

Ultradeep N-glycoproteome Atlas of Mouse Reveals Spatiotemporal Signatures of Brain Aging and Neurodegenerative Diseases

Corresponding Author: Dr Huali Shen

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The study proposed by Pang et al. presents the largest mouse N-glycoproteome dataset to date with different enzymes and enrichment methods, offering a diverse resource for glycoproteomics research. This is certainly useful for computational glycoproteomics at the AI age.

Here are my major comments:

1. Although the authors claimed the largest dataset in glycoproteomics, it is still unknown that if this can improve the AI models for MS/MS or retention time prediction. My suggestion is to retrain existing models such as GlycoDeep (Yi et al. Nat. Commun. 2024) or DeepGP (Yu et al. BioRxiv 2024), and compare the prediction performance. This is the most important concern.
2. Comparing four software tools for glycoproteomics is insightful. However, discussing the potential biases of each software and how these might affect data interpretation would add depth to the analysis. For example, do the search engines easier to lose the IDs of peptide or glycans? What about the biases of peptide lengths, glycan lengths, glycan types and so on.
3. The creation of the N-GlycoMiner database is a valuable contribution. However, details about the user interface and accessibility for researchers unfamiliar with glycoproteomics would help readers understand its usability.
4. It will be useful to include AlphaFold2's predicted structures into N-GlycoMiner database for visualization.
5. The study mentions that more than 88.4% of GPSMs remained unidentified, indicating limitations in current methods. A more detailed discussion of how future improvements in MS or AI-based spectral recognition could overcome these limitations can be discussed and would provide a balanced perspective.

Minor comments:

1. In line 501 page 17: pGlyco-N-Human.gdb was selected but for mouse samples, the glycan database must be mouse-related.
2. In Fig. 4d, there are two "NeuGc containing" items.
3. In Fig. 1 second panel, it may be not adequate to say it is "spatial" specificity. Tissue specificity is definitely OK.

Reviewer #2

(Remarks to the Author)

This is a mixed bag of high quality large glycoproteomic datasets generated by a group dedicated to the development and

applications of glycoproteomics platform. Notably, they compared the glycoproteomics data derived from 5 different mouse tissues, digested with three tryptic combinations, two enrichment methods, and data analyzed using 4 different software. This is further accompanied by glycoproteomics data on the mouse brains from control vs Parkinson and Alzheimer disease models, at young vs old stages. A website holding these data was set-up and presented.

Overall, it is, however, somewhat unfocused, attempting to highlight a few technical aspects based on the comparisons made on different sample prep and data analysis strategies but did not go far enough to provide better insights. Similarly, it also tried to highlight the differences found in the glycoproteomes from different biological sources, including tissue-specific and disease-related changes, but what biological insights can be inferred is not apparent. This work is rich in data and different variations of bioinformatics analysis. It may indeed serve its purpose by generating a large dataset that can facilitate current endeavors in machine learning and developing AI-driven software. The N-GlycoMiner website constructed from this work and literature is a useful ref source worth publicizing. However, from glycobiology perspective, as well as advancement in glycoproteomics itself, there is very little to savor.

A few examples of why this Reviewer thinks the authors did not take their work a step further and seemingly content instead to just describe their bioinformatic analysis and diagram generated.

i) three tryptic combinations and 2 enrichment methods. These seem just to increase the number of IDs by multiple runs instead of trying to gain insight into their underlying preference or bias.

ii) 4 different software for glycopeptide ID - just provide the number of IDs and reinforce the ideas that MS-Fragger and GlycoDecipher would give much higher number of IDs relative to pGlyco and StrcuGP, but did not provide an analysis of why. Confidence is assumed when all 4 identified the same glycopeptide, whereas when each identified a different glycopeptide from the same gPSM, it is considered ambiguous. Moderate level of confidence when the same glycopeptide was identified by 2 or 3 of the software but not consistently so across the board. This is all good and provides a basis for grading the confidence level in generating the database but it does not provide a further analysis of the bias or blindspots of each software for others to learn. It is unclear why the authors then decided to use pGlyco3 - is it a rationale decision made on the basis that it seems to be the most reliable despite its lower number of ID?

iii) It is interesting to show that only a single glycosite and 2 glycoforms were ID for CD36 from brain but that from heart has 7 glycosites and over 200 glycoforms ID. However, it is unclear the biological or technical reasons contributing to this difference. Only one glycosite identified does not mean the other sites are necessarily not occupied. The complexity of the glycopeptides at each sampling elution timepoint may differ significantly between the two sources of samples, which may also have very different protein expression level. What exactly were the respective glycoforms identified was not provided. The authors simply made a statement that "The distinct differences of glycosylation on CD36 implied the importance of glycosylation for the tissue specific functions of CD36".

Other minor issues:

Fig 2C. Unclear what is co-spectra identification ratios for different levels of protein structure - please clarify. What's the significance?

Fig 4, what's the difference between panel c and e? This is not clear from the description provided in the Figure legend although one can infer that from the manuscript. In panel d, one of the two NeuGc containing label should be NeuAc containing. Percentage of high mannose seems to be exceptionally low. 0% for brain?

The Suppl Tables, esp Suppl Table 2, should be better described.

Reviewer #3

(Remarks to the Author)

The manuscript from Shen and co-workers provides a very large collection of N-linked glycoproteomics, as well as a searchable database. The authors use multiple methods and both the biochemical and data analysis steps to arrive at a very robust set of N-linked glycans and glycoproteins. They do a very thorough job characterizing these data and apply their approach to find some differences in aging mice, as well as AD and PD models. I believe that this dataset will be a useful one for the community, as well as for training better data analysis programs. I support its publication.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

This reviewer thanks the authors' efforts to address all the concerns. There is a minor issue: there are no data in the iProx repo.

Reviewer #2

(Remarks to the Author)

The authors have addressed most of the concerns and inadequacies. The revised manuscript has greatly increased its impact by including significant additional data and analysis.

Reviewer #3

(Remarks to the Author)

I was supportive of publishing the first version of this manuscript and am still supportive.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Point-to-point response to reviewers' comments

We thank the Reviewers for your constructive comments that greatly helped us to improve our manuscript. We have revised our manuscript and fully addressed all of your concerns. For more details, please see our point-by-point responses to all of your questions below. All revision changes in the manuscript are marked in highlight yellow color for your convenience.

Reviewer #1 (Remarks to the Author):

The study proposed by Pan et al. presents the largest mouse N-glycoproteome dataset to date with different enzymes and enrichment methods, offering a diverse resource for glycoproteomics research. This is certainly useful for computational glycoproteomics at the AI age.

Here are my major comments:

1. Although the authors claimed the largest dataset in glycoproteomics, it is still unknown that if this can improve the AI models for MS/MS or retention time prediction. My suggestion is to retrain existing models such as DeepGlyco (Yi et al. Nat. Commun. 2024) or DeepGP (Yu et al. BioRxiv 2024), and compare the prediction performance. This is the most important concern.

Responses: We appreciate the reviewer's insightful suggestion regarding the potential of our dataset to enhance AI models for MS/MS spectra and retention time predictions. According to your suggestion, we have retrained the existing models and compared the prediction performance:

The glycoproteomic spectra library and dataset produced in this study not only significantly expanded the scale of the existing database, but also classified the glycopeptide spectra and their identification results into different quality/confidence levels according to the consistency of identification results of different search software. The spectra with consistent identification results obtained by multiple search software are relatively of higher quality, and their identification results are more reliable. This study provides not only a larger amount of glycopeptide data but also high-quality glycopeptide spectra and high-confidence identification results, which can partially make up for the current lack of glycopeptide standard libraries and provide reliable, high-quality training data for AI model training.

We first evaluated whether the different confidence datasets generated in this study could enhance the prediction performance of the DeepGP model¹. Since the search results reported by pGlyco3, which provide possible glycan structural annotations (PlausibleStruct), were compatible with DeepGP, we selected pGlyco3 subsets and partitioned them based on

confidence levels for model training. We trained the model with datasets of ambiguous, low, moderate, 90% of high-confidence data (high_train), as well as a combination of moderate and high_train, with remaining 10% of high-confidence data reserved for testing. We ensured that no glycopeptides overlapped between the training and testing sets. The results demonstrated that data with moderate confidence and above significantly improved the cosine similarity between the predicted and experimental mass spectra, with the median for high_train + moderate reaching 0.95, exceeding previously reported performance (**Figure 1a**). In contrast, training with low and ambiguous datasets resulted in lower performance, with median cosine similarity of only 0.88 and 0.89, respectively. The performance of the moderate-confidence dataset is comparable to that of the high-confidence dataset, likely due to its larger size (60k vs. 26k glycopeptides) and relatively high data quality, as it was co-identified by two or three different search engines. In addition, the model trained on moderate+high_train data, in the absence of RT calibration, achieved higher Pearson correlation coefficients (PCC) of iRT prediction than the previously reported performance (**Figure 1b**). This comparison shows that both the size and quality of the training dataset have important impacts on the training effect of the model.

We also evaluated our data using DeepGlyco model, performing 5-fold cross-validation on the high-confidence dataset, where four parts were used for training and one part for validation in each iteration (**Figure 1c**)². The median dot product (DP) for the whole glycopeptides was consistently above 0.986, and exceeded 0.992 for the glycan parts across all subsets. Furthermore, the spectral angle loss (SA) for glycopeptides ranged from 0.102 to 0.108, while for the glycan parts, it ranged from 0.078 to 0.083. The results were stable, indicating superior data quality and representativeness. Next, we trained the model using the low and moderate datasets, testing on the same high-confidence data aligning with shown in **Figure 1d**. Although the similarity was slightly lower than high-confidence results mentioned above, it still surpassed the previously reported benchmarks, with SA of 0.150 and DP of 0.972 at the whole spectrum level for the holdout set².

Furthermore, we performed cross-validation partitioning the high-confidence dataset based on different enzymatic patterns. Specifically, we trained the model using data from any two enzymes and validated it against the third one. The similarities dropped marginally, yet the performance remained strong (DP of glycopeptide: 0.983-0.992, SA: 0.081-0.116), reflecting the model's generalization ability (**Figure 1e**).

In summary, we demonstrated that using a larger dataset generated in this study, particularly the high-quality dataset with consistent identifications from multiple search engines, can aid in training AI models. Moderate and above datasets significantly improve the

performance of AI models for MS/MS or retention time prediction. We demonstrated that it is essential to classify the spectra based on confidence using multiple software.

We have incorporated these results into the manuscript (lines 174-201), updated Figures 2e, f, g, and Figure S10 accordingly, and added Supplementary note 3 and Supplementary data 3.

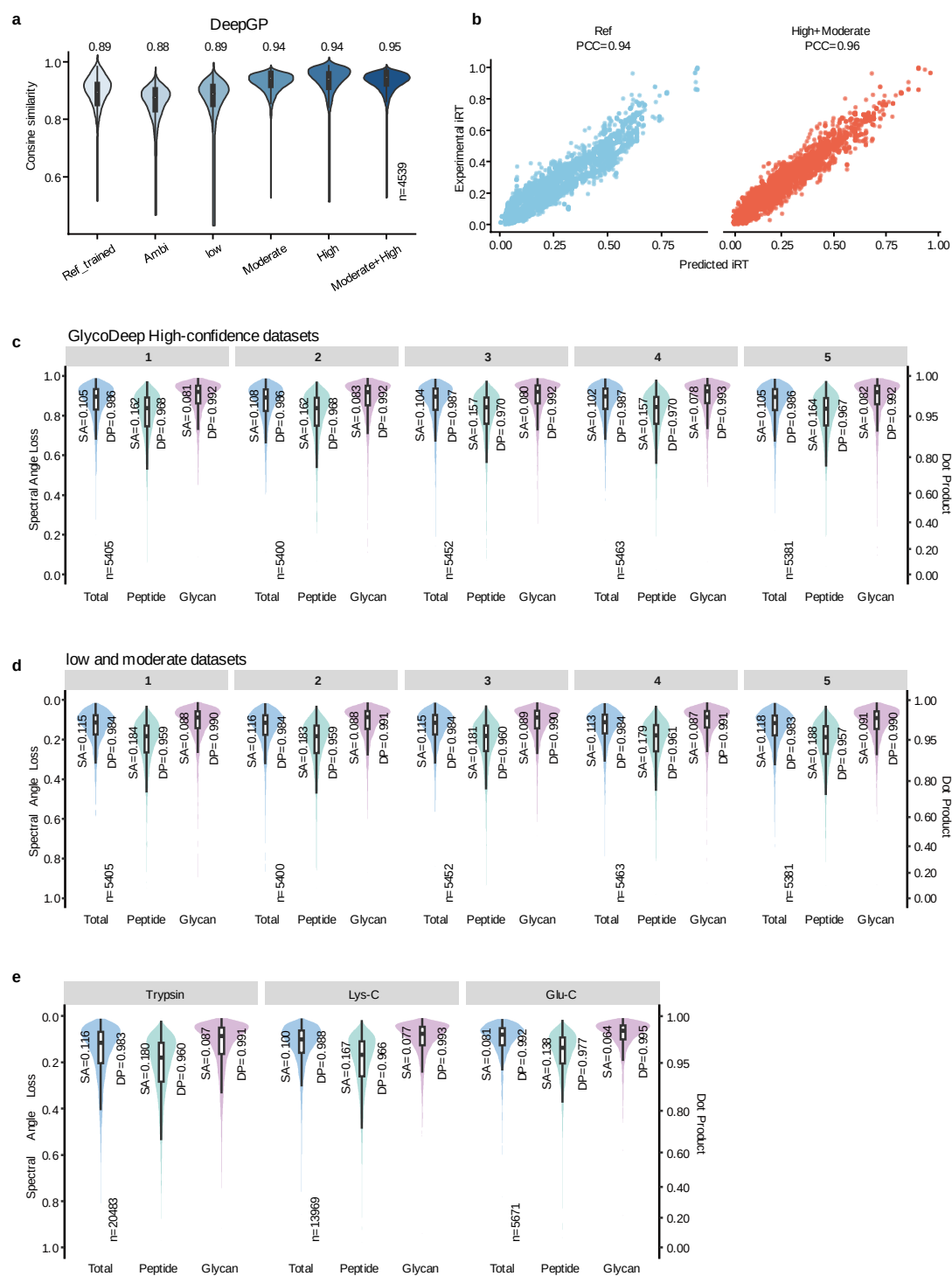


Figure 1. Enhancement of AI Training Performance with Datasets of Varying Quality (a) Distribution of cosine similarities between the predicted and experimental spectra for different training datasets using DeepGP model: literature-reported results, ambiguous, low, moderate, high_train, and high_train+moderate subsets. The remaining high-confidence subset (4539 spectra) were utilized for testing. The median cosine similarities for each group are indicated above. (b) Correlation between the predicted and experimental iRT is presented, with Pearson correlation coefficients (PCC) shown above. Predictions were made using model provided in the literature and the model trained on high and moderate-confidence data, without RT calibration. (c) Distribution of spectral similarities for high-confidence dataset in 5-fold cross validation using DeepGlyco model. (d) Distribution of spectral similarities using DeepGlyco trained on the low and moderate datasets and tested on the same dataset as in Fig C. (e) Distribution of spectral similarities in cross-validation, with the high-conf dataset partitioned according to different enzymatic patterns.

2. Comparing four software tools for glycoproteomics is insightful. However, discussing the potential biases of each software and how these might affect data interpretation would add depth to the analysis. For example, do the search engines easier to loss the IDs of peptide or glycans? What about the biases of peptide lengths, glycan lengths, glycan types and so on.

Responses: Thank you for your valuable comment. We agree that understanding the biases inherent in glycoproteomics software tools is essential for accurate data interpretation. Below, we tried to discuss some of the potential biases and their impact on glycoproteomic analyses:

(1) Bias in identification numbers of the four software across different ID levels. At the GPSM, precursor, and glycoform levels (**Figure 2**), the ranking of identification numbers is almost consistent across all tools, with Glyco-decipher achieving the highest, followed by MSFragger-Glyco, pGlyco3, and finally StrucGP, which has the lowest. However, at other identification levels, including glycosites, glycoproteins and glycan compositions, the results vary significantly. pGlyco3 exhibits the lowest identification numbers for glycosites and glycoproteins but performs moderately well in other categories, notably surpassing MSFragger-Glyco in glycan identification. StrucGP, on the other hand, performs comparably to MSFragger-Glyco and Glyco-decipher at the glycosite and glycoprotein levels. This analysis underscores the trade-offs between the tools in their ability to identify glycoproteins, glycosites, and glycans, highlighting the influence of their design and analytical focus on identification performance.

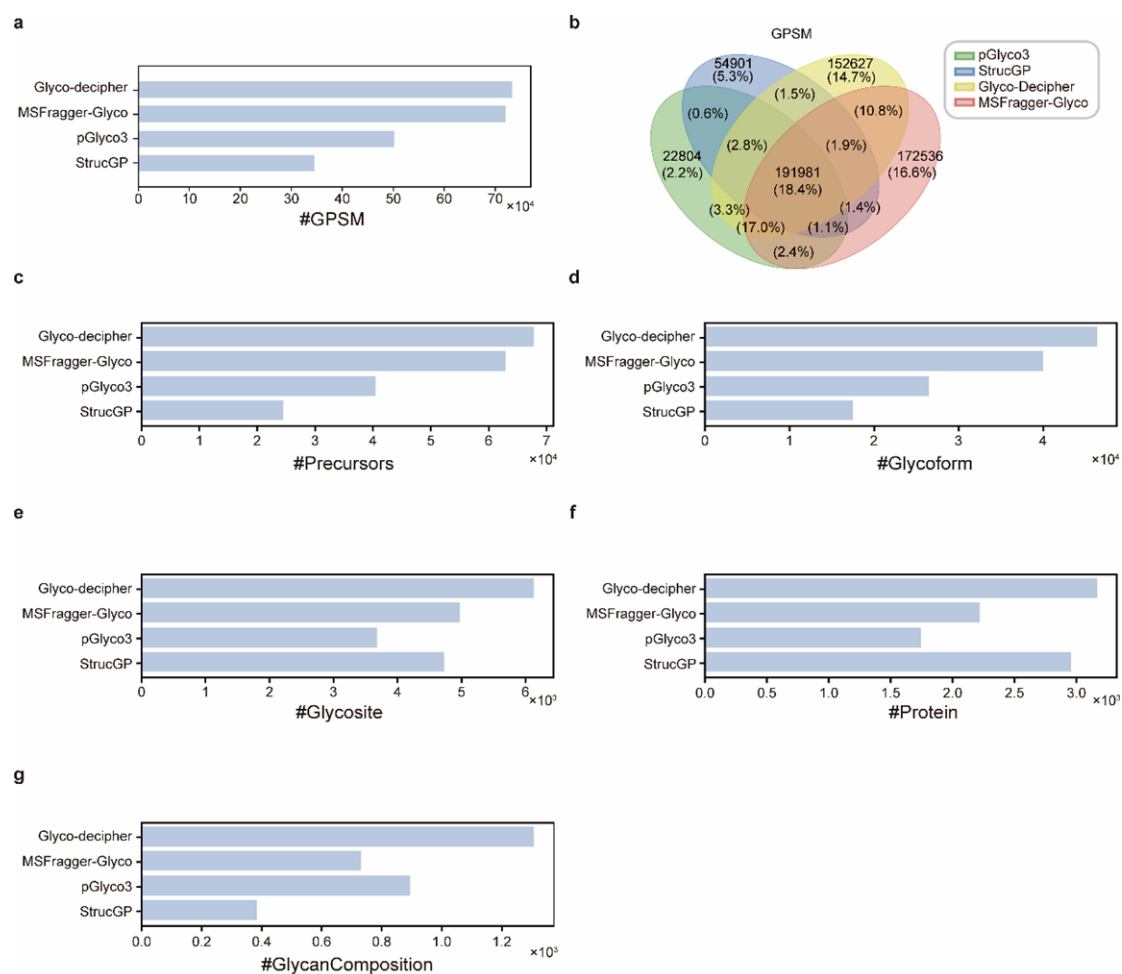


Figure 2. Comparison of Identification Numbers from Different Glycoproteomics Software Platforms. (a) Bar plot showing the total number of glycopeptide-spectrum matches (GPSM) identified by each of the four software tools: Glyco-Decipher, MSFragger-Glyco, pGlyco3, and StrucGP. (b) Venn diagram showing the overlap in the number of GPSMs identified by the four software tools, with percentages indicating the shared and unique identifications. (c, d, e, f, g) Bar plot showing the number of precursors, glycoforms, glycosites, proteins and glycan compositions identified by each software.

(2) The analysis of inconsistent identifications showed the bias in identification confidence across the four software tools. For the spectra identified by all four tools (191,981 in Figure 2b), 160,928 spectra were consistently identified as the same precursors by all tools, while 31,053 spectra showed inconsistencies. MSFragger-Glyco contributed to the highest proportion of inconsistent identifications (35.6%), followed by Glyco-decipher (18.1%) and StrucGP (12.8%). In contrast, pGlyco3 had the lowest inconsistency rate (4.2%) (Figure 3a). For the spectra identified by any three of the four tools, MSFragger-Glyco still exhibited the largest inconsistency (13.3%), followed by Glyco-decipher (3.7%) and pGlyco3 (1.7%) (Figure 3b).

These results suggest that pGlyco3 is the most reliable tool for consistent spectrum identification. On the other hand, MSFragger-Glyco identifies a larger number of spectra with greater inconsistency. The high inconsistency rate of MSFragger-Glyco indicates that while it has higher sensitivity, it may also lead to more false positives. These differences highlight the varying strengths and biases of these tools, with pGlyco3 excelling in reliability and MSFragger-Glyco providing a broader range of identifications at the expense of consistency.

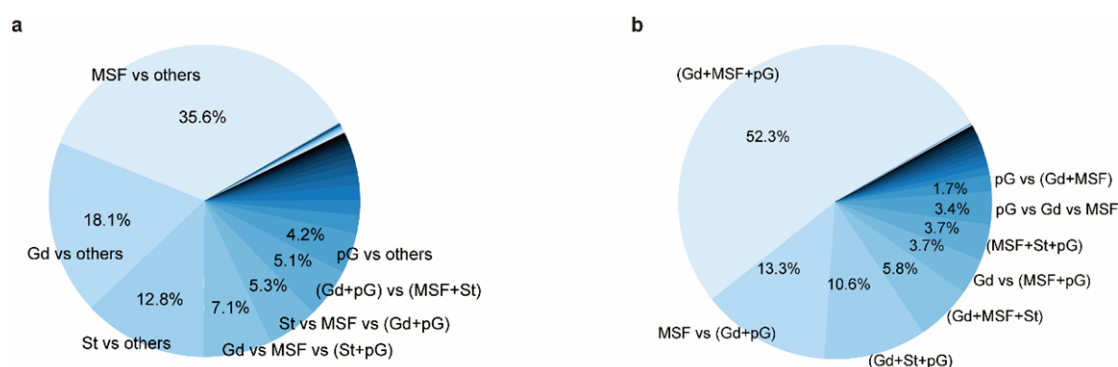


Figure 3. Analysis of Inconsistent Precursor Identifications Across Four Software Tools. (a) The pie chart illustrates the distribution of spectra with inconsistent precursor identifications among four glycoproteomics software tools: MSFragger-Glyco (MSF), pGlyco3 (pG), Glyco-decipher (Gd), and StrucGP (St). It highlights spectra where one tool reports a unique identification or where groups of tools produce distinct results, excluding the cases where all four tools agree. For example, "MSF vs others" indicates spectra where MSFragger-Glyco reported a unique result compared to the other tools, which agreed on a different precursor. (b) The pie chart shows the precursor identification analysis for spectra identified by any three of the four tools. Each section represents the percentage of spectra with matching or differing precursor identifications. For example, "Gd+MSF+pG" represents spectra where Glyco-decipher, MSFragger-Glyco, and pGlyco3 all identified the same precursor, while "MSF vs (Gd + pG)" shows spectra where MSFragger-Glyco identified a different precursor compared to Glyco-decipher and pGlyco3, which agreed on the same one.

(3) The preferences or bias of peptide and glycan features among the four tools. We specifically analyzed the distribution of peptide length (number of amino acids), peptide molecular weight, glycan length (number of monosaccharides) and glycan types across the four software tools (**Figure 4**). StructGP consistently identifies shorter peptides and glycans, whereas MSFragger-Glyco shows a preference for longer peptides, and Glyco-decipher tends to identify longer glycans. Notably, pGlyco3's glycan-length bias was influenced by the exclusion of glycan structures above 20 monosaccharides, as we used the pGlyco-N-Mouse database instead of pGlyco-N-Mouse-Large. In terms of glycan types, StructGP is biased toward high-mannose and pauci-mannose glycans, with the lowest identification of sialylated

glycans, indicating its limited ability to detect complex glycan structures. On the other hand, MSFragger-Glyco exhibits the highest sialic acid content in its identifications. Meanwhile, both pGlyco3 and Glyco-decipher demonstrate a stronger focus on fucosylated glycan identifications. Collectively, these observations highlight distinct differences in the analytical focus and limitations of the tools, emphasizing the need to carefully select software based on the specific requirements of a glycoproteomics study.

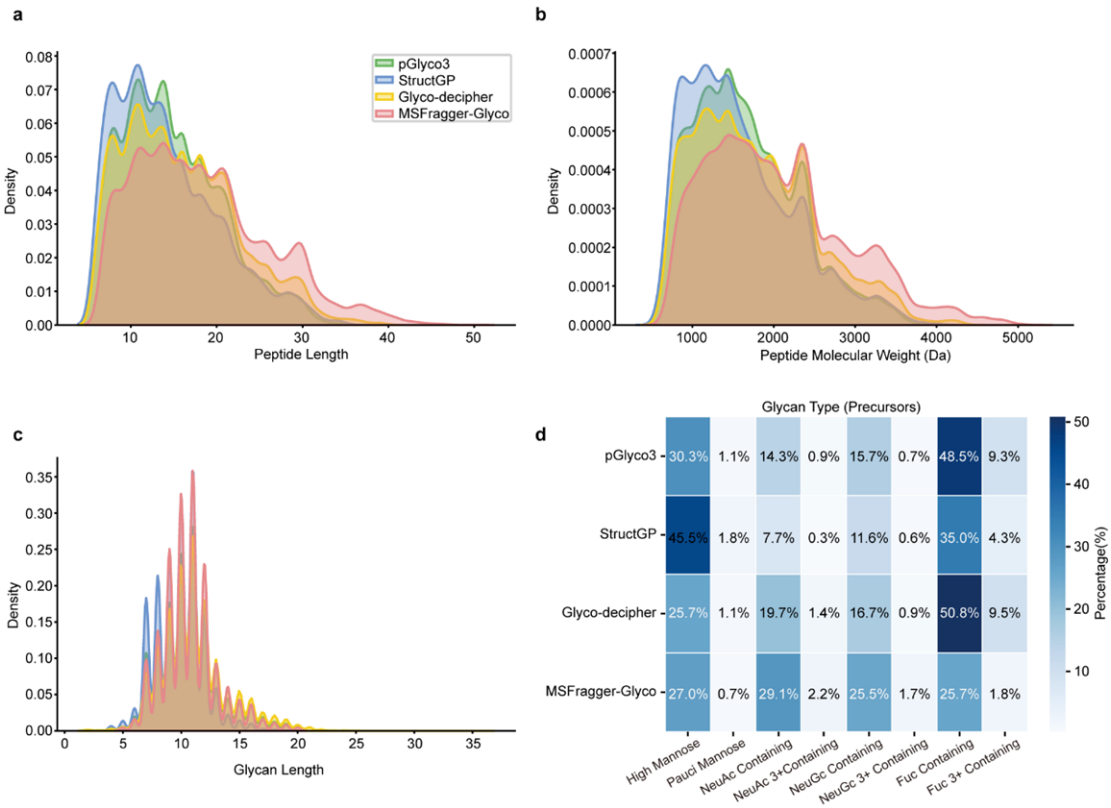


Figure 4. Comparison of glycoproteomics analysis results across different software tools. (a) Density distribution of peptide lengths identified by pGlyco3, StructGP, Glyco-decipher, and MSFragger-Glyco. The x-axis represents peptide length, while the y-axis indicates density. (b) Density distribution of peptide molecular weights (Da) identified by the four tools. (c) Density distribution of glycan lengths identified by the four tools. (d) Proportion of different glycan types (precursors) identified by the four tools. The glycan types are categorized into high mannose, pauci mannose-containing, NeuAc-2+ containing, NeuAc-3+ containing, NeuGc-2+ containing, NeuGc-3+ containing, Fuc-2+ containing, and Fuc-3+ containing. The intensity of the blue color represents the percentage of each glycan type identified by each tool.

(4) The bias of software tools in quantifying glycoproteomics data: a comparison between heart and brain samples. We further compared the quantification results from pGlyco3,

Glyco-decipher, and MSFragger as these three software tools are the only ones offering label-free quantification. We found that the differentially expressed glycoproteins and glycosites showed a high overlap in up-regulated and down-regulated ones, indicating a high consistency in quantification between the three software tools at these levels (**Figure 5a**). However, at the glycan and site-specific glycoform levels, the overlap in quantification results was much lower compared to glycoproteins and glycosites, indicating lower consistency in quantification at these levels.

We further conducted a correlation analysis of quantification values among the three software tools using their commonly identified and quantified glycopeptides (**Figure 5b**). The highly positive correlations in the quantification of glycoproteins and glycosites (median PCC above 0.78) were revealed, suggesting that the three software platforms agree well in quantifying these levels of data. However, the correlation analysis for glycans showed more variability, with lower median PCC (median PCC ranges 0.52-0.67). This suggests that the software tools are generally consistent in quantifying glycoproteins and glycosites, but less agreement at the glycans levels. This inconsistency could be due to differences in how each tool processes and interprets glycan-related data. This lower consistency in glycans and glycoforms may lead to divergent interpretations in subsequent studies of glycosylation's biological functions.

Overall, these findings emphasize the need for caution when comparing glycan and glycoform data across different software platforms, particularly in studies focused on the functional implications of glycosylation. Further improvements in glycan quantification algorithms and validation with non-omics techniques could help address these discrepancies.

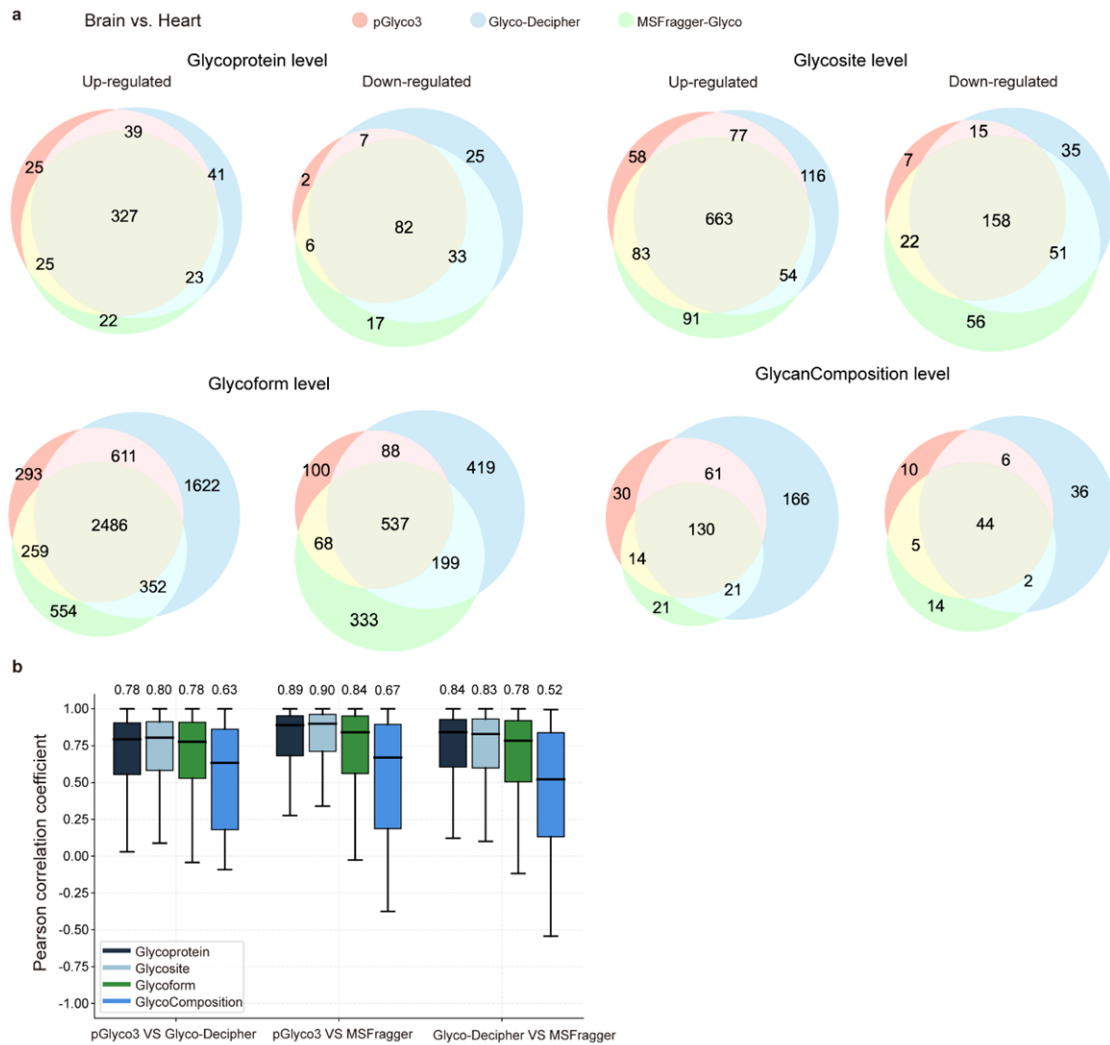


Figure 5. Comparison of Quantitative Glycoproteomics Data Between Brain and Heart Samples Using Different Software Platforms. (a) Venn diagrams showing the overlap of up-regulated and down-regulated glycoproteins, glycoforms, glycosites, and glycan compositions between brain and heart samples, as quantified by three software tools: pGlyco3 (red), Glyco-Decipher (blue), and MSFragger-Glyco (green). (b) Boxplots of Pearson correlation coefficients comparing the glycoprotein, glycosite, glycoform, and glycan composition data commonly identified and quantified by the three software platforms. The comparisons include pGlyco3 vs. Glyco-Decipher, pGlyco3 vs. MSFragger, and Glyco-Decipher vs. MSFragger. Higher correlation values indicate better agreement in the quantification values between the platforms.

(5) Comparison of the four software tools from their identification strategies, glycan and peptide scoring and false positive control. Each software tool utilizes a distinct search strategy for glycoproteomic identification, which impacts their performance in glycan and peptide identification and false positive control.

pGlyco3 adopts a glycan-first search strategy, which prioritizes the identification of N-glycopeptides by first identifying glycans using ion-indexing techniques, followed by peptide identification^{3, 4}. This approach is integrated with scoring systems for glycans, peptides, and glycopeptides. A notable feature of pGlyco3 is its rigorous false discovery rate (FDR) control, which applies across all levels—glycans, peptides, and glycopeptides. pGlyco3 used glycan database combining N-glycans from Glycome-DB, as well as the three largest substructures of high-mannose, hybrid, and complex N-glycans. This glycan-first approach and FDR control allow pGlyco3 to outperform the other tools in terms of accuracy though the strict glycan-level quality control may reduce sensitivity (**Figure 2 and 3**).

StrucGP uses a modular approach, dividing N-glycans into three modules, with each identified based on distinct Y ion patterns or combinations of B/Y ions⁵. This methodology facilitates detailed glycan structure determination. FDR estimations for both the peptide and glycan portions of glycopeptides are performed using both decoy database and decoy spectrum approaches. While StrucGP performs fewer glycan identifications compared to other tools, it offers similar performance to MSFragger-Glyco and Glyco-Decipher in terms of glycosite and glycoprotein identification (**Figure 2**). This is likely because StrucGP requires more ion information for detailed glycan structure analysis, which makes it more prone to identifying shorter peptides and glycans, especially high-mannose glycan types (**Figure 4**).

MSFragger utilizes an open mass offset search strategy, where glycopeptide spectra are analyzed based on possible glycan compositions derived from mass offsets. It performs deisotoping and decharging of spectra prior to analysis, improving the sensitivity of the tool, particularly for high-mass glycopeptides (**Figure 4**). However, MSFragger does not place a strong emphasis on glycan-level quality control, which may lead to less precise glycan identification compared to other tools⁶. The high sensitivity of MSFragger often results in a larger number of identifications, but it has the lowest consistency in identifications compared to the other software (**Figure 2 and 3**).

Glyco-Decipher employs a glycan database-independent peptide matching strategy, which uses shared fragmentation patterns of peptide backbones in glycopeptides to improve spectrum interpretation⁷. This technique enables Glyco-Decipher to extend its glycopeptide identification capabilities through the analysis of peptide backbone fragments. Glyco-Decipher has an advantage in identification numbers due to its spectrum expansion method.

In summary, the differences in search strategy, glycan identification and quality control across software tools have a substantial impact on their overall performance, particularly in terms of identification consistency and accuracy.

For this revision, we have added these analyses into the manuscript (lines 123-152, 160-164), updated Figure S6-9 accordingly, and added Supplementary note 2.

3. The creation of the N-GlycoMiner database is a valuable contribution. However, details about the user interface and accessibility for researchers unfamiliar with glycoproteomics would help readers understand its usability.

Responses: We thank the reviewer for recognizing the value of the N-GlycoMiner database and for highlighting the need for more information on its user interface and accessibility. We agree that providing a clearer description of these aspects will improve the reader's understanding of its usability. In response, we have prepared a step-by-step PowerPoint presentation to guide users in navigating the website. Please refer to the attached supplementary data 8.

4. It will be useful to include AlphaFold2's predicted structures into N-GlycoMiner database for visualization.

Responses: Based on the reviewers' suggestions, we have linked the protein structures predicted by AlphaFold2 to our website for visualization. Additionally, the glycoprotein sites and glycan compositions have also been visualized on the website (**Figure 6**). This was mentioned in manuscript line 398-404 and Supplementary Data 8.

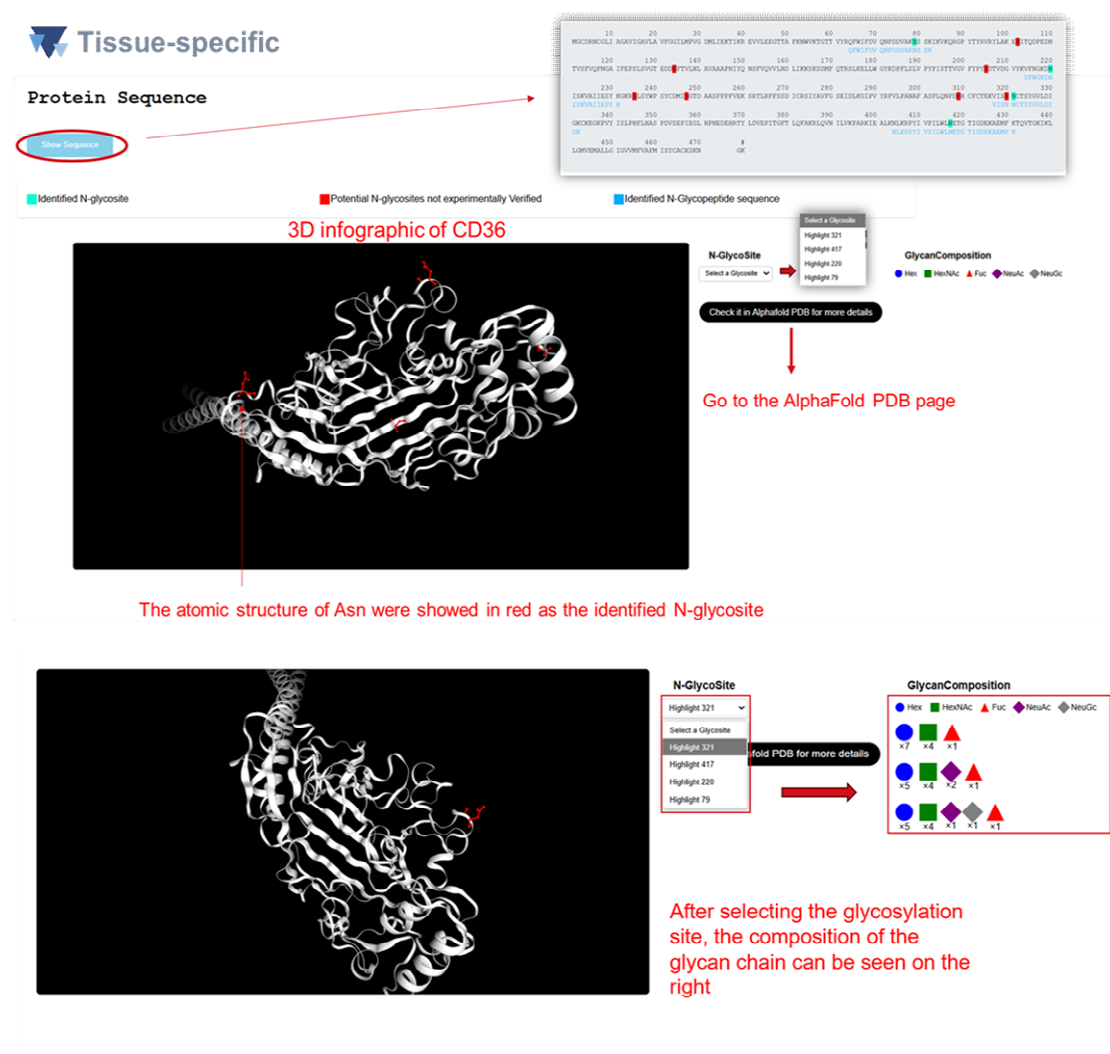


Figure 6. Visualization of glycoprotein CD36 using the N-GlycoMiner database, incorporating AlphaFold2-predicted structures. The top panel shows the protein sequence with identified N-glycosites (green), potential N-glycosites not experimentally verified (red), and identified N-glycopeptide sequences (blue). Users can access this view by selecting "Show Sequence." A 3D infographic of CD36 (middle panel) highlights identified asparagine (Asn) residues in red, corresponding to N-glycosites. The dropdown menu allows users to select specific glycosylation sites for further exploration. By linking to the AlphaFold PDB page, users can retrieve detailed structural information. The bottom panel demonstrates that after selecting a specific N-glycosite (e.g., Highlight 321), the glycan chain composition is displayed on the right. Glycan compositions are represented with color-coded symbols for Hex, HexNAc, Fuc, and NeuAc, enabling a clear visualization of site-specific glycosylation patterns. This integration facilitates detailed exploration of glycoprotein structures and their glycosylation features. This feature-rich visualization enhances the interpretation of glycosylation patterns and supports deeper insights into glycoprotein function.

5. The study mentions that more than 88.4% of GPSMs remained unidentified, indicating limitations in current methods. A more detailed discussion of how future improvements in MS or AI-based spectral recognition could overcome these limitations can be discussed and would provide a balanced perspective.

Responses: We first examined the quality of the unidentified spectra to better understand the challenges. Through manual inspection, we found that approximately 50% of the unidentified glycopeptide spectra exhibited high-quality fragmentation patterns with abundant ion information. Despite their rich data, these spectra were not successfully identified by all the software (**Figure 7a**). The inability of current algorithms to identify these high-quality spectra suggests that improvements are needed in the software's ability to interpret complex fragmentation data. On the other hand, around 50% of the unidentified spectra were of low quality, with minimal fragment ion information. These low-quality spectra are often the result of insufficient fragmentation efficiency, either in the glycan or peptide backbone, during the MS/MS process (**Figure 7b**).

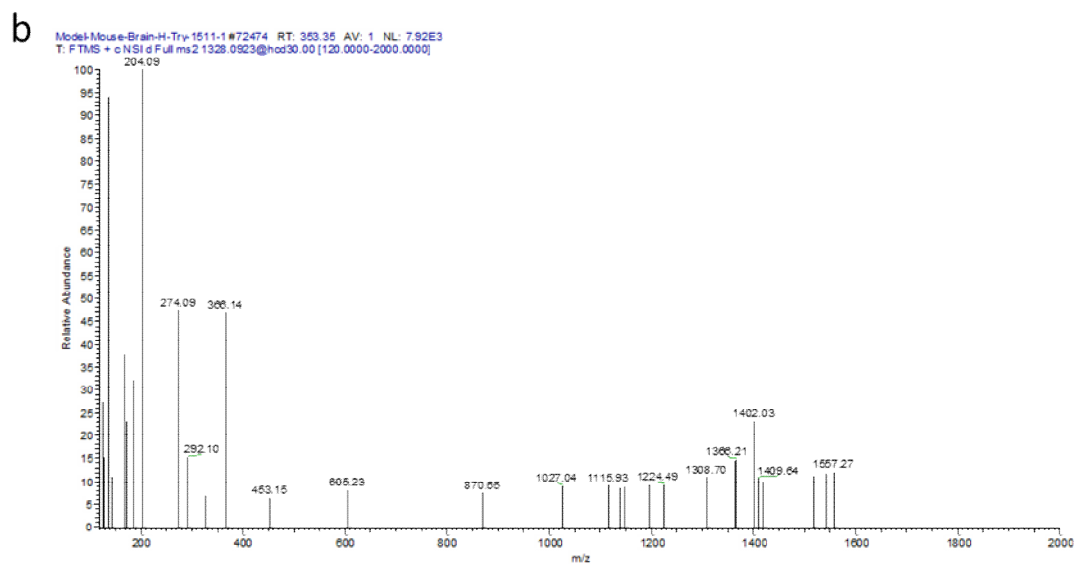
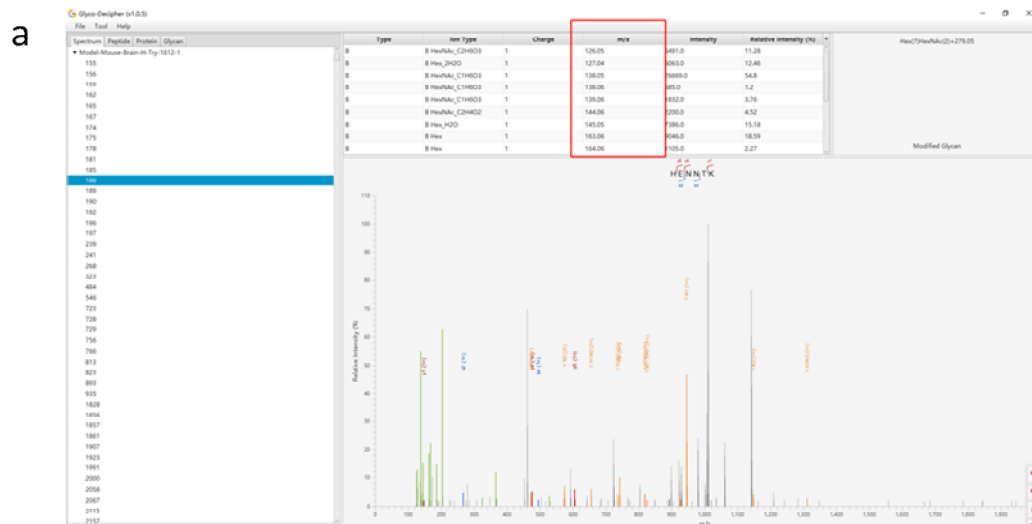


Figure 7. Example spectra for unidentified spectra. (a) An example spectrum with abundant ion information, but the glycopeptide was not identified by any software. (b) A low-quality spectrum showing limited ion information, which makes it difficult to accurately identify the glycopeptide.

For the glycopeptide spectra with rich fragment ion information that remained unidentified, we believe the inability to identify them could stem from several factors:

(1) Modifications on glycans: The presence of specific glycan modifications, such as high-mannose phosphorylated modifications, sulfation, O-acetylation of sialic acid, or methylation, could complicate the accurate identification of the glycan composition by software⁸. These modifications may alter the spectral profile in ways that make it difficult for existing algorithms to correctly identify the glycan structures. Additionally, new glycan modifications are continually being discovered, and there may be other unknown modifications contributing to the challenges in identification.

(2) Complex or rare glycan structures: Certain glycan structures, such as Galili-type polysialic acid (polysia) and poly-LacNAc, are particularly complex or rare^{9, 10}. These large or intricate structures may not be included in some glycan databases used by the search software, leading to missed identifications.

(3) Mixed Spectra: Another possible explanation for unidentified spectra is mixed spectra, where fragment ions originate from multiple glycopeptides rather than a single glycopeptide⁴. This mixed fragmentation pattern could confuse the software and prevent accurate identification of the glycopeptide sequence.

(4) O-Glycopeptides: There is also the possibility that a significant number of O-glycopeptides are present in the sample. Since O-glycan removal was not performed during sample preparation, the identification of O-glycopeptides is expected. We did search for O-glycopeptides using pGlyco3 and successfully identified a significant number of O-glycopeptide spectra, even though HCD fragmentation could not pinpoint the glycan attachment sites.

To improve the identification rate of these remaining unidentified spectra, several strategies can be considered:

(1) Optimization of fragmentation conditions: Although the stepped collision energy HCD fragmentation employed in this study is effective for most glycopeptides, optimizing the HCD energy for certain complex glycopeptides could lead to more complete fragmentation, thus improving identification rates. For glycopeptides where the peptide portion is not detected efficiently by HCD, alternative fragmentation methods may provide better results. These include activated ion Electron Transfer

Dissociation (AI-ETD), Electron Transfer Higher-energy Collision Dissociation (ET_hCD), Electron Transfer-Collision Induced Dissociation (ET_{ci}D), as well as advanced instruments like TimsTOF Pro equipped with Trapped Ion Mobility Spectrometry (TIMS), ZonoTOF 7600 equipped with ECD and Orbitrap Astral instruments with fast speed.

(2) Assisted identification using predicted spectral libraries: As demonstrated by Yu et al¹, the use of spectral libraries predicted by deep learning methods (e.g., DeepGP) can significantly enhance identification accuracy. In their study, the predicted library contributed to the identification of 7,472 additional spectra, demonstrating the potential of such approaches to improve identification rates.

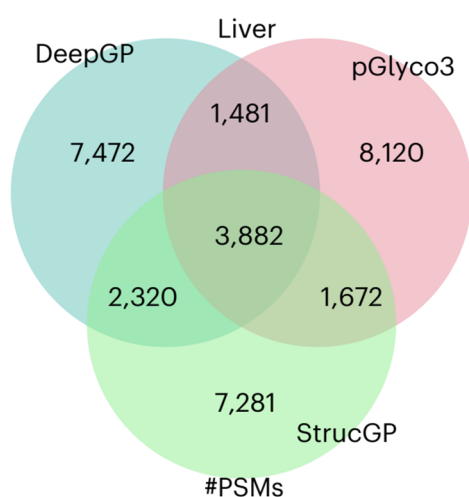


Figure 8. Venn diagram from literature demonstrating that AI-predicted spectral libraries can improve the identification of previously unidentified spectra by current glycopeptide identification software¹.

We have incorporated these discussions into the manuscript (lines 202-211,438-455), and added Figure S11.

Minor comments:

1. In line 501 page 17: pGlyco-N-Human.gdb was selected but for mouse samples, the glycan database must be mouse-related.

Responses: We sincerely thank the reviewer for pointing out the potential discrepancy regarding the glycan database used for the mouse samples. The incorrect mention of pGlyco-N-Human.gdb in the manuscript has been corrected to reflect the accurate database (pGlyco-

N-Mouse.gdb). Upon review of our search data, we have verified that the correct mouse-specific glycan database (pGlyco-N-Mouse.gdb) was indeed used during the analysis. We carefully re-examined the glycan compositions identified in our dataset and confirmed the presence of Neu5Gc-containing glycans, which are characteristic of mouse glycoproteins but absent in humans.

2. In Fig. 4d, there are two “NeuGc containing” items.

Responses: We thank the reviewer for pointing out the issue with Fig. 4d, where two items were labeled as “NeuGc containing.” Upon review, we identified this as a labeling error in the figure. One of the two “NeuGc containing” labels was intended to represent “NeuAc containing.” We have corrected this mistake in the revised version of Fig. 4d. The revised figure now accurately distinguishes between “NeuGc containing” and “NeuAc containing” glycans.

3. In Fig. 1 second panel, it may be not adequate to say it is "spatial" specificity. Tissue specificity is definitely OK.

Responses: We thank the reviewer for pointing out the potential inadequacy of using the term “spatial specificity” in the second panel of Fig. 1. We agree that “tissue specificity” is a more precise and appropriate description for this context.

Reviewer #2 (Remarks to the Author):

This is a mixed bag of high quality large glycoproteomic datasets generated by a group dedicated to the development and applications of glycoproteomics platform. Notably, they compared the glycoproteomics data derived from 5 different mouse tissues, digested with three tryptic combinations, two enrichment methods, and data analyzed using 4 different software. This is further accompanied by glycoproteomics data on the mouse brains from control vs Parkinson and Alzheimer disease models, at young vs old stages. A website holding these data was set-up and presented.

Overall, it is, however, somewhat unfocused, attempting to highlight a few technical aspects based on the comparisons made on different sample prep and data analysis strategies but did not go far enough to provide better insights. Similarly, it also tried to highlight the differences found in the glycoproteomes from different biological sources, including tissue-specific and disease-related changes, but what biological insights can be inferred is not apparent.

Responses: Thank you for your constructive feedback on our manuscript. We have revised the manuscript to address your concerns and provide clearer insights.

To improve clarity, we have expanded on the comparison of different sample preparation and data analysis strategies. This includes the generation of graded datasets that facilitate AI model training (**Figure 2 in this response letter**), a thorough evaluation of digestion methods, enrichment techniques (**Figure 3 and 4**), and a more detailed comparison of different software tools (**Figure 5, 6, 7 and 8**). Additionally, we have provided further validation of tissue-specific glycoprotein expression using western blot and explored the biological insights derived from these datasets (**Figure 1 and 9**). These revisions aim to offer a more comprehensive understanding of both the technical advancements in glycoproteomics and the biological implications of tissue-specific and disease-related changes.

Regarding the biological insights, we have now explicitly highlighted the tissue-specific changes in our revised manuscript. Our dataset offers a rich resource for understanding glycoproteins across different tissues. For example, the expression patterns of CD36 and L1CAM demonstrate the capability of our dataset to reflect tissue-specific expression profiles (**Figure 1**). These proteins show distinct expression patterns in different tissues, supporting the idea of tissue-specific glycoproteome changes. The data visualization on the website now includes 4,563 glycoproteins with tissue-specific expression profiles.

Furthermore, we have highlighted clear disease- and age-specific glycosylation changes across different brain regions in the manuscript (Figure 5 in manuscript). For example, we observed that aging mice showed a significant decline in glycosylation levels across many proteins compared to younger mice. This finding provides insight into the role of glycosylation in aging and its potential implications for neurodegenerative diseases. Additionally, the co-regulation network analysis of N-glycosites and N-glycans across brain regions and disease conditions identified distinct modules that were associated with specific biological processes. For example, module M5, which was significantly downregulated in the substantia nigra of young subjects but upregulated in PD, was enriched in biological processes such as axonogenesis and glutamatergic synapse function. Since the substantia nigra is the most important pathological region in PD, our findings suggest that brain region-specific glycosylation changes may play a role in the neurodegenerative processes in PD. These findings provide new perspectives into glycosylation's role in aging and diseases and suggest potential biomarkers and therapeutic targets for future study.

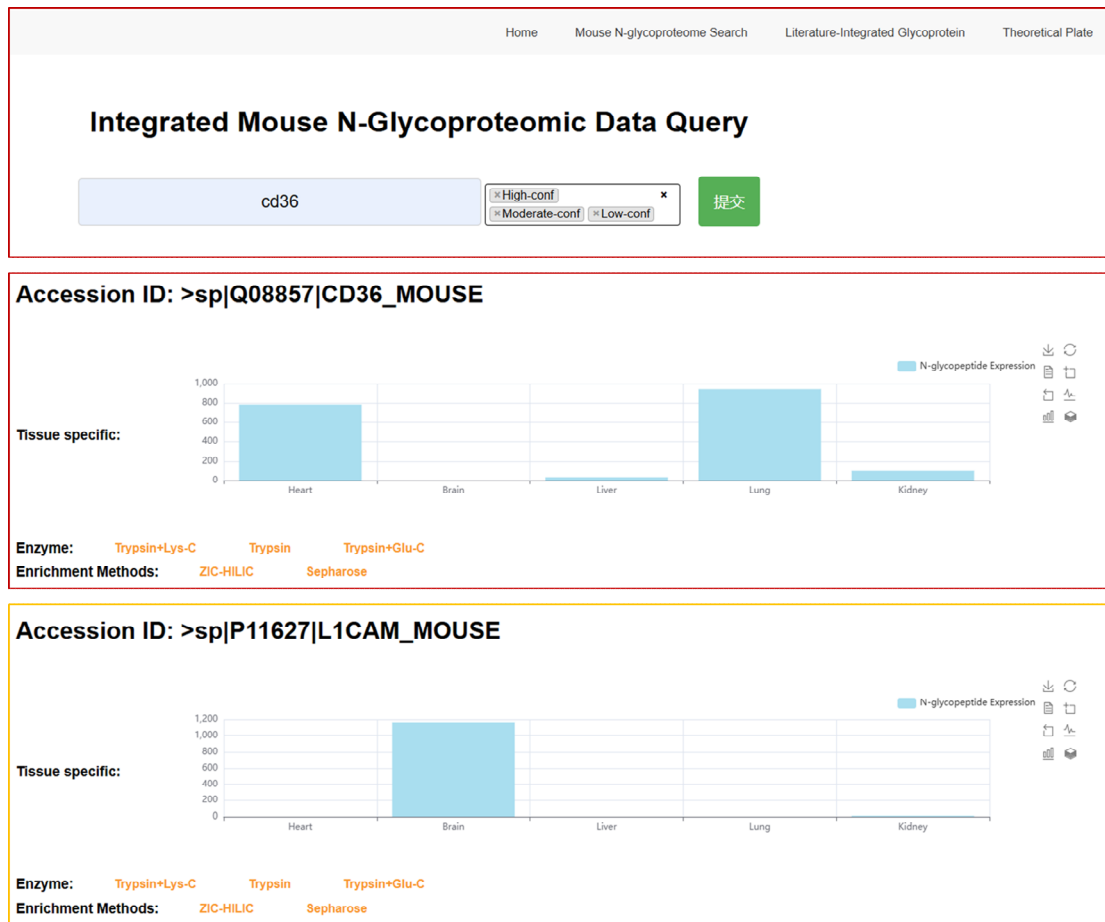


Figure 1. Integrated mouse N-Glycoproteomic data query showing tissue-specific expression of glycoproteins.

This work is rich in data and different variations of bioinformatics analysis. It may indeed serve its purpose by generating a large dataset that can facilitate current endeavors in machine learning and developing AI-driven software.

Responses: Thank you for your comments. We greatly appreciate your recognition of the potential of our dataset to facilitate in AI models. We have addressed this concern in detail in response to review1 and would like to provide a summary of our findings below.

In this study, we constructed a large-scale glycoproteomic spectra library, not only expanding existing databases but also classifying spectra based on confidence levels assessed by multiple search engines. Our evaluation demonstrated that training AI models with moderate- and high-confidence data significantly improved performance (**Figure 2**). Specifically, the DeepGP model trained with these datasets achieved a median cosine similarity of 0.95, surpassing previous models (**Figure 2a and b**), while the DeepGlyco model consistently exhibited superior predictive accuracy ($DP > 0.986$, $SA < 0.083$) (**Figure 2c and d**)^{1, 2}. Prediction of glycopeptide fragment mass spectra by deep learning. Additionally, cross-

validation across different enzymatic digestion patterns confirmed the model's generalization ability (**Figure 2e**). Our study highlights the importance of both dataset size and quality in improving AI-driven glycopeptide tandem spectra and retention time prediction, providing a valuable resource for AI applications in glycoproteomics.

We have incorporated these results into the manuscript (lines 174-201), updated Figures 2e, f, g, and Figure S10 accordingly, and added Supplementary note 3 and Supplementary data 3.

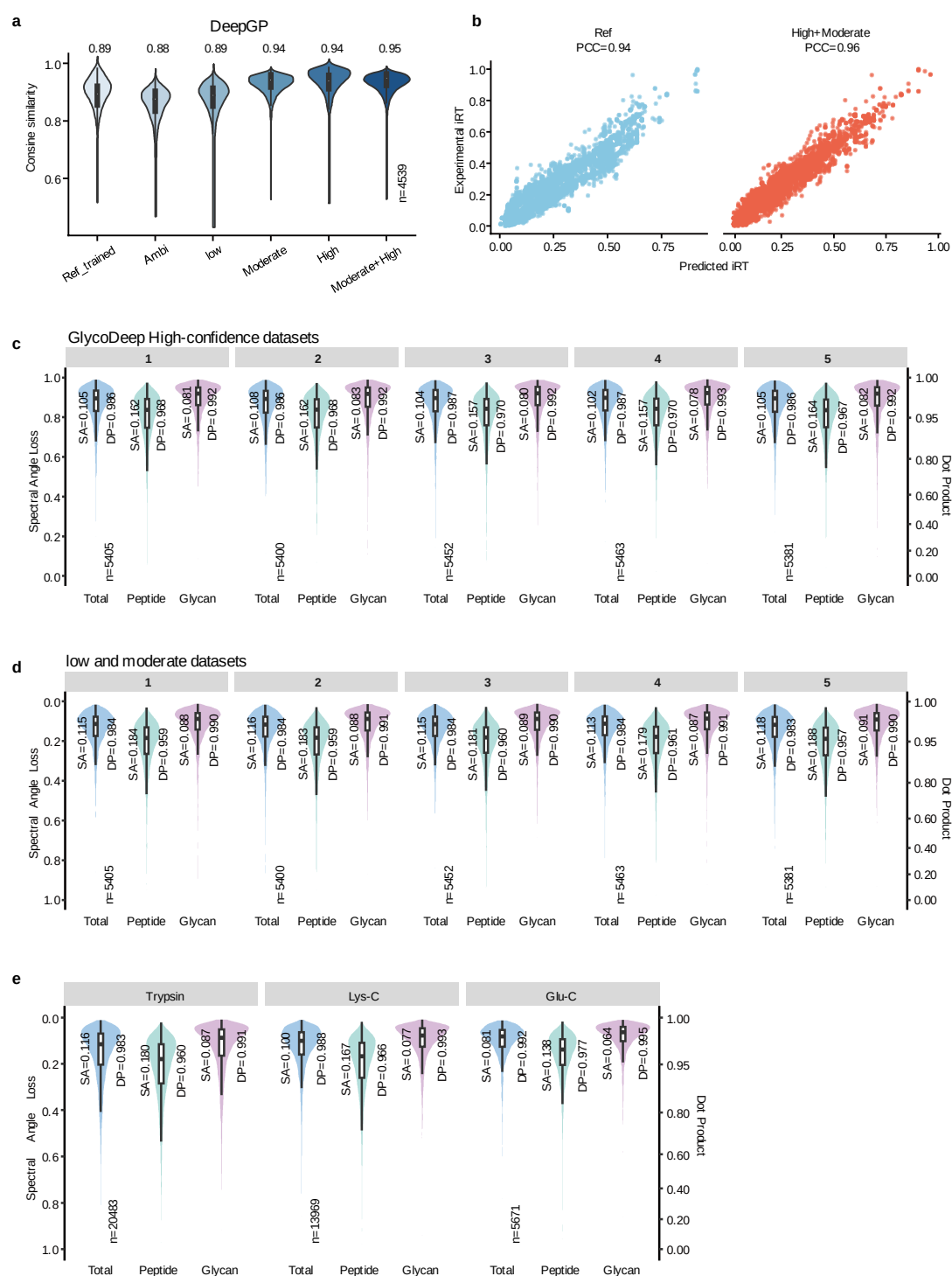


Figure 2. Enhancement of AI training performance with datasets of varying quality (a) Distribution of cosine similarities between the predicted and experimental spectra for different training datasets using DeepGP model. The 1/5 high-confidence subset (4539 spectra) were utilized for testing. The median cosine similarities for each group are indicated above. (b) Correlation between the predicted and experimental iRT is presented with Pearson correlation coefficients (PCC) shown above. Predictions were made using model provided in the literature

and the model trained on high and moderate-confidence data, without RT calibration. (c) Distribution of spectral similarities for high-confidence dataset in 5-fold cross validation using DeepGlyco model. (d) Distribution of spectral similarities using DeepGlyco trained on the low and moderate datasets. (e) Distribution of spectral similarities in cross-validation, with the high-conf dataset partitioned according to different enzymatic patterns.

The N-GlycoMiner website constructed from this work and literature is a useful ref source worth publicizing.

Thank you for your positive feedback on the N-GlycoMiner website. In response to Reviewer 1's suggestion, we have prepared a step-by-step PowerPoint guide to help users navigate the website (Supplementary Data 8). We have integrated protein structures predicted by AlphaFold2 for visualization, providing users with a more interactive and comprehensive view of the proteins. The visualizations of the glycoprotein sites and glycan compositions were also added on the website. These updates aim to improve the website's usability and utility, allowing researchers to better explore and analyze the glycoproteomics data we have made available.

We believe these enhancements will make the N-GlycoMiner website a more useful and accessible resource for the glycoproteomics community, and we look forward to its broader use and impact.

However, from glycobiology perspective, as well as advancement in glycoproteomics itself, there is very little to savor. A few examples of why this Reviewer thinks the authors did not take their work a step further and seemingly content instead to just describe their bioinformatic analysis and diagram generated.

i) three tryptic combinations and 2 enrichment methods. These seem just to increase the number of IDs by multiple runs instead of trying to gain insight into their underlying preference or bias.

Responses: Thank you for your insightful comment on the use of three tryptic combinations and two enrichment methods. We analyzed their complementarity and biases in depth to better understand their impact on glycopeptide identification.

1. Enzymatic digestion strategies and their complementarity

(1) Complementarity of trypsin and Trypsin+Lys-C: In our large-scale mouse dataset, identification numbers for each enzyme plateaued after multiple runs, indicating a saturation limit (**Figure 3a**). While both enzymes cleave at lysine (K) and arginine (R), their complementarity is evident: trypsin uniquely identified 8503 glycopeptides, while Trypsin+Lys-C identified 5526 (**Figure 3b**). Trypsin alone yielded slightly more glycopeptides

than the trypsin + Lys-C combination but had a higher missed cleavage ratio (**Figure 3c**). Both enzymes showed a similar glycosylation site distribution, with a preference near the N-terminal (**Figure 3d**).

(2) Contribution of Trypsin + Glu-C: This combination provided an additional 6151 unique identifications compared to trypsin alone, particularly enriching peptides ending with glutamic acid (E) or aspartic acid (D) (**Figure 3b**). Although its total identifications were slightly lower, likely due to shorter peptides generated by Glu-C, it contributed valuable complementary coverage.

(3) Peptide molecular weight distribution: Trypsin and trypsin + Lys-C produced nearly identical molecular weight distributions, while trypsin + Glu-C yielded shorter peptides due to Glu-C's cleavage specificity (**Figure 3e, f**), supporting the robustness of our approach.

(4) Necessity of multi-enzyme digestion: In our previous published work¹¹, we systematically compared the performance of multiple proteases for glycosite identification by removing glycans with PNGase F and identifying deglycosylated peptides based on the 0.98 Da mass shift. This study demonstrated that the combination of three enzymes (trypsin, Lys-C, and Glu-C) provided the highest glycosite coverage. Other proteases, such as Chymotrypsin, Glu-C, and Pepsin, yielded significantly fewer identifications. Based on these results, we chose these three enzymes, and the current study further validates their complementarity and reliability for comprehensive and large-scale glycoproteomics analysis.

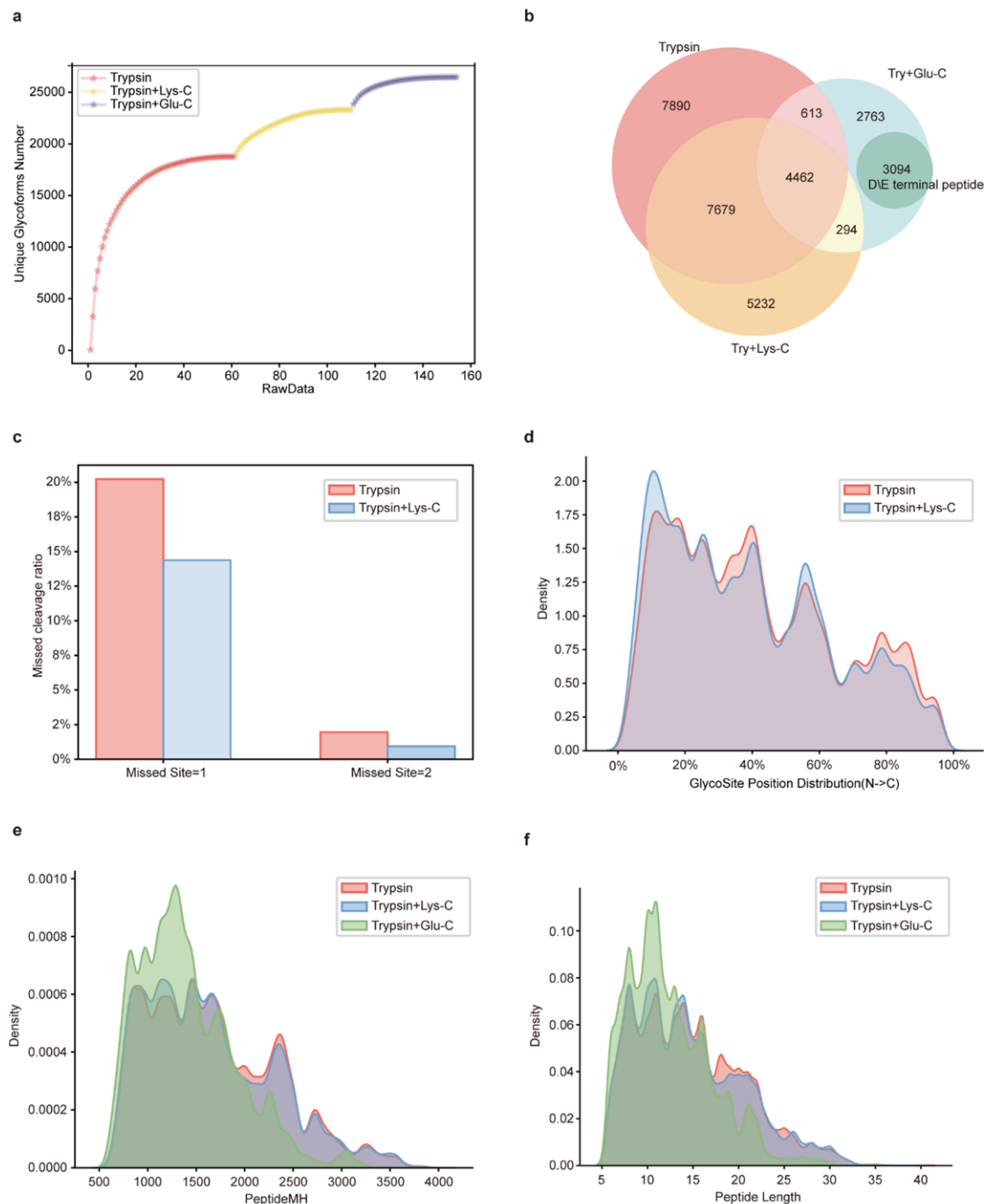


Figure 3: Comparison of glycopeptide analysis using different proteases. (a) Accumulation of unique glycoforms from all raw data for Trypsin, Trypsin+Lys-C, and Trypsin+Glu-C. (b) Venn diagram showing the overlap of glycopeptides identified by Trypsin, Trypsin+Lys-C, and Trypsin+Glu-C. (c) Bar chart showing the ratio of missed cleavage sites for the different proteases. (d) Glycosite position distribution (N→C) for the glycopeptides identified by Trypsin and Trypsin+Lys-C. (e) Peptide mass distribution (PeptideMH) for the glycopeptides identified by each protease. (f) Distribution of peptide lengths for the glycopeptides identified by each protease.

2. Comparative analysis of glycopeptide enrichment methods

(1) Differences in enrichment principles: ZIC-HILIC employs a mixed-mode mechanism for separating and enriching glycopeptides, involving hydrophilic partitioning, electrostatic interactions, dipole interactions, adsorption, and hydrogen bonding. This mixed-mode mechanism provides broader coverage for glycopeptide enrichment, as demonstrated in previous studies¹². Sepharose CL-4B primarily operates based on hydrophilic interaction chromatography (HILIC), where glycopeptides are enriched through hydrogen bonding with polar groups in glycopeptides. These distinct mechanisms contribute to differences in their enrichment performance.

(2) Glycopeptide identification: In our ultra-deep dataset, both enrichment methods achieved saturation in glycopeptide identification (**Figure 4a**), enabling a comprehensive comparison of the differences between the two methods. This saturation demonstrates that while both methods are effective, they provide complementary benefits for glycopeptide enrichment. ZIC-HILIC yielded higher glycopeptide identification overall but showed significant complementarity with Sepharose CL-4B (**Figure 4b**). This complementarity underscores the importance of combining multiple enrichment methods to capture the full range of glycopeptides present in the sample. Using only one method cannot fully enrich all glycopeptides, highlighting the necessity of complementary approaches.

(3) Enrichment efficiency: The enrichment efficiency of ZIC-HILIC was consistently high (greater than 70%), as shown in **Figure 4c**, while Sepharose CL-4B exhibited more variable efficiency, ranging from 30% to 80%. This variability suggests that while ZIC-HILIC is more reliable for large-scale analyses, Sepharose CL-4B may still capture certain glycopeptides that are not efficiently enriched by ZIC-HILIC.

(4) Glycan identification: ZIC-HILIC covered a wider range of glycan compositions, whereas Sepharose CL-4B contributed only a small number of unique identifications. Further analysis of glycan types revealed that Sepharose CL-4B showed a slightly higher proportion of high-mannose and paucimannose glycans (**Figure 4e**). For sialylated glycans, the two methods were comparable, but ZIC-HILIC enriched a higher proportion of fucosylated glycans than Sepharose CL-4B.

Our study demonstrates that both ZIC-HILIC and Sepharose CL-4B have distinct advantages and complement each other in glycopeptide enrichment. The combined use of these methods enables more comprehensive glycopeptide coverage, enhancing the depth and reliability of glycosylation analysis.

In conclusion, we have conducted a detailed analysis of the three enzymes and two enrichment methods, uncovering their individual and complementary contributions to large-scale glycosylation studies. These findings add significant depth and value to our study. We appreciate the reviewer's comments, which have allowed us to refine and expand our discussion.

We have mentioned these discussions in the manuscript (lines 111-116), and added Supplementary note 1, Figure S3 and Figure S4.

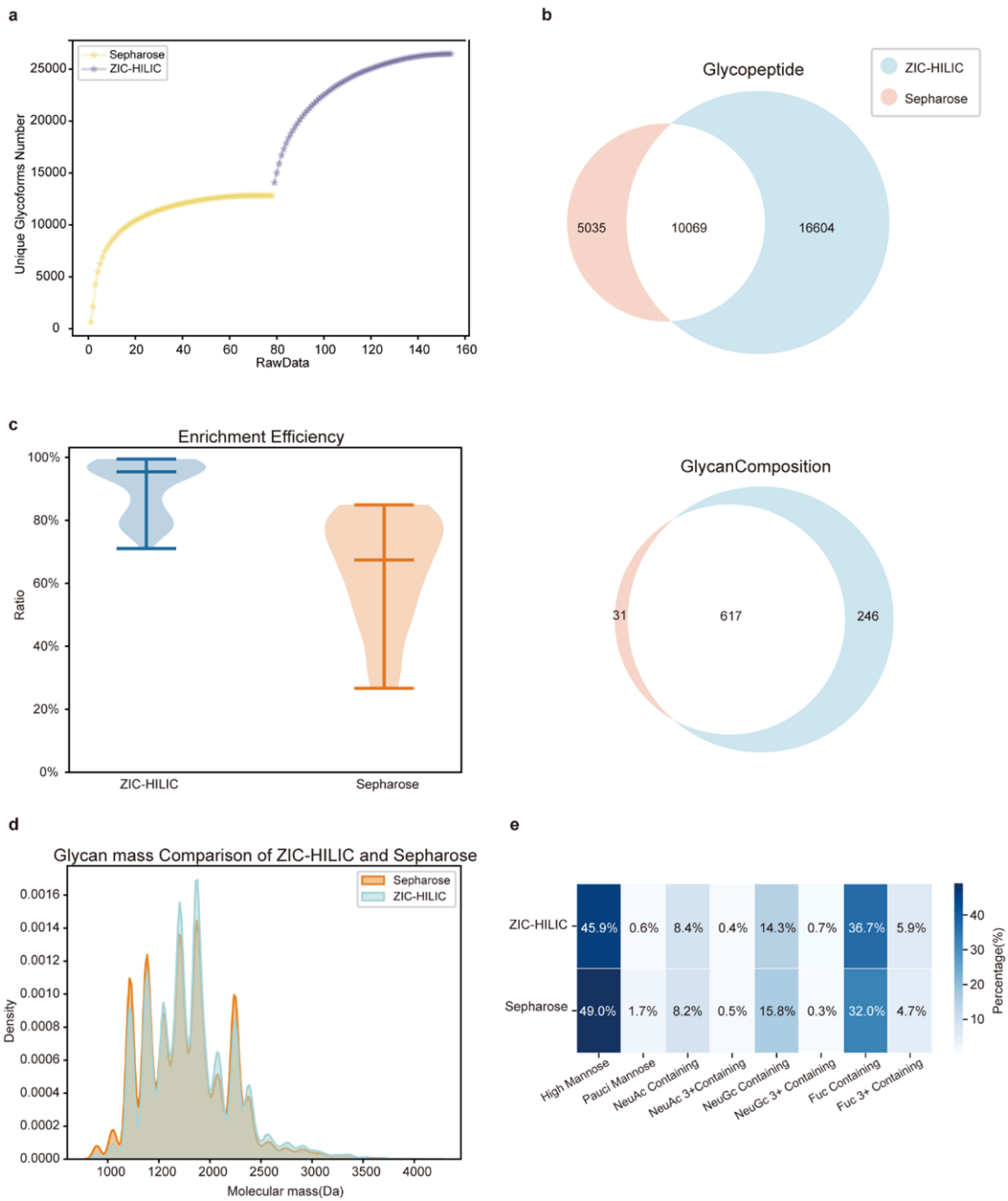


Figure 4. Comparison of glycopeptide and glycan compositions between Sepharose and ZIC-HILIC enrichment methods. (a) Accumulation of unique glycoforms as a function of raw data. (b) Venn diagrams showing the overlap of glycopeptides and glycan compositions identified by ZIC-HILIC and Sepharose. (c) Enrichment efficiency comparison between ZIC-HILIC and Sepharose. (d) Glycan mass comparison of ZIC-HILIC and Sepharose, displaying the molecular mass distribution of glycans. (e) Heatmap showing the percentage distribution of different glycan types.

ii) 4 different software for glycopeptide ID - just provide the number of IDs and reinforce the ideas that MS-Fragger and GlycoDecipher would give much higher number of IDs relative to pGlyco and StrucGP, but did not provide an analysis of why. Confidence is assumed when all 4 identified the same glycopeptide, whereas when each identified a different glycopeptide from the same gPSM, it is considered ambiguous. Moderate level of confidence when the same glycopeptide was identified by 2 or 3 of the software but not consistently so across the board. This is all good and provides a basis for grading the confidence level in generating the database but it does not provide a further analysis of the bias or blindspots of each software for others to learn. It is unclear why the authors then decided to use pGlyco3 - is it a rationale decision made on the basis that it seems to be the most reliable despite its lower number of ID?

Responses: Thank you for your valuable comment. We agree that understanding the biases inherent in glycoproteomics software tools is essential for accurate data interpretation. Below, we tried to discuss some of the potential biases and their impact on glycoproteomic analyses:

(1) Bias in identification across different levels. At the GPSM, precursor, and glycoform levels (Figure 5), the rankings of identification numbers are almost consistent across all tools, with Glyco-decipher achieving the highest, followed by MSFragger-Glyco, pGlyco3, and StrucGP, which has the lowest. However, at other identification levels, including glycosites, glycoproteins and glycan compositions, the results vary significantly. pGlyco3 exhibits the lowest identification numbers for glycosites and glycoproteins but performs moderately well in other categories, notably surpassing MSFragger-Glyco in glycan identification. StrucGP, on the other hand, performs comparably to MSFragger-Glyco and Glyco-decipher at the glycosite and glycoprotein levels. This analysis underscores the trade-offs between the tools in their ability to identify glycoproteins, glycosites, and glycans, highlighting the influence of their design and analytical focus on identification performance.

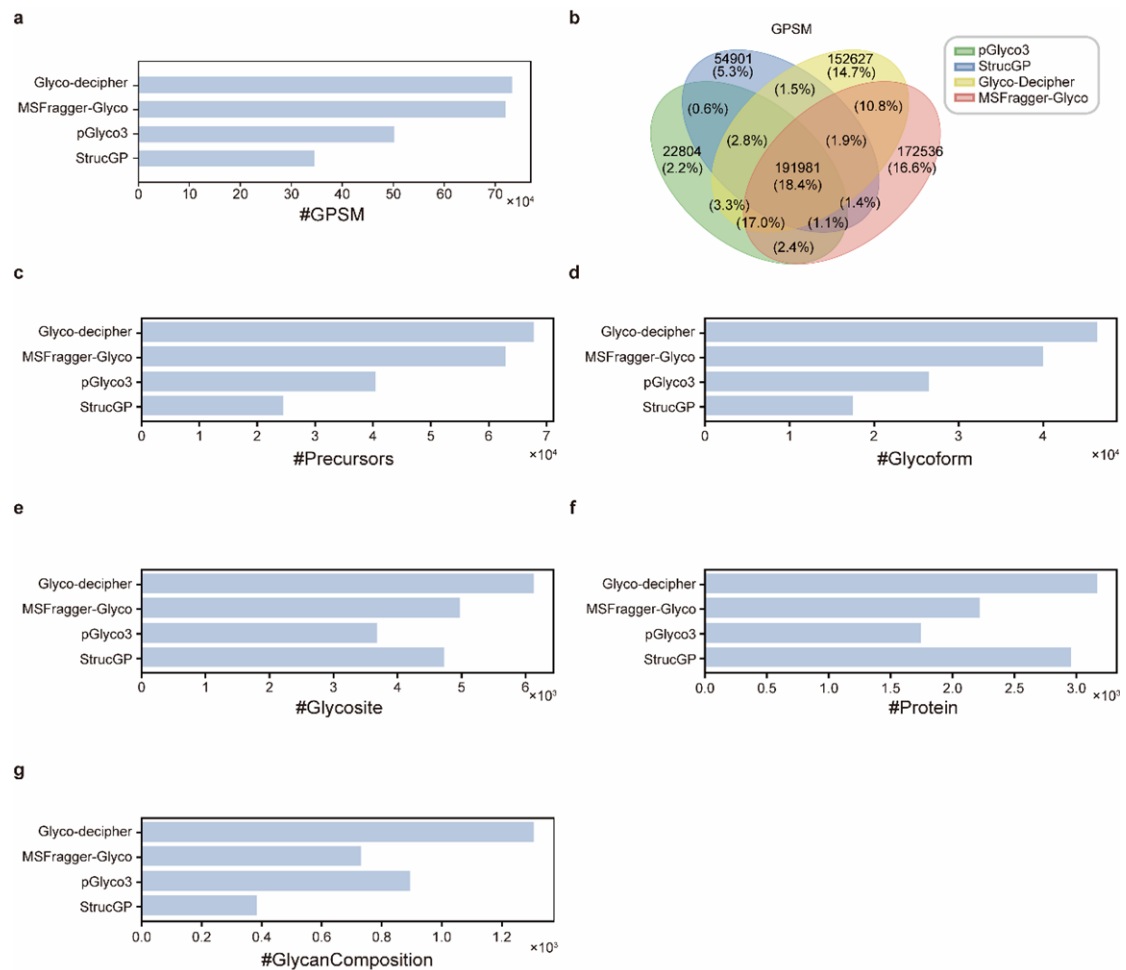


Figure 5. Comparison of identification numbers from different glycoproteomics software platforms. (a) Bar plot showing the total number of glycopeptide-spectrum matches (GPSM) identified by each of the four software tools. (b) Venn diagram showing the overlap in the number of GPSMs identified by the four software tools, with percentages indicating the shared and unique identifications. (c, d, e, f, g) Bar plot showing the number of precursors, glycoforms, glycosites, proteins and glycan compositions identified by each software.

(2) Identification consistency analysis. For the spectra identified by all four tools (191,981 in Figure 5b above), 160,928 spectra were consistently identified as the same precursors by all tools, while 31,053 spectra showed inconsistencies. MSFragger-Glyco contributed to the highest proportion of inconsistent identifications (35.6%), followed by Glyco-decipher (18.1%) and StrucGP (12.8%). In contrast, pGlyco3 had the lowest inconsistency rate (4.2%) (Figure 6a). For the spectra identified by any three of the four tools, MSFragger-Glyco still exhibited the largest inconsistency (13.3%), followed by Glyco-decipher (3.7%) and pGlyco3 (1.7%) (Figure 6b). These results suggest that pGlyco3 is the most reliable tool considering consistent spectrum identification. On the other hand, MSFragger-Glyco identifies a larger number of spectra with greater inconsistency. The high inconsistency rate of MSFragger-Glyco indicates

that while it has higher sensitivity, it may also lead to more false positives. These differences highlight the varying strengths and biases of these tools, with pGlyco3 excelling in reliability and MSFragger-Glyco providing a broader range of identifications at the expense of consistency.

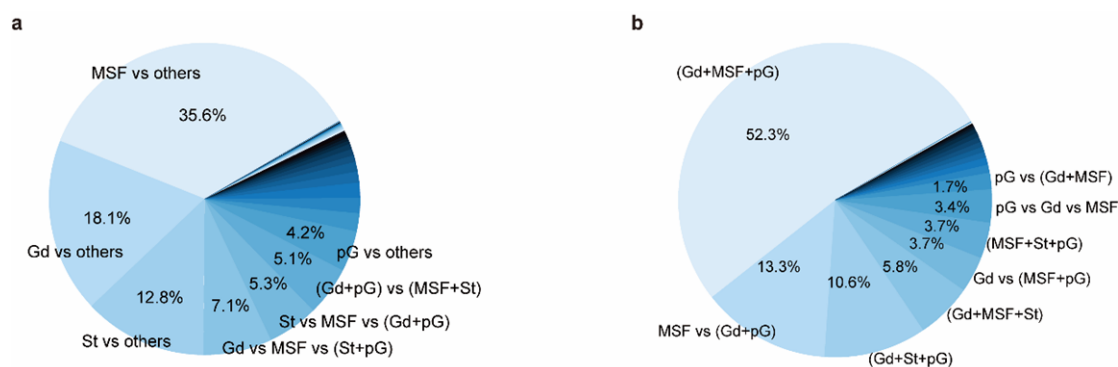


Figure 6. Analysis of inconsistent precursor identifications across four software tools. (a) The pie chart illustrates the distribution of spectra with inconsistent precursor identifications among four glycoproteomics software tools: MSFragger-Glyco (MSF), pGlyco3 (pG), Glyco-decipher (Gd), and StrucGP (St). It highlights spectra where one tool reports a unique identification or where groups of tools produce distinct results, excluding the cases where all four tools agree. For example, "MSF vs others" indicates spectra where MSFragger-Glyco reported a unique result compared to the other tools, which agreed on a different precursor. (b) The pie chart shows the precursor identification analysis for spectra identified by any three of the four tools. Each section represents the percentage of spectra with matching or differing precursor identifications. For example, "Gd+MSF+pG" represents spectra where Glyco-decipher, MSFragger-Glyco, and pGlyco3 all identified the same precursor, while "MSF vs (Gd + pG)" shows spectra where MSFragger-Glyco identified a different precursor compared to Glyco-decipher and pGlyco3, which agreed on the same one.

(3) Bias in peptide and glycan features. We specifically analyzed the distribution of peptide length (number of amino acids), peptide molecular weight, glycan length (number of monosaccharides) and glycan types across the four software tools (Figure 7). StructGP consistently identifies shorter peptides and glycans, whereas MSFragger-Glyco shows a preference for longer peptides, and Glyco-decipher tends to identify longer glycans. Notably, pGlyco3's glycan-length bias was influenced by the exclusion of glycan structures above 20 monosaccharides, as we used the pGlyco-N-Mouse database instead of pGlyco-N-Mouse-Large. In terms of glycan types, StructGP is biased toward high-mannose and pauci-mannose glycans, with the lowest identification of sialylated glycans, indicating its limited ability to detect

complex glycan structures. On the other hand, MSFragger-Glyco exhibits the highest sialic acid content in its identifications. Meanwhile, both pGlyco3 and Glyco-decipher demonstrate a stronger focus on fucosylated glycan identifications. Collectively, these observations highlight distinct differences in the analytical focus and limitations of the tools, emphasizing the need to carefully select software based on the specific requirements of a glycoproteomics study.

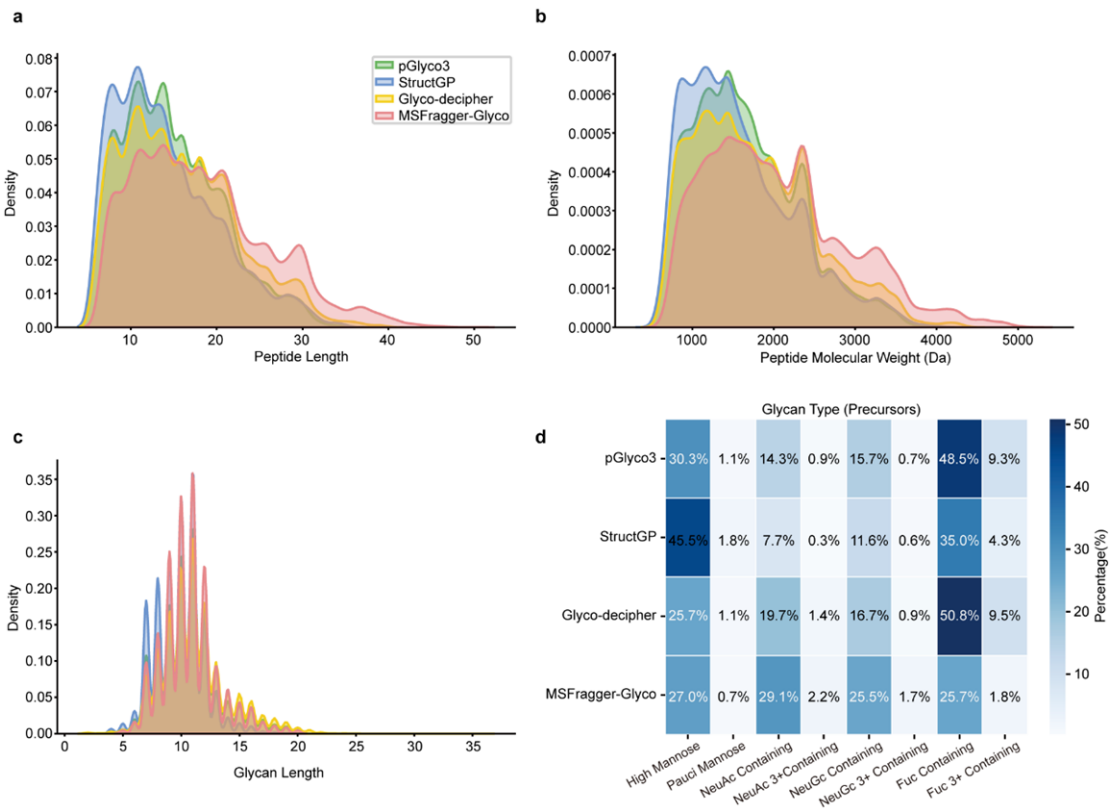


Figure 7. Comparison of glycoproteomics analysis results across different software tools. (a) Density distribution of peptide lengths identified by pGlyco3, StructGP, Glyco-decipher, and MSFragger-Glyco. The x-axis represents peptide length, while the y-axis indicates density. (b) Density distribution of peptide molecular weights (Da) identified by the four tools. The x-axis represents peptide molecular weight, and the y-axis indicates density. (c) Density distribution of glycan lengths identified by the four tools. The x-axis represents glycan length, while the y-axis indicates density. (d) Proportion of different glycan types (precursors) identified by the four tools. The glycan types are categorized into high mannose, pauci mannose-containing, NeuAc-2+ containing, NeuAc-3+ containing, NeuGc-2+ containing, NeuGc-3+ containing, Fuc-2+ containing, and Fuc-3+ containing. The intensity of the blue color represents the percentage of each glycan type identified by each tool.

(4) Quantification Bias: Heart vs. Brain Samples. For quantification evaluation, we compared pGlyco3, Glyco-decipher, and MSFragger-Glyco, as these are the only tools supporting quantitative analysis. We found that the differentially expressed glycoproteins and

glycosites showed a high overlap in up-regulated and down-regulated molecules, indicating a high consistency in quantification between the three software tools at these levels (**Figure 8a**). However, at the glycan and site-specific glycoform levels, the overlap in quantification results was much lower compared to glycoproteins and glycosites, indicating lower consistency in quantification at these levels.

We further conducted a correlation analysis of the quantification values among the three software tools using their commonly identified and quantified glycopeptides (**Figure 8b**). The highly positive correlations in the quantification of glycoproteins and glycosites (median PCC above 0.78) were revealed, suggesting that the three software platforms agree well in quantifying these levels of data. However, the correlation analysis for glycans showed more variability, with lower median PCC (median PCC ranges 0.52-0.67). This suggests that the software tools are generally consistent in quantifying glycoproteins and glycosites, but less agreement at the glycans levels. This inconsistency could be due to differences in how each tool processes and interprets glycan-related data. This lower consistency in glycans' and glycoforms' quantification may lead to divergent interpretations in subsequent studies of glycosylation's biological functions.

Overall, these findings emphasize the need for caution when comparing glycan and glycoform data across different software platforms, particularly in studies focused on the functional implications of glycosylation. Further improvements in glycan quantification algorithms and validation with non-omics techniques could help address these discrepancies.

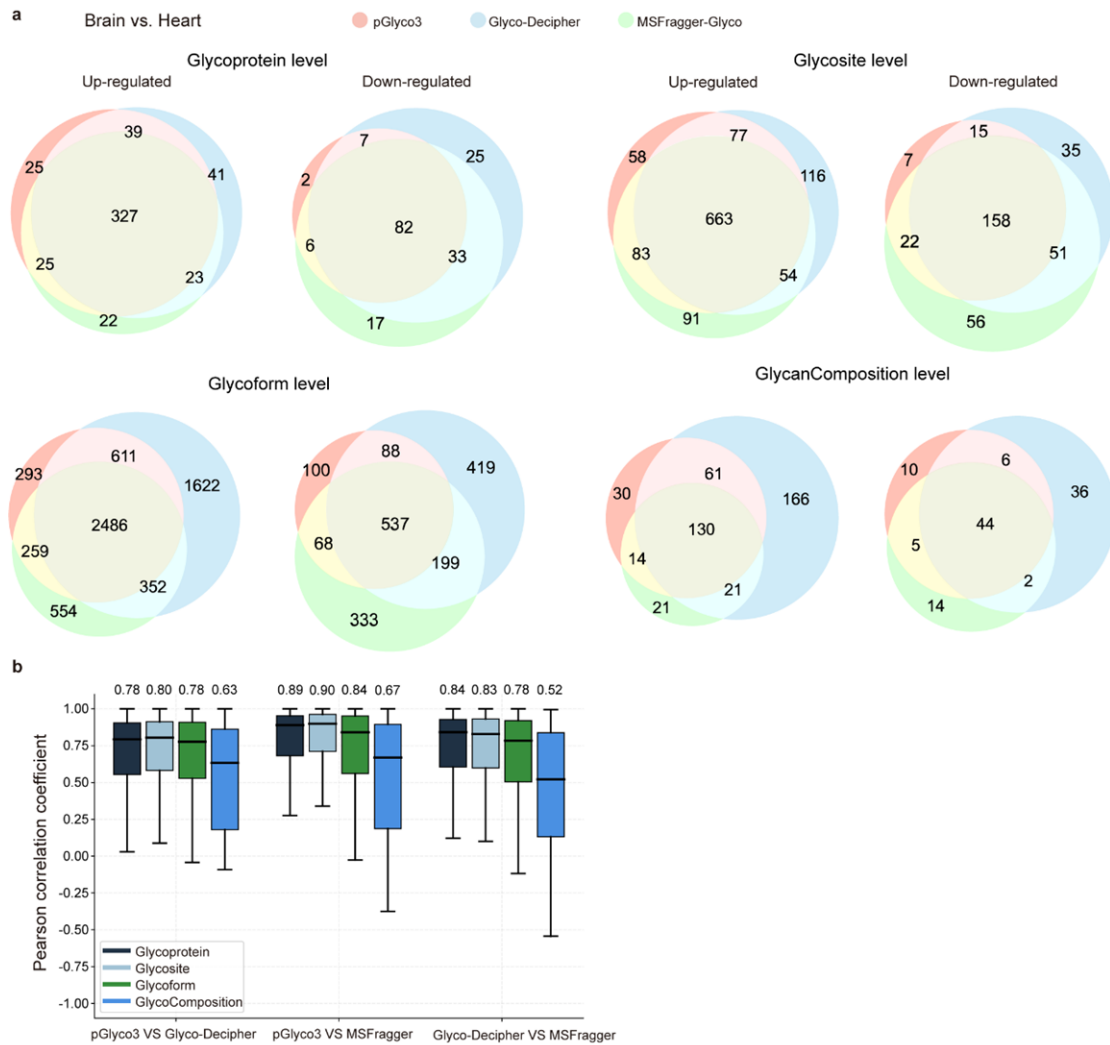


Figure 8. Comparison of Quantitative Glycoproteomics Data Between Brain and Heart Samples Using Different Software Platforms. (a) Venn diagrams showing the overlap of up-regulated and down-regulated glycoproteins, glycoforms, glycosites, and glycan compositions between brain and heart samples, as quantified by three software tools: pGlyco3 (red), Glyco-Decipher (blue), and MSFragger-Glyco (green). (b) Boxplots of Pearson correlation coefficients comparing the glycoprotein, glycosite, glycoform, and glycan composition data commonly identified and quantified by the three software platforms. The comparisons include pGlyco3 vs. Glyco-Decipher, pGlyco3 vs. MSFragger, and Glyco-Decipher vs. MSFragger. Higher correlation values indicate better agreement in the quantification values between the platforms.

(5) Comparison of the four software tools from their identification strategies, glycan and peptide scoring and false positive control. Each software tool utilizes a distinct search strategy for glycoproteomic identification, which impacts their performance in glycan and peptide identification and false positive control.

pGlyco3 adopts a glycan-first search strategy, which prioritizes the identification of N-glycopeptides by first identifying glycans using ion-indexing techniques, followed by peptide identification^{3, 4}. This approach is integrated with scoring systems for glycans, peptides, and glycopeptides. A notable feature of pGlyco3 is its rigorous false discovery rate (FDR) control, which applies across all levels—glycans, peptides, and glycopeptides. pGlyco3 used glycan database combining N-glycans from Glycome-DB, as well as the three largest substructures of high-mannose, hybrid, and complex N-glycans. This glycan-first approach and FDR control allow pGlyco3 to outperform the other tools in terms of accuracy though the strict glycan-level quality control may reduce sensitivity (Figure 5 and 6).

StrucGP uses a modular approach, dividing N-glycans into three modules, with each identified based on distinct Y ion patterns or combinations of B/Y ions⁵. This methodology facilitates detailed glycan structure determination. FDR estimations for both the peptide and glycan portions of glycopeptides are performed using both decoy database and decoy spectrum approaches. While StrucGP performs fewer glycan identifications compared to other tools, it offers similar performance to MSFragger-Glyco and Glyco-Decipher in terms of glycosite and glycoprotein identification (Figure 5). This is likely because StrucGP requires more ion information for detailed glycan structure analysis, which makes it more prone to identifying shorter peptides and glycans, especially high-mannose glycan types (Figure 7).

MSFragger utilizes an open mass offset search strategy, where glycopeptide spectra are analyzed based on possible glycan compositions derived from mass offsets¹³. It performs deisotoping and decharging of spectra prior to analysis, improving the sensitivity of the tool, particularly for high-mass glycopeptides (Figure 7). However, MSFragger does not place a strong emphasis on glycan-level quality control, which may lead to less precise glycan identification compared to other tools³. The high sensitivity of MSFragger often results in a larger number of identifications, but it has the lowest consistency in identifications compared to the other software (Figure 5 and 6).

Glyco-Decipher employs a glycan database-independent peptide matching strategy, which uses shared fragmentation patterns of peptide backbones in glycopeptides to improve spectrum interpretation⁷. This technique enables Glyco-Decipher to extend its glycopeptide identification capabilities through the analysis of peptide backbone fragments. Glyco-Decipher has an advantage in identification numbers due to its spectrum expansion method.

Our analysis highlights the trade-offs in glycoproteomics software tools regarding identification sensitivity, consistency, and quantification accuracy. pGlyco3 is the most consistent but less sensitive, MSFragger-Glyco is highly sensitive but prone to false positives, Glyco-decipher achieves the highest identification numbers through spectrum expansion, and

StrucGP excels in structural characterization. These findings underscore the need for careful software selection depending on the specific objectives of a glycoproteomics study, particularly in glycan-focused research where quantification consistency remains a challenge.

In summary, the differences in glycan identification and quality control across software tools have a substantial impact on their overall performance, particularly in terms of identification consistency and accuracy. We chose pGlyco3 for the analysis due to its highest consistent identification rate and comprehensive FDR control.

For this revision, we have added these analysis into the manuscript (lines 123-152, 160-164), updated Figure S6-9 accordingly, and added Supplementary note 2.

iii) It is interesting to show that only a single glycosite and 2 glycoforms were ID for CD36 from brain but that from heart has 7 glycosites and over 200 glycoforms ID. However, it is unclear the biological or technical reasons contributing to this difference. Only one glycosite identified does not mean the other sites are necessarily not occupied. The complexity of the glycopeptides at each sampling elution timepoint may differ significantly between the two sources of samples, which may also have very different protein expression level. What exactly were the respective glycoforms identified was not provided. The authors simply made a statement that “The distinct differences of glycosylation on CD36 implied the importance of glycosylation for the tissue specific functions of CD36”.

Responses: Thank you for your comments. Based on the complementary techniques and data analysis methods employed in this study, we have identified as many N-glycosylation sites as possible across multiple mouse organs, and we believe this data accurately reflects the glycosylation status in vivo.

To further validate this, we performed additional western blot analysis of CD36 under native (**Figure 9a**) and denatured conditions (**Figure 9b**) with or without PNGase F treatment. Under native conditions, PNGase F treatment did not work to remove the N-glycans from CD36, indicating that the protein's glycosylation state was intact. Upon applying PNGase F treatment under denatured conditions, the N-glycans were completely removed, confirming that CD36 mainly undergoes N-glycosylation. In both conditions, we observed that the protein expression of glycosylated CD36 in the heart was much higher than in the brain. Under native conditions, heart and lung tissues exhibited very high levels of N-glycosylation of CD36, while the brain and liver showed much lower levels of glycosylation.

These results from the western blot analysis were consistent with the identification of glycoforms for these tissues, as observed in both pGlyco3 and the combined software analysis

(Figure 9c). The differences in glycosylated CD36 between heart and brain mainly stemmed from the protein expression levels in these two tissue types, which were higher in the heart and lower in the brain. This additional data reinforces our conclusion that the observed differences in glycosylation are largely attributed to tissue-specific expression patterns of CD36, rather than technical limitations in glycoform identification. We have incorporated these results into the manuscript (lines 268-283, Figure 3g, h), added Figure S15 and Supplementary Data 5.

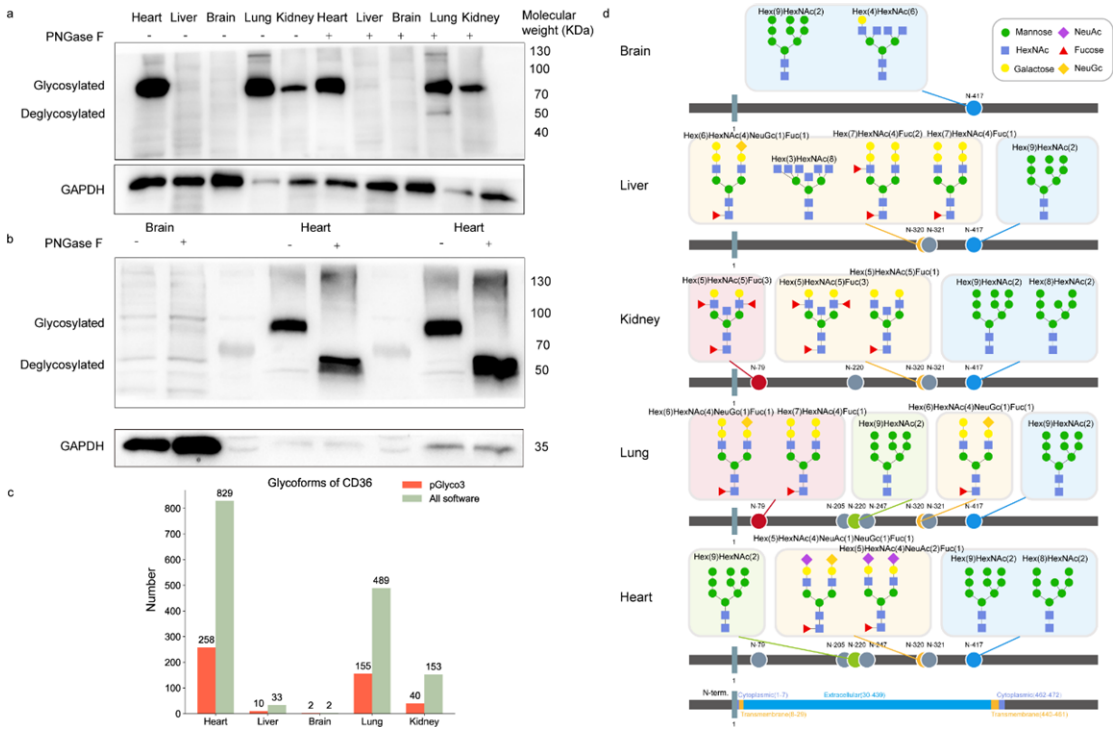


Figure 9. Western blot analysis and glycoform identification of CD36 glycoprotein across different tissues. (a) Western blot analysis of CD36 glycoprotein under native conditions, demonstrating the presence of intact glycoprotein in various tissues. Under native conditions, PNGase F treatment did not remove the N-glycans from the CD36. (b) Western blot analysis of CD36 glycoprotein under denatured conditions, where PNGase F treatment was applied to assess the effect of glycan removal on protein migration. The removal of N-linked glycans by PNGase F resulted in a shift in the molecular weight of the CD36 protein, highlighting the impact of glycosylation on the protein's electrophoretic behavior. (c) Comparison of glycoforms numbers of CD36 identified in different tissues. (d) The top five CD36 glycoforms, ranked by their intensities in pGlyco3, are shown across various tissues. The glycan composition and possible structures are displayed, illustrating the diversity of glycoforms in different tissue types.

Other minor issues:

Fig 2C. Unclear what is co-spectra identification ratios for different levels of protein structure - please clarify. What's the significance?

Responses: Thanks for pointing out this. We also found that the "co-spectra identification ratios" in Figure 2C was not very easy for reader to understand. The Co-spectra" was removed. In this figure 2C, the common identified GPSMs by the four software can result in consistent (Identical ID for 4 software) and inconsistent identifications (ID as the same by 2 or 3 software, ID mutually distinct) in precursors and other identifications. The ratios of consistent and inconsistent identifications were showed. This explanation was added to the figure legend.

This classification helps illustrate the reliability and consistency of the identification results across different software tools. The greater the overlap (light and dark blue areas), the more robust and consistent the identification process, whereas a higher proportion of ambiguous (red) results suggests the need for further refinement or validation of the identification methods.

Fig 4, what's the difference between panel c and e? This is not clear from the description provided in the Figure legend although one can infer that from the manuscript.

Responses: Thank you for pointing this out. Panel c represents the hierarchical clustering of glycan compositions and Panel e, in contrast, depicts the hierarchical clustering of glycoproteins across the same organs. We corrected this error in the figure legend to accurately describe the distinction between panels c and e in the revised manuscript.

In panel d, one of the two NeuGc containing label should be NeuAc containing.

Responses: We appreciate your attention to detail. This was an oversight during figure preparation, and we corrected this error in the revised version to ensure accuracy.

Percentage of high mannose seems to be exceptionally low. 0% for brain?

Responses: Figure 3D presents glycan composition clusters derived from unsupervised hierarchical clustering in Figure 3C, demonstrating tissue-specific partitioning of glycosylation patterns. Notably, quantitative analysis indicates that high-mannose glycans were not detected (0%) in brain-associated Cluster 1, suggesting their relative lower expression in C1(brain) rather than an absence.

The Suppl Tables, esp Suppl Table 2, should be better described.

Responses: Thank you for your valuable feedback. We acknowledge that the description of the supplementary tables, particularly Supplementary Table 2, is insufficiently detailed in the current manuscript. We added sheet descriptions for each supplementary table.

Reviewer #3 (Remarks to the Author):

The manuscript from Shen and co-workers provides a very large collection of N-linked glycoproteomics, as well as a searchable database. The authors use multiple methods and both the biochemical and data analysis steps to arrive at a very robust set of N-linked glycans and glycoproteins. They do a very thorough job characterizing these data and apply their approach to find some differences in aging mice, as well as AD and PD models. I believe that this dataset will be a useful one for the community, as well as for training better data analysis programs. I support its publication.

Responses to Reviewer #3:

We sincerely appreciate your positive evaluation and support for the publication of our manuscript. We are delighted that you recognize the robustness of our N-linked glycoproteomics dataset and the thoroughness of our biochemical and data analysis approaches.

We appreciate your recognition of the dataset's potential for training better data analysis programs. We agree that the dataset can serve as a benchmark for developing and refining algorithms in glycoproteomics research. To underline this point, we have added reanalysis and discussion in the manuscript about potential applications and the value of this dataset in advancing computational tools (see revised section). Thank you for supporting the publication of our work.

References

1. Zong, Y., Wang, Y., Qiu, X., Huang, X. & Qiao, L. Deep learning prediction of glycopeptide tandem mass spectra powers glycoproteomics. *Nature Machine Intelligence* **6**, 950-961 (2024).
2. Yang, Y. & Fang, Q. Prediction of glycopeptide fragment mass spectra by deep learning. *Nature Communications* **15**, 2448 (2024).
3. Zeng, W.-F., Cao, W.-Q., Liu, M.-Q., He, S.-M. & Yang, P.-Y. Precise, fast and comprehensive analysis of intact glycopeptides and modified glycans with pGlyco3. *Nature Methods* **18**, 1515-1523 (2021).
4. Liu, M.Q. et al. pGlyco 2.0 enables precision N-glycoproteomics with comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification. *Nat Commun* **8**, 438 (2017).
5. Shen, J. et al. StrucGP: de novo structural sequencing of site-specific N-glycan on glycoproteins using a modularization strategy. *Nat Methods* **18**, 921-929 (2021).
6. Zeng, W.F., Cao, W.Q., Liu, M.Q., He, S.M. & Yang, P.Y. Precise, fast and comprehensive analysis of intact glycopeptides and modified glycans with pGlyco3. *Nat Methods* **18**, 1515-1523 (2021).
7. Fang, Z. et al. Glyco-Decipher enables glycan database-independent peptide matching and in-depth characterization of site-specific N-glycosylation. *Nat Commun* **13**, 1900 (2022).
8. Muthana, S.M., Campbell, C.T. & Gildersleeve, J.C. Modifications of glycans: biological significance and therapeutic opportunities. *ACS Chem Biol* **7**, 31-43 (2012).
9. Villanueva-Cabello, T.M., Gutiérrez-Valenzuela, L.D., Salinas-Marín, R., López-Guerrero, D.V. & Martínez-Duncker, I. Polysialic Acid in the Immune System. *Front Immunol* **12**, 823637 (2021).
10. Lee, C.G., Oh, M.J., Park, S.Y., An, H.J. & Kim, J.H. Inhibition of poly-LacNAc biosynthesis with release of CMP-Neu5Ac feedback inhibition increases the sialylation of recombinant EPO produced in CHO cells. *Sci Rep* **8**, 7273 (2018).
11. Fang, P. et al. In-depth mapping of the mouse brain N-glycoproteome reveals widespread N-glycosylation of diverse brain proteins. *Oncotarget* **7**, 38796-38809 (2016).
12. Guo, Y. Recent progress in the fundamental understanding of hydrophilic interaction chromatography (HILIC). *Analyst* **140**, 6452-6466 (2015).
13. Polasky, D.A., Yu, F., Teo, G.C. & Nesvizhskii, A.I. Fast and comprehensive N- and O-glycoproteomics analysis with MSFragger-Glyco. *Nat Methods* **17**, 1125-1132 (2020).