

SCIENTIFIC REPORTS

There are amendments to this paper

OPEN

Intact glycosphingolipidomic analysis of the cell membrane during differentiation yields extensive glycan and lipid changes

Maurice Wong , Gege Xu, Dayoung Park , Mariana Barboza & Carlito B. Lebrilla 

Glycosphingolipids (GSLs) are found in cellular membranes of most organisms and play important roles in cell-cell recognition, signaling, growth, and adhesion, among others. A method based on nanoflow high performance liquid chromatography-chip-quadrupole-time-of-flight mass spectrometry (nanoHPLC Chip-Q-TOF MS) was applied towards identifying and quantifying intact GSLs from a variety of samples, including cultured cell lines and animal tissue. The method provides the composition and sequence of the glycan, as well as variations in the ceramide portion of the GSL. It was used to profile the changes in the glycolipidome of Caco-2 cells as they undergo differentiation. A total of 226 unique GSLs were found among Caco-2 samples from five differentiation time-points. The method provided a comprehensive glycolipidomic profile of a cell during differentiation to yield the dynamic variation of intact GSL structures.

Glycosphingolipids (GSLs) are a group of complex lipids that are comprised of a ceramide (N-acylsphingosine) glycosidically linked to a glycan moiety. These molecules are ubiquitous in cell membranes, where they are known to participate in cellular processes such as signaling, adhesion, and cell differentiation, among other functions¹. Malignancies such as cancer and lysosomal storage diseases are also associated with marked changes in the glycolipidome of affected cells. Tumor cells express aberrant glycosylation in GSLs, displaying either incomplete synthesis leading to an accumulation of precursors, or further addition of glycan residues to form new structures². Furthermore, some cancers have been observed to shed elevated levels of GSLs with immunosuppressive activity, and the rate of shedding is affected by the length of the acyl chain in the ceramide^{3,4}. Thus there is value in determining the levels of intact GSLs and identifying both its glycan and lipid portions.

Caco-2, a human colorectal cancer cell line, undergoes enterocyte-like cellular differentiation *in vitro*. As such, Caco-2 has been used to model drug absorption in the human intestine and to study interactions between gut microbiota and intestinal epithelial cells^{5,6}. Previous studies have shown that intestinal epithelial cells have a varied array of glycolipids, including neutral, sulfated, sialylated, and fucosylated GSLs, as well as globo-type GSLs^{7,8}. Furthermore, these GSLs also express ABO blood group and Lewis epitopes⁹. Clinical studies on colorectal cancer tissue have found alterations in GSL glycosylation during cancer progression, such as increased fucosylation and a shift from type I (lacto-type) to type II (neolacto-type) oligosaccharides^{10,11}. Studies on the N-linked glycosylation of Caco-2 have revealed that undifferentiated cells express abundant high-mannose structures while differentiated cells have mostly complex and hybrid structures decorated with fucose and sialic acid¹². The expression of blood group antigens on the glycoproteins of differentiating Caco-2 has also been explored¹³. However, we do not yet have a clear picture of how the cell surface GSLs of Caco-2 change in relation to differentiation, which may yield insights into the link between cell surface glycosylation and cancer progression.

Despite its potential for drug and biomarker discovery, a comprehensive analysis of the cell surface glycosphingolipidome can often prove challenging due to the structural complexity and heterogeneity of GSL structures, as well as the amphipathic nature of these biomolecules. The first challenge arises from the extraction and purification of GSLs from biological matrices. The simplest and most common approach is through solvent extraction; the procedures described by Folch¹⁴ and Bligh and Dyer¹⁵ are now well-established and widely used. In both procedures, total lipids are first extracted from the sample, and then partitioned between a biphasic

Department of Chemistry, University of California, Davis, 1 Shields Ave., Davis, California, 95616, USA. Correspondence and requests for materials should be addressed to C.B.L. (email: cblebrilla@ucdavis.edu)

Received: 12 April 2018
Accepted: 5 July 2018
Published online: 20 July 2018

solvent system. Charged groups or those with at least four glycan residues mostly partition to the methanol-rich layer, while other less polar lipids remain in the chloroform-rich layer¹⁶. When collecting the methanol-rich layer, recoveries greater than 90% can be achieved for sialylated GSLs, as reported by Svennerholm¹⁷. However, neutral GSLs with less than four glycan residues are not effectively recovered¹⁶. In this report, we utilize a modified Folch procedure and focus mainly on the GSLs that are recoverable from the methanol-rich layer. Although the limited recovery of mono-, di-, and tri-hexosylceramides will cause their abundances to be underreported, we nevertheless include them in our analysis to follow the changes in their abundance relative to the other samples.

The formation of the glycan is a non-templated process, which can lead to glycan structures that vary not just in monosaccharide composition, but also in sequence, linkage, and branching¹⁸. The ceramide moiety can have between two to four hydroxyl groups and varying unsaturation and length in their fatty acyl chain¹⁸. Hence the second challenge is in addressing the heterogeneity of these complicated biomolecules. Early approaches made use of thin-layer chromatography (TLC) for purification and separation, and immunoassay and lectin-binding techniques for quantitative and qualitative analysis of specific glycan motifs^{19,20}. However, these approaches are applicable to only a few glycan epitopes for which there are known lectins, and they are unable to reveal the nature of the ceramide. More recent efforts make use of enzymatic digestion to cleave the lipid portion from the glycan headgroups before analyzing the portions separately, at the expense of information about how these two portions are connected^{21,22}.

Mass spectrometry, especially when coupled with high performance liquid chromatography, has emerged as the method of choice for characterizing large numbers of GSLs^{23,24}. There have been several examples of MS analysis of intact GSLs. Neutral GSLs from human erythrocytes and sialylated GSLs from granulocytes have been analyzed through nano-HPLC/ESI-Q-TOF by Kirsch *et al.*^{25,26}. Gangliosides from murine cell lines and human brain have also been examined extensively with the technique^{27,28}. Intact gangliosides from milk have been characterized with MALDI-FTICR MS and UHPLC-QqQ MS^{29,30}. MALDI-MS has also been performed on intact murine brain gangliosides and permethylated GSLs from differentiated human embryonic stem cells^{31,32}. MS profiling of intact gangliosides was performed by Zamfir *et al.* to identify and elucidate ganglioside structures from human astrocytoma, its surrounding tissue, and normal brain tissue using an Orbitrap MS resulting in approximately 80 unique compositions³³. A study of gangliosides from human serum using nano-HPLC-ESI Q-TOF MS also by Kirsch *et al.* profiled 33 gangliosides and yielded potential biomarkers for pancreatic cancer, illustrating the potential of profiling methods for biomarker discovery³⁴.

In this report, we employ a nanoflow high performance liquid chromatography chip-quadrupole-time-of-flight mass spectrometry (nanoHPLC Chip-Q-TOF MS) method to comprehensively identify and quantitate over 220 intact cell-surface GSLs of Caco-2 during differentiation. The method effectively covered 15 different glycan headgroups, including acidic GSLs with sulfation and sialylation, as well as fucosylated and non-fucosylated neutral GSLs. The glycans of each identified GSL were characterized by their composition and connectivity, while the ceramides were distinguished by their lipid chain length, number of hydroxyl groups, and degrees of unsaturation. Furthermore, we applied observed trends in GSL elution towards predicting the retention times of GSLs and assigning peaks with more confidence thereby increasing the coverage significantly. The study presents the most comprehensive analysis of cell membrane GSLs to date, and exceeds the number of identified GSLs in earlier reports by a factor of 2. It further yields highly reproducible and quantitative information allowing measurement of distinct changes during cellular progression.

Results and Discussion

Caco-2 cells undergo enterocyte-like differentiation upon reaching confluence. Although the cell line comes from colorectal cancer, differentiated Caco-2 cells acquire morphological and functional features that resemble intestinal epithelial cells. During differentiation, they exhibit changes such as bicellular tight junction assembly and remodel the apical surface to form a brush border⁵. As the cells reach confluence, they stop dividing exponentially and begin to differentiate, with cells forming villi and migrating to the tip as they mature³⁵. The cell surface glycosylation of proteins was previously reported to change during the differentiation process, based on the analysis of released N-glycans¹².

In this study, cell surface GSLs were monitored *in vitro* at five different time points: undifferentiated (Day 5), confluent (Day 7), partially differentiated (Day 14), differentiated (Day 21), and post-differentiated (Day 24). Analysis of MS data yielded a total of 227 unique GSLs in all the Caco-2 samples as characterized by a combination of accurate masses, retention times, and tandem MS. Peak areas of identified GSLs are normalized to the total abundance of GSLs in each sample and reported as relative abundances.

LC-MS/MS data analysis. Ganglioside standards GM1a and GD3 from Calbiochem and Avanti respectively were run with the samples to provide the retention times standards. The commercial compounds were found to contain mixtures of different ceramide chain lengths, and their retention times matched those identified from samples. The GM1a and GD3 standards were also run repeatedly during the batch analysis to determine instrument reproducibility over time. The relative standard deviations of GM1a and GD3 were calculated to be 19.2% and 16.5% respectively from 5 injections over 24 h of instrument run time. Serial dilutions of the standards from 10 µg/ml to 200 µg/ml were also injected to determine the linear range of the instrument response. Figure 1 shows the linear plots of the peak area for the GSLs identified from the standards as a function of concentration. The linear range encompasses the response of the reported GSLs, assuming that the response of the different GSLs are comparable.

GSLs were initially identified based on the monoisotopic masses of possible combinations of glycans and ceramides. However, identification of GSL structures from cells by MS data can be complicated by the presence of phospholipids and other lipid species, which cannot be completely separated through Folch extraction. These compounds can have masses that differ by only a few ppm with those of hexosylceramides, lactosylceramides,

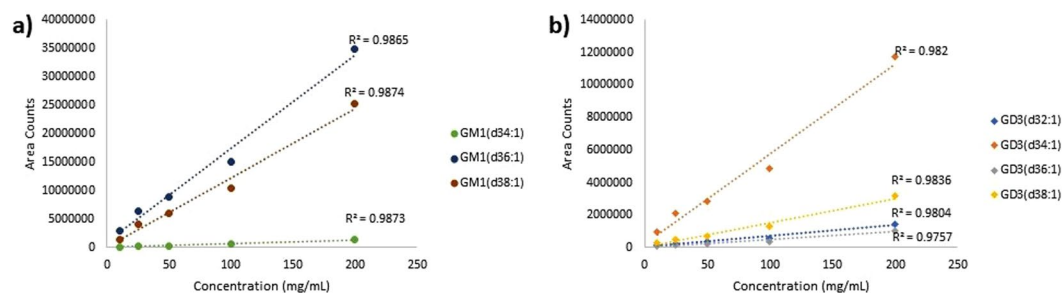


Figure 1. Linear regression plots for the area counts of GSL compounds from (a) GM1a and (b) GD3 standards as a function of concentration. The peak area shows generally linear behavior with concentration. The coefficient of determination for each identified GSL is shown.

and sulfatides. To avoid the misidentification of GSLs, their structures were verified with tandem MS data using collisional-induced dissociation (CID). As was shown previously by others, ions corresponding to combinations of ceramide fragments and glycan residues such as hexose, fucose, N-acetylhexosamine, and N-acetylneuraminic acid were featured in the CID spectra³⁶. Ceramides, which are composed of a sphingoid base and a fatty acid, were found to have mostly 2 to 3 hydroxyl groups, 32 to 44 total carbon atoms, and 0 to 2 hydrocarbon unsaturation based on this analysis. Sphingosine, an 18-carbon aliphatic amine with an unsaturated hydrocarbon chain and 2 hydroxyl groups, is the most common sphingoid base found in mammals, although variations in chain length, unsaturation, and the number of hydroxyl groups are also present¹⁸. Fragment ions for the sphingoid moieties *d18:1*, *d18:0*, and *d20:1* were found at *m/z* 264.3, 266.3, and 292.3 respectively. Annotated tandem mass spectra of two GSLs in Caco-2 are shown in Fig. 2. The fragmentation pattern in Fig. 2a suggests a tetraosylceramide that has a globo-type structure, Gb₄(d18:1/C16:0). Compound 2b has the glycan composition Hex₃HexNAc₁Fuc₁, which can possibly have the Lewis a/x or the H antigen (Lewis x structure is shown). It should be noted that the linkages and absolute configuration of the glycan residues cannot be determined from the mass spectrum alone, and that the proposed structures presented here are putative and based on current knowledge of human intestinal cell surface GSLs⁷.

With our data-dependent acquisition program, we obtained tandem MS spectra for 54% of the reported GSLs. To identify GSLs with greater confidence, especially those that were too low in abundance to obtain tandem MS, we predicted the retention times of GSLs based on observed elution trends of GSLs identified from tandem MS data. These trends were also observed in GM1a and GD3 ganglioside standards which contained a mixture of different lipid chain lengths. The prediction considers the characteristics of both the glycan and the lipid. For GSLs with the same glycan compositions that differ only by the lengths of their hydrocarbon chains in the ceramide, a linear trend was found between the retention times and the number of carbons in the ceramide, as shown in Fig. 3. Lengthening the ceramide by 1 carbon delayed elution by about 1.6 minutes. Those with trihydroxy ceramides eluted 0.6–1 min earlier than comparable dihydroxy ceramides. GSLs with a charged group, such as a carboxylic acid or sulfate moiety, eluted about 1–1.5 mins earlier than a comparable GSL with the same ceramide and the same number of monosaccharide residues.

A library of GSLs that includes their monoisotopic masses, chemical formulae, and estimated retention times was created for rapid identification and semi-quantitation. The library combines identified glycan head groups with other possible configurations of ceramide to have a more comprehensive list of GSLs. Signals that correspond to the GSLs in the library were extracted by the MassHunter Qualitative Analysis software through its targeted Find by Formula function, and quantitated by their peak area. Peaks were extracted and quantitated with a mass tolerance of 10 ppm and a RT tolerance of 0.5 min. The program also accounts for the isotopic distribution of the compounds, their charge states, and the presence of other adducts for peak quantitation. Charge states of 1–2 were included in peak extraction, and both proton and ammonium adducts were considered as well. For data analysis, abundances are normalized to the total abundance of identified GSLs in each sample. Relative abundances are reported for each GSL. Representative extracted compound chromatograms for each of the five differentiation time points are shown in Fig. 4, annotated with schematic representations of the more abundant GSLs. A complete list of the GSL compositions, retention times, and abundances are provided in Supplementary Table 1.

Glycan profile changes. The lipid-bound glycan profile is not only tissue-specific, but is also affected by the condition and development of the cell¹⁸. There is a surprisingly small number of glycan compositions that comprise the glycolipidome. A total of 15 glycan compositions were identified and quantitated. A summary of the glycan structures and their abundances relative to the total abundance of all GSLs found in each sample is shown in Fig. 5. GSLs with isomeric glycan units, such as linkage and epimeric isomers, do not resolve sufficiently on the C18 column, and some representations will likely correspond to several isomeric structures. The putative structures of the glycans shown here are believed to be the major contributors based on previous studies on the GSLs in human intestinal tissue^{7,10}.

GSLs were further grouped according to glycan subtypes to have a more general view of how the cell surface GSLs change during differentiation. The general trends for each subtype are shown in Fig. 6. The subtypes presented here are mostly based on how glycans are decorated with specific monosaccharide moieties, such as sialic acid and fucose, which are frequent features in glycan-based antigen epitopes.

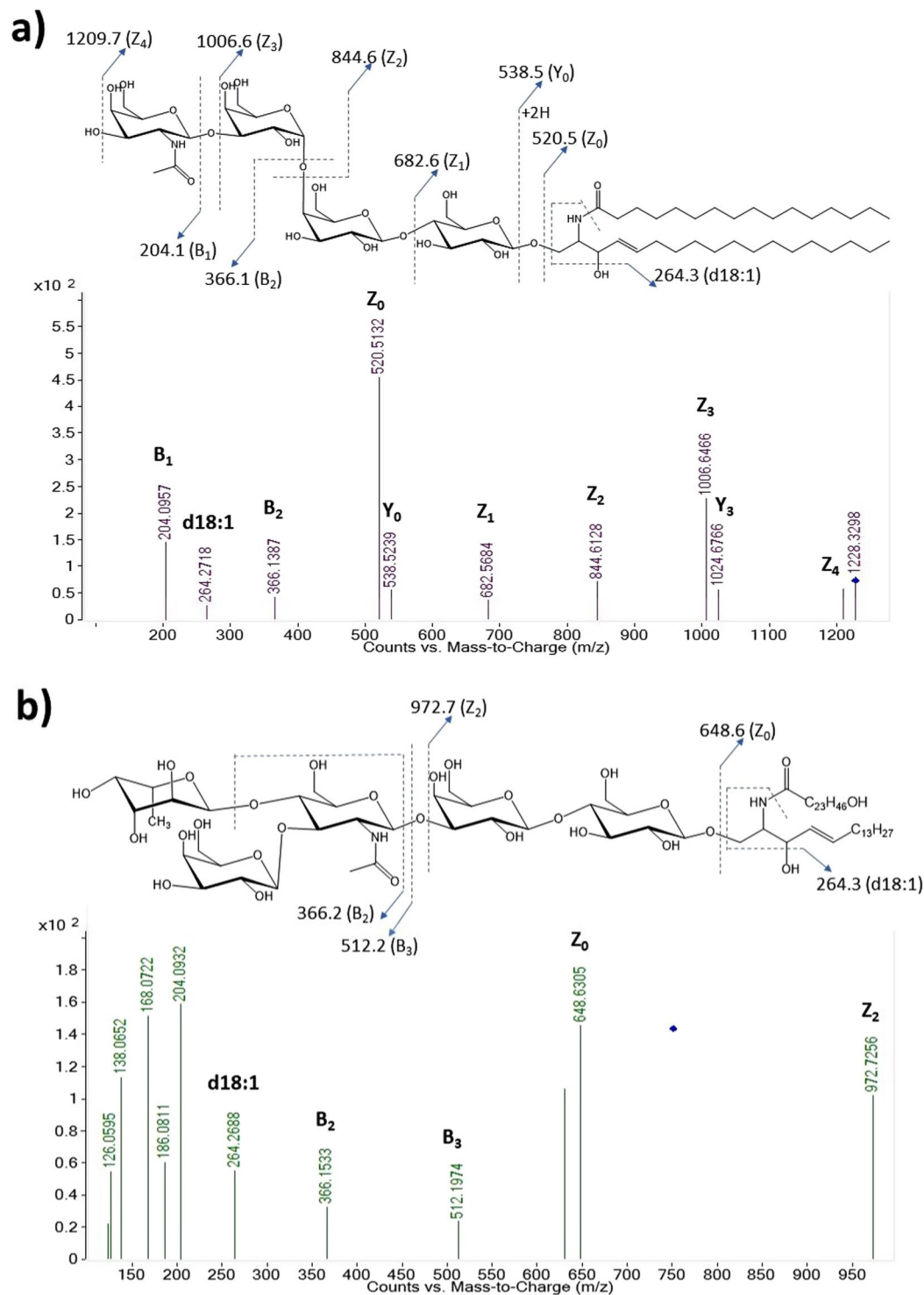


Figure 2. Tandem MS spectra of intact GSLs (a) $Gb_4(d18:1/C16:0)$, and (b) $Fuc-Lc_4(d18:1/C24:0\ OH)$ with the fragmentation pattern annotated for relevant peaks. Putative structures of the compound are also presented with designations for product ions.

Fucosylated structures initially comprised 11% of GSLs in undifferentiated cells. This number fell to 3.2% during confluence, then climbed back up to 12.8% in post-differentiated cells. These structures may express Lewis a/x epitopes, both of which have been found in colon cancer³⁷. Sialylated GSLs steadily increased throughout the differentiation timeline. However, the gangliosides GM3 and GM2 followed a decreasing trend while GM1 increased from confluence until differentiation. GM3 at confluence accounted for 5.9% of total detected GSLs and tapered off to 4.3% post-differentiation, while GM1 increased from 2.3% to 7.6% over the same period. Caco-2 has been reported to express mostly GM3 on its cell surface. The expression of GM1 may be due to a passage-dependent cell transformation

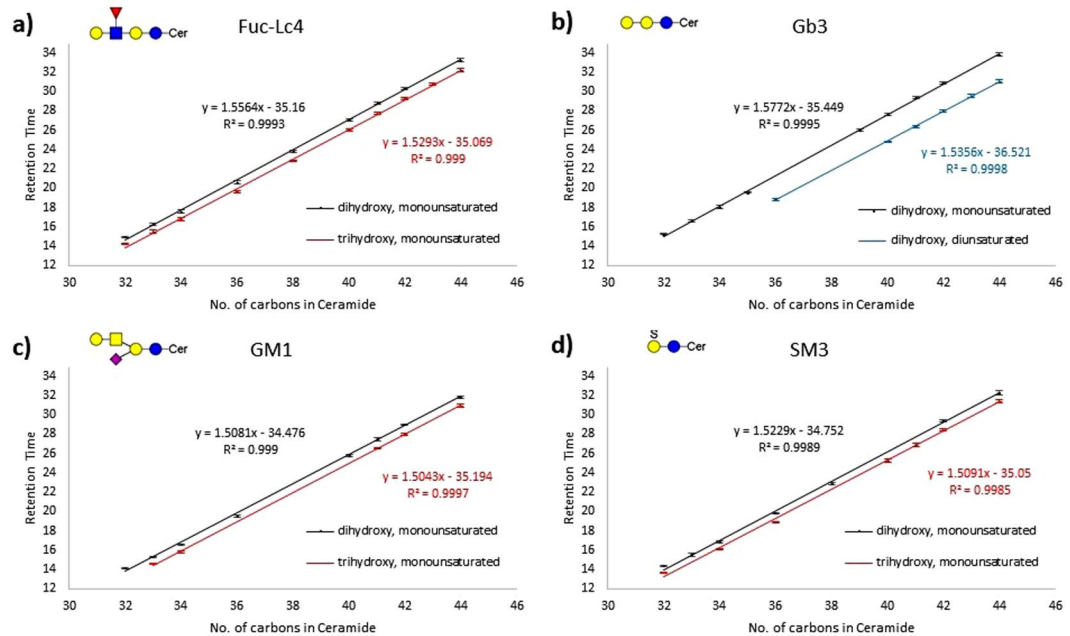


Figure 3. Graphs of retention time (mins) vs the number of carbons in the ceramide for 4 different glycan headgroups. Putative structures are shown for each glycan headgroup. The linear plots of di- and tri- hydroxy monounsaturated ceramides are compared in fuc-Lc₄ (a), GM1 (c), and SM3 (d), while mono- and di-unsaturated dihydroxy ceramides are compared in Gb₃ (b). These relationships allowed the identification of low abundance GSL with accurate masses, but where tandem MS could not be obtained. This predictive method was used to increase the number of identified compounds.

that promotes the epithelial to mesenchymal transition (EMT) of the cultured cells³⁸. This concomitant but undesirable transformation can further complicate reproducibility studies between passages as it can drastically change the lipid-bound glycosylation pattern. GSLs that were decorated with both fucose and sialic acid also showed a steady increase from 0.12% (undifferentiated) to 1.32% (post-differentiated). The fragmentation spectra suggest that the GSLs with Hex₃HexNAc₁Fuc₁NeuAc₁ (labeled Fuc-Sia-Lc₄ in Fig. 5) mostly express sialyl-Lewis a/x epitopes.

Globosides were observed to predominate as the Caco-2 cells reached confluence, accounting for more than 50% relative abundance on day 7 then decreasing to 11% in post-differentiated cells. Both Gb₃ and Gb₄ followed the same trend, being at their most abundant during confluence. The marked increase in globosides at confluence depresses the relative abundance of other decorated subtypes, as shown in Fig. 6, after which there is a steady increase in GSLs decorated with fucose or sialic acid. The trend is analogous to the switching of GSL core structures in human embryonic stem cells from globo-type to ganglio-type GSLs upon differentiation³¹. The globoside Gb₃, which is among the most abundant glycans found in the cells, has been associated with increased metastasis in colon cancer³⁹.

Sulfatides are sulfated GSL structures with one or two hexose units. Sulfatides play an important role in cell adhesion, and their elevated expression has been linked with metastatic potential of colorectal cancer⁴⁰. Two sulfatides were monitored: monohexosyl sulfatide, SM4, and dihexosyl sulfatide, SM3. Around 77% of total sulfatide had SM4 structure at all time points except at confluence, when SM4 accounted for 89%. Overall, the subtype decreased from 23.0% to 9.1% upon confluence, then climbed steadily to 36.7% post-differentiation.

Neutral GSLs include monohexosylceramides (HexCer), lactosylceramides (LacCer), and elongated poly-lactosamines. Folch extraction of HexCer (both glucosyl- and galactosyl- ceramides) and LacCer have low recoveries because of their more nonpolar nature¹⁶. However, here the changes to their relative amounts can be compared across the differentiation timeline because the solvent ratios were consistent throughout this analysis batch. LacCer is the common precursor for more complex GSLs. Both HexCer and LacCer were at their most abundant during partial differentiation, and declined as GSLs became increasingly decorated during differentiation. Poly-lactosamines were not consistently found in all samples because of their low abundance.

Ceramide profile changes. Variations in composition of the ceramide portion, such as the number of hydroxyl groups, the degree of unsaturation, and the chain lengths of the sphingoid and fatty acid, were also observed. The ceramides mostly expressed the d18:1 sphingoid moiety, with low levels of d20:1, and fatty acids with 14 to 24 carbons. Ceramides with 34 and 42 total carbons were the most abundant among the GSLs, as shown in Fig. 7; the ceramides are mostly comprised of the sphingoid d18:1 that are N-acylated with C16:0, C24:0, or C24:1 fatty acids. Ceramides with C16:0 acylation increased from 30% in undifferentiated cells to around 60% in fully differentiated cells, while those with C24:0 or C24:1 decreased from 43% to 25% over the same period. Ceramides with an odd number of carbons were also observed at low abundances. Trihydroxyceramides were also abundant, in which the fatty acid has an additional α -hydroxyl group aside from the two found in sphingosine¹⁸.

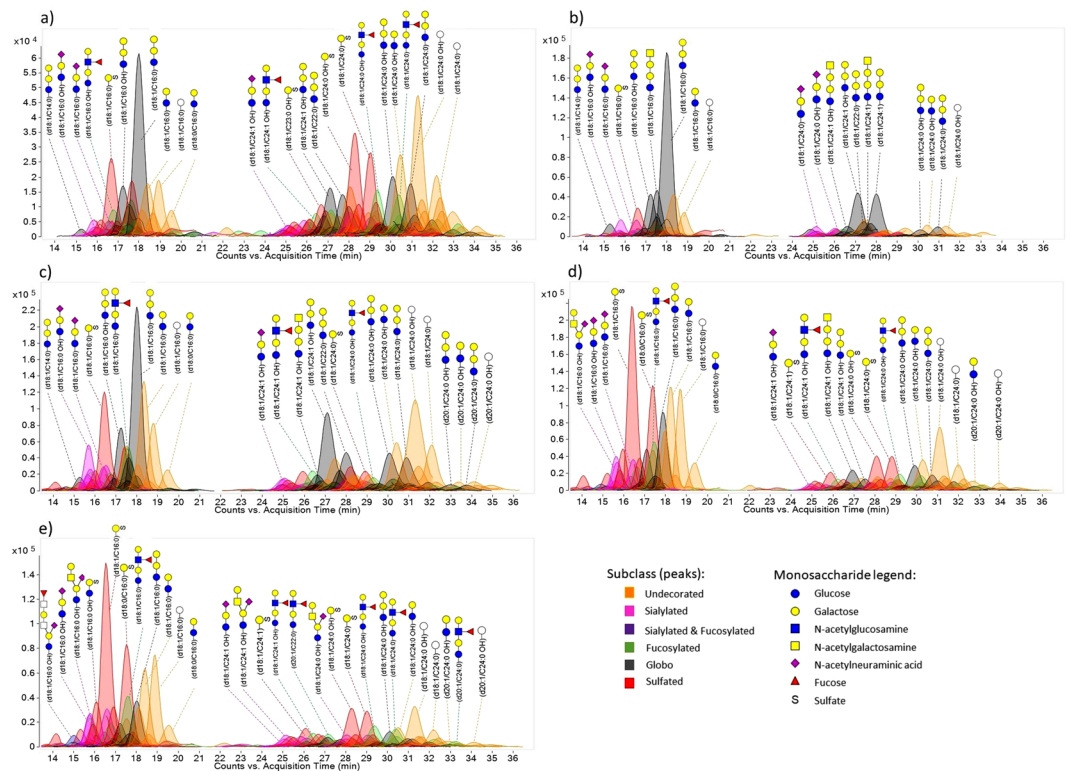


Figure 4. Representative chromatograms of GSLs in (a) undifferentiated (Day 5), (b) confluent (Day 7), (c) partially differentiated (Day 14), (d) differentiated (Day 21), and (e) post-differentiated (Day 24) Caco-2 cells. A different peak color is assigned to each glycan subtype: Sialylated – pink, Fucosylated – green, Fucosylated & Sialylated – blue, Globoside – gray, Sulfated – red, Undecorated – orange. Peaks of selected GSLs are annotated with schematic representations of their glycan structures while the ceramide structures are indicated above the glycan. The y-axis of each graph is adjusted to the most abundant compound in each time point. Structures were drawn with GlycoWorkbench⁴².

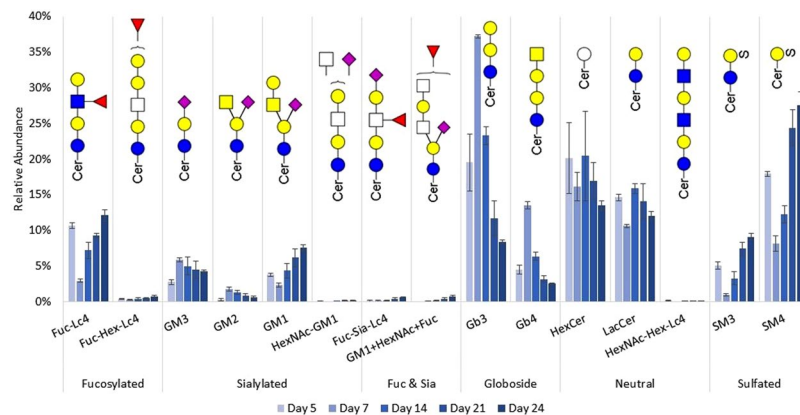


Figure 5. Relative abundances of GSL glycan subclasses. Error bars indicate the standard deviation for 3 biological replicates at each differentiation time point.

No clear trends were found between the relative abundances of dihydroxy- and trihydroxy-ceramides. On average, dihydroxyceramides accounted for 55.9% (± 4.7) of all GSLs. Low levels (<1%) of tetrahydroxyceramides were also found in some samples.

The changes to the attached lipid for each glycan head group can be tracked as well. The length of the acyl chain is correlated with the rate of shedding and GSL turnover, with shorter acyl chains exhibiting more rapid turnovers than longer ones⁴¹. Figure 8 shows a heatmap of the average number of carbons in ceramide associated with the different glycan headgroups. In general, the ceramides became shorter until full differentiation on Day 21, then lengthened again as they approached post-differentiation. This suggests that GSL turnover was fastest when the cells are fully differentiated.

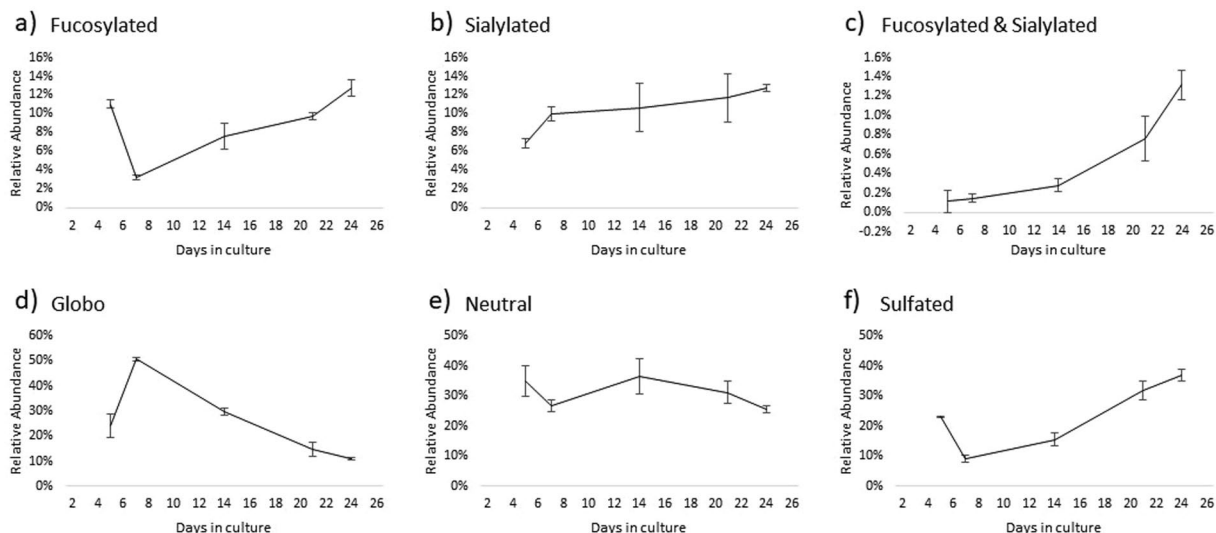


Figure 6. Relative abundances of GSLs based on glycan subtype as a function of growth time. Error bars indicate the standard deviation for 3 biological replicates.

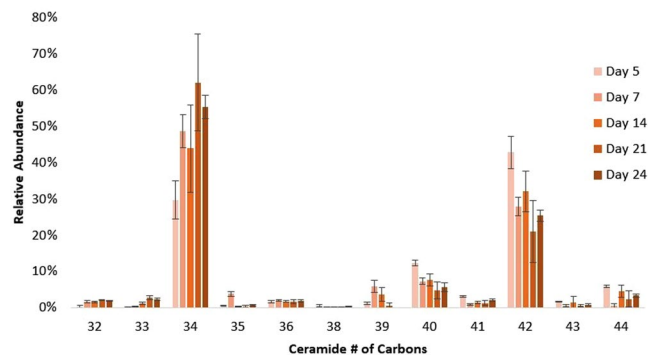


Figure 7. Relative abundances of GSLs according to the total number of carbons in the ceramide. Error bars indicate the standard deviation for 3 biological replicates.

Materials and Methods

Cell culture. Caco-2 cells were obtained from American Type Culture Collection and cultured in Eagle's Minimum Essential Medium, supplemented with 10% (v/v) fetal bovine serum, 100 U/ml penicillin, and 100 ug/ml streptomycin. The medium was replaced every 2–3 days, and cells were subcultured on reaching around 80% confluence using 0.05% trypsin (Gibco). Cells were grown at 37 °C in a humidified incubator with 5% CO₂. Cells were collected in biological triplicates by scraping after 5, 7, 14, 21, and 24 days, corresponding to undifferentiated, confluent, partially differentiated, differentiated, and post-differentiated cells respectively. All cells were harvested at passage 5. Cell counts of at least 3×10^5 were used for each replicate.

Membrane extraction. Harvested cells were suspended in homogenization buffer containing 1:100 protease inhibitor cocktail set V, 0.25 M sucrose, and 20 mM HEPES, with the pH adjusted to 7.5 using KOH. The cells were lysed with a probe sonicator. Nuclear fractions were pelleted and discarded through centrifugation at $2,000 \times g$ for 10 min. Membrane fractions were then pelleted and separated from other subcellular components through a series of ultracentrifugation steps at $200,000 \times g$ for 45 min.

Folch extraction. Pelleted membrane fractions were resuspended in 150 ul water. Methanol and chloroform were then added stepwise to form a mixture of 3:8:4 (v/v/v) water/methanol/chloroform. Samples were homogenized with sonication for 10 min, and centrifuged at $20,000 \times g$ for 1 min. Pellets, which contain mostly precipitated proteins, can be set aside for other analytical procedures, while supernatant portions were collected for GSL analysis. To the supernatant, 100 ul of 0.1 M potassium chloride was added to cause phase separation between the water- and methanol-rich upper layer, which contains mostly GSLs, and the chloroform-rich lower layer, which contains mostly other lipids such as cholesterol and phospholipids. The upper Folch layer was collected and dried under vacuum.

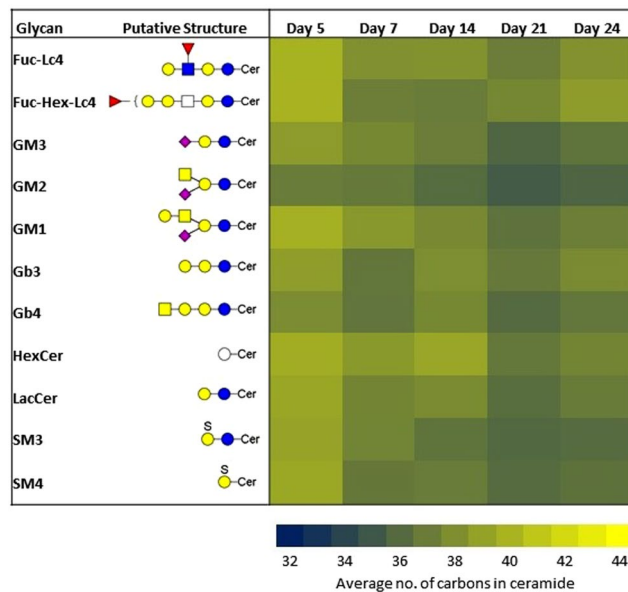


Figure 8. Heat map of the average length of ceramide lipids associated with different glycan headgroups during differentiation.

Solid phase extraction. Samples were further purified using C8 solid phase extraction (SPE) plate. SPE columns were rinsed with 1:1 methanol/water and 1:1 methanol/isopropanol, then conditioned with 1:1 methanol/water. Samples were loaded and washed with 1:1 methanol/water to remove salts. Elution was done with 1:1 methanol/isopropanol, after which the eluted sample was dried.

GSL standards. Ganglioside standards GM1a (Cat. no. 345724-1MG, Lot no. D00167174) from bovine brain and GD3 (Cat. No. 860060P, Lot. No. GD3-10) from bovine milk were obtained from Calbiochem and Avanti Polar Lipids, respectively. Standards were diluted to the desired concentrations with 1:1 methanol/water and injected into the LC-MS instrument without further processing.

NanoHPLC Chip-Q-TOF MS analysis. Samples were reconstituted in 50 μ l solution of 1:1 methanol/water and analyzed with an Agilent 6520 Accurate Mass Q-TOF LC/MS equipped with a C18 microfluidic chip, which incorporates an enrichment column, an analytical column, and a nanoelectrospray tip in a single assembly. A binary gradient consisting of (A) 20 mM ammonium acetate and 0.1% acetic acid in water, and (B) 20 mM ammonium acetate and 0.1% acetic acid in 85:15 (v/v) methanol/isopropanol was used to separate the GSLs at a flow rate of 0.3 μ l/min. The acquisition method was programmed to linearly increase the percentage of elution solvent B from 80% to 100% over 36 min. One MS and six tandem MS spectra were acquired every 3.1 s on positive mode through data-dependent acquisition. Fragmentation was achieved by collision-induced dissociation (CID) with nitrogen gas at collision energies of around 25 V. Data was processed with Agilent MassHunter B.07 software. GSL structures that were identified through their m/z and MS/MS fragmentation pattern were added to a library, which includes their chemical formulae, retention times (RT), and accurate masses. Compounds which match expected GSL masses, but for which there is no MS/MS data, were included when RT can be inferred from confirmed GSL compounds. Configurational isomers were disambiguated based on expected RT and MS/MS data. Compounds were then extracted using the “Find by Formula” program, allowing for 10 ppm mass tolerance and 0.5 min RT tolerance. Charge states of 1–2 and both proton and ammonium adducts were included in the extraction.

Accession codes. Caco-2: <https://www.atcc.org/Products/All/HTB-37.aspx>.

References

- Hakomori, S. Structure, organization, and function of glycosphingolipids in membrane. *Current Opinion in Hematology* **10**, 16–24 (2003).
- Hakomori, S.-i. Aberrant Glycosylation in Cancer Cell Membranes as Focused on Glycolipids: Overview and Perspectives. *Cancer Research* **45**, 2405–2414 (1985).
- Chang, F., Li, R. & Ladisch, S. Shedding of Gangliosides by Human Medulloblastoma Cells. *Experimental Cell Research* **234**, 341–346, <https://doi.org/10.1006/excr.1997.3619> (1997).
- Li, R. & Ladisch, S. Shedding of human neuroblastoma gangliosides. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* **1083**, 57–64, [https://doi.org/10.1016/0005-2760\(91\)90124-Z](https://doi.org/10.1016/0005-2760(91)90124-Z) (1991).
- Hidalgo, I. J., Raub, T. J. & Borchardt, R. T. Characterization of the human colon carcinoma cell line (Caco-2) as a model system for intestinal epithelial permeability. *Gastroenterology* **96**, 736–749 (1989).
- Zhang, L., Li, N., Caicedo, R. & Neu, J. Alive and Dead Lactobacillus rhamnosus GG Decrease Tumor Necrosis Factor- α -Induced Interleukin-8 Production in Caco-2 Cells. *The Journal of Nutrition* **135**, 1752–1756 (2005).

7. Breimer, M. E., Hansson, G. C., Karlsson, K.-A., Larson, G. & Leffler, H. Glycosphingolipid composition of epithelial cells isolated along the villus axis of small intestine of a single human individual. *Glycobiology* **22**, 1721–1730, <https://doi.org/10.1093/glycob/cws115> (2012).
8. Schnabl, K. L., Field, C. & Clandinin, M. T. Ganglioside composition of differentiated Caco-2 cells resembles human colostrum and neonatal rat intestine. *British Journal of Nutrition* **101**, 694–700, <https://doi.org/10.1017/S0007114508048289> (2008).
9. Holgersson, J., Jovall, P., Aring, K. & Breimer, M. E. Glycosphingolipids of Human Large Intestine: Detailed Structural Characterization with Special Reference to Blood Group Compounds and Bacterial Receptor Structures. *The Journal of Biochemistry* **110**, 120–131 (1991).
10. Holst, S. *et al.* Investigations on Aberrant Glycosylation of Glycosphingolipids in Colorectal Cancer Tissues Using Liquid Chromatography and Matrix-Assisted Laser Desorption Time-of-Flight Mass Spectrometry (MALDI-TOF-MS). *Molecular & Cellular Proteomics* **12**, 3081–3093, <https://doi.org/10.1074/mcp.M113.030387> (2013).
11. Misonou, Y. *et al.* Comprehensive clinico-glycomic study of 16 colorectal cancer specimens: elucidation of aberrant glycosylation and its mechanistic causes in colorectal cancer cells. *Journal of proteome research* **8**, 2990–3005 (2009).
12. Park, D. *et al.* Characteristic Changes in Cell Surface Glycosylation Accompany Intestinal Epithelial Cell (IEC) Differentiation: High Mannose Structures Dominate the Cell Surface Glycome of Undifferentiated Enterocytes. *Molecular & cellular proteomics: MCP* **14**, 2910–2921, <https://doi.org/10.1074/mcp.M115.053983> (2015).
13. Amano, J. & Oshima, M. Expression of the H type 1 blood group antigen during enterocytic differentiation of Caco-2 cells. *Journal of Biological Chemistry* **274**, 21209–21216 (1999).
14. Folch, J., Lees, M. & Sloane-Stanley, G. A simple method for the isolation and purification of total lipids from animal tissues. *The Journal of biological chemistry* **226**, 497–509 (1957).
15. Bligh, E. G. & Dyer, W. J. A Rapid Method of Total Lipid Extraction and Purification. *Canadian Journal of Biochemistry and Physiology* **37**, 911–917, <https://doi.org/10.1139/o59-099> (1959).
16. Smith, D. F. & Prieto, P. A. Special considerations for glycolipids and their purification. *Current protocols in molecular biology* **22**, 17.13. 11–17.13. 13 (1993).
17. Svennerholm, L. & Fredman, P. A procedure for the quantitative isolation of brain gangliosides. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* **617**, 97–109, [https://doi.org/10.1016/0005-2760\(80\)90227-1](https://doi.org/10.1016/0005-2760(80)90227-1) (1980).
18. Merrill, A. H. Sphingolipid and Glycosphingolipid Metabolic Pathways in the Era of Sphingolipidomics. *Chemical Reviews* **111**, 6387–6422, <https://doi.org/10.1021/cr2002917> (2011).
19. Higashi, H. *et al.* Sensitive Enzyme-Immunostaining and Densitometric Determination on Thin-Layer Chromatography of N-Glycolylneuraminic Acid-Containing Glycosphingolipids, Hanganutzii-Deicher Antigens1. *The Journal of Biochemistry* **95**, 1517–1520, <https://doi.org/10.1093/oxfordjournals.jbchem.a134760> (1984).
20. Cummings, R. D. [6] Use of lectins in analysis of glycoconjugates. *Methods in Enzymology* **230**, 66–86, [https://doi.org/10.1016/0076-6879\(94\)30008-9](https://doi.org/10.1016/0076-6879(94)30008-9) (1994).
21. Albrecht, S. *et al.* Comprehensive Profiling of Glycosphingolipid Glycans Using a Novel Broad Specificity Endoglycosidase in a High-Throughput Workflow. *Analytical Chemistry* **88**, 4795–4802, <https://doi.org/10.1021/acs.analchem.6b00259> (2016).
22. Fujitani, N. *et al.* Total cellular glycomics allows characterizing cells and streamlining the discovery process for cellular biomarkers. *Proceedings of the National Academy of Sciences* **110**, 2105–2110, <https://doi.org/10.1073/pnas.1214233110> (2013).
23. Merrill, A. H. *et al.* Sphingolipidomics: a valuable tool for understanding the roles of sphingolipids in biology and disease. *Journal of Lipid Research* **50**, S97–S102 (2009).
24. Kirsch, S. & Bindila, L. Nano-LC and HPLC-chip–ESI–MS: an emerging technique for glycobioanalysis. *Bioanalysis* **1**, 1307–1327, <https://doi.org/10.4155/bio.09.110> (2009).
25. Kirsch, S. *et al.* On-Line Nano-HPLC/ESI QTOF MS and Tandem MS for Separation, Detection, and Structural Elucidation of Human Erythrocytes Neutral Glycosphingolipid Mixture. *Analytical Chemistry* **80**, 4711–4722, <https://doi.org/10.1021/ac702175f> (2008).
26. Kirsch, S., Müthing, J., Peter-Katalinić, J. & Bindila, L. In *Biological Chemistry* Vol. 390, 657 (2009).
27. Zarei, M., Müthing, J., Peter-Katalinić, J. & Bindila, L. Separation and identification of GM1b pathway Neu5Ac- and Neu5Gc gangliosides by on-line nanoHPLC-QTOF MS and tandem MS: toward glycolipidomics screening of animal cell lines. *Glycobiology* **20**, 118–126, <https://doi.org/10.1093/glycob/cwp154> (2010).
28. Vukelić, Ž., Zarei, M., Peter-Katalinić, J. & Zamfir, A. D. Analysis of human hippocampus gangliosides by fully-automated chip-based nano-electrospray tandem mass spectrometry. *Journal of Chromatography A* **1130**, 238–245, <https://doi.org/10.1016/j.chroma.2006.05.033> (2006).
29. Lee, H. *et al.* Rapid profiling of bovine and human milk gangliosides by matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry. *International Journal of Mass Spectrometry* **305**, 138–150, <https://doi.org/10.1016/j.ijms.2010.10.020> (2011).
30. Lee, H., German, J. B., Kjelden, R., Lebrilla, C. B. & Barile, D. Quantitative Analysis of Gangliosides in Bovine Milk and Colostrum-Based Dairy Products by Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectrometry. *Journal of Agricultural and Food Chemistry* **61**, 9689–9696, <https://doi.org/10.1021/jf402255g> (2013).
31. Liang, Y.-J. *et al.* Switching of the core structures of glycosphingolipids from globo- and lacto- to ganglio-series upon human embryonic stem cell differentiation. *Proceedings of the National Academy of Sciences* **107**, 22564–22569, <https://doi.org/10.1073/pnas.1007290108> (2010).
32. Zarei, M. *et al.* A sialylation study of mouse brain gangliosides by MALDI a-TOF and o-TOF mass spectrometry. *Journal of Mass Spectrometry* **43**, 716–725, <https://doi.org/10.1002/jms.1367> (2008).
33. Zamfir, A. D. *et al.* Profiling and sequence analysis of gangliosides in human astrocytoma by high-resolution mass spectrometry. *Analytical and Bioanalytical Chemistry* **405**, 7321–7335, <https://doi.org/10.1007/s00216-013-7173-x> (2013).
34. Stephan Kirsch, J. S. M. M. L. B. & JasnaPeter, K. Ceramide Profiles of Human Serum Gangliosides GM2 and GD1a exhibit Cancer-associated Alterations. *Journal of Glycomics & Lipidomics* **2**, 1–10, <https://doi.org/10.4172/2153-0637.S2-005> (2012).
35. Rousset, M. The human colon carcinoma cell lines HT-29 and Caco-2: two *in vitro* models for the study of intestinal differentiation. *Biochimie* **68**, 1035–1040 (1986).
36. Costello, C. E. & Vath, J. E. In *Methods in Enzymology* Vol. 193, 738–768 (Academic Press, 1990).
37. Korekane, H. *et al.* Novel fucogangliosides found in human colon adenocarcinoma tissues by means of glycomic analysis. *Analytical Biochemistry* **364**, 37–50, <https://doi.org/10.1016/j.ab.2007.01.034> (2007).
38. Jahn, K. A., Biazik, J. M. & Braet, F. GM1 expression in caco-2 cells: Characterisation of a fundamental passage-dependent transformation of a cell line. *Journal of Pharmaceutical Sciences* **100**, 3751–3762, <https://doi.org/10.1002/jps.22418> (2011).
39. Kovbasnjuk, O. *et al.* The glycosphingolipid globotriaosylceramide in the metastatic transformation of colon cancer. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 19087–19092, <https://doi.org/10.1073/pnas.0506474102> (2005).
40. Morichika, H., Hamanaka, Y., Tai, T. & Ishizuka, I. Sulfatides as a predictive factor of lymph node metastasis in patients with colorectal adenocarcinoma. *Cancer* **78**, 43–47 (1996).
41. Skotland, T. *et al.* Determining the Turnover of Glycosphingolipid Species by Stable-Isotope Tracer Lipidomics. *Journal of Molecular Biology* **428**, 4856–4866, <https://doi.org/10.1016/j.jmb.2016.06.013> (2016).
42. Ceroni, A. *et al.* GlycoWorkbench: A Tool for the Computer-Assisted Annotation of Mass Spectra of Glycans. *Journal of Proteome Research* **7**, 1650–1659, <https://doi.org/10.1021/pr7008252> (2008).

Acknowledgements

Research reported in this publication was supported by the National Institutes of General Medicine of the National Institutes of Health under award number (2R01GM049077).

Author Contributions

M.W., D.P., M.B., and C.L. designed the project. M.W. and G.X. cultured the cells. M.W. performed the experiments, analysed the results, and wrote the manuscript in consultation with C.L. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-29324-7>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018