

Large-Scale Serum N-glycomics Tracks N-Glycosylation Dynamics in Hepatocellular Carcinoma Progression and Enables Early Diagnosis

Corresponding Author: Professor Ying Zhang

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)

In this study, the authors performed a large-scale investigation of serum N-glycomics in hepatocellular carcinoma (HCC), analyzing 1,074 individuals. They first used MALDI-TOF to analyze serum N-glycome profiles and detected 100–200 N-glycan compositions. Next, they found specific N-glycan patterns associated with poor liver function. They then identified Glycan Co-expression Modules (CCMs) that classify clinical liver conditions. To further understand the serum N-glycome profiles, they integrated N-glycoproteomics data, inferred glyco-enzyme activities, and compared the results to mRNA data from TCGA. Finally, they developed machine learning models for diagnostic purposes.

Overall, this study is comprehensive, technically solid, and well-conducted. It will be of interest to the proteomics, glycomics, and HCC communities. We therefore believe this is a strong candidate for Nature Communications. That said, we present below some concerns and comments—mostly minor—that the authors could address or respond to improve the manuscript's presentation.

Major

1. Even without Figure 6 and Figure 7, we believe the present work is already qualified for Nature Communications. However, the inclusion of Figures 6 and 7—especially Figure 6, which introduces the NglyBAI algorithm—complicates the logical flow (we spent a long time reading and trying to understand this part). We therefore suggest that the authors consider moving the NglyBAI figure to the supplementary materials (or consider publishing it in a separate paper) and downplaying the associated results for the following reasons: a. The N-glycome and N-glycoproteomic datasets are all derived from serum or blood, where N-glycan profiles can be very different from those in cells. Inferring glyco-enzyme activities from such data is theoretically challenging and would require extensive validation, which goes beyond the scope of this study. b. The manuscript already contains 8 main figures. c. It is not clear how increased/decreased mRNA expression in Figure 7 supports the results inferred in Figure 6.

Minor

1. Pseudo-Glycome: Although the results between the pseudo-glycome and the real glycome are highly correlated at the end (Figure 5A), the authors should include a cautionary note in the Results and perhaps in the Discussion as well, because it can be problematic to sum mass spec signals across different proteins and different IGP (Figure 5). Emphasizing the “surprising” consistency is encouraged, but the limitations should also be acknowledged.

2. Line 88: Please briefly mention how quantification was done for MALDI.

3. Throughout the figures, the authors should provide the number of individuals included in their analysis and plots (e.g., Figure 2a, Figure 3a–k, etc.).

4. In all supplementary figure legends, the authors should provide more information and descriptions—e.g., about colors, numbers, statistical tests, and abbreviations (e.g., Supplementary Figures 2 and 9).

5. Currently, there are 34 supplementary figures. The authors are advised to combine some into multi-panel figures to reduce

this number.

6. Figure 3g: The authors should follow standard glycomics nomenclature here (i.e., to be consistent with Figure 3a—e.g., H3N5).
 7. The authors should consider adding a supplementary table with patient-specific information (instead of only the summary table currently shown as Table 1), to allow future analysis of their data.
 8. Line 716: Please specify whether DIA or DDA analyses were used for total proteome and glycoproteomic analysis, respectively.
 9. Line 850: If the authors decide to include NglyBAI in this manuscript, they should provide the code or package on GitHub.
- Overall, we want to reiterate that this is an excellent contribution in our opinion, with the above caveats otherwise noted.

Reviewer #2

(Remarks to the Author)

This manuscript by Lu and colleagues presents a comprehensive study examining serum N-glycosylation dynamics during hepatocellular carcinoma (HCC) progression. Through meticulous multi-cohort analysis and sophisticated multi-omics approaches, the authors have systematically characterized serum N-glycosylation associated with HCC development and assessed their potential as early diagnostic biomarkers. The scale of data collection and analytical rigor demonstrated throughout this work are impressive.

The diagnostic machine learning models developed from N-glycome data represent a significant advancement, notably outperforming alpha-fetoprotein (AFP), the current clinical standard. The authors' efforts to elucidate underlying mechanisms through deconstructing glycomic changes into "carrier protein abundance-driven" versus "glycosylation process-driven" modes demonstrates exceptional scientific depth. Furthermore, the novel NGlyBAI algorithm for enzyme activity inference represents a methodological innovation with broad potential applications.

While the findings are compelling, several points require clarification to strengthen the manuscript's conclusions:

1. The 'Mean Eigenvalue' parameter used to characterize HCC subtypes in Figure 4c lacks methodological clarity. A detailed description of its calculation methodology should be included in the Methods section to ensure reproducibility and facilitate interpretation.
2. Figure 5j presents an apparent contradiction requiring resolution: H5N3S1 glycan modification on A2MG peaks in the CHB group, while total A2MG protein levels are elevated in LC and HCC groups. This observation contradicts a protein-driven mechanism for H5N3S1 elevation in CHB and requires explicit discussion in the Figure 5 narrative.
3. The manuscript would be strengthened by reporting whether the NGlyBAI algorithm identified glycotransferases with significantly different activity scores specifically between LC and HCC groups. Such findings would substantially enhance the algorithm's clinical utility by demonstrating its capacity to detect malignancy-specific enzymatic alterations beyond general liver disease-associated dysregulation.
4. A compelling paradox emerges between glycomic and transcriptomic findings: decreased overall galactosylation in HCC (Fig. 3k) contrasts with the association between high B4GALT2 and B4GALT3 expression and poorer prognosis (Fig. 7d-g). This intriguing discrepancy between gene expression and resulting glycan phenotype warrants thorough mechanistic discussion.
5. For completeness and methodological clarity, comprehensive survival analyses for both B4GALT2 and B4GALT3 across all four endpoints (OS, DSS, PFI, DFI) should be provided, either in supplementary materials or with a clear rationale for the current selective presentation.
6. Model 3's external validation performance (AUC=0.76) for distinguishing HCC from LC, while still impressive, is lower than other models. This reflects the well-documented clinical challenge of differentiating HCC from its cirrhotic background and merits acknowledgment in the Discussion section.
7. The correlation between Model 4 probability score (PP4) and disease severity raises important translational questions that should be addressed: Could this quantitative metric enable longitudinal risk monitoring in cirrhotic patients? Moreover, could a streamlined model using fewer N-glycan features maintain diagnostic efficacy while enhancing clinical implementability?

Reviewer #3

(Remarks to the Author)

This study represents a solid glycan analysis of over 1000 patients, the majority of which have liver disease. The mix of patients is good and also consist of over 400 patients with hepatocellular carcinoma.

The major issue is that very similar work, in both serum and tissue has been done by numerous groups. Nothing in the work presented is novel or alters our understanding of liver disease. For example, the changes observed in branching with the development of HCC have been observed by a many groups, at both the tissue level, in serum and on specific glycoproteins. The alterations in fucosylation with LC and HCC again have been observed by many groups and in the case of LC, linked with specific immunoglobulins and in the case of HCC, with other well known proteins. Again, nothing new is presented here.

In addition, as the results in Figure 8, are not impressive when compared to what is either already clinically available or what is in the literature.

Lastly, a much better job of reviewing the literature should be done to improve the manuscript. The implication that this is the first such study of this kind is disingenuous.

Reviewer #4

(Remarks to the Author)

In this manuscript, Fu et al. presents a large-scale, multi-omics investigation into serum N-glycosylation changes during hepatocellular carcinoma (HCC) development. Using a cohort of over a thousand (Chinese) individuals spanning healthy, chronic hepatitis B, liver cirrhosis and HCC, systematically profiled serum N-glycans via MALDI-TOF MS and LC-MS/MS. The analytical and bioinformatic methodologies used are well established in the field. Using this combination of methods, Fu et al identified 48 N-glycans as significantly dysregulated across disease groups and demonstrated correlations with liver function (e.g. the ALBI score), reflecting hepatic synthetic capacity. In general, the manuscript is very well written, and clearly an impressive amount of work has gone into this work and is represented on the pages of the manuscript. The authors also make their data available. However, I do have a few comments on the manuscript in its current form.

While there are statistically significant differences in the distributions of glycan types between the various groups and controls (e.g. Figure 2c, Figure 3k), there are also considerable overlaps between the distributions, which means these global glycan type differences in themselves are not useful as diagnostic markers. Neither are the HCC subtypes well separated by PCA (Figure 4b) based on the expression of glycan types. This does not match the titular claim that “Large-Scale and multi-cohort Serum N-glycomics Tracks N-Glycosylation Dynamics in Hepatocellular Carcinoma Progression and Enables Early Diagnosis”. I am not convinced the authors have shown that their method enables early diagnosis (which by definition is of individual patients, not cohorts).

Lines 367-368: The authors state that “all three models maintained comparable performance in the external validation cohorts”. What is meant by “comparable” is not defined, but it is clear that all models perform worse (or less well) in both external validation cohorts. This should also be stated, even though it of course is expected.

Minor comments :

Figure 1: Typo “recruided”, unusual/obsolete word “dismangles”

Several figure elements, including the patient and fridge (Figure 1), liver icon (Figure 2), and patient and doctor icons (Figure 3) are from BioRender rather than the authors themselves. This must be acknowledged.

The in-house scripts could also be made available.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors addressed all of our previous questions and suggestions. Good job! We are now supportive for its publication.

Reviewer #2

(Remarks to the Author)

The authors have addressed all of the previous comments and incorporate suggestions. Overall it is a nice manuscript, congrats!

Reviewer #4

(Remarks to the Author)

The authors have addressed all of my comments.

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/>

Reviewer #1

In this study, the authors performed a large-scale investigation of serum N-glycomics in hepatocellular carcinoma (HCC), analyzing 1,074 individuals. They first used MALDI-TOF to analyze serum N-glycome profiles and detected 100–200 N-glycan compositions. Next, they found specific N-glycan patterns associated with poor liver function. They then identified Glycan Co-expression Modules (CCMs) that classify clinical liver conditions. To further understand the serum N-glycome profiles, they integrated N-glycoproteomics data, inferred glyco-enzyme activities, and compared the results to mRNA data from TCGA. Finally, they developed machine learning models for diagnostic purposes.

Overall, this study is comprehensive, technically solid, and well-conducted. It will be of interest to the proteomics, glycomics, and HCC communities. We therefore believe this is a strong candidate for Nature Communications. That said, we present below some concerns and comments—mostly minor—that the authors could address or respond to improve the manuscript's presentation.

R: We sincerely thank the reviewer for the overall positive feedback. We greatly appreciate all the suggestions, which have helped us improve the quality and clarity of our manuscript.

1. Even without Figure 6 and Figure 7, we believe the present work is already qualified for Nature Communications. However, the inclusion of Figures 6 and 7—especially Figure 6, which introduces the NglyBAI algorithm—complicates the logical flow (we spent a long time reading and trying to understand this part). We therefore suggest that the authors consider moving the NglyBAI figure to the supplementary materials (or consider publishing it in a separate paper) and downplaying the associated results for the following reasons: a. The N-glycome and N-glycoproteomic datasets are all derived from serum or blood, where N-glycan profiles can be very different from those in cells. Inferring glyco-enzyme activities from such data is theoretically challenging and would require extensive validation, which goes beyond the scope of this study. b. The manuscript already contains 8 main figures. c. It is not clear how increased/decreased mRNA expression in Figure 7 supports the results inferred in Figure 6.

R: We sincerely thank the reviewer for this thoughtful suggestion. We agree that the NGlyBAI analysis is exploratory, and risks complicating the main narrative. To improve logical flow and avoid over-interpretation, we have removed all NGlyBAI-related content (previously Figure 6) from the manuscript. This algorithm will be further refined and comprehensively validated in a subsequent, dedicated study.

Regarding the transcriptomic analysis (previously Figure 7), we believe it provides useful

mechanistic context for the glycomic alterations observed in our study. We therefore retain this analysis for completeness. We hope these revisions address the reviewer's concerns and strengthen the clarity and focus of the manuscript.

2. Pseudo-Glycome: Although the results between the pseudo-glycome and the real glycome are highly correlated at the end (Figure 5A), the authors should include a cautionary note in the Results and perhaps in the Discussion as well, because it can be problematic to sum mass spec signals across different proteins and different IGP (Figure 5). Emphasizing the “surprising” consistency is encouraged, but the limitations should also be acknowledged.

R: We appreciate the reviewer's insightful comment. We agree that the pseudo-glycome represents an estimation strategy rather than a rigorous quantitative approach, since summing signals across different proteins and glycopeptides can introduce potential biases. In response, we have revised both the Results section to explicitly acknowledge this limitation. At the same time, we emphasize the high consistency observed between the pseudo-glycome and the real glycome, while clarifying that these results should be interpreted with caution.

We have added this paragraph to the Results section:

“It should be noted that the pseudo-glycome represents an estimation strategy by aggregating signals across different proteins and glycopeptides, rather than a strict quantitative approach. While the high consistency with the real glycome is indeed remarkable, this agreement should be interpreted with caution given the potential biases introduced by signal summation.”

As for the Discussion section, since we did not directly draw conclusions from the pseudo-glycome analysis in that part, we have chosen not to reiterate it to avoid redundancy.

3. Line 88: Please briefly mention how quantification was done for MALDI.

R: We thank the reviewer for pointing this out. In response, we have added a brief description at Line 88 to clarify the quantification method:

“Identification was done using in-house tool GlyHunter and quantitative information was obtained by applying median quotient normalization to peak intensities after removing glycans with over 50% missing values (see Methods for details).”

4. Throughout the figures, the authors should provide the number of individuals included in their analysis and plots (e.g., Figure 2a, Figure 3a–k, etc.).

R: We thank the reviewer for this helpful suggestion. We have revised the figures and their captions to ensure that each main figure now includes at least one sub-figure with clear annotation of the number of individuals analyzed. This has been applied to Figure 2a, Figure 3a–c, and Figure 4a, as suggested.

5. In all supplementary figure legends, the authors should provide more information and descriptions—e.g., about colors, numbers, statistical tests, and abbreviations (e.g., Supplementary Figures 2 and 9).

R: We thank the reviewer for this valuable suggestion. We have revised the supplementary figure legends (including previous Supplementary Figures 2 and 9) to provide more detailed information. Specifically, we added explanations for color schemes, sample sizes, statistical annotations, and all abbreviations not mentioned in the main text.

6. Currently, there are 34 supplementary figures. The authors are advised to combine some into multi-panel figures to reduce this number.

R: We appreciate the reviewer's suggestion. In response, we have combined several supplementary figures into multi-panel formats, reducing the total number from 34 to 25.

7. Figure 3g: The authors should follow standard glycomics nomenclature here (i.e., to be consistent with Figure 3a—e.g., H3N5).

R: We thank the reviewer for pointing this out. We have revised Figure 3g to follow the standard glycomics nomenclature (e.g., H3N5), ensuring consistency with other figures and the community standard.

8. The authors should consider adding a supplementary table with patient-specific information (instead of only the summary table currently shown as Table 1), to allow future analysis of their data.

R: We thank the reviewer for this helpful suggestion. In response, we have added a supplementary table (Supplementary Data 8) containing detailed patient-specific clinical information, in addition to the summary provided in Table 1.

9. Line 716: Please specify whether DIA or DDA analyses were used for total proteome and glycoproteomic analysis, respectively.

R: We thank the reviewer for this comment. Both the total proteome and glycoproteomic analyses were performed in DDA mode, and this information has now been added to the corresponding section in the Results.

10. Line 850: If the authors decide to include NglyBAI in this manuscript, they should provide the code or package on GitHub.

R: We thank the reviewer for this suggestion. As noted above, we have removed the NGlyBAI-related content from the manuscript, and therefore no code release is required.

Reviewer #2

This manuscript by Lu and colleagues presents a comprehensive study examining serum N-glycosylation dynamics during hepatocellular carcinoma (HCC) progression. Through meticulous multi-cohort analysis and sophisticated multi-omics approaches, the authors have systematically characterized serum N-glycosylation associated with HCC development and assessed their potential as early diagnostic biomarkers. The scale of data collection and analytical rigor demonstrated throughout this work are impressive.

The diagnostic machine learning models developed from N-glycome data represent a significant advancement, notably outperforming alpha-fetoprotein (AFP), the current clinical standard. The authors' efforts to elucidate underlying mechanisms through deconstructing glycomic changes into "carrier protein abundance-driven" versus "glycosylation process-driven" modes demonstrates exceptional scientific depth. Furthermore, the novel NGlyBAI algorithm for enzyme activity inference represents a methodological innovation with broad potential applications.

R: We sincerely thank the reviewer for the overall positive feedback. We greatly appreciate all the suggestions, which have helped us improve the quality and clarity of our manuscript.

1. The 'Mean Eigenvalue' parameter used to characterize HCC subtypes in Figure 4c lacks methodological clarity. A detailed description of its calculation methodology should be included in the Methods section to ensure reproducibility and facilitate interpretation.

R: We thank the reviewer for pointing out this issue. In Figure 4c and its legend, the label should have read eigenglycans rather than eigenvalues. We have corrected this to ensure consistency with Figure 3f. In addition, we have added a detailed description in the Methods section to clarify the calculation of eigenglycans, as follows:

“We adapted the concept of “eigengenes” from the well-established weighted gene correlation network analysis (WGCNA) framework to define “eigenglycans” for each glycan co-expression module (GCM). Specifically, eigenglycans were calculated as the first principal component of the z-score-transformed abundances of all glycans within a GCM, thereby representing the overall expression pattern of the module.”

2. Figure 5j presents an apparent contradiction requiring resolution: H5N3S1 glycan modification on A2MG peaks in the CHB group, while total A2MG protein levels are elevated in LC and HCC groups. This observation contradicts a protein-driven mechanism for H5N3S1 elevation in CHB and requires explicit discussion in the Figure 5 narrative.

R: We thank the reviewer for carefully pointing out this issue and apologize for not making it sufficiently clear in the original manuscript. The observation in Figure 5j does not represent a contradiction, but rather illustrates the glycosylation-dominant mechanism. In this mechanism, specific glycomic alterations arise from dysregulation of upstream glycosylation pathways, rather than changes in the overall abundance of the carrier protein.

To clarify this point, we have revised the relevant section of the Results as follows:

“By integrating glycomics and glycoproteomics data, we observed two distinct patterns underlying glycomic alterations: 1) glycosylation-dominant pattern and 2). protein-dominant pattern. The glycosylation-dominant pattern indicates that certain glycomic alterations arise from dysregulation of upstream glycosylation pathways. For instance, H5N3S1, a mono-sialylated hybrid glycan elevated in glycomic data across all liver diseases, particularly in CHB (Fig. 3f, 3g), was predominantly linked to the N869 site on alpha-2-macroglobulin (A2MG) ($\log_2FC = 7.73$, $FDR = 0.00032$) (Fig. 5i). Further analysis of this site revealed a marked increase in H5N3S1 in CHB, despite only a modest elevation in global A2MG levels across liver disease progression (Fig. 5j). This exemplifies how dysregulation of N-glycosylation itself contributes to phenotypic N-glycome changes. In the protein-dominant pattern, alterations in the expression levels of certain highly abundant glycoproteins can exert a non-negligible impact on the observed glycome, even in the absence of changes in the upstream glycosylation regulatory pathways, as demonstrated by the increased levels of tetra-antennary, fucosylated, multisialylated glycans (H7N6F1S2, H7N6F1S3, and H7N6F1S4), primarily driven by elevated A1AG1, an acute-phase immune protein (Supplementary Fig. 18). This highlights an important perspective for future glycomics research, namely that glycomic data alone may not be sufficient to provide regulatory insights.”

We hope this clarification resolves the reviewer’s concern and makes the underlying mechanisms more explicit.

3. The manuscript would be strengthened by reporting whether the NGlyBAI algorithm identified glycotransferases with significantly different activity scores specifically between LC and HCC groups. Such findings would substantially enhance the algorithm's clinical utility by demonstrating its capacity to detect malignancy-specific enzymatic alterations beyond general liver disease-associated dysregulation.

R: We sincerely thank the reviewer for this constructive suggestion. We agree that distinguishing malignancy-specific enzymatic alterations from general liver disease-associated dysregulation would be critical for demonstrating the clinical utility of the NGlyBAI algorithm. However, after careful consideration and in line with another reviewer's advice, we have decided to remove the NGlyBAI analysis from the current manuscript.

4. A compelling paradox emerges between glycomic and transcriptomic findings: decreased overall galactosylation in HCC (Fig. 3k) contrasts with the association between high B4GALT2 and B4GALT3 expression and poorer prognosis (Fig. 7d-g). This intriguing discrepancy between gene expression and resulting glycan phenotype warrants thorough mechanistic discussion.

R: We thank the reviewer for raising this important point. We agree that the apparent discrepancy between decreased overall galactosylation in HCC (Fig. 3k) and the association of higher B4GALT2/3 expression with poorer prognosis (Fig. 7d–g) requires clarification.

In fact, these findings are not contradictory. In the human liver, B4GALT1 is the predominant enzyme responsible for β 1,4-galactosylation, whereas B4GALT2 and B4GALT3 play only minor roles under physiological conditions in specific organs like brain. In our transcriptomic analysis, B4GALT1 was not detected as significantly expressed, but previous studies have reported its downregulation in HCC tissues¹, which is consistent with our observation of reduced overall galactosylation.

The upregulation of B4GALT2 and B4GALT3 may represent a compensatory feedback mechanism in response to the loss of B4GALT1 activity. However, because these enzymes exhibit more restrictive substrate specificities and have low enzymatic activities in most organs, they cannot fully compensate for the function of B4GALT1. As a result, despite increased B4GALT2/3 expression, the global galactosylation level still declines in HCC.

We have added the following paragraph to relevant position in the Results section to clarify this point:

“It is worth noting that B4GALT1, the predominant β 1,4-galactosyltransferase in human liver, was not significant in our transcriptomic dataset but has been reported to be downregulated in

HCC tissues¹. Therefore, despite the observed upregulation of B4GALT2 and B4GALT3, their more restrictive substrate specificities and narrow organ distributions may not fully compensate for the loss of B4GALT1 activity, which is consistent with the overall decrease in galactosylation observed in our glycomics data.”

1. Nwosu, Z. C. *et al.* Identification of the Consistently Altered Metabolic Targets in Human Hepatocellular Carcinoma. *Cellular and Molecular Gastroenterology and Hepatology* 4, 303-323.e1 (2017).

5. For completeness and methodological clarity, comprehensive survival analyses for both B4GALT2 and B4GALT3 across all four endpoints (OS, DSS, PFI, DFI) should be provided, either in supplementary materials or with a clear rationale for the current selective presentation.

R: We thank the reviewer for this constructive suggestion. In our initial submission, we selectively presented the survival curves for B4GALT2 and B4GALT3 because they showed the strongest associations (smallest *p*-values) with the four survival endpoints examined (OS, DSS, PFI, and DFI). Specifically, B4GALT2 was most strongly associated with OS and DSS, while B4GALT3 was most strongly associated with PFI and DFI.

To ensure completeness and transparency, we have now provided the comprehensive survival analysis results for all four endpoints across all glycogenes in Supplementary Data 9.

6. Model 3's external validation performance (AUC=0.76) for distinguishing HCC from LC, while still impressive, is lower than other models. This reflects the well-documented clinical challenge of differentiating HCC from its cirrhotic background and merits acknowledgment in the Discussion section.

R: We thank the reviewer for this important point. We agree that the relatively lower performance of Model 3 (AUC = 0.76 in external validation) highlights the well-recognized clinical difficulty of distinguishing HCC from cirrhosis. We have acknowledged this limitation in the Discussion section by adding the following sentence:

“Notably, the diagnostic performance for distinguishing HCC from cirrhosis (Model 3) was lower than for other comparisons, reflecting the well-documented clinical challenge of separating malignancy from its cirrhotic background^{1,2}.”

1. Vogel, A., Meyer, T., Sapisochin, G., Salem, R. & Saborowski, A. Hepatocellular carcinoma. *Lancet* 400, 1345–1362 (2022).
2. Johnson, P., Zhou, Q., Dao, D. Y. & Lo, Y. M. D. Circulating biomarkers in the diagnosis

and management of hepatocellular carcinoma. *Nat Rev Gastroenterol Hepatol* 19, 670–681 (2022).

7. The correlation between Model 4 probability score (PP4) and disease severity raises important translational questions that should be addressed: Could this quantitative metric enable longitudinal risk monitoring in cirrhotic patients? Moreover, could a streamlined model using fewer N-glycan features maintain diagnostic efficacy while enhancing clinical implementability?

R: We sincerely thank the reviewer for raising these important translational questions.

1. Regarding the potential of PP4 for longitudinal risk monitoring in cirrhotic patients, we fully agree that this is a highly relevant and clinically valuable direction. Unfortunately, after discussion with our collaborating hospitals, we confirmed that most of the patients could no longer be contacted because longitudinal follow-up was not part of the original consent and study design. As a result, we were unable to obtain sufficient longitudinal data to directly evaluate this possibility within the current study. We have acknowledged this limitation in the revised manuscript.
2. In response to the reviewer's second point on clinical implementability, we retrained the diagnostic models using only the six most important glycans. As shown in Supplementary Fig. 22 or Fig. R1, the streamlined models achieved nearly identical performance across all three cohorts, with only a slight, non-significant decrease compared to the full 26-glycan models (DeLong test, $p > 0.05$). We believe this strongly supports the translational potential of glycan-based models for practical clinical use.

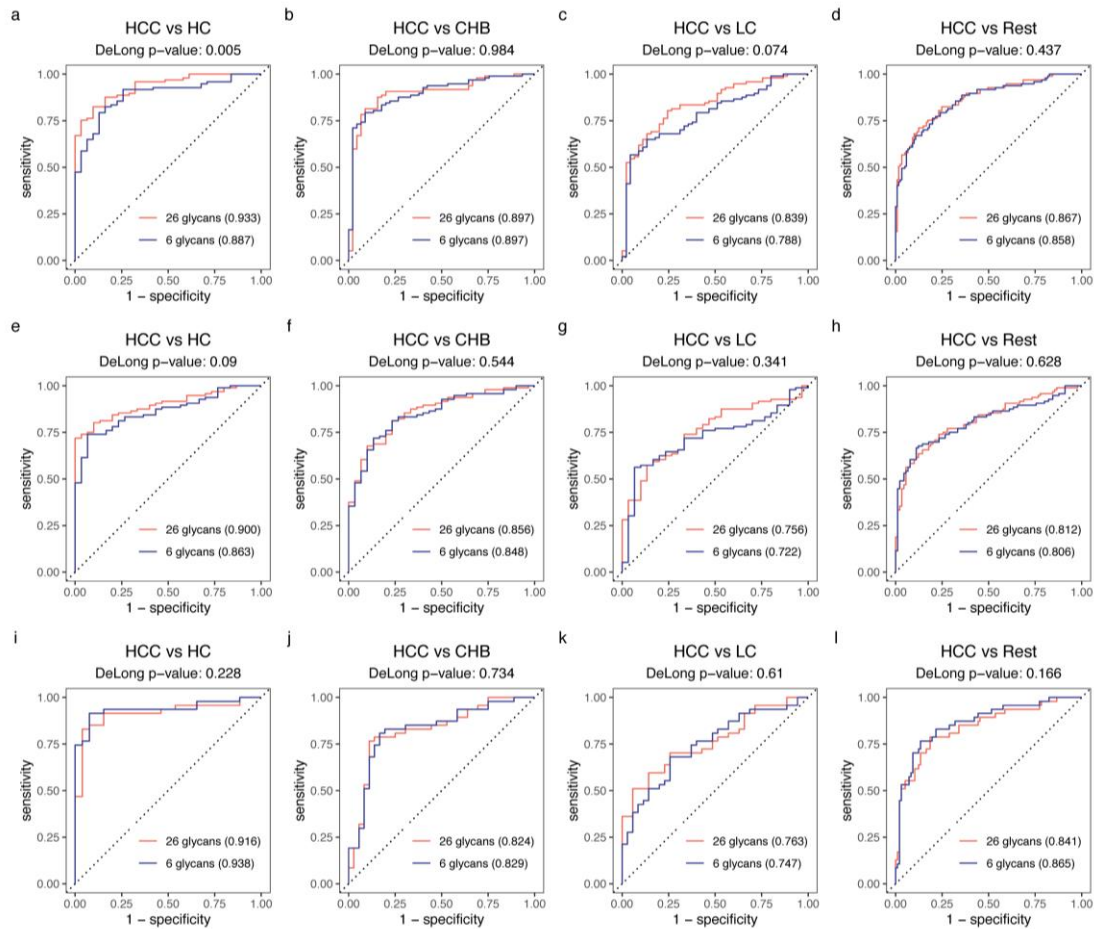


Figure R1 (Supplementary Figure 22). ROC curves comparing models trained with all 26 glycans or the six most important glycans. Models were compared through in all three cohorts: (a-d) the test cohort, (e-h) the external validation cohort 1, and (i-l) the external validation cohort 2. ROC curves were compared using DeLong test.

We have revised the Results section accordingly to highlight these findings, adding the following description:

“Furthermore, the importance of these glycans was validated by training ML models using only the six most important glycans (Supplementary Fig. 22). The models maintained nearly the same performance across all three cohorts, with only a slight, non-significant decrease (DeLong test, $p > 0.05$).”

We have also added the following limitation to the Discussion section:

“Another limitation of our study is the lack of longitudinal follow-up data, which prevented us from directly assessing the utility of PP4 for risk monitoring in cirrhotic patients.”

Reviewer #3

This study represents a solid glycan analysis of over 1000 patients, the majority of which have liver disease. The mix of patients is good and also consist of over 400 patients with hepatocellular carcinoma.

R: We sincerely thank the reviewer for the careful evaluation of our manuscript and for providing constructive feedback. We fully appreciate the reviewer's concerns regarding novelty, clinical relevance, and contextualization within existing literature. We take these comments very seriously and have substantially revised the manuscript to address them. In particular, we have (i) expanded our review of prior serum glycomics studies to better acknowledge existing knowledge, (ii) clarified how our work builds upon and extends these findings through larger cohorts, multi-omics integration, and rigorous external validation, and (iii) revised the Discussion and Introduction to ensure that the novelty and contribution of this study are appropriately framed.

1. The major issue is that very similar work, in both serum and tissue has been done by numerous groups. Nothing in the work presented is novel or alters our understanding of liver disease. For example, the changes observed in branching with the development of HCC have been observed by a many groups, at both the tissue level, in serum and on specific glycoproteins. The alterations in fucosylation with LC and HCC again have been observed by many groups and in the case of LC, linked with specific immunoglobulins and in the case of HCC, with other well known proteins. Again, nothing new is presented here.

R: We sincerely thank the reviewer for this critical comment. We agree that branching and fucosylation changes have been reported previously in both tissue and serum studies, and we are grateful for the opportunity to clarify how our work advances beyond earlier findings (Table R1).

Table R1. Comparison of published serum N-glycomics studies for HCC diagnosis and the present study.

Year	DOI	Techniques	Sia.	Sia. Iso.	#G	Sample Group	Sample Size	Omics	Ex. Val.	ROC
2025	Ours	MALDI-TOF	Yes	Yes	201	HC, CH, LC, HCC	1074	T, G, GP	Yes	HCC vs HC: 0.93; HCC vs CHB: 0.90; HCC vs LC: 0.84
2025	10.1097/HEP.0000000000001316	DSA-FACE	No	No	9	LC, HCC	397	G	No	HCC vs LC: 1.00
2024	10.1186/s12964-024-01569-y	MALDI-TOF	Yes	No	36	HC, CH, LC, HCC	80	T, G, GP	No	/
2024	10.1016/j.cca.2024.117827	LC-MS	Yes	No	55	HC, HCC	60	G	No	HCC vs HC: 0.91
2022	10.21037/jgo-22-93	DSA-FACE	No	No	9	HC, CH, LC, HCC	81	G	No	HCC vs LC: 0.87
2021	10.1016/j.clinbiochem.2020.12.005	DSA-FACE	No	No	/	HC, CH, LC, HCC	80	G	No	HCC vs LC: 0.88
2020	10.3390/ijms21238913	HPLC-FLR	No	No	22	HC, CH, LC, HCC	182	G	No	HCC vs LC: 0.96
2020	10.1002/ijc.33564	NGFP	No	No	13	HC, HCC	1451	T, G	Yes	HCC vs HC: 0.87
2014	10.1111/hepr.12441	MALDI-TOF	Yes	No	36	CH, LC, HCC	219	G	No	HCC vs CH/LC: 0.81
2013	10.1002/prca.201200125	HPLC-FLR	No	No	27	HC, LC, HCC	60	G	No	HCC vs LC: 0.85
2010	10.1002/ijc.25030	DSA-FACE	No	No	9	HC, LF, HCC	393	G	No	HCC vs LF: 0.87
2009	10.1158/1078-0432.CCR-07-5261	MALDI-TOF	Yes	No	83	HC, CH, HCC	202	G	No	HCC vs CH: 0.96
2007	10.1002/hep.21855	DSA-FACE	No	No	9	HC, LF, LC, HCC	580	G	No	HCC vs LC: 0.81

We would like to emphasize several aspects where our study provides novelty and broader impact:

1) Cohorts and study design

Our study represents the largest mass spectrometry (MS)-based serum N-glycomics investigation of HCC to date, encompassing 1,074 individuals recruited from three independent hospitals. The study design included one discovery cohort and two external validation cohorts, ensuring robust evaluation across different clinical centers and demographic backgrounds. Importantly, we covered the entire liver disease progression spectrum—healthy controls (HC), chronic hepatitis B (CHB), liver cirrhosis (LC), and hepatocellular carcinoma (HCC). This comprehensive design allowed us not only to track dynamic N-glycosylation remodeling along hepatocarcinogenesis, but also to separate disease-type-specific effects from those related to progressive liver function deterioration. Such coverage and validation have not been achieved in previous glycomics studies.

2) Methods and analytical framework

We leveraged our previously developed high-throughput N-glycan sample preparation platform, enabling the most extensive serum N-glycan characterization so far, including sialylated structures and uniquely linkage-specific sialic acid isomers in case-control cohorts.

Methodologically, we extended beyond descriptive glycomics by introducing several innovations:

1. Serum glycomics-based molecular subtyping of HCC, providing a proof-of-concept that serum glycome profiles can reflect the heterogeneity of pathological processes within HCC.
2. Integration with glycoproteomics, which enabled us to distinguish two mechanistic drivers of glycan alterations: (i) the glycosylation-dominant mode, in which enzymatic dysregulation alters glycan occupancy (e.g., H5N3S1 elevation at A2MG N869 in CHB), and (ii) the protein-dominant mode, in which protein abundance drives apparent glycan changes (e.g., A1AG1-driven tetra-antennary glycans).
3. Incorporation of transcriptomics, which added mechanistic depth and clarified paradoxes, such as the decrease in overall galactosylation in HCC despite poor prognosis associated with B4GALT2/3 upregulation—highlighting instead the dominant role of B4GALT1 and its downregulation in HCC.
4. Rigorous statistical modeling controlling for liver function, which allowed us to separate glycomic changes reflecting hepatic functional deterioration from those that are truly HCC-specific.

For translation, we built machine learning diagnostic models trained on 26 glycans and further

developed a streamlined model based on the six most informative glycans. The simplified model retained comparable performance with no significant decrease (DeLong test, $p > 0.05$), demonstrating feasibility for clinical application.

3) Key findings and conclusions

Our integrated analyses yielded several important discoveries:

1. We demonstrated that glycomic alterations previously attributed to HCC—such as increased bisected glycans and core fucosylation—are in fact more strongly associated with declining liver function (as measured by ALBI score), rather than being HCC-specific.
2. We identified tetra-antennary glycan upregulation as a unique feature of HCC, independent of liver function deterioration. This represents a novel serum-based insight into HCC-specific glycosylation dynamics.
3. We established a two-mechanism conceptual framework (glycosylation-dominant vs. protein-dominant) that provides a new paradigm for interpreting serum glycomics data in mechanistic terms, moving the field beyond “black box” descriptions.
4. Our machine learning models achieved high diagnostic accuracy, significantly outperforming the clinical biomarker AFP, including in AFP-negative patients and in early-stage (TNM I) HCC, and their robustness was confirmed across two independent external cohorts.

In summary, while it is true that some glycan alterations such as branching and fucosylation have been reported before, our study goes far beyond replication. It contributes novelty in scale, technical depth, multi-cohort validation, and integrative mechanistic analysis, while also offering new biological insights (HCC-specific tetra-antennary glycan upregulation, molecular subtyping). Together, these advances expand our understanding of N-glycosylation in liver disease progression and provide a stronger foundation for future translational applications.

2. In addition, as the results in Figure 8, are not impressive when compared to what is either already clinically available or what is in the literature.

R: We thank the reviewer for this critical comment and the opportunity to clarify.

1. Comparison with AFP

We directly compared our four machine learning models against the widely used clinical biomarker AFP. As shown in Supplementary Fig. 21c, all four models significantly outperformed AFP across diagnostic tasks. Moreover, combining AFP with glycan features did not result in a significant improvement (Supplementary Fig. 21d), indicating that our glycomic

models not only capture the diagnostic information contained in AFP but also provide additional discriminatory power beyond it.

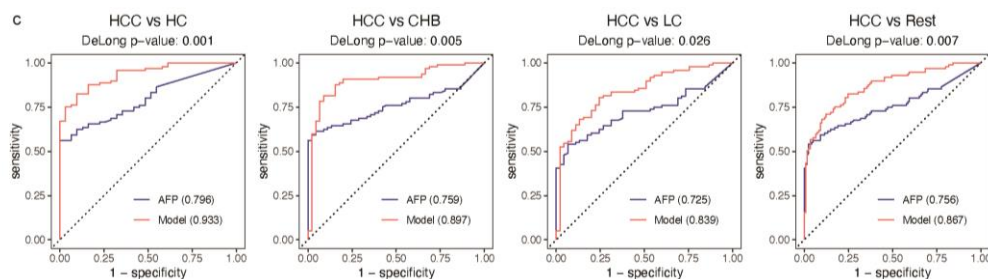


Figure R2 (Supplementary Figure 21c). ROC curves of machine learning models and AFP for diagnosing HCC. ROC curves were compared using DeLong test.

2. Validation across multiple cohorts

A key strength of our study is that model performance was validated in one internal and two external validation cohorts, spanning three hospitals from different regions and demographics. To our knowledge, only one previous study¹ on serum glycomic biomarkers for HCC has included two external validation cohorts like us, and its reported diagnostic performance was lower (AUC = 0.87 for HCC vs HC), compared with AUC = 0.93 achieved by our model. For biomarker studies—particularly machine learning models based on multiple features—lack of external validation is a well-recognized source of inflated performance estimates². This limitation has been highlighted in recent high-profile discussions, including a Science (2024) article³. By using large sample sizes and validating across three independent cohorts, we substantially reduce the risk of overfitting and demonstrate that our models generalize well across different clinical settings.

Taken together, while some glycomic alterations have indeed been reported previously, our study provides stronger and more generalizable diagnostic evidence than prior literature, with performance that is both statistically superior to AFP and validated across large, demographically diverse patient populations.

To emphasize this point, we added the following conclusive sentence to Results, after comparing the performance of our models and AFP:

“These results highlight the high diagnostic performance and robustness of our N-glycome-based models, which not only outperform the current clinical biomarker AFP but also demonstrate consistent effectiveness across two independent external validation cohorts.”

1. Higashi, M. et al. A Potential Serum N-glycan Biomarker for Hepatitis C Virus-Related Early-Stage Hepatocellular Carcinoma with Liver Cirrhosis. *IJMS* 21, 8913 (2020).
2. Nakayasu, E. S. et al. Tutorial: best practices and considerations for mass-spectrometry-based protein biomarker discovery and validation. *Nat Protoc* 16, 3737–3760 (2021).

3. Chekroud, A. M. et al. Illusory generalizability of clinical prediction models. *Science* 383, 164–167 (2024).

3. Lastly, a much better job of reviewing the literature should be done to improve the manuscript. The implication that this is the first such study of this kind is disingenuous.

R: We thank the reviewer for this important comment and apologize if our wording gave the impression that no similar studies had been conducted before. We acknowledge that several groups have previously reported glycomic alterations in HCC and related liver diseases, particularly changes in branching, fucosylation, and sialylation in both serum and tissue contexts. These pioneering works have provided valuable biological insights and laid the foundation for our study.

To address this point, we have substantially revised the Introduction to more thoroughly review the relevant literature and clearly position our work in relation to prior studies. In particular, we now highlight that while earlier studies have made important contributions, they often suffered from limitations such as small sample sizes, lack of covariate control (e.g., liver function), absence of external validation, limited characterization of sialylation linkage isomers, and minimal integration with glycoproteomics or transcriptomics.

We then clarify that our study does not claim to be the first of its kind, but rather seeks to extend prior work through several key advances: (i) the largest serum N-glycomics cohort to date (n = 1,074) covering the full hepatocarcinogenesis spectrum; (ii) the most comprehensive glycan profiling, including linkage-specific sialylation isomers; (iii) integration with glycoproteomics and transcriptomics to provide mechanistic explanations; and (iv) robust machine learning models validated in two independent external cohorts.

We hope these revisions make it clear that our work builds upon and complements prior literature, while also providing novel insights and methodological advances.

Attached below are the revised last two paragraphs of the Introduction:

“Several serum N-glycomics studies have previously reported alterations associated with hepatocellular carcinoma (HCC) and related liver diseases, including branching, fucosylation, and sialylation changes in both serum and tissue contexts^{13,18–21}. These pioneering studies have provided important biological insights and laid the foundation for considering glycans as potential biomarkers. However, despite these advances, current glycomic studies still face several critical limitations: (i) relatively small sample sizes that limit the ability to capture patient heterogeneity, (ii) lack of rigorous covariate control (e.g., liver function) in statistical analyses, (iii) limited or absent external validation, which may inflate performance estimates, (iv) incomplete characterization of sialylation patterns and linkage-specific differences across

disease progression stages, and (v) limited integration with glycoproteomics to resolve site-specific contributions of key glycans.

This study advances HCC glycomics research by addressing these gaps. First, we established a large-scale N-glycomic profiling cohort encompassing 1,074 Chinese subjects across the hepatocarcinogenesis spectrum, from healthy controls through chronic hepatitis and cirrhosis to HCC. Systematic trajectory analysis revealed dynamic glycosylation remodeling patterns closely correlated with progressive hepatic functional deterioration, while also distinguishing malignancy-specific glycomic traits. Second, transcending conventional single-omics approaches, we implemented a multi-dimensional analytical framework integrating quantitative glycomics, site-specific glycoproteomics, and transcriptomic profiling to provide mechanistic depth. Third, leveraging AutoML and explainable AI, we constructed robust diagnostic models validated across geographically distinct cohorts, demonstrating strong generalizability for clinical translation. Together, these innovations build upon prior work while substantially extending its scope, providing not only stage-specific glyco-biomarkers but also mechanistic insights into their biosynthetic regulation.”

Reviewer #4

In this manuscript, Fu et al. presents a large-scale, multi-omics investigation into serum N-glycosylation changes during hepatocellular carcinoma (HCC) development. Using a cohort of over a thousand (Chinese) individuals spanning healthy, chronic hepatitis B, liver cirrhosis and HCC, systematically profiled serum N-glycans via MALDI-TOF MS and LC-MS/MS. The analytical and bioinformatic methodologies used are well established in the field. Using this combination of methods, Fu et al identified 48 N-glycans as significantly dysregulated across disease groups and demonstrated correlations with liver function (e.g. the ALBI score), reflecting hepatic synthetic capacity. In general, the manuscript is very well written, and clearly an impressive amount of work has gone into this work and is represented on the pages of the manuscript. The authors also make their data available. However, I do have a few comments on the manuscript in its current form.

R: We sincerely thank the reviewer for the overall positive feedback. We greatly appreciate all the suggestions, which have helped us improve the quality and clarity of our manuscript.

1. While there are statistically significant differences in the distributions of glycan types between the various groups and controls (e.g. Figure 2c, Figure 3k), there are also considerable overlaps between the distributions, which means these global glycan type differences in themselves are not useful as diagnostic markers. Neither are the HCC

subtypes well separated by PCA (Figure 4b) based on the expression of glycan types. This does not match the titular claim that “Large-Scale and multi-cohort Serum N-glycomics Tracks N-Glycosylation Dynamics in Hepatocellular Carcinoma Progression and Enables Early Diagnosis”. I am not convinced the authors have shown that their method enables early diagnosis (which by definition is of individual patients, not cohorts).

R: We thank the reviewer for this important observation. We agree that the exploratory analyses based on individual glycan traits—such as statistical comparisons, boxplots, and PCA—are not intended to reflect diagnostic capacity. Rather, these analyses were designed to illustrate global glycosylation dynamics across disease progression. As the reviewer notes, such global shifts cannot serve as reliable standalone diagnostic markers due to substantial overlaps among groups.

Our main diagnostic contribution is presented in the final section of the manuscript, where we employed machine learning (ML) models trained on multi-glycan profiles. These models demonstrated excellent diagnostic performance and, critically, their robustness was confirmed in two independent external validation cohorts.

We agree that “early diagnosis” ultimately refers to identifying disease at the individual patient level. However, the evaluation of biomarkers—whether single markers or multi-feature machine learning models—necessarily relies on cohort-level analyses such as ROC curves, sensitivity, and specificity. These cohort-based metrics are the standard means to assess whether a biomarker has the statistical power to be translated into individual-level early detection.

To further strengthen this point during revision, we added three additional analyses or revisions:

1. We evaluated the diagnostic utility of the ML models across different TNM stages of HCC (Supplementary Fig. 21e or Fig. R3). Importantly, the models achieved decent performance compared to AFP, even in TNM stage I patients (ROC AUC = 0.86 for TNM-I HCC vs HC), thereby supporting their potential for early diagnosis at the individual-patient level. As a revision, we added the following paragraph to the Results section: *“To further assess the potential for early diagnosis, we evaluated the performance of the four models across different TNM stages (Supplementary Fig. 21e). All four models achieved ROC AUCs above 0.80 for stages II, III, and IV, and they retained acceptable diagnostic power even for stage I HCC patients, with AUC values reaching 0.859 (HCC vs HC). These findings demonstrate that our models hold strong promise for the early detection of HCC.”*
2. We constructed a simplified model using only the six most important glycans identified by SHAP analysis (Supplementary Fig. 22 or Fig. R4). This compact model achieved comparable performance to the full 26-glycan model, underscoring the robustness and diagnostic power of the glycan-based approach. These sentences were added to the Results section: *“Furthermore, the importance of these glycans was validated by training ML*

models using only the six most informative glycans (Supplementary Fig. 22). The models maintained nearly the same performance across all three cohorts, with only a slight, non-significant decrease (DeLong test, $p > 0.05$)."

3. We added in the Discussion section that "while global shifts in glycan types offer valuable biological insights, their diagnostic utility is limited; the true diagnostic strength emerges from integrative machine learning models that leverage the full multi-dimensional glycan profile."

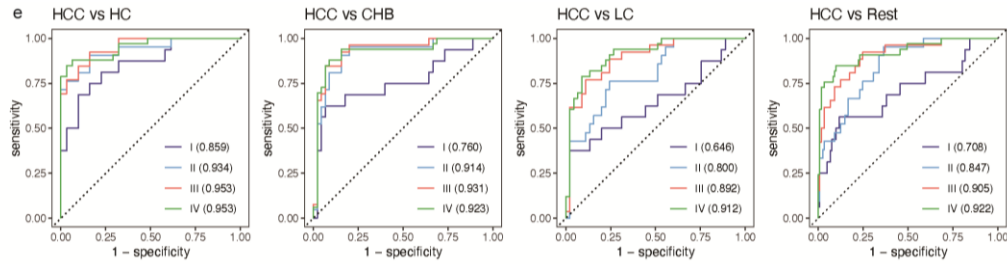


Figure R3 (Supplementary Figure 21e). ROC curves for models differentiating HCC with different TNM stages (I-IV) from controls.

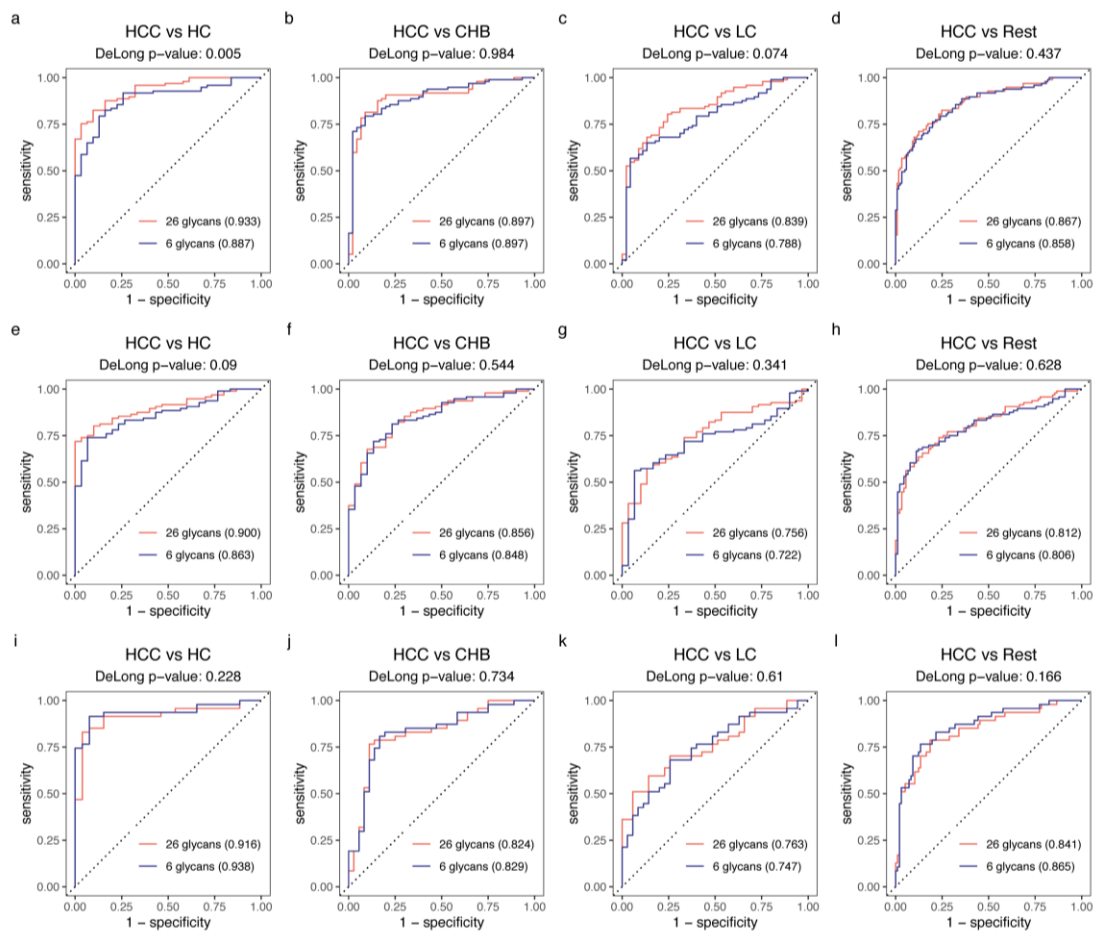


Figure R4 (Supplementary Figure 22). ROC curves comparing models trained with all 26

glycans or the six most important glycans. Models were compared through in all three cohorts: (a-d) the test cohort, (e-h) the external validation cohort 1, and (i-l) the external validation cohort 2. ROC curves were compared using DeLong test.

To avoid overstatement, we have clarified the manuscript text accordingly, ensuring that our claim of “early diagnosis” is firmly grounded in the ML results and not in the exploratory analyses of individual glycans.

2. Lines 367-368: The authors state that “all three models maintained comparable performance in the external validation cohorts”. What is meant by “comparable” is not defined, but it is clear that all models perform worse (or less well) in both external validation cohorts. This should also be stated, even though it of course is expected.

R: We thank the reviewer for pointing out this ambiguity. In the revised manuscript, we have clarified this statement to acknowledge the expected performance drop in external validation. The sentence now reads:

“Notably, all three models maintained consistent performance in the external validation cohorts, with only a slight decrease (GZ-II and XZ; Fig. 7c–e, Supplementary Fig. 21a), despite potential differences in population characteristics and medical facilities, demonstrating robust generalization and resilience to distribution shifts.”

3. Figure 1: Typo “recruided”, unusual/obsolete word “dismangles”

R: We thank the reviewer for noticing these typographical errors. We have corrected them in Figure 1: “recruided” → “recruited” and “dismangles” → “disentangles”.

4. Several figure elements, including the patient and fridge (Figure 1), liver icon (Figure 2), and patient and doctor icons (Figure 3) are from BioRender rather than the authors themselves. This must be acknowledged.

R: We thank the reviewer for pointing this out. We have revised the corresponding figure captions (Figures 1-3) to explicitly acknowledge that certain elements were created with BioRender.

5. The in-house scripts could also be made available.

R: We thank the reviewer for this suggestion. All in-house scripts were version-controlled using

Git and GitHub. We have now made the repository public and it is accessible at <https://github.com/fubin1999/HCC-N-glycome>. The Code Availability section has been updated accordingly.