

15 Chiba-Falek O, Kerem E, Shoshani T, Aviram M, Augarten A, Bentur L, Tal A, Tullis E, Rahat A, Kerem B. The molecular basis of disease variability among cystic fibrosis patients carrying the 3849 + 10 kb C–>T mutation. *Genomics* 1998; **53**: 276–83.

16 Nissim-Rafinia M, Kerem B. The splicing machinery is a genetic modifier of disease severity. *Trends Genet* 2005; **21**: 480–3.

17 Nguyen LS, Wilkinson MF, Gecz J. Nonsense-mediated mRNA decay: inter-individual variability and human disease. *Neurosci Biobehav Rev* 2014; **46**(Pt 2): 175–86.