Non-targeted isomer-sensitive N-glycome analysis reveals new layers of organ-specific diversity in mice.

Johannes Helm[1*], Stefan Mereiter[2,3*], Tiago Oliveira[2,3], Anna Gattinger[3,4], David M. Markovitz[5], Josef M. Penninger[2,3,6,7], Friedrich Altmann[1], Johannes Stadlmann[1,8].

Affiliations

1. Institute of Biochemistry, Department of Chemistry, University of Natural Resources and Life Sciences (BOKU), Muthgasse 18, 1190 Vienna, Austria.
2. Eric Kandel Institute, Department of Laboratory Medicine, Medical University of Vienna, Spitalgasse 23, 1090 Vienna, Austria.
3. Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna BioCenter (VBC), Dr. Bohr-Gasse 3, 1030 Vienna, Austria.
4. Bioinformatics Research Group, University of Applied Sciences Upper Austria, Softwarepark11, 4232 Hagenberg, Austria.
5. Division of Infectious Diseases, Department of Internal Medicine, and the Programs in Immunology, Cellular and Molecular Biology, and Cancer Biology, University of Michigan, Ann Arbor, MI 48109, USA.
6. Department of Medical Genetics, Life Sciences Institute, University of British Columbia, Vancouver Campus, 2350 Health Sciences Mall, Vancouver, BC Canada V6T 1Z3.
7. Helmholtz Centre for Infection Research, Braunschweig, Germany
8. BOKU Core Facility Mass Spectrometry, University of Natural Resources and Life Sciences (BOKU), Muthgasse 18, 1190 Vienna, Austria.

*These authors contributed equally

Correspondence

Johannes Stadlmann

Institute of Biochemistry, Department of Chemistry,

University of Natural Resources and Life Sciences (BOKU)

Muthgasse 18, A-1190 Vienna, Austria.

email: j.stadlmann@boku.ac.at

Summary

N-glycosylation is one of the most common protein modifications in eukaryotes, with immense importance at the molecular, cellular, and organismal level. Accurate and reliable N-glycan analysis is essential to obtain a systems-wide understanding of fundamental biological processes. Due to the structural complexity of glycans, their analysis is still highly challenging. Here we make publicly available a consistent N-glycome dataset of 20 different mouse tissues and demonstrate a multimodal data analysis workflow that allows for unprecedented depth and coverage of N-glycome features. This highly scalable, LC-MS/MS data-driven method integrates the automated identification of N-glycan spectra, the application of non-targeted N-glycome profiling strategies and the isomer-sensitive analysis of glycan structures. Our delineation of critical sub-structural determinants and glycan isomers across the mouse N-glycome uncovered tissue-specific glycosylation patterns, the expression of non-canonical N-glycan structures and highlights multiple layers of N-glycome complexity that derive from organ-specific regulations of glycobiological pathways.

Introduction

Protein glycosylation, the covalent attachment of simple or complex sugar structures to amino-acid side chains of polypeptides, affects virtually all aspects of biology. Over 50% of all human proteins are subject to post-translational modifications by glycans[1], which alter their functions in fundamental biological processes, such as cell adhesion, signal transduction, intracellular trafficking, essential immune functions[2] or host-pathogen interactions[3].

In the process of mammalian protein glycosylation, glycan structures are co- or post-translationally linked to nascent polypeptide chains, which then become extensively processed along the secretory pathway, throughout the Golgi-network to the cell surface. In protein N-glycosylation, arguably the currently best understood form of protein glycosylation, a large, highly conserved N-glycan precursor structure is *en bloc* transferred and covalently linked to the side chains of specific asparagine-residues within the endoplasmic reticulum (ER). Subsequent N-glycan processing is a non-template driven process which critically depends on the coordinated action of dedicated enzymes, a highly specific set of glycosyltransferases and glycosidases[1]. Perturbations in any of these tightly interconnected and subtly tuned cellular processes hold the potential to result in differentially processed glycoproteins that exhibit aberrant glycan structures.

Owing to its close resemblance to the human physiology, its short generation time and its visible phenotypic variants, the mouse (*Mus musculus*) represents the most common mammalian model organism to study fundamental biological processes, including glycosylation[4]. Nevertheless, striking differences between human and murine glycosylation have been described. Most importantly, humans lack the Gal-α1,3-Gal epitope (due to an inactivation mutation in the α1,3-galactosyl transferase (*Ggta1*)[5]) and the biosynthetic capabilities for the generation of N-glycolylneuraminic acid (Neu5Gc) due to a mutation in the CMP-NeuAc hydroxylase[6] (*Cmah*). Further differences between human and murine glycosylation have recently been reported for the brain[7] and immune cells[8].

The comprehensive analysis of protein-linked N-glycan structures is a highly challenging task. Due to compositional, structural, positional, and anomeric isomers[9], this analytical challenge has long been tackled using a variety of glycomics techniques, which all profile different aspects of glycosylation and all come with specific analytical advantages and disadvantages. Importantly however, the use of this wide range of different methodologies, including lectin-arrays[10], MALDI-TOF mass-spectrometry (MS) analysis of native[11] , permethylated[12] or otherwise derivatized glycans[13,14], fluorescence labelling of glycans followed by either reversed- or normal-phase HPLC analysis[15,16], various LC-MS and LC-tandem MS (MS/MS) analysis workflows[17,18], causes low comparability between studies.

Generation of glycobiological ground truth was pioneered by the Consortium of Functional Glycomics glycan profiling initiative (CFG, www.functionalglycomics.org) who generated, annotated, and made publicly available the hitherto most comprehensive, consistent, and coherent mammalian glycome datasets. From 2001 to 2011, the CFG Analytical Glycotechnology Core analyzed the compositions of N- and O-linked glycans from 11 human and 16 murine tissues, primarily using MALDI-TOF MS analysis. While the CFG's analytical activities have come to an end, their seminal glycan profiling datasets still represent a vital and fundamental resource to the field of mammalian glycobiology. More recently, new glyco-analytical developments and the paucity in modern, comprehensive N-glycome datasets prompted several studies aiming at the systemic profiling of tissue-specific N-glycosylation patterns using MALDI-TOF MS[19–24]. Despite the remarkable compositional variations that were detected by such ("single-stage") MS approaches, these methods intrinsically fall short in capturing the unique level of N-glycan structure micro-heterogeneity. Instead, glycan structures are tentatively deduced from composition and based on pre-established knowledge of glyco-biosynthetic pathways.

Here, we mapped the N-glycomes of 20 mouse tissues by performing PGC-LC-MS/MS (porous graphitic carbon liquid-chromatography coupled to a tandem mass-spectrometer). We used PGC-LC to chromatographically separate closely related N-glycan structure isomers and an Orbitrap mass-analyzer for high resolution MS/MS data-acquisition. Conventionally, the analysis of such LC-MS(/MS)-N-glycome datasets is based on the targeted, selective extraction for known or anticipated glycan-compositions and -masses [20,22,25]. Subsequent glycan-structure assignment incorporates additional information on pre-established, relative chromatographic retention times[9,26,27], complementary MS/MS data and expert knowledge on glycan biosynthetic pathways. Despite the breath of glycan-information that can be retrieved by such -largely manual- approaches[9,25,27–30], they are hardly scalable to the size of the present dataset. Instead, we developed an automated, non-targeted and scalable MS/MS-centric N-glycomics data-analysis workflow, largely independent of prior glycobiological knowledge and anticipated glycan-compositions. We showcase this approach by characterizing 20 different mouse tissues, providing a consistent and complete N-glycome atlas. Our analyses reveal tissue specific N-glycome signatures and glycan-structural features, highlighting organ-intrinsic regulations of glycobiological pathways.

Results

Precursor-independent MS/MS-based N-glycome profiling of 20 mouse tissues

To generate a consistent N-glycome dataset, we collected tissues, as well as serum, in duplicates from age-matched C57BL/6J mice, and processed all samples using identical protein extraction, enzymatic N-glycan release (i.e. PNGase F), chemical reduction and glycan clean-up protocols. All samples were analyzed using porous graphitic carbon liquid-chromatography (PGC-LC) coupled to a high-resolution tandem (MS/MS) mass-spectrometer (i.e. Orbitrap Exploris 480).

To survey this data, we first assessed the MS/MS data for the occurrence of diagnostic, N-glycan derived fragment ions, independent of intact glycan precursor mass information. For this, we automatically extracted all MS/MS spectra and retained only those that contained N-glycan specific fragment ions (i.e. 224.1 amu, diagnostic for reduced N-acetylhexosamine). This first breakdown of our dataset confirmed that, overall, more than 40 percent of all MS/MS spectra generated in this study (i.e. 220506 out of 509283 MS/MS spectra) contained information on chemically reduced glycan precursor ions. Importantly, however, we also found considerable differences in the relative proportion of N-glycan-derived MS/MS spectra between tissues, ranging from approx. 20% for ileum to more than 80% for seminal vesicle (SI Fig. 1.A). This observation highlighted important, tissue-dependent differences in the dynamic range and the structural complexity of the individual N-glycomes and prompted us to further explore tissue-specific N-glycome features as sialylation or fucosylation, independent of intact glycan precursor information.

Sialylation is essential to mammalian development[31] and results from the attachment of either N-acetyl-neuraminic acid (Neu5Ac) or N-glycolyl-neuraminic acid (Neu5Gc). To dissect this important compositional heterogeneity of N-glycans, we extended our initial MS/MS data-filtering criteria to sialylation-specific diagnostic fragment ions (i.e. Neu5Ac fragment ion mass 292.1 amu, Neu5Gc fragment ion mass 308.1 amu) and counted the number of N-glycan derived MS/MS spectra tissue specifically. As expected, the number of Neu5Ac- and/or Neu5Gc-containing spectra showed substantial differences between tissues, ranging from approx. 80% in serum, lung, and heart, to only approx. 18% in seminal vesicle (Fig. 1., Supplementary data). Differential analysis of the numbers of Neu5Ac- versus Neu5Gc-containing MS/MS spectra, on the other hand, showed that most tissues presented with comparable levels for both sialic acid variants (SI Fig. 1.B) (mean ratio Neu5Gc:Neu5Ac = 1.5), except for murine serum which is vastly dominated by Neu5Gc (Neu5Gc:Neu5Ac = 11.2), as well as colon (Neu5Gc:Neu5Ac = 0.3), and brain (Neu5Gc:Neu5Ac = 0.1), which both presented with particularly high levels of Neu5Ac. This data corroborated previous reports on the murine brain N-sialome being vastly dominated by Neu5Ac and to only contain trace amounts of Neu5Gc[28,32,33]. In general, heart, kidney, liver, lung, mammary gland, serum, skin, spleen, testis, and thymus exhibited a higher numbers of sialic acid containing spectra compared to the mean value across all tissues, while bladder, brain, colon, duodenum, ileum, jejunum, lymph node, seminal vesicle, pancreas, and white adipose tissue resulted in lower numbers of sialic acid containing MS/MS spectra (SI Fig. 2.A, 2.B, and 2.C).

Similar to sialylation, fucosylation is a vital, developmentally controlled modification of N-glycans, which has been implicated in numerous cell-cell interactions[34] and massively expands the structural heterogeneity of the N-glycome. Fucose can be linked to the most proximal core-GlcNAc residue, originally connected to the protein (i.e. core-fucose). Additionally, fucose residues have been found linked to either galactoses or N-acetylhexosamines in multiple positions of the distal parts of N-glycan antenna (i.e. distal fucosylation). Interestingly, distal fucosylation also encodes a series of essential,

immune-reactive glycan-epitope isomers, comprising one (e.g. Lewis X, sialyl Lewis X, Lewis A, sialyl Lewis A, blood-group H type 1 and type 2) or more fucose residues (e.g. Lewis Y, Lewis B), in humans. Fucosyltransferase 3 (FUT3), the enzyme which catalyzes the addition of the Lewis A and Lewis B epitope in humans was found to be a pseudogene in mice. The murine system is thus believed to lack these two important fuco-epitopes[35,36]. Unfortunately, most of these fucosylated N-glycan isomers cannot be resolved by MS/MS alone [37–39]. As a consequence, our precursor independent MS/MS profiling approach, which is based on fragment ions that are indicative[37–39] of fucosylated N-glycans (i.e. one fucose linked to one HexNAc and to one hexose; fragment ion mass = 512.2 amu), merely reflects on the combined expression patterns of core fucose, Lewis X, sialyl Lewis X, Lewis Y, blood-group H (bgH) type 1 and type 2, but not Lewis A or Lewis B, in murine tissues. Our analysis showed that fucosylated N-glycans were indeed present in all tissues, with exceptionally high levels in seminal vesical (85%)[40], kidney (65%), and brain (50%) (Fig 1., SI Fig. 2.E). This suggested generally high expression levels of fucosyltransferases in these tissues, such as *Fut9*, which is highly expressed in kidney and brain[41], *Fut2* and *Fut4*, which are highly expressed in colon epithelial cells, or *Fut8*, which is globally expressed in mouse[42].

Another distinctive structural feature of N-glycans is bisecting N-acetylglucosamine (GlcNAc). Bisecting GlcNAc residues are β1,4-linked to the core β-mannose residue by MGAT3 and have been reported vital to fetal development[43], immunity, and cell adhesion[44]. Expression of this critical structural modification of N-glycans was profiled across all tissues, based on the diagnostic fragment ion of mass 792.3 amu (i.e. one reduced GlcNAc-residue, two GlcNAc-residues, and one hexose residue). As expected, we found bisecting GlcNAc expression levels to be highly tissue specific, with the highest expression levels in brain (4%), kidney (2.5%), and colon (2%) (Fig. 1., SI Fig. 2.D), all in good agreement with moue *Mgat3* gene expression data[45].

Moreover, an extensive screening of all tissues was conducted to identify fragment ions indicative of the Sda antigen. This antigen is characterized by a Neu5Ac residue α2,3-linked and a GalNAc residue β1,4-linked to the galactose within a LacNAc motif (860.3 amu)[46]. Intriguingly, this specific structural modification was predominantly expressed in the colon (9%), followed by the jejunum (3.5%) and duodenum (1%) (SI Fig. 1.C). Notably, the ileum exhibited only minimal expression, accounting for less than 0.25% of all N-glycan associated MS/MS spectra. In previous studies, the Sda antigen has been identified in the colon of healthy humans, on the Tamm-Horsfall glycoprotein of Sda^+ individuals, and in the serum of patients with gastric cancer[46]. In humans, the biosynthesis of this antigen is driven by

the *B4GALNT2* gene, which orthologs' (*B4Galnt2*) expression indeed appears to be restricted to the intestinal organs in mice[42].

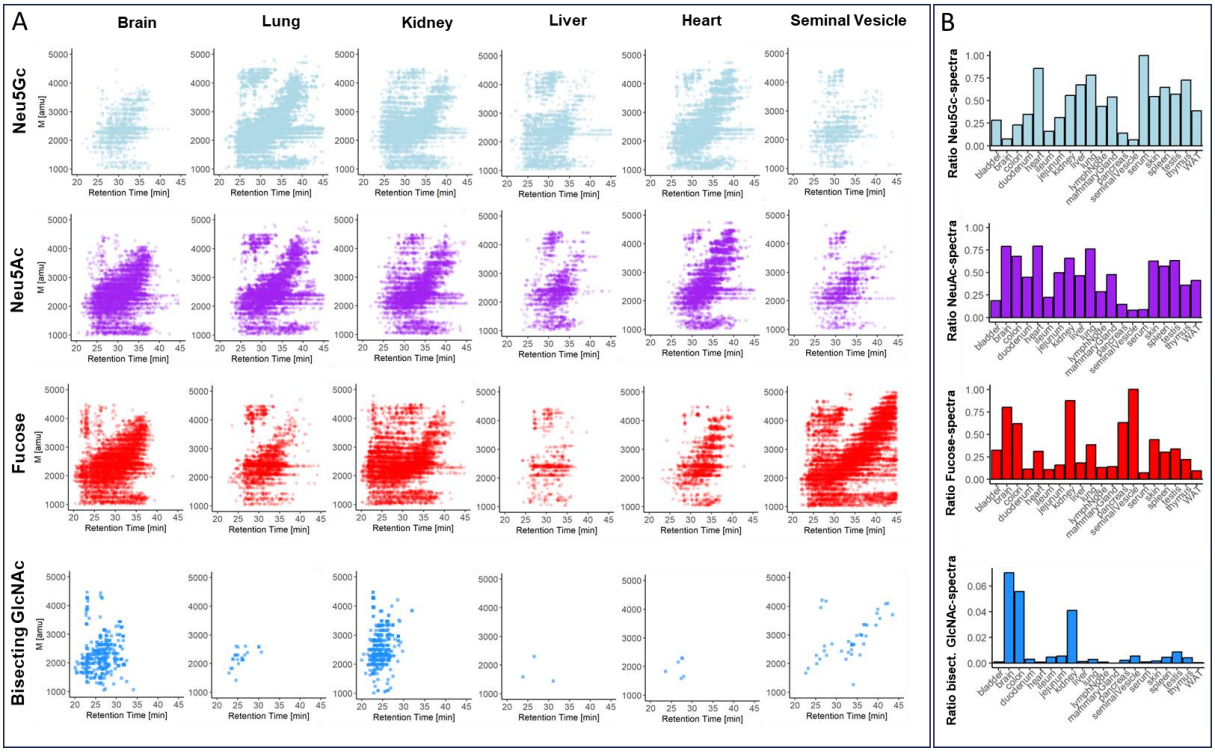

Figure 1. Precursor-independent MS/MS-based N-glycome profiling. ***(A)*** *Precursor mass over chromatographic retention time plots MS/MS spectra. Each data point represents a single N-glycan derived MS/MS spectrum (i.e containing diagnostic fragment ions at 224.1 amu and one of the following signals: 274.1 amu for Neu5Ac-sialylation, 290.1 amu for Neu5Gc-sialylation, 512.2 amu for fucosylation, and 792.3 amu for bisecting GlcNAc).* ***(B)*** *Relative spectral count. Data were normalized to the total number of N-glycan derived MS/MS spectra (i.e. containing 224.1 amu) in the respective tissues. "WAT" - white adipose tissue.*

Automated MS/MS-data-driven reconstruction of the mouse N-glycome.

Next, we asked which N-glycan precursors contributed to the N-glycome signatures revealed by our precursor independent MS/MS-based profiling. For this we automatically reconstructed the N-glycosylation patterns of the different mouse tissues at the precursor level, by implementing a data-aggregation workflow that efficiently reduces PGC-LC-MS data complexity, leverages MS/MS information to identify N-glycan precursors, and extracts quantitative precursor information.

The extraction of precursor mass information from LC-MS/MS data is often complicated by the imperfection of mass-spectrometric data, due to e.g. incorrect mono-isotopic peak-picking or charge-state assignments by the mass-spectrometer, unintended precursor ion co-isolation, in-source fragmentation, or excessive adduct-ion formation. To control for such experimental errors and imprecisions, we post-analysis charge-deconvoluted and deisotoped all MS spectra (using DeCon2[47], Supplemental Information), assigned the precursor mass and intensity values to mass bins (i.e. range of 1000 - 4000 Da; bin-width = ±0.05 amu), and summed the individual precursor mass-to-intensity values

across the entire chromatographic time-range (i.e. retention time 0 – 50 min.), using custom code. This transformation of time-resolved LC-MS data into two-dimensional precursor mass-to-intensity arrays not only improved the robustness of mono-isotopic precursor mass assignments, but also efficiently reduced the dimensionality of our dataset. Additionally, these data-arrays allowed for the construction of tissue-specific, quantitative histograms (Fig. 2.A, Supplementary data). These histograms closely resemble the MALDI-TOF MS-spectra frequently communicated in the fields of glycomics, and thus provide a convenient data-visualization format, that seamlessly integrates with current glycome data-repositories (e.g. CFG, www.functionalglycomics.org; Fig. 2.A).

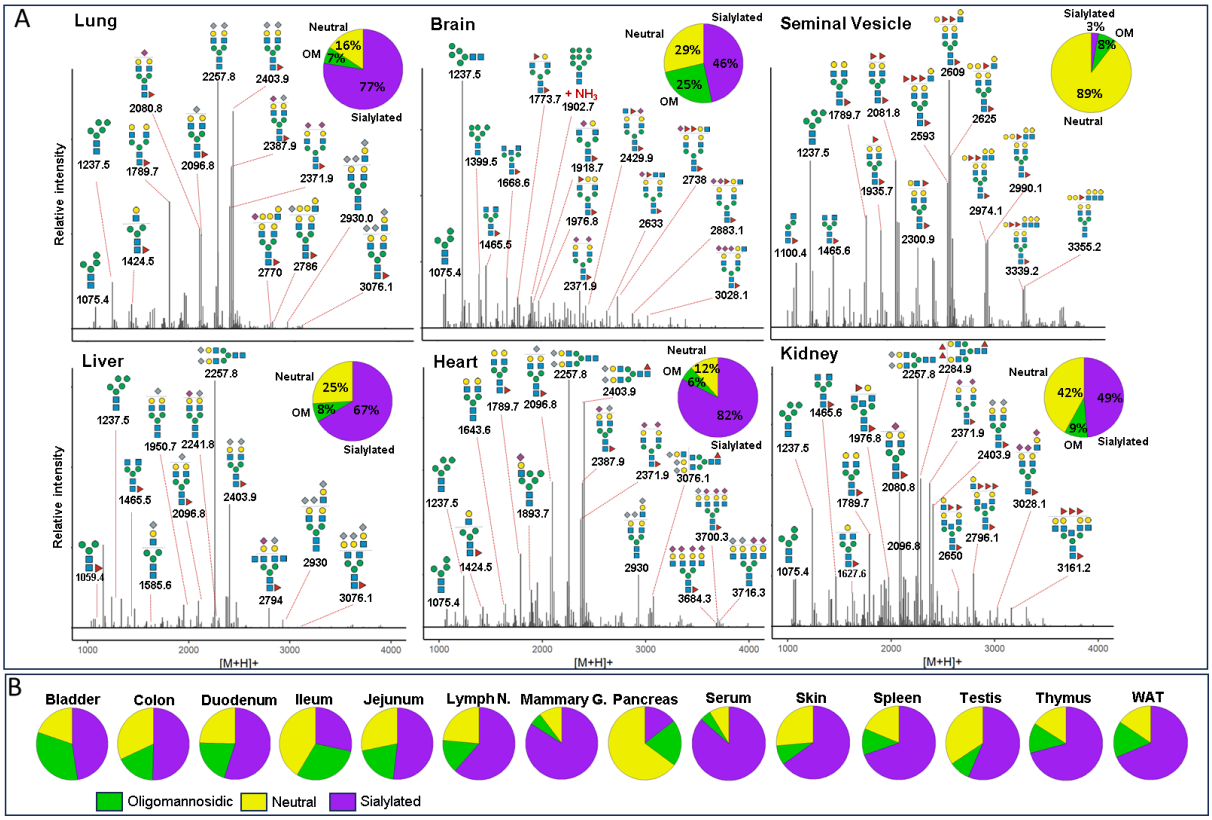

Figure 2. Comparative N-glycome analysis. (A) *Semi-quantitative histograms of SNOG-filtered LC-MS data of selected tissues. LC-MS data were charge-deconvoluted, deisotoped, and summed across the chromatographic time-range using custom code. Precursor mass-to-intensity arrays are shown in the range of 1000 - 4000 Da (bin-width = ±0.05 amu). Representative, tissue-specific N-glycan compositions are indicated by their precursor mass (i.e. $[M+H]^+$ in amu) and tentative structure assignments using the „Symbol Nomenclature for Glycans" (SNFG). Signals below a cumulative intensity threshold of 5E+6 and a SNOG-score of 0.03 were removed. Inset pie charts show the N-glycan compositional fractions of the respective N-glycome (green: oligomannosidic- (OM), purple: sialylated (including both Neu5Ac and Neu5Gc-sialylated glycans), yellow: undecorated/neutral N-glycans).* ***(B)*** *N-glycan compositions of tissues not included in Panel A. "WAT"- white adipose tissue, "Lymph N" - lymph nodes, "Mammary G." - mammary gland.*

In a next step, raw MS/MS data were refined (i.e. mono-isotopic peak picking and charge state assignment re-evaluation) and converted into the generic .mgf file-format using the proteomics software PEAKS[48]. From this, to identify MS/MS spectra that derived from N-glycan precursors and to stringently control for unintended precursor ion co-isolation events, we calculated spectrum-specific

Score-values (i.e. SNOG-score) from the intensity of an N- (and O-) Glycan-specific fragment ion (i.e. oxonium ion of the reduced-end monosaccharide GlcNAc; $[M+H]^+$ = 224.1 amu), using custom code (Supplemental data). Empirically, we determined that MS/MS spectra of SNOG-score greater than 0.03 derived from actual N-glycan precursors (SI Fig. 3., SI Fig. 4., and SI Fig. 5.). Eventually, we aligned the precursor mass information of SNOG-scored MS/MS spectra with our initial two-dimensional, MS precursor mass-to-intensity arrays, and only retained MS signals of true N-glycan precursors above a cumulative intensity threshold of 5E+6.

The number of automatically identified glycan precursor masses greatly varied between tissues. For example, while SNOG-filtering reduced the total number of glycan-derived precursor mass-bins by approx. 30% in brain, it removed approx. 75 % of all precursor mass-bins in liver (SI Fig. 4.B). This, again, highlighted important tissue-dependent differences in the dynamic range and the structural complexity of N-glycomes, and suggested sample-specific background signals at the precursor level. Manual inspection of MS/MS spectra of low-scoring mass-bins (i.e. rejected) confirmed, for example, exceptionally high levels of non-N-glycan signals that derived from hexose-oligomers (e.g. dextrans or maltodextrins degradation products of glycogen[49]) in liver, which were efficiently removed by our approach (SI Fig. 5.).

The correlation and cluster analysis of our SNOG-filtered dataset (Fig. 3.) validated robust reproducibility among sample replicates and revealed distinctive tissue-specific clustering patterns. Notably, we found clustering even among distantly related mouse tissues, such as the exocrine organs, i.e. seminal vesicles and pancreas, clustering with brain, or the central organs of the immune system, spleen and thymus, coalescing with mammary glands, for which we observed convergent N-glycan signatures. Intriguingly, while spleen and thymus share a cluster, lymph nodes exhibit a glycosylation pattern more akin to white adipose tissue, potentially influenced by spatial/histological proximity.

It is noteworthy that a distinctive glycosylation pattern sets apart the tissues of the small intestine, particularly the jejunum and duodenum, constituting a distinct cluster that is separate from the ileum and colon. The duodenum and jejunum, positioned directly after the stomach in the digestive tract, showcase a unique glycosylation pattern, presenting a contrast to the sequential arrangement observed in the ileum and colon. Conversely, the ileum and colon form a cluster with the bladder and skin, suggesting a potential association rooted in shared epithelium glycosylation patterns.

Additionally, our observations extend to the highly blood-perfused organs, where the liver and lung share a cluster with serum, while heart, kidney, and testis form another distinct cluster. This clustering pattern provides insights into the glycosylation variations within these organ groups, potentially reflecting their functional relationships or physiological roles.

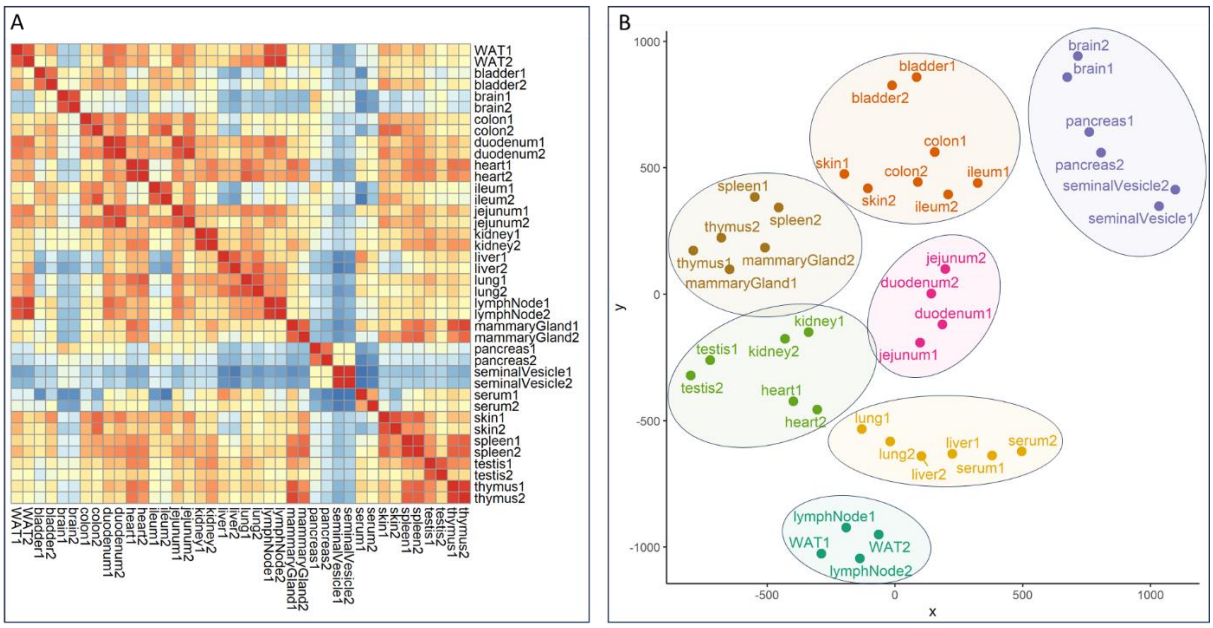

Figure 3. Correlation and cluster analysis of SNOG-filtered LC-MS data. (A) *Correlation heatmap was calculated using Pearson´s correlation coefficients.* ***(B)*** *t-SNE plot. Hierarchical clustering was based on Euclidean distances of the calculated t-SNE values. Clustering was calculated by Ward´s method ("ward.D2"). Prior to analysis transformed LC-MS data were SNOG-filtered and intensity values were normalized to TIC. Signals below a cumulative intensity threshold of 5E+6 and a SNOG-score of 0.03 were removed. "WAT" - white adipose tissue.*

Dissecting the mouse N-glycome based on sub-structural determinants.

To systematically query and stratify N-glycan precursors based on fragment ion data, in a next step, we extended our initial SNOG-scoring scheme by additional N-glycan specific fragment ions (i.e. eSNOG). MS/MS spectrum-specific eSNOG-scores were thus calculated from the relative intensities of sub-structure specific diagnostic fragment ions and empirically determined, sub-structure specific eSNOG cut-off values were used for down-stream data-filtering (SI Fig. 6.). Efficiently limiting the impact of potential gas-phase rearrangements[50], this versatile data-filtering approach allowed for automated stratification of our reconstructed N-glycome data based on sub-structure specific features.

At first, we used eSNOG values to automatically stratify the N-glycome of mouse kidney (Fig. 4.) based on signals that contained: (I) distal-fucosylated, (II) Neu5Gc-sialylated, (III) NeuAc5-sialylated, (IV) alpha-galactosylated, (V) oligomannosidic N-glycans, and (VI) compositions which do not fall into any of these categories (i.e. predominantly undecorated or unusually decorated structures). In total, our eSNOG approach automatically uncovered 227 fucose-, 180 Neu5Gc-, 207 Neu5Ac-, 31 alpha-Gal-containing N-glycan precursor mass-bins, as well as 82 N-glycan precursors that do not fall into any of the six categories. When accumulating the MS signal intensities of all N-glycan precursors that fell into these six different categories, distal fucosylated N-glycans made up ~34% of the total ion current (TIC), Neu5Gc-sialylated mass-bins ~30%, Neu5Ac-sialylated mass-bins ~23%, alpha-Gal containing mass-bins ~4%, and oligomannose-associated mass-bins ~8%. Glycan compositions that did not confirm with any of these categories (i.e. undecorated structures), represented ~16% of the TIC in kidney.

Manual inspection of the stratified kidney N-glycome dataset (Fig. 4.) confirmed that our approach accurately classified N-glycan precursor mass-bins. For example, based on their precursor masses, biantennary, core-fucosylated and di-sialylated N-glycans were correctly assigned to either contain only Neu5Ac (2371.9 amu, $Hex_5HexNAc_4Fuc_1Neu5Ac_2$), only Neu5Gc (2403.9 amu, $Hex_5HexNAc_4Fuc_1Neu5Gc_2$), or one Neu5Ac as well as one Neu5Gc-residue (2387.9 amu, $Hex_5HexNAc_4Fuc_1Neu5Ac_1Neu5Gc_1$) (Fig. 4.A). Similarly, the mass-bins of 2257.8 amu and 2113.8 amu were correctly assigned to represent Neu5Gc sialylated ($Hex_5HexNAc_4Neu5Gc_2$) and alpha-galactosylated ($Hex_7HexNAc_4Fuc_1$) N-glycan precursor compositions, respectively. Furthermore, the mass-bin of 2039.7 amu was classified to contain Neu5Ac, which allowed to accurately define its composition to $Hex_6HexNAc_3Fuc_1Neu5Ac_1$, corresponding to a Neu5Ac-sialylated hybrid-type N-glycan. The Neu5Gc-capped analogue of the same glycan at 2055.7 amu (i.e. $Hex_6HexNAc_3Fuc_1Neu5Gc_1$) was correctly classified as Neu5Gc-containing N-glycan. The most abundant fucosylated N-glycan precursor in mouse kidney corresponded to a presumably bisected, bi-antennary N-glycan with three fucose residues (2284.9 amu) (Fig. 4.A). Other fucosylated precursor masses corresponded to tri- or tetra-antennary extensions of this N-glycan composition, with or without bisecting GlcNAc and with 3 to 5 fucose residues (e.g. $Hex_6HexNAc_5Fuc_{3/4}$, $Hex_6HexNAc_6Fuc_{3/4}$, $Hex_7HexNAc_6Fuc_{3/4/5}$, $Hex_7HexNAc_7Fuc_{3/4/5}$), confirming previous studies[51,52]. Abundant N-glycan precursor masses that were classified as “undecorated”, included the mass of 1789.7 amu, which corresponded to the composition $Hex_5HexNAc_4Fuc_1$, as well as smaller, truncated N-glycans, including the compositions $Hex_3HexNAc_2Fuc_1$ and $Hex_3HexNAc_3Fuc_1$ (Fig. 4.A), which are indeed not modified by distal fucose or sialic acid residues.

Capitalizing on our automated data-analysis workflows, we next stratified all other N-glycome datasets into the same six N-glycan categories and compared their relative abundances across the 20 murine tissues analyzed (Fig. 4.B). This comparative analysis allowed to further dissect the structural complexity of the mouse N-glycome at the precursor level and revealed a remarkable diversity in glycosylation patterns across tissues.

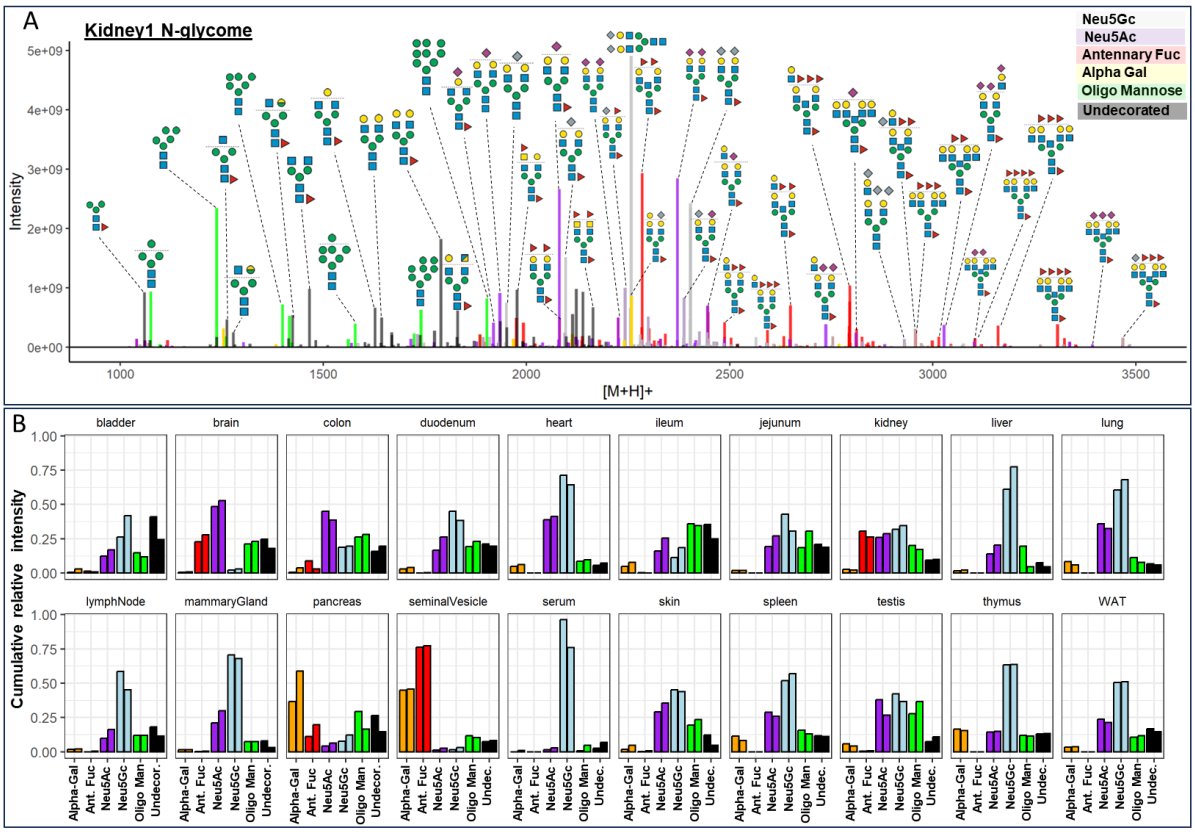

Figure 4. Sub-structural stratification of the mouse N-glycome. (A) *eSNOG-based classification of the kidney N-glycome. LC-MS signals were automatically categorized based on diagnostic fragment ions using sub-structure specific eSNOG cut-off values and classified as: distal-fucosylated (red), Neu5Gc-sialylated (light blue), NeuAc5-sialylated, (purple), alpha-galacotsylated (yellow), oligomannosidic N-glycans (green), and glycan compositions that do not fall into any of the other categories (black). N-glycan cartoons depict tentative structure assignments based on composition, using the Symbol Nomenclature for Glycans (SNFG).* ***(B)*** *Relative abundance of sub-structural determinants across tissues. Both replicates for each tissue are shown. "Ant. Fuc" - antennary fucosylation, "Oligo Man" - oligomannosidic, "Undecor." – undecorated (i.e. do not fall into any of the other categories).*

From our initial precursor-independent profiling analyses we found that serum was largely dominated by Neu5Gc-bearing N-glycans, representing up to 96% of the TIC[53]. Additionally, our N-glycan precursor-informed data now revealed that this important trait of the murine serum N-glycome was essentially derived from only two biantennary, non-fucosylated N-glycans ($Hex_5HexNAc_4Neu5Gc_1$ and $Hex_5HexNAc_4Neu5Gc_2$), totaling ~57% of all its Neu5Gc-containing structures. The sialylated fraction of the brain N-glycome, on the other hand, consisted almost entirely of Neu5Ac-decorated N-glycan species, in good agreement with previous studies[25,54]. Based on our eSNOG-stratified precursor information, the N-sialome of the brain presented as highly diverse, comprising of truncated (e.g. $Hex_4HexNAc_3Fuc_1Neu5Ac_1$), hybrid-type (e.g. $Hex_6HexNAc_3Fuc_1Neu5Ac_1$), bisecting GlcNAc-containing (e.g. $Hex_5HexNAc_5Fuc_1Neu5Ac_2$) and highly complex tetra-antennary N-glycans (e.g. $Hex_7HexNAc_6Fuc_3Neu5Ac_2$). Additionally, we found approx. 50% of the sialylated N-glycan species in brain to also carry distal fucoses. From this data we calculated that almost 50% of the brain N-glycome is sialylated. Of note, the degree of sialylation reported for the brain varies vastly between studies, with

previous estimates ranging from only ~3%[25], over 20%[29], to up to ~40% of sialylation[55], depending on methodology used.

Next to brain, kidney and seminal vesicle stood out by their comparably high abundance of distal fucosylated N-glycans. Different to brain and kidney, however, seminal vesicles showed extremely low levels of sialylation, with only ~3% of the total N-glycome being either Neu5Ac- or Neu5Gc-sialylated. Instead, the seminal vesicle N-glycome was dominated by distal fucosylated (~75%) N-glycans that were at the same time alpha-galactosylated (~45%). The most abundant precursor compositions of this type corresponded to a series of tri-antennary, core-fucosylated N-glycans with permutational additions of alpha-galactose and/or distal fucose (i.e. $Hex_6HexNAc_5Fuc_4$ – $Hex_7HexNAc_5Fuc_3$ – $Hex_8HexNAc_5Fuc_2$), suggesting that fucosylation and alpha-galactosylation on the same antenna are mutually exclusive. Similarly high levels of alpha-galactosylated N-glycans were only found in the pancreatic N-glycome (~47%), which also presented with rather low levels of sialylation (approx. 10%) and distal fucosylation (approx. 15%). In pancreas, the most abundant alpha-galactosylated N-glycans ranged from biantennary, with or without core-fucose, with one or two alpha-galactosylated or fucosylated antennae (e.g. $Hex_6HexNAc_4Fuc_1$, $Hex_7HexNAc_4$, $Hex_6HexNAc_4Fuc_2$, $Hex_7HexNAc_4Fuc_1$), to tri- and tetra-antennary antennary N-glycans with alpha-galactosylated and/or fucosylated antenna (e.g. $Hex_7HexNAc_5Fuc_2$, $Hex_8HexNAc_5Fuc_1$, $Hex_7HexNAc_5Fuc_3$, $Hex_8HexNAc_5Fuc_2$, $Hex_9HexNAc_5Fuc_1$, $Hex_{10}HexNAc_6Fuc_2$, $Hex_{11}HexNAc_6Fuc_1$). The exceptionally high levels of alpha-Gal in the two exocrine organs, seminal vesicle and pancreas, have remained unobserved in previous mouse studies.

Differentiating the N-glycomes of seminal vesicle and pancreas, we next screened our data for fragment ions that are diagnostic for doubly fucosylated antennae (i.e. $Hex_1HexNAc_1Fuc_2$, 658.3 amu), hence the Lewis Y-epitope. Our analysis revealed that this N-glycan motif was most specific for seminal vesicle (~1.2%) (SI Fig. 7.A) and barely detected in the pancreas of mice. Lewis Y-containing precursors included a series of partially core-fucosylated, bi- (e.g. $Hex_5HexNAc_4Fuc_4$, $Hex_5HexNAc_4Fuc_5$,) and multi-antennary N-glycans (e.g. $Hex_6HexNAc_5Fuc_5$, and $Hex_8HexNAc_7Fuc_6$). This unique glycosylation landscape of the seminal vesicle stands out among all tissues, and it is, to our knowledge, the first glycobiological description of this organ in mammals. Furthermore, the high expression levels of distally fucosylated N-glycans, including those with the Lewis Y epitope, correlated with gene expression data of the respective fucosyltransferases (i.e. *Fut2*, *Fut4*), as well as previous reports on the glycome of human seminal plasma [40,56]. The functional implications of these exceptionally high levels of distal fucose and alpha-galactose in exocrine organs remain to be explored.

Deep mining the mouse N-glycome uncovers unusual structural modifications.

To comprehensively annotate the results of our non-targeted data-analysis approach we used a *in silico* constructed mouse N-glycome database, holding an extensive list of canonical mouse N-glycans (i.e.

glycoDB; 1429 unique N-glycan compositions, 960 precursor ion mass bins; SI). Surprisingly, the automated annotation of our N-glycome histograms with this glycoDB highlighted several precursor masses, which, despite their consistent and reproducible detection, could not be explained by known, conventionally computed or anticipated N-glycan compositions. We hypothesized that this remarkably large population of unknown N-glycan precursor masses (i.e. approx. 15% of TIC across all samples; ranging from approx. 6% TIC in ileum, to up to approx. 37% in liver; SI Fig. 4.C and 4.D) would either result from experimental noise in the MS raw-data or represent unusual N-glycan compositions. Manual inspection of underlying MS/MS spectra uncovered and confirmed a series of unusual diagnostic fragment ions that were indicative for rare or non-canonical N-glycan structures, such as those bearing the HNK-1 epitope, sulfated HexNAc residues, doubly sialylated antenna, fucosylated LacdiNAc structures, or acetylated sialic acids (i.e. Ac-Neu5Ac and Ac-Neu5Gc); not covered by our *in-silico* N-glycome model database. Systematic mining of our dataset across all tissues using our eSNOG approach revealed striking tissue-specificity for most of these non-canonical N-glycan modifications.

HNK-1 is a unique trisaccharide epitope that consists of glucuronic acid linked to galactose, which is further linked to a GlcNAc residue. Additionally, HNK-1 has been reported to exist in a sulfated and a non-sulfated form (i.e. HSO3-3GlcA-beta1,3-Gal-beta-1,4-GlcNAc or GlcA-beta-1,3-Gal-beta-1,4-GlcNAc)[57,58]. This usually unregarded modification was previously identified in brain[28,57] and kidney[58], mainly in biantennary complex type N-glycan structures. The attachment of glucuronic acid is known to be catalyzed by two, highly tissue-specific, homologous enzymes, namely GlcAT-P (*B3gat1*) in brain and GlcAT-S (*B3gat2*) in kidney. Intriguingly, while GlcAT-P activity is suppressed by bisecting GlcNAc residues, the activity of GlcAT-S is not affected[59]. Mining the entire mouse N-glycome for diagnostic fragments that relate to HNK-1, either sulfated or non-sulfated (i.e. SO4-HexA-Hex-HexNAc; HexA-Hex-HexNAc), confirmed that HNK-1 was indeed only expressed in kidney (~1.5%)[58] and brain (~0.5%) (Fig. 5.A). The sulfated form of HNK-1, however, was exclusively found in the brain. In kidney, the most abundant HNK-1 N-glycan precursor corresponded to a triply fucosylated, tri-antennary structure, carrying a bisecting GlcNAc, with the composition $Hex_6HexNAc_6Fuc_3HexA_1$. Other HNK-1 containing N-glycans of this tissue corresponded to tri- and tetra-antennary structures with 1 – 3 antennary fucose-residues (e.g. $Hex_6HexNAc_6Fuc_4HexA_1$, $Hex_7HexNAc_7Fuc_2HexA_1$, $Hex_7HexNAc_7Fuc_3HexA_1$, $Hex_7HexNAc_7Fuc_4HexA_1$). Noteworthy, from its composition, and corroborated by the underlying MS/MS spectra, we found that the $Hex_6HexNAc_6Fuc_4HexA_1$ N-glycan carried a fucosylated HNK-1 antenna (SI Fig. 8.). To the best of our knowledge, this fucosylated form of HNK-1 has never been found in a natural source and was only artificially synthesized in a previous study[60].

Sulfation is considered the most diverse glycan modification with 35 sulfotransferases involved in the process of glycan sulfation[61]. Although many of these sulfotransferases are thought to serve in decorating O-linked glycosaminoglycan chains, also sulfated N-glycans have been reported for porcine

and human pancreas, with the highest levels found within the islets of Langerhans[62]. Systematic MS/MS-based screening of the different tissues for the relevant fragment ion (HexNAc-SO4 – 284 amu) revealed that almost all tissues exhibited at least low levels of N-glycan sulfation. In line with previous reports[32,62], also in our data the highest amount of HexNAc sulfation was found in pancreas (~5%), bladder (~2.5%), skin (~2%), and brain (~1%) (Fig. 5.F). Remarkably, however, the population of sulfated N-glycans in the pancreas is comprised exclusively of bi-antennary glycans, with and without core-fucose (e.g. $Hex_5HexNAc_4Fuc_1SO4_1$). Other sulfated structures in the pancreas were found to also carry one or two additional alpha-Gal residues (e.g. $Hex_6HexNAc_4Fuc_1SO4_1$), which aligned well with the high expression levels of alpha-Gal in the pancreas found by our study.

The expression of the fucosylated LacdiNAc N-glycan epitope was previously associated with different forms of cancer[63], and previous data has shown that LacdiNAc precursors can be fucosylated by FUT9[64], making it a candidate gene for regulating the tissue-specific biosynthesis of fucosylated LacdiNAc. Recently, fucosylated LacdiNAc containing structures have been identified in the N-glycome of the human brain[27] and of HEK-293 cell-line[65,66]. Systematic screening of all tissues for glycan precursors that generated the respective diagnostic fragment ion ($HexNAc_2Fuc_1$ – 553.2 amu) revealed that this epitope is expressed in a highly tissue-specific manner, with the highest levels found in kidney (~3%), colon (~2%), skin (~0.5%), and brain (~0.1%) (Fig. 5.E). These observations were further corroborated by previous reports of *Fut9* expression in the proximal tubule of kidney, in intestinal epithelial cell and in neurons[42].

Neu5Ac and Neu5Gc can both be installed in either α2,3-, α2,6- or α2,8-linkage, to either galactose or -very rarely in N-glycans- to N-acetyl-hexosamines, such as N-acetyl-glucosamine (i.e. 6-sialyl Lewis C; as reported previously in small amounts in bovine fetuin[67]) or N-acetyl-galactosamine (e.g. sialyl LacDiNAc). We thus compared the relative abundances of the two sialic acid variants linked to hexoses, the canonical acceptor sites on N-glycans, or to N-acetylhexosamines (i.e. N-acetylglucosamine or N-acetylgalactosamine; 6-sialyl Lewis C, sialyl LacdiNAc, respectively), across all tissues. Again, we eSNOG-filtered all N-glycan derived MS/MS spectra for those that contained fragment ions diagnostic for sialic acids (i.e. fragment ion mass 292.1 amu and 308.1 amu) and for sialylated HexNAc (i.e. fragment ion mass 495.2 amu and 511.2 amu, respectively). Expectedly, the relative proportions of precursor signals of canonically sialylated N-glycans greatly exceeded those that were generated from the unusual sialyl-HexNAc structures, in all tissues. Also, the relative incorporation rates of Neu5Ac and Neu5Gc were essentially independent of the acceptor monosaccharide across all tissues. Remarkably, in brain Neu5Ac-HexNAc was found on approximately 10% of the sialylated N-glycan fraction.

Furthermore, we systematically screened all tissues for fragment ions that are diagnostic for the doubly Neu5Ac-capped di-sialyl Lewis C epitope (i.e. 948.3 amu), which has previously been reported for brain[32,55], and its compositional analog, the Neu5Gc-decorated di-sialyl Lewis C epitope (i.e. 980.3

amu). Biosynthesis of di-sialyl Lewis C has been associated with the expression of the sialyltransferases ST6GALNAC3, ST6GALNAC5 and ST6GALNAC6, which are all able to catalyze the addition of α2,6 sialic acid to the GlcNAc residue within LacNAc motifs[68,69]. Interestingly, high expression level of ST6GALNAC5, which in healthy mouse and human is only found in the brain[70,71], has been strongly linked to brain-tropism of breast cancer metastasis[70]. This suggests that di-sialyl Lewis C may be involved in brain microenvironmental interactions facilitating metastatic niche establishment. Our cross-tissue analysis revealed that the expression of Neu5Ac-containing di-sialyl Lewis C is strictly limited to a small number of tissues, with the most prominent being brain (~3% of the TIC) and heart (~0.7%) (Fig. 5.B). Only minor levels (< 0.2%) of di-sialyl Lewis C N-glycans were found in kidney, testis, and lung. The associated precursor masses in the brain correspond to compositions ranging from potential hybrid-type (e.g. $Hex_5HexNAc_3Fuc_1Neu5Ac_2$, $Hex_6HexNAc_3Fuc_1Neu5Ac_2$) over biantennary complex-type N-glycans with up to four Neu5Ac-residues (e.g. $Hex_5HexNAc_4Fuc_1Neu5Ac_4$), to triantennary N-glycans with four Neu5Ac-residues (e.g. $Hex_6HexNAc_6Fuc_1Neu5Ac_4$). Neu5Gc-decorated variants of the di-sialyl Lewis C-epitope were detected, albeit at low levels, in multiple tissues. The highest expression levels were found in liver (ranging from ~2% to ~0.5%), heart (~1%), testis (ranging from ~2% to ~0.8%), lung (~0.8%), kidney (~0.4%), spleen (~0.2%), thymus (~0.4%), and serum (ranging from ~1% to ~0.1%) (Fig. 5.D). Interestingly, the Neu5Gc-carrying di-sialyl Lewis C N-glycans exhibited a diminished structural complexity when compared to the Neu5Ac-capped variants found in the brain (Fig 5.). In contrast to their Neu5Ac-capped homologs, N-glycan compositions containing Neu5Gc-based di-sialyl Lewis C were almost identical across all relevant tissues, comprising essentially two biantennary N-glycans (i.e. $Hex_5HexNAc_4Neu5Gc_3$ and $Hex_5HexNAc_4Fuc_1Neu5Gc_3$).

Further dissecting the structural complexity of sialylated N-glycans, we also investigated the relative abundance of O-acetylated neuraminic acids across tissues. O-Acetylation of sialic acids is correlated with the circulatory half-life of glycoproteins in the human serum and can be crucial for their biological activities[72]. Most importantly, O-acetylated sialic acids are critical entry receptors to many respiratory viruses, including Influenza C virus, human coronavirus OC43 and the murine coronavirus[73]. Murine coronaviruses often spread to the liver, an organ topism that has been suggested to be partially explained by the expression pattern of Ac-Neu5Ac and Ac-Neu5Gc in both lung and liver[74]. Quantifying the signals of all N-glycan precursors with O-acetylated neuraminic acids (i.e. Ac-Neu5Ac, fragment ion mass = 334.1 amu and/or Ac-Neu5Gc, fragment ion mass = 350.1 amu) allowed us to confirm expression of these important modifications almost exclusively in five tissues, namely lung (~4.5%), heart (~2 %), kidney (~1.5%), and, at very low levels, spleen (< 0.1%) (Fig. 5.C). Liver gave ambiguous results, as liver 1 showed very high, and liver 2 lower expression of acetylated sialic acids (~3.5% and ~0.2%, respectively). The respective compositions in the lung are all partially core-fucosylated, bi-antennary N-glycans with one, two, or three Neu5Gc- or Neu5Ac-residues, of which one sialic acid residue was acetylated (e.g. $Hex_5HexNAc_4Fuc_1Neu5Gc_1Ac_1$). So far, CASD1 is the only mammalian enzyme known to catalyze the acetylation of sialic acid resulting in the formation of Neu5,9Ac2[75]. In mouse, like human,

CASD1 appears to be widely expressed among most organs[42]. Our findings suggest additional unknown regulatory mechanisms that restrict expression of O-acetylated neuraminic acids to specific organs. Furthermore, we predominantly observed O-acetylation on Neu5Gc and only to a low degree on Neu5Ac residues (SI Fig. 7.H and SI Fig. 7.I). This suggests important differences in the biosynthesis, stability, or incorporation of the two O-acetylated sialic acid variants into N-glycans. As it is unclear whether CASD1 also catalyzes the O-acetylation of Neu5Gc, this warrants further investigations into the substrate specificities of CASD1 and the identification of additional O-acetyltransferases[76].

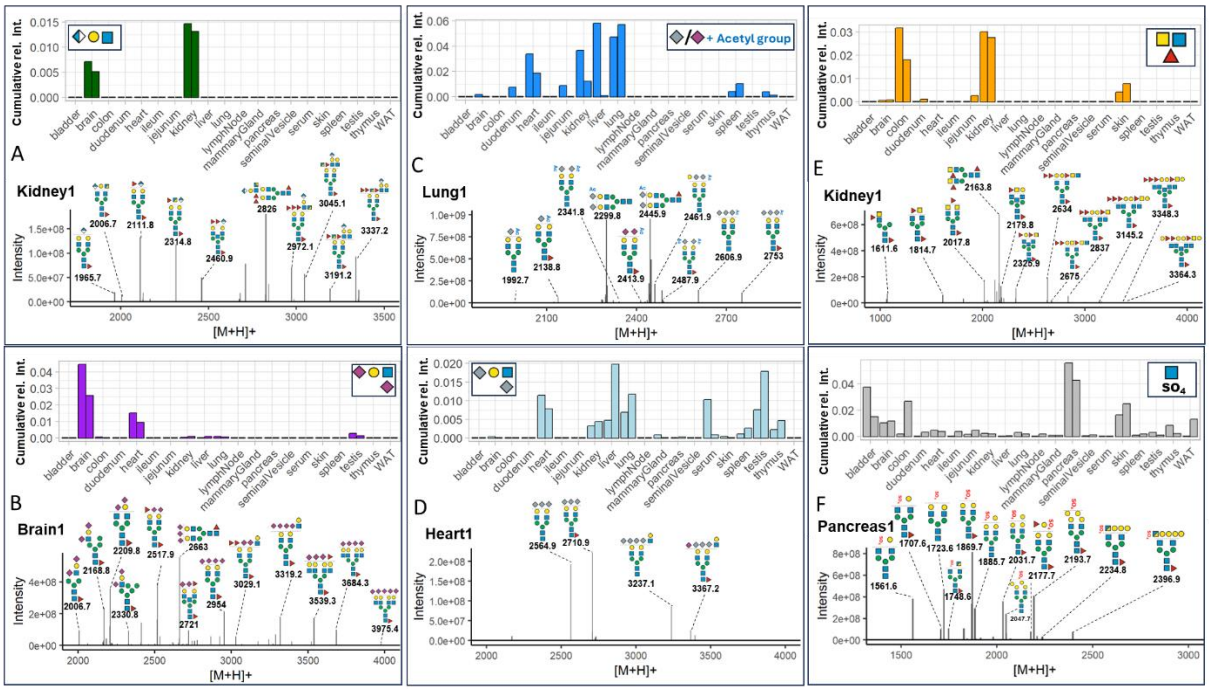

Figure 5. Tissue specific expression of unusual N-glycans*. Comparative analysis of N-glycans carrying* ***(A)*** *HNK-1-epitope* ***(B)*** *Neu5Ac-containing di-sialyl Lewis C-epitope* ***(C)*** *O-Acetylated sialic acids* ***(D)*** *Neu5Gc-containing di-sialyl Lewis C-epitope* ***(E)*** *Fucosylated LacdiNAc and* ***(F)*** *Sulfated HexNAc. Representative N-glycan compositions bearing the respective modification are shown for selected tissues. N-glycan cartoons depict tentative structure assignments based on composition, using the Symbol Nomenclature for Glycans (SNFG).“WAT–” - white adipose tissue.*

Profiling the isomeric structural complexity of the murine N-glycome.

Next to compositional variations, N-glycans exhibit a unique level of micro-heterogeneity that derives from structural, positional, and anomeric isomers. The number of unique isobaric N-glycan structures that can be constructed from a given mono-saccharide composition (hence of identical molecular mass; i.e. isobaric) adds a critical layer of complexity to the N-glycome[9]. Importantly, the co-existence of multiple isobaric N-glycan isomers within a single tissue cannot be captured by MS (or even MS/MS) alone and can only be resolved by either chromatographic or ion-mobility based separation techniques. To generate isomer-sensitive information, in this study, all samples were analyzed using a highly isomer-selective stationary phase (PGC-LC). Previously established N-glycan retention libraries combined with MS/MS data were used to identify the exact structures of respective isomers[9,26,27,77].

Expanding our N-glycome analyses by the integration of chromatographic information revealed the staggering structural complexity of the mouse N-glycome and provided insight into vital, organ-intrinsic regulations of glycobiological pathways. To showcase the granularity of our N-glycan isomer-sensitive dataset, we first compared the retention times of a specific, single precursor mass, that holds all doubly Neu5Ac-sialylated, core-fucosylated, biantennary, complex type N-glycan structures of composition $Hex_5HexNAc_4Fuc1Neu5Ac_2$, across tissues (Fig. 6.A). As sialic acids are usually found in either α2,3- or α2,6-linkage to terminal galactose residues of N-glycans, up to four different structural isomers of unique retention times are expected for this single glycan composition: both antennae carrying α2,3-linked, both antennae capped with α2,6-linked, and one of the antennae with α2,6-linked while the other antenna bearing a α2,3-linked Neu5Ac. To compensate for experimental chromatographic elution-time shifts between samples, all retention times within a given analytical run were normalized to those of the consistently detected Man5 N-glycan structure. Elution profiles of N-structures comprising of $Hex_5HexNAc_4Fuc_1Neu5Ac_2$ showed that almost all tissues were dominated by α2,3-linked Neu5Ac (Fig. 6.A, SI Fig. 9.D). Interestingly, however, brain, lung, and testis presented with a balanced ratio of α2,3- and α2,6-linked Neu5Ac isomers. Moreover, the brain displayed a notable presence of distinct N-glycan structure isomers due to the occurrence of branching Neu5Ac and/or antennary fucose. Remarkably, these specific glycan structures were exclusive to the brain tissue and were not identified in any other organ within the mouse.

Next, we compared the elution profile of the corresponding Neu5Gc-sialylated N-glycan structures (i.e. $Hex_5HexNAc_4Fuc_1Neu5Gc_2$) across all tissues. In stark contrast to Neu5Ac-sialylated structures, Neu5Gc-sialylated structures were predominantly found in α2,6-linkage across tissues, except for skin (i.e. ratio of α2,3- to α2,6-linked Neu5Gc approx 50%) (Fig. 6.A, SI Fig. 9.E). This entirely different construction of the Neu5Gc N-sialome compared to its Neu5Ac-terminated counterpart raised the question of the origin of these differently sialylated structures.

The elution profile of the mixed Neu5Ac/Neu5Gc composition ($Hex_5HexNAc_4Fuc_1Neu5Ac_1Neu5Gc_1$), closely resembled the elution profile of the Neu5Ac/Neu5Ac-sialylated structures in all tissues (Fig. 6.A), suggesting a shared origin for these structures. By contrast, the structural profile of entirely Neu5Gc-sialylated structures (i.e. $Hex_5HexNAc_4Fuc_1Neu5Gc_2$) markedly deviated from the Neu5Ac/Neu5Gc and Neu5Ac/Neu5Ac patterns. Notably, the elution profile of Neu5Gc/Neu5Gc structures exhibited minimal variation across tissues and closely resembled the serum elution profile, suggesting that these structures were actually derived from (contaminant) serum glycoproteins.

PGC-LC also allowed us to discriminate pivotal positional differences in distal fucosylation (i.e. α1,2-linked to galactose or α1,3-linked to GlcNAc-residues) that give rise to the important, glycan-associated immune-determinants bgH and Lewis X. While Lewis X determinants are mainly synthesized by FUT4 or FUT9, bgH-epitopes are synthesized by FUT1 or FUT2[78]. To discern these critical fucose structures, we compared the retention time profiles of a multiply fucosylated precursor mass of the compositions

$Hex_5HexNAc_4Fuc_3$, across all relevant samples (Fig. 6.B). The N-glycan structures that may be deduced from this single composition comprise asialo, core-fucosylated, biantennary, complex type N-structures with two distal fucoses, either in bgH- or Lewis X-related linkage. Based on our normalized elution profile data we found unexpected differences in antennary fucose between brain, kidney, seminal vesicles, and pancreas (Fig 6.B). While brain and kidney were found to contain almost exclusively the Lewis X containing isomer[79], pancreas exhibited only the bgH-epitope containing isomer, which correlates with the high expression of FUT1 in the pancreas[80]. Seminal vesicle, however, exhibited signals corresponding to both the Lewis X and the bgH-eptitope containing isomer. This finding aligns well with the occurrence of the Lewis Y-epitope, as described above, and the high expression level of FUT1, FUT2 and FUT4 in seminal vesicle[40]. Interestingly, a major carrier of Lewis X and Lewis Y in human seminal plasma has been linked to Glycodelin isoform S (GdS)[81]. As Lewis X and Y are known ligands to the immune receptor DC-SIGN[82] and glycodelins are potent immunosuppressors, they have been suggested to be important for feto-embryonic defense[83] in the early stages of gestation against adverse immune reactions.

Chromatography also provided stratification of arm/branch-isomers of tri-antennary, complex-type structures. Complex-type N-glycans are characterized by the substitution of both (i.e. α1,3- and α1,6-) terminal core-mannoses by the addition of GlcNAc residues by *Mgat1* and *Mgat2*, respectively. Additional branching of such complex-type structures results from secondary GlcNAc substitutions on the (i.e. α1,3- and α1,6-) terminal core-mannoses by *Mgat4* and *Mgat5*, respectively, or the installation of a "bisecting GlcNAc" to the core-central mannose by *Mgat3*. Our isomer-sensitive analysis thus allowed us to discern between the catalytic activities of the five different glycosyl-transferases involved, across all tissues. To explore the occurrence of biologically distinct N-glycan structures of identical composition, we compared the retention times of the precursor composition $Hex_3HexNAc_5Fuc_1$ across all tissues which could either be multiple tri-antennary structural isomers or a single biantennary, bisected complex type structure (Fig. 6.C, SI Fig. 9.C). The respective elution profiles indicated three dominant peaks. Intriguingly, brain, kidney, colon, ileum, duodenum, jejunum, and serum were strongly dominated by biantennary, bisected structure isomers. While this confirmed the results of our MS/MS based profiling, it also suggests a high occurrence of bisecting GlcNAc in brain, kidney, and colon, corroborated by *Mgat3* expression data[84].

Apart from oligomannose-type, N-glycans are often classified into hybrid-type or complex-type structures. Hybrid-type N-glycans result from the incomplete action of alpha-mannosidase II, giving rise to unsubstituted mannose residues on the α1,6-arm and a potentially substituted GlcNAc residue linked to the α1,3-mannose residue[1]. Interestingly, many N-glycome studies infer N-glycan classification based on composition. We found that this approach is highly oversimplifying the experimental data presented here. For example, the composition $Hex_5HexNAc_3Fuc_1$, which is often considered an "archetypical" hybrid type N-glycan structure, separated into several chromatographic peaks across tissues (SI Fig

9.A). Although most tissues exhibited highly similar elution profiles, pancreas, seminal vesicle, thymus, and spleen showed distinct dominant structures, which eluted much later. Manual inspection of the associated MS/MS spectra indicated core-fucosylated, truncated N-glycan structure with an alpha-galactosylated antenna. This was further corroborated by high levels of alpha-galactose in pancreas, seminal vesicle, thymus, and spleen, as were observed in our initial MS/MS based N-glycome profiling, and implied elevated levels of hexosaminidase in these tissues.

The classification of N-glycan structures, solely relying on composition, is further complicated by the incorporation of bisecting GlcNAc into hybrid-type N-glycans. Many hybrid-type structures with bisecting GlcNAc have counterparts in the class of complex-type N-glycans, adding an additional layer of complexity to their delineation based on composition (e.g. $Hex_5HexNAc_4Fuc_1$, $Hex_6HexNAc_4Fuc_2$, $Hex_5HexNAc_4$, $Hex_6HexNAc_4$). While the classification of the composition $Hex_5HexNAc_4Fuc_1$ as complex-type was true for most tissues we analyzed, we found that brain exhibited at least two hybrid-type structures within this composition[9] (Fig. 6.D, SI Fig. 9.A). Similarly, while the composition $Hex_6HexNAc_4Fuc_1$ consisted of two hybrid-type structures containing bisecting GlcNAc in the brain, it represented alpha-galactosylated complex-type structures when present in other tissues (SI Fig. 10.A). The same was found for the composition $Hex_6HexNAc_4Fuc_2$, which consists of hybrid-type structures with a bisecting GlcNAc in the brain, yet complex-type structures when present in other tissues (SI Fig. 10.B). While these findings suggest that hybrid-type structures with bisecting GlcNAc are yet another highly distinctive feature of the brain N-glycome, they boldly underscore the merits of isomer-sensitive N-glycome analyses to uncover unexpected structural signatures.

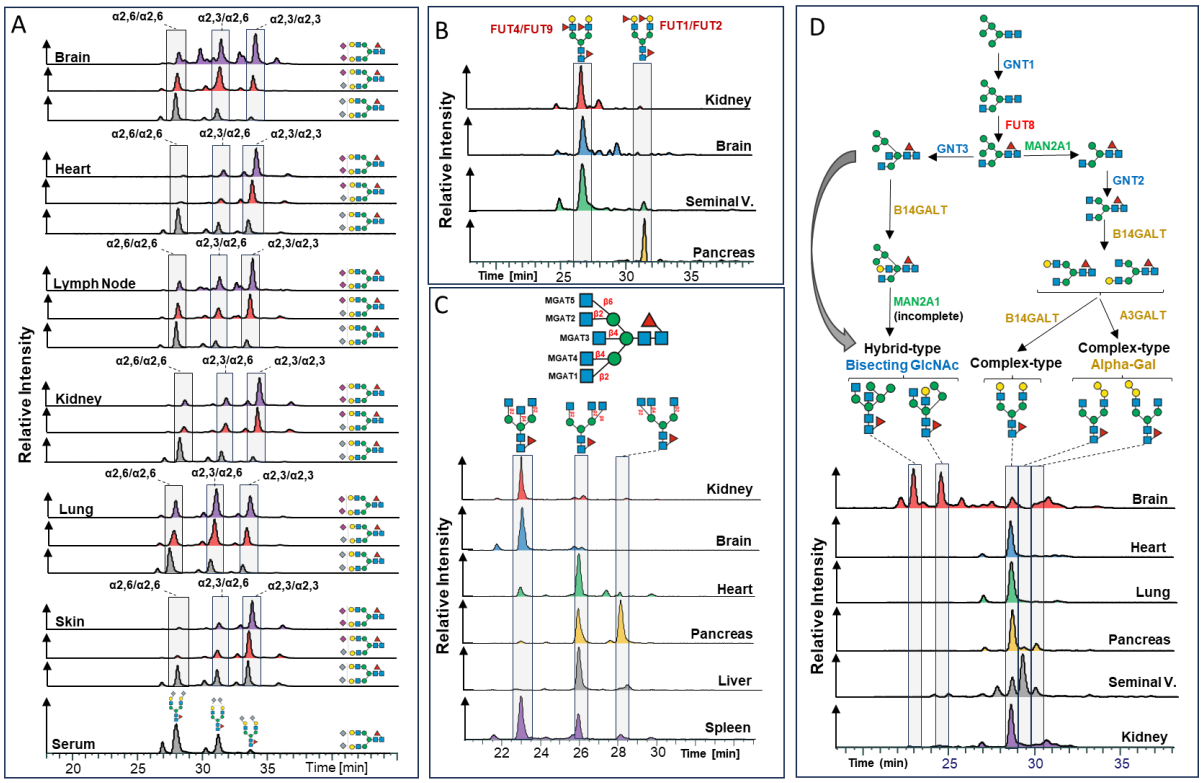

Figure 6. Isomer specific profiling of the murine N-glycome. *Closely related structural N-glycan isomers were separated with PGC-LC. N-glycan retention libraries in conjunctions with MS/MS data were used to identify the exact structure of the respective isomers. All retention times were normalized to the Man5 N-glycan.* ***(A)*** *Elution profiles for the compositions* $Hex_5HexNAc_4Fuc_1Neu5Ac_2$ *(purple),* $Hex_5HexNAc_4Fuc_1Neu5Ac_1Neu5Gc_1$ *(red), and* $ex_5HexNAc_4Fuc_1Neu5Gc_2$ *(light blue) across selected tissues.* ***(B)*** *Elution profiles for the composition* $Hex_3HexNAc_5Fuc_1$ *across kidney (red), brain (blue), heart (green), pancreas (orange), liver (grey), and spleen (purple).* ***(C)*** *Elution profiles of the heavily fucosylated composition* $Hex_5HexNAc_4Fuc_3$ *in kidney (red), brain (blue), seminal vesicle (green), and pancreas (orange).* ***(D)*** *Biosynthetic pathway leading to complex-type and hybrid-type N-glycans and elution profiles of the composition* $Hex_5HexNAc_4Fuc_1$ *in brain (red), heart (blue), lung (green), pancreas (orange), seminal vesicles (grey), and kidney (purple).*

Discussion

Here we make publicly available a consistent and structure sensitive N-glycome dataset of 20 different mouse tissues, essential for modern integrative system biology efforts. We used PGC-LC to chromatographically separate even closely related N-glycan isomers, an Orbitrap mass-analyzer for high-resolution tandem mass-spectrometric data-acquisition and an unconventional, non-targeted data analysis approach. This entirely data-driven, analytical approach provides means to automatically extract and query overall glycosylation patterns and allows for (semi-) quantitative N-glycomics.

Our workflow provides an unprecedented level of depth and structural fidelity, allows for the identification of multiple times more glycan features than previous studies and highlighted organ-intrinsic regulations of glycobiological pathways of not yet fully understood functionality. Additionally,

using our approach, we identified and consistently compared the expression of rare, non-canonical modifications, some of them even reported for the first time.

Unsupervised data analysis found clustering and convergent N-glycan signatures even among distantly related mouse tissues. These clustering patterns provide impartial insights into the glycosylation variations within organ groups, which potentially reflect their functional relationships or physiological roles.

Furthermore, our data suggests that different tissues use different biocatalytic strategies to specify their glycobiological identity, and thereby generate the staggering complexity revealed by our study. For example, while brain N-glycan diversity mainly results from its unique capacity to generate highly unusual and tissue-specific monosaccharide linkages to complex- and hybrid-type N-glycan core structures (e.g. "branching" Neu5Ac or HNK-1; Fig. 5.), the structural complexity of lung N-glycans largely draws from the incorporation and modification of different sialic acid variants (e.g. differently O-acetylated Neu5Gc and Neu5Ac). By contrast, exocrine organs, such as pancreas or seminal vesicle, shape their diversity of N-glycan structures primarily via a uniquely tuned interplay of alpha-galactosylation and antennary fucosylation. Although large parts of such tissue-specific N-glycome signatures can be aligned with gene expression data, especially the expression of non-canonical N-glycan structures imply additional, hitherto unknown, regulatory mechanisms that restrict the catalytic activity of glyco-enzymes to specific organs.

Taken together, using a non-targeted isomer-sensitive workflow to characterize a consistent and essentially complete mouse N-glycome atlas, our analyses reveal tissue specific N-glycan-structural features and highlight important organ-intrinsic regulations of glycobiological pathways. We anticipate that both our dataset and our analytical workflow, will be instructive for fundamental glycobiological research, glycan analytical benchmarking, the development of new glycome data analysis tools, and integrative systems glycobiology.

Methods

Preparation of mouse tissues

Two Male and two female C57BL/6J mice were obtained from the Jackson Laboratory (PRID: IMSR_JAX:000664, Bar Harbor, ME). C57BL/6J mice were bred in the licensed breeding facility of the Institute of Molecular Biotechnology (IMBA, Vienna, AT), under a 14hrs/10hrs light/dark cycle. Food and water were available ad libitum. All mice were euthanized with carbon dioxide at the age of 13 weeks. All tissues were collected in triplicates from either female (F) of male (M) mice, and included: Skin from ear (F), whole pancreas (F), middle section of duodenum (F), middle section of jejunum (F), middle section of ileum (F), middle section of colon (F), whole spleen (F), left kidney (F), whole liver (F), whole heart (F), whole lung (F), whole brain (F), epididymal white adipose tissue (F), inguinal mammary fat pad including mammary glands and excluding inguinal lymph node (F), inguinal lymph node (F), whole thymus (F), left testicle (M), whole bladder (M), serum (submandibular bleeding, M), and left seminal vesicle (M). Prior to freezing, duodenum, jejunum, ileum, colon and bladder were opened and cleaned with PBS from food, fecal matter and/or urine. Serum was collected from blood with Microtainer SST tubes (BD Biosciences, 365968). All samples were immediately after collection snap frozen in liquid nitrogen and stored at -80°C until further processing.

N-glycan extraction

Each mouse tissue was transferred into a Falcon tube and mixed with 100 mM ammonium bicarbonate buffer containing 20 mM ditiothreitol and 2 % sodium dodecyl sulfate in a total volume of 2 mL. After homogenization with an Ultra-Turrax T25 disperser, the homogenate was incubated at 56 °C for 30 min. After cooling down, the solution was brought to 40 mM iodoacetamide and incubated at room temperature in the dark for 30 min. After centrifugal clarification, chloroform-methanol extraction of the proteins was carried out using standard protocols[66]. In brief, the supernatant was mixed with four volumes of methanol, one volume of chloroform and three volumes of water in the given order. After centrifugation of the mixture for 3 min at 2500 rpm, the upper phase was removed, 4 mL of methanol were added, and the pellet was resuspended. The solution was centrifuged, the supernatant was removed, and the pellet was again resuspended in 4 mL methanol. The last step was repeated two more times. After the last methanol washing step, the pellet was dried at room temperature. The dried pellet was taken up in 50 mM ammonium acetate o (pH 8.4). For N-glycan release 2.5 U N-glycosidase F were added and the resulting mixture was incubated overnight at 37°C. The reaction mixture was acidified with some drops of glacial acetic acid and centrifuged at 4000×g. The supernatant was loaded onto a Hypersep C18 cartridge (1000 mg; Thermo Scientific, Vienna) that had been previously primed with 2 mL of methanol and equilibrated with 10 mL water. The sample was applied, and the column was washed with 4 mL water. The flow-through and the wash solution were collected and subjected to centrifugal evaporation. Reduction of the glycans was carried out in 1 % sodium borohydride in 50 mM NaOH at room temperature overnight. The reaction was quenched by the addition of two drops glacial acetic acid. Desalting was performed using HyperSep Hypercarb solid-phase extraction cartridges (25 mg) (Thermo Scientific, Vienna). The cartridges were primed with 450 µL methanol followed by 450 µL 80% acetonitrile in 10 mM ammonium bicarbonate. Equilibration was carried out by the addition of three times 450 µL water. The sample was applied and washed three times with 450 µL water. N-glycans were eluted by the addition of two times 450 µL 80% acetonitrile in 10 mM ammonium bicarbonate. The eluate was subjected to centrifugal evaporation and the dried N-glycans were taken up in 20 µL HQ-water. 5 µL of each sample were subjected to LC-MS/MS analysis.

LC-MS/MS analysis

LC-MS analysis was performed on a Dionex Ultimate 3000 UHPLC system coupled to an Orbitrap Exploris 480 Mass Spectrometer (Thermo Scientific). The purified glycans were loaded on a Hypercarb column (100 mm × 0.32 mm, 5 µm particle size, Thermo Scientific, Waltham, MA, USA) with 10 mM ammonium bicarbonate as the aqueous solvent A and 80 % acetonitrile in 50 mM ammonium bicarbonate as solvent B. The gradient was as follows: 0 – 4.5 min 1 % B, from 4.5 – 5.5 min 1 – 9 % B, from 5.5 – 30 min 9 - 20 % B, from 30 – 41.5 min 20 - 35 % B, from 41.5 – 45 min 35 – 65 % B, followed by an equilibration period at 1 % B from 45 –55 min. The flowrate was 6 µL/min. MS analysis was performed in data-dependent acquisition (DDA) mode with positive polarity from 500 – 1500 m/z using the following parameters:

resolution was set to 60,000 with a normalized AGC target of 300%. The 10 most abundant precursors (charge states 2-6) within an isolation window of 1.4 m/z were selected for fragmentation. Dynamic exclusion was set at 20 s (n=1) with a mass tolerance of ± 10 ppm. Normalized collsion energy (NCE) for HCD was set to 20, 25 and 30% and the precursor intensity threshold was set to 8000. MS/MS spectra were recorded with a resolution of 15,000, using a normalized AGC target of 100% and a maximum accumulation time of 100 ms.

Data Analysis

For the MS/MS based profiling, MS2 data (in the .raw file format) were extracted and converted into .mgf files using the "vendor"-specific peak-picking algorithm implemented in MSConvert (version 3.0.22133-5eed1a6) and removing zero-samples. MS2 spectra were filtered for (N-)glycan specific oxonium-ions (i.e. 167.0914, 183.0863, 204.0867, 222.0972, 224.1118, 243.0264, 274.0921, 284.0435, 290.087, 292.1027, 308.0976, 312.1289, 316.1027, 328.1238, 332.0976, 334.1133, 350.1082, 366.1395, 370.1697, 407.1661, 495.1821, 511.177, 512.1974, 528.1923, 542.1716, 553.224, 569.21887, 622.1284, 657.2349, 658.2553, 673.2298, 698.2615, 699.2455, 714.2564, 715.2404, 715.27677, 792.3234, 803.2928, 819.2877, 860.31427, 876.30917, 948.3303, 980.3201, 446.09627, 649.17567, 731.27167, 877.32957, 1023.38747, 1169.44537; mass-tolerance +/- 0.05 amu) using a custom-made perl-script ("MS2OxoPlot.pl"), binned according to retention time (bin size = 10s), counted and visualized as scatter-plots using Excel365 (Microsoft).

MS1 data were extracted, charge-deconvoluted, and deisotoped using Decon2[47] with custom parameters adjusted for N-glycomics data (Supplementary data). In a next step, we assigned the precursor mass and intensity values to mass bins (i.e. in the range of 1000 - 5000 Da; bin-width = ±0.05 amu), summed the individual precursor mass-to-intensity values across the entire chromatographic time-range (i.e. retention time 0 – 50 min.) and eventually removed signals below a cumulative intensity threshold of 5E+6, using custom code.

Raw MS/MS data were refined (i.e. mono-isotopic peak picking and charge state assignment) and converted into the generic .mgf file-format using PEAKS[48]. Analogous to MS1-data processing, the refined precursor-information was deconvoluted and assigned into mass-bins (bin-width = ±0.05 amu). From this, to automatically identify precursors that derived from N-glycans and to stringently control for unintended precursor ion co-isolation events, we calculated mass-bin-specific score-values (SNOG-score) from the intensity value of an N- (and O-) glycan-specific fragment ion (i.e. oxonium ion of the reduced-end monosaccharide GlcNAc; [M+H]+ = 224.1 amu) using custom-code. To remove signals derived from non-N-glycan-associated structures, all mass-bins with a SNOG-score lower than 0.03 were rejected. This step was performed sample-specifically. The resulting "inclusion list" was used to retrieve the precursor-intensity information from the processed MS1 data.

To stratify the dataset based on sub-structural modifications (e.g. sialylation, fucosylation, sulfation, …), we used diagnostic fragment ions known from previous studies[27,53,85] and screened the dataset for so-far unknown reporter-ions. The sub-structure-specific "extended SNOG-score" (eSNOG) is calculated analogous to the SNOG-score, but instead of using the N-glycan specific fragment-ion 224.1, sub-structure-specific fragment ions are used (SI Table 1). As the different diagnostic fragment ions exhibit vastly differing relative intensities within a MS/MS-spectrum, the cut-off values for the eSNOGs are empirically determined for each of the fragment ions. The resulting sample- and sub-structural feature-specific "inclusion list" was used to retrieve the precursor-intensity information from the processed MS1 data.

All data visualization and statistical analyses were performed with R 4.3.1 using the packages "tidyverse"[86] (2.0.0), "pheatmap" (1.0.12), "RColorBrewer" (1.1.3), "dendextend "(1.17.1)[87], "corrr" (0.4.4), "ggrepel" (0.9.3), and "Rtsne" (0.17).

Data Records

The raw mouse N-glycome PGC-LC-MS/MS raw-data (.raw") and processed datasets referred to in this manuscript (in the ".csv" file format) are available for review via the GlycoPOST repository[88] (URL below) and the GitHub repository (URL below). Tissue and sample replicate numbers are indicated in the respective file names.

https://glycopost.glycosmos.org/preview/138960454965cccd1144ee6 (PIN CODE:7074).

https://github.com/holandir/Mouse-N-glycome-analysis.

Code Availability

All custom code, R scripts, and software for MS and MS/MS data pre-processing is freely available at: http://homepage.boku.ac.at/jstadlmann and https://github.com/holandir/Mouse-N-glycome-analysis.

Acknowledgements

J.H. was partly funded by the European Commission (Newcotiana 760331). S.M. is funded by the ESPRIT-Programme of the Austrian Science Fund (FWF, Project number: ESP 166). J.M.P. received funding from the T. von Zastrow foundation, the FWF Wittgenstein award (Z 271-B19), the Austrian Academy of Sciences, the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 101005026, and the Canada 150 Research Chairs Program F18-01336 as well as the Canadian Institutes of Health Research COVID-19 grants F20–02343 and F20-02015. This project was supported by MS equipment kindly provided the EQ-BOKU VIBT GmbH and the BOKU Core Facility Mass Spectrometry.

1. Varki, A., Cummings, R.D., Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N.H., Prestegard, J.H., et al. (2022). Essentials of Glycobiology. Essentials of Glycobiology. 10.1101/9781621824213.

2. Ugonotti, J., Chatterjee, S., and Thaysen-Andersen, M. (2021). Structural and functional diversity of neutrophil glycosylation in innate immunity and related disorders. Mol Aspects Med *79*, 100882. 10.1016/J.MAM.2020.100882.

3. Zhang, X.L., and Qu, H. (2021). The Role of Glycosylation in Infectious Diseases. Adv Exp Med Biol *1325*, 219–237. 10.1007/978-3-030-70115-4_11/COVER.

4. Orr, S.L., Le, D., Long, J.M., Sobieszczuk, P., Ma, B., Tian, H., Fang, X., Paulson, J.C., Marth, J.D., and Varki, N. (2013). A phenotype survey of 36 mutant mouse strains with gene-targeted defects in glycosyltransferases or glycan-binding proteins. Glycobiology *23*, 363–380. 10.1093/GLYCOB/CWS150.

5. Macher, B.A., and Galili, U. (2008). The Galalpha1,3Galbeta1,4GlcNAc-R (alpha-Gal) epitope: a carbohydrate of unique evolution and clinical relevance. Biochim Biophys Acta *1780*, 75–88. 10.1016/J.BBAGEN.2007.11.003.

6. Varki, A. (2001). Loss of N-glycolylneuraminic acid in humans: Mechanisms, consequences, and implications for hominid evolution. Am J Phys Anthropol *Suppl 33*, 54. 10.1002/AJPA.10018.ABS.

7. C, A., RA, K., and B, S. (2001). Do rodent and human brains have different N-glycosylation patterns? Biol Chem *382*, 187–194. 10.1515/BC.2001.026.

8. Antonopoulos, A., North, S.J., Haslam, S.M., and Dell, A. (2011). Glycosylation of mouse and human immune cells: insights emerging from N-glycomics analyses. Biochem Soc Trans *39*, 1334–1340. 10.1042/BST0391334.

9. Helm, J., Grünwald-Gruber, C., Thader, A., Urteil, J., Führer, J., Stenitzer, D., Maresch, D., Neumann, L., Pabst, M., and Altmann, F. (2021). Bisecting Lewis X in Hybrid-Type N-Glycans of Human Brain Revealed by Deep Structural Glycomics. Anal Chem, acs.analchem.1c03793. 10.1021/ACS.ANALCHEM.1C03793.

10. Nagai-Okatani, C., Zou, X., Fujita, N., Sogabe, I., Arakawa, K., Nagai, M., Angata, K., Zhang, Y., Aoki-Kinoshita, K.F., and Kuno, A. (2021). LM-GlycomeAtlas Ver. 2.0: An Integrated Visualization for Lectin Microarray-based Mouse Tissue Glycome Mapping Data with Lectin Histochemistry. J Proteome Res *20*, 2069–2075. 10.1021/acs.jproteome.0c00907.

11. Lee, J.W., Lee, K., Ahn, S.H., Son, B.H., Ko, B.S., Kim, H.J., Chung, I.Y., Kim, J., Lee, W., Ko, M.S., et al. (2020). Potential of MALDI-TOF-based serum N-glycan analysis for the diagnosis and surveillance of breast cancer. Scientific Reports 2020 10:1 *10*, 1–8. 10.1038/s41598-020-76195-y.

12. Lin, Z., and Lubman, D.M. (2013). Permethylated N-glycan analysis with mass spectrometry. Methods Mol Biol *1007*, 289–300. 10.1007/978-1-62703-392-3_12.

13. Zhang, Y., Wang, B., Jin, W., Wen, Y., Nan, L., Yang, M., Liu, R., Zhu, Y., Wang, C., Huang, L., et al. (2019). Sensitive and robust MALDI-TOF-MS glycomics analysis enabled by Girard's reagent T on-target derivatization (GTOD) of reducing glycans. Anal Chim Acta *1048*, 105–114. 10.1016/J.ACA.2018.10.015.

14. Reiding, K.R., Blank, D., Kuijper, D.M., Deelder, A.M., and Wuhrer, M. (2014). High-throughput profiling of protein N-glycosylation by MALDI-TOF-MS employing linkage-specific sialic acid esterification. Anal Chem *86*, 5784–5793. 10.1021/AC500335T.

15. Kinoshita, M., Saito, A., Yamamoto, S., and Suzuki, S. (2020). A practical method for preparing fluorescent-labeled glycans with a 9-fluorenylmethyl derivative to simplify a fluorimetric HPLC-based analysis. J Pharm Biomed Anal *186*, 113267. 10.1016/J.JPBA.2020.113267.

16. Zhang, L., Luo, S., and Zhang, B. (2016). Glycan analysis of therapeutic glycoproteins. MAbs *8*, 205–215. 10.1080/19420862.2015.1117719.

17. Veillon, L., Huang, Y., Peng, W., Dong, X., Cho, B.G., and Mechref, Y. (2017). Characterization of isomeric glycan structures by LC-MS/MS. Electrophoresis *38*, 2100–2114. 10.1002/elps.201700042.

18. Peng, W., Gutierrez Reyes, C.D., Gautam, S., Yu, A., Cho, B.G., Goli, M., Donohoo, K., Mondello, S., Kobeissy, F., and Mechref, Y. (2021). MS-based glycomics and glycoproteomics methods enabling isomeric characterization. Mass Spectrom Rev. 10.1002/MAS.21713.

19. Yamakawa, N., Vanbeselaere, J., Chang, L.Y., Yu, S.Y., Ducrocq, L., Harduin-Lepers, A., Kurata, J., Aoki-Kinoshita, K.F., Sato, C., Khoo, K.H., et al. (2018). Systems glycomics of adult zebrafish identifies organ-specific sialylation and glycosylation patterns. Nat Commun *9*. 10.1038/s41467-018-06950-3.

20. Wallace, E.N., West, C.A., McDowell, C.T., Lu, X., Bruner, E., Mehta, A.S., Aoki-Kinoshita, K.F., Angel, P.M., and Drake, R.R. (2024). An N-glycome tissue atlas of 15 human normal and cancer tissue types determined by MALDI-imaging mass spectrometry. Sci Rep *14*. 10.1038/S41598-023-50957-W.

21. Drake, R.R., McDowell, C., West, C., David, F., Powers, T.W., Nowling, T., Bruner, E., Mehta, A.S., Angel, P.M., Marlow, L.A., et al. (2020). Defining the human kidney N-glycome in normal and cancer tissues using MALDI imaging mass spectrometry. Journal of Mass Spectrometry *55*, e4490. 10.1002/JMS.4490.

22. Otaki, M., Hirane, N., Natsume-Kitatani, Y., Nogami Itoh, M., Shindo, M., Kurebayashi, Y., and Nishimura, S.I. (2022). Mouse tissue glycome atlas 2022 highlights inter-organ variation in major N-glycan profiles. Scientific Reports 2022 12:1 *12*, 1–15. 10.1038/s41598-022-21758-4.

23. Liu, Y., Han, Y., Zhu, W., Luo, Q., Yuan, J., and Liu, X. (2023). Characterization of N-glycome profile in mouse brain tissue regions by MALDI-TOF/MS. Anal Bioanal Chem *415*, 5575–5588. 10.1007/S00216-023-04848-8.

24. Grgic, A., Krestensen, K.K., and Heeren, R.M.A. (2023). Optimized protocol for MALDI MSI of N-glycans using an on-tissue digestion in fresh frozen tissue sections. Scientific Reports 2023 13:1 *13*, 1–7. 10.1038/s41598-023-29560-6.

25. Williams, S.E., Noel, M., Lehoux, S., Cetinbas, M., Xavier, R.J., Sadreyev, R.I., Scolnick, E.M., Smoller, J.W., Cummings, R.D., and Mealer, R.G. (2022). Mammalian brain glycoproteins exhibit diminished glycan complexity compared to other tissues. Nat Commun *13*. 10.1038/s41467-021-27781-9.

26. Abrahams, J.L., Campbell, M.P., and Packer, N.H. (2018). Building a PGC-LC-MS N-glycan retention library and elution mapping resource. Glycoconj J *35*, 15–29. 10.1007/s10719-017-9793-4.

27. Helm, J., and Hirtler, L. (2022). Towards Mapping of the Human Brain N-Glycome with Standardized Graphitic Carbon Chromatography. 1–20.

28. Lee, J., Ha, S., Kim, M., Kim, S.W., Yun, J., Ozcan, S., Hwang, H., Ji, I.J., Yin, D., Webster, M.J., et al. (2020). Spatial and temporal diversity of glycome expression in mammalian brain. Proc Natl Acad Sci U S A *117*, 28743–28753. 10.1073/pnas.2014207117.

29. Barboza, M., Solakyildirim, K., Knotts, T.A., Luke, J., Gareau, M.G., Raybould, H.E., and Lebrilla, C.B. (2021). Region-Specific Cell Membrane N-Glycome of Functional Mouse Brain Areas Revealed by nanoLC-MS Analysis. Molecular & Cellular Proteomics, 100130. 10.1016/j.mcpro.2021.100130.

30. Ji, I.J., Hua, S., Shin, D.H., Seo, N., Hwang, J.Y., Jang, I.S., Kang, M.G., Choi, J.S., and An, H.J. (2015). Spatially-resolved exploration of the mouse brain glycome by tissue glyco-capture (TGC) and nano-LC/MS. Anal Chem *87*, 2869–2877. 10.1021/ac504339t.

31. Ghosh, S. (2020). Sialic acid and biology of life: An introduction. Sialic Acids and Sialoglycoconjugates in the Biology of Life, Health and Disease, 1. 10.1016/B978-0-12-816126-5.00001-9.

32. Klarić, T.S., and Lauc, G. (2022). The dynamic brain N-glycome. Glycoconj J *39*, 443–471. 10.1007/S10719-022-10055-X.

33. Davies, L.R.L., and Vark, A. (2015). Why Is N-Glycolylneuraminic Acid Rare in the Vertebrate Brain? Top Curr Chem *366*, 31. 10.1007/128_2013_419.

34. Becker, D.J., and Lowe, J.B. (2003). Fucose: biosynthesis and biological function in mammals. Glycobiology *13*, 41R-53R. 10.1093/GLYCOB/CWG054.

35. Gersten, K.M., Natsuka, S., Trinchera, M., Petryniak, B., Kelly, R.J., Hiraiwa, N., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., and Lowe, J.B. (1995). Molecular Cloning, Expression, Chromosomal Assignment, and Tissue-specific Expression of a Murine α-(1,3)-Fucosyltransferase Locus Corresponding to the Human ELAM-1 Ligand Fucosyl Transferase. Journal of Biological Chemistry *270*, 25047–25056. 10.1074/JBC.270.42.25047.

36. Wang, H., Morales-Levy, M., Rose, J., MacKey, L.C., Bodary, P., Eitzman, D., and Homeister, J.W. (2013). α(1,3)-Fucosyltransferases FUT4 and FUT7 control murine susceptibility to thrombosis. Am J Pathol *182*, 2082–2093. 10.1016/J.AJPATH.2013.02.010.

37. Mucha, E., Lettow, M., Marianski, M., Thomas, D.A., Struwe, W.B., Harvey, D.J., Meijer, G., Seeberger, P.H., von Helden, G., and Pagel, K. (2018). Fucose Migration in Intact Protonated Glycan Ions: A Universal Phenomenon in Mass Spectrometry. Angew Chem Int Ed Engl *57*, 7440–7443. 10.1002/ANIE.201801418.

38. Lettow, M., Greis, K., Mucha, E., Lambeth, T.R., Yaman, M., Kontodimas, V., Manz, C., Hoffmann, W., Meijer, G., Julian, R.R., et al. (2023). Decoding the Fucose Migration Product during Mass-Spectrometric Analysis of Blood Group Epitopes. Angewandte Chemie International Edition. 10.1002/ANIE.202302883.

39. Lettow, M., Mucha, E., Manz, C., Thomas, D.A., Marianski, M., Meijer, G., von Helden, G., and Pagel, K. (2019). The role of the mobile proton in fucose migration. Anal Bioanal Chem *411*, 4637–4645. 10.1007/S00216-019-01657-W.

40. Chunmei, W., Huang, C., Gu, Y., Zhou, Y., Zhu, Z., and Zhang, Y. (2011). Biosynthesis and distribution of lewis x- And lewis y-containing glycoproteins in the murine male reproductive system. Glycobiology *21*, 225–234. 10.1093/glycob/cwq152.

41. Cailleau-Thomas, A., Coullin, P., Candelier, J.J., Balanzino, L., Mennesson, B., Oriol, R., and Mollicone, R. (2000). FUT4 and FUT9 genes are expressed early in human embryogenesis. Glycobiology *10*, 789–802. 10.1093/GLYCOB/10.8.789.

42. Schaum, N., Karkanias, J., Neff, N.F., May, A.P., Quake, S.R., Wyss-Coray, T., Darmanis, S., Batson, J., Botvinnik, O., Chen, M.B., et al. (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. Nature 2018 562:7727 *562*, 367–372. 10.1038/s41586-018-0590-4.

43. Chen, Q., Zhang, Y., Zhang, K., Liu, J., Pan, H., Wang, X., Li, S., Hu, D., Lin, Z., Zhao, Y., et al. (2022). Profiling the Bisecting N-acetylglucosamine Modification in Amniotic Membrane via Mass Spectrometry. Genomics Proteomics Bioinformatics *20*, 648–656. 10.1016/J.GPB.2021.09.010.

44. Chen, Q., Tan, Z., Guan, F., and Ren, Y. (2020). The Essential Functions and Detection of Bisecting GlcNAc in Cell Biology. Front Chem *8*, 511. 10.3389/FCHEM.2020.00511.

45. Priatel, J.J., Sarkar, M., Schachter, H., and Marth, J.D. (1997). Isolation, characterization and inactivation of the mouse Mgat3 gene: the bisecting N-acetylglucosamine in asparagine-linked oligosaccharides appears dispensable for viability and reproduction. Glycobiology *7*, 45–56. 10.1093/GLYCOB/7.1.45.

46. Groux-Degroote, S., Vicogne, D., Cogez, V., Schulz, C., and Harduin-Lepers, A. (2021). B4GALNT2 Controls Sda and SLex Antigen Biosynthesis in Healthy and Cancer Human Colon. ChemBioChem *22*, 3381–3390. 10.1002/CBIC.202100363.

47. Jaitly, N., Mayampurath, A., Littlefield, K., Adkins, J.N., Anderson, G.A., and Smith, R.D. (2009). Decon2LS: An open-source software package for automated processing and visualization of high resolution mass spectrometry data. BMC Bioinformatics *10*, 1–15. 10.1186/1471-2105-10-87/TABLES/3.

48. Zhang, J., Xin, L., Shan, B., Chen, W., Xie, M., Yuen, D., Zhang, W., Zhang, Z., Lajoie, G.A., and Ma, B. (2012). PEAKS DB: De Novo Sequencing Assisted Database Search for Sensitive and Accurate Peptide Identification. Mol Cell Proteomics *11*. 10.1074/MCP.M111.010587.

49. Burock, R., Cajic, S., Hennig, R., Buettner, F.F.R., Reichl, U., and Rapp, E. (2023). Reliable N-Glycan Analysis–Removal of Frequently Occurring Oligosaccharide Impurities by Enzymatic Degradation. Molecules *28*. 10.3390/MOLECULES28041843/S1.

50. Campos, D., Girgis, M., Yang, Q., Zong, G., Goldman, R., Wang, L.X., and Sanda, M. (2023). "Ghost" Fragment Ions in Structure and Site-Specific Glycoproteomics Analysis. Anal Chem *95*, 10145–10148. 10.1021/ACS.ANALCHEM.3C02207/ASSET/IMAGES/LARGE/AC3C02207_0003.JPEG.

51. Gustafsson, O.J.R., Briggs, M.T., Condina, M.R., Winderbaum, L.J., Pelzing, M., McColl, S.R., Everest-Dass, A. V., Packer, N.H., and Hoffmann, P. (2015). MALDI imaging mass spectrometry of N-linked glycans on formalin-fixed paraffin-embedded murine kidney. Anal Bioanal Chem *407*, 2127–2139. 10.1007/S00216-014-8293-7/FIGURES/3.

52. Powers, T.W., Neely, B.A., Shao, Y., Tang, H., Troyer, D.A., Mehta, A.S., Haab, B.B., and Drake, R.R. (2014). MALDI imaging mass spectrometry profiling of N-glycans in formalin-fixed paraffin embedded clinical tissue blocks and tissue microarrays. PLoS One *9*. 10.1371/JOURNAL.PONE.0106255.

53. Gao, W.N., Yau, L.F., Liu, L., Zeng, X., Chen, D.C., Jiang, M., Liu, J., Wang, J.R., and Jiang, Z.H. (2015). Microfluidic Chip-LC/MS-based Glycomic Analysis Revealed Distinct N-glycan Profile of Rat Serum. Scientific Reports 2015 5:1 *5*, 1–14. 10.1038/srep12844.

54. Naito-Matsui, Y., Davies, L.R.L., Takematsu, H., Chou, H.H., Tangvoranuntakul, P., Carlin, A.F., Verhagen, A., Heyser, C.J., Yoo, S.W., Choudhury, B., et al. (2017). Physiological exploration of the long term evolutionary selection against expression of N-Glycolylneuraminic Acid in the Brain. Journal of Biological Chemistry *292*, 2557–2570. 10.1074/jbc.M116.768531.

55. Zamze, S., Harvey, D.J., Chen, Y.J., Guile, G.R., Dwek, R.A., and Wing, D.R. (1998). Sialylated N-glycans in adult rat brain tissue - A widespread distribution of disialylated antennae in complex and hybrid structures. Eur J Biochem *258*, 243–270. 10.1046/j.1432-1327.1998.2580243.x.

56. Pang, P.C., Tissot, B., Drobnis, E.Z., Morris, H.R., Dell, A., and Clark, G.F. (2009). Analysis of the human seminal plasma glycome reveals the presence of immunomodulatory carbohydrate functional groups. J Proteome Res *8*, 4906–4915. 10.1021/PR9001756/ASSET/IMAGES/MEDIUM/PR-2009-001756_0008.GIF.

57. Morise, J., Takematsu, H., and Oka, S. (2017). The role of human natural killer-1 (HNK-1) carbohydrate in neuronal plasticity and disease. Biochim Biophys Acta Gen Subj *1861*, 2455–2461. 10.1016/J.BBAGEN.2017.06.025.

58. Tagawa, H., Kizuka, Y., Ikeda, T., Itoh, S., Kawasaki, N., Kurihara, H., Onozato, M.L., Tojo, A., Sakai, T., Kawasaki, T., et al. (2005). A non-sulfated form of the HNK-1 carbohydrate is expressed in mouse kidney. Journal of Biological Chemistry *280*, 23876–23883. 10.1074/JBC.M501728200.

59. Kawade, H., Morise, J., Mishra, S.K., Tsujioka, S., Oka, S., and Kizuka, Y. (2021). Tissue-Specific Regulation of HNK-1 Biosynthesis by Bisecting GlcNAc. Molecules *26*. 10.3390/MOLECULES26175176.

60. Tsvetkov, Y.E., Burg-Roderfeld, M., Loers, G., Ardá, A., Sukhova, E. V., Khatuntseva, E.A., Grachev, A.A., Chizhov, A.O., Siebert, H.C., Schachner, M., et al. (2012). Synthesis and molecular recognition studies of the HNK-1 trisaccharide and related oligosaccharides. the specificity of monoclonal anti-HNK-1 antibodies as assessed by surface plasmon resonance and STD NMR. J Am Chem Soc *134*, 426–435. 10.1021/JA2083015/SUPPL_FILE/JA2083015_SI_002.PDF.

61. Schjoldager, K.T., Narimatsu, Y., Joshi, H.J., and Clausen, H. (2020). Global view of human protein glycosylation pathways and functions. Nat Rev Mol Cell Biol *21*, 729–749. 10.1038/S41580-020-00294-X.

62. McDowell, C.T., Klamer, Z., Hall, J., West, C.A., Wisniewski, L., Powers, T.W., Angel, P.M., Mehta, A.S., Lewin, D.N., Haab, B.B., et al. (2021). Imaging Mass Spectrometry and Lectin Analysis of N-Linked Glycans in Carbohydrate Antigen–Defined Pancreatic Cancer Tissues. Molecular & Cellular Proteomics *20*, 100012. 10.1074/MCP.RA120.002256.

63. Hirano, K., and Furukawa, K. (2022). Biosynthesis and Biological Significances of LacdiNAc Group on N- and O-Glycans in Human Cancer Cells. Biomolecules 2022, Vol. 12, Page 195 *12*, 195. 10.3390/BIOM12020195.

64. Kawar, Z.S., Haslam, S.M., Morris, H.R., Dell, A., and Cummings, R.D. (2005). Novel Poly-GalNAcβ1–4GlcNAc (LacdiNAc) and Fucosylated Poly-LacdiNAc N-Glycans from Mammalian Cells Expressing β1,4-N-Acetylgalactosaminyltransferase and α1,3-Fucosyltransferase. Journal of Biological Chemistry *280*, 12810–12819. 10.1074/JBC.M414273200.

65. Yan, S.B., Chao, Y.B., and Van Halbeek, H. (1993). Novel Asn-linked oligosaccharides terminating in GalNAc beta (1-->4)[Fuc alpha (1-->3)]GlcNAc beta (1-->.) are present in recombinant human protein C expressed in human kidney 293 cells. Glycobiology *3*, 597–608. 10.1093/GLYCOB/3.6.597.

66. Costa, J., Gatermann, M., Nimtz, M., Kandzia, S., Glatzel, M., and Conradt, H.S. (2018). N-Glycosylation of Extracellular Vesicles from HEK-293 and Glioma Cell Lines. Anal Chem *90*, 7871–7879. 10.1021/ACS.ANALCHEM.7B05455.

67. Green, E.D., Adelt, G., Baenziger, J.U., Wilson, S., and Van Halbeek, H. (1988). The asparagine-linked oligosaccharides on bovine fetuin. Structural analysis of N-glycanase-released oligosaccharides by 500-megahertz 1H NMR spectroscopy. Journal of Biological Chemistry *263*, 18253–18268. 10.1016/s0021-9258(19)81354-6.

68. Tsuchida, A., Okajima, T., Furukawa, K., Ando, T., Ishida, H., Yoshida, A., Nakamura, Y., Kannagi, R., Kiso, M., and Furukawa, K. (2003). Synthesis of disialyl Lewis a (Le(a)) structure in colon cancer cell lines by a sialyltransferase, ST6GalNAc VI, responsible for the synthesis of alpha-series gangliosides. J Biol Chem *278*, 22787–22794. 10.1074/JBC.M211034200.

69. Senda, M., Ito, A., Tsuchida, A., Hagiwara, T., Kaneda, T., Nakamura, Y., Kasama, K., Kiso, M., Yoshikawa, K., Katagiri, Y., et al. (2007). Identification and expression of a sialyltransferase responsible for the synthesis of disialylgalactosylgloboside in normal and malignant kidney cells: downregulation of ST6GalNAc VI in renal cancers. Biochem J *402*, 459–470. 10.1042/BJ20061118.

70. Bos, P.D., Zhang, X.H.F., Nadal, C., Shu, W., Gomis, R.R., Nguyen, D.X., Minn, A.J., Van De Vijver, M.J., Gerald, W.L., Foekens, J.A., et al. (2009). Genes that mediate breast cancer metastasis to the brain. Nature *459*, 1005–1009. 10.1038/NATURE08021.

71. Okajima, T., Fukumoto, S., Ito, H., Kiso, M., Hirabayashi, Y., Urano, T., Furukawa, K., and Furukawa, K. (1999). Molecular Cloning of Brain-specific GD1α Synthase (ST6GalNAc V) Containing CAG/Glutamine Repeats. Journal of Biological Chemistry *274*, 30557–30562. 10.1074/JBC.274.43.30557.

72. Llop, E., Gutiérrez-Gallego, R., Segura, J., Mallorquí, J., and Pascual, J.A. (2008). Structural analysis of the glycosylation of gene-activated erythropoietin (epoetin delta, Dynepo). Anal Biochem *383*, 243–254. 10.1016/J.AB.2008.08.027.

73. Matrosovich, M., Herrler, G., and Klenk, H.D. (2015). Sialic Acid Receptors of Viruses. SialoGlyco Chemistry and Biology II *367*, 1. 10.1007/128_2013_466.

74. Paidas, M.J., Mohamed, A.B., Norenberg, M.D., Saad, A., Barry, A.F., Colon, C., Kenyon, N.S., and Jayakumar, A.R. (2021). Multi-Organ Histopathological Changes in a Mouse Hepatitis Virus Model of COVID-19. Viruses *13*. 10.3390/V13091703.

75. Baumann, A.M.T., Bakkers, M.J.G., Buettner, F.F.R., Hartmann, M., Grove, M., Langereis, M.A., De Groot, R.J., and Muhlenhoff, M. (2015). 9-O-Acetylation of sialic acids is catalysed by CASD1 via a covalent acetyl-enzyme intermediate. Nature Communications 2015 6:1 *6*, 1–12. 10.1038/ncomms8673.

76. Visser, E.A., Moons, S.J., Timmermans, S.B.P.E., de Jong, H., Boltje, T.J., and Büll, C. (2021). Sialic acid O-acetylation: From biosynthesis to roles in health and disease. J Biol Chem *297*. 10.1016/J.JBC.2021.100906.

77. Pabst, M., and Altmann, F. (2008). Influence of electrosorption, solvent, temperature, and ion polarity on the performance of LC-ESI-MS using graphitic carbon for acidic oligosaccharides. Anal Chem *80*, 7534–7542. 10.1021/ac801024r.

78. Javaud, C., Dupuy, F., Maftah, A., Julien, R., and Petit, J.M. (2003). The fucosyltransferase gene family: An amazing summary of the underlying mechanisms of gene evolution. Genetica *118*, 157–170. 10.1023/A:1024101625214/METRICS.

79. Noro, E., Togayachi, A., Sato, T., Tomioka, A., Fujita, M., Sukegawa, M., Suzuki, N., Kaji, H., and Narimatsu, H. (2015). Large-scale identification of N-glycan glycoproteins carrying lewis x and site-specific N-glycan alterations in fut9 knockout mice. J Proteome Res *14*, 3823–3834. 10.1021/ACS.JPROTEOME.5B00178/ASSET/IMAGES/LARGE/PR-2015-00178Y_0008.JPEG.

80. Domino, S.E., Hiraiwa, N., and Lowe, J.B. (1997). Molecular cloning, chromosomal assignment and tissue-specific expression of a murine alpha(1,2)fucosyltransferase expressed in thymic and epididymal epithelial cells. Biochemical Journal *327*, 105. 10.1042/BJ3270105.

81. Morris, H.R., Dell, A., Easton, R.L., Panico, M., Koistinen, H., Koistinen, R., Oehninger, S., Patankar, M.S., Seppala, M., and Clark, G.F. (1996). Gender-specific glycosylation of human glycodelin affects its contraceptive activity. J Biol Chem *271*, 32159–32167. 10.1074/JBC.271.50.32159.

82. Appelmelk, B.J., van Die, I., van Vliet, S.J., Vandenbroucke-Grauls, C.M.J.E., Geijtenbeek, T.B.H., and van Kooyk, Y. (2003). Cutting edge: carbohydrate profiling identifies new pathogens that interact with dendritic cell-specific ICAM-3-grabbing nonintegrin on dendritic cells. J Immunol *170*, 1635–1639. 10.4049/JIMMUNOL.170.4.1635.

83. Pang, P.C., Haslam, S.M., Dell, A., and Clark, G.F. (2016). The human fetoembryonic defense system hypothesis: Twenty years on. Mol Aspects Med *51*, 71–88. 10.1016/J.MAM.2016.06.002.

84. Bhaumik, M., Seldin, M.F., and Stanley, P. (1995). Cloning and chromosomal mapping of the mouse Mgat3 gene encoding N-acetylglucosaminyltransferase III. Gene *164*, 295–300. 10.1016/0378-1119(95)00260-D.

85. Vos, G.M., Hooijschuur, K.C., Li, Z., Fjeldsted, J., Klein, C., de Vries, R.P., Toraño, J.S., and Boons, G.J. (2023). Sialic acid O-acetylation patterns and glycosidic linkage type determination by ion mobility-mass spectrometry. Nature Communications 2023 14:1 *14*, 1–13. 10.1038/s41467-023-42575-x.

86. Wickham, H., Averick, M., Bryan, J., Chang, W., D', L., Mcgowan, A., François, R., Grolemund, G., Hayes, A., Henry, L., et al. (2019). Welcome to the Tidyverse. J Open Source Softw *4*, 1686. 10.21105/JOSS.01686.

87. Galili, T. (2015). dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. Bioinformatics *31*, 3718–3720. 10.1093/BIOINFORMATICS/BTV428.

88. Watanabe, Y., Aoki-Kinoshita, K.F., Ishihama, Y., and Okuda, S. (2021). GlycoPOST realizes FAIR principles for glycomics mass spectrometry data. Nucleic Acids Res *49*, D1523–D1528. 10.1093/NAR/GKAA1012.