Supplementary Information for

B3GALT5 Knockout Alters Glycosphingolipid Profile and Facilitates Transition to human Naive Pluripotency

Ruey-Jen Lin¹, Ming-Wei Kuo¹, Bei-Chia Yang¹, Hsiu-Hui Tsai¹, Kowa Chen¹, Jing-Rong Huang¹, Yun-Shien Lee², Alice L. Yu¹, ³, ⁴, John Yu¹, ⁵*

*corresponding author:

John Yu, MD, PhD
Distinguished Chair Professor & Director, Institute of Stem Cell and Translational Cancer Research, Chang Gung Memorial Hospital at Linkou, Taoyuan 333, Taiwan,
Tel: 886-3-328-1200#7803; Email: johnyu@gate.sinica.edu.tw

This PDF file includes:
Supplementary Methods
Figures S1 to S9
Tables S1 to S3
SI References
Supplementary Methods

Establishment of B3GALT5 knockout clones in hESCs

B3GALT5-KO clones were developed by introducing a CRISPR-Cas9 vector containing B3GAL5 guide sequences into H9 or HUES6 cells by transfection, as described in detail below.

CRISPR design: CRISPRs targeting the B3GALT5 locus were designed using E-CRISP (http://www.e-crisp.org/E-CRISP/). The CRISPRs were generated following the protocol from a previous study (1) in a single plasmid containing both sgRNA and Cas9 (pX330-U6-DR-BB-DR-Cbh-NLS-hSpCas9-NLS-H1-shorttratrcr-PGK-puro was a gift from Feng Zhang, Addgene plasmid # 42229). All constructs were checked using Sanger sequencing (Genomics BioSci & Tech). The following single-stranded oligonucleotides were used: hB3GALT5-CRISPR-gRNA-Px330_1F: CACCGGTATATTTGCCTTCTGGTTC, hB3GALT5-CRISPR-gRNA-Px330_1R: AAACCCAGAACCAGAAGGCAAATAT, hB3GALT5-CRISPR-gRNA-Px330_2F: CACCGGTCACCAGCAGGAGGAA, hB3GALT5-CRISPR-gRNA-Px330_2R: AAACCCCTTCTCCTGCTGTGGTG, hB3GALT5-CRISPR-gRNA-Px330_3F: CACCGGACCTCCATCCCCACAAACAGT, hB3GALT5-CRISPR-gRNA-Px330_3R: AAACCCAACTGTTTGTGGGATGAGG.

T7E1 assay: 293T cells were transfected with 10 μg of Cas9-encoding plasmid using the TransIT-LT1 transfection reagent (Mirus Bio) according to the manufacturer’s protocol. The rate of Cas9-mediated gene disruption was measured by T7 endonuclease I (T7E1) digestion of hybridized PCR products. Cellular genomic DNA from 293T cells was extracted using a KAPA Mouse Genotyping Kit (KAPA BIOSYSTEMS) after transfection for 2 d. The first PCR product was amplified with the first primers set (5’-AACTTGCCGAGGATAGCAGA-3’ and 5’-ATGAGTTTCCCATCCGCAGC-3’). The nested PCR products were amplified with diluted first PCR products and the second primers set (5’-TACAATGCGTGGTTTCCACA-3’ and 5’-CATAGAATGGGCTCAGCGTCCATCGC-3’). Nested PCR product in 1 x NEB Buffer 2 was hybridized in a thermocycler under the following conditions: 95°C for 5 min, 95 - 85°C at -2°C/sec, 85 - 25°C at -1°C/sec, and held at 4°C. Ten units of T7E1 (NEB, M0302) were added and the samples were incubated at 37°C for 30 min. The samples were then immediately run on a 2% agarose gel. PCR product without T7E1 treatment was loaded as a control. Band intensities were quantified using ImageJ software. T7E1 can recognize and cleave DNA mismatches in
those heteroduplexes. Running the cleavage products on the agarose gel, can resolve full length and cleavage products. The intensity of the respective bands allows to calculate the gene-editing percentage (also present as Indel percentage) that has occurred. Indel percentage was calculated using the following equation: (cut product intensities/uncut + cut product intensities) x 100 %.

**Transfection and drug selection:** For transfection, hESCs were dissociated into single cells with Accutase and plated onto 6-well Matrigel-coated plates (5x10^5 cells per well) in chemically defined E8 media with 10 μM ROCK inhibitor, Y-27632 (Tocris Ellisville). The following day, the media was replaced with 2 ml fresh E8 without ROCK inhibitor. Subsequently, 2.5 μg of Crispr-px330-B3GALT5 gRNA plasmid and 6.25 μl of TransIT-LT1 were mixed with 250 μl of OptiMEM® I Reduced-Serum Medium for 20 min added to each well of 6-well plates. 0.5 μg/ml puromycin was used to select cells for 2 d, and then the medium was changed to E8 medium without puromycin until clones were picked.

**Genotyping and sequencing:** Genomic DNA from the single colony was extracted using a KAPA Mouse Genotyping Kit (KAPA BIOSYSTEMS). The first PCR product was amplified with the first primers set (5’-AACCTTGCCGAGGATAGCAGA-3’ and 5’-ATGAGTTTCCCATCAGGCAG-3’). The nested PCR products were amplified with diluted first PCR products and the second primers set (5’-TACAATGCGTGTTTCCAGA-3’ and 5’-CATAGAATGGGTCCATCGCT-3’). The PCR product was digested with T7E1 to confirm the genotype. Then, the selected PCR products were further purified using a NucleoSpin Gel and PCR Clean-up kit (MACHEREY-NAGEL) and TA cloning with the pGEM-T Easy vector (Promega). The plasmids from selected clones were extracted with a TOOLS Plasmid Mini kit (TOOLS Biotech) and analyzed by Sanger sequencing (Genomics BioSci & Tech).

**Off-target analysis:** Potential off-target sites in the human genome that might be recognized by RNA-guided endonuclease specific for the B3GALT5 gene were examined with Cas-OFFinder (http://www.rogenome.net/cas-offinder/) and Benchling (https://benchling.com/signin) softwares. There are 36 and 49 potential off-target sites (with up to four mismatches); and their genome location sites and mismatch information were shown in Fig. S9A-B. Among these, a total of five potential off-target sites which were located in functional genes (PRKD1, SYNE2, EVC, MYO16 and NBEAL1) were listed in Fig. S9D. The gDNA from genome-editing cells were further examined in agarose gel in T7E1 assay.
after PCR analysis with primers specific for each gene surrounding the target locus. No indels were detected at any of the putative off-target sites (Fig. S9E). In addition, the lack of off-target effect was also verified by whole genome sequencing.

**Differentiation**

To access the multi-lineage potential of H9 cells with B3GALT5 knockout, the methods described below were used for differentiation.

**For neural (ectodermal) lineages:** The protocol for differentiation of hESCs into neurons was modified from previous reports (2, 3). For inducing neural differentiation, hESCs were passaged and placed (2x10^6 cells/well) in ultra-low attachment 6-well plates (Costar) and grown as floating cell aggregates of embryoid bodies (EBs) for 4 d with a daily change of differentiation medium, consisting of DMEM/F12: Neural basal (1:1), 1% Glutamax, 1% N2 supplement, and 2% B-27 supplement, including 100 ng/ml Noggin (d 1-4) (PeproTech), 10 μM SB431542 (d 1-4) (Sigma-Aldrich) and 0.7 μM CHIR (d 1-4) (Cayman). The EBs then were replated on Matrigel-coated plates and switched to a medium composed of DMEM/F12: Neural basal (1:1), 1% GlutaMAX, 0.5% N2 supplement, and 1% B-27 supplement, including 100 ng/ml Noggin (d 4-11) and 0.7 nM CHIR (d 4-11) to derive neuroepithelial cells. Neuroepithelial cells, manifesting as columnar cells organized into neural tube-like rosettes, developed on approximately d 7-9. The neural rosettes were isolated and expanded in the same medium to drive neural progenitors. Finally, the neural progenitors were then passaged using Accutase and replated (2x10^5 cells/cm^2) on Matrigel-coated plates in Neural Basal medium, containing 1% GlutaMAX, and 2% B-27 supplement, with 10 ng/ml BDNF (PeproTech), 10 μM GDNF (PeproTech), and 200 μM ascorbic acid (Sigma-Aldrich) for 14 d to induce mature neurons. Cells were harvested or fixed on d 11 and d 25 for immunofluorescence staining and flow cytometry analysis.

**For cardiac (mesodermal) lineages:** The procedure for differentiation of cardiomyocytes from hESCs was modified from previous reports (4, 5). Briefly, cardiac differentiation was induced by treating the hESCs with RPMI1640, containing the following to induce pre-cardiac cells: 6 μM CHIR99021 (d 1-2) (Cayman), 213 μg/ml AA2P (d 1-4) (Sigma-Aldrich), 2 μM Wnt-C59 (d 3-4) (Selleck Chemicals), 0.25% BSA (d 2-4) (Sigma-Aldrich); pre-cardiac cells
were then treated with RPMI1640, AA2P (d 5-20) (Sigma-Aldrich), 0.25% BSA (d 5-20) (Sigma-Aldrich) to induce cardiomyocytes. Cells were harvested or fixed on d 5 and d 20 for immunofluorescence staining and flow cytometry analysis.

For hepatic (endodermal) lineages: The procedure for differentiation of hepatocyte-like cells from hESCs was previously described (6). For the first phase of hepatic differentiation (d 1-3), hESCs were treated with RPMI 1640, 1% GlutaMAX, and 2% B-27 supplement, containing the following to induce DE (definitive endoderm) cells: 3 μM CHIR99021 (d 1 only) (Cayman), 100 μg/ml and Activin A (d 1-3) (PeproTech). For the second phase of differentiation (d 4-8), the DE cells were switched to the same basal medium containing 5 ng/ml BMP4 (d 4-8) (PeproTech) and 10 ng/ml FGF2 (d 4-8) (PeproTech) to derive hepatic progenitor cells. For final differentiation (d 9-17), the hepatic progenitor cells were treated with an L-15 based-basal medium [L-15 (GIBCO), 1% GlutaMAX, 1% ITS (Invitrogen), 8.3% FBS (Hyclone), and 8.3% TPB (Sigma-Aldrich)] supplemented with the following to induce hepatocyte-like cells : 100 nM dexamethasone (d 9-17) (Sigma-Aldrich), 200 μM ascorbic acid (d 9-17) (Sigma-Aldrich), 20 ng/ml human oncostatin M (d 9-17) (PeproTech), 10 ng/ml HGF (d 9-17) (PeproTech). Cells were harvested or fixed on d 8 and d 17 for immunofluorescence staining and flow cytometry analysis.

Matrix-assisted laser desorption ionization-time of flight (MALD-TOF) profiling and MALDI collision-induced dissociation MS/MS analysis for permethylated GSLs

MALDI-MS profiling of permethylated GSLs from each cell line was carried out as previously described (7, 8). Briefly, MALDI-MS profiling was performed with an ABI 4700 Proteomics Analyzer (Applied Biosystems) using a 2,5-dihydroxybenzoic acid matrix (10 mg/ml in water). Low- and high-energy collision-induced dissociation MALDI-tandem mass spectrometry (MS/MS) sequencing was performed on a Quadrupole/Time-of-Flight Ultima MALDI (Waters Micromass) with α-cyano-4-hydrocinamic acid and a 4700 Proteomics Analyzer using the 2,5-dihydroxybenzoic acid matrix, respectively.

Frozen cell pellets (approximately 3X10^7) were homogenized and extracted with methanol and chloroform and bath-sonicated for 30 min. After centrifugation, the supernatant was collected, and the pellet was extracted three times with the same solvent. The supernatants
containing neutral and acidic GSLs were pooled. GSLs were further separated with a 1cc cartridge for the total extraction (Oasis HLB Method, Waters) or for anion-exchange chromatography (Oasis MAX Method, Waters) to separate the acid partition from the non-acid partition. The purified GSLs were subjected to permethylation according to a NaOH/DMSO method (9). MS analyses were carried out on a SCIEX 5800 MALDI-TOF/TOF mass spectrometer using 2,5-dihydroxybenzoic acid as matrix. MS spectra were acquired in positive-ion mode and accumulated in 4000 laser shots with a random sampling mode.

**Analysis of B3GALT5 expression from public data**

The RNA-seq data from Chan et al. (E-MTAB-2031) (10) and Sperber et al. (GSE60955) (11), and microarray data from Gafni et al. (GSE52824) (12) and Theunissen et al. (GSE59435) (13) were also analyzed with GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r). Single cell RNA-seq data of E-MTAB-6819(14) derived from H9 hESCs were analyzed by Single Cell Expression Atlas (https://www.ebi.ac.uk/gxa/sc/experiments/E-MTAB-6819/results/tsne).

**Karyotype analysis**

Standard G-banding chromosome and cytogenetic analysis were performed at the Genetic Consultation Center, Department of Pediatrics (CGMH, Taiwan). Confluent cells were treated with 0.5 µg/mL of KaryoMAX Colcemid (Gibco) for 20 min and digested into single cells with Accutase. The single cells were resuspended in hypotonic 0.075M potassium chloride for 30 min at 37 °C and fixed in methanol/glacial acetic acid 3:1, then stained in Giemsa stain. Metaphase spreads for each sample were analyzed. A total of 20 pictures per slides were automatically acquired, chromosomes counted and at least 10 karyotypes for each cell line were classified.
Fig. S1. Validation of Primed and Naïve State hESCs. (related to Fig. 1)

Human primed and naïve (cultured in 5iLAF) H9 hESCs were profiled with multiple antibodies. (A) Phase-contrast microscopic images of primed and naïve state H9 hESCs. Flat and large colonies were observed in the primed state, whereas dome shape and small colonies were observed in naïve state hESCs. (B) Immunofluorescence staining shows expression of naïve-specific markers (KLF4, TFCP2L1, STELLA, TBX3 and TFE3) only in naïve hESCs. (C) Immunofluorescence staining shows expression of pluripotency markers (NANOG, OCT3/4, SOX2, TRA-1-60 and TRA-1-81) in both primed and naïve hESCs. Cell marker-specific antibodies are stained green or red, and nuclei are stained blue with DAPI. Scale bar, 200 μm.
**Fig. S2.** *B3GALT5* expression is decreased in naïve hESCs (related to Fig. 1)

(A) The expression of *B3GALT5* in naïve and primed hESCs in single-cell RNA-seq was analyzed from E-MTAB-6819(14) as supplied by online website, Single Cell Expression Atlas (https://www.ebi.ac.uk/gxa/sc/experiments/E-MTAB-6819/results/tsne). The data obtained from single-cell RNA seq of H9 cells cultured, separately, in primed or naïve conditions were plotted using t-SNE (t-distributed stochastic neighbor embedding). In upper panel, the cells cultured in naïve condition were colored as orange and in primed as green. The t-SNE plot shows the unique clustered distributions of 1,344 individual cells from the naïve or primed culture conditions respectively. Lower panel further exhibited the expression pattern of a selected gene, *B3GALT5*, in the single cells within the t-SNE plot. These expression levels are shown in blue color scale according to counts per million (CPM). The darker blue color in the scale refers to higher gene expression and the grey color indicates lower expression with the expression level below cutoff, 0.1 CPM considered to be undetectable. (B) The expression of *B3GALT5* is decreased in naïve hESCs as detected by RNA seq (E-MTAB-2031(10) and GSE60955(11)) and microarrays (GSE59435(13) and GSE52824(12)). Data are presented in relative log2-fold change for *B3GALT5* expression in naïve as compared to primed hESCs. (C)
The expression of *B3GALT5* is decreased in the naïve state of HUES6 cells. Expression of GSL-related glycosyltransferase genes was analyzed by qRT-PCR. Relative quantities are presented as the ratios of naïve/primed hESCs. Results are shown as mean ± SEM of three independent experiments. *p < 0.05*, two-tailed t-test.
Fig. S3. Decreased GSL complexity in B3GALT5-KO hESCs using mass spectrometry analysis (related to Fig. 2)
(A-B) MALDI-MS profiles of the permethylated GSLs from WT hESCs (upper panels) of H9 (A) and HUES6 (B) is dominated by Glc/Gal-ceramide, Gb3-ceramide, Gb4/Lc4- ceramide, SSEA-3, Fuc-Lc4Cer and Globo H among the neutral GSLs, and by GM3-ceramide and SSEA-4 among the acidic GSLs. Lower panels: Knockout of B3GALT5 (KO) decreases the neutral GSLs SSEA-3, Fuc-Lc4Cer and Globo H, and the acidic GSL SSEA-4 in H9 and HUES6.
Fig. S4. mRNA Expression Levels of Glycosyltransferases in WT and B3GALT5-KO hESCs (related to Fig. 2)

Expression of glycosyltransferase genes was analyzed by qRT-PCR. Relative quantities are presented as the ratios of B3GALT5-KO/WT H9 hESCs. Results are shown as means ± SEM of three independent experiments. None of the changes were statistically significant. Multiples of change of glycosyltransferase genes are summarized and shown in the diagrams.
Fig. S5. Characterization of B3GALT5-KO hESCs (related to Fig. 2, 3)

(A) Immunofluorescence staining for expression of pluripotency markers (NANOG, OCT3/4 and SOX2) in KO (primed state) and KO+2iLAF cells of H9. Scale bars, 200 μm. (B-C) qRT-PCR analysis for expression of genes previously reported as naïve or primed markers in indicated cells. Values were normalized to ACTB and then compared to WT in the primed medium. Error bars indicate SEM of three biological replicates.
Fig. S6. Further Validation of primed and naïve state in WT and B3GALT5-KO using HUES6 hESCs
(A) Representative morphology of wildtype (HUES6-WT) and B3GALT5 knockout cells (HUES6-KO) cultured in the indicated medium. The HUES6-WT and HUES6-KO cells in 5iLAF, and HUES6-KO cells in 2iLAF medium have naïve-specific dome-like morphology. Scale bar, 200 μm. (B) The expression of SSEA-3, SSEA-4, Globo H, Fuc-Lc4Ce and TRA-1-60 was not observed in HUES6-KO hESCs by immunofluorescence staining. The indicated specific antibodies are stained green or red, and nuclei are stained blue with DAPI. Scale bar, 100 μm. (C) Immunofluorescence staining for the expression of SSEA-3 and SSEA-4 in KO cells of HUES6, as compared with the control cells, after rescue via overexpression of B3GALT5 (OE B3GALT5-GFP and vector only (OE Ctrl-GFP) in the cells. GFP* cells indicate Ctrl-GFP- or B3GALT5-GFP-overexpressing cells. Scale bars, 100 μm. (D) HUES6-KO can differentiate into ectoderm, mesoderm and endoderm in vitro. Left panel: representative image of neuronal morphology. Middle panels: immunofluorescence staining of lineage markers include α-Actinin, NKX2.5 and TNNT2 for cardiomyocytes (mesoderm). Specific antibodies are stained green or red, and nuclei are stained blue with DAPI. Scale bar, 100 μm. Right panels: flow cytometry analysis of expression of Albumin and HNF4a for hepatocytes (endoderm). Isotype control (blue histogram) and the binding of specific antibodies (red histogram) are indicated. The gate on each histogram indicates the percentage of positive staining cells in each sample. (E) Representative RNA-FISH images detecting HUWE1 nascent transcripts (red) in DAPI-stained nuclei (blue) for the indicated cells of HUES6 hESCs. Scale bar, 20 μm. Bar charts show the quantification of X chromosome activation status (XaXi or XaXa) based on HUWE1 from samples of at least 100 cells.
Fig. S7. KO cells cultured in primed and naïve (2iLAF) medium display a normal female karyotype (46, XX).

Representative karyotypes of KO cells of H9 and HUES6 in the primed and naïve conditions. KO cells have a normal female karyotype (46, XX) as evaluated after 12 passages of H9 and 10 passage of HUES6 in the primed media whereas after 7 passages of H9 and HUES6 in the naïve culture condition.
Fig. S8. B3GALT5 knockout leads to an increase in intracellular Ca2+ required for naïve pluripotency in HUES6 (related to Fig. 5)

(A) Comparison of Intracellular Ca2+ in H9-WT cells in 5iLAF and 2iLAF medium. (B) Intracellular Ca2+ was visualized in HUES6 primed hESCs and naïve cells generated from WT+5iLAF and KO+2iLAF cells at 8 days. Cells were loaded with the cell-permeable fluorescent calcium indicator Fluo-4 AM. Ca2+ imaging was conducted using a fluorescent
microscope. The fluorescence intensities of the cells with Fluo-4 AM displayed in fluorescence image (left panels) and spectrum (right panels) were shown. Color-coded bar for intensity calibration is shown on the right. Scale bars, 50 μm. (C) Representative pictures for morphology and naïve makers (KLF4 and TFE3) of the WT and KO cells in the naïve culture condition (with 5iLAF and 2iLAF, respectively, for 8 days) and with or without BAPTA-AM (0 to 10 μM). Specific antibodies were stained green or red, and nuclei were stained blue with DAPI. Scale bars, 100 μm. Each experiment replicates three times.
Potential off-target sites in the human genome: Details of off-target analysis were described in method details. (A-B) A total of 36 and 49 potential off-target sites were found,
respectively, with Cas-OFFinder and Benchling softwares. (C) The number of potential off-target sites in the human genome after search for 2 to 4 mismatches with gRNA. (D) The list of functional genes carrying the potential off-target sites and the primers surrounding these target locus used for T7E1 assay. (E) Mutation frequencies of these 5 functional genes carrying the potential off-target sites. T7E1 assays for genomic DNA from WT- and B3GALT5-KO H9 cells using each of gene-specific primers were performed. In this assay, no Insertions and deletions (indels) were detected at measurable frequency at any of these putative off-target sites we had examined. * indicates 100 bp DNA ladder marker.
| Antibodies                                           | Source               | Identifier                     |
|-----------------------------------------------------|----------------------|--------------------------------|
| Mouse anti-Nestin (clone 196908)                    | R&D systems          | Cat# MAB1259; RRID: AB_2251304 |
| Goat anti-Sox1 (clone C-20)                         | Santa Cruz           | Cat# sc-17318; RRID: AB_2195365 |
| Goat anti-Sox2 (clone Y-17)                         | Santa Cruz           | Cat# sc-17320; RRID: AB_2286684 |
| Mouse anti-tubulin beta III isoform (clone TU-20)   | Millipore Corp.      | Cat# MAB1637; RRID: AB_2210524 |
| Mouse anti-Gata4 (clone G-4)                        | Santa Cruz           | Cat# sc-25310; RRID: AB_627667  |
| Mouse anti-SMA (clone 1A4)                          | DAKO                 | Cat# M0851; RRID: AB_2223500   |
| Mouse anti-a Actinin (clone EA-53)                  | Sigma-Aldrich        | Cat# A-7811; RRID: AB_476766   |
| Rabbit anti-Nkx2.5 (clone H-114)                    | Santa Cruz           | Cat# sc-14033; RRID: AB_650281  |
| Mouse anti-TNNT2 (clone 13-11)                      | ThermoFisher Scientific | Cat# MA5-12960; RRID: AB_11000742 |
| Mouse anti-AFP (clone 1G7)                          | Calbiochem           | Cat# ST1673; RRID: AB_10697987  |
| Rabbit anti-FoxA2                                   | Millipore Corp.      | Cat# 07-633; RRID: AB_390153  |
| Rabbit anti-HNF-4a (clone H-171)                    | Santa Cruz           | Cat# sc-8987; RRID: AB_2116913  |
| Mouse anti-human serum albumin (clone 188835)       | R&D systems          | Cat# MAB1455; RRID: AB_2225797  |
| Mouse anti-Oct3/4 (clone C-10)                      | Santa Cruz           | Cat# sc-5279; RRID: AB_628051   |
| Rabbit anti-Sox2                                    | GeneTex              | Cat# GTX101507; RRID: AB_2038021 |
| Rat anti-SSEA3 (clone MC-631)                       | Millipore Corp.      | Cat# MAB4303; RRID: AB_177628  |
| Mouse anti-SSEA4 (clone MC-813-70)                  | Millipore Corp.      | Cat# MAB4304; RRID: AB_177629  |
| Mouse anti-SSEA5 (clone 8e11)                       | GeneTex              | Cat# GTX70019; RRID: AB_11165050 |
| Mouse anti-TRA-1-60                                 | Millipore Corp.      | Cat# MAB4360; RRID: AB_2119183  |
| Mouse anti-TRA-1-81                                 | Millipore Corp.      | Cat# MAB4381; RRID: AB_177638  |
| Rabbit anti-TFE3                                    | Sigma-Aldrich        | Cat# HPA023881; RRID: AB_1857931 |
| Mouse anti-Globo H (clone VK9)                      | Alice L. Yu Lab      | N/A                            |
| Mouse anti-human CD173 (clone BRIC231)              | Bio-Rad              | Cat# MCA1984; RRID: AB_1100920  |
| Mouse anti-SSEA1 (clone MC-480)                     | Santa Cruz           | Cat# sc-101462; RRID: AB_1568826 |

RRID: Research Resource Identifiers
Table S2: Primers used in this study

| name       | sequence (5'-->3')                      |
|------------|-----------------------------------------|
| OCT4_F     | CCTGAAGCAGAGAGGATCAC CT                  |
| OCT4_R     | AAAGCGGCAGATGCTGTTTGG                   |
| SOX2_F     | GCTACAGCATGCAGGACCA                     |
| SOX2_R     | TCTGCGAGCTGTCATGGAGTT                   |
| NANOG_F    | CTCCAAACATCTGAACCTGACGC                 |
| NANOG_R    | CGTCACACCATTGTCTATTCTTCG                |
| KLF4_F     | CATCTCAAGGCACACCTGTGAA                  |
| KLF4_R     | TCGTGCTGTACACCTGGACT                    |
| TFCP2L1_F  | TTCACCGCAACAGGTTCCTGCA                  |
| TFCP2L1_R  | GCCCTTGATGGCGTGAAGAC                   |
| TBX3_F     | GGAACCTGGAAATGGCCGAAGA                  |
| TBX3_R     | GCTGCTTTGTCCATGGAGGACT                  |
| ACTB_F     | AGAAATCTGGCACCACACC                    |
| ACTB_R     | AGAGGCTACAGGGATAGCA                    |
| GAPDH_F    | GTCTCCTCTGACTCCAACAGCG                 |
| GAPDH_R    | ACCACCTGTGCTGATGCACA                   |
| KLF2_F     | CAAGAAGAGTGCGCATCTGAAGG                |
| KLF2_R     | TACATGTGCGTTTCCATGTGCAG                |
| DPPA3_F    | AGACAAACAAACAGAGATCCT                  |
| DPPA3_R    | CCCATCCATTAGACAGCGA                    |
| DPPA5_F    | ACATCGAGCGGTGAAGCAAGG                  |
| DPPA5_R    | CATGGCTTGCGAAGTGG                    |
| DUSP6_F    | TTCCCTGAGGCGTATCTT                    |
| DUSP6_R    | AGTGACTGAAGCGCTAATG                   |
| ZIC2_F     | GATGTGCGACAAGTGCTCCTACAC              |
| ZIC2_R     | TGGAGCAGACTCATAGCCCGGA                 |
| DNMT3L_F   | GGGACAAGCTGAAGCATGTGTT                |
| DNMT3L_R   | AAGATCGAAGGGTCCCTA CT                 |

F: forward; R: reverse
| TaqMan Probe                             | Source                  | Cat#          |
|-----------------------------------------|-------------------------|---------------|
| B3GALT5 TaqMan Gene Expression Assay    | ThermoFisher Scientific | Hs00707757_s1|
| FUT1 TaqMan Gene Expression Assay      | ThermoFisher Scientific | Hs00355741_m1|
| FUT2 TaqMan Gene Expression Assay      | ThermoFisher Scientific | Hs00382834_m1|
| ST3GALT2 TaqMan Gene Expression Assay  | ThermoFisher Scientific | Hs00911835_m1|
| A4GALT TaqMan Gene Expression Assay    | ThermoFisher Scientific | Hs00213726_m1|
| B3GALNT1 TaqMan Gene Expression Assay  | ThermoFisher Scientific | Hs00364202_s1|
| B3GNT5 TaqMan Gene Expression Assay    | ThermoFisher Scientific | Hs01935580_s1|
| UGCG TaqMan Gene Expression Assay      | ThermoFisher Scientific | Hs00234293_m1|
| GAPDH TaqMan Gene Expression Assay     | ThermoFisher Scientific | Hs99999905_m1|
| GUSB TaqMan Gene Expression Assay      | ThermoFisher Scientific | Hs99999908_m1|
| UBC TaqMan Gene Expression Assay       | ThermoFisher Scientific | Hs05002522_g1|
SI Reference

1. Cong L, et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121):819-823.
2. Kirkeby A, et al. (2012) Generation of regionally specified neural progenitors and functional neurons from human embryonic stem cells under defined conditions. *Cell Rep.* 1(6):703-714.
3. Li W, et al. (2011) Rapid induction and long-term self-renewal of primitive neural precursors from human embryonic stem cells by small molecule inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 108(20):8299-8304.
4. Lian X, et al. (2013) Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/beta-catenin signaling under fully defined conditions. *Nat. Protoc.* 8(1):162-175.
5. Burridge PW, et al. (2014) Chemically defined generation of human cardiomyocytes. *Nat. Methods* 11(8):855-860.
6. Brolén G, et al. (2010) Hepatocyte-like cells derived from human embryonic stem cells specifically via definitive endoderm and a progenitor stage. *J. Biotechnol.* 145(3):284-294.
7. Liang YJ, et al. (2011) Changes in glycosphingolipid composition during differentiation of human embryonic stem cells to ectodermal or endodermal lineages. *Stem cells* 29(12):1995-2004.
8. Liang YJ, et al. (2010) Switching of the core structures of glycosphingolipids from globo- and lacto- to ganglio-series upon human embryonic stem cell differentiation. *Proc. Natl. Acad. Sci. U.S.A.* 107(52):22564-22569.
9. Smith DF & Prieto PA (2001) Special considerations for glycolipids and their purification. *Current protocols in molecular biology* Chapter 17:3.1-3.13.
10. Chan YS, et al. (2013) Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. *Cell stem cell* 13(6):663-675.
11. Sperber H, et al. (2015) The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition. *Nature cell biology* 17(12):1523-1535.
12. Gafni O, et al. (2013) Derivation of novel human ground state naive pluripotent stem cells. *Nature* 504(7479):282-286.
13. Theunissen TW, et al. (2014) Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency. *Cell stem cell* 15(4):524-526.
14. Messmer T, et al. (2019) Transcriptional Heterogeneity in Naive and Primed Human Pluripotent Stem Cells at Single-Cell Resolution. *Cell reports* 26(4):815-824 e814.