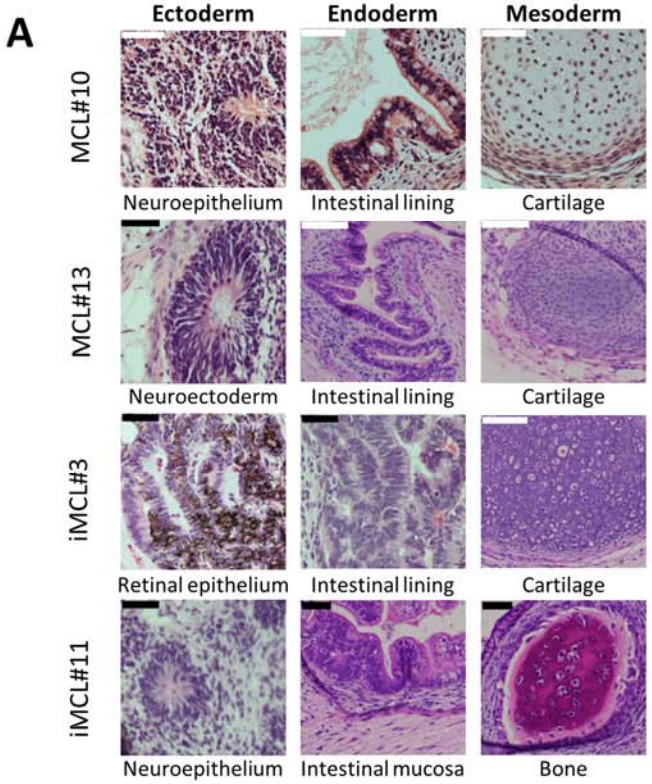
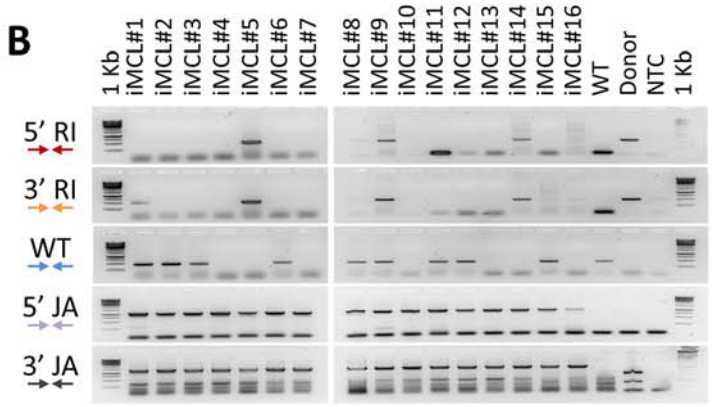
Stem Cell Reports, Volume *5*

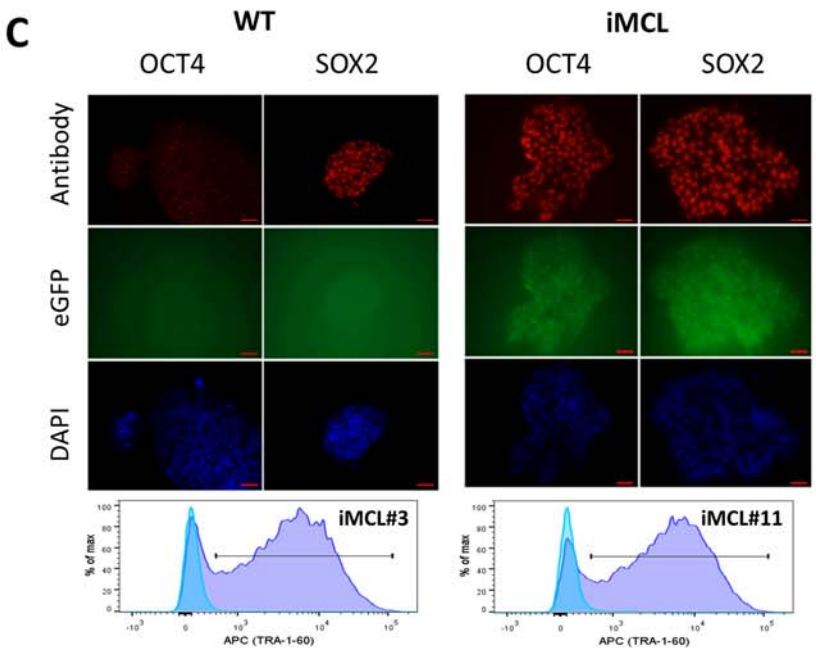
Supplemental Information

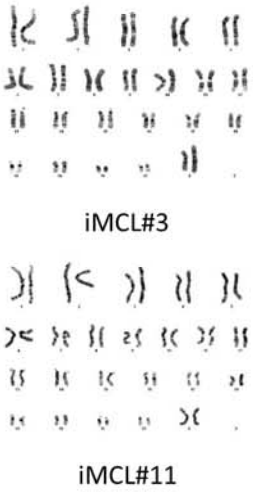
Efficient Recombinase-Mediated Cassette Exchange in hPSCs to Study the Hepatocyte Lineage Reveals *AAVS1* Locus-Mediated Transgene Inhibition

Laura Ordovás, Ruben Boon, Mariaelena Pistoni, Yemiao Chen, Esther Wolfs, Wenting Guo, Rangarajan Sambathkumar, Sylwia Bobis-Wozowicz, Nicky Helsen, Jolien Vanhove, Pieter Berckmans, Qing Cai, Kim Vanuytsel, Kristel Eggermont, Veerle Vanslembrouck, Béla Z. Schmidt, Susanna Raitano, Ludo Van Den Bosch, Yaakov Nahmias, Toni Cathomen, Tom Struys, and Catherine M. Verfaillie


A
Ectoderm
Endoderm
Mesoderm
MCL#10
Neuroepithelium
Intestinal lining
Cartilage
MCL#13
Neuroectoderm
Intestinal lining
Cartilage
iMCL#3
Retinal epithelium
Intestinal lining
Cartilage
iMCL#11
Neuroepithelium
Intestinal mucosa
Bone


B
1 Kb
iMCL#1
iMCL#2
iMCL#3
iMCL#4
iMCL#5
iMCL#6
iMCL#7
iMCL#8
iMCL#9
iMCL#10
iMCL#11
iMCL#12
iMCL#13
iMCL#14
iMCL#15
iMCL#16
WT
Donor
NTC
1 Kb
5′ RI
3′ RI
WT
5′ JA
3′ JA


C
WT
iMCL
OCT4
SOX2
OCT4
SOX2
Antibody
eGFP
DAPI
iMCL#3
iMCL#11
% of max
APC (TRA-1-60)


D
iMCL#3
iMCL#11

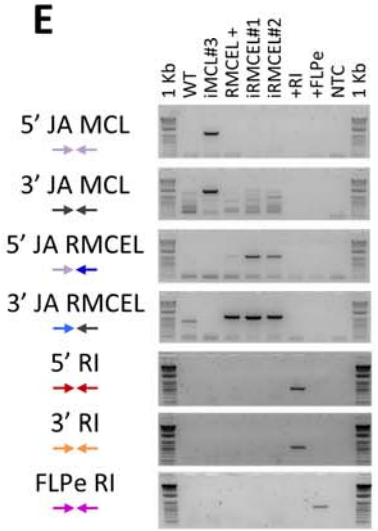

E
1 Kb
WT
iMCL#3
RMCEL +
iRMCEL#1
iRMCEL#2
+RI
+FLPe
NTC
1 Kb
5′ JA MCL
3′ JA MCL
5′ JA RMCEL
3′ JA RMCEL
5′ RI
3′ RI
FLPe RI

Figure S1, related to Figure 1. Generation and characterization of FRT-containing hiPSC master cell lines and validation of RMCE. (**a**) Histologic analysis of the teratoma demonstrating tissue derivatives representative of all three germ layers in two ESC (MCL) and two iPSC (iMCL) master cell line clones. Scale bars 50 µm (black) and 100 µm (white). (**b**) PCR genotyping of hygromycin resistant clones using primer sets depicted in Figure 1A. (**c**) Expression of pluripotency markers (OCT4 and SOX2) and GFP in wild type iPSC (WT) and a representative clone (iMCL). Scale bar 20 µm. Bottom, FACS quantification of TRA-1-60 expression in two master cell line (MCL) clones of hiPSC.). (**d**) Karyogram of two representative MCL clones. (**e**) PCR characterization of RMCE (n=2) of all $Puro^R$/$FIAU^R$ cells without colony selection using primer sets depicted in Figure 2a for 5'/3' junction assays (JA) of the MCL and RMCE lines (RMCEL), 5'/3' random integration (RI) of the donor or FPLe expressing vector. Wild type (WT), a MCL clone (iPSC#3), a positive control for RMCE (H9-RMCEL +), random integration (+RI) and pFLPe integration (+FLPe) as well as a negative control template (NTC) were included.

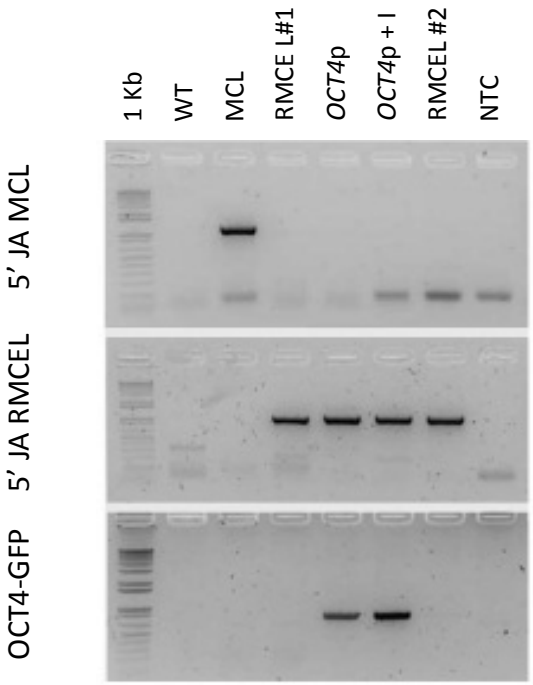


Figure S2, related to Figure 3. PCR characterization of the *OCT4*p-GFP RMCE lines. PCR amplification of the 5' junction of the master cell line (5' JA MCL) and the RMCE line (5' JA RMCEL) and the *OCT4-GFP* junction in wild type (WT), MCL, RMCEL and *OCT4*p-GFP lines without (OCT4p) or with insulators (OCT4p + I). Negative template control (NTC) is included.

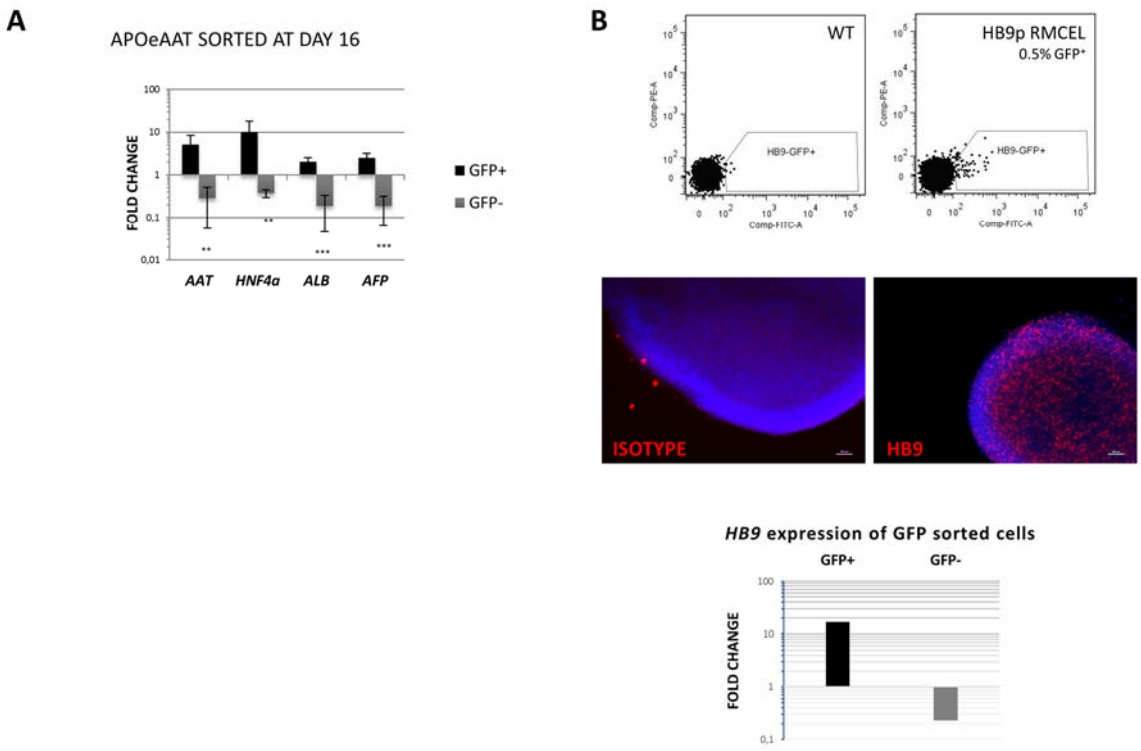


Figure S3, related to Figure 4. Characterization of lineage-specific promoters. a) mRNA expression profile of *APOeAAT*p-GFP^{+} and *APOeAAT*p-GFP^{-} sorted cells on day 16 represented as relative gene expression to the unsorted cell population (n=3). **b**) Top: GFP reporter activity during motor neuron differentiation of the *HB9*p RMCE line. Middle: immunocytochemistry of HB9 expression of day 35 motor neuron progeny (scale bar 100 µm). Bottom: *HB9* mRNA expression relative to the unsorted cell population of GFP+ and GFP- sorted cells (n≤2 IE).

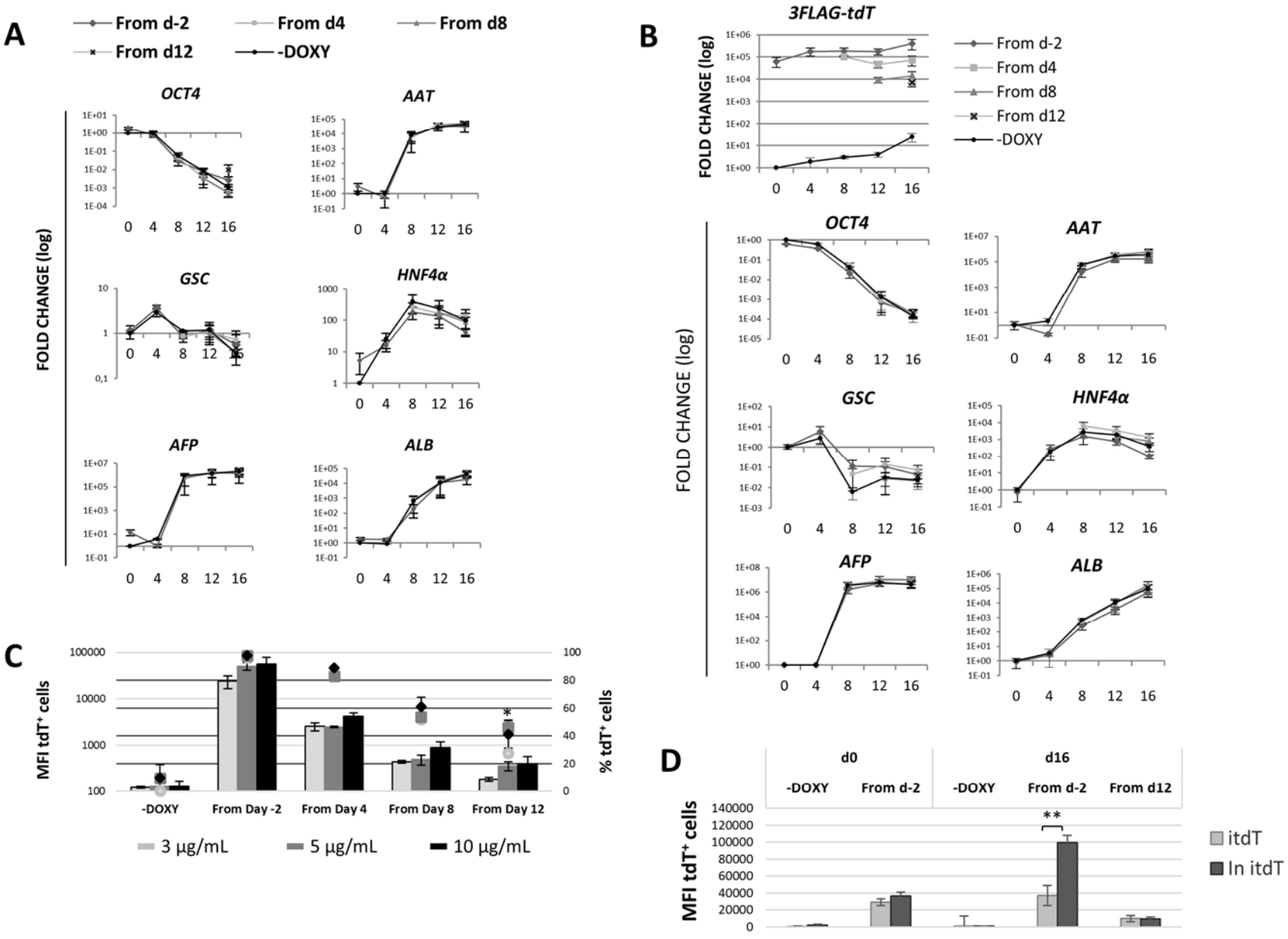

A
From d-2
From d4
From d8
From d12
-DOXY
OCT4
AAT
GSC
HNF4α
AFP
ALB
FOLD CHANGE (log)
B
3FLAG-tdT
C
MFI tdT+ cells
% tdT+ cells
-DOXY
From Day -2
From Day 4
From Day 8
From Day 12
3 µg/mL
5 µg/mL
10 µg/mL
D
d0
d16
-DOXY
From d-2
From d12
itdT
In itdT

Figure S4, related to Figure 6. Inducible expression from the *AAVS1* locus during hepatocyte differentiation. (**a, b**) Gene expression profile, relative to undifferentiated ESC (day 0), of differentiation markers and *3FLAG-tdT* during hepatocyte commitment without (-DOXY) or with 3 µg/mL (a) or 10 µg/mL (b) DOXY starting from the indicated time points. (**c**) Quantification of tdT expression in day 16 hepatocyte progeny in the absence (-DOXY) or presence of 3, 5 or 10 µg/mL DOXY from the indicated time points. Results are represented as MFI of tdT^+ cells (bars) and percentage of tdT^+ cells (markers). (**d**) Response to 3 µg/mL DOXY, initiated at the indicated time points, of the RMCELs without (itdT) and with insulators (In itdT) on day 0 or day 16 determined by FACS and expressed as mean fluorescence intensity (MFI). Data represent the mean ± sem of n=3 and *$p<0.05$ by Student's t-test.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES.

HUMAN ESC/iPSC CULTURE AND DIFFERENTIATION

The hESC line H9 (WA09) was purchased from WiCell Research Institute and maintained on inactivated mouse embryonic fibroblasts (iMEF) in hESC medium (DMEM/F12 (Invitrogen) supplemented with 20% KnockOut Serum Replacement (Invitrogen), 1% nonessential amino acids (Invitrogen), 1mM L-glutamine (Sigma-Aldrich), 0.1 mM β-mercaptoethanol (Sigma-Aldrich) and 4 ng/ml bFGF (Peprotech). The hiPSC line was generated by transduction of BJ1 fibroblasts with retroviral vectors encoding for OKSM (Takahashi et al., 2007) and was maintained in feeder free conditions using mTeSR1 medium (Stemcell Technologies) on hESC-qualified matrigel (Becton Dickinson).

Hepatic differentiation was sequentially induced using a combination of cytokines. Briefly, cells were induced to differentiate to definitive endoderm with Activin A 50 ng/mL for 4 days and Wnt3a 50 ng/mL for the initial 2 days, to hepatic endoderm with BMP4 (50 ng/mL) for 4 days, to hepatoblasts with FGF1 50 ng/mL for 4 days and finally to hepatocyte-like cells with HGF 20 ng/mL (Figure 3d). All growth factors were purchased from Peprotech.

GENERATION OF THE MASTER CELL LINES AND RMCE LINES

For the generation of the master cell line, the *pZ:F3-CAGGS GPHTK-F* gene targeting vector was constructed using the *pZ* donor *AAVS1* vector (CompoZr Targeted Integration Kit, Sigma Aldrich), which contains the homology arm (HA) regions to the *AAVS1* locus used for recombination and a reporter-selectable cassette constituted of a CAGGS promoter driving a GFP-F2A-Hygromycin resistance/Thymidine kinase unit, flanked by heterotypic FRT sequences (Figure 1a).

The RMCE vectors were generated using the pZ donor *AAVS1* puromycin vector (Sigma Aldrich) modified to contain the heterotypic F3-F FRT sequences and to maintaining the homology regions to the *AAVS1* locus (intermediate cloning vector *pZ:F3-P F*). In this manner the vector is suitable for RMCE but maintains its utility for gene targeting becoming a dual-purpose vector. This is of interest when functionality of a construct has to be tested in multiple cell lines, not all suitable for RMCE (Raitano et al., 2015). The vector is constituted by two units: the first one is constant for all RMCE vectors and harbors a splicing acceptor (SA), a T2A self-cleaving peptide and the puromycin resistance gene followed by a polyadenylation signal. The second unit consist of the CAGGS promoter driving a P2A-linked tdT and hygromycin resistance genes followed by a polyadenylation sequence.

The RMCE vector *pZ:F3-P CAGGS tdTPH-F* used to optimize RMCE consists of two functional cassettes (Figure 2a). The FLPe expressing (p*FLPe)* plasmid (Thermo Scientific, Open Biosystems) was modified to remove the IRES-Puromycin resistance gene included downstream of the *FLPe* coding sequence.

Lineage-tracing RMCE vectors were first constructed for the *OCT4* promoter (*OCT4*p) using the *pZ:F3-P F* vector with or without two inverted tandem repeat copies of the cHS4 insulator, *pZ:F3-P (cHS4)X4 OCT4p-GFP-F* and *pZ:F3-P OCT4p-GFP-F*, respectively (Figure 3a). The *OCT4*p followed by GFP was obtained from the ph*OCT4*-EGFP vector (Addgene 38776) (Gerrard et al., 2005) and the cHS4 insulator sequences were kindly

provided by Prof. Brindley (Suttiprapa et al., 2012) (George Washington University, Washington DC, USA). Additional promoters for *GSC* (fragment -2.2 Kb from translation start codon, Supplemental Information), *APO*e*AAT*p, *AAT*p, *AFPep* and *HB9*, were incorporated into the *pZ:F3-P (cHS4)X4 OCT4p-GFP-F* vector replacing the *OCT4p*. *APO*e*AAT*p (Lam et al., 2007), *AAT*p (Duan et al., 2007) and *AFPep* (Ishii et al., 2008) were kindly provided by Prof. Lam (National Cancer Centre, Singapore), Prof. Zern (University of California Davis Medical Center, California, USA) and Prof. Uemoto (Kyoto University, Kyoto, Japan), respectively. The *HB9* promoter was acquired from the lenti-HB9-GFP vector (Addgene 37080) (Marchetto et al., 2008). All the promoter-GFP units were cloned in antisense orientation with respect to the expression direction of the *PPP1R12C* gene to minimize possible transcriptional interference effects with the locus or with the puromycin resistance gene.

The NF-κB sensor was generated by cloning of 4 tandem copies of the NF-κB transcription response element GGGAATTTCC upstream of the CMV minimal promoter to control 3flag-tdT expression-polyA into the pZ:F3-P F vector (*pZ:F3-P* NF-κB tdT-*F,* Figure 5a). The *pZ:F3-P (cHS4)X4* NF-κB tdT-*F* vector (Figure 5f) was generated by flanking the NF-κB sensor with insulators in the *pZ:F3-P* NF-κB tdT-*F* vector.

The inducible expression vector *pZ:F3-P TetOn 3f-tdT-F* was generated into the *pZ:F3-P F* vector to contain the Tetracycline Response Element (TRE) from pTRIPZ (Open Biosystems) (Figure 6a). The "all-in-one" vector contains the *TRE* driving the expression of *3flag-tdT* in reverse orientation to the *CAGGS m2rtTA* cassette. The *pZ:F3-P (cHS4)X4 TetOn 3f-tdT-F* was generated by insertion of cHS4 insulators flanking the TetOn tdT cassette in the *pZ:F3-P TetOn 3f-tdT-F* vector (Figure 6e).

In general, no large blocks of prokaryotic sequences, known to be affected by epigenetic inactivation, are included in the vectors. When used, codon adapted sequences for appropriate expression in eukaryotic cells were utilized.

RMCE was performed by nucleofection of the dual RMCE vectors and 2.5-3 µg *pFLPe* in a 2:1 RMCE donor:p*FLPe* molecular ratio as described above. hESC were plated on iMEF from the DR4 mouse strain which possess puromycin resistance. Selection was started 2-3 days after transfection with increasing doses of puromycin (100-200 ng/mL) for 6-7 days. Three-four days after the start of puromycin, 0.5 µM 1-(2-deoxy-2-fluoro-beta-D-arabinofuranosyl)-5-iodouracil (FIAU) was added for 7 days. $Puro^R/FIAU^R$ colonies, without single colony selection, were expanded for further characterization (Figure 2b).

SEQUENCE OF THE HUMAN *GSC* PROMOTER

ctcgagCTGGGAAAGATTGGTGGGGGAGTATCTGCCCTGGAGCATGGTGACTTAAGTCTCTGAGAACGGAACAACAAT
AAAAAAGACTTCATCCTAGTCAGTGGGAGATGGAAGGTGCCTCACTTCTTACCCTAGACCTGACCTCTCTAAGGGGAG
GGTCACACTCCCTCTTTGTTTTTATTGTTAAAAAAGAAAAAAAGGAAAACTCTATTCACTTCCCCAAATTGAGACCCATC
TCCGTCATGTCCTAATCCCCATTTACAACCACCTCTAGTTTGCGCAGCACTGGAAGGAGCAACTCTCAGGCCACAAGCT
GGTGGCCGTGAGGACACAGAGGTCCACGCTGGGAGGCCTAGGAGGAGAAACGTCGCCGCCTGGCTCCACTGGGCCG
CTTCTGCGCGCCGCCACAGTAGTGCCCTCGGACAGGACTAAACCTCCGTGCAATGGCTCTTTGAAACCCGAACAGAAA
TTGACAAGTGGGCTTTCCCCTTTGCCCCTGAGCCGGGGAGTGGGCAGCCCACTTACTTAGCCTTACGGCGGCGACTTC

AGACAACTTTGCGGGGTCCCTTGGTTAATTAGCCTGAGAGAAATCGCGAAAAATAGGTGTGAGTTAAAGGCGGACTT
CCTTCAGGAAACCCGAGGTCTCCCGCAACCCCATCTCCCTCTCCGGCTCTCGGGGGTATCGCCTGGGTCGGGTACTGC
CCCTTCCCCTGGAGTGGCCTGAGTTGCAAGCCGGGAGCTTGCACCCGGTCCACAGGATCTGGGGGACTGCTTTTGTTC
CGCGCGTACCGTCGCCTACCTTCTCGAAGGGCGCTCGGCCGCCTGCTTGCCGTGAGTGGTGGGCTGGCTGCGCCTCCC
GTGGCTCGGGCAGGCTGGCCGTGCCATGCCGAGAATCCCGCAAATCATCGCGGCCGCCTGCCGCTATCAGACCCGGC
CATCGTCCAGGCCACTGCGGCCCAGGGAATCCACTCTGGGCCTGCCAGGCGGAAGGGCAGGCCTACCGGCAGCCAG
CAGGTCGTAAGGAAAAATGCGACCCAATTATTGGGGAAGAAGATAGGGAAGTCCCAAAGACCCCATCTGGATTGAA
AATCAAAAAAATAGTTTGGGAGAGAAGGCACCACATTTCCAAAACTAATTTTTTAAAAGCACAAAAACTACCGACACC
CACGCTGGGAGCAGCTAGAAGCACCGCTTCAGACTGCAACTCTTGGTGGCGAATTTTAAAGAGCGGCTGAATTTGGT
CTTCGCTGTGTAGGACCTAGGGACTGCCATGGTAGTTTTCAGACCCAGAGTGGCTGGAAAAAAAAATCTTTTTAAGAA
GCACGATGCACCCCTCCTTTAGTCGTTTTTGCCTTTACAAAATTTGGATTTTACAAAACCAAAAGTTTTTTGGGGGGGTT
AACAAGTTCCATTCGGGAGAGAAACTCAAGATTTAAAAAAGAGGTAAATCAGGCGGTGAGTACGTGGGGGCGCGGA
GGAGTGGTGGAATTTTCAATCCTGCTTGAGAGGGGATGCAGCCCCTCACATTTTCTAAAACAAAACCAATTCCGAAAG
GGGGAAAAAAGTAGCCCCAAAGTAAAATCGTCCACAATGAATTTGAGCTACAGGCAGAGGAAATCGCACCCAATAAT
AGGACCCAATCTGAGAAAAAAGGGTGGGGGTGGAGAGCTCGGGCGGAGGGAGTTGTTAACTGCGGCGGCACAAAG
GAATTAATGAGATTAACCAAGGCAATTAGGCCGCCCGCCCAGCTCGGCCGGCCCGCGGCCCGGCTGCCGAATGGAAA
AGATTAGGTTAATTTCATTAATTCGCAATCCACAATCTCTTTTCAGGCCTTGTAACCCCCTCCCCTTGGCACCTCTCCCCC
TCCCCCCAACCCTCTTCCCCCACCCAAAGGAAAAGAAAGGGCCAAATCTGGTTCTGTTTGTCATCTGCATATTACCAGG
AACTAAATCCAGGATGACGTCGACTCAGTATAAAACCAACAAGAGGTTCAGCTGGTCTGAGCTCCGTCCTACCCGCGG
GTTGAGTTCAGCGAACGCTGCGGCTAGGGGAGGGCGGGAGGAGGGAGAGCGGACGCAGGGGGCGGGGAGGGGC
GCAGGGCTGCGCGCTCGCCGGCGCTCTCTTTCGGTTTGGTCGGCGGCTGGAGGAGAGTGGACCCCCCCACTTTAAGG
CTCTGTCCTCGGCGCGTTCCCGCCGCCCCCCGGTCCCGACGCGGGGCTCGGGGaccggt

PCR GENOTYPING AND SOUTHERN BLOT.

PCR genotyping was done using 40 ng of genomic DNA with Go Taq DNA polymerase (Promega) in 10 µl reactions. Primer sequences and conditions of PCRs carried out on Figures 1, 2, S1 and S2 are listed in Supplementary Table 1.

Southern blot was done using 5' internal, 3' external and puromycin resistance probes labelled by PCR with Digoxigenin-dNTPs (Roche) and Go Taq DNA polymerase (Promega), using primers and conditions listed in Supplementary Table 2. Hybridization and development was done using the DIG High Prime DNA Labelling and Detection starter Kit II (Roche) according to manufacturer's instructions.

Supplemental Table 1. Primers sets used for PCR genotyping.

Assay	Forward	Reverse	Amplicon length	PCR Cycle
5'JA MCL	CACTTTGAGCTCTACTGGCTTC	CGTTACTATGGGAACATACGTCA	1.1 Kb	95ºC, 5' – [95ºC, 30'' – 68ºC (-0.5ºC/cycle), 1' 30'']X15 – [95ºC, 30'' – 58ºC, 30'' – 72ºC, 1' 30'']X25 – 72ºC, 5'
3'JA MCL	TAACTGAAACACGGAAGGAG	AAGGCAGCCTGGTAGACA	1.4 Kb	95ºC, 5' – [95ºC, 30'' – 68ºC (-0.5ºC/cycle), 1' 30'']X15 – [95ºC, 30'' – 58ºC, 30'' – 72ºC, 1' 30'']X25 – 72ºC, 5'
5'JA RMCEL	CACTTTGAGCTCTACTGGCTTC	CATGTTAGAAGACTTCCTCTGC	1.1 Kb	95ºC, 5' – [95ºC, 30'' – 68ºC (-0.5ºC/cycle), 1' 30'']X15 – [95ºC, 30'' – 58ºC, 30'' – 72ºC, 1' 30'']X25 – 72ºC, 5'
3'JA RMCEL	TTCACTGCATTCTAGTTGTGG	AAGGCAGCCTGGTAGACA	1.5 Kb	95ºC, 5' – [95ºC, 30'' – 68ºC (-0.5ºC/cycle), 1' 30'']X15 – [95ºC, 30'' – 58ºC, 30'' – 72ºC, 1' 30'']X25 – 72ºC, 5'
WT AAVS1	TTCGGGTCACCTCTCACTCC	GGCTCCATCGTAAGCAAACC	0.5 Kb	94ºC, 5' – [94ºC, 30'' - 60ºC, 30'' - 72ºC, 30''] X28-30 – 72ºC, 5'
5'RI DONOR	GTACTTTGGGGTTGTCCAG	TTGTAAAACGACGGCCAG	0.5 Kb	95ºC, 5' – [95ºC, 30'' - 60ºC, 30'' - 72ºC, 30''] X25 – 72ºC, 5'
3'RI DONOR	CCTGAGTTCTAACTTTGGCTC	ACACAGGAAACAGCTATGAC	0.5 Kb	95ºC, 5' – [95ºC, 30'' - 60ºC, 30'' - 72ºC, 30''] X25 – 72ºC, 5'
RI FLPe	CCTAGCTACTTTCATCAATTGTG	GTATGCTTCCTTCAGCACTAC	0.65 Kb	95ºC, 5' – [95ºC, 30'' - 60ºC, 30'' - 72ºC, 30''] X25 – 72ºC, 5'
OCT4-GFP	TTCCACAGACACCATTGCC	GCGGATCTTGAAGTTCACC	0.76 Kb	95ºC, 5' – [95ºC, 30'' - 60ºC, 30'' - 72ºC, 30''] X25 – 72ºC, 5'

Supplemental Table 2. Primers sets used for southern blot probe generation by PCR-DIG labeling.

Probe	Forward	Reverse	Amplicon length	Cycle PCR labeling	Detection
5' Internal	CTTTCTCTGACCTGCATTCTC	CTGCCCAAATGAAAGGAGT	411 pb	95ºC, 2' – [95ºC, 30'' –58ºC, 30'' – 72ºC, 45'']X30 – 72ºC, 7'	On-target and RI
3' External	GGCCAGGACTCCTGGCTCTGAAGG	TCCGACTCGGCCAGGTCCA	695 bp	95ºC, 2' – [95ºC, 30'' – 60ºC, 30'' – 72ºC, 45'']X30 – 72ºC, 7'	On-target integration
Puromycin	ATGACCGAGTACAAGCCCA	TCAGGCACCGGGCTTGCGGGTCA	600 bp	95ºC, 2' – [95ºC, 30'' – 72ºC, 1' 30'']X30 – 72ºC, 7'	RI

IMMUNOCYTOCHEMISTRY.

Cells were fixed either with 10% neutral-buffered formalin (NBF) or 4% Paraformaldehyde (PFA) and permeabilized with PBST (PBS 0.1% Triton) or TNT buffer in Tyramide amplification protocols. Blocking was carried out with 5% donkey serum or TNB in amplification protocols. Antibodies as well as isotype controls were used in the appropriate dilutions (Supplementary Table 3) and incubated overnight at 4°C. Secondary antibodies were used at 1:500 dilution. The staining were analyzed using the AxioImagerZ.1 fluorescence microscope and the Axiovision software (Zeiss).

Supplemental Table 3. List of antibodies used for immunocytochemistry.

Antibody	Catalog number	Company	Fixation	Dilution	Blocking	Secondary Antibody
OCT4	Sc-8628	Santa Cruz	10% NBF	1:200	5% Donkey serum	Donkey anti Rb-Alexa Fluor 488 or 555 (Invitrogen)
SOX2	AB5603	Chemicon	10% NBF	1:500	5% Donkey serum	Donkey anti Rb-Alexa Fluor 555 (Invitrogen)
NANOG	AF1997	R&D	10% NBF	1:400	5% Donkey serum	Donkey anti Gt-Alexa Fluor 555 (Invitrogen)
TRA-1-60	MAB4360	Millipore	10% NBF	1:1000	5% Donkey serum	Donkey anti Ms-Alexa Fluor 555 (Invitrogen)
FOXA2	Sc-6554	Santa Cruz	4% PFA	1:200	TNB/5% Donkey serum	Donkey anti Gt-Alexa Fluor 555 (Invitrogen)
SOX17	AF1924	R&D	4% PFA	1:200	TNB/5% Donkey serum	Donkey anti Gt-Alexa Fluor 555 (Invitrogen)
AAT	A0012	Dako	4% PFA	1:2000	TNB/5% Donkey serum	Donkey anti Rb-Alexa Fluor 555 (Invitrogen)
GFP	Ab5450	Abcam	4% PFA	1:500	TNB/5% Donkey serum	Anti-chicken biotin (Jackson) and Tyr-FITC (Perkin Elmer)
HB9	81.5C10	DSHB	4% PFA	1:50	5% Donkey serum	Donkey anti Ms-Alexa Fluor 555 (Invitrogen)

TERATOMA FORMATION ASSAY AND KARYOTYPING.

Five to ten million cells pre-incubated with 10 µM ROCKi for 1 hour were collected using Accutase (Sigma), washed and resuspended in a 1:1 solution of PBS-matrigel (BD) for injection under the skin of $Rag2^{-/-}\gamma c^{-/-}$ mice, using a protocol approved by the Institutional Ethics Committee at KU Leuven. Cells were used between passages 46 and 50. After 6-8 weeks, animals were sacrificed and teratomas harvested, fixed in 10% NBF overnight, stored in 70% ethanol and routinely embedded in paraffin. Animals that received hESC containing the inducible tdT expression cassette (Figure 6a), were fed with or without 500 µg/mL DOXY 10 days before teratoma dissection. Sections were stained with hematoxylin-eosin using standard protocols and analyzed by expert histologists of the Group of Morphology (Hasselt University, Belgium) to identify the presence of derivatives of the three different germ layers. Two clones of each ESC and iPSC were analyzed.

Immunohistochemistry performed against GFP and 3flag was performed by blocking tissue sections in 10% normal goat serum and using antibodies in their appropriate dilutions – GFP (ab290, Abcam) at 1:500 and 3flag (F7425, Sigma Aldrich) at 1:200 - and incubated overnight at 4°C. Secondary goat anti-rabbit horseradish peroxidase-conjugated antibody was incubated for 30 min at room temperature (Dako, EnVision+ System, HRP). Cells were counterstained with Mayer's hematoxylin (supplier if relevant) and staining was analyzed using a Nikon Eclipse 80i microscope (Nikon Co., Japan).

For karyotype analysis, metaphases of cells from two clones of each FRT-containing hESC and iPSC (passage number between 41 and 55) were harvested 3 days after plating in feeder free conditions. Karyotype analysis was done using standard cytogenetic procedures and results were evaluated at the Center of Human Genetics of KU Leuven.

RNA ISOLATION AND QUANTITATIVE REAL TIME PCR (QRT-PCR).

Total RNA was extracted using the GenElute Mammalian Total RNA Kit (Sigma) and cDNA was synthesized with the SuperScript III Reverse Transcriptase (Life Technologies). Quantitative RT-PCR (qRT-PCR) was performed in either Viia7 or Step One plus instruments using specific primers (listed in Supplementary Table 4) and the Platinum SyBRGreen qPCR supermix-UDG (Invitrogen). GAPDH was used as housekeeping gene. Relative expression to GAPDH was calculated as $2^{-\Delta Ct}$ and relative gene expression as fold change was calculated as $2^{-\Delta\Delta Ct}$.

Supplemental Table 4. Primers sets used for gene expression analysis by qRT-PCR.

Gene	Forward	Reverse	Amplicon length	Primer efficiency (%)
CXCL3	TAGCCACACTCAAGAATGGGAA	TCTCTCCTGTCAGTTGGTGC	110 bp	102
NFKBIα	CAGAGAGTGAGGATGAGGAGAG	TCATCATAGGGCAGCTCGTC	79 bp	97
NAF1	GCCTACAGAACCAGAGTCTCC	GGCTCTGCAAGATGAAGGTG	92 bp	102
OCT4	GATGGCGTACTGTGGGCCC	TGGGACTCCTCCGGGTTTTG	195 bp	109
GSC	TCTCAACCAGCTGCACTGTC	CCAGACCTCCACTTTCTCCTC	169 bp	106
AFP	TGAGCACTGTTGCAGAGGAG	GTGGTCAGTTTGCAGCATTC	123 bp	103
AAT	AGGGCCTGAAGCTAGTGGAT	TCCTCGGTGTCCTTGACTTC	250 bp	106
ALB	ATGCTGAGGCAAAGGATGTC	AGCAGCAGCACGACAGAGTA	85 bp	96
HNF4α	ACTACGGTGCCTCGAGCTGT	GGCACTGGTTCCTCTTGTCT	126 bp	101
3FLAG-tdT	TGATTACAAGGATGACGATGACG	CTCGCCCTTGCTCACCAT	67 bp	98
m2rtTA	*TCAAAAGCTGGGAGTTGAGC*	*CTGTCCAGCATCTCGATTGG*	93 bp	97
HB9	TGCCTAAGATGCCCGACTT	AGCTGCTGGCTGGTGAAG	91 bp	not determined

STUDY OF DNA METHYLATION BY BISULFITE SEQUENCING.

Genomic DNA of the cells was extracted using the QIAamp DNA Mini kit (QIAGEN) and 500ng DNA were used to perform the bisulfite conversion with EpiTect Bisulfite Kit (QIAGEN), according to the manufacturer's instructions and the following cycling process: [95°C, 30 sec – 50ºC, 1 hour]X16 – 10ºC as previously described (Izzi et al., 2014). Regular PCR was used to amplify specific fragments using 10-20 ng of converted DNA and

primers sets listed in Table S5. The PCR products were cloned into a p-GEM easy vector (Promega) and 15-20 clones were sequenced. The methylation rate of the CpG pairs was quantified using QUMA software.

Supplemental Table 5. Primers sets used for bisulfite sequencing

Region	Forward	Reverse	Amplicon length	CpGs	PCR Cycle
DHR	AAATTTAGTTGGGTATGGTGGTTTA	CTTCCCAAATTCAAACAATTCTAAC	215 bp	10	94°C, 3' – [94°C, 30'' - 60°C, 45'' - 72°C, 30'']X45 - 72°C, 10'
IS	GTTATTTGGGGTATTTTTTTT	AACTAAACCCAAAACCAATTAAAAC	169 bp	9	94°C, 3' – [94°C, 30'' - 54°C, 45'' - 72°C, 30'']X45 - 72°C, 10'
ZFN SITE	GTGGAAAATTTTTTTTGTGAGAAT	TCCTAAACTTACCAAAAACTCAAAC	204 bp	5	94°C, 3' – [94°C, 30'' - 52°C, 45'' - 72°C, 30'']X45 - 72°C, 10'
OCT4p-GFP	TAGATGAATTTTAGGGTTAGTTTGT	AAAAAAAATTAAATAATCCCTTC	272 bp	20	94°C, 3' – [94°C, 30'' - 52°C, 45'' - 72°C, 30'']X45 - 72°C, 10'

DIFFERENTIATION TO MOTOR NEURONS AND ENDOTHELIAL CELLS/PERICYTES.

hESC were differentiated to motor neurons following the protocol described by (Hu and Zhang, 2009). Directed differentiation of hESC towards endothelial cells and pericytes was carried out following a described protocol with some modifications (Orlova et al., 2014). Briefly, basal differentiation medium was maintained common to the hepatic differentiation protocol. Mesoderm induction was obtained by addition of 50 ng/ml of BMP4, ActivinA and Wnt3A till day 4. Endothelial specification was induced by addition of BMP4 (50ng/ml), VEGFA (50 ng/ml) and bFGF (20 ng/ml) at day 4 till day 6. Further endothelial maturation was obtained by the addition of VEGFA (50ng/ml) and bFGF (20 ng/ml) supplemented with the TGFβ inhibitor SB431542 (Sigma) from day 6 till d12 or 20 of differentiation.

SUPPLEMENTAL REFERENCES

Duan, Y., Catana, A., Meng, Y., Yamamoto, N., He, S., Gupta, S., Gambhir, S.S., and Zern, M.A. (2007). Differentiation and enrichment of hepatocyte-like cells from human embryonic stem cells in vitro and in vivo. Stem Cells *25*, 3058-3068.

Gerrard, L., Zhao, D., Clark, A.J., and Cui, W. (2005). Stably transfected human embryonic stem cell clones express OCT4-specific green fluorescent protein and maintain self-renewal and pluripotency. Stem Cells *23*, 124-133.

Hu, B.Y., and Zhang, S.C. (2009). Differentiation of spinal motor neurons from pluripotent human stem cells. Nat Protoc *4*, 1295-1304.

Ishii, T., Fukumitsu, K., Yasuchika, K., Adachi, K., Kawase, E., Suemori, H., Nakatsuji, N., Ikai, I., and Uemoto, S. (2008). Effects of extracellular matrixes and growth factors on the hepatic differentiation of human embryonic stem cells. Am J Physiol Gastrointest Liver Physiol *295*, G313-321.

Izzi, B., Binder, A.M., and Michels, K.B. (2014). Pyrosequencing Evaluation of Widely Available Bisulfite Conversion Methods: Considerations for Application. Med Epigenet *2*, 28-36.

Lam, P.Y., Sia, K.C., Khong, J.H., De Geest, B., Lim, K.S., Ho, I.A., Wang, G.Y., Miao, L.V., Huynh, H., and Hui, K.M. (2007). An efficient and safe herpes simplex virus type 1 amplicon vector for transcriptionally targeted therapy of human hepatocellular carcinomas. Mol Ther *15*, 1129-1136.

Marchetto, M.C., Muotri, A.R., Mu, Y., Smith, A.M., Cezar, G.G., and Gage, F.H. (2008). Non-cell-autonomous effect of human SOD1 G37R astrocytes on motor neurons derived from human embryonic stem cells. Cell Stem Cell *3*, 649-657.

Orlova, V.V., van den Hil, F.E., Petrus-Reurer, S., Drabsch, Y., Ten Dijke, P., and Mummery, C.L. (2014). Generation, expansion and functional analysis of endothelial cells and pericytes derived from human pluripotent stem cells. Nat Protoc *9*, 1514-1531.

Raitano, S., Ordovas, L., De Muynck, L., Guo, W., Espuny-Camacho, I., Geraerts, M., Khurana, S., Vanuytsel, K., Toth, B.I., Voets, T.*, et al.* (2015). Restoration of progranulin expression rescues cortical neuron generation in an induced pluripotent stem cell model of frontotemporal dementia. Stem Cell Reports *4*, 16-24.

Suttiprapa, S., Rinaldi, G., and Brindley, P.J. (2012). Prototypic chromatin insulator cHS4 protects retroviral transgene from silencing in Schistosoma mansoni. Transgenic Res *21*, 555-566.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.