Supporting Information

Highly efficient enrichment of O-GalNAc glycopeptides by using immobilized metal ion affinity chromatography

Xuyang Yue[1, 2, 3, 4], Hongqiang Qin*[2, 4], Yao Chen[2,3], Zheng Fang[2,3], Luyao Liu[2,3], He Zhu[2,3], Xiaoyan Liu[2], Jiahua Zhou[2,3], Kailu Tian[1, 2, 3], Xiaoqiang Qiao*[1] and Mingliang Ye* [2]

1 College of Pharmacy, Hebei Province Key Laboratory of Analytical Science & Technology, MOE Key Laboratory of Medicinal Chemistry & Molecular Diagnosis, Hebei University, Baoding 071002, China

2 CAS Key Laboratory of Separation Science for Analytical Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Science, Dalian, 116023, China

3 University of Chinese Academy of Sciences, Beijing 100049, China

4 These authors contributed equally to this paper.

Corresponding authors: qinhq@dicp.ac.cn;
xiaoqiao@hbu.edu.cn;
mingliang@dicp.ac.cn ;

Table of Contents

Supplementary methods and data processing

Figure S1-8.

Other supplementary materials not included in this file: Supplementary Information Table S1-S7 as Excel files.

Supplementary methods and data processing

Reagents and materials

Trypsin was purchased from Sigma (St. Louis, MO). PNGase F and GlycoBuffer 2 were purchased from New England Biolabs (Ipswich, MA, USA). Chemical reagents of iodacetamide (IAA), 1,4-dithiothreitol (DTT), and trifluoroacetic acid (TFA) were obtained from Sigma (St. Louis, MO). Formic acid (FA) was obtained from Fluka (Buches, Germany). Ammonium bicarbonate (NH_4HCO_3) and urea were obtained from Bio Basic Inc. (Ontario, Canada). Acetonitrile (ACN, HPLC grade) and ammonium hydroxide (25 wt % solution in water) were purchased from Merck (Darmstadt, Germany). Pure water used in all experiments was purified with a Milli-Q system (Millipore, Milford, MA). The centrifugal filter units (Amicon Ultra-0.5 mL) were purchased from Millipore (Milford, MA). Other chemicals used were either of analytical grade or of better grade. C18 AQ beads (1.9 μm) reversed phase chromatographic packing was purchased from Dr. Maisch GmbH. The control serum and hepatocellular carcinoma (HCC) serum samples were collected from 10 volunteers and HCC patients pooled together with equal-volume, respectively. The samples were stored at -80 ℃ until usage. The human serum for the experiment was obtained from the Second Affiliated Hospital of Dalian Medical University (Dalian, China). All the experiments were approved by the Ethics Committees of the Second Affiliated Hospital of Dalian Medical University according to the moral standards of the Declaration of Helsinki. Click maltose-HILIC beads[1] and Ti-IMAC(IV) beads[2] were prepared in house as reported.

Digestion of human serum proteins

Human serum or cancer patient serum (50 μL) was mixed with 450 μL of 8 M urea/100 mM NH_4HCO_3. The resultant solution was treated with 20 mM DTT for 2 h at 37 ℃. Then, 40 mM IAA was added, and the mixture was incubated in the dark for 30 min at room temperature. Trypsin was added at the ratio of

1:25 (enzyme: protein, w/w, 37 ℃ overnight) into the reaction buffer after the concentration of urea was diluted to <2 M. Then the digest peptides were desalted by Oasis HLB SPE column (Waters, 30 mg). After lyophilization, the peptides were re-dissolved in Glycobuffer 2 (pH=7.5), and 3000 units of PNGase F was added to release N-linked glycans at 37 ℃ overnight. Ultimately, the protein digest was desalted by Oasis HLB SPE column and lyophilized for enrichment of glycopeptides.

Enrichment of O-GalNAc glycopeptides using HILIC method

The O-GalNAc glycopeptides were enriched by using the centrifugation assisted click maltose-HILIC approach according to our previous report.[3] Firstly, 5 mg click maltose material were washed three times with 80% ACN/1% TFA, and the digested proteins were re-dissolved in the loading buffer (80% ACN/1% TFA). The digests were mixed with the HILIC materials and the resulting suspension was shaken for 45 minutes. Then, the mixture was pipetted into a HILIC tip and washed with 40 μL of loading buffer twice. Finally, the enriched O-glycopeptides were eluted with 100 μL of 30% ACN/1% FA twice. After that the elution was combined and lyophilized for MS analysis.

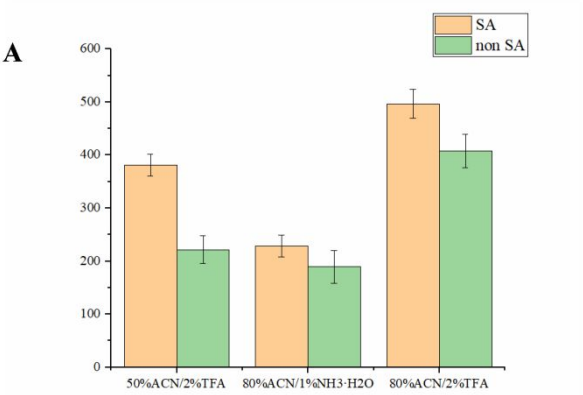
Fractionation of enriched O-GalNAc glycopeptides

In order to further in-depth analysis of the O-GalNAc glycosylation in human serum samples, fractionation was performed after the glycopeptide enrichment. The O-GalNAc glycopeptides enriched from serum digest equivalent to 10.0 μL initial serum (7.2 μL initial serum for MS detection) was fractioned by using a column (150 mm × 2.1 mm i.d.) packed with C18 particles 5 μm, 150 Å, Agela). Mobile phases A (5 mM NH_4HCO_3/H_2O, adjusted pH to 9.5 using $NH_3 \cdot H_2O$) and B (80% ACN/5 mM NH_4HCO_3, adjusted pH to 9.5 using $NH_3 \cdot H_2O$) were used to develop a gradient. The solvent gradient was set as follows: 100-95% A, 1 min; 95-80% A, 7 min; 80-50% A, 16.5 min; 50-10% A, 0.5 min; 10% A, 8 min. The column was then equilibrated by phase A for 7 min. Fractions were collected in every 1 minute from 6 to 28 min, and the 22 collected samples were combined to 11 fractions.[4] The samples were dried

under vacuum for subsequent MS analysis.

Figure S1

A



B

Buffer	SA	non SA
50% ACN/2% TFA	-23.29%	-45.58%
80% ACN/1% $NH_3 \cdot H_2O$	-53.96%	-53.52%

Figure S1 (A) The number of sialylated and non-sialylated glycopeptides identified using three different loading buffers (detail results shown in **Table S3**); (B) The percentage of the sialylated and non-sialylated peptides decreased compared to the result using 80% ACN/2% TFA as the loading buffer.

Figure S2

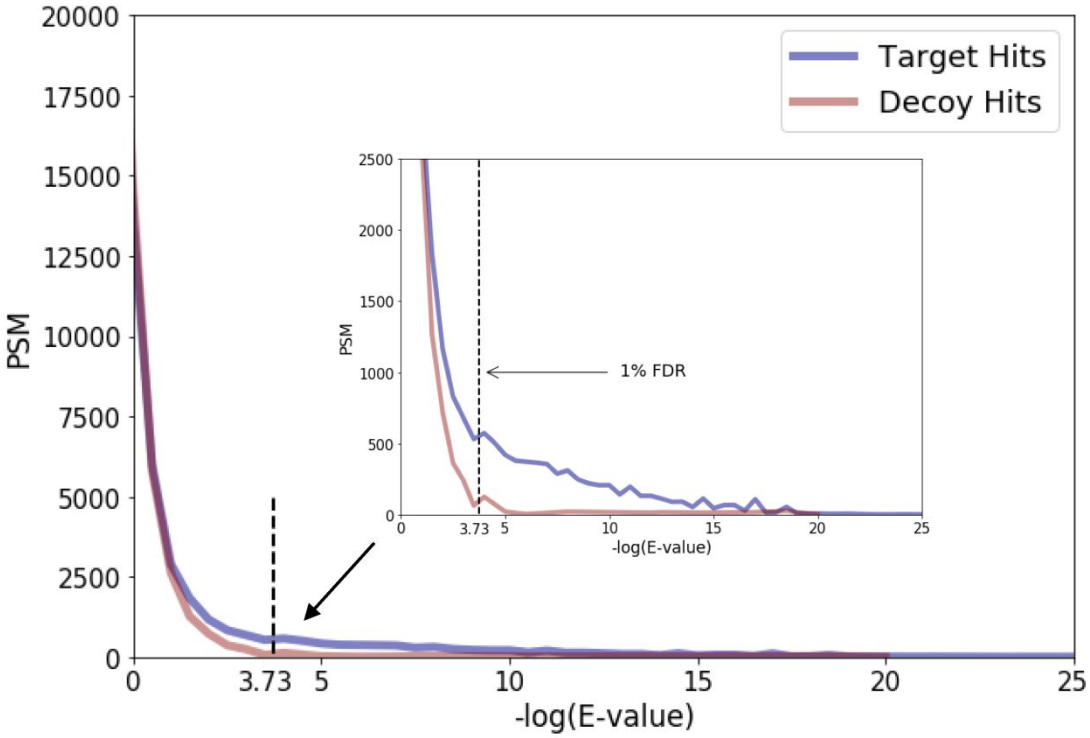


Figure S2 The e-value distribution of the target (blue) and decoy (red) PSMs of O-GalNAc glycopeptides using O-Search.[5]

Figure S3

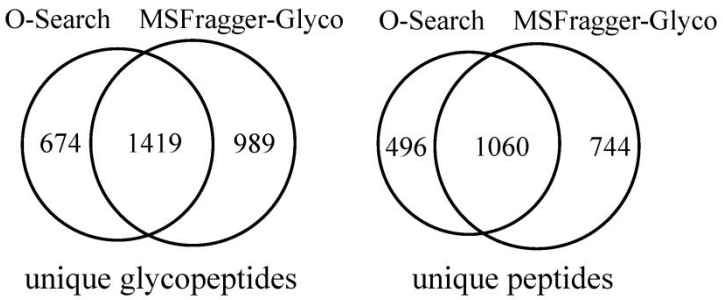


Figure S3 Overlap of O-GalNAc glycopeptides and peptide sequences identified from human serum 2D LC-MS/MS data by using O-Search[5] and MSFragger-Glyco[6], respectively (Detail sequences in **Table S6**).

Figure S4

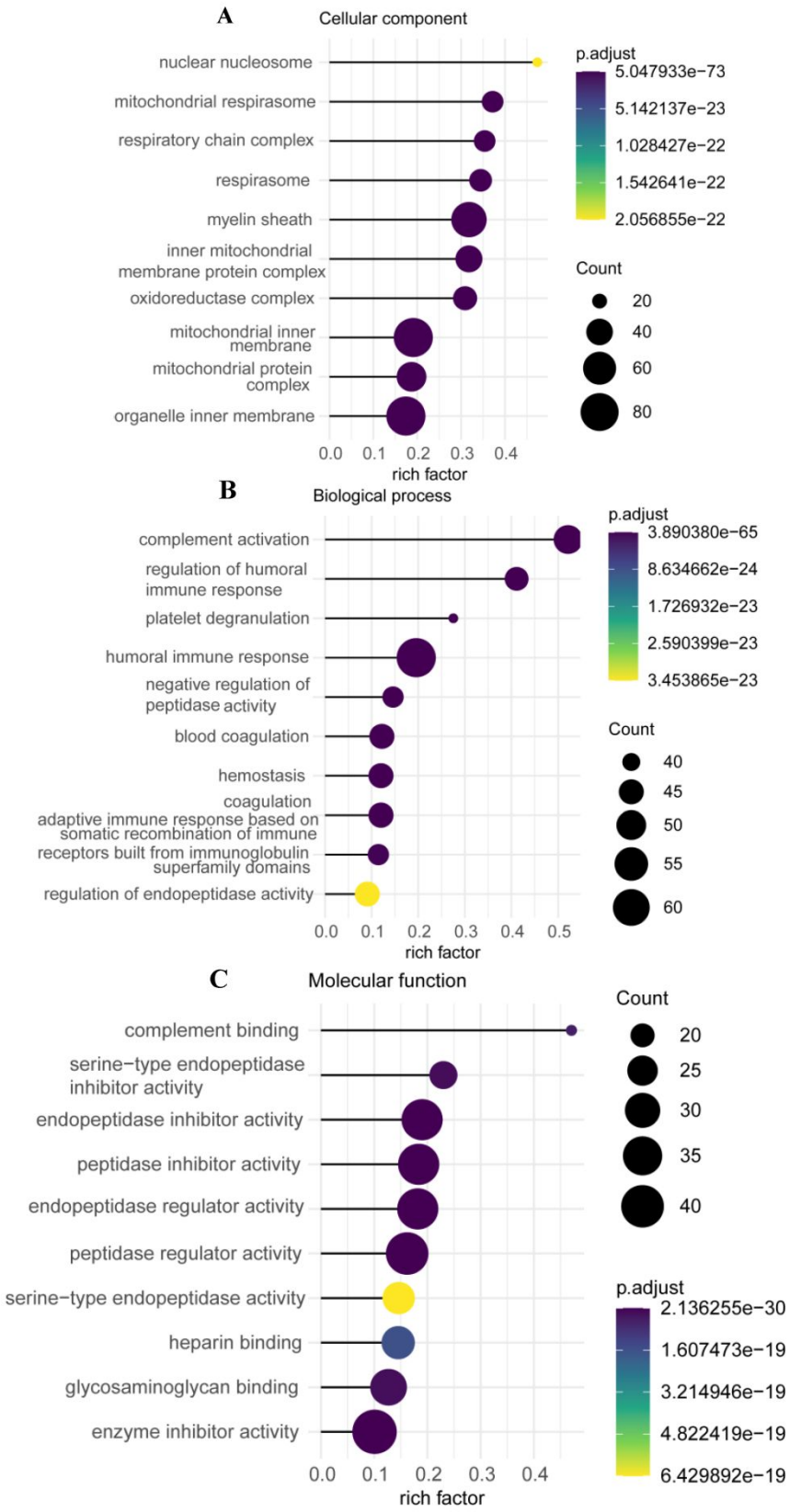


Figure S4 The cellular components (A), biological process (B) and molecular functions (C) of the identified O-GalNAc glycoproteins from 2D LC-MS/MS human serum samples by using O-Search.

Figure S5

Precursor Mass : 4501.0157 Precursor m/z : 1126.2612 (4+)
Peptide Mass : 4135.8826
Delta Mass : 365.1322 Glycan compositions: Hex(1)HexNAc(1)

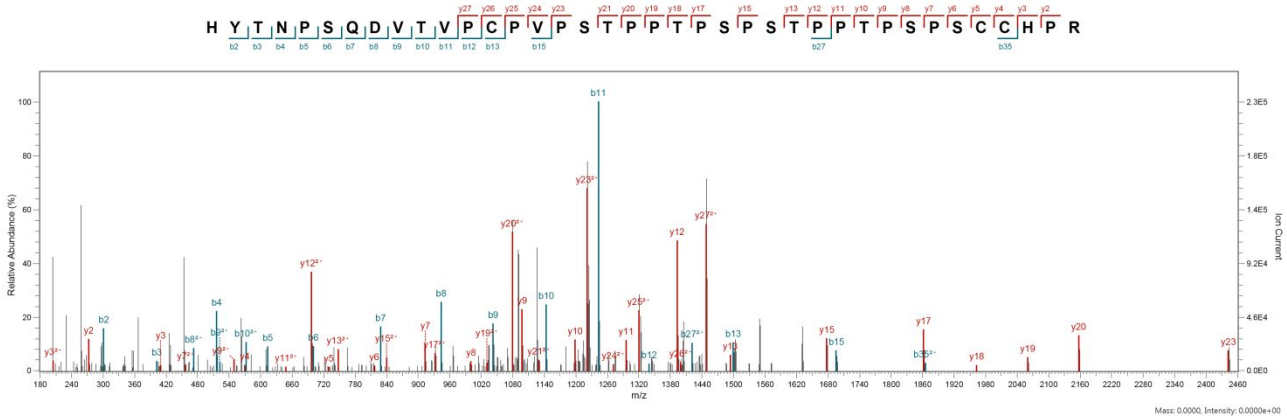


Figure S5 The MS/MS spectra of the O-GalNAc glycopeptide backbone sequence in protein IGHA1.

Figure S6

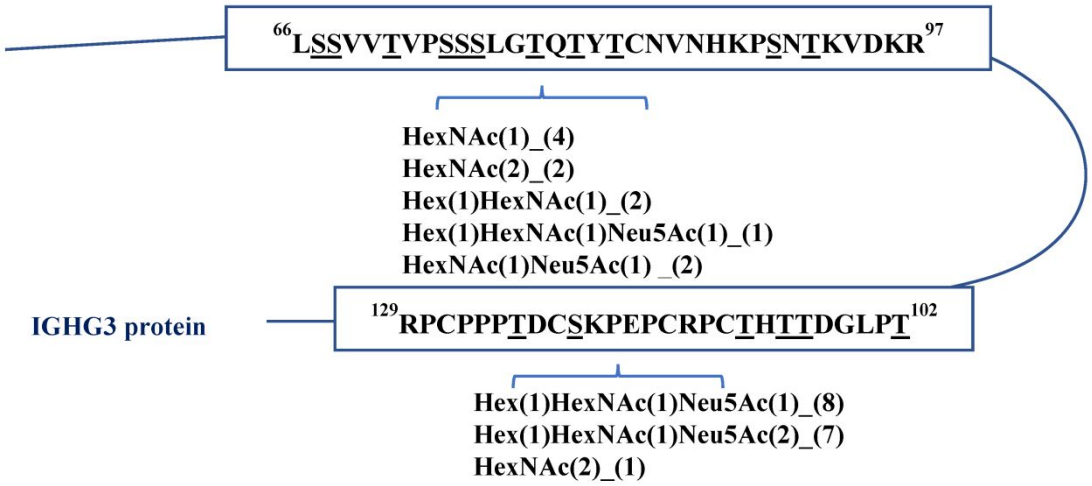
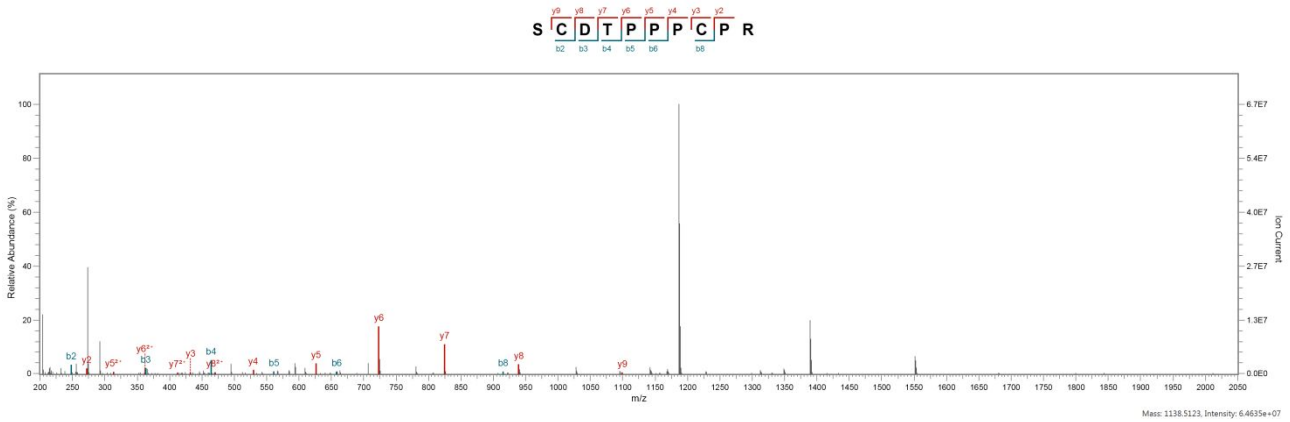


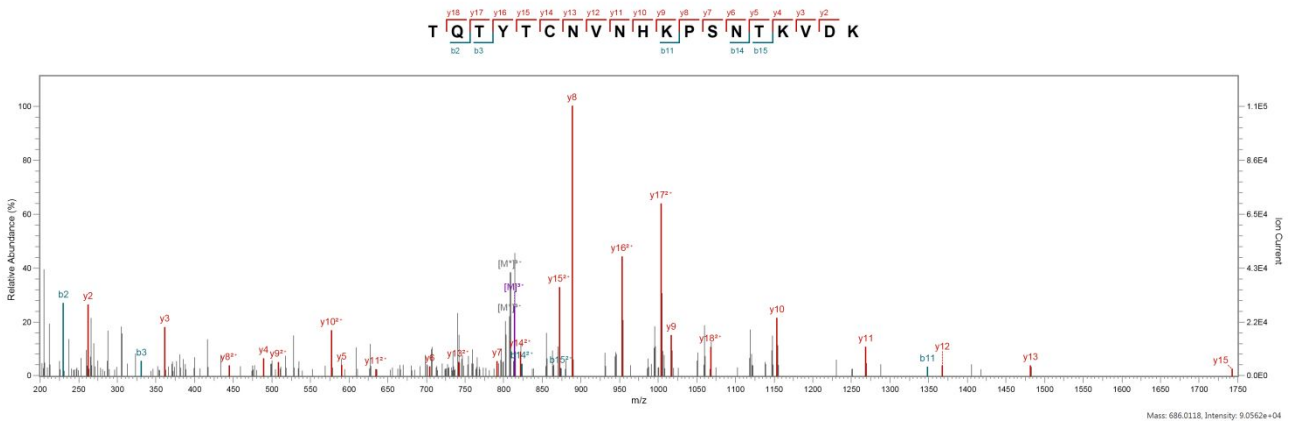
Figure S6 Schemes for the O-GalNAc glycan compositions attached onto the peptide regions of IGHG3 identified from human serum.

Figure S7

Precursor Mass: 2132.8180 Precursor m/z: 1067.4163 (2+)
Peptide Mass: 1185.4907
Delta Mass: 947.3230 Glycan compositions: Hex(1)HexNAc(1)Neu5Ac(2)



Precursor Mass: 2439.1522 Precursor m/z: 814.0580 (3+)
Peptide Mass: 2235.0590
Delta Mass: 203.0794 Glycan compositions: HexNAc(1)



Precursor Mass: 2010.8572 Precursor m/z: 671.2930 (3+)
Peptide Mass: 1354.6299
Delta Mass: 656.2276 Glycan composition: Hex(1)HexNAc(1)Neu5Ac(1)

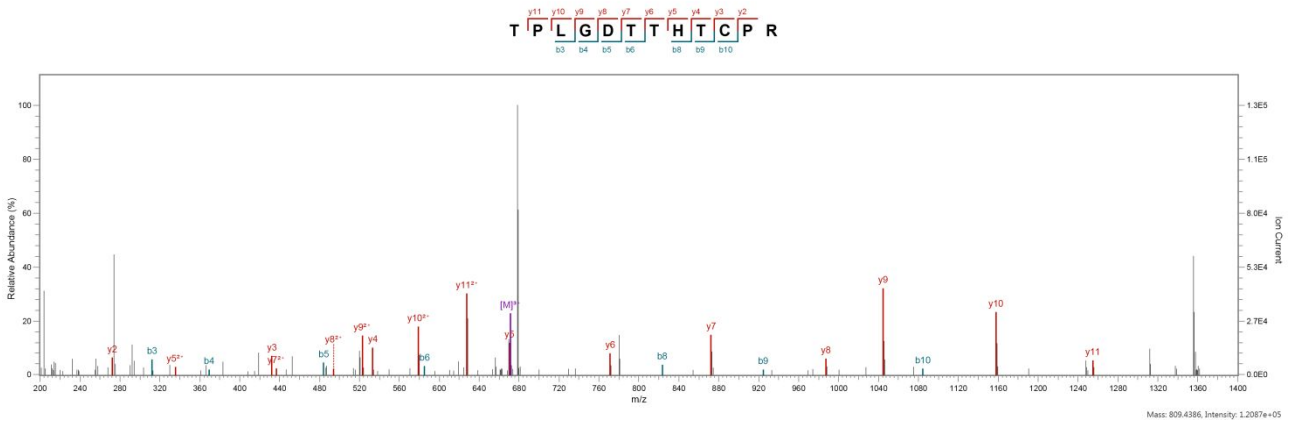


Figure S7 The MS/MS spectra of the peptide sequences of O-GalNAc glycospeptides from analysis of IGHG3 protein.

Figure S8

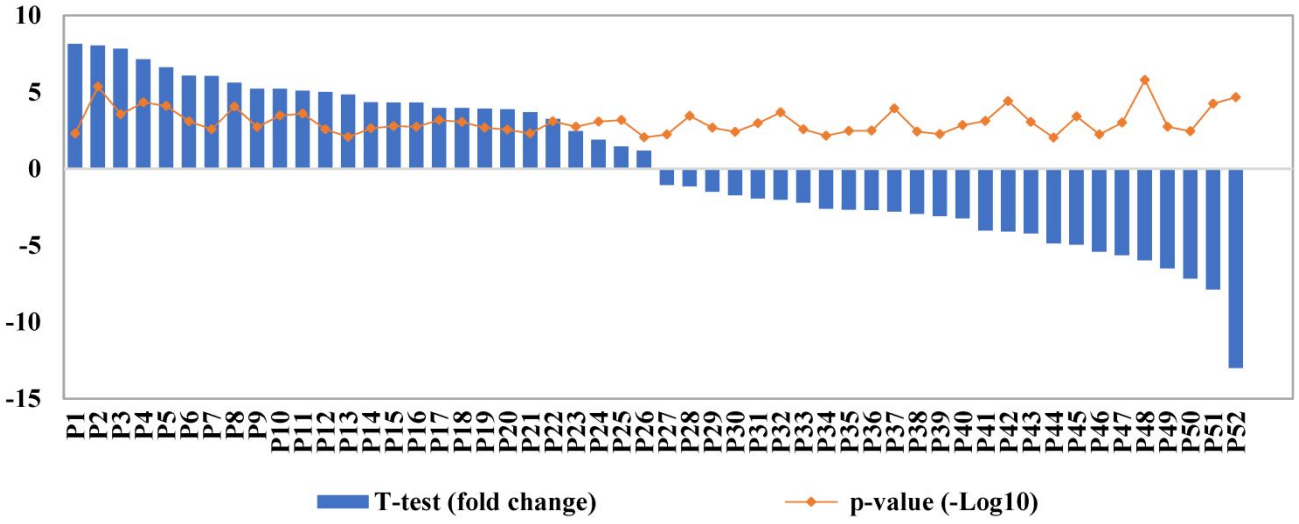


Figure S8 The 52 O-GalNAc glycopeptides with significant change between control and HCC serum samples (Detail sequences shown in **Table S7**).

References：

(1) Yu, L.; Li, X.; Guo, Z.; Zhang, X.; Liang, X. *Chemistry* **2009**, *15*, 12618-12626.

(2) Huang, J.; Dong, J.; Shi, X.; Chen, Z.; Cui, Y.; Liu, X.; Ye, M.; Li, L. *Anal. Chem.* **2019**, *91*, 11589-11597.

(3) Qin, H.; Cheng, K.; Zhu, J.; Mao, J.; Wang, F.; Dong, M.; Chen, R.; Guo, Z.; Liang, X.; Ye, M.; Zou, H. *Anal. Chem.* **2017**, *89*, 1469-1476.

(4) Song, C.; Ye, M.; Han, G.; Jiang, X.; Wang, F.; Yu, Z.; Chen, R.; Zou, H. *Anal. Chem.* **2010**, *82, 53–56*.

(5) Mao, J.; You, X.; Qin, H.; Wang, C.; Wang, L.; Ye, M. *Anal. Chem.* **2019**, *91*, 3852-3859.

(6) Polasky, D. A.; Yu, F.; Teo, G. C.; Nesvizhskii, A. I. *Nat. Methods* **2020**, *17, 1125-1132*.