DEFINING ROLES FOR POLYSIALYLATION IN THE EARLY DIFFERENTIATION OF HUMAN PLURIPOTENT STEM CELLS

by

RYAN PATRICK BERGER

(Under the Direction of STEPHEN DALTON)

ABSTRACT

Glycosylation is a carbohydrate-based post-translational modification attached to proteins and lipids to alter their functional properties and increase structural diversity. The products of this modification, known as glycans, are known to participate in numerous biological processes but their specific roles in human embryonic development have remained elusive. Through this research I have explored the roles of cell surface glycans in the regulation of pluripotency and differentiation of human pluripotent stem cells (hPSCs) as well as identified a novel functional role for the glycan polymer polysialic acid (PSA). Investigation has revealed that polysialylation occurs during the initial differentiation of hESCs to progenitors of all three embryonic germ layers through an enzyme-driven process that is tightly regulated by well-defined developmental networks and is critical for proper differentiation. These findings indicate that glycans are playing active roles during embryonic development and serve to broaden the understanding of developmental glycosylation and the mechanisms driving early cell fate decisions.

INDEX WORDS: Glycosylation, human pluriptotent stem cells, glycobiology, polysialic acid, neural cell adhesion molecule, differentiation.

DEFINING ROLES FOR POLYSIALYLATION IN THE EARLY DIFFERENTIATION OF HUMAN PLURIPOTENT STEM CELLS

by

RYAN PATRICK BERGER

BS, EASTERN MICHIGAN UNIVERSITY, 2009

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2016

DEFINING ROLES FOR POLYSIALYLATION IN THE EARLY DIFFERENTIATION OF HUMAN PLURIPOTENT STEM CELLS

by

RYAN PATRICK BERGER

Major Professor:	Stephen Dalton
Committee:	Michael Pierce
	Richard Steet
	Lianchun Wang

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2016

DEDICATION

This work is dedicated to my parents Tom and Donna Berger. All of my academic achievements are the result of your endless encouragement and support throughout my life. You pushed me to achieve greater things when I was young and without that none of this would have been possible. Thanks Mom and Dad.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my family for their support over these last six years. It was difficult to move across the country to attend graduate school but I always felt like my family was there for me. That was so helpful in getting through everything.

I also am very grateful to my wife Caitee Jean Berger for her love and support. Since we met she's been incredibly encouraging and supportive of everything that I do and has made me incredibly happy every day that I'm with her.

I would like to thank Dr. Steve Dalton for taking me into his lab and providing guidance over these past several years. He has been a great mentor throughout my time in graduate school who has taught me how to think critically and design experiments to address scientific questions.

I feel extremely privileged to have spent my time at UGA working with my fellow members of the Dalton Lab. My colleagues have made it fun to go to work every day and I think what I will miss most about leaving Athens is the group camaraderie that we have all developed over the years and all the good times we've had together.

I would also like to thank Dr. Mike Pierce, Dr. Rich Steet, and Dr. Lianchun Wang for serving as members of my graduate committee and providing assistance and valuable input into my research project. Additionally, I am grateful to the members of the CCRC with whom we collaborated on this project including Kelley Moremen, Alison Nairn, and Jin Kyu Lee. Lastly I am thankful to the National Institute of Health for funding this research.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Glycosylation

Glycosylation is a post-translational modification consisting of the addition of carbohydrate structures to protein or lipid substrates. The glycosylation process occurs intracellularly in the endoplasmic reticulum and Golgi where a series of glycosyltransferase enzymes build carbohydrate structures via transfers of monosaccharides from activated sugar donors onto acceptor proteins (Varki et al. 2009). Each glycosyltransferase attaches a single monosaccharide in a specific glycosidic linkage onto substrates and multiple enzymes work together through patterns of sequential addition to generate long and complex carbohydrate polymers. The resulting glycans are characterized into two main groups depending on their amino acid attachment; N-linked glycans are attached to asparagine (Asn) and O-linked glycans are attached to serine or threonine (Ser/Thr). All N-glycans are generated from a common oligosaccharide precursor structure of $Glc_3Man_9(GlcNAc)_2$ that is assembled on a lipid donor molecule and transferred en masse onto substrates. Upon transferring, N-glycans can be subsequently trimmed, modified, and processed to yield more complex structures. In contrast, O-glycans do not all originate from a common precursor but are assembled independently on each target protein. The primary regulation of glycosylation occurs through transcriptional control of individual glycosyltransferases whose expression levels

vary widely between cell types. This process is critically important for human health and development as defective glycosylation causes a wide range of congenital disorders that have been recently characterized (Freeze 2006).

POLYSIALYLATION: PROCESS AND COMPONENTS

Polysialic acid

Polysialic acid (PSA) is a glycan linear homopolymer consisting of repeating *N*-acetylneuraminic acid (Neu5Ac or sialic acid) residues joined by α-2,8 glycosidic bonds. PSA synthesis is carried out by glycosyltransferases ST8SIA4 (PST) and ST8SIA2 (STX), which are structurally similar type II transmembrane proteins consisting of a short N terminal cytoplasmic tail and large catalytic subunit that faces the lumen of the trans Golgi compartment (Mühlenhoff et al. 2013). Each enzyme can synthesize PSA chains independently and have been observed to autopolysialylate their own N-glycans, a modification seen to enhance activity but not required for function (Close et al. 2001).

PSA was first observed in developing rat brains (Finne 1982) and subsequent studies indicated that PSA expression is also elevated during the development of many tissues in mice including heart, kidney, muscle, and brain (Close et al. 2001; Lackie et al. 1994; Angata et al. 1997; Phillips et al. 1997; Ong et al. 1998). In contrast, adult mice show very little expression of PSA except for certain regions of the brain characterized by a high level of developmental plasticity and synapse formation (Rutishauser 2008). PSA expression has only been observed in vertebrates and in some pathogenic bacteria that utilize PSA on their surface as a means of host mimicry (Stummeyer et al. 2004).

The *in vivo* mechanism of PSA removal in vertebrates is unknown, however a PSA degrading enzyme known as endoneuraminidase N (EndoN) has been isolated from bacteriophages (Rutishauser et al. 1985). EndoN specifically cleaves α-2,8 sialic acid linkages and efficiently removes PSA chains with a degree of polymerization of at least five to eight residues (Stummeyer et al. 2004).

PSA-NCAM

The neural cell adhesion molecule (NCAM) has been the most studied PSA modified protein and appears to be the primary substrate of polysialylation (Mühlenhoff et al. 2013). A member of the immunoglobulin (Ig) family of adhesion proteins, the NCAM structure consists of an intracellular domain, transmembrane domain, and extracellular domain made up of two fibronectin type III (FN III) repeats and five Ig domains (Close et al. 2003). Three main splicing isoforms exist and differ by truncation of the intracellular and transmembrane domains while each retain the same extracellular domain. These isoforms are specified by their molecular weight in kDa as NCAM-180 (full length), NCAM-140 (shortened intracellular domain), and NCAM-120 (truncated transmembrane domain, GPI anchored) (Fryer and Hockfield 1996). All three isoforms of NCAM can be polysialylated *in vitro,* however it has been observed during mouse brain development that polysialylated NCAM-180 and NCAM-140 are the prevailing forms expressed and expression over time switches to unpolysialylated NCAM-120, suggesting that isoforms 180 and 140 are more common targets of developmental PSA modifications *in vivo* (Galuska et al. 2007).

The polysialylation of NCAM begins with an initial recognition and docking of PST or STX to an α-helix and acidic surface patch on the first FN III repeat (Close et al. 2003; Mendiratta and Sekulic 2006; Mendiratta et al. 2005) followed by initiation of the first α-2,8 linkage onto terminal sialic acid residues on N-glycans in the fifth Ig domain (Figure 1.2). The process continues by elongation of the PSA chain by tens to hundreds of repeating residues (Close et al. 2003) until termination, a step that is not fully understood. Although there are six N glycan sites on NCAM that are potential targets for PSA addition, only the two sites on the fifth Ig domain are polysialylated (Close et al. 2003).

Additional polysialylated substrates

Although NCAM is the classical and most prevalent protein modified by PSA, additional polysialylated proteins have been identified. Neuropilin-2 (NRP-2), a transmembrane co-receptor involved in semamorphin and VEGF signaling, has been shown to be polysialylated on O-glycans by PST in dendritic cells (Curreli et al. 2007; Colley et al. 2014). Interestingly, the addition of PSA to NRP-2 in these cells was shown to decrease dendritic cell-induced T cell activation and proliferation to specify a novel role for PSA in immune system function (Curreli et al. 2007). Structural investigation of the interaction between PST and NRP-2 has shown that the Meprin-A5 antigen-μ tyrosine phosphatase (MAM) domain and a linker region containing the polysialylated O-glycans are the required structural elements for NRP-2 polysialylation (Bhide et al. 2016). Analysis of the MAM domain revealed an acidic surface similar to the acidic patch required on the FNII domain of NCAM used for PST recognition and docking (Figure

1.2) and suggests that the mechanisms of polysialylation are similar between substrates (Bhide et al. 2016).

Additionally, the synaptic cell adhesion molecule 1 (SynCAM1) has also been identified as a substrate of polysialylation. SynCAM1 is a cell surface adhesion protein involved in synapse formation and has been shown to be modified with PSA in polydendrocytes (NG2 cells) in the mouse brain (Galuska et al. 2010). Despite the observation that SynCAM1 can be polysialyated by both PST and STX *in vitro,* results indicate that only STX activity is present *in vivo* (Rollenhagen et al. 2012). PSA is attached to an N-glycan site in the first Ig domain of SynCAM1 (Figure 1.2), the terminal extracellular domain, and yields a complete loss of homophilic binding interactions when attached (Galuska et al. 2010). This anti-adhesive effect is similar to the loss of NCAM adhesion after PSA addition and highlights PSA as a modulator of cellular adhesion in multiple biological contexts.

A recent report has also indentified the chemokine receptor CCR7 as a carrier of PSA in lymph nodes (Kiermaier et al. 2016). The addition of PSA to CCR7 is facilitated by PST and was shown to be critical for proper dendritic cell trafficking and lymph node homeostasis (Kiermaier et al. 2016).

Other mammalian proteins reported to be PSA modified include the voltage sensitive sodium channel α-subunit(James and Agnew 1987) and the CD36 scavenger receptor in human milk (Yabe et al. 2003).

BIOLOGICAL FUNCTIONS OF PSA

Polysialylated NCAM disrupts cell-cell contacts

NCAM expression on the cell surface functions to form epithelial contacts with neighboring cells via *cis* and *trans* binding interactions. NCAM is commonly involved in homophilic binding between all five extracellular Ig domains on adjoining molecules but can also participate in heterophilic binding with other adhesion molecules and extracellular matrix components (Rønn et al. 2000). When PSA is attached to NCAM, it takes the form of a large helical structure that is highly negatively charged due to carboxyl groups of sialic acid and is very hydrated (Figure 1.3A) (Angata and Fukuda 2003). This large volume occupancy of PSA increases the hydrodynamic radius of NCAM and interferes with its ability to participate in binding interactions through short range repulsion between neighboring PSA molecules (Johnson et al. 2005). In addition, the force caused by accumulation of PSA-NCAM on the cell surface disrupts cadherin homophilic binding and is capable of completely breaking cell-cell contacts (Johnson et al. 2005).

It is through this mechanism that PSA seems to facilitate migration by breaking intercellular contacts (Figure 1.3B), which is a key component of the epithelial to mesenchymal transition (EMT). Also it lends additional function of NCAM to perform opposing mechanisms on the cell surface by adhesion and maintenance of epithelial character in one instance and repulsion and gain of mesenchymal character in another. This duality of NCAM function is tightly controlled developmentally by the expression of PST and STX through largely unknown mechanisms.

NCAM and PSA in FGF signaling

In addition to its roles in cell adhesion, NCAM has been observed to interact with signaling receptors with the most notable example being the fibroblast growth factor receptor (FGFR) (Walsh and Doherty 1997). Studies have shown direct binding of NCAM with FGFR1 and FGFR2 and structural investigations have mapped the interaction to be between the second FN III repeat of NCAM and the second and third Ig domains of FGFR1 (Kiselyov et al. 2003; Christensen et al. 2006). This binding has been shown to have a negative effect on FGF signaling pathways as phosphorylated Erk1/2 (p-Erk), a downstream target of activated FGFR, is reduced in NIH-3T3 cells overexpressing NCAM and increased upon NCAM knockdown via siRNA (Francavilla et al. 2007).

Additional studies have been done to investigate the role of PSA in FGF signaling and have found that PSA is capable of binding with FGF2 (Ono et al. 2012). A minimum chain length of 17 is required for PSA-FGF2 binding interactions and comparisons to heparan sulfate (HS), a glycosaminoglycan capable of binding FGF2, show that PSA forms larger complexes with FGF2 than HS (Ono et al. 2012). To investigate the effects of PSA-NCAM in FGF signaling, stable NIH-3T3 cell lines overexpressing NCAM and STX or NCAM and PST were generated and results showed that PSA-NCAM expressing cells displayed increased Erk activation upon stimulation with FGF2 compared to control cells (Ono et al. 2012). Although the mechanism by which PSA-NCAM activates signaling is not fully understood, it is likely that PSA acts as a reservoir of FGF2 and activates FGF signaling by directly passing ligands to receptors or by maintaining a high local concentration of FGF2 on the cell surface (Ono et al. 2012).

Polysialylation in neural development

Perhaps the most well-studied biological utilization of sialylation to date involves the developing nervous system of mice where it has been shown to play a critical role (Schwarzkopf et al. 2002). One aspect of neonatal development that is facilitated by PSA-NCAM is the migration of neuronal progenitor cells from the subventricular zone (SVZ) along the rostral migratory stream (RMS) to the olfactory bulb in the brain (Figure 1.4). Removal of PSA using EndoN yields abnormal accumulation of cells in the SVZ and drastically reduced migration to the olfactory bulb (Figure 1.4) (Cremer et al. 1994; Hu et al. 1996). Additionally, polysialylated NCAM is critical for muscular axon innervation as cells lacking PSA are unable to separate from each other and effectively form mature branch structures (Landmesser et al. 1990). Overall it appears that neuronal precursors express PSA on the cell surface to migrate to their end destination and begin synapse formation and then lose expression when they become mature neurons (Bruses et al. 1995).

PSA is critical for mouse development

The importance of PSA and the individual components of polysialylation have been illustrated through the use of PST, STX, and NCAM knockout mice. Cremer et. al published the first study to utilize a homozygous knockout of the NCAM gene and results showed a relatively mild phenotype. The mice were healthy and fertile but had a 10% reduction in overall brain weight and a 36% reduction in size of the olfactory bulb, reinforcing the importance of PSA-NCAM in the migration from the SVZ (Cremer et al. 1994). They also exhibited slight disabilities in spatial learning but had normal activity

and motor abilities (Cremer et al. 1994). Individual knockouts of PST and STX also showed mild phenotypes including learning disabilities, reduced response to fear, impairment of hippocampal long-term potentiation, and depression (Eckhardt et al. 2000; Angata et al. 2004). In both models there remained PSA expression and normal migration from the SVZ, indicating that the roles of each enzyme may overlap or that each is capable of compensating for the loss of the other.

Despite individual knockouts of NCAM, PST, and STX each yielding a similar phenotype, investigations into a double knockout of PST and STX generated much more severe defects. Mice lacking both polysialyltransferases showed a high rate of lethality as only 50% survived to 3 weeks of age and only 20% to 4 weeks (Weinhold et al. 2005). Those that survived exhibited growth retardation, small olfactory bulbs, and severe deficits in motor coordination, strength, and balance (Weinhold et al. 2005). Further studies examining the effects on brain development in greater detail found gross migratory defects throughout the brain as well as aberrant positioning of cells (Angata et al. 2007).

Interestingly, mice lacking NCAM, PST, and STX showed a similar mild phenotype as the NCAM individual knockout and did not display the harsh phenotypes or lethality of the PST/STX null mice (Weinhold et al. 2005). These results suggest that in addition to the defects caused by lack of polysialylation, the severe phenotype and lethality is caused by aberrant expression of unpolysialylated NCAM. This changes the paradigm of PSA-NCAM in a way that implies that PSA expression is critical for masking NCAM during development to allow key events to occur that are otherwise occluded with unmodified NCAM.

PSA IN CANCER

One aspect of polysialylation research that is of particular interest for clinical purposes is the abundant expression of PSA on the surface of cancer cells. Consistent with the ability to facilitate migration, PSA is commonly found in highly metastatic cancers such as neuroblastoma (Takahashi et al. 2007; Livingston et al. 1988), pancreatic carcinoma (Schreiber et al. 2008), small cell lung carcinoma (Sumi et al. 2007; Komminoth et al. 1991; Ohtsuka and Dalton 2007), and Wilms' tumor (Roth et al. 1988). An investigation into pancreatic cancer supports the hypothesis that PSA expression is utilized to drive metastasis by showing that expression of PSA-NCAM disrupts E-cadherin mediated junctions and enhances transwell migration ability and that enzymatic removal of PSA by EndoN restores E-cadherin junctions and causes cells to aggregate (Schreiber et al. 2008). Elevated levels of both PST and STX have been found in cancer and it has been observed in many cancer subtypes that PSA expression correlates with tumor progression and is an adverse prognosis factor (Suzuki 2005; Amoureux et al. 2010; Tanaka et al. 2000).

The mechanisms for activation of PST or STX in cancer are not well understood but may be driven by aberrant expression of the transcription factors used in development. Recent evidence has shown cancer re-expression of EMT developmental regulator genes and this expression has been shown to promote metastasis (Hartwell et al. 2006; Fernando et al. 2010; Kilic et al. 2011). In particular, the T-box transcription factor Brachyury (T) and Spemann organizer gene Goosecoid (GSC), both important in early differentiation to mesendoderm lineages, are highly expressed in epithelial cancers of the lung, colon, breast, and prostate (Fernando et al. 2010; Hartwell et al. 2006). Moreover the *in vitro* overexpression of T or GSC in cancer cell lines induced EMTs and

significantly enhanced cell migration and invasion ability (Fernando et al. 2010; Hartwell et al. 2006), indicating that re-expression of these developmental genes confers greater metastatic potential. However, these studies did not investigate the expression of PSA on the cell surface but it is plausible to speculate considering that PSA is expressed in embryonic stages similar to T and GSC.

Given this information it is of great interest to understand the molecular mechanisms that regulate expression of the polysialyltransferases as doing so would not only provide insight into developmental processes but could potentially identify novel anti-cancer drug targets. Using this approach to target PSA expression and EMT ability would be very useful in treatments considering that cancer lethality is due in large part to its ability to metastasize to distant sites throughout the body (Klein 2009).

PLURIPOTENT STEM CELLS

Human embryonic stem cells as a model system for studying PSA in development

Since the establishment of the first *in vitro* cell lines in the late 1990s, human embryonic stem cells (hESCs) have been a valuable tool for studying human development and disease. They are derived from the inner cell mass of pre-implantation blastocysts and are characterized by their ability to self-renew indefinitely and ability to differentiate into cell types of all three germ layers (Thomson et al. 1998). More recently the generation of induced pluripotent stem cells (iPSCs) from adult fibroblasts that share the same characteristics as hESCs has unveiled vast potential for stem cell technology in regenerative medicine and disease modeling (Takahashi et al. 2007).

In culture, hESCs or iPSCs can be maintained in their pluripotent state using chemically defined media containing recombinant cytokines that stimulate the FGF, PI3 kinase, and Activin/Nodal signaling pathways (Wang et al. 2007). The pluripotent state is regulated intracellularly by a core of transcription factors including Nanog, Sox2, Oct4, and c-myc (Sumi et al. 2007; Ohtsuka and Dalton 2007). Additionally, the *in vitro* differentiation of hESCs has been well studied and various protocols have been generated to generate cell types of all three germ layers (Close and Colley 1998; McLean et al. 2007; Menendez et al. 2011; Singh et al. 2012). Understanding the mechanisms of pluripotency and differentiation represents a key step in the potential use of hESCs for therapeutics and regenerative medicine.

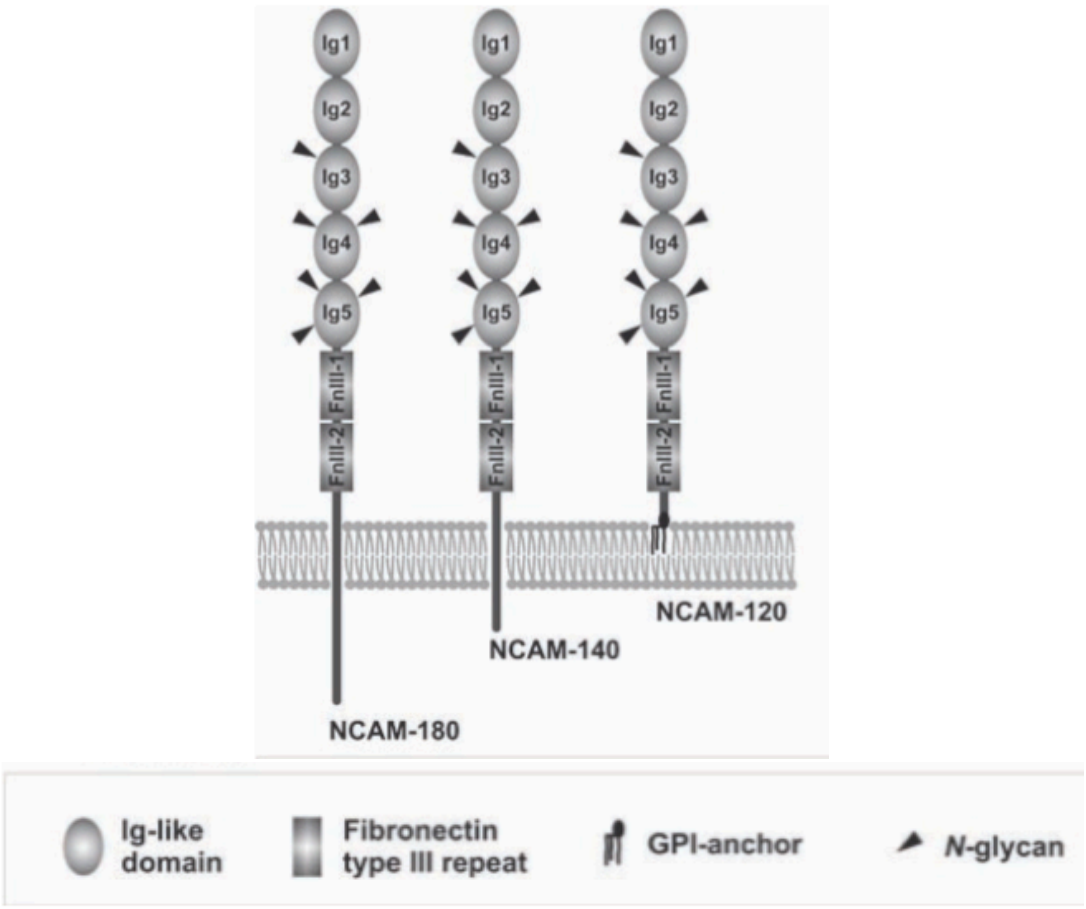


Figure 1.1 – NCAM isoforms. Illustration of the three splice isoforms of NCAM showing the same extracellular domain but a variable transmembrane domain. *(Adapted from Hildebrandt et al. 2009).*

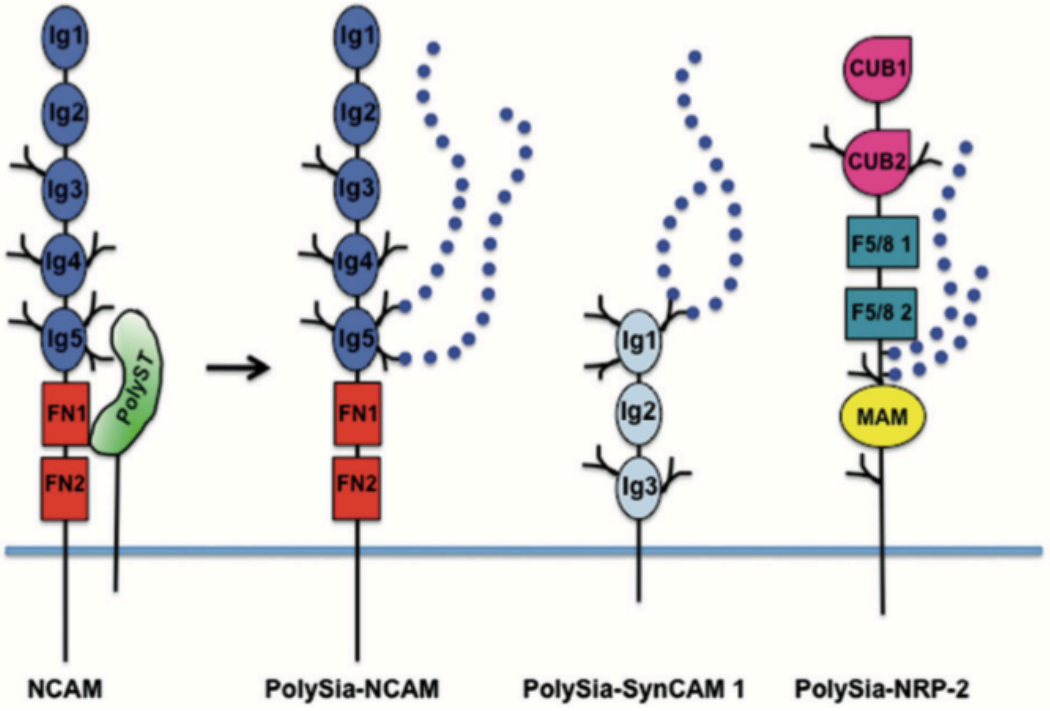


Figure 1.2 – Substrates of polysialylation. Illustration of PSA modified proteins NCAM, SynCAM1, and NRP-2 on the cell surface. *(Adapted from Colley et al., 2014)*.

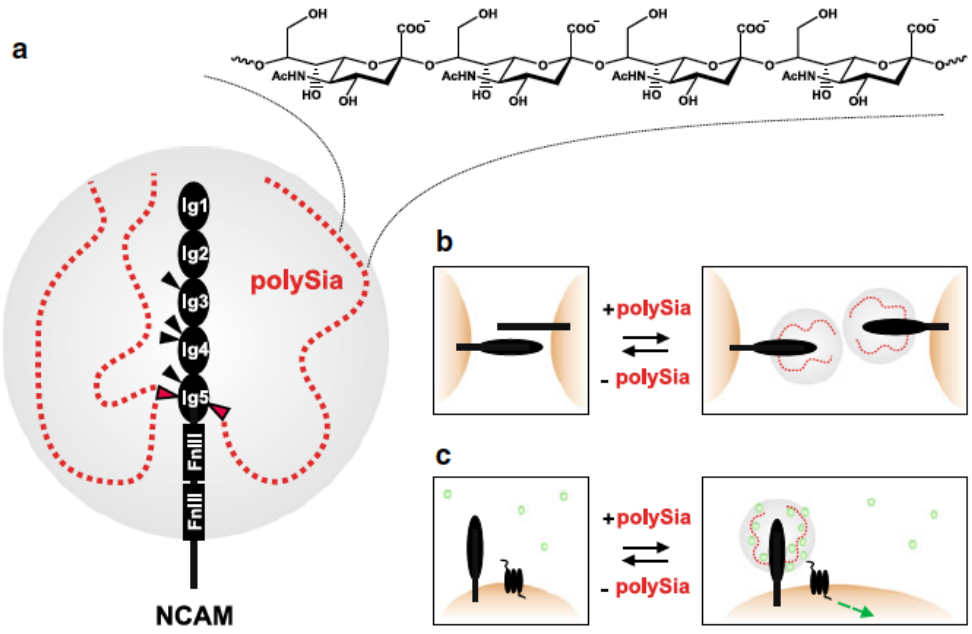


Figure 1.3 – Representation of PSA-NCAM in intercellular binding. (A) Schematic illustrating the larger hydrodynamic radius (gray shaded) of NCAM due to PSA addition (red squares). (B) NCAM homophilic binding interactions in the absence (left) or presence (right) of PSA modifications. (C) FGF signaling hindered by NCAM and activated by PSA-NCAM. *(Adapted from Mühlenhoff et al., 2013).*

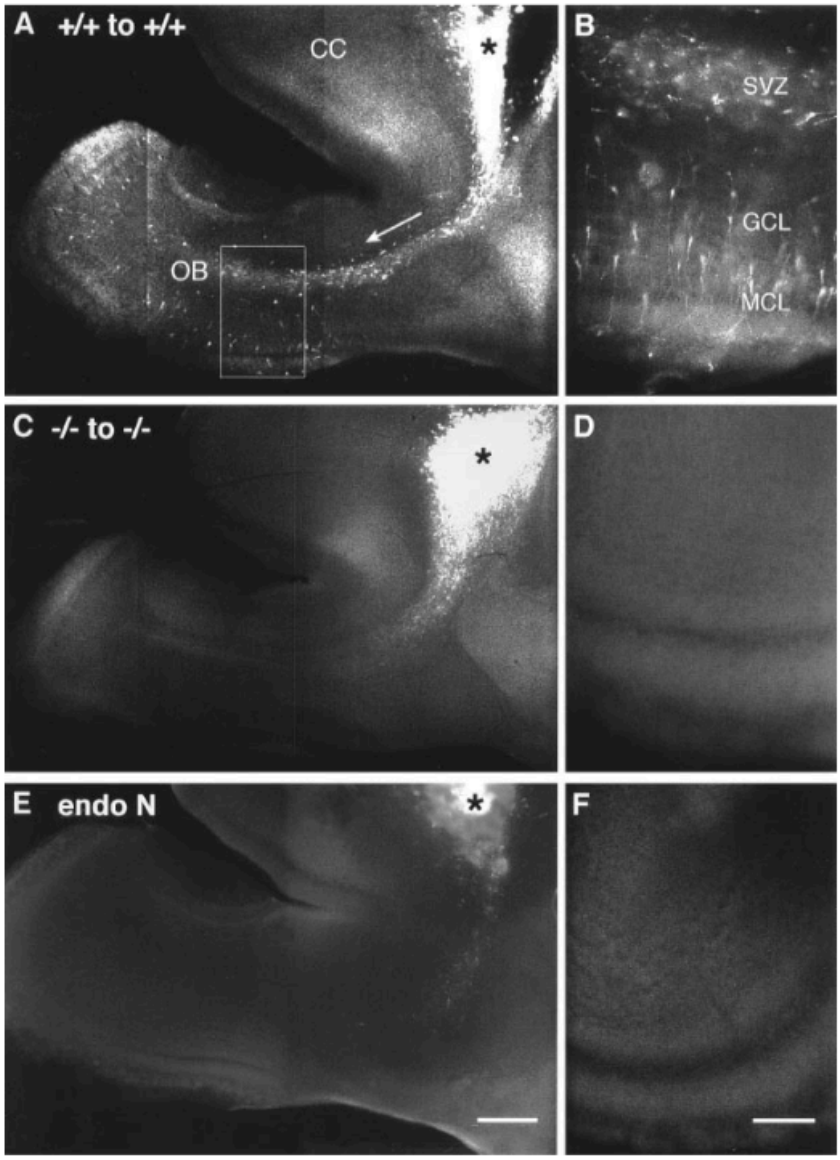


Figure 1.4 – PSA is required for neuronal migration to sub-ventricular zone. (A, C, E) Sections of P3 mouse brain showing DiI labeled cells 7d after transplant into the SVZ. Comparison shown is wild type cells (A-B), NCAM deficient cells (C-D), and wild type cells treated with Endo N (E-F). Higher magnification is shown for each (B, D, F). CC, cerebral cortex; OB, olfactory bulb; GCL, granule cell layer; MCL, mitral cell layer. Scale bar represents 250 μm for (A, C, E) and 50 μm for (B, D, F). *(Adapted from Hu et al., 1996).*

REFERENCES

Amoureux M-C, Coulibaly B, Chinot O, Loundou A, Metellus P, Rougon G, Figarella-Branger D. 2010. Polysialic Acid Neural Cell Adhesion Molecule (PSA-NCAM) is an adverse prognosis factor in glioblastoma, and regulates olig2 expression in glioma cell lines. *BMC Cancer* **10**: 91.

Angata K, Fukuda M. 2003. Polysialyltransferases: major players in polysialic acid synthesis on the neural cell adhesion molecule. *Biochimie* **85**: 195–206.

Angata K, Huckaby V, Ranscht B, Terskikh A, Marth JD, Fukuda M. 2007. Polysialic Acid-Directed Migration and Differentiation of Neural Precursors Are Essential for Mouse Brain Development. *Molecular and Cellular Biology* **27**: 6659–6668.

Angata K, Long JM, Bukalo O, Lee W, Dityatev A, Wynshaw-Boris A, Schachner M, Fukuda M, Marth JD. 2004. Sialyltransferase ST8Sia-II assembles a subset of polysialic acid that directs hippocampal axonal targeting and promotes fear behavior. *Journal of Biological Chemistry* **279**: 32603–32613.

Angata K, Nakayama J, Fredette B, Chong K, Ranscht B, Fukuda M. 1997. Human STX polysialyltransferase forms the embryonic form of the neural cell adhesion molecule - Tissue-specific expression, neurite outgrowth, and chromosomal localization in comparison with another polysialyltransferase, PST. *Journal of Biological Chemistry* **272**: 7182–7190.

Bhide GP, Fernandes NRJ, Colley KJ. 2016. Sequence requirements for neuropilin-2 recognition by ST8SiaIV and polysialylation of its O-glycans. *J Biol Chem* jbc.M116.714329.

Bruses JL, Oka S, Rutishauser U. 1995. NCAM-associated polysialic acid on ciliary ganglion neurons is regulated by polysialytransferase levels and interaction with muscle. *The Journal of Neuroscience* **15**: 8310–8319.

Christensen C, Lauridsen JB, Berezin V, Bock E, Kiselyov VV. 2006. The neural cell adhesion molecule binds to fibroblast growth factor receptor 2. *FEBS Letters* **580**: 3386–3390.

Close BE, Colley KJ. 1998. In vivo autopolysialylation and localization of the polysialyltransferases PST and STX. *Journal of Biological Chemistry* **273**: 34586–34593.

Close BE, Mendiratta SS, geiger KM, broom LJ, Ho L-L, Colley KJ. 2003. The Minimal Structural Domains Required for Neural Cell Adhesion Molecule Polysialylation by PST/ST8Sia IV and STX/ST8Sia II. *Journal of Biological Chemistry* **278**: 30796–30805.

Close BE, Wilkinson JM, Bohrer TJ, Goodwin CP, broom LJ, Colley KJ. 2001. The polysialyltransferase ST8Sia II/STX: posttranslational processing and role of autopolysialylation in the polysialylation of neural cell adhesion molecule. *Glycobiology* **11**: 997–1008.

Colley KJ, Kitajima K, Sato C. 2014. Polysialic acid: Biosynthesis, novel functions and applications. *Critical Reviews in Biochemistry and Molecular Biology* **49**: 498–532.

Cremer H, Lange R, Christoph A, Plomann M, Vopper G, Roes J, Brown R, Baldwin S, Kraemer P, Scheff S. 1994. Inactivation of the N-CAM gene in mice results in size reduction of the olfactory bulb and deficits in spatial learning.

Curreli S, Arany Z, Gerardy-Schahn R, Mann D, Stamatos NM. 2007. Polysialylated Neuropilin-2 Is Expressed on the Surface of Human Dendritic Cells and Modulates Dendritic Cell-T Lymphocyte Interactions. *Journal of Biological Chemistry* **282**: 30346–30356.

Eckhardt M, Bukalo O, Chazal G, Wang L, Goridis C, Schachner M, Gerardy-Schahn R, Cremer H, Dityatev A. 2000. Mice deficient in the polysialyltransferase ST8SiaIV/PST-1 allow discrimination of the roles of neural cell adhesion molecule protein and polysialic acid in neural development and synaptic plasticity. *The Journal of Neuroscience* **20**: 5234–5244.

Fernando RI, Litzinger M, Trono P, Hamilton DH, Schlom J, Palena C. 2010. The T-box transcription factor Brachyury promotes epithelial-mesenchymal transition in human tumor cells. *Journal of clinical investigation* **120**: 533–544.

Finne J. 1982. Occurrence of unique polysialosyl carbohydrate units in glycoproteins of developing brain. *Journal of Biological Chemistry* **257**: 11966–11970.

Francavilla C, Loeffler S, Piccini D, Kren A, Christofori G, Cavallaro U. 2007. Neural cell adhesion molecule regulates the cellular response to fibroblast growth factor. *Journal of Cell Science* **120**: 4388–4394.

Freeze HH. 2006. Genetic defects in the human glycome. *Nat Rev Genet* **7**: 537–551.

Fryer HJ, Hockfield S. 1996. The role of polysialic acid and other carbohydrate polymers in neural structural plasticity. *Current opinion in neurobiology* **6**: 113–118.

Galuska SP, Geyer R, Gerardy-Schahn R, Muhlenhoff M, Geyer H. 2007. Enzyme-dependent Variations in the Polysialylation of the Neural Cell Adhesion Molecule (NCAM) in Vivo. *Journal of Biological Chemistry* **283**: 17–28.

Galuska SP, Rollenhagen M, Kaup M, Eggers K, Oltmann-Norden I, Schiff M, Hartmann M, Weinhold B, Hildebrandt H, Geyer R. 2010. Synaptic cell adhesion molecule SynCAM 1 is a target for polysialylation in postnatal mouse brain. *Proceedings of the National Academy of Sciences* **107**: 10250–10255.

Hartwell KA, Muir B, Reinhardt F, Carpenter AE, Sgroi DC, Weinberg RA. 2006. The Spemann organizer gene, Goosecoid, promotes tumor metastasis. *Proceedings of the National Academy of Sciences* **103**: 18969–18974.

Hu H, Tomasiewicz H, Magnuson T, Rutishauser U. 1996. The role of polysialic acid in migration of olfactory bulb interneuron precursors in the subventricular zone. *Neuron* **16**: 735–743.

James WM, Agnew WS. 1987. Multiple oligosaccharide chains in the voltage-sensitive Na channel from electrophorus electricus: evidence for alpha-2,8-linked polysialic acid. *Biochemical and Biophysical Research Communications* **148**: 817–826.

Johnson CP, Fujimoto I, Rutishauser U, Leckband DE. 2005. Direct evidence that neural cell adhesion molecule (NCAM) polysialylation increases intermembrane repulsion and abrogates adhesion. *Journal of Biological Chemistry* **280**: 137–145.

Kiermaier E, Moussion C, Veldkamp CT, Gerardy-Schahn R, de Vries I, Williams LG, Chaffee GR, Phillips AJ, Freiberger F, Imre R, et al. 2016. Polysialylation controls dendritic cell trafficking by regulating chemokine recognition. *Science* **351**: 186–190.

Kilic N, Feldhaus S, Kilic E, Tennstedt P, Wicklein D, Wasielewski von R, Viebahn C, Kreipe H, Schumacher U. 2011. Brachyury expression predicts poor prognosis at early stages of colorectal cancer. *European Journal of Cancer* **47**: 1080–1085.

Kiselyov VV, Skladchikova G, Hinsby AM, Jensen PH. 2003. Structural basis for a direct interaction between FGFR1 and NCAM and evidence for a regulatory role of ATP. *Structure* **11**: 691–701.

Klein CA. 2009. Parallel progression of primary tumours and metastases. *Nature Reviews Cancer* **9**: 302–312.

Komminoth P, Roth J, Lackie PM, Bitter-Suermann D, Heitz PU. 1991. Polysialic acid of the neural cell adhesion molecule distinguishes small cell lung carcinoma from carcinoids. *The American Journal of Pathology* **139**: 297.

Lackie PM, Zuber C, Roth J. 1994. Polysialic acid of the neural cell adhesion molecule (N-CAM) is widely expressed during organogenesis in mesodermal and endodermal derivatives. *Differentiation* **57**: 119–131.

Landmesser L, Dahm L, Tang J, Rutishauser U. 1990. Polysialic acid as a regulator of intramuscular nerve branching during embryonic development. *Neuron* **4**: 655–667.

Livingston BD, Jacobs JL, Glick MC, Troy FA. 1988. Extended polysialic acid chains (n > 55) in glycoproteins from human neuroblastoma cells. *Journal of Biological Chemistry* **263**: 9443–9448.

McLean AB, D'Amour KA, Jones KL, Krishnamoorthy M, Kulik MJ, Reynolds DM, Sheppard AM, Liu H, Xu Y, Baetge EE, et al. 2007. Activin A Efficiently Specifies Definitive Endoderm from Human Embryonic Stem Cells Only When Phosphatidylinositol 3-Kinase Signaling Is Suppressed. *Stem Cells* **25**: 29–38.

Mendiratta SS, Sekulic N. 2006. A novel α-helix in the first fibronectin type III repeat of the neural cell adhesion molecule is critical for N-glycan polysialylation. *Journal of Biological Chemistry*.

Mendiratta SS, Sekulic N, Lavie A, Colley KJ. 2005. Specific amino acids in the first fibronectin type III repeat of the neural cell adhesion molecule play a role in its recognition and polysialylation by the polysialyltransferase ST8SiaIV/PST. *Journal of Biological Chemistry*.

Menendez L, Yatskievych TA, Antin PB, Dalton S. 2011. Wnt signaling and a Smad pathway blockade direct the differentiation of human pluripotent stem cells to multipotent neural crest cells. *Proceedings of the National Academy of Sciences* **108**: 19240–19245.

Mühlenhoff M, Rollenhagen M, Werneburg S, Gerardy-Schahn R, Hildebrandt H. 2013. Polysialic acid: versatile modification of NCAM, SynCAM 1 and neuropilin-2. *Neurochemical Research* **38**: 1134–1143.

Ohtsuka S, Dalton S. 2007. Molecular and biological properties of pluripotent embryonic stem cells. *Gene Ther* **15**: 74–81.

Ong E, Nakayama J, Angata K, Reyes L, Katsuyama T, Arai Y, Fukuda M. 1998. Developmental regulation of polysialic acid synthesis in mouse directed by two polysialyltransferases, PST and STX. ....

Ono S, Hane M, Kitajima K, Sato C. 2012. Novel Regulation of Fibroblast Growth Factor 2 (FGF2)-mediated Cell Growth by Polysialic Acid. *Journal of Biological Chemistry* **287**: 3710–3722.

Phillips GR, Krushel LA, Crossin KL. 1997. Developmental expression of two rat sialyltransferases that modify the neural cell adhesion molecule, N-CAM. *Developmental Brain Research* **102**: 143–155.

Rollenhagen M, Kuckuck S, Ulm C, Hartmann M, Galuska SP, Geyer R, Geyer H, Muhlenhoff M. 2012. Polysialylation of the Synaptic Cell Adhesion Molecule 1 (SynCAM 1) Depends Exclusively on the Polysialyltransferase ST8SiaII in Vivo. *Journal of Biological Chemistry* **287**: 35170–35180.

Roth J, Zuber C, Wagner P, Blaha I, Bitter-Suermann D, Heitz PU. 1988. Presence of the long chain form of polysialic acid of the neural cell adhesion molecule in Wilms' tumor. Identification of a cell adhesion molecule as an oncodevelopmental antigen and implications for tumor histogenesis. *The American Journal of Pathology* **133**: 227.

Rutishauser U. 2008. Polysialic acid in the plasticity of the developing and adult vertebrate nervous system. *Nature Reviews Neuroscience* **9**: 26–35.

Rutishauser U, Watanabe M, Silver J, Troy FA, Vimr ER. 1985. Specific alteration of NCAM-mediated cell adhesion by an endoneuraminidase. *The Journal of cell biology* **101**: 1842–1849.

Rønn LC, Berezin V, Bock E. 2000. The neural cell adhesion molecule in synaptic plasticity and ageing. *International Journal of Developmental Neuroscience* **18**: 193–199.

Schreiber SC, Giehl K, Kastilan C, Hasel C, Mühlenhoff M, Adler G, Wedlich D, Menke A. 2008. Polysialylated NCAM Represses E-Cadherin-Mediated Cell-Cell Adhesion in Pancreatic Tumor Cells. *Gastroenterology* **134**: 1555–1566.

Schwarzkopf M, Knobeloch K-P, Rohde E, Hinderlich S, Wiechens N, Lucka L, Horak I, Reutter W, Horstkorte R. 2002. Sialylation is essential for early development in mice. *Proceedings of the National Academy of Sciences* **99**: 5267–5270.

Singh AM, Reynolds D, Cliff T, Ohtsuka S, Mattheyses AL, Sun Y, Menendez L, Kulik M, Dalton S. 2012. Signaling Network Crosstalk in Human Pluripotent Cells: A Smad2/3-Regulated Switch that Controls the Balance between Self-Renewal and Differentiation. *Cell Stem Cell* **10**: 312–326.

Stummeyer K, Dickmanns A, Mühlenhoff M, Gerardy-Schahn R, Ficner R. 2004. Crystal structure of the polysialic acid–degrading endosialidase of bacteriophage K1F. *Nat Struct Mol Biol* **12**: 90–96.

Sumi T, Tsuneyoshi N, Nakatsuji N, Suemori H. 2007. Apoptosis and differentiation of human embryonic stem cells induced by sustained activation of c-Myc. *Oncogene* **26**: 5564–5576.

Suzuki M. 2005. Polysialic acid facilitates tumor invasion by glioma cells. *Glycobiology* **15**: 887–894.

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. 2007. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* **131**: 861–872.

Tanaka F, Otake Y, Nakagawa T, Kawano Y, Miyahara R, Li M, Yanagihara K, Nakayama J, Fujimoto I, Ikenaka K. 2000. Expression of polysialic acid and STX, a human polysialyltransferase, is correlated with tumor progression in non-small cell lung cancer. *Cancer Research* **60**: 3072–3080.

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. 1998. Embryonic stem cell lines derived from human blastocysts. *Science* **282**: 1145–1147.

Varki A, RD C, JD E, HH F, P S, CR B, GW H, ME E. 2009. *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press.

Walsh FS, Doherty P. 1997. Neural cell adhesion molecules of the immunoglobulin superfamily: role in axon growth and guidance. *Annual review of cell and developmental biology* **13**: 425–456.

Wang L, Schulz TC, Sherrer ES, Dauphin DS, Shin S, Nelson AM, Ware CB, Zhan M, Song CZ, Chen X, et al. 2007. Self-renewal of human embryonic stem cells requires insulin-like growth factor-1 receptor and ERBB2 receptor signaling. *Blood* **110**: 4111–4119.

Weinhold B, Seidenfaden R, Röckle I, Mühlenhoff M, Schertzinger F, Conzelmann S, Marth JD, Gerardy-Schahn R, Hildebrandt H. 2005. Genetic Ablation of Polysialic Acid Causes Severe Neurodevelopmental Defects Rescued by Deletion of the Neural Cell Adhesion Molecule. *Journal of Biological Chemistry* **280**: 42971–42977.

Yabe U, Sato C, Matsuda T, Kitajima K. 2003. Polysialic Acid in Human Milk: CD36 IS A NEW MEMBER OF MAMMALIAN POLYSIALIC ACID-CONTAINING GLYCOPROTEIN. *J Biol Chem* **278**: 13875–13880.

CHAPTER 2

DYNAMIC GLYCOSYLATION PATTERNS PLAY KEY ROLES IN THE PLURIPOTENCY AND DIFFERENTIATION OF STEM CELLS

Berger RP, Dookwah M, Steet R, Dalton S. 2016. To be submitted to *Bioessays*.

ABSTRACT

Glycosylation is a common post-translational modification occurring on over half of human proteins. Most notably, the surface of cells is densely covered by a heterogeneous array of complex glycan structures that varies significantly between cell types. Recently, the surface glycosylation patterns of pluripotent stem cells (PSCs) have been documented and their functional impacts on pluripotency and differentiation have been revealed. This review aims to highlight the major findings in the field of PSC glycosylation. Main points include: (1) PSCs express simpler N-glycans compared to differentiated cells as evidenced by an unusual abundance of high mannose structures, (2) α1-2 fucosylation is common in PSCs and appears to be an identifying mark recognized by many stem cell markers, and (3) dynamic sialylation patterns play critical roles in maintaining pluripotency. In addition, recent studies utilizing induced pluripotent stem cells (iPSCs) to study congenital diseases of glycosylation are discussed.

INTRODUCTION

Glycosylation is a post-translational modification (PTM) utilized extensively to increase the structural diversity of proteins and lipids through the addition of carbohydrates onto target molecules. Glycans are highly expressed on the cellular surface of living organisms and represent the most common PTM as it has been estimated that more than half of all proteins are glycosylated (Apweiler et al. 1999). The process of glycosylation is performed intracellularly by hundreds of carbohydrate transferring enzymes, termed glycosyltransferases, which work together in a sequential and coordinated manner to yield an extremely wide-ranging array of glycan products that confer differing properties and biological functions onto their target molecules. The importance of carbohydrate structures for organismal survival is evident by the observation that the surface of virtually all living cells are densely covered by heterogeneous arrays of complex carbohydrates in what is known as the glycocalyx and suggests that carbohydrate modifications have provided a selective advantage. In addition, carbohydrate structures have been shown to perform a variety of cellular functions critical for development and survival. Unfortunately, the immense heterogeneity of glycans has made studying and elucidating their functions a difficult task for researchers using past techniques. Recent high-throughput approaches, however, have opened up the field and allowed for a better understanding of the complex biological roles that glycosylation plays in living organisms.

One area of research that has advanced greatly in recent years is the study of glycans in embryonic development. Complex carbohydrate structures have long been observed during developmental stages and many have been used as stage-specific cell

markers but determining their potential functions in pluripotency and germ layer specification has remained elusive. Furthermore, previous studies have primarily examined glycan structures on an individual basis and have not observed the effects of the entire cellular glycome on developmental processes. Because of this, the use of human pluripotent stem cells (hPSCs) as a model system coupled with modern high-throughput glycomics technologies has become an increasing trend and has allowed for a more thorough analysis of the expression and impact of glycan structures during embryonic development. In this review, we highlight the recent findings in hPSC glycobiology and summarize the use of induced pluripotent stem cells (iPSCs) as an *in vitro* model of congenital diseases of glycosylation.

GLYCANS AS MARKERS OF PLURIPOTENT STEM CELLS

Given their abundant expression on the cell surface and vast structural diversity between cell types, carbohydrate structures are the epitope for many antibodies commonly used as cell markers and have been used extensively to identify pluripotent cells. Stage specific embryonic antigen 1 (SSEA-1) was among the first markers of mouse embryonic carcinoma cells and recognizes the lacto-series Lewis X antigen (Galβ1-4(Fucα1-3)GlcNAcβ1-) (Gooi et al. 1981; Solter and Knowles 1978). mESCs express Lewis X highly at the 8 cell stage and through the blastocyst stage but lose expression upon differentiation to defined lineages (Solter and Knowles 1978; Pennington et al. 1985). In contrast, hPSCs do not express the Lewis X antigen and are negative for SSEA-1 but are instead identified by SSEA-3 and SSEA-4, which recognize similar globo-series ganglioside antigens (Gooi et al. 1981; Kannagi et al. 1983; Damjanov et al. 1982).

Monoclonal antibodies against tumor rejection antigens 1-60 and 1-81 (TRA-1-60 and TRA-1-81) are also commonly used for identifying hPSCs through recognition of terminal type 1 terminal lactosamine structures (Galβ1–3GlcNAc) (Andrews et al. 1984; Natunen et al. 2011). These epitopes are carried by the heavily glycosylated membrane protein podocalyxin and are down-regulated upon differentiation (Schopperle and DeWolf 2006). Each of these antibodies identifies glycan epitopes that are highly expressed in pluripotent cells as compared to differentiated and somatic cells; together, they have been used as the major PSC markers for decades.

Despite their utility, however, the glycan epitopes commonly used to identify PSCs are not required for the pluripotent state and have also shown expression in other cell types (Brimble et al. 2006). This is potentially problematic for downstream therapeutic applications in which the identification and removal of PSCs prior to transplantation is critically important. The issue lies in the ability of pluripotent cells to proliferate *in vivo* and differentiate into teratomas, benign tumors composed of cells from multiple germ layers, after cells are transplanted into a patient. To prevent this, residual PSCs must be completely removed from a cell population before they are used for cellular therapy.

To address these concerns, recent efforts have utilized high throughput techniques to identify novel glycan epitopes that are more sensitive and specific to PSCs than the traditional reagents. One such example is SSEA-5, a monoclonal antibody identified from a hybridoma library that highly labels undifferentiated cells and the blastocyst inner cell mass (Tang et al. 2011). SSEA-5 recognizes the H-type 1 epitope, a trisaccharide O-glycan (Fucα1-2Galβ1-3GlcNAc), that is highly expressed in pluripotent cells but lost

rapidly upon differentiation and is likely carried by multiple proteins on the cell surface. SSEA-5 showed a high dynamic range of expression as differentiation resulted in 2 to 3 order of magnitude reduction in signal, which was much greater than traditional markers SSEA-3, SSEA-4, and TRA-1-81 (Tang et al. 2011). Additionally, SSEA-5 was used along with markers CD9 and CD90 to completely eliminate residual pluripotent cells prior to transplantation in a murine model and prevented teratoma formation whereas removal using SSEA-3 and TRA-1-81 did not (Tang et al. 2011).

Similar efforts have recently been made to find PSC markers using lectins: highly specific carbohydrate binding proteins. Lectins are an ideal reagent for identifying cells due to their glycan sensitivity, compatibility between species, reversible binding to the cell surface, and much lower expense compared to antibodies.

The lectin rBC2LCN was identified using a high-density glycan microarray and was found to bind highly to over 100 PSC lines while not binding to any somatic cells tested (Tateno et al. 2011). Originally derived from the bacterium *Burkholderia cenocepacia,* rBC2LCN recognizes the epitope of SSEA-5 (H-type 1) but also binds to the similar epitopes H-type 3 and 4 *(*Fucα1-2Galβ1-3GlcNAc/GalNAc) (Tateno et al. 2011; SulAk et al. 2010). These glycan structures were found to be attached to podacalyxin on the cell surface, the same protein carrier of TRA-1-60 and TRA-1-81, suggesting that rBC2LCN recognizes multiple elements commonly used to identify PSCs (Tateno et al. 2013; Schopperle and DeWolf 2006). Most importantly, rBC2LCN was shown to completely eliminate pluripotent cells from a mixed population and may prove useful for therapeutic applications (Tateno et al. 2015).

In a similar experiment, Wang et al. identified a set of three lectins that could specifically identify PSCs via recognition of fucosylated and sialylated glycans (Wang et al. 2011). One of the lectins, fucose binding UEA-1, showed very low reactivity toward differentiated progenitors and was capable of removing 99.5% of PSCs from a mixed population of three cell types (Wang et al. 2011). Additionally, isolation of PSCs using UEA-1 conjugated to magnetic beads was shown to be effective and yield viable cells capable of differentiating to all three germ layers after elution.

GLYCAN PROFILE OF PLURIPOTENT STEM CELLS

Pluripotency is marked by simpler N-glycans

The global surface glycan expression profiles vary considerably between cell types and recent reports have shown that PSCs express a characteristic pattern of surface glycosylation. Among the most prominent features of the hPSC glycan signature is the abundant expression of high mannose N-glycans on the cell surface (An et al. 2012; Satomaa et al. 2009; Fujitani et al. 2013; Hasehira et al. 2012). High mannose structures are the core building block for all N-glycans that become processed enzymatically into more complex structures in the Golgi. The expression of high mannose glycans in PSCs is surprising because they are not typically expressed in high quantities on the surface of mammalian cells yet represent as high as 85% of the total N-glycome in PSCs (An et al. 2012). This contrasts considerably with previous studies that have identified complex and hybrid types as representing the vast majority of N-glycan structures in adult cell lineages and human serum fractions (Tao et al. 2008; Chu et al. 2009; An et al. 2012; Hasehira et al. 2012). Additionally, hPSCs express a high level of terminal glucose on the termini of

high mannose structures (An et al. 2012). This is also unusual as terminal glucose is not typically found on the surface of mammalian cells but is often expressed in pathogens and recognized by the innate immune system (An et al. 2012).

Fucosylated glycans are a strong indicator of pluripotency in PSCs

Fucose is a deoxyhexose monosaccharide that is involved in a variety of biological processes in eukaryotic organisms including cell adhesion, signaling, and embryonic development (Becker and Lowe 2003). Fucosylation occurs in the Golgi by a family of 13 fucosyltransferases which catalyze the addition of fucose onto N- and O- glycan structures and also, to a lesser extent in the endoplasmic reticulum, directly onto protein targets (Ma et al. 2006). Unlike other monosaccharides which form the core elements of carbohydrate structures, fucose is primarily utilized as a terminal modification to alter the properties of cell surface glycans with the distinction between the ABO blood groups being the most prominent example (Becker and Lowe 2003).

The expression of fucosylated glycans has been one of the most observed differences between pluripotent and differentiated cells. Specifically, hPSCs have been shown to express high levels of α1-2 fucosylated glycans on the cell surface compared to differentiated cells (Hasehira et al. 2012; Wang et al. 2011; Satomaa et al. 2009; Liang et al. 2010; Ojima et al. 2015). Increased expression of genes FUT1 and FUT2, the enzymes that catalyze the transfer of fucose in α1-2 linkage, has also been observed in pluripotent cells compared to somatic cells and is likely responsible for the abundant fucosylated glycan expression (Tateno et al. 2011; Wang et al. 2011; Liang et al. 2010; Ojima et al. 2015). These findings contrast with a previous study, however, that indicated that α1-2

fucosylation is not prevalent in mESCs and highlights glycosylation differences between species (Nairn et al. 2012).

Interestingly, stem cell markers SSEA-5, rBC2LCN, and UEA-1 all recognize α1-2 fucosylated glycan structures (Tang et al. 2011; Tateno et al. 2011). Given that these markers were derived from high throughput comparisons of global cell surface glycosylation between PSCs and adult cells suggests that α1-2 fucosylation may be a defining carbohydrate moiety of pluripotency. This raises interesting questions about the importance of fucose and genes FUT1 and FUT2 for maintenance of the pluripotent state, however previous studies using mice deficient in FUT1 and FUT2 have shown that α1-2 fucosylation is not critical for normal development (Domino et al. 2001). To date there have been no published reports confirming these results in human cells so it is unclear whether this is conserved between species. Taken together, these observations indicate that fucosylation is an abundant identifying mark in pluripotent cells that is lost upon differentiation but may not be required for proper function.

Cell surface sialylation is critical for maintenance of pluripotency

N-acetyl-neuraminic acid (Neu5Ac), commonly known as sialic acid, is a nine-carbon sugar with an acidic carboxyl group attached to the anomeric carbon. Similar to fucose, sialic acid is also a terminal modification of complex carbohydrate structures and often serves as a cap on the non-reducing ends of mature N- and O- glycans. Because of this, sialic acid residues are extensively displayed on the surface of mammalian cells with tens of millions residues present on a typical cell (Collins et al. 2004). Sialylation occurs in the Golgi by a family of sialyltransferases which most commonly catalyze reactions in

α2-3 or α2-6 linkage to galactose or α2-8 linkage to sialic acid in oligo or polysialic acid structures (Varki et al. 2009). This process is critical for mammalian development as sialylated glycans coat the surface of cells and perform a multitude of biological functions (Schwarzkopf et al. 2002; Varki et al. 2009; Weidemann et al. 2009). In addition, sialylated glycans on the surface play major roles in the immune system and in host recognition by viruses, which often recognize specific sialic acid linkages to gain entry into cells (Varki and Gagneux 2012).

Recent studies examining global glycan expression have revealed major changes in sialylation as cells transition from pluripotency to differentiated progenitors. Among the findings is the observation that hPSCs highly express sialic acid in the α2-6 linkage on the cell surface (Alisson-Silva et al. 2014; Hasehira et al. 2012; Wang et al. 2011; Satomaa et al. 2009). In contrast, differentiated and somatic cells primarily express α2-3 linked sialic acid on the surface (Hasehira et al. 2012; Wang et al. 2011) to indicate that a dramatic shift in sialylation patterns takes place upon lineage commitment. Similarly, reprogramming of human dermal fibroblasts to iPSCs dramatically reverted the sialic acid expression back to α2-6 linked and suggests that α2-6 sialylated glycans are an integral part of pluripotency (Hasehira et al. 2012). Expression profiles of sialyltransferases matches the pattern of sialylation as ST6Gal1 is highly expressed in PSCs and likely generates the α2-6 linked structures (Tateno et al. 2011; Wang et al. 2015).

Interestingly, new reports have shown that expression of α2-6 sialic acid plays functional roles in pluripotency. Alisson-Silva et al. have demonstrated that enzymatic removal of sialic acid from the cell surface caused cells to spontaneously differentiate to

ectoderm progenitors (Alisson-Silva et al. 2014). In addition, another recent study has shown that loss of ST6Gal1 activity via RNA interference or pharmacological inhibition results in down regulation of the core pluripotency factor OCT4, increased expression of developmental genes, and reduced reprogramming efficiency (Wang et al. 2015). Together, these observations point toward α2-6 sialylation playing major roles in maintaining the pluripotent state. The mechanisms behind these roles are currently unknown but represent major areas of interest moving forward.

Polysialylation is required for germ layer commitment from hPSCs

Polysialic acid (PSA) is an unusual glycosylation reaction that is characterized by repeating α2-8 sialic acid residues on the end of N-glycan chains and can form extended polymer structures hundreds of residues in length. Two enzymes, ST8SIA2 and ST8SIA4, are capable of synthesizing PSA in the Golgi and show a high degree of substrate specificity as PSA has only been observed on a small subset of proteins with the neural cell adhesion molecule (NCAM) being the most common and most studied (Angata and Fukuda 2003; Hildebrandt et al. 2009). Polysialylation is critical for development and has been shown to play extensive roles in cell signaling, migration, and neurogenesis (Colley et al. 2014; Angata and Fukuda 2010; Fujimoto 2001; Ono et al. 2012).

Although the majority of studies regarding polysialylation have focused on later developmental pathways, our research group has found that PSA is intricately involved in the earliest cell fate decisions. PSCs are epithelial cells and do not express polysialylated glycans on the surface, however PSA expression dramatically increases as cells

differentiate to progenitors of all three germ layers (See Chapter 3). This effect is derived from cell type specific expression of polysialyltransferases in which ST8SIA4 is prevalent in endoderm and mesoderm while ST8SIA2 is increased in ectoderm. Interestingly, the expression of ST8SIA4 is regulated by well-known developmental pathways and is required for proper differentiation (See Chapter 3). These observations indicate that polysialic acid expression is critically important for germ layer specification in the embryo. The mechanism of this requirement is currently unknown but likely involves polysialylated glycans playing a role in cell signaling on the surface similar to previous reports (Kiermaier et al. 2016; Eggers et al. 2011; Li et al. 2011).

INDUCED PLURIPOTENT STEM CELLS AS MODELS FOR GLYCOSYLATION DEFECTS

Over the past decade, induced pluripotent stem cells (iPSCs) have proven valuable for modeling a wide array of diseases and genetic disorders (Rashid et al. 2010; Cherry and Daley 2013; Marchetto et al. 2011; Cramer and MacLaren 2013). The use of iPSCs is an expanding field of research that allows for the study of cellular and molecular mechanisms of diseases using cells directly from a patient. The differentiation capabilities of iPSCs allow for these mechanisms to be elucidated in a cell-type specific manner and their proliferative capacity permits unlimited expansion of cells from disease samples that are very rare and difficult to obtain. Additionally, the ability to manipulate iPSCs using emerging genome editing technologies like CRISPR-Cas9 provides opportunities to determine the direct impact of a patient's genetic profile on disease pathology and potentially correct the disorder *in vitro*. All of these benefits of iPSC

technology make them an ideal tool for studying diseases, including those that are caused by defects of glycosylation.

Congenital disorders of glycosylation (CDGs) and lysosomal storage diseases (LSDs) are two broad classes of disorders that are caused by either defects in glycan biosynthesis or degradation, respectively (Varki et al. 2009). These disorders occur rarely and are not well understood but advances in technology have led to the discovery of over 100 glycosylation related disorders in the last decade (Freeze et al. 2014). Some challenges in studying glycosylation defects lie in the fact that glycan expression varies between tissues, therefore a mutation in a certain biosynthetic enzyme may have a greater affect in one cell type than in another. Another challenge is that the genetic mutations that cause a particular glycosylation disorder can vary from patient to patient. iPSCs provide the ability to study multiple cell types from a single individual while allowing for consistency with regards to genetic background. This approach allows researchers to investigate the cell-type specific changes that may result in the disorder's pathology. The use of iPSCs as disease models is relatively new technology and so the number of published studies modeling either CDGs or LSDs is limited. This represents an area of new opportunities in the field of glycosylation research that should be explored in the coming years.

Disease modeling of PMM2-CDG with iPSCs

The most common and most studied neurological CDG is phosphomannomutase 2 deficiency, or PMM2-CDG (Jaeken 2013). PMM2-CDG, also known as CDG-1a, is caused by a mutation in *PMM2* that codes for the enzyme phophomannomutase, which

catalyzes the conversion of mannose-6-phosphate to mannose-1-phosphate prior to synthesis of GDP-mannose. As a result, mutations in PMM2 limits the amount of GDP-mannose available and causes a deficiency in the supply of lipid-linked oligosaccharides (LLOs) available for the production of N-linked glycans (Westphal et al. 2001). Complete loss of N-linked glycosylation is lethal; therefore, studies of PMM2-CDG cannot be done using complete knockout of the gene (Varki et al. 2009). There are several animal models of the disease, including drosophila, zebrafish and mouse (Cline et al. 2012; Parkinson et al. 2016; Schneider et al. 2011), however the hypomorphic mouse model exhibits embryonic lethality unless the pregnant dam is fed mannose supplementation (Schneider et al. 2011; Thiel et al. 2006).

To alleviate these technical concerns, multiple research groups have generated iPSCs to study PMM2-CDGs. Losfeld et al. utilized iPSCs generated from two PMM2-CDG patients to test a fluorescent marker for changes in N-glycan site occupancy. Their results indicate that the PMM2 iPSCs express similar levels of decreased glycosylation as the patient fibroblasts, suggesting that the iPSCs accurately reflect the disease phenotype *in vitro* (Losfeld et al. 2012). More recently, Thiesler et al. generated iPSC lines from PMM2-patient fibroblasts to characterize how PMM2-CDG affects stem cell glycosylation and pluripotency. Their results show that the iPSCs exhibit decreased PMM2 activity like the patient fibroblasts and confirm previous studies that report reduced N-glycan site occupancy in PMM2 patient cells, but no change in the pattern of N-glycan expression (Thiesler et al. 2016; Freeze 2006). Interestingly, decreased PMM2 activity in the iPSCs yielded up to 40% reduction in high mannose glycan expression compared to WT PSCs but this reduction did not affect the cells ability to differentiate to

all three germ layers (Thiesler et al. 2016). PMM2-iPSCs also exhibited similar patterns of sialylation to WT PSCs as N-glycans were α2-6 sialylated and not α2-3 (Thiesler et al. 2016).

Disease modeling of lysosomal storage disorders with iPSCs

Another emerging area of iPSC disease modeling involves lysosomal storage disorders (LSDs) including Hurler syndrome (Tolar et al. 2011), Pompe disease (Higuchi et al. 2014; Huang et al. 2011; Raval et al. 2015), Gaucher disease (Panicker et al. 2012), Fabry disease (Kawagoe et al. 2013), and Niemann-Pick Type C (Maetzel et al. 2014). Each of these storage disorders result from mutations in proteins that facilitate the degradation of particular glycan products. Absence of these proteins results in a build-up of glycan waste products and leads to significant developmental and physiological effects (Grabowski 2012). Due to the variance of glycan expression between cell lineages, storage disorders often exhibit cell-type specific phenotypes that can be elucidated by differentiating iPSCs towards the most afflicted cell-type (Varki et al. 2009).

Using this approach, an investigation of Pompe disease using patient-derived iPSCs recently revealed a novel glycosylation deficit in cardiomyocytes that may contribute to Pompe cardiomyopathy (Raval et al. 2015). This finding is significant because it provides a potential mechanism to explain how accumulation of glycogen in lysosomes, a hallmark feature of Pompe disease, results in physiological effects and may highlight new therapeutic targets. Similarly, Maetzel et al. utilized iPSCs to study the effects of lysosomal cholesterol accumulation in hepatic and neural cells from Niemann-Pick Type C (NPC) disease patients (Maetzel et al. 2014). In their report, they found that

defects in autophagic flux were associated with mutations in the NPC1 gene and identified an autophagy-inducing compound via small molecule screen that led to increased cell viability in NPC1 deficient cells. These recent works, among others, highlight the immense potential of iPSCs as a model system for *in vitro* disease modeling of CDGs.

Challenges of modeling CDGs and LSDs with iPSCs:

Despite their utility, multiple research groups have discussed significant challenges associated with using iPSCs from patient fibroblasts for modeling CDGs (Thiesler et al. 2016; Huang et al. 2011). To date, the biggest issue lies in reprogramming efficiency. Thiesler et al. reported 0.008% efficiency of reprogramming PMM2-fibroblasts using standard methods, much lower than 0.01-1.0% rates typically achieved using wild type human fibroblasts. Similarly, efforts to generate iPSCs from Pompe disease patient fibroblasts were unsuccessful without temporarily rescuing the defective enzyme, acid alpha-glucosidase (Huang et al. 2011). This effect is likely associated with perturbations of glycosylation pathways caused by the mutated glyco-enzymes and suggests that proper glycosylation is important for reprogramming. The degree of CDG reprogramming difficulty likely varies depending on the affected gene and its significance in pluripotency.

DISCUSSION

The global expression of glycans on the cell surface appears to be a very dynamic process that undergoes significant changes as cells develop from PSCs to defined adult lineages.

Among the most prominent features observed in pluripotent cells is the unusual expression of high mannose N-glycan structures in stark contrast with adult tissues that express a higher percentage of complex and hybrid N-glycans. These results suggest that the N-glycan processing pathway is severely down-regulated in pluripotent cells as removal of terminal glucose and trimming of mannose represents the first steps to forming more complex structures (Varki et al. 2009). It also appears that pluripotent cells favor a simpler N-glycome compared to adult cells and may indicate that PSCs adopt a "blank slate" that is to be specialized upon commitment to defined lineages. Alternatively, expression of high mannose glycans could be playing functional roles important for maintenance of the pluripotent state through a signaling or cell adhesion mechanism. The significance of this unusual expression of high mannose carbohydrates on the surface of PSCs is yet to be determined and should be the subject of future investigation.

Additionally, terminal fucosylation and sialylation expression patterns show interesting trends in pluripotency and differentiation. The high expression of α1-2 fucosylated glycans on the surface and their recognition by PSC specific reagents points toward fucosylation as a defining mark of pluripotency. It is unknown whether this expression is playing a role in the maintenance of pluripotency in hPSCs although a fucosyltransferase deficient mouse model suggests that it may be dispensable (Domino et al. 2001). Sialylation, however, is perhaps the most interesting case of surface glycosylation in human embryonic development as there is abundant expression of specific sialic acid linkages for different cell types; pluripotent cells express α2-6 sialylated glycans, the initial differentiation stages express α2-8 polysialylated NCAM,

and defined adult lineages display α2-3 surface sialylation. Recent studies have illustrated that these sialylation patterns play functional roles in regulating both pluripotency and early differentiation processes (Wang et al. 2015; See Chapter 3). The mechanisms through which sialylation is acting in these processes is unknown but represents an area for future study.

The use of iPSCs as a model for CDGs and LSDs is an emerging field of research that offers great promise in elucidating the mechanisms of glycosylation disorders. Although limited in number, the iPSC disease modeling studies to date have yielded useful information in the pathogenesis of multiple diseases and provide a useful platform for continued research. Additionally, the use of CRISPR-Cas9 genome editing tools makes possible the ability to correct a genetic defect in these cells *in vitro* and perhaps serve as a therapeutic option for a subset of disorders in the future.

REFERENCES

Alisson-Silva F, de Carvalho Rodrigues D, Vairo L, Asensi KD, Vasconcelos-dos-Santos A, Mantuano NR, Dias WB, Rondinelli E, Goldenberg RCDS, Urmenyi TP, et al. 2014. Evidences for the involvement of cell surface glycans in stem cell pluripotency and differentiation. *Glycobiology* **24**: 458–468.

An HJ, Gip P, Kim J, Wu S, Park KW, McVaugh CT, Schaffer DV, Bertozzi CR, Lebrilla CB. 2012. Extensive Determination of Glycan Heterogeneity Reveals an Unusual Abundance of High Mannose Glycans in Enriched Plasma Membranes of Human Embryonic Stem Cells. *Molecular & Cellular Proteomics* **11**: M111.010660–M111.010660.

Andrews PW, Banting G, Damjanov I, Arnaud D, Avner P. 1984. Three monoclonal antibodies defining distinct differentiation antigens associated with different high molecular weight polypeptides on the surface of human embryonal carcinoma cells. *Hybridoma* **3**: 347–361.

Angata K, Fukuda M. 2003. Polysialyltransferases: major players in polysialic acid synthesis on the neural cell adhesion molecule. *Biochimie* **85**: 195–206.

Angata K, Fukuda M. 2010. *Roles of Polysialic Acid in Migration and Differentiation of Neural Stem Cells*. 1st ed. Elsevier Inc.

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.

Becker DJ, Lowe JB. 2003. Fucose: biosynthesis and biological function in mammals. *Glycobiology* **13**: 41R–53R.

Brimble SN, Sherrer ES, Uhl EW, Wang E, Kelly S, Merrill AH Jr., Robins AJ, Schulz TC. 2006. The Cell Surface Glycosphingolipids SSEA-3 and SSEA-4 Are Not Essential for Human ESC Pluripotency. *Stem Cells* **25**: 54–62.

Cherry ABC, Daley GQ. 2013. Reprogrammed Cells for Disease Modeling and Regenerative Medicine. *Annu Rev Med* **64**: 277–290.

Chu CS, Niñonuevo MR, Clowers BH, Perkins PD, An HJ, Yin H, Killeen K, Miyamoto S, Grimm R, Lebrilla CB. 2009. Profile of native N-linked glycan structures from human serum using high performance liquid chromatography on a microfluidic chip and time-of-flight mass spectrometry. *Proteomics* **9**: 1939–1951.

Cline A, Gao N, Flanagan-Steet H, Sharma V, Rosa S, Sonon R, Azadi P, Sadler KC, Freeze HH, Lehrman MA, et al. 2012. A zebrafish model of PMM2-CDG reveals altered neurogenesis and a substrate-accumulation mechanism for N-linked glycosylation deficiency. *Molecular Biology of the Cell* **23**: 4175–4187.

Colley KJ, Kitajima K, Sato C. 2014. Polysialic acid: Biosynthesis, novel functions and applications. *Critical Reviews in Biochemistry and Molecular Biology* **49**: 498–532.

Collins BE, Blixt O, DeSieno AR, Bovin N, Marth JD, Paulson JC. 2004. Masking of CD22 by cis ligands does not prevent redistribution of CD22 to sites of cell contact. *Proc Natl Acad Sci USA* **101**: 6104–6109.

Cramer AO, MacLaren RE. 2013. Translating induced pluripotent stem cells from bench to bedside: application to retinal diseases. *Curr Gene Ther* **13**: 139–151.

Damjanov I, Fox N, Knowles BB, Solter D, Lange PH, Fraley EE. 1982. Immunohistochemical localization of murine stage-specific embryonic antigens in human testicular germ cell tumors. *The American Journal of Pathology* **108**: 225–230.

Domino SE, Zhang L, Gillespie PJ, Saunders TL, Lowe JB. 2001. Deficiency of Reproductive Tract alpha (1,2)Fucosylated Glycans and Normal Fertility in Mice with Targeted Deletions of the FUT1 or FUT2 alpha (1,2)Fucosyltransferase Locus. *Molecular and Cellular Biology* **21**: 8336–8345.

Eggers K, Werneburg S, Schertzinger A, Abeln M, Schiff M, Scharenberg MA, Burkhardt H, Muhlenhoff M, Hildebrandt H. 2011. Polysialic acid controls NCAM signals at cell-cell contacts to regulate focal adhesion independent from FGF receptor activity. *Journal of Cell Science* **124**: 3279–3291.

Freeze HH. 2006. Genetic defects in the human glycome. *Nat Rev Genet* **7**: 537–551.

Freeze HH, Chong JX, Bamshad MJ, Ng BG. 2014. Solving Glycosylation Disorders:Fundamental Approaches Reveal Complicated Pathways. *The American Journal of Human Genetics* **94**: 161–175.

Fujimoto I. 2001. Regulation of Cell Adhesion by Polysialic Acid. EFFECTS ON CADHERIN, IMMUNOGLOBULIN CELL ADHESION MOLECULE, AND INTEGRIN FUNCTION AND INDEPENDENCE FROM NEURAL CELL

ADHESION MOLECULE BINDING OR SIGNALING ACTIVITY. *Journal of Biological Chemistry* **276**: 31745–31751.

Fujitani N, Furukawa J-I, 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. *Proceedings of the National Academy of Sciences* **110**: 2105–2110.

Gooi HC, Feizi T, Kapadia A, Knowles BB, Solter D, Evans MJ. 1981. Stage-specific embryonic antigen involves alpha 1 goes to 3 fucosylated type 2 blood group chains. *Nature* **292**: 156–158.

Grabowski GA. 2012. Gaucher disease and other storage disorders. *Hematology Am Soc Hematol Educ Program* **2012**: 13–18.

Hasehira K, Tateno H, Onuma Y, Ito Y, Asashima M, Hirabayashi J. 2012. Structural and Quantitative Evidence for Dynamic Glycome Shift on Production of Induced Pluripotent Stem Cells. *Molecular & Cellular Proteomics* **11**: 1913–1923.

Higuchi T, Kawagoe S, Otsu M, Shimada Y, Kobayashi H, Hirayama R, Eto K, Ida H, Ohashi T, Nakauchi H, et al. 2014. The generation of induced pluripotent stem cells (iPSCs) from patients with infantile and late-onset types of Pompe disease and the effects of treatment with acid-α-glucosidase in Pompe's iPSCs. *Mol Genet Metab* **112**: 44–48.

Hildebrandt H, Mühlenhoff M, Gerardy-Schahn R. 2009. Polysialylation of NCAM. In ... *Adhesion Molecule NCAM*, Vol. 663 of *Advances in Experimental Medicine and Biology*, pp. 95–109, Springer New York, New York, NY.

Huang H-P, Chen P-H, Hwu W-L, Chuang C-Y, Chien Y-H, Stone L, Chien C-L, Li L-T, Chiang S-C, Chen H-F, et al. 2011. Human Pompe disease-induced pluripotent stem cells for pathogenesis modeling, drug testing and disease marker identification. *Human Molecular Genetics* **20**: 4851–4864.

Jaeken J. 2013. Chapter 179 - Congenital disorders of glycosylation. In *Handbook of Clinical Neurology* (ed. M.L.A.H.B.S. Olivier Dulac), Vol. Volume 113 of *Handbook of Clinical Neurology*, pp. 1737–1743, Elsevier.

Kannagi R, Cochran NA, Ishigami F, Hakomori S. 1983. Stage-specific embryonic antigens (SSEA-3 and-4) are epitopes of a unique globo-series ganglioside isolated from human teratocarcinoma cells. *EMBO J*.

Kawagoe S, Higuchi T, Otaka M, Shimada Y, Kobayashi H, Ida H, Ohashi T, Okano HJ, Nakanishi M, Eto Y. 2013. Morphological features of iPS cells generated from Fabry disease skin fibroblasts using Sendai virus vector (SeVdp). *Mol Genet Metab* **109**: 386–389.

Kiermaier E, Moussion C, Veldkamp CT, Gerardy-Schahn R, de Vries I, Williams LG, Chaffee GR, Phillips AJ, Freiberger F, Imre R, et al. 2016. Polysialylation controls dendritic cell trafficking by regulating chemokine recognition. *Science* **351**: 186–190.

Li J, Dai G, Cheng YB, Qi X, Geng MY. 2011. Polysialylation promotes neural cell adhesion molecule-mediated cell migration in a fibroblast growth factor receptor-dependent manner, but independent of adhesion capability. *Glycobiology* **21**: 1010–1018.

Liang YJ, Kuo HH, Lin CH, Chen YY. 2010. Switching of the core structures of glycosphingolipids from globo-and lacto-to ganglio-series upon human embryonic stem cell differentiation.

Losfeld M-E, Soncin F, Ng BG, Singec I, Freeze HH. 2012. A sensitive green fluorescent protein biomarker of N-glycosylation site occupancy. *FASEB J* **26**: 4210–4217.

Ma B, Simala-Grant JL, Taylor DE. 2006. Fucosylation in prokaryotes and eukaryotes. *Glycobiology* **16**: 158R–184R.

Maetzel D, Sarkar S, Wang H, Abi-Mosleh L, Xu P, Cheng AW, Gao Q, Mitalipova M, Jaenisch R. 2014. Genetic and chemical correction of cholesterol accumulation and impaired autophagy in hepatic and neural cells derived from Niemann-Pick Type C patient-specific iPS cells. *Stem Cell Reports* **2**: 866–880.

Marchetto MC, Brennand KJ, Boyer LF, Gage FH. 2011. Induced pluripotent stem cells (iPSCs) and neurological disease modeling: progress and promises. *Human Molecular Genetics* **20**: R109–R115.

Nairn AV, Aoki K, Rosa dela M, Porterfield M, Lim JM, Kulik M, Pierce JM, 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. *Journal of Biological Chemistry* **287**: 37835–37856.

Natunen S, Satomaa T, Pitkanen V, Salo H, Mikkola M, Natunen J, Otonkoski T, Valmu L. 2011. The binding specificity of the marker antibodies Tra-1-60 and Tra-1-81 reveals a novel pluripotency-associated type 1 lactosamine epitope. *Glycobiology* **21**: 1125–1130.

Ojima T, Shibata E, Saito S, Toyoda M, Nakajima H, Yamazaki-Inoue M, Miyagawa Y, Kiyokawa N, Fujimoto J-I, Sato T, et al. 2015. Glycolipid dynamics in generationand differentiation of inducedpluripotent stem cells. *Nature Publishing Group* 1–13.

Ono S, Hane M, Kitajima K, Sato C. 2012. Novel Regulation of Fibroblast Growth Factor 2 (FGF2)-mediated Cell Growth by Polysialic Acid. *Journal of Biological Chemistry* **287**: 3710–3722.

Panicker LM, Miller D, Park TS, Patel B, Azevedo JL, Awad O, Masood MA, Veenstra TD, Goldin E, Stubblefield BK, et al. 2012. Induced pluripotent stem cell model recapitulates pathologic hallmarks of Gaucher disease. *Proceedings of the National Academy of Sciences* **109**: 18054–18059.

Parkinson WM, Dookwah M, Dear ML, Gatto CL, Aoki K, Tiemeyer M, Broadie K. 2016. Neurological roles for phosphomannomutase type 2 in a new Drosophila congenital disorder of glycosylation disease model. *Disease Models & Mechanisms*.

Pennington JE, Rastan S, Roelcke D, Feizi T. 1985. Saccharide structures of the mouse embryo during the first eight days of development. Inferences from

immunocytochemical studies using monoclonal antibodies in conjunction with glycosidases. *J Embryol Exp Morphol* **90**: 335–361.

Rashid ST, Corbineau S, Hannan N, Marciniak SJ, Miranda E, Alexander G, Huang-Doran I, Griffin J, Ahrlund-Richter L, Skepper J, et al. 2010. Modeling inherited metabolic disorders of the liver using human induced pluripotent stem cells. *Journal of clinical investigation* **120**: 3127–3136.

Raval KK, Tao R, White BE, De Lange WJ, Koonce CH, Yu J, Kishnani PS, Thomson JA, Mosher DF, Ralphe JC, et al. 2015. Pompe disease results in a Golgi-based glycosylation deficit in human induced pluripotent stem cell-derived cardiomyocytes. *J Biol Chem* **290**: 3121–3136.

Satomaa T, Heiskanen A, Mikkola M, Olsson C, Blomqvist M, Tiittanen M, Jaatinen T, Aitio O, Olonen A, Helin J, et al. 2009. The N-glycome of human embryonic stem cells. *BMC Cell Biol* **10**: 42.

Schneider A, Thiel C, Rindermann J, DeRossi C, Popovici D, Hoffmann GF, Gröne H-J, Körner C. 2011. Successful prenatal mannose treatment for congenital disorder of glycosylation-Ia in mice. *Nature Medicine* **18**: 71–73.

Schopperle WM, DeWolf WC. 2006. The TRA-1-60 and TRA-1-81 Human Pluripotent Stem Cell Markers Are Expressed on Podocalyxin in Embryonal Carcinoma. *Stem Cells* **25**: 723–730.

Schwarzkopf M, Knobeloch K-P, Rohde E, Hinderlich S, Wiechens N, Lucka L, Horak I, Reutter W, Horstkorte R. 2002. Sialylation is essential for early development in mice. *Proceedings of the National Academy of Sciences* **99**: 5267–5270.

Solter D, Knowles BB. 1978. Monoclonal antibody defining a stage-specific mouse embryonic antigen (SSEA-1). *... of the National Academy of Sciences*.

SulAk O, Cioci G, Delia M, Lahmann M, Varrot A, Imberty A, WimmerovA M. 2010. A TNF-like Trimeric Lectin Domainfrom Burkholderia cenocepacia with Specificityfor Fucosylated Human Histo-Blood Group Antigens. *Structure/Folding and Design* **18**: 59–72.

Tang C, Lee AS, Volkmer J-P, Sahoo D, Nag D, Mosley AR, Inlay MA, Ardehali R, Chavez SL, Pera RR, et al. 2011. An antibody against SSEA-5 glycan on human pluripotent stem cells enables removal of teratoma-forming cells. *Nature biotechnology* **29**: 829–834.

Tao SC, Li Y, Zhou J, Qian J, Schnaar RL, Zhang Y, Goldstein IJ, Zhu H, Schneck JP. 2008. Lectin microarrays identify cell-specific and functionally significant cell surface glycan markers. *Glycobiology* **18**: 761–769.

Tateno H, Matsushima A, Hiemori K, Onuma Y, Ito Y, Hasehira K, Nishimura K, Ohtaka M, Takayasu S, Nakanishi M, et al. 2013. Podocalyxin Is a Glycoprotein Ligand of the Human Pluripotent Stem Cell-Specific Probe rBC2LCN. *Stem Cells Translational Medicine* **2**: 265–273.

Tateno H, Onuma Y, Ito Y, Minoshima F, Saito S, Shimizu M, Aiki Y, Asashima M, Hirabayashi J. 2015. Stem Cell Reports. *Stem Cell Reports* **4**: 811–820.

Tateno H, Toyota M, Saito S, Onuma Y, Ito Y, Hiemori K, Fukumura M, Matsushima A, Nakanishi M, Ohnuma K, et al. 2011. Glycome Diagnosis of Human Induced Pluripotent Stem Cells Using Lectin Microarray. *Journal of Biological Chemistry* **286**: 20345–20353.

Thiel C, Lübke T, Matthijs G, Figura von K, Körner C. 2006. Targeted disruption of the mouse phosphomannomutase 2 gene causes early embryonic lethality. *Molecular and Cellular Biology* **26**: 5615–5620.

Thiesler CT, Cajic S, Hoffmann D, Thiel C, van Diepen L, Hennig R, Sgodda M, Weißmann R, Reichl U, Steinemann D, et al. 2016. Glycomic characterization of induced pluripotent stem cells derived from a patient suffering from phosphomannomutase 2 congenital disorder of glycosylation. *Molecular & Cellular Proteomics*.

Tolar J, Park I-H, Xia L, Lees CJ, Peacock B, Webber B, McElmurry RT, Eide CR, Orchard PJ, Kyba M, et al. 2011. Hematopoietic differentiation of induced pluripotent stem cells from patients with mucopolysaccharidosis type I (Hurler syndrome). *Blood* **117**: 839–847.

Varki A, Gagneux P. 2012. Multifarious roles of sialic acids in immunity. *Annals of the New York Academy of Sciences* **1253**: 16–36.

Varki A, RD C, JD E, HH F, P S, CR B, GW H, ME E. 2009. *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press.

Wang Y-C, Nakagawa M, Garitaonandia I, Slavin I, Altun G, Lacharite RM, Nazor KL, Tran HT, Lynch CL, Leonardo TR, et al. 2011. Specific lectin biomarkers for isolation of human pluripotent stem cells identified through array-based glycomic analysis. *Cell Research* **21**: 1551–1563.

Wang Y-C, Stein JW, Lynch CL, Tran HT, Lee C-Y, Coleman R, Hatch A, Antontsev VG, Chy HS, O'Brien CM, et al. 2015. Glycosyltransferase ST6GAL1 contributes to the regulation of pluripotency in human pluripotent stem cells. *Scientific reports* 1–13.

Weidemann W, Klukas C, Klein A, Simm A, Schreiber F, Horstkorte R. 2009. Lessons from GNE-deficient embryonic stem cells: sialic acid biosynthesis is involved in proliferation and gene expression. *Glycobiology* **20**: 107–117.

Westphal V, Peterson S, Patterson M, Tournay A, Blumenthal A, Treacy EP, Freeze HH. 2001. Functional significance of PMM2 mutations in mildly affected patients with congenital disorders of glycosylation Ia. *Genet Med* **3**: 393–398.

CHAPTER 3

ST8SIA4-DEPENDENT POLYSIALYLATION IS PART OF A DEVELOPMENTAL PROGRAM REQUIRED FOR GERM LAYER FORMATION FROM HUMAN PLURIPOTENT STEM CELLS

Berger RP, Sun YH, Kulik M, Lee JK, Nairn AV, Moremen KW, Pierce M, Dalton S. 2016. Accepted by *Stem Cells.* Reprinted here with permission of the publisher.

ABSTRACT

Polysialic acid (PSA) is a carbohydrate polymer of repeating α-2,8 sialic acid residues that decorates multiple targets, including neural cell adhesion molecule (NCAM). PST and STX encode the two enzymes responsible for PSA modification of target proteins in mammalian cells, but despite widespread polysialylation in embryonic development, the majority of studies have focused strictly on the role of PSA in neurogenesis. Using human pluripotent stem cells (hPSCs), we have revisited the developmental role of PST and STX and show that early progenitors of the three embryonic germ layers are polysialylated on their cell surface. Changes in polysialylation can be attributed to lineage-specific expression of polysialyltransferase genes; PST is elevated in endoderm and mesoderm, while STX is elevated in ectoderm. In hPSCs, PST and STX genes are epigenetically marked by overlapping domains of H3K27 and H3K4 trimethylation, indicating that they are held in a 'developmentally-primed' state. Activation of PST transcription during early mesendoderm differentiation is under control of the T-Goosecoid transcription factor network, a key regulatory axis required for early cell fate decisions in the vertebrate embryo. This establishes polysialyltransferase genes as part of a developmental program associated with germ layer establishment. Finally, we show by shRNA knockdown and CRISPR-Cas9 genome editing that PST-dependent cell surface polysialylation is essential for endoderm specification. This is the first report to demonstrate a role for a glycosyltransferase in hPSC lineage specification.

INTRODUCTION

Glycosylation is a carbohydrate modification that increases structural diversity and function of proteins and lipids. The significance of glycosylation in mammalian cell biology is emphasized by its contribution to the glycocalyx- a dense, glycan-rich structure on the outer surface of cells. The glycocalyx is composed of a complex array of glycans that vary in a cell type-specific manner but deciphering specific roles for carbohydrates in the function of specific cell surface targets has been problematic. The broad impact of cell surface glycans on human health, however, has become increasingly important due to strong clinical links between cell surface glycosylation defects and human disease (Varki et al. 2009). From a developmental perspective, glycosylation is an integral and dynamic process associated with cell migration, cell fate specification, proliferation, and organogenesis in the embryo (Haltiwanger and Lowe 2004). Specific roles for glycans in these early developmental processes are, however, poorly understood.

One aspect of embryonic development that has been well documented in mice, but limited in humans, is the expression of polysialic acid (PSA) on the surface of developing cells (Rutishauser 2008; Mühlenhoff et al. 2013). PSA is a linear glycan homopolymer composed of repeating *N*-acetylnueraminic acid (Neu5Ac or sialic acid) residues linked by α-2,8 glycosidic bonds (Mühlenhoff et al. 2013). Two polysialyltransferases, ST8SIA2 (STX) and ST8SIA4 (PST), reside in the Golgi and add polysialic acid chains to target proteins or to themselves through autopolysialylation (Close and Colley 1998). Polysialylated epitopes are elevated during the development of many tissues in mice including heart, kidney, pancreas and brain with cell type-specific expression of PST and STX (Finne 1982; Phillips et al. 1997; Hildebrandt et al. 1998;

Lackie et al. 1994) . Little is known, however, about the function and regulation of polysialylation during early stages of development.

Although neuropilin-2 and SynCAM1 have been reported to be polysialylated (Curreli et al. 2007; Galuska et al. 2010), neural cell adhesion molecule (NCAM) is the most abundant and best characterized substrate for PSA modification (Mühlenhoff et al. 2013). NCAM is an immunoglobulin (Ig) adhesion protein residing on the cell surface and plays roles in the formation of intercellular contacts through homophilic and heterophilic binding, in addition to intracellular signaling (Mühlenhoff et al. 2013). NCAM is polysialylated by either PST or STX on the fifth Ig domain, forming a polymer chain typically between 60-90 sialic acid residues (Galuska et al. 2007; Close et al. 2003). The resulting modification of NCAM (PSA-NCAM) is a large, highly-negatively charged side chain that disrupts binding interactions between neighboring PSA-NCAM molecules to enhance cellular motility (Johnson et al. 2005; Li et al. 2011; Rutishauser et al. 1985). This property is utilized in the developing mouse nervous system in which PSA-NCAM is instrumental in neuronal migration, neurite outgrowth and synaptic plasticity (Hu et al. 1996; Landmesser et al. 1990). Evidence also suggests that PSA impacts NCAM-directed signaling events (Ono et al. 2012; Eggers et al. 2011; Seidenfaden et al. 2006). PSA is widely expressed on the surface of metastatic cancer cells including neuroblastoma, pancreatic ductal adenocarcinomas, small cell lung carcinoma and Wilms' tumor (Livingston et al. 1988; Schreiber et al. 2008; Tanaka et al. 2000; Roth et al. 1988), highlighting its clinical significance. Consistent with this, the two polysialyltransferases (PST and STX) are expressed at elevated levels in tumor cells and the resulting polysialylation is an adverse prognosis factor (Amoureux et al. 2010).

Understanding the role of cell surface glycosylation and how it is regulated during human development is very limited. In this report, we address this by identifying broad roles for polysialylation in cell fate specification of human pluripotent stem cells and provide the first report describing a role for a specific glycosyltransferase in hPSC lineage specification.

RESULTS

Early PSC-derived progenitors are polysialylated on their cell surface

hPSCs are known to express characteristic glycans on their cell surface such as SSEA3 but do not present PSA (Fig. 3.1A). However, when differentiated towards early progenitor cells of the three embryonic germ layers, WA09 hPSCs transition to a PSA^+ state (Fig. 3.1). This observation is apparent in multiple hPSC lines (WA01, WA07, TE03) as well as mESCs (R1) (Supplemental Fig. 3.1). By immuno-blotting of definitive endoderm (DE) lysates, PSA-associated immuno-reactive bands are observed in the ~150-250 kDa range (Fig. 3.1B), characteristic of polysialylation. Within 2-4 days of Activin A-directed differentiation, $CXCR4^+$ DE acquires cell surface (Fig. 3.1A, C, E, H and Supplemental Fig. 3.1B,C) and Golgi-associated, α-mannosidase II co-localized PSA^+ epitopes (Fig. 3.1H). $T^+/Isl1^+$ early mesoderm progenitors generated by BMP4 treatment also acquire cell surface and intracellular polysialylated epitopes within 4 days (Fig. 3.1C, F). Finally, early $AP2^+$ ectoderm progenitors also become cell surface polysialylated but in contrast to mesoderm and endoderm, Golgi-associated ($Giantin^+$) reactivity is less obvious (Fig. 3.1C, G, I). In all three germ layers, cell surface

polysialylation is acquired as cells lose pluripotency markers, such as NANOG and SSEA3 (Fig. 3.1A, F, G).

Polysialic acid can be attached to target proteins by two enzymes, PST and STX. To establish if polysialylation is regulated at the transcriptional level and whether PST or STX is involved in polysialylation during germ layer induction, transcript levels for these two genes were determined. This analysis showed germ layer specific differences in PST/STX transcript accumulation at the time when polysialylation increased (Fig. 3.2A-C). During early endoderm and mesoderm differentiation, PST transcript levels increase significantly while STX mRNA levels remain low (Fig. 3.2A,B and Supplemental Fig. 3.1D). In contrast, STX transcript levels increase during early ectoderm differentiation, while PST mRNA remains low (Fig. 3.2C). NCAM transcript levels increase during the early stages of differentiation in all three germ layers (Fig. 3.2A-C). To confirm that PST was responsible for polysialylation in DE, shRNA was used to target PST transcripts. This resulted in ~85-90% reduction of PST transcript levels and loss of cell surface and Golgi-associated polysialylation (Fig. 3.2D and Supplemental Fig. 3.2).

We next addressed the issue of whether NCAM is a major target of polysialylation in endoderm and mesoderm differentiation. This is suggested by flow cytometry showing that PSA^+ cells are also $NCAM^+$ (Fig. 3.2E). This possibility is supported by a series of shRNA knockdown experiments showing that loss of NCAM eliminates cell surface polysialylation, but not Golgi-associated PSA reactivity (Fig. 3.2F,G). Next, NCAM was shown to be polysialylated in multiple PSC-derived cell types of endoderm, mesoderm and ectoderm origin by immunoprecipitation-immunoblot experiments using whole cell lysates (Fig. 3.2H,I). Finally, although ectopic PST

expression in PSCs increased Golgi-dependent polysialylation, presumably through autocatalysis (see ref (Close and Colley 1998)), its co-expression with NCAM was required for decorating the cell surface with polysialylated substrates (Supplemental Fig. 3.3). Taken together, these data indicate that NCAM is the major acceptor of polysialic acid on the surface of hPSC-derived progenitor cells.

Developmentally-regulated transcription factors control PST transcription and polysialylation

Overlapping domains of activating (H3K4me3) and repressive (H3K27me3) histone modifications mark developmental genes in pluripotent cells (Bernstein et al. 2006). These epigenetic marks hold developmental genes in a 'poised' state, so they can be rapidly activated under differentiation conditions. Genes encoding components of the glycomics network have not been implicated in this developmental pathway before; however, based on our previous observations, it seemed reasonable that PST, STX, and NCAM could be part of this broad developmental program. ChIP-seq data from the UCSC genome database indicate bivalent regions of H3K4me3 and H3K27me3 marks overlapping the transcription start sites of PST, STX, and NCAM genes in WA01 hPSCs (Fig. 3.3A-C) (Rosenbloom et al. 2012). To confirm this in WA09 hPSCs and to determine whether these histone marks change during differentiation, quantitative chromatin immunoprecipitation (qChIP) assays were performed. In agreement with ChIP-seq data, qChIP of the PST, STX, and NCAM promoters shows significant enrichment for both bivalent histone marks (Supplemental Fig. 3.4), confirming them to be bivalent genes in hPSCs. During early germ layer formation, bivalent domains resolve

in a pattern consistent with the transcription patterns for each gene as cells differentiate into endoderm, mesoderm and ectoderm (Supplemental Fig. 3.4). These data indicate that genes implicated in cell surface polysialylation during germ layer formation are regulated by epigenetic mechanisms known to control a broader group of developmental regulators.

The kinetics of PST transcript accumulation in endoderm differentiation follows that of Brachyury (T) and Goosecoid (GSC) and coincides with the endoderm markers SOX17 and FOXA2 (Fig. 3.2A,B and Fig. 3.3D). Since the polysialyltransferase genes are bivalently marked, they are likely to be part of a broader developmental program associated with lineage specification. Therefore, we hypothesized potential roles for T and GSC in PST/STX regulation because they are developmentally regulated transcription factors known to control early cell fate decisions, including the epithelial to mesenchymal transition (EMT) associated with early PSC differentiation (Showell et al. 2004; Fernando et al. 2010; Hartwell et al. 2006). Developmental transcription factors have not previously been implicated in regulation of cell surface glycans, making this a potentially important possibility. To investigate this, we first tested the possibility that T directly regulates the PST gene during endoderm differentiation. Ectopic activation of T using a T-glucocorticoid receptor (T-GR) fusion protein (Carey et al. 1996) elevates cell surface polysialylation in hPSCs within 4 days treatment with dexamethasone (Dex) (Fig. 3.3E). Transcript analysis of cells expressing T-GR shows an increase in PST and NCAM expression following Dex induction with no change in STX expression (Fig. 3.3F). This is consistent with potential roles for T and PST in mesoderm/endoderm differentiation, but not ectoderm formation, which is more closely associated with STX expression (Fig. 3.2A-C). Dex treatment also increased endoderm markers SOX17 and FOXA2 (Fig.

3.3F), mesoderm markers ISL1 and GATA6 (Supplemental Fig. 3.5A), the EMT marker SNAI1 (Fig. 3.3E), and PSA (Fig. 3.3E) within 4 days. Although T-GR increased PST and NCAM transcript levels, indicating that T lies upstream of these genes, there was a significant delay between Dex addition and increased cell surface polysialylation. This points towards the effects of T-GR induction on PST transcription and polysialylation as being indirect or perhaps, requiring additional other factor(s).

Interestingly, we noticed that transcript levels for the developmental transcription factor GSC peaked slightly earlier than that for PST following T-GR induction (Fig. 3.3G). This is consistent with the respective kinetics of T, GSC and PST during Activin A-induced endoderm differentiation (Supplemental Fig. 3.5B). This indicates that GSC is downstream of T and that GSC may act directly on PST. To investigate this possibility, GSC was ectopically expressed in hPSCs to establish if it could activate the PST gene and increase cell surface polysialylation. Within 24 hours of GSC expression, cell surface polysialylation was significantly elevated (Fig. 3.3H,I). Furthermore, GSC transfected cells showed increased PST, GATA6 and SOX17 mRNA levels (Fig. 3.3J), and an inducible GSC-GR construct activated a PST-luciferase reporter (PST-LUC) (Supplemental Fig. 3.5C) within 6 hours (Fig. 3.3K) of Dex addition, whereas T-GR (+Dex) had no effect on the reporter (Supplemental Fig. 3.5D). This reporter faithfully reproduces induction kinetics of the endogenous PST gene during differentiation (Fig. 3.2A-B and Supplemental Fig. 3.5E), confirming it to be a suitable tool for these studies. Other luciferase reporter assays showed that GSC-GR strongly activated the PST promoter, whereas T-GR had a weaker effect and neither GSC-GR nor T-GR activated an STX-luciferase reporter (Supplemental Fig. 3.5F). GSC was shown to bind the PST

promoter, but not the STX promoter, by ChIP assays (Fig. 3.3L), confirming that it directly regulates PST transcription. Finally, shRNA knockdown of GSC blocks the up-regulation of PST during endoderm differentiation and suppresses the down-regulation of T (Fig. 3.3M and Supplemental Fig. 3.5G). This latter observation is consistent with a previous report showing that GSC is a repressor of T transcription (Artinger et al. 1997). Taken together, these data indicate that T promotes expression of GSC, which then represses T and directly activates transcription of PST resulting in cell surface polysialylation (Fig. 3.3N).

PST-dependent polysialylation is required for early cell fate specification

Earlier in this report, we showed that shRNA knockdown of PST activity significantly reduces polysialylation under endoderm differentiation conditions (Fig. 3.2D). qPCR and immunofluorescence staining analysis also showed that up-regulation of endoderm markers FOXA2 and SOX17 was blocked under these conditions (Fig. 3.4A and Supplemental Fig. 3.6A-B), indicating that polysialylation is essential for normal differentiation. shRNA knockdown of NCAM, however, did not affect the ability to form DE despite a loss of PSA on the surface and its restriction to the Golgi (Fig. 3.2G and Supplemental Fig. 3.6C-E).

To establish by an independent approach that PST is required for endoderm differentiation, we engineered hPSC lines carrying a homozygous deletion of the PST gene using CRISPR-Cas9 genome editing (Fig. 3.4B and Supplemental Fig. 3.7). Culture of these cells with Activin A confirmed results obtained by shRNA knockdown of PST. After 4 days of Activin A-treatment of $PST^{-/-}$ cells, endoderm markers assayed by qPCR,

Western blot, flow cytometry, and immunofluorescence are markedly reduced (Fig. 3.4C-F) relative to hPSCs, indicating a major disruption to the differentiation program. Additionally, forced expression of either PST or STX restores cell surface polysialylation and rescues the DE differentiation defect of $PST^{-/-}$cells (Supplemental Fig. 3.8). These results suggest that polysialylation by PST is required for hPSC differentiation to DE and PSA plays an important role in hPSC cell fate determination.

In contrast to endoderm experiments, $PST^{-/-}$ cells differentiate to mesoderm and are comparable to WT cells (Supplemental Fig. 3.9A-C). Flow cytometry indicates $PST^{-/-}$ mesoderm cells express a high level of cell surface PSA (Supplemental Fig. 3.9A), but STX expression is elevated and appears to compensate for loss of PST activity (Supplemental Fig. 3.9C). These results are surprising considering both endoderm and mesoderm express PST exclusively in WT cells, suggesting that the mechanisms controlling polysialylation are regulated differently. In addition, $PST^{-/-}$ cells were also able to generate NCCs and neural progenitors similar to WT cells (Supplemental Fig. 3.9D-F); this result was expected given that STX is expressed in early neuroectoderm (Fig. 3.2A). Analysis of the bivalent marks on the STX promoter during WT mesoderm differentiation shows that there is a high level of activating H3K4me3 despite STX transcript not being expressed (Supplemental Fig. 3.4). This is also true for $PST^{-/-}$ mesoderm (Supplemental Fig. 3.9G) and may indicate that STX is unrepressed and capable of activation in the absence of PST.

To test if polysialylation is sufficient to promote exit from pluripotency and for transition towards endoderm, PST alone or PST with NCAM was over-expressed in hPSCs under self-renewal conditions. When polysialylated NCAM was presented on the

cell surface following ectopic expression of PST and NCAM, hPSCs maintained the expression of pluripotency marker OCT4 (Supplemental Fig. 3.10A,D), retained an epithelial morphology (Supplemental Fig. 3.10B-D), and did not up-regulate mesendoderm markers such as T (Supplemental Fig. 3.10C) or DE markers such as SOX17 and FOXA2 (Supplemental Fig. 3.10A-B). Although cell surface polysialylation is a characteristic of early hPSC differentiation towards the three germ layers, ectopic expression of PST and ectopic cell surface polysialylation is not sufficient for exit from pluripotency.

DISCUSSION

In this report, we show that as pluripotent cells commit towards the three germ layers, they become polysialylated on their cell surface. The major target for this polysialylation in endoderm, ectoderm and mesoderm differentiation appears to be the cell adhesion molecule NCAM. Polysialylation of cell surface NCAM is critical for endoderm differentiation but is not sufficient to trigger exit from pluripotency and is not directly involved in the EMT associated with transition through the mesendoderm state. This is surprising because polysialylation of NCAM has been reported to play role in cell migration by promoting an EMT (Lehembre et al. 2008). Loss of NCAM eliminates surface polysialylation in our experiments, similar to that seen in PST loss of function experiments, but it does not impact endoderm differentiation. This may suggest that additional polysialylated targets control differentiation and that NCAM is only part of the general mechanism. Our preferred explanation however, relates to previous reports in mice where loss of NCAM rescues neurological defects seen in PST and STX mutants

(Weinhold et al. 2005). This has led to the idea that unpolysialylated NCAM inhibits neural development and that polysialylated NCAM is compatible with neurogenesis. The role of PST and STX in this context would be to suppress negative signaling inputs generated by NCAM. This model is consistent with our observations but the details of how unpolysisalylated NCAM blocks differentiation needs to be explored further.

Although cell surface polysialylation is synonymous with early differentiation, it is regulated differently in the three germ layers. In endoderm polysialylation is exclusively controlled by PST, while in ectoderm, STX is the primary enzyme for catalytic activity. Mesoderm is perhaps the most interesting case. Here PST is the principle polysialyltransferase under normal conditions, but this switches to STX when PST activity is eliminated. This raises some interesting questions about how the regulatory mechanisms of polysialyltransferase genes are connected and how under specific conditions they can compensate for one another.

Many reports have been published describing the role of polysialylation in mouse development, particularly during neurogenesis. However, there have been no reports describing roles for polysialylation as cells exit pluripotency and as they form the three embryonic germ layers. We found that genes encoding the polysialyltransferases (PST and STX) and the main PSA acceptor NCAM are epigeneticaly marked by bivalent domains in pluripotent cells, and under specific conditions, these marks become resolved depending on the lineage chosen. These epigenetic marks implicated PST, STX and NCAM as part of a broader developmental program required for germ layer formation. This was subsequently confirmed by shRNA and CRISPR-Cas9 genome editing based experiments. Despite showing an essential role for PST in early differentiation, its exact

function has not yet been resolved. We know, however, that its function is required after the EMT associated with loss of pluripotency but before key lineage specification genes, such as SOX17 and FOXA2, are activated. We speculate that a signaling defect is associated with loss of PST activity and this will be the subject of further investigation.

In endoderm, PST activation is dependent on a well-characterized transcriptional regulatory network required for lineage specification. Most notably, T is a T-box transcription factor with downstream effectors, such as GSC, that are expressed during primitive streak formation. In our experiments, we find that T is required for GSC activation, and GSC directly binds and transcriptionally activates the PST gene. Regulation of polysialylation through PST is therefore part of the developmental program required for endoderm and probably mesoderm and ectoderm formation. The regulatory network required for STX activation in ectoderm still needs to be defined but is likely to involve suppression of Smad2,3 activity (Chambers et al. 2009) and/or activation of Wnt activity (Menendez et al. 2011).

T and GSC have been previously implicated as transcriptional drivers of metastasis in a range of cancers (Fernando et al. 2010; Hartwell et al. 2006). It is intriguing to speculate that reactivation of the T/GSC transcriptional network in tumor progression could also involve activation of polysialyltransferase genes such as PST to facilitate migration and metastasis. This is particularly interesting because there are numerous reports describing cell surface polysialyation on metastatic, migratory cells (Livingston et al. 1988; Schreiber et al. 2008; Tanaka et al. 2000; Roth et al. 1988; Amoureux et al. 2010). Reactivation of the T/GSC-driven polysialyltransferase activity could therefore be an important component of T/GSC-driven tumor development.

Interestingly, our observations of PST requirement for endoderm specification in hPSCs contrast with previous reports showing that PST null mice are viable with only mild cognitive impairments (Eckhardt et al. 2000). Additional studies showed that mice lacking STX shared a similar phenotype (Angata et al. 2004) and that double knockout of PST and STX led to severe developmental defects and precocious death (Weinhold et al. 2005), indicating that polysialylation is required for development but that PST and STX are functionally redundant and capable of compensating for each other. The discrepancy between our findings and previous reports are intriguing as it could represent a differential requirement of individual polysialyltransferase genes between species. Furthermore, this observation of an essential human gene being inessential in mice has been documented previously (Liao and Zhang 2008). Liao and Zhang demonstrated that null mutations of mouse genes orthologous to essential human genes often show mild or normal phenotypes and that this phenomenon is frequently observed in genes involved in glycosylation and carbohydrate metabolism (Liao and Zhang 2008). Their work coupled with our observations implies that considerable differences exist between species in regard to single gene mutation and the subsequent phenotypic effect and that this gap may be particularly evident in genes involved in glycosylation.

Alternatively, this discrepancy may be the result of a non-cell autonomous rescue of early endoderm *in vivo* by neighboring cells utilizing STX in the absence of PST. Our results show that $PST^{-/-}$ cells form mesoderm and ectoderm similar to WT and it could be that these cells assist in the formation of endoderm in the embryo through mechanisms not present during *in vitro* differentiation.

The impact of glycosylation in embryonic development is not well understood and is an increasingly important topic. In addition to our findings there have been a number of recent reports documenting potential roles for sialylation and sialyltransferases in hPSC pluripotency and differentiation (Wang et al. 2015; Alisson-Silva et al. 2014; Swindall et al. 2013; Wang et al. 2011; Tateno et al. 2011). The full impact of these processes is to be determined but may point toward sialylated glycans playing key roles in early development.

Conclusion

Our observations indicate that the initial differentiation of hPSCs toward lineages of all three germ layers is marked by abundant cell surface polysialylation stemming from increases in polysialytransferase expression. This process is regulated epigenetically and transcriptionally via well-established developmental pathways and appears to be critical for cell fate specification.

MATERIALS AND METHODS

Stem Cell Culture and Differentiations

WA09, WA01, WA07, and TE03 hESCs (http://grants.nih.gov/stem_cells/registry/current.htm) were cultured on Geltrex (Life Technologies) coated plates and sustained using complete defined medium (CDM) containing recombinant Heregulin β-1 (10 ng/mL), Activin A (10 ng/mL), bFGF2 (8 ng/mL), and IGF-1 (200 ng/mL) as described previously (Wang et al. 2007). Cells were plated at a density of 50,000 cells/cm^2 and grown to 90% confluency before passaging with Accutase (ICT).

Mesoderm cells were generated by culturing WA09 cells for 4 days in CDM supplemented with WNT3a (25 ng/mL) and BMP4 (100 ng/mL). Endoderm and ectoderm differentiations followed previously established protocols (McLean et al. 2007; Menendez et al. 2011).

qRT-PCR.

RNA was collected using the E.Z.N.A. RNA isolation kit (Omega) and cDNA made from 1 μg RNA using the Iscript cDNA synthesis kit (Bio-Rad). Genes were assayed using Taqman primers (Life Technologies) on a ViiA 7 Real-Time PCR System (Life Technologies). Assays were performed in triplicate and normalized to GAPDH expression. Error bars indicate ± SEM.

Western blots

Protein was prepared using RIPA lysis buffer and Western blots done using 30 μg of protein lysate loaded into NuSep Tris-Glycine Gels. Protein was then transferred onto a nitrocellulose membrane and probed using antibodies shown in Table 1 with detection via HRP.

Immunohistochemistry

Cells were grown on Lab-Tek 4-well chamber slides, fixed using 4% paraformaldehyde treatment for 10 minutes, and permeabilized using 0.2% Triton in PBS. Antibodies and concentrations used for staining are shown in Table 1. Images were obtained using a Leica DM6000 B microscope and a Zeiss LSM 710 confocal microscope.

Flow Cytometry

Cells were harvested using Accutase and 1 million cells were incubated with antibodies or isotype control in 20 μL PBS containing 0.2% BSA. Antibodies and concentrations used are shown in Table 1. Results were obtained using a Beckman Coulter Cyan ADP analyzer with data analysis using FlowJo software. Cell sorting was done using a Beckman Coulter MoFlo XDP and a Bio-Rad S3.

ChIP-PCR assays

10 million cells were cross-linked using 2% formaldehyde and quenched using 2.5M glycine. Lysis was performed using the Agilent Mammalian ChIP protocol and DNA sonicated to ~500 bp using a Covaris S220 Focused-ultrasonicator. Immunoprecipitation was done by overnight incubation of cell lysate at 4°C with Dynabeads Protein G magnetic beads (Life Technologies) conjugated with 10 μg of antibody (Table 1). qRT-PCR of ChIP lysate was done using genomic primers on a ViiA 7 Real-Time PCR System with comparison to whole cell extract.

Luciferase assays

Reporter plasmids were constructed by cloning the 5kb upstream genomic region of PST and STX amplified using Expand Long Range dNTPack (Roche) with ligation into a pGL4.1 Luciferase plasmid (Promega). Cells were plated onto 24-well plates and transfected in triplicate with 1 μg reporter plasmid and 100 ng Renilla control plasmid (Promega) using the Dual-Luciferase Reporter Assay System (Promega). Cells were

collected and analyzed 48h after transfection and analyzed for luminescence by a Biotek Synergy 2 plate reader. Data are shown in triplicate with error bars indicating ± SD.

shRNA knockdowns and CRISPR genome editing

shRNA knockdowns were done using TRC lentiviral shRNA plasmids (Dharmacon) targeting human PST (Cat. RHS4533-EG7903), NCAM (RHS4533-EG4684), and GSC (Cat. RHS4533-EG145258). Lentiviral particles were generated using a Trans-Lentiviral Packaging Kit (Thermo Scientific). Virus was concentrated upon collection using Lenti-X Concentrator (Clontech) and titer determined via qPCR titer kit (Mellgen Labs). Cells were plated in 24-well plates, transduced at MOI=1-5 in the presence of 6 μg/uL polybrene, and selected using 3 μg/mL Puromycin. CRISPR sgRNAs were designed using CRISPR Design software (Ran et al. 2013) and cells were transfected with sgRNAs, hCas9 (Mali et al. 2013), and repair plasmid. Cells were selected using Zeocin (100μg/mL) and grown as single cell clones in 96-well flat bottom plates with addition of 10μM Y-27632 (Tocris). $PST^{-/-}$ cells were verified via genomic PCR.

ACKNOWLEDGEMENTS

We thank Dr. Karen J. Colley (University of Illinois at Chicago) for providing PST-v5 and STX-v5 plasmids and IP positive control lysate and Dr. Rita Gerardy-Schahn (Hannover Medical School) for providing mAb735 antibody. This work was supported by NIH grants P01 GM75334 and P41 GM103490.

Figure 3.1. Polysialic acid is expressed upon differentiation to all three germ layers. (A) Flow cytometry of hPSC differentiation to definitive endoderm showing surface expression of PSA along with endoderm marker CXCR4 (Top) and pluripotency marker SSEA3 (Bottom). (B) Immunoblot of PSA expression in hPSC and DE. (C) Flow cytometry of PSA expression (Top) in hPSC (Shaded gray), DE (Red), mesoderm (Blue), and NCC (Green). PSA^+ fractions are 2.9%, 79.7%, 72.4%, and 81.1% respectively. Isotype control is shown for comparison (Bottom). (D-H) Immunostaining of PSA expression in hPSC differentiation. Scale bar, 100μm. (D) hPSC: PSA and DE marker SOX17 (Top), PSA and Golgi marker α-mannosidase II (Bottom). (E) DE: PSA and SOX17 (Top), PSA and α-mannosidase II (Bottom). (F) Mesoderm: PSA and T (Top), PSA and NANOG (Bottom). (G) NCC: PSA and NANOG (Top), PSA and Golgi marker Giantin (Bottom). (H) Zoom of PSA and α-mannosidase II in DE. Arrows point to overlap of PSA and α-mannosidase II. (I) Time course immunoblot of PSA expression during hPSC differentiation to NCC (Top). CDK2 shown as loading control (Bottom). Abbreviations: hPSC, human pluripotent stem cell; DE, definitive endoderm; NCC, neural crest cell.

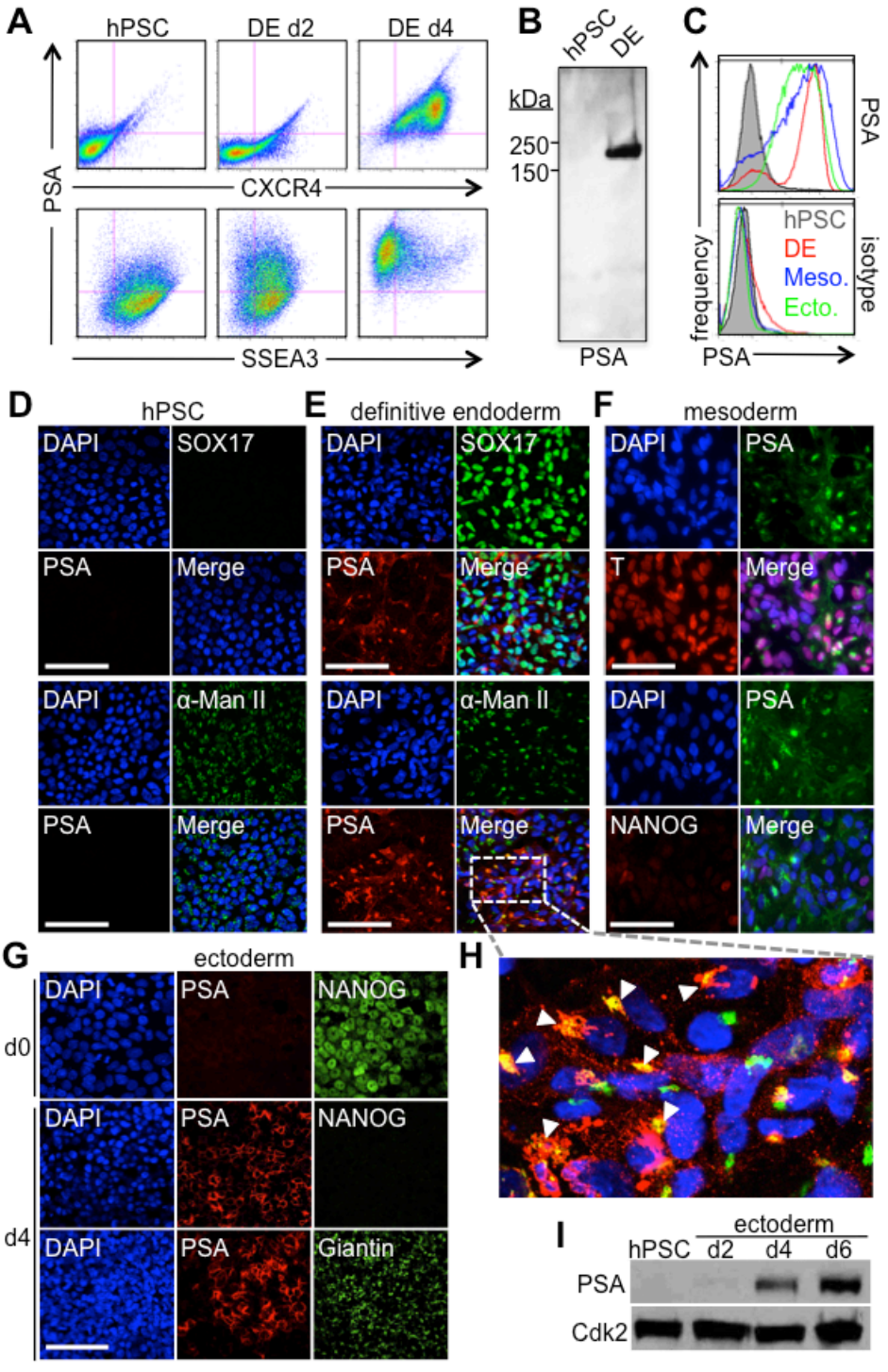

A
hPSC
DE d2
DE d4
PSA
CXCR4
SSEA3
B
hPSC
DE
kDa
250
150
PSA
C
PSA
hPSC
DE
Meso.
Ecto.
isotype
frequency
PSA
D
hPSC
DAPI
SOX17
PSA
Merge
α-Man II
E
definitive endoderm
F
mesoderm
T
NANOG
G
ectoderm
d0
d4
Giantin
H
I
ectoderm
hPSC
d2
d4
d6
Cdk2

Figure 3.2. Expression of polysialyltransferases is lineage specific and NCAM is the substrate of polysialylation. (A-C) qPCR expression of PST, STX, and NCAM (Top) and lineage markers (Bottom) in hPSC differentiation to DE (A), mesoderm (B), and NCC (C). (D) Immunostaining of PSA expression in DE day 4 cells transduced with shRNA targeting PST (Right) compared to control shRNA (Left). Scale bar, 100μm. (E) Flow cytometry of PSA and NCAM surface expression in hPSC, DE, and mesoderm. (F) Immunostaining of shRNA knockdowns of PST and NCAM in DE day 4. Scale bar, 100μm. (G) Immunostaining of PSA and α-mannosidase II expression in DE cells transduced with NCAM shRNA. Arrows point to overlap in expression. Scale bar, 100μm. (H) Immunoblot of NCAM (Top) and PSA (Middle) in all three germ layers. β-Actin shown as loading control (Bottom). (I) Immunoprecipitation of NCAM and PSA in hPSC and DE. IP lysate was then immunoblotted for PSA (Left) and NCAM (Right). Lysate from COS-1 cells overexpressing PST, STX, and NCAM was used as positive control. Abbreviations: hPSC, human pluripotent stem cell; DE, definitive endoderm; CTL, control.

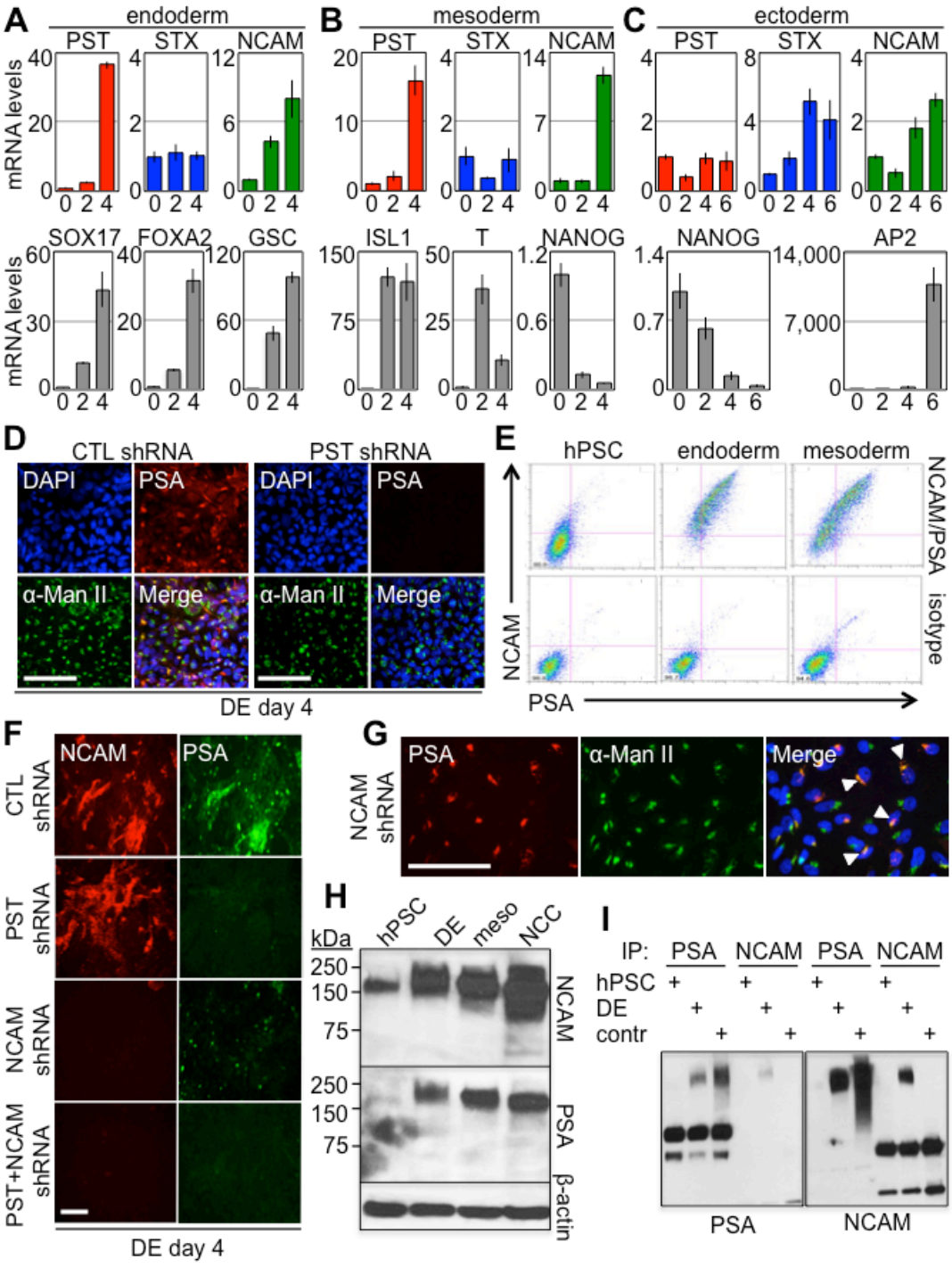

A
endoderm
PST
STX
NCAM
mRNA levels
SOX17
FOXA2
GSC
B
mesoderm
ISL1
T
NANOG
C
ectoderm
NANOG
AP2
D
CTL shRNA
PST shRNA
DAPI
PSA
α-Man II
Merge
DE day 4
E
hPSC
endoderm
mesoderm
NCAM/PSA
isotype
NCAM
PSA
F
CTL shRNA
PST shRNA
NCAM shRNA
PST+NCAM shRNA
DE day 4
G
NCAM shRNA
H
kDa
hPSC
DE
meso
NCC
β-actin
I
IP:
hPSC
DE
contr

Figure 3.3. Polysialylation is bivalently regulated with PST under the control of the T/GSC network. (A-C) Graphical representation of H3K4me3, H3K27me3, and CpG islands at the PST (A), STX (B), and NCAM (C) genomic loci in WA01 hPSCs (Data from the UCSC genome browser database (Rosenbloom et al. 2012)). (D) Graphical illustration of gene kinetics during DE differentiation. (E) Immunostaining comparison of T-GR expression of PSA and SNAI1 after 4 day addition of Dex compared to –Dex. Scale bar, 100μm. (F) qPCR of genes in T-GR cells –Dex, +Dex 2d, and +Dex 4d. (G) Transcript analysis of GSC and PST over 4 day time course addition of Dex in T-GR cells. (H) Flow cytometry for surface PSA expression in hPSCs, hPSCs+GFP, and hPSCs+GSC-GFP 24h post electroporation. (I) Analysis of the percentage of GFP^+/PSA^+ cells from Fig. 3.4H. (J) qPCR of samples from Fig. 3.4H. (K) Luciferase assay of hPSCs transfected with GSC-GR, PST-LUC, and LUC-Control after Dex addition for 0, 6, and 12h. (L) GSC ChIP assay in hPSCs transfected with GSC-GFP. Genomic primers used probe the TSS of the PST promoter and the STX -5kb upstream region shown as negative control. (M) qPCR of DE d4 cells transduced with shRNA targeting GSC compared to control shRNA. (N) Diagram of proposed interaction of T, GSC, and PST. Abbreviations: hPSC, human pluripotent stem cell; Dex, dexamethasone; GR, glucocorticoid receptor; DE, definitive endoderm; LUC, luciferase; TSS, transcription start site; CTL, control.

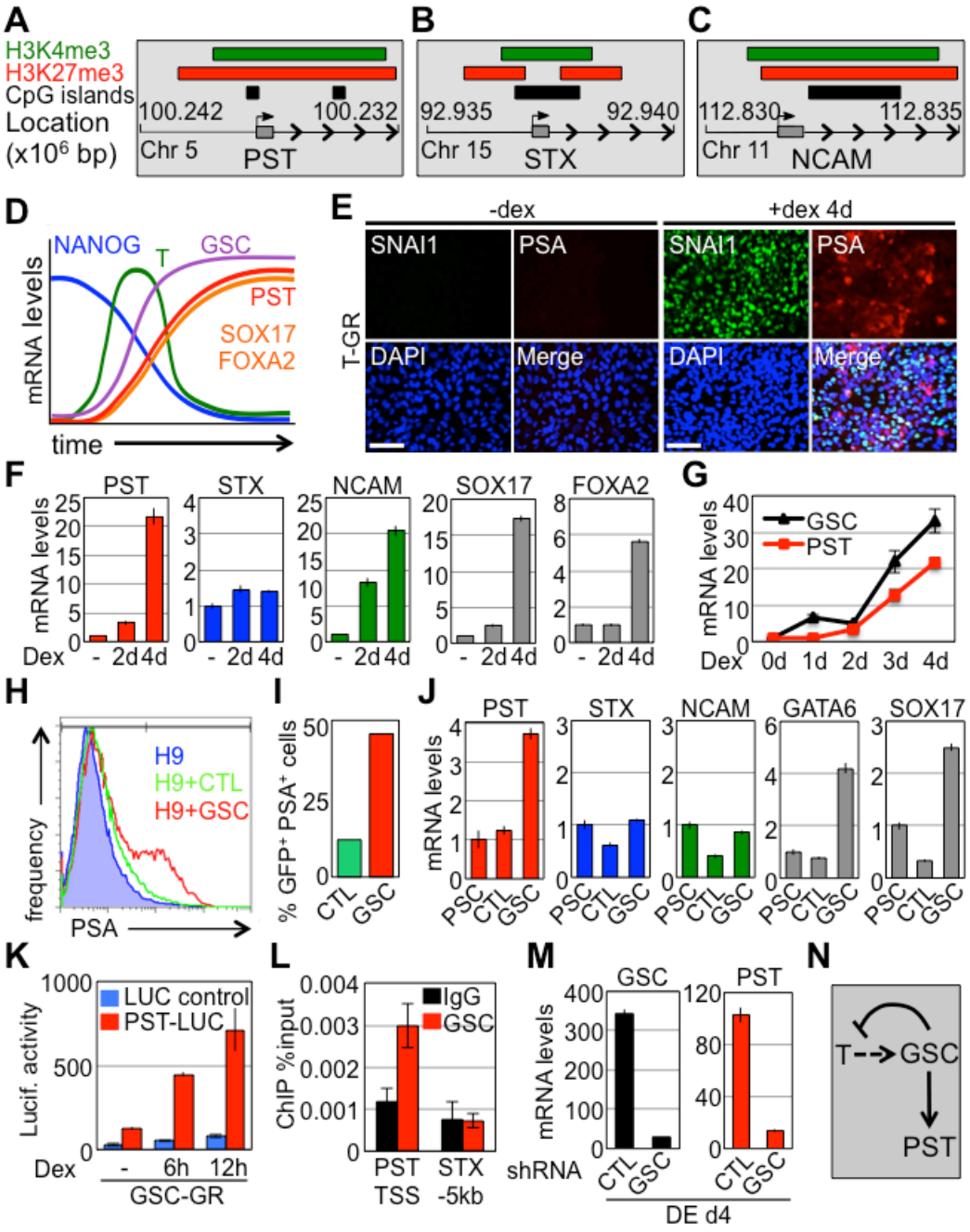
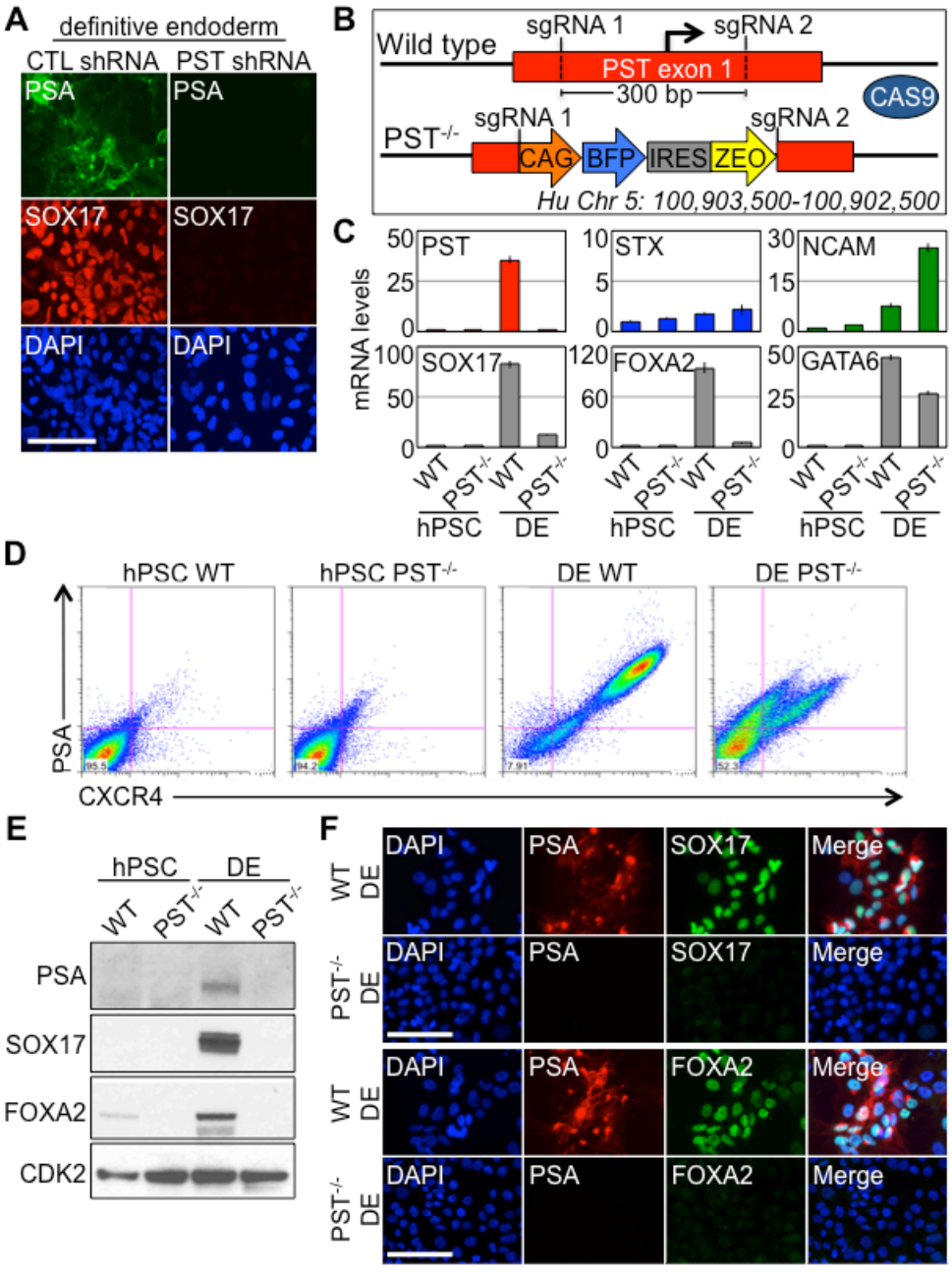
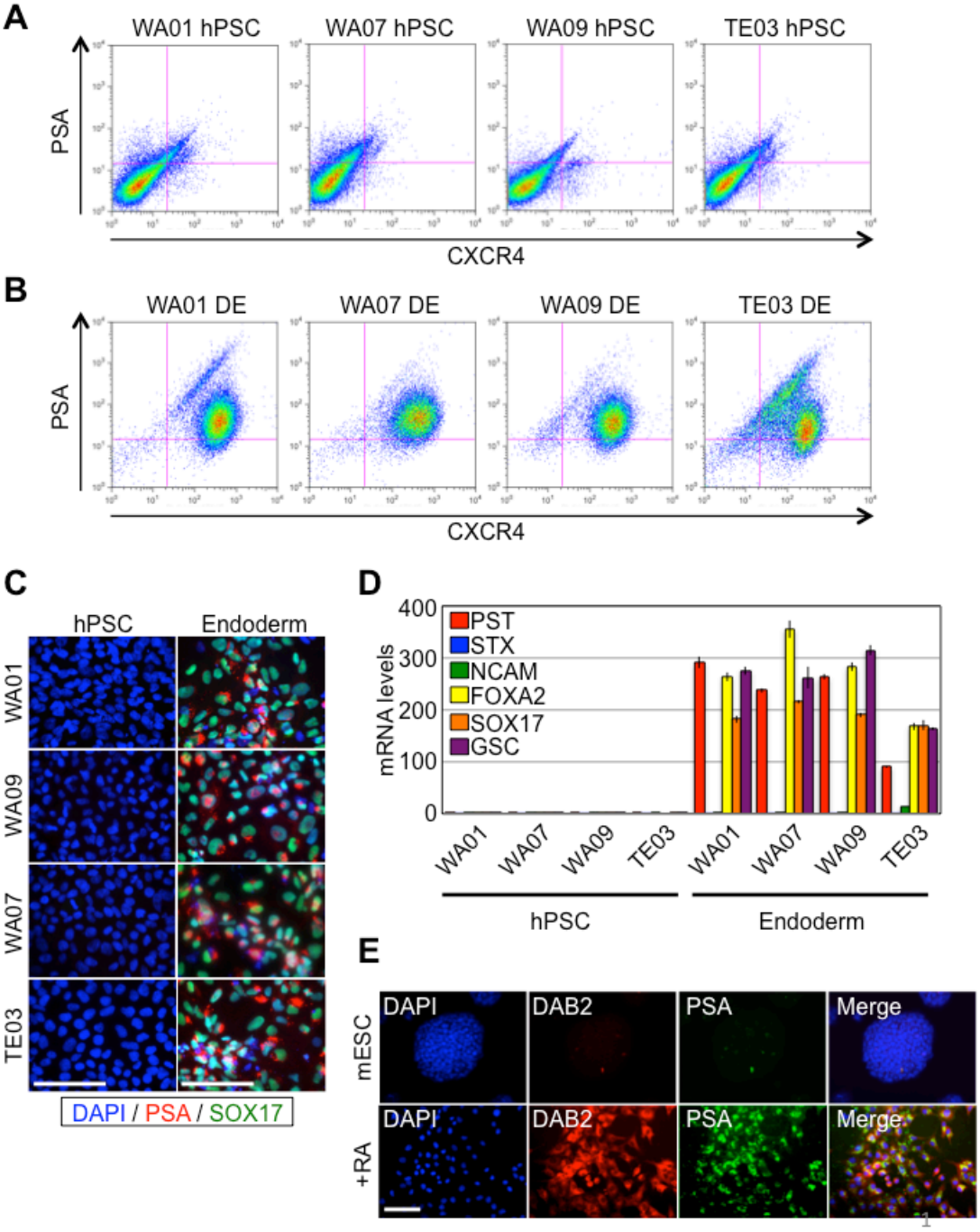
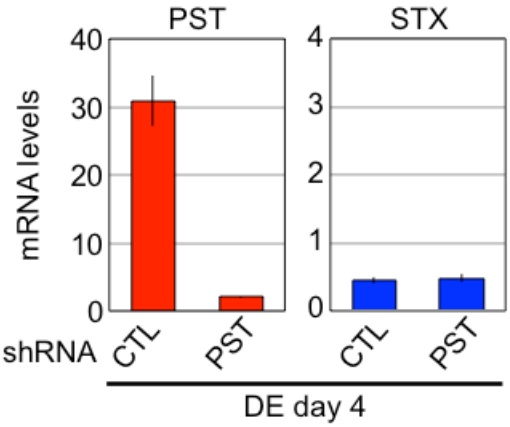

A
H3K4me3
H3K27me3
CpG islands
Location
(x10^6 bp)
100.242
100.232
Chr 5
PST
B
92.935
92.940
Chr 15
STX
C
112.830
112.835
Chr 11
NCAM
D
NANOG
T
GSC
PST
SOX17
FOXA2
mRNA levels
time
E
-dex
+dex 4d
T-GR
SNAI1
PSA
DAPI
Merge
F
PST
STX
NCAM
SOX17
FOXA2
mRNA levels
Dex
- 2d 4d
G
GSC
PST
mRNA levels
Dex
0d 1d 2d 3d 4d
H
H9
H9+CTL
H9+GSC
frequency
PSA
I
% GFP+ PSA+ cells
CTL
GSC
J
PST
STX
NCAM
GATA6
SOX17
mRNA levels
PSC CTL GSC
K
LUC control
PST-LUC
Lucif. activity
Dex
- 6h 12h
GSC-GR
L
IgG
GSC
ChIP %input
PST TSS
STX -5kb
M
GSC
PST
mRNA levels
shRNA
CTL GSC
DE d4
N
T
GSC
PST

Figure 3.4. PST is required for efficient hPSC differentiation to endoderm. (A) Immunostaining of PSA and SOX17 expression in DE cells transduced with control shRNA (Left) and PST shRNA (Right). Scale bar, 100μm. (B) Schematic of CRISPR gene editing approach of the PST first exon. The 300 bp region containing the PST TSS was cleaved and replaced by a BFP-IRES-Zeo cassette via homology directed repair. (C) qPCR of genes in WT and $PST^{-/-}$ cells differentiated to DE. (D) Flow cytometry of CXCR4 and PSA expression in WT and $PST^{-/-}$ cells differentiated to DE. (E) Immunoblot comparisons of WT and $PST^{-/-}$ cells differentiated to DE. (F) Immunostaining of PSA with SOX17 (Top) and FOXA2 (Bottom) in WT and $PST^{-/-}$ cells differentiated to DE. Abbreviations: CTL, control; hPSC, human pluripotent stem cell; DE, definitive endoderm; TSS, transcription start site; sgRNA, single guide RNA.


A
definitive endoderm
CTL shRNA
PST shRNA
PSA
SOX17
DAPI
B
Wild type
sgRNA 1
sgRNA 2
PST exon 1
300 bp
CAS9
PST-/-
CAG
BFP
IRES
ZEO
Hu Chr 5: 100,903,500-100,902,500
C
mRNA levels
PST
STX
NCAM
SOX17
FOXA2
GATA6
WT
hPSC
DE
D
hPSC WT
hPSC PST-/-
DE WT
DE PST-/-
PSA
CXCR4
E
hPSC
DE
PSA
SOX17
FOXA2
CDK2
F
DAPI
PSA
SOX17
FOXA2
Merge

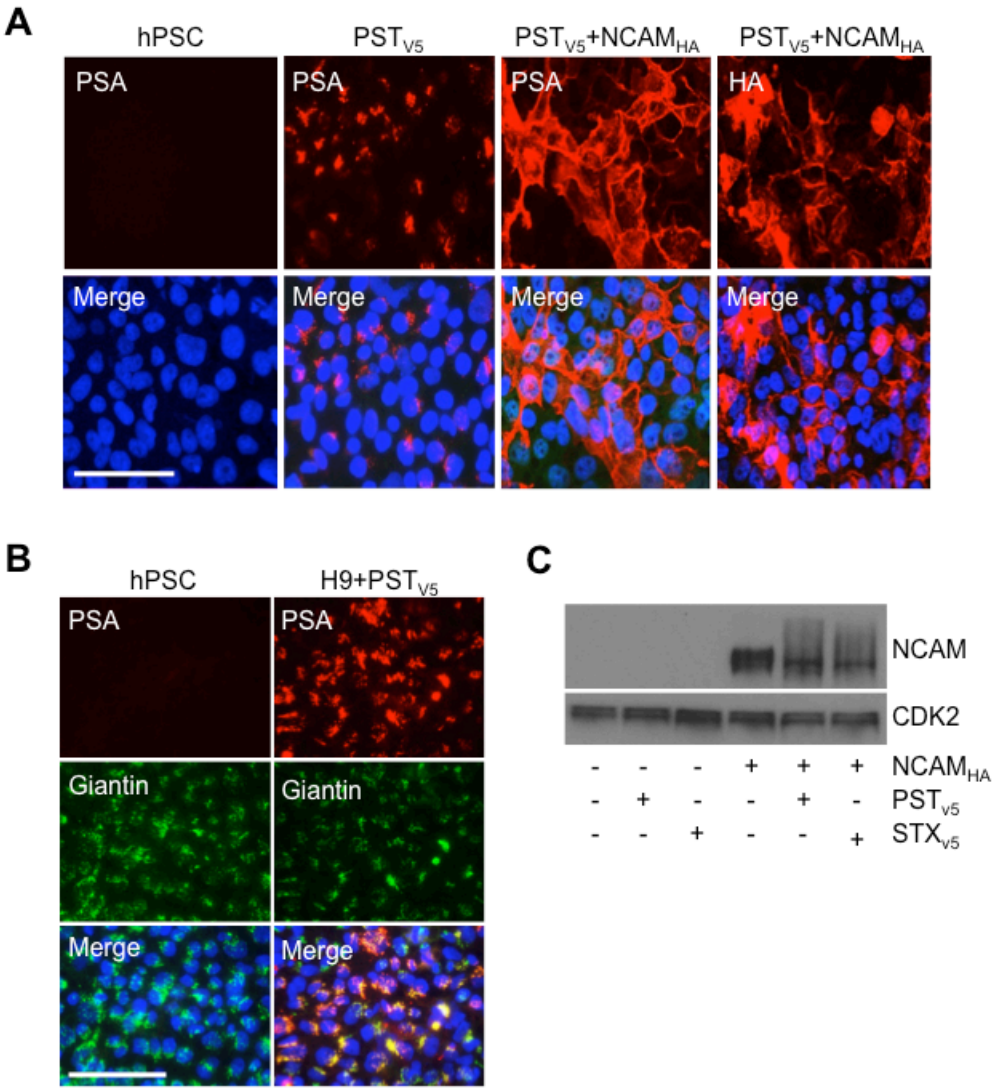
Supplemental Figure 3.1. Polysialic acid is expressed upon differentiation in multiple hPSC lines. (A-B) Flow cytometry showing surface expression of PSA and CXCR4 in hPSC (A) and DE (B) in WA01, WA07, WA09, and TE03 cell lines. (C) Immunostaining of PSA and SOX17 in DE cells. Scale bar, 100μm. (D) qPCR expression profile of hPSC endoderm differentiation. (E) Immnunostaining of mESC differentiation with retinoic acid. Abbreviations: mESC, mouse embryonic stem cells; RA, retinoic acid.

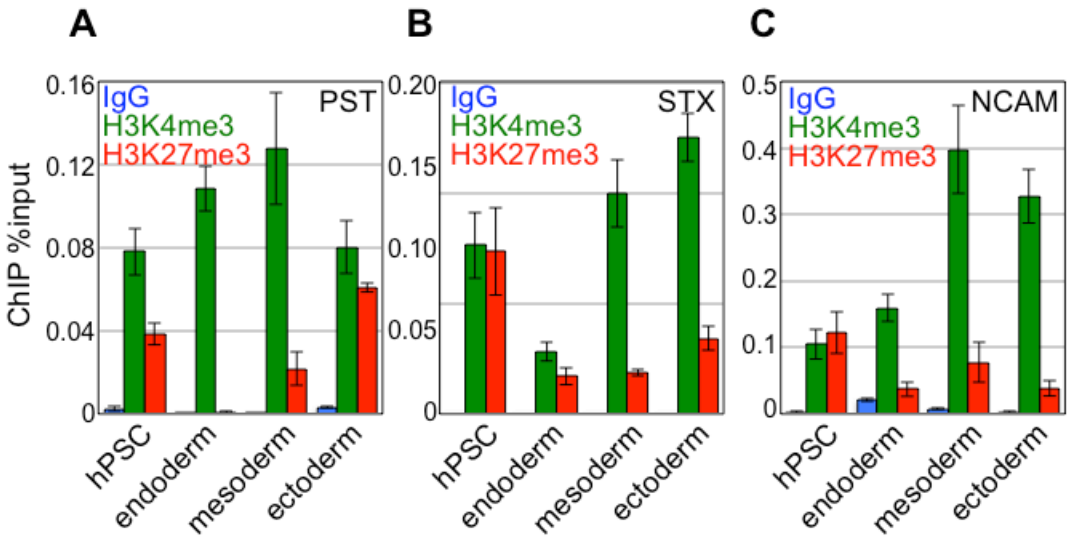

A
WA01 hPSC
WA07 hPSC
WA09 hPSC
TE03 hPSC
PSA
CXCR4
B
WA01 DE
WA07 DE
WA09 DE
TE03 DE
PSA
CXCR4
C
hPSC
Endoderm
WA01
WA09
WA07
TE03
DAPI / PSA / SOX17
D
mRNA levels
400
300
200
100
0
PST
STX
NCAM
FOXA2
SOX17
GSC
WA01
WA07
WA09
TE03
WA01
WA07
WA09
TE03
hPSC
Endoderm
E
mESC
+RA
DAPI
DAB2
PSA
Merge
DAPI
DAB2
PSA
Merge



Supplemental Figure 3.2. Transcript levels of PST and STX during PST shRNA knockdown in DE day 4 cells. Data are normalized to WA09 hPSC. Abbreviations: CTL, control; DE, definitive endoderm.

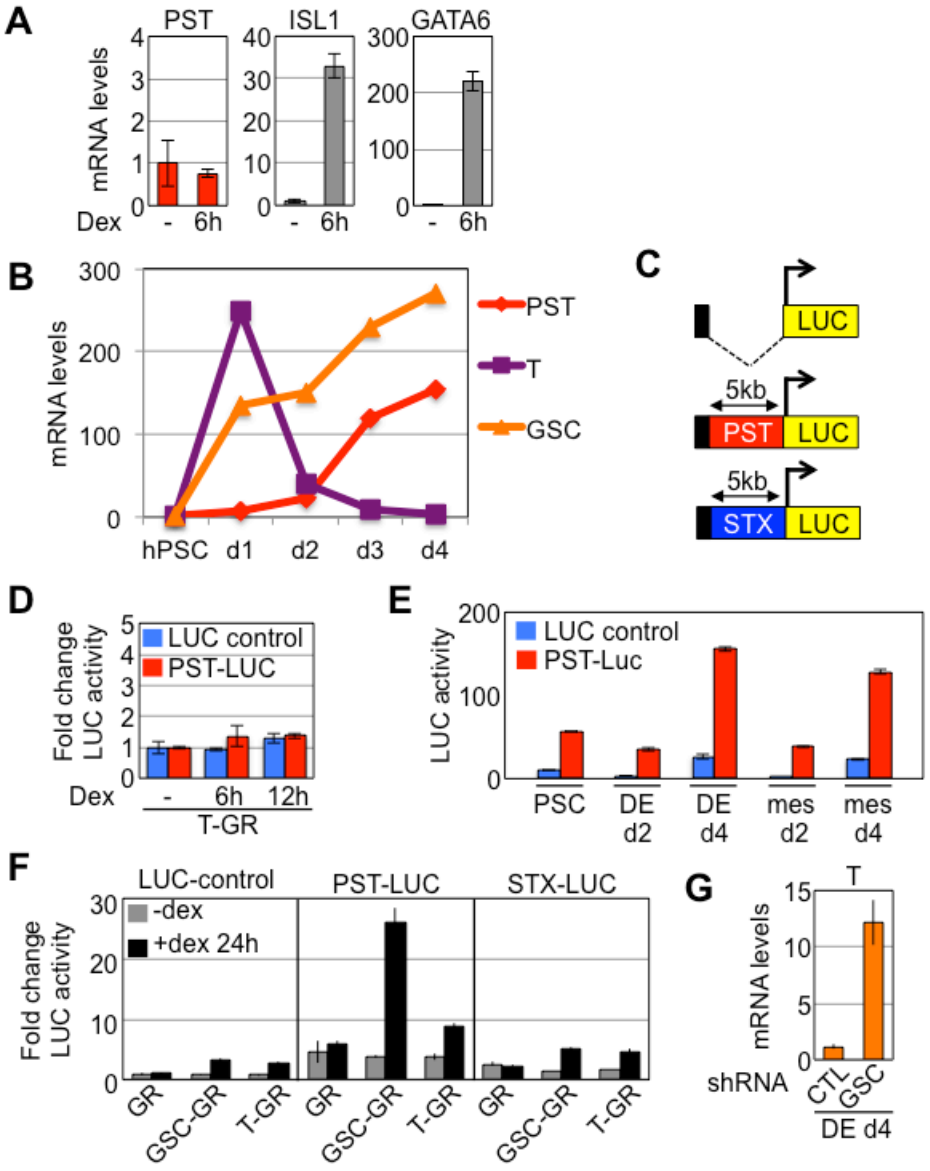
Supplemental Figure 3.3. Cell surface polysialylation requires expression of both PST/STX and NCAM. (A) Immunostaining of PSA expression hPSCs transfected with PST and NCAM overexpression constructs. Scale bar, 100μm. (B) Immunostaining of PSA and Golgi marker Giantin in hPSCs overexpressing PST. Scale bar, 100μm. (C) Immnunoblot of NCAM expression in hPSCs transfected with PST, STX, and NCAM overexpression constructs. CDK2 is shown as control. Abbreviations: hPSC, human pluripotent stem cell.


A
hPSC
PST$_{V5}$
PST$_{V5}$+NCAM$_{HA}$
PST$_{V5}$+NCAM$_{HA}$
PSA
PSA
PSA
HA
Merge
Merge
Merge
Merge
B
hPSC
H9+PST$_{V5}$
PSA
PSA
Giantin
Giantin
Merge
Merge
C
NCAM
CDK2
- - - + + + NCAM$_{HA}$
- + - - + - PST$_{v5}$
- - + - - + STX$_{v5}$

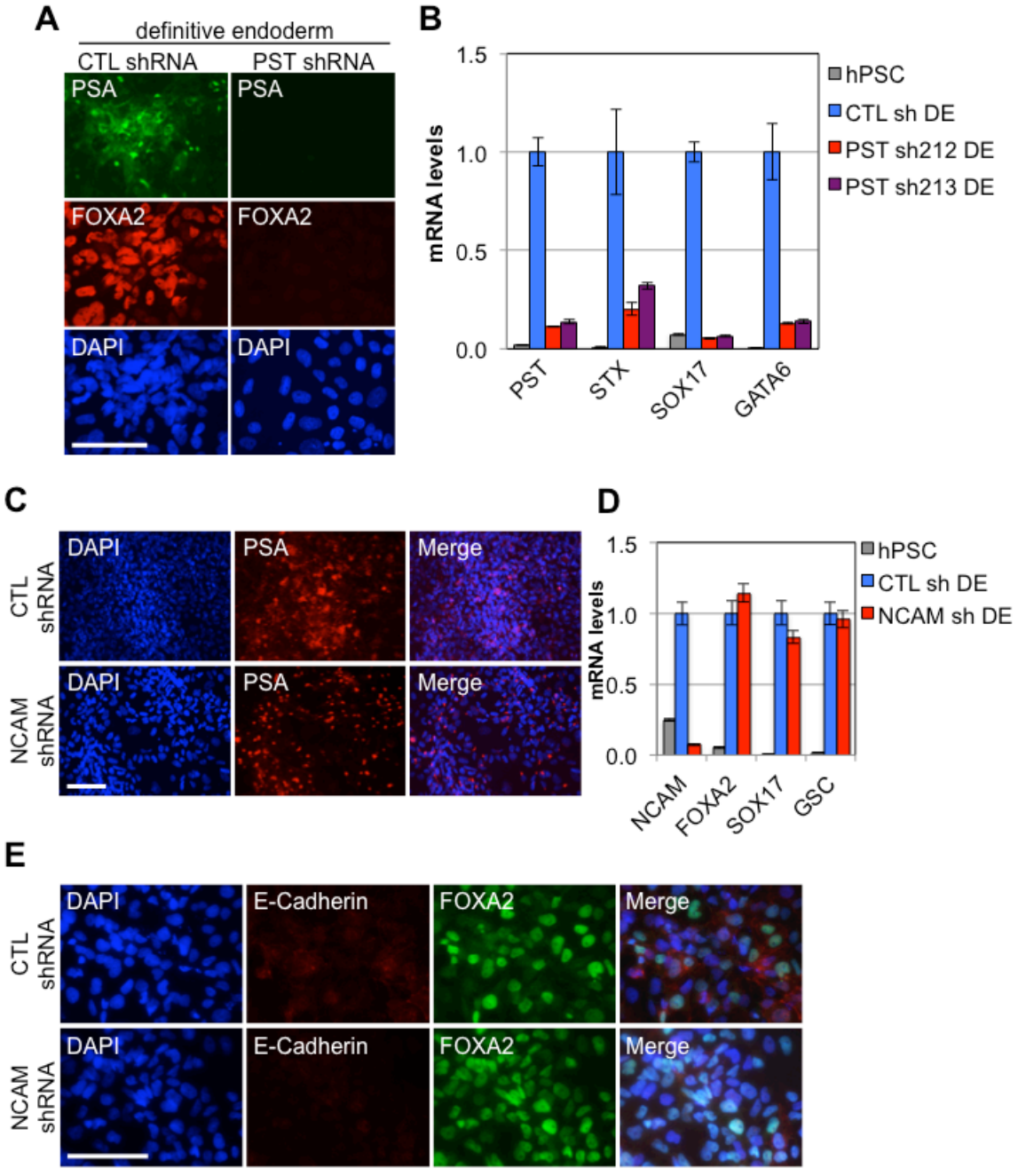


Supplemental Figure 3.4. Polysialylation is regulated by bivalent histone modifications. (A-C) ChIP PCR showing changes in H3K4me3 and H3K27me3 on the PST (A), STX (B), and NCAM (C) loci during differentiation to DE, mesoderm, and NCC. Abbreviations: hPSC, human pluripotent stem cell.

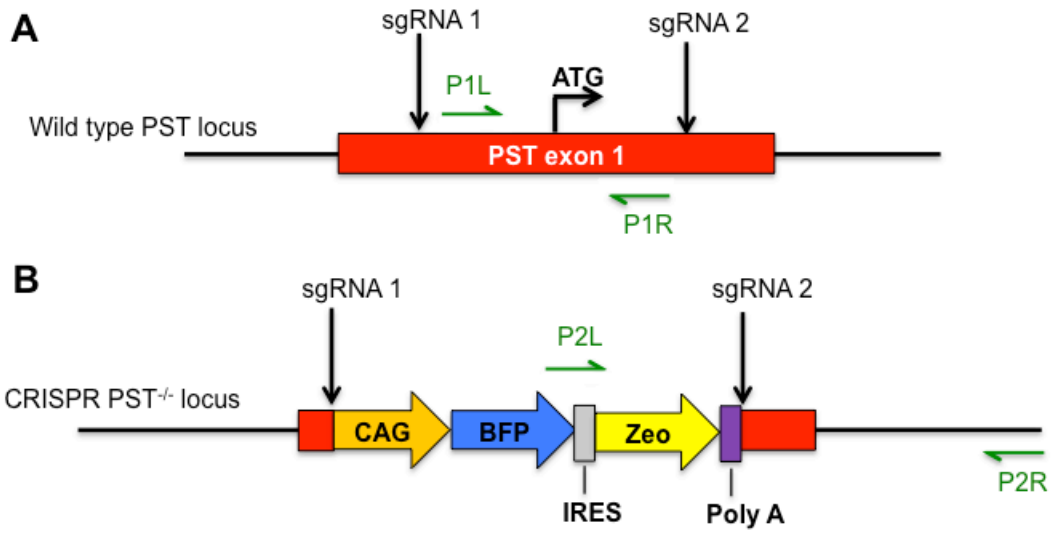
Supplemental Figure 3.5. PST is under the control of the T/GSC regulatory network in mesendoderm. (A) Transcript levels of PST, ISL1, and GATA6 in T-GR cells after 6h Dex addition. (B) Time course expression levels of PST, T, and GSC trancsripts during 4 day differentiation to DE. (C) Illustration of luciferase constructs. LUC-control (Top) has no promoter while PST-LUC (Middle) and STX-LUC (Bottom) contain the 5kb upstream promoters of the respective genes. (D) Luciferase assay comparing PST-LUC activity to LUC-control in T-GR cells induced with Dex for 6-12h. (E) Luciferase assay comparing PST-LUC activity to LUC-control in hPSC differentiation to DE and mesoderm. (F) Luciferase assay showing activity of LUC-control (Left), PST-LUC (Middle), and STX-LUC (Right) in GSC-GR and T-GR cells induced with Dex for 24h. (G) Transcript levels of T in shRNA knockdown of GSC in DE day 4 cells. Abbreviations: Dex, dexamethasone; hPSC, human pluripotent stem cell; LUC, luciferase; DE, definitive endoderm; mes, mesoderm; CTL, control.

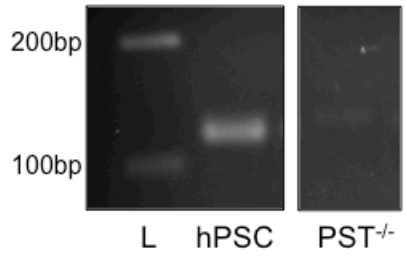

A
PST
ISL1
GATA6
mRNA levels
Dex
-
6h
B
mRNA levels
PST
T
GSC
hPSC
d1
d2
d3
d4
C
LUC
5kb
PST
STX
D
Fold change LUC activity
LUC control
PST-LUC
Dex
6h
12h
T-GR
E
LUC activity
LUC control
PST-Luc
PSC
DE d2
DE d4
mes d2
mes d4
F
LUC-control
PST-LUC
STX-LUC
-dex
+dex 24h
Fold change LUC activity
GR
GSC-GR
T-GR
G
T
mRNA levels
shRNA
CTL
GSC
DE d4

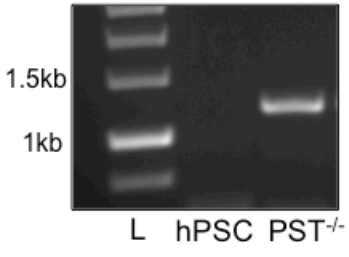
Supplemental Figure 3.6. shRNA knockdown of PST inhibits DE differentiation while NCAM knockdown does not. (A) Immunostaining of PSA and FOXA2 expression in DE cells transduced with control shRNA (Left) and PST shRNA (Right). Scale bar, 100μm. (B) qPCR of DE cells transduced with control shRNA and two shRNAs targeting PST. (C) Immunostaining of PSA expression in DE cells transduced with control shRNA (Top) and NCAM shRNA (Bottom). Scale bar, 100μm. (D) qPCR of NCAM and endoderm markers in DE cells transduced with control shRNA and NCAM shRNA. (E) Immunostaining of E-cadherin and FOXA2 expression in DE cells transduced with GFP control shRNA and NCAM shRNA. Scale bar, 100μm. Abbreviations: hPSC, human pluripotent stem cell; CTL, control; DE, definitive endoderm.


A
definitive endoderm
CTL shRNA
PST shRNA
PSA
PSA
FOXA2
FOXA2
DAPI
DAPI
B
mRNA levels
1.5
1.0
0.5
0.0
hPSC
CTL sh DE
PST sh212 DE
PST sh213 DE
PST
STX
SOX17
GATA6
C
CTL shRNA
NCAM shRNA
DAPI
PSA
Merge
DAPI
PSA
Merge
D
mRNA levels
1.5
1.0
0.5
0.0
hPSC
CTL sh DE
NCAM sh DE
NCAM
FOXA2
SOX17
GSC
E
CTL shRNA
NCAM shRNA
DAPI
E-Cadherin
FOXA2
Merge
DAPI
E-Cadherin
FOXA2
Merge

Supplemental Figure 3.7. Disruption of the PST gene using CRISPR-Cas9 genome editing. (A) Diagram of PST WT genomic locus. sgRNAs were designed to target the PST peptide start codon in the first exon. (B) Diagram of the PST CRISPR knockdout genomic locus. The region between sgRNAs has been replaced with a resistance cassette via homology directed repair. (C-D) genomic PCR of PST locus in hPSC and $PST^{-/-}$. Primers were designed to amplify the WT region excised by CRISPR-Cas9 (C) and the knockout locus containing the resistance cassette (D). (E-F) Expected PCR product sizes for WT PCR test (E) and CRISPR-Cas9 knockout PCR test (F). Abbreviations: sgRNA, single guide RNA; L, ladder.







E

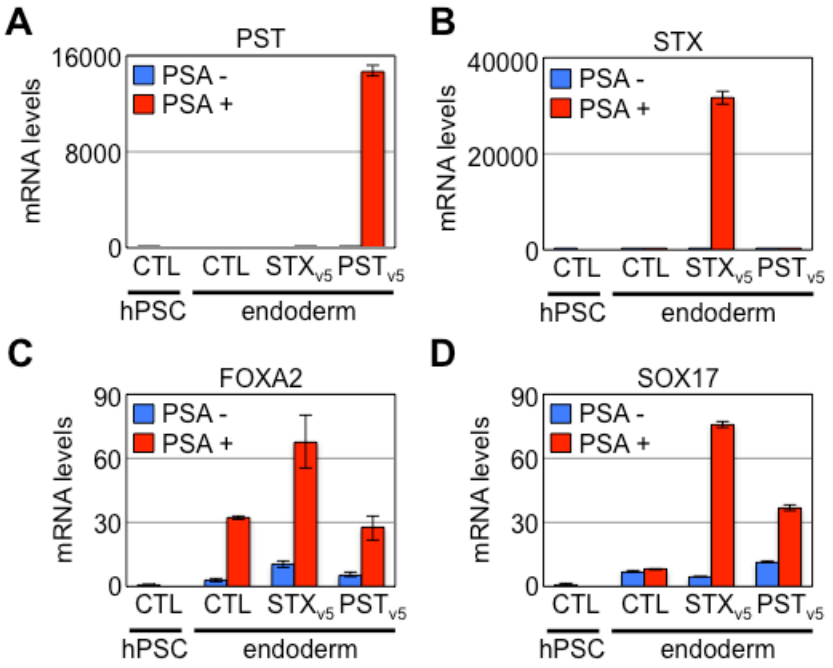
Wild type PST locus PCR

Primer set	hPSC (bp)	PST$^{-/-}$ (bp)
P1L P1R	119	0

F

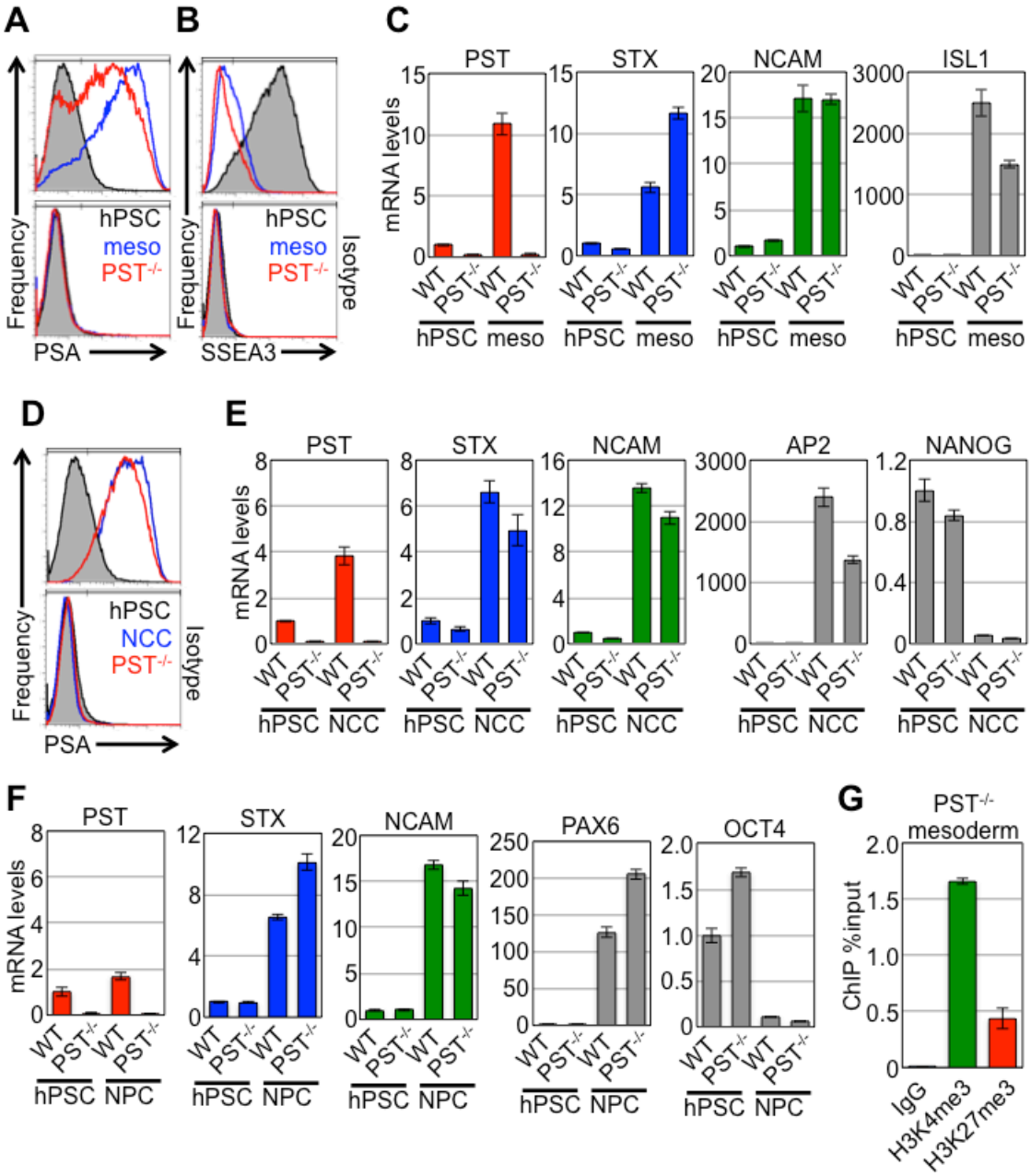
CRISPR PST locus PCR

Primer set	hPSC (bp)	PST$^{-/-}$ (bp)
P2L P2R	0	1269

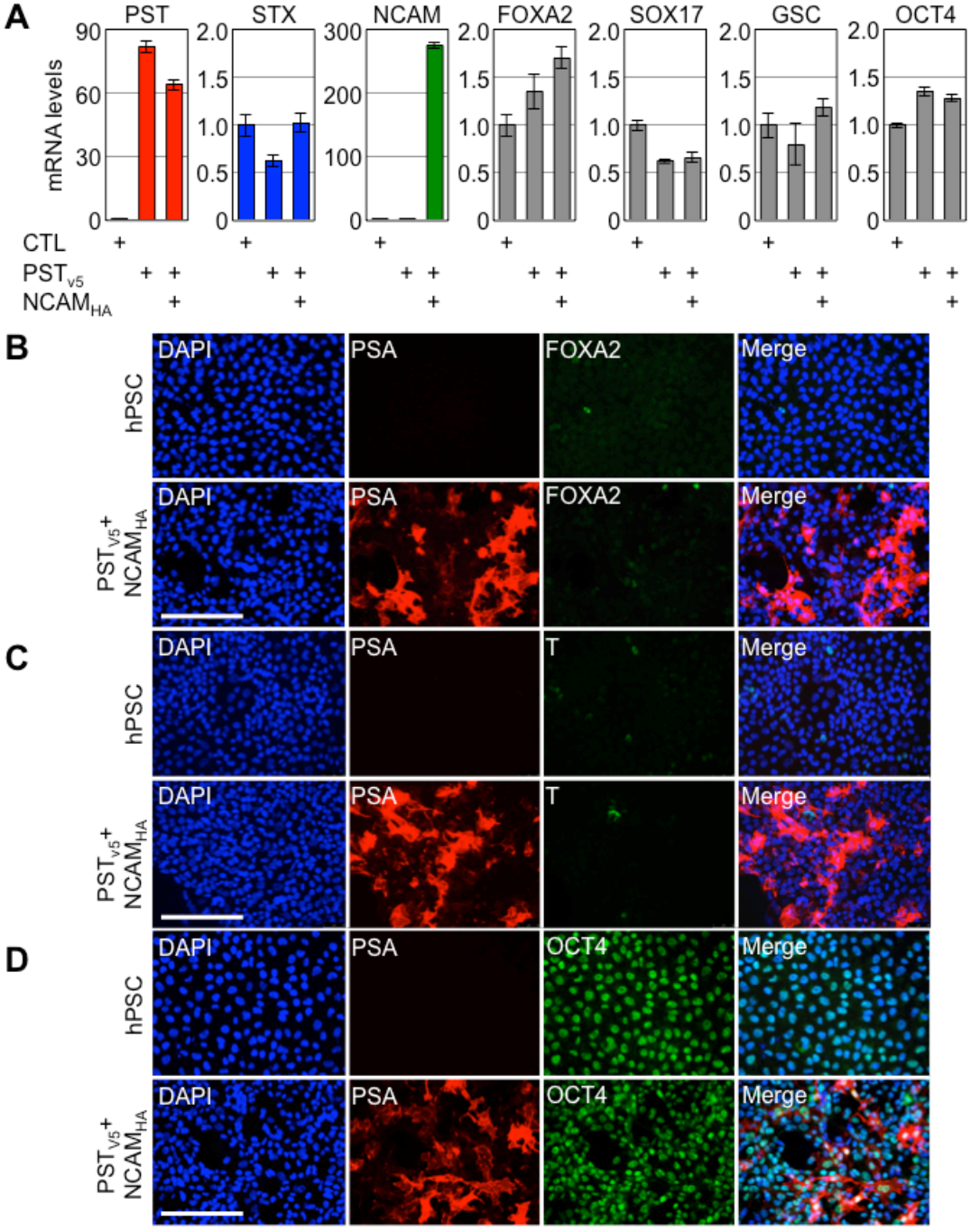


Supplemental Figure 3.8. Overexpression of polysialyltransferases restores endoderm marks in PST$^{-/-}$ DE. (A-D) qPCR of endoderm marks in PST$^{-/-}$ DE cells overexpressing PST$_{v5}$ or STX$_{v5}$ expression constructs. Cells were sorted for PSA^{+}/PSA^{-} 3 days after transfection.

Supplemental Figure 3.9. PST$^{-/-}$ cells are capable of forming mesoderm and ectoderm. (A-B) Flow cytometry showing surface expression of PSA (A) and SSEA3 (B) in hPSCs and PST$^{-/-}$ cells differentiated to mesoderm. (C) qPCR comparison of hPSCs and PST$^{-/-}$ cells differentiated to mesoderm. (D) Flow cytometry showing PSA expression in hPSCs and PST$^{-/-}$ cells differentiated to NCC. (E-F) qPCR comparison of hPSCs and PST$^{-/-}$ cells differentiated to ectoderm lineages NCC (E) and neural progenitor cells (F). (G) ChIP PCR of PST$^{-/-}$ mesoderm cells showing levels of H3K4me3 and H3K27me3 epigenetic marks on the STX promoter. Abbreviations: hPSC, human pluripotent stem cell; meso, mesoderm; NCC, neural crest cell; NPC, neural progenitor cell.


A
B
C
Frequency
hPSC
meso
PST-/-
PSA
SSEA3
Isotype
mRNA levels
PST
STX
NCAM
ISL1
WT
PST-/-
hPSC
meso
D
E
NCC
AP2
NANOG
F
NPC
PAX6
OCT4
G
PST-/- mesoderm
ChIP %input
IgG
H3K4me3
H3K27me3

Supplemental Figure 3.10. Ectopic expression of PSA in hPSCs is not sufficient to drive differentiation. (A) qPCR of hPSCs electroporated with PST and NCAM compared to a GFP control plasmid. (B-D) Immunostaining of PSA (B-D), FOXA2 (B), T (C), and OCT4 (D) in hPSCs electroporated with PST and NCAM overexpression construct. Scale bar, 100μm. Abbreviations: CTL, control; hPSC, human pluripotent stem cell.


A
PST
STX
NCAM
FOXA2
SOX17
GSC
OCT4
mRNA levels
CTL
PST$_{v5}$
NCAM$_{HA}$
B
C
D
hPSC
PST$_{v5}$+ NCAM$_{HA}$
DAPI
PSA
FOXA2
T
OCT4
Merge

REFERENCES

Alisson-Silva F, de Carvalho Rodrigues D, Vairo L, Asensi KD, Vasconcelos-dos-Santos A, Mantuano NR, Dias WB, Rondinelli E, Goldenberg RCDS, Urmenyi TP, et al. 2014. Evidences for the involvement of cell surface glycans in stem cell pluripotency and differentiation. *Glycobiology* **24**: 458–468.

Amoureux M-C, Coulibaly B, Chinot O, Loundou A, Metellus P, Rougon G, Figarella-Branger D. 2010. Polysialic Acid Neural Cell Adhesion Molecule (PSA-NCAM) is an adverse prognosis factor in glioblastoma, and regulates olig2 expression in glioma cell lines. *BMC Cancer* **10**: 91.

Angata K, Long JM, Bukalo O, Lee W, Dityatev A, Wynshaw-Boris A, Schachner M, Fukuda M, Marth JD. 2004. Sialyltransferase ST8Sia-II assembles a subset of polysialic acid that directs hippocampal axonal targeting and promotes fear behavior. *Journal of Biological Chemistry* **279**: 32603–32613.

Artinger M, Blitz I, Inoue K, Tran U, Cho KW. 1997. Interaction of *goosecoid* and *brachyury* in *Xenopus* mesoderm patterning. *Mechanisms of Development* **65**: 187–196.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. 2006. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* **125**: 315–326.

Carey KL, Richards SA, Lounsbury KM, Macara IG. 1996. Evidence using a green fluorescent protein-glucocorticoid receptor chimera that the Ran/TC4 GTPase

mediates an essential function independent of nuclear protein import. *The Journal of cell biology* **133**: 985–996.

Chambers SM, Fasano CA, Papapetrou EP, Tomishima M, Sadelain M, Studer L. 2009. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature biotechnology* **27**: 275–280.

Close BE, Colley KJ. 1998. In vivo autopolysialylation and localization of the polysialyltransferases PST and STX. *Journal of Biological Chemistry* **273**: 34586–34593.

Close BE, Mendiratta SS, geiger KM, broom LJ, Ho L-L, Colley KJ. 2003. The Minimal Structural Domains Required for Neural Cell Adhesion Molecule Polysialylation by PST/ST8Sia IV and STX/ST8Sia II. *Journal of Biological Chemistry* **278**: 30796–30805.

Curreli S, Arany Z, Gerardy-Schahn R, Mann D, Stamatos NM. 2007. Polysialylated Neuropilin-2 Is Expressed on the Surface of Human Dendritic Cells and Modulates Dendritic Cell-T Lymphocyte Interactions. *Journal of Biological Chemistry* **282**: 30346–30356.

Eckhardt M, Bukalo O, Chazal G, Wang L, Goridis C, Schachner M, Gerardy-Schahn R, Cremer H, Dityatev A. 2000. Mice deficient in the polysialyltransferase ST8SiaIV/PST-1 allow discrimination of the roles of neural cell adhesion molecule protein and polysialic acid in neural development and synaptic plasticity. *The Journal of Neuroscience* **20**: 5234–5244.

Eggers K, Werneburg S, Schertzinger A, Abeln M, Schiff M, Scharenberg MA, Burkhardt H, Muhlenhoff M, Hildebrandt H. 2011. Polysialic acid controls NCAM signals at cell-cell contacts to regulate focal adhesion independent from FGF receptor activity. *Journal of Cell Science* **124**: 3279–3291.

Fernando RI, Litzinger M, Trono P, Hamilton DH, Schlom J, Palena C. 2010. The T-box transcription factor Brachyury promotes epithelial-mesenchymal transition in human tumor cells. *Journal of clinical investigation* **120**: 533–544.

Finne J. 1982. Occurrence of unique polysialosyl carbohydrate units in glycoproteins of developing brain. *Journal of Biological Chemistry* **257**: 11966–11970.

Galuska SP, Geyer R, Gerardy-Schahn R, Muhlenhoff M, Geyer H. 2007. Enzyme-dependent Variations in the Polysialylation of the Neural Cell Adhesion Molecule (NCAM) in Vivo. *Journal of Biological Chemistry* **283**: 17–28.

Galuska SP, Rollenhagen M, Kaup M, Eggers K, Oltmann-Norden I, Schiff M, Hartmann M, Weinhold B, Hildebrandt H, Geyer R. 2010. Synaptic cell adhesion molecule SynCAM 1 is a target for polysialylation in postnatal mouse brain. *Proceedings of the National Academy of Sciences* **107**: 10250–10255.

Haltiwanger RS, Lowe JB. 2004. Role of glycosylation in development. *Annual review of biochemistry* **73**: 491–537.

Hartwell KA, Muir B, Reinhardt F, Carpenter AE, Sgroi DC, Weinberg RA. 2006. The Spemann organizer gene, Goosecoid, promotes tumor metastasis. *Proceedings of the National Academy of Sciences* **103**: 18969–18974.

Hildebrandt H, Becker C, Mürau M, Gerardy-Schahn R, Rahmann H. 1998. Heterogeneous expression of the polysialyltransferases ST8Sia II and ST8Sia IV during postnatal rat brain development. *Journal of Neurochemistry* **71**: 2339–2348.

Hu H, Tomasiewicz H, Magnuson T, Rutishauser U. 1996. The role of polysialic acid in migration of olfactory bulb interneuron precursors in the subventricular zone. *Neuron* **16**: 735–743.

Johnson CP, Fujimoto I, Rutishauser U, Leckband DE. 2005. Direct evidence that neural cell adhesion molecule (NCAM) polysialylation increases intermembrane repulsion and abrogates adhesion. *Journal of Biological Chemistry* **280**: 137–145.

Lackie PM, Zuber C, Roth J. 1994. Polysialic acid of the neural cell adhesion molecule (N-CAM) is widely expressed during organogenesis in mesodermal and endodermal derivatives. *Differentiation* **57**: 119–131.

Landmesser L, Dahm L, Tang J, Rutishauser U. 1990. Polysialic acid as a regulator of intramuscular nerve branching during embryonic development. *Neuron* **4**: 655–667.

Lehembre F, Yilmaz M, Wicki A, Schomber T, Strittmatter K, Ziegler D, Kren A, Went P, Derksen PWB, Berns A, et al. 2008. NCAM-induced focal adhesion assembly: a functional switch upon loss of E-cadherin. *EMBO Journal* **27**: 2603–2615.

Li J, Dai G, Cheng YB, Qi X, Geng MY. 2011. Polysialylation promotes neural cell adhesion molecule-mediated cell migration in a fibroblast growth factor receptor-dependent manner, but independent of adhesion capability. *Glycobiology* **21**: 1010–1018.

Liao B-Y, Zhang J. 2008. Null mutations in human and mouse orthologs frequently result in different phenotypes. *Proceedings of the National Academy of Sciences* **105**: 6987–6992.

Livingston BD, Jacobs JL, Glick MC, Troy FA. 1988. Extended polysialic acid chains (n > 55) in glycoproteins from human neuroblastoma cells. *Journal of Biological Chemistry* **263**: 9443–9448.

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-Guided Human Genome Engineering via Cas9. *Science* **339**: 823–826.

McLean AB, D'Amour KA, Jones KL, Krishnamoorthy M, Kulik MJ, Reynolds DM, Sheppard AM, Liu H, Xu Y, Baetge EE, et al. 2007. Activin A Efficiently Specifies Definitive Endoderm from Human Embryonic Stem Cells Only When Phosphatidylinositol 3-Kinase Signaling Is Suppressed. *Stem Cells* **25**: 29–38.

Menendez L, Yatskievych TA, Antin PB, Dalton S. 2011. Wnt signaling and a Smad pathway blockade direct the differentiation of human pluripotent stem cells to multipotent neural crest cells. *Proceedings of the National Academy of Sciences* **108**: 19240–19245.

Mühlenhoff M, Rollenhagen M, Werneburg S, Gerardy-Schahn R, Hildebrandt H. 2013. Polysialic acid: versatile modification of NCAM, SynCAM 1 and neuropilin-2. *Neurochemical Research* **38**: 1134–1143.

Ono S, Hane M, Kitajima K, Sato C. 2012. Novel Regulation of Fibroblast Growth Factor 2 (FGF2)-mediated Cell Growth by Polysialic Acid. *Journal of Biological Chemistry* **287**: 3710–3722.

Phillips GR, Krushel LA, Crossin KL. 1997. Developmental expression of two rat sialyltransferases that modify the neural cell adhesion molecule, N-CAM. *Developmental Brain Research* **102**: 143–155.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. 2013. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* **8**: 2281–2308.

Rosenbloom KR, Sloan CA, Malladi VS, Dreszer TR, Learned K, Kirkup VM, Wong MC, Maddren M, Fang R, Heitner SG, et al. 2012. ENCODE Data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Research* **41**: D56–D63.

Roth J, Zuber C, Wagner P, Blaha I, Bitter-Suermann D, Heitz PU. 1988. Presence of the long chain form of polysialic acid of the neural cell adhesion molecule in Wilms' tumor. Identification of a cell adhesion molecule as an oncodevelopmental antigen and implications for tumor histogenesis. *The American Journal of Pathology* **133**: 227.

Rutishauser U. 2008. Polysialic acid in the plasticity of the developing and adult vertebrate nervous system. *Nature Reviews Neuroscience* **9**: 26–35.

Rutishauser U, Watanabe M, Silver J, Troy FA, Vimr ER. 1985. Specific alteration of NCAM-mediated cell adhesion by an endoneuraminidase. *The Journal of cell biology* **101**: 1842–1849.

Schreiber SC, Giehl K, Kastilan C, Hasel C, Mühlenhoff M, Adler G, Wedlich D, Menke A. 2008. Polysialylated NCAM Represses E-Cadherin-Mediated Cell-Cell Adhesion in Pancreatic Tumor Cells. *Gastroenterology* **134**: 1555–1566.

Seidenfaden R, Krauter A, Hildebrandt H. 2006. The neural cell adhesion molecule NCAM regulates neuritogenesis by multiple mechanisms of interaction. *Neurochemistry International* **49**: 1–11.

Showell C, Binder O, Conlon FL. 2004. T-box genes in early embryogenesis. *Developmental dynamics* **229**: 201–218.

Swindall AF, Londono-Joshi AI, Schultz MJ, Fineberg N, Buchsbaum DJ, Bellis SL. 2013. ST6Gal-I Protein Expression Is Upregulated in Human Epithelial Tumors and Correlates with Stem Cell Markers in Normal Tissues and Colon Cancer Cell Lines. *Cancer Research* **73**: 2368–2378.

Tanaka F, Otake Y, Nakagawa T, Kawano Y, Miyahara R, Li M, Yanagihara K, Nakayama J, Fujimoto I, Ikenaka K. 2000. Expression of polysialic acid and STX, a human polysialyltransferase, is correlated with tumor progression in non-small cell lung cancer. *Cancer Research* **60**: 3072–3080.

Tateno H, Toyota M, Saito S, Onuma Y, Ito Y, Hiemori K, Fukumura M, Matsushima A, Nakanishi M, Ohnuma K, et al. 2011. Glycome Diagnosis of Human Induced Pluripotent Stem Cells Using Lectin Microarray. *Journal of Biological Chemistry* **286**: 20345–20353.

Varki A, RD C, JD E, HH F, P S, CR B, GW H, ME E. 2009. *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press.

Wang L, Schulz TC, Sherrer ES, Dauphin DS, Shin S, Nelson AM, Ware CB, Zhan M, Song CZ, Chen X, et al. 2007. Self-renewal of human embryonic stem cells requires insulin-like growth factor-1 receptor and ERBB2 receptor signaling. *Blood* **110**: 4111–4119.

Wang Y-C, Nakagawa M, Garitaonandia I, Slavin I, Altun G, Lacharite RM, Nazor KL, Tran HT, Lynch CL, Leonardo TR, et al. 2011. Specific lectin biomarkers for isolation of human pluripotent stem cells identified through array-based glycomic analysis. *Cell Research* **21**: 1551–1563.

Wang Y-C, Stein JW, Lynch CL, Tran HT, Lee C-Y, Coleman R, Hatch A, Antontsev VG, Chy HS, O'Brien CM, et al. 2015. Glycosyltransferase ST6GAL1 contributes to the regulation of pluripotency in human pluripotent stem cells. *Scientific reports* 1–13.

Weinhold B, Seidenfaden R, Röckle I, Mühlenhoff M, Schertzinger F, Conzelmann S, Marth JD, Gerardy-Schahn R, Hildebrandt H. 2005. Genetic Ablation of Polysialic Acid Causes Severe Neurodevelopmental Defects Rescued by Deletion of the Neural Cell Adhesion Molecule. *Journal of Biological Chemistry* **280**: 42971–42977.

CHAPTER 4

DETAILED EXPERIMENTAL PROCEDURES

MATERIALS AND METHODS

Stem cell culture

Human embryonic stem cells (hESCs) were grown on polystyrene culture dishes (Thermo Scientific) coated with Geltrex LDEV-Free hESC qualified reduced growth factor basement membrane matrix (Thermo Scientific) diluted at 1:200 from manufacturer stock in DMEM/F-12 (Corning). A chemically defined medium (DM) was used as a base medium and consisted of DMEM/F-12 (Corning) supplemented with 0.2% Probumin (Millipore), 1X Antibiotic Antimycotic solution (Corning), 1X MEM nonessential amino acids (Corning), 2mM glutagro supplement (Corning), 1X trace elements A/B/C (Corning), 50 μg/mL L-ascorbic acid, 10 μg/mL Transferrin (Athens Research & Technology), and 100 μM 2-Mercaptoethanol (gibco). Addition of 10 ng/mL Heregulin β-1, 10 ng/mL Activin A, 8 ng/mL bFGF2, and 200 ng/mL IGF-1 to DM constituted complete defined medium (CDM) that was used to maintain hESCs. Cells were grown in an incubator at constant conditions of 37°C and 5% CO_2 and were passaged every 4-5 days when they reached approximately 90% confluency by treatment with Accutase (ICT) at 37°C for 10 minutes, followed by centrifugation at 1000 rpm for 4 minutes, resuspension in CDM, quantification using a hemocytometer, and plating at a density of 50,000 cells/cm^2 onto Geltrex coated plates.

Differentiations

Definitive endoderm (DE) cells were generated from hESCs following a previously established protocol (McLean et al. 2007) by culturing in DM supplemented with 100 ng/mL Activin A and 8 ng/mL bFGF2 for 4 days with 25 ng/mL WNT3a present for the first 24 hours.

Mesoderm (ISL1) cells were generated from hESCs by culturing in CDM supplemented with 25 ng/mL WNT3a and 100 ng/mL BMP4 for 4 days.

Ectoderm cells were generated following a neural crest protocol (Menendez et al. 2011) by culturing in DM supplemented with 10 ng/mL Heregulin β-1, 8 ng/mL bFGF2, and 200 ng/mL IGF-1, 20 μM SB 43152 (Tocris), and 2 μM BIO (GSK3 inhibitor IX, Calbiochem) for 8 days.

Neural progenitor cells were generated from hESCs by culturing in DM supplemented with 10 ng/mL Heregulin β-1, 200 ng/mL IGF-1, 20 μM SB 43152 (Tocris), and 500 nM LDN-193189 (Sigma) for 6 days.

qRT-PCR

Cultured cells were harvested using Accutase and centrifuged at 1000 rpm for 4 minutes prior to RNA isolation using the E.Z.N.A. RNA isolation kit (Omega). RNA was quantified using a Biotek Synergy 2 plate reader and 1 μg of RNA was used to make cDNA with the Iscript cDNA synthesis kit (Bio-Rad) followed by dilution in molecular grade water to a total volume of 500 μL for analysis. Genes were assayed using Taqman primers (Life Technologies) in a 5 μL reaction consisting of 2 μL cDNA, 2.5 μL PCR Mastermix (Life Technologies), 0.25 μL Taqman primer, and 0.25 μL molecular grade

water. Samples were aliquoted into a 384 well plate and analyzed using a ViiA 7 Real-Time PCR System (Life Technologies). Reactions were performed in triplicate with expression normalized to reference genes GAPDH or GUSB and data were shown as mean ± SEM.

Flow cytometry

Cells for analysis were harvested using Accutase for 10 minutes at 37°C, quantified using a hemocytometer, and aliquoted into 5 mL round bottom tubes. 1 million cells were used per tube and were filtered by pipetting through the 35 μm cell strainer cap. Cells were washed once using 3 mL of PBS containing 0.2% BSA, centrifuged at 1000 rpm for 4 minutes, and the PBS aspirated leaving a small pellet. Antibodies were diluted in the PBS + 0.2% BSA solution and a volume of 20 μL was used to resuspend the pellet. Cells were then incubated in antibody solution for 20 minutes, washed with 3 mL of PBS + 0.2% BSA, centrifuged at 1000 rpm for 4 minutes, and aspirated. This process was repeated until cells had been labeled with all primary and secondary antibodies. Each cell sample for analysis had an isotype control sample that was labeled using the corresponding isotype fluorophores or secondary antibodies. After the final wash step, cells were resuspended in 0.5 mL of PBS + 0.2% BSA and analyzed using a Beckman Coulter Cyan or Hypercyan cytometer followed by analysis using FlowJo software.

Western blotting

Cells for Western blotting were typically grown in 60 mm plates and were collected by aspirating the culture medium from the plate, adding 1 mL of cold PBS, scraping the

plate surface with a cell scraper, and pipetting the cell mixture into a 1.5 mL tube. Cells were centrifuged at 1000 rpm for 4 minutes, PBS aspirated off the cell pellet, snap frozen in liquid Nitrogen for 1 minute, and stored at -80°C. Protein was isolated from the cells by resuspension in 20-50 μL of RIPA lysis buffer (Sigma) containing protease inhibitor (Roche), phosphatase inhibitor (Calbiochem), and 100 mM dithithreitol. Cells were incubated in lysis buffer on ice for 30 minutes followed by centrifugation at 20,000 x g and extraction of the protein lysate supernatant into a new tube. Protein was quantified using a Bradford assay measuring absorbance at 595nm using a Biotek Synergy 2 plate reader. Sample buffer was made using 950 μL Laemmli buffer supplemented with 50 μL of 2-mercaptoethanol and was added to the samples in equal volume to the protein lysate.

Electrophoresis was done using 20-40 μg of protein loaded into each well of Bolt bis-Tris precast gels (Life Technologies) and followed by transfer onto nitrocellulose membranes. The membranes were blocked using 2% nonfat milk dissolved in 0.5% TBST for 1 hour at room temperature. Primary antibodies were diluted (Table 1) in the blocking solution and were incubated with the membrane overnight at 4°C. The following day membranes were washed 3 times for 5 minutes each in 0.5% TBST on a shaker to remove excess antibody. HRP secondary antibodies were diluted 1:2000 in blocking solution and incubated on the membrane for 1 hour at room temperature followed by washing 3 times with 0.5% TBST. Amersham ECL detection reagent (GE) was then added to the membrane for 1 minute and removed prior to detection using Amersham hyperfilm (GE).

shRNA knockdowns

TRC shRNA lentiviral plasmid sets (GE Dharmacon) were ordered for PST (Cat. RHS4533-EG7903), NCAM (RHS4533-EG4684) and GSC (Cat. RHS4533-EG145258). Each set contained 4-6 plasmids targeting different sequences within the gene. Lentiviral particles were generated following the Thermo Scientific trans-lentiviral packaging kit using 5.5 million HEK293-T cells plated on a 10 cm culture dish in high glucose DMEM supplemented with 10% FBS and 2mM L-glutamine. One day after plating, cells were transfected with 42 μg of shRNA plasmid, 30 μL of Trans-lentiviral packaging mix, 105 μL of $CaCl_2$, and 1050 μL of 2X HBSS brought to a total volume of 2 mL with molecular grade water. 16 hours following transfection cells were switched to 14 mL of reduced serum medium consisting of high glucose DMEM, 5% FBS, and 2mM L-glutamine and incubated for 48 hours. During the incubation period the cells were visually inspected on a microscope for morphology and cell death; high levels of cell death and poor morphology were indicative of efficient viral production. The supernatant was collected 48 hours after addition of reduced serum medium and was centrifuged at 1600 x g for 10 minutes followed by filtration using a 0.22 μm Steriflip tube. Lenti-X concentrator (Clontech) was then added at 1:3 the volume of the viral supernatant and incubated at 4 °C overnight. The samples were centrifuged at 1500 x g for 45 minutes at 4 °C the next day yielding a small pellet of concentrated viral particles. The supernatant was discarded and the pellet was resuspended in DMEM/F12 according to 100 μL per 10 cm plate used and single use 20 μL aliquots were frozen and stored at -80°C. Titering of the concentrated virus was done using a qRT-PCR based lentiviral titering kit (Mellgen Labs) using SYBR Green (Bio-Rad) and 1 μL of viral supernatant.

Knockdown experiments were done in WA09 hESCs plated in 24 well plates at a density of 40,000 cells/cm^2. Cells were transduced 24 hours after plating by first adding growth media supplemented with 6 ng/μL polybrene, incubating at 37°C for 30 minutes, followed by addition of concentrated viral particles at 1-5 MOI. 24 hours after transduction the culture medium was supplemented with 3 μg/mL of Puromycin to eliminate non-transduced cells. Cells were cultured with media containing Puromycin until they were collected (typically 3-5 days). shRNA knockdown efficiency was calculated as the percent reduction of mRNA levels assayed by qRT-PCR compared to a control shRNA.

ChIP assays

Cells were harvested using Accutase treatment for 10 minutes at 37°C, quantified using a hemocytometer, and 10 million cells were aliquoted into 15 mL tubes and washed with PBS. The cells were resuspended in 1 mL of PBS and then cross-linked by adding 27.5 μL of 37% formaldehyde and incubating for 10 minutes followed by quenching with 50 μL of 2.5M glycine for 5 minutes. Cells were then washed with PBS, snap frozen in liquid nitrogen, and stored at -80°C.

Lysis of the cells was done following the Agilent mammalian ChIP protocol using lysis buffers (LB) 1, 2, and 3. Cells were resuspended in 5 mL of LB1, rocked for 10 minutes at 4°C, and centrifuged at 1350 x g at 4°C for 4 minutes. This was repeated using LB2 and after centrifugation the pellet was resuspended in 1 mL of LB3. The genomic DNA was then sonicated to approximately 500 bp using a Covaris S220 focused

ultrasonicator for 8 minutes at 5% duty cycle. Following sonication the lysate was stored at -80°C.

Protein G magnetic beads (Life Technologies) were incubated overnight at 4°C with 10 μg of antibody, washed 3 times with 1 mL RIPA wash buffer, and resuspended in 100 μL RIPA wash buffer. Sonicated lysate was then thawed and aliquoted into the tubes containing antibody-bound magnetic beads and was placed in a rotator overnight at 4°C. 50 μL of sonicated lysate was aliquoted into a separate tube and frozen at -20°C to serve as the input control for each sample. After overnight immunoprecipitation, samples were washed 5 times each with 1 mL RIPA wash buffer and eluted using 210 μL of Agilent elution buffer. The elution was done by incubating the magnetic beads at 65°C for 15 minutes with vortexing every 2 minutes and cross-links were reversed by incubating overnight at 65°C.

200 μL of TE was added to the elution samples the next day followed by addition of 4 μL of RNAse A and incubation at 37 °C for 2 hours. Protein was then degraded using Proteinase K for 30 minutes at 55 °C and DNA was purified by phenol chloroform extraction and ethanol precipitation. The precipitated DNA pellet was then resuspended in 100 μL molecular grade water and stored at -20 °C. ChIP-PCR was done using SYBR Green (Bio-Rad) and genomic primers with detection from a Viaa7 real time PCR system (Life Technologies).

CRISPR-CAS9 GENOME EDITING

The overall strategy for CRISPR-Cas9 genome editing (Figure 4.1) involved (1) designing and cloning single guide RNAs (sgRNAs) to target a specific locus, (2)

generation of a donor plasmid to facilitate homology directed repair after Cas9-induced double strand breaks, (3) electroporation of Cas9, sgRNAs, and donor plasmid into cells, (4) drug selection of a polyclonal population of CRISPR edited cells, (5) clonal isolation and expansion of CRISPR edited cells, and (6) identification and validation of properly edited clonal cell lines. Each of these steps will be discussed in detail with the knockout of ST8SIA4 being an example.

Guide RNA design and construction

Genes targeted for CRISPR-Cas9 knockout were first analyzed by uploading the genomic locus from the NCBI database (www.ncbi.nlm.nih.gov) to SnapGene software to determine the region within the gene to be edited. The strategy for knocking out ST8SIA4 was designed to delete a 300 base pair segment of the first exon containing the start codon for translation of the protein using two sgRNAs on each side (Figure 4.2). Generation of sgRNAs was done by copying the genomic region of interest and probing for potential sgRNAs using the Optimized CRISPR Design web-based portal (crispr.mit.edu/), which analyzes genomic sequences for protospacer adjacent motifs (PAMs) and scores them based on frequency of off target binding throughout the genome (Ran et al. 2013). sgRNA 20 base pair sequences (without PAMs) with off target scores above 90 were selected and input into a spreadsheet along with their respective reverse complement sequences. sgRNA sequences were ordered as DNA oligos and were designed to be cloned into a pSpCas9(BB)-2A-GFP plasmid (Addgene plasmid #48138, Figure 4.3), referred to here as pCas9-GFP, using BbsI compatible overhangs (Figure 4.3). The sequences of the oligos followed the form

CACCGGNNNNNNNNNNNNNNNNNNN where the Ns represent the sgRNA sequence and CCNNNNNNNNNNNNNNNNNNNNCAAA where the Ns represent the sgRNA reverse complement. The oligos were then annealed and phosphorylated using T4 PNK (New England Biolabs) and ligated into BbsI digested pCas9-GFP (Figure 4.3B) at 16°C for 1 hour using a ligation mix (Takara). Ligated plasmid was then transformed into High Efficiency competent cells (New England Biolabs) and clones generated. Initial validation of the sgRNA plasmids was performed by BbsI digestion and comparison to pCas9-GFP lacking sgRNA insert. pCas9-GFP is digested by BbsI and yields one band on an agarose gel while sgRNA-pCas9-GFP plasmids lack the BbsI site as a consequence of cloning and yield two bands on a gel from supercoiled and linear plasmid. Final validation was done by Sanger sequencing to confirm the proper sgRNA sequence was present in sgRNA-pCas9-GFP plasmids.

Donor plasmid design

Generation of a donor plasmid was done to provide both a fluorescence marker and drug resistance to select for CRISPR edited cells more easily. The donor plasmid for ST8SIA4 knockout was designed by using two 1 kb homology arms consisting of the genomic regions upstream and downstream of the sgRNA sequences beginning within 25 bp of the Cas9 induced double strand breaks (Figure 4.2). A 20 bp spacer region containing restriction sites compatible for subcloning a pCAG-BFP-Zeo cassette into the donor construct was designed in between the homology arms. This sequence block was ordered from Genscript as a custom gene synthesis project inserted into a pUC57 construct. The final donor plasmid was completed by ligation of the pCAG-BFP-Zeo cassette into the

custom pUC57 construct and validation was carried out by restriction fragment length polymorphism analysis.

Electroporation of sgRNA-Cas9 and donor plamids

Once the sgRNA-pCAS9-GFP and donor plasmids were generated they were electroporated together into WA09 hESCs using a Neon electroporation system (Life Technologies). The electroporation reactions were carried out using 1.5 million WA09 cells and 7.5 μg total plasmid DNA (1.87 μg sgRNA-Cas9-GFP plasmid 1, 1.87 μg sgRNA-Cas9-GFP plasmid 2, 3.75 μg pCAG-BFP-Zeo donor plasmid) with instrument parameters of 1050V, pulse width 30, and 2 pulses. Each reaction was done in 100 μL of electroporation buffer R and cells were prepared according to the manufacturers instructions including a single wash with PBS (without Calcium or Magnesium) prior to resuspension in buffer R. After electroporation cells were plated into a single well of a 6-well plate containing CDM lacking antibiotic and supplemented with 10 μM Rho-associaed kinase inhibitor (ROCK) Y-27632 (Tocris) to reduce apoptosis and increase plating efficiency. This reaction was repeated a total of 12 times yielding two 6-well plates to allow for a large starting population.

Selection for CRISPR polyclonal cell population

Selection began 24 hours after electroporation cells using CDM supplemented with100 μg/mL Zeocin (Sigma) to select for cells that stably integrated the pCAG-BFP-Zeo cassette from the donor construct. When the effects of Zeocin selection were apparent, after 4-6 days of selection, the cells from a 6 well plate were passaged into a 35 mm plate

to maintain proper cell density. This process of passaging into smaller wells continued every 3-4 days until cell death diminished and drug resistant cells began amplifying. Cells were then amplified to >5 million cells to be frozen down into stocks of polyclonal Zeocin resistant cells labeled as "PST-CR." The PST-CR cells were then assayed for BFP expression using flow cytometry and showed >99% BFP expression. Genomic DNA was extracted from the PST-CR cells and genomic PCR was done using primer sets designed to identify the wild type (WT) ST8SIA4 locus and the properly CRISPR edited ST8SIA4 locus lacking the 300 bp first exon region but containing the CAG-BFP-Zeo cassette (Supplemental Figure 3.7). The results of the genomic PCR indicated the presence of both WT and CRISPR edited ST8SIA4 loci. This result was expected and indicated that properly CRISPR edited cells could be isolated through clonal expansion.

Clonal cell line isolation

The generation of clonal cell lines was done by pre-treating the polyclonal cell culture for 30 minutes with CDM containing 10 μM ROCK inhibitor, releasing cells from the plate using Accutase for 10 minutes, filtering cells using a 35 μm cell strainer, and sorting single cells using a Beckman Coulter MoFLo XDP into each well of a 96 well plate coated with Geltrex (1:200 dilution in DMEM/F12). The media used for plating and culturing on 96 well plates was a 50/50 mixture of CDM and 24-hour hESC conditioned media (CM) supplemented with 10 μM ROCK inhibitor. The conditioned media was collected from a 10 cm plate of hESCs 24 hours after addition, filtered using a Steriflip 0.22 μm filter unit, and supplemented with 8 ng/mL bFGF2. Cells were plated in 125 μL per well of the CDM/CM 50/50 mix and media was changed every other day using a

volume of 80 μL. Wells were examined for colonies under a microscope 1-2 weeks after sorting and wells containing colonies were marked. The number of colonies formed was variable between experiments and cell types and ranged between 5-40 wells of a 96 well plate. Colonies were expanded in the 96 well plate until the colony density became very high; typically the cells would not expand to fill the entire well but would begin to grow on top of each other within the colony. Once this occurred, cells were passaged using Accutase and plated in a single well of a 24 well plate in the same CDM/CM mix media. These cells were then expanded until they filled the 24 well plate and were passaged using Accutase and plated into a 35 mm plate containing CDM. The media was switched at this point as cell density reached normal levels and no longer necessitated conditioned media.

Validation of CRISPR cell lines

Screening of CRISPR clonal cell lines was performed to identify those that had properly edited the ST8SIA4 locus according to the design. The first step was a genomic PCR based assay to identify the presence of the BFP-Zeo cassette in the genomic locus. To do this, primers specific to regions in the BFP-Zeo sequence were designed along with primers from the genomic regions beyond the homology arm sequence (Supplemental Figure 3.7B). Cells were collected during clonal isolation culture from either a 24 well or 6 well plate and genomic DNA was extracted using a Wizard genomic DNA purification kit (Promega) and genomic PCR was done using a Phusion high fideligy PCR kit (New England Biolabs). The products of the PCR were then analyzed using a 1.0% agarose gel and WT WA09 cells did not generate a PCR product for this reaction but correctly edited

cell lines yielded the predicted band of 1269 bp (Supplemental Figure 3.7D). Once it was determined that a cell line contained the BFP-Zeo cassette in the proper location, a second PCR assay was done to confirm that the wild type locus is lost in both alleles. To accomplish this, genomic primers were designed for the region to be excised between the gRNAs (Supplemental Figure 3.7A). The results of this assay yielded a small 119 bp band for the WA09 control but no band for CRISPR cell lines (Supplemental Figure 3.7C). Cell lines that satisfied the PCR criteria were considered to be ST8SIA4 knockout cell lines. qRT-PCR for PST also confirmed this as the knockout cell lines showed a loss of PST transcript due to the gene disruption.

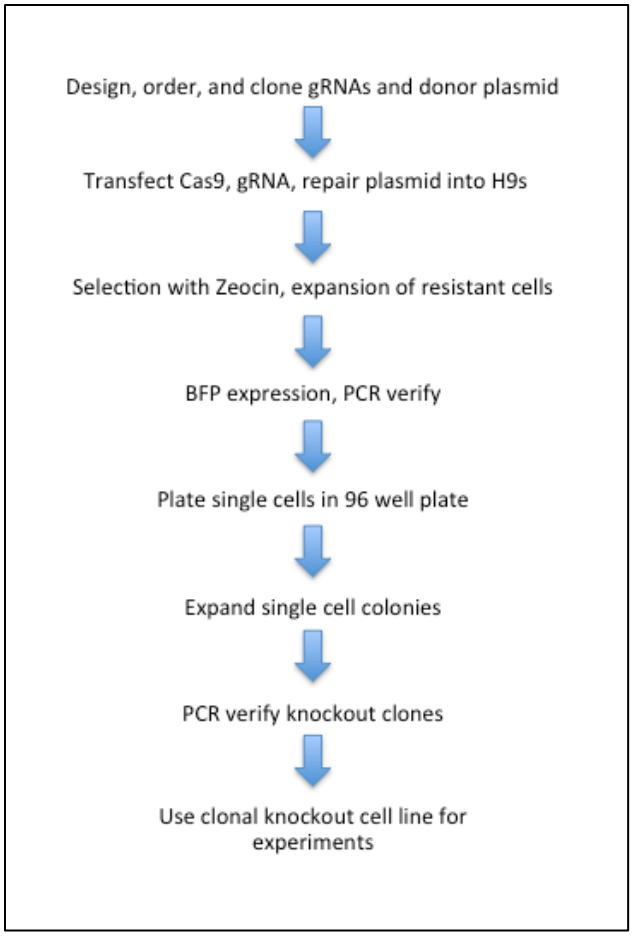


Figure 4.1. CRISPR-Cas9 genome editing workflow.

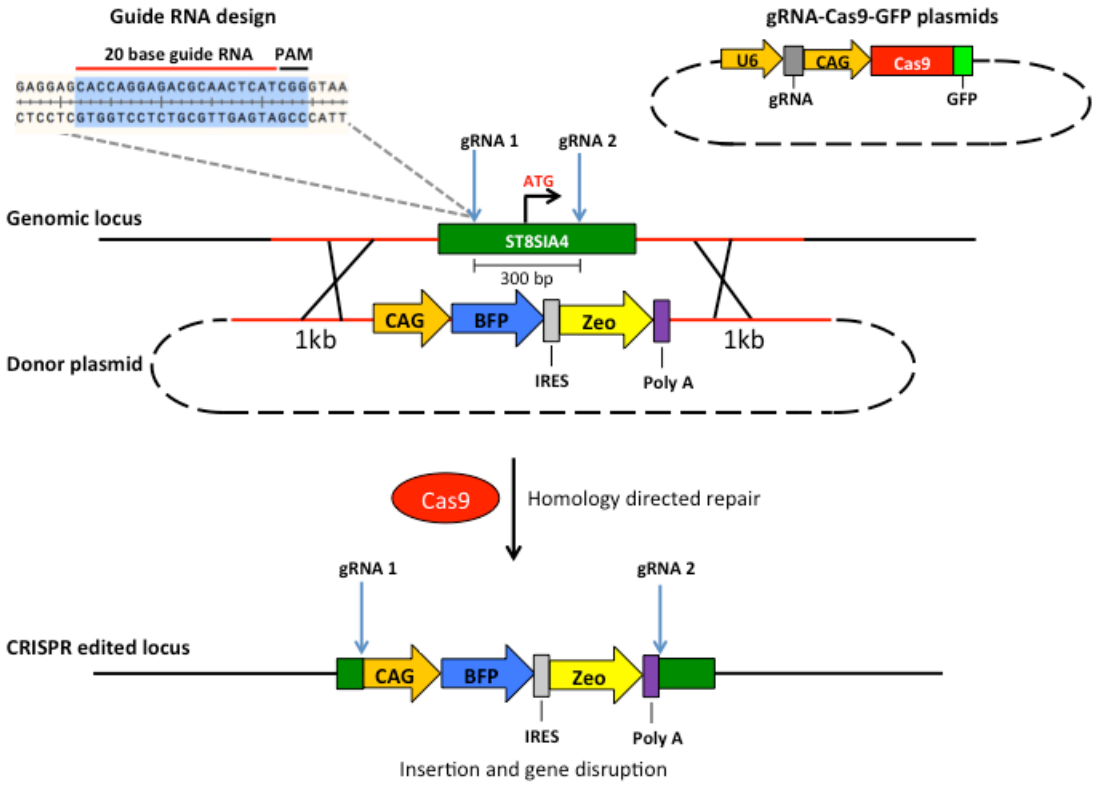


Figure 4.2. CRISPR editing strategy of the ST8SIA4 genomic locus. ST8SIA4 was targeted with two guide RNAs flanking the translation start codon in the first exon. The 300 bp region was excised and replaced with a CAG-BFP-Zeo resistance cassette through homology directed repair.

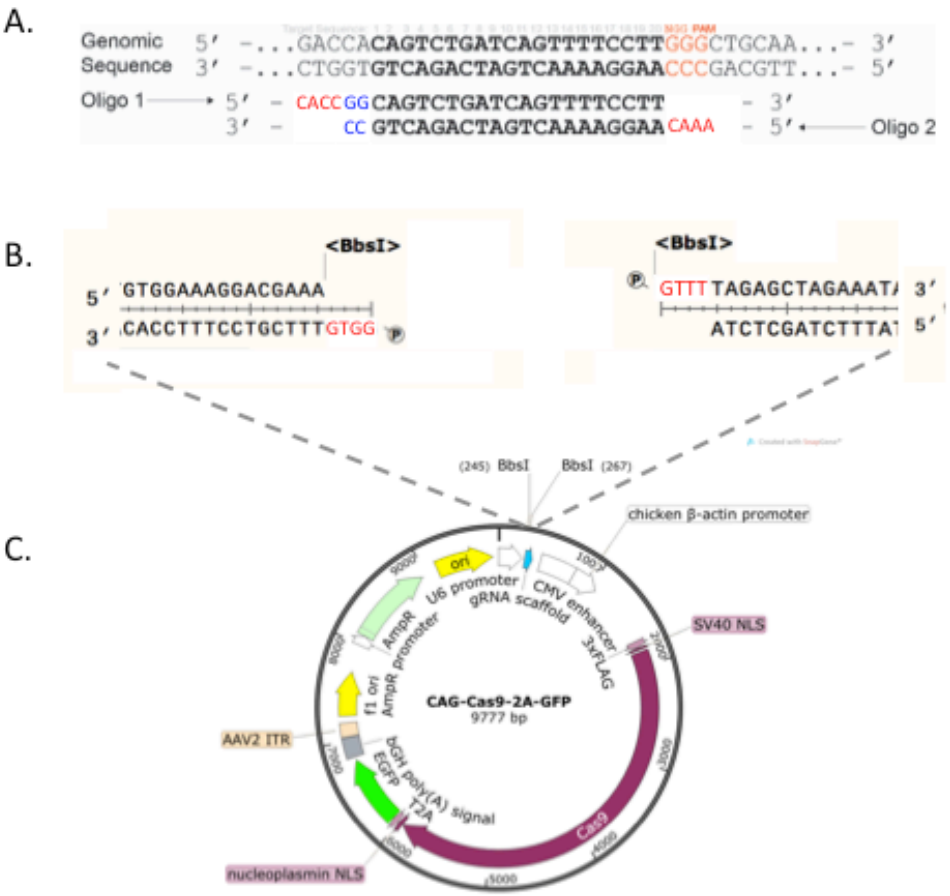


Figure 4.3. Design of oligos for cloning into pCAG-Cas9-GFP. (A) Guide RNAs were designed from genomic sequences and were ordered as oligos containing overhangs compatible for cloning into BbsI restriction sites. (B) Illustration of BbsI digested sequence from pCAG-Cas9-GFP. (C) Plasmid map of pCAG-Cas9-GFP.

Table 1 – ANTIBODIES USED

Antibody	Source	Catalog no.	Uses*	Dilution (IB)	Dilution (FC)	Dilution (IHC)
PSA		mAb 735	IB, IHC, FC	1:1000	1:100	1:250
SOX17	R&D Systems	AF1924	IB, IHC	1:1000	-	1:100
T	R&D Systems	AF2085	IHC	-	-	1:100
NANOG	Genetex	GTX100863	IHC	-	-	1:100
Giantin	abcam	ab24586	IHC	-	-	1:100
α-Mannosidase II	abcam	ab12277	IHC	-	-	1:100
CDK2	Santa Cruz	SC-163	IB	1:2000	-	-
NCAM	Santa Cruz	123A8	IB, IHC, FC	1:500	1:50	1:50
SNAI1	R&D Systems	AF3639	IHC	-	-	1:100
FOXA2	Millipore	07-633	IB, IHC	1:1000	-	1:100
CXCR4	eBioscience	17-9999	FC	-	1X	-
SSEA3	eBioscience	53-8833	FC	-	1X	-
V5	Invitrogen	46-0705	IHC	-	-	1:100
HA	Sigma	H3663-200UL	IHC	-	-	1:100
GSC	R&D Systems	AF4086	ChIP	-	-	-
H3K4me3	Abcam	Ab8580	ChIP	-	-	-
H3K27me3	Abcam	Ab6002	ChIP	-	-	-
IgG	Abcam	Ab46540	ChIP	-	-	-
E-Cadherin	R&D Systems	AF648	IHC	-	-	1:100

*Uses key: IB – Immunoblot, IHC – Immunohistochemistry, FC – Flow cytometry, ChIP

– chromatin immunoprecipitation

REFERENCES

McLean AB, D'Amour KA, Jones KL, Krishnamoorthy M, Kulik MJ, Reynolds DM, Sheppard AM, Liu H, Xu Y, Baetge EE, et al. 2007. Activin A Efficiently Specifies Definitive Endoderm from Human Embryonic Stem Cells Only When Phosphatidylinositol 3-Kinase Signaling Is Suppressed. *Stem Cells* **25**: 29–38.

Menendez L, Yatskievych TA, Antin PB, Dalton S. 2011. Wnt signaling and a Smad pathway blockade direct the differentiation of human pluripotent stem cells to multipotent neural crest cells. *Proceedings of the National Academy of Sciences* **108**: 19240–19245.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. 2013. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* **8**: 2281–2308.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

The presence of polysialic acid on the surface of cells throughout embryonic development has been observed for decades in rodents yet has not been described in detail in human development. The vast majority of these studies focused on the role of PSA in the developing nervous system where it is crucial for proper migratory patterns (Rutishauser 2008) but have paid little attention to the abundant expression of PSA occurring in early embryogenesis (Lackie et al. 1994; Angata et al. 1997; Phillips et al. 1997; Ong et al. 1998). Because of this, the existing paradigm of polysialylation is limited to its anti-adhesive effects, facilitation of migration and invasion, and cell signaling potential.

In our studies we have outlined a specific role for polysialylation in the germ layer specification of hPSCs. This work definitively shows PSA expression occurs at the earliest stages of lineage specification as pluripotent cells first begin to differentiate into progenitors of all three germ layers. Our results also demonstrate that this is a tightly regulated process that is under the control of previously described developmental transcriptional networks. The observation that loss of PST gene expression leads to negative effects is perhaps the most interesting finding as it suggests that polysialylation is a critical component of the differentiation program and represents the first observation of a carbohydrate structure participating in early cell fate decisions.

The exact mechanism through which PSA is acting to influence differentiation is not currently known but we speculate that it is via altered cell signaling caused by aberrant, unmodified NCAM on the cell surface. Investigations into PST, STX, and NCAM null mice have previously shown that NCAM in its unpolysialylated state generates severe developmental phenotypes that are rescued by mice that lack NCAM (Weinhold et al. 2005), and it is intriguing to speculate whether the same effect is being observed in hPSCs.

Our studies to date have focused on the genetic knockout of PST and its effects on endoderm differentiation but this raises additional questions of how knockouts of STX and NCAM would affect other lineages. From our results it would be expected that STX knockout would show differentiation defects in ectoderm cells given that it is the sole polysialyltransferase expressed in that cell type. Additionally, it would be interesting to see if cells completely devoid of PSA through PST and STX double knockout would be able to differentiate properly and if the adverse effects would be alleviated by a knockout of NCAM similar to murine models (Weinhold et al. 2005). These studies would help elucidate role of PSA in embryonic development and perhaps gain greater perspective on the studies done in knockout mice.

Interestingly, our results coincide with other studies outlining specific roles for sialic acid in stem cells. α2-6 sialylation has been shown by multiple groups to be highly expressed on the surface of hPSCs (Alisson-Silva et al. 2014; Hasehira et al. 2012; Wang et al. 2011; Satomaa et al. 2009), and a recent study has shown that this sialylation pattern is required for the maintenance of pluripotency (Wang et al. 2015). Taken together with our findings it suggests that varying sialic acid expression may represent a

dynamic carbohydrate-based mechanism that functions during early embryonic development to control cell fate decisions in pluripotency and differentiation.

REFERENCES

Alisson-Silva F, de Carvalho Rodrigues D, Vairo L, Asensi KD, Vasconcelos-dos-Santos A, Mantuano NR, Dias WB, Rondinelli E, Goldenberg RCDS, Urmenyi TP, et al. 2014. Evidences for the involvement of cell surface glycans in stem cell pluripotency and differentiation. *Glycobiology* **24**: 458–468.

Angata K, Nakayama J, Fredette B, Chong K, Ranscht B, Fukuda M. 1997. Human STX polysialyltransferase forms the embryonic form of the neural cell adhesion molecule - Tissue-specific expression, neurite outgrowth, and chromosomal localization in comparison with another polysialyltransferase, PST. *Journal of Biological Chemistry* **272**: 7182–7190.

Hasehira K, Tateno H, Onuma Y, Ito Y, Asashima M, Hirabayashi J. 2012. Structural and Quantitative Evidence for Dynamic Glycome Shift on Production of Induced Pluripotent Stem Cells. *Molecular & Cellular Proteomics* **11**: 1913–1923.

Lackie PM, Zuber C, Roth J. 1994. Polysialic acid of the neural cell adhesion molecule (N-CAM) is widely expressed during organogenesis in mesodermal and endodermal derivatives. *Differentiation* **57**: 119–131.

Ong E, Nakayama J, Angata K, Reyes L, Katsuyama T, Arai Y, Fukuda M. 1998. Developmental regulation of polysialic acid synthesis in mouse directed by two polysialyltransferases, PST and STX. ....

Phillips GR, Krushel LA, Crossin KL. 1997. Developmental expression of two rat sialyltransferases that modify the neural cell adhesion molecule, N-CAM. *Developmental Brain Research* **102**: 143–155.

Rutishauser U. 2008. Polysialic acid in the plasticity of the developing and adult vertebrate nervous system. *Nature Reviews Neuroscience* **9**: 26–35.

Satomaa T, Heiskanen A, Mikkola M, Olsson C, Blomqvist M, Tiittanen M, Jaatinen T, Aitio O, Olonen A, Helin J, et al. 2009. The N-glycome of human embryonic stem cells. *BMC Cell Biol* **10**: 42.

Wang Y-C, Nakagawa M, Garitaonandia I, Slavin I, Altun G, Lacharite RM, Nazor KL, Tran HT, Lynch CL, Leonardo TR, et al. 2011. Specific lectin biomarkers for isolation of human pluripotent stem cells identified through array-based glycomic analysis. *Cell Research* **21**: 1551–1563.

Wang Y-C, Stein JW, Lynch CL, Tran HT, Lee C-Y, Coleman R, Hatch A, Antontsev VG, Chy HS, O'Brien CM, et al. 2015. Glycosyltransferase ST6GAL1 contributes to the regulation of pluripotency in human pluripotent stem cells. *Scientific reports* 1–13.

Weinhold B, Seidenfaden R, Röckle I, Mühlenhoff M, Schertzinger F, Conzelmann S, Marth JD, Gerardy-Schahn R, Hildebrandt H. 2005. Genetic Ablation of Polysialic Acid Causes Severe Neurodevelopmental Defects Rescued by Deletion of the Neural Cell Adhesion Molecule. *Journal of Biological Chemistry* **280**: 42971–42977.