ISCI, Volume *16*

Supplemental Information

HNF4A Haploinsufficiency in MODY1 Abrogates Liver and Pancreas Differentiation from Patient-Derived Induced Pluripotent Stem Cells

Natasha Hui Jin Ng, Joanita Binte Jasmen, Chang Siang Lim, Hwee Hui Lau, Vidhya Gomathi Krishnan, Juned Kadiwala, Rohit N. Kulkarni, Helge Ræder, Ludovic Vallier, Shawn Hoon, and Adrian Kee Keong Teo

Supplemental Figure S1: Ng et al.

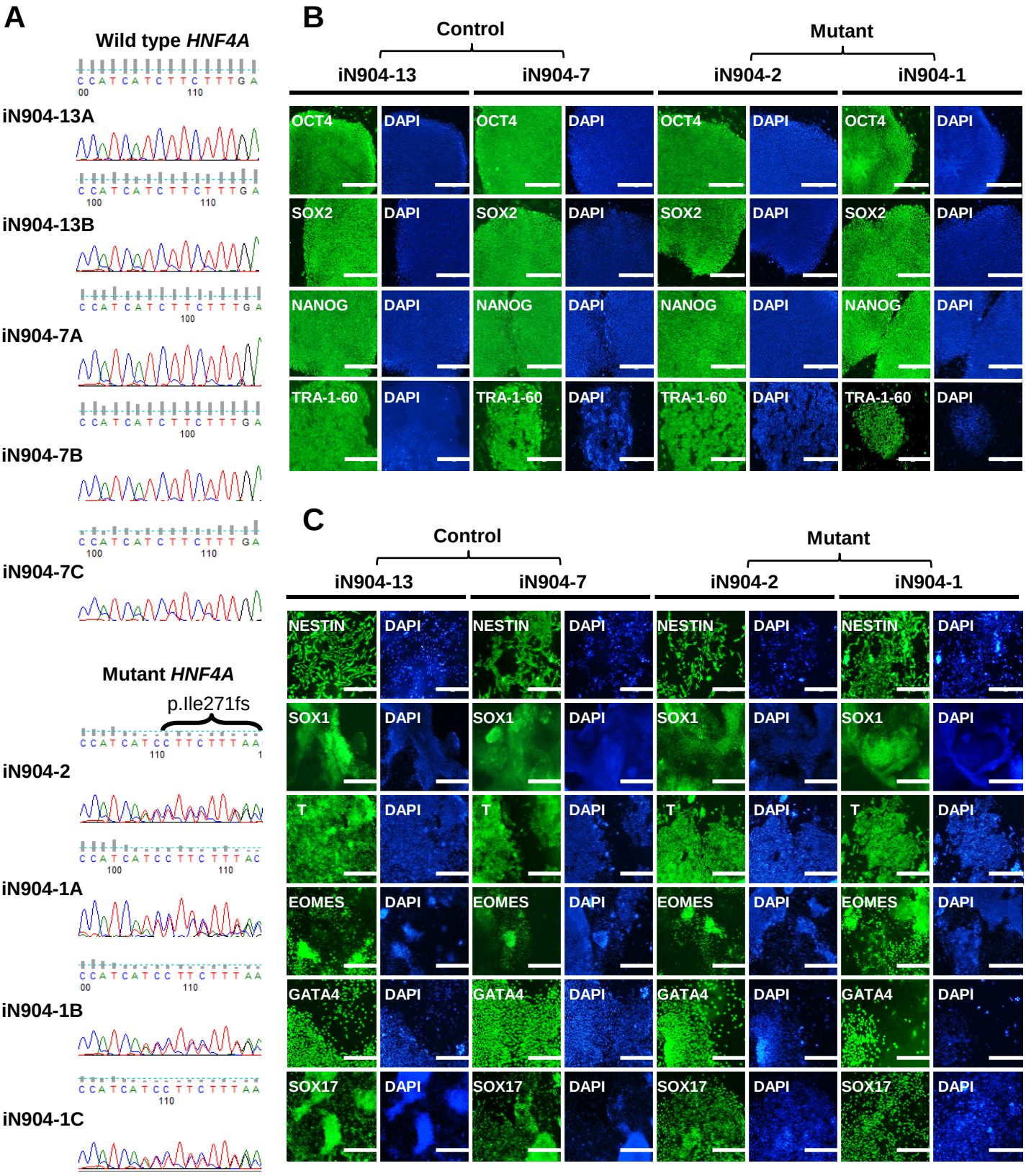
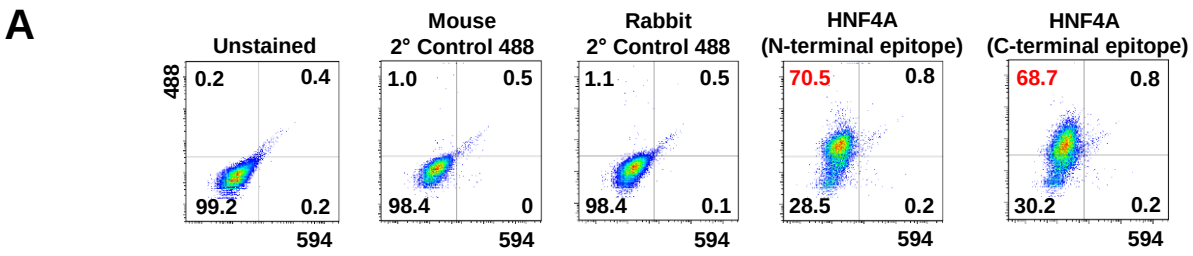
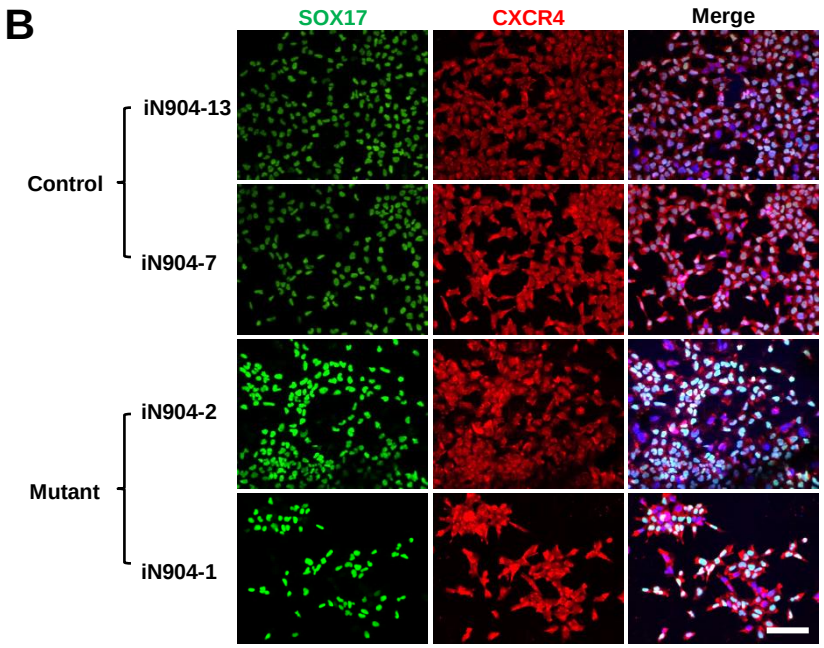
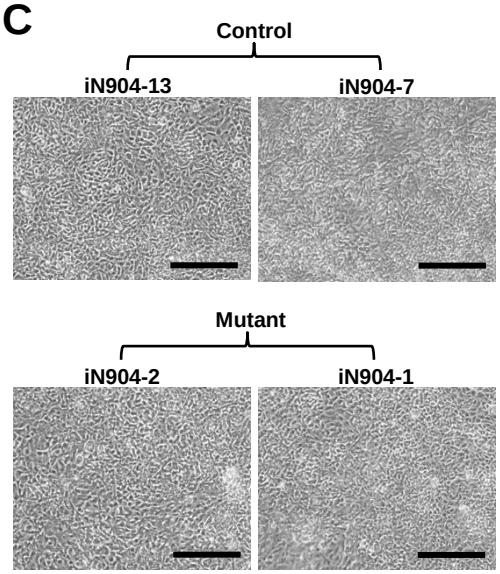


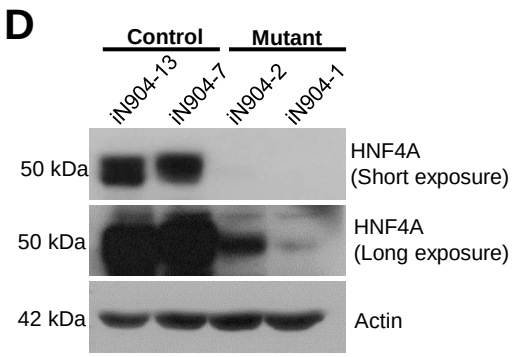
Figure S1, related to Figure 1. Characterization of MODY1-hiPSCs. (**A**) DNA sequencing to verify absence or presence of *HNF4A* p.Ile271fs mutation in control- and MODY1-hiPSCs. Control- and MODY1-hiPSCs (**B**) express pluripotency markers and (**C**) can give rise to cell types making up the three germ layers. Scale bars: 400 µm.

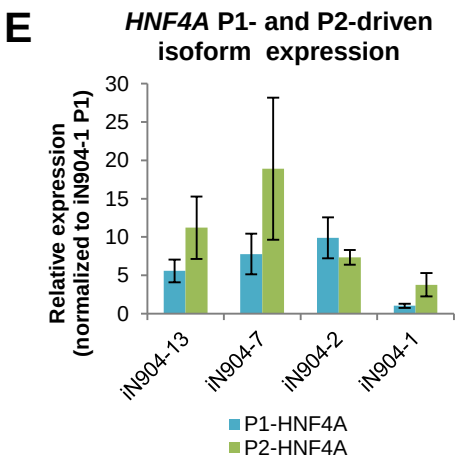
Supplemental Figure S2: Ng et al.











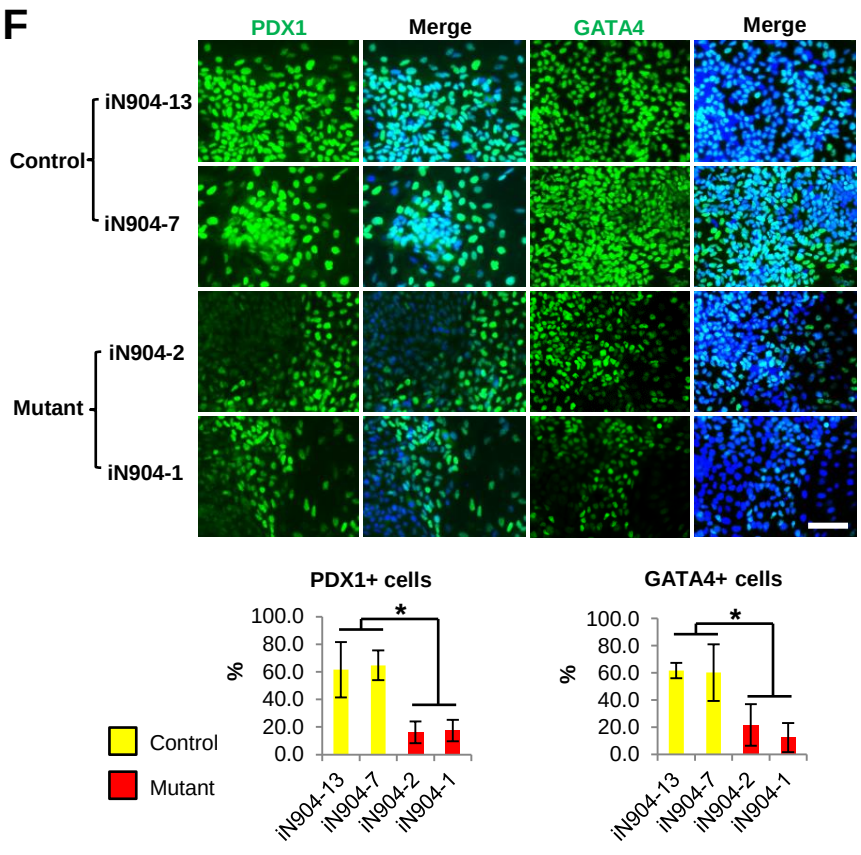
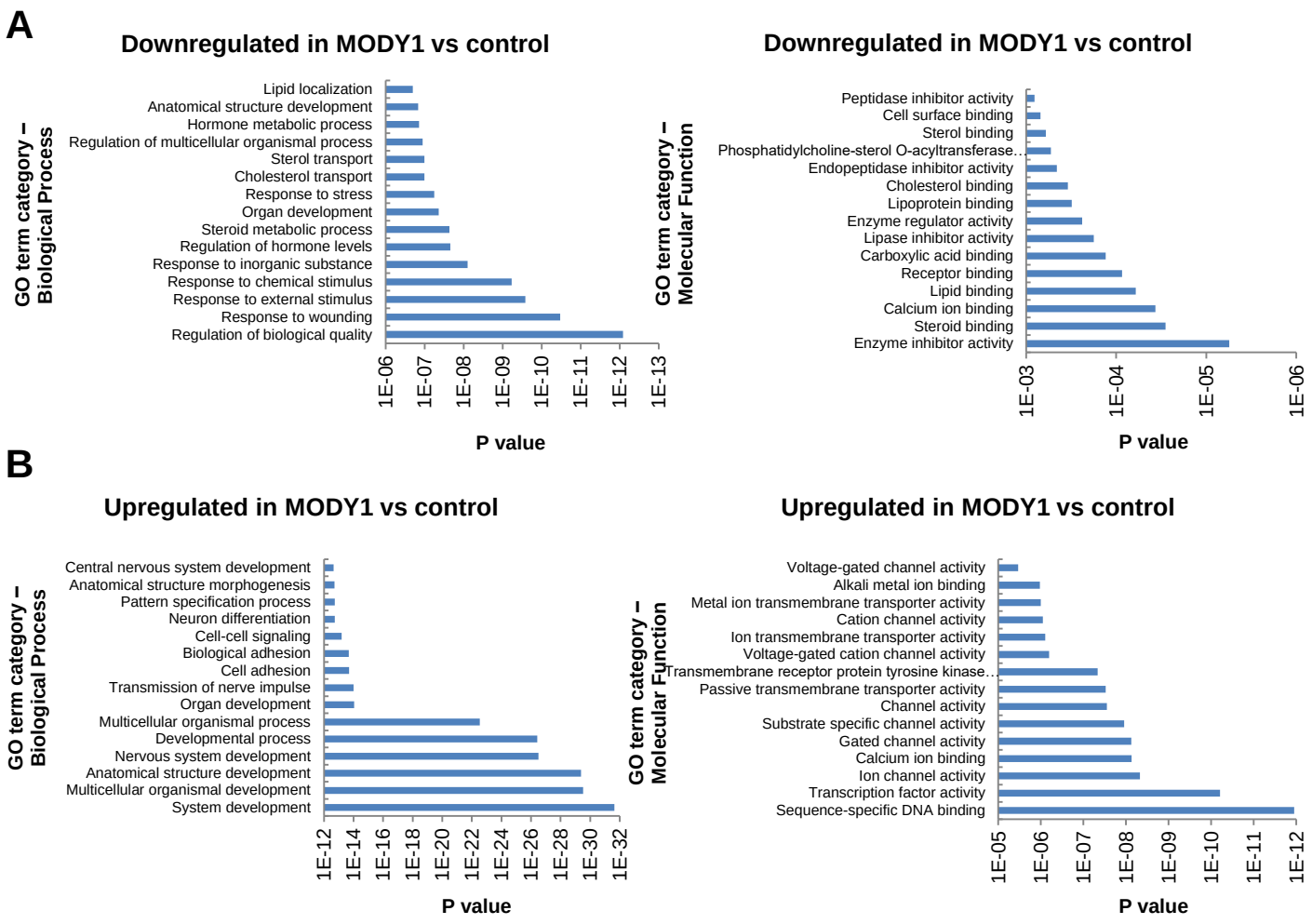
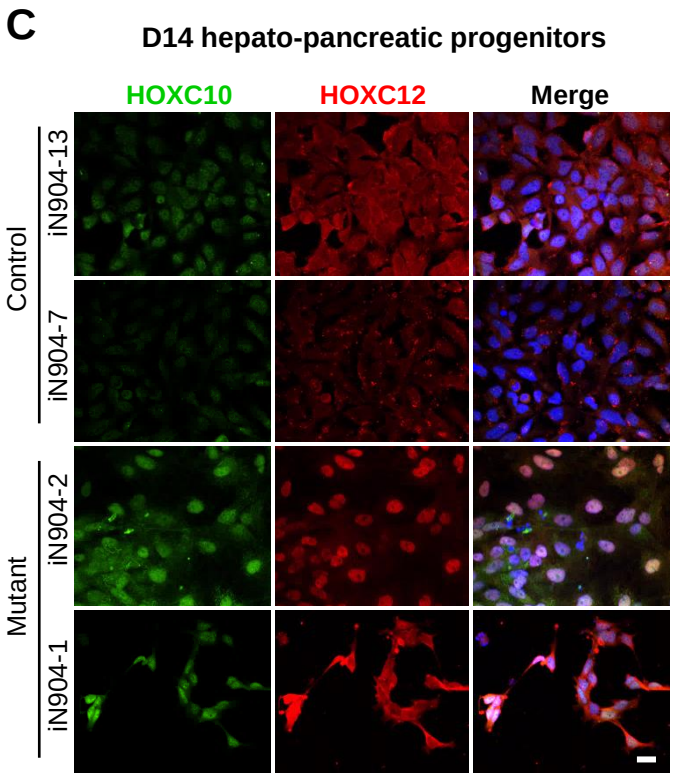


Figure S2, related to Figure 2. Characterization of the differentiation of iPSCs into D14 HPPs.

(**A**) ~70% of control-hPSC-derived cells are HNF4A^{+} based on FACS analysis. (**B**) Co-expression of definitive endoderm markers SOX17 (Green, R&D AF1924) and CXCR4 (Red, BD 555976) was determined at day 3 of differentiation. (**C**) No obvious morphological differences were observed between control- and MODY1-HPPs after 14 days of differentiation. (**D**) Western blot analysis using an antibody against the HNF4A C-terminal epitope (that recognizes full-length but not C-terminally-truncated protein) shows markedly reduced levels of HNF4A WT protein in MODY1-HPPs. (**E**) Both *HNF4A* P1 and P2 promoter-driven transcripts are expressed by isoform-specific qPCR analysis. Data are represented as mean ± SD of n=3, representative of 3 independent experiments. (**F**) Expression of pancreatic progenitor markers PDX1 (R&D AF2419) and GATA4 (Thermo Fisher 6H10) was downregulated in MODY1-HPPs compared to controls. The percentage of PDX1- or GATA4-positive cells were quantified from at least 3 different images, each with 189-594 cells counted. Data are represented as mean ± SD of n=3. *$p < 0.05$ versus controls as indicated by Student's t test. Scale bars: 100 μm.

Supplemental Figure S3: Ng et al.





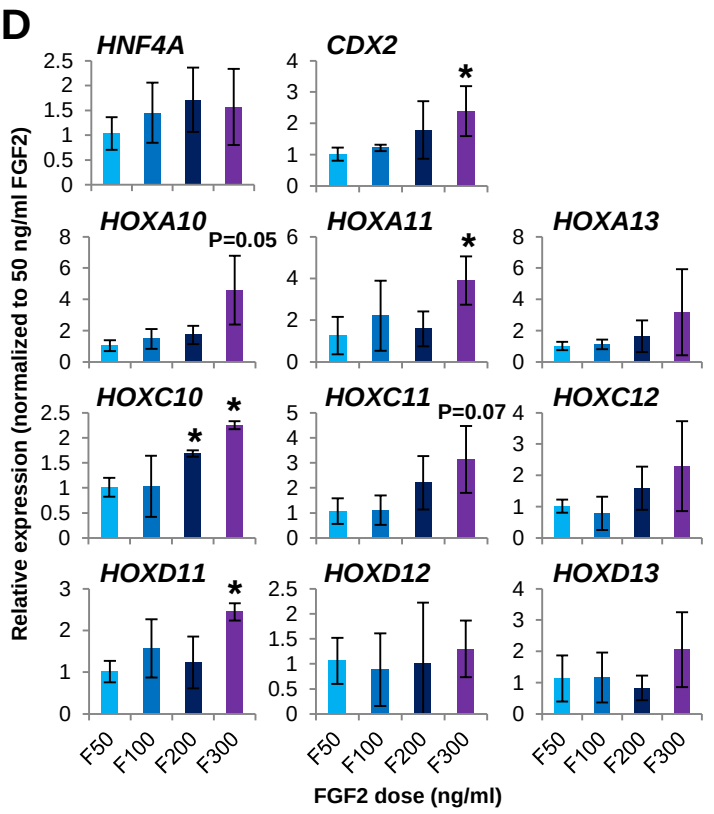
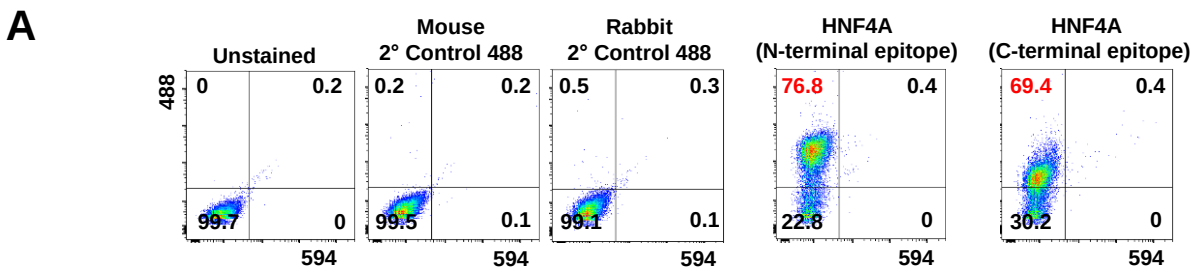
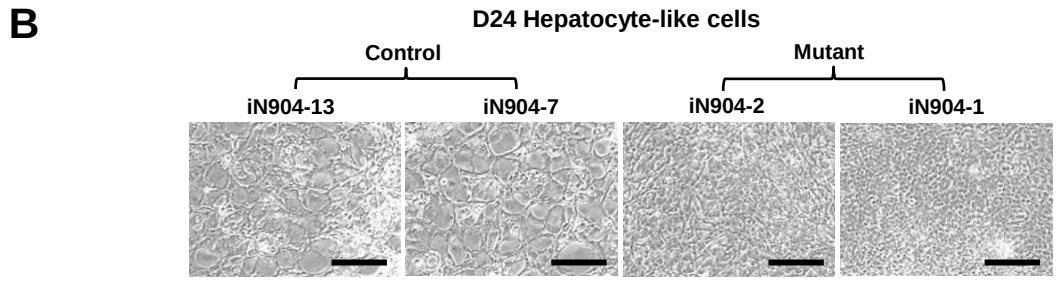


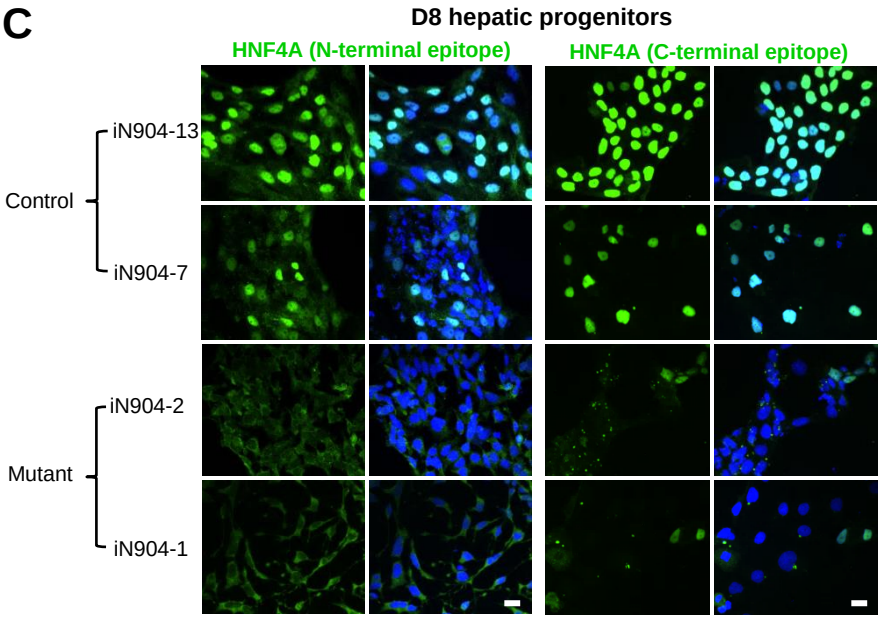
Figure S3, related to Figure 2 and Table S1. Gene ontology analyses of global transcriptional changes induced by the *HNF4A* (p.Ile271fs) mutation in MODY1-HPPs.

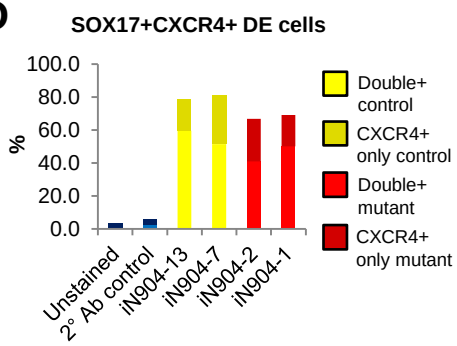
(**A**) Downregulated genes were involved in biological processes related to steroid metabolism and sterol transport, and molecular functions related to enzyme inhibitor activity, lipoprotein and sterol binding. (**B**) Upregulated genes were important for developmental processes and molecular functions related to DNA binding and channel activity. (**C**) Expression of representative hindgut markers HOXC10 and HOXC12 were evaluated by immunofluorescent confocal microscopy in D14 HPPs. Blue: DAPI, Green: HOXC10, Red: HOXC12. Scale bar: 50 µm. (**D**) High dose of FGF2 treatment increases the expression of caudal *HOX* genes. Data are represented as mean ± SD of n=3, representative of 3 independent experiments. *P<0.05 versus 50 ng/ml FGF2 (F50) samples by Student's *t*-test.

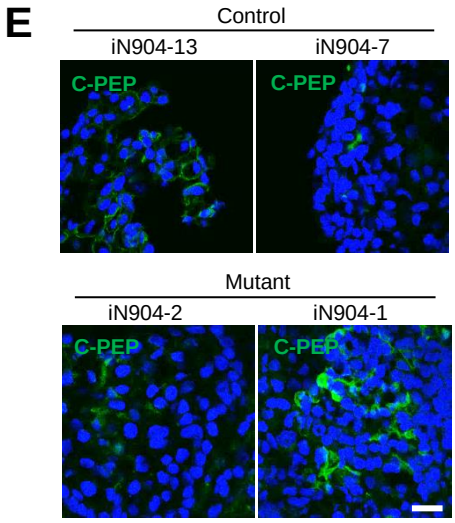
Supplemental Figure S4: Ng et al.

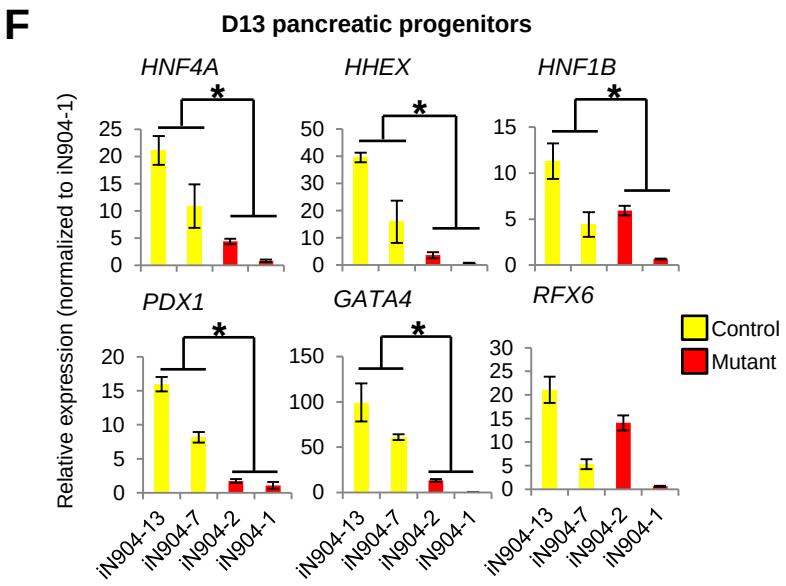












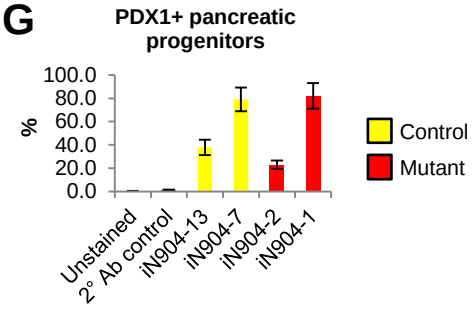


Figure S4, related to Figure 3. Characterization of hiPSC-derived hepatic and pancreatic cells.

(**A**) FACS analysis using HNF4A antibodies targeting the N- or C-terminal regions showing that 70 – 80% of control-hPSC-derived hepatic progenitors are HNF4A^{+} at D8. (**B**) Brightfield images showing that hepatocyte-like cells derived from MODY1-hiPSCs lack a polygonal morphology after 24 days of differentiation. Scale bar: 100 µm. (**C**) Immunofluorescent confocal images showing that HNF4A protein is predominantly localized to the nuclei in control D8 hepatic progenitors but is largely sequestered in the cytoplasm of MODY1 hepatic progenitors, based on antibodies targeting the N- or C-terminal regions of HNF4A. Blue: DAPI, Green: HNF4A, Scale bars: 50 µm. Confocal images were acquired using similar scan settings across samples. (**D**) FACS analysis showed co-expression of definitive endoderm (DE) markers SOX17-PE (R&D IC19241P) and CXCR4-APC (BD 555976) indicated as 'double+', and expression of CXCR4 alone ('CXCR4+ only') in control- and MODY1-derived cells after 5 days of differentiation in suspension cultures, using the β cell differentiation protocol adapted from Pagliuca et al. Data obtained from one experiment. (**E**) Immunofluorescence staining revealed expression of C-peptide (Green, DSHB GN-ID4) in both control- and MODY1-derived day 35 β-like cells. Blue: DAPI; scale bar: 50 µm. (**F**) QPCR analyses of D13 pancreatic progenitors generated using the β cell differentiation protocol were used to compare gene expression changes between control- and MODY1-derived cells. Data are represented as mean ± SD of n=3, representative of up to 3 independent experiments. *P<0.05 vs controls by Student's *t*-test. (**G**) FACS analysis showed expression of PDX1 (Abcam ab47308) in Day 13 pancreatic progenitors obtained using the β cell differentiation protocol. Data are represented as mean ± SD of n=2 independent experiments.

Supplemental Figure S5: Ng et al.

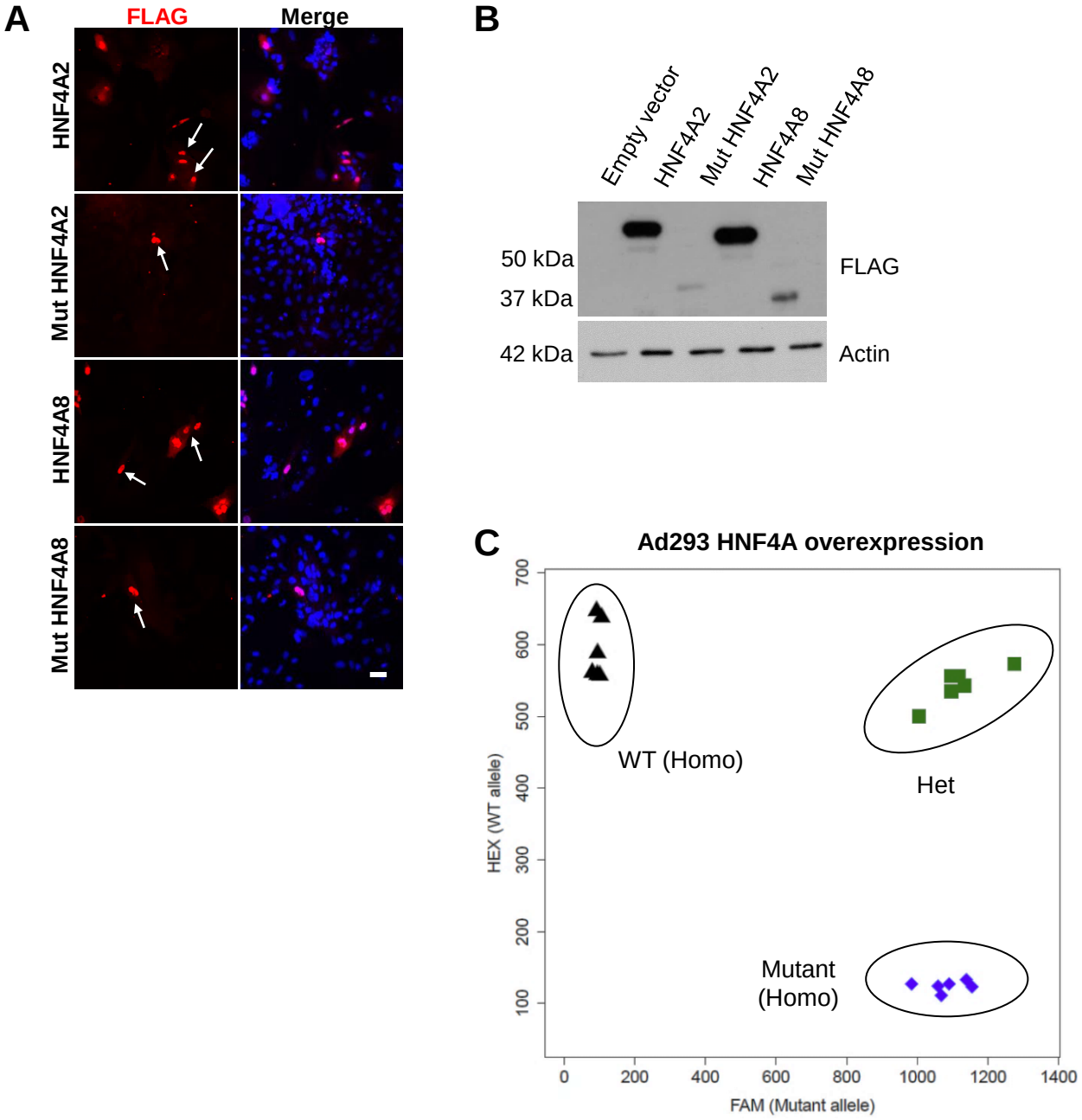


Figure S5, related to Figures 4 and 5. Validation of HNF4A WT and mutant overexpression.

Transient overexpression of WT and mutant (Mut) HNF4A2 and HNF4A8 constructs (N-terminal FLAG-tagged) are validated by (**A**) immunostaining in hiPSC-derived D16 hepatic progenitors and (**B**) western blot analysis in Ad293 cells. Blue: DAPI, Red: FLAG-tagged HNF4A, Scale bar: 50 µm. (**C**) Allele-specific qPCR analyses of mRNA from Ad293 transfected with HNF4A2 or HNF4A8 WT and/or mutant constructs, in triplicates, showed clear segregation among the genotypes. Axes show relative fluorescence units for each allele-specific TaqMan probe.

Table S1 (Excel table, related to Figures 2 and S3). Expression values (RPKM) of hiPSC-derived HPPs at D14 from RNA-Seq data. The first 2 columns indicate the gene ID and official gene symbol for all protein-coding genes in the dataset. FPKM values are shown for control-hiPSC-derived cells (13A, 13B, 7A, 7B, 7C) and MODY1-hiPSC-derived cells (2, 1A, 1B, 1C).

Table S2 (Related to Methods, Figures 1, 2 and 3). Quantitative real-time PCR primers used in this study.

Gene	Accession Number	Forward Primer (5' to 3')	Reverse Primer (5' to 3')
ABCC8	NM_001287174.1	GGAGCAGCAGCCTTCCTGACA	TGCGAAGCATAGGCCACGGG
ACTIN	NM_001101.3	TTGCCGATCCGCCGCCCGTC	CCCATGCCCACCATCACGCCCTGG
AFP	NM_001134.2	TGCTTCCAAACAAAGGCAGCAACAG	TGTACATGGGCCACATCCAGGAC
ALB	NM_000477.6	CCTGCCTGTTGCCAAAGCTCGAT	TGGCTCAGGCGAGCTACTGC
APOA2	NM_001643.1	AGCCTTGAAGGAGCTTTGGTTCGG	ACTGGGTGGCAGGCTGTGTTC
APOB	NM_000384.2	CCTGGGGCAGTGTGATCGCT	TGACAAGGGGCGGGTCATGC
APOC1	NM_001645.4	TCCTGTCGCTCCCGGTCCTG	GAAAACCACTCCCGCATCTTGGC
APOE	NM_001302688.1	CCAGGAGCCGGTGAGAAGCG	TGTGATTGGCCAGTCTGGAGGC
CDX2	NM_001265.4	CGGCAGCCAAGTGAAAACCAGG	TCGGCTTTCCTCCGGATGGTG
CYP3A4	NM_017460.5	CCCAGCAAAGAGCAACACAGAGC	CAGAGGTGTGGGCCCTGGAAT
GATA4	NM_001308093.1	GCAGAGAGTGTGTCAACTGTGGGG	TGGGGACCCCGTGGAGCTT
HHEX	NM_002729.4	ACACGCACGCCCTGCTCCGC	TGGCCAGACGCTTCCTCTCGGGC
HLXB9	NM_005515.3	GCGTCCACCGCGGGCATGATCC	AAGCGCTTGGGCCGCGACAGG
HNF1A	NM_001306179.1	CTTCTGCAGGAGGACCCGTGGCGT	GGCGGCCCGCTTCTGCGTCT
HNF1B	NM_000458.3	GGGGCCCGCGTCCCAGCAAA	GGCCGTGGGCTTTGGAGGGGG
HNF4A	NM_000457.4, NM_175914.4	GGACGACCAGGTGGCCCTGCTCAGA	GCTCCGGGCAGTGCCGAGGGA
HNF4A (P1)	NM_000457.4	GTGTTGACGATGGGCAATGACACG	CATGCCAGCCCGGAAGCATT
HNF4A (P2)	NM_175914.4	CAGTGGAGAGTTCTTACGACACG	CTTCTTCGCCCGAATGTCGC
HOXA10	NM_018951.3	TCCCTGGGCAATTCCAAAGGTGAAA	AGGTGGACGCTGCGGCTAAT
HOXA11	NM_005523.5	TTCCGGCCACACTGAGGACAAG	GTTGAGCATGCGGGACAGTTGC
HOXA13	NM_000522.4	CACTCTGCCCGACGTGGTCTC	ACCTTGGTATAAGGCACGCGCT
HOXC10	NM_017409.3	CCTCGGATAACGAAGCGAAAGAGGA	TCTTGCTAATCTCCAGGCGGCG
HOXC11	NM_014212.3	CGCAGATTTCGGCGAGCGAG	TGGTGCCACTTGCCGGATGG
HOXC12	NM_173860.1	GGGGCCGCTGGTAAACATCCA	GGCGTGGGTAGGACAGCGAA
HOXD11	NM_021192.2	ATGCTCAACCTCACTGACCGGC	GGCGCTTCCTGGAGCTCTCAAA
HOXD12	NM_021193.3	AGCGGAAACCCTACACGAAGCA	TGACTTGCTGGTCGCTGAGGTTC
HOXD13	NM_000523.3	GCTACCACTTCGGCAACGGCTAC	GGCACGTGCTGGTAAGGGCTC
INS	NM_000207.2	CCTGCAGGTGGGGCAGGTGGAGC	CGGGTGTGGGGCTGCCTGCG
KCNJ11	NM_000525.3	AGTGGGACCCAGGTGGAGGT	GTGGCCTAGGGCCTCACTGC
MAFA	NM_201589.3	GCCCGCTGGCCATCGAGTACGTCA	GAGGACAGCGAGCCTGGCGGC
NKX6.1	NM_006168.2	ACGCACGCCTGGCCTGTACCCC	CCCTCTCGGGCCCCGCCAAGTA
PDX1	NM_000209.3	CCTTCCCGGAGGGAGCCGAGCC	GTAGGCCGTGCGCGTCCGCT
RFX6	NM_173560.3	GCGGCTTGGAACAAGAGGCCA	ACGAGTGAAGCCACCCTCATTCTTT
SOX9	NM_000346.3	ACCAGCCGCGGCGGAGGAAGT	GGGATTGCCCCGAGTGCTCGCC

TRANSPARENT METHODS

Generation of hiPSCs

5 x 10^5 human fibroblast cells were seeded on a 6-well plate prior to transduction. After 24 hours, cells were replenished with medium containing non-integrative Sendai reprogramming vectors (polycistronic Klf4–Oct3/4–Sox2, L-Myc, and Klf4) according to the CytoTune-iPS 2.0 Sendai Reprogramming Kit (Thermo Scientific). All experiments with hiPSCs were approved by the Regional Committee of Medical and Health Research Ethics (REK 2010/2295), and all methods were performed in accordance with the Helsinki Declaration.

Human PSC culture and differentiation

HiPSCs generated from human fibroblast cells and H9 cells were at cultured at 37°C with 5% CO_2 in DMEM/F-12 with 15 mM HEPES (STEMCELL Technologies), 20% KnockOut™ serum replacement (KOSR), L-glutamine, NEAA (Life Technologies) and supplemented with 10 ng/ml FGF2 (Miltenyi Biotec). HiPSCs were seeded on irradiated CF-1 mouse embryonic fibroblasts (MEFs) and hiPSC media was replaced every 24 hours.

HiPSCs or H9 cells were differentiated into HPPs as described previously (Teo et al., 2012), with some modifications. Briefly, cells were incubated with Dispase (STEMCELL Technologies) and Collagenase IV (Life Technologies) and collected by mechanical scraping. MEFs were removed by passing the suspension through a 70 µm cell strainer. Plated cells were differentiated 2 days later in RPMI-1640/2% B-27 (no vitamin A; serum-free chemically-defined medium) supplemented with 100 ng/ml Activin A (R&D Systems), 3 µM CHIR99021 (Tocris) and 10 µM LY294002 (LC

Labs). On day 3, differentiation medium containing 50 ng/ml Activin A was added. On day 5, differentiation medium containing 50 ng/ml FGF2 (Miltenyi Biotec), 3 μM all-trans-retinoic acid (RA) (WAKO), and 10 mM nicotinamide (Sigma) was added, followed by subsequent media changes and addition of 20 μM DAPT (Abcam) on days 10 and 12.

HiPSCs and H9 cells were differentiated into hepatocyte-like cells as described previously (Hannan et al., 2013), with some modifications. The basal differentiation media used was the same as that for the HPP differentiation protocol described above during the first 9 days. The same supplements were also used for the first 4 days of differentiation. On day 5, differentiation medium containing 50 ng/ml Activin A was added. From days 6-9, differentiation media supplemented with 20 ng/ml BMP4 (Miltenyi Biotec) and 10 ng/ml FGF10 (Miltenyi Biotec) was added and replaced daily. From days 10-24, the HCM Bulletkit (Lonza) differentiation media supplemented with 30 ng/ml Oncostatin M (Miltenyi Biotec) and 50 ng/ml HGF (Miltenyi Biotec) was added and replaced every other day.

HiPSCs and H9 cells were differentiated into β-like cells as described previously (Pagliuca et al., 2014), with some modifications. Cells were dissociated using Dispase (STEMCELL Technologies Inc) and Collagenase IV (Life Technologies), followed by mechanical scraping. MEFs were removed by passing the suspension through a 40 μm cell strainer. TrypLE Express (Life Technologies) was added for further dissociation into single cells. 1 x 10^6 single cells in mTeSR1 (STEMCELL Technologies Inc) containing 10 μM of Rho-Kinase Inhibitor (Y27632) (STEMCELL Technologies Inc) were seeded into each well of a Corning® CoStar® ultra-Low attachment 6-well plate. After 24 hours, media was replaced with mTeSR1 without

Y27632 and cells were cultured for another 48h before initiation of differentiation (Pagliuca et al., 2014).

In general, 1–3 independent hiPSC lines per subject were used in experiments, with biological triplicates analysed for each line. Experiments were repeated at least thrice.

All hPSC lines used were tested mycoplasma-negative.

FGF2 dose-response treatments

H9 cells were differentiated into day 5 cells following the HPP differentiation protocol as described above. On day 5 onwards, cells were treated with differentiation medium containing increasing doses of FGF2 (Miltenyi Biotec) – 50 ng/ml, 100 ng/ml, 200 ng/ml, 300 ng/ml till day 14 before being harvested for analysis (Ameri et al., 2010).

Clonal cell line culture

HepG2 cells were purchased from ATCC (HB-8065) and cultured in DMEM/Low glucose (Hyclone) with 10% heat-inactivated FBS (Life Technologies) and 1% NEAA (Life Technologies). EndoC-βH1 cells (Ravassard et al., 2011) were purchased from Univercell Biosolutions and cultured in DMEM/Low glucose (Life Technologies) supplemented with BSA (Sigma Aldrich), penicillin/streptomycin, 2 mM L-glutamine, 50 µM 2-mercaptoethanol, 10 mM nicotinamide (Sigma Aldrich), 5.5 µg/ml transferrin (Sigma Aldrich) and 6.7 ng/ml sodium selenite (Sigma Aldrich) on plates coated with 2 µg/ml fibronectin (Sigma Aldrich) and 1% ECM (Sigma Aldrich). Ad293 cells were cultured in DMEM/High glucose (Hyclone) with 10% heat-inactivated FBS (Life Technologies) and 1% NEAA (Life Technologies).

All cell lines used were tested mycoplasma-negative.

Immunofluorescence staining

Cell clumps were collected at the end of the differentiation and cryo-embedded in tissue freezing medium (Leica Biosystems). Cryo-embedded cells were sectioned and mounted onto glass slides. Sectioning was performed by the Advanced Molecular Pathology Laboratory (AMPL), A*STAR and stored at -80 °C. Cell sections or cells in monolayer culture were fixed with 4% paraformaldehyde for 20 minutes, permeabilised with 0.1% Triton X-100 in DPBS, then blocked with 5% donkey serum or 5% bovine serum albumin in DPBS containing 0.1% Triton X-100, before overnight incubation with primary antibodies at 4^{o}C. Primary antibodies used are for the detection of HNF4A (C-terminal epitope, Santa Cruz, sc-6556), HNF4A (N-terminal epitope, Abcam, ab181604), FLAG tag (Sigma Aldrich, F1804), HOXC10 (Abcam, ab153904), HOXC12 (Life Technologies, MA5-19125), SOX17 (R&D, AF1924), CXCR4 (BD Biosciences, 555976), PDX1 (R&D, AF2419), GATA4 (Thermo Fisher Scientific, MA5-15532) and C-peptide (DSHB, GN-ID4). Secondary antibody incubation was carried out followed by staining with DAPI. The secondary antibodies used were Alexa Fluor® 488 (Invitrogen, A11055), Alexa Fluor ® 488 (Invitrogen, 21202), Alexa Fluor® 594 (Invitrogen, A21203) or Alexa Fluor® 488 (Invitrogen, A21270). Brightfield images were acquired with the Axiovert 200M inverted microscope using the Axiovision LE software. Confocal images were acquired with the Olympus FV1000 inverted confocal microscope using the Olympus Fluoview v3.1 software.

Fluorescence-activated cell sorting (FACS)

Cells were harvested by mechanical scraping and dissociated into single cells following incubation with 0.25% Trypsin/EDTA at 37°C and passed through a 40 µm cell strainer. Cell clumps were dissociated into single cells using TrypLE Express (Life Technologies) at 37°C and passed through a 40 µm cell strainer. Single cells were fixed with 4% paraformaldehyde on ice for 1 hour, then blocked in FACS buffer (5% FBS in DPBS) containing 0.1% Triton X-100 on ice for 1 hour, followed by incubation with primary antibodies for 1 hour at 4°C. Primary antibodies used are for the detection of HNF4A (C-terminal epitope, Santa Cruz, sc-6556), HNF4A (N-terminal epitope, Abcam, ab181604), SOX17-PE (R&D, IC19241P), CXCR4-APC (BD Biosciences, 555976), PDX1 (Abcam, ab47308) and INS (Abcam, ab7842). Cells were washed with FACS buffer containing 0.1% Triton X-100 cells and incubated with secondary antibodies in the dark for 1 hour at 4°C. The secondary antibodies used were Alexa Fluor® 488 (Invitrogen, A11055), Alexa Fluor ® 488 (Invitrogen, 21202) or Alexa Fluor® 594 (Invitrogen, A11076). Finally, cells were washed in FACS buffer, resuspended in DPBS and analysed with the BD™ LSR II Flow Cytometer. Data analysis was performed using the FlowJo 7.0 software.

RNA extraction, reverse transcription and quantitative PCR

PrepEase RNA Spin Kit (Affymetrix) was used to extract total RNA from differentiated hiPSCs according to the manufacturer's instructions. To remove genomic DNA from the preparation, DNase treatment was carried out for 15 minutes at room temperature. Purified RNA was reverse transcribed using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). QPCR was performed on the CFX384 Touch™ Real-Time PCR Detection System with iTaq™ Universal

SYBR® Green Supermix (Bio-Rad). Reported fold changes are based on relative expression values calculated using the $2^{-\Delta\Delta C(T)}$ method with normalization to actin expression for each sample. QPCR primers were custom-designed to span exon-exon junctions, wherever possible, using Primer-BLAST (NCBI). Sequences of primers used are listed in Table S2.

Immunoblotting

Cells were harvested by mechanical scraping on ice and lysed in M-PER (Thermo Scientific) in the presence of protease and phosphatase inhibitors (Sigma Aldrich). Protein lysates were quantified using the BCA Assay (Thermo Scientific), separated with sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) using the Mini-PROTEAN Tetra Cell system (Bio-Rad) and transferred to PVDF membranes (Bio-Rad). Primary antibodies against endogenous HNF4A protein (Cell Signaling, 3113), FLAG tag (Sigma Aldrich) or actin (Sigma Aldrich) were used, followed by HRP-conjugated secondary antibodies (Santa Cruz). Chemiluminescent signals were detected using Super Signal West Dura Extended Duration substrate (Thermo Scientific).

Generation of expression constructs

The pCDH plasmid (System Biosciences) containing the CMV promoter, N-terminal FLAG tag coding sequence and ampicillin resistance gene was used as the expression vector for HNF4A2 and HNF4A8. HNF4A2 coding sequence was amplified from cDNA obtained from control D14 HPPs using the following primers: Forward primer 5’ ATGCGACTCTCCAAAACCCTC 3’; Reverse primer: 5’ CTAGATAACTTCCTGCTTGGTGA 3’. HNF4A8 coding sequence was amplified

from cDNA obtained from human islets (University of Alberta) using Phusion polymerase (Thermo Scientific) and the following primers: Forward primer 5' ATGGTCAGCGTGAACGCG 3'; Reverse primer: 5' CTAGATAACTTCCTGCTTGGTGA 3'. PCR products were inserted into the pCDH vector using the Quick Ligation Kit (NEB). The ligated plasmid was used to transform STBL3 competent cells (Thermo Fisher Scientific). Inserted sequences were verified by DNA sequencing. For site-directed mutagenesis, the following primers were used to generate the p.Ile271fs mutation by introducing an additional adenine base into the HNF4A coding sequence through a PCR using the Phusion polymerase (Thermo Scientific): Forward primer 5' CTCAAAGCCATCAATCTTCTTTGACC 3'; Reverse primer: 5' GTAGGCATACTCATTGTCATCGATC 3'. The parental strand was digested following incubation with *Dpn1* (NEB). Introduced mutations were verified by DNA sequencing.

Luciferase reporter assays

The h*HNF1A* promoter (-605 to +18) and h*APOB* promoter (-664 to +7) were cloned into the pGL4.10 vector. The h*AFP* enhancer region (-4730 to -3666) containing the HNF4A binding site was cloned into pGL4.23-Promoter vector. The h*HNF4A* P1 and P2 promoter constructs in the pGL4.10 vector were described previously (Teo et al., 2016). Cell lines or differentiated hiPSCs were co-transfected with the respective promoter construct, pRL-TK renilla vector, and an overexpression vector (Empty pCDH or pCDH-HNF4A WT or mutant) using Fugene 6 transfection reagent (Promega) or Lipofectamine 2000 (Life Technologies). Cells were transfected in triplicate wells and each experiment was independently performed at least twice. For *HNF4A* knockdown in EndoC-βH1 cells, siRNA-mediated RNA interference was

carried out using 25 nM non-targeting (D-001810-10) or *HNF4A*-targeting (J-003406-09) ON-TARGETplus human siRNA (Dharmacon, GE Healthcare) with Lipofectamine RNAiMAX (Life Technologies) for 48 hours before transfection for luciferase experiments. Cells were harvested 24-48h after transfection, and luciferase activity was measured using the Dual Luciferase Assay System (Promega). Firefly luciferase activity was normalized to Renilla luciferase activity for each well.

Chromatin immunoprecipitation (ChIP)

EndoC-βH1 cells from 2 confluent 10 cm plates were cross-linked with 3.3 mg/ml of dimethyl 3,3'-dithiobispropionimidate and 1 mg/ml of 3,3’-dithiodipropionic acid di(N-hydroxysuccinimide ester) (both Sigma Aldrich) for 30 minutes at room temperature and with 1% formaldehyde (Amresco) for 15 minutes. The cross-linking reaction was quenched with 0.125 M glycine and cells were first lysed in cell lysis buffer (10 mM Tris-HCl pH 8, 10 mM NaCl and 0.2% NP-40) and then in nuclear lysis buffer (50 mM Tris-HCl pH 8, 10 mM EDTA and 1% SDS) on ice in the presence of protease inhibitors on ice. Nuclear lysates were diluted in IP dilution buffer (20 mM Tris-HCl pH 8, 2 mM EDTA, 150 mM NaCl, 0.01% SDS and 1% Triton X-100) and sonicated for 30s on/45s off for 10 cycles using a Q500 sonicator (QSonica) with microtip probes at 30% power. Sonicated samples were pre-cleared using 10 µg rabbit IgG (Santa Cruz) and Protein A/G agarose beads. Agarose beads were removed by centrifugation and a portion of the supernatant was collected as the input control. Samples were divided equally and incubated with 10 µg of HNF4A antibody (Santa Cruz, sc-8987) or rabbit IgG overnight at 4°C. The following day, samples were incubated with Protein A/G agarose beads and the beads were recovered and

washed with IP wash buffer and Tris-EDTA buffer. The immunoprecipitated DNA was eluted from the beads using IP elution buffer (100 mM NaHCO3, 1% SDS, 100 mM DTT). Samples were successively treated with RNaseA, NaCl and Proteinase K. DNA was extracted by phenol/chloroform extraction. Finally, qPCR was carried out on the input, HNF4A pulldown and IgG samples using SYBR green (Bio-Rad), targeting the *HNF1A* promoter or a control region in *GAPDH*. QPCR data were quantitated using a standard curve based on the input DNA, and normalized against *GAPDH*. Results are expressed as fold change for *HNF4A* pulldown relative to IgG control.

RNA sequencing and differential expression analysis

Poly-A mRNA was enriched from 1 µg of total RNA with oligo-dT beads (Invitrogen). Up to 100 ng of poly-A mRNA recovered was used to construct multiplexed strand-specific RNA-seq libraries as per manufacturer's instruction (NEXTflex^{TM} Rapid Directional RNA-SEQ Kit, dUTP-Based, v2). Individual library quality was assessed with an Agilent 2100 Bioanalyzer and quantified with a QuBit 2.0 fluorometer before pooling for sequencing on a HiSeq 2000 (1 x 101 bp read). The pooled libraries were quantified using the KAPA quantification kit (KAPA Biosystems) prior to cluster formation. Adapter sequences and low quality bases in Fastq read sequences were trimmed using Trimmomatic (v.0.33) (parameters: LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36). The quality filtered Fastq sequence reads were then aligned to the human genome (hg19) using Tophat (v.2.0.14) (parameters: --no-coverage-search --library-type=fr-firststrand) and annotated with Ensembl gene IDs. The resulting bam files were used to generate feature read counts using the Python package-based htseq-count of HTSeq (v.0.6.1p1) (parameters: default union-

counting mode, --stranded=reverse). The read count matrix output from HTSeq was used to perform differential expression analysis using the edgeR package (available in R (v.3.1.3)) in both 'classic' and generalized linear model (glm) modes to contrast patient versus control. Procedures described in edgeR documentation were followed to calculate P-values, FDR adjusted p-values (q-values) and fold-changes. A false discovery rate (FDR) cutoff of 0.05 was used to filter significantly differentially expressed genes. These genes with Ensembl IDs were mapped to gene symbols.

Allele-specific qPCR

A custom TaqMan® assay (Applied Biosystems) was designed to target the human *HNF4A* mRNA sequence surrounding the p.Ile271fs mutation (Assay ID: ANRWJVF). QPCR was performed on the CFX384 Touch™ Real-Time PCR Detection System (Bio-Rad) using the TaqMan® SNP Genotyping MasterMix (Applied Biosystems), according to the manufacturer's protocol. An allelic discrimination plot was generated on R using relative luciferase units (RFU) from the HEX (WT allele) and FAM (Mutant allele) probes.

Statistical analysis

Statistical parameters, number of replicates (n) and independent experiments conducted are indicated in the figure legends. Data represent mean ± SD. Gene expression and ChIP qPCR data were analyzed using two-tailed unpaired Student's *t*-test. Luciferase data were expressed as normalized relative Firefly/Renilla luciferase activity. Mean differences in relative activity were analyzed using two-tailed paired Student's *t*-test for clonal cell line studies, while two-way ANOVA with Bonferroni post tests (GraphPad Prism) was used for hiPSC-based studies to

analyze differences in relative activity segregated by MODY1 status and overexpression condition. Immunofluorescence data was quantified by cell counting using ImageJ across at least three representative images per condition. Results were considered to be significant at $P<0.05$.

Data and software availability

The accession number for the RNA-Seq data reported in this paper is GEO: GSE106335.

SUPPLEMENTAL REFERENCES

Ameri, J., Stahlberg, A., Pedersen, J., Johansson, J.K., Johannesson, M.M., Artner, I., and Semb, H. (2010). FGF2 specifies hESC-derived definitive endoderm into foregut/midgut cell lineages in a concentration-dependent manner. Stem Cells *28*, 45-56.

Hannan, N.R., Segeritz, C.P., Touboul, T., and Vallier, L. (2013). Production of hepatocyte-like cells from human pluripotent stem cells. Nature protocols *8*, 430-437.

Pagliuca, F.W., Millman, J.R., Gurtler, M., Segel, M., Van Dervort, A., Ryu, J.H., Peterson, Q.P., Greiner, D., and Melton, D.A. (2014). Generation of Functional Human Pancreatic beta Cells In Vitro. Cell *159*, 428-439.

Ravassard, P., Hazhouz, Y., Pechberty, S., Bricout-Neveu, E., Armanet, M., Czernichow, P., and Scharfmann, R. (2011). A genetically engineered human pancreatic beta cell line exhibiting glucose-inducible insulin secretion. J Clin Invest *121*, 3589-3597.

Teo, A.K., Lau, H.H., Valdez, I.A., Dirice, E., Tjora, E., Raeder, H., and Kulkarni, R.N. (2016). Early Developmental Perturbations in a Human Stem Cell Model of MODY5/HNF1B Pancreatic Hypoplasia. Stem Cell Reports *6*, 357-367.

Teo, A.K.K., Ali, Y., Wong, K.Y., Chipperfield, H., Sadasivam, A., Poobalan, Y., Tan, E.K., Wang, S.T., Abraham, S., Tsuneyoshi, N*., et al.* (2012). Activin and BMP4 Synergistically Promote Formation of Definitive Endoderm in Human Embryonic Stem Cells. STEM CELLS *30*, 631-642.