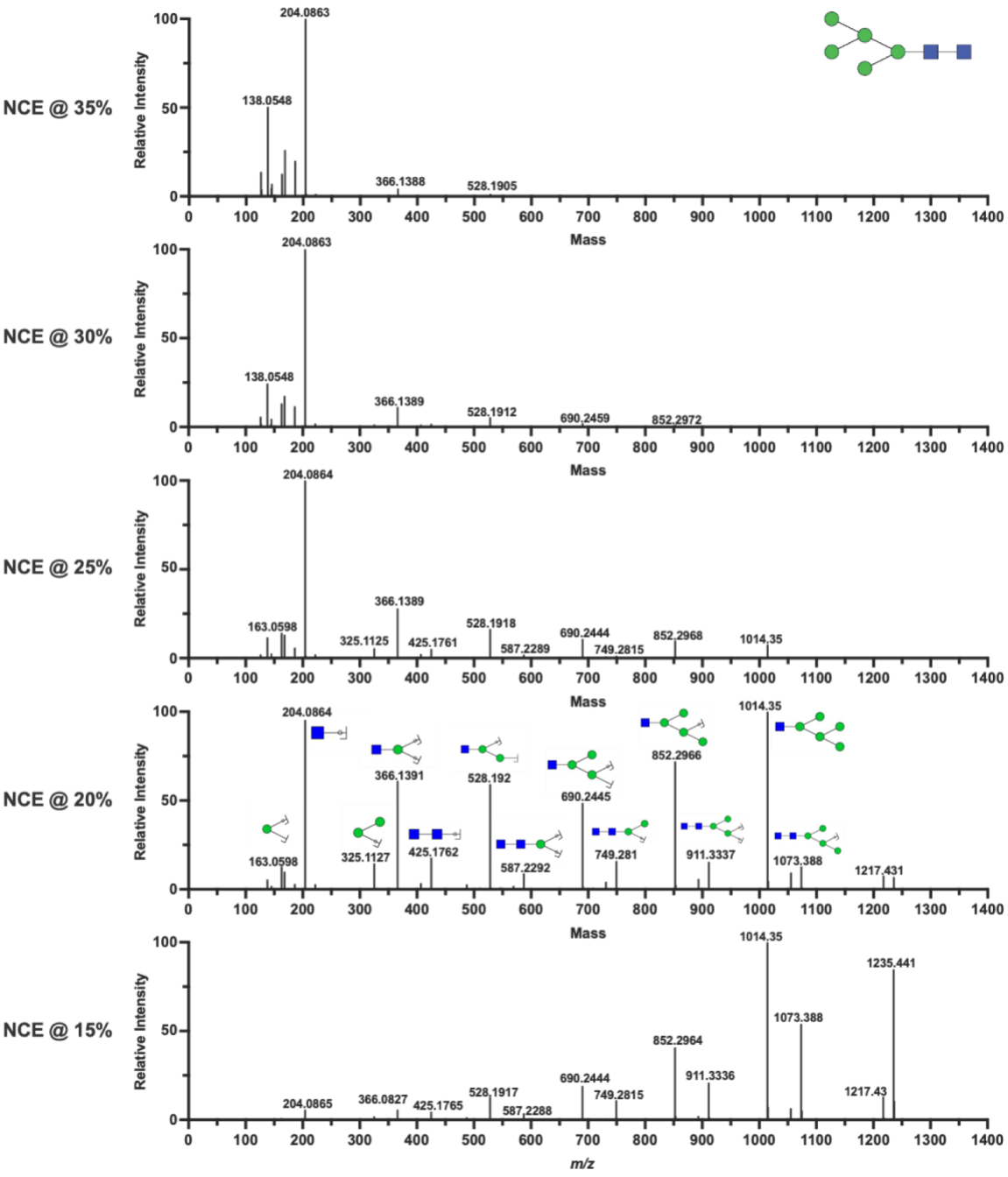
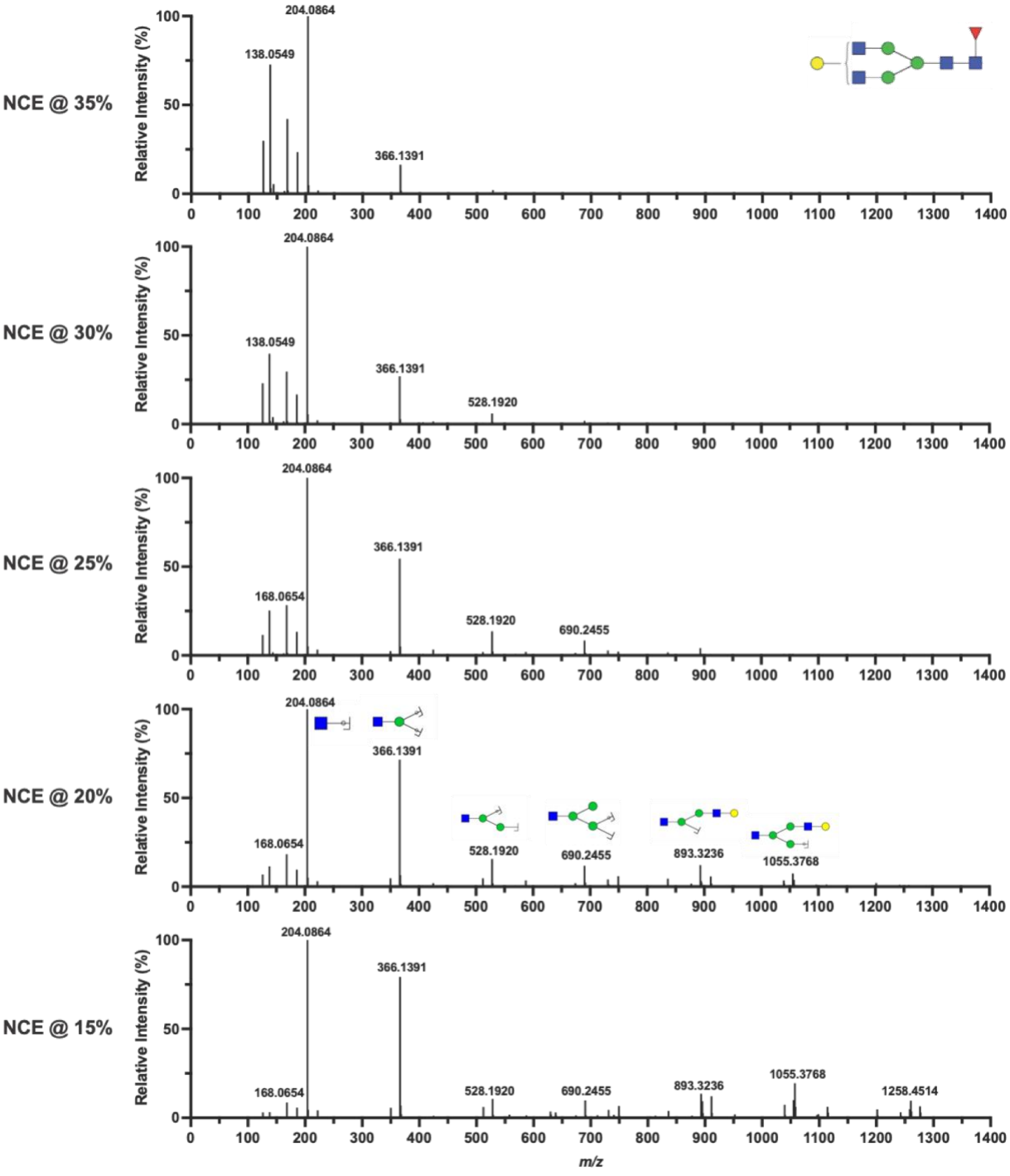
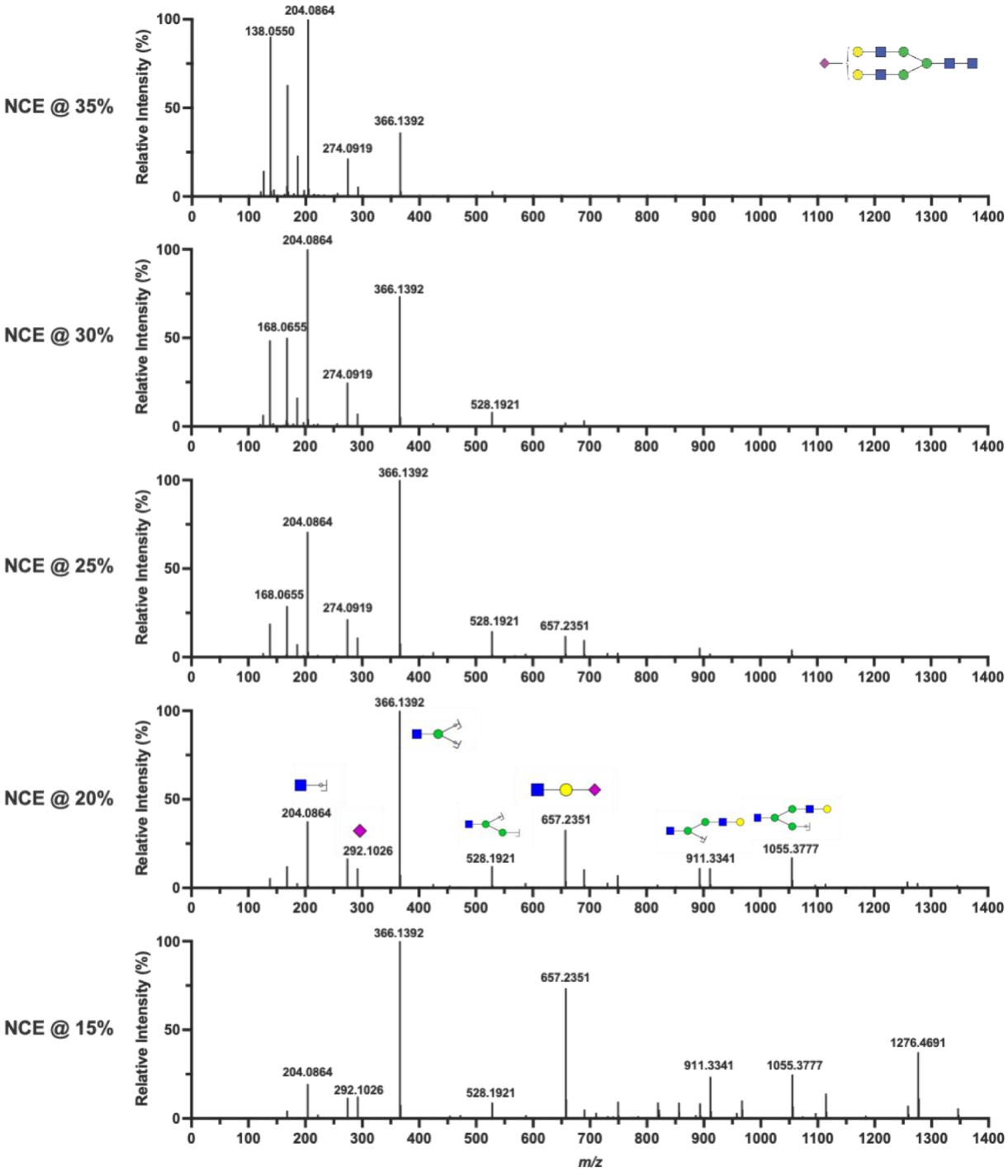
Supplementary Figures.



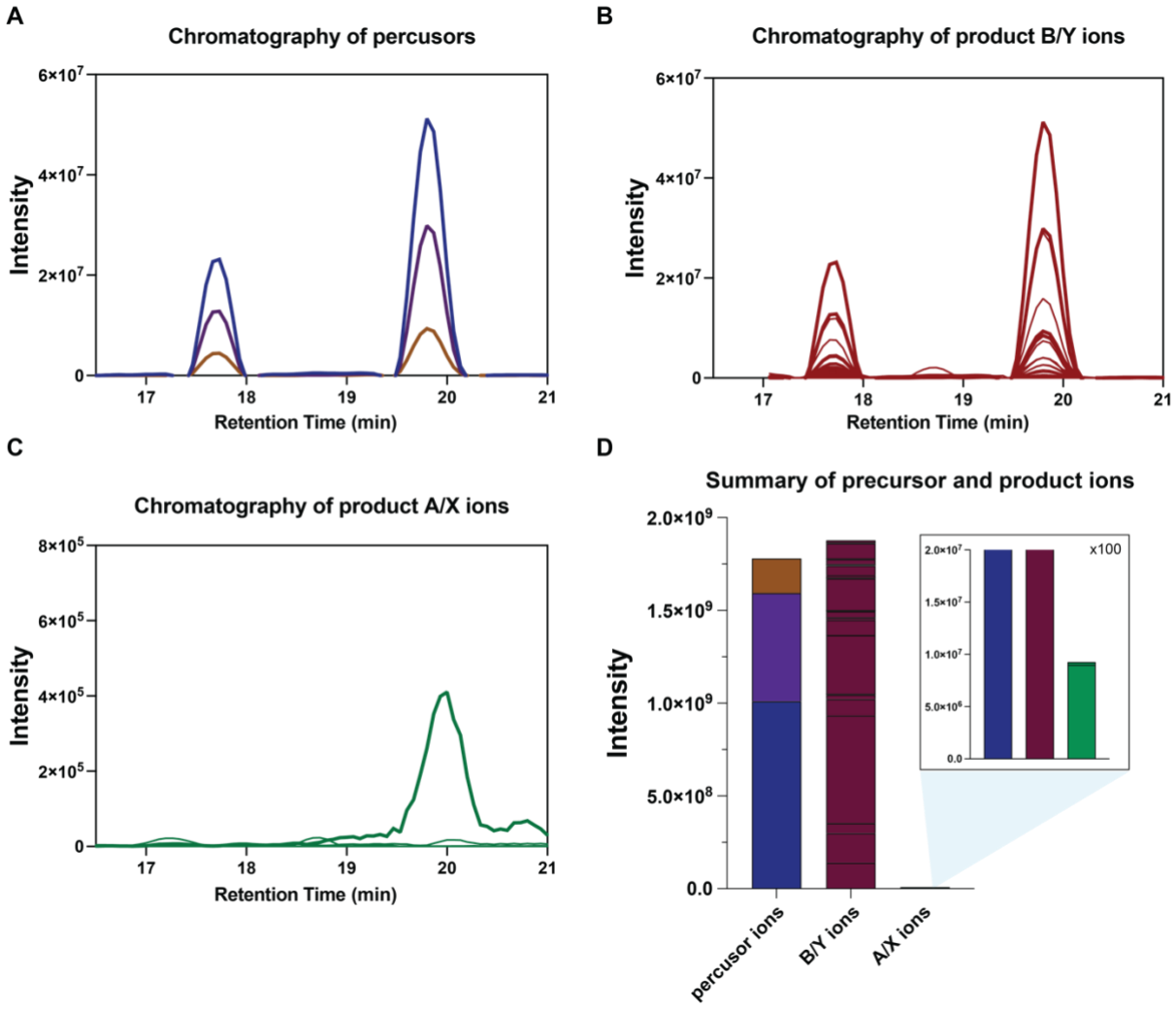
Supplementary Figure 1. Representative tandem MS/MS spectra of high mannose glycan, Hex(5)HexNAc(2), at different fragmentation energies. The abundance of generated fragments that can be used for sequencing the example glycan increased with the NCE boost, while the sequencing information was lost after the collision energy was greater than 25% due to the over-fragmentation.



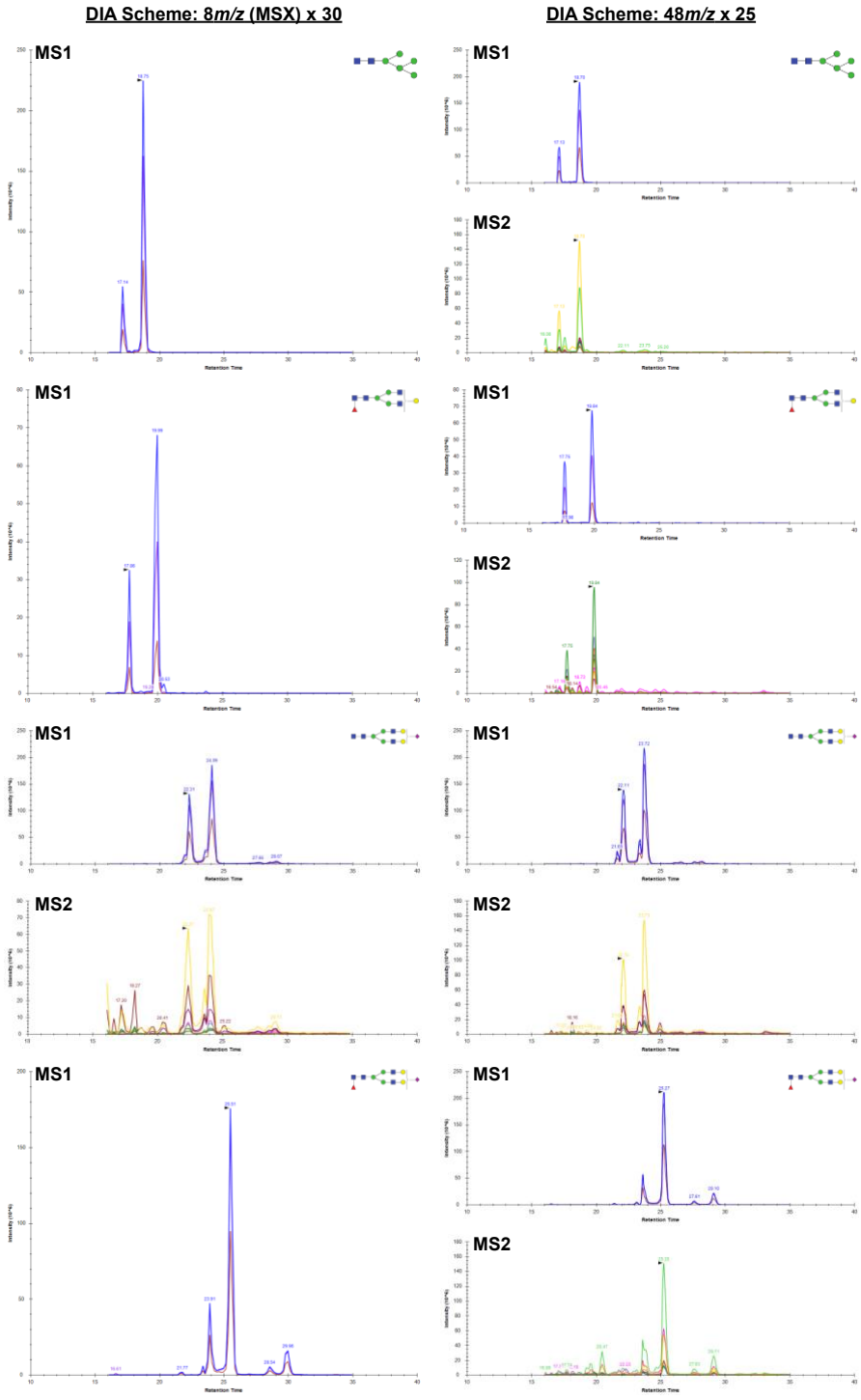
Supplementary Figure 2. Representative tandem MS/MS spectra of fucosylated glycan, Hex(4)HexNAc(4)Fuc(1), at different fragmentation energies. The abundance of generated fragments that can be used for sequencing the example glycan increased with the NCE boost, while the sequencing information was lost after the collision energy was greater than 25% due to the over-fragmentation.



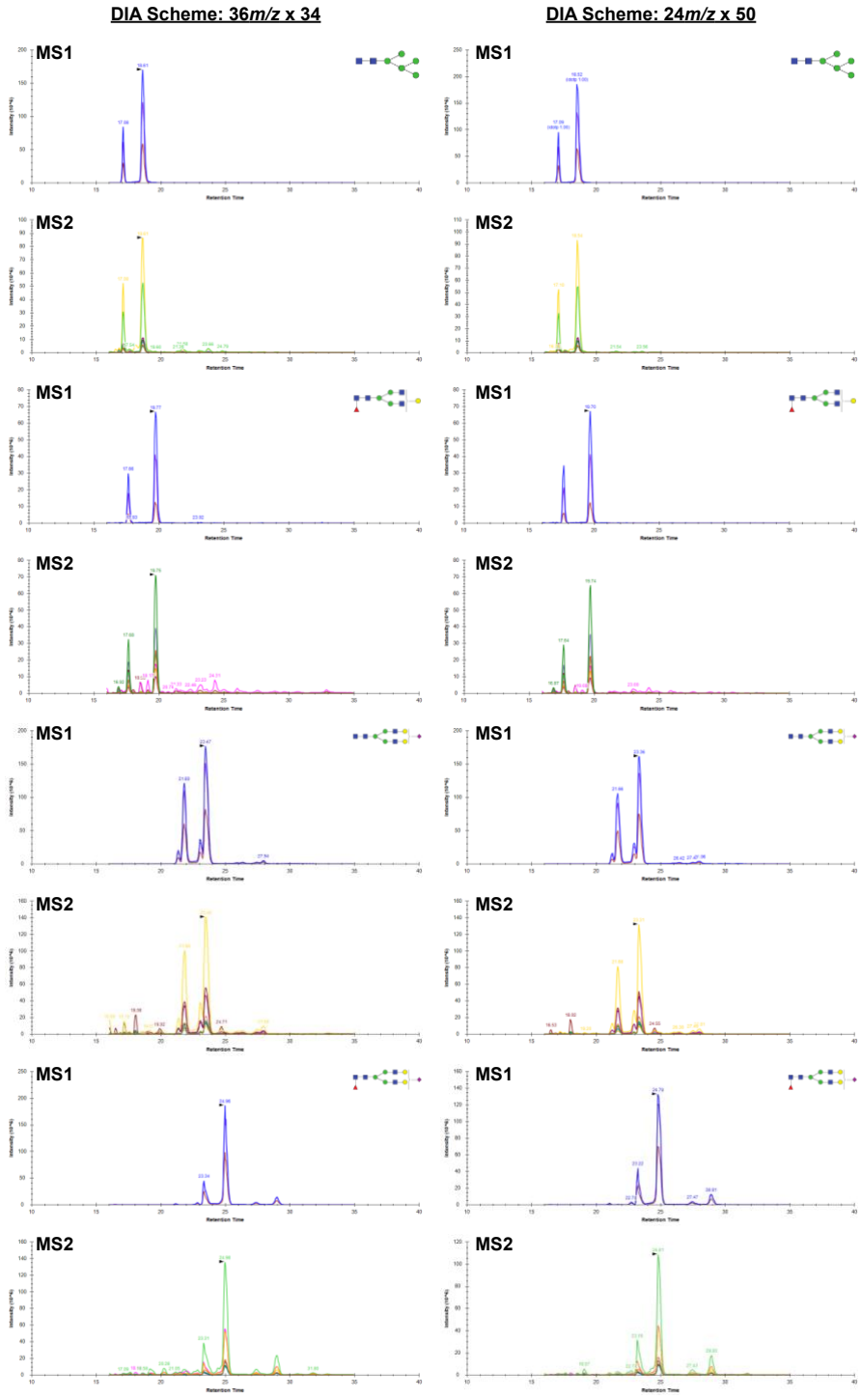
Supplementary Figure 3. Representative tandem MS/MS spectra of sialylated glycan, Hex(5)HexNAc(4)Neu5Ac(1), at different fragmentation energies. The abundance of generated fragments that can be used for sequencing the example glycan increased with the NCE boost, while the sequencing information was lost after the collision energy was greater than 25% due to the over-fragmentation.



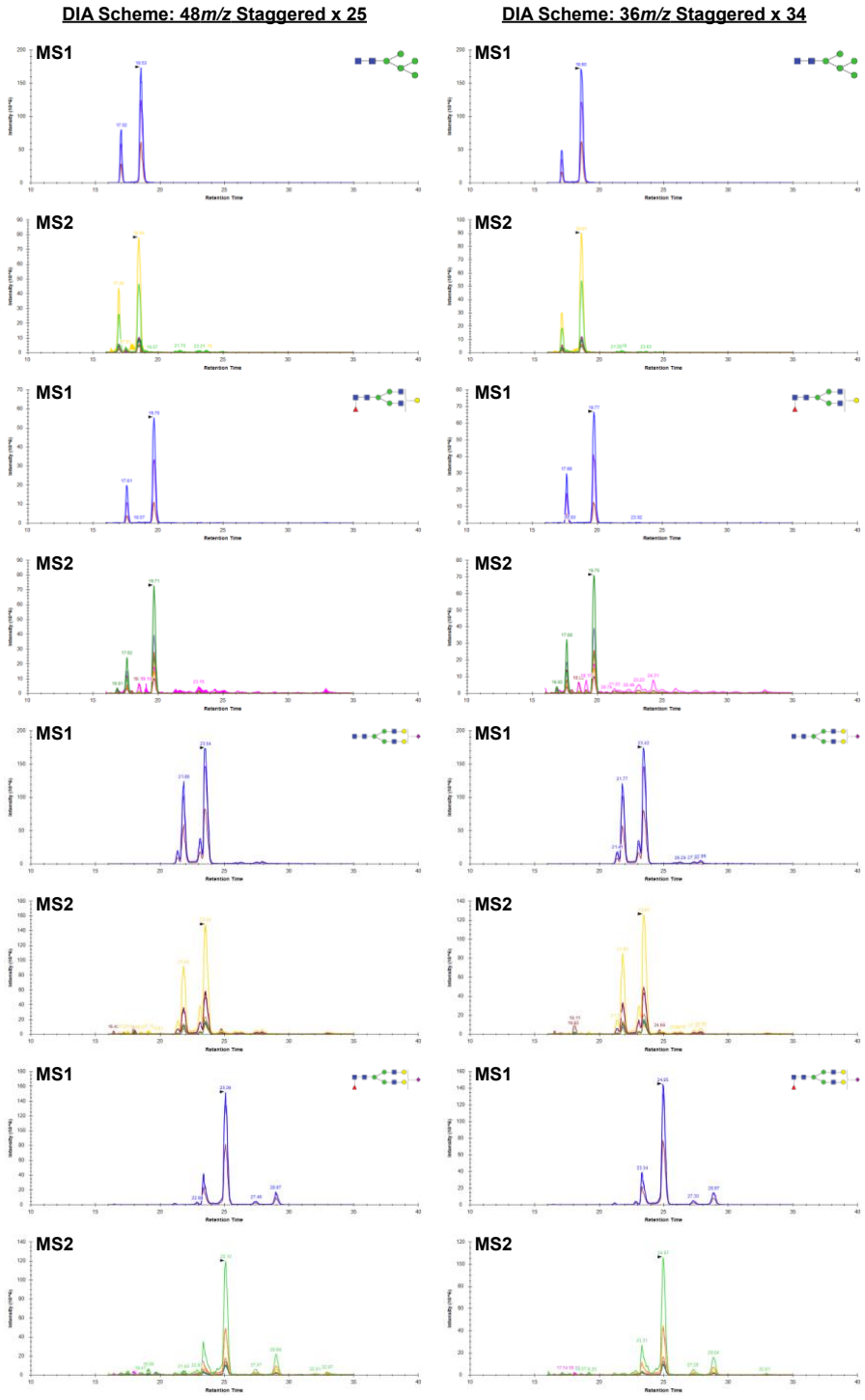
Supplementary Figure 4. The chromatograms of example Hex(5)HexNAc(2) glycan. A. The chromatogram of precursor ions, including [M+0] (blue), [M+1] (purple), and [M+2] (brown). **B.** The chromatogram of generated B and Y product ions from glycosidic bond cleavage. **C.** The chromatogram of generated A and X product ions from cross-ring dissociation. **D.** The summary of different ions from the example glycan showed more than 99.5% of product ions were generated from glycosidic bond cleavages at HCD NCE 20%.



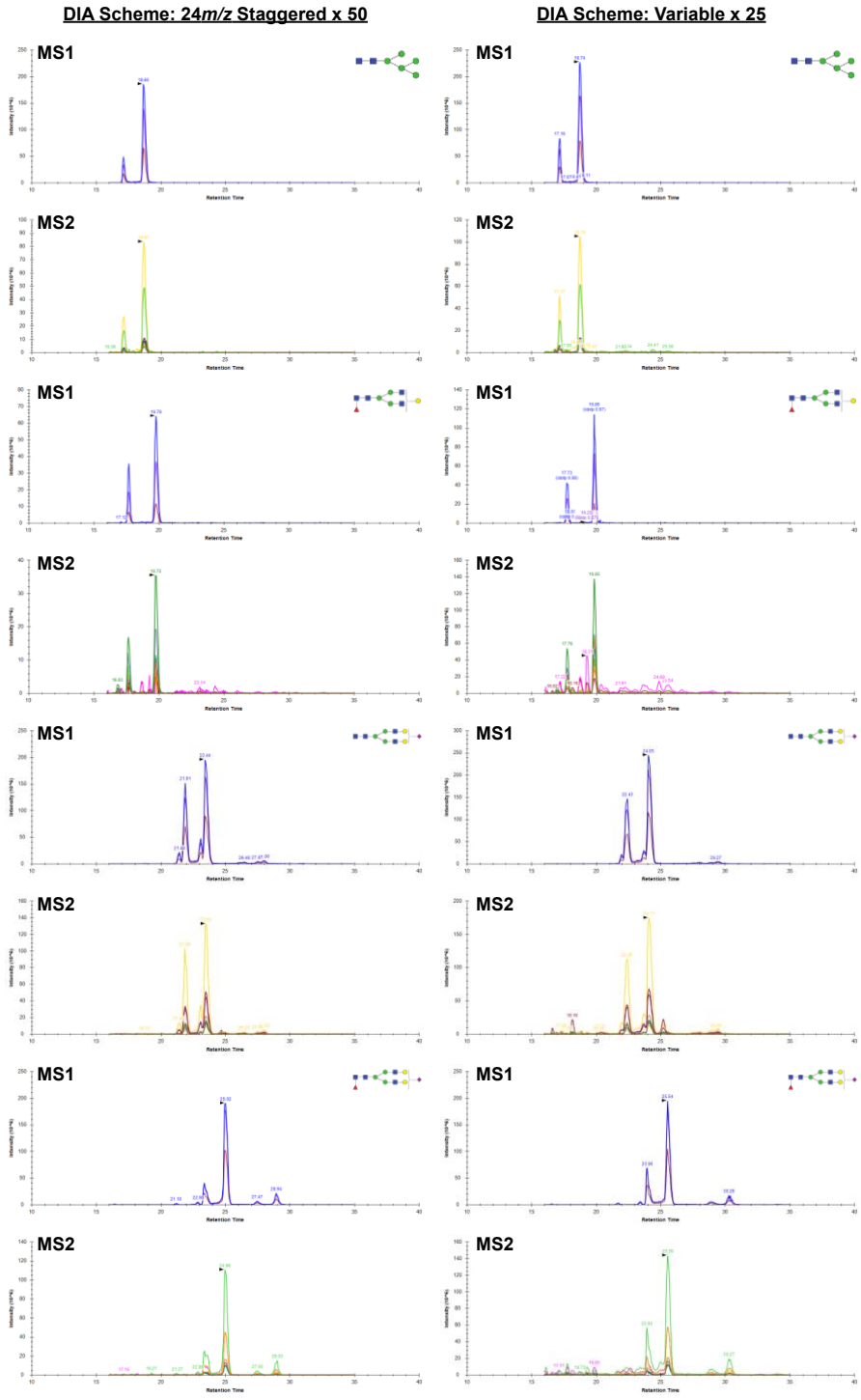
Supplementary Figure 5. Evaluation of different DIA strategies and methods (part 1). Left, multiplexed (MSX) DIA strategy with 8 *m/z* isolation and 30 windows. Some of the glycan MS/MS information was missing due to the randomness of MSX selection. Right, fixed DIA strategy with 48 *m/z* isolation and 25 windows.



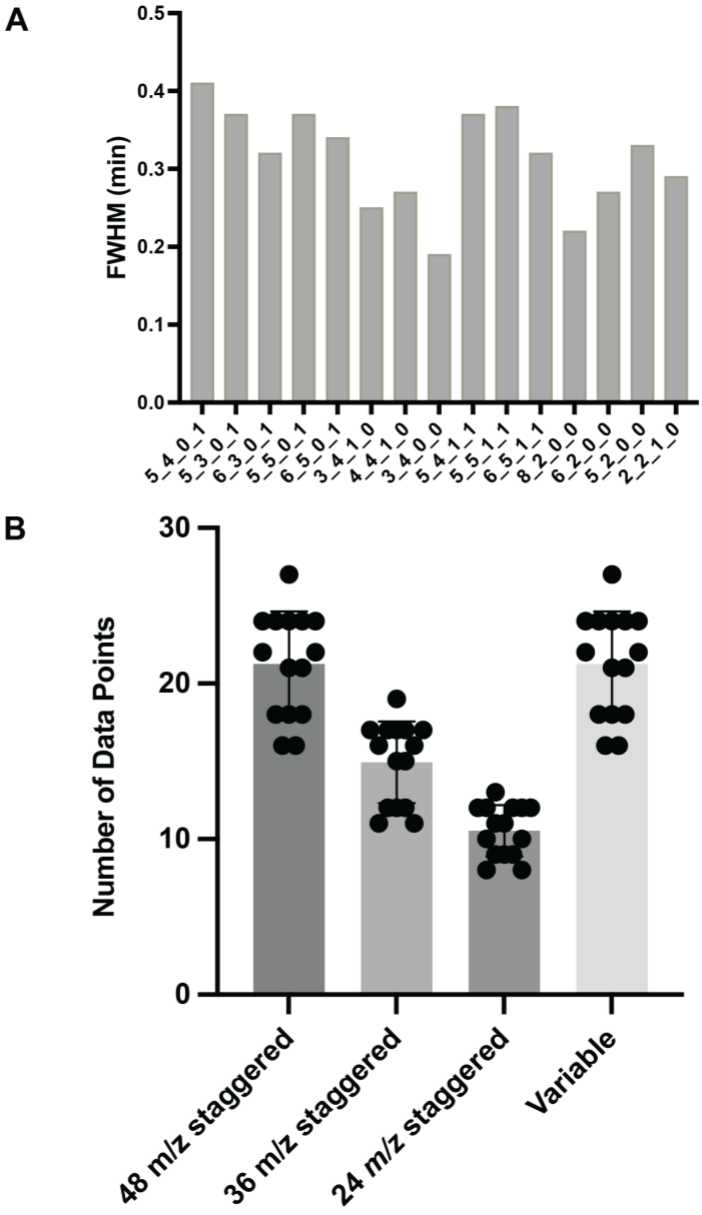
Supplementary Figure 6. Evaluation of different DIA strategies and methods (part 2). Left, fixed DIA strategy with 36 *m/z* isolation and 34 windows. Right, fixed DIA strategy with 24 *m/z* isolation and 50 windows.



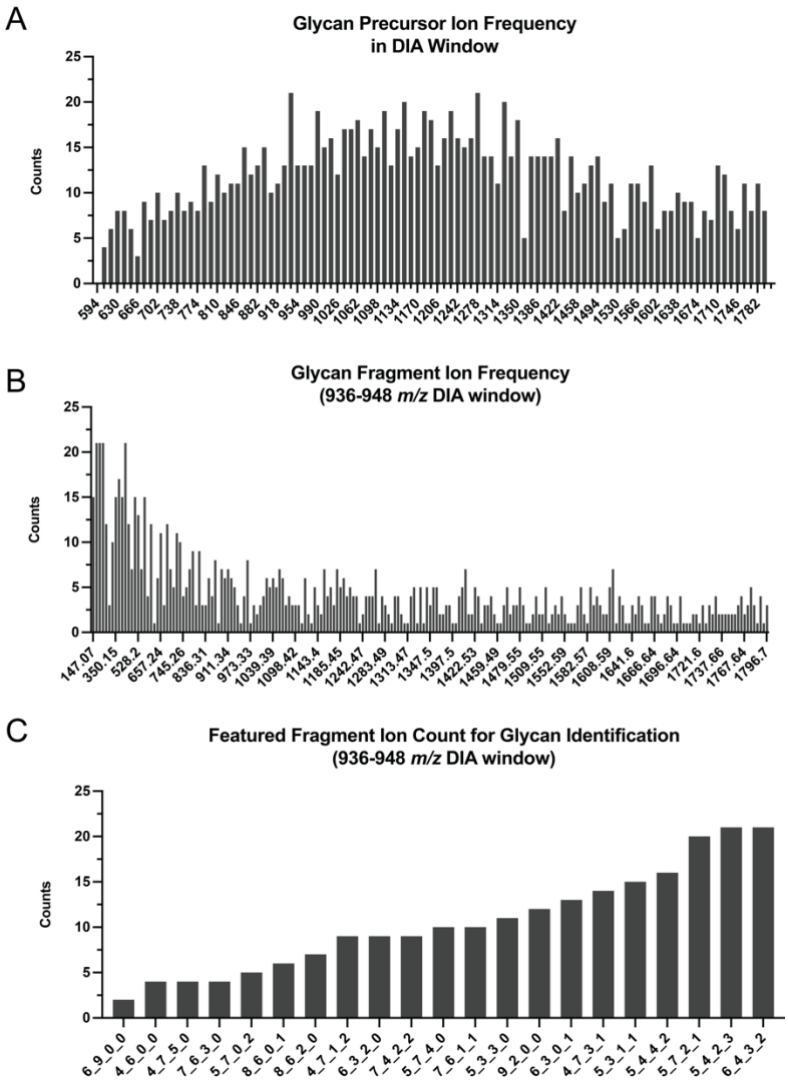
Supplementary Figure 7. Evaluation of different DIA strategies and methods (part 3). Left, stagger DIA strategy with 48 *m/z* isolation and 25 windows. Right, stagger DIA strategy with 36 *m/z* isolation and 34 windows.



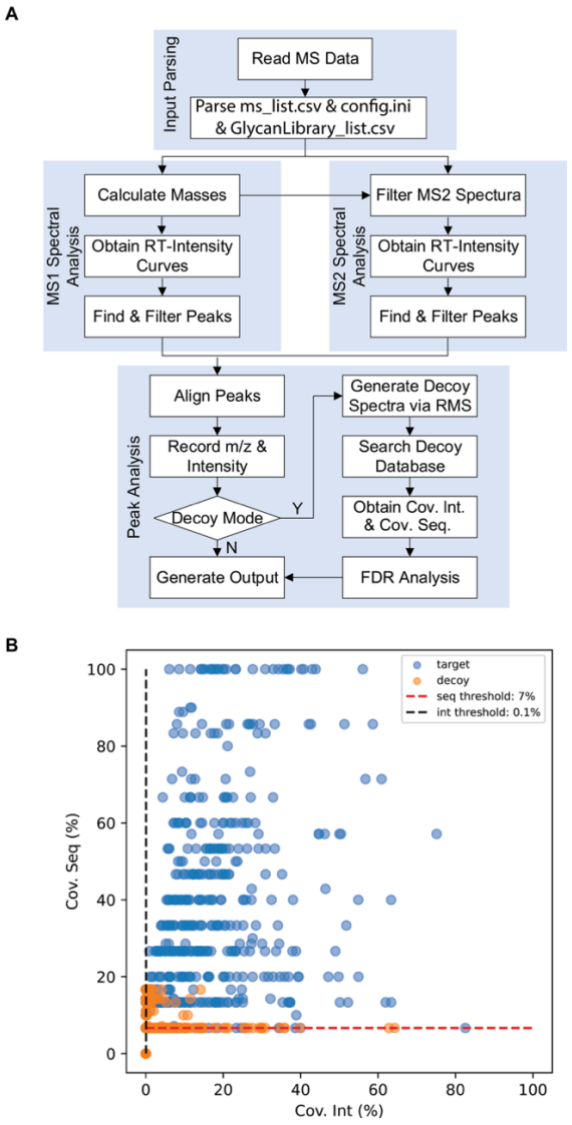
Supplementary Figure 8. Evaluation of different DIA strategies and methods (part 4). Left, stagger DIA strategy with 24 *m/z* isolation and 50 windows. Right, variable DIA strategy with 25 windows, and the isolation window was calculated based on the ion distribution of analyzed glycans.



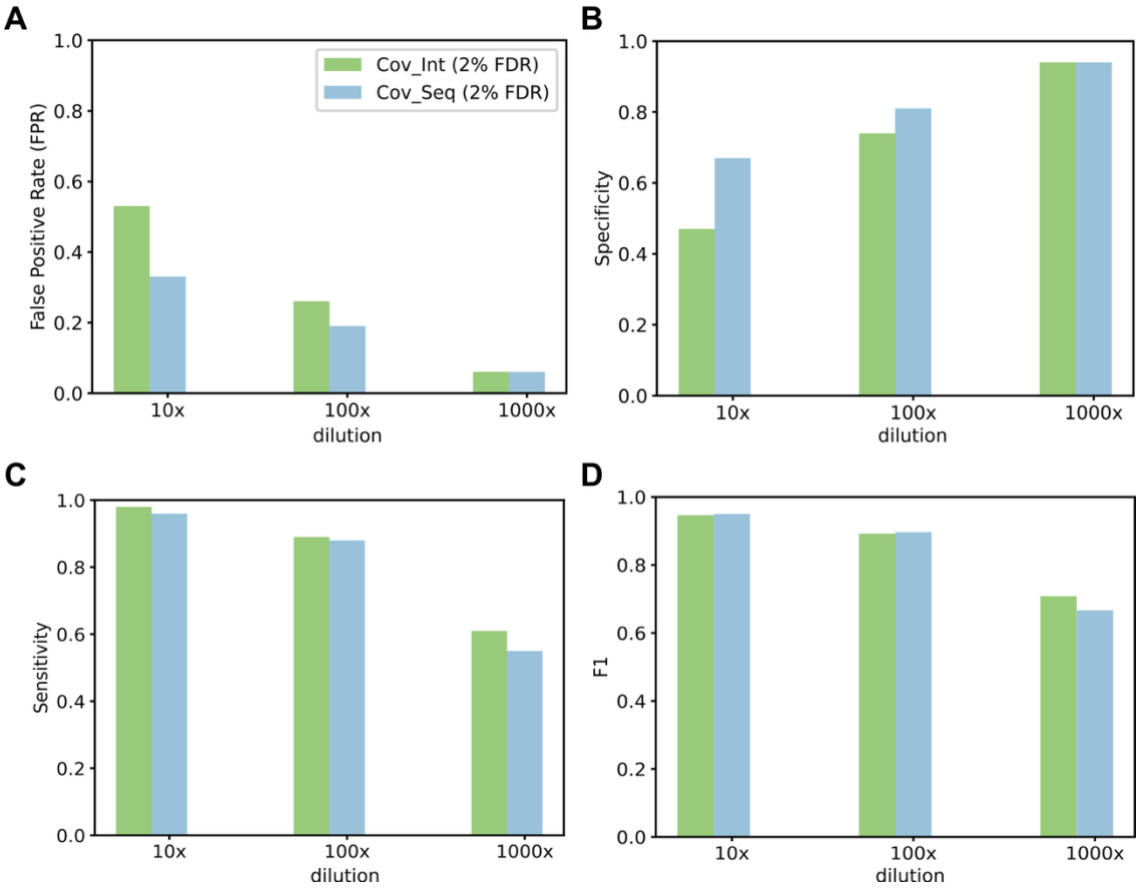
Supplementary Figure 9. The full width at half maximum (FWHM) and data points across glycan peaks. A. In the PGC column, the average FWHM of glycans is more than 0.3 minutes, which provides enough length for DIA scans. **B.** Due to the retention of glycans in PGC materials, the GlycanDIA workflow enables more than 10 data points for different glycan peaks, which is optimal for accurate quantification. Each point represents different glycans, and data are presented as mean ± SD. Source data are provided as a Source Data file.



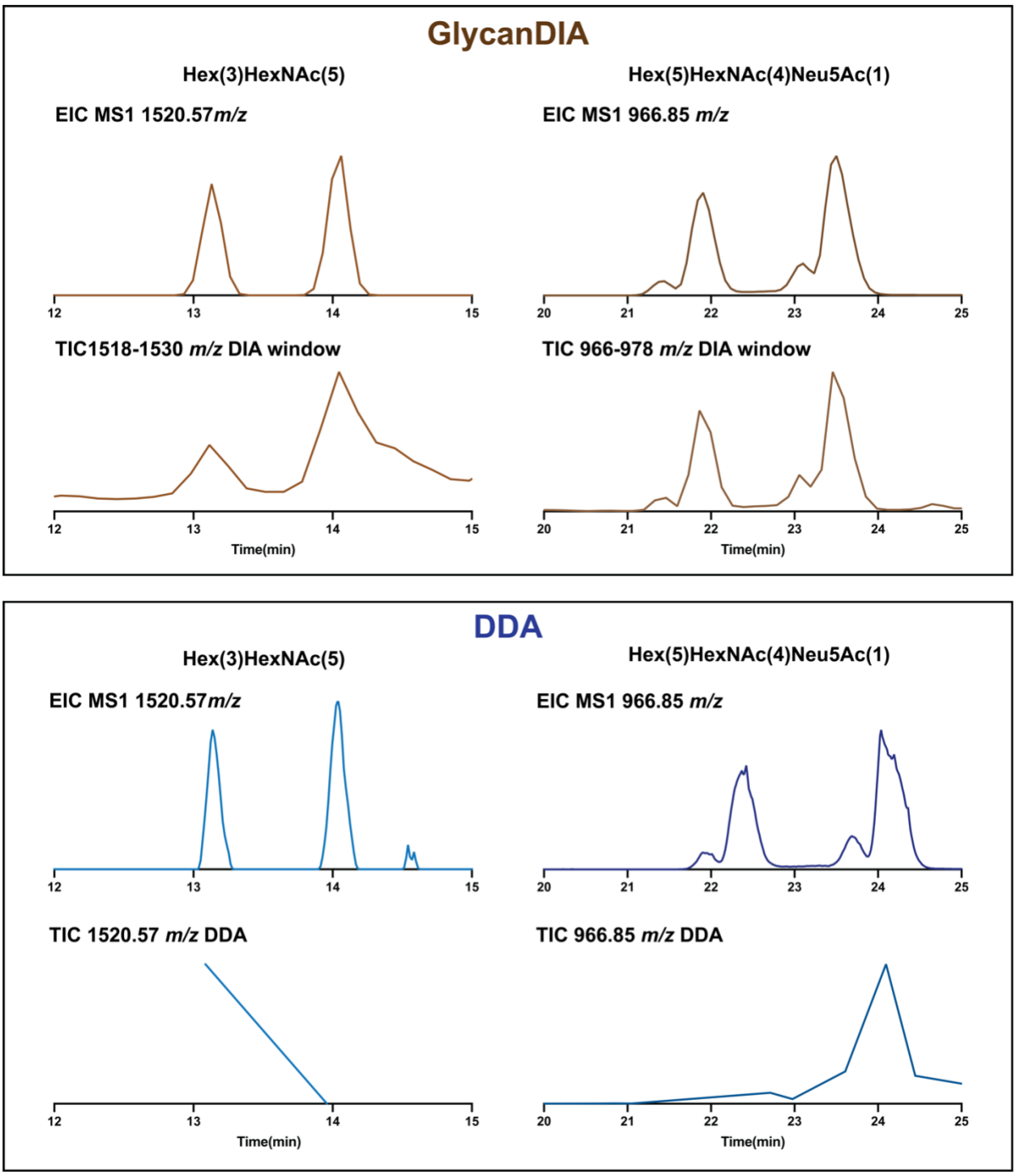
Supplementary Figure 10. Evaluation of precursor and fragment ion distributions in a 24 *m/z* staggered DIA window. A. Precursor ion density across *m/z* 600–1800, with highest peak density observed at 936–948 *m/z* (21 precursors). **B.** Fragment ion distribution in the 936–948 *m/z* window, showing limited sharing of fragments >500 *m/z* (shared by ≤5 precursors). **C.** Featured fragment ions per glycan (≥2 ions/glycan, shared by ≤3 precursors). Collectively, these findings validate the 24 *m/z* staggered window as optimal for GlycanDIA. Source data are provided as a Source Data file.



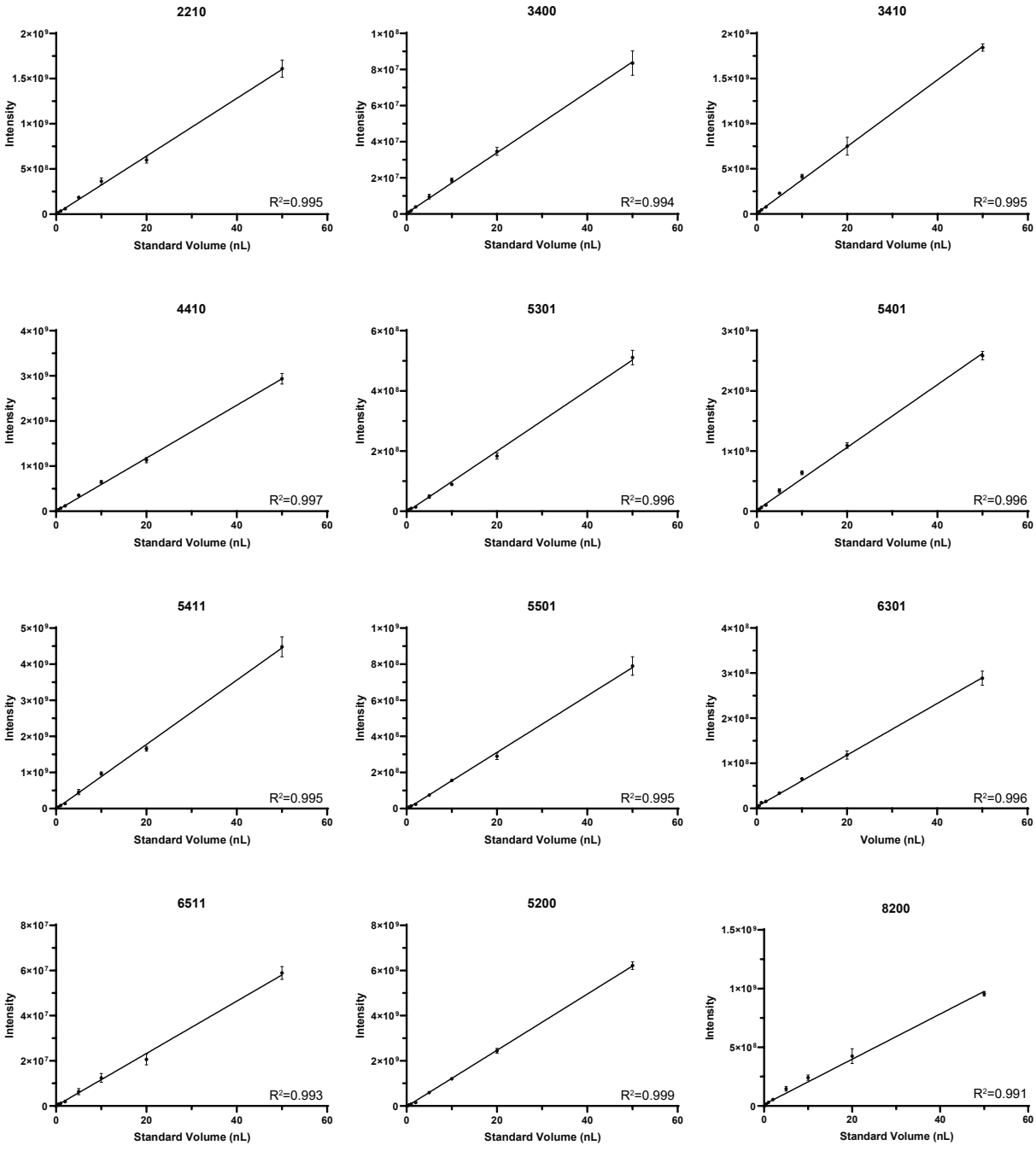
Supplementary Figure 11. Data analysis using GlycanDIA Finder. A. GlycanDIA Finder is an open-source software tool built upon Python and MatchMS library to automatically search glycans and align peaks between the MS1 and MS2 data in a batch of MS files. The configuration settings, including input/output paths, polarity, adduct, charges, mass error, relative height and minimum height of peaks, the time window, and minimum matched count, control the behaviors of the software. The software also incorporated FDR calculation to remove the false identified glycans. A ± 50% mass-shift range is used as the default parameters in the target–decoy-based quality control: 1) the precursor mass and the number of MS/MS peaks remain the same; 2) all original peaks were deleted, and then randomly selected peaks were added to the spectrum. **B.** Score distribution of target and decoy searches. The FDR was derived from **Supplementary Data 6**. Target search (blue) distributed in most parts of the density map, indicating that there are both high-quality and low-quality spectra. At the same time, spectra of decoy search(orange) distributed only in the low-left parts of the density map, indicating the distribution of false positives. The threshold is with 2% FDR. More details about the software can be found in **Supplementary Information 2.**



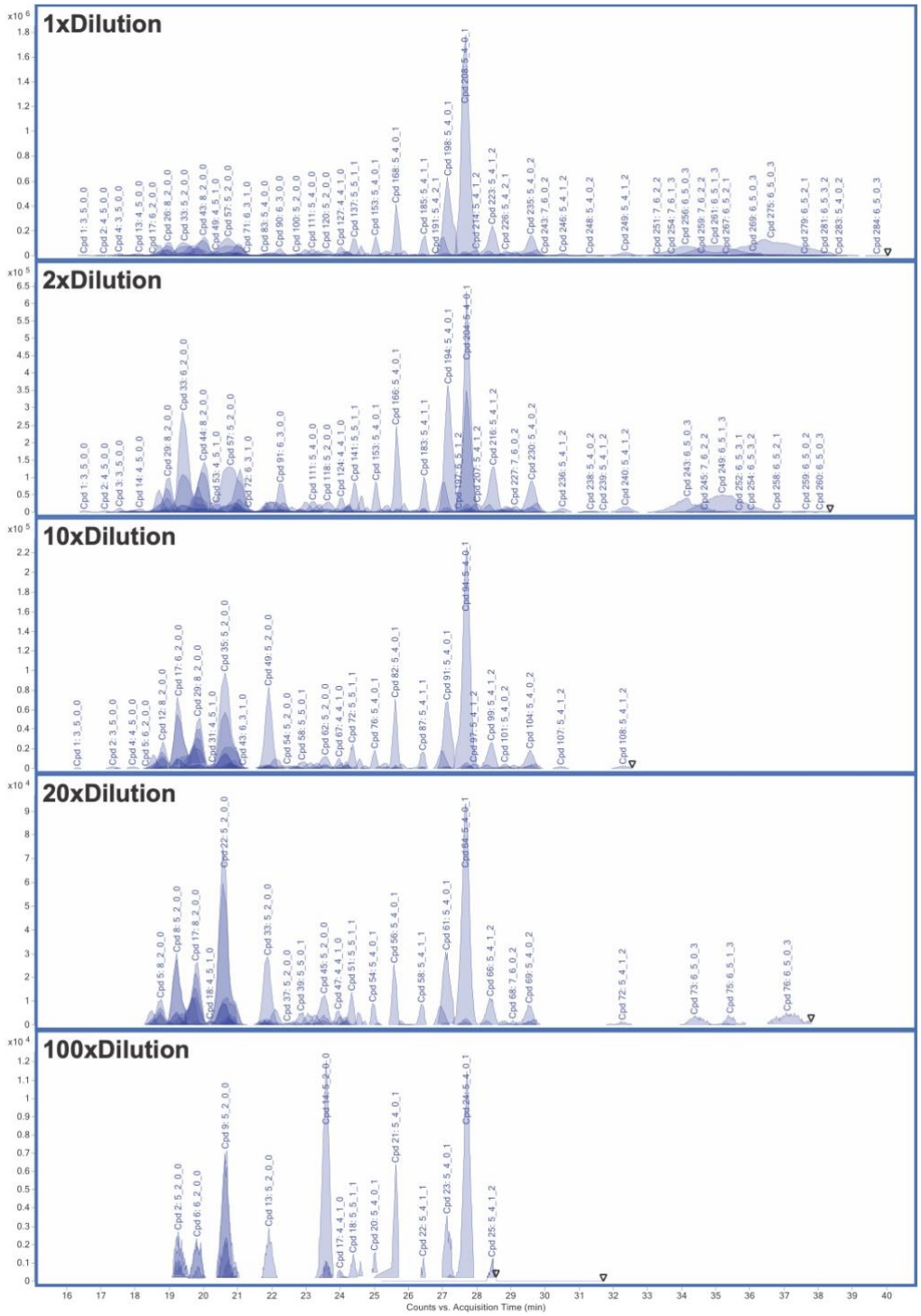
Supplementary Figure 12. Robustness of Con_Int and Con_Seq scoring metrics across dilution series in GlycanDIA Finder. The analysis was conducted across three dilution levels (10×, 100×, and 1000×) and evaluated using four key performance metrics: **A.** False Positive Rate, **B.** Specificity, **C.** Sensitivity, and **D.** F1 Score. Con_Int reflects the agreement between observed and theoretical fragment ion intensities, while Con_Seq measures sequence-level alignment. Panel A shows that the false positive rate (FPR) decreases progressively with dilution, dropping below 0.3 at 100× and approaching zero at 1000×, indicating improved discrimination under lower signal conditions. Panel B highlights a substantial increase in specificity, reaching ~0.9 at 1000× dilution. Panel C shows that sensitivity remains high at all dilutions but gradually declines at higher dilutions. Panel D demonstrates that the F1 score remains consistently high at 10× and 100× dilutions, with a modest decline at 1000×. These results demonstrate that both Con_Int and Con_Seq maintain robust performance under dilution and that the 100× dilution condition offers an optimal trade-off between specificity and sensitivity. Source data are provided as a Source Data file.



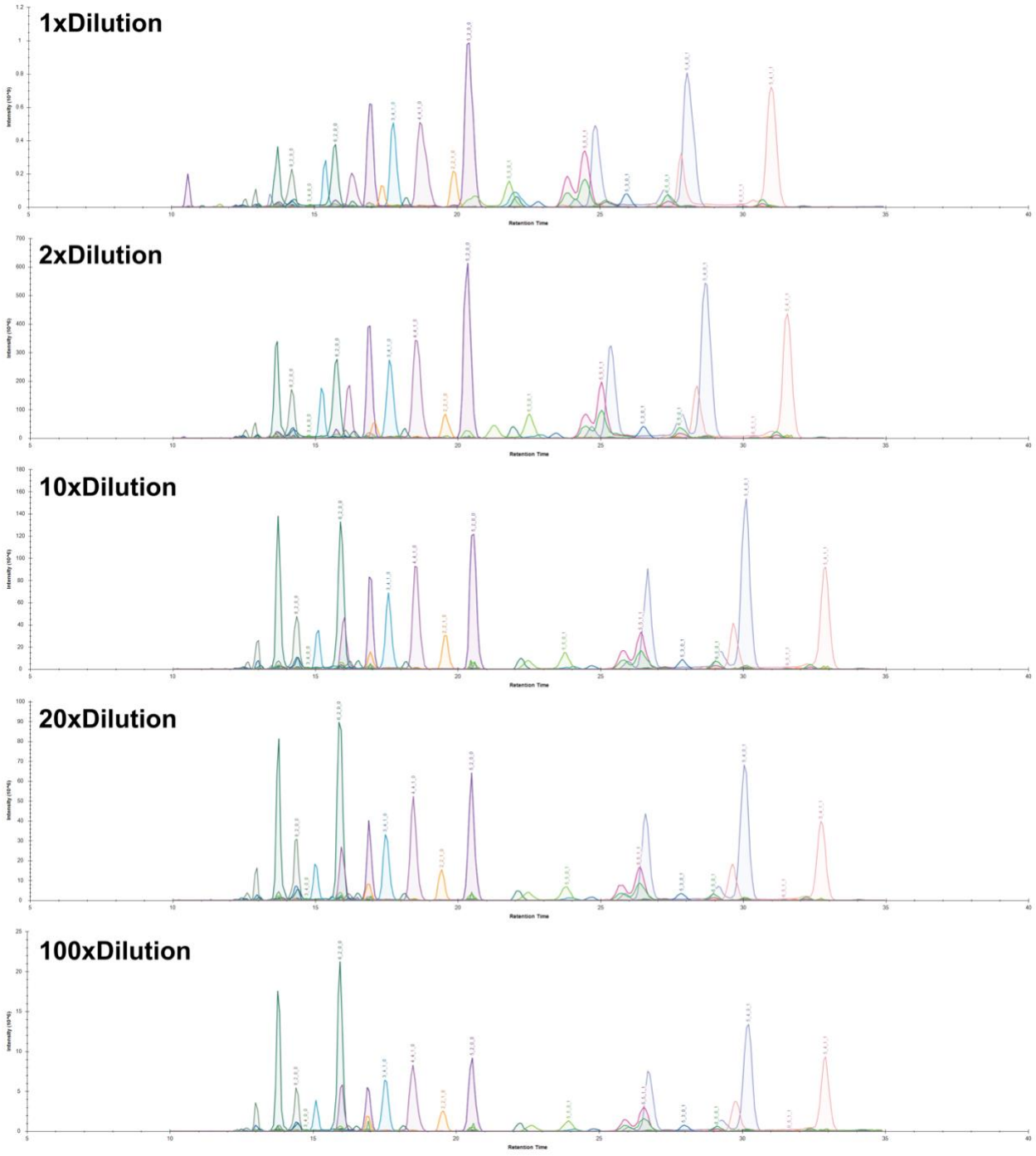
Supplementary Figure 13. GlycanDIA enables comprehensive MS/MS coverage for isomers. Top: Tandem MS/MS spectra for Hex(3)HexNAc(5) (left) and Hex(5)HexNAc(4)Neu5Ac(1) (right) generated by GlycanDIA across all chromatographic peaks. **Bottom:** DDA spectra for the same glycans, with only 1 and 3 MS2 events triggered, respectively.



Supplementary Figure 14. Calibration curves of different glycans using DIA-based nanoLC-orbitrap system. Unfilled points were not included in linear curves. All curves yielded superior R^2 values (>0.99, and decent linearities were obtained after a 1000-time dilution for most glycans. These results demonstrated the sensitivity of the GlycanDIA workflow. Annotation is as follows: Hex_HexNAc_Fuc_Neu5Ac. Source data are provided as a Source Data file.

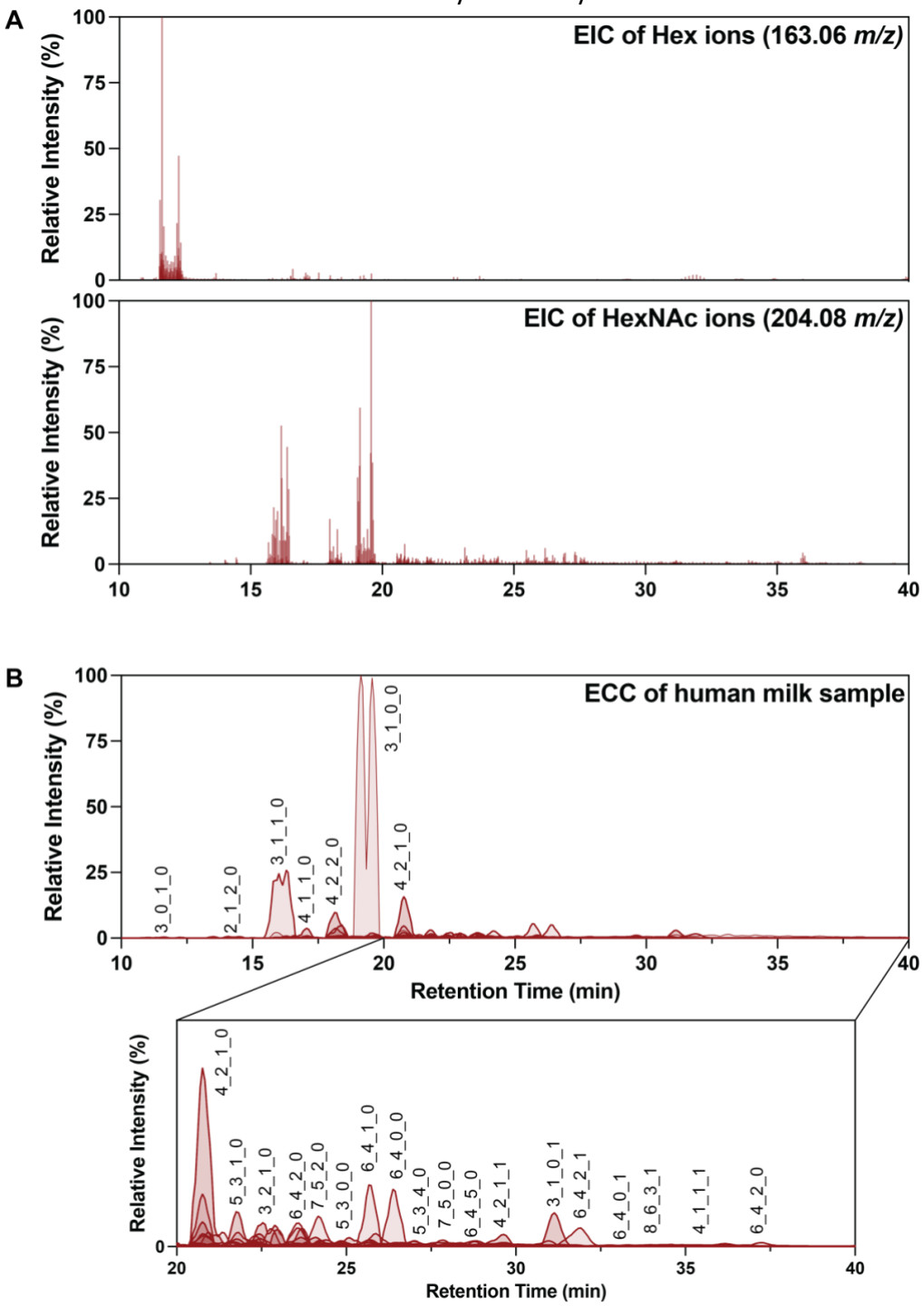


Supplementary Figure 15. Characterization of glycans with different dilution factors using the nanoLC-ToF system. Each peak in the chromatogram represents a unique glycan compound (Cpd). The glycan peak can barely be quantified after a 100-time dilution of the glycan standard.

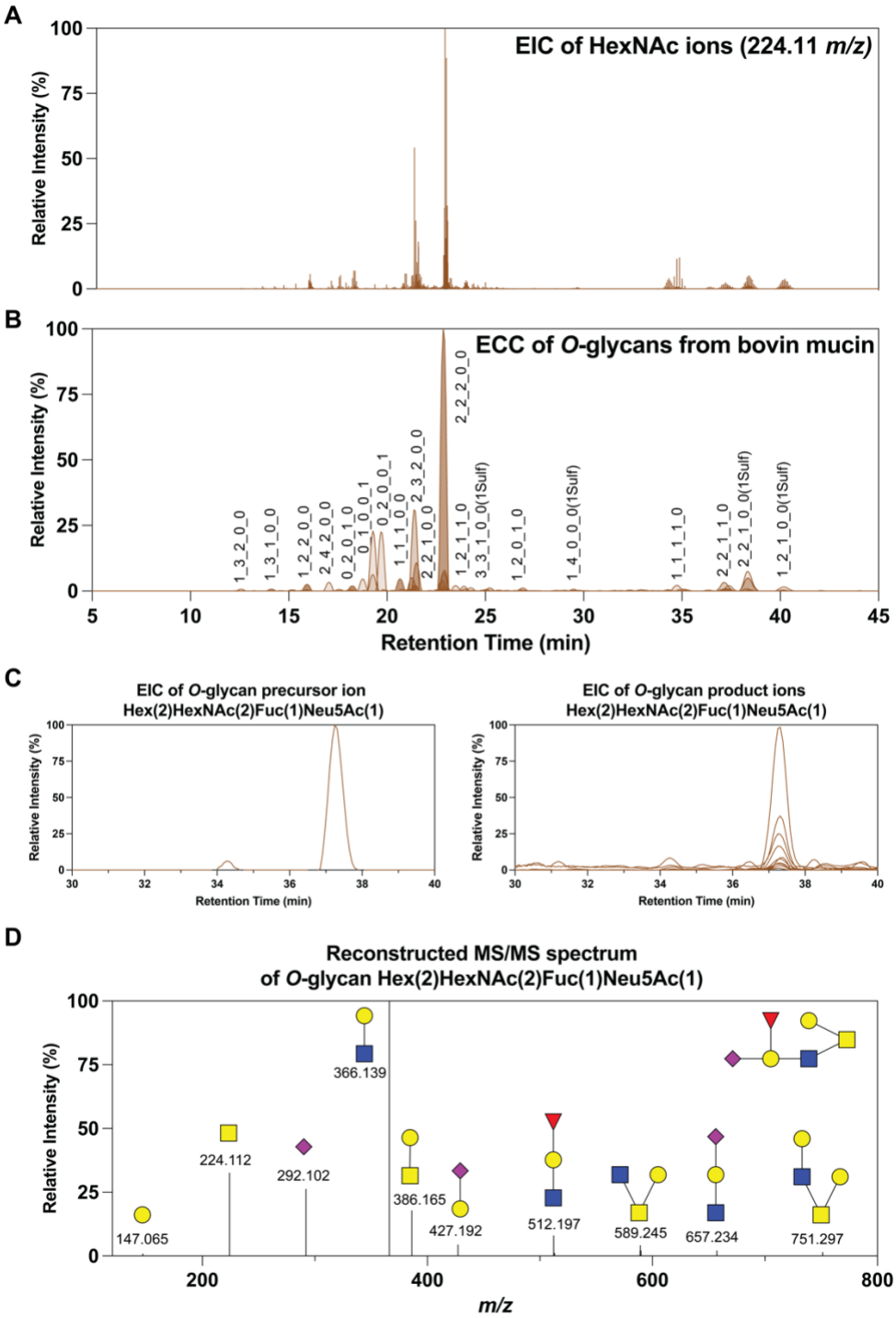


Supplementary Figure 16. Characterization of glycans with different dilution factors using GlycanDIA. Each peak in the chromatogram represents a unique glycan compound. The glycan peak identification and quantification can be preserved after a 100-time dilution of the glycan standard.

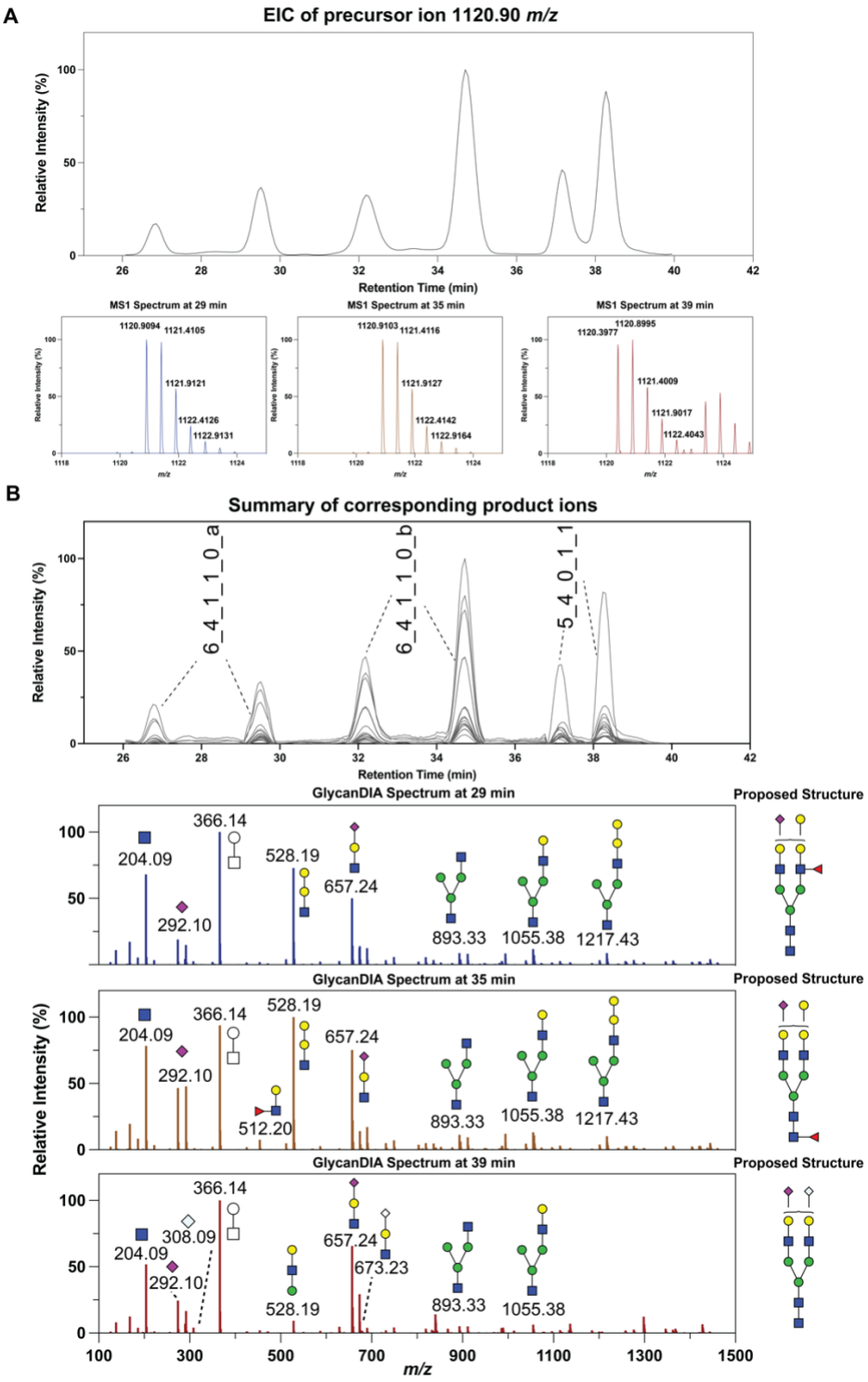
demonstrated the sensitivity of the GlycanDIA workflow.



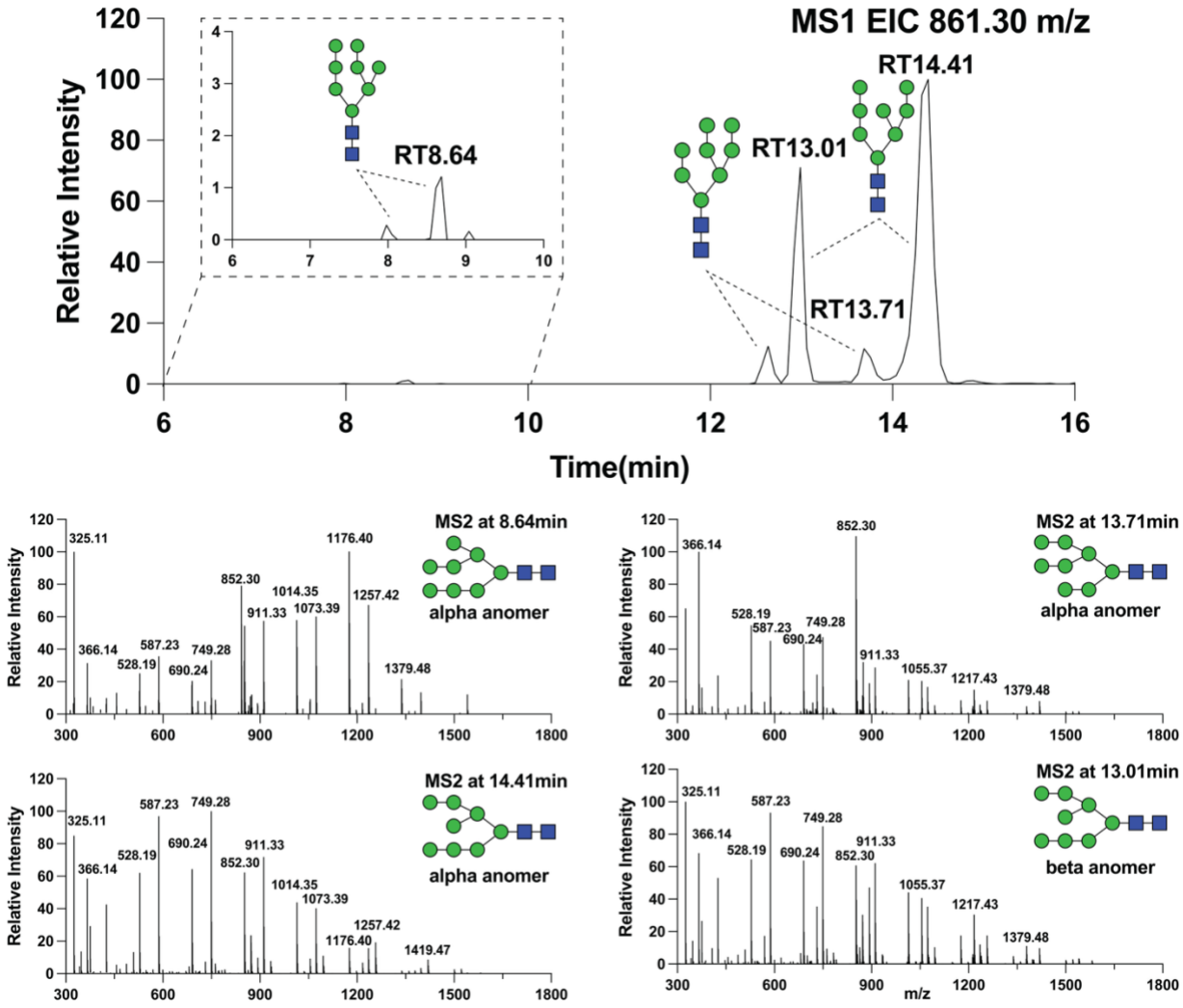
Supplementary Figure 17. The profile of HMO from human milk sample using GlycanDIA. A. Extracted ion chromatogram (EIC) of 163.06 *m/z* and 204.08 *m/z* shows the signal of hexose- and N-acetylhexosamine-containing HMOs from the sample. **B.** Extracted compound chromatogram (ECC) shows the example elution profile of HMO from human milk. Annotation is as follows: Hex_HexNAc_Fuc_Neu5Ac.



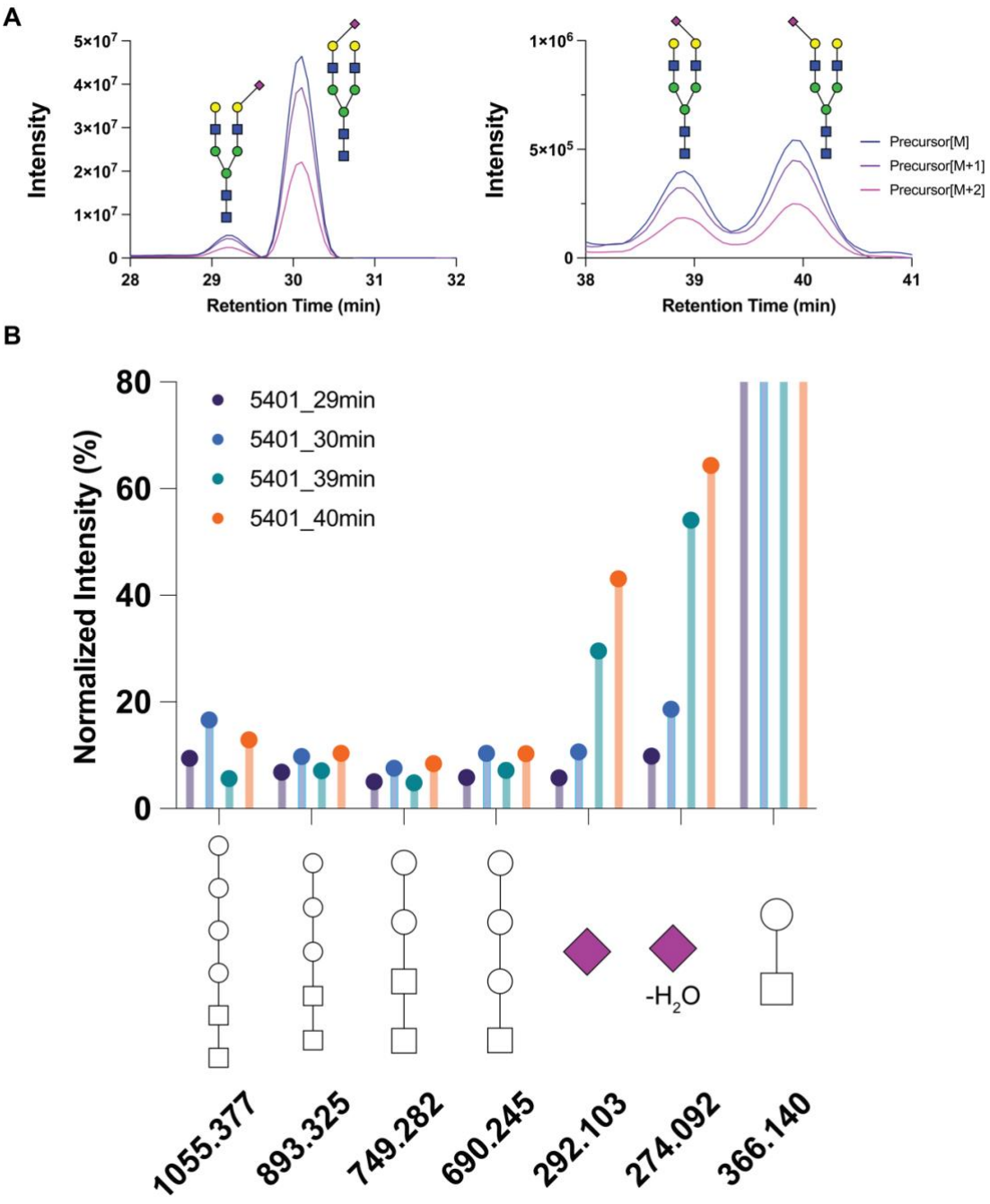
Supplementary Figure 18. The profile of *O*-glycans from bovine mucin using GlycanDIA. A. EIC of 224.11 *m/z* shows the signal of reduced GalNAc. **B.** ECC shows the example elution profile of *O*-glycans from bovine mucin. Annotation is as follows: Hex_HexNAc_Fuc_Neu5Ac_Neu5Gc. **C.** The EIC of example *O*-glycan Hex(2)HexNAc(2)Fuc(1)Neu5Ac(1) precursor (left) and product (right) ions. **D.** The example annotation of reconstructed tandem mass spectrum of the example *O*-glycan. The fragment ions confirmed its structure, while the linkage specification is missed because HCD cannot reveal the general linkage information.



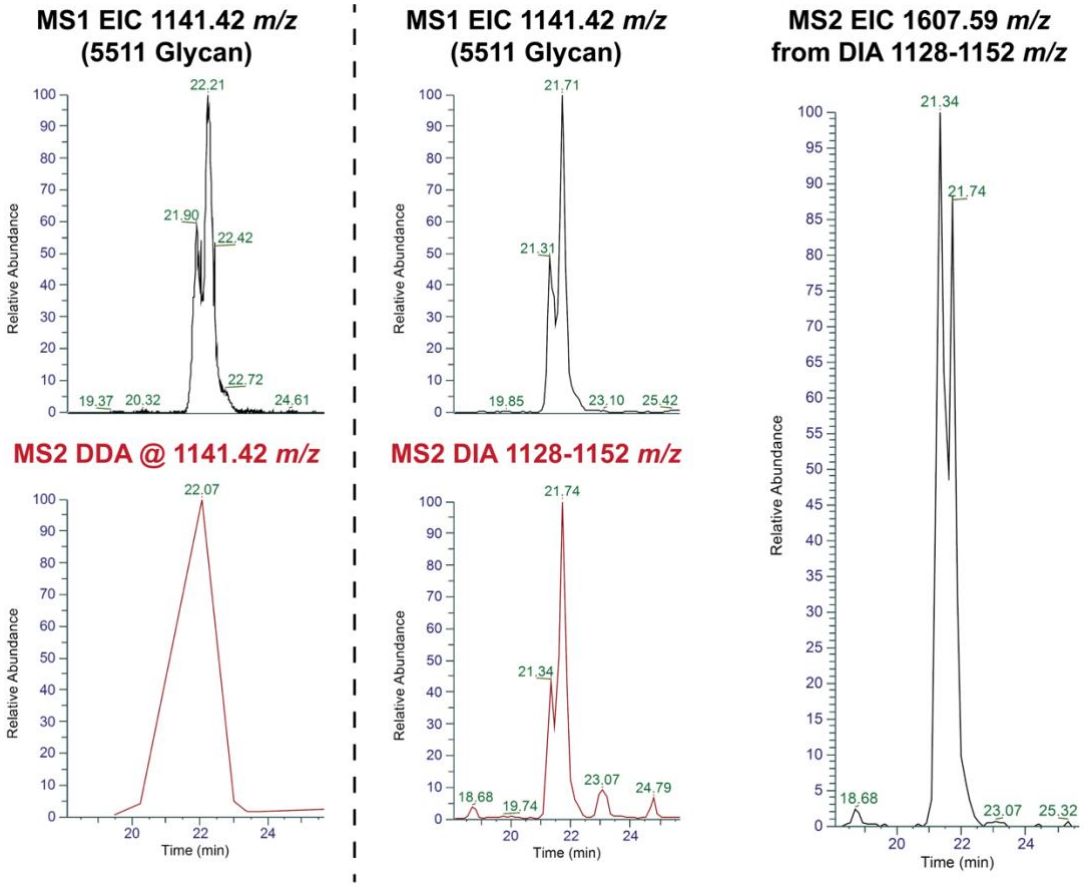
Supplementary Figure 19. The example of identification of glycan composition isomers using GlycanDIA. A. Extracted ion chromatogram (EIC) of 1120.90 *m/z* at MS1 level, and six peaks corresponding to three different glycans and their anomers were noted. **B.** The extracted MS2 chromatogram of identified glycans and the representative glycan MS2 spectra at three different retention times, 29, 35, and 39min, with annotation and proposed structures.



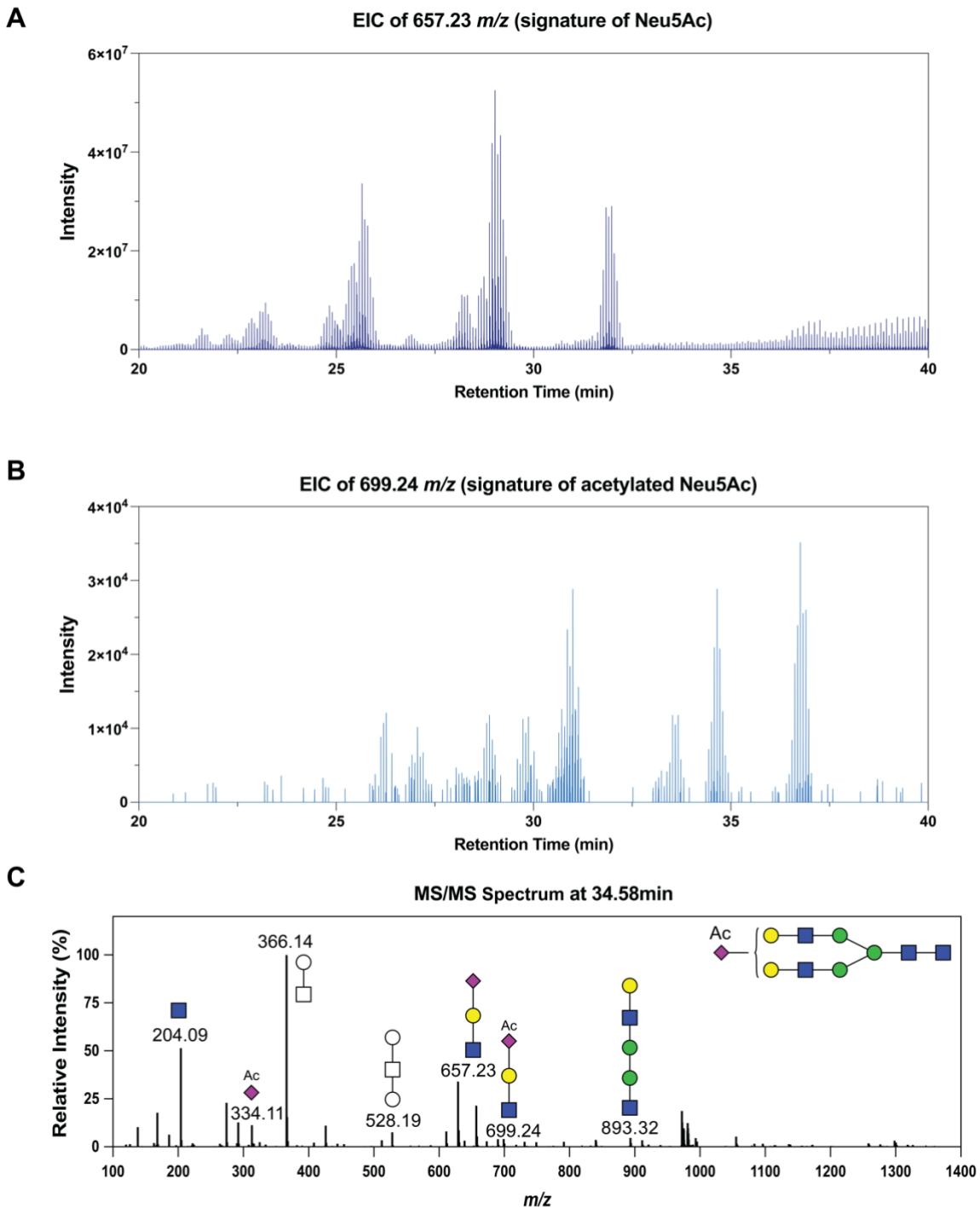
Supplementary Figure 20. Identification and quantification of man8 glycan isomers and anomers using tandem ms/ms. Characteristic fragmentation patterns enabled differentiation of six Man8 isomers/anomers, confirming the method's reliability for isomer-resolved glycomic analysis. MS/MS fragmentation with B-type ions containing two consecutive GlcNAc residues indicates the β-anomer of the less abundant glycan, while B-fragments with only one innermost GlcNAc correspond to the α-anomer of the more abundant glycan.



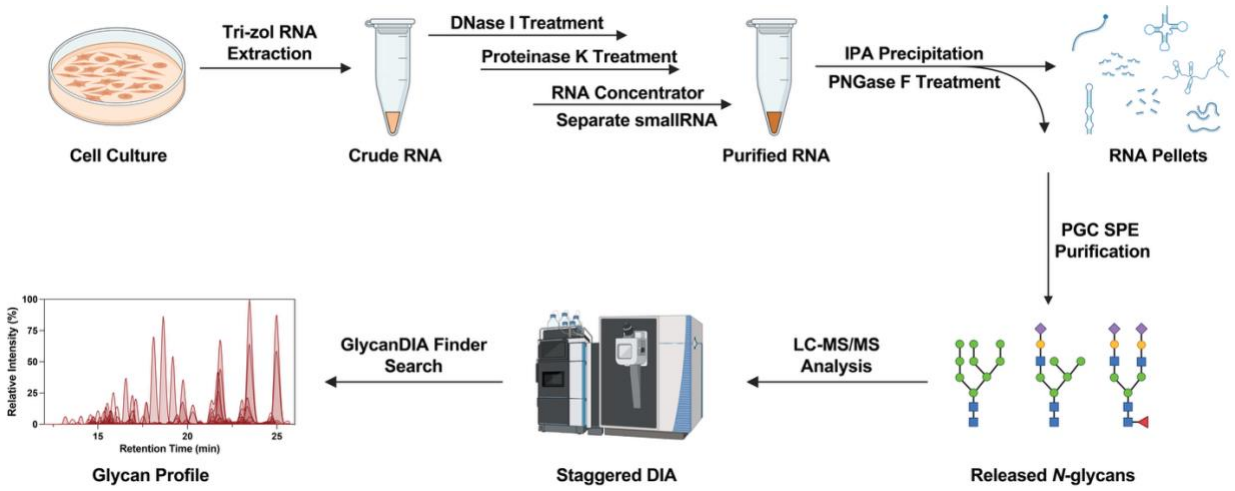
Supplementary Figure 21. The example of identification of sialylated isomers using GlycanDIA. **A.** EIC of the glycan Man(3)Gal(2)GlcNAc(4)Neu5Ac(1) isomer precursor and its isotopic peaks at the MS1 level. **B.** The normalized intensity of different fragments from the example glycans. The product ion intensity was normalized to the fragment of 1Hexose+1HexNAc (366.14 *m/z*). Different product ion distributions reveal the potential capability to determine glycan linkage isomers using GlycanDIA. *Note: the specific glycan structure was based on the previous results reported by Palmisano *et al.* (*RSC Advances*, **2013**). Source data are provided as a Source Data file.



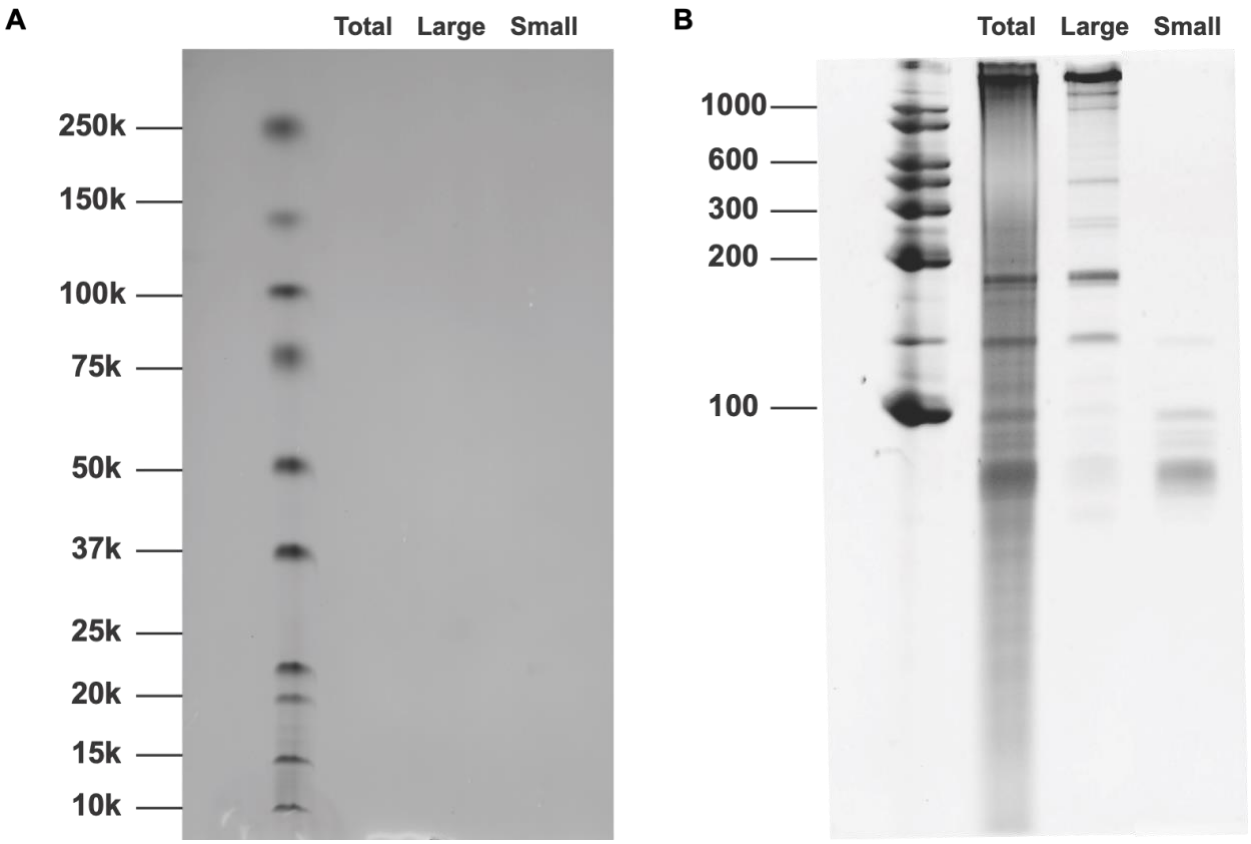
Supplementary Figure 22. Resolving co-eluting glycan isomers via GlycanDIA. GlycanDIA acquires tandem MS/MS spectra across both apex regions of partially resolved Hex(5)HexNAc(5)Fuc(1)Neu5Ac(1) isomers (e.g., 1607.59 *m/z* diagnostic ions). In contrast, DDA triggers only one tandem MS/MS scan between poorly separated peaks, yielding ambiguous identifications.



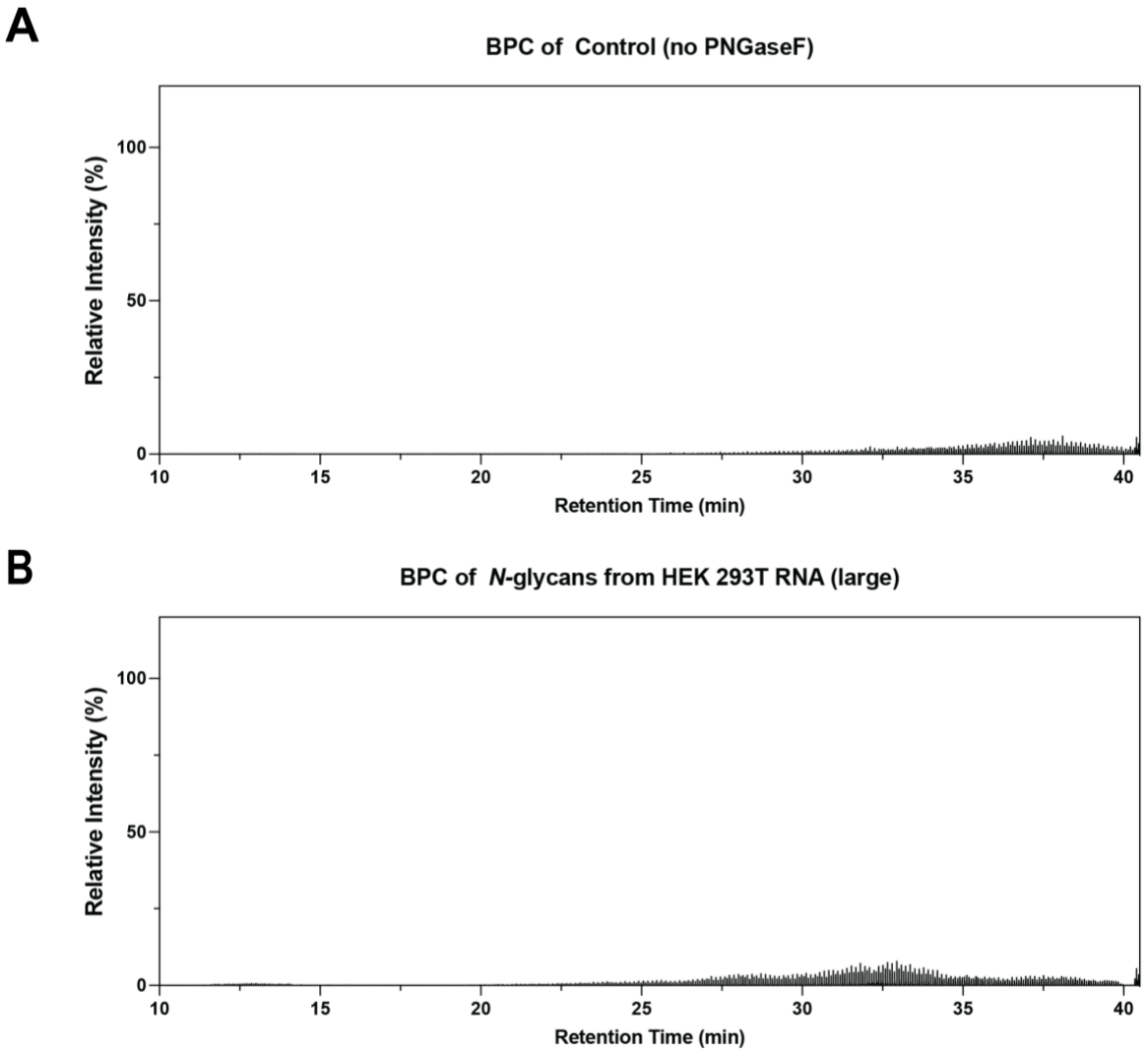
Supplementary Figure 23. The identification of acetyl-modified glycans using GlycanDIA. A. EIC of 657.23 *m/z* representing the signature fragment from Neu5Ac-containing glycans. **B.** EIC of 699.24 *m/z* representing the signature fragment from acetylated Neu5Ac-containing glycans. **C.** The MS2 spectrum at 34.58min corresponding to an acetylated glycan, Hex(5)HexNAc(4)Neu5Ac(1), was shown as an example. The peak observed from the EIC demonstrates the low-abundant modified glycans can be identified using GlycanDIA.



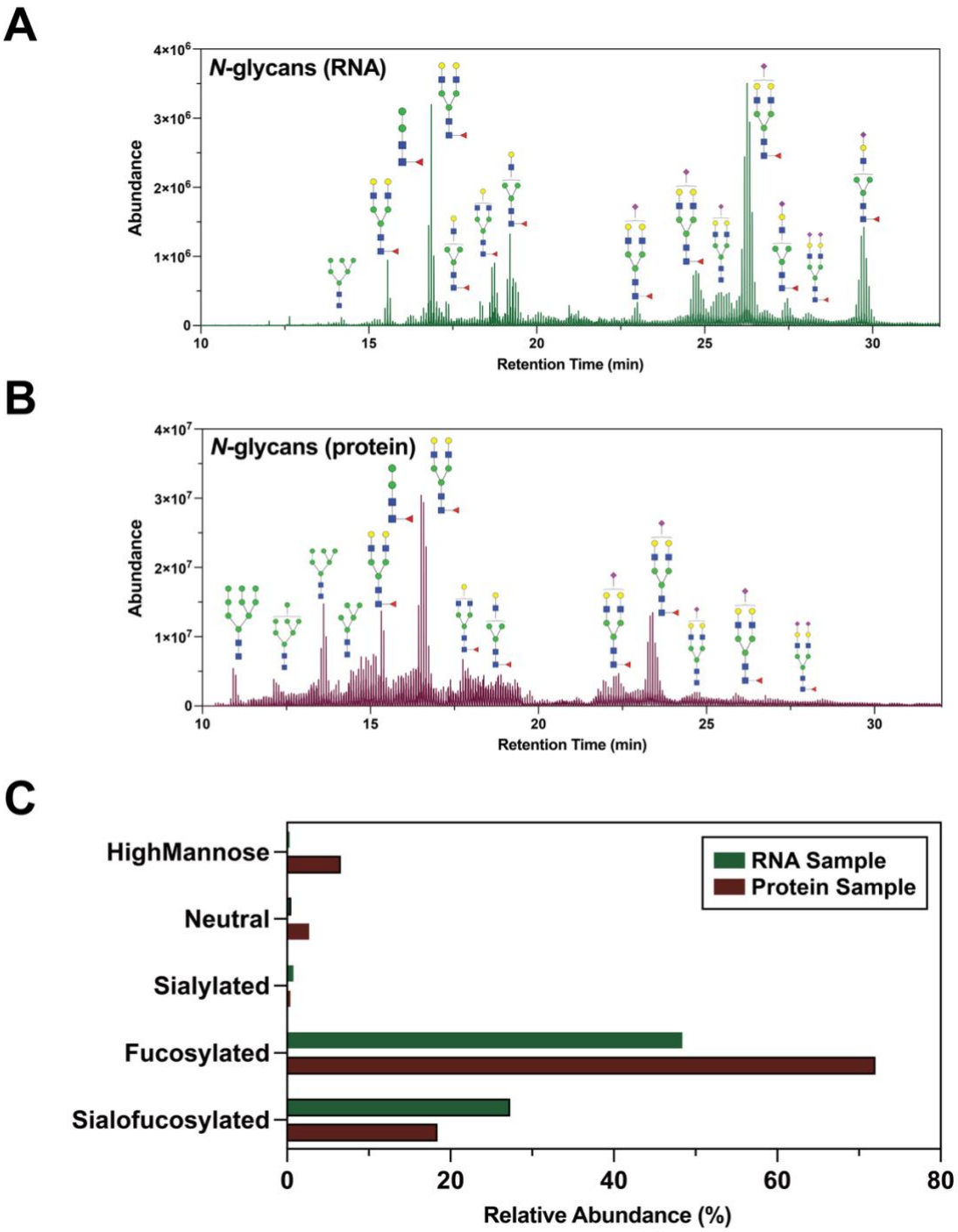
Supplementary Figure 24. The workflow for analyzing *N*-glycans from RNAs using GlycanDIA. RNAs were extracted from cells using Trizol reagents, and small RNAs were separated using an RNA concentrator after the treatment of DNase and protease. The extracted RNAs were subject to the PNGase F treatment to release N-glycans. Glycans were purified using a PGC SPE plate, analyzed using the DIA-based high-resolution mass spectrometry, and identified using GlycanDIA Finder. Created in BioRender. Xie, A. (2025) https://BioRender.com/2tfycxp.



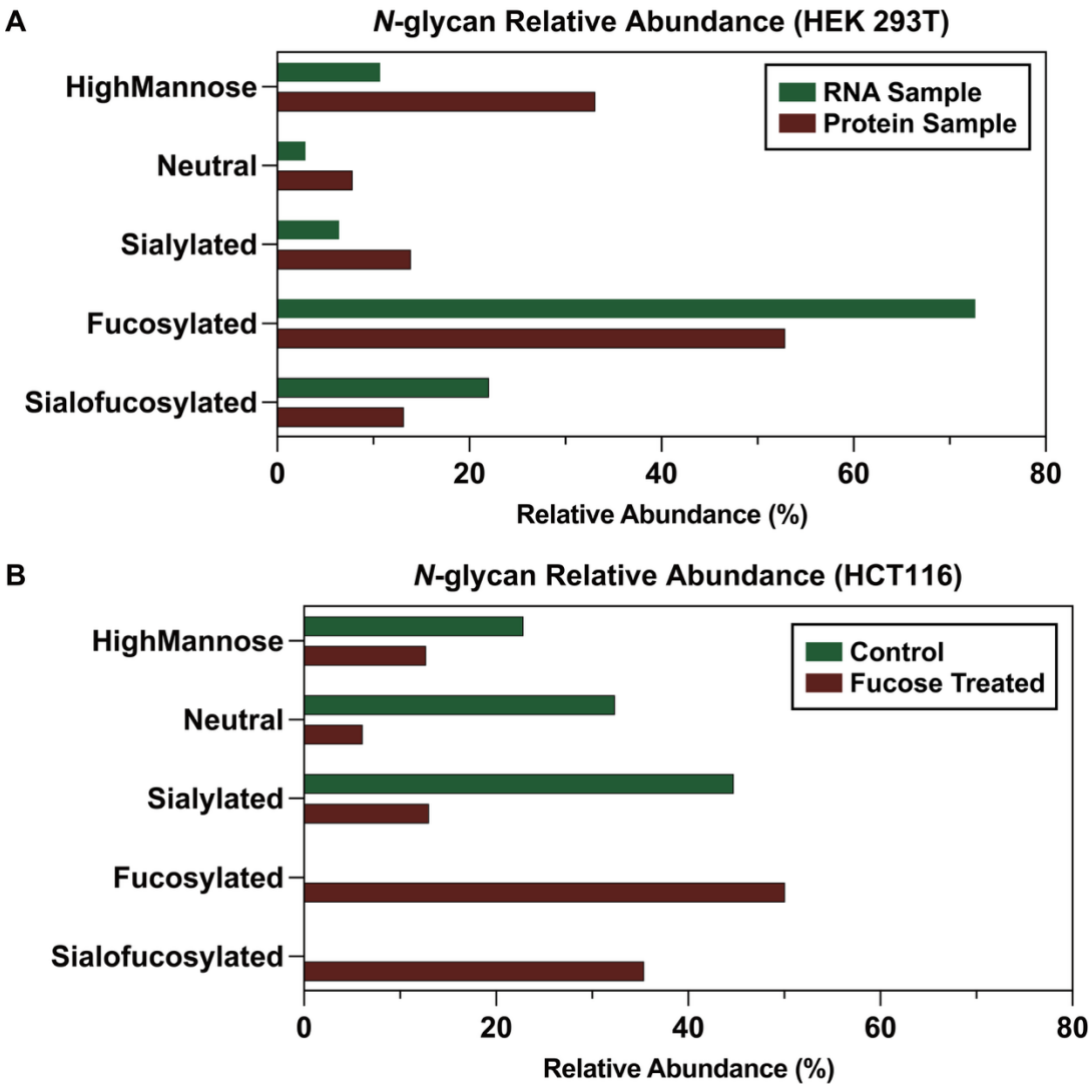
Supplementary Figure 25. The protein and RNA gels of RNA samples. A. protein SDS-PAGE gel of the extracts. No protein band was observed from the RNA samples, confirming the purity of RNA extracts. **B.** RNA TBE-urea gel of the extracts, the bands suggested successful extraction of RNAs and the separation of large and small RNAs.



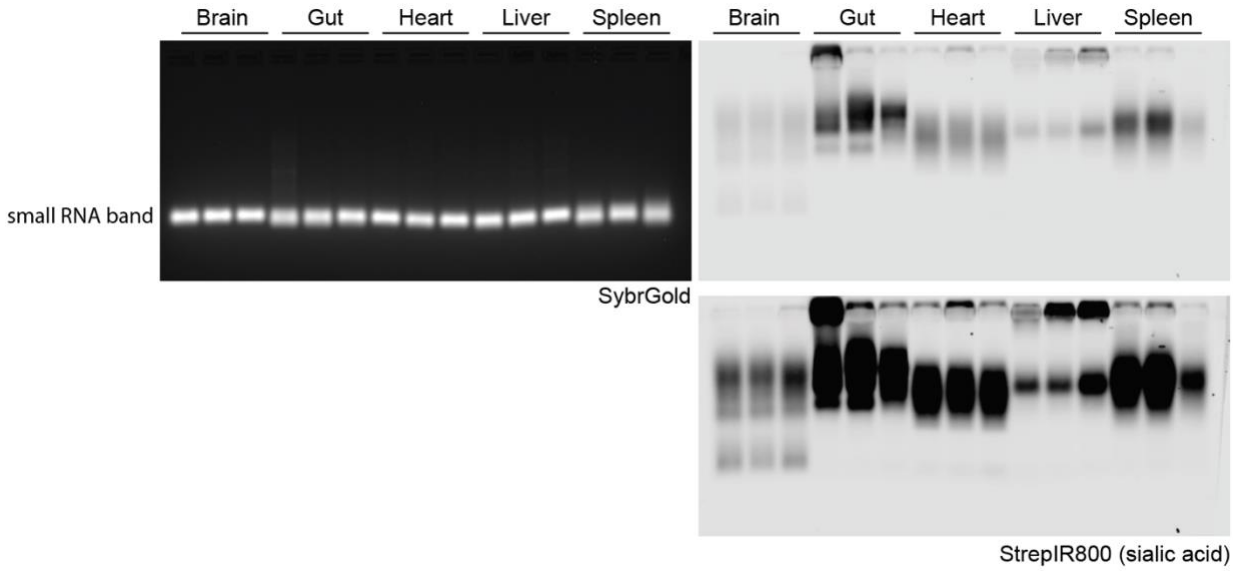
Supplementary Figure 26. The chromatogram of *N*-glycans from HEK 293T cell line. A. The background signal from RNA extracts showed limited free glycans in the sample. **B.** Only background signals were noted from large RNA fractions, suggesting the glycosylation on RNA was barely in large RNAs.



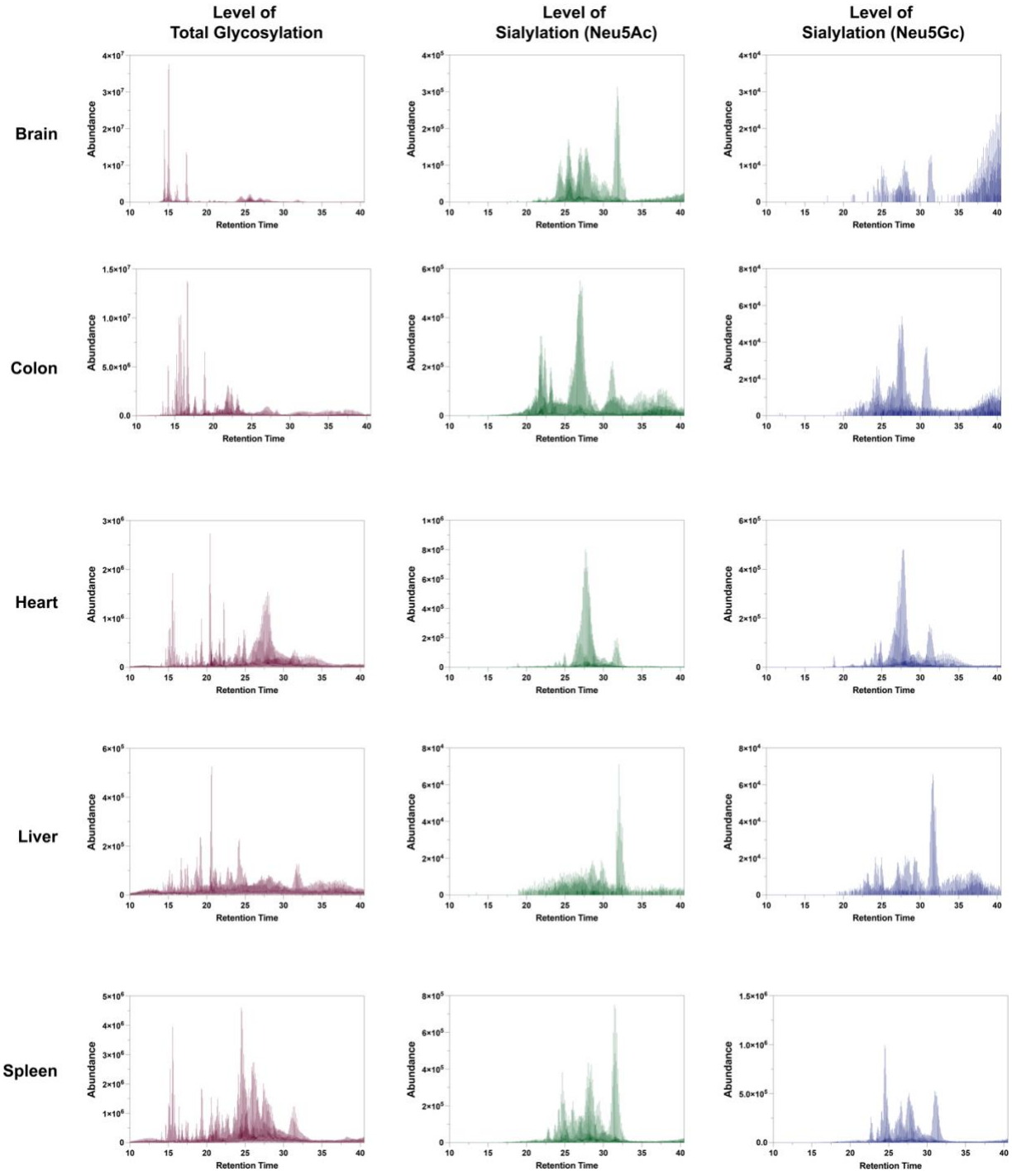
Supplementary Figure 27. The chromatogram of *N*-glycans from HeLa cell line. A. Glycans from protein extracts showed higher abundance in both high mannose type and sialic acid-containing *N*-glycans. **B.** Glycans from small RNA fractions only showed higher abundance in the glycans with sialic acid. **C.** Only background signals were noted from large RNA fractions, suggesting the glycosylation on RNA was mainly in small RNAs. **C.** The relative abundance of RNA and protein *N*-glycans showed distinct glycan profiles. Source data are provided as a Source Data file.



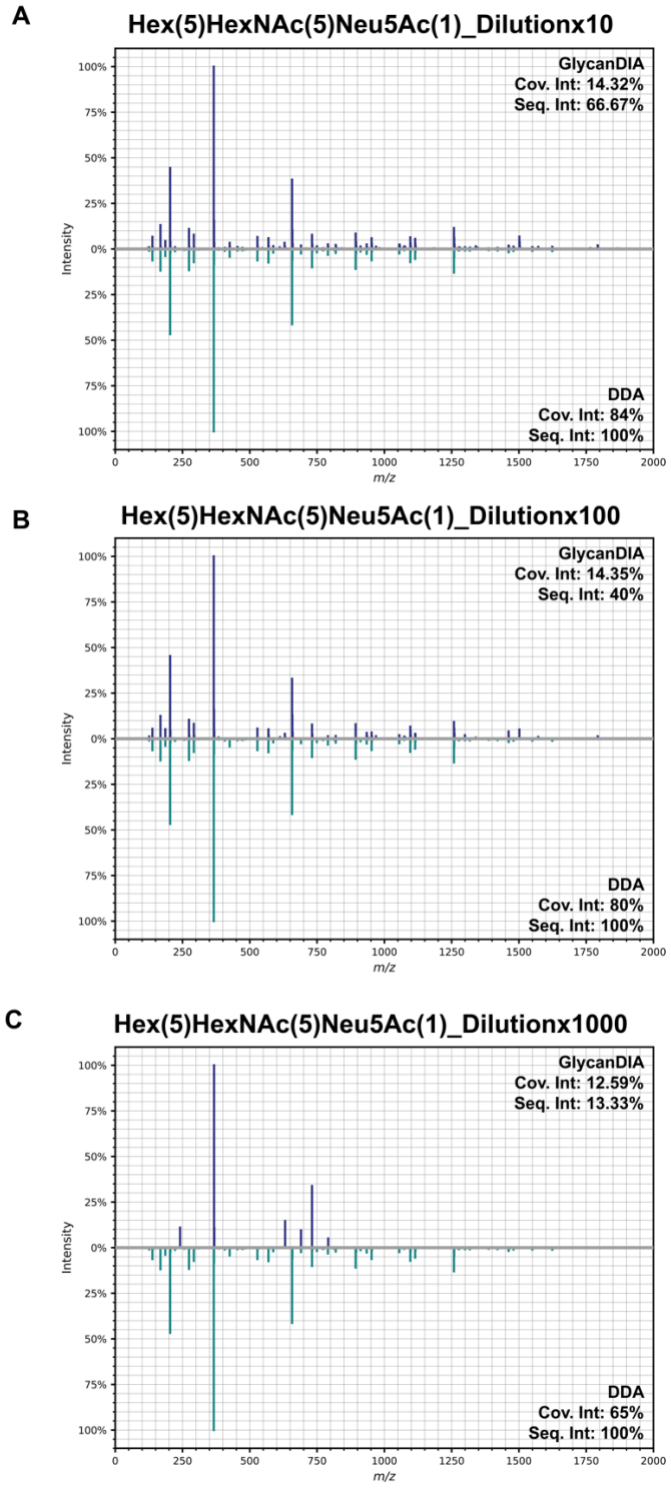
Supplementary Figure 28. The relative abundance of different *N*-glycan subtypes. A. HEK 293T RNA and protein N-glycans abundance **B.** HCT116 N-glycans abundance before and after treatment of fucose. Source data are provided as a Source Data file.



Supplementary Figure 29. RNA blotting of small RNA from indicated mouse tissues labeled to detect the native sialic acids with rPAL. In gel detection of total RNA with SybrGold (left) and on membrane detection of biotin (Streptavidin-IR800, right) is shown. Two exposures of the biotin signal are displayed on the right. Each lane represents one biological replicate.



Supplementary Figure 30. The level of Neu5Ac and Neu5Gc were expressed differently in mouse tissues. The glycan fragments of total glycosylation (red, 204.08 *m/z*), Neu5Ac-containing glycans (green, 292.10 *m/z*), and Neu5Gc-contaning glycans (blue, 308.09 *m/z*).

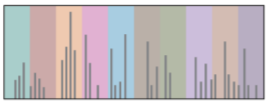
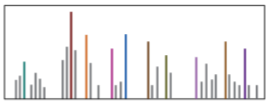
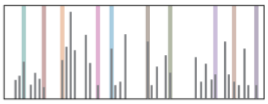


Supplementary Figure 31. Head-to-tail comparison of sialylated glycan, Hex(5)HexNAc(5)Neu5Ac(1) aligned spectra using GlycanDIA method against conventional DDA method. The aligned spectra showed a perfect match to the reference spectra up to 100-fold dilution, while user defined peaks were detected by GlycanDIA even after a 1000-fold dilution with a Cov. Int. 12.6% and Cov. Seq. 13.33%. The robustness of the alignment algorithm also indicates a potential library-free DIA analysis method for future development.

Supplementary Tables.

	48 *m/z* Staggered	36 *m/z* Staggered	24 *m/z* Staggered	Variable
Average Number of Data Points	21	15	10	21
Cycle Time (sec)	2	2.8	4	2
Loop Count	25	34	50	25
Average FWHM (min)	0.31			
NCE	20%			
Precursor Range	600-1800 *m/z*			
Resolution	MS1: 60,000; MS2: 30,000			

Supplementary Table 1. Summary of the details of LC-MS/MS parameters in GlycanDIA workflow.

	 GlycanDIA	 Conventional DDA	 Targeted Analysis
Limited Sample	✓		✓
Easy Setup	✓	✓✓	
High-throughput	✓	✓	
Glycan Identification	✓✓	✓	
MS2-based Quantification	✓		✓
Automated Data Analysis	✓	✓	
Discovery	✓✓	✓	
Data Remining	✓		
Example	**This Manuscipt**	**Chu *et al.*, 2009 She *et al.*, 2020**	**Zhou *et al.*, 2015 Cho *et al.*, 2022**

Supplementary Table 2. Features of MS-based methods in glycomic analysis. References of the example: Chu *et al.*, 9, 7, 1939–1951, *Proteomics* 2009; She *et al.*, 92, 20, 14038–14046, *Anal. Chem.* 2020; Zhou *et al.*, 26, 4, 596-603, *J. Am. Soc. Mass Spectrom.* 2015; Cho *et al.*, 94, 44, 15215–15222, *Anal. Chem.* 2022.

Compound Name	Abbreviation	Vendor	Catalog Number
Human Sera	/	Sigma-Aldrich, MO	S7023
Bovine Submaxillary Mucin (Type I-S)	/	Sigma-Aldrich, MO	M3895
RNase B (Bovine Pancreas)	/	New England Biolabs, MA	P7817S
Lacto-N-difusohexaose I	LNDFH I	Biosynth Carbosynth, CA	OL01664
Lacto-N-difusohexaose II	LNDFH II	Biosynth Carbosynth, CA	OL06826
Lacto-N-fucopentaose II	LNFP II	Biosynth Carbosynth, CA	OL03876
Lacto-N-fucopentaose III	LNFP III	Biosynth Carbosynth, CA	OL05682
2’-Fucosyllactose	2'FL	Biosynth Carbosynth, CA	OF06739
3-Fucosyllactose	3FL	Biosynth Carbosynth, CA	OF05673
3’-Sialyllactose	3'SL	Biosynth Carbosynth, CA	OS04397
6’-Sialyllactose	6'SL	Biosynth Carbosynth, CA	OS04398
Lacto-N-tetraose	LNT	Biosynth Carbosynth, CA	OL162665
Lacto-N-neotetraose	LNnT	Biosynth Carbosynth, CA	OL05765
Sialyl-lacto-N-tetraose a	LSTa	Biosynth Carbosynth, CA	OL58490
Sialyl-lacto-N-tetraose b	LSTb	Biosynth Carbosynth, CA	OL03877
Sialyl-lacto-N-tetraose c	LSTc	Biosynth Carbosynth, CA	OL06570
Lacto-N-neohexaose	LNnH	Biosynth Carbosynth, CA	OL05765
Lactodifucosyllactose	LDFT	Biosynth Carbosynth, CA	OL06567
^{13}C Biantennary G2F	^{13}C-G2F	Asparia Glycomics, Spain	/
^{13}C Paucimannose Man5	^{13}C-Man5	Asparia Glycomics, Spain	/
^{13}C Biantennary G2	^{13}C-G2	Asparia Glycomics, Spain	/

Supplementary Table 3. Product information of standard glycoproteins, *N*-glycans, and HMOs.

Unit Name	Abbreviation	Legend	Type
Mannose	Man		Hexose
Galactose	Gal		Hexose
Glucose	Glu		Hexose
N-Acetylglucosamine	GlcNAc		HexNAc
N-Acetylgalactosamine	GalNAc		HexNAc
Fucose	Fuc		Deoxyhexose
N-Acetylneuraminic acid	Neu5Ac		Sialic acid
N-Glycolylneuraminic acid	Neu5Gc		Sialic acid

Supplementary Table 4. Legends and annotations of glycan units. The legend and annotation followed the rules of Symbol Nomenclature for Glycans (SNFG), and the annotation for glycans in this manuscript is as follows: Hex_HexNAc_Fuc_Neu5Ac_Neu5Gc.

Glycan Name	Precursor	Product 1	Product 2	Product 3	Product 4	Product 5	Product 6
Man5_light	1235.4407	853.2900	690.2399	528.1877	366.1368	204.0849	138.0535
G2_light	821.3034	893.3123	690.2399	528.1877	366.1368	204.0849	138.0535
G2F_light	894.3323	893.3123	690.2399	528.1877	366.1368	204.0849	138.0535
Man5_heavy	1239.4541	855.2964	692.2463	530.1941	368.1432	206.0913	140.0599
G2_heavy	825.3168	897.3251	692.2463	530.1941	368.1432	206.0913	140.0599
G2F_heavy	898.3457	897.3251	692.2463	530.1941	368.1432	206.0913	140.0599

Supplementary Table 5. The precursor and product ions in determining the standard glycan.

Supplementary Methods

Materials.

Sodium carbonate (Na_2CO_3), guanidinium chloride (GuHCl), sodium hydroxide (NaOH), sodium borohydride ($NaBH_4$), ammonium acetate (NH_4OAc), sodium periodate ($NaIO_4$), Magnesium sulfate ($MgSO_4$), sodium sulfite, Tween-20, and mucinase StcE were purchased from Sigma-Aldrich, MO. 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer, SybrGold, Proteinase K, acetonitrile (ACN), trifluoroacetic acid (TFA), and formic acid (FA) were purchased from Thermo Fisher Scientific, MA. The source of standard glycoproteins, *N*-glycans, and HMOs, were listed in **Table S2**.

Animals. All mouse procedures and protocols were approved by the Animal Care and Use Committee of Boston Children's Hospital and followed all relevant guidelines and regulations. C57BL/6 mice were crossed and bred in house. Male C57BL/6 mice (24-28 weeks old) were euthanized, and their liver, spleen, colon, heart, and brain were harvested. The organs were directly added to 1 mL of RNAzol RT (Molecular Research Center, Inc.) in 2 mL tubes containing Zirconium oxide beads (Thomas Scientific). Tissues were homogenized using a Bead mill 24 Homogenizer (Fisher Scientific); 4-8 cycles of 20 seconds on, 10 seconds off at 25°C. After homogenization was complete, samples were stored at -80°C or processed to extract RNA.

Cell lines. Immortalized human embryonic kidney HEK 293T cells, immortalized human cervix HeLa cells, and human colorectal carcinoma HCT116 cells were obtained from American Type Culture Collection (ATCC, Manassas, VA). The cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) medium supplemented with 10% (*v*/*v*) fetal bovine serum (FBS), 1% (*v*/*v*) non-essential amino acids, and 1% (*v*/*v*) GlutaMAX. Cells were maintained in a humidified incubator at 37 °C with 5% CO_2 and subcultured at 80% confluency.

rPAL labeling from mouse tissues. we followed the procedure as described.[38] We started by lyophilizing 1 µL of enzymatically treated small RNA from above. First, we prepared blocking buffer: for each reaction we mixed 1 µL 16 mM mPEG3-Ald (BroadPharm, CA), 15 µL 1 M $MgSO_2$ and 12 µL 1 M NH_4OAc pH 5 (with HCl). 28 µL of the blocking buffer is added to the lyophilized RNA, mixed completely by vertexing, and then incubated for 45 minutes at 35 °C to block. The samples are briefly allowed to cool to room temperature (2-3 minutes), then working quickly, 1

µL 30 mM aminooxy-biotin (Cayman Chemicals, MI) is added first, then 2 µL of 7.5 mM $NaIO_4$ (periodate, stock made in water) is added. The periodate is allowed to perform oxidation for exactly 10 minutes at room temperature in the dark. The periodate is then quenched by adding 3 µL of 22 mM sodium sulfite. The quenching reaction is allowed to proceed for 5 minutes at 25 °C. Both the sodium periodate and sodium sulfite stocks were made fresh within 20 minutes of use. Next, the reactions are moved back to the 35 °C heat block, and the ligation reaction is allowed to occur for 90 minutes. The reaction is then cleaned up using a Zymo-I column. 19 µL of water is added in order to bring the reaction volume to 50 µL, and the Zymo protocol is followed as per the above details. RNA was eluted from the column using 2x 6.2 µL water (final volume approximately 12 µL).

RNA blotting from mouse tissues. In order to visualize the periodate labeled RNA, it is run on a denaturing agarose gel, transferred to a nitrocellulose (NC) membrane, and stained with streptavidin as described.[38] RNA is combined with 12 µL of Gel Loading Buffer II (GLBII, 95% formamide, 18 mM EDTA, 0.025% SDS) with a final concentration of 2X SybrGold and denatured at 55 °C for 10 minutes. It is important to not use GLBII with dyes. Immediately after this incubation, the RNA is placed on ice for at least 2 minutes. The samples are then loaded into a 1% agarose, 0.75% formaldehyde, 1.5X MOPS Buffer (Lonza Bioscience, NC) denaturing gel. Precise and consistent pouring of these gels is critical to ensure a similar thickness of the gel for accurate transfer conditions; we aim for approximately 1 cm thick of solidified gel. RNA is electrophoresed in 1x MOPS at 115V for between 34 or 45 minutes, depending on the length of the gel. Subsequently, the RNA is visualized on a UV gel imager, and excess gel is cut away; leaving ~0.75 cm of gel around the outer edges of sample lanes will improve transfer accuracy. The RNA is transferred using transfer buffer (3 M NaCl solution at pH 1, with HCl) to a NC membrane for 90 minutes at 25 °C. Post transfer, the membrane is rinsed in 1x PBS and dried on Whatman Paper (GE Healthcare, MA). Dried membranes are rehydrated in Intercept Protein-Free Blocking Buffer, TBS (Li-Cor Biosciences, NE), for 30 minutes at 25 °C. After the blocking, the membranes are stained using Streptavidin-IR800 (Li-Cor Biosciences, NE) diluted 1:5,000 in Intercept blocking buffer for 30 minutes at 25 °C. Excess Streptavidin-IR800 was washed from the membranes using three washes with 0.1% Tween-20 in 1x PBS for 3 minutes each at 25 °C. The membranes were

then briefly rinsed with PBS to remove the Tween-20 before scanning. Membranes were scanned on a Li-Cor Odyssey CLx scanner (Li-Cor Biosciences, NE).

Supplementary Note: Instructions for using GlycanDIA Finder

GlycanDIA Finder is an open-source software tool built on Python and MatchMS library to automatically search and align peaks between the MS1 and MS2 data for several compounds in a batch of MS files. To start GlycanDIA Finder, input data is needed, consisting of (1) a set of MS files to be processed iteratively, (2) a .csv form glycan library, including glycan compound name for MS1, notes, add-on mass, and masses for MS2, and (3) a configuration file named **config.ini** including input/output paths, polarity, adduct, charges, mass error, relative height and minimum height of peaks, the time window, minimum matched count, and maximum aligned rcord ms2 to control the behaviors of the software.

The example data and results are available at:

https://github.com/ChenfengZhao/GlycanDIAFinder

Analysis setup:

input_path*:* folder contains MS data files (.mzXML).

output_path: folder to save the GlycanDIA search results (.mzXML)

ms_list_name: glycan library that used for searching

polarity: instrument polarity in MS analysis

max_charge: maximum glycan charge that is calculated for searching

charge_range: (alternatively to max_charge) a range of glycan charges that can be calculated

adduct: adduct in MS analysis

ms1_mass_error_ppm: mass tolerance for MS1 searching

ms2_mass_error_ppm: mass tolerance for MS2 searching

min_rel_height: (optional) relative intensity threshold for glycan searching

min_height: (optional) absolute intensity threshold for glycan searching

min_mass: (optional) minimum *m/z* that is considered for glycan searching

max_mass: (optional) maximum *m/z* that is considered for glycan searching

min_time_min: (optional) minimum retention time that is considered for glycan searching

max_time_min = (optional) maximum retention time that is considered for glycan searching

min_matched_counts: minimum MS2 ions that are required for glycan identification

max_aligned_record_ms2: the number of MS2 ions that are used for quantification
flex_mode: (optional) disable of glycan monoisotopic distribution reqiuremnts

Analysis Workflow:

Step 1 **Input Parsing:** GlycanDIA Finder reads and parses each MS data file for the analysis of all the compounds listed in ms_list.csv. Note that although GlycanDIA Finder sets .mzXML as the default data format of MS data in GlycanDIA Finder, it is easy to be extended to other formats using loading functions provided by MatchMS. Each MS data file is divided into an MS1 file and an MS2 file, which are processed separately in later steps. The MS1 *m/z* are calculated based on compound numbers, add-on masses, abduct, and charges. Note that we can either set the maximum charge to make the software search from charge 1 to the maximum one or define the range of charges to be searched.

Step 2 **MS1 Spectral Analysis:** After reading and parsing the input data files, GlycanDIA Finder calculates the three MS1 *m/z* based on compound numbers, add-on masses, abduct, and charges. Note that we can either set the maximum charge to make the software search from charge 1 to the maximum one or define the range of charges to be searched. If the calculated masses lie in the predefined mass range in config.ini, we will consider these masses as MS1 *m/z* and search these masses in MS1 data. Then GlycanDIA Finder calculates ion distributions based on MS1 masses and predefined MS1 mass error to check the existence of MS1 masses in a spectrum. By default, if all three masses exist in the same spectrum (i.e., the difference between the MS1 mass and the *m/z* of a peak in the spectrum is smaller than the margin), we extract the retention time and scan number of the spectrum, and intensity of the matched peaks to obtain the RT-intensity curves for all the three MS1 masses.
To find the outstanding peaks of the RT-intensity curve of the first MS1 mass, GlycanDIA Finder first performs a Gaussian filter and records all the possible peaks regardless of absolute intensity. Then it calculates the surrounding baseline for every peak (i.e., the maximum of the left and right baseline) with a prominence calculation algorithm. If the signal-to-noise ratio of a compound peak is less than a predefined threshold (2 by default), then the peak is regarded as

trivial and will be filtered. Additionally, GlycanDIA Finder checks the RT of two adjacent peaks as well to filter the peaks getting too close to each other. Users can also define the minimum absolute intensity and normalized maximum threshold to filter peaks further with more flexibility.

Step 3 **MS2 Spectral Analysis:** GlycanDIA Finder finds the nearest precursor to the first MS1 *m/z* and extracts all the spectra of the same precursor from the MS2 data. Similar to MS1 spectral analysis, we count the number of masses that exist in the same spectrum. If it is greater than the predefined minimum matched count, the retention time and scan number of the spectrum and the intensity of the matched peaks are recorded to obtain the RT-intensity curve for each MS2 *m/z*. GlycanDIA Finder finds peaks for each MS2 RT-intensity curve in the same way as MS1.

Step 4 **Peak Analysis**: GlycanDIA Finder counts the number of aligned MS2 peaks for each MS1 peak by checking the difference between their scan numbers. The results, including the intensity of peaks, matched count, etc., are output as .csv forms and then combined for all the compounds and datasets.

Equation:
The target glycan database defined by the user:

$\{(G_1)i_1,(G_2)i_2,\ldots,(G_k)i_k\}$

$\boldsymbol{G_1}, \ldots, \boldsymbol{G}_k$ represents a type of fragmentation, $\boldsymbol{i_1}, \ldots, \boldsymbol{i}_k$ represents maximum number of peaks.

The aligned MS1 and MS2 by GlycanDIAfinder are:

$$A_{align} = \begin{bmatrix} M1 & I1 \\ \ldots & \ldots \\ M_a & I_a \end{bmatrix}$$

M is the m/z of a fragment ion and I is the intensity of each fragment ion.
The two scores, Cov. Int and Cov. Seq are calculated using number of *n* matched peaks to the user define database. We aim to demonstrate that the Coverage of Intensity (Cov. Int) is derived from the summation of intensities of aligned MS2 peaks. In contrast, the Coverage of Sequence (Cov. Seq) is determined by the total count of user-defined library fragments. As a

result, it's important to note that the relationship between these two scores is not necessarily linear.

$$A_{match} = \begin{bmatrix} M1 & I1 \\ \dots & \dots \\ M_m & I_m \end{bmatrix}$$

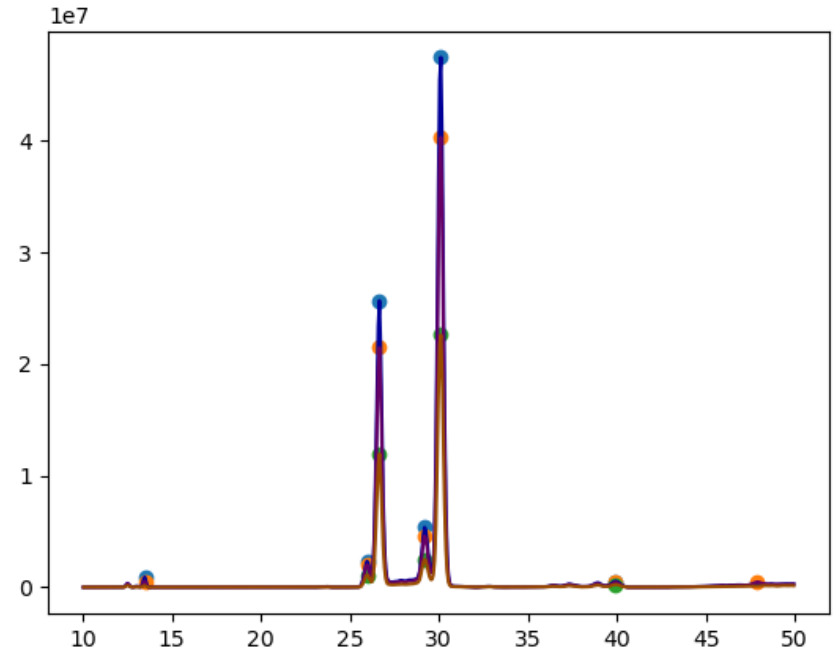
$$Cov_{int} = \frac{\sum_{t=1}^{m} I_t}{\sum_{t=1}^{a} I_t}$$

$$Cov_{seq} = \frac{m}{i_k}$$

Result files:

GlycanDIA Finder generates files, including the figures to visualize identified glycan peaks and .csv files contain the intensity of different glycan isomers, glycan compositions, and glycan subtypes.

Example MS1 peak:



Example MS2 peak: