Title: Overlapping and unique substrate specificities of ST3GAL1 and 2 during hematopoietic and megakaryocytic differentiation

Short Title: Novel glycoprotein substrates for ST3GAL2

Authors: Nanyan Zhang, Siying Lin, Weiguo Cui, and Peter J. Newman

Supplemental methods and figures
Supplemental Figure 1

Figure S1. mRNA expression of major α2,3- and α2,6-sialyltransferases and molecules responsible for T antigen synthesis in human iPSC-derived HPCs and MKs, and human bone marrow HSCs and MKs. (A) Analysis of the published RNA sequencing data set of human iPSC-derived HPCs and MKs (GSE119828). (B) UMAP embedding single-cell RNA sequencing data isolated from bone marrow from 12 subjects (GSE120221). 14 distinct clusters were identified. Cell types in each cluster were identified based on differential gene expression (specified in the dot plot), and clusters 9 was identified as HSC while cluster 11 was identified as MK. (C) UMAP embedding single-cell RNA sequencing data of HSC and MK. To further quantify differential gene expression of HSC and MK clusters, clusters 9 and 10 were singled out from other cell types and re-analyzed by principle-component analysis, resulted in pure clusters of MK and HSC. (D) Dot plot showing gene expression for HSC and MK identification. (E) Sialyltransferases and T-synthase expression in bone marrow HSCs and MKs.
Figure S2. PCR genotyping of ST3GAL1 and ST3GAL2 KO clones. (A) Schematic illustration of the ST3GAL1 and ST3GAL2 loci, showing the location of the gRNA binding sites (red arrows) and primer binding sites (blue arrows) for PCR genotyping. (B) Genomic DNA, isolated from iPSCs that had been transfected with px459 V2.0-gRNA1, px459 V2.0-gRNA2 targeting ST3GAL1 or ST3GAL2 was PCR amplified to confirm bi-allelic deletion at the targeted locus. Red arrowheads indicate the expected fragment sizes of a WT clone versus a KO clone.
Supplemental Figure 3

Figure S3. Disruption of ST3GAL1 and ST3GAL2 significantly reduces $\alpha_2,3$-, but not $\alpha_2,6$-, sialylation during iPSC differentiation. (A) Lectin binding to iPSCs. PNA lectin is specific for the non-sialylated epitope Galβ1-3GalNAc. RCAI lectin binds terminal galactose. MALII lectin is specific for $\alpha_2,3$-linked sialic acid. SNA lectin is specific for $\alpha_2,6$-linked sialic acid. (B) Lectin binding to iPSC-derived HPCs. (C) Lectin binding to iPSC-derived MKs. Values represent the means ± SEM from three or four independent experiments. *P<0.05, **P<0.01 by one-way ANOVA with Dunnett’s test compared with WT cells.
Figure S4. Disruption of ST3GAL1 and ST3GAL2 does not affect hematopoietic mesoderm development. Flow cytometric analysis shows development of $\text{KDR}^+ / \text{CD31}^+$ hematopoietic mesoderm from the indicated cell lines on Day 6.
Figure S5. Disruption of ST3GAL1 and ST3GAL2 results in exposure of altered O-glycans on the surface of iPSC-derived HPCs. (A) Western blot demonstrating that the loss of ST3GAL1 or ST3GAL2 does not cause dramatic compensatory upregulation of the other enzyme in iPSC-derived HPCs. (B) Quantitated PNA binding to iPSC-derived HPCs in Figure 3C. Values represent the means ± SEM from three independent experiments. (C) Flow cytometric analysis of PNA binding to HPCs derived from different KO clones as that used in Figure 3C to rule out the possibility that the phenotype is caused by an off-target effect of CRISPR. (D) Quantitated PNA blot of WCL from iPSC-derived HPCs in Figure 3D. Values represent the means ± SEM from three independent experiments. **P<0.01 , *P<0.05 by one-way ANOVA with Dunnett’s test compared with WT cells.
Supplemental Figure 6

Figure S6. Disruption of ST3GAL1 and ST3GAL2 does not affect MK maturation, but results in exposure of altered O-glycans on the surface of iPSC-derived MKs. (A) and (B) Ploidy analysis of iPSC-derived MKs. (C) Quantitated PNA binding to iPSC-derived MKs in Figure 4C. Values represent the means ± SEM from three independent experiments. (D) Flowcytometric analysis of PNA binding to MKs derived from different KO clones as that used in Figure 4C to rule out the possibility that the phenotype is caused by an off-target effect of CRISPR. (E) Quantitated PNA blot of WCL from iPSC-derived MKs in Figure 4D. Values represent the means ± SEM from three independent experiments. **P<0.01, *P<0.05 by one-way ANOVA with Dunnett’s test compared with WT cells. (F) Coimmunoprecipitation of GPIIb-IIIa complex with anti-GPIIIa monoclonal antibody from lysates of different KO clones as that used in Figure 4F to rule out the possibility that the phenotype resulted from an off-target effect of CRISPR. Immunoblot was done with either PNA or rabbit anti-GPIIb polyclonal antibody.

**Zhang et al.**
Figure S7. Disruption of ST3GAL1 and ST3GAL2 does not affect thrombin-induced GPIIb-IIIa activation in MKs. (A) and (B) PAC1 binding to iPSC-derived MKs upon thrombin (0.1 U/ml) stimulation. Values represent the means ± SEM from three independent experiments. (C) and (D) Fibrinogen binding to iPSC-derived MKs upon thrombin (0.1 U/ml) stimulation. Values represent the means ± SEM from three independent experiments.
Figure S8. Partially redundant role of ST3GAL1 and ST3GAL2 in erythropoiesis.

(A) Flow cytometric analysis showed generation of transferrin receptor (CD71)-positive erythroblasts from the indicated iPSC lines. (B) Quantitated yield of CD71^+ erythroblasts from the indicated iPS cell lines. Values represent the means± SEM from three independent experiments. (C) Flow cytometric analysis of PNA binding to iPSC-derived erythroblasts. Numbers indicate median fluorescence intensities.
Figure S9. ST3GAL1 and ST3GAL2 function as Core 1 sialyltransferases in human cells. The Core 2 glycan shown here is the predominant glycan emanating from the mucin-rich core of GPIbα, and contains two branches that can be sialylated, but by distinct sialyltransferases. ST3GAL1 and ST3GAL2 appear to favor sialylating Galβ1,3GalNAc (Core 1) as a substrate, while ST3GAL4 favors transferring sialic acid residues to Galβ1,4GlcNAc.
Supplemental methods

Cell culture and transfection
Human OT1-1 iPSCs were cultured on Matrigel (Corning, Corning, NY)-coated plates in StemFlex Medium (Thermo Fisher Scientific) at 37 °C in 4% O₂/5% CO₂. After incubation with 10 µM ROCK inhibitor Y27632 (StemRD Inc., Burlingame, CA), 2x10⁵ cells were transfected with 1 µg of guide plasmid pairs targeting ST3GAL1 or ST3GAL2 using the Amaxa P3 primary cell 4D Nucleofector Kit (Lonza, Allendale, NJ) and Nucleofector Program CB-150. The cells were then plated on Matrigel-coated plates with 10 µM Y27632. 24-hour post-transfection puromycin was applied at a concentration of 1 µg/ml for 48 hr. Single clones were harvested at 12 to 14 days post-puromycin-selection and re-plated on Matrigel-coated plates. To generate ST3GAL1/2 KO cells, ST3GAL1 KO cells were transfected with guide plasmid pairs targeting ST3GAL1 using the same procedure.

Genotyping
Genomic DNA was extracted from each iPSC clone using the QuickExtract DNA Extraction Solution (Epicenter) following the manufacture's protocol. The targeted regions were amplified by PCR using the pairs of primers: ST3GAL1 for: 5’- CATCCCCCAGCAGGAGTGAAGT -3’, ST3GAL1 rev: 5’- GACCCCATTTACAGAAGCCCCC -3’, ST3GAL2 for: 5’- TGCTCATGAACCACAAGGAC -3’ and ST3GAL2 rev: 5’- AGAAAGGGGCTCTGTCTTCTCA - 3’.

Differentiation of iPS cells
CRISPR-edited iPS cell lines were differentiated to HPCs as previously described.1,2 Briefly, cells were plated on Matrigel for differentiation. Media and cytokine changes were followed as described except that the GSK-3β inhibitor, CHIR99021 (Tocris, Minneapolis, MN) (1 µM) was used instead of Wnt3a. Cells were cultured at 37°C with 4% O₂/5% CO₂ for 9 days, and loosely adherent HPCs were collected by carefully removing the supernatant. Cells were analyzed by flow cytometry to confirm surface expression of CD34 and CD43. The HPCs were further differentiated to MKs in serum-free differentiation (SFD) medium, which is comprised of Iscove’s Modified Dulbecco’s Medium (IMDM) (Thermo Fisher Scientific) containing 25% Ham’s F12 (Corning), 0.5% N2, 1% B27 without Vitamin A (Thermo Fisher Scientific), 0.05% BSA (Sigma, St. Louis, MO), 2mM L-glutamine and penicillin/streptomycin supplemented with 50 ng/ml SCF and 50 ng/ml TPO (R&D systems, Minneapolis, MN) at 37 °C, 5% CO₂ for 6 days. MKs were analyzed by flow cytometry to confirm the surface expression of CD41 and CD42b. To generate EBs, the HPCs were cultured in SFD medium supplemented with 50 ng/ml SCF and 2 U/ml EPO (R&D systems) at 37 °C, 5% CO₂ for 6 days. EBs were analyzed by flow cytometry to confirm the surface expression of CD71 and CD235.

Western blot analysis
Following electrophoresis, the samples were electrotransferred onto PVDF membrane (EMD Millipore, Billerica, MA) and immunoblotted with either biotinylated-PNA, biotinylated-STL (Vector Laboratories Inc.) or primary antibodies including sheep anti-ST3GAL1, sheep anti-ST3GAL2, sheep anti-CD34 and goat anti-CD43 (R&D system), mouse anti-β-actin (Sigma),
rabbit anti-GPIIb (Sew 8) or rabbit anti-GPIbα (LSBio, Seattle, WA) antibodies. Bound lectins and primary antibodies were visualized using HRP-streptavidin (BioLegend) and species-specific peroxidase-conjugated secondary antibodies: goat anti-rabbit IgG, goat anti-mouse IgG, donkey anti-goat IgG and donkey anti-sheep IgG (Jackson ImmunoResearch Laboratories).

**PAC-1 and fibrinogen binding**

iPSC-derived MKs were re-suspended in modified Tyrode's buffer containing 20 mM HEPES, 134 mM NaCl, 2.9 mM KCl, 0.34 mM Na2HPO4, 12 mM NaHCO3, 1 mM MgCl2 (pH 7.3). The cells were incubated either with PE-conjugated anti-CD41a and FITC-conjugated anti-PAC-1 (BD Biosciences), or PE-conjugated anti-CD42a (BD Biosciences) and FITC-conjugated fibrinogen (Thermo Fisher Scientific) upon stimulation with 0.1 U/ml Thrombin (Roche, Indianapolis, IN). After 20 min incubation at 37°C, the cells were fixed with 2% PFA before analyzed by flow cytometry.

**DNA content analysis**

iPSC-derived MKs were incubated with PE-conjugated anti-CD41a at room temperature for 30 min before fixing with 70% ethanol at -20°C overnight. The cells were then stained with 2 μM DAPI with 50 μg/ml RNase A (Sigma-Aldrich) in PBS buffer at room temperature for 1 hour. Cellular DNA content was analyzed by flow cytometry.

**Single-cell RNA sequencing analysis**

Bone marrow, single-cell RNA sequence data from 12 human subjects were obtained from The National Center for Biotechnology Information/Gene Expression Omnibus. Data analysis was performed in R (v 4.0.2) using the package Seurat (v 4.0.3), with the package tidyverse (v 1.2.1) used to organize data, and the package ggplot2 (v 3.2.1) used to generate figures. scRNA-seq data was filtered to keep cells with a low percentage of mitochondrial genes in the transcriptome (< 5%) and between 200 and 3000 unique genes to exclude apoptotic cells, low quality cells, and doublets. Cell cycle scores were regressed when scaling gene expression values during scTransform integration and clustering, which was performed using 50 principal components with the Louvain algorithm and visualized with UMAP plots.

References

1. Mills JA, Paluru P, Weiss MJ, Gadue P, French DL. Hematopoietic differentiation of pluripotent stem cells in culture. In: Qu KDBaC-K, ed. Methods in Molecular Biology. Vol. 1185. New York: Springer Science+Business Media; 2014:181-194.
2. Paluru P, Hudock KM, Cheng X, et al. The negative impact of Wnt signaling on megakaryocyte and primitive erythroid progenitors derived from human embryonic stem cells. Stem Cell Res. 2014;12(2):441-451.
3. Oetjen KA, Lindblad KE, Goswami M, et al. Human bone marrow assessment by single-cell RNA sequencing, mass cytometry, and flow cytometry. JCI Insight. 2018;3(23).
4. Stuart T, Butler A, Hoffman P, et al. Comprehensive Integration of Single-Cell Data. Cell. 2019;177(7):1888-1902 e1821.

Zhang et al.
5. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* 2019;20(1):296.