Supplementary Materials:

Supplementary tables

Table S1. Primer and oligo (sgRNA & ssDNA) sequence used for qPCR and CRISPR/Cas9 with *HNF1A* sgRNA off-target sites characterization.

Table S2. Sanger sequencing for *HNF1A* sgRNAs off-target genes in different cell lines.

Table S3. List of down/up-regulated genes from bulk RNAseq transcriptome of $INS^{GFP/wt}$ sorted cells *in vitro.*

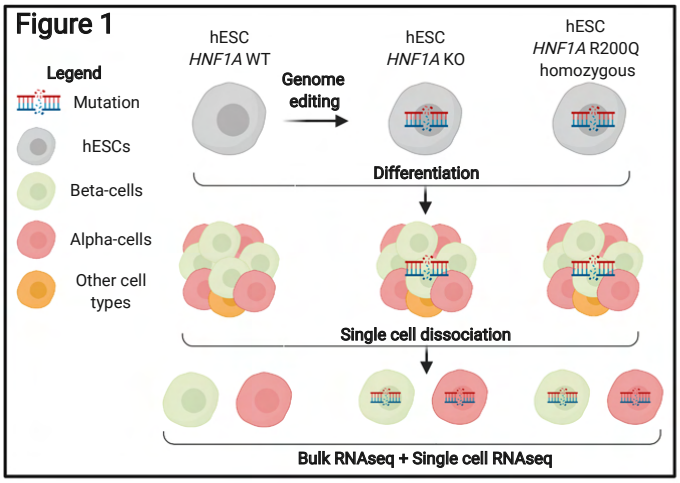
Table S4. GOTERM and KEGG analysis from single cell RNAseq transcriptome of $INS^{GFP/wt}$ sorted cells *in vitro.*

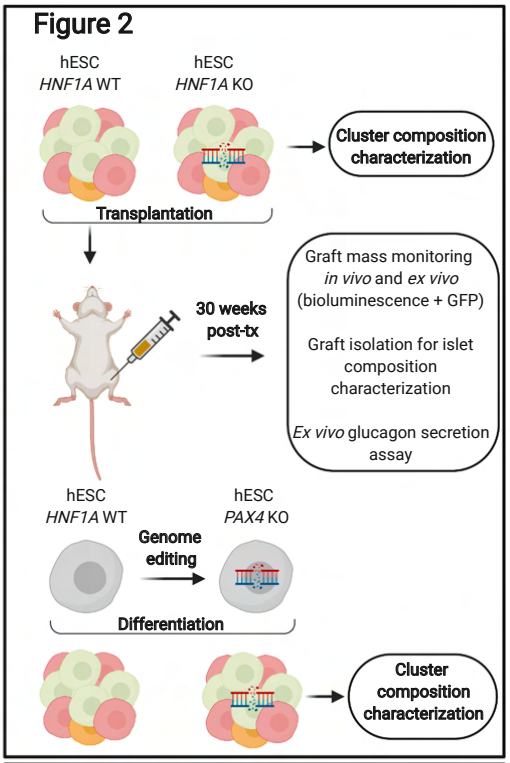
Table S5. GOTERM and KEGG analysis from single cell RNAseq transcriptome of unsorted SC-islet-like cells *in vitro.*

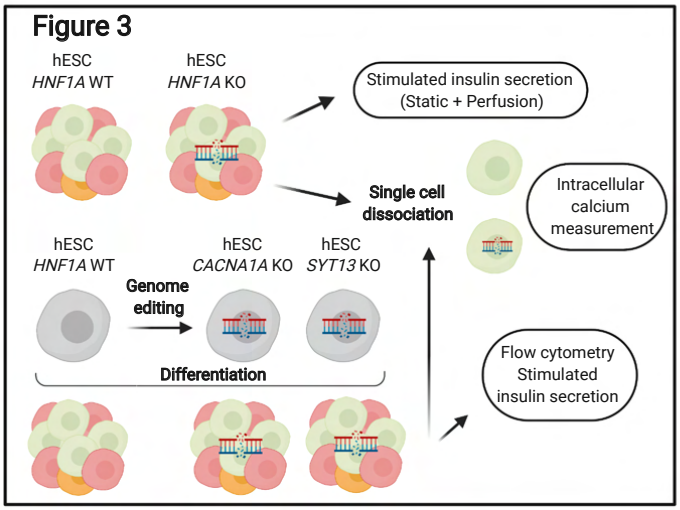
Table S6. Genotypes and ClinVar analysis of three MODY3 patients. MODY3 patient 2 (Pt2), mother of patient 3 (Pt3) and segregating for same *HNF1A* (+/R200Q) mutation had a history of gestational diabetes in 2 pregnancies and was diagnosed as having T2D in 2003 at age 38. She was managed with oral agents until 2015 for increasing hyperglycemia (292 mg/dl and HbA1c of 10.3%) when insulin was initiated and sulfonylurea discontinued. Since her last visit in 2016, fasting glucose levels remained elevated (216 mg/dl and HbA1c of 11.9%). MODY3 patient 3 (Pt3) was misdiagnosed as having T1D at age 13 and was found to have a mutation in *HNF1A* (+/R200Q) at age 15. Insulin was discontinued and sulfonylurea treatment prescribed. Fasting glucose levels at age 19 were 177 mg/dl with an HbA1C of 7.8%, for which the sulfonylurea dosage was increased. Since then, Pt3's fasting glucose and HbA1C levels have risen progressively, reaching 338 mg/dl and 6.9% by 2017 at age 22. During the patient's last visit in 2017, non-fasting C-peptide levels were within the normal range (724 pM) but insulin was undetectable (0 pM).

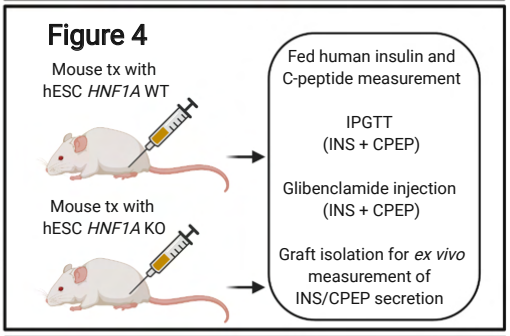
Table S7. Primary and secondary antibodies used for immunohistochemistry

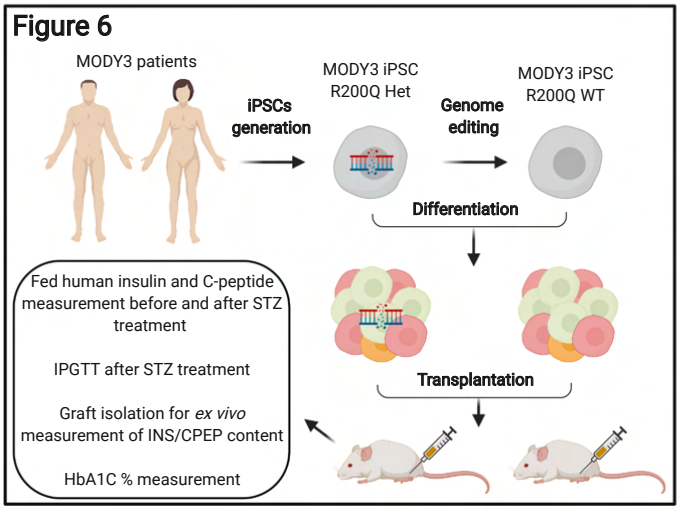
Supplemental Figure Legends











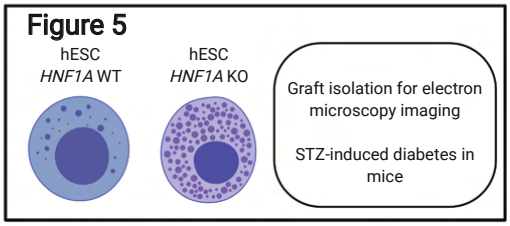
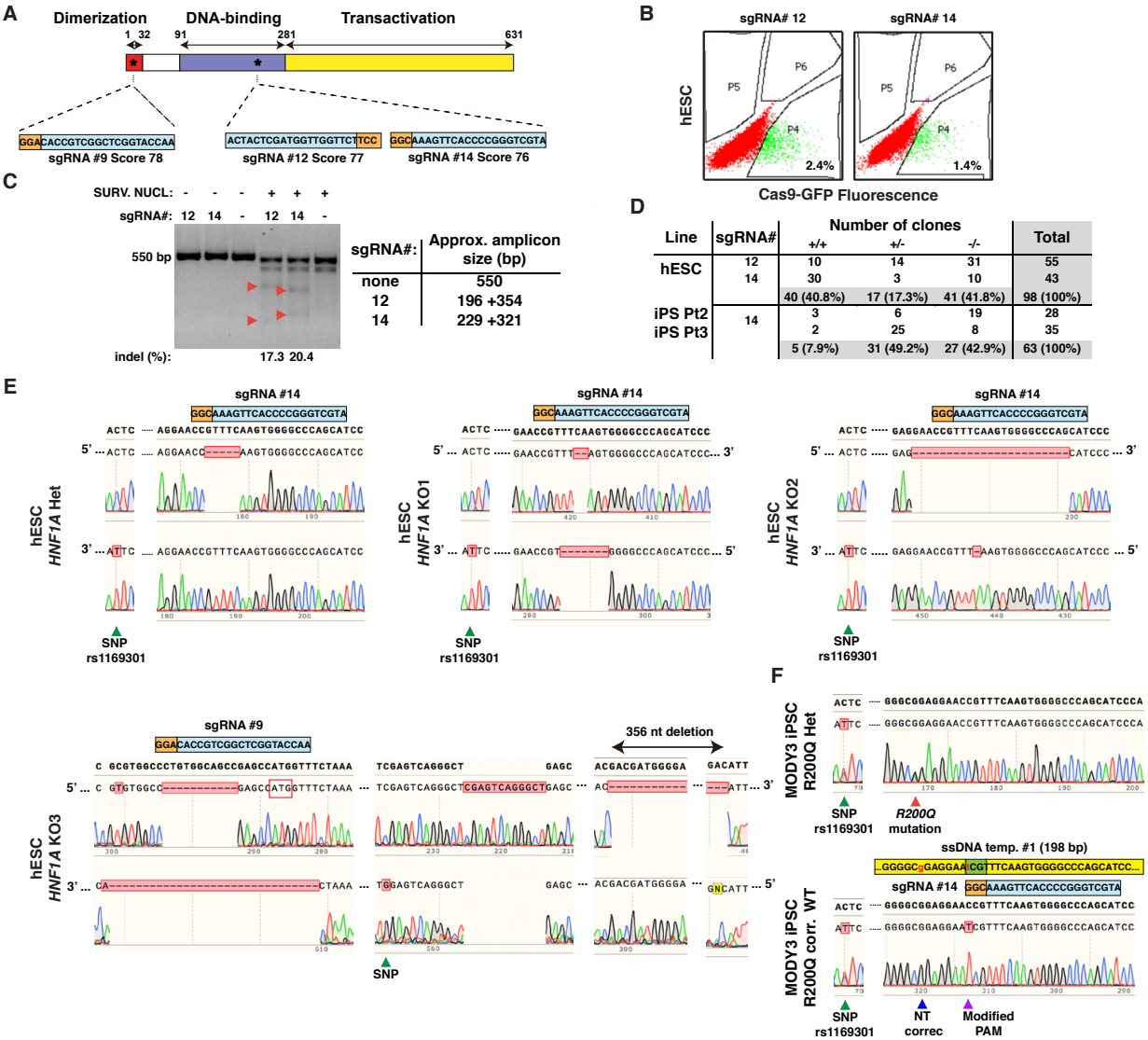


Fig. S1. Schematic overview of main figures. IPGTT = intraperitoneal glucose tolerance test, STZ = streptozotocin, IHC = immunohistochemistry, MODY = maturity onset diabetes of the young, Het = heterozygous, WT = wild type. Designed using Biorender.


A
Dimerization
DNA-binding
Transactivation
1 32
91
281
631
GGACACCGTCGGCTCGGTACCAA
sgRNA #9 Score 78
ACTACTCGATGGTTGGTTCTTCC
sgRNA #12 Score 77
GGCAAAGTTCACCCCGGGTCGTA
sgRNA #14 Score 76
B
sgRNA# 12
sgRNA# 14
hESC
P5
P6
P4
2.4%
1.4%
Cas9-GFP Fluorescence
C
SURV. NUCL: - - - + + +
sgRNA#: 12 14 - 12 14 -
550 bp
indel (%): 17.3 20.4
sgRNA#: | Approx. amplicon size (bp)
none | 550
12 | 196 +354
14 | 229 +321
D
Line | sgRNA# | Number of clones +/+ | +/- | -/- | Total
hESC | 12 | 10 | 14 | 31 | 55
14 | 30 | 3 | 10 | 43
40 (40.8%) | 17 (17.3%) | 41 (41.8%) | 98 (100%)
iPS Pt2 iPS Pt3 | 14 | 3 | 6 | 19 | 28
2 | 25 | 8 | 35
5 (7.9%) | 31 (49.2%) | 27 (42.9%) | 63 (100%)
E
hESC HNF1A Het
hESC HNF1A KO1
hESC HNF1A KO2
hESC HNF1A KO3
sgRNA #14
sgRNA #9
356 nt deletion
SNP rs1169301
SNP
F
MODY3 iPSC R200Q Het
MODY3 iPSC R200Q corr. WT
SNP rs1169301
R200Q mutation
ssDNA temp. #1 (198 bp)
sgRNA #14
NT correc
Modified PAM

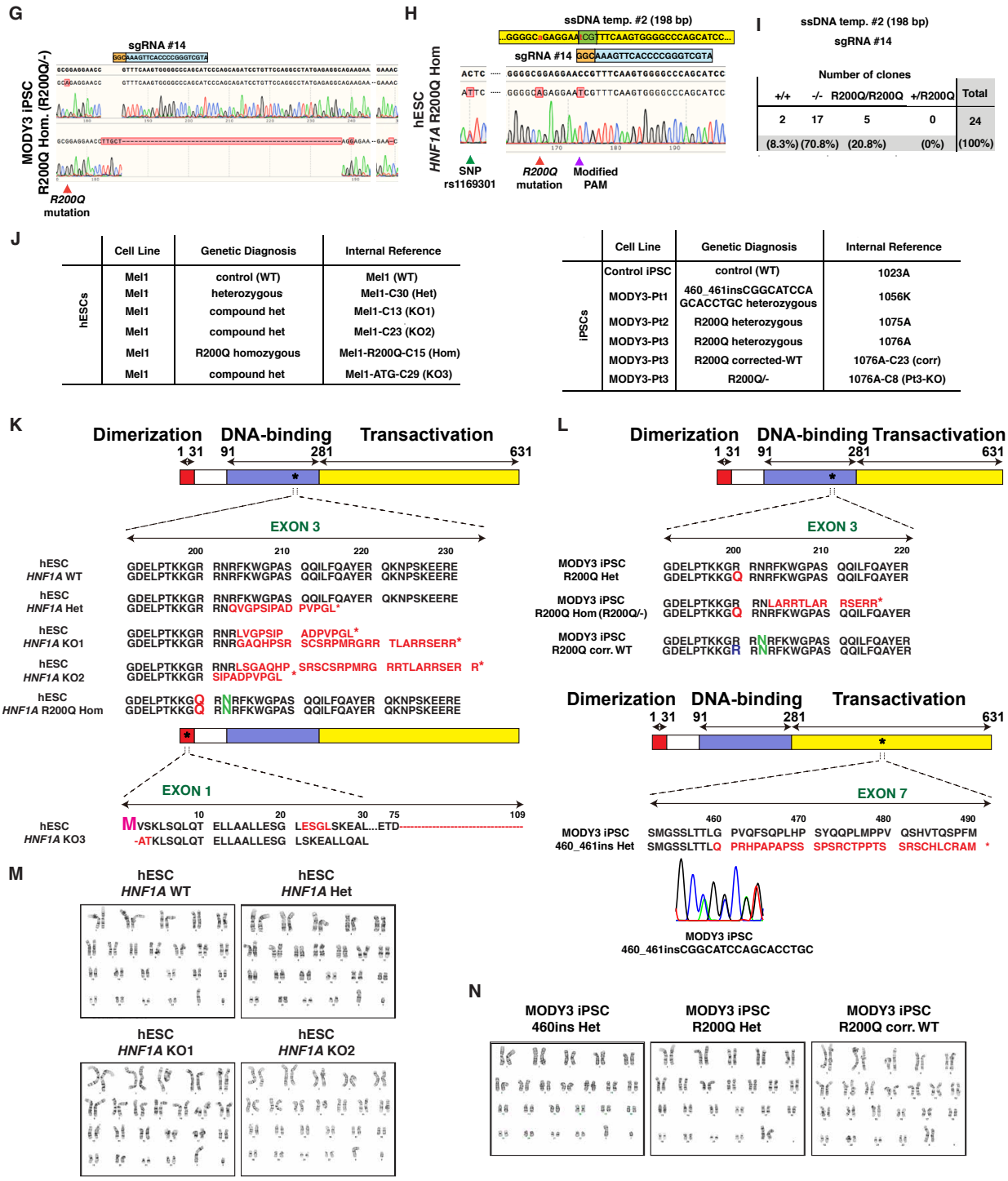


Fig. S2. Generation of isogenic cell lines with *HNF1A* mutations in hESCs and MODY3 iPSCs. (A) Diagram illustrating functional domains of HNF1A protein with sgRNAs #12 (on-target score 77) and #14 (on-target score 76) targeting the DNA-binding domain (exon 3). **(B)**

FACs sorting of hESCs nucleofected with Cas9-GFP and sgRNA#12 or #14 with 2.4% and 1.4% trasnfection efficiency. **(C)** Surveyor assay to assess sgRNA indel efficiency in *HNF1A* gene. *HNF1A* PCR products with (+) or without (-) surveyor nuclease or sgRNA indicating cutting efficiency of 17.3% and 20.4% for sgRNA#12 and sgRNA#14, respectively, with indicated amplicon sizes (Table). The surveyor assay cleaves DNA at the site of heterozygosity. Bands with amplicon size of ~500bp are due to the linked heterozygous SNPs rs1169301, creating cleavage. **(D)** Table showing the percentage of mutant lines generated after CRISPR/Cas9 in hESCs leading to 17.3% heterozygous clones (17 clones out of 98 total) and 41.8% compound heterozygous clones (41 clones out of 98 total), indicating a 59.2% of mutation efficiency (58 clones out of 98 total). 4 clones of the 98 (4%) appeared homozygous by Sanger sequencing, but also lacked two linked SNPs, rs1169301 (102 bp from PAM sequence of sgRNAs#14 to SNP) and rs2071190 (130 bp from PAM sequence of sgRNAs#14 to SNP), which are both located in the second intron. These homozygous lines (4/98, 4%) could be caused by a large deletion, and were not used for further studies. In MODY3 iPSC, 92.1% of clones were (+/-, heterozygous) and (-/-, compound heterozygous). Mutation-correction efficiency of MODY3 iPSC Pt2/Pt3 was 7.9% (+/+, wild type) from total number of clones. **(E)** Sanger sequencing results showing heterozygous (Het) and knockout frameshift mutations (KO1 and KO2) from both alleles in exon3 and knockout frameshift mutations (KO3) in exon1 near the start codon (red box). **(F)** Sanger sequencing showing heterozygous R200Q mutation (red arrow) in MODY3 iPSC line, after correction of the mutation (blue arrow) in MODY3 iPSC line (MODY3 iPSC R200Q-corrected WT) using a ssDNA template donor #1 (198 bp). **(G)** Sanger sequencing after introduction of a frameshift mutations in the wild type allele of MODY3 iPSC line (R200Q/-). **(H)** Sanger sequencing results showing R200Q homozygous mutation introduction (red arrow) in hESCs using a ssDNA template donor #2 (198

bp). PAM sequence in template donor was modified (C>T, silent mutation) to increase recombination efficiency (purple arrow). Heterozygosity at the upstream rs1169301 SNP C>T in *HNF1A* (green arrow; 102 bp from PAM sequence) indicates that both alleles are detected. All Sanger sequencing results were verified by TOPO® cloning of at least six clones per genotype. **(I)** Table showing the percentage of clones with R200Q homozygous mutation generated (20.8%) in hESCs from total number of clones. **(J)** Genetic information of cell lines (hESCs and iPSCs) included in the study**. (K)** Frameshift mutations in exon 3 of hESC cell lines leading to premature stop-codons (*) and generation of truncated versions of HNF1A protein (different amino acids from WT in red). For *HNF1A* KO3, mutation is at start codon in exon 1, indicated in red are different amino acids from WT sequence, red dash are deletions and in magenta the start codon. **(L)** Mutations in exon 3 of iPSC cell lines leading to heterozygous amino acid substitution (R200Q heterozygous), generation of R200Q homozygous (R200Q/-) and isogenic R200Q-corrected WT lines (+/+). For MODY3 iPSC 460_461ins Het line, CGGCATCCAGCACCTGC insertion is in exon 7 as shown by Sanger Sequencing, leading to frameshift mutations and premature stop codon (*). Different amino acids from WT in red, amino acid correction in bleu and amino acid silent mutation in green. **(M)** Cytogenic analysis performed on 20 G-banded metaphase cells from hESC and **(N)** iPSC lines indicating normal karyotypes.

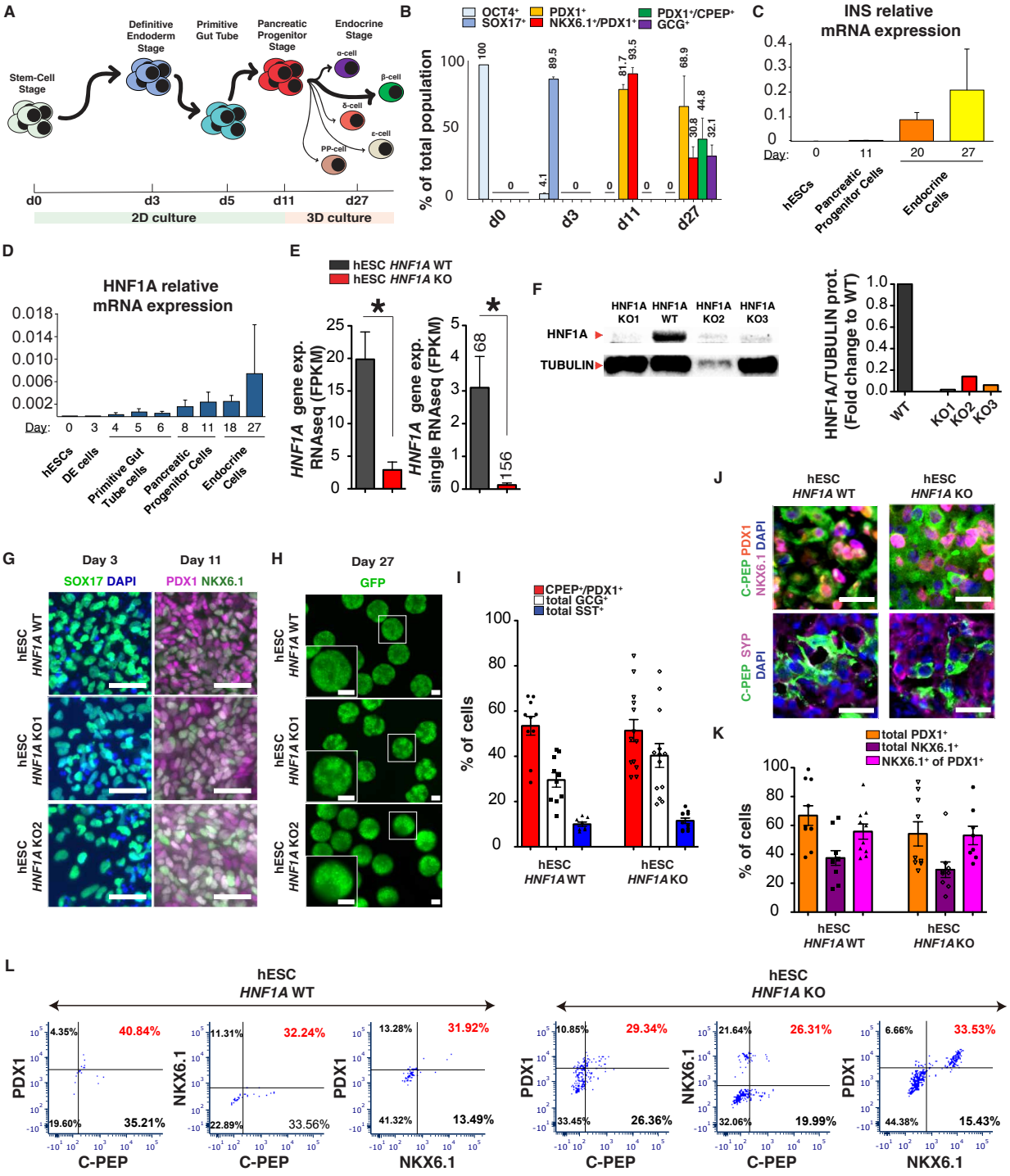


Fig. S3. *HNF1A* is not required to generate pancreatic endocrine cells *in vitro*. (A) Overview of the different stages of stem cell differentiation to pancreatic progenitor cells. The initial stages of differentiation were conducted in planar culture from day 0 to day 11 followed by 3D clustering

on day 12 to form aggregates or clusters of endocrine cells on day 27. **(B)** Percentage of total cell population by immunohistochemistry (IHC) expressing key markers of pluripotency stage (OCT4, d0), definitive endoderm stage (SOX17, d3), pancreatic progenitor stage (PDX1 and NKX6.1, d11) and endocrine stage (CPEP, c-peptide; GCG, glucagon, d27). **(C)** *INS* and **(D)** *HNF1A* relative mRNA expression levels determined by qPCR throughout the differentiation from hESC *HNF1A* WT cells (n=3-7). Each independent biological replicate (n) consists of 2-3 technical replicates for all experiments. **(E)** *HNF1A* gene expression levels (FPKM) determined by RNAseq after FACs sorting of $INS^{GFP/wt}$ positive cells from hESC-derived endocrine cells (WT n=3 and KO n=3) on day 27 of differentiation *in vitro*. Numbers on top of histogram denotes total number of single cells used for analysis. **(F)** Western blot for HNF1A (67 kDa) and β-Tubulin-III (55 kDa) from hESC-derived endocrine cells with respective quantification on day 27 of differentiation *in vitro*. **(G)** Representative IHC staining of hESC-derived cells *in vitro* at definitive endoderm stage with SOX17 (day 3) and pancreatic progenitor stage with PDX1 and NKX6.1 (day 11). Scale bars: 50 µm. White cells are PDX1/NKX6.1 double positive cells. **(H)** Representative GFP images. Scale bars at 200 µm. **(I)** Percentage of scβ-like-cells ($CPEP^+$/$PDX1^+$ in red), scα-like-cells (GCG^+ in white) and scδ-like-cells (SST^+ in blue) by IHC (day 27). **(J)** Representative IHC image of hESC-derived endocrine cell lines with for indicated markers with **(K)** quantification of $PDX1^+$ and $NKX6.1^+$ cells (day 27). (**L**) $CPEP^+$, $PDX1^+$ and $NKX6.1^+$ populations by flow cytometry at day 27 of differentiation *in vitro*. Scale bars: 20 µm. 20 clusters (~10k cells per cluster) of endocrine cells were used flow cytometry. For scatter plots, each point in plots represents an independent biological experiment (n). Data are represented as mean ± SEM. p-values: *p<0.05, **p<0.01, ***p<0.001, two-tailed t-test.

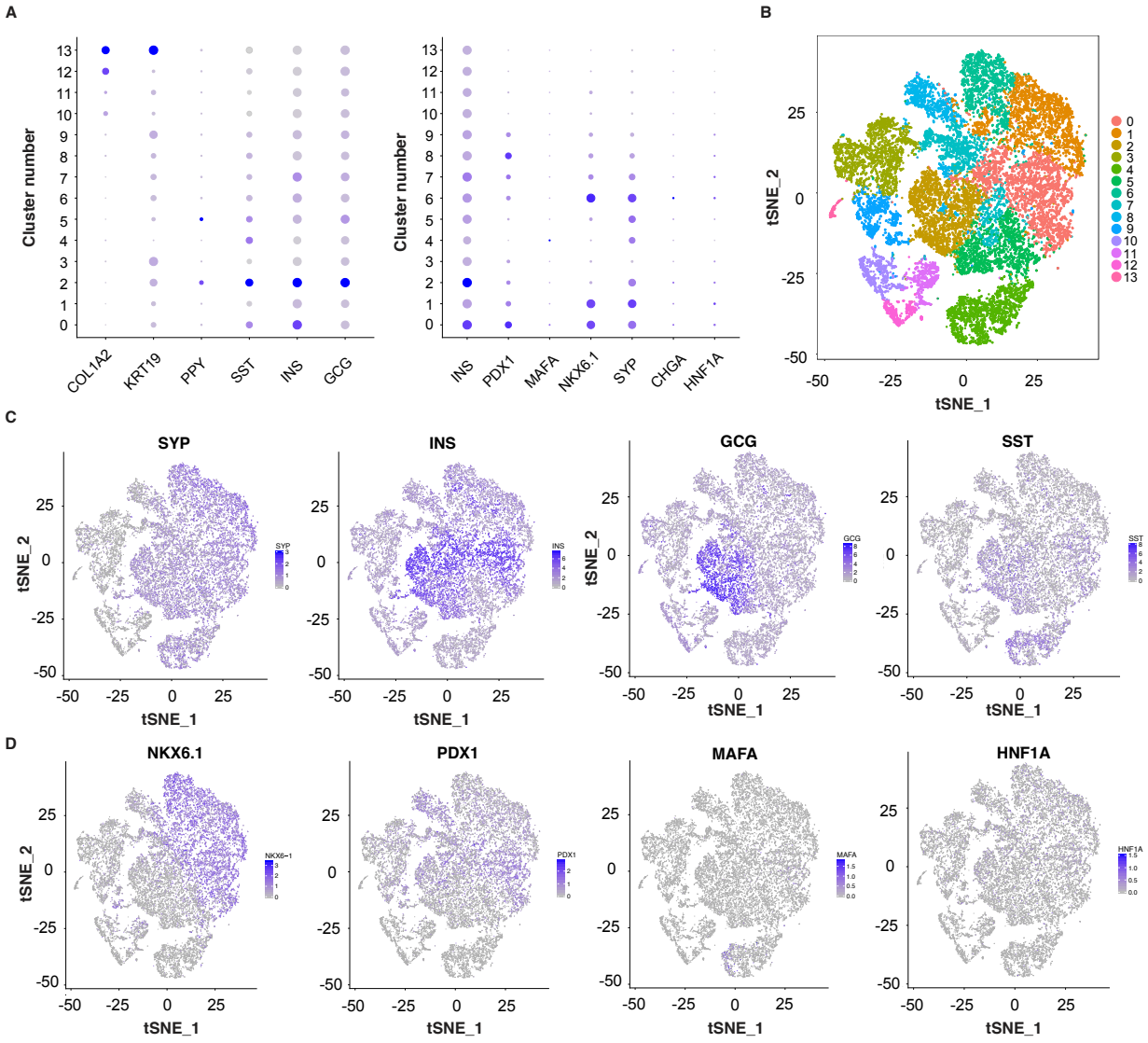


Fig. S4. Identification of thirteen stem cell-derived cell populations by single cell RNAseq *in vitro*. (A-D) Single cell RNA sequencing of 22164 (all genotypes combined) unsorted hESC-derived endocrine cells (n=3 for each genotype) *in vitro*. **(A)** Cell type clustering of cells based on their normalized expression of different pancreatic markers: *COL1A2* (pancreatic stellate cell), *KRT19* (ductal cell), *PPY* (γ-cell), *SST* (δ-cell), *INS* (β-cell), and *GCG* (α-cell). A total of 13 clusters were identified. Size is the percentage of cells expressing the indicated marker and color intensity the average expression level of the indicated marker, where blue indicates the highest expression level and gray the lowest. **(B)** Feature plot based on tSNE projection of cells where the

colors denote different cell type clusters identified (total of 13 identified) by *HNF1A* genotype line via Louvain algorithm performed by Seurat. **(C)** Feature plot based on tSNE projection of cells based on the normalized expression of endocrine markers *SYP* (synaptophysin), *INS* (insulin), *GCG* (glucagon) and *SST* (somatostatin) and **(D)** pancreatic transcription factors (*NKX6.1*, *PDX1*, *MAFA* and *HNF1A*) where blue indicates the highest expression level and gray the lowest. Genes expressed at low levels will have fewer labeled cells because of dropout. All stem cell differentiations were done for 27-30 days.

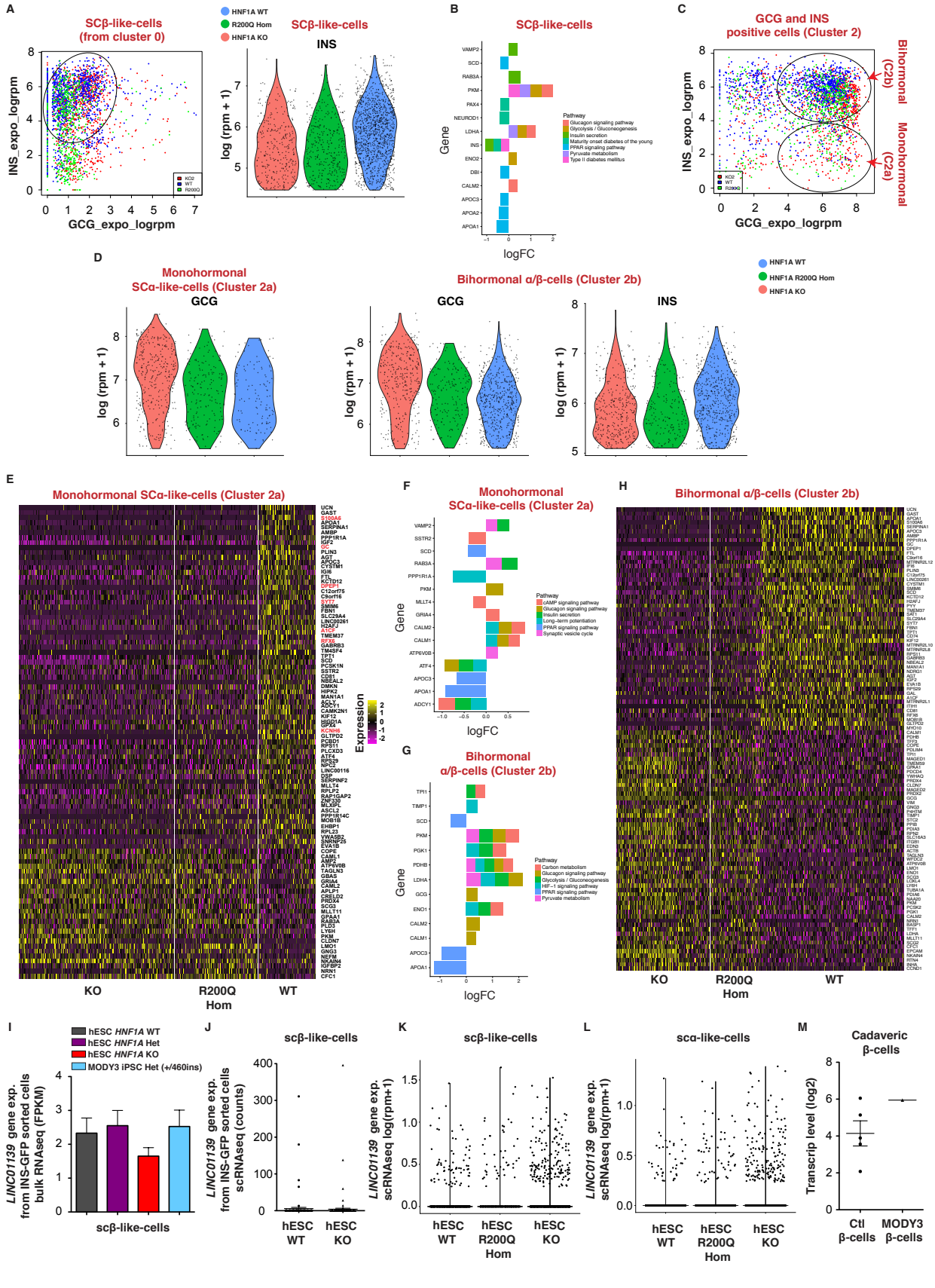

A
SCβ-like-cells (from cluster 0)
INS_expo_logrpm
GCG_expo_logrpm
HNF1A WT
R200Q Hom
HNF1A KO
SCβ-like-cells
INS
log (rpm + 1)
B
SCβ-like-cells
Gene
logFC
Pathway
Glucagon signaling pathway
Glycolysis / Gluconeogenesis
Insulin secretion
Maturity onset diabetes of the young
PPAR signaling pathway
Pyruvate metabolism
Type II diabetes mellitus
C
GCG and INS positive cells (Cluster 2)
Bihormonal (C2b)
Monohormonal (C2a)
D
Monohormonal SCα-like-cells (Cluster 2a)
GCG
Bihormonal α/β-cells (Cluster 2b)
INS
E
Monohormonal SCα-like-cells (Cluster 2a)
Expression
KO
R200Q Hom
WT
F
Monohormonal SCα-like-cells (Cluster 2a)
Pathway
cAMP signaling pathway
Glucagon signaling pathway
Insulin secretion
Long-term potentiation
PPAR signaling pathway
Synaptic vesicle cycle
G
Bihormonal α/β-cells (Cluster 2b)
Carbon metabolism
Glucagon signaling pathway
Glycolysis / Gluconeogenesis
HIF-1 signaling pathway
PPAR signaling pathway
Pyruvate metabolism
H
Bihormonal α/β-cells (Cluster 2b)
I
hESC HNF1A WT
hESC HNF1A Het
hESC HNF1A KO
MODY3 iPSC Het (+/460ins)
LINC01139 gene exp. from INS-GFP sorted cells bulk RNAseq (FPKM)
scβ-like-cells
J
scβ-like-cells
LINC01139 gene exp. from INS-GFP sorted cells scRNAseq (counts)
hESC WT
hESC KO
K
scβ-like-cells
LINC01139 gene exp. scRNAseq log(rpm+1)
hESC WT
hESC R200Q Hom
hESC KO
L
scα-like-cells
LINC01139 gene exp. scRNAseq log(rpm+1)
hESC WT
hESC R200Q Hom
hESC KO
M
Cadaveric β-cells
Transcrip level (log2)
Ctl β-cells
MODY3 β-cells

Fig. S5. Identification of scβ- and scα-like-cells and their gene regulatory network *in vitro*. **(A)** Scatter plot of 3133 hESC-derived β-like-cells from cluster 0 displayed (**Fig. S4**) based on *GCG* and *INS* expression (expo_logrpm). Total cell number from all genotypes combined. Encircled are scβ-like-cells (total of 1846 cells) with violin plot displayed based on *INS* expression (log1(rpm+1)). **(B)** Barplot showing genes differentially expressed as log2FC and enriched in specific pathway from scβ-like-cells. **(C)** Scatter plot of 2670 hESC-derived α-like-cells from cluster 2 (**Fig. S4**) displayed based on *GCG* and *INS* expression (expo_logrpm). Total cell number from all genotypes combined. Encircled are monohormonal scα-like-cells (Cluster 2a) (total of 641 cells) and bihormonal scα-like-cells (Cluster 2b) (total of 1054 cells). **(D)** Violin plot of indicated cluster of cells displayed based on *GCG* and *INS* expression (log1(rpm+1)). Monohormonal scβ-like-cells (cluster 0) and bihormonal scα-like-cells (cluster 2b) were identified and displayed as |logFC|>0.35 and adjusted p-value $<1e^{-4}$. Monohormonal scα-like-cells (cluster 2a) displayed as |logFC|>0.25 and adjusted p-value $<1e^{-2}$. **(E)** Heatmap showing differentially expressed genes from monohormonal scα-like-cells (Cluster 2a) by *HNF1A* genotypes. Total of 641 scα-like-cells (cluster 2a) (all genotypes combined) were identified and displayed as |logFC|>0.25 and adjusted p-value $<1e^{-2}$. (**F and G)** Barplot showing genes differentially expressed as log2FC and enriched in specific pathway from indicated endocrine cell type. **(H)** Heatmap showing top differentially expressed genes from bihormonal α/β-like-cells (Cluster 2b) across genotypes. Genes are listed in decreasing order of $\log_2$ fold change between *HNF1A* WT and *HNF1A* mutant genotypes. n=3 for each genotype. **(I-M)** Gene expression profile of *LINC01139* in scβ-like-cells, scα-like-cells and cadaveric β-cells. **(I)** *LINC01139* gene expression levels (FPKM) determined by bulk RNAseq and after FACs sorting of $INS^{GFP/wt}$ positive cells from hESC-derived endocrine cells (hESC WT n=3, hESC Het n=3, hESC KO n=3 and MODY3 iPSC

460ins Het n=2). **(J)** *HNF1A* gene expression levels (FPKM) determined by single-cell RNAseq after FACs sorting of $INS^{GFP/wt}$ positive cells from hESC-derived endocrine cells (hESC WT n=113 cells and hESC KO n=158 cells). **(K)** Violin plot of scβ-like-cells and **(L)** scα-like-cells based on *LINC01139* expression (log1(rpm+1)) by genotype. **(M)** *LINC01139* transcript levels of sorted pancreatic β-cells from non-diabetic donors (n=5) and MODY3 donor (n=1) (Haliyur et al. 2019). All stem cell differentiations were done for 27 days.

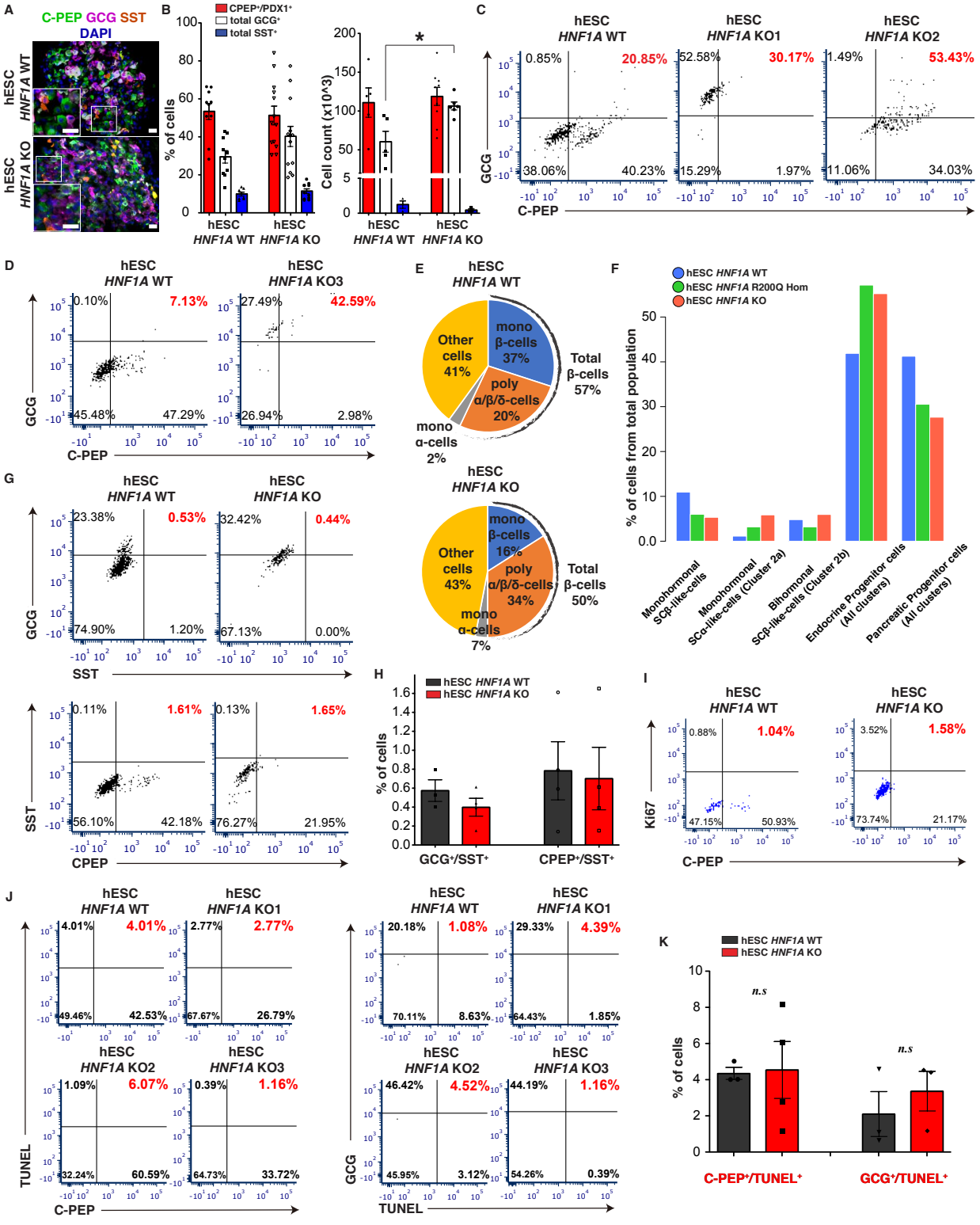


Fig. S6. *HNF1A* deficiency causes a developmental bias towards the α-cell fate *in vitro*. (A)

Representative IHC images of hESC-derived endocrine cell lines with indicated markers and **(B)**

quantification of percentage and total cell count of scβ-like-cells ($CPEP^+/PDX1^+$ in red), scα-like-cells (GCG^+ in white) and scδ-like-cells (SST^+ in blue). Cell count performed from 20 clusters. Scale bars: 20 µm. (**C and D)** Representative $CPEP^+$ and GCG^+ populations by flow cytometry from *HNF1A* KO1, KO2 and KO3. **(E)** Endocrine cells as percentage from total cells from hESC-derived clusters based on IHC staining. $CPEP^+/GCG^-/SST^-$ cells indicate monohormonal β-cells and $CPEP^+/GCG^+/SST^+$ cells indicate polyhormonal α/β/δ-cells. **(F)** Barplot showing the percentage (%) of indicated endocrine cell types from total cells (22164 cells) analyzed by single cell RNAseq from hESC-derived endocrine cells by *HNF1A* genotype (n=3 for each genotype). Endocrine progenitor cells are *SYP* expressing cells and pancreatic progenitor cells are *PDX1* expressing cells. **(G)** Representative GCG^+, SST^+ and $CPEP^+$ populations by flow cytometry with **(H)** respective quantification. **(I and J)** Flow cytometry for indicated markers with **(K)** respective quantification. All stem cell differentiations were done for 27 days. 20 clusters (~10k cells per cluster) of endocrine cells were used flow cytometry. For scatter plots, each point in plots represents an independent biological experiment (n). Data are represented as mean ± SEM. p-values: *p<0.05, **p<0.01, ***p<0.001; Mann-Whitney test. n.s: non-significant.

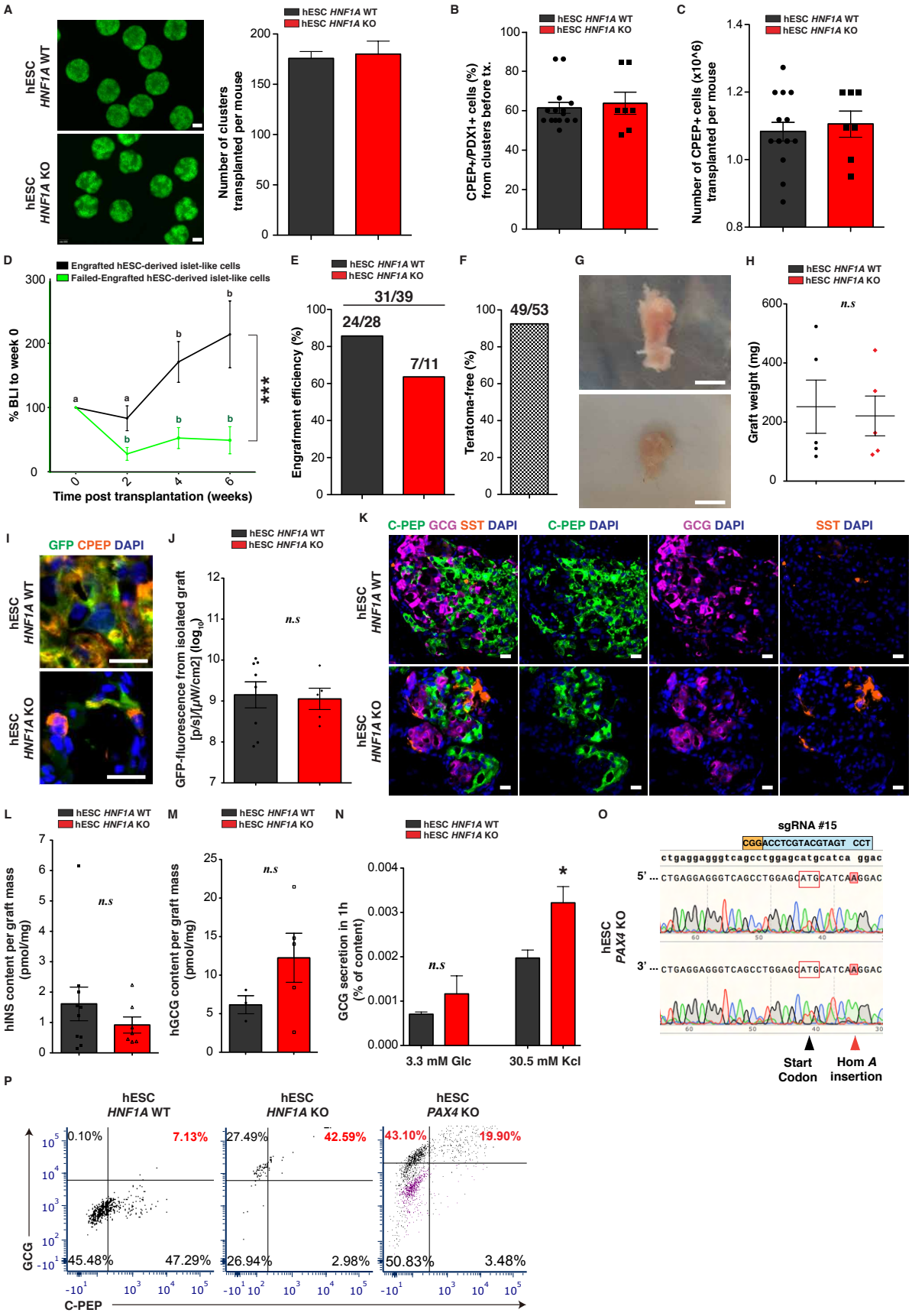

A
hESC HNF1A WT
hESC HNF1A KO
Number of clusters transplanted per mouse
B
CPEP+/PDX1+ cells (%) from clusters before tx.
C
Number of CPEP+ cells (x10^6) transplanted per mouse
D
Engrafted hESC-derived islet-like cells
Failed-Engrafted hESC-derived islet-like cells
% BLI to week 0
Time post transplantation (weeks)
***
E
31/39
24/28
7/11
Engraftment efficiency (%)
F
49/53
Teratoma-free (%)
G
H
n.s
Graft weight (mg)
I
GFP CPEP DAPI
J
GFP-fluorescence from isolated graft [p/s]/[µW/cm2] (log10)
K
C-PEP GCG SST DAPI
C-PEP DAPI
GCG DAPI
SST DAPI
L
hINS content per graft mass (pmol/mg)
M
hGCG content per graft mass (pmol/mg)
N
GCG secretion in 1h (% of content)
3.3 mM Glc
30.5 mM Kcl
O
sgRNA #15
CGGACCTCGTACGTAGT CCT
hESC PAX4 KO
Start Codon
Hom A insertion
P
hESC HNF1A WT
hESC HNF1A KO
hESC PAX4 KO
0.10%
7.13%
45.48%
47.29%
27.49%
42.59%
26.94%
2.98%
43.10%
19.90%
50.83%
3.48%
GCG
C-PEP

Fig. S7. *HNF1A* deficiency causes a developmental bias towards the α-cell fate *in vivo* (A) Representative image and number (WT n=17 and KO n=7) of pancreatic islet-like clusters transplanted per mouse and per genotype. Scale bars at 200 µm. **(B)** Percentage (%) of $CPEP^+/PDX1^+$ cells from clusters before transplantation with **(C)** total number of $CPEP^+$ cells (10x6) transplanted per mouse and per genotype. **(D)** % Bioluminescence intensity (BLI) normalized to week 0 (transplantation day) of mice transplanted with $GAPDH^{Luciferase/wt}$ reporter hESC lines 0, 2, 4 and 6 weeks post-transplantation. Engrafted hESC-derived islet-like cells (WT n=11 and KO n=4) resulted in >50 pM of circulating human c-peptide after 6 weeks. Non-engrafted hESC-derived islet-like cells (WT n=4 and KO n=6) resulted in no or <5 pM of circulating human c-peptide after 30 weeks. Results are presented as mean of mice transplanted with both hESC genotypes. p-values were b: $p<0.05$. **(E)** Engraftment efficiency (%) from all mice transplanted with islet-like cells. **(F)** Teratoma-free mice transplanted with hESC-derived islet-like cell lines. Graft tissue was considered a teratoma when larger than 2 cm and by IHC of the germ layers. **(G)** Representative image of graft tissue 30 weeks post-transplantation with hESC-derived endocrine cells after explant from quadriceps muscle. Scale bars: 0.5 cm. **(H)** Graft weight in mg from explants. **(I)** Representative IHC image from isolated grafts for indicated markers. Scale bars: 20 µm. **(J)** GFP fluorescence ($[p/s]/[\mu W/cm^2](\log_{10})$) from isolated grafts. **(K)** IHC image from isolated graft for SCβ-cells (CPEP: c-peptide), scα-cells (GCG: glucagon) and scδ-cells (SST: somatostatin). (**L and N**) Endocrine hormone content. **(L)** Human insulin and **(M)** human glucagon content per graft mass (pmol/mg) *ex vivo*. **(N)** Human glucagon secretion in 1h as % of content in static assay in response to indicated secretagogues from hESC-derived endocrine cells (WT n=4, KO n=7) *in vitro*. All protein concentrations were measured by ELISA**.** **(O)** Sanger sequencing results showing knockout frameshift mutation (A insertion) in hESC *PAX4*

KO from both alleles in exon3 near start codon (red box). **(P)** Representative flow cytometry of $CPEP^+$ and CGC^+ populations. Gating for GCG and CPEP negative cells (magenta) was determined by incubating cells without primary antibodies and with secondary antibodies. 20 clusters (~10k cells per cluster) of endocrine cells were used flow cytometry. All stem cell differentiations were done for 27 days for *in vitro* assay and for transplantation. All grafts were isolated 30 weeks post-transplantation for *ex vivo* analysis. For scatter plots, each point in plots represents an independent biological experiment (n). Data are represented as mean ± SEM. Different letters designate significant differences within group. p-values: *p<0.05, **p<0.01, ***p<0.001; Mann-Whitney test. n.s: non-significant.

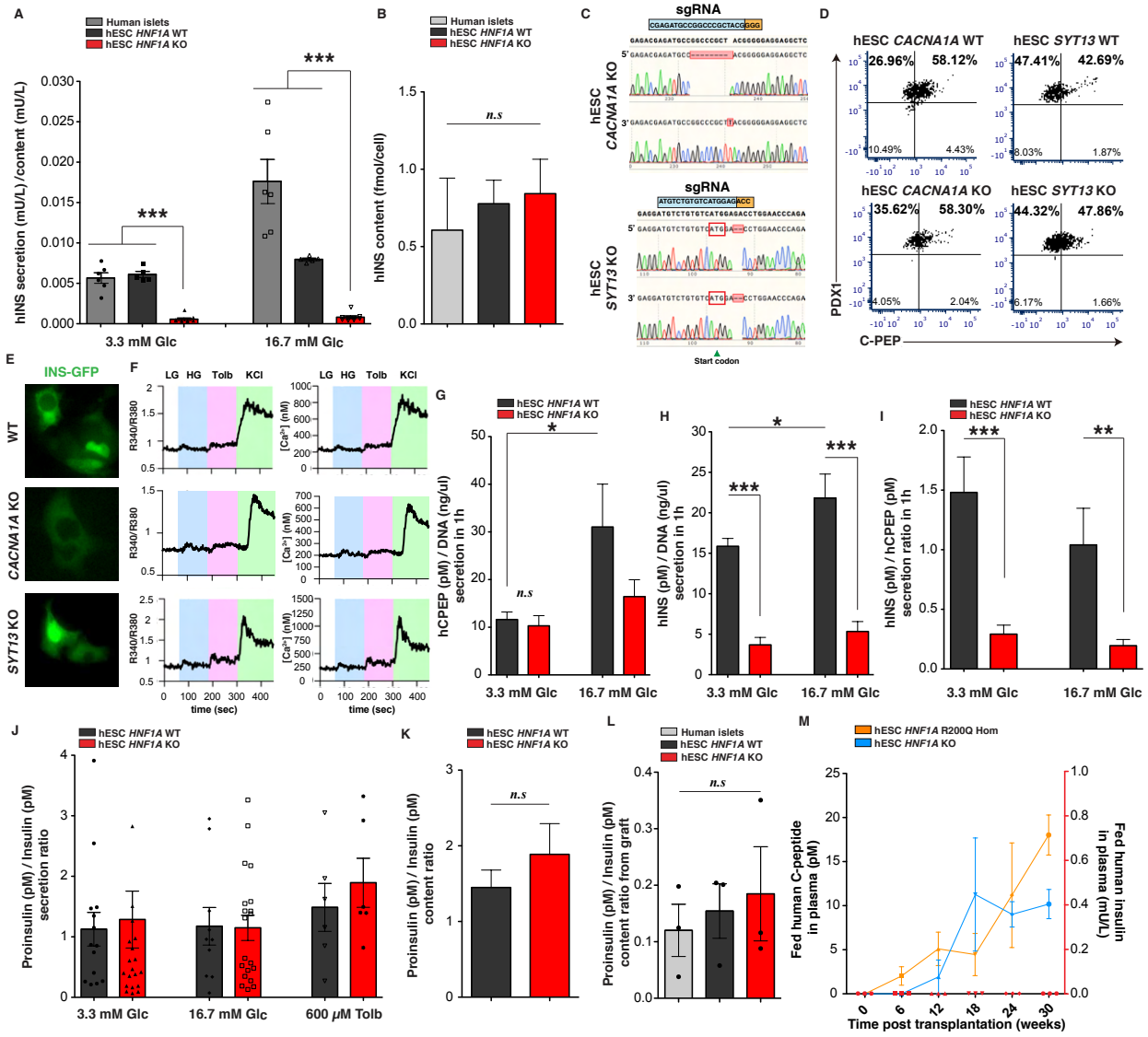


Fig. S8. *HNF1A* deficiency affects insulin secretion in association with *CACNA1A* and *SYT13* down-regulation *in vitro*. (A) Human insulin secretion (mU/L) in 1h normalized to content (mU/L) in response to low glucose (LG, 3.3 mM) and high glucose (HG, 16.7 mM) stimulation in static assay. **(B)** Human insulin protein content per cell (fmol/cell) (Human islets n=6, WT n=17, KO n=21). **(C)** Sanger sequencing results showing knockout frameshift mutations in hESC *CACNA1A* KO and hESC *SYT13* KO line from both alleles near start codon (red box). **(D)** Representative flow cytometry of $CPEP^{+}$ and $PDX1^{+}$ populations. **(E)** dispersed scβ-like-cells

(INS-GFP) with **(F)** representative intracellular calcium levels measured as R340/R380 and [Ca^{2+}](nM) over time (sec) with indicated secretagogues. **(G)** Human c-peptide (pM) and **(H)** human insulin secretion (pM) in 1h normalized to DNA (ng/µl) in response to basal (3.3 mM) and high (16.7 mM) glucose stimulation in static assay (WT n= 7 and KO n=20). **(I)** hINS(pM)/hCPEP(pM) secretion ratio in response to basal (3.3 mM) and high (16.7 mM) glucose stimulation (WT n=9 and KO n=14). **(J-L)** Proinsulin(pM)/Insulin(pM) ratio from **(J)** secretion and **(K)** content from *in vitro* cells (WT n=8 and KO n=15) and **(L)** *ex vivo* grafts. (**M**) Plasma human C-peptide (pM) (left y-axis) and human insulin secretion (mU/L) (right y-axis) in plasma of *ad libitum-fed* mice transplanted with hESC *HNF1A* KO1 and *HNF1A* R200Q homozygous-derived endocrine cells (n=3). All stem cell differentiations were done for 27 days. All protein concentrations were measured by ELISA. For scatter plots, each point in plots represents an independent biological experiment (n). Data are represented as mean ± SEM. p-values: *p<0.05, **p<0.01, ***p<0.001; Mann-Whitney test. n.s: non-significant.

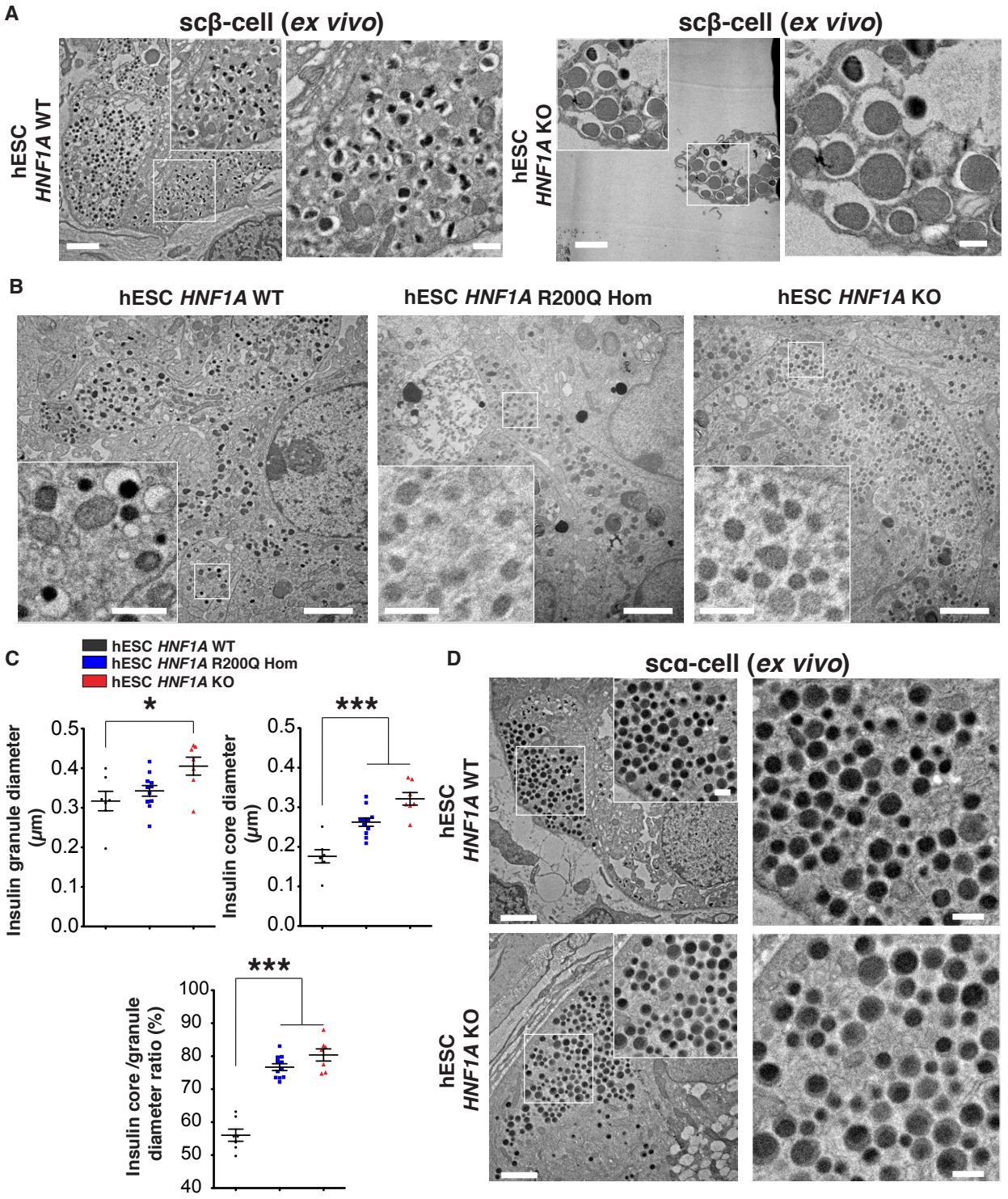


Fig. S9. ***HNF1A*** **deficiency causes abnormal insulin granule structure.** (**A**) Representative electron microcopy (EMC) image of scβ-cells *ex vivo* from isolated graft 30 weeks post-transplantation. Explants are from euglycemic mice. **(B)** Representative EMC images of scβ-like-

cells *in vitro* with **(C)** quantification of insulin granule diameter (µm), insulin granule core diameter (µm) and insulin granule core diameter to insulin granule diameter ratio (%) per cell. **(D)** Representative EMC images of scα-cells. For all images, scale bars: 2 µm in low and 0.5 µm in high magnification. Each point in plots is the average of insulin granules per scβ-like-cells. All stem cell differentiations were done for 27 days for *in vitro* assay and for transplantation. For scatter plots, each point in plots represents an independent biological experiment (n). Data are represented as mean ± SEM. p-value: *p<0.05, **p<0.01, ***p<0.001; two-tailed t-test. n.s: non-significant.

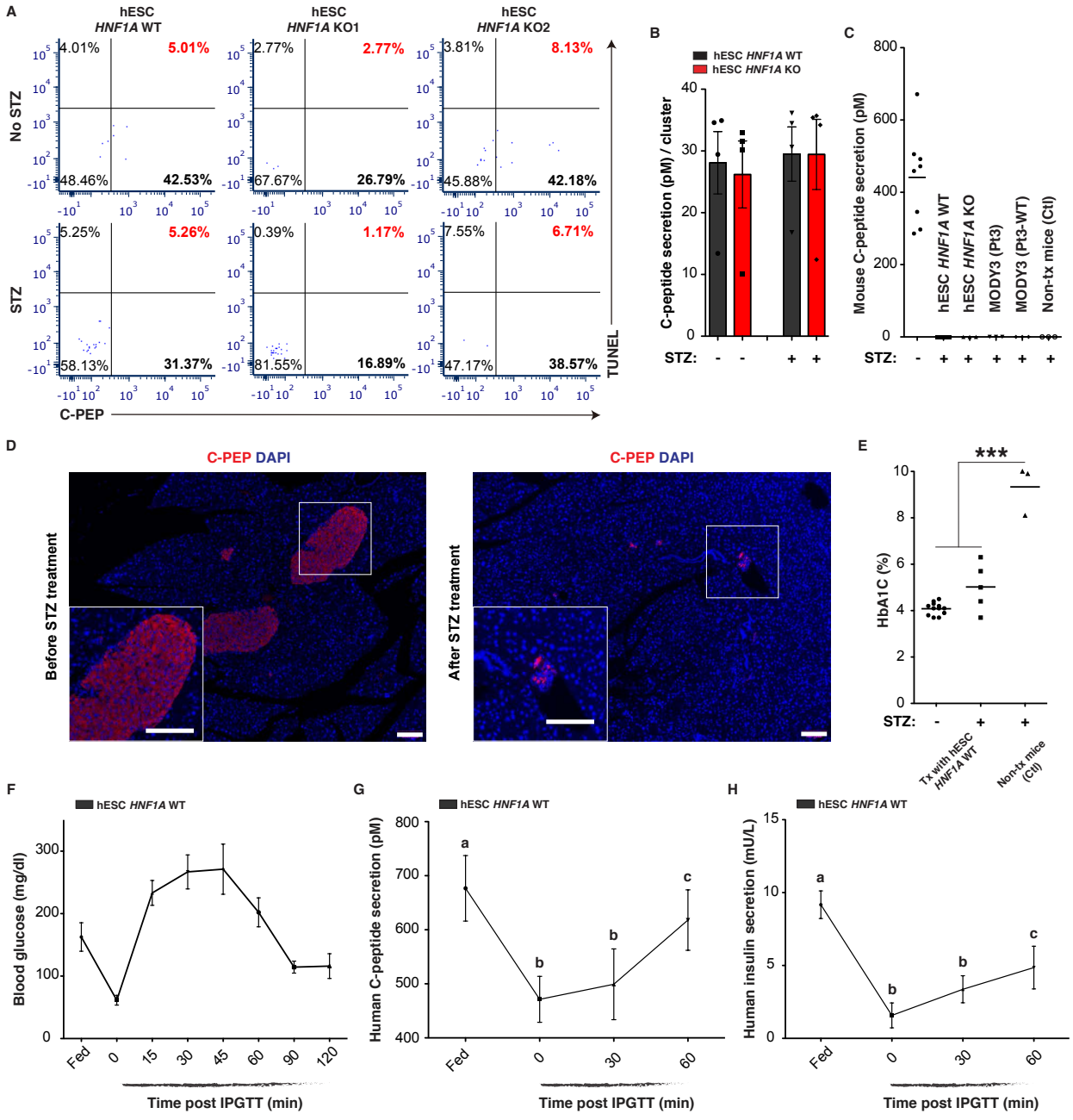


Fig. S10. *HNF1A* deficient scβ-cells are unable to maintain glucose homeostasis in diabetic mice. (A and B) hESC-derived endocrine cells were treated with (4mg/ml; 15mM) or without STZ *in vitro*. **(A)** Apoptosis quantification after flow cytometry of TUNEL^{+} and CPEP^{+} cells. 10 clusters (~10k cells per cluster) of endocrine cells were used flow cytometry. **(B)** Human c-peptide

secretion (pM) per cluster measured before and after STZ *in vitro*. **(C)** Mouse C-peptide secretion (pM) in plasma of *ad libitum-fed* mice transplanted with indicated cell lines or non-transplanted mice (Ctl) before (-) or >3 weeks after STZ injection (+). **(D)** Representative immunofluorescent staining of mouse pancreas showing islets (CPEP; red) before and after STZ treatment. Scale bars: 20 µm. **(E)** HbA1C (%) in mice transplanted with hESC *HNF1A* WT-derived endocrine cells or without cells (Ctl) before (-) and 5-15 weeks after (+) STZ treatment. (**F-H**) IPGTT in mice transplanted with hESC *HNF1A* WT-derived endocrine cells (n=6) several weeks (>3 weeks) after STZ treatment in *ad libitum-fed* state and during an iPGTT (t0, t30 and t60). **(F)** Blood glucose concentrations (mg/dl), **(G)** human C-peptide (pM) and **(H)** human insulin secretion (mU/L) in plasma. p-values were b: $p<0.001$, c: $p<0.05$. All stem cell differentiations were done for 27 days. For scatter plots, each point in plots represents an independent biological experiment (n). Data are represented as mean ± SEM. Different letters designate significant differences within group. p-values: *$p<0.05$, **$p<0.01$, ***$p<0.001$; two-tailed t-test.

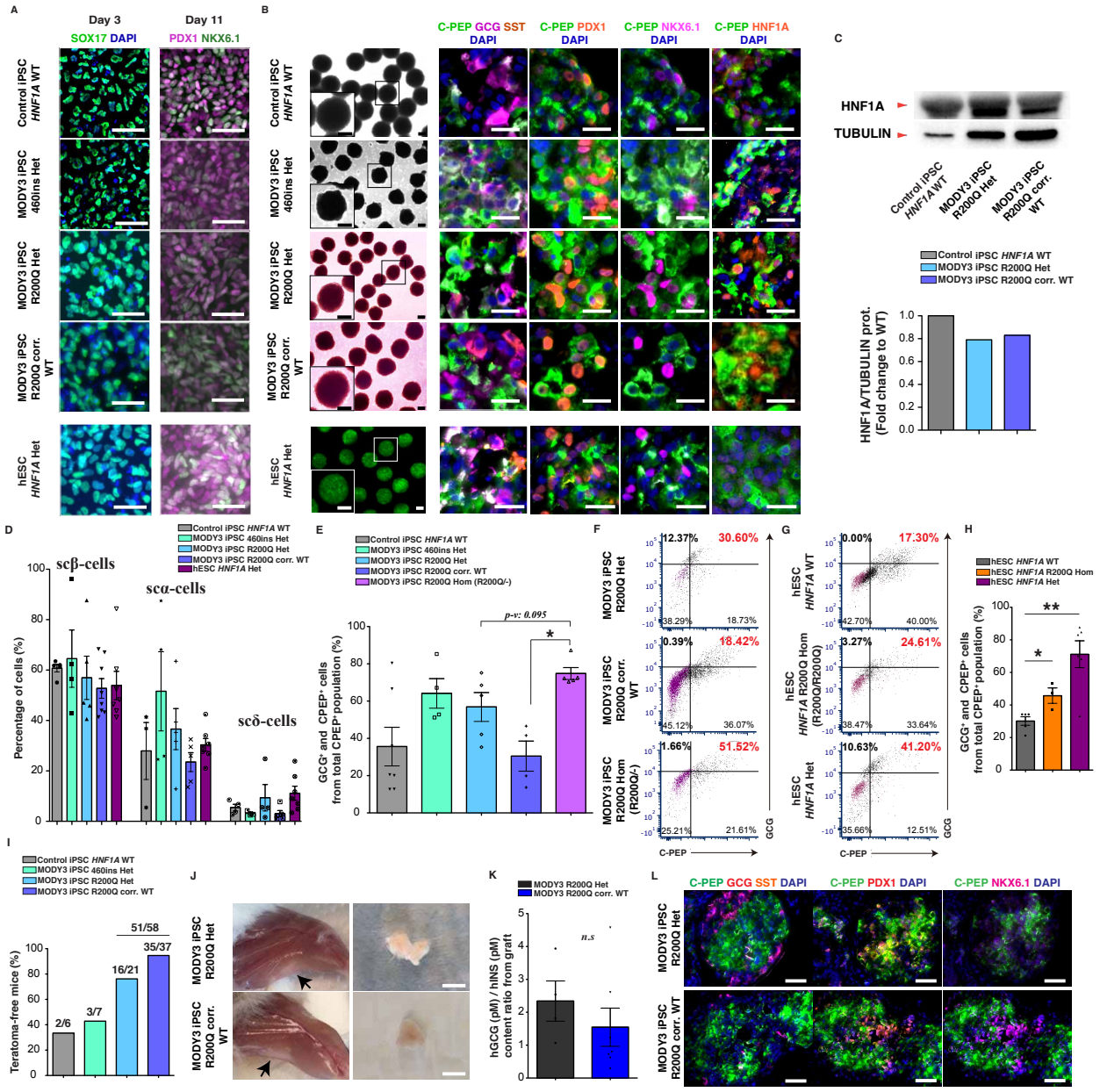


Fig. S11. *HNF1A* R200Q mutation is pathogenic and causes developmental bias towards the α-cell fate *in vitro*. (A) Representative IHC staining of stem cell-derived cell lines *in vitro* at definitive endoderm stage with SOX17 (day 3), pancreatic progenitor stage with PDX1 and NKX6.1 (day 11) and endocrine cell stage (day 27). Scale bars: 50 µm. White cells are PDX1/NKX6.1 double positive cells. **(B)** Representative bright-field images from MODY3 iPSC-derived organoids and GFP image from hESC Het-derived organoids with IHC staining for

indicated markers. White cells are GCG/CPEP double positive cells. Scale bars: 20 µm. **(C)** Western blot for HNF1A (67 kDa) and β-Tubulin-III (55 kDa) from iPSC-derived endocrine cells with respective quantification on day 27 of differentiation *in vitro*. **(D)** Percentage of scβ-like-cells ($CPEP^+/PDX1^+$), scα-like-cells (GCG^+) and scδ-like-cells (SST^+) determined by IHC. **(E)** Quantification of GCG and CPEP positive cells from total CPEP positive cell population determined by IHC. (**F)** $CPEP^+$ and GCG^+ populations from iPSC-derived endocrine cells and **(G)** hESC-derived endocrine ells by flow cytometry. Gating for GCG and CPEP negative cells (magenta) was determined by incubating cells without primary antibodies and with secondary antibodies. **(H)** Quantification of GCG and CPEP positive cells from total CPEP positive cell population (%) determined by flow cytometry. **(I)** Teratoma-free mice transplanted with MODY3 iPSC-derived endocrine cells. Graft tissue was considered a teratoma when larger than 2 cm and by IHC of the germ layers. **(J)** Representative image of graft tissue (black arrows) 30 weeks post-transplantation with MODY3 iPSC-derived endocrine cells before and after explant from quadriceps muscle. Scale bars: 0.5 cm. **(K)** Human glucagon content per graft mass (pmol/mg) *ex vivo* measured by ELISA. **(L)** Representative IHC images showing MODY3 iPSC-derived endocrine cells from isolated graft. Scale: 50 µm. All stem cell differentiations were done for 27 days for *in vitro* assay and for transplantation. All grafts were isolated 30 weeks post-transplantation for *ex vivo* analysis. For scatter plots, each point in plots represents an independent biological experiment (n). Data are represented as mean ± SEM. Different letters designate significant differences within group. p-values: *p<0.05, **p<0.01, ***p<0.001; Mann-Whitney test. n.s: non-significant.

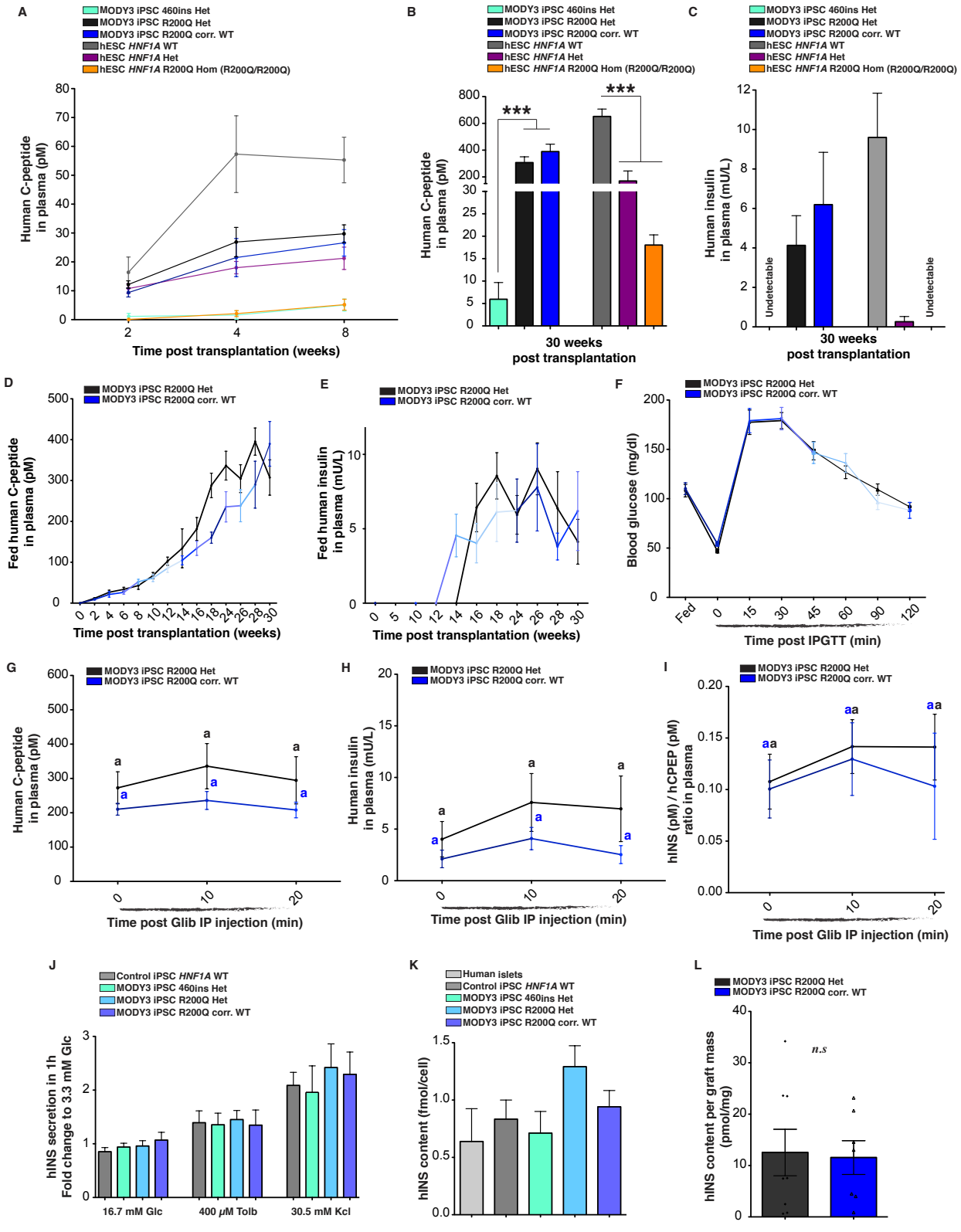

A
MODY3 iPSC 460ins Het
MODY3 iPSC R200Q Het
MODY3 iPSC R200Q corr. WT
hESC HNF1A WT
hESC HNF1A Het
hESC HNF1A R200Q Hom (R200Q/R200Q)
Human C-peptide in plasma (pM)
Time post transplantation (weeks)
B
***
***
Human C-peptide in plasma (pM)
30 weeks post transplantation
C
Human insulin in plasma (mU/L)
Undetectable
Undetectable
30 weeks post transplantation
D
Fed human C-peptide in plasma (pM)
Time post transplantation (weeks)
E
Fed human insulin in plasma (mU/L)
Time post transplantation (weeks)
F
Blood glucose (mg/dl)
Fed
Time post IPGTT (min)
G
Human C-peptide in plasma (pM)
Time post Glib IP injection (min)
H
Human insulin in plasma (mU/L)
Time post Glib IP injection (min)
I
hINS (pM) / hCPEP (pM) ratio in plasma
Time post Glib IP injection (min)
J
Control iPSC HNF1A WT
MODY3 iPSC 460ins Het
MODY3 iPSC R200Q Het
MODY3 iPSC R200Q corr. WT
hINS secretion in 1h Fold change to 3.3 mM Glc
16.7 mM Glc
400 μM Tolb
30.5 mM Kcl
K
Human islets
Control iPSC HNF1A WT
MODY3 iPSC 460ins Het
MODY3 iPSC R200Q Het
MODY3 iPSC R200Q corr. WT
hINS content (fmol/cell)
L
hINS content per graft mass (pmol/mg)
n.s

Fig. S12. MODY3 iPSC-derived β-cells are initially glucose responsive. (**A-C**) Human C-peptide (pM) and human insulin levels in plasma of *ad libitum-fed* mice transplanted with MODY3 iPSC- and hESC-derived endocrine cells**. (A)** Plasma human C-peptide (pM) two, four and eight weeks post transplantation in mice transplanted with MODY3 iPSC lines (460ins Het n=4, R200Q Het n=12 and R200Q corr. WT n=12) and hESC lines (WT n=17, Het n=10 and R200Q Hom n=6). **(B)** Plasma human C-peptide (pM) thirty weeks post transplantation in mice transplanted with MODY3 iPSC lines (460ins Het n= 5, R200Q Het n=5 and R200Q corr. WT n=6) and hESC lines (WT n=7, Het n=4 and R200Q Hom n=3) and **(C)** human insulin (mU/L) thirty weeks post transplantation in mice transplanted with MODY3 iPSC lines (460ins Het n= 3, R200Q Het n=4 and R200Q corr. WT n=6) and hESC lines (WT n=8, Het n=7 and R200Q Hom n=3). **(D)** Plasma human C-peptide (pM) and **(E)** human insulin secretion (mU/L) in plasma of *ad libitum-fed* mice transplanted with MODY3 iPSC-derived endocrine cells (R200Q Het n=14 and R200Q corr. WT n=12). **(F)** Blood glucose concentrations (mg/dl) in *ad libitum-fed* state and during an iPGTT (t0, t30 and t60) in mice with MODY3 iPSC-derived endocrine cells (R200Q Het n=12 and R200Q corr. WT n=10). **(G-I**) Glibenclamide injection in *ad libitum-fed* mice transplanted with MODY3 iPSC-derived endocrine cells (R200Q Het n=5 and R200Q corr. WT n=9). **(G)** Human C-peptide secretion (pM), **(H)** human insulin secretion (mU/L) and **(I)** hINS (pM)/hCPEP (pM) secretion ratios. **(J)** Human insulin secretion in 1h from static assay as a fold change to 3.3 mM Glc in response to indicated secretagogues *in vitro* (iPSC WT n=5, 460ins Het n=5, R200Q Het n=10 and R200Q corr. WT n=7). **(K)** Human insulin protein content per cell (fmol/cell) *in vitro* (Human islets n=6, iPSC WT n=4, 460ins Het n=4, R200Q Het n=7 and R200Q corr. WT n=5). **(L)** Human insulin content per graft mass (pmol/mg). All stem cell differentiations were done for 27 days for *in vitro* assay and for transplantation. All grafts were isolated 30 weeks post-transplantation for *ex*

vivo analysis. All protein concentrations were measured by ELISA. For scatter plots, each point in plots represents an independent biological experiment (n). Data are represented as mean ± SEM. Different letters designate significant differences within group. p-values: *p<0.05, **p<0.01, ***p<0.001; two-tailed t-test. n.s: non-significant.

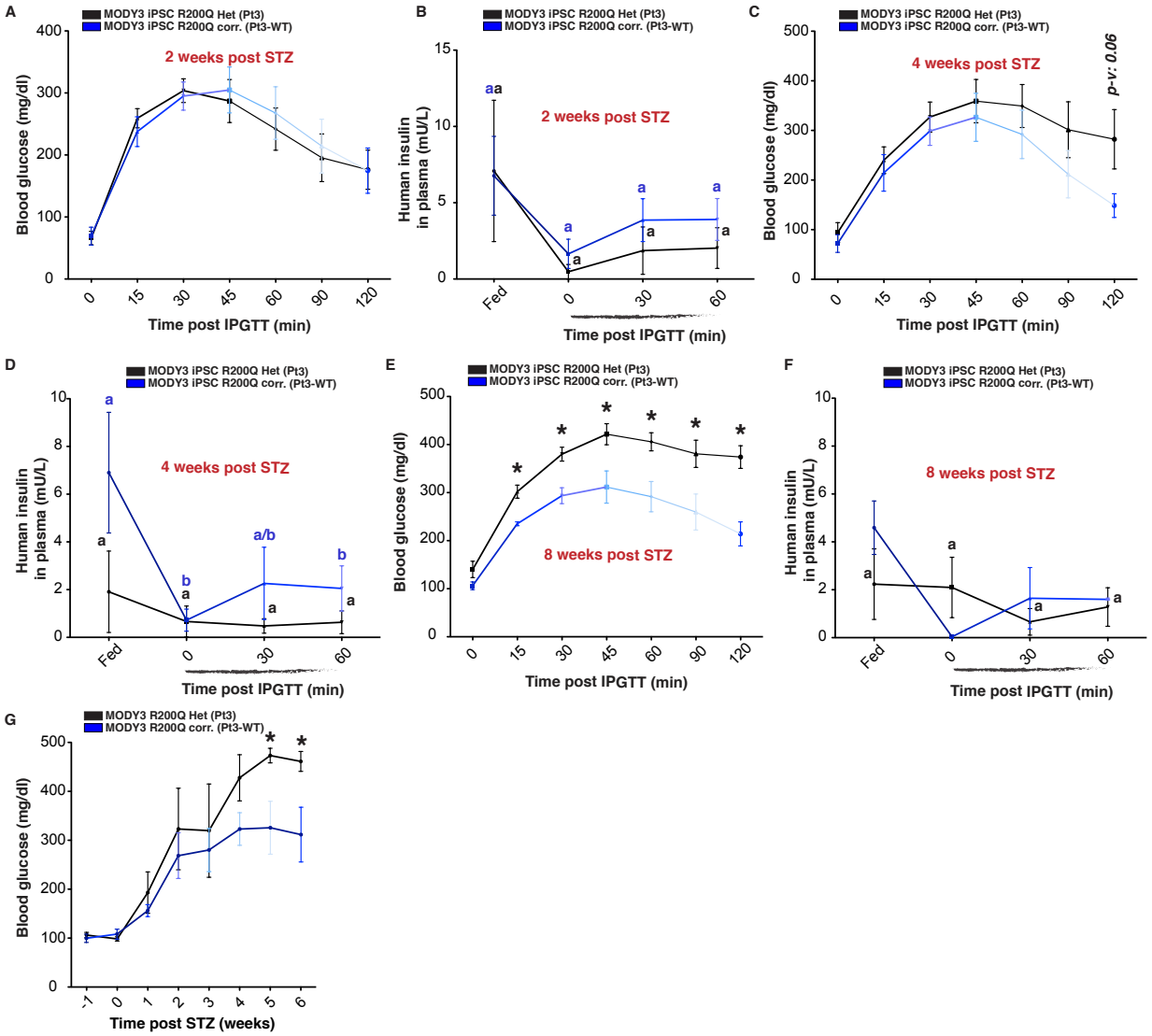


Fig. S13. ***HNF1A*** **haploinsufficiency gradually impairs scβ-cell function** ***in vivo.*** (**A-F**) IPGTT in mice transplanted with MODY3 iPSC-derived endocrine cells in *ad libitum-fed* state and during an iPGTT (t0, t30 and t60) 2, 4 and 8 weeks post STZ treatment. **(A)** Blood glucose concentrations (mg/dl) and **(B)** human insulin secretion (mU/L) in plasma of mice after 2 weeks post STZ (R200Q Het n=12; R200Q corr. WT n=10). **(C)** Blood glucose concentrations (mg/dl) and **(D)** human insulin secretion (mU/L) in plasma of mice after 4 weeks post STZ (R200Q Het n=5; R200Q corr. WT n=4). **(E)** Blood glucose concentrations (mg/dl) and **(F)** human insulin secretion (mU/L) in plasma of mice after 8 weeks post STZ (R200Q Het n=5; R200Q corr. WT n=2). p-values were b

and c: $p<0.05$. **(G)** Blood glucose levels (mg/dl) monitored in *ad libitum-fed* mice weeks before (-1) or after (0 to 6) STZ injection (R200Q Het n=10; R200Q corr. Wt n=7). All protein concentrations were measured by ELISA. For scatter plots, each point in plots represents an independent biological experiment (n). Data are represented as mean ± SEM. Different letters designate significant differences between fed, t0, t30 and t60 for each genotype. p-values: *$p<0.05$, **$p<0.01$, ***$p<0.001$; two-tailed t-test.