DISSERTATION

CHARACTERIZATION AND FUNCTIONAL ANALYSIS OF GLYCOSYLATION IN MOUSE PLURIPOTENT STEM CELLS

2021 MARCH

SOKA UNIVERSITY

GRADUATE SCHOOL OF ENGINEERING

PECORI FEDERICO

Table of contents

1. Introduction

Mammalian embryonic development is a complex stepwise process. Upon cleavage and compaction, the embryo organizes itself into the inner cell mass (ICM) and the trophoblast cells. At this stage, the first ICM commitment occurs. The segregation of cells within the ICM give rise to the epiblast (EPI) and the primitive endoderm. Cells generating the EPI can differentiate into the three germ layers and its associated extraembryonic tissues, thus are pluripotent stem cells (PSCs). Following segregation, the embryo implants into the uterine epithelium, initiating significant changes in (PSCs) identity and embryo morphology (Bedzhov et al., 2014). Embryonic stem cells (ESCs) are PSCs derived from pre-implantation embryos at E3.5-E4.5 (Evans and Kaufman, 1981). EPI-like cells (EpiLCs), which differentiate from ESCs in culture, are PSCs which resemble the embryo post-implantation stage at E5.5 (Hayashi et al., 2011). ESCs and EpiLCs reflect two distinct pluripotent states known as the naïve state and the primed state (formative) (Smith, 2017), respectively, providing a useful model system to examine the pluripotent state transition occurring at embryo implantation in *vitro* (Nichols and Smith, 2009) (Fig. 1).

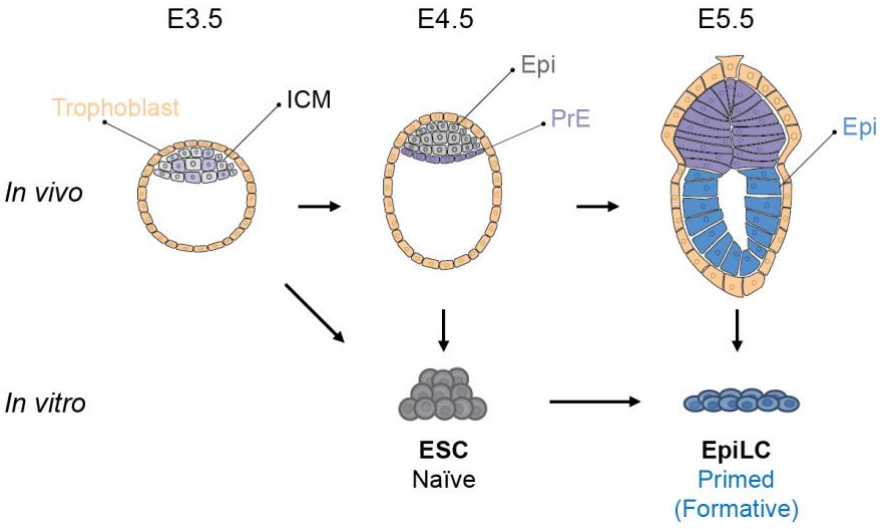


Fig. 1. Mouse embryonic development in *vitro* vs in *vitro.*

ESCs are naïve PSCs derived from pre-implantation embryos, whereas EpiLCs, which can be differentiated from ESCs in culture, are primed PSCs resembling the embryo post-implantation stage.

To date, several signaling pathways involved in pluripotency regulation have been identified. Leukemia inhibitory factor (LIF) and bone morphogenic factor (BMP) signaling are strongly involved in the maintenance of ESC pluripotency via the LIF receptor (Lifr) and BMP receptor (Bmpr) families, respectively (Niwa et al., 2009; Ying et al., 2003). In contrast, fibroblast growth factor (FGF) signaling triggers ESC differentiation into EpiLCs via the FGF receptor (Fgfr) family (Lanner and Rossant, 2010). Wingless-type (Wnt) signaling maintains the pluripotent state while at the same time priming cells for differentiation via interaction of Wnt ligands with Frizzled (Fzd) receptor family (ten Berge et al., 2011). Indeed, induction of Wnt signaling by CHIR99021 (CHIR) together with the FGF signaling inhibitor PD0325901 (PD) is commonly used to culture ESCs in a ground state of pluripotency. However, CHIR alone induces ESC differentiation (Ying et al., 2008) (Fig. 2). PSCs represent a promising tool to dissect the mechanisms underpinning mammalian development and advance regenerative medicine. Nevertheless, a significant number of attributes remain unexplored, hampering PSC exploitation.

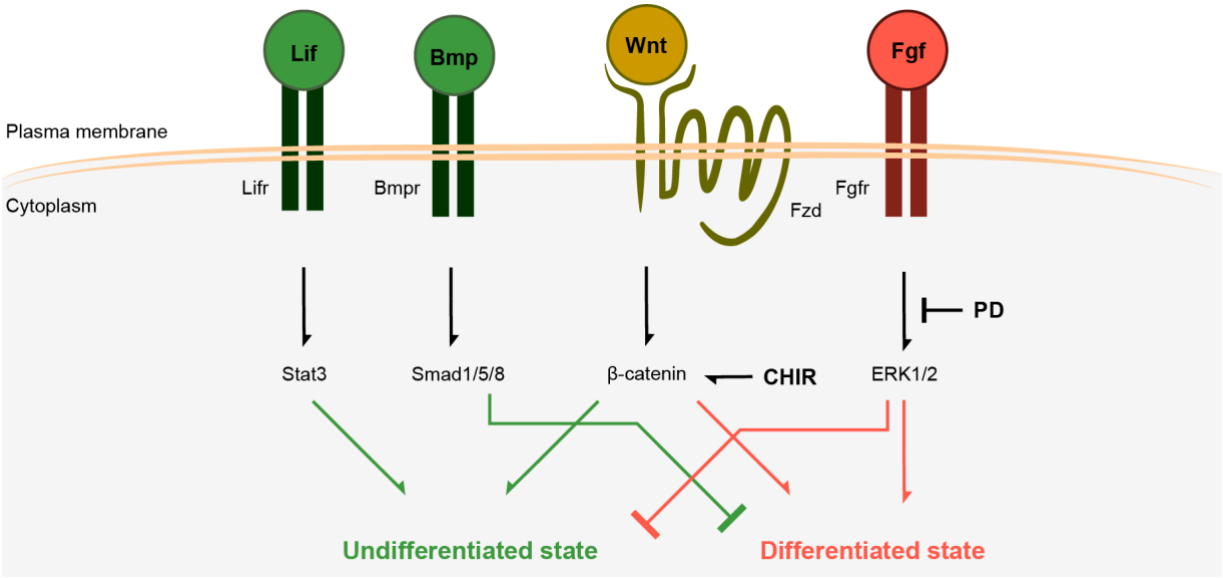


Fig. 2. Signaling pathways involved in PSC pluripotency network.
Schematic representation of main signaling pathways involved in PSC pluripotency network.

Glycosylation is one of the most abundant post-translational modification which is predicted to be present on more than half of all mammalian proteins (Apweiler et al., 1999) and is exerted by over 200 distinct glycosyltransferases and related enzymes in mammals. Characterization of glycomic profile has been performed in mouse and conventional human ESCs (hESCs), human induced pluripotent stem cells (hiPSCs), tumors and late differentiating cells (Nairn et al., 2012; Fujitani et al., 2013; Homan et al., 2019). Nonetheless, the glycosylation dynamics during pluripotency state transition remain unexplored. Glycosylation is involved in a wide range of cellular processes, such as adhesion, signaling regulation, and endocytosis (Varki, 2017). To date, numerous studies have reported the critical role of glycosylation in development and stem cell regulation across different species (Haltiwanger and Lowe, 2004; Nishihara, 2018). However, the molecular mechanisms and its relationship with signaling in ESCs remains poorly understood. Mucin-type *O*-glycosylation is an evolutionarily conserved protein modification and is one of the most abundant forms of glycosylation present on membrane proteins and secreted proteins (Tran and Ten Hagen, 2013). Mucin-type *O*-glycosylation is characterized by the initial addition of N-acetylgalactosamine (GalNAc) to a serine or threonine residues of the target protein by a family of 19 polypeptide α-*N*-acetylgalactosaminyltransferases (Galnts) in mouse; the attachment of GalNAc forms the so-called Tn antigen. Galactose (Gal), sialic acid (Neu5Ac), or N-acetylglucosamine (GlcNAc) can be successively added to Tn antigen by C1galt1, St6galnac1, and B3gnt6, to form three main structures: T antigen (Galβ1-3GalNAc), sialyl Tn antigen (NeuAcα2-6GalNAc), and Core 3 structure (GlcNAcβ1-3GalNAc) (Bennet et al., 2012). The importance of this mucin-type O-glycosylation during embryonic development is highlighted by the fact that knockout of the *C1galt1* gene is embryonic lethal in mice (Xia et al, 2004). Moreover, a novel culture comprising the primitive growth factor $NME7_{AB}$, which binds to the extracellular domain of the cleaved form of the mucin-type *O*-glycosylated protein MUC1, was observed to maintain hESCs in an undifferentiated state (Carter et al., 2016), indicating that mucin-type *O*-glycosylation plays a key role in ESC pluripotency network. In the present study I characterized the glycomic dynamics during mouse pluripotent state transition and clarified the role mucin-type *O*-glycosylation in ESC pluripotency network.

2. Materials and Methods

2-1. Cell culture

R1 (Nagy et al, 1993) and E14TG2a (Smith and Hooper, 1987) ESC lines were maintained on mouse embryonic fibroblasts (MEF) that were prepared from embryos at embryonic day 14.5 and inactivated with 10 μg mitomycin C (Sigma-Aldrich) at 37°C, 5% CO_2. ESCs were maintained in ESC medium consisting of DMEM (Gibco) supplemented with 15% fetal bovine serum (FBS) (Nichirei Biosciences), 1% penicillin/streptomycin (Gibco), 0.1 mM 2-mercaptoethanol (Gibco), 0.1 mM non-essential amino acids (Gibco) and 1000 U/mL LIF (Chemicon International). R1 ESC line was used for the experiments unless stated otherwise. R1 ESCs were tested for mycoplasma contamination and authenticated by ATCC cell line authentication service – mouse STR profiling.

To analyze the effect of PRC2 inhibition on ESCs' glycome, ESCs were treated with 10 μM EED226 in ESC medium containing LIF for 48h.

For alkaline phosphatase assay, 3-5 x 10^4 cells were plated on gelatin-coated 60-mm dishes in ESC medium containing LIF. The cells were fixed and stained for ALP with 5-bromo-4-chloro-3-indolyl phosphate-nitroblue tetrazolium (Nacalai Tesque) 4 days after replating. The colonies were counted by microscopy and scored manually.

To analyze the effect of *O*-glycosylation inhibition, ESCs were treated with 2 mM GalNAc-Bn in ESC medium containing LIF for 48h.

For galectin dissociation, ESCs were cultured in ESC medium (- FBS) containing 50 mM lactose monohydrate (Kanto Chemical) for 30 minutes. For galectin-3 addition, ESCs were cultured in ESC medium (- FBS) for 2 hours before addition of 15 μg/mL Lgals3 for 30 minutes in ESC medium (- FBS).

2-2. ESCs differentiation

EpiLC differentiation was performed following a previously established protocol (Hayashi et al., 2011) with slight modifications. ESCs maintained on MEF feeder cells were passaged on gelatin-coated 60-mm dishes (TrueLine), and cultured for 30 min in ESC medium containing LIF to completely remove MEF feeder cells 24h prior induction. Next, ESCs were seeded at low density in gelatin-coated 60-mm dishes containing ESC medium in the presence of 1 μM PD0325901 (Wako), 3 μM CHIR99021 (Wako) and LIF. The following day, EpiLC induction was performed: ESC were cultured at 1 x 10^6 in gelatin-coated 60-mm dishes containing ESC medium without LIF for 24h. Subsequently, EpiLC medium, consisting of DMEM/F12 (Gibco) supplemented with 20% knockout serum

replacement (KSR) (Gibco), 2 mM L-glutamine (Invitrogen), 1% penicillin/streptomycin (Gibco), 0.1 mM 2-mercaptoethanol (Gibco), 0.1 mM nonessential amino acids (Gibco), 30 ng/ml FGF2 (Wako) and 0.6 μM JAK inhibitor (JAKi) (Santa Cruz Biotechnology), was added. EpiLCs were collected for further analysis 72h post-induction.

To induce embryoid body (EB) formation, ESCs were seeded in 35-mm Low Cell Binding dishes (Nunc) and cultured in ESC medium without LIF for 2.5, 4, 8, or 12 days. EB assay from transient *C1galt1* KD ESCs was performed 2 days post transfection and collected at day 0 (ESCs), 2, 3, and 4.

2-3. Vectors construct and transfection

For knockdown (KD) of the *C1galt1* in ESCs, I generated small hairpin RNA (shRNA) expression vectors targeting two different regions (and also enhanced green fluorescent protein (*Egfp*) as control) by inserting the double-stranded DNA between the BamH1 and HindIII sites of pSilencer 3.1-H1 (Ambion) or p.SUPER.retro.puro (OligoEngine), for transient and stable KD, respectively. shRNA sequences were designed as described previously (Sasaki et al., 2009) by using siDirect:

Egfp (5’ – 3’):

GATCCCGCCACAACGTCTATATCATGGGGAAAATCCATGATATAGACGTTGTGGCTTTTTTGGAAA

C1galt1 KD 1 (5’ – 3’):

GATCCCGGGCCCAGCGTTGTAATAAAGGCTTCCTGTCACCTTTATTACAACGCTGGGCCCTTTTTTA

C1galt1 KD 2 (5’ – 3’):

GATCCCGGCTTCTATCAAAGTATAACCGCTTCCTGTCACGGTTATACTTTGATAGAAGCCTTTTTTA

For transient KD ESCs were plated prior to transfection in gelatin-coated feeder-free culture dishes in ESC medium with LIF. The next day, the cells were transfected with 2 or 4 μg of expression vector using Lipofectamine 2000 (Invitrogen). From one day after transfection, the cells were selected with 2 μg/mL puromycin (Sigma-Aldrich). *C1galt1* KD 1, and *C1galt1* KD 2 transfected cells were cultured for 4 days post transfection. For stable *C1galt1* KD, p.SUPER.retro.puro vector containing *C1galt1* shRNA were transfected into ecotropic virus-packaging (PLAT-E) cells. Supernatants containing retrovirus were collected 48 hours following PLAT-E transfection, mixed with 8 μg/mL polybrene (Sigma-Aldrich), and incubated with ESCs for 24 hours. The following day, ESCs were replated with ESC medium containing LIF and 2 μg/mL puromycin, and selected for 7 days before use.

2-4. Real-time PCR and RNA-seq

Total RNA was extracted from cells using TRIzol reagent (Invitrogen) and reverse-transcribed using a Superscript II First Strand Synthesis Kit (Invitrogen) and oligo-dT primers. Real-time PCR was performed with an ABI PRISM 7700 Sequence Detection System (Applied Biosystems) and SYBR Green Master Mix (Roche). The relative amount of each mRNA was normalized against the amount of *β-actin* or *Gapdh* mRNA. The primer sets used are listed in Table 1.

Copy number was calculated as previously described (Whelan et al, 2003). In brief, a region of 80-1000 bp of the *St6galnac1*, *C1galt1, B3gnt6*, and *Gapdh* genes was inserted into a pGEM®-T easy vector (Promega). The vectors were linearized and a concentration ranging 100-0.01 pg/μL diluted in 1 μg/mL yeast tRNA was obtained by serial dilution to create a standard curve for each gene.

The RNA-seq libraries were prepared from total RNA using the TruSeq Stranded mRNA Prep kit (Illumina). 100 bp single-read sequencing was performed by NovaSeq6000 (Illumina).

2-5. Fluorescence-activated cell sorting (FACS)

A single cell suspension for cell surface molecules staining was obtained using 0.02% EDTA/PBS. Subsequently, 2–3 x 10^5 cells were harvested and washed in FACS buffer (0.5% BSA (Iwai), 0.1% sodium azide (Sigma-Aldrich) in PBS). After washing, the cell suspension was incubated with Abs or lectins in FACS buffer. For internal molecules analysis, cells were collected with 0.25% trypsin/EDTA (Thermo Fisher Scientific) and fixed/permeabilized for 30 min with 100% methanol (Wako) before staining. Samples were then analyzed using a BD FACSAria III Cell Sorter (Becton Dickinson). Cells were gated to exclude debris, dead cells (identified by propidium iodide staining; Sigma-Aldrich), and doublets. The primary and secondary Abs and lectins used are listed in Table 2.

2-6. Glycome analysis

Cultured ESCs and EpiLCs (> 1 × 10^6 cells) were washed 5 times with phosphate buffered saline (PBS) and collected using a scraper. Collected cells were resuspended in 100 mL of PBS and homogenized using an Ultrasonic Homogenizer (TAITEC). Ethanol (400 μL) was added to the lysate and then placed at -30°C. After ethanol precipitation, the proteinous pellet and supernatant fractions were separated by centrifugation, then the resulting pellet was dissolved in 100 μL of H2O and their cellular protein concentration was measured

using a BCA protein assay kit (Thermo Fisher). The pellet fractions corresponding to 25 μg, 50 μg, and 100 μg proteins were used to *N*-glycans, *O*-glycans, and GAG analyses, respectively. The supernatant fraction corresponding to 100 μg proteins was concentrated for GSL and FOS analyses. Glycomic analyses of *N*-glycans, GSL, and FOS were performed by glycoblotting methods combined with SALSA procedure (Furukawa et al., 2008; Hanamatsu et al., 2018). *O*-glycome analysis was performed by BEP method and GAGs were measured by HPLC as previously described (Fujitani et al., 2013).

2-7. Immunostaining

Cells were fixed with 4% paraformaldehyde in PBS (PFA/PBS) and washed in PBS. The fixed cells were blocked with 10% Block Ace (Dainihon Pharmaceutical) in PBS (Block Ace/PBS), 10% Block Ace/PBS or 1%BSA/PBS with 0.3% Triton X-100 (Sigma-Aldrich) for the analysis of cell surface or intracellular molecules, respectively. For primary labeling, the cells were incubated with anti-Oct3/4 Ab (Santa Cruz Biotechnology; sc-5279; 1:100), anti-Nanog Ab (ReproCELL; RCAB001P; 1:100), anti-Otx2 Ab (Millipore; AB9566; 1:100), PNA-biotin (Cosmo Bio; J214; 1:100), anti-Fzd5 Ab (Sigma-Aldrich; SAB4503132; 1:100), anti-β-catenin Ab (Cell Signaling; #9562; 1:100), anti-Rab5 Ab (Cell Signaling; #46449; 1:100), or anti-Lgals3 Ab (Santa Cruz Biotechnology; sc-32790; 1:100). Subsequently, the cells were stained with Alexa Fluor 555-conjugated streptavidin (Life Technologies; S32355; 1:300), goat anti-rabbit IgG Alexa Fluor 488-conjugated (Life Technologies; A-11008; 1:300), goat anti-mouse IgG Cy5-conjugated (Life Technologies; A-10524; 1:300), anti-mouse IgG Alexa Fluor 488-conjugated (Life Technologies; A-11001; 1:300) and Hoechst 33342 (Invitrogen; H3570; 5 μg/mL). Images were obtained using an LSM 700 confocal laser microscope (Carl Zeiss). Puncta staining and nuclear/cytoplasmic ratio was quantified using Fiji (ImageJ) software. 3D reconstructions were created using Imaris version 9.3.1 (https://imaris.oxinst.com).

2-8. Western and lectin blotting

Cells were lysed with lysis buffer (50 mM Tris-HCl [pH 7.4], 150 mM NaCl, 1% Triton X-100, 5 mM EDTA, 1 mM Na_3VO_4, 10 mM NaF and protease inhibitors). Protein samples were separated on an SDS polyacrylamide gel and transferred onto polyvinylidene fluoride membranes (Millipore). For western blotting, membranes were blocked using 1% BSA and incubated with the following primary antibodies: anti-Sox2 Ab (R&D Systems; MAB2018; 1:1000), anti-Oct3/4 Ab (Santa Cruz Biotechnology; sc-5279; 1:1000), anti-β-actin Ab

(Sigma-Aldrich; A5441; 1:10000), anti-p-β-catenin Ab (Cell Signaling; #9561; 1:1000), anti-β-catenin Ab (Cell Signaling; #9562; 1:1000), anti-Fzd5 Ab (Sigma-Aldrich; SAB4503134; 1:1000), anti-ERK1/2 Ab (Cell Signaling; #9102L; 1:1000), anti-p-ERK1/2 Ab (Cell Signaling; #9101L; 1:1000), anti-Akt Ab (BD Biosciences; 610836; 1:1000), anti-p-Akt Ab (Cell Signaling; #9275; 1:1000), anti-LRP5 Ab (Cell Signaling; #5731; 1:500), anti-LRP6 Ab (Cell Signaling; #3395; 1:1000), anti-Nanog Ab (ReproCELL; RCAB001P; 1:1000), anti-Otx2 Ab (Millipore; AB9566; 1:1000), anti-H3 Ab (Cell Signaling; #9715; 1:1000), anti-H3K27me3 Ab (Cell Signaling; #9733; 1:1000), or anti-Gapdh Ab (Santa Cruz Biotechnology; sc-32233; 1:1000). The membranes were then incubated with anti-rabbit IgG (Cell Signaling; #7074; 1:10000), anti-rabbit IgG light chain specific (Jackson Immunoresearch; 211-032-171; 1:10000), or anti-mouse IgG (Cell Signaling; #7076; 1:10000) HRP-conjugated secondary Ab. For lectin analysis, membranes were probed with PNA conjugated to horseradish peroxidase (HRP) (Cosmo Bio; J414; 1:10000). Membranes were then washed and developed with ECL Plus reagents (GE Healthcare).

2-9. Luciferase assay

A shRNA expression vector targeting *C1galt1* (2 μg) was co-transfected using Lipofectamine 2000 with the reporter plasmid TOPFLASH (Upstate Biotechnology; 2 μg) or FOPFLASH (Upstate Biotechnology; 2 μg); pCH110 (GE Healthcare; 0.2 μg) was used as a control for transfection efficiency. Luciferase activity was measured 4 days after transfection by Dual-Light System (Applied Biosystems). Luminescence was measured with a Lumat LB9501 luminometer (Berthold).

2-10. Immunoprecipitation and pull-down assay

Cells were lysed with lysis buffer (50 mM Tris-HCl [pH 7.4], 150 mM NaCl, 1% Triton X-100, 5 mM EDTA, 1 mM Na_3VO_4, 10 mM NaF, and protease inhibitors). A 100–500 μg of proteins was diluted 10-fold with wash buffer (lysis buffer w/o Triton X-100). For immunoprecipitation, anti-Fzd5 Ab (Sigma-Aldrich; SAB4503134; 1 μg), anti-LRP5 Ab (Cell Signaling; #5731; 1 μg), anti-LRP6 Ab (Cell Signaling; #3395; 1 μg), or normal rabbit IgG (R&D Systems; AB-105-C; 1 μg) was added. Protein G Magnetic Beads (New England Biolabs) were then added. For the pull-down assay, the diluted cell lysate was incubated with PNA-biotin (Cosmo Bio; J214; 20 μg); streptavidin magnetic beads (Bio-Rad) were then added. Surface biotinylation assay was performed using the Pierce cell surface protein isolation kit (Thermo Scientific) following the manufacturer's instruction.

2-11. Immunoelectron microscopy

ESCs (5 x10^5) were plated on gelatin-coated 35-mm plastic dishes in ESC medium containing LIF. At 4 days after transfection, cells were fixed with 4% PFA/0.01% glutaraldehyde/PBS for 1 hour at room temperature and washed in PBS. The fixed cells were permeabilized with 0.5% saponin in PBS for 5 minutes and unreacted aldehyde functional groups were quenched using 10 mM glycine in PBS for 5 minutes. After PBS washing, the cells were blocked with 5% donkey serum (Jackson Immunoresearch) in PBS and incubated with anti-Fzd5 Ab against the N-terminal region (Sigma-Aldrich; SAB4503132; 1:100) or anti-Fzd5 Ab against the C-terminal region (Sigma-Aldrich; SAB4503134; 1:100) in 5% donkey serum, PBS, and 0.1% Tween 20 at 4°C overnight. Subsequently, the cells were stained with ultrasmall gold-DARabbit IgG (Nanoprobes; 1:100) for 1 hour at room temperature and washed with PBS. The cells were then post-fixed with 1% glutaraldehyde and 0.5% tannic acid for 30 minutes. Following silver-enhancement of the gold particles, the cells were postfixed with 0.5% OsO_4 in 100 mM phosphate buffer (pH7.3) for 30 minutes and dehydrated in a graded series of ethanol. After passage through propylene oxide, the cells were embedded in Epon 812. Ultrathin sections were cut, stained with uranyl acetate and lead citrate, and observed with a transmission electron microscope (JEM-1010C; JEOL).

2-12. Statistical analysis

An unpaired two-tailed Student's *t*-test was used to calculate *P* values between two group of samples. Statistical significance is denoted by asterisks: * $P < 0.05$; ** $P < 0.01$; and *** $P < 0.001$. Error bars represent s.e.m.

Table 1.

	Gene	**Forward Primer** 5' - 3'	**Reverse Primer** 5' - 3'
	Gapdh	TGCACCACCAACTGCTTAGC	GGCATGGACTGTGGTCATGAG
	β-actin	GCTCTGGCTCCTAGCACCAT	GCCACCGATCCACACAGAGT
	Nanog	CAGAAAAACCAGTGGTTGAAGACTAG	GCAATGGATGCTGGGATACTC
	Esrrb	TTTCTGGAACCCATGGAGAG	AGCCAGCACCTCCTTCTACA
	Klf2	GCCTGTGGGTTCGCTATAAA	AAGGAATGGTCAGCCACATC
	Klf4	GGAAGGGAGAAGACACTGCG	ATGTGAGAGAGTTCCTCACGCC
	Rex1	GCTCCTGCACACAGAAGAAA	GTCTTAGCTGCTTCCTTCTTGA
	Tbx3	AGGAGCGTGTCTGTCAGGTT	GCCATTACCTCCCCAATTTT
	Sox2	ACCAGAAGAACAGCCCGGA	CCCGGGACCATACCATGA
	Oct3/4	CTCACCCTGGGCGTTCTCT	AGGCCTCGAAGCGACAGA
	Fgf5	GCAGCCCACGGGTCAA	CGGTTGCTCGGACTGCTT
	Otx2	CATGATGTCTTATCTAAAGCAACCG	GTCGAGCTGTGCCCTAGTA
	Cdx2	GAGCTGGCTGCCACACTTG	GCTTCTTCTTGATTTTCCTCTCCTT
	Gata3	CCATTACCACCTATCCGCCC	TCGACTTACATCCGAACCCG
	Gata4	TCCATGTCCCAGACATTCAGACT	AGCAGACAGCACTGGATGGAT
	Gata6	CCCCTCATCAAGCCACAGAA	GTGACAGTTGGCACAGGACAGT
	Brachyury	TGCTGCAGTCCCATGATAACTG	ATGACTCACAGGCAGCATGCT
	Mixl1	GCACGTCGTTCAGCTCGGAGCAGC	AGTCATGCTGGGATCCGGAACGTGG
	FoxA2	AGCCGTGAAGATGGAAGGG	CTCCGCGTAGTAGCTGCTCC
	Sox17	GCACAACGCAGAGCTAAGCA	CTGCCAAGGTCAACGCCT
	Pax6	AACCTGGCTAGCGAAAAGCA	CCCGTTCAACATCCTTAGTTTATCA
	Nestin	TGCAGACACCTGGAAGAAGTTC	CCCAAGGAAATGCAGCTTCA
	Axin2	GGGAGCAGTTTTGTGGCAGCA	AGGGTCCTGGGTAAATGGGTGAG
	β-catenin	GTTAAACTCCTGCACCCACCA	AAAGGGCAAGGTTTCGAATCA
***N*-glycan**	*Dpm1*	TGACCTGAAGTCCATGGCAG	CTTAACTGCTTCCCAGCCCA
	Dpm2	AGAGGCTGAACATTGCTGGT	CTTGTGACGGGACAGGACTT
	Dpm3	CGGAGCCAACATGACGAAGT	CGTCCTCGCAGTCGTGAAAT
	Glt28d2	CGGCTCAGTCCTCTGTTGTG	TATGACTTCTGCGTGCGGC
	Dpagt1	CCTGGTGGAACTGGAAGGTG	GGTCTTGCTGAAGTGTCCCA
	Alg1	CCTGTGTGTGACCAATGCTATG	GAGAAGTCCTCGTCCTCTGTCC
	Alg2	GAAGATATGGACCGCGCACT	ACGTAGAGCGCCAGAAAGAC
	Alg3	TACGTACCCATCCACGTCCT	TCTTAGGGAAGGACTCGGGG
	Alg5	TGGTTGACGATGGCAGTGAA	CAGCGTTATCACCCGGACTT
	Alg6	TGCAGTGTTTGTTTCTTTCTTCAGT	TCTGAGGAGGGTCCAGTGTT
	Alg8	CCTGCTGTGGAACTTTGGGT	AAAGAGCAACGCCCCTTCAA
	Alg9	GTGCTCCACTCCACCTTGTT	AAGGCCTGCTAGAAGTTGCC
	Alg10b	CCTGTATCTGGTGTCGGTCG	GCCTTGTTTCTGGGTTGCAC
	Alg11	CCAGGGCATTTGCTGGTTTC	CCGATGGTCGCTTCAGACAA
	Alg12	CAGCCACGTCCCTCTATGTG	GGACTCGATGTGTGTCCCTG

	Alg13	AGAGCGTTTGTGACTGTCGG	TCTCGAGGATCTGAACGCAG
	Alg14	CTATCTGTGTGTCCGCCCTG	ACAATTCGCCCCAGGTACAC
	Stt3a	AAGAATGAGGTGGCGAGTGG	CACTGGTCACCCAAGTCGAA
	Stt3b	TGGTGGGACTACGGCTATCA	TTTTCCGACCAGTGCGATGT
	Uggt1	TCCCATGACTAAGGAGCCCA	TACACCTGCCCAATCCGTTC
	Uggt2	TACAGAACATCGCGGGCATT	ACATCATAGCGAGAGGCACG
	Gnptab	CCTGATTGCGAAGGGCTAGT	ATGGACACCATCACTCGCTG
	Gnptg	AGAAACCACTCACTCTCAGCAC	TAGTCCTGGGTCACTCCGC
	Nagpa	GATCCCGTGAATCCGTTCGT	CTCGTCACACTCGATGGCTT
	Mgat1	AGGACGGGCTTGTATTCGTC	CAGGTCCAACTGGGTGAAGG
	Mgat2	TGACCCTAGAGATTGCCCCA	GCCTCTCTGTAATGGCCGAA
	Mgat3	GTTCTGTATGGCTGGCCTGT	GCTGGATACGAGGTTAGGGC
	Mgat4a	CGCGTCCTCTTCTGTTTGGA	ATGGGAAGAGTCCCTGTCCG
	Mgat4b	CCTAACAGAGGATCCGCGAC	AAGACGGTGGGCAGATGAAG
	Mgat4c	TCTCACGATCTCCCACGTCT	ATAGCCAATCGCAGGGCATT
	Mgat4d	TTAGGAGCTGTTTGCAGCGA	TTGAGGTGCTGTTCTCGTCC
	Mgat5	CCGGGTGCTAGACTCCTTTG	AACTGCTGCGGGTTCAGATT
	Fut8	ATTCGTCCACAACCTTGGCT	TTCCCACTTTGTCTGTGCGT
O-glycan	*Galnt1*	GTGCTTCCCTCTTCTGTGGG	CTTTGCCGCTGCTCTTGATG
	Galnt2	TTACTGTGGTGGACCGTTCG	CCACTTCTGTCTGCTGTCGT
	Galnt3	ACGCAGGTGATTGCTCGTAA	AGGTCTGGCACATACGCTTC
	Galnt4	ACCCTACGGGGGCTAATCTT	CCGTCTTTGGGGCAGTTTTG
	Galnt5	GGTTGCTGAGGTTTGGCTTG	GGTAAGGTTGCCCACATCCA
	Galnt6	GCTGTACCAGGGTTCCTTCC	TGTGGTCTCCCATCTCGCTA
	Galnt7	CTGGGGAACCCGTCTGTTAC	GAATCTTCAAGCCCCACGGA
	Galnt9	TCCCCATGACAAACCCAGGA	GGCATTTTAACTTCTCGCGCA
	Galnt10	CGAGGCTGCTTGGAACAGTA	GGAGGGATTGCAGGTGTTCA
	Galnt11	TTCTGCCTTACTGCGGACTG	GGGGCTGTAACCACATCACA
	Galnt12	AAGCAAGCTCCCTATTCCCG	GGGGATTGCGGTGGTAGTAG
	Galnt13	TCCGATGGTATCCTGTCCCT	CCTCCACATTGCCAGATCCT
	Galnt14	GGGGTGGGGAGAACTTTGAAAT	TCCTCAGGTTCAGCCTGTTCTC
	Galnt15	CGAGTGGGACACATCTACCG	AGATGGGTACAGCTCAGGGT
	Galnt16	CCGGACTGACCCCACTAAAC	TGCCACCACACATCCATACC
	Galnt17	AATTGGTCTGCTGGACCCTG	ACGGTTTCTTCTTCCGCTCA
	Galnt18	TTGGTGTCCACTTGCGTGAT	GCTGGCGATGTAGTTGGTGA
	Galntl5	CATCGAGAGAGAGGTGCCAG	GAGGTTCACCACACTCGACA
	Galntl6	GGAGCCTTCGACTGGGAAAT	GAGACCTCCAGCCATAACCG
	C1galt1	GCAAGGCATTCAGATGACAA	ATGTTGGCTGGAATCTGCAT
	C1galt1c1	GCAAGCCCCTAACAAAGACG	CCGGAAACTCTTACTGAGCTCC
	Gcnt1	TTCCAACGGTGCCCCTTATC	AGCATCCAGCTCAAGTCACC
	Gcnt3	CACCCTCCAGCGTGCC	TCAAATTTTCGATGCAAGGGAT
	Gcnt4	ACGGAGCATCATCGACTTGG	TGTGGACGACCAGCGAATAG

	A4gnt	AAAGACTTCCACGGGTTGGG	CCACAAATGCAGTGCGTAGG
	B3gnt6	AGTCCCACGACACTGGCTTTC	CCTGCCTGTGTTCCTGGAGG
	Pomt1	TATGAGAATGGCCGTGGCAG	GGGCAGGCATGGAGATGTTA
	Pomt2	TGGGAAAGATGCTGATCGGT	TCGTATCTGTCCCCAGGCTT
	Pomgnt1	TGGACACTCGGAGAGCAATC	GTATGTGTCAAACACGCGCT
	Pomgnt2	GTGCCTTACATGCTGACCCT	AACTGTGGACCTGGGAAAGC
	B3galnt2	TCGGTGTTCTTAGGTCCCCT	ACCAGCCAGTTTCGCATTGG
	Pomk	TCGCCAGGCATTTTGTGAAAG	TACTGGCCTTCTACAGCCGA
	Fktn	GTGGCAACTACCTCTGGCAT	CGCCCGAGCTTCCTTATACC
	Fkrp	CTCTAGTGCCTGATGGAGCG	CCGAGAGGTTGAAGAGGTCG
	Rxylt1	GGAGGCTTTGTGGATCTGCTT	TGTCCAGCTGGCTTCAACTAC
	B4gat1	GGTGCCTGCATTTGAAATCC	TGCACCAGCTCGTTCTTGTT
	Large1	CCAGGACCCACTGCACTTAG	CCCTCATAATACTTTCTTGGCCTTG
	Large2	ACCTGCCTACTCCCTCTACG	GCACCACCAAAGCAGTCTTG
	Pofut1	AGGCAGTCGTCCTTCTTTGG	GGAGGTGGGCTGCTTTGTAT
	Rfng	ACTCGCACCTGGAAAACCTG	GGACCCCCATAGCTCAAGGT
	Mfng	ATTCCGCTCCCTCCATTGTC	ACACCAACTCAAGCGTTCCT
	Lfng	TGTCCATTGCCACCTGTACC	AGTTAGGGACAGACCTGCCA
	Pofut2	CGAAGAAACACGGGGTCTGA	TCTTATCAGCGGCATCCGTC
	B3glct	GGAGGCGGGATGGTCTTTAG	TGGGGTAATCTACTGGCCGA
	Eogt	GCTGGCCTTACCCACTTACTT	CAGCGCTCATCTTCACAGTTG
	Ogt	GAAAGAGGCACGCATTTTTGA	ACCACCGCATGATTTGGG
	Poglut1	TTGTGCGGTTCACTGGTCTT	GGTCGGTCTTGACTGGGATG
	Poglut2	TGGTGAGCACGTTGGCTTTA	AGTCAGGTCGTAGGTAGGCA
	Poglut3	TCCTCAAGCAGGAGTCACCA	GCTGGGGAACAAGCTCCATT
	Gxylt1	CACGTACTGGGGCTCATGTT	TGGGCGCAGATCCTTTTGAT
	Gxylt2	ACGACGACAAACAACCCACT	CTTCTCGTAAGCCCTTCGCA
	Xxylt1	CCGTCATCCAAGGACAGAGG	GACTCACAAAGCCCGTCTCA
GAG	*Xylt1*	CTGACCGATTCCAGGGCTTT	GGGACTTGCGACCTTGAAGA
	Xylt2	GGACAACACCTTCACTCCGA	GGGCACAGACTTGGGCATTA
	B4galt7	CATTCTGCTGCTGTCCAAACA	GCCCCAGCCCCAAAAG
	B3galt6	CCCAGAGGACATGTTGGAGAA	AACTGCACCTCATGCTTGCA
	B3gat3	CCTACAGCCGGGAGCTCTTT	CCCACAGGCCACACTGAGA
	Ext1	CAGGCTTGGGTCCTTCAGATT	CCATCCGTTGCTGAGCATT
	Ext2	CCCACAGAGGCAGATTGAAGA	TGGCTTTAATGGACTGGAAGTATG
	Extl1	AGCCTTCAAGACTTCCCCTTTT	GATCACGGCGCTGAAACTG
	Extl2	TGAGAAACCGACTCCAGGTCTT	TGAGTGTGTCATCGTCTACCATCA
	Extl3	ACCTTTTGTGGCCAGACATTG	AACCGATTGTTCAAACTGTTCT
	Hs2st1	TGGCGTATCTGTACCCCTGA	ACCCACATCACAGCTCACTG
	Hs3st1	CTACAAGGCTCTCAACCGCA	CGGCCTTTGGACTCGTGTAA
	Hs3st2	CCACAGAGCGTCCCCAGAAA	AGCATCTCCCGGTATTAAGCG
	Hs3st3a1	CCTTACTTCTGGACGAGGGC	TCTGGGCATCAAATCCCGGT

	Hs3st3b1	AGCTACCACAAGGGTCTTGC	CATGGTGATCTGTCCCTTGA
	Hs3st4	GTGACGAATGAGGCTCCCAA	GAGTCGCTCACCACTGACAA
	Hs3st5	CCAGAGTTGGGAGCTTGG	CACCACCAAATCGACTTTCA
	Hs3st6	CTGTTTGTAAGCGGTGAGCG	AGGGGAAGCCCTTAGTAGCA
	Hs6st1	CTGACTGGACCGAACTCACC	GAACTTTCTGGGCGAACGCA
	Hs6st2	ACCTGACTCTAGTGGGATGCT	AACAGGTACTGGGTCTTGCG
	Hs6st3	TCAACATCAAAGGGCGGGAT	CATGGCTGCTCCAGACGAAT
	Ndst1	GTATGAGGCCTGGAAGCAAGT	GCCGGCTTCAGATGTGGGTA
	Ndst2	CGCTGGCTGACTTACTATCCCT	CCGGGTTGACACGAAGCT
	Ndst3	TCCCAAGACAAACCTGTGGG	AAGAAGAGCTCCCCTCCATGA
	Ndst4	ATTCCGATCAACCTGGGCTAC	GGTGACCTGAATACCCCAGAC
	Glce	AGGAACAACAGAAGGCACCC	AGCCCTAACACCTTGCTTCC
	Csgalnact1	ACAGCCAGGGAAGAAGGTATTTT	TCGTGATGGCCGTAGATTACG
	Csgalnact2	GACCTGTCAATATCAATCGGATTTC	TTCTCCACCCCAGCCTTTTA
	Chpf	GCCGCTGCATCCTTGATG	CTCAGTTCCAGGTAGTTGTAGTGCAT
	Chpf2	AGGGTGAGGGAGAAGATCCC	ATACCGAGTCCTGAGCACCT
	Chsy1	CCGGGAAGGTGGAGTTCTTC	TCTTCTGGGGTGGGTAGGAG
	Chsy3	GAAGGTGTGTTCGCCGTTTC	GCTGTTGTGGAGGTCTTGGA
	Dse	CCCACAGCTTCTCCTTCTTG	CTCCTCAAAAGGGACATCCA
	Dsel	GAGCTTCTGGATGTATGGAG	CTGAACCAGGGAGTGAGG
	Chst3	CTGACCCAGCCCTGCTCTT	AAAAGGTGGAATCCAACTCGC
	Chst7	TGGGCGAACTCTTCAACCA	GATACAGTGCCTGCCACAGATG
	Chst11	CCAGAGATCAACCACCGCTT	TTCTGCGTGAACTTGTTGCG
	Chst12	TCGGTGAGTGAGGCTTTCAG	GCCGTCCTGTTCCGATAACT
	Chst13	CCAGGAGGTCTGGAACGAAC	AGCAACCTGTACCCTACCCA
	Chst14	GCCTGCTCTAACTGGAAACG	CTGCCAGAAACACCAAGTCA
	Chst15	TATGACAACAGCACAGACGG	TGCAGATTTATTGGAACTTGCGAA
	Ust	TTGGGGTCATTGGGAGATGC	GTGCCTTGTGGTAGAGCCAT
	B3gnt7	GGCCACGAGGGCTTCAA	CCTTGTTCATCCGGCTGCT
	Chst1	TGTGCGCATGTGTGGCTT	TGACTGCGCTCACGACAAG
	Has1	CATGGTGCTGCTGTCACTCT	GCAACACGGGGACGTAGTTA
	Has2	GGCCGGTCGTCTCAAATTCA	ACAATGCATCTTGTTCAGCTCC
	Has3	GAAAACTGCTGGCTGGTGTG	CAACTGCCTGTCACAAACGG
GSL	*Ugt8a*	TATCGGCAGAGGGCTCAGAA	TGATGGACAGCAGAACGGAG
	Ugcg	GAGCCAGCCCATTTGCCTAA	GGAAATGCTGTTTGGCCCAG
	B4galt5	CCATCGAGGAGAGGTCCAGTT	CTTGCCGCTCCTTTGACTTC
	B4galt6	GGGTCTCCAATCGCTCTCTG	AGAGGTACGTGTTGGCGATG
	A4galt	GAGGAGACCATGTCCAAGCC	TGAAGAAGGTGAAGACCCGC
	B3galnt1	CCTTGGAGGATGAGCACGTT	ACTTGGCATTGGGGCAAAAC
	Gbgt1	TGAACCGCCACTTCATCTGG	CCGAGGCTTCCTTTCATCCC
	B4galnt1	TGTACTCCAGCACCCGAAAC	GATCCTGACCGGGATGTGTG
	B3galt4	CTCAGAGTCAGCAGCACACA	GATGTAGCGGGCCATAGGAC

	B3gnt5	GCTCTGTGCAGCCTCACTAC	AGTTCTCATTGCCCCACGTT
Elongation/branching	*B4galt1*	GCTGGAGACACCATGTTCAATC	ACAAAGCAGTTGTAATCATAGTCCT
	B4galt2	GAGACCTTCAACCGAGCCAA	ACAACGGTACAGATTGCGGT
	B4galt3	CCATAATTGTGCCCCATCGT	GGTGCAGGTGATAGAGCAGCA
	B4galt4	TCTTTACCCTGTGCCTGACG	GTCCTTTGCCTTGGGGATCA
	B3gnt2	GGGTCTTCCTGTTGGGCAA	AAACTTAAGCATGTCCGAAAGGTC
	B3gnt3	TACGGCGACATTCTCCAGTG	CACATCGTCGTCCCCATTGA
	B3gnt4	TCTGAAGGAGCGCATACCAG	CTATGGTGGGACCTTGGCAG
	B3gnt8	GACTGTCAGACTTGGGGGAC	AGTGGAAAGCTCCTACAGGC
	B3gnt9	CGACAGAGTGGATGAAGGAAGGATA	CTCCTCCTCCTCCTCATGTCCA
	B4galnt3	GGGCCATATCCCACAGACAG	CACTCGAAAGAGGGTGTCCC
	B4galnt4	GGACCCACATGGTTACTGGG	GTGGTGGTAGAAGTGTCGCA
	B3galt1	CAATCAGGGATGTCCGCAGT	AGAACGGTGGGTAGTTGCTG
	B3galt2	AGCTCCTCAGACTCTGAGGCC	TGTGACTCCCTGTGGTGATAGC
	B3galt5	CACTCACCGGCTGCTCTTTA	ATGTGAGCCATCTTTGCCGA
	Gcnt2	CTCCATGCCACCAAACGCA	CTTCGCTTCCATGTCCATCCAC
	Gcnt7	GCCCACCCCACAATCTAACA	ACGGGGATCAGTTAGGACGA
	Mgat5b	GGCCGTCCTCAACAAGTACA	CCGGAAGAACTCGTGGTTGA
	Ggta1	ACCTATGAGAGGCGGGAACT	ATGGCCGCGTGGTAGTAAAA
	A3galt2	CGTGCAGCAAGTACCACCTA	CACTGCCAAGGGCTAAGGAA
Capping/terminal modifications	*St3gal1*	AACCCAACAACCTGAGCGAC	GGATCCACATTGCCAGGC
	St3gal2	ACTGGCACCACTACTGGGAGA	ATGCACGCCTGTCTTCCG
	St3gal3	GACTCTAAACTGCCTGCCGA	TGCGAAAGGAGTCATCCAGG
	St3gal4	AGCCTCGAGTGTCGTCGCT	AGCGAGCTGTTCCGCAAC
	St3gal5	GCTTTGGCTACGACCTCAGT	GTTTGCCGTGTTCCGAGTTC
	St3gal6	TCCCAAGGATCAGAAACCCA	CACTGCATATGTGAAAGGCCA
	St6gal1	TCAAGTTCAGCGTAGAGGCG	AAGACACGACGGCACACTTA
	St6gal2	CCACCCTGAGAACCCTTTCC	GCGCTTTGTGATTTCCTCCC
	St6galNAc1	GAGAGGCAGTCCAAGGAGAGC	TGAGGATTCTCTGGTGCTGGC
	St6galNAc2	GAGCCGACTTTGCCAACACTC	GGGTGGAGAGGTCAAACAGGC
	St6galNAc3	CCAGTGTCCCTCTCTTGCTG	TCCTCATGTTGCGGAAAGGG
	St6galNAc4	CAGGCAAGAACCGGAGACAA	TAGACCACGATCTCCTCGCA
	St6galNAc5	AAGGGCAGTCATCACCGTTT	ACAGCGGGTGATTCTGGTTT
	St6galNAc6	CCCGAACATGGAGGCCTAT	AAGAGGTCGTCAAACTGCTGC
	St8sia1	GTGACCCTGCCCATCTCTTTG	CCATCGTACCACAGGCTCTTC
	St8sia2	AATTCTGGAGGCAGAGGTACAAT	GCGTCTGGTTGTGTCTCCA
	St8sia3	GGGCGCCCCGAATG	CTGGTCTAGAAACTTCAGTGCAAACT
	St8sia4	TGGCTCCACCATCTTCCAAC	TTTAGTGTCCGGCGTCTGTC
	St8sia5	AATACATGCTGGACGACTTCCA	TGCTCAGGTACTGCGGGTG
	St8sia6	TTGTCCACACTCGTGAAGGG	AAAGGTGTCGGTGGGACATT
	Fut1	CCCGACACAAAGACCCCATC	AAGTGCAAAGTCCTTCCCCG
	Fut2	AGACCTGTCCCCACTCCTTA	ACCAAGTGACTCCTTCTGCC

	Sec1	CCTTCCACACGGTTTTCAAGC	CTACTGTGTTTGGCTGTCCCA
	Fut4	ACGTGTCTGTGGACGTGTTT	AGCCAGGGAAGCAGCATTAG
	Fut7	CAAACCCGGCAATCTCTCCT	AGTTGAAGATGCCCCGGAAG
	Fut9	GCCCTGGATCTTTCCTCACC	TCTAAAGCGATCCCAGCGAG
	Fut10	CCCTTTGTCAGTCATCCGGG	GGCTCCAGGTGTTTAGGCTC
	Fut11	CGCTACATCCCGGTGGATTC	CCGTGGCTGTGTCCTGTAAT
	Abo	ACTCGATGGCAGGATGTGTC	CCACCTCACGTAGAAAGCGT
	B3gat1	GTTGGTGGCCTTCGGTATGA	CCCTTCACACCACGTAGCTT
	B3gat2	GCTGTATTTAAGCGCCGTGG	ACGAGAACCTTGGTGCAGTT
	B4galnt2	CTTGACGGCTTTCCCGG	TGTGAGCCAAGAAGAAGTTCACC
	Chst2	TGCCCACCTCAACGCCT	CCTTGATAACCAGCGTGCG
	Chst4	CGTCCACAGACACCTTTCCC	AAGACAGCACCAGCACATGC
	Chst5	CGGAGCTTGAGCAAAGGG	CCACTTGAAACTGTCCATGCCT
	Chst8	CTGACCAACCTCGACCTCAC	GTTTCGGACGAGTTGTTGCC
	Chst9	TCACAAAGGGAGAACGGAGC	ATGACCGACGCTTCTCTTGG
	Chst10	GTACCGGAAGAACCGGACAG	ATCTGCCTCCAGGGTCTCAT
	Gal3st1	AACCACATGCGCTTCCACTA	GGCACCACGGATCCAAAGTA
	Gal3st2	AGCCCAAGAACCTGACACAC	AGAATATTGGCCTTCCCACCC
	Gal3st3	CACCATGCCACCTATCCTCC	AACTTGGGGTACCAACTGAGC
	Gal3st4	CTGGGTCCCTTCACTGGGAT	CCAAACTGGTAGTGGGCAGG
NST	*Slc35a1*	GCGAAGTTGAGTGTGCCATC	AAGCAGTGCAGGGGATCTTC
	Slc35a2	GGCTGTCCTGGTCCAATATGT	ACTGGAGGTTATTCTGCAAGGT
	Slc35a3	TGGATTCTCACAGTGGTGCC	GACGGACACACTTCATCCTGA
	Slc35b1	CATCATCACCACAACCCGGA	TTGGCATCAAGACCGAGACC
	Slc35b2	GGACAGAGACAGCGGAATCC	ACCCGAGGCACAGAAGACC
	Slc35b3	CTGGTTTACCCTCGCTGACA	CCGAGTTGGAAGCATTGTGC
	Slc35b4	TCGCACTTGTTGCTACCGAT	GAATGTTGGCTGTGCTGCAA
	Slc35c1	GCCACCATCCTGTTCTCCAA	GAACAGAGGCTGAGTGGACC
	Slc35c2	GTCTCCACTTGTCCACCTCTG	AACTCGGAGAAGCCCAAACC
	Slc35d1	TGTTTGCTGAAGGCGCTCTA	GCTACGAAGGCTCCGATGAT
	Slc35d2	AGACCAGAGGTCCCACACAT	GCTCATTGTGCAGTGACGTG

Table 2.

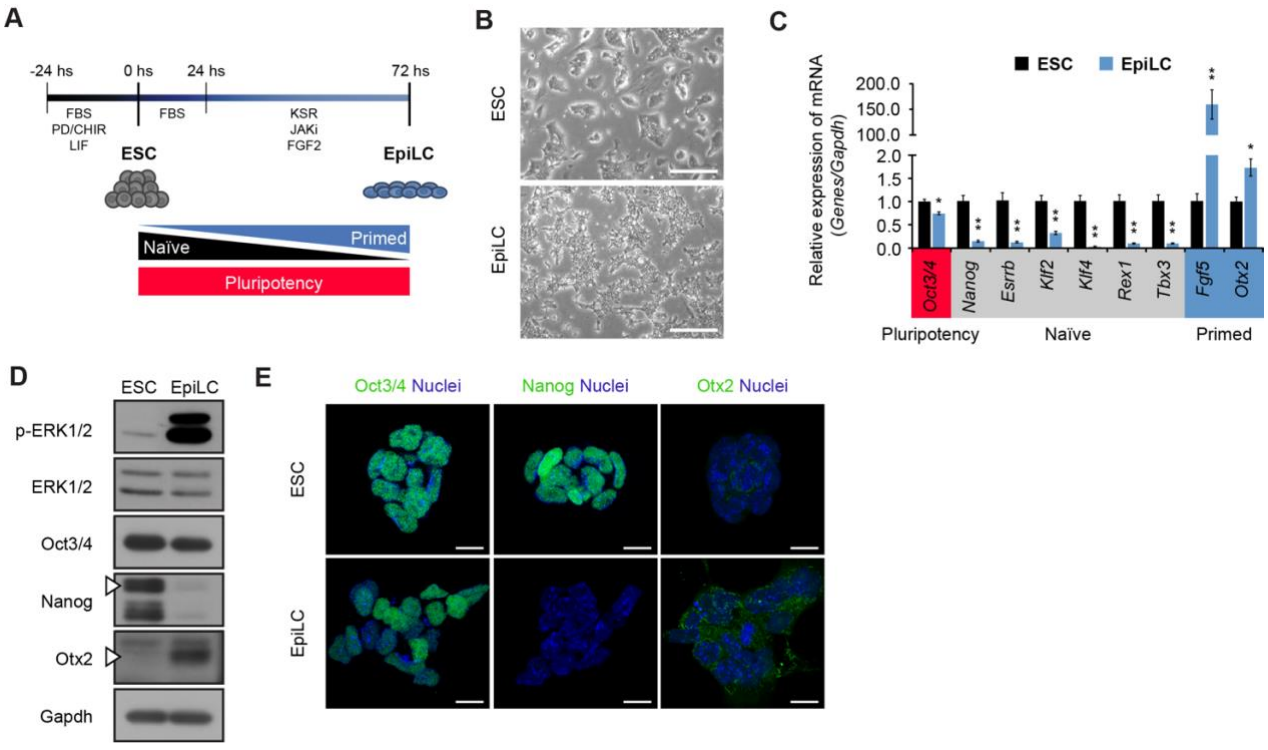
Class	Ab/Lectin	Source	Dilution
	H3	Cell Signaling; #9715	1:100
	H3K27me3	Cell Signaling; #9733	1:100
***N*-glycan**	GNA	EY Laboratories; BA-7401-2	1 µg
	LCA	VECTOR Laboratories; B-1045	1 µg
	ConA	Cosmo Bio; J203	1 µg
	PHA-E4	VECTOR Laboratories; B-1125	1 µg
	PHA-L4	VECTOR Laboratories; B-1115	1 µg

	DSL	VECTOR Laboratories; B-1185	1 µg
O-glycan	VVA-B4	Kindly gifted by Prof. Yamamoto	1 µg
	HPA	EY Laboratories; F-3601-1	10 µg
	PNA	Cosmo Bio; J514	10 µg
	PNA	Cosmo Bio; J214	1 µg
	MAH	Kindly gifted by Prof. Yamamoto	1 µg
	MECA-79	Santa Cruz Biotechnology; sc-19602	1:5
	RL-2	Santa Cruz Biotechnology; sc-59624	1:100
GAG	JM403	Seikagaku Corporation; 370730	1:100
	Hepss-1	Kindly gifted by Prof. Yoshie	1:50
	mCochlin-Fc	Kindly gifted by Prof. Yamamoto	1:10
	CS-56	Sigma Aldrich; C8035	1:100
	MO-225	TCI chemicals; A2872	1:50
	LY111	TCI chemicals; A3143	1:10
	R-10G	Kindly gifted by Prof. Kawasaki	1:100
	5D4	Seikagaku Corporation; 270427	1:100
GSL	CD77	BD biosciences 551352	1:50
	Choleratoxin subunit B	Sigma Aldrich; C1655	1 µg
	SSEA-3	Abcam; ab16286	1:10
	Forsmann antigen	Santa Cruz Biotechnology; sc-23939	1:100
	SSEA-4	Santa Cruz Biotechnology; sc-21704	1:100
Elongation/capping	ECorL	Kindly gifted by Prof. Yamamoto	1 µg
	WFA	Kindly gifted by Prof. Yamamoto	1 µg
	MAL-I	Kindly gifted by Prof. Yamamoto	1 µg
	SNA	EY Laboratories; F-6802-1	10 µg
	SNA	VECTOR Laboratories; B-1305	1 µg
	735	Absolute antibody; Ab00240-1.6	1:100
	AAL	VECTOR Laboratories; B-1395	1 µg
	SSEA-5	MACS Miltenyi Biotec; 130-106-129	1:11
	SSEA-1	Immunotech; 1954	1:3
	SSEA-1	Beckman Coulter; IM1423U	1:3
	Alexa Fluor 488 anti-mouse IgG	Life Technologies; A-11001	1:300/10000
	FITC anti-mouse IgM	Sigma Aldrich; F9259	1:300
	Alexa Fluor 488 anti-rabbit IgG	Life Technologies; A-11008	1:300/10000
	Alexa Fluor 488 anti-rat IgM	Invitrogen; A-21212	1:300
	FITC Streptavidin	Biolegend; 405201	1:300
	FITC anti-human IgG Fc	Biolegend; 409310	1:300

3. Results

3-1. ESC differentiation to EpiLCs

To examine the glycosylation dynamics occurring during the transition from naïve to primed pluripotency, we differentiated EpiLCs from ESCs. ESCs and EpiLCs maintain the pluripotent state by LIF and FGF signaling, respectively (Niwa et al., 2009; Lanner and Rossant, 2010). A previously established protocol (Hayashi et al., 2011) with slight modifications was used to induce EpiLCs differentiation (Fig. 3A). Compared to ESCs, EpiLCs exhibits a typical flattened morphology after 72h of induction (Fig. 3B). Real-time PCR analysis showed a negligible change in the pluripotency marker *Oct3/4*, whereas genes associated with the naïve state, such as *Nanog, Esrrb, Klf2, Klf4, Rex1* and *Tbx3* were substantially downregulated, together with a striking increase in the primed markers *Fgf5*, and *Otx2* (Fig. 3C), consistently with previous results (Hayashi et al., 2011; Du et al., 2018). Oct3/4 protein was retained in EpiLCs at levels similar to those in ESCs, while the Nanog decreased and Otx2 increased. Moreover, the phosphorylation level of ERK1/2, a downstream kinase involved in FGF signaling (Lanner and Rossant, 2010), was significantly higher in EpiLCs (Fig. 3D). The expression of these markers was further assessed by immunostaining: Oct3/4 staining slightly decreased during the transition from naïve to primed, whereas Nanog was not detected and Otx2 was strongly expressed in EpiLCs, confirming the successful differentiation of EpiLCs from ESCs (Fig. 3E).



(see legend in the next page)

Fig. 3. ESCs differentiation into EpiLCs.

(A): Schematic representation of EpiLC differentiation protocol. **(B):** Cell morphology of ESCs (upper panel), and EpiLCs (lower panel). Scale bar, 200 μm. **(C):** Transcriptional analysis of pluripotency (red), naïve (grey), and primed (blue) markers normalized against *Gapdh* in ESCs and EpiLCs, and shown as a fold change relative to ESCs. **(D):** Representative image of a western blot analysis of p-ERK1/2, ERK1/2, Oct3/4, Nanog, Otx2 and Gapdh. **(E):** Representative image of permeabilized ESCs and EpiLCs using anti-Oct3/4, anti-Nanog and anti-Otx2 Abs. Nuclei were stained with Hoechst. Scale bar, 10 μm. Values are shown as means ± s.e.m. of three independent experiments.

3-2. Glycosylation dynamics in pluripotent stem cells

3-2-1. N-linked glycosylation

N-linked glycosylation involves the transfer of a tetradecasaccharide core unit from a lipid-linker donor to an asparagine of nascent proteins at the endoplasmic reticulum (ER). Following transfer, the *N*-glycan undergoes trimming and a quality control process. Fully folded *N*-glycans are transported to the Golgi where they are subjected to processing before transportation to the plasma membrane, whereas misfolded proteins are recycled, resulting in the release of free oligosaccharides (FOS) (Weerapana and Imperiali, 2006; Hirayama et al., 2010) (Fig. 4A). *N*-glycans and FOS composition was measured by mass spectrometry (MS) analysis. The total amount of *N*-glycans was similar between ESCs and EpiLCs. Among the detected *N*-glycan subclasses, namely high mannose-type (HM), pauci-mannose (PM), and complex/hybrid (C/H), HM structures were the most abundant in both ESCs and EpiLCs, confirming previous results obtained in conventional hESCs and hiPSCs (Fujitani et al., 2013). The total amount of fucosylated *N*-glycans was higher in EpiLCs compared to ESCs and was characterized by increased levels of both fucosylated PM and C/H structures. Total sialylated *N*-glycans, mainly present in an α2,6 configuration, were enhanced in EpiLCs (Fig. 4B; Fig. 5). FOS amount were reported to increase under ER stress condition (Hirayama et al., 2010). A substantial reduction in the total amount and relative percentage of FOS was observed in EpiLCs, indicating a reduction of ER stress upon transition from ESC to EpiLC (Fig. 4C; Fig. 6). Notably, a recent study showed that ER stress promotes interleukin-6 family expression, including LIF, in mouse astrocytes and macrophages (Sanchez et al., 2019), suggesting a connection between FOS, ER stress, and LIF expression in ESCs. Real-time PCR analysis of *N*-glycosylation pathway-specific genes well correlated with MS data. Indeed, *Fut8*, the sole enzyme responsible for

catalyzing *N*-glycans core fucosylation (Wang et al., 2005), showed an increased expression. Interestingly, a significant enhancement of *Uggt2* was observed in EpiLCs, suggesting that a reduction in FOS-mediated ER-stress is partially mediated by increased expression of enzymes involved in *N*-glycans quality control and folding (Fig. 4D). To obtain more details, a FACS profiling using a set of lectins that recognize different *N*-glycan structures was performed. As a result, we observed a shift to shorter HM by *galanthus nivalis* agglutinin (GNA). Moreover, the increase in core fucosylated *N*-glycans observed at the structural and transcriptional level in EpiLCs was further confirmed by *lens culinaris* agglutinin (LCA) staining. In addition, *phaseolus vulgaris* erythroagglutinin (PHA-E4) and *phaseolus vulgaris* leucoagglutinin (PHA-L4) staining indicated an enhancement in *N*-glycans carrying bisecting GlcNAc and 2,6-branching in EpiLCs, suggesting a relevant role for these structures during EpiLC differentiation (Fig. 4E; Fig. 7). Collectively, HM structures and core fucosylated *N*-glycans not carrying bisecting GlcNAc are observed in ESCs, whereas short HM structures and core fucosylated *N*-glycans carrying bisecting GlcNAc and 2,6-branching are present in EpiLCs (Fig. 4F).

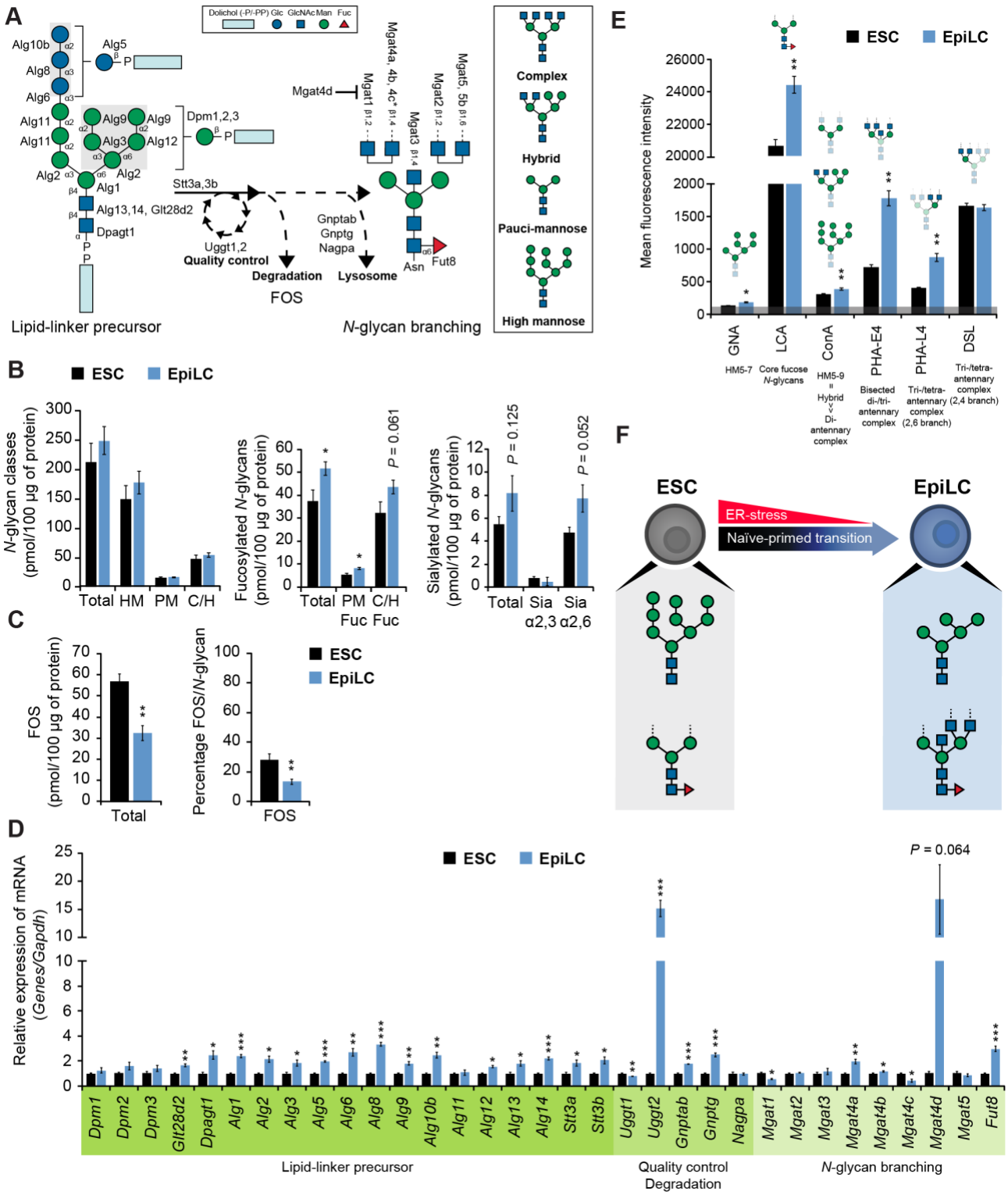


Fig. 4. *N*-glycosylation comparative analysis in ESCs and EpiLCs.

(A): Schematic diagram of the *N*-glycosylation pathway. The asterisk denotes enzyme putative activity. **(B):** Absolute amount of *N*-glycans detected by MS. High mannose (HM), pauci-mannose (PM), and complex/hybrid (C/H). Fucose: Fuc; sialylation: Sia. **(C):** Absolute amount of FOS detected by MS (left panel) and FOS percentage relative to total

(legend continues in the next page)

N-glycans (right panel). **(D):** Transcriptional analysis of *N*-glycosylation-specific enzymes normalized against *Gapdh* and shown as a fold change relative to ESCs. **(E):** *N*-glycans profiling by FACS using specific lectins. Lectins specificities are stated below the labels and schematically represented above each histogram. The grey line at the bottom represents the negative control staining. **(F):** Model of *N*-glycome alterations during naïve to primed transition. Values are shown as means ± s.e.m. of three independent experiments in **(D), (E)** and four independent experiments in **(B), (C)**.

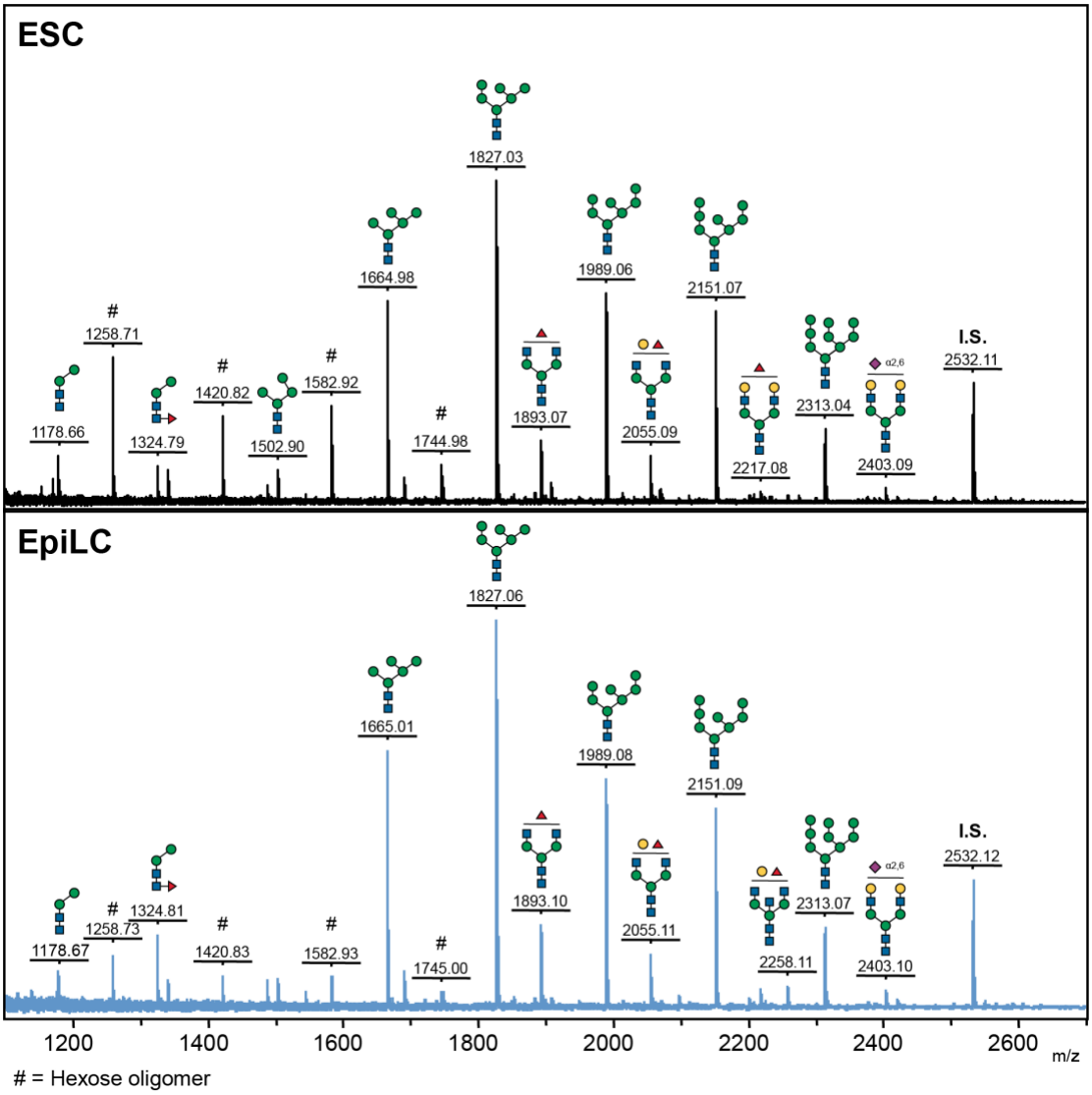


Fig. 5. *N*-glycosylation MS spectra.

Representative *N*-glycans MS spectra of ESCs and EpiLCs. To allow intensity comparison across ESCs and EpiLCs, MS signals in ESCs and EpiLCs spectra are scaled to the same I.S. (internal standard) signal area. Representative estimated structures are shown.

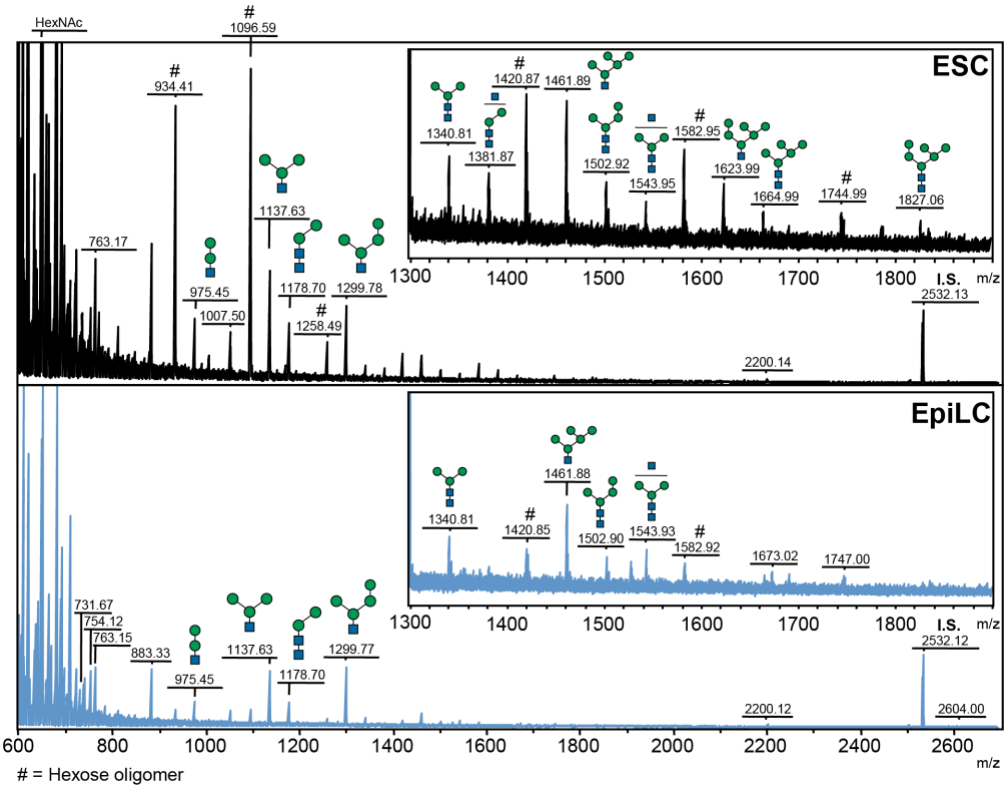


Fig. 6. FOS MS spectra.

Representative FOS MS spectra of ESCs and EpiLCs. To allow intensity comparison across ESCs and EpiLCs, MS signals in ESCs and EpiLCs spectra are scaled to the same I.S. (internal standard) signal area. Representative estimated structures are shown.

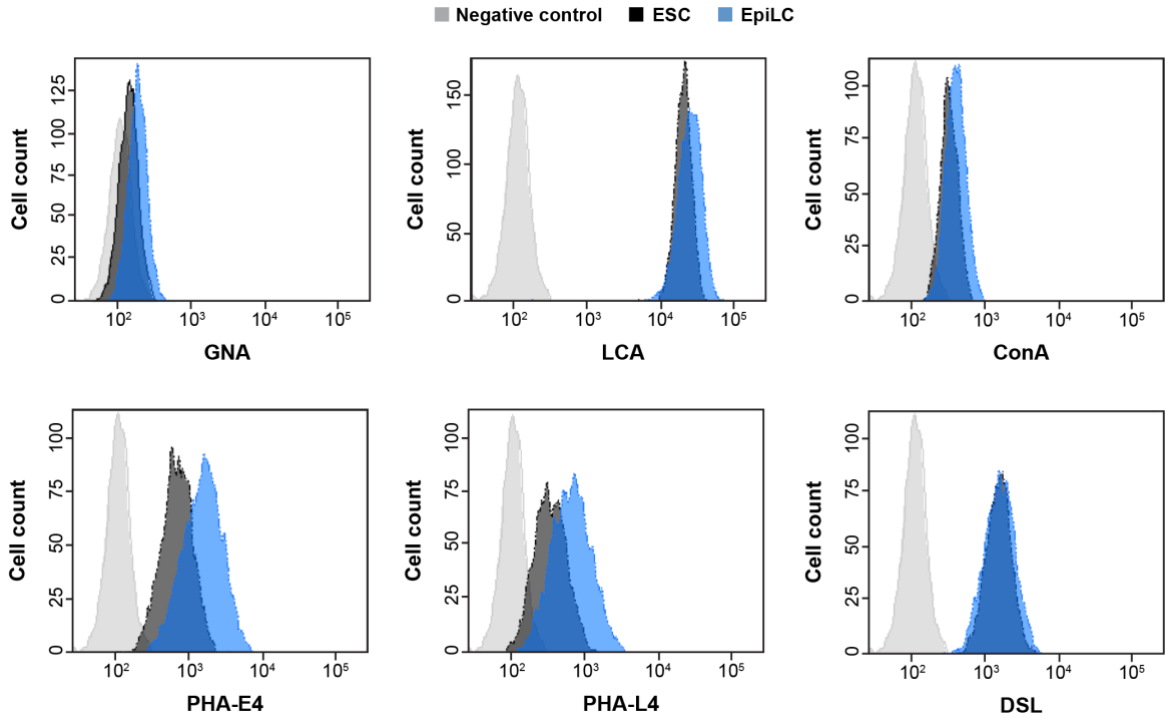


Fig. 7. FACS profiling of *N*-glycan structures.

Representative FACS histograms. Negative control: grey; ESC: black; EpiLC: blue.

3-2-2. O-linked glycosylation

O-linked glycosylation is characterized by the initial addition of a monosaccharide to the hydroxyl group of a serine/threonine residue on the target polypeptide. Various monosaccharides can be added to serine/threonine giving rise to different two major *O*-glycosylation subclasses: mucin-type *O*-glycosylation and not mucin-type *O*-glycosylation. Mucin-type *O*-glycosylation is initiated in the Golgi by a large family of up to 19 transferases in mouse that add *N*-acetylgalactosamine (GalNAc) on a serine/threonine residue to form Tn antigen. Tn antigen can be further elongated into one of 4 distinct core extensions: T antigen (Core 1), Core 2, Core 3 and Core 4 (Bennett et al., 2012). Not mucin-type *O*-glycosylation takes place in the ER, except for the *O*-linked β-*N*-acetylglucosylation (*O*-GlcNAcylation) catalyzed by Ogt, the sole glycosylation that occurs in the cytoplasm and nucleus (Torres and Hart, 1984) (Fig. 8A). *O*-glycome profiling by MS revealed that HexNAc (Tn antigen or *O*-GlcNAc) was the most prominent *O*-glycan structure in both ESCs and EpiLCs. Moreover, T antigen and its modified structures were the only detected *O*-glycans among the mucin-type *O*-glycosylation core structures in both ESCs and EpiLCs (Fig. 8B; Fig. 9), indicating that the T antigen elongation pathway is the most abundant during early embryonic development. Transcriptionally, a substantial upregulation of both mucin-type *O*-glycan and not mucin-type *O*-glycan-specific genes was observed in EpiLCs. In particular, the expression of *Galnt3*, one of the enzymes involved in the formation of Tn antigen, was dramatically increased (Fig. 8C), suggesting its possible major involvement in the formation of Tn antigen in EpiLCs, similarly to previous observations in trophoblast stem cells (Raghu et al., 2019). FACS profiling showed a shift from short mucin-type *O*-glycans, as indicated by *Vicia villosa* agglutinin-B4 (VVA-B4) and *peanut* agglutinin (PNA), which recognize Tn and T antigens, respectively, to elongated or branched mucin-type *O*-glycan structures, as shown by MECA-79 antibody (Ab) staining in EpiLCs. Intracellular *O*-GlcNAc was present in both ESCs and EpiLCs, despite it increased in EpiLCs, reflecting *Ogt* expression upregulation (Fig. 8D,E; Fig. 10).

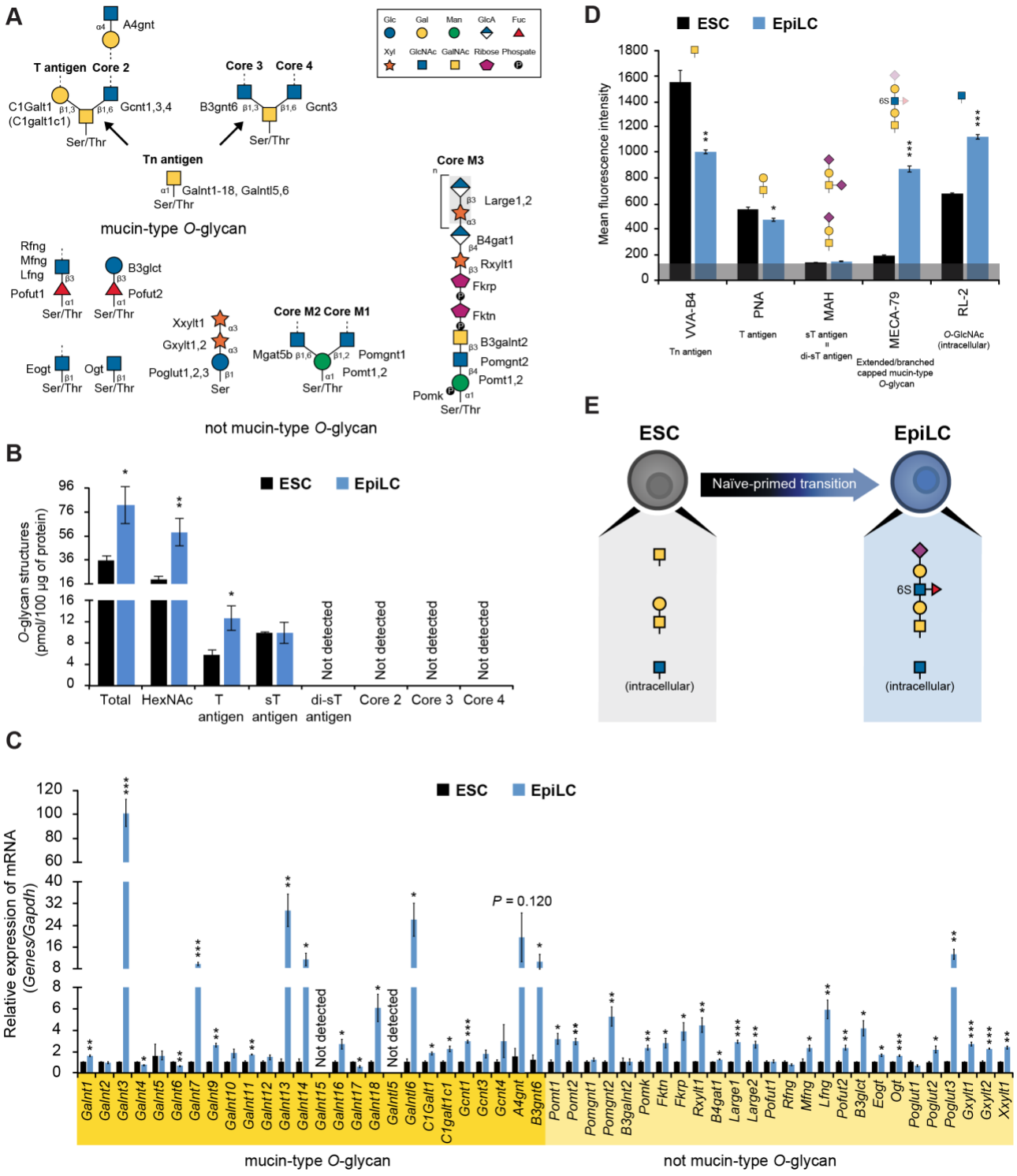


Fig. 8. *O*-glycosylation comparative analysis in ESCs and EpiLCs.

(A): Schematic diagram of the *O*-glycosylation pathway. **(B):** Absolute amount of *O*-glycans detected by MS. HexNAc: GalNAc or GlcNAc. **(C):** Transcriptional analysis of *O*-glycosylation-specific enzymes normalized against *Gapdh* and shown as a fold change relative to ESCs. **(D):** *O*-glycans profiling by FACS using specific lectins/Abs. Lectin/Ab specificities are stated below the labels and schematically represented above each histogram. The grey line at the bottom represents the negative control staining. **(E):** Model

of *O*-glycome changes during naïve to primed transition. Values are shown as means ± s.e.m. of three independent experiments in **(C), (D)** and four independent experiments in **(B)**.

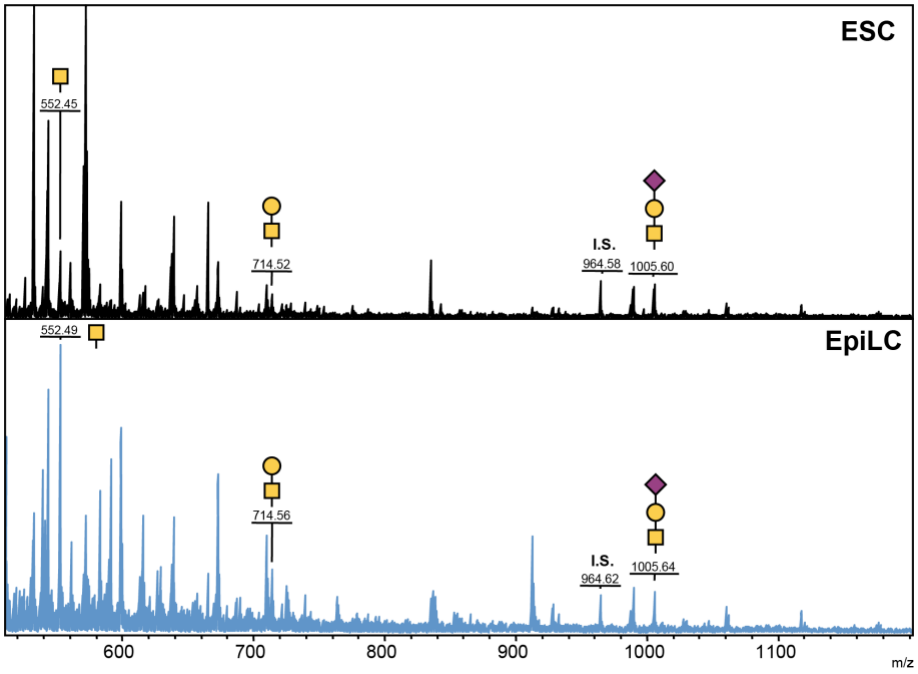


Fig. 9. *O*-glycosylation MS spectra.

Representative *O*-glycome MS spectra of ESCs and EpiLCs. To allow intensity comparison across ESCs and EpiLCs, MS signals in ESCs and EpiLCs spectra are scaled to the same I.S. (internal standard) signal area. Representative estimated structures are shown.

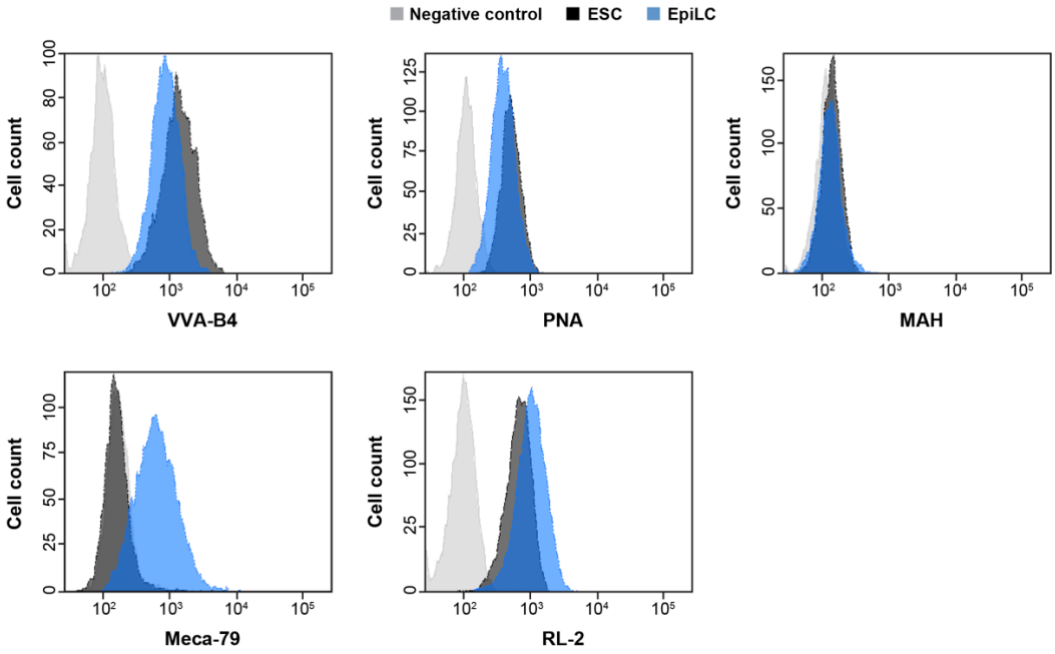


Fig. 10. FACS profiling of *O*-glycan structures.

Representative FACS histograms. Negative control: grey; ESC: black; EpiLC: blue.

3-2-3. Glycosaminoglycans

Glycosaminoglycans (GAG) are polysaccharides consisting of repeating disaccharide units attached to a core protein. Heparan sulfate (HS), chondroitin sulfate (CS) and dermatan sulfate (DS) are connected to a serine of the core protein via a tetrasaccharide linker and are categorized by their different disaccharide unit composition, sulfation and epimerization. Keratan sulfate (KS) is a sulfated polylactosamine chain extended from a GlcNAc on *N*-glycans or *O*-glycans. Conversely, unlike all the other GAG classes that are synthesized in the Golgi on core proteins, hyaluronan (HA) polymerization occurs at the cell membrane and is not linked to any protein (Lindahl et al., 2015-2017) (Fig. 11A). GAG quantification was performed by high-performance liquid chromatography (HPLC). A sharp increase in the total amount of GAG in EpiLCs was observed. Especially, HS and CS/DS, which represent the vast majority of detected GAG, were strongly upregulated, whereas HA levels did not differ between ESCs and EpiLCs; KS and 3-*O*-sulfation of HS could not be detected due to technical limitations. The absolute amount of HS sulfation was substantially higher in EpiLCs, with a specific increase in mono-sulfated 6S and di-sulfated 2SNS, indicating that variations of the HS sulfation pattern occur from early developmental stages. An overall higher level of CS/DS structures was observed in EpiLCs, with predominancy of Unit A, followed by Unit C and unsulfated Unit O (Fig.11B; Fig. 12). Real-time PCR analysis of GAG-specific genes well correlated with HPLC data. Indeed, 2SNS and 6S sulfation enhancement in EpiLCs was followed by a significant enhancement of *Ndst2-4* and *Hs2st1*, and *Hs6st2* expression, respectively (Fig. 11C). FACS staining confirmed the results obtained by HPLC. In addition, increased staining of di-sulfated CS, high-sulfated KS and 3-*O*-sulfated HS were observed in EpiLCs, consistently with the striking enhancement of *Ust* and *Chst15*, *Chst1*, and *Hs3st4-6* expression (Fig. 11D; Fig. 13). The HS sulfation pattern has been shown to be critical for growth factors binding and downstream signaling. Consistent with the present observations in EpiLCs, *N*-sulfation and 3-*O*-sulfation were reported to be required for exit from the naïve pluripotent state via FGF and first apoptosis signal receptor (Fas), respectively (Nishihara, 2018). In conclusion, the ESC GAG profile is characterized by unsulfated HS, mono-sulfated CS and low-sulfated KS. Sulfation of HS, CS and KS markedly increases upon naïve to primed transition (Fig. 11E).

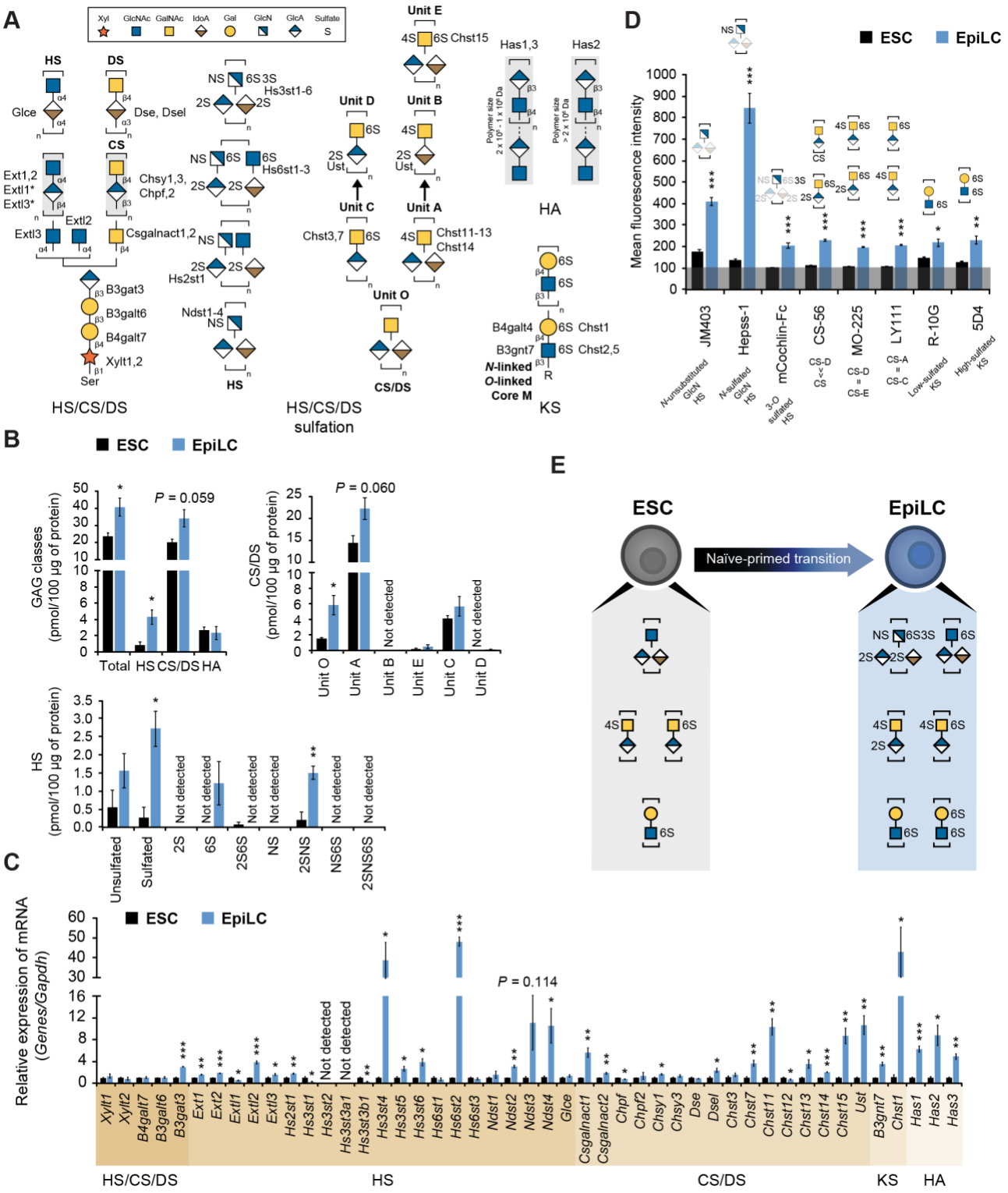


Fig. 11. GAG comparative analysis in ESCs and EpiLCs.

(A): Schematic diagram of the GAG synthetic pathway. Asterisks denote enzymes whose contribution to biosynthesis remains unclear. **(B):** Absolute amount of GAG detected by HPLC. Heparan sulfate (HS), chondroitin sulfate/dermatan sulfate (CS/DS) and hyaluronan (HA). **(C):** Transcriptional analysis of GAG-specific enzymes normalized against *Gapdh* and shown as a fold change relative to ESCs. **(D):** GAG profiling by FACS

using specific Abs. Ab specificities are stated below the labels and schematically represented above each histogram. The grey line at the bottom represents the negative control staining. **(E):** Model of GAG modifications during naïve to primed transition. Values are shown as means ± s.e.m. of three independent experiments.

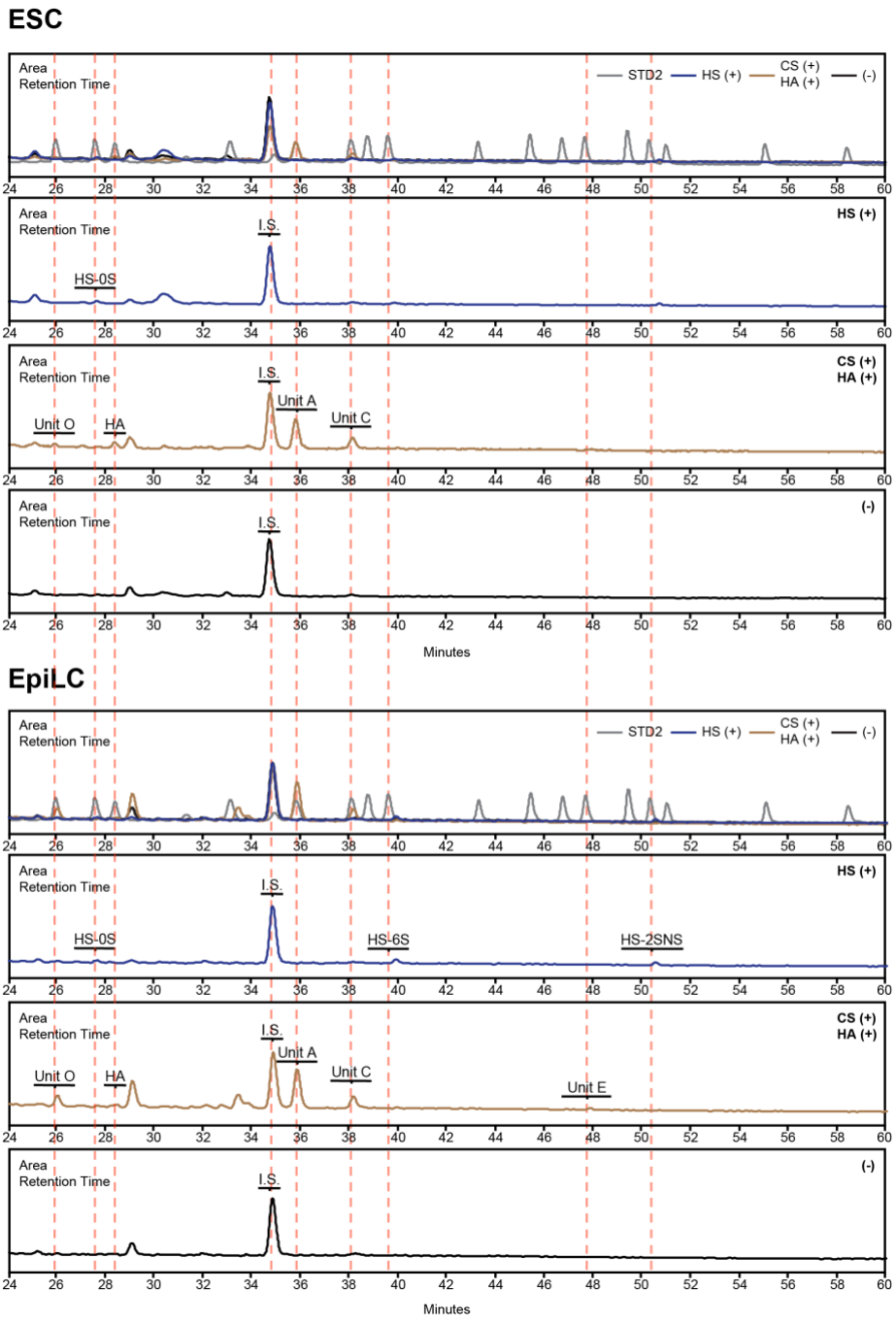


Fig. 12. GAG HPLC chromatogram.

Representative GAG HPLC chromatogram of ESCs and EpiLCs. To allow intensity comparison across ESCs and EpiLCs, HPLC signals are scaled to the same I.S. (internal standard) signal area. Estimated structures are labeled on the chromatogram.

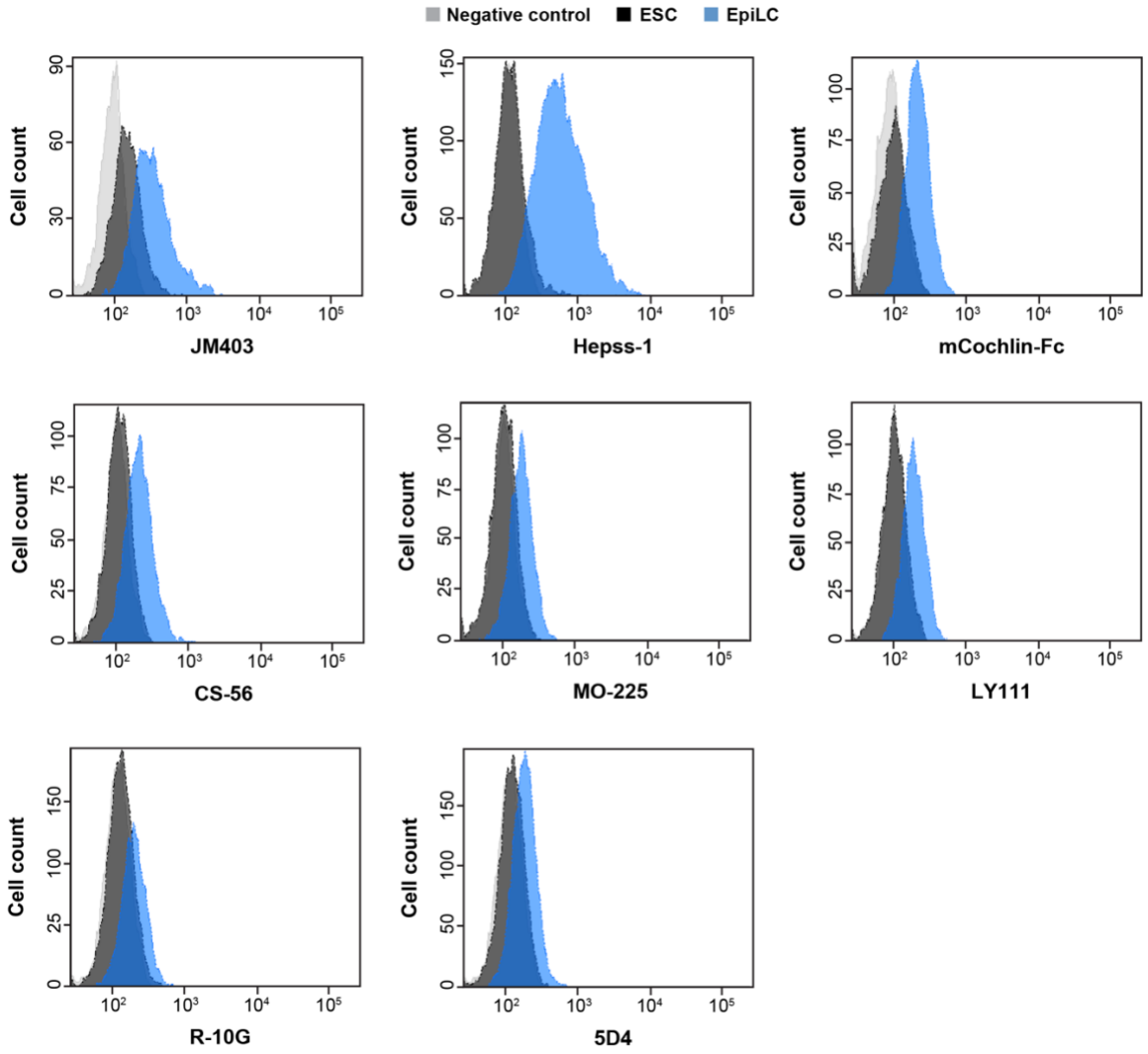


Fig. 13. FACS profiling of GAG structures.

Representative FACS histograms. Negative control: grey; ESC: black; EpiLC: blue.

3-2-4. Glycosphingolipids

Glycosphingolipids (GSL) are characterized by the initial addition of glucose or galactose to a ceramide unit to form glucosylceramide (GlcCer) or galactosylceramide (GalCer), respectively. GlcCer synthesis occurs in the Golgi, where it is further processed to lactosylceramide (LacCer), which is the branching point for the formation of the globo (Gb), ganglio (Gg) and neolacto/lacto ((n)Lc) series. Conversely, GalCer is formed in the ER and is the precursor of the gala-series (Schnaar and Kinoshita, 2015-2017) (Fig. 14A). GSL composition quantified by MS showed a marked reduction of total GSL in EpiLCs. A shift from Gb and (n)Lc to Gg series was observed during ESC differentiation into EpiLCs (Fig. 14B; Fig. 15; Table 3). GlcCer and GalCer could not be detected due to the inherent enzymatic specificity of the EGCase used to release the glycan moieties (EGCase I). The Gb to Gg series switch in EpiLCs observed at the structural level was consistent also at

the expression level. A reduction in the expression of the Gb series-specific enzyme *B3galnt1* was followed by a robust upregulation of the Gg series-specific enzyme *B4galnt1* (Fig. 14C). Moreover, FACS profiling further confirmed the shift from Gb to Gg observed (Fig. 14D), demonstrating that the Gb to Gg series switch previously observed in neurons and embryoid bodies derived from mouse and conventional human ESCs, respectively (Liang et al., 2010; Russo et al., 2018), occurs specifically during the naïve to primed transition. The GSL profile dynamically changes during embryonic development; as a result, specific GSL structures, such as stage-specific embryonic antigen (SSEA)-3, SSEA-4 and Forssman antigen, are used as developmental stage markers (Russo et al., 2018). FACS profiling showed that SSEA-3,4 and Forssman antigen staining mildly increased in EpiLCs (Fig. 14D; Fig. 16). However, SSEA-3,4 and Forssman antigen are structures of the Gb series (Fig. 14A), which we showed to be drastically reduced and undergo a switch to Gg series upon ESC differentiation to EpiLCs, thus suggesting that Gg series structures might be more suitable developmental stage markers. Together, these data show that the GSL composition shifts from Gb and (n)Lc to Gg series during ESC differentiation into EpiLCs, and demonstrate the presence of SSEA-3,4 and Forsmann antigen structures in both ESCs and EpiLCs (Fig. 14E).

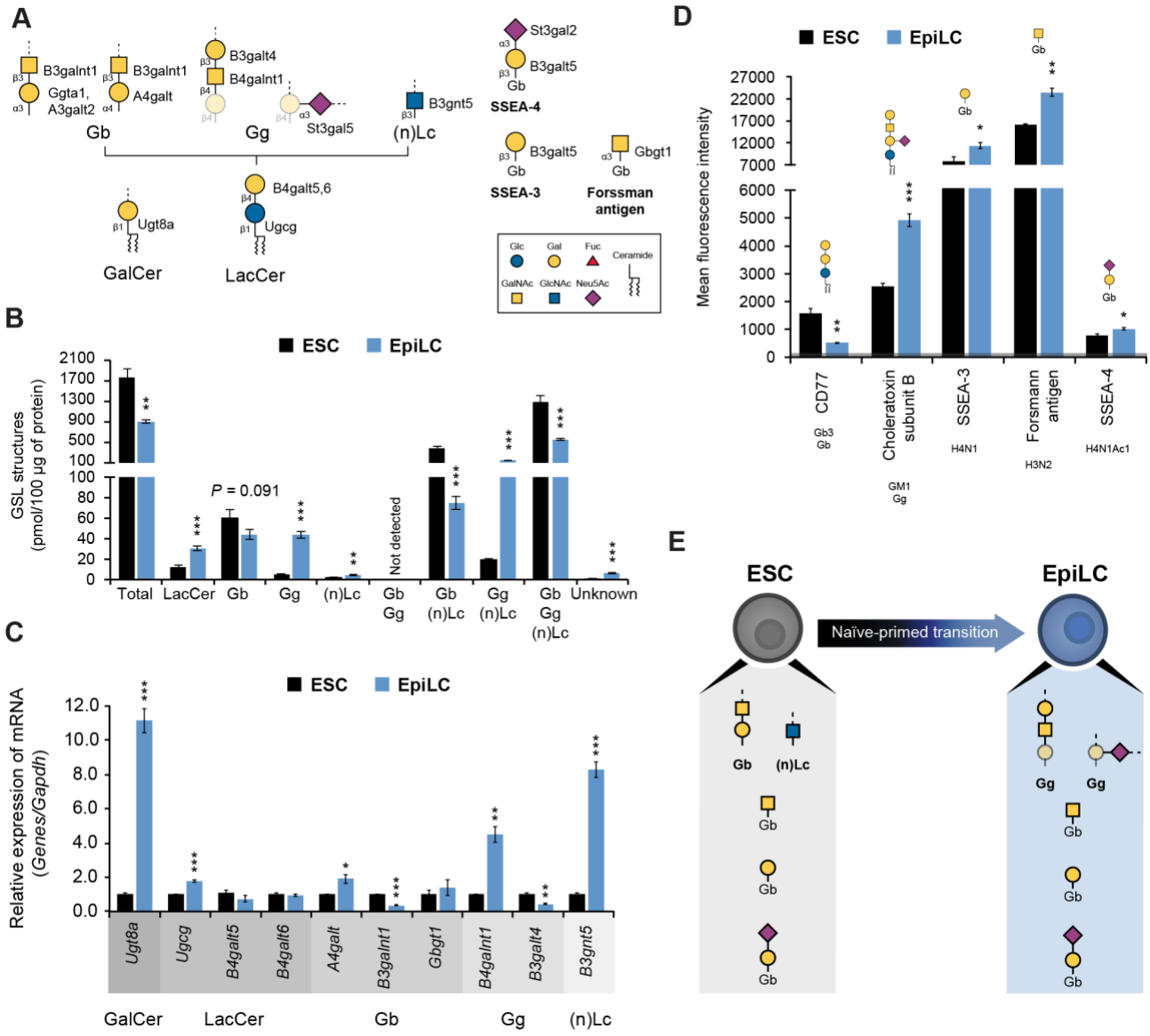


Fig. 14. GSL comparative analysis in ESCs and EpiLCs.

(A): Schematic diagram of the GSL synthetic pathway. **(B):** Absolute amount of GSL detected by MS. Galactosylceramide (GalCer), lactosylceramide (LacCer), globo (Gb), ganglio (Gg) and neolacto/lacto ((n)Lc). **(C):** Transcriptional analysis of GSL-specific enzymes normalized against *Gapdh* and shown as a fold change relative to ESCs. **(D):** GSL profiling by FACS using specific Abs. Ab specificities are stated below the labels and schematically represented above each histogram. The grey line at the bottom represents the negative control staining. **(E):** Model of GSL modifications during naïve to primed transition. Values are shown as means ± s.e.m. of three independent experiments **(C), (D)** and four independent experiments in **(B)**.

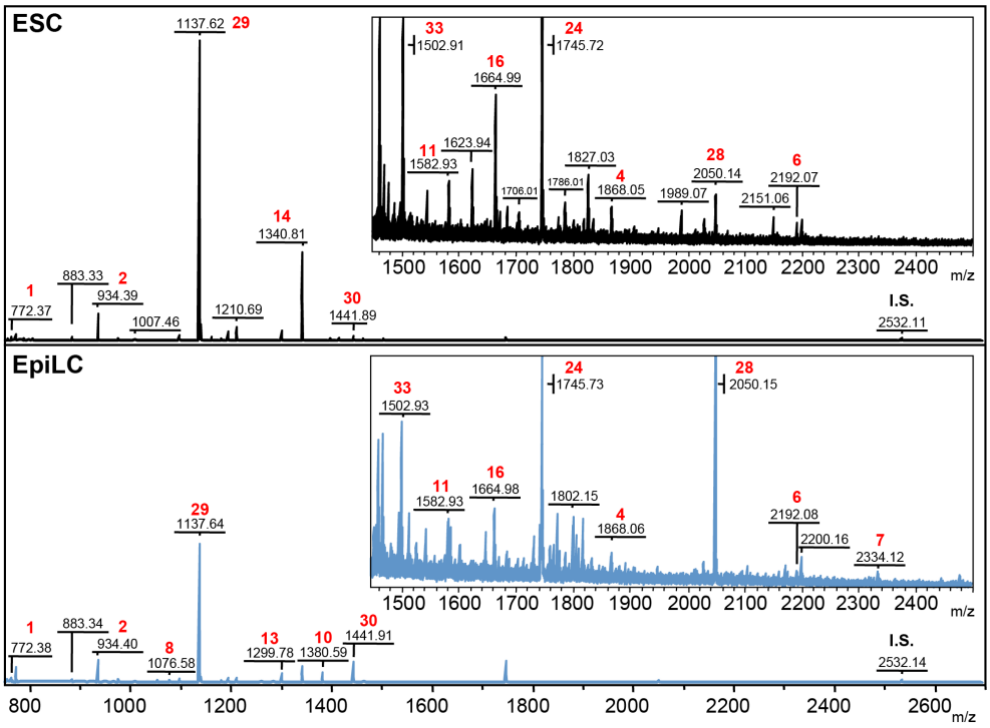


Fig. 15. GSL MS spectra.

Representative GSL MS spectra of ESCs and EpiLCs. To allow intensity comparison across ESCs and EpiLCs, MS signals in ESC and EpiLC spectra are scaled to the same I.S. (internal standard) signal area. Estimated structures are numbered and listed in Table 3.

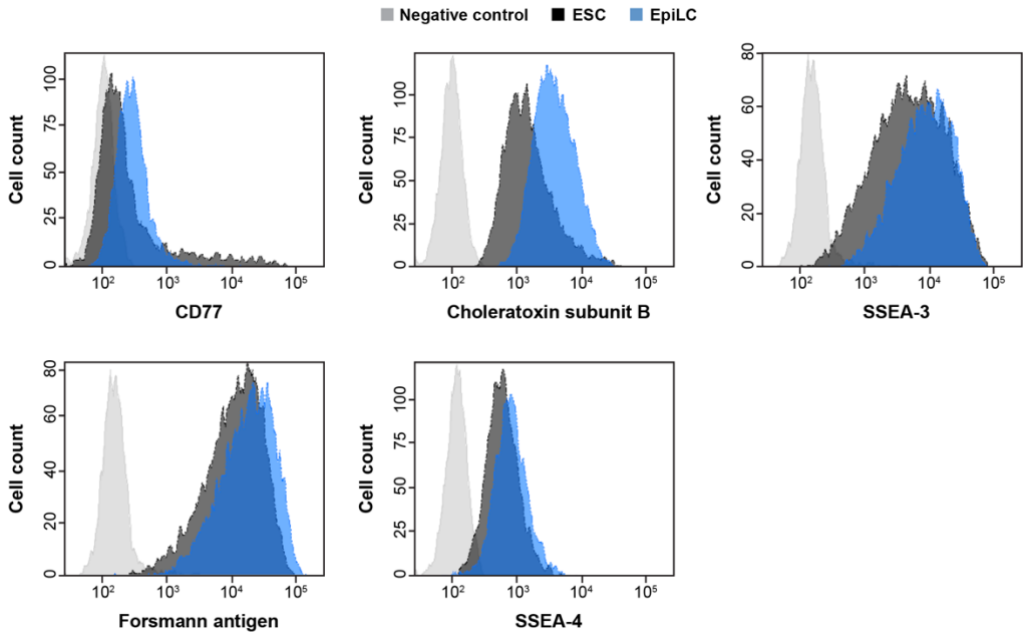


Fig. 16. FACS profiling of GSL structures.

Representative FACS histograms. Negative control: grey; ESC: black; EpiLC: blue.

Table 3

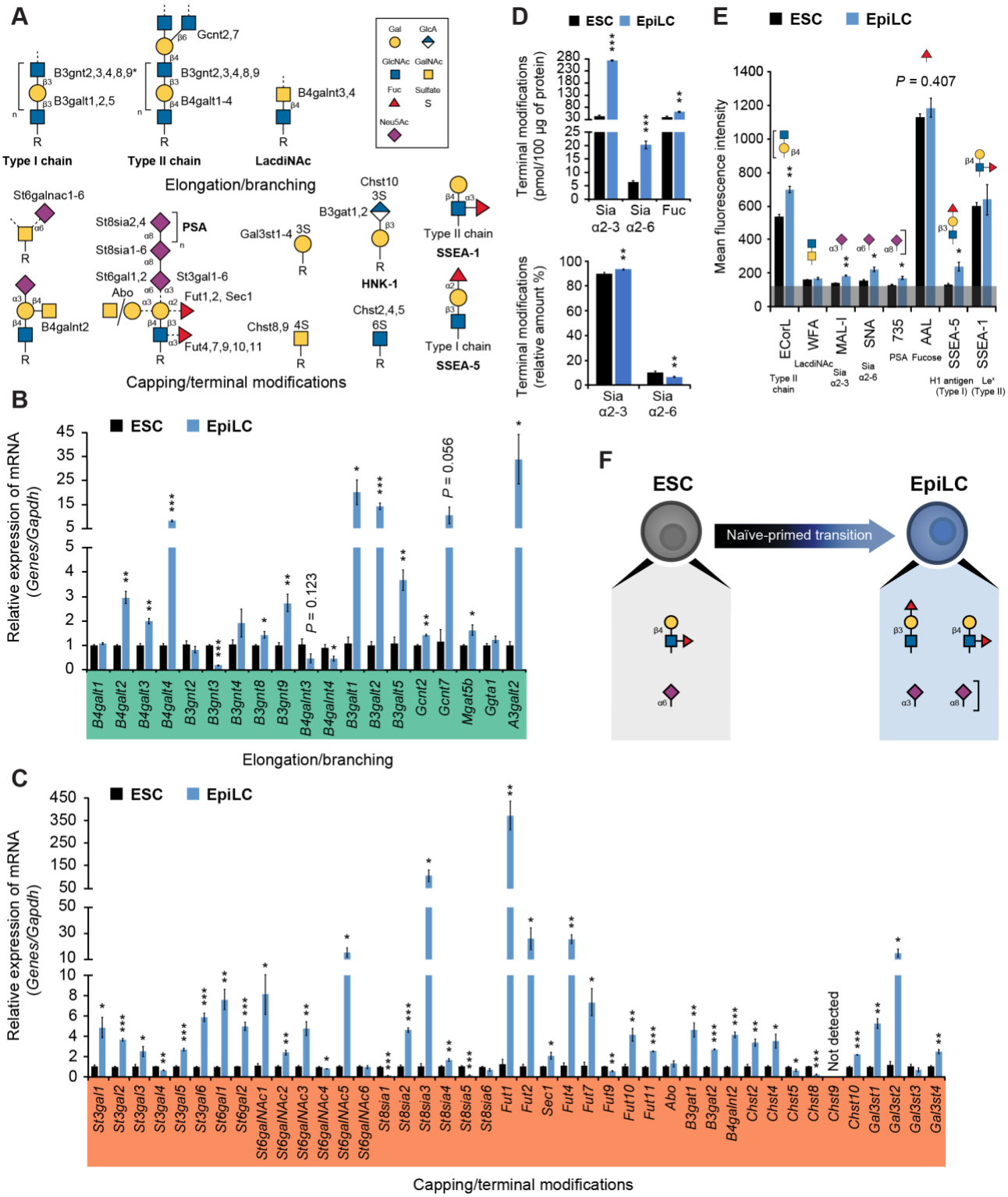
No.	**deduced composition**	**class**
1	(Hex)2	LacCer
2	(Hex)3	Gb
3	(Hex)4(HexNAc)2(2,6NeuAc)1	(n)Lc
4	(Hex)5(HexNAc)3	(n)Lc
5	(Hex)5(HexNAc)3(2,3Neu5Ac)1	(n)Lc
6	(Hex)7(HexNAc)3	(n)Lc
7	(Hex)6(HexNAc)3(2,3NeuAc)1	(n)Lc
8	(Hex)2(2,3NeuAc)1	Gg
9	(Hex)2(HexNAc)1(2,3NeuAc)1	Gg
10	(Hex)2(2,3/2,3NeuAc)2	Gg
11	(Hex)2(HexNAc)1(2,3/2,3NeuAc)2	Gg
12	(Hex)2(2,3/2,3/2,3NeuAc)3	Gg
13	(Hex)4(HexNAc)1	Gb/(n)Lc
14	(Hex)3(HexNAc)2	Gb/(n)Lc
15	(Hex)4(HexNAc)1(2,3NeuAc)1	Gb/(n)Lc
16	(Hex)5(HexNAc)2	Gb/(n)Lc
17	(Hex)4(HexNAc)3	Gb/(n)Lc
18	(Hex)5(HexNAc)2(Fuc)1	Gb/(n)Lc
19	(Hex)2(HexNAc)1	Gg/(n)Lc
20	(Hex)2(HexNAc)2	Gg/(n)Lc
21	(Hex)3(HexNAc)1(Fuc)1	Gg/(n)Lc
22	(Hex)3(HexNAc)1(2,6NeuAc)1	Gg/(n)Lc
23	(Hex)3(HexNAc)1(Fuc)1(2,3NeuAc)1	Gg/(n)Lc
24	(Hex)3(HexNAc)1(2,3/2,3NeuAc)2	Gg/(n)Lc
25	(Hex)3(HexNAc)1(2,3/2,6NeuAc)2	Gg/(n)Lc
26	(Hex)3(HexNAc)1(2,3NeuAc)1(2,3NeuGc)1	Gg/(n)Lc
27	(Hex)4(HexNAc)2(2,3NeuAc)1	Gg/(n)Lc
28	(Hex)3(HexNAc)1(2,3/2,3/2,3NeuAc)3	Gg/(n)Lc
29	(Hex)3(HexNAc)1	Gb/Gg/(n)Lc
30	(Hex)3(HexNAc)1(2,3NeuAc)1	Gb/Gg/(n)Lc
31	(Hex)4(HexNAc)1(Fuc)1	Gb/Gg/(n)Lc
32	(Hex)3(HexNAc)1(2,3NeuGc)1	Gb/Gg/(n)Lc
33	(Hex)4(HexNAc)2	Gb/Gg/(n)Lc
34	(Hex)4(HexNAc)2(Fuc)1	Gb/Gg/(n)Lc
35	(Hex)3(HexNAc)3	unknown
36	(Hex)3(HexNAc)1(2,6/2,6NeuAc)2	unknown
37	(Hex)4(HexNAc)1(Fuc)1(2,6NeuAc)1	unknown
38	(Hex)3(HexNAc)2(Fuc)1(2,6NeuAc)1	unknown
39	(Hex)6(HexNAc)3	unknown

3-2-5. Elongation/branching and capping/terminal modifications

Glycosylation is a complex process involving more than 200 distinct glycosyltransferases and related enzymes in mammals. These can be classified as pathway-specific and pathway-non-specific, which generally include enzymes involved in biosynthetic steps overlapping different glycosylation classes (Fig. 17A). Real-time PCR analysis showed that *B3galt1,2,5* and *B4galt2-4,* involved in type I chain (Galβ1-3GlcNAc) and type II chain (Galβ1-4GlcNAc) structures formation, respectively, were increased, whereas LacdiNAc (GlcNAcβ1-4GalNAc)-specific enzymes *B4galnt3,4*, which was previously reported to positively regulate LIF signaling in ESCs (Sasaki et al., 2011), were reduced in EpiLCs. Consistently with a higher level of KS (Fig. 11D), *B4galt4*, required for KS elongation (Lindahl et al., 2015-2017), was markedly increased. SSEA-3 upregulation observed by MS (Fig. 14D), correlated with a significantly higher expression of *B3galt5*, the enzyme involved in SSEA-3 synthesis on Gb (Schnaar and Kinoshita, 2015-2017) (Fig. 17B). An general enhancement of sialyltransferases, including *St8sia2,4,* the genes involved in polysialic acid (PSA) formation (Takashima and Tsuji, 2011), was observed in EpiLCs. *St6gal1,2*, the only sialyltransferases involved in *N*-glycans sialylation (Takashima and Tsuji, 2011), substantially increased in EpiLCs, demonstrating a correlation between the observed enhancement of α2,6 sialic acid structures on *N*-glycans by MS (Fig. 4B) and the enzyme expression. Fucosyltransferases expression strongly increased in EpiLCs, except for that of *Fut9*, the enzyme synthesizing SSEA-1 (Kudo et al., 2004). Remarkably, *Fut1,2*, involved in the formation of SSEA-5, showed a striking increase. Moreover, sulfotransferases mRNA levels were higher in EpiLCs. In particular, a significant increase in *Chst2,4* (Fig. 17C)*,* the sole sulfotransferases that catalyze the formation of extended or branched capped mucin-type *O*-glycan structure (Uchimura et al., 2005), confirming at the transcriptional level the data obtained by MECA-79 staining (Fig. 8D).

To obtain further insights into the major pathway-non-specific structures, MS collected data were combined with a FACS profiling using a combination of different lectins and Abs. Structures characterized by MS and FACS strongly correlated with the expression of the relative enzymes. An increase in type II chain structures in EpiLCs, as indicated by *erythrina corallodendron* lectin (ECorL) staining was observed, despite Lewis x (Le^x) type II structure (SSEA-1) was unchanged, likely because of decreased *Fut9* expression. Conversely, H1 antigen, a type I structure known as SSEA-5, was significantly enhanced in EpiLCs and only detected at a negligible level in ESCs. Overall fucosylation and sialylation levels were enhanced in EpiLCs, correlating with the observed increased

expression of the fucose and sialic acid nucleotide transporters (Fig. 18). Notably, sialic acid configuration shifted from an α2,6 to an α2,3 configuration in EpiLCs (Fig. 17D). Furthermore, PSA structure was specifically enhanced in EpiLCs (Fig. 17E,F; Fig. 19).



(see legend in the next page)

Fig. 17. Pathway-non-specific comparative analysis in ESCs and EpiLCs.
(A): Schematic diagram of major pathway-non-specific structures. The asterisk denotes enzyme putative activity. **(B):** Transcriptional analysis analysis of elongation/branching enzymes normalized against *Gapdh* and shown as a fold change relative to ESCs. **(C):** Transcriptional analysis of capping/terminal modification enzymes normalized against *Gapdh* and shown as a fold change relative to ESCs. **(D):** Total absolute amount of terminal modifications detected by MS (upper panel). Relative amount of total sialic acid in α2,3 and α2,6 configuration (lower panel). Fucose: Fuc; sialylation: Sia. **(E):** Pathway-non-specific profiling by FACS using specific lectins/Abs. Lectin/Ab specificities are stated below the labels and schematically represented above each histogram. The grey line at the bottom represents the negative control staining. **(F):** Model of pathway-non-specific alterations during naïve to primed transition. Values are shown as means ± s.e.m. of three independent experiments in **(B), (C), (E)** and four independent experiments in **(D)**.

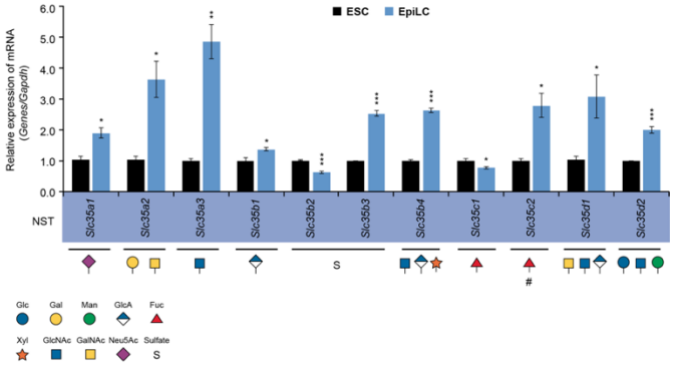


Fig. 18. NST transcriptional analysis in ESCs and EpiLCs.
Transcriptional analysis analysis of NST normalized against *Gapdh* and shown as a fold change relative to ESCs.

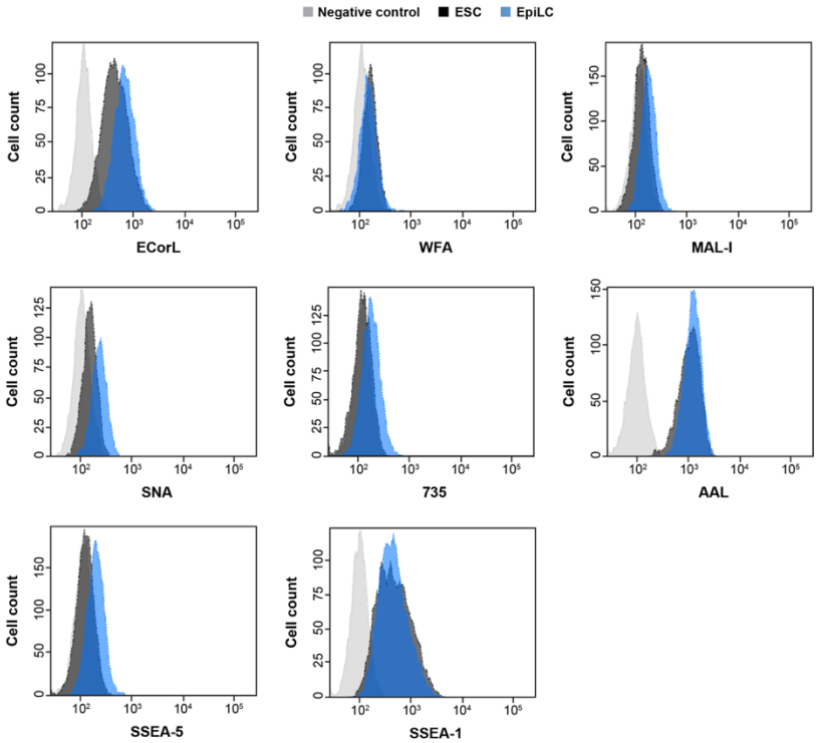
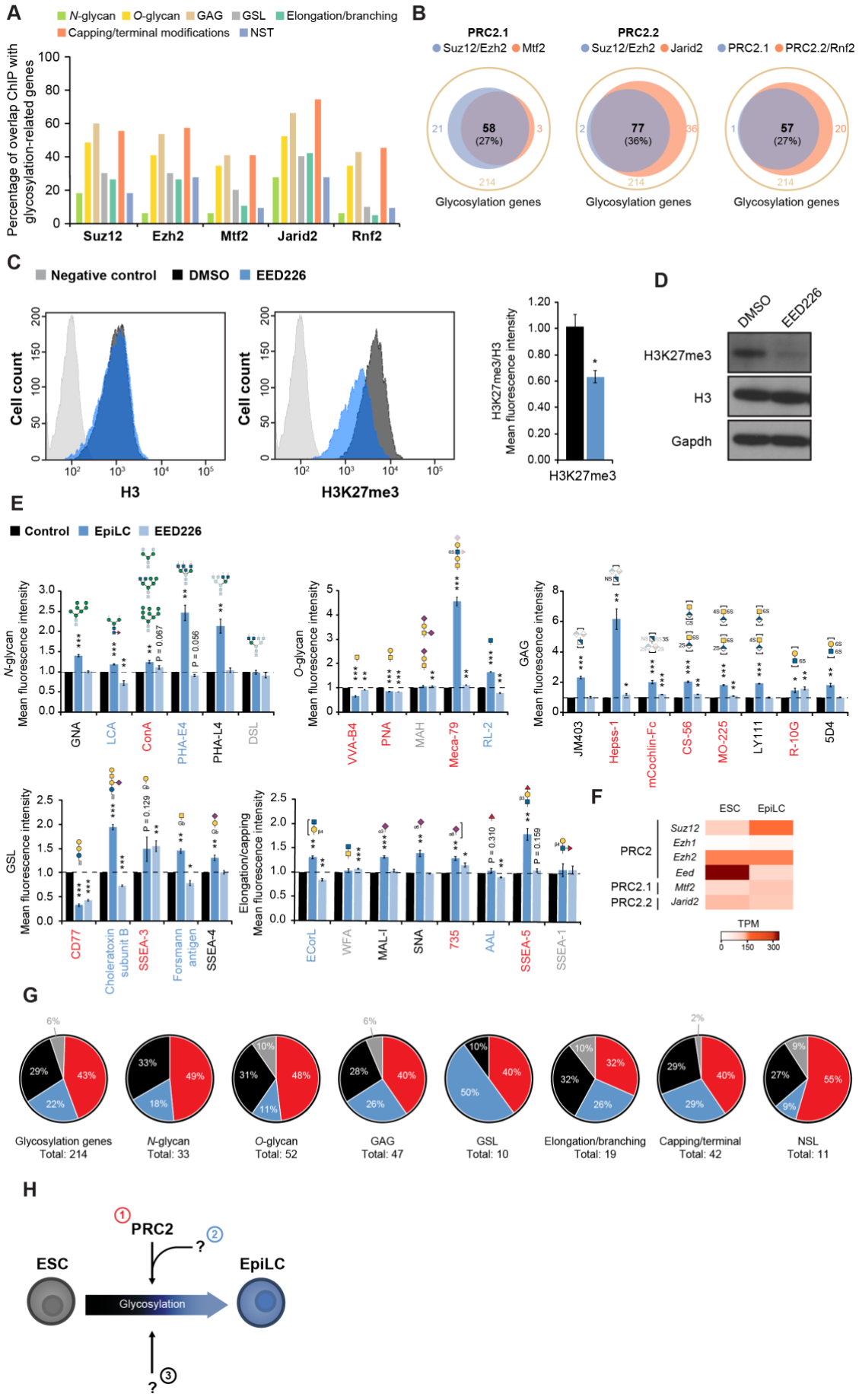


Fig. 19. FACS profiling of pathway-non-specific structures.
Representative FACS histograms. Negative control: grey; ESC: black; EpiLC: blue.

3-3. PRC2 regulates glycosylation alterations during ESCs to EpiLCs differentiation

Since the glycome drastically changes during ESC differentiation to EpiLCs, we hypothesized the presence of a defined regulatory network. To identify putative candidates, an in-depth analysis of previously published chromatin immunoprecipitation sequencing (ChIP-seq) datasets obtained in ESCs using the ChIP-Atlas comprehensive database was performed (Oki et al., 2018) (https://chip-atlas.org). As a result, PRC2 components were observed to occupy a great variety of glycosylation-related genes across all glycosylation classes in ESCs (Fig. 20A; Fig. 21). PRC2, whose core components are Suz12, Eed, and either Ezh2 or Ezh1, is a chromatin-remodeling complex which catalyzes the H3K27me3 modification to promote gene repression. PRC2 can associate with other factors that regulate its chromatin recruitment, such as Mtf2 or Jarid2, to form two different subtypes of PRC2 named PRC2.1 and PRC2.2, respectively. In addition, PRC2 can synergically repress gene expression together with PRC1, which is composed by core components, such as Rnf2 (Deevy and Bracken, 2019, Gökbuget and Blellock, 2019). Further analysis of ChIP-seq data indicated that PRC2.1, PRC2.2 and Rnf2 (PRC1) act cooperatively to regulate around 27% of the glycosylation-related genes in ESCs (Fig. 20B), suggesting that PRC2 is involved in glycosylation changes occurring during ESC differentiation to EpiLCs. To test this hypothesis, ESCs were treated with the PRC2 inhibitor EED226 (Qi et al., 2017) and examined the effects on the glycome. EED226 treatment resulted in a significant reduction of the H3K27me3 modification (Fig. 20C,D). Next, we compared the glycosylation alterations observed by FACS during the ESC differentiation to EpiLCs and in ESCs after EED226 treatment. Strikingly, many structures were increased or decreased both in EpiLCs and EED226-treated ESCs, indicating a direct modulation by PRC2 (red in Fig. 20E). Conversely, other structures showed an opposite trend in EpiLCs and EED226-treated ESCs, suggesting the presence of other regulatory component(s) (blue in Fig. 20E). Finally, some structures were altered in EpiLCs but unchanged in EED226-treated ESCs, implying the presence of PRC2-independent pathway(s) (black in Fig. 20E). Transcriptional analysis of PRC2 core components by RNA-seq in ESCs and EpiLCs showed a decrease at the transcriptional level of the PRC2 core component Eed (Fig. 20F). Moreover, global H3K27me3 has been previously reported to be drastically reduced and redistributed upon ESC differentiation to EpiLCs (Kurimoto et al., 2015), providing further evidence supporting our PRC2-mediated glycosylation regulatory network model during the naïve to primed transition. To further confirm that PRC2 is involved in glycan changes during the ESC to EpiLC transition we

examined the glycosyltransferase and related enzyme expression difference in ESCs vs EpiLCs and EED226 untreated vs EED226 treated samples by RNA-seq. Similar to the changes observed at the structural level, we observed three different patterns of expression: expression increased or decreased both in EpiLCs and EED226- treated ESCs (43% of the glycosyltransferases and related genes); or following an opposite trends of expression (22% of the glycosyltransferases and related genes); or unaffected by PRC2 treatment and changed during the naïve-to-primed transition (29% of the glycosyltransferases and related genes) (Fig. 20G). Collectively, these data indicate the presence of at least three coordinated pathways that control glycosylation dynamics during EpiLC differentiation (Fig. 20H). In conclusion, these findings demonstrate for the first time that PRC2 mediates the glycosylation dynamics during the naïve to primed pluripotency state transition (Fig. 22).



(see legend in the next page)

Fig. 20. PRC2 mediates glycosylation dynamics during ESC to EpiLC transition. **(A):** Previously published ChIP-seq datasets obtained from wild-type/untreated ESCs precipitated: anti-Suz12 Ab (SRX1372665) (Kloet et al., 2016), anti-Ezh2 Ab (SRX2528911) (Perino et al., 2018), anti-Mtf2 Ab (SRX2776968) (Li et al., 2017), anti-Jarid2 Ab (SRX3738839) (Perino et al., 2018) and anti-Rnf2 Ab (SRX191072) (He et al., 2013). Percentages of ChIP occupancy per glycosylation class were determined within a range of ± 5kb and with a threshold for statistical significance set as 50 (1 < 1E-05) calculated by peak-caller MACS2. **(B):** Venn diagrams showing the ChIP occupancies of PRC components overlap on glycosylation genes. **(C):** FACS analysis of H3 and H3K27me3 after EED226 treatment for 48h and relative fluorescence mean intensity shown as a fold change relative to that of DMSO-treated cells. Negative control: grey; DMSO: black; EED226: blue. **(D):** Representative image of a western blot of H3K27me3, H3 and Gapdh in EED226-treated ESCs. **(E):** FACS profiling using specific lectins/Abs in EpiLCs and EED226-treated ESCs, and shown as a fold change relative to ESCs and DMSO-treated ESCs (control), respectively. Lectin/Ab specificities are schematically represented above each histogram. The dotted line indicates the control. **(F):** Heat map of PRC2 core components, PRC2.1 and PRC2.2 main accessory subunits detected in ESCs and EpiLCs by RNA-seq. **(G):** Pie charts showing percentage of glycosylation related genes in each expression pattern by comparing ESCs/EpiLCs and EED226 untreated/EED226 treated samples expression by RNA-seq. Genes following a similar expression pattern during EpiLC differentiation and EED226 treatment (red), opposite patterns (blue), and changed in EpiLCs but not upon EED226 treatment (black) (the grey label indicates genes which expression was unchanged). RNA-seq data were obtained from a single experiment. **(H):** Schematic model of the glycosylation regulatory network during ESC to EpiLC transition: PRC2 direct regulation (1, red), PRC2 regulation together with other unidentified factor(s) (2, blue), and PRC2-independent pathway(s) (3, black) (the grey label indicates unchanged structures). Values are shown as means ± s.e.m. of three independent experiments.

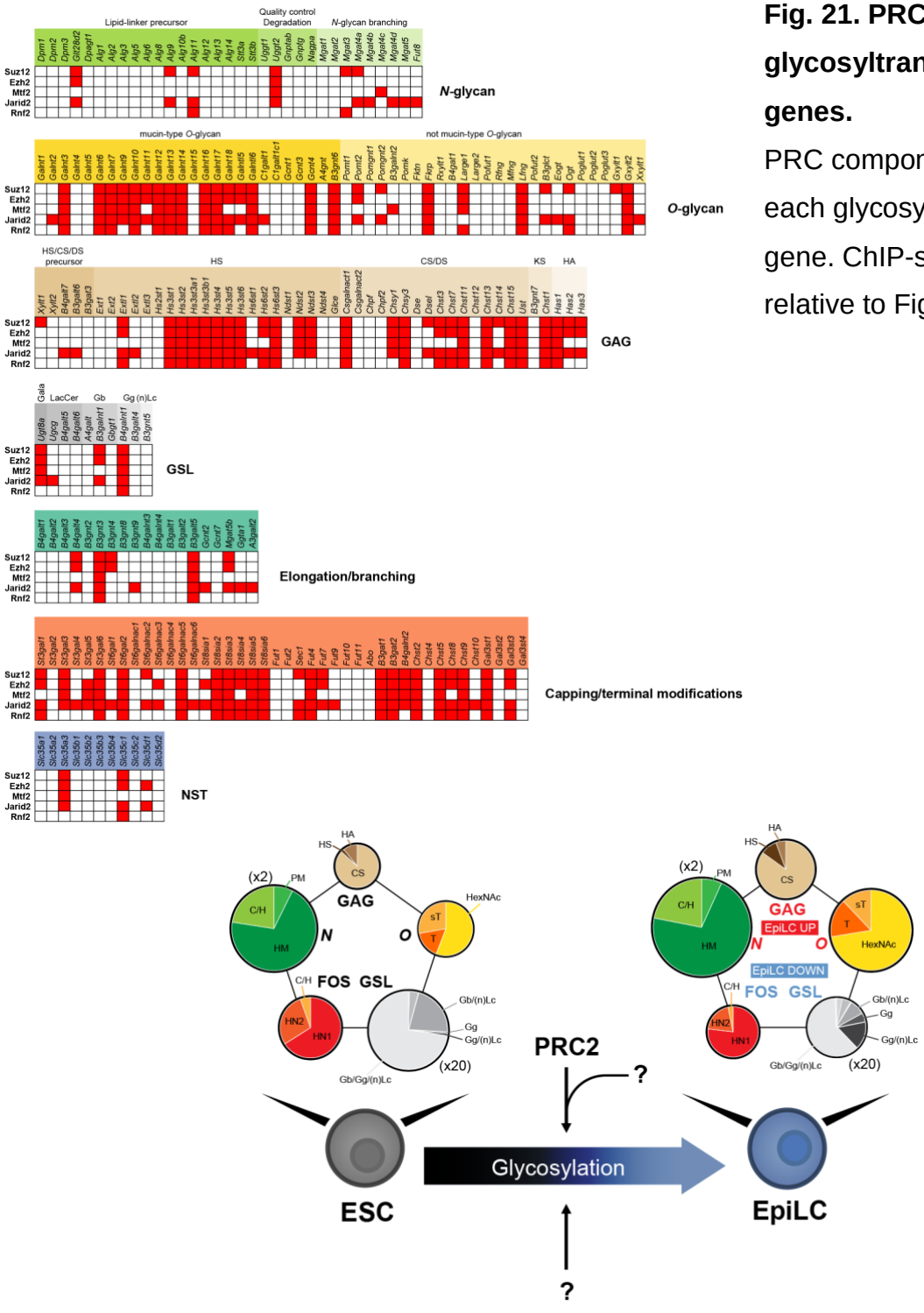


Fig. 21. PRC binding to glycosyltransferases genes.

PRC components binding to each glycosyltransferase gene. ChIP-seq datasets relative to Fig. 20.

Fig. 22. Schematic model of glycosylation dynamics during ESC to EpiLC transition.

Glycome undergoes dramatic alterations during the naïve to primed transition and these changes are partially mediated by PRC2. The size of each pie chart reflects the absolute mean quantity of glycans (pmol/100 µg of protein). *N*-linked and GSL pie charts size is scaled 2 and 20 times, respectively.

3-4. C1galt1 is the most abundant pathway in ESC and during early differentiation

Next, mucin-type *O*-glycosylation role in ESC pluripotency network was explored. Mucin-type *O*-glycosylation is characterized by the initial addition of GalNAc to a serine or threonine residues of the target protein by a family of 19 transferases (Galnts) in mouse to form the so-called Tn antigen. Tn antigen is further elongated by St6galnac1, C1galt1, or B3gnt6 to synthesize sTn antigen, T antigen (Core 1), or Core 3 structure, respectively (Fig. 23A). The partial functional redundancy of Galnts makes analysis of their function highly complex (Bennet et al., 2012); we therefore focused on the catalyzation step following Tn antigen formation that is selectively performed by St6galnac1, C1galt1, or B3gnt6. In addition to sTn antigen, T antigen, and Core 3 structure, the Tn antigen can also be extended to form Core 5 to 8 structures. However, Core 5 to 8 structures have an extremely restricted occurrence and the glycosyltransferases involved in their formation remain to be fully characterized (Brockhausen et al., 2009); thus, they have not been considered in the present study.

Total RNA expression analysis in ESCs showed that *C1galt1* was the most highly expressed among the three enzymes (Fig. 23B). MS analysis of *O*-glycans in ESCs identified T antigen, and its further modifications, as the only detectable structures among sTn antigen, T antigen, and Core 3 structure (Fig. 23C). Notably, the absolute amount of T antigen, and C1galt1-mediated elongation pathway modifications, were consistent among different ESC lines, suggesting that the C1galt1 elongation pathway plays a crucial role in ESCs (Fig. 24A,B). Investigation of the *O*-glycosyltransferase dynamics during early and late differentiation by using an embryoid body (EB) assay was carried out. *C1galt1* showed the highest level of expression among the three *O*-glycosyltransferases between day 0 and day 4, suggesting that the C1galt1-mediated elongation pathway have a role during early differentiation of ESCs. However, at later differentiation stages, expression of all three enzymes increased dramatically, indicating that ESC cells surface is decorated by several types of *O*-glycan structure at late differentiation stages (Fig. 24C-E).

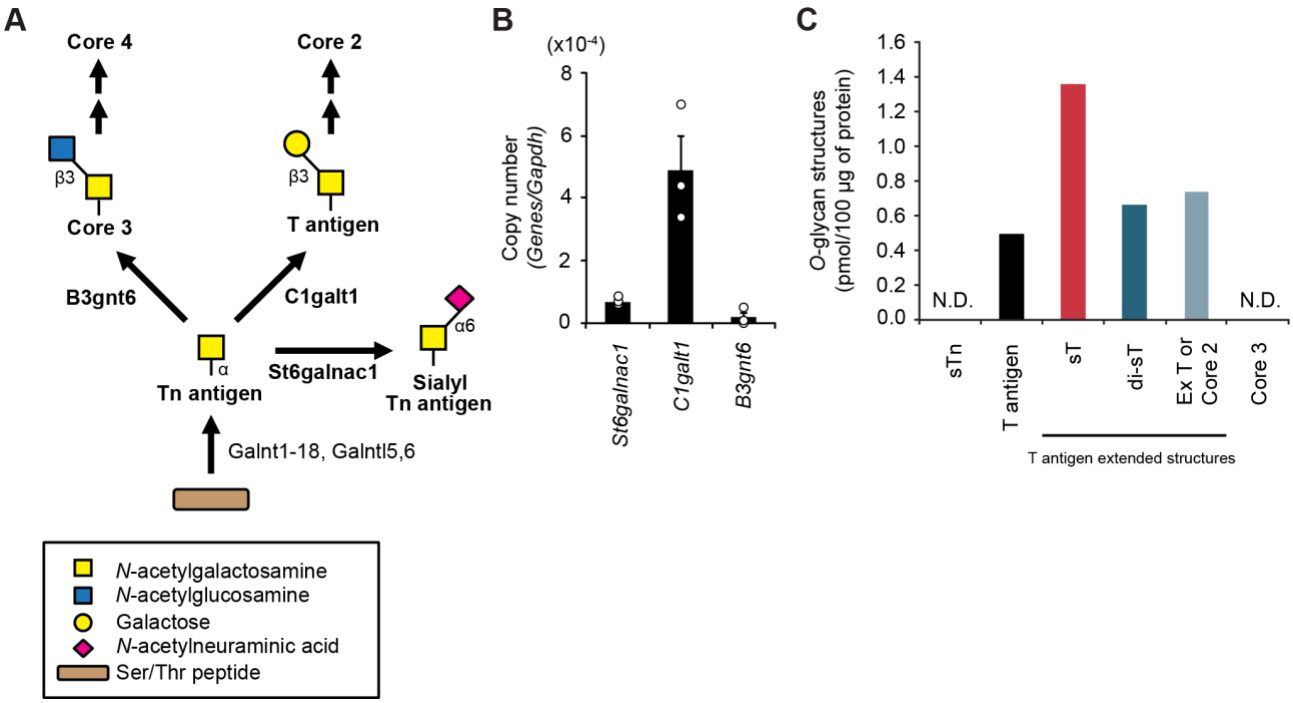


Fig. 23. C1galt1 elongation pathways is the most abundant pathway in ESCs.

(A): Schematic diagram of mucin-type *O*-glycosylation pathway initial step. **(B):** *Genes* absolute quantification by real-time PCR. The genes copy number is normalized to that of *Gapdh*. The values are shown as means ± s.e.m. of three independent experiments. **(C):** Absolute amount of *O*-glycan structures detected by MS. sTn antigen (sTn), Core 3 structure (Core 3), T antigen, and C1galt1-mediated elongation pathway modifications; sT antigen (sT), disialyl T antigen (di-sT), and extended T antigen or Core 2 structure (Ex T or Core 2). The data were obtained from a single technical and biological replicate. N.D. not detected.

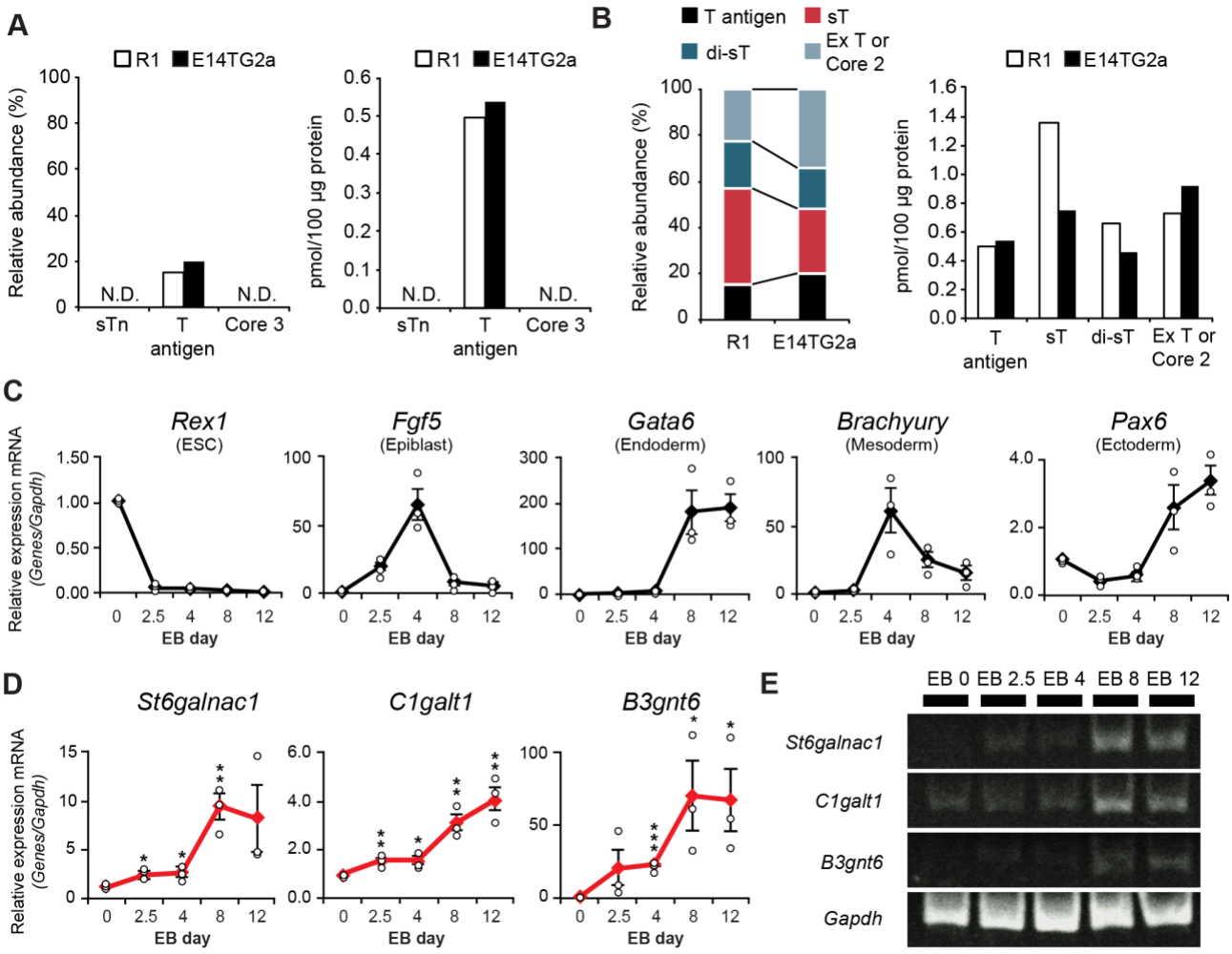


Fig. 24. C1galt1 elongation pathways is the most abundant pathway during early differentiation.

(A): Relative and absolute amount of sTn antigen (sTn), T antigen, and Core 3 structure (Core 3) by MS in R1 (white box) and E14TG2a cell lines (black box). **(B):** Relative and absolute amount of *O*-glycan structures detected by MS in R1 and E14TG2a cell lines. T antigen (black box), and C1galt1-mediated elongation pathway modifications; sT antigen (sT, red box), disialyl T antigen (di-sT, dark blue box), and extended T antigen or Core 2 structure (Ex T or Core 2, light blue box). The data were obtained from a single technical and biological replicate. N.D. not detected. **(C), (D):** Transcriptional analysis of *genes* in embryoid bodies (EB) at day 0 (ESC), 2.5, 4, 8, and 12 normalized against *Gapdh*. **(E):** *St6galnac1, C1galt1, B3gnt6,* and *Gapdh* mRNA quantification by PCR. mRNA was amplified by PCR using 25 cycles to avoid saturation and separated on polyacrylamide gel. The values are shown as means ± s.e.m. from three independent experiments**.**

3-5. C1galt1 knockdown results in ESCs differentiation via canonical Wnt signaling

To clarify the role of the C1galt1 elongation pathway, we performed a RNAi knockdown (KD) of *C1galt1*. At 4 days after transfection, *C1galt1* KD cells showed reduced *C1galt1* expression (Fig. 25A) followed by a decrease of T antigen at the surface and internally as shown by PNA staining (Fig. 25B-F). *C1galt1* KD resulted in a drastic increase in Tn antigen and a mild decrease in sTn antigen at the cell surface (Fig. 26A-D). *C1galt1* KD cells exhibited a flat shape, which is typical of differentiated cells, compared to the dome-shape morphology of undifferentiated cells (Fig. 27A). Next, we analyzed the expression of the core pluripotency markers Oct3/4 and Sox2 and of SSEA-1, a pluripotent cell surface marker (Nakai-Futatsugi and Niwa, 2013). The expression of *Oct3/4* and *Sox2* was significantly decreased in *C1galt1* KD cells compared to the control; a similar decrease at the protein level was also observed (Fig. 27B-D). The SSEA-1^{+} cell population was reduced in *C1galt1* KD cells compared to the control (Fig. 27E). To assess the ability of *C1galt1* KD cells to self-renew we performed an alkaline phosphatase (ALP) assay 4 days post transfection. The number of ALP positive colonies was markedly decreased in *C1galt1* KD cells, demonstrating that the self-renewing ability of *C1galt1 KD* cells was compromised (Fig. 27F). ESC pluripotency and self-renewal loss following *C1galt1* KD was further confirmed by using stable *C1galt1* KD ESCs (Fig. 28A-D). *C1galt1* KD resulted in an upregulation of the trophoblast markers *Cdx2* and *Gata3*, suggesting that *C1galt1* KD induces ESC transdifferentiation toward the trophectoderm (Fig. 27G). Consistently, an EB assay showed an enhanced differentiation potential in *C1galt1* KD ESCs (Fig. 28E,F). These findings demonstrate that *C1galt1* KD cells spontaneously exit from pluripotency, even in the presence of LIF.

Previous studies reported that mucins, which are heavily *O*-glycosylated proteins, interact with β-catenin, the key mediator of the canonical Wnt signaling, in many tumors (Pai et al., 2016). Therefore, we hypothesized that *C1galt1* KD might alter canonical Wnt signaling in ESCs. Activation of Wnt signaling results in inhibition of β-catenin phosphorylation, thereby preventing its degradation. Stabilized β-catenin accumulates in the cytoplasm and translocates to the nucleus where, together with Tcf/Lef, it triggers Wnt specific target genes transcription (Nusse and Clevers, 2017). *C1galt1* KD cells exhibited a reduction in phospho-β-catenin (p-β-catenin) and an increase in total β-catenin (Fig. 29A,B). *β-catenin* expression was reduced, indicating that *C1galt1* KD resulted in increased β-catenin protein stability (Fig. 29C). A luciferase assay using a β-catenin/Tcf responsive element (TOP) showed a marked increase of Wnt signaling (Fig. 29D), followed by an increased

expression of the Wnt specific target gene *Axin2* (Fig. 29E). Moreover, immunostaining of β-catenin in permeabilized *C1galt1* KD cells showed an increased localization of β-catenin in the nucleus compared to the control (Fig. 29F,G), demonstrating that *C1galt1* KD promotes canonical Wnt signaling activation in ESCs. Mucins have also been reported to activate PI3K/Akt and Fgf signaling in breast cancer cells (Woo et al, 2012; Hiraki et al, 2016). However, *C1galt1* KD did not result in activation of PI3K/Akt and Fgf signaling pathways in ESCs (Fig. 30A-D).

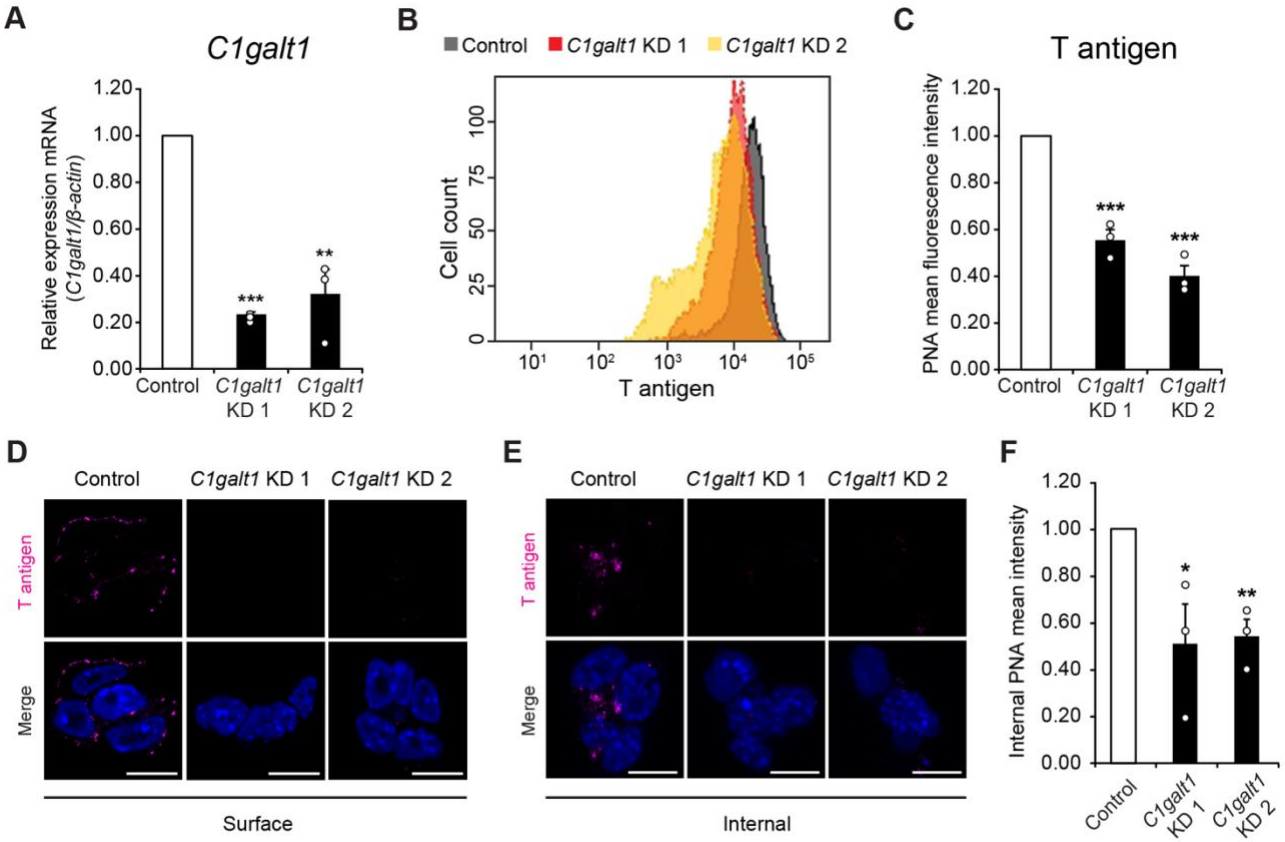


Fig. 25. *C1galt1* KD results in T antigen reduction.

(A): Transcriptional analysis of *C1galt1* KD cells normalized against the amount of *β-actin*. **(B):** FACS staining by PNA of *C1galt1* KD. **(C):** Fluorescence mean intensity of image (**B**). The fold change is relative to that of control cells. **(D):** Representative image of cell surfaces of *C1galt1* KD cells after immunostaining using PNA. Nuclei were stained with Hoechst. Scale bar, 10 μm. **(E):** Representative image of a maximum intensity projection of internal molecules in *C1galt1* KD cells after immunostaining using PNA. Nuclei were stained with Hoechst. Scale bar, 10 μm. **(F):** PNA mean intensity in image **(E)**. The mean intensity is shown as fold change relative to that of the control. The values are shown as means ± s.e.m. of three independent experiments.

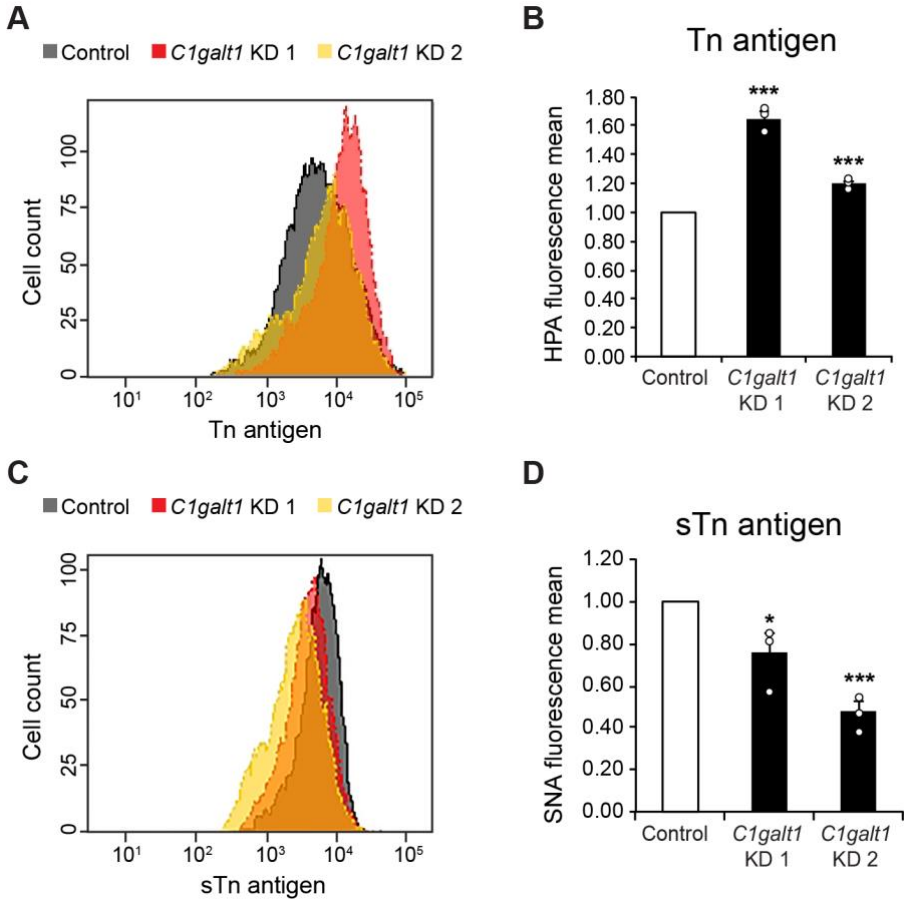


Fig. 26. *C1galt1* KD results in Tn antigen increase.

(A): FACS staining by HPA in *C1galt1* KD cells. **(B):** Fluorescence mean intensity of image (**A**). **(C):** FACS staining by SNA in *C1galt1* KD cells. **(D):** Fluorescence mean intensity of image (**C**). The fold change is relative to that of control cells. The values are shown as means ± s.e.m. of three independent experiments.

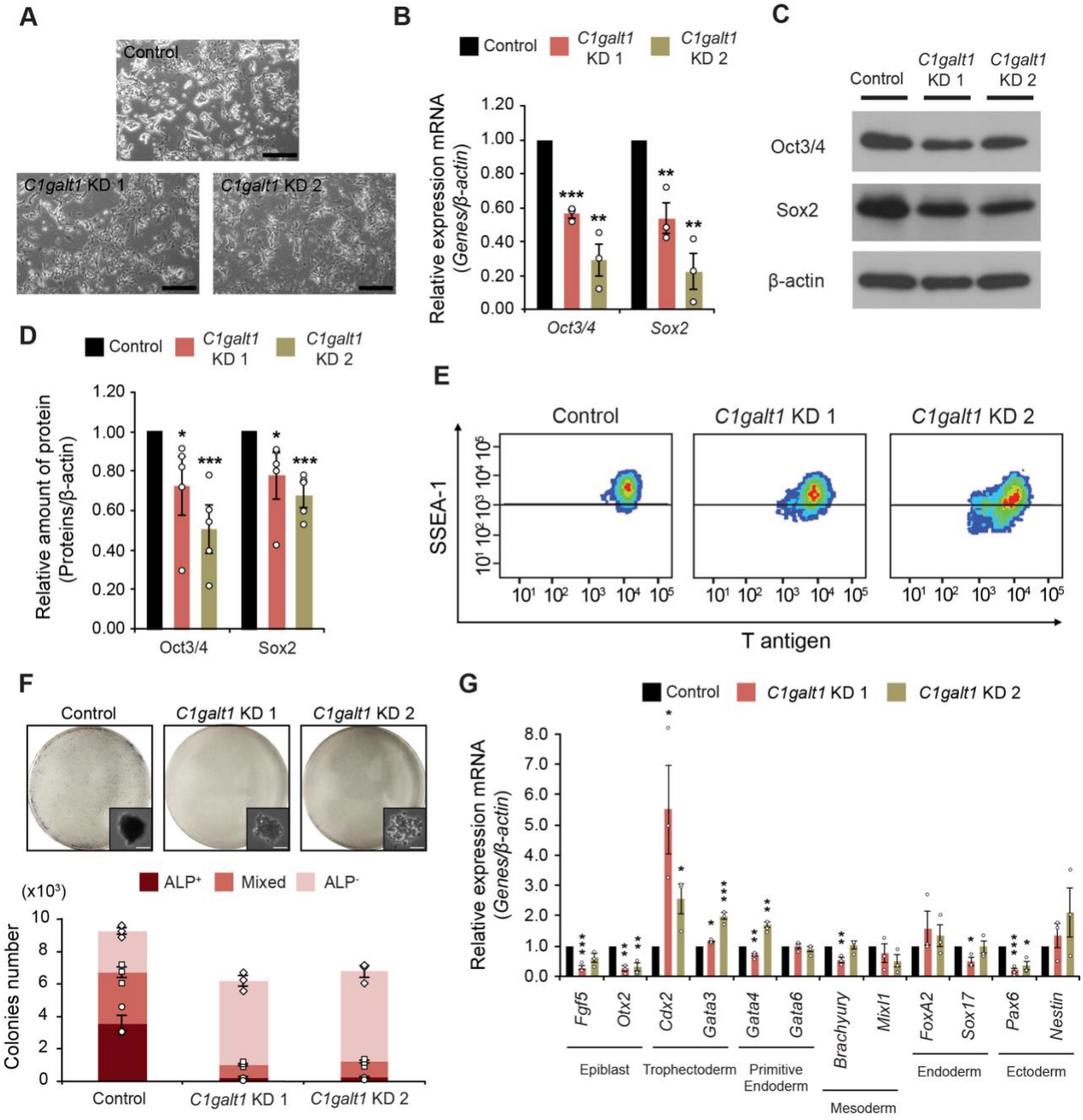


Fig. 27. *C1galt1* transient KD results in exit from the pluripotent state.

(A): *C1galt1* KD ESCs morphology. Scale bar, 200 μm. **(B):** Transcriptional analysis of the pluripotent markers in *C1galt1* KD cells normalized against *β-actin*. **(C):** Representative image of a western blot analysis of pluripotent markers in *C1galt1* KD cells. **(D):** Oct3/4 and Sox2 in image **(C)** were normalized to β-actin. **(E):** Density plot by FACS of *C1galt1* KD cells stained with anti-SSEA-1 Ab and PNA. The dark line separates the SSEA-1^{+} population and the SSEA-1^{-} population. **(F):** ALP assay of *C1galt1* KD cells. Scale bar, 25 μm. **(G):** Transcriptional analysis of differentiation markers in *C1galt1* KD cells normalized against *β-actin* and shown as a fold change relative to control cells. The values are shown as means ± s.e.m. from three or five (**(C)** and **(D)**) independent experiments**.**

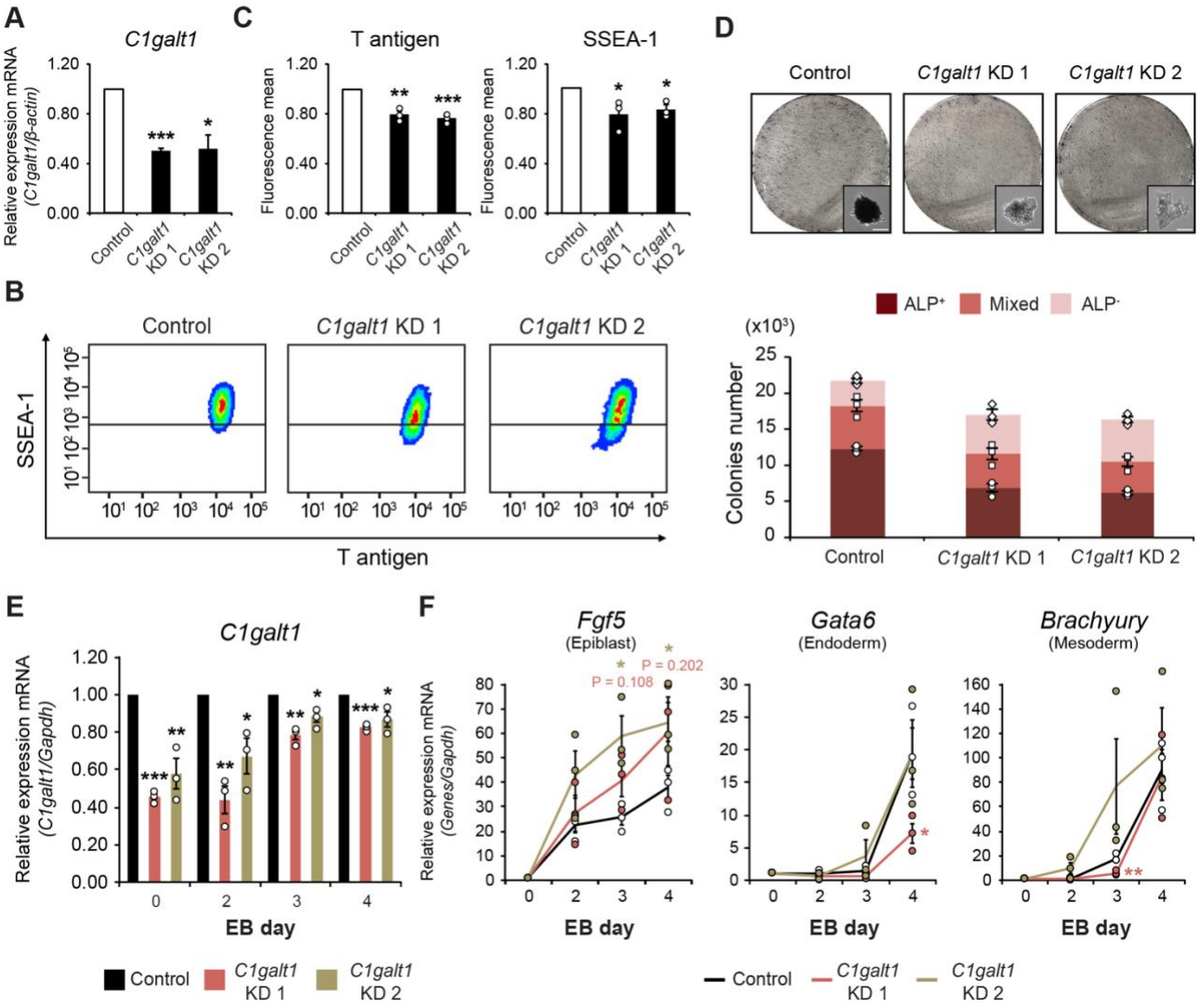


Fig. 28. *C1galt1* stable KD results in exit from the pluripotent state.

(A): Transcriptional analysis of the pluripotent markers in *C1galt1* KD cells normalized against *β-actin*. **(B):** Density plot by FACS of *C1galt1* KD cells stained with anti-SSEA-1 Ab and PNA. The dark line separates the SSEA-1^{+} population and the SSEA-1^{-} population. **(C):** PNA and SSEA-1 fluorescence mean intensity of image (**B**). **(D):** ALP assay of stable *C1galt1* KD cells. Scale bar, 25 μm. **(E), (F):** Transcriptional analysis of *C1galt1* in embryoid bodies (EB) from transient *C1galt1* KD ESCs plated 2 days post transfection and collected at day 0 (ESC), 2, 3, and 4 normalized to *Gapdh*. The values are shown as means ± s.e.m. of three independent experiments.

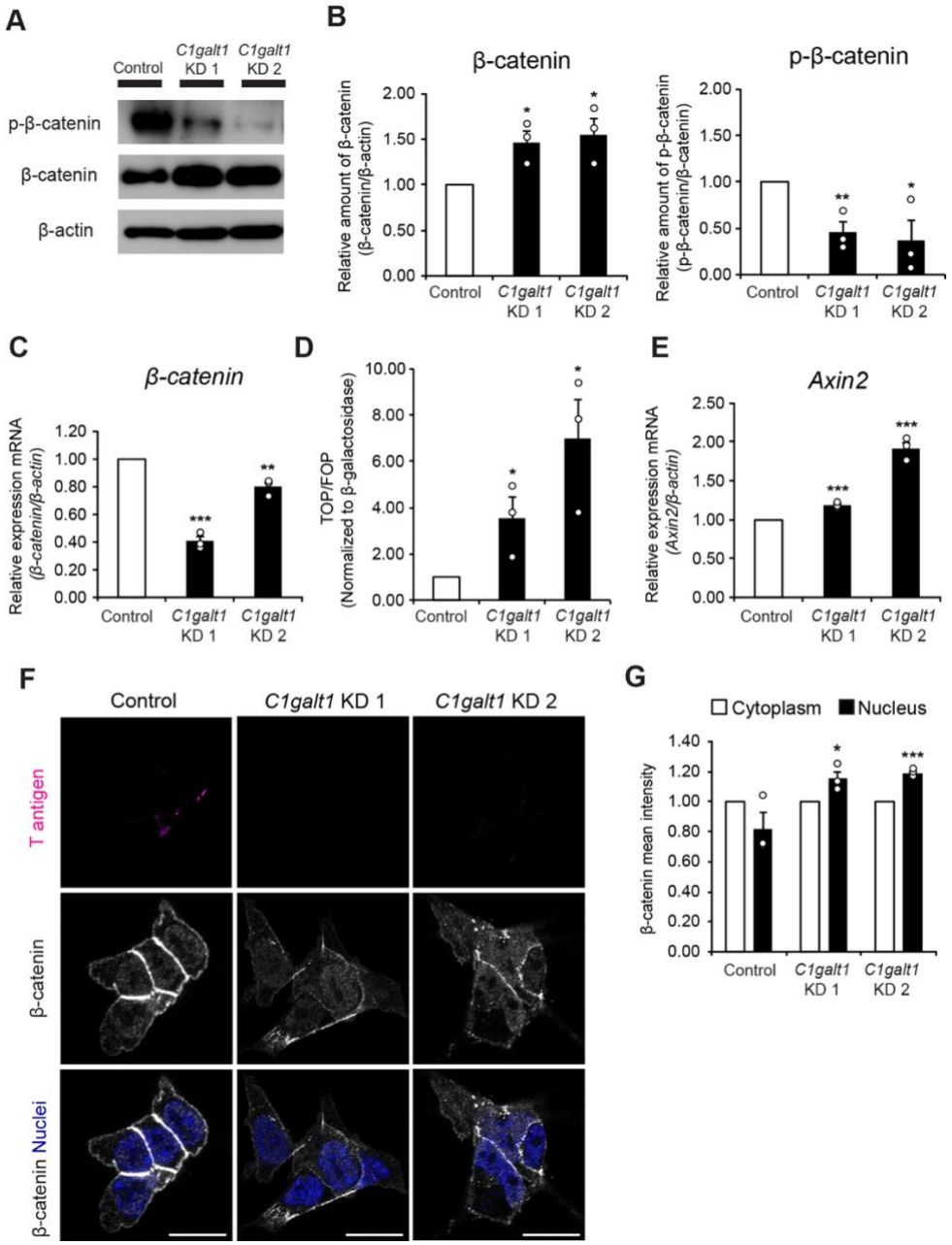


Fig. 29. *C1galt1* KD results in Wnt signaling activation.

(A): Representative image of western blot analysis using p-β-catenin Ab and β-catenin Ab in *C1galt1* KD cells. **(B):** β-catenin and p-β-catenin western blot normalized against β-actin and β-catenin, respectively. **(C):** Transcriptional analysis of *β-catenin* in *C1galt1* KD cells normalized against *β-actin*. **(D):** β-catenin reporter assay of *C1galt1* KD cells. The relative light unit amount is shown as a ratio of TOPFLASH/FOPFLASH normalized against β-galactosidase. **(E):** *Axin2* expression in *C1galt1* KD cells normalized against *β-actin*. **(F):** Representative image of intracellular molecules using PNA and an anti-β-catenin Ab in *C1galt1* KD cells. Nuclei were stained with Hoechst. Scale bar, 10 μm. **(G):** Cytoplasmic and nuclear β-catenin mean intensity of *C1galt1* KD cells. The values are shown as means ± s.e.m. of three independent experiments.

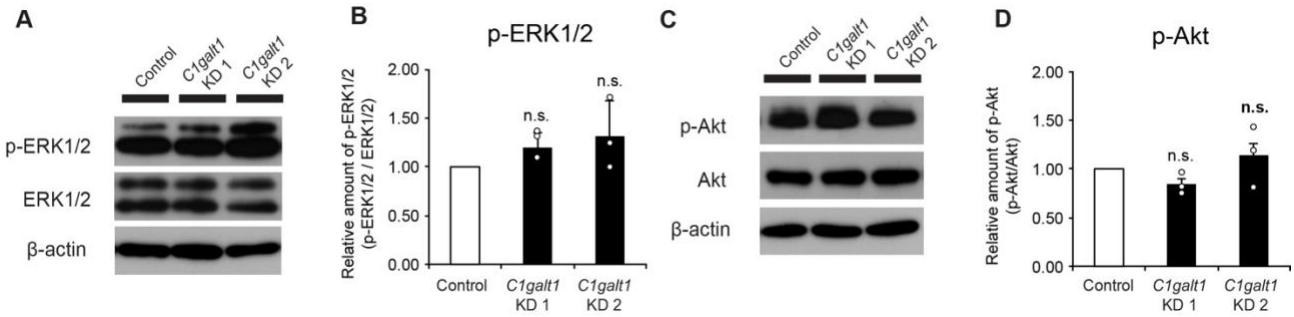


Fig. 30. *C1galt1* KD do not affect FGF and Akt signaling.

(A): Representative image of western blot analysis using p-Erk1/2 Ab and Erk1/2 Ab in *C1galt1* KD cells. **(B):** p-Erk1/2 western blot band normalized against Erk1/2. **(C):** Representative image of western blot analysis using p-Akt Ab and Akt Ab in *C1galt1* KD cells. **(D):** p-Akt western blot normalized against Akt. The values are shown as means ± s.e.m. of three independent experiments; n.s. not significant.

3-6. Frizzled-5 endocytosis is mediated by Galectin-3 binding to T antigen

Canonical Wnt signaling is initiated by the binding of the Wnt ligand to its receptor Fzd, which in turn dimerizes with the co-receptor LRP5/6 and results in downstream signaling (Nusse and Clevers, 2017). Among the Fzd receptor family involved in the canonical Wnt signaling *Fzd5* is the most highly expressed in ESCs and plays an essential role during development (Hao et al., 2006). A pull-down assay using biotinylated PNA showed that Fzd5 was successfully precipitated, demonstrating that Fzd5 carries T antigen (Fig. 31A). Immunoprecipitation on an ESC lysate using an anti-Fzd5 antibody (Ab) followed by a lectin blot analysis using PNA confirmed that Fzd5 was modified by T antigen (Fig. 31B). Moreover, the amount of Fzd5 protein precipitated in *C1galt1* KD cells decreased compared to the control, further confirming that Fzd5 is *O*-glycosylated (Fig. 31C,D). Notably, T antigen was not detected by lectin blot analysis for LRP5 and LRP6 immunoprecipitates (Fig. 32). Collectively, these data demonstrate for the first time that Wnt receptor Fzd5 carries T antigen.

Mucin-type *O*-glycosylation has multiple functions, such as in protein-protein interaction, and in trafficking and turnover of cell surface proteins (Razawi et al., 2013; Karabasheva et al., 2014). Immunostaining using an anti-Fzd5 Ab in permeabilized *C1galt1* KD cells showed a striking reduction of internalized Fzd5 (Fig. 33A-C). In contrast, Fzd5 staining at the cell surface increased in *C1galt1* KD cells (Fig. 33D). A surface biotinylation assay demonstrate that Fzd5 was strikingly increased on the surface of *C1galt1* KD cells (Fig. 33E). Furthermore, ultrastructural localization of Fzd5 by immunoelectron microscopy

showed that colloidal gold labeling was scarcely present on the plasma membrane of control cells compared to *C1galt1* KD cells (Fig. 33F), further confirming that Fzd5 was retained at the cell membrane after *C1galt1* KD. Upon ESCs treatment with the *O*-glycosylation inhibitor benzyl 2-acetoamido-2-deoxy-α-D-galactopyranoside (GalNAc-Bn), a similar effect to that in *C1galt1* KD cells was observed (Fig. 34A-D), confirming that T antigen on Fzd5 is involved in its internalization. The ultrastructural localization of Fzd5 was investigated using two different Abs against Fzd5. Colloidal gold labeling was observed inside and outside the endosomes of control cells using anti-Fzd5 Abs against the N-terminal and C-terminal regions, respectively (Fig. 33G). In addition, immunostaining showed that the early endosome marker Rab5 colocalized with Fzd5 in ESCs (Fig. 33H), demonstrating that Fzd5 is cleared from the plasma membrane by endocytosis.

Galectins (Lgals) belong to a family of proteins that bind to β-galactose (Johannes et al., 2018). Previous studies reported that Lgals, in particular Lgals3, promote endocytosis (Merlin et al., 2011; Gao et al., 2012; Lepur et al., 2012; Lakshminarayan et al., 2014). Furthermore, frontal affinity chromatography analysis demonstrated that T antigen is a ligand for Lgals9, followed by Lgals2, Lgals3, and Lgals4 (Iwaki and Hirabayashi, 2018). Thus, we hypothesized that Lgals are involved in Fzd5 endocytosis. Immunostaining in permeabilized ESCs showed a marked reduction of internalized Fzd5 in ESCs treated with a high concentration of lactose, demonstrating that Lgals regulate Fzd5 endocytosis (Fig. 35A,B). RNA-seq analysis showed that *Lgals3* was the most highly expressed in ESCs (Fig. 35C). Consistently, addition of Lgals3 resulted in an increase of internalized Fzd5, confirming that Lgals3 is involved in Fzd5 endocytosis in ESCs (Fig. 35D,E). In conclusion, our findings demonstrate that reduction of T antigen results in disruption of Lgals3-mediated endocytosis of the Wnt receptor Fzd5. As a result, Fzd5 is retained at the plasma membrane, thereby prolonging the activation of Wnt signaling and culminating in ESCs exit from pluripotency (Fig. 36).

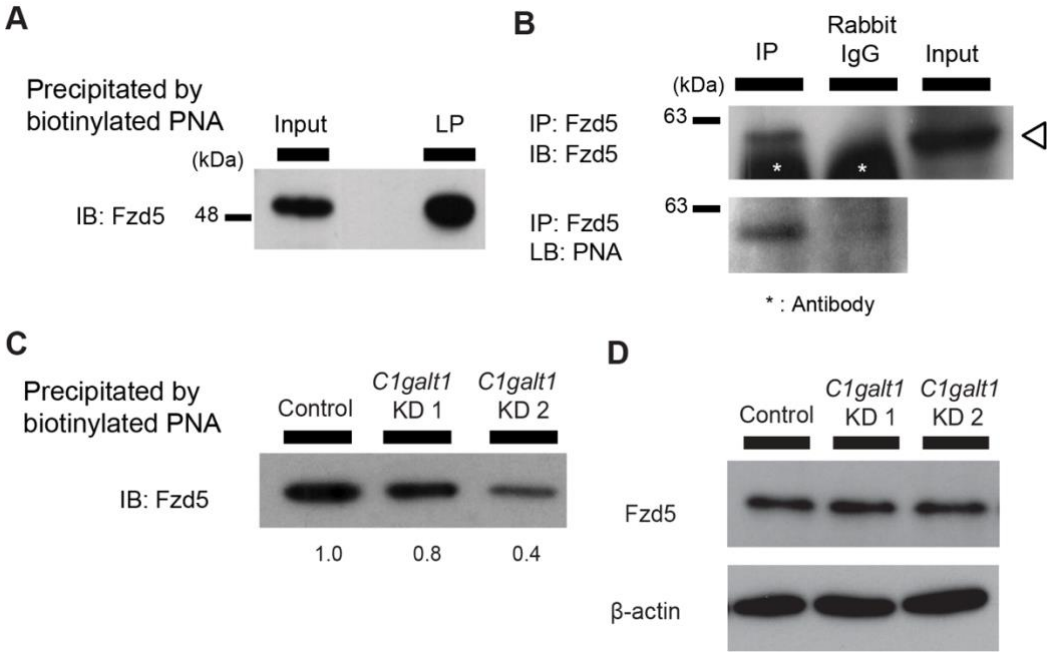


Fig. 31. Wnt receptor Frizzled-5 carries T antigen.

(A): Representative image of a western blot (IB) using an anti-Fzd5 Ab on the lectin precipitated fraction (LP) precipitated with biotinylated PNA. The input represents the total ESC lysate. Similar results were obtained in four independent experiments. **(B):** Representative image of a western blot (IB) using an anti-Fzd5 Ab and lectin blot (LB) by PNA-HRP on the immunoprecipitated fraction (IP) precipitated with an antibody against Fzd5. The input represents the total ESC lysate. The arrowhead indicates Fzd5. Similar results were obtained in three independent experiments. **(C):** Representative image of a western blot (IB) using an anti-Fzd5 Ab on the lectin precipitated fraction precipitated with biotinylated PNA in *C1galt1* KD cells. Similar results were obtained in three independent experiments. **(D):** Representative image of western blot using an anti-Fzd5 Ab in *C1galt1* KD cells. Similar results were obtained in three independent experiments.

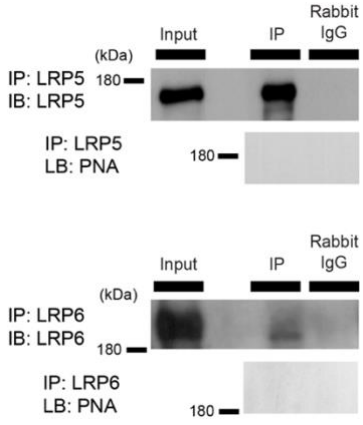


Fig. 32. LRP5/6 are not mucin-type *O*-glycosylated.

Representative images of western blot (IB) using an anti-LRP5 Ab (upper panel), or an anti-LRP6 Ab (lower panel), and lectin blot (LB) using PNA, on the immunoprecipitated fraction (IP) precipitated with an antibody against LRP5 (upper panel), or LRP6 (lower panel). The input represents the total ESC lysate. Similar results were obtained from three independent experiments.

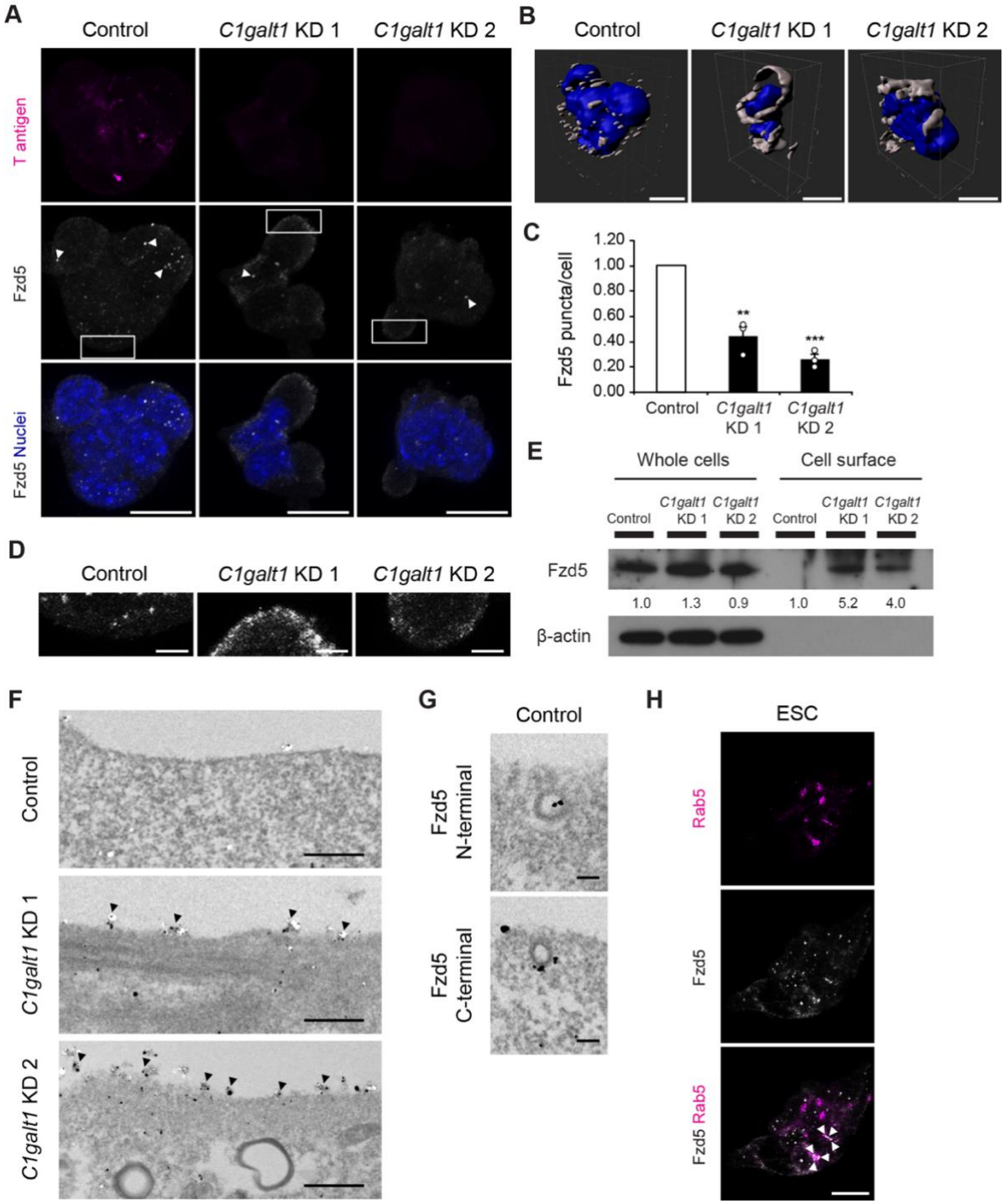


Fig. 33. T antigen regulates Frizzled-5 endocytosis.

(A): Representative image of a maximum intensity projection of intracellular molecules using PNA and an anti-Fzd5 Ab in *C1galt1* KD cells. Nuclei were stained with Hoechst. Arrowheads indicate Fzd5 puncta staining. Scale bar, 10 µm. **(B):** 3D reconstruction of images in **(A)** using Imaris version 9.3.1. Scale bar, 10 µm. **(C):** Quantification of Fzd5 puncta staining normalized against the number of nuclei. **(D):** Magnification of areas highlighted in image **(A)**. Scale bar, 2.5 µm. **(E):** Fzd5 surface protein in *C1galt1* KD cells using a surface biotinylation assay. **(F):** Representative transmission electron micrographs

from *C1galt1* KD cells. Arrowheads indicate Fzd5 staining at the plasma membrane. Scale bar, 500 nm. The micrographs were obtained from a single biological replicate. **(G):** Representative transmission electron micrographs from control cells showing the ultrastructural localization of Fzd5 in the endosome. Colloidal gold labeling is inside the endosome using an anti-Fzd5 Ab against the N-terminal region (upper panel), and outside the endosome using an anti-Fzd5 Ab against the C-terminal region (lower panel). Scale bar, 10 nm. The micrographs were obtained from a single biological replicate. **(H):** Representative image of a maximum intensity projection of intracellular molecules using an anti-Fzd5 Ab and an anti-Rab5 Ab in ESCs. Arrowheads indicate colocalization of Fzd5 puncta and Rab5. Scale bar, 10 μm. Values are shown as means ± s.e.m. of three independent experiments.

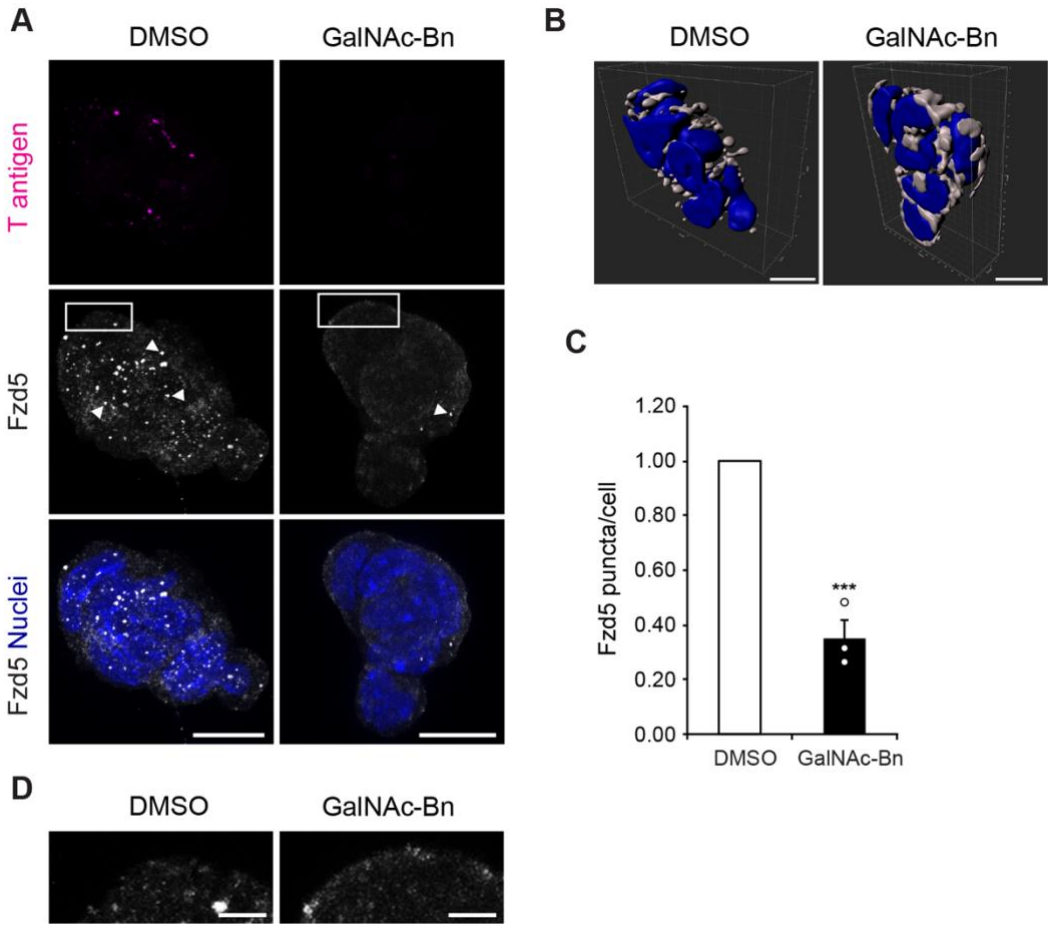


Fig. 34. Mucin-type *O*-glycosylation inhibitor reduces Frizzled-5 internalization. **(A):** Representative image of a maximum intensity projection of intracellular molecules using PNA and an anti-Fzd5 Ab in the presence or absence of 2 mM GalNAc-Bn for 48 hr. Nuclei were stained with Hoechst. Arrowheads indicate Fzd5 puncta staining. Scale bar, 10 μm. **(B):** 3D reconstruction of images in **(A)** using Imaris version 9.3.1. Scale bar, 10

µm. **(C):** Quantification of Fzd5 puncta staining normalized against the number of nuclei and shown as a fold change relative to control. **(D):** Magnification of highlighted areas in image **(A)**. Scale bar, 2.5 µm. The values are shown as means ± s.e.m. of three independent experiments.

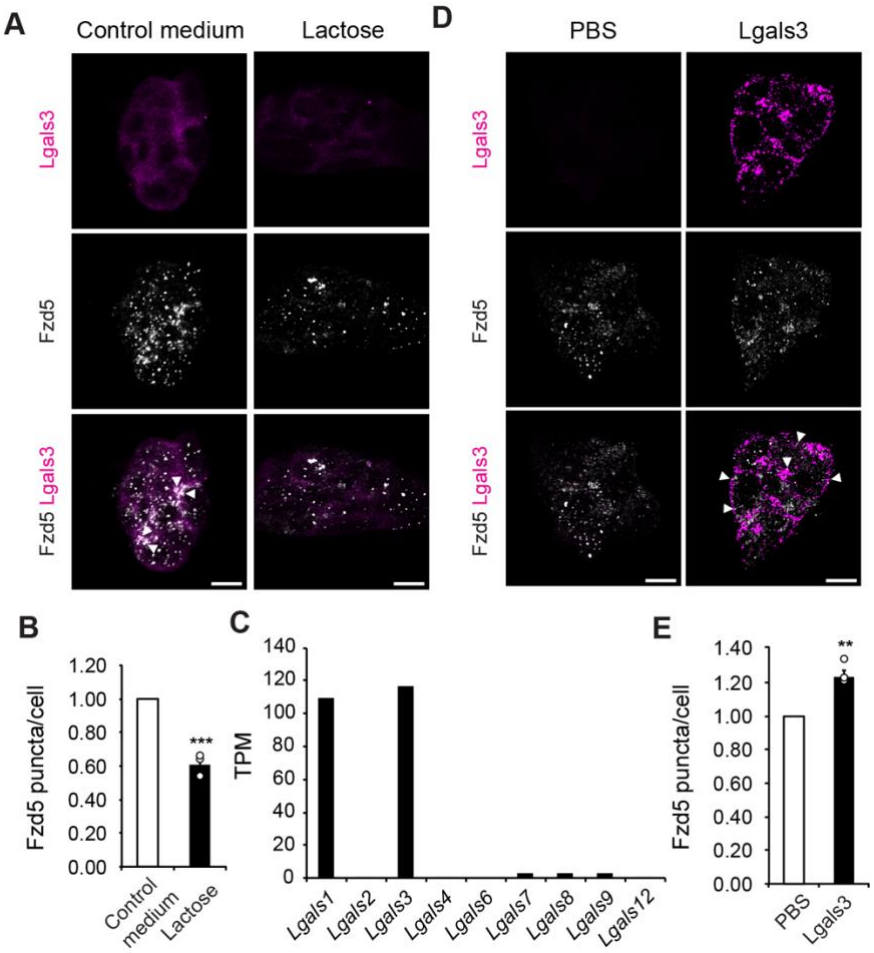


Fig. 35. Frizzled-5 endocytosis is mediated by galectin-3 binding to T antigen. **(A):** Representative image of a maximum intensity projection of intracellular molecules using an anti-Fzd5 Ab and an anti-Lgals3 Ab in ESCs treated with 50 mM lactose monohydrate. Arrowheads indicate colocalization of Fzd5 puncta and Lgals3. Scale bar, 10 µm. **(B):** Quantification of Fzd5 puncta staining of image **(A)** normalized against the number of nuclei. **(C):** Galectins expression in ESCs analyzed by RNA-seq. RNA-seq data obtained from a single technical and biological replicate. **(D):** Representative image of a maximum intensity projection of intracellular molecules using an anti-Fzd5 Ab and an anti-Lgals3 Ab in ESCs treated with 15 µg/mL of recombinant Lgals3. Arrowheads indicate colocalization of Fzd5 puncta and Lgals3. Scale bar, 10 µm. **(E):** Quantification of Fzd5 puncta staining of image **(D)** normalized against the number of nuclei. Values are shown as means ± s.e.m. of three independent experiments.

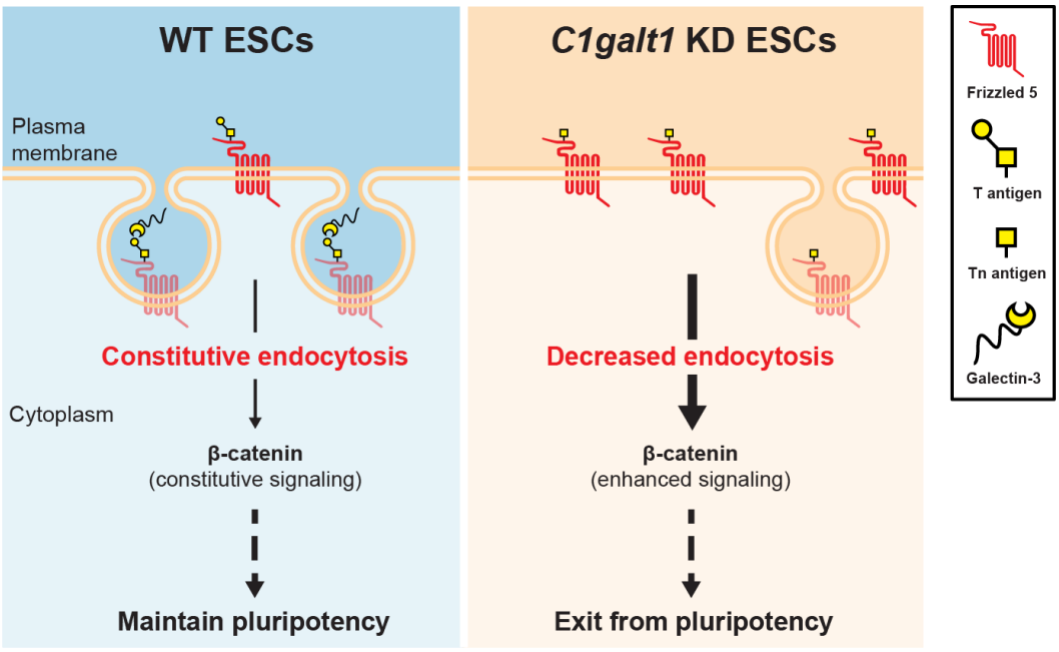


Fig. 36. Schematic representation of ESC pluripotency regulation by mucin-type *O*-glycosylation.

Mucin-type *O*-glycosylated Wnt receptor Frizzled-5 on wild-type (WT) ESCs is cleared from the plasma membrane by endocytosis via the binding of galectin-3 to T antigen, resulting in β-catenin constitutive signaling which promotes the maintenance of ESC pluripotency. By contrast, *C1galt1* KD ESCs show disrupted galectin-3 mediated endocytosis of Frizzled-5 and its retention at the plasma membrane. As a result, Wnt signaling, mediated by β-catenin, is enhanced and culminates in the loss of ESC pluripotency.

4. Discussion

The pivotal role of glycosylation during development and in determining stem cell identity across different species is becoming increasingly clear (Nishihara, 2018). Previous studies reported the glycomic profiles of mouse and conventional human ESCs, hiPSCs, tumors and late differentiating cells (Nairn et al., 2012; Fujitani et al., 2013; Homan et al., 2019). In the present study, a comparative and comprehensive analysis to investigate the glycosylation dynamics during the pluripotency state transition from ESCs to EpiLCs was performed. As a result, we demonstrate that glycosylation undergoes dramatic alterations from early stages of development, and we identified PRC2 as a key component involved in these changes (Fig. 22).

PSCs in the primed state exhibit substantial differences compared to their naïve counterpart, such as a flat morphology, glycolytic metabolism, slow proliferation, and closer chromatin (Weinberger et al, 2016). During the pluripotency state transition, RAS mediates the epithelial-to-mesenchymal transition (EMT), characterized by the switch from epithelial cadherin (Ecad) to neural cadherin (Ncad) (Altshuler et al, 2018); a process similar to cancer progression and tightly regulated by glycosylation (Pinho et al, 2011; Pinho and Reis, 2015). In fact, it has been reported that *N*-glycosylation 1-6 branching structure on Ecad promotes its clearance from the cell surface and inhibits Ncad-mediated cell-cell adhesion. Furthermore, core fucosylation weakens Ecad cell-cell adhesion in lung cancer (Pinho et al, 2011). Moreover, an increased amount of sialylated glycans (Pinho and Reis, 2015) and reduced GSL (Guan et al, 2009) were documented to correlate with EMT progression. Similarly, we detected enhanced 2,6 branching, core fucosylation (Fig. 4B,E), and total sialylated glycans (Fig. 17D), followed by a sharp reduction in GSL (Fig. 14B), indicating a common EMT regulation by glycosylation across different biological contexts. RAS activation was also reported to be linked to a metabolic shift from oxidative phosphorylation to glycolysis, and subsequent closer chromatin during the naïve-to-primed transition (Altshuler et al, 2018). Notably, RAS is downstream of FGF signaling, which requires *N*-sulfation of HS. Moreover, Myc amount, which controls the ESCs' proliferative program thus regulating the proliferation speed, is inversely correlated with FGF-ERK activation (Díaz-Díaz et al, 2017), highlighting the functional importance of FGF signaling regulation by HS. Accordingly, we observed a substantial increase in *Ndst2-4* expression and *N*-sulfated HS staining (Fig. 11C,D). Our data emphasize that glycosylation dynamic changes contribute and partially drive the major phenotypical alterations occurring during

the pluripotency state transition, thus underlining the importance of functionally dissecting the role of glycosylation during developmental transitions *in vitro* and *in vivo*.

Comprehensive analysis of total glycomes allows the identification of pluripotency biomarker candidates (Fujitani et al., 2013). Indeed, our data confirmed previously established pluripotency markers, such as SSEA-1,3,4 and Forsmann antigen, which are expressed at a similar level in the naïve and primed state. Moreover, we demonstrated that SSEA-5 is specifically expressed in EpiLCs, augmenting previous studies performed in conventional hESCs and hiPSCs (Tang et al., 2011). In addition, we identified a wide range of novel structures across all glycosylation classes that are specifically enhanced in the primed state but not detected or detected at very low levels in the naïve state, providing additional markers to distinguish between the different pluripotent states: bisecting GlcNAc and 2,6 branched tri-/tetra-antennary complex *N*-glycans, extended or branched capped mucin-type *O*-glycan structure, *N*-unsubstituted GlcN, *N*-sulfated GlcN, and 3-*O*-sulfated HS, CS-A, C, D and E units, and PSA.

The recent establishment of culture conditions to reprogram mouse ESCs into expanded potential stem cells (EPSCs) or 2-cell-embryo-like cells (2CLCs) allowed the *in vitro* investigation of the totipotent state. Despite the most suitable *in vitro* system to model the totipotent state is still under debate (Posfai et al., 2020 preprint), these models allowed the identification of some molecular features of totipotent cells (Yang et al., 2017; Hendrickson et al., 2017; Posfai et al., 2020 preprint), providing an invaluable tool to characterize the totipotent state. It will be of great interest to examine the glycosylation dynamics during the reprogramming process from ESCs to EPSCs and 2CLCs to establish novel totipotency biomarkers and obtain mechanistic insights into the transition from the totipotent to the pluripotent state.

PRC2 regulates early embryonic specification processes by repressing crucial developmental genes (Boyer et al., 2006). Indeed, deficiency in PRC2 core components *Eed*, *Suz12*, or *Ezh2* results in embryo lethality around E7.5-E8.5 due to gastrulation defects (Deevy and Bracken, 2019). Here, we showed for the first time that PRC2 contributes to overall glycosylation alterations occurring during the naïve to primed transition. Recently, ISY1 has been reported to modulate the biogenesis of a large subset of crucial miRNAs during the transition from ESCs to EpiLCs (Du et al., 2018). Moreover, PRC2 represses around 512 developmental regulators in ESCs (Boyer et al., 2006). Thus, the glycosylation changes we observed in EpiLCs are likely to be the result of the synergic action of PRC2 on glycosylation-related genes expression together with other

component(s) involved both directly and indirectly in the glycosylation pathway and PRC2-independent pathway(s). In addition, PRC1 and PRC2 activity is directly modulated by *O*-GlcNAc (Decourcelle et al., 2019), suggesting a metabolically regulated network controlling the glycomic profile.

Glycosylation is a developmentally coordinated post-translational modification (Haltiwanger and Lowe, 2004; Nishihara, 2018). Previous studies have identified glycosylation class-specific key regulators. For example, hepatocyte nuclear factor 1α (HNF1α) was shown to control *N*-glycan fucosylation of human plasma proteins (Lauc et al., 2010). More recently, autism susceptibility candidate 2 (Auts2) was found to drive GSL metabolic switch during neural differentiation from mouse ESCs (Russo et al., 2018). Nonetheless, regulation of overall glycosylation dynamics has remained unknown. Here, PRC2 was shown to contribute to the glycosylation changes occurring during naïve to primed transition, demonstrating that the glycome complex alterations occurring during developmental transitions are orchestrated by a defined regulatory network. Consequently, it will be important to characterize the glycomic dynamics in a variety of developmental stages and cell types in order to identify transition-specific glycosylation regulatory components.

It is becoming clear that glycosylation acts as a pivotal regulatory switch of pluripotency in a range of cell types in different organisms (Nishihara, 2018). In the present study, the function of mucin-type *O*-glycosylation in the pluripotency network and the relationship between *O*-glycosylation and signaling in ESCs was characterized. In brief, *C1galt1* knockdown results in a decrease of T antigen on the Wnt receptor Fzd5, reducing the level of Lgals3-mediated Fzd5 endocytosis. The retention of Fzd5 on the surface, in turn, promotes excessive canonical Wnt signaling activation via β-catenin stabilization, resulting in the exit from pluripotency (Fig. 36).

A recent study showed that mucin-type *O*-glycosylation plays a critical role in maintaining the epithelial state of trophoblast stem cells, which are derived from the first embryo lineage commitment (Raghu et al., 2019). Furthermore, *C1galt1* KD induced expression of the trophectoderm markers *Cdx2* and *Gata3* (Fig. 27G). Thus, it will be of particular interest to investigate the role of mucin-type *O*-glycosylation in the early commitment of ESCs and in early embryo development *in vivo*.

In the past decade, Wnt signaling has been shown to be a key factor in the maintenance of the undifferentiated state. The addition of Wnt3a and LIF is sufficient to support self-renewal and allows derivation of ESCs from non-permissive strains (ten Berge

et al., 2011). The induction of Wnt signaling by CHIR together with the FGF signaling inhibitor PD is commonly used to maintain the undifferentiated state in cultured ESCs (Ying et al., 2008; Ying and Smith, 2017). However, Wnt signaling outcome depends on its interactions with other signaling pathways. CHIR alone induces ESC differentiation (Ying et al., 2008). In addition, other groups have reported that β-catenin promotes the expression of genes associated with both pluripotency and differentiation (Kelly et al., 2011; Price et al., 2013; Zhang et al., 2013). Therefore, Wnt signaling maintains the undifferentiated state while at the same time priming cells for differentiation. Here, Wnt signaling dysregulation led to a loss of ESC pluripotency.

Mucin-type *O*-glycosylation plays multiple roles: protection from shedding; protein-protein interactions; and, protein turnover and trafficking (Tian and Ten Hagen, 2009; Razawi et al., 2013; Karabasheva et al., 2014). In this study, T antigen on the Wnt receptor Fzd5 was found to modulate endocytosis of Fzd5 via Lgals3 (Fig. 36). Fzd endocytosis modulates Wnt signaling positively or negatively depending on the cellular context (Brunt and Scholpp, 2018). Here, a reduction in Fzd5 endocytosis resulted in Wnt signaling activation. Abnormalities in mucin-type *O*-glycosylation and Wnt signaling are hallmarks of tumorigenesis (Pinho and Reis, 2015; Zhan et al., 2017). The present study demonstrated for the first time a direct connection between *O*-glycosylation and Wnt signaling. In conclusion, these data provide a solid groundwork for further investigations in basic research and translational medicine.

5. Acknowledgments

I would like to express my deepest appreciation to Prof. Shoko Nishihara for the continuous support throughout my Ph.D. course. I am convinced that her guidance and training laid the foundation for my future career in the research field. Beside my advisor, I would also like to thank the rest of my dissertation committee: Prof. Sayaka Takase and Prof. Akira Togayachi for their insightful comments and advices. Moreover, I would like to express my deepest gratitude to the Makiguchi foundation for education International Students for the enormous financial support throughout my studies in Soka University. Finally, I would like to express my gratitude to the following research collaborators: Dr. Jun-ichi Furukawa, Ms. Ikuko Yokota, Dr. Hisatoshi Hanamatsu and Mr. Kazue Okada (Department of Advanced clinical glycobiology, Hokkaido University) for glycomic analyses by MS and HPLC, Prof. Yamamoto Kazuo (Department of Integrated Biosciences, The University of Tokyo) for providing materials, for deep insights regarding the lectins used in the FACS profiling and for RNA-seq analysis, Prof. Oki Shinya (Department of Drug Discovery Medicine, Kyoto University) for ChIP Atlas software training and supervision, Prof. Osamu Yoshie (Health and Kampo Institute), Dr. Makoto Matsuyama (Shigei Medical Research Institute) and Prof. Masanori Taira (Department of Biological Sciences, The University of Tokyo) for Hepss-1 antibody, and Prof. Yoshihiro Akimoto, Ms. Sachie Matsubara, and Ms. Hayakawa (Department of Anatomy, Kyorin University School of Medicine) for immune TEM analysis. Also, I express my gratitude to the following researchers in our laboratory: Dr. Taichi Miura, Ms. Chika Ogura and Mr. Hayato Ota for contribution to transcriptional analysis of ESCs and EpiLCs and the constructive discussion and advice. Lastly, my greatest appreciation goes to my parents, my family and all the people who constantly encouraged me and have been a source of inspiration during this beautiful journey at Soka University of Japan.

6. References

Apweiler, R., Hermjakob, H., Sharon, N. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* **1473**, 4-8.

Altshuler, A., Verbuk, M., Bhattacharya, S., Abramovich, I., Haklai, R., Hanna, J. H., Kloog, Y., Gottlieb, E., Shalom-Feuerstein, R. (2018). RAS Regulates the Transition from Naive to Primed Pluripotent Stem Cells. *Stem Cell Reports* **10**, 1088–1101.

Bedzhov, I., Graham, S. J., Leung, C. Y., Zernicka-Goetz, M. (2014). Developmental plasticity, cell fate specification and morphogenesis in the early mouse embryo. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **369**, pii: 20130538.

Bennet, E. P., Mandel, U., Clausen, H., Gerken, T. A., Fritz, T. A., Tabak, L. A. (2012). Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology* **22,** 736-756.

Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K. et al. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349-353.

Brockhausen, I., Schachter, H., Stanley, P. (2009). O-GalNAc Glycans. In *Essentials of Glycobiology 2nd ed* (ed. A. Varki, R. D. Cummings, J. D. Esko et al.), pp. 1-21. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

Brunt, L., Scholpp, S. (2018). The function of endocytosis in Wnt signaling. *Cell. Mol. Life Sci.* **75**, 785-795.

Carter, M. G., Smagghe, B. J., Stewart, A. K., Rapley, J. A., Lynch, E., Bernier, K. J., Keating, K. W., Hatziioannou, V. M., Hartman, E. J., Bamdad, C. C. (2016). A primitive growth factor, NME7AB is sufficient to induce stable naïve state human pluripotency; reprogramming in this novel growth factor confers superior differentiation. *Stem Cells* **34**, 847-859.

Deevy, O., Bracken, A. P. (2019). PRC2 functions in development and congenital disorders. *Development* **146**, pii: dev181354.

Díaz-Díaz, C., Fernandez de Manuel, L., Jimenez-Carretero, D., Montoya, M. C., Clavería, C., Torres, M. (2017). Pluripotency Surveillance by Myc-Driven Competitive Elimination of Differentiating Cells. *Dev. Cell* **42**, 585–599.

Du, P., Pirouz, M., Choi, J., Huebner, A. J., Clement, K., Meissner, A., Hochedlinger, K., Gregory, R. I. (2018). An intermediate pluripotent state controlled by microRNAs is required for the naive-to-primed stem cell transition. *Cell stem cell* **22,** 851–864.

Evans, M. J., Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156.

Fujitani, N., Furukawa, J., Araki, K., Fujioka, T., Takegawa, Y., Piao, J., Nishioka, T., Tamura, T., Nikaido, T., Ito, M. et al. (2013). Total cellular glycomics allows characterizing cells and streamlining the discovery process for cellular biomarkers. *Proc. Natl. Acad. Sci. USA* **110**, 2105-2110.

Furukawa, J., Shinohara, Y., Kuramoto, H., Miura, Y., Shimaoka, H., Kurogochi, M., Nakano, M., Nishimura, S. (2008). Comprehensive approach to structural and functional glycomics based on chemoselective glycoblotting and sequential tag conversion. *Anal. Chem.* **80,** 1094-1101.

Gao, X., Liu, D., Fan, Y., Li, X., Xue, H., Ma, Y., Zhou, Y., Tai, G. (2012). The two endocytic pathways mediated by the carbohydrate recognition domain and regulated by the collagen-like domain of galectin-3 in vascular endothelial cells. *PLoS One* **7**, e52430.

Gökbuget, D., Blellock, R. (2019). Epigenetic control of transcriptional regulation in pluripotency and early differentiation. *Development* **146**, pii: dev164772.

Guan, F., Handa, K., Hakomori, S. I. (2009). Specific glycosphingolipids mediate epithelial-to-mesenchymal transition of human and mouse epithelial cell lines. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 7461–7466.

Haltiwanger, R. S., Lowe, J. B. (2004). Role of glycosylation in development. *Annu. Rev. Biochem.* **73**, 491-537.

Hanamatsu, H., Nishikaze, T., Miura, N., Piao, J., Okada, K., Sekiya, S., Iwamoto, S., Sakamoto, N., Tanaka, K., Furukawa, J. I. (2018). Sialic acid linkage specific derivatization of glycosphingolipid glycans by ring-opening aminolysis of lactones. *Anal. Chem.* **90**, 13193-13199.

Hao, J., Li, T. G., Qi, X., Zhao, D. F., Zhao, G. Q. (2006). WNT/beta-catenin pathway up-regulates Stat3 and converges on LIF to prevent differentiation of mouse embryonic stem cells. *Dev. Biol.* **290**, 81-91.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519-532.

He, J., Shen, L., Wan, M., Taranova, O., Wu, H., Zhang, Y. (2013). Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. *Nat. Cell. Biol*. **15**, 373-384.

He, S., Pant, D., Schiffmacher, A., Meece, A., Keefer, C. L. (2008). Lymphoid enhancer factor 1-mediated Wnt signaling promotes the initiation of trophoblast lineage differentiation in mouse embryonic stem cells. *Stem Cells* **26**, 842-849.

Hendrickson, P. G., Doráis, J. A., Grow, E. J., Whiddon, J. L., Lim, J. W., Wike, C. L., Weaver, B. D., Pflueger, C., Emery, B. R., Wilcox, A. L. et al. (2017). Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **49**, 925-934.

Hiraki, M., Suzuki, Y., Alam, M., Hinohara, K., Hasegawa, M., Jin, C., Kharbanda, S., Kufe, D. (2016). MUC1-C stabilizes MCL-1 in the oxidative stress response of triple-negative breast cancer cells to BCL-2 inhibitors. *Sci Rep.* **6**, 26643.

Hirayama, H., Seino, J., Kitajima, T., Jigami, Y., Suzuki, T. (2010). Free oligosaccharides to monitor glycoprotein endoplasmic reticulum-associated degradation in Saccharomyces cerevisiae. *J. Biol. Chem.* **285***,* 12390-12404.

Homan, K., Hanamatsu, H., Furukawa, J. I., Okada, K., Yokota, I., Onodera, T., Iwasaki, N. (2019). Alteration of the Total Cellular Glycome during Late Differentiation of Chondrocytes. *Int. J. Mol. Sci.* **20,** pii: E3546.

Iwaki, J., Hirabayashi, J. (2018). Carbohydrate-binding specificity of human galectins: an overview by frontal affinity chromatography. *Trends Glycosci. Glycotechnol.* **172**, 137-153.

Johannes, L., Jacob, R., Leffler, H. (2018). Galectins at a glance. *J Cell Sci.* **131**, jcs208884.

Karabasheva, D., Cole, N. B., Donaldson, J. G. (2014). Roles for trafficking and *O*-linked glycosylation in the turnover of model cell surface proteins. *J. Biol. Chem.* **289**, 19477-19490.

Kelly, K. F., Ng, D. Y., Jayakumaran, G., Wood, G. A., Koide, H., Doble, B. W. (2011). β-catenin enhances Oct-4 activity and reinforces pluripotency through a TCF-independent mechanism. *Cell Stem Cell* **8**, 214-227.

Kloet, S. L., Makowski, M. M., Baymaz, H. I., van Voorthuijsen, L., Karemaker, I. D., Santanach, A., Jansen, P., Di Croce, L., Vermeulen, M. (2016). The dynamic interactome and genomic targets of Polycomb complexes during stem-cell differentiation. *Nat. Struct. Mol. Biol.* **23,** 682-690.

Kudo, T., Kaneko, M., Iwasaki, H., Togayachi, A., Nishihara, S., Abe, K., Narimatsu, H. (2004). Normal embryonic and germ cell development in mice lacking alpha 1,3-fucosyltransferase IX (Fut9) which show disappearance of stage-specific embryonic antigen 1. *Mol. Cell. Biol.* **24**, 4221-4228.

Kurimoto, K., Yabuta, Y., Hayashi, K., Ohta, H., Kiyonari, H., Mitani, T., Moritoki, Y., Kohri, K., Kimura, H., Yamamoto, T., Katou, Y., Shirahige, K., Saitou, M. (2015). Quantitative Dynamics of Chromatin Remodeling during Germ Cell Specification from Mouse Embryonic Stem Cells. *Cell stem cell*. **16**, 517–532.

Lakshminarayan, R., Wunder, C., Becken, U., Howes, M. T., Benzing, C., Arumugam, S., Sales, S., Ariotti, N., Chambon, V., Lamaze, C. et al. (2014). Galectin-3 drives glycosphingolipid-dependent biogenesis of clathrin-independent carriers. *Nat. Cell. Biol.* **16**, 595-606.

Lanner, F., Rossant, J. (2010). The role of FGF/Erk signaling in pluripotent cells. *Development* **137**, 3351-3360.

Lauc, G., Essafi, A., Huffman, J. E., Hayward, C., Knežević, A., Kattla, J. J., Polašek, O., Gornik, O., Vitart, V., Abrahams, J. L. et al. (2010). Genomics meets glycomics-the first GWAS study of human N-glycome identifies HNF1α as a master regulator of plasma protein fucosylation. *PLoS Genet*. **6**, e1001256.

Lepur, A., Carlsson, M. C., Novak, R., Dumić, J., Nilsson, U. J., Leffler, H. (2012). Galectin-3 endocytosis by carbohydrate independent and dependent pathways in different macrophage like cell types. *Biochim. Biophys. Acta.* **1820**, 804-818.

Li, H., Liefke, R., Jiang, J., Kurland, J. V., Tian, W., Deng, P., Zhang, W., He, Q., Patel, D. J., Bulyk, M. L. et al. (2017). Polycomb-like proteins link the PRC2 complex to CpG islands. *Nature* **549**, 287-291.

Liang, Y. J., Kuo, H. H., Lin, C. H., Chen, Y. Y., Yang, B. C., Cheng, Y. Y., Yu, A. L., Khoo, K. H., Yu, J. (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. USA* **107**, 22564-22569.

Lindahl, U., Couchman, J., Kimata, K., Esko, J. D. (2015-2017). Proteoglycans and Sulfated Glycosaminoglycans in *Essential of Glycobiology* [Internet] *3rd edition Ch. 17* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor).

Merlin, J., Stechly, L., de Beaucé, S., Monté, D., Leteurtre, E., van Seuningen, I., Huet, G., Pigny, P. (2011). Galectin-3 regulates MUC1 and EGFR cellular distribution and EGFR downstream pathways in pancreatic cancer cells. *Oncogene* **30**, 2514-2525.

Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W., Roder, J. C. (1993). Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **90**, 8424-8428.

Nairn, A. V., Aoki, K., dela Rosa, M., Porterfield, M., Lim, J. M., Kulik, M., Pierce, J. M., Wells, L., Dalton, S., Tiemeyer, M. et al. (2012). Regulation of glycan structures in murine embryonic stem cells: combined transcript profiling of glycan-related genes and glycan structural analysis. *J. Biol. Chem.* **287**, 37835-37856.

Nakai-Futatsugi, Y., Niwa, H. (2013). Transcription factor network in embryonic stem cells: heterogeneity under the stringency. *Biol. Pharm. Bull.* **36**, 166-170.

Nichols, J., Smith, A. (2009). Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487-492.

Nishihara, S. (2018). Glycans in stem cell regulation: from *Drosophila* tissue stem cells to mammalian pluripotent stem cells. *FEBS Lett.* **592**, 3773-3790.

maintains pluripotency of mouse ES cells. *Nature* **460**, 118-122.

Niwa, H., Ogawa, K., Shimosato, D., Adachi, K. (2009). A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature* **460**, 118-122.

Nusse, R., Clevers, H. (2017). Wnt/β-catenin signaling, disease, and emerging therapeutic modalities. *Cell* **169**, 985-999.

Oki, S., Ohta, T., Shioi, G., Hatanaka, H., Ogasawara, O., Okuda, Y., Kawaji, H., Nakaki, R., Sese, J., Meno, C. (2018). ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. *EMBO Rep.* **19**, e46255.

Pai, P., Rachagani, S., Dhawan, P., Batra, S. K. (2016). Mucins and Wnt/β-catenin signaling in gastrointestinal cancers: an unholy nexus. *Carcinogenesis* **37**, 223-232.

Perino, M., van Mierlo, G., Karemaker, I. D., van Genesen, S., Vermeulen, M., Marks, H., van Heeringen, S. J., Veenstra, G. (2018). MTF2 recruits Polycomb Repressive Complex 2 by helical-shape-selective DNA binding. *Nat. Genet*. **50**, 1002-1010.

Pinho, S. S., Reis, C. A. (2015). Glycosylation in cancer: Mechanisms and clinical implications. *Nat. Rev. Cancer* **15**, 540–555.

Pinho, S. S., Seruca, R., Gärtner, F., Yamaguchi, Y., Gu, J., Taniguchi, N., Reis, C. A. (2011). Modulation of E-cadherin function and dysfunction by N-glycosylation. *Cell. Mol. Life Sci.* **68**, 1011–1020.

Posfai, E., Schell, J. P., Janiszewski, A., Rovic, I., Murray, A., Bradshaw, B., Pardon, T., Bakkali, M. E., Talon, I., De Geest, N. et al. (2020). Defining totipotency using criteria of increasing stringency. *bioRxiv* doi: https://doi.org/10.1101/2020.03.02.972893.

Price, F. D., Yin, H., Jones, A., van Ijcken, W., Grosveld, R., Rudnicki, M. A. (2013). Canonical Wnt signaling induces a primitive endoderm metastable state in mouse embryonic stem cells. *Stem Cells* **31**, 752-764.

Qi, W., Zhao, K., Gu, J., Huang, Y., Wang, Y., Zhang, H., Zhang, M., Zhang, J., Yu, Z., Li, L. et al. (2017). An allosteric PRC2 inhibitor targeting the H3K27me3 binding pocket of EED. *Nat. Chem. Biol.* **13**, 381-388.

Raghu, D., Mobley, R. J., Shendy, N. A. M., Perry, C. H., Abell, A. N. (2019). GALNT3 maintains the epithelial state in trophoblast stem cells. *Cell Rep.* **26**, 3684-3697.

Razawi, H., Kinlough, C. L., Staubach, S., Poland, P. A., Rbaibi, Y., Weisz, O. A., Hughey, R. P., Hanisch, F. G. (2013). Evidence for core 2 to core 1 *O*-glycan remodeling during the recycling of MUC1. *Glycobiology* **23**, 935-945.

Russo, D., Capolupo, L., Loomba, J. S., Sticco, L., D'Angelo, G. (2018). Glycosphingolipid metabolism in cell fate specification. *J. Cell. Sci.* **131**, pii: jcs219204.

Russo, D., Della Ragione, F., Rizzo, R., Sugiyama, E., Scalabrì, F., Hori, K., Capasso, S., Sticco, L., Fioriniello, S., De Gregorio, R. et al. (2018). Glycosphingolipid metabolic reprogramming drives neural differentiation. *EMBO J.* **37**, pii: e97674.

Sanchez, C. L., Sims, S. G., Nowery, J. D., Meares, G. P. (2019). Endoplasmic reticulum stress differentially modulates the IL-6 family of cytokines in murine astrocytes and macrophages. *Sci. Rep*. **9**, 14931.

Sasaki, N., Hirano, T., Ichimiya, T., Wakao, M., Hirano, K., Kinoshita-Toyoda, A., Toyoda, H., Suda, Y., Nishihara, S. (2009). The 3'-phosphoadenosine 5'-phosphosulfate transporters, PAPST1 and 2, contribute to the maintenance and differentiation of mouse embryonic stem cells. *PLoS One* **4**, e8262.

Sasaki, N., Shinomi, M., Hirano, K., Ui-Tei, K., Nishihara, S. (2011). LacdiNAc (GalNAcβ1-4GlcNAc) contributes to self-renewal of mouse embryonic stem cells by regulating leukemia inhibitory factor/STAT3 signaling. *Stem Cells* **29**, 641-650.

Schnaar, R. L., Kinoshita, T. (2015-2017). Glycosphingolipids in *Essential of Glycobiology* [Internet] *3rd edition Ch. 11* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor).

Smith, A. (2017). Formative pluripotency: the executive phase in a developmental continuum. *Development* **144**, 365-373.

Smith, A. G., Hooper, M. L. (1987). Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells. *Dev. Biol.* **121**, 1-9.

Takashima, S., Tsuji, S. (2011). Functional Diversity of Mammalian Sialyltransferases *Trends Glycosci. Glycotechnol.* **23**, 178–193.

Tang, C., Lee, A. S., Volkmer, J. P., Sahoo, D., Nag, D., Mosley, A. R., Inlay, M. A., Ardehali, R., Chavez, S. L., Pera, R. R. et al. (2011). An antibody against SSEA-5 glycan on human pluripotent stem cells enables removal of teratoma-forming cells. *Nat. Biotechnol.* **29**, 829-834.

ten Berge, D., Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R. K., Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. *Nat. Cell. Biol.* **13**, 1070-1075.

Tian, E., Ten Hagen, K. G. (2009). Recent insights into the biological roles of mucin-type *O*-glycosylation. *Glycoconj. J.* **26**, 325-334.

Torres, C. R., Hart, G. W. (1984). Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. *J. Biol. Chem.* **259**, 3308-3317.

Tran, D. T., Ten Hagen, K. G. (2013). Mucin-type *O*-glycosylation during development. *J. Biol. Chem.* **288**, 6921-6929.

Uchimura, K., Gauguet, J. M., Singer, M. S., Tsay, D., Kannagi, R., Muramatsu, T., von Andrian, U. H., Rosen, S. D. (2005). A major class of L-selectin ligands is eliminated in mice deficient in two sulfotransferases expressed in high endothelial venules. *Nat Immunol.* **6**, 1105-1113.

Varki, A. (2017). Biological roles of glycans. *Glycobiology* **27**, 3-49.

Wang, X., Inoue, S., Gu, J., Miyoshi, E., Noda, K., Li, W., Mizuno-Horikawa, Y., Nakano, M., Asahi, M., Takahashi, M. et al. (2005). Dysregulation of TGF-beta1

receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucose-deficient mice. *Proc. Natl. Acad. Sci. USA* **102**, 15791-15796.

Weerapana, E., Imperiali, B. (2006). Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology* **16**, 91R-101R.

Weinberger, L., Ayyash, M., Novershtern, N., Hanna, J. H. (2016). Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat Rev Mol Cell Biol.* **17**, 155-169.

Whelan, J. A., Russell, N. B., Whelan, M. A. (2003). A method for the absolute quantification of cDNA using real-time PCR. *J. Immunol. Methods* **278**, 261-269.

Woo, J. K., Choi, Y., Oh, S. H., Jeong, J. H., Choi, D. H., Seo, H. S., Kim, C. W. (2012). Mucin 1 enhances the tumor angiogenic response by activation of the AKT signaling pathway. *Oncogene* **31**, 2187-2198.

Xia, L., Ju, T., Westmuckett, A., An, G., Ivanciu, L., McDaniel, J. M., Lupu, F., Cummings, R. D., McEver, R. P. (2004). Defective angiogenesis and fatal embryonic hemorrhage in mice lacking core 1-derived O-glycans. *J. Cell. Biol*. **164**, 451-459.

Yang, Y., Liu, B., Xu, J., Wang, J., Wu, J., Shi, C., Xu, Y., Dong, J., Wang, C., Lai, W. et al. (2017). Derivation of Pluripotent Stem Cells with In Vivo Embryonic and Extraembryonic Potency. *Cell* **169**, 243-257.

Ying, Q. L., Nichols, J., Chambers, I., Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* **115**, 281-292.

Ying, Q. L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-523.

Ying, Q. L., Smith, A. (2017). The art of capturing pluripotency: creating the right culture. *Stem Cell Reports* **8**, 1457-1464.

Zhan, T., Rindtorff, N., Boutros, M. (2017). Wnt signaling in cancer. *Oncogene* **36**, 1461-1473.

Zhang, X., Peterson, K. A., Liu, X. S., McMahon, A. P., Ohba, S. (2013). Gene regulatory networks mediating canonical Wnt signal-directed control of pluripotency and differentiation in embryo stem cells. *Stem Cells* **31**, 2667-2679.