# Supporting Information

### **Fujitani et al.**

### **SI Materials and Methods**

**Cell culture.** The following cell lines were provided by RIKEN Bio Resource Center in Japan: Chinese hamster ovary cells (CHO-K1), human promyeloleukemic cells (HL60), human cervical carcinoma cells (HeLa), human embryonic carcinoma (EC) cells(NEC8), human lung carcinoma cells (A549), human pancreatic cancer cells (KLM-1), human colon carcinoma cells (Caco-2), human hepatocyte carcinoma cells (HepG2), human embryonic kidney cells (HEK293), and normal embryonic lung fibroblasts (MRC-5). Lec1 and Lec8 cells (Chinese hamster ovary (CHO) mutants with a deficiency in N-acetylglucosaminetransferase I (GlcNAc-TI) and the UDP-galactose transporter, respectively) were provided by American Type Culture Collection (Manassas, VA). CHO-K1, Lec1, Lec8, HL60, NEC8 and KLM-1 cells were cultured in RPMI1640 medium (Sigma-Aldrich, St. Louis, MO). HeLa, Caco-2, HepG2 and HEK293 cells were cultured in Minimum Essential Medium (MEM; Sigma-Aldrich). A549 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM; Sigma-Aldrich). MRC-5 cells were cultured in MEM Alpha (Sigma-Aldrich). The medium for the culture of Caco-2 cells was supplemented with 20% fetal calf serum (FCS) (Nichirei Biosciences, Tokyo, Japan), and the medium for the culture of the other cell types was supplemented with 10% FCS. In addition, a mixture of non-essential amino acids (0.1 mM) was added to the medium for the culture of Caco-2, HepG2 and HEK293 cells. Human ES cells (KhES1, 3, 4 and 5) and iPS cells (HiPS-RIKEN-1A, 2A and 12A derived from human umbilical cord, or iPS1A, 2A and 12A; HiPS-RIKEN-3A (iPS3A) derived from human decidual tissue, and HiPS-RIKEN-11A (iPS11A) derived from amniotic tissue) were maintained in an undifferentiated state in the presence of

mouse embryonic fibroblasts (MEFs) as described previously ( 1 ). Antibiotic reagents (Invitrogen, Carlsbad, CA) containing penicillin (100 U/mL) and streptomycin (100 µg/mL) were used for all cultures except for stem cells and MEFs. All cells were cultured in a 10 cm dish with 10 mL medium at 37 $^{\circ}$ C and 5% CO<sub>2</sub> until they reached approximately 100% confluence. After washing the culture dish with cold phosphate buffered saline (PBS), the cells were scraped (except for HL60 cells, which are non-adherent) into cold PBS with 10 mM EDTA and collected by centrifugation at 800 × *g* for 5 min. The cells were stored at −30°C as pellets containing  $1-2 \times 10^6$  cells/pellet. The total cellular protein concentration was determined by using the bicinchoninic acid (BCA) method (Pierce Biotechnology, Rockford, IL).

**Expression and purification of** *Rhodococcus* **endoglycoceramidase I (EGCase I).**  EGCase I found in *Rhodococcus equi* (2) was expressed in *R. erythropolis* strain JCM3201 L88, which was transformed with a pTip LCH 2.2 vector (3) containing *egcase I* as a Cterminal His-tagged protein (4). In brief, transformed JCM3201 L88 cells were cultured in Luria-Bertani medium with 17 µg/mL chloramphenicol. Pre-cultured host cells were transferred into a 10-fold volume of medium containing 1 µg/mL thiostreptone for the induction of EGCase I expression. Cells were cultured at 26ºC for 24 h and collected by centrifugation (20,000 × *g*

-

<sup>1.</sup> Suemori H, *et al.* (2006) Efficient establishment of human embryonic stem cell lines and long-term maintenance with stable karyotype by enzymatic bulk passage. *Biochem Biophys Res Commun* 345: 926-932.

<sup>2.</sup> Ito M, Yamagata T (1986) A novel glycosphingolipid-degrading enzyme cleaves the linkage between the oligosaccharide and ceramide of neutral and acidic glycosphingolipids. *J Biol Chem* 261: 14278- 14282.

<sup>3.</sup> Nakashima N, Tamura T (2004) A novel system for expressing recombinant proteins over a wide temperature range from 4 to 35 degrees C. *Biotechnol Bioeng* 86: 136-148.

<sup>4.</sup> Ishibashi Y, *et al.* (2012) Preparation and characterization of endoglycoceramdiase I, applicable to the comprehensive analysis of glycosphingolipids, using a rhodococcal expression system. *J Lipid Res* 53: 2242-2251.

for 20 min at 4ºC). The cell pellet was resuspended in 20 mM of sodium phosphate buffer (20 mL, pH 7.4) containing 0.2 mg/ml lysozyme and sonicated on ice for 30 min. The supernatant containing recombinant EGCase I was collected by centrifugation (20,000 × *g* for 20 min at 4ºC) and purified via affinity chromatography with Ni-Sepharose 6 Fast Flow (GE Healthcare, Buckinghamshire, UK). The fractions containing EGCase I were eluted by phosphate buffer containing 400 mM imidazole and dialyzed against 20 mM sodium acetate buffer (pH 5.5). Protein quantification was performed by using the BCA protein assay (Pierce).

**Extraction of (glyco)proteins and FOSs.** For N-glycan and FOS analyses, cell pellets consisting of 1  $\times$  10<sup>6</sup> cells were homogenized by using an Ultrasonic Homogenizer (Taitec Corp. Saitama, Japan) in 100 mM Tris-acetate buffer (100 µL, pH 7.4) supplemented with 2% sodium dodecyl sulfate as a surfactant for the complete dissolution of cell pellets. In the presence of Benzonase® Nuclease (50 units; Sigma-Aldrich), the reductive alkylation of the cellular proteins was performed by the addition of 500 mM tris(2-carboxyethyl)phosphine (Sigma-Aldrich) at room temperature for 60 min, and subsequently by the addition of 200 mM iodoacetamide (Sigma-Aldrich) at room temperature for 30 min. After reductive alkylation, ethanol precipitation was carried out by the addition of a 4-fold volume of cold ethanol and incubation for 3 h at -30°C. Supernatants and precipitated proteins were separated by centrifugation at 20,000  $\times$  g for 10 min at 4<sup>o</sup>C, and precipitates were again washed with cold ethanol. Supernatants containing cellular FOSs were carefully transferred to fresh tubes and completely desiccated with a centrifugal evaporator. Desiccated samples were dissolved in deionized water and were directly subjected to glycoblotting, as described below. Collected precipitates containing glycoproteins/N-glycans were dried at 37°C for 10 min.

**Extraction of GSLs and GAGs.** The streamlined GAG and GSL extraction procedures were essentially the same as previously described (5, 6). Briefly, delipidation was performed by the addition of chloroform/methanol (C/M) solution (2/1, v/v; 450 µL) to the cell pellets, followed by sonication at room temperature in the same manner as described above for the extraction of glycoproteins/N-glycans. Methanol (150 µL) was then added, yielding a solvent composition of  $C/M = 1/1$  (v/v). Sonication was repeated in the same manner. Finally, methanol (300  $\mu$ L) was added (C/M =  $1/2$ ,  $v/v$ ), and sonication was repeated once again. The resulting extracts were subjected to centrifugation at 20,000 ×*g* for 10 min, and the supernatants and the pellets were subjected to GSL-glycan and GAG analysis, respectively. For GSL glycan analysis, supernatants containing crude cellular lipids were completely dried with a centrifugal evaporator and subjected to digestion with ceramidase. For GAG analysis, pellets were subjected to an overnight digestion at  $37^{\circ}$ C with non-specific protease by the addition of 100 µg Pronase (Calbiochem, San Diego, CA) in 50mM Tris-acetate buffer (pH 7.5) containing 10 mM CaCl<sub>2</sub>. After Pronase digestion, ethanol precipitation was carried out in the same manner as described for glycoprotein/FOS preparation to precipitate peptides linked to GAG chains.

**Release of N-glycans.** Ethanol-precipitated proteins were dissolved in 100 mM ammonium bicarbonate and digested overnight at  $37^{\circ}$ C by the addition of trypsin (Sigma-Aldrich). Following deglycosylation by overnight treatment with peptide: N-glycanase F (PNGase F;

 $\overline{a}$ 

<sup>5.</sup> Fujitani N, *et al.* (2011) Qualitative and quantitative cellular glycomics of glycosphingolipids based on rhodococcal endoglycosylceramidase-assisted glycan cleavage, glycoblotting-assisted sample preparation, and matrix-assisted laser desorption ionization tandem time-of-flight mass spectrometry analysis. *J Biol Chem* 286: 41669-41679.

<sup>6.</sup> Takegawa Y, *et al.* (2011) Simultaneous analysis of heparan sulfate, chondroitin/dermatan sulfates, and hyaluronan disaccharides by glycoblotting-assisted sample preparation followed by single-step zwitter-ionic-hydrophilic interaction chromatography. *Anal Chem* 83, 9443-9449.

Roche) (2 U), the resulting samples were dried by using a centrifugal evaporator and dissolved in deionized water.

**Release of glycans from GSLs.** Crude cellular lipids were suspended in 50 mM sodium acetate buffer pH 5.5 (50 µL) containing 0.2% Triton X-100 (Sigma-Aldrich) as a surfactant, followed by the addition of EGCase I and/or II (Takara Bio Inc. Shiga, Japan) (25 mU) to release intact glycans from GSLs. The enzymatic digestions were performed at 37°C for 24 h. To distinguish GSL-derived glycans from contaminating FOSs, the crude cellular lipids were also suspended in sodium acetate buffer/Triton X-100 in the absence of EGCases. The latter served as a negative control.

**Enzymatic digestion of GAGs.** The ethanol-precipitated sample was completely digested to disaccharides using 5 mU each of a mix of heparinase, heparitinase isolated from *Flavobacteriumheparinum* (Sigma-Aldrich), hyaluronidase SD isolated from *Streptococcus dysgalactiae*, and chondroitinase ABC isolated from Proteus vulgaris (Seikagaku Kogyo Co., Ltd., Tokyo, Japan). The sample and the enzymes were incubated overnight at 37°C in 100 mM ammonium acetate (100 μL) containing 5 mM calcium acetate (pH 7.5). The digested sample was subjected to centrifugation, and the supernatant was collected, dried *in vacuo*, and dissolved in deionized water (25 µL) containing a known amount (typically 50 pmol) of an internal standard, isomaltotriose (Sigma-Aldrich).

**Glycoblotting.** N-glycans, FOSs and GSL glycans were subjected to glycoblotting as previously described (7), with minor modifications. In brief, the sample solution (20–50 µL) was

 $\overline{a}$ 

<sup>7.</sup> Furukawa J, *et al.* (2008) A comprehensive approach to structural and functional glycomics based on chemoselective glycoblotting and sequential tag conversion. *Anal Chem* 80: 1094-1101.

directly applied to the well of a filter plate (MultiScreen Solvinert 0.45 μm Low-Binding Hydrophilic PTFE, Millipore, Billerica, MA) containing BlotGlyco® beads (5.0 mg; Sumitomo Bakelite Co. Ltd., Tokyo, Japan). Glycans were captured in 2% acetic acid (AcOH) in acetonitrile (ACN, 450 μL) and incubated at 80°C for 45 min. On-bead acetyl capping of unreacted hydrazide groups was performed by using 10% acetic anhydride in MeOH for 30 min at room temperature. Next, on-bead methyl esterification of the carboxyl groups in glycanderived sialic acid was carried out by incubation with 150 mM 3-methyl-1-p-tolyltriazene in dioxane at 60°C. The trapped and esterified glycans on the beads were subjected to transiminization by incubation with a mixture of 2% AcOH in ACN (180 μl) and 20 mM aminooxy-WR (aoWR, 20 μL), a dipeptidic aminooxy compound that was synthesized as described previously(8). The aoWR-labeled glycans were recovered with distilled water (100 μL), and the collected solution was purified by using a hydrophilic interaction liquid chromatography (HILIC) purification plate (MassPrep HILIC μElution plate, Waters, MA) to remove the excess aoWR. To concentrate the aoWR-labeled glycans, the purified solution was desiccated by using a rotational evaporator and subsequently dissolved in distilled water (10 µL). To isolate GAG disaccharides, the polymer surface-bound glycans were recovered by adding 180 µL 2% AcOH/ACN and incubating the mixture at 80°C for 45 min. Fifty microliters of 2-aminobenzamide (2AB, Nacalai Tesque, Kyoto, Japan) solution (0.35 M  $2AB/1M$  NaBH<sub>3</sub>CN in 30% AcOH/DMSO) was added and incubated with the mixture at 60°C for 2 h. The 2AB-labeled glycans were recovered by filtration, and excess reagent was removed by using a silica gel column composed of Iatrobeads (Mitsubishi Chemical Medience, Tokyo, Japan). Briefly, the reaction mixture was diluted with ACN (final concentration 95%) and then applied to  $\sim$ 25 mg of latrobeads silica gel packed in a disposable filter column pre-

-

<sup>8.</sup> Uematsu, R. *et al.* (2005) High throughput quantitative glycomics and glycoform-focused proteomics of murine dermis and epidermis. *Mol. Cell Proteomics* **4**, 1977-1989.

equilibrated with 1 M AcOH and ACN. The column was washed first with ACN and then with 96% ACN in water, followed by 2% AcOH in ACN. Glycosaminoglycan disaccharides labeled with 2AB were eluted from the Iatrobeads with 50% aqueous ACN.

#### **Cellular O-glycomic analysis by β-elimination in the presence of pyrazolone analogues**

**(BEP).** The extracted cellular (glyco)proteins were subjected to BEP as previously described (9), with modifications. Briefly, the extracted (glyco)proteins from cell pellets were subjected to the same procedure described above for N-glycan preparation. Collected proteins were concentrated by using an Amicon Ultra Centrifugal Filter (30K) (Millipore). The concentrated proteins were subjected to BEP reaction. The solution was neutralized with 1.0 M hydrochloric acid. Chloroform was added, and the mixture was shaken vigorously. The chloroform layer was discarded to remove excess reagents, and the resultant aqueous layer was subjected to purification on a graphitized carbon column(10) and an Iatrobeads silica gel column. More detailed technical procedure for cellular O-glycome will be reported elsewhere.

**MALDI-TOF/TOF MS (matrix-assisted laser desorption/ionization-time of flight mass spectrometry) analysis.** Purified N-glycans, O-glycans, FOSs and GSL glycan solutions were mixed with 2,5-dihydrobenzoic acid solution (10 mg/mL in 30% ACN) and subsequently subjected to MALDI-TOF MS analysis as previously described, with minor modifications. Briefly, all measurements were performed by using an Ultraflex II TOF/TOF mass spectrometer equipped with a reflector and controlled by the FlexControl 3.0 software package

-

<sup>9.</sup> Furukawa J, *et al.* (2011) A versatile method for analysis of serine/threonine posttranslational modifications by β-elimination in the presence of pyrazolone analogues. *Anal Chem* 83: 9060-9067.

<sup>10.</sup> An, H.J. *et al.* Extensive determination of glycan heterogeneity reveals an unusual abundance of high mannose glycans in enriched plasma membranes of human embryonic stem cells. *Mol Cell Proteomics* in press.

(Bruker Daltonics GmbsH, Bremen, Germany) according to general protocols. All spectra were obtained in reflectron mode with an acceleration voltage of 25 kV, a reflector voltage of 26.3 kV, and a pulsed ion extraction of 160 ns in the positive ion mode. Masses were annotated by using the FlexAnalysis 3.0 software package (BrukerDaltonics GmbsH, Bremen, Germany). The GlycoSuiteDB (http://glycosuitedb.expasy.org/glycosuite/glycodb) and SphinGOMAP (http://www.sphingomap.org/) online databases were used for structural identification of GSL glycans. Absolute quantification was performed by comparative analyses between the areas of the MS signals derived from each glycan and 10pmol of the internal standard NeuAc<sub>2</sub>Gal<sub>2</sub>GlcNAc<sub>2</sub> + Man<sub>3</sub>GlcNAc<sub>1</sub>(A2GN1), which was added to the sample solution prior to glycoblotting. In TOF/TOF mode measurements for fragment ion analysis, precursor ions were accelerated to 8 kV and selected in a timed ion gate. Fragment ions generated by laserinduced decomposition of the precursor were further accelerated by 19 kV in the LIFT cell. Fragment masses were analyzed after passing through the ion reflector.

### **Exoglycosidase-assisted structure analysis**

GSL-glycans: aoWR-labelled glycans were digested with β1,3-galactosidase for distinguishing the isomer structures possessing galactose residue(s) at nonreducing termini. Recombinant β1,3-galactosidase (100 U) derived from *Xanthomonas manihotis* (Calbiochem, Darmstadt, Germany) was added to 5 µL of the final solution obtained from glycoblotting procedure. The solution was incubated at 37oC for 1 hr, and the solution  $(1 \mu L)$  was directly subjected to MS analysis without further purification.

O-glycans: The extracted (glyco)proteins from ES1cell pellets were treated with either β1,3 galactosidase (*Xanthomonas manihotis*, Calbiochem) or β1,4-galactosidase (*Streptococcus pneumoniae*, Calbiochem). Galactosidase digested samples were subjected to O-glycan analysis as described.

**High performance liquid chromatography (HPLC) analysis.** Separation and detection of 2AB-labeled GAG disaccharides were performed as described previously (6). Briefly, 2ABlabeled GAG disaccharides were separated on a ZIC®-HILIC column (3.5  $\times$  µm  $\times$  2.0 mm  $\times$ 150 mm; SeQuant, Umeå, Sweden) in a column oven at 35°C. A gradient elution was applied at a flow rate of 200 µL/min using water (solvent A), ACN (solvent B) and 200 mM ammonium acetate buffer (solvent C) (A/B/C =  $5/90/5$  (0 min) $\rightarrow$ 30/65/5 (60 min). Disaccharides were detected by in-line fluorescence (excitation at 330 nm and emission at 420 nm). The peak identification was confirmed through coinjection of a mixture of unsaturated ∆-disaccharide standards (Seikagaku Kogyo Co., Ltd., Tokyo, Japan). Quantitation was performed using isomaltotriose (IM3, Sigma-Aldrich, Dorset, UK) as the internal standard.

**Statistical analysis.** Cluster analyses were performed with software Cluster 3.0, developed by the Michael Eisen laboratory (http://www.eisenlab.org/eisen/) with a hierarchical clustering algorithm (11). The calculated dendrogram and correlation matrix were visualized by using TreeView 1.6 software downloaded from the same site. Pearson product-moment correlation coefficient *r* between all cells was calculated based on the absolute amounts of all detected glycans (amount/100 µg of protein). Correlation matrices were produced by aligning *r* with the cells by using Microsoft Excel. Box- and whisker plots were generated using TIBCO Spotfire software (TIBCO Software Inc., Göteborg, Sweden).

-

<sup>11.</sup> Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genomewide expression patterns. *Proc Natl Acad Sci USA* 95: 14863–14868.

#### **SI Results**

-

**Validation of cellular glycomics as a tool for cellular characterization/description.** The usefulness of total glycomic analysis for the characterization/description of cells was validated using the wild-type CHO cell line and its lectin-resistant mutants (Lec1 and Lec 8). Lec1 CHO cell mutants lack GlcNAc-TI activity and do not synthesize complex or hybrid N-glycans. The lec8 mutation inhibits translocation of UDP-Gal into the Golgi lumen and thereby dramatically reduces galactosylation of all glycoconjugates (12). Perturbed glycan synthetic pathways such as those observed in the mutant CHO cells can cause unexpected glycan expression profiles, which may be difficult to predict in the absence of glycomics.

Upon analysis of the total cellular glycome of these cell lines, we could quantitatively detect 80 N-glycans, 4 O-glycans, 34 GSL glycans, 19 FOSs and 12 kinds of GAG disaccharides (Supplementary Table S1, Supplementary Fig. S15). The total glycomic profile of each cell line is illustrated based on the absolute amount of each type of glycan and each glycan substructure, as shown in Supplementary Fig. S15a. Pie charts at the vertices of the pentagon correspond to the glycan expression profiles of N-glycans, FOSs, GAGs, GSL glycans and Oglycans. The size of each circle and its constituent colors reflect the absolute glycan quantity (pmol/100 µg protein) and the glycan substructures, respectively. As such, the relative abundance of each class of glycoconjugate and its diversity is obvious at a glance by the pentagonal cellular glycomic notations. GAGs (~302 pmol), as assessed by the total amount of disaccharides, were found to be the most abundant glycan type in CHO cells, followed by Nglycans (~195 pmol), GSL glycans (~55 pmol), FOSs (~51 pmol), and O-glycans (~49pmol). The total glycomic profiles of Lec1 and Lec8 cells were revealed in the same manner (Fig.14b,

<sup>12.</sup> Stanley P, Sundaram S, Sallustio S (1991) A subclass of cell surface carbohydrates revealed by a CHO mutant with two glycosylation mutations. *Glycobiology* 1: 307-314.

c). As predicted and as previously reported (13), the amounts of complex or hybrid N-glycans were drastically reduced in Lec1 compared with CHO cells. In case of Lec8 cells, the amounts of O-glycans and GSLs were both substantially reduced compared with CHO cells. This result is to be expected because galactosylation is essential for the synthesis of T-antigen and lactosylceramide, which are the precursor structures for most mucin-type O-glycans and GSL glycans, respectively.

Interestingly, unlike CHO and Lec1 cells, most of the N-glycans in Lec8 cells corresponded to complex type glycans terminating with GlcNAc, while those in CHO and Lec1 cells corresponded to high-mannose type glycans. This may be explained by an attempt on the part of Lec8 cells to compensate for their reduced ability to synthesize GSL glycans and O-glycans compared with CHO and Lec1 cells. We also observed an increased expression of GAGs, Oglycans and GSL glycans in Lec1 compared with CHO cells by a factor of 2.4-, 4.0- and 1.7 fold, respectively, as well as a 2.7-fold increase in GAG expression in Lec8 compared with CHO cells.

-

<sup>13.</sup> North SJ, *et al.* (2010) Glycomics profiling of Chinese hamster ovary cell glycosylation mutants reveals N-glycans of a novel size and complexity. *J Biol Chem* 285: 5759-5775.



Fig.S1 General scheme for the analysis of N-glycans, glycosphingolipid (GSL) glycans, glycosaminoglycans (GAGs), and free oligosaccharides (FOSs) (a) and O-glycans (b).



Fig.S2 Optimization of conditions for sample preparation for the analysis of N-glycans and FOSs. A model cells (A549) were homogenized by using an Ultrasonic Homogenizer (Taitec Corp. Saitama, Japan) in 100 mM Tris-acetate buffer (100 µL, pH 7.4) supplemented with 1% Triton X-100 (P1), 1% CHAPS with 5mM EDTA (P2) or 2% sodium dodecyl sulfate (P3) as a surfactant for the complete dissolution of cell pellets. N-glycans and FOSs were recovered and quantified as described in Experimental section (N=3). It was found that the glycan recovery was not very affected by the type of surfactant as far as samples were mechanically homogenized. We chose to employ SDS as a surfactant due to relatively better reproducibility. Reductive alkylation and protease treatment prior to PNGase F digestion were employed based on our previous study to optimize efficiency of enzymatic release of N-glycans from serum glycoproteins (Kita Y, et al. (2007) *Mol Cell Proteomics* 6: 1437-1445, 2007). Optimization of conditions for the preparation of GSL glycans and GAG disaccharides was described previously (Fujitani N, *et al.* (2011) *J Biol Chem* 286: 41669-41679, Takegawa Y, *et al.* (2011) *Anal Chem* 83: 9443-9949).



Fig.S3 Reproducibility of the glycomic analysis for cellular N-glycans, free oligosaccharides (FOSs) and O-glycans. Chinese hamster ovary cells (CHO-K1) were used as a model and the absolute quantitation of N-glycans, FOSs and O-glycans was performed in triplicate. Each value shown is the mean  $\pm$  the S.D. (n =3). The relatively high S.D. value observed for Hex1HexANc1 (O-glycan) is most likely attributable to the fact that MALDI-TOF MS spectra often have higher baseline in low m/z region. Reproducibility of the analysis of glycosaminoglycans (GAGs) and glycosphingolipid (GSL) glycans was described previously(Fujitani N, *et al.* (2011) *J Biol Chem* 286: 41669-41679, Takegawa Y, *et al.* (2011) *Anal Chem* 83: 9443-9949).



Fig.S4 Separation of free oligosaccharides (FOSs) from glycoproteins. Upon ethanol precipitation, the pellet (precipitated proteins, a) and the supernatant (b) were separated by centrifugation and subjected to glycan analysis via glycoblotting with (upper panel of a/b)or without (lower panel of a/b) PNGase F digestion. No FOSs were detected among the precipitated proteins (lower panel of a), indicating that the FOSs were quantitatively recovered from the supernatant. No N-glycans were detected in the supernatant (upper panel of b), indicating that the glycoproteins were quantitatively recovered from the pellet. Estimated structures are shown.









19

























a-2)

1.0x107

cells









Fig.S5 MS spectra and chromatograms showing the glycomic profiles obtained for human cells analyzed in this study: (a) N-glycans, (b) glycosphingolipid (GSL) glycans, (c) O-glycans, (d) FOSs, and (e) glycosaminoglycans (GAGs). Estimated structures are shown. Green circle, Man; yellow circle, Gal; blue square, GlcNAc; yellow square, GalNAc; red triangle, Fuc; purple diamond, Neu5Ac; open diamond, Neu5Gc. The m/z values (both theoretical experimental) are shown in Supplementary Table S3.



Fig.S6 Pentagonal cellular glycomic notations showing the entire glycome of four hESC lines (ES1, 3, 4 and 5); five hiPSC lines (iPS1A, 2A and 12A derived from human umbilical cord, iPS3A derived from human decidual tissue, and iPS11A derived from amniotic tissue). Pie charts at the vertices of the pentagon correspond to the glycan expression profiles of Nglycans, FOSs, GAGs, GSL glycans and O-glycans. The size of each circle and its constituent colors reflect the absolute quantity of glycans (pmol/100 µg protein) and the glycan substructures, respectively. The size of each circle corresponding to FOS content is enlarged by 40-fold (area ratio). Estimated structures are shown.



Fig.S7 Correlation matrix of the 18 kinds of cells. For simple and clear-cut analysis, the correlation coefficient (r) value was used to distinguish the different cell types from each other. The comparison of r values based on total glycomic analysis ( $N + GSL + GAG + FOS +$ O) displayed clear separation of stem cells (ESCs and iPSCs) from the other cancer-related and normal cells. Quantitative glycomic analysis of each glycan class also distinguished the stem cells from the other cells, although the results were not as clear-cut as those for the total glycomic analysis. In fact, only the N-glycan profile, which revealed a high r value between all cells except for NEC8 and MRC5 cells, was capable of distinguishing the ESCs and iPSCs from the other cells.



Fig.S8 Comparison of the ratio of total cellular FOSs and total cellular N-glycans among carcinoma cells, ESCs and iPSCs.



Fig.S9 Glycomic profiles of complex/hybrid type N-glycans of various human cells. Four-digit numbers represent the number of hexose, N-acetyl hexosamine, fucose and sialic acid residues. G designates that thesialic acid(s) is/are present as N-glycolylneuraminic acid (Neu5Gc).



Fig.S10 Representative fluorescence-activated cell-sorting (FACS) plots demonstrating high Tra-1-60, Tra-1-81, SSEA-3, 4 expressions while low expression of SSEA-1 on pluripotent hESCs (ES-1).



Fig.S11 Box-and-whisker plots showing the total cellular levels of various glycans in stem cells (hESCs and hiPSCs) and non-stem cells (p<0.001 and p<0.01). The encircled outside vales indicate that they are originated from mouse embryonic fibroblasts (MEFs), feeder cells used for the culture of hESCs and hiPSCs. Glycans shown in italics indicate that they are mostly if not all derived from MEFs.

# $(a)$

## a-1) N-glycans



### a-2) GSL-glycans



## a-3) O-glycans



$$
(b)
$$

b-1) GSL-glycans



Partial MALDI-TOF spectra of GSL-glycans released from ES1 cell before (Upper) and after (Lower) digestion with 61.3 galactosidase.

Linkage specific 61,3-galactosidase treatment of aoWR-tagged GSL-glycans allowed nearly complete digestion of non-reducing terminal  $B1,3$ -linked galactose, resulting in distinguishing and quantifying GM1 and sialyl nLc4 ((Hex)3(HexNAc)1(NeuAc)1). With decreasing of GM1, the area of the signal corresponding to GM2 increased in the spectrum of galactosidase treated sample (lower). Signals marked with asterisk corresponds to hexose oligomer (Hex)6 contaminated from the sample.

## b-2) O-glycans



Partial MALDI-TOF spectra of O-glycans released from ES1(a). Mass spectra after β1,3 galactosidase digestion (b) and after β1,4-galactosidase digestion (c). Both Hex3HexNAc3 (right) and Hex2HexNAc2 (left) were sensitive to β1,3-galactosidase digestion (b) but not to β1,4-galactosidase digestion (c). Upon β1,3-galactosidase digestion, with decreasing of the signals of Hex3HexNAc3 (right) and Hex2HexNAc2 (left), signals corresponding to Hex2HexANc3 and Hex1HexNAc2, respectively, increased in the spectra of 81,3galactosidase treated samples. The incomplete digestion even after treatment of β1,3 galactosidase may be explained that the β1,3-galactosidase activity on mucins may be incomplete.

Fig S12 Structural evaluation of pluripotency biomarker candidates by MALDI-TOF/TOF MS analysis (a) and exoglycosidase digestion analysis (b).



Fig S13 Pentagonal cellular glycomic notations showing the entire glycome of mouse embryonic fibroblasts (MEFs). Pie charts at the vertices of the pentagon correspond to the glycan expression profiles of N-glycans, FOSs, GAGs, GSL glycans and O-glycans. The size of each circle and its constituent colors reflect the absolute quantity of glycans (pmol/100 µg protein) and the glycan substructures, respectively. The size of each circle corresponding to FOSs content is enlarged by 40-fold (area ratio).



Fig.S14 MS spectra and chromatograms showing the glycomic profiles obtained for CHO-K1: (a) N-glycans analyzed followed by aoWR labeling, (b) N-glycans analyzed followed by permethylation, (c) glycosphingolipid (GSL) glycans, (d) O-glycans, (e) FOSs, and (f) glycosaminoglycans (GAGs). Estimated structures are shown. The N-glycomic profile obtained followed by aoWR labeling and permethylation yielded comparable results both qualitatively and quantitatively. Green circle, Man; yellow circle, Gal; blue square, GlcNAc; yellow square, GalNAc; red triangle, Fuc; purple diamond, Neu5Ac; open diamond, Neu5Gc.



Fig.S15 Pentagonal cellular glycomic notations showing the entire glycome of the Chinese hamster ovary (CHO) cell line (a) and its lectin-resistant mutants (Lec1 (b) and Lec 8 (c)). Pie charts at the vertices of the pentagon correspond to the glycan expression profiles of Nglycans, FOSs, GAGs, GSL glycans, and O-glycans. The size of each circle and its constituent colors reflect the absolute quantity of glycans (pmol/100 µg protein) and the glycan substructures, respectively. Estimated structures are shown.

# Table S1 List of glycans quantified in in wild-type and mutant CHO cells



# Table S2 List of glycans quantified in hESCs, hiPSCs and other human cells









