The Landscape of Glycosyltransferase Gene Signature for Overall Survival Prediction in Hepatocellular Carcinoma

Qiang Cai
Affiliated Hospital of Yunnan University

Shizhe Yu
First Affiliated Hospital of Zhengzhou University

Jian Zhao
Affiliated Hospital of Yunnan University

Duo Ma
Affiliated Hospital of Yunnan University

Long Jiang
Affiliated Hospital of Yunnan University

Xinyi Zhang
Affiliated Hospital of Yunnan University

Zhiyong Yu (xsrsyby@outlook.com)
Affiliated Hospital of Yunnan University

Research Article

Keywords: Hepatocellular carcinoma, glycosyltransferase gene, prognosis, overall survival, risk score

DOI: https://doi.org/10.21203/rs.3.rs-114747/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background: Hepatocellular carcinoma (HCC) is a heterogeneous disease occurring in the background of chronic liver diseases. The role of glycosyltransferase (GT) genes has recently been the focus of research associating with the development of tumors. However, the prognostic value of GT genes in HCC remains not elucidated. This study aimed to demonstrate the GT genes related to the prognosis of HCC through bioinformatics analysis.

Methods: The GT genes signatures were identified from the training set of The Cancer Genome Atlas (TCGA) dataset using univariate and the least absolute shrinkage and selection operator (LASSO) Cox regression analyses. Then, we analyzed the prognostic value of GT genes signatures related to the overall survival (OS) of HCC patients. A prognostic model was constructed, and the risk score of each patient was calculated as formula, which divided HCC patients into high- and low-risk groups. Kaplan-Meier (K-M) and Receiver operating characteristic (ROC) curves were used to assess the OS of HCC patients. The prognostic value of GT genes signatures was further investigated in the validation set of TCGA database. Univariate and multivariate Cox regression analyses were performed to demonstrate the independent factors on OS. Finally, we utilized the gene set enrichment analysis (GSEA) to annotate the function of these genes between the two risk categories.

Results: In this study, we identified and validated 4 GT genes as the prognostic signatures. The K-M analysis showed that the survival rate of the high-risk patients was significantly lower than that of the low-risk patients. The risk score calculated with 4 gene signatures could predict OS for 3-, 5-, and 7-year in patients with HCC, revealing the prognostic ability of these gene signature. In addition, Multivariate Cox regression analyses indicated that the risk score was an independent prognostic factor for HCC. Functional analysis further revealed that immune-related pathways were enriched, and immune status in HCC were different between the two risk groups.

Conclusion: In conclusion, a novel GT genes signature can be used for prognostic prediction in HCC. Thus, targeting GT genes may be a therapeutic alternative for HCC.

Introduction

Liver cancer ranks the sixth most common cancer with approximately 800,000 new cases and 780,000 deaths per year[1]. Hepatocellular carcinoma (HCC) is the world's second leading cause of cancer mortality worldwide, always occurring in the background of chronic liver diseases[2, 3]. It is reported that the development of HCC was correlated with several risk factors, mainly including nonalcoholic fatty liver disease, chronic viral infection, and alcohol abuse[1, 4, 5]. Despite the great development in the diagnosis and treatment of HCC, the majority of patients with HCC are diagnosed at an advanced stage, leading to a poor prognosis[6]. Thus, it is necessary to explore novel prognostic factors for HCC patients.

Glycosylation is the common process of modifications mainly comprised of N- and O-glycosylation that exhibit an important role in various physiological and pathological processes, which requires for the
participation of glycosyltransferases (GTs)\([7-10]\). Over the past decades, a growing number of studies have revealed that the GTs involved in glycosylation are essential in the progression of various of disease such as joint diseases, inflammatory diseases, cancers and liver diseases\([11-13]\). Differentially expressed GTs and their corresponding target proteins have been demonstrated to be acted as tumor biomarkers and therapeutic targets in specific cancer\([14, 15]\). An additional evidence revealed a novel role of GLANT3 and B3GT3 in the self-renewal of pancreatic cancer stem cells\([16]\). On the other hand, Taniguchi et al. discovered that N-glycan is directly associated with cancers, which provided new biomarkers for the progression, metastasis and therapeutics of cancer\([17]\). However, whether these GT genes are involved in the prognosis of HCC patients remains largely not elucidated.

In our study, we comprehensively explored the expression profiles of GT genes using the RNA sequencing (RNA-seq) data from The Cancer Genome Atlas (TCGA) database, and identified 4 gene signatures for the prognosis of HCC. Then, the prognostic value of GT genes was evaluated and validated in the training set and validation set of TCGA database of HCC. Intriguingly, these gene signatures can effectively predict the prognosis of HCC patients. Additionally, we also performed the Gene Set Enrichment Analysis (GSEA) to study the potential mechanisms.

**Materials And Methods**

**Data acquisition**

The RNA sequencing (RNA-seq) data and paired clinical information comprising of 371 primary HCC patients samples and 50 normal samples were downloaded from The Cancer Genome Atlas (TCGA) (Supplementary Table.1). The principal component analysis (PCA) was performed to reduce the dimensionality in the TCGA database. Subsequently, the patients with HCC were randomly assigned to the training set (normal samples: n=35, tumor tissues: n=260) and validation set (normal samples: n=15, tumor tissues: n=111) according to the ratio of 7:3. The detailed clinical characteristics of HCC patients between the training set and validation set of TCGA database was listed in Table.1. The 210 human glycosyltransferase (GT) genes were obtained from literature research (Supplementary Table.2)\([18]\). The workflow in the present study of bioinformatics analysis from both 210 GT genes and TCGA database was presented in Fig.1.

**Construction and validation of a prognostic GT gene signature**

Differentially expressed genes (DEGs) were screened between tumor group and normal group with the threshold of $|\log_2 FC| \geq 1$ and $P$-value $\leq 0.05$ by using the "limma" R package. Univariate Cox analyses were applied for the identification of GT genes related to the overall survival (OS) of HCC patients ($P < 0.05$). 4 gene signatures were identified with the LASSO Cox regression algorithm using "glmnet" package in the R, and used to construct a prognostic model. Then, the risk scores of HCC patients were calculated according to the formula of score$= \frac{e^{\sum (\text{each gene's expression levels} \times \text{corresponding coefficient})}}{e^{\sum (\text{each gene's mean expression levels} \times \text{corresponding coefficient})}}$. Based in the median value of risk score, the patient with HCC
were classified by two groups, including high-risk and low-risk groups. PCA was performed to reveal the distribution of samples in the two risk groups. In addition, ROC curve of time-dependent was conducted using the "survivalROC" R package.

**Functional enrichment analysis**

Gene Set Enrichment Analysis (GSEA) was utilized to performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses based on the DEGs identified with the cut-off of $|\log_2 FC| \geq 1$, $P$-value < 0.05 between the high-risk and low-risk groups. The enrichment scores of immune infiltration cells were calculated by single-sample gene set enrichment analysis (ssGSEA), which was performed in the "gsva" package of R. Immune score, ESTIMATE score and stromal score was generated by ESTIMATE algorithm according to the ssGSEA.

**Statistical analysis**

All bioinformatics analyses of our study were performed with R software. The student’s t-test was performed to compare the tumor tissues and normal tissues. The correlation between risk score and clinical characteristics was compared by Chi-squared test. Univariate, multivariate Cox, and LASSO regression analyses in the current study were used to identify prognostic factors and independent predictors for the OS of HCC patients. We compared the overall survival (OS) between the different groups using Kaplan-Meier analysis with the log-rank test. A $P$-value < 0.05 was considered statistically as significant, and all $P$ values were two-tailed.

**Results**

**Identification of co-differentially expressed GT genes in HCC**

A PCA of 421 samples in the TCGA database was performed, and the results revealed that there was an obvious difference between tumor tissues and normal tissues samples, which was applied for the subsequent analyses (Fig.2A). To further explore which GT genes may act as prognostic factors in the progression of HCC, we first comprehensively analyzed the training set in the TCGA database of HCC. As presented in Fig.2B, a total of 2264 DEGs were screened based on the threshold of $|\log_2 FC| \geq 1$ and $P$-value $\leq 0.05$, of which 1765 were up-regulated and 499 were down-regulated compared with the tumor group (Supplementary Table.3). The heatmap showed the expressions of DEGs between tumor group and normal group (Fig.2C). Importantly, the expressions of the top10 DEGs were statistically significant differences between the two groups (Fig.2D, $P < 0.05$). Furthermore, these DEGs were combined with 210 GTs downloaded from literature, and 28 co-differentially expressed GT genes were obtained in the TCGA database of HCC (Fig.2E).

**Identification of gene signatures associated with OS in HCC**
To further identify the gene signatures in HCC prognosis, we undertook a univariate Cox regression analysis on the 28 co-differentially expressed GT genes. As a result, 11 of them were obtained with the cut-off of \( P \)-value < 0.05 (Fig.3A). After the least absolute shrinkage and selection operator (LASSO) Cox regression algorithm in the TCGA dataset, we obtained 4 gene signatures, such as ALG3, B3GAT3, GLA, ST6GALNAC4 (Fig.3B-C). To better demonstrate the prognostic role of these 4 gene signatures in HCC, we firstly used K-M analysis in the TCGA dataset in which samples were divided into high and low expression based on the gene expression level (Supplementary Figure.1). All 4 of these genes were significantly associated with the OS of HCC patients, which indicated that these gene signatures may be acted as prognostic factors in HCC. Intriguingly, PPI network revealed that there was no direct interactions between 4 gene signatures (Fig.3D).

**Prognostic value of gene signatures in the training dataset of TCGA database**

Having shown that these gene signatures may act as prognostic factors in the progression of HCC, we attempted to investigate their prognostic value in HCC. Therefore, we constructed the prognostic model of gene signatures in the training dataset of TCGA database. The risk score was calculated according to the gene expression levels and their corresponding coefficients. Based on the median value of risk score, the patients with HCC were divided into high- and low-risk groups (Fig.4A). The results showed that the number of patients died of HCC was significantly increased with the increasing risk score. The OS of HCC patients had an obvious difference between the two risk groups (Fig.4B, \( P=0.00095 \)). Thereafter, ROC curve was performed to investigate the effectiveness of the prognostic model (Fig.4C). The area under the curve (AUC) reached 0.676 at 3 years, 0.631 at 5 years, and 0.539 at 7 years, which indicated that the risk score of prognostic model had high accuracy. A nomogram of 3-year and 5-year survival probability was displayed in Fig.4D. In addition, the calibration plots of this risk prognostic model were further constructed (Fig.4E-F, \( c\text{-index}=0.66328 \), calibrated \( c\text{-index}=0.64057 \)

**Validation of gene signatures in the validation set of TCGA database**

To further seek additional evidence that the gene signatures are prognostic factors, we generated a prognostic model in the validation dataset of TCGA database. Interestingly, survival analyses of 4 gene signatures confirmed that two of these genes were correlated with poor OS of HCC (Supplementary Figure.2). To demonstrate the robustness of the validation model from the TCGA database, the patients with HCC were divided into high- and low-risk groups according to the median value of risk score (Fig.5A). Consistent with the results obtained from the training dataset of TCGA database (Fig.5B, \( P=0.0039 \)). Besides, the ROC curve showed that the AUC of 4 gene signatures reached 0.708 at 3 years, 0.712 at 5 years, and 0.594 at 7 years, further revealing the accuracy of risk score in the validation prognostic model (Fig.5C).

**The correlation of the risk scores and clinicopathological characteristics in patients with HCC**

The expressions of the 4 gene signatures between the two risk groups and clinical characteristics in the TCGA database were visualized by heatmap (Fig.6A-B). We found that the expressions of these gene
signatures were significantly increased as the increasing risk score in the TCGA database. Furthermore, we performed univariate and multivariate Cox regression analyses after adding clinical characteristics to investigate whether the risk score of prognostic model was an independent factor for the HCC prognosis (Fig.6C-D). In the univariate Cox regression analyses, the risk score was correlated with the OS of HCC both in the training set and validation set of TCGA database (TCGA training set: HR = 2, 95% CI = 1.5-2.8, P=2.2e-05; TCGA validation set: HR = 2, 95% CI = 1.4-2.8, P= 0.0034), respectively. The multivariate Cox regression analysis showed that the risk score of prognostic model was still proved to be an independent factor for the OS of HCC patients (TCGA training set: HR =2.2, 95% CI = 1.5-3.2, P=5.4e-05; TCGA validation set: HR=1.7, 95% CI = 1.1-2.7, P = 0.018).

Functional analyses in the TCGA database of HCC

Based on these observations, we next attempted to elucidate the biological progresses related to the risk score. Therefore, we focused on the DEGs between the two risk groups to perform GSEA. As a result, GO analysis of the DEGs were involved in several immune-related biological processes, such as GO:0002263 cell activation involved in immune response, GO: 0002275 myeloid cell activation involved in immune response, GO:0002283 neutrophil activation involved in immune response and GO:0002520 immune system development, which may be highly associated with the development of HCC (Fig.7A, P-value < 0.05). All significantly enriched GO terms were shown in Supplementary Table.4.

In order to investigate whether the risk score was correlated with the immune status in HCC, we analyzed the differences between the high-risk and low-risk groups by quantifying the scores of immune cells enrichment. Interestingly, activated CD4 T cell, activated dendritic cell, CD56dim natural killer cell, central memory CD8 T cell, Eosinophil, MDSC, Natural killer T cell showed obvious significantly differences between the two risk groups (Fig.7B, *P < 0.05, **P < 0.01, ***P < 0.001, ****, P< 0.0001). Moreover, the immune score calculated by ESTIMATE algorithm in high risk group was higher than those of low risk group, which further revealed that the risk score of prognostic model was strongly correlated with the immune status of HCC patients (Fig.7C, **P < 0.01). Interestingly, the stromal score and ESTIMATE score showed no significant difference between high-risk and low-risk groups (Fig.7D-E).

Discussion

HCC is the most common malignancy worldwide with a poor prognosis, which always occurred at an advanced stage[19, 20]. The critical role played by glycosylation modifications has been demonstrated in several cell biology processes occurring in cancer, such as immune modulation and metastasis, tumor cell invasion, cell signaling and communication, and tumor angiogenesis[10, 21, 22]. Accordingly, this suggests that GTs participated in glycosylation modification process have a potential application in the diagnosis and prognosis of HCC.

Over the past years, various biomarkers has been identified in many cancers with the great progress in epigenetics and metabolomics[23-25]. In the current study, we focused on whether GTs involved in the progression of HCC as well as being associated with the prognosis of HCC. By analyzing the expression
profiles of HCC patients in the TCGA database, we identified gene signature that could predict the OS of HCC patients. Four differentially expressed GT genes correlated with the prognosis of HCC were screened out from 210 GT genes, including ALG3, B3GAT3, GLA, and ST6GALNAC4. We constructed 4 gene signatures that divided HCC patients into high- and low-risk groups for the prediction of HCC patients prognosis. Furthermore, the prognostic value of the 4 gene signatures was assessed by ROC curve and K-M analysis in the training set and validation set, respectively. Most importantly, the risk score generated from 4 gene signatures were demonstrated to be an independent prognostic factors in the patients with HCC by univariate and multivariate Cox regression analyses. Additionally, GSEA result revealed the differentially expressed genes (DEGs) between the high- and low-risk groups was involved in several immune-related biological processes and KEGG pathways. The immune status between the two groups was further evaluated.

Although previous study has demonstrated that the expressions of GT genes might regulated aberrant glycosylation in many cancers[18], their correlation with the OS of HCC patients still remains unknown. Surprisingly, 28 GT genes were differentially expressed between tumor tissues and normal tissues, and 4 of them were demonstrated to be acted as gene signatures, which indicated that the potential prognostic value of GT genes in HCC.

Of 4 GT genes, several genes have been investigated to be involved in the progression of tumors. ALG3 is an oncoprotein correlated with multiple maligancies, which is located at 3q27.1 of the chromosomal region[26]. In non-small cell lung cancer (NSCLC), the expression of ALG3 in tumor tissues was significantly increased compared with that of normal tissues[27]. Up-regulation of ALG3 promoted the metastasis of lymph node in esophageal squamous cell cabcer and the proliferation of cervical cancer cells[26, 28]. ALG3 was found to promote the proliferation and metastasis of breast cancer cells, which indicated that over-expression of ALG3 had poor prognosis[29]. In accordance with previous studies, our results revealed that ALG3 was up-regulated in HCC tissues. To our knowledge, we first discovered the potential prognostic role of ALG3 for the patients with HCC.

It has been reported altered genes expression of ST6GALNAC family in several cancers. ST6GALNAC4, as a member of ST6GALNAC family, promoted the metastasis of lung cancer via adhesion to galectins. Aberrant glycosylation caused by the changes of ST6GALNAC4 expression levels promoted the lung cancer metastasis through adhesion to galectins in the metastatic niche[30]. MiR-4299 mediated the invasive properties and tumorigenicity of human follicular thyroid carcinoma via targeting ST6GALNAC4[31]. B3GAT3, as a glycosyltransferase, is attached to the serine residues of key proteins[32, 33]. Moreover, the research conformed that B3GAT3 was highly expressed in liver cancer tissues, and had poor prognostic value, which was consistent with our present study[34]. With respect to GLA, there was no reports on the development of cancers and related mechanisms. In our study, we evaluated the prognostic role of GT genes and correlations between the GT genes and clinical characteristics in HCC. Besides, we also explored potential molecular mechanism in HCC, providing new insights for the treatment.
For a comprehensive analysis, we first demonstrated GT genes signatures in the training set and validation set of TCGA database, respectively, and observed their prognostic value for predicting the OS of HCC patients. However, there was obvious difference on the OS of each gene, and this may be due to the deficiency and clinical information of samples, leading to the conflicting or meaningless results. In addition, functional analysis revealed that these DEGs between the high-risk and low-risk groups mainly involved in several important immune biological processes and pathways, which may provided a new direction for the mechanism of HCC progression. It is unclear how GT genes involve in HCC prognosis, and this requires for subsequent study.

**Conclusion**

In conclusion, we first revealed the prognostic value of GT genes in the progression of HCC. A GT genes signature was constructed which could effectively divide HCC patients into high- and low-risk in order to accurately predict their OS. Our study opened a new understanding of the molecular mechanisms of GT genes involved in the occurrence and development of HCC, and provides a unique approach to explore the biomarkers for targeted therapy in HCC.

**Declarations**

**Acknowledgments**

We thank all the participants who were involved in this study, and gratefully acknowledge contributions from the TCGA Network, and www.yiqishiyan.com.

**Authors’ contributions**

Qiang Cai and Shizhe Yu designed the research. Jian Zhao, Duo Ma, Long Jiang, and Xinyi Zhang contributed to the collection of data and analysis, and were involved in the writing of this manuscript. Qiang Cai and Shizhe Yu contributed to the conducted the experiments. Zhiyong Yu performed the correction of the language and revision. All authors proofread and approved the final manuscript.

**Funding**

This research was financially supported by the Academician (Expert) Workstation Construction Project of Yunnan Province (2018IC107).

**Availability of data and materials**

The data used to support the findings of this study is included in the article, and the datasets analysed during the current study are available in The Cancer Genome Atlas database (https://portal.gdc.cancer.gov/projects/TCGA-LIHC).

**Ethics approval and consent to participate**
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1 Department of Hepatobiliary and pancreatic Surgery, Affiliated Hospital of Yunnan University, Kunming, China, 650021
2 Department Surgery, First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, 400052

References
1. Freddie, et al., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. Ca A Cancer Journal for Clinicians, 2018.
2. Degasperi, E. and M. Colombo, Distinctive features of hepatocellular carcinoma in non-alcoholic fatty liver disease. Lancet Gastroenterology & Hepatology, 2016. 1(2): p. 156-164.
3. Sim, H.W. and J. Knox, Hepatocellular carcinoma in the era of immunotherapy. Curr Probl Cancer, 2017: p. S0147027217301101.
4. Llovet, J.M., et al., Hepatocellular carcinoma. Nature Reviews Disease Primers, 2016. 2: p. 16018.
5. Chen and Chien-Jen, Risk of Hepatocellular Carcinoma Across a Biological Gradient of Serum Hepatitis B Virus DNA Level. Jama, 2006. 295(1): p. 65-73.
6. Sarcopenia, intramuscular fat deposition, and visceral adiposity independently predict the outcomes of hepatocellular carcinoma. Journal of Hepatology, 2015. 63(1): p. 131-140.
7. Christina, S. and M. Paul, Emerging facets of prokaryotic glycosylation. Fems Microbiology Reviews, (1): p. fuw036.
8. Wadzinski, T.J., et al., Rapid phenolic O-glycosylation of small molecules and complex unprotected peptides in aqueous solvent. Nature Chemistry, 2018. 10(6).
9. Yang, Y., X. Zhang, and B. Yu, O-Glycosylation methods in the total synthesis of complex natural glycosides. Cheminform, 2015. 32(9): p. 1331-1355.
10. Pinho, S.S. and C.A. Reis, Glycosylation in cancer: mechanisms and clinical implications. Nature Reviews Cancer, 2015.
11. Blomme, B., et al., Alteration of protein glycosylation in liver diseases. Journal of Hepatology, 2009. 50(3): p. 592-603.
12. Morrow, A., L. and L. Denson, A., Use of glycans and glycosyltransferases for diagnosing/monitoring inflammatory bowel disease.

13. Dong, S., et al., Bioinformatics insight into glycosyltransferase gene expression in gastric cancer: POFUT1 is a potential biomarker. Biochemical & Biophysical Research Communications, 2016. 483(1): p. 171.

14. Meany, D.L. and D.W. Cha, Aberrant glycosylation associated with enzymes as cancer biomarkers. Clinical Proteomics, 2011. 8(1).

15. Drake, R.R., Glycosylation and Cancer: Moving Glycomics to the Forefront. Advances in Cancer Research, 2015. 126: p. 1-10.

16. Barkeer, S., et al., Novel role of O-glycosyltransferases GALNT3 and B3GNT3 in the self-renewal of pancreatic cancer stem cells. Bmc Cancer, 2018. 18(1).

17. Taniguchi, N. and Y. Kizuka, Glycans and cancer: role of N-glycans in cancer biomarker, progression and metastasis, and therapeutics. Advances in Cancer Research, 2015. 126: p. 11.

18. Ashkani, J. and K.J. Naidoo, Glycosyltransferase Gene Expression Profiles Classify Cancer Types and Propose Prognostic Subtypes. entific Reports, 2016. 6: p. 26451.

19. Marquardt, J.U. and S.S. Thorgeirsson, SnapShot: Hepatocellular carcinoma. 2014.

20. Llovet, et al., Prognosis of Hepatocellular Carcinoma: The BCLC Staging Classification. Seminars in Liver Disease, 1999.

21. Bhatia, R., et al., Cancer-associated mucins: role in immune modulation and metastasis. Cancer and metastasis reviews, 2019.

22. Chen, Y., L. Ding, and H. Ju, In Situ Cellular Glycan Analysis. Accounts of Chemical Research, 2018: p. acs.accounts.7b00617.

23. Mou, Y., et al., The Landscape of Iron Metabolism-Related and Methylated Genes in the Prognosis Prediction of Clear Cell Renal Cell Carcinoma. Frontiers in Oncology, 2020. 10: p. 788.

24. Liang, J.Y., et al., A Novel Ferroptosis-related Gene Signature for Overall Survival Prediction in Patients with Hepatocellular Carcinoma. International journal of biological ences, 2020. 16(13): p. 2430-2441.

25. Wang, P., et al., Identification of RNA: 5-Methylcytosine Methyltransferases-Related Signature for Predicting Prognosis in Glioma. Frontiers in Oncology, 2020. 10: p. 1119.

26. Shi, Z.-Z., et al., Identification of putative target genes for amplification within 11q13.2 and 3q27.1 in esophageal squamous cell carcinoma. Clinical\s&\stranslational Oncology, 2013.

27. Ke, S.B., et al., ALG3 Contributes to the Malignancy of Non-small Cell Lung Cancer and Is Negatively Regulated by MiR-98-5p. Pathology - Research and Practice, 2019. 216(3): p. 152761.

28. Gene expression profiles in squamous cell cervical carcinoma using array-based comparative genomic hybridization analysis. International Journal of Gynecological Cancer, 2010. 17.

29. De Leoz, M.L.A., et al., High-mannose glycans are elevated during breast cancer progression. Molecular & Cellular Proteomics, 2011. 10(1): p. M110.002717.
30. Reticker-Flynn, N.E. and S.N. Bhatia, Aberrant Glycosylation Promotes Lung Cancer Metastasis through Adhesion to Galectins in the Metastatic Niche. Cancer Discovery, 2015. 5(2): p. 168.

31. Miao, X., et al., miR-4299 mediates the invasive properties and tumorigenicity of human follicular thyroid carcinoma by targeting ST6GALNAC4. Iubmb Life, 2016. 68(2): p. 136-144.

32. Job, F., et al., Functional validation of novel compound heterozygous variants in B3GAT3 resulting in severe osteopenia and fractures: expanding the disease phenotype. Bmc Medical Genetics, 2016. 17(1).

33. Pedersen, L.C., et al., Heparan/chondroitin sulfate biosynthesis. Structure and mechanism of human glucuronyltransferase I. Journal of Biological Chemistry, 2000. 275(44): p. 34580.

34. Zhang, Y.L., C. Ding, and L. Sun, High Expression B3GAT3 Is Related with Poor Prognosis of Liver Cancer. Open medicine (Warsaw, Poland). 14(1): p. 251.

Tables

Due to technical limitations, table 1 is only available as a download in the Supplemental Files section.

Supplementary Information

Additional file1. Figure.S1 Survival analyses grouped by the optimal cut-off expression value of each gene signature in the training dataset of TCGA database.

Figure.S2 Survival analyses grouped by the optimal cut-off expression value of each gene signature in the validation dataset of TCGA database.

Table S1 The clinical information comprising of 371 primary HCC patients samples and 50 normal samples were downloaded from The Cancer Genome Atlas (TCGA) Table S2 The clinical information and 210 human glycosyltransferase (GT) genes were obtained.

Table S3 DEGs of 371 samples in the TCGA database.

Table S4 All significantly enriched GO terms.