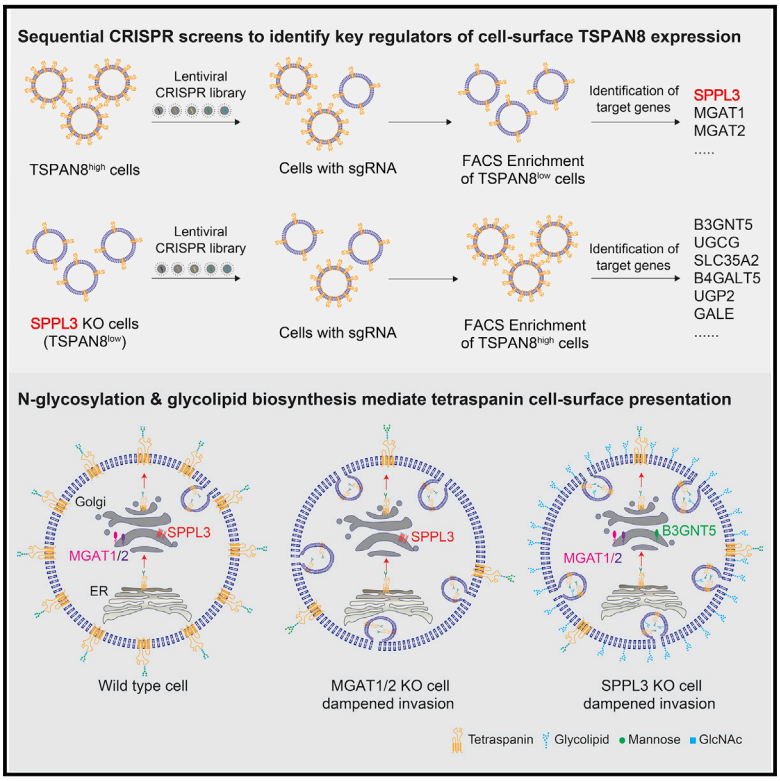
Cell Reports

Article

Sequential genome-wide CRISPR-Cas9 screens identify genes regulating cell-surface expression of tetraspanins

Graphical abstract



Highlights

- Identifying key regulators of tetraspanin trafficking by CRISPR-Cas9 screens
- The role of N-glycosylation in cell-surface presentation of tetraspanins
- Control of cell-surface tetraspanin expression by a SPPL3/B3GNT5 axis
- MGAT1/2 and SPPL3 are potential targets for blocking cancer cell invasion

Authors

Jicheng Yang, Fusheng Guo, Hui San Chin, ..., Qingsong Lin, Wanjin Hong, Nai Yang Fu

Correspondence

naiyang.fu@duke-nus.edu.sg

In brief

Using sequential genome-wide loss-of-function CRISPR-Cas9 screens, Yang et al. identify key molecular regulators of cell-surface tetraspanin presentation. They unravel that N-glycosylation modification and biosynthesis of lacto-series glycolipids control the internalization of tetraspanins. Their findings constitute an avenue for modulating the diverse functions of tetraspanins, including cancer cell invasion.

Yang et al., 2023, Cell Reports *42*, 112065
February 28, 2023
https://doi.org/10.1016/j.celrep.2023.112065

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Sequential genome-wide CRISPR-Cas9 screens identify genes regulating cell-surface expression of tetraspanins

Jicheng Yang,[1] Fusheng Guo,[1] Hui San Chin,[1] Gao Bin Chen,[1] Chow Hiang Ang,[1] Qingsong Lin,[2] Wanjin Hong,[3] and Nai Yang Fu[1,4,5,6,7,*]

[1]Cancer and Stem Cell Biology Program, Duke-NUS Medical School, Singapore 169857, Singapore
[2]Department of Biological Sciences, National University of Singapore, Singapore 117543, Singapore
[3]Institute of Molecular and Cell Biology, Agency for Science, Technology, and Research (A*STAR), Singapore 138673, Singapore
[4]Department of Physiology, National University of Singapore, Singapore 117593, Singapore
[5]Stem Cells and Cancer Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia
[6]Department of Medicine, University of Melbourne, Parkville, VIC 3010, Australia
[7]Lead contact
*Correspondence: naiyang.fu@duke-nus.edu.sg
https://doi.org/10.1016/j.celrep.2023.112065

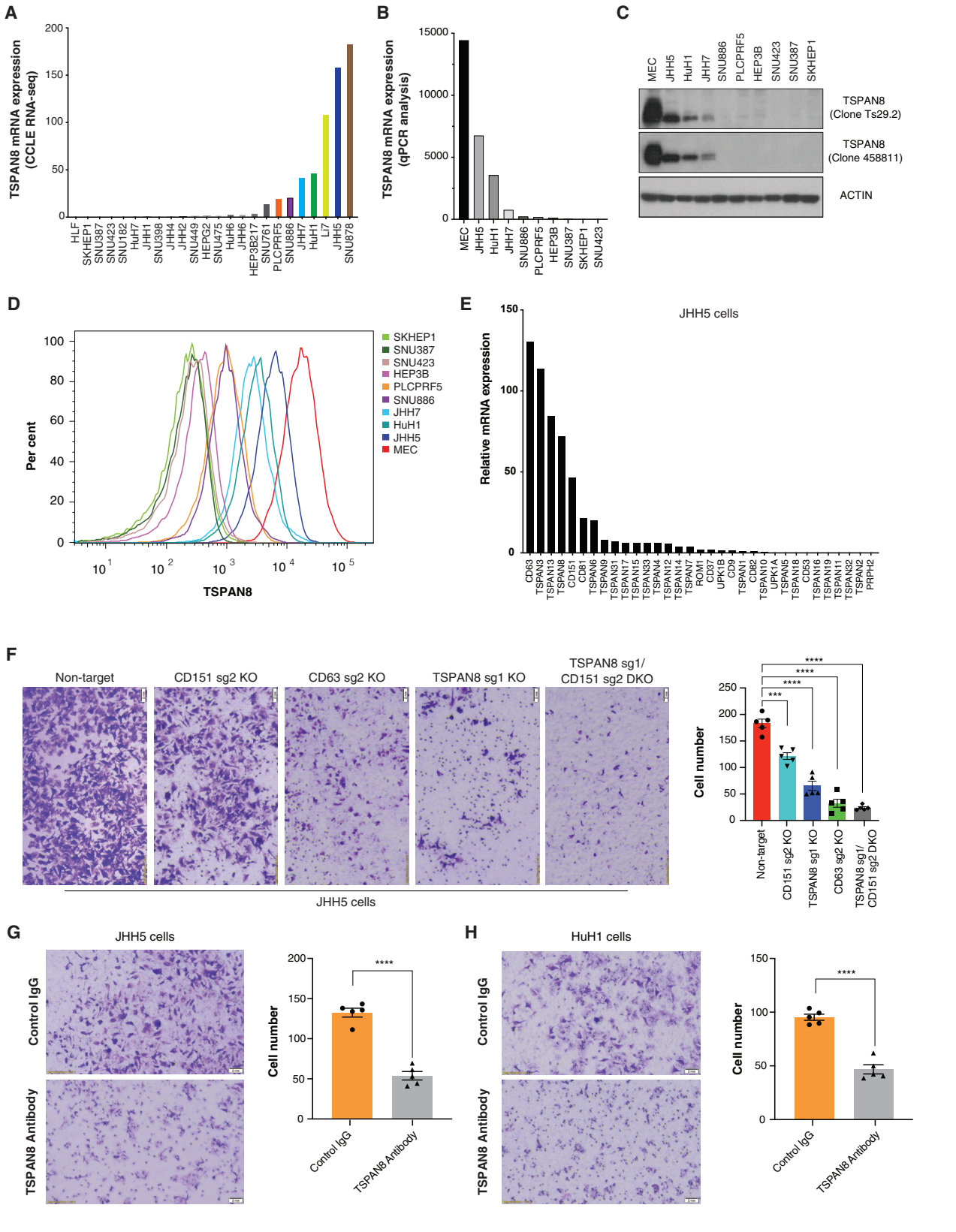
SUMMARY

Tetraspanins, a superfamily of membrane proteins, mediate diverse biological processes through tetraspanin-enriched microdomains in the plasma membrane. However, how their cell-surface presentation is controlled remains unclear. To identify the regulators of tetraspanin trafficking, we conduct sequential genome-wide loss-of-function CRISPR-Cas9 screens based on cell-surface expression of a tetraspanin member, TSPAN8. Several genes potentially involved in endoplasmic reticulum (ER) targeting, different biological processes in the Golgi apparatus, and protein trafficking are identified and functionally validated. Importantly, we find that biantennary N-glycans generated by MGAT1/2, but not more complex glycan structures, are important for cell-surface tetraspanin expression. Moreover, we unravel that SPPL3, a Golgi intramembrane-cleaving protease reported previously to act as a sheddase of multiple glycan-modifying enzymes, controls cell-surface tetraspanin expression through a mechanism associated with lacto-series glycolipid biosynthesis. Our study provides critical insights into the molecular regulation of cell-surface presentation of tetraspanins with implications for strategies to manipulate their functions, including cancer cell invasion.

INTRODUCTION

The tetraspanin superfamily consists of a large group of membrane proteins containing four putative intramembrane domains together with two extracellular loops (one small and one large), a small inner loop, and short N- and C-terminal inner tails.[1] It is important to note that many other proteins also contain four transmembrane domains, but they are not tetraspanin members because they lack the characteristic amino acid sequences of tetraspanins, including 4–8 cysteine residues in the large extracellular loop, with two in a highly conserved cysteine-cysteine-glycine (CCG) triplet motif, and polar residues within transmembrane domains.[1–4] Moreover, the crystal structure of the full-length human CD9 and CD81 proteins revealed a special organization of the four transmembrane domains of tetraspanins.[5,6] They form two separated pairs of helices that create an intramembrane cavity for docking of a cholesterol molecule to regulate its function.[6] Although tetraspanins are conserved throughout evolution,[7] they were first discovered in mammals during the search for novel cell-surface antigens of cancer cells.[8–11] Since then, 33 members of the tetraspanin superfamily in humans have been found. While the original names for some of them are still widely used in the community, such as CD63, CD81, CD82, they all have been accordingly named TSPAN1–TSPAN33. The functional importance of tetraspanins was highlighted by identification of genetic alterations of multiple tetraspanin members in human patients and by genetically engineered laboratory animal models. Accumulating genetic studies in the past 3 decades have indicated that tetraspanins are involved in a multitude of biological processes, including fertilization, parasite and viral infection, mental health, immune response, and tumor development.[1,2]

Emerging evidence suggests that the tetraspanin-enriched microdomain (TEM) (also known as the tetraspanin web) in the plasma membrane is important for diverse biological processes under physiological and pathological conditions. In this unique microdomain, multiple tetraspanins interact with each other as well as other membrane proteins and lipids to form networks with context-dependent functions.[2,12] Of note, a study based on super-resolution microscopy suggests that nanoclusters are normally formed by up to 10 molecules of a single member of



(legend on next page)

tetraspanins instead of multiple different members.[13] Clusters of different tetraspanin members generally do not overlap, but they may interact with each other dynamically and form a TEM in the plasma membrane. Great efforts have also been made to address how tetraspanins function through their physical interactions with other molecular partners in TEMs. More than 60 proteins have been reported to interact with different tetraspanin members.[1,2,14] For instance, interactions of tetraspanins with integrins within the TEM indicate involvement of tetraspanins in cellular adhesion, migration, and invasion.[15] In addition, their interactions with membrane receptors, such as growth factor receptors and other receptors mediating intracellular signaling pathways, imply that certain tetraspanin members may play a role in cell signal transduction.[16,17]

In mammalian cells, like most transmembrane proteins, biosynthesis of tetraspanins begins with translation by ribosomes. After initiation of translation at the ribosome, the polypeptide encoded by a tetraspanin gene is believed to be targeted to the endoplasmic reticulum (ER) for further synthesis. However, tetraspanins lack a defined signal peptide sequence for ER targeting. Very little is known about how tetraspanins are fed into the ER during ribosomal translation. When tetraspanins arrive at the Golgi apparatus from the ER, distinct and sequential post-translational modifications may occur immediately. For example, multiple tetraspanins have been reported to be palmitoylated in the Golgi.[18,19] However, loss of palmitoylation sites appeared to have little effect on trafficking of tetraspanin proteins to the plasma membrane.[18] Maturation of N-glycosylation modification at the Golgi apparatus has also been suggested to play a role in regulation of tetraspanins.[20,21] N-glycosylation of a few tetraspanin members, including CD63, TSPAN1, and CD82, has been demonstrated in different studies.[21–24] However, the molecular details and functional consequences of palmitoylation, glycosylation, and other post-translational modifications of tetraspanins in regulation of their trafficking to the plasma membrane and their cell-surface expression are poorly understood. Here, we explored the molecular regulation of cell-surface presentation of tetraspanins and its impacts on their role in cancer cell invasion.

RESULTS

Cell-surface tetraspanins regulate invasion of cancer cells

Many tetraspanin proteins were originally identified as human tumor-associated antigens and shown to play significant roles in multiple stages of cancer development.[14,25] Among them, TSPAN8 was found to be overexpressed in different epithelial cancers.[26–34] Moreover, cell-surface TSPAN8 was recently identified as a marker and important regulator for quiescent mammary stem cells.[35,36] To establish *in vitro* culture systems to investigate the regulation and function of *TSPAN8* in cancer cells, we analyzed *TSPAN8* expression in the RNA sequencing (RNA-seq) dataset from the Cancer Cell Line Encyclopedia (CCLE), where detailed genetic characterizations of more than 1,000 established human cancer cell lines have been conducted. We found that *TSPAN8* mRNA is highly expressed in a subset of cancer cell lines derived from different tissues. For instance, among the 24 liver cancer cell lines, half of them do not express *TSPAN8*. However, JHH5, SNU878, HuH1, and Li7 display high *TSPAN8* expression (Figure 1A). The relative expression of *TSPAN8* mRNA analyzed by qPCR is largely consistent with the RNA-seq data in the CCLE (Figure 1B) and the western blot results (Figure 1C). We then carried out fluorescence-activated cell sorting (FACS) analysis of cell-surface TSPAN8 expression by using a rat monoclonal antibody that recognizes the extracellular domain of human TSPAN8 (Figure 1D). Among all liver cancer cell lines tested, MEC cells expressed the highest levels of TSPAN8 at the mRNA and protein levels. Of note, MEC is a liver cancer cell line derived from a patient with cholangiocarcinoma but is not included in the CCLE.[37] The expression levels of TSPAN8 analyzed by western blotting (detecting total expression) and FACS (detecting cell-surface expression) correlate well in the different cell lines tested, implying that the vast majority of TSPAN8 proteins in these cell lines are likely localized on the cell surface. Interestingly, we found that multiple tetraspanin members are often co-expressed in the same liver cancer cell lines. For example, mRNA expression of *CD63*, *CD151*, *TSPAN3*, *TSPAN8*, and *TSPAN13* are readily detected in JHH5, SNU878, and HuH1 cells (Figures 1E, S1A, and S1B). To test the role of different tetraspanins in regulation of cell invasion, we generated a series of JHH5 knockout cell lines lacking either one or two tetraspanin members by using a CRISPR-Cas9 approach based on a doxycycline-inducible sgRNA (single-guide RNA) expression system (Figures S2A–S2C). Notably, all knockout cell lines used in this study were a pool of cells generated by FACS based on expression of the corresponding cell-surface marker(s). We found that deletion of any tested tetraspanin members could significantly impair the invasion activity of JHH5 cells, with CD63 and TSPAN8 knockout displaying more profound effects (Figures 1F and S1C). The important role of TSPAN8 in promoting invasion was also confirmed in HuH1 cells (Figures S1D–S1F). Conversely, stable expression of exogenous TSPAN8 significantly promoted invasion of M213 cells, a liver

Figure 1. Cell-surface expression of tetraspanins mediates cancer cell invasion

(A) RNA-seq data from the CCLE (Cancer Cell Line Encyclopedia), showing a high TSPAN8 mRNA level in a subset of human liver cancer cell lines.

(B) qPCR analysis of *TSPAN8* mRNA expression in liver cancer cell lines.

(C) Western blot analysis of the total TSPAN8 level in liver cancer cell lines. Two different monoclonal anti-TSPAN8 antibodies were used. n = 3.

(D) Representative FACS plot showing the cell-surface expression of TSPAN8 in the indicated liver cancer cell lines. n = 3.

(E) RNA-seq data from the CCLE, showing co-expression of multiple members of tetraspanins in JHH5 cells.

(F) Representative images of the Matrigel invasion assay, showing that the absence of different tetraspanins profoundly impairs invasion of JHH5 cells. ***$p < 0.0005$, ****$p < 0.0001$.

(G and H) Inhibition of cell invasion by an anti-TSPAN8 antibody. HuH1 and JHH5 cells, which express high levels of TSPAN8, were pre-incubated with TSPAN8 antibody or control immunoglobulin G (IgG), respectively. ****$p < 0.0001$.

cancer cell line lacking endogenous TSPAN8 expression (Figures S1G–S1I). Most importantly, a monoclonal anti-TSPAN8 antibody significantly impaired invasion of JHH5 and HuH1 cells (Figures 1G and 1H), suggesting that cell-surface expression of TSPAN8 is critical for promoting cancer cell invasion.

Genome-wide CRISPR-Cas9 screens to identify genes regulating cell-surface TSPAN8 expression

To identify the potential regulators of cell-surface expression of tetraspanins, we performed unbiased genome-scale loss-of-function CRISPR-Cas9 screens on MEC and JHH5 cells (Figures 1B–1D) based on detection of cell-surface TSPAN8 expression levels. A large lentivirus-based sgRNA library[38] was used to perform the screens on ~300 millions cells to achieve good coverage (Figure 2A). To enrich cells with decreased cell-surface TSPAN8 levels, we conducted multiple rounds of FACS for 5%–10% TSPAN8^{low} cells, with 5–7 days of culture expansion of sorted cells for each round, until a TSPAN8^{low} population clearly appeared (Figures 2B and S3A). Notably, all 10 and 6 sgRNAs targeting TSPAN8 in the library were enriched in the final TSPAN8^{low} population in the screens on MEC and JHH5 cells, respectively. This strongly indicates that our screens were robust (Figures 2C, 2D, S3B, and S3C). While the top candidate genes, such as SPPL3, MGAT1, and MGAT2, were commonly identified in the two independent screens on MEC and JHH5 cells, the data derived from the screening on MEC cells appears to be more robust. This is likely due to the higher cell-surface expression of TSPAN8 in MEC cells to allow better separation of TSPAN8^{low} cells (Figures 2B and S3A). We hence mainly focused on candidate target genes with $p < 0.01$ from the screening on MEC cells for subsequent analyses (Table S2). Gene Ontology (GO) enrichment analysis showed that genes in a couple of defined pathways, including "protein N-linked glycosylation via asparagine" and "catalytic activity, acting on a glycoprotein," are involved in regulation of cell-surface TSPAN8 expression (Figure 2E).

N-glycosylation regulates the cell-surface presentation of TSPAN8 and other tetraspanins

MGAT1 (α-1,3-mannosyl-glycoprotein 2-β-N-acetylglucosaminyltransferase) and MGAT2 (α-1,6-mannosyl-glycoprotein 2-β-N-acetylglucosaminyltransferase) are two essential enzymes at the initial steps of maturation of the N-glycosylation process at the Golgi apparatus[39,40] (Figure 3A). Multiple sgRNAs targeting these two genes were highly enriched in the TSPAN8^{low} cells in both screens on MEC and JHH5 cells (Figures 2C, 2D, S3B, and S3C). We validated the impact of their deletion on cell-surface TSPAN8 expression by CRISPR-Cas9 (Figure S2). We then established MGAT1 and MGAT2 knockout MEC and JHH5 cell lines by FACS of TSPAN8^{low} cells (Figure 3B). The decreased cell-surface expression of TSPAN8 in MEC MGAT1 and MGAT2 knockout (KO) cells was further confirmed by an additional monoclonal anti-TSPAN8 antibody for FACS analysis (Figure S3D). Notably, deficiency of MGAT1 showed more profound effects on attenuating cell-surface TSPAN8 expression, which is in line with the fact that MGAT1 is the most upstream N-acetylglucosaminyltransferase (Figure 3A). We next performed a rescue experiment to confirm that re-expression of wild-type exogenous MGAT1, but not the enzyme-dead mutants (C121R/D289N and C121R/R413K),[41,42] was able to completely restore cell-surface TSPAN8 expression in MEC MGAT1 KO cells (Figure 3C). Interestingly, the MGAT1 mutant (C121R) with dampened enzymatic activity could partially restore cell-surface TSPAN8 expression. We also examined whether MGAT1 and MGAT2 regulate cell-surface expression of other tetraspanin members. Indeed, their disruption reduced the cell-surface expression of all tested tetraspanin members to various extents in MEC (Figure 3D) and JHH5 cells (Figure S3E).

To validate N-glycosylated modification of a protein in cells, a classic approach is to subject cell lysates to enzymatic de-glycosylation by peptide-N-glycosidase F (PNGase F) and analysis by western blotting.[43] The migration of TSPAN8 protein in SDS gels is much faster in total cell lysates treated with PNGase F (Figure S3F). We also sought to conduct PNGase F treatment on intact and live cells to test whether the N-glycans are attached to the extracellular domains of TSPAN8. The glycans of TSPAN8 could be completely removed (Figure S3G) without any effect on detection of cell-surface TSPAN8 expression by FACS (Figure S3H). These data not only indicated that the potential N-glycosylation site(s) of TSPAN8 are at its extracellular regions but also provided evidence suggesting that the binding affinity of the anti-TSPAN8 antibody used for CRISPR-Cas9 screens is independent of its N-glycosylation modification. Unlike many other tetraspanin members, CD9 is predicted to contain a single potential N-glycosylation site in its small extracellular loop, but this remains to be validated.[44] Interestingly, we did not detect N-glycosylation modification of CD9 in the MEC and SNU878 cell lines (Figures S3I and S3J), while its cell-surface expression is also regulated by MGAT1/2 (Figure 3D).

To further evaluate the MGAT1 and MGAT2 mediated N-glycosylated status of TSPAN8 and CD63, the total lysates of the MEC non-target control, MGAT1 KO, and MGAT2 KO cells were subjected to PNGase F and endoglycosidase H (Endo H) and analyzed by western blotting.[43] Of note, Endo H does not cleave complex glycans (Figure 3A).Treatment of cell lysates from control cells with PNGase F, but not Endo H, facilitated migration of TSPAN8 and CD63 proteins in SDS-PAGE gels, indicating that complex N-glycosylation occurs on TSPAN8 and CD63 proteins (Figure 3E). In line with this, the N-glycosylation patterns of TSPAN8 and CD63 proteins were altered in MEC MGAT1 KO and MGAT2 KO cells, and the removal of their glycans remained resistant to Endo H in MGAT2-KO cells but became sensitive in MGAT1 KO cells (Figure 3E). We also confirmed that kifunensine, an inhibitor of mannosidases, completely blocked N-glycosylation of TSPAN8 and CD63 at 2 μM (Figure 3F) and decreased the cell-surface expression level of TSPAN8 (Figure S3K). Moreover, KO of MGAT1 and kifunensine almost completely blocked the invasion activity of JHH5 cells, while MGAT2 KO only showed around 50% of inhibition (Figures 3G and 3H).

We next sought to identify the N-glycosylation site(s) in TSPAN8 protein by mutagenesis. Two Asn-X-Ser/Thr (X is any amino acid except proline) consensus sequences (37NDS and 118NET) are present in human TSPAN8 protein. However, the N118 residue was found to be the only site for N-glycosylation of human TSPAN8 protein because mutation of N to A at the site completely destroyed its N-glycosylation (Figures S4A and

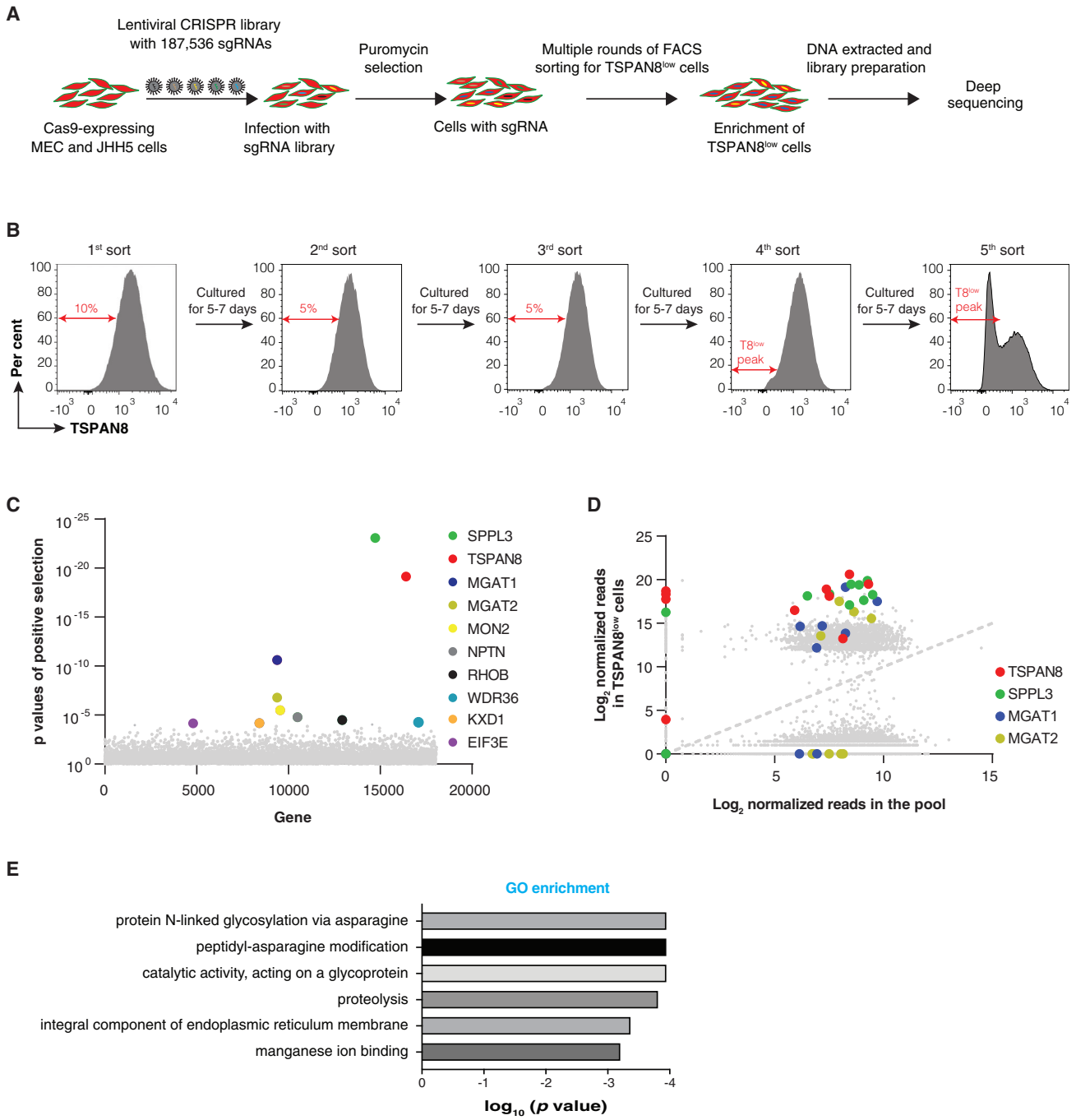


Figure 2. Genome-wide CRISPR-Cas9 screening to identify genes regulating the cell-surface expression of TSPAN8

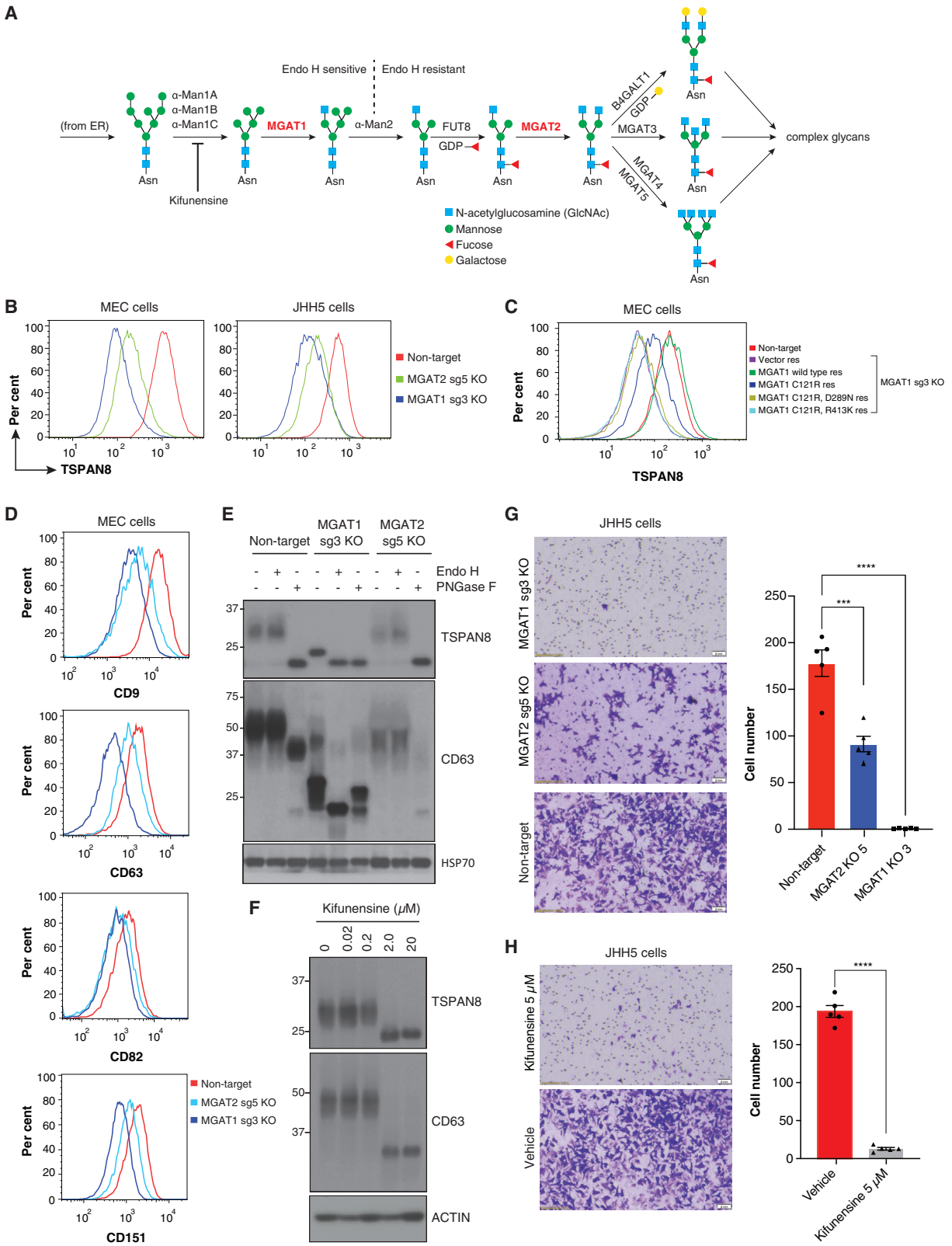
(A) Schematic illustrating the workflow of genome-wide loss-of-function CRISPR-Cas9 screening. JHH5 and MEC cells stably expressing Cas9 were infected with the lentivirus-based sgRNA library. The TSPAN8^{low} cell population was enriched by multiple rounds of sorting and culture. The sgRNAs enriched in the TSPAN8^{low} cell population were identified by next-generation sequencing (NGS).

(B) FACS plots showing multiple rounds of sorting to enrich cells with decreased TSPAN8 levels. The 5%–10% population of total cells with low TSPAN8 expression was sorted for each round and cultured for expansion before next sorting until a distinct TSPAN8^{low} population could be readily detected.

(C) Identification of gene hits in the screen of MEC cells. The top 10 candidate genes, ranked by p values, are highlighted.

(D) Scatterplot showing enrichment of individual sgRNAs of the selected candidate genes.

(E) Gene Ontology (GO) term analysis showing that genes that are targeted by the enriched sgRNAs are involved in the protein N-linked glycosylation pathway. Candidate genes with $p < 0.01$ were selected for the analysis.



(legend on next page)

S4B). This mutation also profoundly impaired the cell-surface expression level of TSPAN8 (Figure S4C). Of note, the N-glycosylation site identified in the human TSPAN8 protein is evolutionally conserved across different species (Figure S4D). To obtain further evidence to support the critical role of N-glycosylation at N118 for TSPAN8-mediated cell invasion, we made stable lines expressing exogenous wild-type TSPAN8 or the N118A mutant in JHH5 TSPAN8 KO cells. While the total expression levels of the wild type and mutant were similar, cell-surface expression of the N118A mutant was dramatically reduced (Figures S4E and S4F). Accordingly, the N118 mutant was not able to restore the invasion activity of JHH5 TSPAN8 KO cells (Figure S4G).

More complex glycan modification and core fucosylation are dispensable for cell-surface expression of tetraspanins

While MGAT1 and MGAT2 were among the top hits in our CRISPR-Cas9 screens on MEC and JHH5 cells, other MGAT enzymes mediating more complex glycan modification in the N-glycosylation maturation cascade at the Golgi apparatus in mammalian cells, such as MGAT3 and MGAT5, were not identified in our screens (Figure 3A). Moreover, FUT8, the enzyme for core fucosylation, was not identified in our screens either. Nevertheless, we sought to test whether these genes may play a role in regulating the cell-surface expression of tetraspanins. It has been well characterized that these genes mediate the cell-surface levels of their glycan species that are specifically recognized by a cognate lectin. MEC MGAT3, MGAT5, and FUT8 KO cells were generated by the CRISPR-Cas9 approach, followed by FACS based on the cell-surface levels of their cognate N-glycan species (Figures S5A–S5C). Interestingly, KO of these genes indeed showed no effect on cell-surface expression of any tested tetraspanin members in MEC cells (Figure S5D).

Validation of other identified candidate genes regulating cell-surface tetraspanin expression

A number of other genes predicted to locate at the Golgi apparatus and regulate Golgi apparatus functions and protein trafficking were also identified in our CRISPR-Cas9 screens. Among them, PI4KB (phosphatidylinositol 4-kinase beta) and GORASP2 (Golgi reassembly stacking protein 2) have been shown to play a role in phosphatidylinositol phosphate biosynthetic process at the Golgi apparatus and in the stacking of Golgi apparatus cisternae, respectively.[45,46] C10orf76, also known as ARMH3 (Armadillo-like helical domain-containing protein 3), has been reported to be involved in regulation of Golgi apparatus organization and may form a complex with PI4KB, but its function remains largely unexplored.[47] MON2 has been shown to be a regulator of trafficking between endosomes and the Golgi apparatus.[48] We confirmed that KO of these genes in MEC cells significantly reduced cell-surface expression of TSPAN8 and other tetraspanin members (Figures S6A–S6E). Tetraspanin proteins do not contain a signal peptide sequence for membrane targeting. How they are targeted to the ER for further modification remains unknown. TMEM208 (Transmembrane Protein 208) protein has recently been reported to be a key molecule for a novel ER-targeting pathway in yeast, while its function in mammals remains to be investigated.[49,50] Interestingly, TEME208 was identified as a hit in our screen, and its deletion significantly reduced the cell-surface expression level of TSPAN8. KO of TEME208 and GORASP2 (Golgi Reassembly Stacking Protein 2) in cells with reduced TSAPN8 expression was confirmed by western blotting (Figures S6F and S6G).

SPPL3 regulates the cell-surface expression of tetraspanins

Signal peptide peptidase-like 3 (SPPL3) is a Golgi apparatus-localized, intramembrane-cleaving aspartic protease that cleaves type II membrane protein substrates in or close to its luminal transmembrane domain boundary.[51] It was identified as a top hit in our CRISPR-Cas9 screens on MEC and JHH5 cells (Figures 2C, 2D, S3B, and S3C). Indeed, disruption of SPPL3 dramatically diminished the cell-surface TSPAN8 expression levels in these two cell lines (Figures 4A and S7A). Using a cell-surface biotinylation assay, the decreased TSPAN8 cell-surface expression levels were further validated in MGAT1 and SPPL3 KO cells compared with non-target control cells, while the total TSPAN8 protein levels remained constant in these cells (Figure S7B). We then carried out a rescue experiment by introducing wild-type and a known enzyme-dead mutant of SPPL3[52] into MEC SPPL3 KO cells. We found that only wild-type SPPL3, but not the

Figure 3. Deletion of MGAT1 and MGAT2 impairs cell-surface expression of tetraspanins

(A) Schematic showing the N-linked glycosylation process in the Golgi apparatus. N-linked glycosylation is initiated at the ER to generate a high-mannose type of glycans, followed by subsequent modification by Golgi apparatus-localized glycosyltransferases, including MGAT1 and MGAT2 (red), to form complex glycans. Kifunensine disrupts N-glycosylation by blocking α-mannosidase I activity. Removal of glycan in glycoproteins is resistant to Endo H after modification by α-mannosidase II.

(B) Representative FACS plots showing that deletion of MGAT1 and MGAT2 by CRISPR-Cas9 profoundly decreases the cell-surface expression of TSPAN8 in MEC and JHH5 cells. n = 3.

(C) Representative FACS plot showing that re-expression of wild-type MGAT1, but not the enzyme-dead mutants (C121R/D289N and C121R/R413K), completely restores TSPAN8 cell-surface expression in MEC MGAT1 KO cells. The single mutant (C121R) with reduced enzymatic activity only partially restores TSPAN8 expression. n = 3.

(D) Representative FACS plot showing the critical role of MGAT1 and MGAT2 in regulation of cell-surface expression of other tetraspanin members. n = 3.

(E) Western blot analysis of TSPAN8 and CD63 proteins in MEC MGAT1 and MGAT2 KO cells. The total cell lysates were subjected to enzymatic de-glycosylation by PNGase F or Endo H before western blot analysis. n = 3.

(F) Western blot analysis showing that kifunensine blocks N-glycosylation of TSPAN8 and CD63 in MEC cells. n = 3 independent experiments.

(G) Inhibition of JHH5 cell invasion by KO of MGAT1 and MGAT2. Values represent mean ± SEM of cell numbers. ***$p < 0.0005$, ****$p < 0.0001$.

(H) Inhibition of JHH5 cell invasion by kifunensine. Cells were pretreated with kifunensine for 72 h before the invasion assay was carried out. Values represent mean ± SEM of cell numbers. ****$p < 0.0001$.

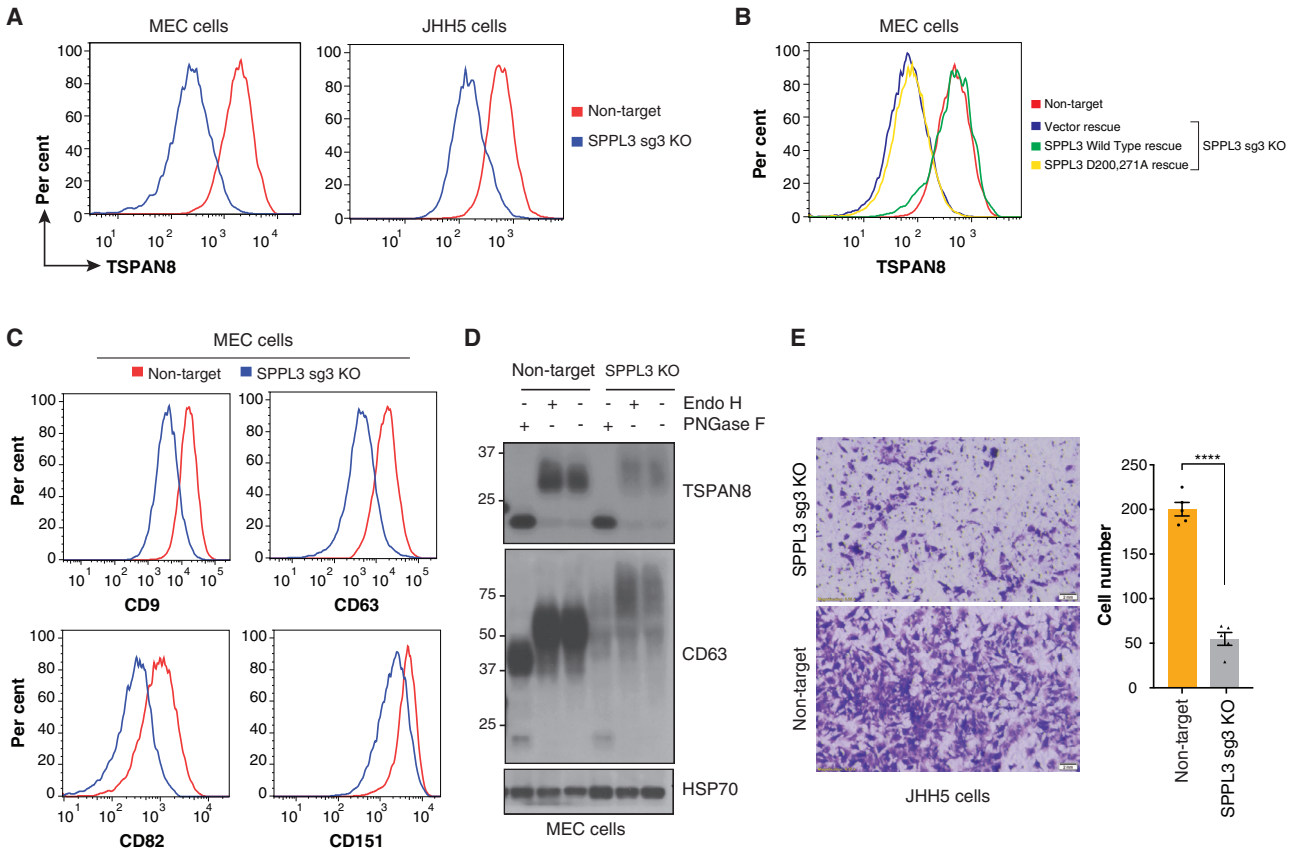


Figure 4. Cell-surface expression of tetraspanins is profoundly impaired in the absence of SPPL3 expression

(A) Representative FACS plots showing that SPPL3 KO greatly diminishes the cell-surface expression of TSPAN8 in MEC and JHH5 cells. n = 3.

(B) Representative FACS plot showing that wild-type SPPL3, but not the enzyme-dead mutant, restores cell-surface TSPAN8 expression in SPPL3 KO cells. n = 3.

(C) Representative FACS plots showing that cell-surface expression of different tetraspanin members is impaired in SPPL3 KO MEC cells. n = 3.

(D) Alteration of N-glycosylation patterns of TSPAN8 and CD63 proteins in SPPL3 KO MEC cells. n = 3.

(E) Representative images showing impaired invasion of JHH5 cells by SPPL3 KO. Values represent mean ± SEM of cell numbers. ****$p < 0.0001$.

enzyme-activity-dead mutant, restored cell-surface TSPAN8 expression in MEC SPPL3 KO cells, suggesting that the protease activity is necessary for SPPL3-mediated TSPAN8 expression in the plasma membrane (Figure 4B). FACS analysis of multiple tetraspanin members indicated that SPPL3 is a general positive regulator of cell-surface expression of different tetraspanins in MEC, JHH5, and SNU878 cells (Figures 4C, S7C, and S7D). Interestingly, cell-surface expression of the glycoproteins EGFR (Epidermal Growth Factor Receptor) and EPCAM (Epithelial Cell Adhesion Molecule) was not affected by KO of MGAT1 and SPPL3 in MEC and JHH5 cells (Figures S7E and S7F). The vast majority of TSPAN8 and CD63 proteins remained modified by N-glycosylation but showed more smearing and complex patterns in SPPL3 KO cells (Figure 4D). Moreover, KO of SPPL3 profoundly impaired the invasion activity of JHH5 and SNU878 cells (Figures 4E and S7G). SPPL3 has been reported to function as a sheddase to mediate proteolytic release and secretion of glycan-modifying glycosidase and glycosyltransferase enzymes, including MGAT5, ST6GAL1, and B4GALT1.[43,53] However, KO of these proteins was not able to restore cell-surface TSPAN8 expression in SPPL3 KO cells (data not shown).

Genome-wide CRISPR-Cas9 screen to identify genes whose absence restores cell-surface TSPAN8 expression in SPPL3 KO cells

To uncover the underlying mechanism of reduced cell-surface TSPAN8 expression in the absence of SPPL3, we designed and conducted a sequential genome-wide CRISPR-Cas9 screen to identify molecule(s) whose deletion restores cell-surface TSPAN8 expression in SPPL3 KO cells (Figure 5A). Remarkably, a subset of infected cells with complete restoration of cell-surface TSPAN8 expression appeared after multiple rounds of FACS with 1-week culture expansion of sorted cells for each round (Figure 5B). The top 10 genes targeted by sgRNAs enriched in these cells were highlighted (Figure 5C). Of note, 6–10 sgRNAs for B3GNT5 (Lactosylceramide 1,3-N-acetyl-beta-D-glucosaminyltransferase), UGCG (UDP-glucose ceramide glucosyltransferase), SLC35A2 (Solute Carrier Family 35 Member A2), and B4GALT5 (Beta-1,4-galactosyltransferase 5) were highly enriched in these cells (Figure 5D). GO enrichment analysis of the candidate target genes (Table S3) showed that genes in a couple of defined pathways, including "glycolipid metabolic process" and "ceramide biosynthetic process," contributes to SPPL3

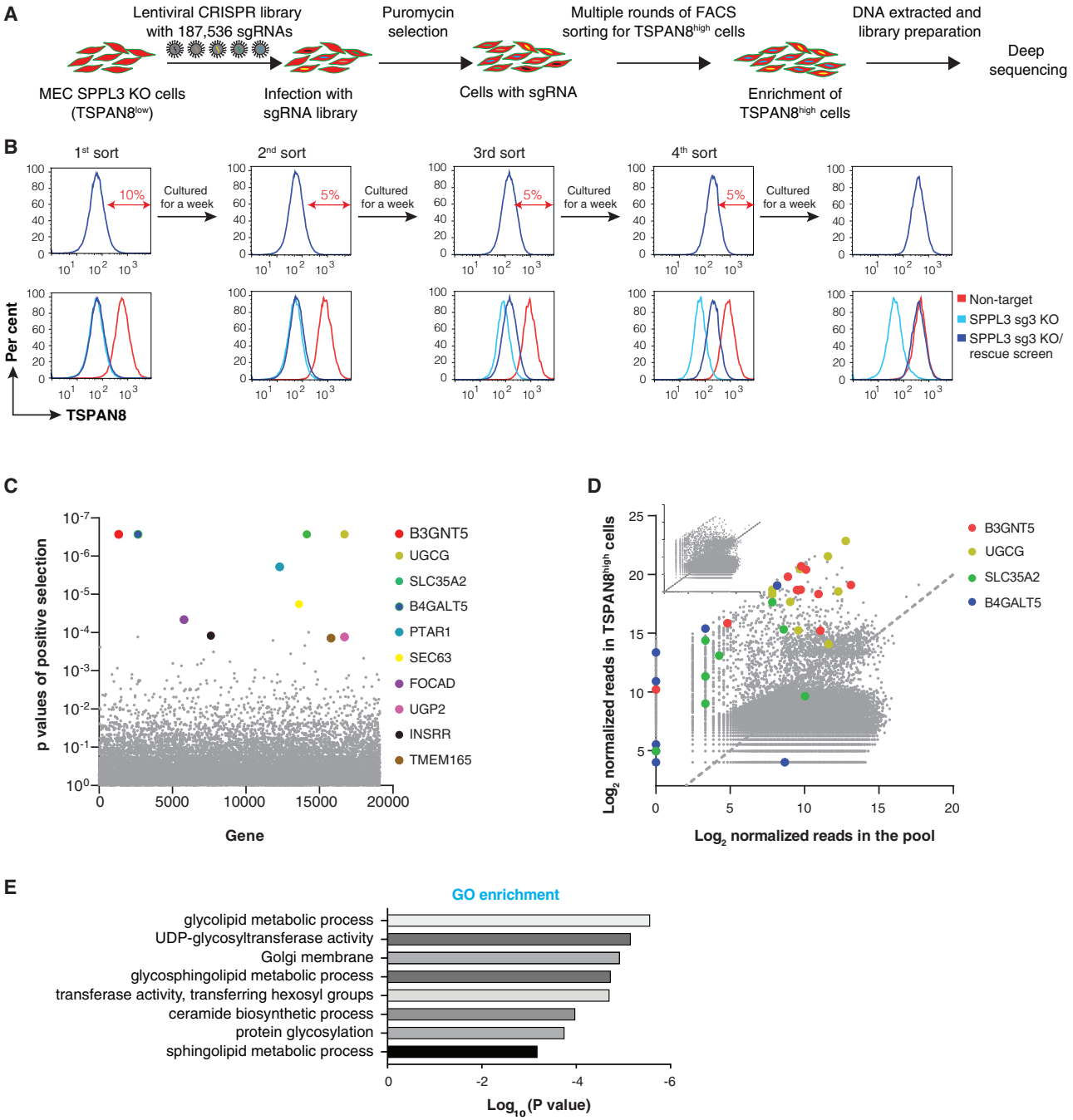


Figure 5. Identification of downstream mediators for the impaired cell-surface expression of tetraspanins caused by SPPL3 deficiency

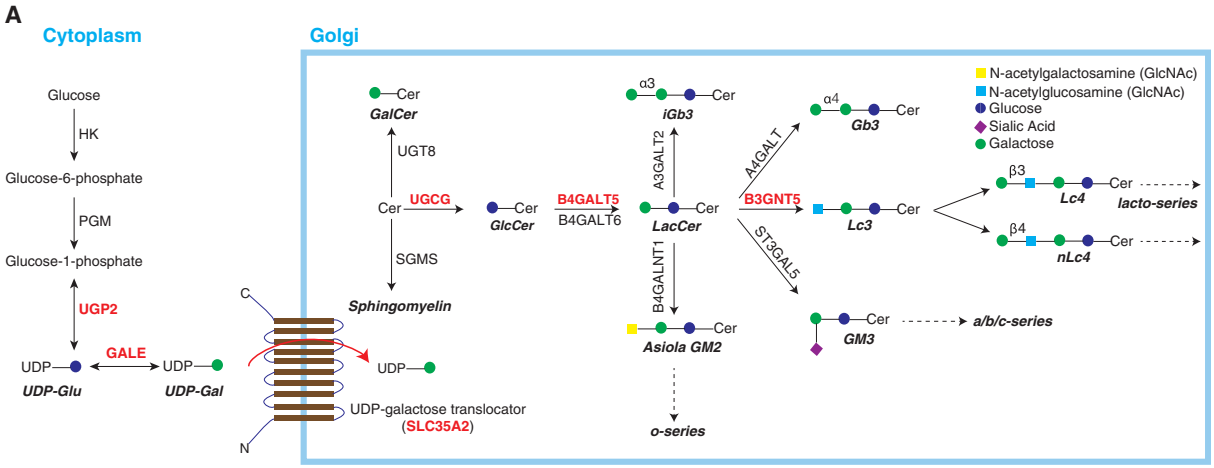
(A) Schematic illustrating the workflow of genome-wide CRISPR-Cas9 screening to identify key downstream effectors for the defect of cell-surface TSPAN8 expression in MEC SPPL3 KO cells.

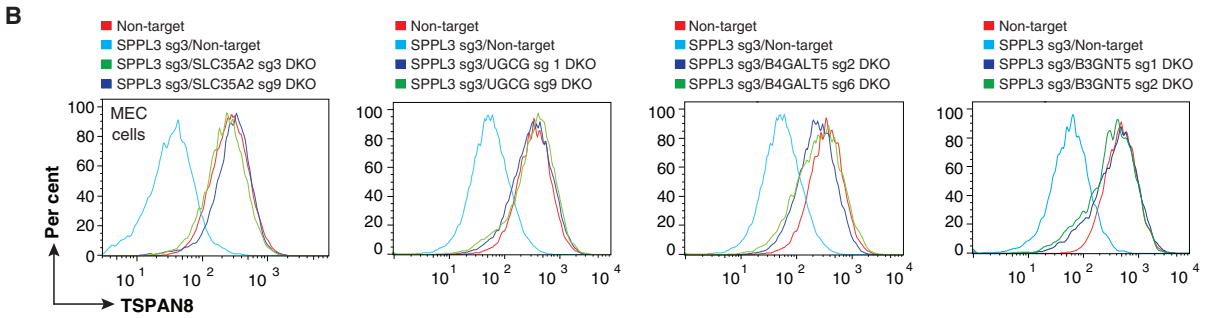
(B) FACS plots showing multiple rounds of sorting to enrich for cells with restored surface expression of TSPAN8 after MEC SPPL3 KO cells were infected with the sgRNA library and selected by puromycin. The 5%–10% population of total cells with high TSPAN8 expression was sorted for each round and cultured for expansion before the next sort until a distinct TSPAN8^{high} population appeared.

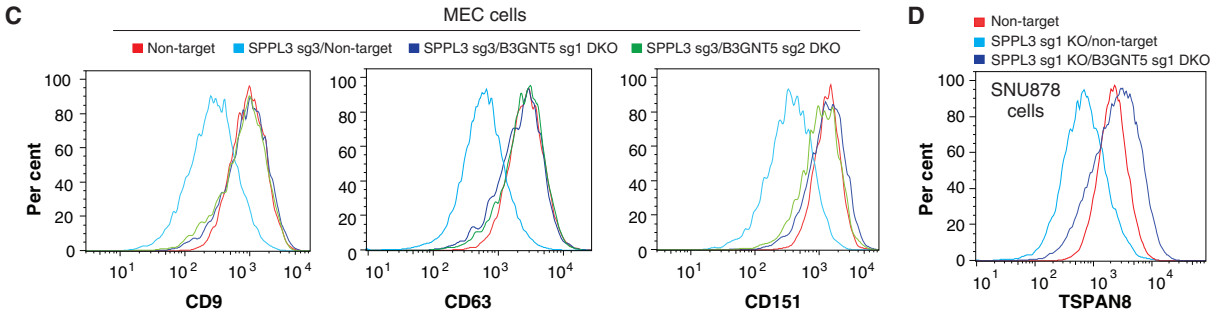
(C) Identification of gene candidates in the screen by comparing the read counts of all sgRNAs for each gene in the TSPAN8^{high} versus pool cell population. The top 10 candidate genes, ranked by p values, are highlighted.

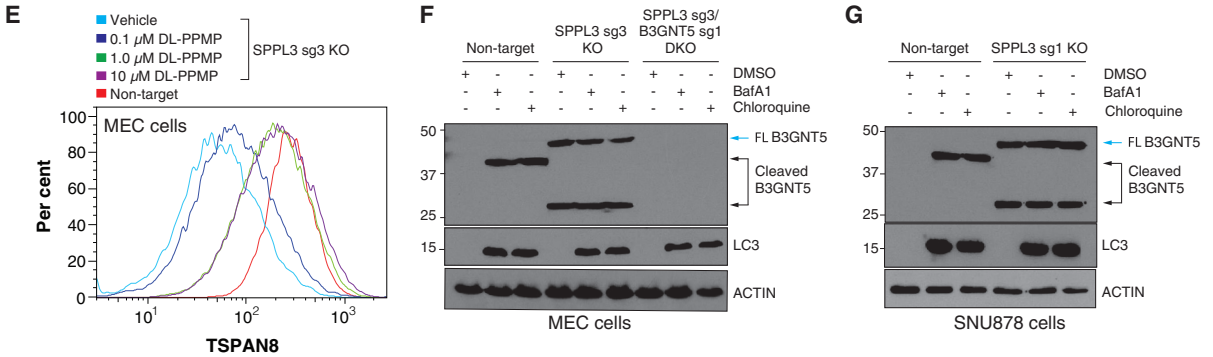
(D) Scatterplot showing enrichment of individual sgRNAs of the selected candidate genes.

(E) GO term analysis showing that genes that are targeted by the enriched sgRNAs are involved in the GSL synthesis and metabolic pathway. The candidate genes with $p < 0.01$ were selected for the analysis.









(legend on next page)

KO-mediated down-regulation of cell-surface TSPAN8 expression (Figure 5E).

Disruption of genes in lacto-series glycolipid synthesis restores cell-surface expression of tetraspanins in SPPL3 KO cells

From the rescue CRISPR-Cas9 screen described above, we noted that UGCG, B4GALT5, and B3GNT5 are enzymes involved in stepwise synthesis of lactotriaosylceramide (Lc3). Interestingly, another candidate gene, SLC35A2, encodes a multi-pass membrane protein that acts as a UDP (uridine diphosphate glucose)-galactose transporter at the Golgi (Figure 6A). Moreover, UGP2 (UDP-glucose pyrophosphorylase 2) and GALE (UDP-galactose-4-epimerase), two enzymes critical for biosynthesis of UDP-galactose in the cytoplasm, were also identified in the rescue screen. KO of each of these six candidate genes using two independent sgRNAs completely restored cell-surface expression of TSPAN8 in MEC SPPL3 KO cells (Figures 6B and S7H). We then established double KO (DKO) cell lines for UGCG/SPPL3, B4GALT5/SPPL3, B3GNT5/SPPL3, and SLC35A2/SPPL3 based on restoration of cell-surface TSPAN8 expression for further analysis (Figure 6B). Remarkably, the cell-surface expression of other tetraspanins, including CD9, CD63, and CD151, was also completely restored in all respective DKO cell lines (Figure 6C). Moreover, complete restoration of cell-surface TSPAN8 expression was also observed in SNU878 SPPL3 KO cells when B3GNT5 was concurrently disrupted, suggesting that the mechanism by which deletion of SPPL3 decreases the cell-surface expression of tetraspanin is likely conserved across different cell lines (Figure 6D). We then determined whether pharmacological inhibition of Lc3 synthesis can restore cell-surface expression of tetraspanins. Indeed, DL-PPMP (DL-threo-1-phenyl-2-palmitoylamino-3-morpholino-1-propanol hydrochloride), a classic inhibitor of UGCG, restored cell-surface TSPAN8 expression in MEC SPPL3 KO cells at 1 μM (Figure 6E).

Because B3GNT5 is the most downstream enzyme in the cascade for the Lc3 synthesis, we hypothesized that SPPL3 may regulate its expression level in cells. qPCR analysis suggested that *B3GNT5* mRNA is highly expressed in MEC cells and is not affected by disruption of SPPL3 (Figure S7I). Moreover, we found that expression of endogenous B3GNT5 protein was barely detectable in the lysates of tested cell lines. We next explored whether B3GNT5 protein is secreted from cells because SPPL3 has been reported to act as a sheddase for several proteins at the Golgi apparatus. However, we did not detect any B3GNT5 protein in the conditioned medium of cultured cells (data not shown). Remarkably, the full-length B3GNT5 protein was found to be highly expressed in MEC and SNU878 SPPL3 KO cells (Figures 6F and 6G). We then examined whether B3GNT5 may be degraded in SPPL3 wild-type cells under normal physiological conditions. We treated the cells with various inhibitors to block different protein degradation pathways. We found that, similar to the key autophagy-associated protein LC3, a truncated form of B3GNT5 protein was massively up-regulated in MEC and SNU878 non-target control cells treated with two different lysosome inhibitors (Figures 6F and 6G). Of note, a smaller form of B3GNT5 at around 27 kDa was also detected in MEC and SNU878 SPPL3 KO cells. It is plausible that additional protease(s) may cleave some B3GNT5 when a large amount of this protein is present in SPPL3 KO cells.

Regulation of internalization of TSPAN8 by N-glycosylation and lacto-series glycolipids

Internalization of transmembrane proteins on the plasma membrane is an important step to control their net cell-surface expression levels. We hence sought to address whether the internalization process of TSPAN8 is affected by disruption of MGAT1 and SPPL3. To measure internalization of cell-surface proteins, the cell-surface antigen-bound antibodies need to be efficiently removed by an optimized stripping buffer. We confirmed that the APC (Allophycocyanin)-conjugated anti-TSPAN8 antibodies bound to cell-surface TSPAN8 were completely removed by the stripping buffer (Figure S7J). We then incubated MEC non-target, MGAT1 KO, and SPPL3 KO cells with the anti-TSPAN8 antibody in complete culture medium for 30 min at 37°C in a CO_2 incubator and carried out the internalization assay. The total level of TSPAN8 (unstripped) was higher in non-target cells compared with MGAT1 KO and SPPL3 KO cells. However, more internalized TSPAN8 antibodies (stripped) were shown in MGAT1 KO and SPPL3 KO cells compared with non-target cells (Figures 7A and 7B). The interplay between glycolipids and galectins has been shown to play a critical role in endocytosis of certain glycoproteins.[54–56] We then explored whether the cell-surface expression of distinct species of glycolipids was altered in MGAT1 KO and SPPL3 KO cells. Remarkably, compared with non-target cells, cell-surface expression of the lacto-series glycolipid nLc4 was markedly up-regulated in SPPL3 KO cells but remained unchanged in MGAT1 KO cells (Figure 7C). Of note, no change in cell-surface expression of the other two tested glycolipid species, GM3 and GD3, was observed in MGAT1 and SPPL3 KO cells

Figure 6. SPPL3 regulates cell-surface expression of tetraspanins by promoting lysosomal degradation of B3GNT5

(A) Schematic illustrating the cascade of GSL biosynthesis and the key enzymes. In the Golgi apparatus membrane, the enzymes UGCG, B4GALT5, and B3GNT5 act consecutively to synthesize the lacto-series GSLs by linkage of sugar residues derived from UDP-glucose, UDP-galactose, and UDP-GlcNAc (N-Acetylglucosamine) donors to ceramide molecules. UDP-galactose is generated in the cytosol and transported into the Golgi apparatus through a translocator encoded by gene *SLC35A2*.

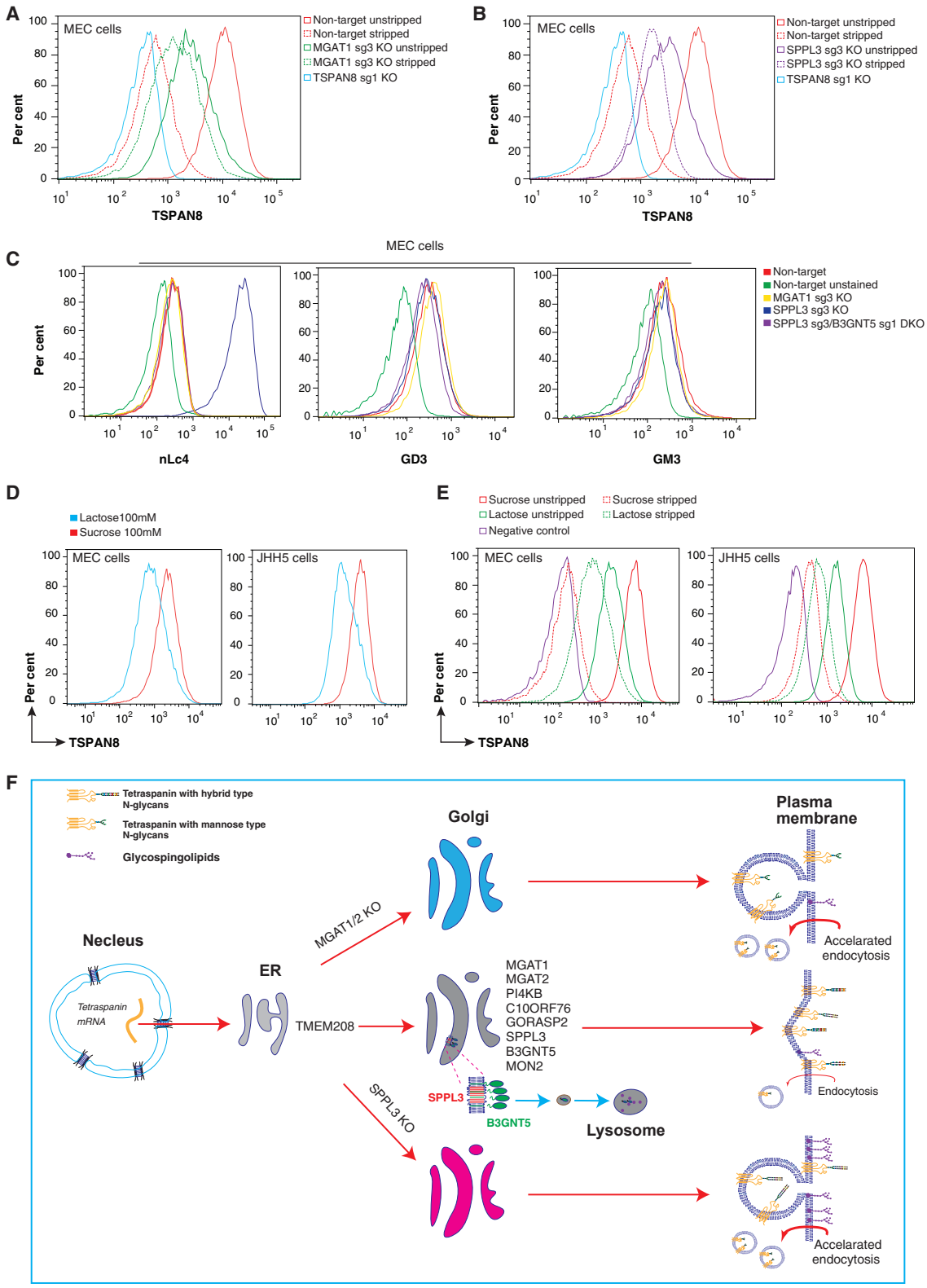
(B) Representative FACS plots showing that deletion of B3GNT5, B4GALT5, UGCG, or SLC35A2 fully restores the cell-surface expression of TSPAN8 in MEC SPPL3-deficient cells. n = 3.

(C) Representative FACS plots showing that deletion of B3GNT5 completely restores the cell-surface expression of other tetraspanins in MEC SPPL3 KO cells. n = 3.

(D) Representative FACS plots showing that deletion of B3GNT5 restores the cell-surface expression of TSPAN8 in SNU878 SPPL3 KO cells. n = 3.

(E) Representative FACS plots showing that DL-PPMP, a ceramide analog and inhibitor of UGCG, restores the cell-surface expression of TSPAN8 in MEC SPPL3 KO cells. n = 3.

(F and G) Western blot analysis showing that SPPL3 mediates B3GNT5 cleavage and subsequent degradation by the lysosome. Bafilomycin A1 (BafA1) and chloroquine were used to inhibit the lysosome. Full-length (FL) B3GNT5 protein is readily detected in MEC and SNU878 SPPL3-deficient cells. The cleaved form of the B3GNT5 protein is readily detected in control cells with lysosome inhibition. Accumulation of LC3 was used as indicator for lysosome inhibition. n = 3.



(legend on next page)

(Figure 7C). Lactose is a general inhibitor of galectin proteins.[57–59] We then investigated the potential effects of lactose treatment on cell-surface presentation of tetraspanins. Interestingly, short-term treatment (30 min) with lactose significantly diminished cell-surface TSPAN8 expression in MEC and JHH5 cells (Figure 7D). We next addressed whether lactose treatment could affect endocytosis of TSPAN8. We incubated MEC and JHH5 cells with the APC-conjugated anti TSPAN8 antibody in the presence of lactose or sucrose at 37°C in a CO_2 incubator for 15 min and conducted the internalization assay. The total amount of TSPAN8 antibodies (unstripped) was lower in cells treated with lactose. However, many more internalized TSPAN8 antibodies (stripped) were shown in lactose-treated cells (Figure 7E). Collectively, these data suggest that the endocytosis process mediated by N-glycosylation and glycolipids of the plasma membrane plays an important role in determining the cell-surface expression of TSPAN8.

DISCUSSION

Previous studies have shown that distinct tetraspanin members play a critical role in regulating cell motility and invasion in various cancer cell lines, likely through modulating the functions of their various binding protein partners that are involved in cell cytoskeletal contraction and polarity, cell-to-cell adhesion and communication, cell-to-ECM (extracellular matrix) adhesion, and ECM modification.[14,25] In this study, we first surveyed the expression of TSPAN8 in established human liver cancer cell lines and showed that it is highly expressed in a subset of the tested cell lines. Interestingly, we also found that a panel of tetraspanin membrane proteins, including CD63, CD151, TSPAN3, TSPAN8, and TSPAN13, tend to be concurrently expressed in liver cancer cell lines, while it remains to be determined whether their transcription is activated by similar factors (Figures 1E, S1A, and S1B). Nevertheless, we confirmed the important role of these tetraspanins in promoting invasion of liver cancer cells by CRISPR-Cas9 KO. Moreover, by blocking TSPAN8 with a specific antibody, we provided evidence demonstrating that cell-surface expression of TSPAN8 is critical for this function (Figures 1G and 1H).

To gain insights into how cell-surface expression of tetraspanins is controlled in mammalian cells, we performed genome-wide loss-of-function CRISPR-Cas9 screens based on cell-surface TSPAN8 expression. To generate robust data with less background noise, we sorted and expanded the cells for multiple rounds for our CRISPR screens. A number of genes that are potentially involved in ER targeting, multiple biological processes in the Golgi apparatus, and protein trafficking were identified and functionally validated. GO term analysis of the candidate genes suggested that the N-glycosylation process could be a key mechanism for regulation of cell-surface expression of TSPAN8. We hence focused on this direction for our study. N-glycosylation of proteins can be predicted based on the presence of the consensus Asn-X-Ser/Thr sequence, where X could be any residue except for proline. Based on this, most tetraspanins in humans are predicted to have at least one N-linked glycosylation site at their large extracellular domains. Our results from the mutagenesis analysis demonstrated that a single N-glycosylation site is present in TSPAN8 and critical for its cell-surface presentation. In mammalian cells, the N-glycosylation process is initiated at the ER, but specific structural elements, such as core fucosylation and further modification/branching of glycans for their maturation, occur in the Golgi apparatus. N-glycans are modified by a cascade consisting of a series of glycosidases and glycosyltransferases that act sequentially to modify glycans in the Golgi apparatus and consequently dictate the glycan status of N-glycoproteins (Figure 3A). Intriguingly, only two of the multiple glycosyltransferases, MGAT1 and MGAT2, were identified in our CRISPR screens, while it is generally believed that the elongated and complex glycans on proteins through stepwise modification of glycans by a series of glycosyltransferases may be critical for the maturation and precise presence of glycoproteins. We showed that the biantennary glycan modification mediated by MGAT1 and MGAT2 can promote the presence of TSPAN8 and other tetraspanins at the cell surface. However, more complex glycan modifications catalyzed by MGAT3 and MGAT5 appear to be dispensable for this process, while they have been shown to regulate the cell-surface expression and trafficking of other glycoproteins.[60,61] FUT8 catalyzes the transfer of core fucose to N-linked complex glycoproteins,[62] and it has been shown to play a critical role in regulation of cell-surface expression of N-glycoproteins, such as PD-1, PD-L2, and EGFR.[63–65] However, our results indicate that the core fucose mediated by FUT8 does not contribute to the cell-surface presentation of tetraspanins. Collectively, these data suggest that the underlying mechanism for regulation of cell-surface expression of tetraspanins could be unique compared

Figure 7. Cell-surface retention and endocytosis of TSPAN8 protein are tightly controlled by N-glycosylation and SPPL3-mediated lacto-series glycolipid synthesis

(A and B) Representative FACS plots showing accelerated internalization of TSPAN8 protein in MEC MGAT1 or SPPL3 KO cells compared with the non-target control. MEC TSPAN8 KO cells were used as a negative control for staining. "Unstripped" and "stripped" represent total and internalized TSPAN8, respectively. n = 3.

(C) Representative FACS plots showing expression of different glycolipid species in the indicated MEC control and KO cell lines. n = 3.

(D) Representative FACS plots showing reduced cell-surface expression of TSPAN8 by treatment with lactose. Sucrose was used as a control. n = 3.

(E) Representative FACS plots showing accelerated internalization of TSPAN8 in the presence of lactose. "Unstripped" and "stripped" represent total and internalized TSPAN8, respectively. n = 3.

(F) Proposed model for regulation of cell-surface expression of tetraspanins. Several genes were identified to play a role in regulation of tetraspanin trafficking from sequential genome-wide loss-of-function CRISPR-Cas9 screens. TMEM208 is localized at the ER, and the rest are resident in the Golgi apparatus. In wild-type cells, tetraspanin proteins are modified by N-glycosylation and traffic to the cell surface. Their internalization into the cells occurs at a slow rate. In the absence of MGAT1/2, the N-glycosylation modification of tetraspanin proteins is blocked, and their internalization is facilitated. Release from the Golgi apparatus membrane and subsequent degradation is an important mechanism for keeping B3GNT5 at a low level in cells. In the absence of SPPL3, a large amount of FL and functional B3GNT5 protein is retained in the Golgi apparatus membrane, which, in turn, leads to high expression of nLc4 in the plasma membrane and accelerates endocytosis of tetraspanin proteins.

with other N-glycoproteins. In line with this, cell-surface presentation of the glycoproteins EGFR and EPCAM was not affected by KO of MGAT1 in the tested cell lines.

SPPL3 was another top-ranked gene identified in our CRISPR screens. Previous studies showed that SPPL3 mainly acts like a sheddase to promote secretion of multiple glycosyltransferases and glycosidases from cells, which, in turn, keeps these proteins at a low level within the cells and plays a role in the N-glycosylation process.[43,53] To gain insights into the mechanism by which SPPL3 regulates TSPAN8 trafficking to the cell surface, we designed a genome-wide CRISPR-Cas9 screen to identify genes whose KO is able to restore cell-surface TSPAN8 expression in SPPL3 KO cells. Remarkably, the genes encoding two biosynthesis enzymes (UGP2 and GALE) in the cytoplasm, a UDP-galactose translocator at the Golgi apparatus membrane (SLC35A2), and all three sequential enzymes (UGCG, B4GALT5, and B3GNT5) for synthesis of lacto-series glycolipid were identified as top hits from the rescue CRISPR screen. By using gene deletion and pharmacological inhibition approaches, we validated that dysregulation in lacto-series glycolipid synthesis is the key mechanism for impaired cell-surface expression of tetraspanins in SPPL3 KO cells. These results also highlight that a sequential CRISPR-Cas9 screening strategy based on rescue of a phenotype can serve as a robust unbiased methodology for decoding the mechanism of a candidate gene identified from the primary functional CRISPR-Cas9 screen. In our efforts to address how SPPL3 regulates the proteins in the cascade of lacto-series glycolipid synthesis, we found that full-length B3GNT5, a single-pass type II membrane protein located at the Golgi, was barely detectable in wild-type cells but remarkably elevated in SPPL3 KO cells. Moreover, we further showed that soluble B3GNT5 generated by SPPL3-mediated cleavage in wild-type cells is rapidly degraded by lysosomes. Based on these results, we proposed here a model where SPPL3 mediates B3GNT5 cleavage and its subsequent lysosomal degradation (Figure 7F). Interestingly, two recent studies reported that SPPL3 deletion impaired detection of cell-surface expression of MHC I (major histocompatibility complex class I) and CD59, and this phenotype was rescued by concurrent disruption of B3GNT5.[66,67] In one of the studies, it was shown that SPPL3 acts as a sheddase to trigger secretion of B3GNT5 from tumor cells.[66] Hence, it is plausible that the mechanism by which soluble B3GNT5 released from the Golgi apparatus membrane through cleavage by SPPL3 is kept at low levels could be cell context dependent.

Glycosphingolipids (GSLs) consist of a heterogeneous group of membrane lipids formed by covalent conjugation of glycans to a lipid (ceramide) core. GSLs are transported to and located primarily in the outer leaflet of the plasma membrane after they are synthesized in the Golgi apparatus. The composition of GSLs is known to undergo remarkable changes during different development stages and cellular statuses, while the underlying mechanisms are not fully understood. GSLs in the plasma membrane have been shown to be required for clustering and membrane bending and further internalization of different bacterial toxins as well as glycosylated cargo glycoproteins generated in mammalian cells through clathrin-independent endocytosis (CIE).[56,57,68,69] The process has been termed the glycolipid-lectin (GL-Lect) hypothesis. However, the detailed mechanisms involving GSLs behind these processes are still largely unknown because it is challenging to isolate and analyze GSLs, considering the high variability in their structure and limited experimental methods.[70] As described above, the full-length B3GNT5 protein was found to be massively upregulated in SPPL3 KO cells. We confirmed that the accumulated B3GNT5 protein is functional in SPPL3 KO cells because nLc4, a GSL derived from Lc3, was dramatically enriched on the cell surface. Interestingly, functional associations of tetraspanins with different GSLs have been reported in previous studies. For instance, inhibition of the GSL biosynthetic pathway destabilizes CD82-containing TEMs.[71] Formation of the GM2 and CD82 complex inhibits Met tyrosine kinase and its crosstalk with integrins.[72] Moreover, interaction of GM3 and CD9 has been shown to play a role in oncogenic transformation and malignancy.[73–75] However, it remains unclear whether the accelerated internalization of tetraspanins in SPPL3 KO cells is mediated through their physical interaction with GSLs. Moreover, while it is likely that expression of nLc4 and other lacto-series glycolipids contribute to the accelerated endocytosis of tetraspanins, which, in turn, reduces their net cell-surface expression levels, it remains unclear which lacto-series glycolipid(s) are most important and how they facilitate internalization of tetraspanins. It is also important to note that the expression of nLc4 was not altered in MGAT1 KO cells, while disruption of MGAT1 was also found to significantly accelerate the internalization process of TSPAN8. Galectins are a subfamily of soluble proteins that specifically bind β-galactoside sugars that are covalently conjugated to the proteins with either N-linked or O-linked glycosylation. In humans, distinct types of galectins encoded by more than 10 different *LGALS* genes with functional redundancy have been identified. Multiple galectins have been reported to organize the cell membrane via formation of large-scale molecular networks called the galectin lattice, which retains cell-surface glycoproteins at the plasma membrane.[54,76–78] Treatment of lactose, a wide-spectrum inhibitor of galectins, enhances internalization of TSPAN8 and reduces its cell-surface expression level. Other than binding to glycoproteins, galectins are able to bind to different GSLs, including lacto-series GSLs, on the cell surface.[79] It remains to be further investigated whether neutralization of galectins by massive upregulation of lacto-series GSLs contributes in part to the accelerated endocytosis of TSPAN8 in SPPL3 KO cells. Nevertheless, regulation of tetraspanin endocytosis could be a good paradigm to better understand the delicate equilibrium between cell-surface retention by the galectin lattice and internalization by glycolipid-lectin interaction.[56]

In summary, using a sequential CRISPR-Cas9 screening approach, we showed here that biantennary N-glycans and expression of a particular GSL series in the plasma membrane play a critical role in regulation of cell-surface expression levels of tetraspanins. Our findings not only provide critical insights into the molecular regulation of the dynamic trafficking process of tetraspanins but also yield strategies, through inhibition of N-glycosylation and synthesis of lacto-series GSLs, to module their critical roles in diverse biological processes, including cell invasion and cancer metastasis, although more precise molecular details remain to be deciphered in the future.

Limitations of the study

In this study, the regulation of cell-surface presentation of several tetraspanin members was explored in cell lines derived from human liver cancer. Antibodies for other tetraspanins that are also expressed in these cell lines are currently not available, which hinders investigation of their cell-surface expression. Many tetraspanin members are not expressed in these cell lines, and their trafficking needs to be addressed in appropriate cell model systems. For example, certain tetraspanin members are highly expressed and play important roles in immune cells. It remains to be addressed whether their cell-surface presentation is controlled via mechanisms similar to those reported here. While our data suggest that N-glycosylation and biogenesis of lacto-series glycolipids regulate endocytosis of tetraspanins, the underlying mechanism remains unclear. Moreover, several other regulators located at the ER and Golgi apparatus, such as TMEM208, PI4KB, C10ORF76, and GORASP2, have also been identified in our genome-wide CRISPR-Cas9 screens, but how these molecules control the abundance of cell-surface tetraspanins is even less clear and requires future research. The critical role of MGAT1/2 and SPPL3 in cancer cell invasion and cancer metastasis also needs to be validated by *in vivo* studies in the future.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Trans-well assay
  - Overexpression plasmids
  - Doxycycline inducible sgRNA lentiviral vectors
  - Transient transfection
  - Production of lentivirus and infection
  - Flow cytometry analysis and FACS sorting
  - Generation of CRISPR-Cas9 knockout (KO) and double knockout (DKO) rescue cells
  - Genome-wide CRISPR/Cas9 knockout screens
  - Western blot analysis
  - Enzymatic deglycosylation of live cells
  - Endocytosis assay
  - Effects of lactose treatment on endocytosis of N-glycoproteins
  - Surface biotinylation assay
  - Reverse transcription and qPCR analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112065.

ACKNOWLEDGMENTS

We are grateful to Z.M. Lim for assistance with FACS analysis and sorting. We thank Dr. J.C. Hu for critical reading and valuable comments on the manuscript. We also thank C. Bouchiex, M. Herold, T.B. Tean, and U. Mandel for kindly sharing reagents. This work was supported by MOE Tier 2 (MOE-T2EP30121-0013), the Gilead Research Scholars Liver Disease Program (Asia) (2020-Gilead-001), Khoo Bridge Funding Awards, and NMRC OF-IRG. H.S.C. (MOH-000546) and C.H.A. (MOH-000935) were supported by NMRC OF-YIRG.

AUTHOR CONTRIBUTIONS

N.Y.F. and J.Y. designed the study. J.Y., F.G., H.S.C., and C.H.A. performed experiments and analyzed data. G.B.C. conducted the bioinformatic analysis. N.Y.F., J.Y., Q.L., and W.H. carried out interpretation of data. N.Y.F. and J.Y. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 22, 2022
Revised: November 16, 2022
Accepted: January 18, 2023
Published: January 31, 2023

REFERENCES

1. Charrin, S., Jouannet, S., Boucheix, C., and Rubinstein, E. (2014). Tetraspanins at a glance. J. Cell Sci. *127*, 3641–3648. https://doi.org/10.1242/jcs.154906.
2. Hemler, M.E. (2005). Tetraspanin functions and associated microdomains. Nat. Rev. Mol. Cell Biol. *6*, 801–811. https://doi.org/10.1038/nrm1736.
3. Robert, J.M.H., Amoussou, N.G., Mai, H.L., Logé, C., and Brouard, S. (2021). Tetraspanins: useful multifunction proteins for the possible design and development of small-molecule therapeutic tools. Drug Discov. Today *26*, 56–68. https://doi.org/10.1016/j.drudis.2020.10.022.
4. Hemler, M.E. (2001). Specific tetraspanin functions. J. Cell Biol. *155*, 1103–1107. https://doi.org/10.1083/jcb.200108061.
5. Umeda, R., Satouh, Y., Takemoto, M., Nakada-Nakura, Y., Liu, K., Yokoyama, T., Shirouzu, M., Iwata, S., Nomura, N., Sato, K., et al. (2020). Structural insights into tetraspanin CD9 function. Nat. Commun. *11*, 1606. https://doi.org/10.1038/s41467-020-15459-7.
6. Zimmerman, B., Kelly, B., McMillan, B.J., Seegar, T.C.M., Dror, R.O., Kruse, A.C., and Blacklow, S.C. (2016). Crystal structure of a full-length human tetraspanin reveals a cholesterol-binding pocket. Cell *167*, 1041–1051.e11. https://doi.org/10.1016/j.cell.2016.09.056.
7. Garcia-España, A., Chung, P.J., Sarkar, I.N., Stiner, E., Sun, T.T., and Desalle, R. (2008). Appearance of new tetraspanin genes during vertebrate evolution. Genomics *91*, 326–334. https://doi.org/10.1016/j.ygeno.2007.12.005.
8. Oren, R., Takahashi, S., Doss, C., Levy, R., and Levy, S. (1990). TAPA-1, the target of an antiproliferative antibody, defines a new family of transmembrane proteins. Mol. Cell Biol. *10*, 4007–4015. https://doi.org/10.1128/mcb.10.8.4007-4015.1990.
9. Szala, S., Kasai, Y., Steplewski, Z., Rodeck, U., Koprowski, H., and Linnenbach, A.J. (1990). Molecular cloning of cDNA for the human tumor-associated antigen CO-029 and identification of related transmembrane antigens. Proc. Natl. Acad. Sci. USA *87*, 6833–6837. https://doi.org/10.1073/pnas.87.17.6833.
10. Demetrick, D.J., Herlyn, D., Tretiak, M., Creasey, D., Clevers, H., Donoso, L.A., Vennegoor, C.J., Dixon, W.T., and Jerry, L.M. (1992). ME491 melanoma-associated glycoprotein family: antigenic identity of ME491, NKI/C-3, neuroglandular antigen (NGA), and CD63 proteins. J. Natl. Cancer Inst. *84*, 422–429. https://doi.org/10.1093/jnci/84.6.422.

11. Takagi, S., Fujikawa, K., Imai, T., Fukuhara, N., Fukudome, K., Minegishi, M., Tsuchiya, S., Konno, T., Hinuma, Y., and Yoshie, O. (1995). Identification of a highly specific surface marker of T-cell acute lymphoblastic leukemia and neuroblastoma as a new member of the transmembrane 4 superfamily. Int. J. Cancer *61*, 706–715. https://doi.org/10.1002/ijc.2910610519.

12. Levy, S., and Shoham, T. (2005). Protein-protein interactions in the tetraspanin web. Physiology *20*, 218–224. https://doi.org/10.1152/physiol.00015.2005.

13. Zuidscherwoude, M., Göttfert, F., Dunlock, V.M.E., Figdor, C.G., van den Bogaart, G., and van Spriel, A.B. (2015). The tetraspanin web revisited by super-resolution microscopy. Sci. Rep. *5*, 12201. https://doi.org/10.1038/srep12201.

14. Hemler, M.E. (2014). Tetraspanin proteins promote multiple cancer stages. Nat. Rev. Cancer *14*, 49–60. https://doi.org/10.1038/nrc3640.

15. Berditchevski, F. (2001). Complexes of tetraspanins with integrins: more than meets the eye. J. Cell Sci. *114*, 4143–4151.

16. Levy, S., and Shoham, T. (2005). The tetraspanin web modulates immune-signalling complexes. Nat. Rev. Immunol. *5*, 136–148. https://doi.org/10.1038/nri1548.

17. Termini, C.M., and Gillette, J.M. (2017). Tetraspanins function as regulators of cellular signaling. Front. Cell Dev. Biol. *5*, 34. https://doi.org/10.3389/fcell.2017.00034.

18. Yang, X., Claas, C., Kraeft, S.K., Chen, L.B., Wang, Z., Kreidberg, J.A., and Hemler, M.E. (2002). Palmitoylation of tetraspanin proteins: modulation of CD151 lateral interactions, subcellular distribution, and integrin-dependent cell morphology. Mol. Biol. Cell *13*, 767–781. https://doi.org/10.1091/mbc.01-05-0275.

19. Kovalenko, O.V., Yang, X., Kolesnikova, T.V., and Hemler, M.E. (2004). Evidence for specific tetraspanin homodimers: inhibition of palmitoylation makes cysteine residues available for cross-linking. Biochem. J. *377*, 407–417. https://doi.org/10.1042/BJ20031037.

20. Tominaga, N., Hagiwara, K., Kosaka, N., Honma, K., Nakagama, H., and Ochiya, T. (2014). RPN2-mediated glycosylation of tetraspanin CD63 regulates breast cancer cell malignancy. Mol. Cancer *13*, 134. https://doi.org/10.1186/1476-4598-13-134.

21. Scholz, C.J., Sauer, G., and Deissler, H. (2009). Glycosylation of tetraspanin Tspan-1 at four distinct sites promotes its transition through the endoplasmic reticulum. Protein Pept. Lett. *16*, 1244–1248. https://doi.org/10.2174/092986609789071234.

22. Ono, M., Handa, K., Withers, D.A., and Hakomori, S. (2000). Glycosylation effect on membrane domain (GEM) involved in cell adhesion and motility: a preliminary note on functional alpha3, alpha5-CD82 glycosylation complex in ldlD 14 cells. Biochem. Biophys. Res. Commun. *279*, 744–750. https://doi.org/10.1006/bbrc.2000.4030.

23. Schäfer, T., Starkl, P., Allard, C., Wolf, R.M., and Schweighoffer, T. (2010). A granular variant of CD63 is a regulator of repeated human mast cell degranulation. Allergy *65*, 1242–1255. https://doi.org/10.1111/j.1398-9995.2010.02350.x.

24. Zhang, H., Li, X.J., Martin, D.B., and Aebersold, R. (2003). Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. Nat. Biotechnol. *21*, 660–666. https://doi.org/10.1038/nbt827.

25. Zöller, M. (2009). Tetraspanins: push and pull in suppressing and promoting metastasis. Nat. Rev. Cancer *9*, 40–55. https://doi.org/10.1038/nrc2543.

26. Yue, S., Mu, W., and Zöller, M. (2013). Tspan8 and CD151 promote metastasis by distinct mechanisms. Eur. J. Cancer *49*, 2934–2948. https://doi.org/10.1016/j.ejca.2013.03.032.

27. Zhang, H.S., Liu, H.Y., Zhou, Z., Sun, H.L., and Liu, M.Y. (2020). TSPAN8 promotes colorectal cancer cell growth and migration in LSD1-dependent manner. Life Sci. *241*, 117114. https://doi.org/10.1016/j.lfs.2019.117114.

28. Zhu, R., Gires, O., Zhu, L., Liu, J., Li, J., Yang, H., Ju, G., Huang, J., Ge, W., Chen, Y., et al. (2019). TSPAN8 promotes cancer cell stemness via activation of sonic Hedgehog signaling. Nat. Commun. *10*, 2863. https://doi.org/10.1038/s41467-019-10739-3.

29. Xie, C., and McGrath, N.A. (2020). TSPAN8 and distant metastasis of nasopharyngeal carcinoma cells. Ann. Transl. Med. *8*, 165. https://doi.org/10.21037/atm.2019.10.102.

30. Wei, L., Li, Y., and Suo, Z. (2015). TSPAN8 promotes gastric cancer growth and metastasis via ERK MAPK pathway. Int. J. Clin. Exp. Med. *8*, 8599–8607.

31. Wang, H., Rana, S., Giese, N., Büchler, M.W., and Zöller, M. (2013). CD44v6 and alpha6beta4 are biomarkers of migrating pancreatic cancer-initiating cells. Int. J. Cancer *133*, 416–426. https://doi.org/10.1002/ijc.28044.

32. Voglstaetter, M., Thomsen, A.R., Nouvel, J., Koch, A., Jank, P., Navarro, E.G., Gainey-Schleicher, T., Khanduri, R., Groß, A., Rossner, F., et al. (2019). Tspan8 is expressed in breast cancer and regulates E-cadherin/catenin signalling and metastasis accompanied by increased circulating extracellular vesicles. J. Pathol. *248*, 421–437. https://doi.org/10.1002/path.5281.

33. Heo, K., and Lee, S. (2020). TSPAN8 as a novel emerging therapeutic target in cancer for monoclonal antibody therapy. Biomolecules *10*, 388. https://doi.org/10.3390/biom10030388.

34. Fang, T., Lin, J., Wang, Y., Chen, G., Huang, J., Chen, J., Zhao, Y., Sun, R., Liang, C., and Liu, B. (2016). Tetraspanin-8 promotes hepatocellular carcinoma metastasis by increasing ADAM12m expression. Oncotarget *7*, 40630–40643. https://doi.org/10.18632/oncotarget.9769.

35. Fu, N.Y., Rios, A.C., Pal, B., Law, C.W., Jamieson, P., Liu, R., Vaillant, F., Jackling, F., Liu, K.H., Smyth, G.K., et al. (2017). Identification of quiescent and spatially restricted mammary stem cells that are hormone responsive. Nat. Cell Biol. *19*, 164–176. https://doi.org/10.1038/ncb3471.

36. Fu, N.Y., Pal, B., Chen, Y., Jackling, F.C., Milevskiy, M., Vaillant, F., Capaldo, B.D., Guo, F., Liu, K.H., Rios, A.C., et al. (2018). Foxp1 is indispensable for ductal morphogenesis and controls the exit of mammary stem cells from quiescence. Dev. Cell *47*, 629–644.e8. https://doi.org/10.1016/j.devcel.2018.10.001.

37. Yoshida, K., Tomizawa, H., Ota, T., Nagashima, T., Kikuchi, H., Watanabe, H., Hashizaki, K., and Yonaha, A. (1990). [Establishment and characterization of human cholaginocarcinoma, MEC, producing carbohydrate antigen 19-9]. Hum. Cell *3*, 346–351.

38. Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., and Sabatini, D.M. (2015). Identification and characterization of essential genes in the human genome. Science *350*, 1096–1101. https://doi.org/10.1126/science.aac7041.

39. Yen, C.L.E., Stone, S.J., Cases, S., Zhou, P., and Farese, R.V., Jr. (2002). Identification of a gene encoding MGAT1, a monoacylglycerol acyltransferase. Proc. Natl. Acad. Sci. USA *99*, 8512–8517. https://doi.org/10.1073/pnas.132274899.

40. Yen, C.L.E., and Farese, R.V., Jr. (2003). MGAT2, a monoacylglycerol acyltransferase expressed in the small intestine. J. Biol. Chem. *278*, 18532–18537. https://doi.org/10.1074/jbc.M301633200.

41. Chen, W., and Stanley, P. (2003). Five Lec1 CHO cell mutants have distinct Mgat1 gene mutations that encode truncated N-acetylglucosaminyltransferase I. Glycobiology *13*, 43–50. https://doi.org/10.1093/glycob/cwg003.

42. Puthalakath, H., Burke, J., and Gleeson, P.A. (1996). Glycosylation defect in Lec1 Chinese hamster ovary mutant is due to a point mutation in N-acetylglucosaminyltransferase I gene. J. Biol. Chem. *271*, 27818–27822. https://doi.org/10.1074/jbc.271.44.27818.

43. Kuhn, P.H., Voss, M., Haug-Kröper, M., Schröder, B., Schepers, U., Bräse, S., Haass, C., Lichtenthaler, S.F., and Fluhrer, R. (2015). Secretome analysis identifies novel signal Peptide peptidase-like 3 (Sppl3) substrates and reveals a role of Sppl3 in multiple Golgi glycosylation pathways. Mol. Cell. Proteomics *14*, 1584–1598. https://doi.org/10.1074/mcp.M115.048298.

44. Boucheix, C., Benoit, P., Frachet, P., Billard, M., Worthington, R.E., Gagnon, J., and Uzan, G. (1991). Molecular cloning of the CD9 antigen. A new family of cell surface proteins. J. Biol. Chem. *266*, 117–122.

45. Baba, T., Alvarez-Prats, A., Kim, Y.J., Abebe, D., Wilson, S., Aldworth, Z., Stopfer, M.A., Heuser, J., and Balla, T. (2020). Myelination of peripheral nerves is controlled by PI4KB through regulation of Schwann cell Golgi function. Proc. Natl. Acad. Sci. USA *117*, 28102–28113. https://doi.org/10.1073/pnas.2007432117.

46. Grond, R., Veenendaal, T., Duran, J.M., Raote, I., van Es, J.H., Corstjens, S., Delfgou, L., El Haddouti, B., Malhotra, V., and Rabouille, C. (2020). The function of GORASPs in Golgi apparatus organization in vivo. J. Cell Biol. *219*, e202004191. https://doi.org/10.1083/jcb.202004191.

47. McPhail, J.A., Lyoo, H., Pemberton, J.G., Hoffmann, R.M., van Elst, W., Strating, J.R.P.M., Jenkins, M.L., Stariha, J.T.B., Powell, C.J., Boulanger, M.J., et al. (2020). Characterization of the c10orf76-PI4KB complex and its necessity for Golgi PI4P levels and enterovirus replication. EMBO Rep. *21*, e48441. https://doi.org/10.15252/embr.201948441.

48. Artan, M., Sohn, J., Lee, C., Park, S.Y., and Lee, S.J.V. (2022). MON-2, a Golgi protein, promotes longevity by upregulating autophagy through mediating inter-organelle communications. Autophagy *18*, 1208–1210. https://doi.org/10.1080/15548627.2022.2039523.

49. Aviram, N., Ast, T., Costa, E.A., Arakel, E.C., Chuartzman, S.G., Jan, C.H., Haßdenteufel, S., Dudek, J., Jung, M., Schorr, S., et al. (2016). The SND proteins constitute an alternative targeting route to the endoplasmic reticulum. Nature *540*, 134–138. https://doi.org/10.1038/nature20169.

50. Talbot, B.E., Vandorpe, D.H., Stotter, B.R., Alper, S.L., and Schlondorff, J.S. (2019). Transmembrane insertases and N-glycosylation critically determine synthesis, trafficking, and activity of the nonselective cation channel TRPC6. J. Biol. Chem. *294*, 12655–12669. https://doi.org/10.1074/jbc.RA119.008299.

51. Papadopoulou, A.A., and Fluhrer, R. (2020). Signaling functions of intramembrane aspartyl-proteases. Front. Cardiovasc. Med. *7*, 591787. https://doi.org/10.3389/fcvm.2020.591787.

52. Makowski, S.L., Wang, Z., and Pomerantz, J.L. (2015). A protease-independent function for SPPL3 in NFAT activation. Mol. Cell Biol. *35*, 451–467. https://doi.org/10.1128/MCB.01124-14.

53. Voss, M., Künzel, U., Higel, F., Kuhn, P.H., Colombo, A., Fukumori, A., Haug-Kröper, M., Klier, B., Grammer, G., Seidl, A., et al. (2014). Shedding of glycan-modifying enzymes by signal peptide peptidase-like 3 (SPPL3) regulates cellular N-glycosylation. EMBO J. *33*, 2890–2905. https://doi.org/10.15252/embj.201488375.

54. Nabi, I.R., Shankar, J., and Dennis, J.W. (2015). The galectin lattice at a glance. J. Cell Sci. *128*, 2213–2219. https://doi.org/10.1242/jcs.151159.

55. Stanley, P. (2014). Galectins CLIC cargo inside. Nat. Cell Biol. *16*, 506–507. https://doi.org/10.1038/ncb2983.

56. Johannes, L., and Billet, A. (2020). Glycosylation and raft endocytosis in cancer. Cancer Metastasis Rev. *39*, 375–396. https://doi.org/10.1007/s10555-020-09880-z.

57. Lakshminarayan, R., Wunder, C., Becken, U., Howes, M.T., Benzing, C., Arumugam, S., Sales, S., Ariotti, N., Chambon, V., Lamaze, C., et al. (2014). Galectin-3 drives glycosphingolipid-dependent biogenesis of clathrin-independent carriers. Nat. Cell Biol. *16*, 595–606. https://doi.org/10.1038/ncb2970.

58. Mathew, M.P., Tan, E., Saeui, C.T., Bovonratwet, P., Sklar, S., Bhattacharya, R., and Yarema, K.J. (2016). Metabolic flux-driven sialylation alters internalization, recycling, and drug sensitivity of the epidermal growth factor receptor (EGFR) in SW1990 pancreatic cancer cells. Oncotarget *7*, 66491–66511. https://doi.org/10.18632/oncotarget.11582.

59. Noll, A.J., Gourdine, J.P., Yu, Y., Lasanajak, Y., Smith, D.F., and Cummings, R.D. (2016). Galectins are human milk glycan receptors. Glycobiology *26*, 655–669. https://doi.org/10.1093/glycob/cww002.

60. Pinho, S.S., Reis, C.A., Paredes, J., Magalhães, A.M., Ferreira, A.C., Figueiredo, J., Xiaogang, W., Carneiro, F., Gärtner, F., and Seruca, R. (2009). The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin. Hum. Mol. Genet. *18*, 2599–2608. https://doi.org/10.1093/hmg/ddp194.

61. Allam, H., Johnson, B.P., Zhang, M., Lu, Z., Cannon, M.J., and Abbott, K.L. (2017). The glycosyltransferase GnT-III activates Notch signaling and drives stem cell expansion to promote the growth and invasion of ovarian cancer. J. Biol. Chem. *292*, 16351–16359. https://doi.org/10.1074/jbc.M117.783936.

62. Pucci, M., Malagolini, N., and Dall'Olio, F. (2021). Glycobiology of the epithelial to mesenchymal transition. Biomedicines *9*, 770. https://doi.org/10.3390/biomedicines9070770.

63. Okada, M., Chikuma, S., Kondo, T., Hibino, S., Machiyama, H., Yokosuka, T., Nakano, M., and Yoshimura, A. (2017). Blockage of core fucosylation reduces cell-surface expression of PD-1 and promotes anti-tumor immune responses of T cells. Cell Rep. *20*, 1017–1028. https://doi.org/10.1016/j.celrep.2017.07.027.

64. Höti, N., Lih, T.S., Pan, J., Zhou, Y., Yang, G., Deng, A., Chen, L., Dong, M., Yang, R.B., Tu, C.F., et al. (2020). A comprehensive analysis of FUT8 overexpressing prostate cancer cells reveals the role of EGFR in castration resistance. Cancers *12*, 468. https://doi.org/10.3390/cancers12020468.

65. Xu, Y., Gao, Z., Hu, R., Wang, Y., Wang, Y., Su, Z., Zhang, X., Yang, J., Mei, M., Ren, Y., et al. (2021). PD-L2 glycosylation promotes immune evasion and predicts anti-EGFR efficacy. J. Immunother. Cancer *9*, e002699. https://doi.org/10.1136/jitc-2021-002699.

66. Jongsma, M.L.M., de Waard, A.A., Raaben, M., Zhang, T., Cabukusta, B., Platzer, R., Blomen, V.A., Xagara, A., Verkerk, T., Bliss, S., et al. (2021). The SPPL3-defined glycosphingolipid repertoire orchestrates HLA class I-mediated immune responses. Immunity *54*, 132–150.e9. https://doi.org/10.1016/j.immuni.2020.11.003.

67. Kawaguchi, K., Yamamoto-Hino, M., and Goto, S. (2021). SPPL3-dependent downregulation of the synthesis of (neo)lacto-series glycosphingolipid is required for the staining of cell surface CD59. Biochem. Biophys. Res. Commun. *571*, 81–87. https://doi.org/10.1016/j.bbrc.2021.06.093.

68. Pezeshkian, W., Hansen, A.G., Johannes, L., Khandelia, H., Shillcock, J.C., Kumar, P.B.S., and Ipsen, J.H. (2016). Membrane invagination induced by Shiga toxin B-subunit: from molecular structure to tube formation. Soft Matter *12*, 5164–5171. https://doi.org/10.1039/c6sm00464d.

69. Cheng, Z.J., Singh, R.D., Sharma, D.K., Holicky, E.L., Hanada, K., Marks, D.L., and Pagano, R.E. (2006). Distinct mechanisms of clathrin-independent endocytosis have unique sphingolipid requirements. Mol. Biol. Cell *17*, 3197–3210. https://doi.org/10.1091/mbc.e05-12-1101.

70. Merrill, A.H., Jr., and Sullards, M.C. (2017). Opinion article on lipidomics: inherent challenges of lipidomic analysis of sphingolipids. Biochim. Biophys. Acta. Mol. Cell Biol. Lipids *1862*, 774–776. https://doi.org/10.1016/j.bbalip.2017.01.009.

71. Odintsova, E., Butters, T.D., Monti, E., Sprong, H., van Meer, G., and Berditchevski, F. (2006). Gangliosides play an important role in the organization of CD82-enriched microdomains. Biochem. J. *400*, 315–325. https://doi.org/10.1042/BJ20060259.

72. Todeschini, A.R., Dos Santos, J.N., Handa, K., and Hakomori, S.I. (2007). Ganglioside GM2-tetraspanin CD82 complex inhibits met and its cross-talk with integrins, providing a basis for control of cell motility through glycosynapse. J. Biol. Chem. *282*, 8123–8133. https://doi.org/10.1074/jbc.M611407200.

73. Miura, Y., Kainuma, M., Jiang, H., Velasco, H., Vogt, P.K., and Hakomori, S. (2004). Reversion of the Jun-induced oncogenic phenotype by enhanced synthesis of sialosyllactosylceramide (GM3 ganglioside). Proc. Natl. Acad. Sci. USA *101*, 16204–16209. https://doi.org/10.1073/pnas.0407297101.

74. Mitsuzuka, K., Handa, K., Satoh, M., Arai, Y., and Hakomori, S. (2005). A specific microdomain ("glycosynapse 3") controls phenotypic conversion and reversion of bladder cancer cells through GM3-mediated interaction of alpha3beta1 integrin with CD9. J. Biol. Chem. *280*, 35545–35553. https://doi.org/10.1074/jbc.M505630200.

75. Satoh, M., Ito, A., Nojiri, H., Handa, K., Numahata, K., Ohyama, C., Saito, S., Hoshi, S., and Hakomori, S.I. (2001). Enhanced GM3 expression, associated with decreased invasiveness, is induced by brefeldin A in bladder cancer cells. Int. J. Oncol. *19*, 723–731.

76. Garner, O.B., and Baum, L.G. (2008). Galectin-glycan lattices regulate cell-surface glycoprotein organization and signalling. Biochem. Soc. Trans. *36*, 1472–1477. https://doi.org/10.1042/BST0361472.

77. Lau, K.S., Partridge, E.A., Grigorian, A., Silvescu, C.I., Reinhold, V.N., Demetriou, M., and Dennis, J.W. (2007). Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. Cell *129*, 123–134. https://doi.org/10.1016/j.cell.2007.01.049.

78. Grigorian, A., Torossian, S., and Demetriou, M. (2009). T-cell growth, cell surface organization, and the galectin-glycoprotein lattice. Immunol. Rev. *230*, 232–246. https://doi.org/10.1111/j.1600-065X.2009.00796.x.

79. Collins, P.M., Bum-Erdene, K., Yu, X., and Blanchard, H. (2014). Galectin-3 interactions with glycosphingolipids. J. Mol. Biol. *426*, 1439–1451. https://doi.org/10.1016/j.jmb.2013.12.004.

80. Young, W.W., Jr., Portoukalian, J., and Hakomori, S. (1981). Two monoclonal anticarbohydrate antibodies directed to glycosphingolipids with a lacto-N-glycosyl type II chain. J. Biol. Chem. *256*, 10967–10972.

81. Aubrey, B.J., Kelly, G.L., Kueh, A.J., Brennan, M.S., O'Connor, L., Milla, L., Wilcox, S., Tai, L., Strasser, A., and Herold, M.J. (2015). An inducible lentiviral guide RNA platform enables the identification of tumor-essential genes and tumor-promoting mutations in vivo. Cell Rep. *10*, 1422–1432. https://doi.org/10.1016/j.celrep.2015.02.002.

82. Chin, H.S., Li, M.X., Tan, I.K.L., Ninnis, R.L., Reljic, B., Scicluna, K., Dagley, L.F., Sandow, J.J., Kelly, G.L., Samson, A.L., et al. (2018). VDAC2 enables BAX to mediate apoptosis and limit tumor development. Nat. Commun. *9*, 4976. https://doi.org/10.1038/s41467-018-07309-4.

83. Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R.A., Liu, J.S., Brown, M., and Liu, X.S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol. *15*, 554. https://doi.org/10.1186/s13059-014-0554-4.

84. Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinf. *10*, 48. https://doi.org/10.1186/1471-2105-10-48.

85. Eden, E., Lipson, D., Yogev, S., and Yakhini, Z. (2007). Discovering motifs in ranked lists of DNA sequences. PLoS Comput. Biol. *3*, e39. https://doi.org/10.1371/journal.pcbi.0030039.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Unconjugated rat monoclonal anti-human TSPAN8	R&D Systems	Cat No. MAB4734; RRID:AB_2257018
Anti-CD9 antibody	Abcam	Clone EPR23105-125 (ab263019)
CD9 Rabbit mAb	Cell Signaling Technology	Cat No. 13174; RRID:AB_2798139
PE-Vio770-conjugated anti-human TSPAN8, REAfinity™	Miltenyi Biotec	Cat No. 130-106-812
APC-conjugated anti-human TSPAN8, REAfinity™	Miltenyi Biotec	Cat No. 130-106-811; RRID:AB_2654251
APC-conjugated anti-human CD9, REAfinity™	Miltenyi Biotec	Cat No. 130-118-808; RRID:AB_2733406
APC-conjugated anti-human CD63, REAfinity™	Miltenyi Biotec	Cat No. 130-118-078; RRID:AB_2733396
APC-conjugated anti-human CD81, REAfinity™	Miltenyi Biotec	Cat No. 130-119-787; RRID:AB_2751844
APC-conjugated anti-human CD82, REAfinity™	Miltenyi Biotec	Cat No. 130-101-282; RRID:AB_2659301
APC-conjugated anti-human CD151, REAfinity™	Miltenyi Biotec	Cat No. 130-103-663; RRID:AB_2655233
PE-conjugated anti-mouse Tspan8	R&D Systems	Cat No. FAB6524P; RRID:AB_10891351
Rabbit anti-B3GNT5	Proteintech Group, USA	Cat No. 20422-1-AP; RRID:AB_10694280
Mouse monoclonal anti- human CD63	Abcam	Cat No. ab59479; RRID:AB_940915
Rabbit anti-LC3B	Cell Signaling Technology	Cat No. 2775; RRID:AB_915950
Mouse monoclonal anti-human TSPAN8	A kind gift from Dr. Claude Boucheix, Inserm/University Paris Sud UA09	Clone Ts29.2, IgG2b
APC -conjugated anti-Biotin	Miltenyi Biotec	Cat No. 130-090-856; RRID:AB_244256
APC anti- human CD326 (EpCAM) Antibody	Biolegend	Cat No. 324208; RRID:AB_756082
APC-conjugated anti-human EGFR, REAfinity™	Miltenyi Biotec	Cat No. 130-110-529; RRID:AB_2651600
Anti-nLc4	A kind gift from Professor Ulla Mandel, University of Copenhagen	Clone 1B2 (Young et al.[80])
Anti-TMEM208	Proteintech	Cat No. 23882-1-AP; RRID:AB_2879349
Anti-GORASP2	Proteintech	Cat No. 10598-1-AP; RRID:AB_2113473
Bacterial and virus strains		
Stbl3™ Chemically Competent E. coli	Invitrogen	Cat# C737303
pMD2.G	Addgene	Addgene, Plasmid No. 12259
pRSV-Rev	Addgene	Addgene, Plasmid No. 12253
pMDLg/pRRE	Addgene	Addgene, Plasmid No. 12251
Biological samples		
Matrigel™ Growth Factor Reduced Basement Membrane Matrix	Corning	Cat No. 354230
Fetal Bovine Serum	Sigma Aldrich	Cat No. 12003C
Chemicals, peptides, and recombinant proteins		
Biotinylated Phaseolus Vulgaris Erythroagglutinin (PHA-E)	Vector Laboratories, Burlingame, USA	Cat No. B11252
Biotinylated Lens Culinaris Agglutinin (LCA)	Vector Laboratories, Burlingame, USA	Cat No. B10455
Biotinylated Phaseolus Vulgaris Leucoagglutinin (PHA-L)	Vector Laboratories, Burlingame, USA	Cat No. B1115
Kifunensine	Caymanchem	Cat No. 10009437
PNGase F	Promega	Cat No. V483A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Endo H	Promega	Cat No. V490A
Puromycin Dihydrochloride	Goldbio	CAS No. 58-58-2; Cat No. P-600-100
4′,6-Diamidino-2-phenylindole dihydrochloride (DAPI)	Invitrogen	CAS No. 28718-90-3; Cat No. D9542;
Deoxyribonuclease I (DNAse I)	Worthington Biochemical Corp	CAS No. 9003-98-9; Cat No. LS002140
IgG from rat serum	Sigma Aldrich	Cat No. I4131
Bafilomycin A1	MedChemExpress	Cat No. HY-100558
Chloroquine	MedChemExpress	Cat No. HY-17589A
TritonTM X-100	Sigma	Cat No. T8787
DMEM media	Hyclone, Cytiva	Cat No. SH30022.01
penicillin/streptomycin	Hyclone, Cytiva	Cat No. SV30010
William's E media	Pan Biotech	Cat No. P04-29510
Trypsin-EDTA (1X) solution	Thermo Fisher Scientific	Cat No. 25200-056
Pfu DNA polymerase	Promega	Cat No. M7745
Opti-MEM	Gibco™	Cat No.31985047
Polyethylenimine	Polysciences	Cat No. 23966-1
7-AAD	Caymanchem	CAS No. 7240-37-1; Cat No. 11397
protease and phosphatase inhibitor cocktail	Roche, Mannheim, Germany	Cat No. C762Q77
Gibco™ Trypsin-EDTA	Life Technologies Corporation	Cat No. 25200-056
1x non-enzymatic cell dissociation solution	Merck	MDL# MFCD00282844, Cat No. C5914
Lactose	Sigma-Aldrich	Cat No. L3750
Sucrose	Sigma-Aldrich	Cat No. S0389
DL-PPMP	Santa Cruz Biotechnology,	Cat No. sc-205655; CAS 149022-18-4
EZ-Link Sulfo-NHS-SS-Biotin	Thermo Scientific	Cat No. A44390
Critical commercial assays		
Pierce™ Cell Surface Biotinylation and Isolation Kit	The Thermo Scientific	Cat No. A44390
GoScriptTM Reverse Transcription Kit	Promega	Cat No. A5000
RNeasy Micro Kit	Qiagen	Cat No. 74004
One-4-All Genomic DNA Miniprep Kit	Bio Basic	Cat No. BS88504
SYBR™ Green PCR Master Mix	Thermo Fisher Scientific	Cat No. 4309155
Experimental models: Cell lines		
KKU-M213	JCRB	JCRB1557
MEC	CCR, Tohoku University	TKG 0629
Hep 3B	ATCC	HB-8064
JHH5	JCRB	JCRB1029
JHH7	JCRB	JCRB1031
SK-HEP1	ATCC	HTB-52
SNU-387	ATCC	CRL-2237
HuH1	JCRB	JCRB0199
PLC/PRF/5	ATCC	CRL-8024
SNU878	KCLB	KCLB# 00878
SNU886	KCLB	KCLB# 00886
SNU423	ATCC	CRL-2238
HEK 293	ATCC	CRL-1573

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oligonucleotides		
Refer to the Table S1 for all sequences of sgRNAs; Refer to the relevant method sections for all other oligonucleotides.	This paper	N/A
Recombinant DNA		
pFU_T2A_Cherry	A gift from Dr. Marco Herold at WEHI	Addgene (Plasmid No. 70182)
pFU_T2A_GFP	In house	It was engineered by replacing mCherry in pFU_T2A_Cherry with eGFP.
pFU_T2A_TagBFP	In house	It was engineered by replacing mCherry in pFU_T2A_Cherry with TagBFP.
FgH1tUTG	A gift from Dr. Marco Herold at WEHI	Addgene (Plasmid No. 70183). It is used for doxycycline-inducing expression of sgRNA for CRISPR/Cas9 knockout in this study.
FgH1tUTB	In house	It was engineered by replacing eGFP in FgH1UTG with TagBFP. It is used for doxycycline-inducing expression of sgRNA for CRISPR/Cas9 knockout in this study.
pFU_Cas9_T2A_mCherry	In house	Expression of Cas9 and mCherry through the T2A system
pFU_MGAT1_T2A_TagBFP	In house	Expression of wild-type MGAT1 and BFP through the T2A system
pFU_MGAT1-C121R _T2A_TagBFP	In house	Expression of MGAT1 C121R mutant & BFP through the T2A system
pFU_MGAT1-C121R/D289N_T2A_TagBFP	In house	Expression of MGAT1 C121R/D289N mutant & BFP through the T2A system
pFU_MGAT1-C121R&R413K_T2A_TagBFP	In house	Expression of MGAT1 C121R/R413K mutant & BFP through the T2A system
pFU_SPPL3_T2A_TagBFP	In house	Expression of wild-type SPPL3 & BFP through the T2A system
pFU_SPPL3-D200,271A_T2A_TagBFP	In house	Expression of SPPL3 D200,271A & BFP through the T2A system
pFU_TSPAN8_T2A_GFP	In house	Expression of wild-type TSPAN8 & GFP through the T2A system
pFU_TSPAN8-N118A_T2A_GFP	In house	Expression of TSPAN8 N118A mutant & GFP through the T2A system
pFU_TSPAN8-N37A_T2A_GFP	In house	Expression of TSPAN8 N37A mutant & GFP through the T2A system
pFU_TSPAN8-N37,118A_T2A_GFP	In house	Expression of TSPAN8 N37, 118A mutant & GFP through the T2A system
pFU_TSPAN8-N37,118,185A_T2A_GFP	In house	Expression of TSPAN8 N37, 118, 185A mutant & GFP through the T2A system
Software and algorithms		
ImageJ	Fiji	RRID: SCR_003070
Gene information	NCBI	ftp://ftp.ncbi.nlm.nih.gov
Cancer Cell Line Encyclopedia (CCLE)	The Broad Institute of MIT & Harvard	https://portals.broadinstitute.org/ccle
Leica Application Suite X	Leica Microsystems GmbH	RRID:SCR_013673
GraphPad Prism 7	GraphPad Software	RRID: SCR_002798
Flowjo	Tree Star	RRID: SCR_008520

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Other		
24-well permeable support plate	Corning	Cat No. 354578
Medical X-ray Film	FUJI	Cat No. 4741019289

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Nai Yang Fu (naiyang.fu@duke-nus.edu.sg or fu@wehi.edu.au).

Materials availability

All reagents will be made available by the lead contact author after completion of a Materials Transfer Agreement.

Data and code availability

Data reported in this paper will be shared by the lead contact upon request.

This paper does not report original code.

Any additional information required to re-analyse the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Detailed information on all the commercially available established cell lines used in this study is listed in key resources table. The MEC cell line was established from pleural effusion of a patient with cholagiocarcinoma.[37] All cell lines were cultured at 37°C tissue culture incubator with humidified atmosphere of 5% CO_2. All cells were cultured in DMEM media (Hyclone, Cytiva) supplemented with 10% (v/v) fetal bovine serum (FBS, Sigma-Aldrich) and 1% (v/v) penicillin/streptomycin (Hyclone, Cytiva) except for JHH5 cells, which were cultured in William's E media (Pan Biotech). Cultured cells were passaged once they reached 80–90% confluency. For drug treatment experiments, cells were cultured in normal medium with the indicated drug concentration for the desired duration as described in the figure legends for each experiment before they were harvested for downstream analysis. For inhibition of mannosidases, cultured cells were treated with Kifunensine at different concentrations for 72 h before harvesting for western blot and FACS analysis. For inhibition of UGCG, cultured cells were treated with DL-PPMP at different concentrations for 72 h before harvesting for FACS analysis. For lysosome inhibition, cultured cells were treated with Bafilomycin A1 and Chloroquine at the concentration of 0.16 μM and 25 μg/mL respectively for 8 h before harvesting for western blotting.

METHOD DETAILS

Trans-well assay

Cells were cultured in the culture media with low FBS (2%) overnight prior to being added into the trans-well permeable supports (Corning, Cat# 354578). Cell suspension prepared in serum-free culture media was incubated in cell migration or invasion chambers, where normal complete media was placed in the lower well, for 24 h in a 37°C tissue culture incubator with humidified atmosphere of 5% CO_2. After the complete removal of non-migrated or non-invaded cells from apical side of the permeable support, cells were then fixed with 100% methanol and stained with crystal violet. Stained cells were imaged under a microscope and counted in five fields for each of the triplicate wells. The values of the counts shown in the figures represent mean ± SEM of cell numbers in five fields from three independent experiments. For the preparation of invasion chamber, Matrigel (Matrigel Growth Factor Reduced Basement Membrane Matrix, ref. 354,230, Corning) was diluted to a final concentration of 2 mg/mL with coating buffer (0.01 M Tris, pH 8.0, 0.7% NaCl) and added into the 24-well permeable support plate (Corning, Cat# 354578). For antibody blocking experiments, cells were pre-incubated with the 10 μg/mL Anti-TSPAN8 (R&D Systems, Monoclonal rat anti-hTSPAN8 antibody, clone #458811, Cat# MAB4734) or 10 μg/mL control IgG isotype (Sigma, Rat IgG, Cat# I4131) for 30 min before adding into the invasion chambers.

Overexpression plasmids

The cDNAs for wild-type human *TSPAN8*, *MGAT1* and *SPPL3* were PCR-amplified by high-fidelity Pfu DNA polymerase (Promega, Cat# M7745) from the human MEC cell line and subsequently cloned into the pFU-T2A lentivirus vector (The plasmid was a gift from Marco Herold, Addgene plasmid #70182). Specifically, the *SPPL3* and *MGAT1* genes were cloned and integrated into the vector at EcoR I/Nhe I sites with the following primers (5′-3′): hSPPL3 Forward (GAGCGAGCAAGCAAGCAAGCA); hSPPL3 Reverse (CGCAGGAAGAGAAGGAACCAAACA), hMGAT1 Fwd (TCCTAATCCCATAGTCCAGAGG) hMGAT1 Rvs (ATGC ACCTAAGAGGGAAACAC). The *TSPAN8* gene was cloned and integrated into the vector at EcoRI/BamHI sites with the following

primers (5′-3′): hTSPAN8 Forward (CTAGAGGATCTGAATTCGCCGCCACCATGGCAGGTGTGAGTGC) hTSPAN8 Reverse (CTGCCCTCACCGGATCCTTTGTTCCCGATC). Substitution mutants of the N-glycosylation sites of TSPAN8 and the key sites for the enzymatic activity of MGAT1 and SPPL3 were generated using standard site-directed mutagenesis. For the rescue experiments in CRISPR knockout cells, expression plasmids of the corresponding genes with silent mutations in the sgRNA target sequence were generated and used. The entire coding region of the target gene in all the plasmids was sequenced to ensure that no error was introduced during plasmid construction.

Doxycycline inducible sgRNA lentiviral vectors

For generation of sgRNA plasmids for CRISPR/Cas9 knockout, the Doxycycline-inducible sgRNA vector system Fgh1tUT was used.[81] We used two versions of the vector the Fgh1tUTG and Fgh1tUTB. FgH1tUTG was a gift from Marco Herold (Addgene, Plasmid #70183). Fgh1tUTB was generated in house from Fgh1tUTG by replacing the eGFP gene with a TagBFP gene from pU6-sgRNA EF1Alpha-puro-T2A-BFP (Addgene, Plasmid #60955). To clone a target sgRNA sequence into Fgh1tUT vector, two complementary oligos were synthesized in the form of: 5' –TCCCNNNNNNNNNNNNNNNNNNN-3′ and 3' –NNNNNNNNNNNNNNNNNNNCAAA– 5′ and annealed. All "NNNNNNNNNNNNNNNNNNN" sequences of different sgRNAs are shown in Table S1. The vector was digested by BsmBI restriction enzyme. sgRNA sequences are designed by using Benchling (benchling.com/academic) or selected from the list of sgRNA sequences in the Sabatini/Lander CRISPR pooled library (Addgene, Cat # 1000000095) as stated in Table S1.

Transient transfection

10 μg of plasmid DNA was added to 200 μL of Opti-MEM (Gibco) in a 15 mL tube. 30 μL of transfection reagent (Polysciences, Polyethylenimine) was diluted with 200 μL of Opti-MEM in another tube. The diluted transfection reagent was added to the DNA-containing Opti-MEM, mixed well and incubated at room temperature for at least 10 min. Medium was removed and cells was washed with 3 mL of Opti-MEM, then 5 mL of Opti-MEM was added to the complex with DNA and reagent, which was then transferred to the cells on 6-cm dish for incubation for 6 h at 37°C in a 5% CO_2 incubator. The medium with transfection complex was removed and replaced with fresh full medium. The transfected cells were harvested for analysis 24–48 h after transfection.

Production of lentivirus and infection

Lentiviral particles were made by transient transfection of 293 cells grown in 10-cm dishes at 70% confluence with 3 μg of lentivirus plasmid DNA, 3 μg of pMDLg/pRRE, 2 μg of pMD2.G and 2 μg of pRSV-Rev. Virus containing supernatants were collected at 48 h after transfection and passed through a 0.45 μm filter. The supernatants supplemented with 2 μg/mL polybrene were added to cultured cells, which were centrifuged at 500 g for 30 min. After 24 h of incubation, the infected cells were re-cultured in the fresh medium.

Flow cytometry analysis and FACS sorting

The FACS antibodies and lectins are listed with the detailed information in key resources table. Cells were suspended in 0.2 μg/mL 7-AAD (Caymanchem) prior to analysis to exclude dead cells. Flow cytometry analysis was performed with the Fortessa cell analyser (Becton Dickinson) and cell sorting was performed on the FACS Aria (Becton Dickinson). For the biotinylated lectins staining, the secondary Anti-Biotin-APC (Miltenyi Biotec) monoclonal antibody was applied. FACS data was analyzed by FlowJo software (Tree Star, OR, USA).

Generation of CRISPR-Cas9 knockout (KO) and double knockout (DKO) rescue cells

To knock out genes by CRISPR-Cas9 approach, we used a lentivirus-based inducible sgRNA system, which consists of (1) FuCas9Cherry vector (Marco Herold, WEHI, Addgene#70182) with constitutive expression of Cas9 and mCherry fluorescence marker, and (2) FgH1tUT vector with doxycycline-inducible sgRNA expression and constitutive expression of eGFP or TagBFP fluorescence marker as described above. To make the KO of one single gene, cells infected with pFu-Cas9-mCheery lentivirus were sorted on FACS Aria (Becton Dickinson) to establish the stable Cas9/mCherry expressing cells. The Cas9-mCherry stable cells were subsequently infected with lentivirus of doxycycline-inducible sgRNA (Fgh1tUTG vector) targeting protein coding exons of the desired gene. To induce the Tet-O promoter mediated sgRNA expression, cells were incubated with 1ug/mL doxycycline at 37°C with 5% CO_2 and humidified conditions for 2 days. After that, cells were washed with PBS and subsequently cultured in fresh complete medium without doxycycline for at least 5 days before sorting. Cells positive for GFP and mCherry were gated and FACS sorted to make stable single gene KO cells. To make DKO cells, single gene KO cells were infected with lentivirus of doxycycline-inducible sgRNA (Fgh1tUTB vector) targeting protein coding exons of the other desired gene, followed by a similar Dox-inducing and selecting strategy to sort for cells that are positive for GFP, BFP and mCherry. The expression of appropriate cell-surface markers was used as key readouts of gene knockout for establishing all the KO cell lines, in addition to the fluorescent protein markers as described in Table S4.

Genome-wide CRISPR/Cas9 knockout screens

The Human Two Plasmid Activity-Optimized CRISPR Knockout Library constructed by David Sabatini/Eric Lander and colleagues was obtained from Addgene (Cat# 1000000095).[38] We prepared the lentivirus of the library by virus packaging in HEK 293 cells.

We conducted pilot experiments to decide how much virus is required to achieve MOI = 0.3–0.5 (~70% and 50% cell death induced by puromycin respectively) as described before.[82] To carry out actual screens, cells were seeded and cultured in T150 flasks. About 300 M cells in total were incubated with medium containing lentivirus of the library and 2 μg/mL polybrene at an MOI = 0.3–0.5 for 2 days. After infection, the cells were selected with 10 μg/mL puromycin in fresh media for 5 days and then cultured in fresh medium without puromycin for a few more days before FACS sorting. 10–20% of total infected cells were reserved as the pool control sample and the remaining 80–90% of cells were used for FACS sorting based on cell-surface TSPAN8 expression. The genomic DNA were extracted from both the pool control and sorted cells per manufacturer's instructions (Bio Basic, One-4-All Genomic DNA Miniprep Kit). NGS libraries were prepared with specific barcode primers by PCR. sgRNA barcode PCR primers Forward: CAAGCAGAA GACGGCATACGAGATCnnnnnnTTTCTTGGGTAGTTTGCAGTTTT Reverse: AATGATACGGCGACCACCGAGATCTACACnnnnnnnn CACCGACTCGGTGCCACTTTT. "n" denotes the sample-specified barcode sequence. Multiplex NGS libraries were sequenced on HiSeq4000 to identify the sgRNAs in each sample. Normalization, sgRNA modeling and ranking were done on the MAGeCK algorithm as described previously.[83] Of note, more than 75% of the sgRNAs in the library were identified in the pool control sample for each screen. To identify enriched GO terms in the FACS sorted cells, the GOrilla tool was used to analyze the gene set with pScore <0.01 for each screen.[84,85]

Western blot analysis

The primary western blot antibodies are listed in key resources table. Cells were lysed in ice-chilled cell lysis buffer (150 mM NaCl, 50 mM Tris-HCl, pH 7.3, 0.25 mM EDTA, pH 8.0, 1% sodium deoxycholate, 1% Triton X-100, 0.2% sodium fluoride and 0.1% sodium orthovanadate) with protease and phosphatase inhibitor cocktail (Roche, Mannheim, Germany). The lysates were resolved on 10% SDS-PAGE and transferred onto a PVDF membrane. After blocking with 5% non-fat milk in wash buffer (TBS, 0.01% Triton X-100) at 37°C for 1 h, immunoblotting was done by incubation of the primary antibodies at 4°C overnight. Subsequently, the membrane was washed for three times with wash buffer and incubated with HRP-conjugated secondary antibodies for 1 h at room temperature before developing images (FUJI Medical X-ray Film, Ref: 4741019289) per manufacturer's instruction.

Enzymatic deglycosylation of live cells

For treatment of intact cells with PNGase F, cells were seeded and cultured for 24 h prior to the experiment. The medium was removed and the cells were detached by Trypsin (Life Technologies Corporation, Gibco Trypsin-EDTA, 0.25%, Cat# 25200-056) for 5 min and neutralized with complete medium thereafter. The cell suspension pelleted by centrifugation and washed once by PBS. The cells were re-suspended with serum free media and transferred to two Eppendorf tube with or without PNGase F (final concentration = 200 U/mL) and incubated at 37°C under 5% CO_2 for 2 h. Treated cells were then harvested for western blotting or FACS analysis.

Endocytosis assay

Cells were incubated with 1:40 (v/v) APC-conjugated anti-TSPAN8 antibody (Miltenyi Biotec, REAfinity, Clone REA443, Cat# 130-106-811) in culture medium for 30 min at 37°C CO_2 incubator in humidified atmosphere. Medium was aspirated and the cells were rinsed with cold PBS, and detached with 1x non-enzymatic cell dissociation solution (Merck, Cat No. C5914-100ML) on ice for 5 min. Cells from each well were resuspended and equally separated to two tubes, and washed with PBS (for the total level of APC-anti-TSPAN8 antibodies) or stripping buffer (150 mM NaCl, 0.2 M acetic acid, 5 mM KCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, adjusted to pH 2.5, for the level of internalized APC-anti-TSPAN8 antibodies) for 20 s respectively and repeated washing two more times. Cells in both tubes were washed again with cold PBS and resuspended in cold FACS analysis buffer for FACS analysis. FACS analysis of cells washed with PBS was to measure total level of stained TSPAN8 protein (shown as "unstripped" in figures); FACS analysis of cells washed with stripping buffer was to measure internalized level of stained TSPAN8 protein (shown as "stripped" in figures).

Effects of lactose treatment on endocytosis of N-glycoproteins

Cells were seeded onto two dishes with the culture medium and cultured in at 37°C tissue culture incubator with humidified atmosphere of 5% CO_2. One dish was rinsed with prewarmed 150 mM lactose (Cat No. L3750, Sigma) solution in complete culture medium once and treated with 150 mM lactose for 30 min (the fresh lactose solution was changed every 10mins) at 37°C CO_2 incubator; the other dish was rinsed with prewarmed 150 mM sucrose (Sigma, Cat# S0389) solution in complete culture medium once and treated with 150 mM sucrose for 30 min (the fresh sucrose solution was changed every 10 min) at 37°C CO_2 incubator. The dish of cells treated with lactose were then incubated with 1:40 (v/v) APC-conjugated anti-TSPAN8 antibody in the 150 mM lactose solution (prewarmed) for 15 min at 37°C CO_2 incubator; The cells in the other dish treated with sucrose were incubated with 1:40 (v/v) APC-conjugated anti-TSPAN8 antibody in the 150 mM sucrose solution (prewarmed) for 15 min at 37°C CO_2 incubator. Cells were then washed with ice-cold PBS and detached from dishes with 1x non-enzymatic cell dissociation solution (Merck, Cat #C5914-100ML) on ice for 5 min and resuspended in cold FACS analysis buffer for the analysis of internalized antibody. To measure the effect of lactose on the cell-surface expression level of TSPAN8, the cultured cells were treated with 150 mM lactose or sucrose solution for 30 min at 37°C 5% CO_2 incubator (The lactose or sucrose solution was replaced every 10 min) and immediately stained with APC-conjugated anti-TSPAN8 antibody on ice for 30 min.

Surface biotinylation assay

For surface biotinylation assay, cells were first incubated with EZ-Link Sulfo-NHS-SS-Biotin (Thermo Scientific), a thiol-cleavable amine-reactive biotinylation reagent, for 10 min at room temperature. The biotinylated cells were subsequently lysed with detergent and the supernatants with labeled proteins were then incubated and isolated with NeutrAvidin Agarose resin for 30 min at room temperature. The bound, labeled proteins were released and eluted by reduction of the bisulfide bond with 10 mM DTT and subsequently analyzed by western blotting.

Reverse transcription and qPCR analysis

Total RNA extraction was performed according to the manufacturer's protocol (Qiagen, RNeasy Micro Kit). mRNA was converted to cDNA with GoScriptTM Reverse Transcription Kit (Promega). cDNA was used for gene cloning by PCR amplification or for qPCR. qPCR was carried out by using SYBR Green qPCR Master Mix (Thermo Fisher Scientific) with forward and reverse primers. qPCR was performed using Bio-Rad System. Standard delta-delta Ct ($2^{-\Delta\Delta Ct}$) method was applied to assess mRNA expression relative to *β-actin*, which was used as the house-keeping control. *TSPAN8* primers used for qPCR, Forward: GCAGAGACCATGC CAAAGCTATAATG and Reverse: CGATCTGGCAATACAGGACCATAG. *B3GNT5* primers used for qPCR, Forward: TGGACACC TACTCTACGAAACACG and Reverse: CTCTTCTGCCACTAACCAACATTCTC. *β-actin* primers used for qPCR, Forward: TCCCTGG AGAAGAGCTACG and Reverse: GTAGTTTCGTGGATGCCACA.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data, unless specially stated, are shown as Mean ± standard error of the mean (SEM) and the student's t-test using GraphPad Prism (GraphPad Software) was used where applicable, with $p < 0.05$ considered statistically significant.