Supplementary Information

Supplementary Materials and Methods

Micro-CT imaging

The human foetal control was selected from the Foetal Biobank, located at Amsterdam University Medical Center (UMC), location AMC. This foetus of 13 weeks of development was donated to the Foetal Biobank after termination of pregnancy. All donations to the Foetal Biobank are completely anonymized and require written maternal informed consent. Research is approved by the Medical Ethical Committee (MEC) and Biobank Committee (BC) of the Amsterdam University Medical Centers, location AMC, Amsterdam. After acquisition, the foetus was fixed in 4% Paraformaldehyde (PFA) in PBS for 4 days at 4°C and subsequently stored in 0.2% PFA in PBS at 4°C until usage. The specimens were immersed in Lugol's solution (I2KI; potassium triiodide) to increase soft-tissue visualization on micro-CT. We used a 3.75% weight/volume Lugol's solution and stained for 72 hours, following previously published protocols (Dawood et al., 2020). Immediately prior to imaging, the foetus was removed from the Lugol's solution, rinsed in water, and dried with gauze. The imaging was acquired using an GE Phoenix v|tome|x m tomographer (General Electric, Wunstorf, Germany) micro-CT scanner. Scanning parameters included X-ray energies and beam current at 180 kV and 150 µA, respectively. Exposure time was 250 milliseconds, with in total 1500 projections, with 3 x-ray frames per projection. GE reconstruction software (Wunstorf, Germany) was used to calculate the 3D structure via back projection. The isotropic voxel size was 30 micrometer. For image visualization and animation, the image files were imported in Amira v2019.3 (Thermo Fisher Scientific, Massachussets, U.S.).

Derivation of iPSC from fibroblasts

Fibroblasts isolated from punch skin biopsies were reprogrammed using the CytoTuneTM-iPS 2.0 Sendai Reprogramming Kit (Thermo Fisher Scientific, #A16517) or using retroviral vectors encoding human *KLF4*, *SOX2*, *OCT4*, and *C-MYC* cDNA (Addgene) and grown in knockout DMEM with20% knockout serum replacement, 2 mM L-glutamine, 1% NEAA, 0.1 mM 2-mercaptoethanol, 0.1% P/S, and 4 ng/ml bFGF on γ-irradiated fibroblast feeders. Cells were passaged and plated onto γ-irradiated mouse embryonic feeders 7 days after viral transduction. hiPSC colonies were selected and picked according to morphology between days 17 to 28 and maintained in DMEM/F12 (Sigma, #D6421) supplemented with 20% Knock Out Serum Replacement (Thermo Fisher Scientific, #10828-028), 0.1 mM 2-mercaptoethanol (Thermo Fisher Scientific, #21985-023), 2 mM L-glutamine (Thermo Fisher Scientific, #25030), 0.2 mM NEAA (Thermo Fisher Scientific, #11140-050) and 5 ng/ml bFGF (Peprotech, #100-18B). The following primary antibodies were used to confirm pluripotency: NANOG (1:200, R&D Systems, #AF1997), OCT4 (1:200, Santa Cruz, #111351), SOX2 (1:200, R&D Systems, #MAB2018), SSEA3 (1:50, Millipore, #MAB4303), SSEA4 (1:200, Millipore, #MAB4304), TRA-1-81 (1:200, Millipore, #MAB4381) and TRA-1-60 (1:200, Millipore, #MAB4360,). Alexa fluorophore-conjugated, Donkey secondary antibodies were then used for immunofluorescence (Supplementary table 6). The derivation of the hiPSC lines described in this study was sanctioned and monitored under the protocol entitled "Development of Human Disease Models Using Patients' Primary Somatic Cells and their Derivatives including Human Induced Pluripotent Stem Cells (iPSC)" initially approved by the National University of Singapore Institutional Review Board (NUS IRB 10-051), subsequently overseen by the Agency for Science, Technology and Research (A*STAR) Human Biomedical Research Office, and conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants.

CRISPR/Cas9 genome editing

Construction of Cas9/sgRNA expressing vectors and Targeting Vector: Three single guide RNAs (sgRNAs) targeting exon 19 of the RFX6 locus were identified using the CHOPCHOP web tool (https://chopchop.cbu.uib.no). Briefly, a pair of oligos was annealed, phosphorylated and cloned into BsmB1-linearized hCas9-HF1-IRES-BSD-U6 plasmid. T7E1 assay was carried out to access cleavage efficiency. gRNA 2 targeting at TTCATGGGAACAGCAGCTGG was subsequently chosen for CRISPR gene editing. The donor template was built by PCR from the H9 genomic DNA, introducing a triplicated Ha tag flanked by the homology arms. The PCR product was then cloned into the PCR blunt vector (Life Technologies). To prevent the cleavage by Cas9, site-directed mutagenesis was performed to introduce 4 silent mutations at the target site.

Development • Supplementary information

T7EI assay: H9 were transfected with either 2 µg of empty vector or sgRNA using Lipofectamine STEM (Life Technologies). DNA was harvested using the QIAamp DNA mini kit (Qiagen), 7 days post-transfection. Specific primers were used to amplify genomic regions flanking the sgRNA target sites via PCR, yielding a 571 bp product. PCR products were purified using the QIAQuick PCR purification kit, denatured and allowed to reanneal in NEB buffer 2. The PCR products were then treated with or without T7 endonuclease I (NEB) at 37°C for 30 min before they were resolved on a 2 % agarose gel and visualized using an UVIpro Gel documentation system (UVItec) (supplementary figure 4B and C).

Transfection: For transfection, cells were dissociated into small clumps using gentle cell dissociation reagent (Stem Cell Technologies) and plated onto 6cm Matrigel-coated tissue culture dishes, in mTeSR. Next day, 2 µg of sgRNA 2 and 8 µg donor template were diluted in OptiMEM, mixed gently with 25 ul of Lipofectamine STEM and incubate at room temperature for 15 min before adding dropwise to cultured cells. 24 hr post transfection, cells were refreshed with mTeSR culture medium containing 5.25 µg/ml Blasticidin S (Sigma). Selection continued for two days, after which cells were fed with fresh mTeSR medium daily until colonies were large enough for picking and genotyping.

Genotyping: Genomic DNA was extracted from picked colonies in a 20 µl reaction volume containing 1X detergent (0.05% IGEPAL CA630, 0.05% Tween-20), proteinase K and 1X TE buffer, incubated at 55°C for 1 hour followed by a 5-minute incubation at 95°C. 1 to 2 µl of the reaction was then used in a 20 µl PCR reaction using Primestar GXL (Takara) and 0.25 µM forward and reverse primers. The thermal cycling conditions used were as follows: 3 minutes at 94°C, followed by 40 cycles of 10 seconds at 98°C, 15 seconds at 55°C and 2.5 minutes at 68°C. The primers used for genotyping were designed to bind outside the 5' and 3' homology arms. A positive clone (with insertion of Triplicated HA Tag) would yield a PCR product of 2041 bp while a negative clone (wild-type) would yield a PCR product of 1960 bp. Positive clones were identified and sent for Sanger sequencing to confirm presence of Triple HA Tag without any unwanted mutations (Fig. S4D). Primer sequences can be found in Table S9.

Clonal expansion. Clone 41 with HA Tag correctly integrated into the RFX6 locus were dissociated into single cells using Accutase (Life Technologies) and seeded at very low densities onto Matrigel-coated plates in mTeSR medium supplemented with CloneR (Stem Cell Technologies), following the manufacturer's instructions. mTeSR medium was refreshed daily until colonies emerged and were large enough for picking and genotyping. Homozygous clones were expanded in culture and frozen stocks were made using Cryostor (STEMCELL Technologies).

Immunohistochemistry on Human Embryo Sections. Immunohistochemistry was performed on 5 µm sections as described previously (Jennings et al., 2013), using a guinea pig anti-PDX1 primary antibody (1:500; Abcam [ab47308]).

References for Supplementary Materials and Methods

Dawood, Y., Strijkers, G.J., Limpens, J., Oostra, R.J., de Bakker, B.S., 2020. Novel imaging techniques to study postmortem human fetal anatomy: a systematic review on microfocus-CT and ultra-high-field MRI. Eur Radiol 30, 2280–2292. doi:10.1007/s00330-019-06543-8

Jennings, R.E., Berry, A.A., Kirkwood-Wilson, R., Roberts, N.A., Hearn, T., Salisbury, R.J., Blaylock, J., Piper Hanley, K., Hanley, N.A., 2013. Development of the human pancreas from foregut to endocrine commitment. Diabetes 62, 3514–3522. doi:10.2337/db12-1479

Development • Supplementary information

Supplementary Figures

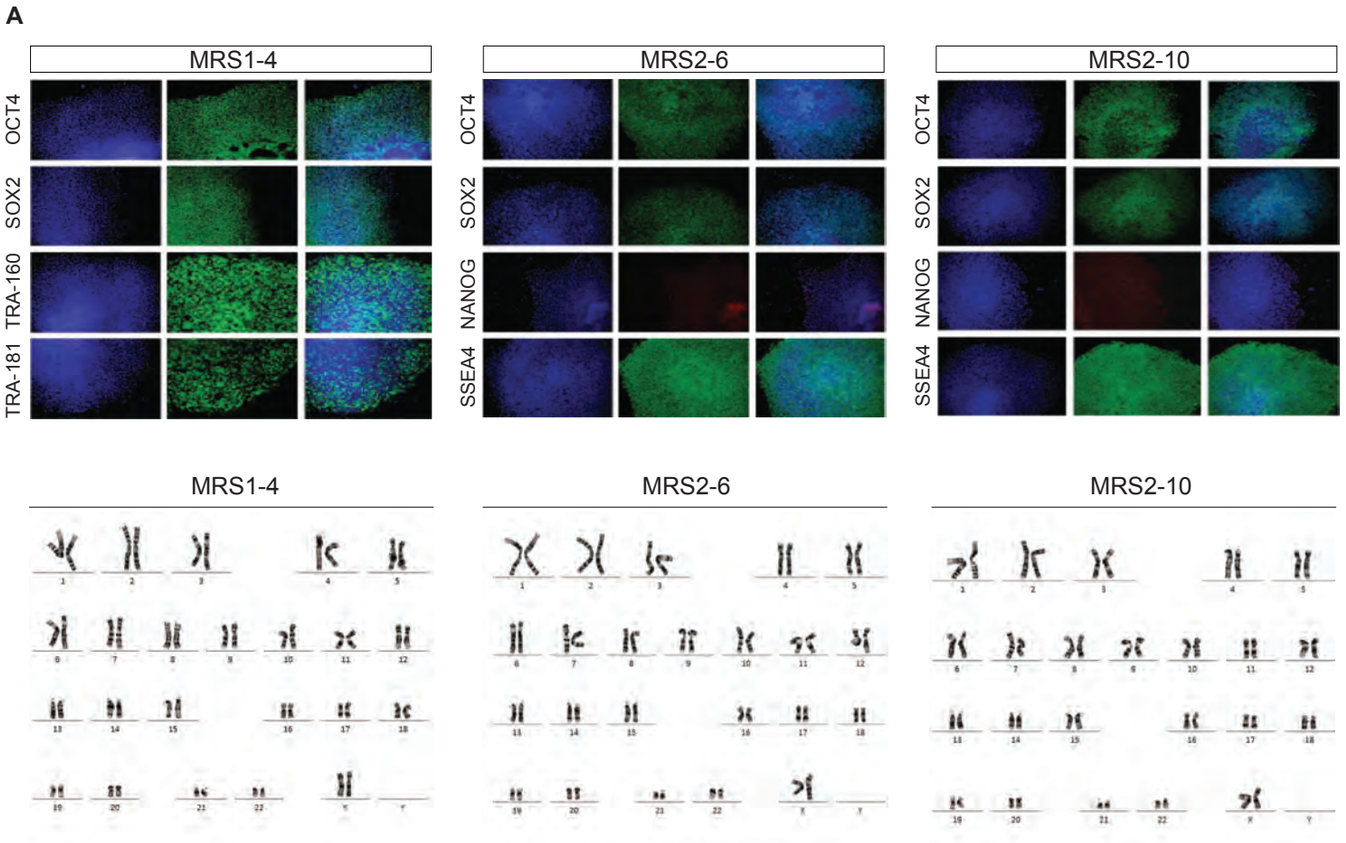
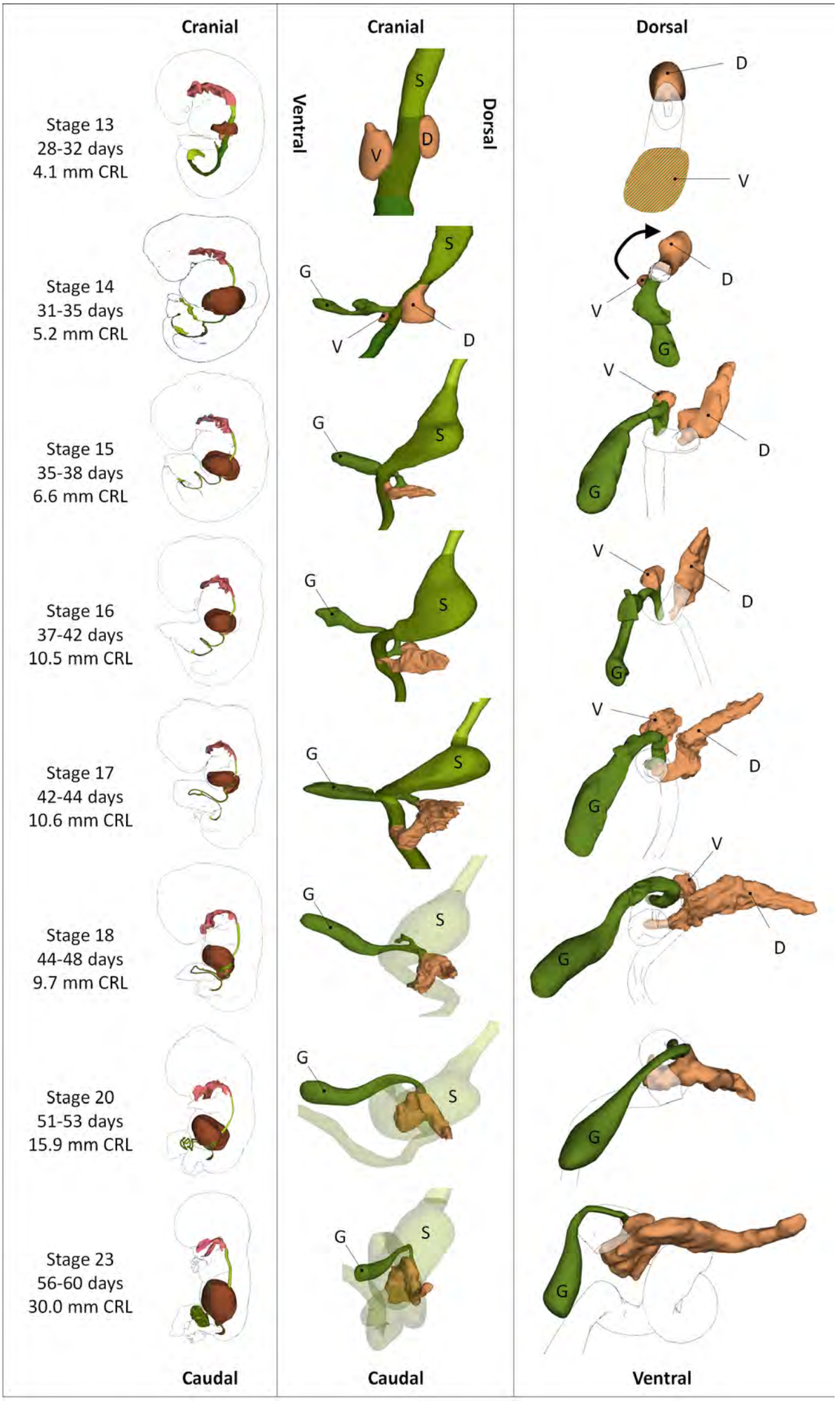


Figure S1. Characterization of patient-derived iPSC. (A) *Top*: Successful derivation of iPSC from patient fibroblasts confirmed by immunofluorescence staining for the indicated markers of pluripotency. MRS1- 4 cells were derived from patient III:3 by lentivirus in 2012, whilst MRS2-6 and MRS2-10 were derived from patient III:4 by Sendai virus in 2014 (see Supplementary Materials and Methods). *Bottom*: Chromosome spreads for the indicated iPSC lines derived from MRS patients. All iPSC lines exhibited a normal karyotype (46+XX). For brevity, data supporting the pluripotency and normal karyotype of the two father lines (F14 and F18) are not shown, but these lines were equally exhaustively characterized by the IMB Stem Cell Bank. (B) Additional micro-CT images of a wild-type control foetus (left) and the 12-week MRS foetus (right). Morphometric analyses reveal that the maximal pancreas length is 7.29 mm in the control and 3.81 mm in the MRS patient.

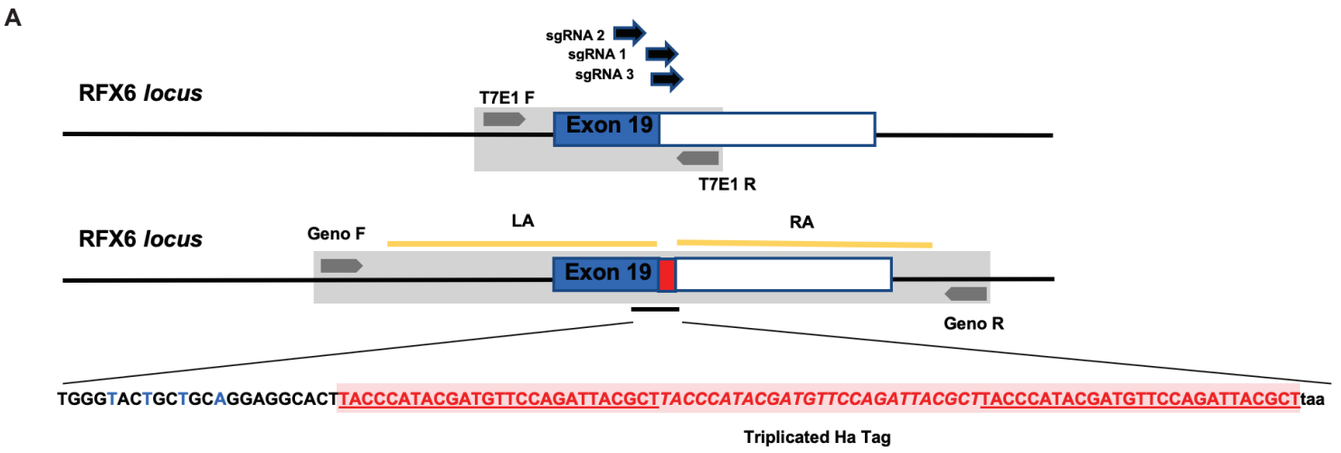
Development • Supplementary information

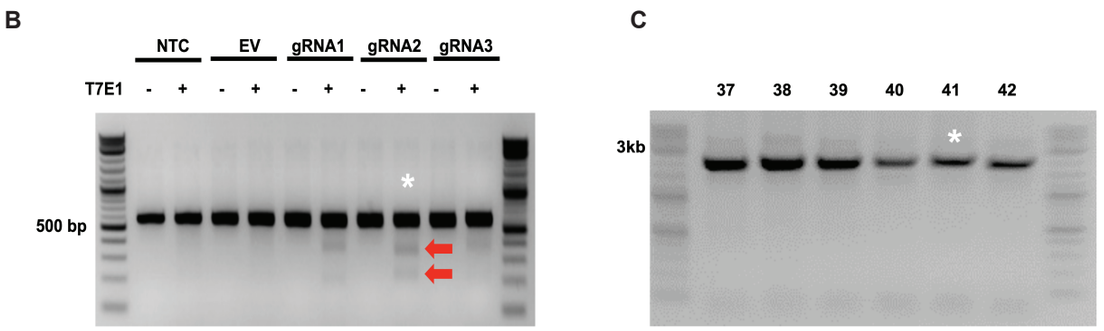


Development • Supplementary information

Figure S2. 3D overview of normal human pancreas development. Left column: Lateral view of 3D reconstructions of human embryos between Carnegie Stages (CS) 13 and 23, based on the *3D Atlas of Human Embryology* (de Bakker et al., 2016). The presented specimen numbers on the horizontal rows are: 836, 8314, 3512, 6517, 6521, 6524, 462 and 9226. The skin of the embryos in the left column is a transparent illustration which provides a view on the embryonic alimentary system including the relatively large liver (brown). As the pancreas is hidden behind the liver, the second and third column show a detailed view on the developing pancreas from the lateral and cranial side respectively. For topographic orientation part of the stomach and gut is also presented in a solid, transparent or illustrated manner. Please note that the ventral pancreatic bud (V) as depicted in stage 13 will also give rise to the gall bladder, but these two entities are at this stage inseparable. In stage 18, the ventral and dorsal pancreas are fused, but the two pancreatic ducts remain present until the foetal period (after CS23). The presented images are not scaled, but the crown-rump-length (CRL) of all embryos is provided in the left column. D: dorsal pancreatic bud, G: gall bladder, V: ventral pancreatic bud, S: stomach. The reader is encouraged to interactively study pancreatic development using **Supplementary Interactive PDF 1**.

Development • Supplementary information





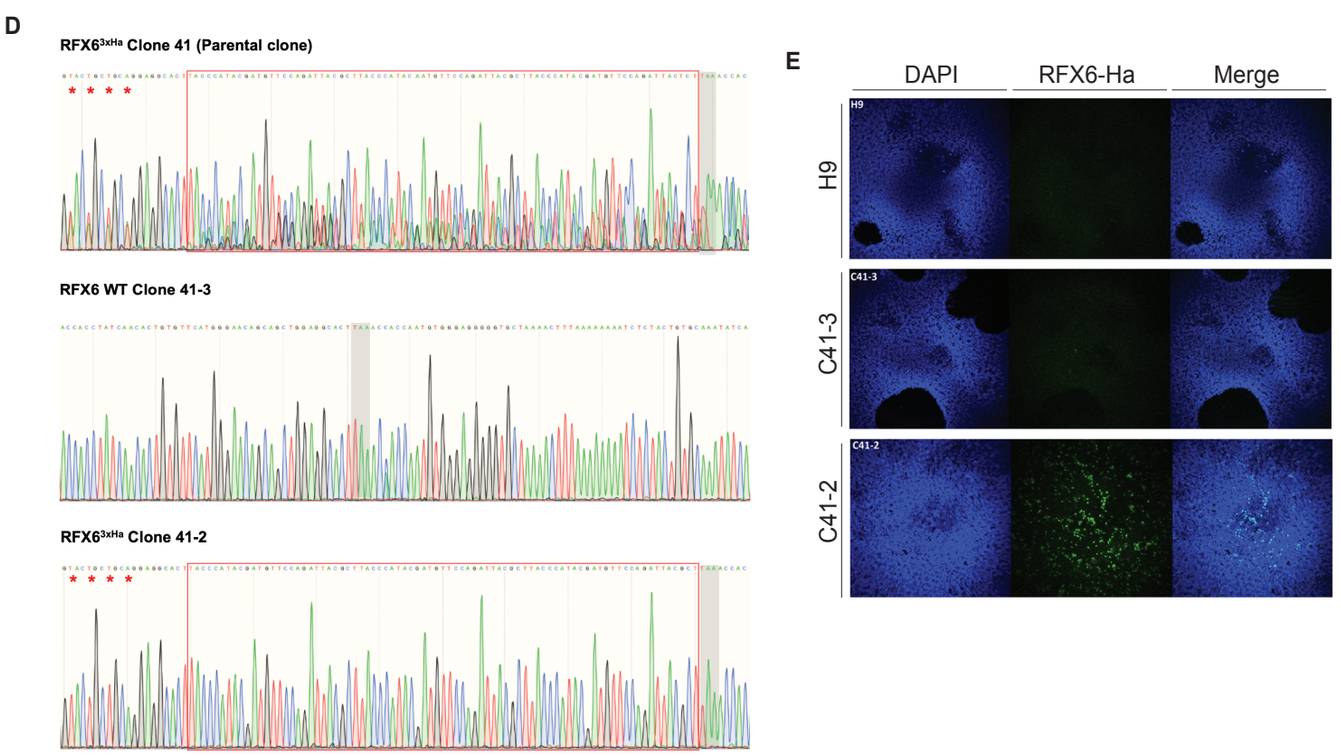


Figure S3. Introduction of a triplicated Human influenza hemagglutinin (HA) epitope into the human *RFX6* gene by CRISPR/Cas9 gene editing coupled with homologous recombination. (A) Schematic showing exon 19 of the *RFX6* gene and the targeting sites of guide (g) RNAs 1, 2 and 3; Forward (F) and Reverse (R) primers used to amplify the region around the target sites for the T7E1 assay; and Geno F and R primers used for genotyping clones. Note that the primer binding sites reside outside the 5' and 3' homology arms (LA, left arm; right arm). The sequence of the oligonucleotide used for gene targeting is indicated. (B) Gel image of T7E1 assay. Red arrows indicate the cleavage products. Wild-type band size is 571 base pairs (bp). NTC, untransfected. EV, empty vector. (C) Gel image of 6 clones. Asterisk indicates the positive clone (Clone 41) with successful integration of the triplicated HA tag (81 bp) into the *RFX6* locus. Note the slight mobility shift. (D) Sanger sequencing chromatograms of parental Clone 41 (heterogenous) and its subclones 41-3 (wild-type) and 41-2 (homozygous HA Knock-in). The asterisk indicates the 4 silent mutations and the box shows the triplicated HA epitope tag introduced before the stop codon. (E) Differentiation of WT and *RFX6*$^{HA/HA}$ H9 cells into pancreatic progenitors using the STEMdiff Pancreatic Progenitor Kit. In this instance, the differentiation was carried out for 15 days. Immunostaining for RFX6 with a rabbit anti-HA monoclonal antibody.

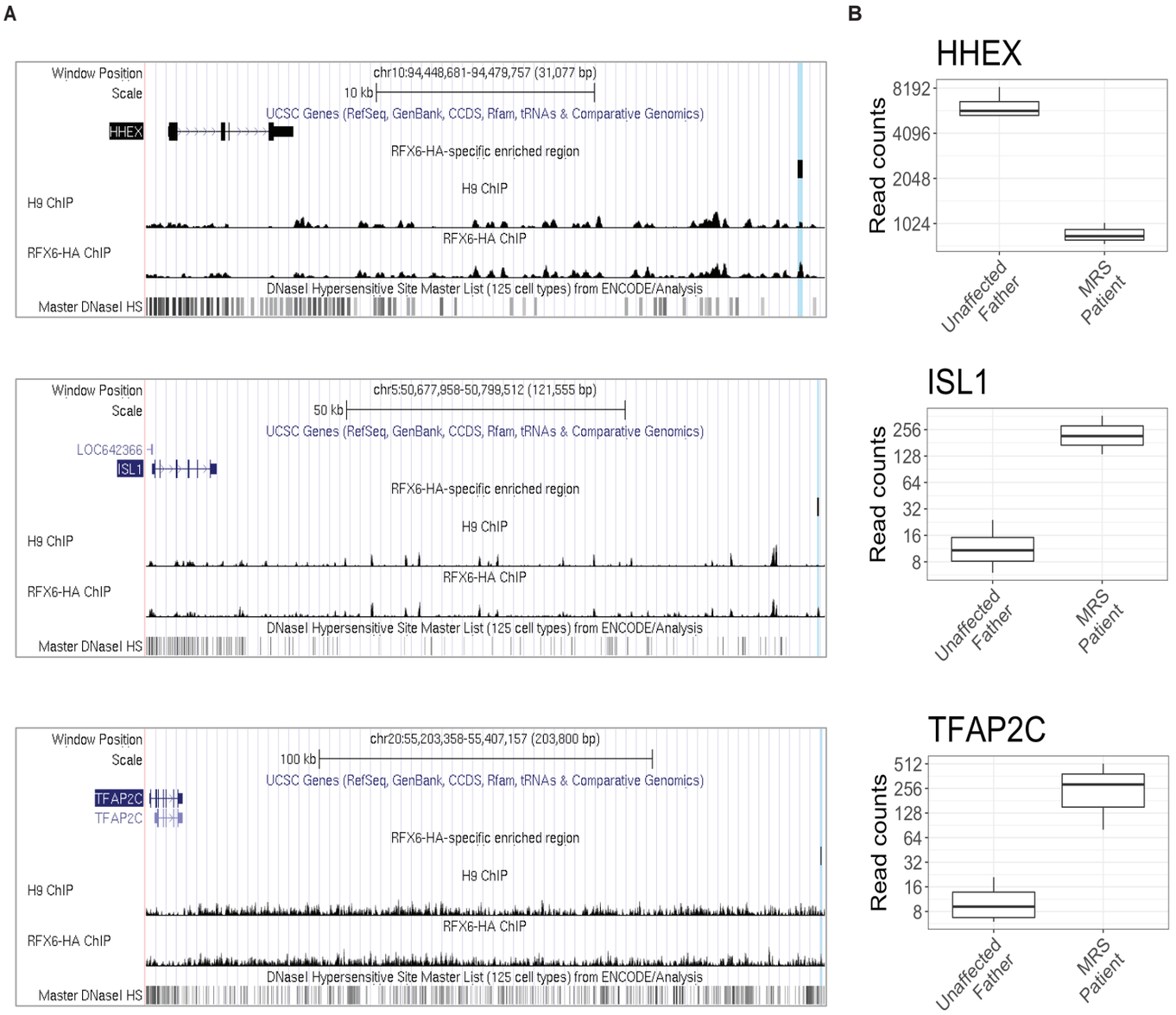
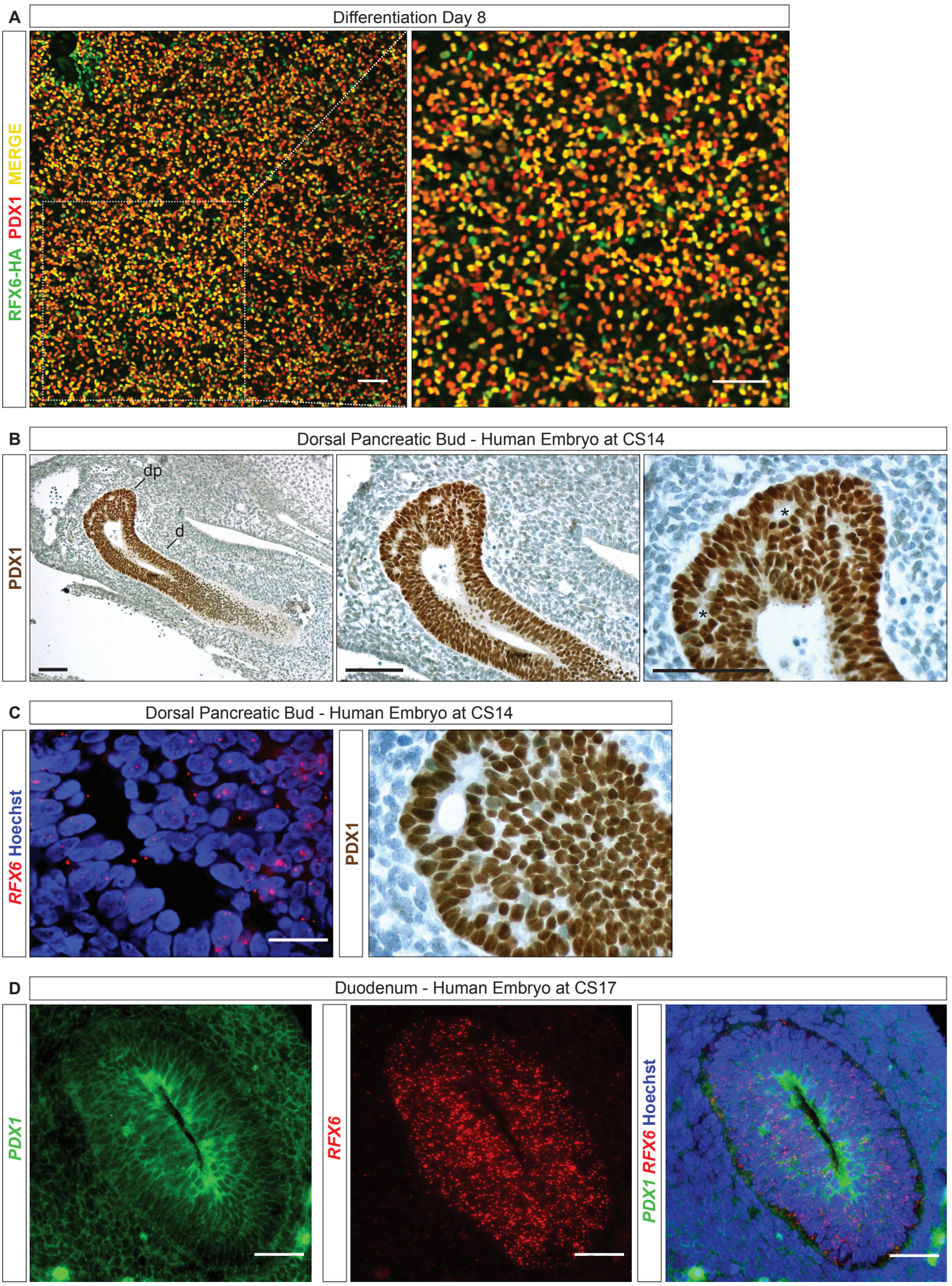


Figure S4. (A) UCSC genome browser visualisation of RHX6-HA-specific peaks (highlighted in blue) associated with the *HHEX*, *ISL1* and *TFAP2C* genes. DNaseI hypersensitivity sites are shown below each track. (B) RNA-seq read counts for genes in (A) in unaffected father and MRS patient samples.



Development • Supplementary information

Figure S5. Additional images of *RFX6* expression *in vitro* and *in vivo* to support Figure 6. (A) PDX1 and RFX6-HA immunofluorescence on day 8 of differentiation. Images in Fig. 6A are enlarged to more easily view PDX1 +;RFX6+ and PDX1+;RFX6- cell populations. (B,C) Sagittal sections of the dorsal pancreatic bud in a human embryo at CS14 and (D) of the duodenum in a human embryo at CS17 were stained with toluidine blue after immunohistochemistry (brown) for PDX1 (B,C) and after RNAScope *in situ* hybridisation for *RFX6* (red; C,D) and *PDX1* (green; D) transcripts, with (C,D) Hoechst fluorescent counter staining for DNA. Note that panel (C) in this Supplementary Figure is the same image as in Figure 6C here shown side-by-side with an adjacent section from the same embryo stained for PDX1. Asterisk (*) highlights microlumen. dp, dorsal pancreatic bud; d, duodenum. Scale bars in (A,B) 100 µm, (C) 20 µm and (D) 40 µm.

Development • Supplementary information

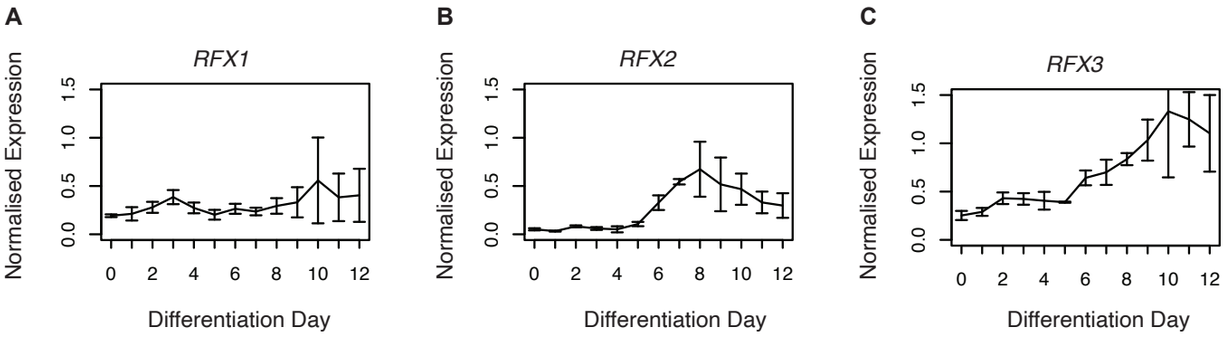


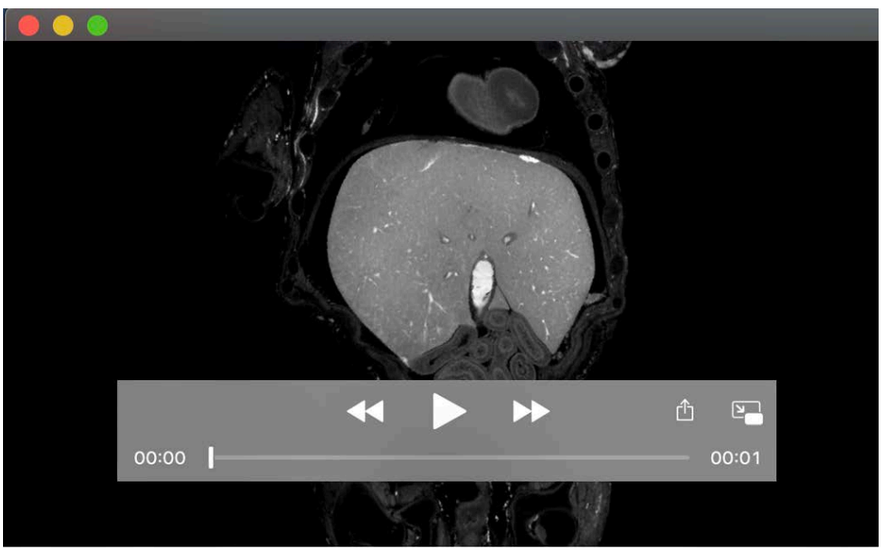
Figure S6. Expression of *RFX1* (A), *RFX2* (B) and *RFX3* (C) measured daily during differentiation of H9 hESC into pancreatic progenitor cells. The number of copies per μL in each sample was determined by reference to a standard curve and plotted relative to the number of copies of the housekeeping gene *TBP*. Error bars represent the standard error of the mean of 3 independent differentiation experiments.

Development • Supplementary information

3D interactive PDF. Interactive 3D models of pancreas development in human embryos. On the first page of the 3D-PDF a user manual is shown. The 3D-PDF file should be viewed in Adobe Reader® X or higher, available from http://www.adobe.com/downloads/ with Javascript enabled. 3D interaction is only possible on MS Windows or Mac OS. Right-click on the 3D model and choose: "Show Model Tree" for more advanced selection options. The 3D reconstructions on which the 3D-PDF is based are derived from the *3D Atlas of Human Embryology* by de Bakker et al. (2016), with permission (de Bakker et al., 2016). All 3D-reconstructions are based on histological sections of real human embryos from the Carnegie collection in Washington DC, USA. The complete procedure for the production of the 3D models and 3D-PDF is described in the original paper (de Bakker et al., 2016). In short, all organ systems were manually segmented in Amira, the resulting 3D models were smoothed and optimized using Blender. Deep Exploration and Adobe Acrobat Pro was used to create the interactive 3D-PDF. For this 3D-PDF, we selected 9 stages and for each stage a subset of 3D structures, mainly the alimentary system. The resulting 3D models were incorporated in one 3D-pdf with a more detailed user interface by using "Adobe Acrobat Pro XI". The following stages are included in the subsequent pages of this 3D-PDF: Stage 13 specimen 836 (28-32 days of development (d.d.)), Stage 14 specimen 8314 (31-35 d.d.), Stage 15 specimen 3512 (35-38 d.d.), Stage 16 specimen 6517 (37-42 d.d.), Stage 17 specimen 6521 (42-44 d.d.), Stage 18 specimen 6524 (44-48 d.d.), Stage 20 specimen 462 (51-53 d.d.), Stage 21 specimen 7254 (53-54 d.d.) and Stage 23 specimen 9226 (56-60 d.d.).

Click here to Download 3D Interactive PDF

Development • Supplementary information



Movie 1. Micro-CT imaging z-stack through the MRS foetus III:4. The pancreas head is highlighted in yellow.

Development • Supplementary information

Supplementary Tables

Table S1. Metadata for RNA-seq samples. RNA Integrity Number (RIN). Raw reads: read counts from the raw data. Clean reads: read counts remaining after removing reads containing: (1) adaptors, (2) greater than 10% undetermined bases, (3) more than 50% of bases with a Q score (Quality value) less than 6. The following values are calculated for clean reads only. Error rate (%): base error rate for all sequences. Q20 and Q30 (%): Percentage of bases with Phred values greater than 20 or 30. GC content (%): The percentage of total bases that are G or C. Uniquely mapped reads: numbers of reads that map to a single position in the genome

Click here to Download Table S1

Table S2. Gene counts and fold-changes for samples harvested from pancreatic endoderm differentiated from MRS patient and unaffected father iPSC. Gene expression measurements were taken from four independent experiments using two different father iPSC lines and compared to those from three independent experiments using two different MRS patient (III:4) iPSC lines. Fold-change and significance values for gene expression differences between father and patient cells were calculated using the DESeq2 package. Raw RNA-seq read files are available for download at ArrayExpress under accession number E-MTAB-9243.

Click here to Download Table S2

Table S3. RFX6-HA-specific peaks in BED format. Raw ChIP-seq read files are available for download at ArrayExpress under accession number E-MTAB-9335.

Click here to Download Table S3

Table S4. ChIP enrichment in RFX6-HA-specific peaks. List of RFX6-HA peaks enriched in RFX6-HA binding compared to H9 control. Enriched binding was defined as Fold change ≥ 4 and p-value ≤ 0.0001

Click here to Download Table S4

Table S5. Differentially expressed genes associated with RFX6-HA-specific peaks. RFX6-HA-specific peaks (FC≥4, p-value≤0.0001) were associated to the nearest protein-coding gene and matched to the differentially expressed genes (log2FC≥1, p-value≤0.05) by their Ensembl ID.

Click here to Download Table S5

Table S6. Antibodies for immunofluorescence staining.

Click here to Download Table S6

Table S7 Antibodies for flow cytometry

Click here to Download Table S7

Table S8. RT-qPCR primers

Click here to Download Table S8

Table S9. PCR primers and gRNA used for CRISPR/Cas9 genome editing.

Click here to Download Table S9

Development • Supplementary information