

N-glycoproteome Analysis of the Secretome of Human Metastatic Hepatocellular Carcinoma Cell Lines Combining Hydrazide Chemistry, HILIC Enrichment and Mass Spectrometry

Xianyu Li^{1,2}, Jing Jiang², Xinyuan Zhao^{1,2}, Jifeng Wang^{1,2}, Huanhuan Han², Yan Zhao², Bo Peng², Rugang Zhong¹, Wantao Ying^{2*}, Xiaohong Qian^{1,2*}

1 The College of Life Science and Bio-engineering, Beijing University of Technology, Beijing, China, **2** State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Radiation Medicine, Beijing, China

Abstract

Cancer cell metastasis is a major cause of cancer death. Unfortunately, the underlying molecular mechanisms remain unknown, which results in the lack of efficient diagnosis, therapy and prevention approaches. Nevertheless, the dysregulation of the cancer cell secretome is known to play key roles in tumor transformation and progression. The majority of proteins in the secretome are secretory proteins and membrane-released proteins, and, mostly, the glycosylated proteins. Until recently, few studies have explored protein N-glycosylation changes in the secretome, although protein glycosylation has received increasing attention in the study of tumor development processes. Here, the N-glycoproteins in the secretome of two human hepatocellular carcinoma (HCC) cell lines with low (MHCC97L) or high (HCCLM3) metastatic potential were investigated with a in-depth characterization of the N-glycosites by combining two general glycopeptide enrichment approaches, hydrazide chemistry and zwitterionic hydrophilic interaction chromatography (zic-HILIC), with mass spectrometry analysis. A total of 1,213 unique N-glycosites from 611 N-glycoproteins were confidently identified. These N-glycoproteins were primarily localized to the extracellular space and plasma membrane, supporting the important role of N-glycosylation in the secretory pathway. Coupling label-free quantification with a hierarchical clustering strategy, we determined the differential regulation of several N-glycoproteins that are related to metastasis, among which AFP, DKK1, FN1, CD151 and TGF β 2 were up-regulated in HCCLM3 cells. The inclusion of the well-known metastasis-related proteins AFP and DKK1 in this list provides solid supports for our study. Further western blotting experiments detecting FN1 and FAT1 confirmed our discovery. The glycoproteome strategy in this study provides an effective means to explore potential cancer biomarkers.

Citation: Li X, Jiang J, Zhao X, Wang J, Han H, et al. (2013) N-glycoproteome Analysis of the Secretome of Human Metastatic Hepatocellular Carcinoma Cell Lines Combining Hydrazide Chemistry, HILIC Enrichment and Mass Spectrometry. PLoS ONE 8(12): e81921. doi:10.1371/journal.pone.0081921

Editor: Joy Marilyn Burchell, King's College London, United Kingdom

Received: July 31, 2013; **Accepted:** October 17, 2013; **Published:** December 4, 2013

Copyright: © 2013 Li et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors are grateful for financial support from the National Key Program for Basic Research of China [2011CB910603, 2013CB911204]; National High Technology Research and Development Program of China [2012AA020203]; International Scientific Cooperation Project of China [2011DFB30370], National Natural Science Foundation of China [21275005, 31100591, 21235001], Beijing Nova Program (Z121107002512014). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: proteomics@126.com (WY); qianxh1@163.com (XQ)

Introduction

Hepatocellular carcinoma (HCC) is a common malignant neoplasm and a major cause of cancer-related deaths in Asian countries. A high mortality rate for HCC is principally caused by uncontrolled tumor invasion and metastasis[1]. Cancer cell metastasis involves intricate, multi-step processes and various cytophysiological changes, including changes in the crosstalk between cells and the components in the extracellular space[2]. Over the past several years, the progress in the analysis of the human plasma proteome has provided a

tremendous opportunity for discovering clinical biomarkers[3]. However, the complicated components of the plasma proteome and the wide dynamic range of concentration of these molecules present challenges for the discovery of new candidates[4].

Tumor cell secretory proteins could be specifically profiled without the depletion of high abundance serum proteins by culturing tumor cells in serum-free conditioned medium for a short duration, collecting the conditioned medium and subjecting it to proteome analysis. Over the past several years, a considerable amount of efforts has been focused on the

analysis of cell secretomes to identify reliable and useful cancer biomarkers. In a representative study, the secretomes of a panel of cancer cell lines were generated, with the detection of 4,600 proteins from 23 cell lines[5]. High quality quantitative analysis of cancer cell secretomes has also been accomplished by combining azidohomoalanine labeling and stable isotope labeling with amino acids in cell culture[6].

Protein glycosylation has been directly linked with cancer development[7]. Almost all of the currently used protein biomarkers are secreted glycoproteins such as carcinoembryonic antigen (CEA), cancer antigen 125 (CA125), prostate specific antigen (PSA) and alpha-fetoprotein (AFP)[8]. Because most of the proteins in the secretory system are glycosylated[9], it is natural to expect that the glycoproteomic analysis of the tumor cell secretome will provide valuable biomarkers. While many studies have been performed for the in-depth profiling of glycoproteins in the plasma, efforts to profile the glycoproteins of secretory proteins from the conditioned medium (CM) are still rather preliminary[10]. Only limited research has been performed to explore the N-glycosylation changes of the secretome that are derived from hepatocellular carcinoma cells, despite great biological and clinical interests[11]. With the recent advances in proteomic technologies, liquid chromatography-mass spectrometry (LC-MS) has become the key tool for analyzing post-translational modifications (PTMs) such as phosphorylation, ubiquitination, acetylation and N-glycosylation[12]. Similar to other PTMs, specific enrichment is essential to capture the often low abundance glycopeptides. Several enrichment methods, including lectin affinity[13,14], hydrophilic interaction[15] and solid phase extraction using hydrazide chemistry[16], have been developed and applied for characterizing N-glycoproteins and N-glycosites[17].

Here, we set out to profile the N-glycoproteins that are secreted by HCC cell lines with low (MHCC97L) or high (HCCLM3) metastatic potential. The two cell lines are derived from the same genetic background[18]. For improved coverage, hydrazide chemistry[19] and zic-HILIC[15] were tested for the enrichment of N-glycopeptides. After duplicate biological analyses, a total of 1,213 unique N-glycosites from 611 N-glycoproteins from both cell lines were confidently identified. A label-free approach was used to quantify the differences in the signal intensity of the MS response between the metastatic cell lines[20,21]. Our study provided clues for the involvement of a few glycoproteins in the metastasis processes, and the differentially regulated proteins may result in the discovery of novel candidates for the measurement of metastasis.

Results

Analytical strategy

Figure 1 is a schematic representation of our experimental approach. To investigate the proteins that are related to liver cancer metastasis, we followed this workflow to profile the difference of N-glycosylated proteins in the secretome of HCC cell lines with different metastatic potential. To reduce the interference from high abundant serum proteins, the culture

medium was not supplemented with serum[10,22,23]. An initial experiment showed that none of the metastatic cell lines suffered from significantly reduced viability after 24 h of culturing in the CM. Thus, a culture time of 24 h and the DMEM CM without serum were selected as the essential conditions for culturing the cells[18]. Figure 1A shows that the procedures for the collection of the secretomes. The CM was centrifuged and precipitated with TCA. Protein digestion was performed using the filter aided sample preparation (FASP) method. After the digestion, N-glycopeptides were captured by hydrazide chemistry and HILIC methods (Figure 1B).

Next, the enriched N-glycopeptides were processed by PNGase F to remove the glycans. The reaction was conducted in H₂¹⁶O for the hydrazide method and in H₂¹⁸O for the HILIC method. For the former, the peptide masses showed a 0.9848 Da increase due to the conversion of asparagine to aspartic acid, and for the latter, a 2.9883 Da increment was detected because of the introduction of an ¹⁸O atom during the PNGase F cleavage[10,17,24]. The deglycosylated peptides were detected using LC-ESI-MS/MS. The LC-ESI-MS/MS data were searched against a human proteome database using Proteome Discoverer software. The matched precursor ion area information was further used for label-free quantification (Figure 1C).

Mapping N-glycopeptides in the secretomes

A total of 100 µg FASP digested peptides (corresponding to 200 µg initial proteins) was used for each experiment with ~100 µL hydrazide-modified agarose beads or 2 mg zic-HILIC material used for the enrichment. Two biological replicates were conducted. The details of the protein identifications are described in Table S1. Figure 2A shows the enrichment specificity, which could be estimated based on the percentage of N-glycopeptides among all of the identified peptides, and the enrichment efficiency, which could be determined based on the number of identified unique N-glycosites. Although the hydrazide chemistry method showed higher specificity than zic-HILIC for the identification of glycopeptides (80 % vs. 30 %, as indicated by the trendline in Figure 2A), the former provided a much lower number of identified N-glycosites than the latter (280 vs. 600 on average). Compared with other studies using similar enrichment methods[10,17,21], higher numbers of N-glycosites and glycoproteins were identified here (526 unique N-glycosites and 320 unique glycoproteins from hydrazide chemistry, and 1,032 unique N-glycosites and 536 unique glycoproteins from the zic-HILIC method), of which 344 N-glycosites and 245 unique glycoproteins were identified by both methods (Figure 2B); the other N-glycosites and glycoproteins were obtained from hydrazide or zic-HILIC alone (181 vs. 687 N-glycosites and 75 vs. 291 glycoproteins, as shown in Figure 2B). Therefore, the two enrichment methods were somewhat complementary.

Figure 2C shows the overlap of unique N-glycosites and unique N-glycoproteins between the MHCC97H and HCCLM3 cell lines. Based on the two enrichment methods, 948 and 907 unique N-glycopeptides were detected from MHCC97L and HCCLM3 cell lines, respectively, and the degree of overlap for the N-glycosites (984 vs. 940) and N-glycoproteins (500 vs.

Figure 1

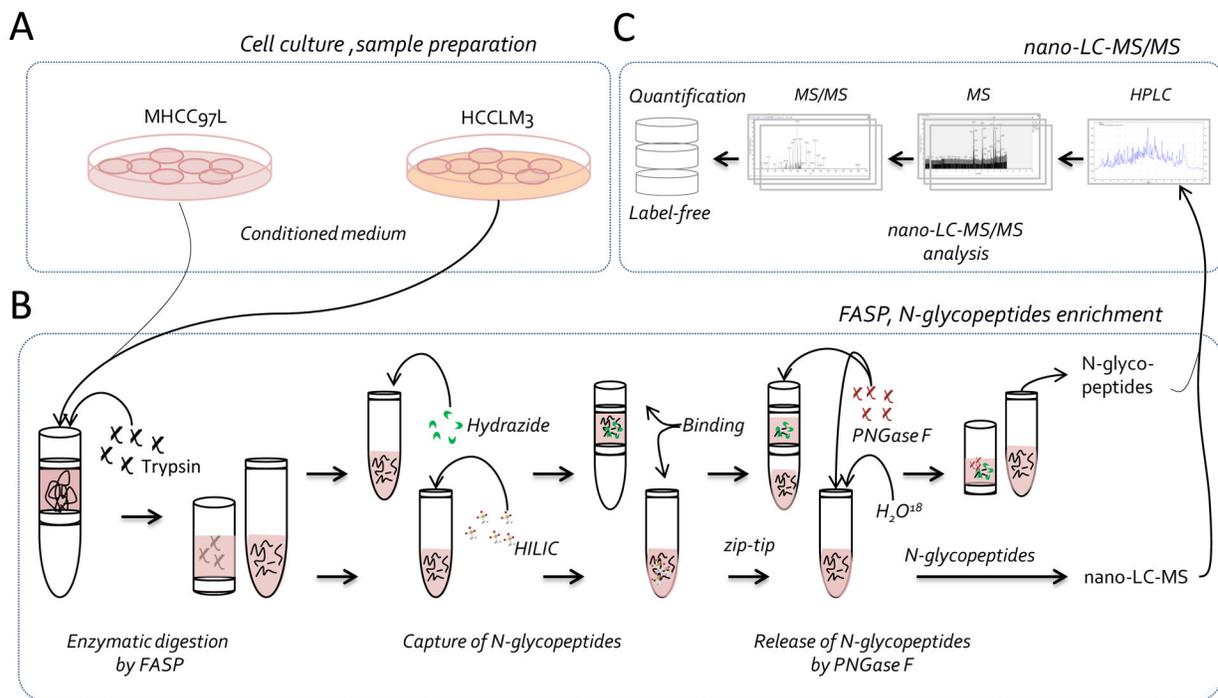


Figure 1. Overview of the experimental workflow. A) The secretome was collected from the conditioned medium. B) N-glycosylated peptides were enriched using hydrazide chemistry and zic-HILIC methods. First, proteins were digested using FASP, and then the N-glycosylated peptides were captured using two methods, followed by de-glycosylation using PNGase F and LC-MS-MS analysis. C) Label-free quantitative analysis.

doi: 10.1371/journal.pone.0081921.g001

499) that were identified from the two metastatic cell lines was ~60 % and ~66 %, which suggests a relatively high degree of similarity between the two cell lines.

A total of 1,165 unique N-glycopeptides and 1,213 N-glycosites were recognized and mapped to 611 glycoproteins (Table S2). Approximately 50 % of the glycoproteins were identified with one N-glycosite, nearly 34 % with two or three N-glycosites, and approximately 14 % with four to nine N-glycosites. The protein with the most N-glycosites was low-density lipoprotein receptor-related protein 1 (LRP1), with 18 identified N-glycosites (Figure 2D).

Biological categorization of identified glycoproteins

To determine what subset of proteins was enriched in the secretome, we performed a prediction analysis of the protein cellular localization using Gene Ontology (GO). Approximately 65 % of the glycoproteins identified were categorized into the extracellular space and plasma membrane in the GO subcellular localization annotation (Figure 3A) (Table S3). Although 27 % of the glycoproteins have a GO subcellular localization term of cytoplasm, 75 % of these proteins were predicted to contain signal peptides. Sixty-four of these cytoplasm-localized proteins contained signal peptides, as

determined by Signal P 4.0, and forty-six of them were predicted to be localized to the exosome by comparing the data against an exosome database[25]. Taken together, 90% of the identified proteins were predicted to be secreted. Moreover, the enrichment scores for the clusters of glycoproteins and secretory proteins were 195.2 and 50.2, respectively, as compared to the human proteome background by DAVID analysis (*p*-values: 2.52×10^{-204} and 2.42×10^{-63} ; Benjamini: 1.26×10^{-201} and 3.03×10^{-61}) (Table S4).

The most significantly enriched biological functions are listed in Figure 3B. The top 10 functional categories, according to their *p*-values, that are enriched in the metastatic cell lines are shown in Figure 3B and Table S5. It can be observed that the top category contains proteins related to cell movement function, which could be involved in cancer cell migration and invasion.

Label-free quantitative analysis of N-glycopeptides

To explore the differentially expressed glycoproteins, label-free quantitative analysis was performed. Peptide quantities were estimated by precursor ions areas, which were shown in Table S1. The differentially expressed glycopeptides between the two cell lines are shown in Table S6, of which 1,096 N-

Figure 2

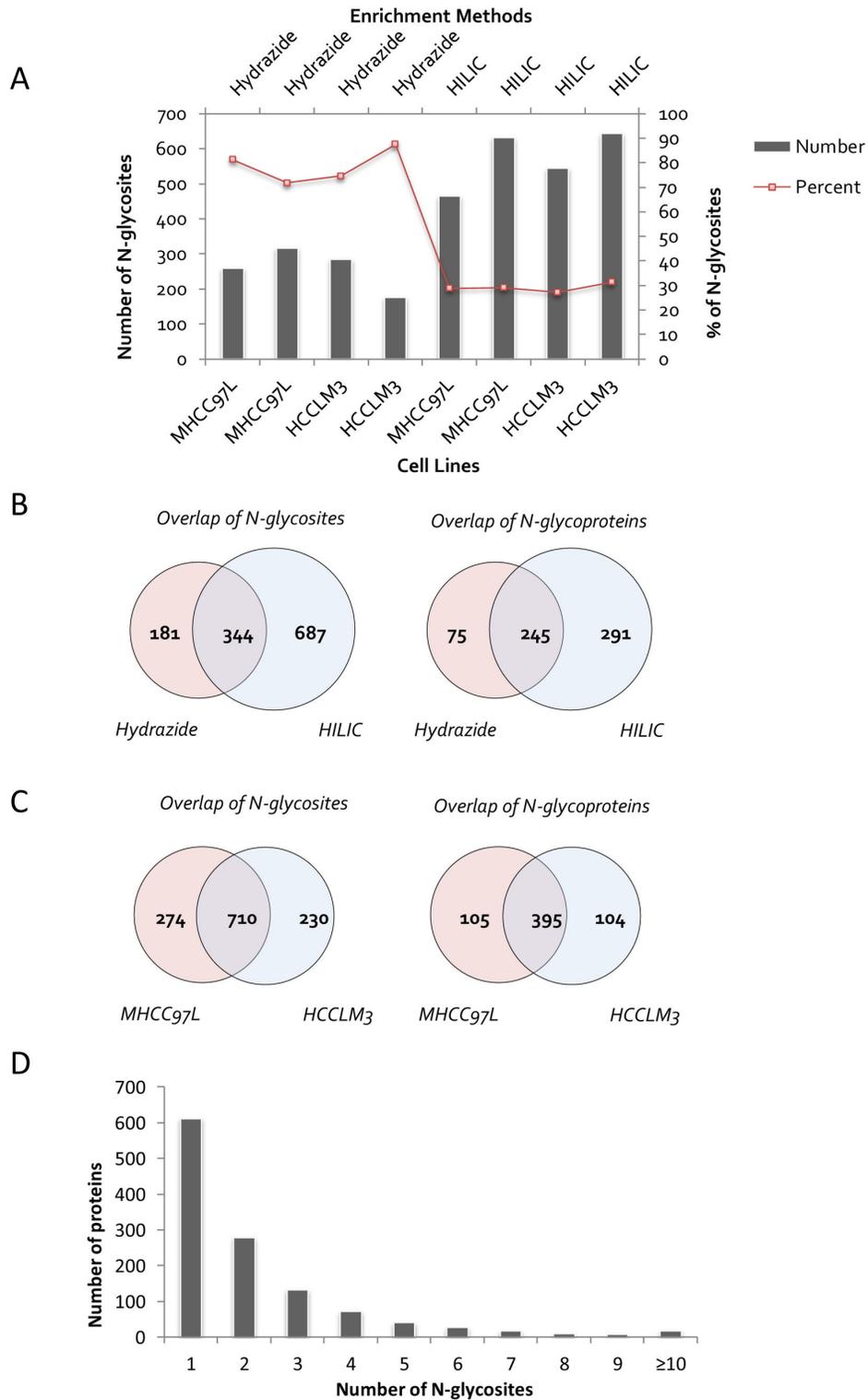


Figure 2. Summary of identified N-glycosites. A) The number of unique N-glycosites identified and the percentage of N-glycopeptides from all of the identified peptides in each cell line. B) Overlap of N-glycosites between the different enrichment methods. C) Overlap of the N-glycosites and proteins between the different cell lines. D) Number of N-glycosites identified per protein.

doi: 10.1371/journal.pone.0081921.g002

Figure 3

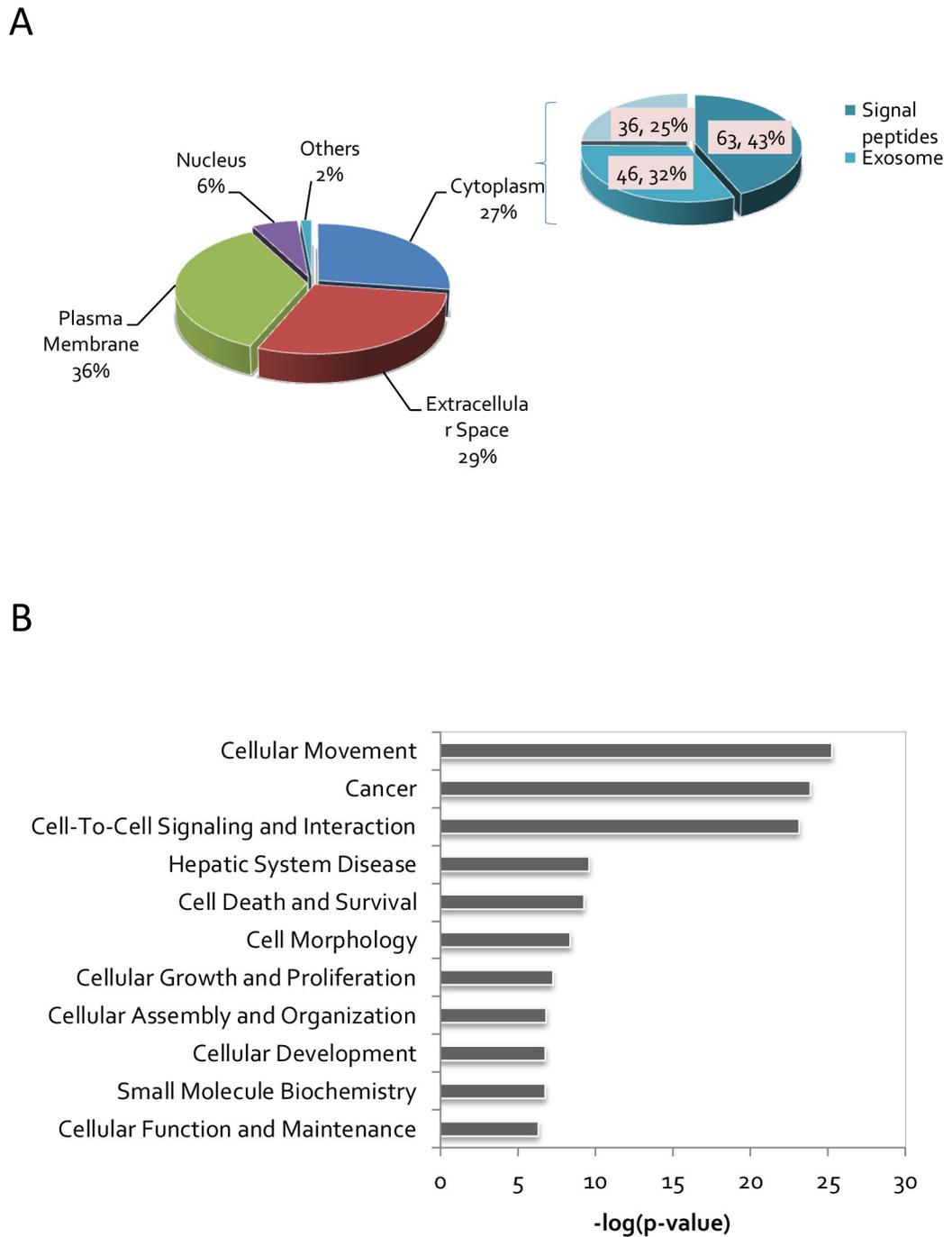


Figure 3. Bioinformatics analysis of identified N-glycoproteins. A) Cellular component annotation of identified N-glycoproteins. B) Biological functions of differentially expressed N-glycoproteins.

doi: 10.1371/journal.pone.0081921.g003

Figure 4

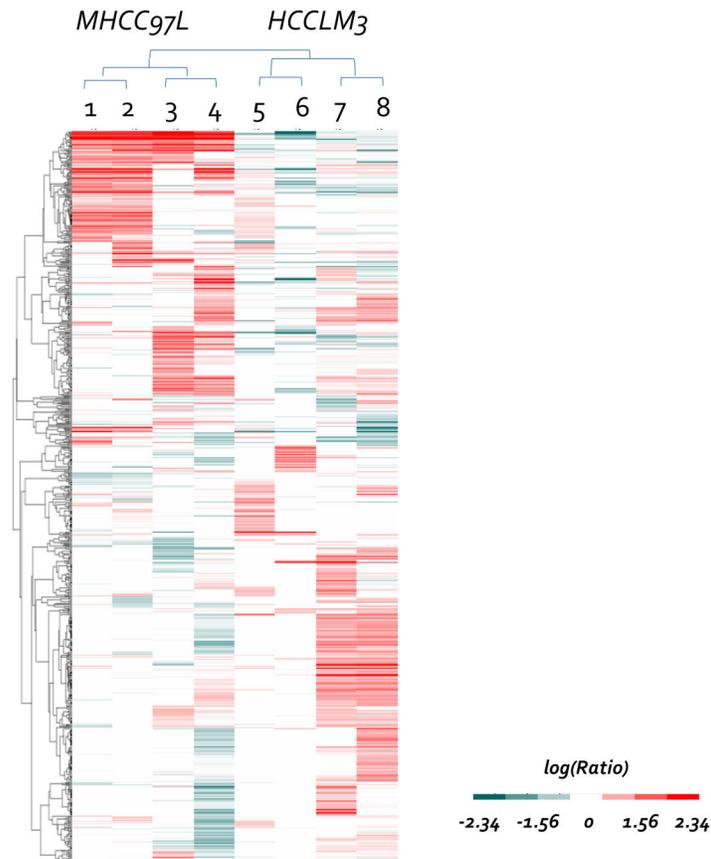


Figure 4. Label-free quantitative analysis. Hierarchical clustering based on the Euclidean distance of the peak area of label-free quantification of N-glycosites showed a significant difference between the MHCC97L and HCCLM3 cell lines.

doi: 10.1371/journal.pone.0081921.g004

glycopeptides had quantification values. To avoid the possible interference from noise, the glycopeptides were normalized by the total precursor area of the identified peptides[20,21]. The quantitative information of the 683 overlapped N-glycopeptides between the two cell lines was filtered by hierarchical clustering (Figure 4) and segregated into 10 distinct clusters, including up-regulated, down-regulated or unchanged glycopeptides (Table 1). Clusters a) and b) contain the glycopeptides with higher relative abundance levels in the MHCC97L cell line, which includes 65 and 86 glycopeptides from the two enrichment methods. Clusters i) and j) contain the glycopeptides that have higher relative abundance in HCCLM3 cells, which includes 101 and 64 glycopeptides from the two enrichment methods (Table S6). Clusters c) - h) include more than 50% of the identified glycopeptides without any significant changes in abundance between the two metastatic cell lines.

In addition to the N-glycopeptides that were identified in both metastatic cell lines, there were 258 N-glycopeptides that were only observed in the MHCC97L cell line and 217 N-glycopeptides that were only observed in the HCCLM3 cell line. These glycopeptides belong to 385 unique N-glycoproteins. Among them, 84 % were identified with only one N-

glycopeptide, while approximately 4 % of them were identified with 3 N-glycosites. Fibronectin 1 (FN1) was identified by 5 and 9 N-glycopeptides from the MHCC97L and HCCLM3 cell lines, respectively, and the quantitative analysis indicated that the relative abundance level of the former was far less than that of the latter. The 4 glycopeptides that were not identified in the MHCC97L cell line may have been of low abundance, or they could have been affected by altered FN1 glycosylation patterns under different metastatic potentials. From a biomarker perspective, the proteins that belong to clusters i) - j) in Table 1 and those that are uniquely expressed in the HCCLM3 cell line would be the most valuable, as these were up-regulated or detected in concordance with the increased HCC metastatic potential. Within these clusters, well-known hepatocellular carcinoma biomarkers such as alpha fetoprotein (AFP) and Dickkopf-related protein 1 (DKK1)[26] were identified, as well as proteins that have been suggested to be related to metastasis in other studies, for example, integrin alpha-3 (ITGA3) and tissue inhibitor of metalloproteinase-1 (TIMP-1) [21]. These glycoproteins are taken as the positive controls for our strategy, and the results indicate that the N-glycosylated

Table 1. Ten major clusters extracted from Figure 4 by K-Means expression clustering.

Description	Cluster	Number	Tendency
The glycopeptides with higher relative abundance levels in the MHCC97L cell line	a)	65	MHCC97L >> HCCLM3
	b)	86	MHCC97L > HCCLM3
More than 50% of the identified glycopeptides without any significant changes in abundance between the two metastatic cell lines.	c)	62	No significant changes
	d)	51	
	e)	26	
	f)	83	
	g)	54	
	h)	86	
The glycopeptides have higher relative abundance in HCCLM3 cells	i)	101	MHCC97L < HCCLM3
	j)	64	MHCC97L << HCCLM3

doi: 10.1371/journal.pone.0081921.t001

secretome of the metastatic cell lines provides a potential source of disease markers.

The network of metastasis-related glycoproteins

To further elucidate the correlation of the identified N-glycoproteins with metastasis and other diseases, we assigned the glycoproteins containing up-regulated N-glycosites in the HCCLM3 cell line to different diseases using the web tool FunDO[27]. As shown in Figure 5A-B, approximately half of the 300 glycoproteins were related to cancer, liver cancer, and neoplasm metastasis (Bonferroni corrected p -values: 7.82×10^{-19} , 6.22×10^{-7} and 2.11×10^{-7}). To explore the key glycoproteins and substrates that are related to HCC metastasis, a metastatic network was constructed using the differentially expressed glycoproteins (Figure 5C). The significantly regulated liver cancer-related glycoproteins that were identified by FunDo and IPA are listed in Table 2. Their gene ontology (GO) information, including cellular component and biological processes, is also shown in Table 2.

Validation of up-regulated glycoproteins

Among these glycoproteins containing up-regulated N-glycosites, FN1 and FAT1 were selected for further validation via western blotting analysis (Figure 6A). The protein alpha-galactosidase A (GLA) was selected as a loading control because it is expressed equally in the cell lines[28]; its ratio was also close to 1 based on the mass spectrometry (MS) data (Table S6) and the western blot results (Figure 6A). To measure the glycosylated status of the two proteins, the same

experiment was repeated with the addition of PNGase F processing (Figure 6B). The results showed that for FAT1, an apparent shift to the low molecular weight region was observed. While for FN1, the shift was not significant. This could be caused by the huge molecular weight of FN1 (260 kDa), thus, the releasing of glycans from the protein sequence did not induce a discriminable shift.

Discussion

Our work represents an in depth glycoproteomics analysis of the secretome of HCC cell lines with different metastatic potentials. By applying hydrazide chemistry and zic-HILIC for glycopeptide enrichment, followed by high mass accuracy LC-MS/MS analysis, N-glycopeptides and specific N-glycosites were identified with high confidence[17,24,28]. Our results demonstrate that the agarose bead-based hydrazide chemistry offers a highly specific means of enriching N-glycosites from the secretome, and the specificity was higher than 80 % [17]. Although many non-glycopeptides were identified, the zic-HILIC enrichment method proved to be a more effective strategy because approximately 600 N-glycosites were identified per experiment. The combination of both methods could significantly improve the identification of N-glycosites.

The expression of the glycoproteins in the secretory system reflected the intracellular regulation of the N-glycosite synthesis, transport and secretory pathway, which could be illustrated through IPA and DAVID analysis[29]. Among these biological functions, the most significant class, cell movement, related to cancer cell migration and invasion was observed, in which, cathepsin B (CTSB), fibronectin (FN1), CD44 antigen (CD44), extracellular matrix protein 1 (ECM1), and protocadherin Fat 1 precursor (FAT1) previously shown to play an important role during the metastasis of other cancer cells[30] were included.

Four hundred forty-seven N-glycosites from 300 glycoproteins were found to be significantly altered with a ratio > 2 (significantly different with $p < 0.05$, as determined by SPSS analysis), including clusters i) and j) in the label-free quantification and the non-overlapping dataset from HCCLM3. Validation by western blot analysis exhibited the same trends as the mass spectrometry intensity (Figure 6), which further supports the observation that most of the glycoproteins could play a role in metastasis. Next, we focused on the function analysis of these up-regulated glycoproteins because they represent potential diagnostic maker candidates for predicting hepatocellular carcinoma metastasis and because they are drug targets for inhibiting metastasis.

The growth factor receptors ErbBs and FGFRs have been reported to be associated with breast cancer development. Two and three of their N-glycosites were identified in this study, respectively[31]. PRSSs (one identified glycosite) and latent TGF-beta binding proteins (LTBPs, five identified glycosites) are involved in growth factor signaling and were included in the detection list. CD151, identified with one N-glycosylation site, is a cell surface glycoprotein that is known to enhance the motility, invasion and metastasis of cancer cells[32]. MET, reported as a hepatocyte metastasis-related protein, is a

Figure 5

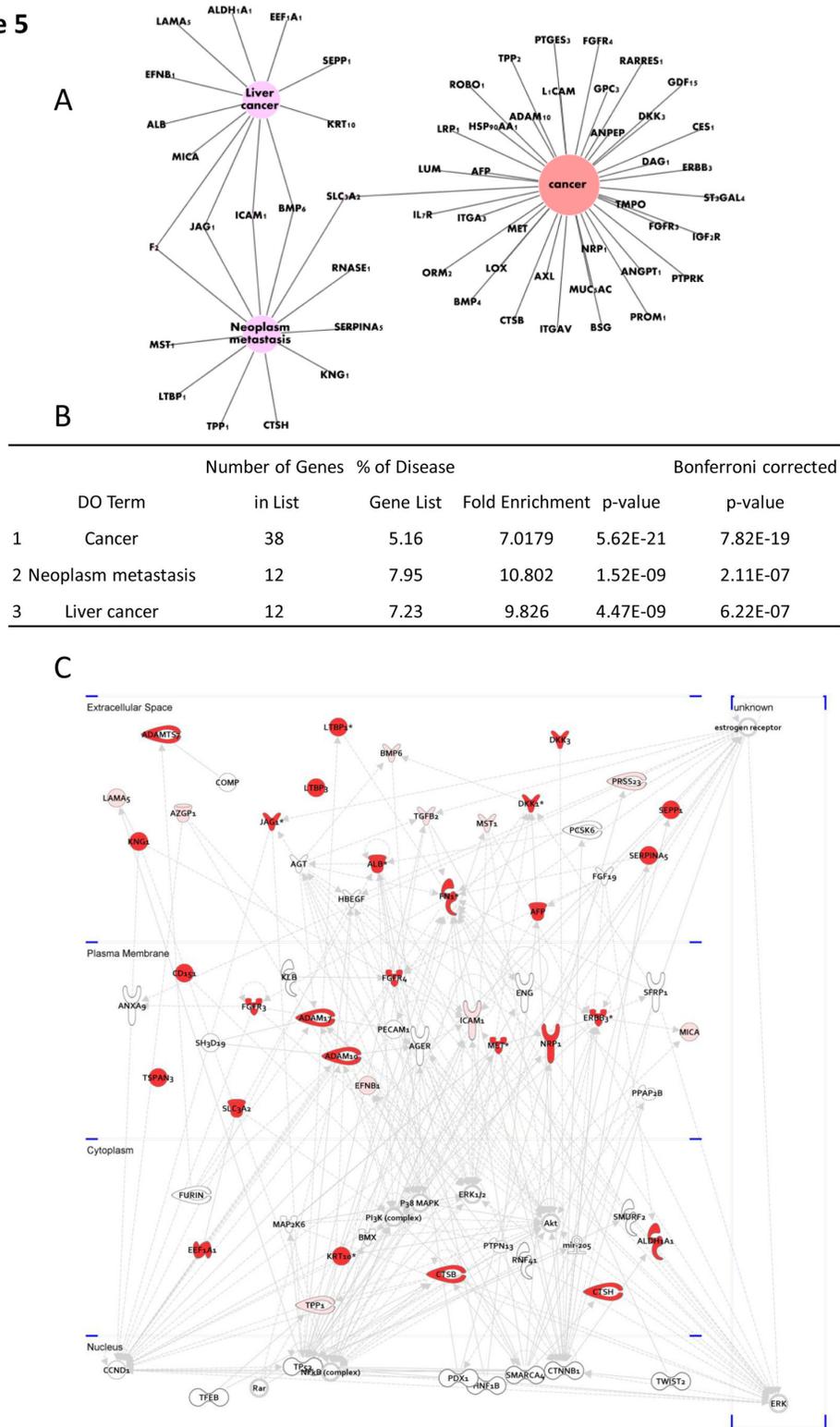


Figure 5. Network view of the up-regulated N-glycoproteins in HCCLM3 cells. A) The networks of the top 3 liver-related diseases. B) The number of related genes and the *p*-value of the top 3 liver-related diseases that are indicated in A). C) Cellular motility network. (The proteins with higher expression in HCCLM3 cells are in red (Ratio > 2), whereas the other proteins that were generated from the IPA database are not colored.).

doi: 10.1371/journal.pone.0081921.g005

Table 2. List of significantly up-regulated N-glycosites related to liver cancer in FunDo and IPA.

Protein Accessions	Peptide Sequence	Protein Name	Gene symbol	Expression level*	Quantitative ratio*	location	Type
40548389	ASSEYNLANLPPSYHn*ETINTDTK	Dickkopf-related protein 3	DKK3	M >> L	∞	√	Cytokine
4502337	DIVEYrn*DSn*GSHVLQGR	Alpha-2-glycoprotein 1, zinc-binding	AZGP1	M > L	2.052	√	Transporter
47132557	DQGVDDITYNVn*DTFHKR	Fibronectin 1	FN1	M > L	2.004	√	Enzyme
18497288	DSCCQGSn*MTLIGENGHSTDILTGSQGR	Latent transforming growth factor beta binding protein 3	LTBP3	M >> L	∞	√	Other
261337165	ECYYNLNDASLCDNVLAPn*VTK	Latent transforming growth factor beta binding protein 1	LTBP1	M >> L	∞	√	Other
62530391	EGYSn*ISYIVVNHQGISSR	Selenoprotein P, plasma, 1	SEPP1	M >> L	∞	√	Other
4502425	Fn*LSQIPEGEVTAAEFR	Bone morphogenetic protein 6	BMP6	M > L	2.024	√	Growth factor
38683827	Fn*LTANQHLLAPGFVSETR	ADAM metalloproteinase with thrombospondin type 1 motif, 7	ADAMTS7	M >> L	∞	√	Peptidase
21265058	GAEYISPLPr*ASAPAAQR	ADAM metalloproteinase with thrombospondin type 1 motif, 15	ADAMTS15	M >> L	∞	√	Peptidase
47132557	GGNSGALCHFPFLYNNHn*YDCTSEGR	Fibronectin 1	FN1	M >> L	64.92	√	Enzyme
4501989	Vn*FTEIQK	Alpha-fetoprotein	AFP	M >> L	∞	√	Transporter
205277383	GTAn*TTTAGVPCQR	Hepatocyte growth factor-like	MST1	M > L	4.103	√	Growth factor
7110719	IQKDHQASn*SSR	Dickkopf-related protein 1	DKK1	M >> L	∞	√	Growth factor
47132557	ISCTIArR	Fibronectin 1	FN1	M >> L	∞	√	Enzyme
4507463	LTSPPEDYPEPEEVPPEVSIYn*STR	Transforming growth factor-beta 2	TGFB2	M > L	4.429	√	Growth factor
4557679	n*CSHLKDHCR	Jagged 1	JAG1	M >> L	∞	√	Growth factor
6005882	QYLSYETLYAr*GSR	Protease, serine, 23	PRSS23	M > L	8.103	√	Peptidase
4506547	Rn*MTQGR	Ribonuclease, RNase A family, 1	RNASE1	M >> L	∞	√	Enzyme
4505821	TFYWDFYTrR	Prolactin-induced protein	PIP	M >> L	∞	√	Other
21264602	TLSELSMSQTGHLGLAn*ASAPSGEQLLR	Laminin, alpha 5	LAMA5	M > L	2.616	√	Other
194018472	VLPSLGISNVFTSHADLSGISn*HSNIQVSEMVHK	Serpin peptidase inhibitor, clade A member 5	SERPINA5	M >> L	∞	√	Other
262050546	YNSQn*QSNQFVLYR	Kininogen 1	KNG1	M >> L	∞	√	Other

L: MHC97L and M: HCCLM3. *, > / >> The expression level of N-glycopeptides in HCCLM3 was more than two/ ten times higher than that in the MHC97L cell line; >>>: only expressed in the HCCLM3 cell line.
doi: 10.1371/journal.pone.0081921.t002

Figure 6

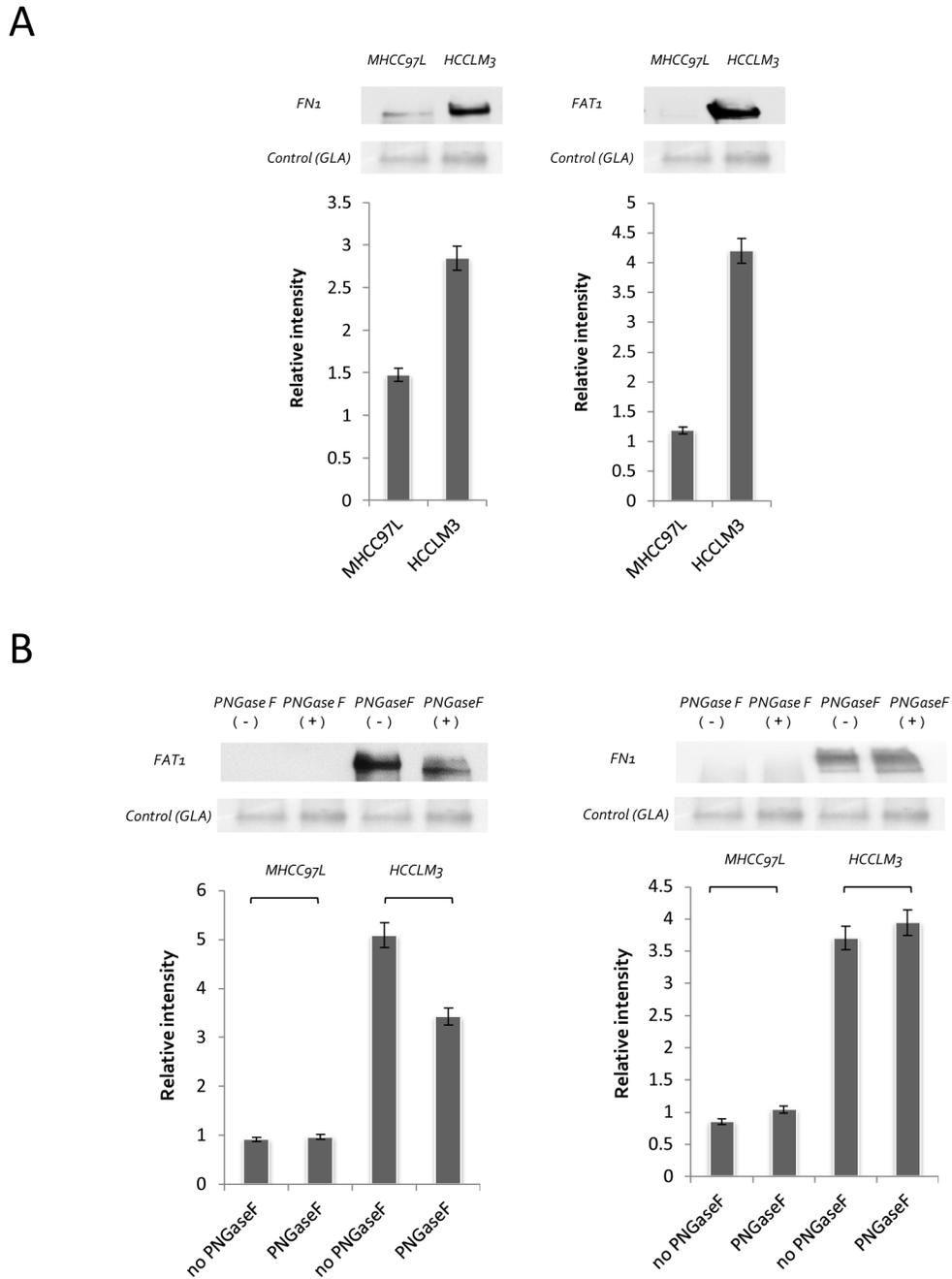


Figure 6. Validation of the differential expression of two selected N-glycoproteins. A) Ten micrograms secretome protein samples were separated on SDS-PAGE gels, transferred to PVDF membranes, and probed with anti-FN1 or FAT1 antibodies. B) The de-glycosylation of the same amount of secreted proteins from MHCC97L and HCCLM3 cells was performed with PNGase F cleavage for 12 h. Proteins were separated on SDS-PAGE and analyzed by western blotting.

doi: 10.1371/journal.pone.0081921.g006

membrane receptor that is essential for embryonic development and wound healing [33]. Seven N-glycosites of MET proteins were identified in our dataset, and three had a higher relative abundance in the HCCLM3 secretome.

Although the aforementioned proteins are involved in tumor invasion and metastasis, they could have low specificity for hepatocellular carcinoma. Combining the prediction results by FunDO[27] that produced a list of liver cancer and metastasis related proteins with IPA network analysis, we obtained 20 glycoproteins that could be related to liver cancer metastasis (Table 2). All of these proteins are localized to the extracellular space. For example, hepatocyte growth factor-like protein in humans is encoded by the MST1 gene that is involved in cell apoptosis and metabolism, and it exhibited a higher relative abundance in the high metastatic cell line[34]. BMP6 is the key regulator of hepcidin, a small peptide secreted by the liver and acted as a major regulator of iron metabolism in mammals[35]. In the IPA network (Figure 5C), the molecule that is upstream of BMP6 is DKK1, and both of them were highly expressed at the N-glycosylation level. Because DKK1 has been reported as an important marker of hepatocellular carcinoma[26], we reasoned that this pathway could be significantly involved in liver cancer metastasis and invasion.

Synergistic secretory proteins were also among the up-regulated proteins in this study. For example, zinc alpha2 glycoprotein (AZGP1) is expressed in epithelial secretory cells[36] and forms a complex with PIP; their co-up-regulation has been detected in well-differentiated tumors[37]. TGF-beta is regulated by its association with LTBP s [38,39]. LTBP1, LTBP3 and TGFβ2 were all up-regulated in the HCCLM3 cell line. ADAMTS[40], selenoprotein (SEPP1)[41], ICAM1[42], JAG1[43], and LAMA5[44] have been demonstrated to have important roles in connective tissue organization, angiogenesis and cell migration. The label-free quantification results implied that they could play an important role in hepatocellular carcinoma metastasis.

Additionally, we found that FN1 had a higher level of expression in HCCLM3 cells (9 N-glycosites) and a lower level in MHCC97L cells (5 N-glycosites), which was further verified by western blot (Figure 6A). FAT1 was only detected in the HCCLM3 cell line (2 N-glycosites), and its expression was also verified by western blot (Figure 6A). The removal of glycans by PNGase F affected the migration of the protein in the gel, indicating its fully glycosylation (Figure 6B).

Conclusion

In summary, we used a glycoproteomics approach to investigate the secretome of hepatocellular carcinoma cell lines with different metastatic potentials. The qualitative and quantitative analysis of N-glycopeptides from the secretomes unveiled a panel of glycoproteins that may be related to hepatocellular carcinoma metastasis. It should be noted that our study focused on the difference of N-glycosites, and the regulation on the intact proteins were not examined, which mean that some of the differences at the N-glycosites level may be resulted from the differential expression of intact proteins. Dedicated experiments should be designed to

consider both the changes at the protein level and glycosylation level. Nevertheless, our results still proved that it is efficient to find potential biomarkers through the capture of glycopeptides and its quantitative application, which was supported by the discovery of several well-known markers that were partially validated in our study. The glycoproteomics approaches we used, and the differentially expressed proteins that we identified, could benefit further research that seeks to discover early diagnostic or therapeutic biomarkers for hepatocellular carcinoma metastasis, and it could also provide a general strategy for other disease-related secretome studies.

Materials and Methods

Cell culture and sample preparation

The human metastatic HCC cell lines MHCC97L and HCCLM3 were obtained from the Liver Cancer Institute of Zhongshan Hospital, which is affiliated with Fudan University (Shanghai, China). Their genetic backgrounds are similar because they are filial-cloned sequentially from a MHCC97 parental cell line[18,45]. The metastatic potential of the HCCLM3 cell line is higher than that of the MHCC97L cell line[46]. After orthotopic implantation of HCCLM3 tumor tissue into nude mouse liver for 35 days, widespread loco-regional and distant metastases were found 100% in lungs and abdominal wall, 80% in intra-abdominal cavity, and 70% in diaphragm[45,47]. While the pulmonary metastatic rate of MHCC97L was only 40 % [48]. Immunocytochemical studies demonstrated that the two clones were positive for AFP and the concentration of serum AFP was higher in HCCLM3-inoculated mice than that of MHCC97L[49]. The cell lines with different metastatic potential provides an important model system for the in vivo and in vitro study of HCC metastasis[50].

Approximately 1×10^7 cells were cultured at 37 °C in 5 % CO₂ in DMEM (Hyclone, USA) supplemented with 10 % fetal bovine serum until reaching 60 ~ 70 % confluence. Cells were washed stringently and gently four times with Dulbecco's phosphate buffered saline with calcium and magnesium (DPBS) and once with serum-free DMEM (Conditioned Medium, CM). Cells were then incubated in the CM at 37 °C for 24 h. The secretory proteins were collected and centrifuged at 2,500 ×g for 10 min (4 °C) to remove the detached cells and large debris. The resulting supernatant was collected and centrifuged for 1 h at 100,000 ×g (4 °C) to remove smaller debris and vesicles. Formic acid (FA, final concentration of 0.1 %) was immediately added to the final supernatant, which was then stored at -80 °C. The addition of FA lowered the pH (pH < 4) of the culture supernatants, thus reducing the activity of many proteases[11,22].

The proteins in the culture supernatants were extracted by trichloroacetic acid (TCA) precipitation. Thirty milliliters of culture supernatants were used for each precipitation. All of the treatments were performed at 4 °C. TCA was added to the CM solution to a final concentration of 12 % [23]. After mixing, the solution was precipitated for 4 h at 4 °C, followed by centrifugation (10,000 ×g, 10 min, 4 °C). The protein pellet was carefully washed twice with 2 mL tetrahydrofuran (THF) (pre-cooled on ice), and each wash was followed by centrifugation

(10,000 ×g, 10 min, 4 °C). The final pellet was re-dissolved in 0.4 mL lysis buffer (4 % SDS, 0.1 M Tris-HCl, pH 7.6, 100 mM DTT) with a sonicator bath (30 min extraction). The concentrations of the extracted proteins were measured by a NanoDrop spectrophotometer (Thermo, USA) at 280 nm with an extinction coefficient of 1.1 absorbance units[51].

Protein digestion by FASP

Detergent was removed from the lysates, and the proteins were digested with trypsin using the FASP protocol[51], which entails the ultrafiltration of spin units with a nominal molecular weight cut-off of 30,000. Briefly, to YM-30 Microcon filter units (Millipore, USA) containing the CM protein concentrates, 200 μL 8 M urea in 0.1 M Tris/HCl, pH 8.5 (UA buffer) was added, and the units were centrifuged at 14,000 ×g at 20 °C for 15 min. This step was repeated twice. Then, 50 μL 0.05 M iodoacetamide in 8 M urea was added to the filters and the samples were incubated in the dark for 20 min. The filters were washed twice with 100 μL 8 M UA buffer followed by three washes with 100 μL 50 mM NH₄HCO₃. Finally, trypsin (Promega, USA) was added in 100 μL 50 mM NH₄HCO₃ to each filter. The protein to enzyme ratio was 100:1. The samples were incubated overnight at 37 °C, and the digested peptides were collected by centrifugation at 14,000 ×g at 20 °C for 15 min.

N-glycopeptide enrichment by hydrazide

N-glycopeptides were captured from tryptic peptide mixtures by hydrazide chemistry following the protocol of Zhang[19] with modifications. Briefly, an aliquot of 100 μg peptides was diluted with coupling buffer (50 % acetonitrile, 0.5 % acetic acid) to a final reaction environment and then oxidized with 20 mM sodium periodate (NaIO₄) at room temperature in the dark for 1 h. The excess NaIO₄ was removed after oxidation by desalting using a Sep-Pak C18 column (Waters, USA). The oxidized peptides in 800 μL coupling buffer (80 % acetonitrile, 0.5 % acetic acid) were then incubated overnight at room temperature with hydrazide beads (100 μL beads for ≤ 100 μg peptides) that were pre-washed twice with 1 mL water. Unbound peptides in the supernatant were removed by centrifugation. The beads were then washed 3 times with 2 mL water, 3 times with 2 mL 1.5 M NaCl and 3 times with 50 mM NH₄HCO₃ to remove the non-glycopeptides. The captured N-glycopeptides were released by incubating the beads with 3 μL PNGase F (500 units/μL, New England Biolabs, USA) in 100 μL 50 mM NH₄HCO₃ buffer (dissolved in ¹⁸O water) at 37 °C for 12 h. The released glycopeptides were collected from the supernatant, and the hydrazide beads were washed twice with 200 μL 80 % ACN and 0.1 % FA. The washing solutions were combined with the supernatant collected from the PNGase F incubation and dried down to a final volume of 20 μL using a SpeedVac concentrator.

N-glycopeptide enrichment by HILIC

The digested peptides were enriched with zic-HILIC media by the procedure reported by Calvano[15] with slight modifications. Briefly, first, the C8 disk was inserted into a 200 μL tip. Approximately 100 μg in-solution digested samples was

dissolved in 80 % ACN, 0.5 % FA and incubated overnight at room temperature with 2 mg zic-HILIC media (Merck; particle size 10 μm) that was pre-washed twice with 1 mL coupling buffer (80 % ACN, 0.5 % FA). Then, zic-HILIC media was loaded into the 200 μL tip that was pre-filled with a C8 disk. The zic-HILIC tip was washed 5 times with 100 μL 80 % ACN, 1 % FA, 19 % H₂O, and the bound peptides were eluted 3 times with 80 μL elution buffer (99 % H₂O, 1 % FA). The eluate was dried, de-glycosylated with 3 μL PNGase F (500 units/μL, New England Biolabs) in 50 mM NH₄HCO₃ (dissolved in ¹⁸O water) at 37 °C overnight and dried down to a final volume of 20 μL using a SpeedVac concentrator.

LC MS/MS analysis

A nanoflow HPLC instrument (EASY-nLC 1000 system, Thermo Fisher Scientific, USA) was coupled on-line to a Q-Exactive mass spectrometer (Thermo Fisher Scientific, USA) with a nanoelectrospray ion source (Thermo Fisher Scientific, USA)[10]. Chromatography columns were packed in-house with Ultimate XB-C18 3 μm resin (Welch Materials, USA). The peptide mixtures were loaded onto the C₁₈-reversed phase column (10 cm length, 75 μm inner diameter) with buffer A (99.5 % water and 0.5 % FA) and separated with a 75 min linear gradient of 3 - 100 % buffer B (99.5 % acetonitrile and 0.5 % FA) at a flow rate of 350 nL/min. Including the loading and washing steps, the total time for an LC MS/MS run was approximately 90 min. The electrospray voltage was 2.0 kV. Peptides were analyzed by data-dependent MS/MS acquisition with a dynamic exclusion duration of 18 s. In MS1, the resolution was 70,000, the AGC target was 3e⁶, and the maximum injection time was 20 ms. In MS2, the resolution was 17,500, the AGC target was 1e⁶, and the maximum injection time was 60 ms. The scan range was 300 - 1400 m/z, and the 20 most intensive precursor ions were selected for MS/MS analysis.

Data analysis

The raw data were processed using the Proteome Discoverer 1.3 proteomics platform. The fragmentation spectra were searched against the RefSeq Human database (20120320) using the Mascot search engine (v 2.2.06) with the precursor and fragment mass tolerances set to 10 ppm and 20 mmu, respectively. Two missed cleavage sites were allowed, and the minimum peptide length was 7 amino acids. Variable modifications: oxidation (M), and acetylation (protein N-terminal). Fixed modification: carbamidomethyl (C). Additionally, the deamidation of asparagine together with the incorporation of an ¹⁸O atom was set as a variable modification for the assignment of N-glycosites. Peptide ions were filtered using the cut-off scores of Percolator based on *p*-values < 0.01. The false discovery rate (FDR) was set to 1 % for peptide identifications. For label-free quantification, precursor ions areas were extracted using Proteome Discoverer 1.3 with a 10 ppm mass precision (the experimental m/z and retention times were recorded for precursor area quantification). Ratios for each peptide were normalized by the total identified peptides, and the ratios were used for N-glycopeptide-level quantification.

Bioinformatics and statistical analysis

To evaluate the trend of protein abundance changes, hierarchical clustering was performed and a distance tree was generated using Multi Experiment Viewer (MeV)[52]. To visualize the significantly regulated N-glycosites, Statistical Product and Service Solutions (SPSS) software was used. To study gene-disease relationships, the differentially regulated genes in metastatic cell lines were assigned to different diseases based on Disease Ontology and peer-reviewed evidence from GeneRIF using the web tool FunDO (<http://django.nubic.northwestern.edu/fundo>)[27]. Then, Cytoscape v. 2.8.2 was used to visualize gene-disease interaction networks[53]. To explore the biological functions, subcellular localization and pathways and networks of the N-glycoproteins involved, Gene Ontology (GO) annotation[54], DAVID Bioinformatics Resources[29] and Ingenuity Pathways Analysis (IPA, Ingenuity Systems, Mountain View) were employed, respectively.

Western-blotting analysis

The secretory proteins of metastatic HCC cell lines were resolved on 10 % SDS-PAGE gels, followed by transferring onto PVDF membranes (Millipore, USA). After incubation in blocking buffer (0.5 % Tween-20 in TBS, 5 % BSA) for 1 h at room temperature, membranes were blotted using antibodies against the targeted proteins for 1 h at room temperature. Membranes were then washed with TBST (TBS with 0.5% Tween-20) and incubated in 1:5,000 diluted HRP-conjugated IgG for 1 h at room temperature. After washing three times with TBST, the bands on the membrane were visualized using an ECL plus detection system[18].

Supporting Information

Table S1. The details of the protein identifications from the 8 replicate experiments, including the peak area information used for label-free quantitative analysis.

(XLSX)

Table S2. A total of 1,165 unique N-glycopeptides and 1,213 N-glycosites were recognized and mapped to 611 glycoproteins.

(XLSX)

Table S3. GO subcellular localization annotation of 611 N-glycoproteins identified.

(XLSX)

Table S4. The enrichment scores for the clusters of glycoproteins and secretory proteins were 195.2 and 50.2, respectively, as compared to the human proteome background by DAVID analysis.

(XLSX)

Table S5. The top 10 functional categories from IPA analysis.

(XLSX)

Table S6. The differentially expressed N-glycopeptides between MHCC97L and HCCLM3 cell lines.

(XLSX)

Author Contributions

Conceived and designed the experiments: XL WY XQ. Performed the experiments: XL JJ XZ. Analyzed the data: XL JW HH. Contributed reagents/materials/analysis tools: YZ. Wrote the manuscript: XL. Helped with expenses: BP. Advised on the writing and revision of the manuscript: RZ. Revised the manuscript: WY XQ.

References

- Chen TY, Li YC, Liu YF, Tsai CM, Hsieh YH et al. (2011) Role of MMP14 gene polymorphisms in susceptibility and pathological development to hepatocellular carcinoma. *Ann Surg Oncol* 18: 2348-2356. doi:10.1245/s10434-011-1574-x. PubMed: 21298348.
- Yeh CB, Hsieh MJ, Hsieh YS, Chien MH, Lin PY, et al. (2012) Terminalia catappa Exerts Antimetastatic Effects on Hepatocellular Carcinoma through Transcriptional Inhibition of Matrix Metalloproteinase-9 by Modulating NF-kappaB and AP-1 Activity. *Evid Based Complement Alternat Med* 2012: 595292
- Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 1: 845-867. doi:10.1074/mcp.R200007-MCP200. PubMed: 12488461.
- Yang G, Cui T, Wang Y, Sun S, Ma T et al. (2013) Selective isolation and analysis of glycoprotein fractions and their glycomes from hepatocellular carcinoma sera. *Proteomics*, 13: 1481–98. PubMed: 23436760.
- Brown KJ, Formolo CA, Seol H, Marathi RL, Duguez S et al. (2012) Advances in the proteomic investigation of the cell secretome. *Expert Rev Proteomics* 9: 337-345. doi:10.1586/epr.12.21. PubMed: 22809211.
- Eichelbaum K, Winter M, Diaz MB, Herzig S, Krijgsveld J (2012) Selective enrichment of newly synthesized proteins for quantitative secretome analysis. *Nat Biotechnol* 30: 984-990. doi:10.1038/nbt.2356. PubMed: 23000932.
- Lau KS, Dennis JW (2008) N-Glycans in cancer progression. *Glycobiology* 18: 750-760. doi:10.1093/glycob/cwn071. PubMed: 18701722.
- Drake PM, Cho W, Li B, Prakobphol A, Johansen E et al. (2010) Sweetening the pot: adding glycosylation to the biomarker discovery equation. *Clin Chem* 56: 223-236. doi:10.1373/clinchem.2009.136333. PubMed: 19959616.
- Ohtsubo K, Marth JD (2006) Glycosylation in cellular mechanisms of health and disease. *Cell* 126: 855-867. doi:10.1016/j.cell.2006.08.019. PubMed: 16959566.
- Boersema PJ, Geiger T, Wisniewski JR, Mann M (2013) Quantification of the N-glycosylated Secretome by Super-SILAC During Breast Cancer Progression and in Human Blood Samples. *Mol Cell Proteomics* 12: 158-171. doi:10.1074/mcp.M112.023614. PubMed: 23090970.
- Cao J, Shen C, Wang H, Shen H, Chen Y et al. (2009) Identification of N-glycosylation sites on secreted proteins of human hepatocellular carcinoma cells with a complementary proteomics approach. *J Proteome Res* 8: 662-672. doi:10.1021/pr800826u. PubMed: 19196183.
- Pan S, Chen R, Aebersold R, Brentnall TA (2011) Mass spectrometry based glycoproteomics—from a proteomics perspective. *Mol Cell Proteomics* 10: 003251 20736408.

13. Zielinska DF, Gnad F, Wiśniewski JR, Mann M (2010) Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell* 141: 897-907. doi:10.1016/j.cell.2010.04.012. PubMed: 20510933.
14. Drake PM, Schilling B, Niles RK, Braten M, Johansen E et al. (2011) A lectin affinity workflow targeting glycosite-specific, cancer-related carbohydrate structures in trypsin-digested human plasma. *Anal Biochem* 408: 71-85. doi:10.1016/j.ab.2010.08.010. PubMed: 20705048.
15. Calvano CD, Zamboni CG, Jensen ON (2008) Assessment of lectin and HILIC based enrichment protocols for characterization of serum glycoproteins by mass spectrometry. *J Proteomics* 71: 304-317. doi:10.1016/j.jprot.2008.06.013. PubMed: 18638581.
16. Zhang H, Li XJ, Martin DB, Aebersold R (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol* 21: 660-666. doi:10.1038/nbt827. PubMed: 12754519.
17. Wang L, Aryal UK, Dai Z, Mason AC, Monroe ME et al. (2012) Mapping N-linked glycosylation sites in the secretome and whole cells of *Aspergillus niger* using hydrazide chemistry and mass spectrometry. *J Proteome Res* 11: 143-156. doi:10.1021/pr200916k. PubMed: 22136231.
18. Shen C, Yu Y, Li H, Yan G, Liu M et al. (2012) Global profiling of proteolytically modified proteins in human metastatic hepatocellular carcinoma cell lines reveals CAPN2 centered network. *Proteomics* 12: 1917-1927. doi:10.1002/pmic.201200027. PubMed: 22623320.
19. Tian Y, Zhou Y, Elliott S, Aebersold R, Zhang H (2007) Solid-phase extraction of N-linked glycopeptides. *Nat Protoc* 2: 334-339. doi:10.1038/nprot.2007.42. PubMed: 17406594.
20. Zhu W, Smith JW, Huang CM (2010) Mass spectrometry-based label-free quantitative proteomics. *J Biomed Biotechnol* 2010: 840518.
21. Sudhir PR, Chen CH, Pavana Kumari M, Wang MJ, Tsou CC et al. (2012) Label-free quantitative proteomics and N-glycoproteomics analysis of KRAS-activated human bronchial epithelial cells. *Mol Cell Proteomics* 11: 901-915. doi:10.1074/mcp.M112.020875. PubMed: 22761399.
22. Cao J, Shen C, Zhang J, Yao J, Shen H et al. (2011) Comparison of alternative extraction methods for secretome profiling in human hepatocellular carcinoma cells. *Sci China Life Sci* 54: 34-38. doi:10.1007/s11427-010-4122-1. PubMed: 21253868.
23. Chevallet M, Diemer H, Van Dorsselaar A, Villiers C, Rabilloud T (2007) Toward a better analysis of secreted proteins: the example of the myeloid cells secretome. *Proteomics* 7: 1757-1770. doi:10.1002/pmic.200601024. PubMed: 17464941.
24. Zielinska DF, Gnad F, Schropp K, Wiśniewski JR, Mann M (2012) Mapping N-glycosylation sites across seven evolutionarily distant species reveals a divergent substrate proteome despite a common core machinery. *Mol Cell* 46: 542-548. doi:10.1016/j.molcel.2012.04.031. PubMed: 22633491.
25. Mathivanan S, Ji H, Simpson RJ (2010) Exosomes: extracellular organelles important in intercellular communication. *J Proteomics* 73: 1907-1920. doi:10.1016/j.jprot.2010.06.006. PubMed: 20601276.
26. Shen Q, Fan J, Yang XR, Tan Y, Zhao W et al. (2012) Serum DKK1 as a protein biomarker for the diagnosis of hepatocellular carcinoma: a large-scale, multicentre study. *Lancet Oncol* 13: 817-826. doi:10.1016/S1470-2045(12)70233-4. PubMed: 22738799.
27. Osborne JD, Flatow J, Holko M, Lin SM, Kibbe WA et al. (2009) Annotating the human genome with Disease Ontology. *BMC Genomics* 10 Suppl 1: S6. doi:10.1186/1471-2164-10-S3-S6. PubMed: 19594883.
28. Chang YH, Lee SH, Liao IC, Huang SH, Cheng HC et al. (2012) Secretomic analysis identifies alpha-1 antitrypsin (A1AT) as a required protein in cancer cell migration, invasion, and pericellular fibronectin assembly for facilitating lung colonization of lung adenocarcinoma cells. *Mol Cell Proteomics* 11: 1320-1339. doi:10.1074/mcp.M112.017384. PubMed: 22896658.
29. Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44-57. PubMed: 19131956.
30. Mantovani A (2009) Cancer: Inflammation metastasis. *Nature* 457: 36-37. doi:10.1038/457036a. PubMed: 19122629.
31. Turner N, Pearson A, Sharpe R, Lambros M, Geyer F et al. (2010) FGFR1 amplification drives endocrine therapy resistance and is a therapeutic target in breast cancer. *Cancer Res* 70: 2085-2094. doi:10.1158/0008-5472.CAN-09-3746. PubMed: 20179196.
32. Zhang XA, Kazarov AR, Yang X, Bontrager AL, Stipp CS et al. (2002) Function of the tetraspanin CD151-alpha6beta1 integrin complex during cellular morphogenesis. *Mol Biol Cell* 13: 1-11. doi:10.1091/mbc.01-10-0481. PubMed: 11809818.
33. Swiercz JM, Wozniak T, Offermanns S (2008) ErbB-2 and met reciprocally regulate cellular signaling via plexin-B1. *J Biol Chem* 283: 1893-1901. PubMed: 18025083.
34. Zhou D, Conrad C, Xia F, Park JS, Payer B et al. (2009) Mst1 and Mst2 maintain hepatocyte quiescence and suppress hepatocellular carcinoma development through inactivation of the Yap1 oncogene. *Cancer Cell* 16: 425-438. doi:10.1016/j.ccr.2009.09.026. PubMed: 19878874.
35. Saueremann U, Meyermann R, Schluesener HJ (1992) Cloning of a novel TGF-beta related cytokine, the vgr, from rat brain: cloning of and comparison to homologous human cytokines. *J Neurosci Res* 33: 142-147. doi:10.1002/jnr.490330118. PubMed: 1453478.
36. Diez-Iltza I, Sánchez LM, Allende MT, Vizoso F, Ruibal A et al. (1993) Zn-alpha 2-glycoprotein levels in breast cancer cytosols and correlation with clinical, histological and biochemical parameters. *Eur J Cancer* 29A: 1256-1260. PubMed: 8343263.
37. Hassan MI, Bilgrami S, Kumar V, Singh N, Yadav S et al. (2008) Crystal structure of the novel complex formed between zinc alpha2-glycoprotein (ZAG) and prolactin-inducible protein (PIP) from human seminal plasma. *J Mol Biol* 384: 663-672. doi:10.1016/j.jmb.2008.09.072. PubMed: 18930737.
38. Koli K, Keski-Oja J (1995) 1,25-Dihydroxyvitamin D3 enhances the expression of transforming growth factor beta 1 and its latent form binding protein in cultured breast carcinoma cells. *Cancer Res* 55: 1540-1546. PubMed: 7882362.
39. Koli K, Saharinen J, Hyytiäinen M, Penttinen C, Keski-Oja J (2001) Latency, activation, and binding proteins of TGF-beta. *Microsc Res Tech* 52: 354-362. doi:10.1002/1097-0029(20010215)52:4. PubMed: 11170294.
40. Apte SS (2004) A disintegrin-like and metalloprotease (repolysin type) with thrombospondin type 1 motifs: the ADAMTS family. *Int J Biochem Cell Biol* 36: 981-985. doi:10.1016/j.biocel.2004.01.014. PubMed: 15094112.
41. Hill KE, Lloyd RS, Burk RF (1993) Conserved nucleotide sequences in the open reading frame and 3' untranslated region of selenoprotein P mRNA. *Proc Natl Acad Sci U S A* 90: 537-541. doi:10.1073/pnas.90.2.537. PubMed: 8421687.
42. Yang L, Froio RM, Sciuto TE, Dvorak AM, Alon R et al. (2005) ICAM-1 regulates neutrophil adhesion and transcellular migration of TNF-alpha-activated vascular endothelium under flow. *Blood* 106: 584-592. doi:10.1182/blood-2004-12-4942. PubMed: 15811956.
43. Shimizu K, Chiba S, Saito T, Kumano K, Takahashi T et al. (2001) Manic fringe and lunatic fringe modify different sites of the Notch2 extracellular region, resulting in different signaling modulation. *J Biol Chem* 276: 25753-25758. doi:10.1074/jbc.M103473200. PubMed: 11346656.
44. Kikkawa Y, Moulson CL, Virtanen I, Miner JH (2002) Identification of the binding site for the Lutheran group glycoprotein on laminin alpha 5 through expression of chimeric laminin chains in vivo. *J Biol Chem* 277: 44864-44869. doi:10.1074/jbc.M208731200. PubMed: 12244066.
45. Tang ZY, Sun FX, Tian J, Ye SL, Liu YK et al. (2001) Metastatic human hepatocellular carcinoma models in nude mice and cell line with metastatic potential. *World J Gastroenterol* 7: 597-601. PubMed: 11819839.
46. Tian J, Tang ZY, Ye SL, Liu YK, Lin ZY et al. (1999) New human hepatocellular carcinoma (HCC) cell line with highly metastatic potential (MHCC97) and its expressions of the factors associated with metastasis. *Br J Cancer* 81: 814-821. doi:10.1038/sj.bjc.6690769. PubMed: 10555751.
47. Li Y, Tang ZY, Ye SL, Liu YK, Chen J et al. (2001) Establishment of cell clones with different metastatic potential from the metastatic hepatocellular carcinoma cell line MHCC97. *World J Gastroenterol* 7: 630-636. PubMed: 11819844.
48. Ji XN, Ye SL, Li Y, Tian B, Chen J et al. (2003) Contributions of lung tissue extracts to invasion and migration of human hepatocellular carcinoma cells with various metastatic potentials. *J Cancer Res Clin Oncol* 129: 556-564. doi:10.1007/s00432-003-0475-1. PubMed: 12942314.
49. Li Y, Tian B, Yang J, Zhao L, Wu X et al. (2004) Stepwise metastatic human hepatocellular carcinoma cell model system with multiple metastatic potentials established through consecutive in vivo selection and studies on metastatic characteristics. *J Cancer Res Clin Oncol* 130: 460-468. PubMed: 15146329.
50. Tang ZY, Ye SL, Liu YK, Qin LX, Sun HC et al. (2004) A decade's studies on metastasis of hepatocellular carcinoma. *J Cancer Res Clin Oncol* 130: 187-196. doi:10.1007/s00432-003-0511-1. PubMed: 14685850.

51. Wiśniewski JR, Zougman A, Nagaraj N, Mann M (2009) Universal sample preparation method for proteome analysis. *Nat Methods* 6: 359-362. doi:10.1038/nmeth.1322. PubMed: 19377485.
52. Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V et al. (2006) TM4 microarray software suite. *Methods Enzymol* 411: 134-193. doi: 10.1016/S0076-6879(06)11009-5. PubMed: 16939790.
53. Zhang F, Xu X, Zhou B, He Z, Zhai Q (2011) Gene expression profile change and associated physiological and pathological effects in mouse liver induced by fasting and refeeding. *PLOS ONE* 6: e27553. doi: 10.1371/journal.pone.0027553. PubMed: 22096593.
54. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25: 25-29. doi:10.1038/75556. PubMed: 10802651.