Proteomics Reveal the Landscape of Gender Disparity in Human Hepatocellular Carcinoma

Zhigang Luo
Third People's Hospital of Sichuan Province
Jian He
Third People's Hospital of Sichuan Province
Shuhai Lin
Xiamen University
Feng Shou
the People's Hospital of Jianyang City
Jingjun Zhang
the People's Hospital of Jianyang City
Lei Chu
the People's Hospital of Jianyang City
Xingcai Zhang (xingcai@mit.edu)
Massachusetts Institute of Technology

Research Article

Keywords: hepatocellular carcinoma, GSTM1, IL-6, iTRAQ, Cytochrome P450

Posted Date: April 6th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1504785/v1

Abstract

The incidence of hepatocellular carcinoma (HCC) in males is about 2 to 3 times higher than that of females. In order to figure out what cause this disparity, we performed isobaric tags for HCC tumor tissues and para-tumor tissues, via relative and absolute quantitation-based quantitative proteomics to obtain data of differential expression proteins (DEPs). The results show that the expressions of glutathione S-transferase Mu 1(GSTM1) was not significantly different between male and female in para-tumor tissues, while the expression of GSTM1 in tumor tissues of female is just 5.6% of that of male, together with further investigation, indicating that the pathway of xenobiotics metabolism by cytochrome P450 may play an essential role in the gender disparity epigenetic on HCC. Furthermore, potential association between GSTM1 and human interleukin 6 (IL-6) was found. Our findings suggest that low expression of GSTM1 increases the vulnerability on HCC in male, which may be the very reason why the HCC incidence in males is higher than females, and GSTM1 may affect occurrence and development of hepatocellular cancer via xenobiotic detoxification and regulation on Interleukin-6 (IL-6) expression, paving a way for HCC theranostics.

Introduction

Hepatocellular carcinoma (HCC) is the most common adult essential liver cancer, and it is one of the highest cancers in the world (4.7%). Its mortality rate (8.3%) is listed as the fourth highest in cancer-related deaths worldwide, with characteristics of poor early detection and prognosis, and with a significant threat to human health [1]. Also, there are incidences that other cancers can cause synchronous hepatic metastases[2],[3].

According to statistical data of morbidity and mortality of liver cancer, there is a significant gender difference, in that males have the fifth highest incidence rate (6.3%) and the second highest mortality rate (10.2%), which is much higher than that of females [1]. Epidemiological studies have also shown that the incidence and mortality of male hepatocellular carcinoma is higher than that of females, and the differences between two genders are significant [4],[5],[6],[7]. Sex hormones are believed to be one of the important reasons why liver cancer incidence and mortality are different between two genders and they also play different roles in the process of liver cancer development. It is known that androgen stimulates liver cancer production, while estrogen protects human being from liver cancer onset [8],[9], [10] .

Isobaric Tags for Relative and Absolute Quantitation (iTRAQ) was applied to analyze our HCC samples. The data shows significantly down-expression on GSTM1 in male patients while up-expression in female patients. GSTM1 is an important member of the glutathione thiotransferase Mu family and is responsible for cell damage and cancer resistance. Gene ontology (GO) function and pathway enrichment network analysis of differential proteins were then performed sequentially. Pathway enrichment analysis revealed that the cytochrome P450- related pathway is the most significant enrichment pathway for differential proteins, and GSTM1 is involved in many metabolic processes in cytochrome P450-related pathways. Therefore, GSTM1 was selected as the key candidate gene for causing gender differences in

hepatocellular carcinoma. The Cancer Genome Atlas (TCGA) is a large-scale cancer genome project which provides researchers with multi- dimensional maps of the key genomic changes and clinicopathological information in 33 types of cancer (http://cancergenome.nih.gov/)[11]. In this study, we first identified the common differentially expressed genes (DEGs) from TCGA LIHC RNA-seq data for proofing GSTM1 expression differences and effects on hepatocellular carcinoma. Then GSTM1 was identified based on TCGA LIHC RNA-seq data and demonstrated good performance for predicting 5-year overall survival. This signature was successfully validated in this independent cohort of patients.

We used the Genemia plug-in of cytoscape software to predict the relationship GSTM1 in many hot-spots signaling pathways, and also found that GSTM1 has a co-expression relationship with the IL-6 pathway. Therefore, we speculate that GSTM1 may become a new HCC signaling molecule, paving a way for HCC theranostics.[12–19] The further GSTM1 and the IL6 pathway regulation, together with nanomedicine and other therapies, may have high potential of overcoming HCC.

Materials And Methods

Tissue sample

Cancer tissues and paracancerous tissues of 60 patients with HCC, including48 male cases and 12 female cases, were from Changhai Hospital, the affiliated hospital of Naval Medical University in Shanghai. (Table 1)

Table 1
Patients Information and the Protein Expression

	Male(n)	Female(n)	*p*
Age			
≤ 40	6	2	
41–60	28	3	0.29
≥ 60	14	7	
Differentiation			
High	4	0	
Middle	41	11	0.385
Low	3	1	
Size			
≤ 3cm	8	2	
3-6cm	15	7	0.156
≥ 6cm	25	3	
Stage			
□/□	25	8	0.081
□/□	23	4	
Lymph node metastasis			
Yes	3	1	0.795
None	45	11	
Hepatitis B history	15	3	0.024
Smoke history	20	0	0.006
CYP8B1	8	3	0.073
MTHFD2	36	5	0.034
NNMT	27	4	0.085
SHMT2	37	2	0.167
SLC27A5	10	3	0.102

Protein preparation

Hepatocellular carcinoma cells from 30 patients, 60 samples totally, were suspended in lysis buffer (7 M Urea, 2 M Thiourea, 4% CHAPS, 40 mM Tris-HCL, pH 8.5, 1mM PMSF, 2mM EDTA) and sonicated in ice. The proteins were reduced with 10 mM DTT (final concentration) at 56°C for 1 hour and then alkylated by 55 mM IAM (final concentration) in the darkroom for 1 hour. The reduced and alkylated protein mixtures were precipitated by adding 4 volume times of chilled acetone at -20°C overnight. After centrifugation at 4°C, 30 000g, the pellet was dissolved in 0.5 M TEAB (Applied Biosystems, Milan, Italy) and sonicated in ice. After centrifuging at 30,000g at 4°C, an aliquot of the supernatant was taken for determination of protein concentration by Bradford [20]. The proteins in the supernatant were kept at -80°C for further analysis.

iTRAQ Labeling and SCX fractionation

Total protein (100μg) was taken out of each sample solution and then the protein was digested with Trypsin Gold (Promega, Madison, WI, USA) with the ratio of protein: trypsin = 30:1 at 37°C for 16 hours. After trypsin digestion, peptides were dried by vacuum centrifugation. peptides were reconstituted in 0.5M TEAB and processed according to the manufacture's protocol for 8-plex iTRAQ reagent (Applied Biosystems). Briefly, one unit of iTRAQ reagent was thawed and reconstituted in 24 μL isopropanol. Samples were labeled with the iTRAQ tags as follow format. The peptides were labeled with the isobaric tags, incubated at room temperature for 2h. The labeled peptide mixtures 28/30were then pooled and dried by vacuum centrifugation. SCX chromatography was performed with a LC-20AB HPLC Pump system (Shimadzu, Kyoto, Japan). The iTRAQ labeled peptide mixtures were reconstituted with 4 mL buffer A (25 mM NaH2PO4 in 25% ACN, pH 2.7) and loaded onto a 4.6×250 mm Ultremex SCX column containing 5-μm particles (Phenomenex). The peptides were eluted at a flow rate of 1mL/min with a gradient of buffer A for 10 min, 5–60% buffer B (25mM NaH2PO4, 1M KCl in 25% ACN, pH 2.7) for 27 min, 60–100% buffer B for 1 min. The system was then maintained at 100% buffer B for 1 min before equilibrating with buffer A for 10 min prior to the next injection. Elution was monitored by measuring the absorbance at 214 nm, and fractions were collected every 1 min. The eluted peptides were pooled into 20 fractions, desalted with a StrataXC18 column (Phenomenex) and vacuum dried.

LC-ESI-MS/MS analysis based on Q EXACTIVE

Each fraction was resuspended in buffer A (2% ACN, 0.1% FA) and centrifuged at 20000g for 10 min, the final concentration of peptide was about 0.5 μg/μl on average. 10μl supernatant was loaded on a LC-20AD nano-HPLC (Shimadzu, Kyoto, Japan) by the autosampler onto a 2cm C18 trap column. Then, the peptides were eluted onto a 10cm analytical C18 column (inner diameter 75 μm) packed in- house. The samples were loaded at 8 μL/min for 4min, then the 44min gradient was run at 300 nL/min starting from 2 to 35% B (98%ACN, 0.1%FA), followed by 2 min linear gradient to 80%, and maintenance at 80% B for 4

min, and finally return to 5% in 1 min. The peptides were subjected to nano-electrospray ionization followed by tandem mass spectrometry (MS/MS) in an Q EXACTIVE (Thermo Fisher Scientific, San Jose, CA) coupled online to the HPLC. Intact peptides were detected in the Orbitrap at a resolution of 70 000. Peptides were selected for MS/MS using high-energy collision dissociation (HCD) operating mode with a normalized collision energy setting of 27.0; ion fragments were detected in the Orbitrap at a resolution of 17500. A data-dependent procedure that alternated between one MS scan followed by 15 MS/MS scans was applied for the 15 most abundant precursor ions above a threshold ion count of 20000 in the MS survey scan with a following dynamic exclusion duration of 15 s. The electrospray voltage applied was 1.6 kV. Automatic gain control (AGC) was used to optimize the spectra generated by the orbitrap. The AGC target for full MS was 3e6 and 1e5 for MS2. For MS scans, the m/z scan range was 350 to 2000 Da. For MS2 scans, the m/z scan range was 100–1800.

Data Analysis

The original data files acquired from the Orbitrap were converted into MGF files using Proteome Discoverer 1.2 (PD 1.2, Thermo), [5600 ms converter] and the MGF file were searched. Proteins identification were performed by using Mascot search engine (Matrix Science, London, UK; version 2.3.02) against database containing 200,000 sequences. For protein identification, a mass tolerance of 0.05Da (10 ppm) was permitted for intact peptide masses and 0.01Da for fragmented ions, with allowance for one missed cleavage in the trypsin digests. Gln->pyro-Glu (N-term Q), Oxidation (M), Deamidated (NQ) as the potential variable modifications, and Carbamidomethyl (C), iTRAQ8plex (N-term), iTRAQ8plex (K) as fixed modifications. The charge states of peptide were set to + 2 and + 3. Specifically, an automatic decoy database search was performed in Mascot by choosing the decoy checkbox in which a random sequence of database is generated and tested for raw spectra as well as the real database. To reduce the probability of false peptide identification, only peptide at the 95% confidence interval by a Mascot probability analysis greater than “identity” were counted as identified. And each confident protein identification involves at least one unique peptide. For protein quantitation, it was required that a protein contains at least two unique peptides. The quantitative protein ratios were weighted and normalized by the median ratio in Mascot. We only used ratios with p-values < 0.05, and only fold changes of > 1.2 were considered as significant.

Function annotation methods

Functional annotations of the identified proteins from our investigation were conducted via Blast2GO program against the non-redundant protein database. Kegg database (http://www.genome.jp/kegg/) and COG database (http://www.ncbi.nlm.nih.gov/COG/) were applied to sort those identified proteins into different categories. As we all known, Gene Ontology (GO) is an international standardization of gene function classification system. it provides a set of dynamic updating controlled vocabulary to describe genes and gene products attributes in organisms, which contains 3 main ontologies context, i.e., molecular function, cellular component, biological process respectively. Cluster of Orthologous Groups of

proteins (COG) is the database for protein orthologous classification. Each protein in COG is derived from the same ancestor. KEGG PATHWAY comprises all the available pathway maps on the molecular interaction and reaction networks.

Public data source

The preprocessed level 3 RNAseq data and corresponding clinical information of LIHC were downloaded from TCGA data portal. The clinical samples from liver were selected. A total of 222 LIHC patients with detailed follow-up time were included for subsequent analysis. Except the progress (overall survival), Pearson statistics between genders and clinical pathological characteristic were analysed, such as the HBsAg, HBV DNA, drinking history, smoking history cirrhosis, tumor size, metastasis, tumor thrombus, alanine aminotransferase, aspartate aminotransferase, ALB (albumin of blood), TBIL (total bilirubin) GGT (Glutamyltransferase), independently.

Immunohistochemistry staining

The expressions of proteins in the specimens were detected by immunohistochemistry assay[13] with the m**onoclonal** or polyclonal antibody against human according to the instructions of the invitrogen. For antibody control, one set of samples was incubated with non-immune rabbit IgG (1:150) instead of primary antibody. Evaluation of proteins staining was independently performed by two experienced pathologists. The intensity of proteins immunostaining was semiquantitatively estimated according to the signal intensity and distribution. Briefly, a mean percentage of positive tumor cells was determined in at least five areas 400 magnification and assigned to one of the five following categories: 0, 5%; 15–25%; 2, 25–50%; 3, 50–75%, and 4, 75%. The intensity of immunostaining was scored as follows: 1 weak; 2 moderate and 3 intense. For tumors that showed heterogeneous staining, the predominant pattern was taken into account for scoring. The percentage of positive tumor cells and the staining intensity were multiplied to produce a weighted score for each case. Tissues with immunohistochemical scoring 4 were considered as having low expression, and scores of 5 to 12 were considered high expression. The proteins tested and the their antibody were NNMT (Nicotine N-methyltransferase, Thermofisher Scientific.co, MA5-15738), CYP8B1(Cytochrome P450 Family 8 Subfamily B Member 1, Thermofisher Scientific.co, PA5-102021), SLC27A5(solute carrier family 27 member 5, Thermofisher Scientific.co, PA5-82292), SHMT2 (Serine Hydroxymethyltransferase 2, Thermofisher Scientific.co, MA5-27057), MTHFD2 (Methylenetetrahydrofolate Dehydrogenase, Thermofisher Scientific.co, PA5- 97986).

mRNA analysis

Expression on mRNA was determined by quantitative reverse-transcription (qRT- PCR) using the LightCycler system (Roche, Mannheim, Germany) 18S was used as an endogenous control to normalize for differences in the amount of total RNA in each sample. The primers used for PCR were as follows:

gene	Primer (sense)	Primer antisense
18S	-5' GGAGTATGGTTGCAAAGCTGA-3'	5'-ATCTGTCAATCCTGTCCGTGT-3’
CYP8B1	5’-GTACACATGGACCCCGACATC-3'	5'-CAGGGTTGAGGAACCGATTG-3’
MTHFD2	5’-GATCCTGGTTGGCGAGAATCC-3'	5'–TCTGGAAGAGGCAACTGAACA-3’
NNMT	5’-TCACCTTCGACGAGGAAGCT-3'	5'-GCTCTGCAGACTTCAGACCA-3’
SHMT2	5’-TTTGCTTCCCCAGTCTGAGT – 3'	5'-TTCTCTTTGTTTTGGGCGG-3’
SLC27A5	5’-CGGGGTACCATGGGTGTCAGG CAACAGTTGGCCTTG-3'	5'-CCCAAGCTTTCAGAGCCTC CAGGTTCCCTCACA-3’

Results

Basic protein identification information

The raw mass spectral data needs to be converted to the mg format first, and then searched in the Uniprot-Human database using the Mascot search engine (Matrix Science, London, UK; version 2.3.02). A total of 4216 proteins were identified (Fig. 1) thereafter.

Differential protein basic identification information

There were four groups of samples, including female tumor tissue (shorter for FTA1), female non-tumor tissue (FNA2), male tumor tissues (MTB1), and male non-tumor tissues (MNB2). Compare the protein abundance ratio between the four groups of samples (Fig. 1E). When the difference fold is greater than 1.5 times and is statistically significant ($p < 0.05$), we consider the protein is a significant differential protein between two sample groups. There were 154 up-regulated proteins and 153 down-regulated proteins between male and female tumor samples, 232 up-regulated and 310 down-regulated proteins between male tumor and non- tumor tissues, 112 up-regulated and 160 down-regulated proteins between female tumor and non-tumor (Fig. 1F). Venn diagram shows the overlaps of differential proteins between samples (Fig. 2).

GO enrichment and Pathway enrichment

Pathway enrichment analysis of differential proteins between male tumor samples and female tumor samples (MTB1 and FTA1) showed that the most relevant pathways for differential proteins were metabolism of xenobiotics by cytochrome P450, drug metabolism cytochrome P450, retinol metabolism, starch-sucrose metabolism, and ascorbic acid-aldehyde acid metabolisms.

The results of GO enrichment analysis (Fig. 3A) showed that the main enrichment biological processes were single-organism metabolic process, carboxylic acid metabolic process, oxoacid metabolic process, small molecule metabolic process and organic acid metabolic process. The main enrichment cellular component are extracellular region, extracellular matrix, basement membrane, extracellular matrix part and endoplasmic reticulum part. The enrichment molecular function are oxidoreductase activity, retinoic acid binding, glucuronosyltransferase activity, retinoid binding and isoprenoid binding.

There are 36 common proteins between MTB1 and FTA1, MTB2 and MNB1, FTA1 and FNA2 (Fig. 2). We are interested in proteins that are differentially expressed between male and female, so these differential proteins are the main targets of screening. Cluster heat map (Fig. 4A) shows significant differences in the expression of GSTM1 and GSTM4 between different genders. GSTM1 and GSTM4 is significantly down-regulated in male and up-regulated in female.

The top 10 proteins were selected as candidate protein of interest, according to their difference in expression between male and female (Table 2). According to the KEGG database, 5 of the 10 dysregulated proteins are related to metabolism of xenobiotics by cytochrome P450, drug metabolism-cytochrome P450 pathway, and the remaining dysregulated proteins are associated with oxidative phosphorylation and bile acid metabolism.

Table 2
List of proteins with expression-level change the most significantly.

Gene Name	Protein Name	Accession	%Cov	%change	*p*
GSTM1	Glutathione S-transferase Mu 1	P09488	62.4	-94.4	*
GSTM4	Glutathione S-transferase Mu 4	Q03013	49.5	-79.7	*
DECR2	Peroxisomal 2,4-dienoyl-CoA reductase	Q4VXZ8	41.8	-68.2	*
KAD4	Adenylate kinase 4, mitochondrial	P27144	24.7	-45.1	*
PGDH	Isoform 5 of 15-hydroxyprostaglandin dehydrogenase [NAD(+)]	P15428-5	29.3	-55.5	*
CYP8B1	7-alpha-hydroxycholest-4-en-3-one 12- alpha-hydroxylase	Q9UNU6	14.4	-72.5	*
UGT2B4	UDP-glucuronosyltransferase 2B4	P06133	17.4	-57.5	*
SULT2A1	Bile salt sulfotransferase	Q06520	52.3	-51.8	*
UGT1A9	Isoform 2 of UDP- glucuronosyltransferase 1–9	O60656-2	31.3	-75.7	*
ACSM5	Acyl-coenzyme A synthetase ACSM5, mitochondrial	Q6NUN0	9.8	-57.6	*

Ten proteins that showed the most significant expression-level change are listed together with gene names, protein names and ExPASy accession numbers. % Cov: percentage of coverage of peptide on identified proteins; % change: percentage change of protein-expression level in J20 mice as compared to that in wild-type mice. * means $P < 0.05$, that is, significant changes in protein- expression level.

Kaplan-Meier survival analysis and GeneMANIA forecast

To verify the low expression of GSTM1 in male HCC and its effects on HCC, we used data from TCGA for Kaplan-Meier survival analysis of cases with different expression of GSTM1. The results show survival analysis showed that there was a significant difference in survival time between the different expression groups of GSTM1 ($p < 0.05$) (Fig. 5). The survival time of the low expression group was lower than that of the high expression group. This indicates that the low expression of GSTM1 will make poor prognosis of HCC and explore GSTM1 has a positive effect on improving the prognosis of HCC.

Using the GeneMANIA (www.genemania.org/) plugin of Cytoscape software to predict the relationship between GSTM1 and IL6, IL6R, PRKAA1, and found there were co-expression relationships between them (Fig. 4B). From the RNA-seq data downloaded by TCGA, the mean value of GSTM1 expression in male patients with liver cancer was significantly lower than that in women (P = 0.015437). In addition, studies have shown that the gene copy number variation of GSTM1 is also significantly correlated with gender [9]. In the data from the TCGA database, it was found that interleukin-6 (IL6), interleukin-6 receptor (IL6R) and

AMP- activated protein kinase alpha 1 (PRKAA1 or AMPKa1) also have gender differences, and studies have been conducted. These proteins were shown to be associated with liver cancer (Fig. 4). Among them, IL6 and IL6R are highly expressed in male HCC and PRKAA1 is highly expressed in female HCC.

Using GeneMANIA plugin of Cytoscape software to predict the relationship of genes. We found there were co-expression relationships between GSTM1 and IL6, IL6R, PRKAA1.

Immunohistochemistry staining

The proteins expression on CYP8B1, MTHFD2, NNMT, SHMT2, SLC27A5 are different. MTHFD2, NNMT, SHMT2 are higher in tumors than in paracancerous while CYP8B1, SLC27A5 are contrary (Fig. 6), which are higher in paracancerous than in tumor. The proteins expression on CYP8B1 are consisted with the RNA. (Table 2)

Discussion

HCC is constantly threatening human health, and its significant difference in incidence between two genders is worth studying. Genomics,[21],[22],[23] and epigenetics including DNA methylation change,[24],[25] were well investigated in the past few years. Sex-dependent DNA methylation events exist in HCC [4]. Proteomics and metabolomics became hot topics of research spot in this area in order to learn the mechanism of the liver cancer, especially between the genders. In recent years, mass spectrometry just become a very good method for differential expression of proteins. And the proteomic data and comparative analysis presented here offer crucial clues for elucidating the mechanisms that underlie the male prevalence in HCC,[26], [27].

We applied iTRAQ method to analyze differential proteins in HCC between different genders in this study and used the data from the public database TCGA to verify conclusions and find more potential differences. A total of 4,216 proteins were identified by the iTRAQ method, including 814 differential proteins. GO enrichment analysis of differential proteins shows that they were mainly involved in the metabolism of ketoacids, carboxylic acids and organic acids, and the binding functions of retinoic acid and retinol. Pathway enrichment analysis was mainly involved in the metabolism of xenobiotics by cytochrome P450, drug metabolism- cytochrome P450, retinol metabolism. This indicates that differential proteins may affect the occurrence, metastasis and invasion of HCC by regulating energy metabolism and metabolism of important substances such as retinoic acid.

The top 10 differential proteins with the largest difference between genders were selected, namely GSTM1, GSTM4, DECR2, KAD4, PGDH, CYP8B1, UGT2B4, SULT2A1, UGT1A9 and ACSM5. Among them, GSTM1, GSTM4, UGT2B4, SULT2A1 and UGT1A9 are all involved in the metabolism of xenobiotics by cytochrome P450, suggesting that this pathway may be a key pathway for the gender difference in the incidence of HCC. We tested expression of proteins and RNAs of CYP8B1, MTHFD2, NNMT, SHMT2, SLC27A5, related to cytochrome P450, in HCC tumors and matched paracancerous tissues. Proteins of

MTHFD2, NNMT, SHMT2 were upregulated while the CYP8B1 (proteins and RNA) was downregulated. Interestingly, bile acid metabolism plays an important role in the development of liver cancer. Among these proteins, CYP8B1 and SULT2A1 are key enzymes in the metabolism of bile acids [10], suggesting that bile acid metabolism has potential research value in the gender difference of HCC.

The gender difference in the expression of GSTM1 between tumor tissues was statistically significant. The expression in male tissues was only 5.6% in female tissues (Table 2), as confirmed by TCGA database data. Previous researches reveal loss expression or low expression of GSTM1 will increase the susceptibility of liver cancer [28], which can explain the reason why the incidence of male liver cancer is much higher than that of female. Kaplan-Meier survival analysis of HCC cases showed that patients with low expression of GSTM1 had a poor prognosis and a five-year survival rate was significantly reduced. On this basis, we believe that GSTM1 could provide new solution for the prevention, treatment and prognosis on essential liver cancer. This is confirmed by many researches with similar results [29],[30].

On the contrary to the lower expression of GSTM1 in males, IL6 and IL6R express higher in males than females. It is well known that IL6 regulates the proliferation, invasion and metastasis of liver cancer by regulating the AKT pathway and the VEGF pathway [12]. Results predicted via GeneMANIA reveals that there are co- expression relationships between GSTM1 and IL6, and also there is a significant gender difference between the correlation factors GSTM5 and FKBP5. In colorectal carcinogenesis, these is an causal connection between the oxidative stress response and inflammatory response and some study show gender disparity in the incidence and prognosis of HCC through sex biased molecular signature, IL6, which serve as a female-specific index for prediction and evaluation of overall survival in liver cancer patients[31].Therefore, GSTM1 may affect the occurrence and development of liver cancer through regulation of the IL6 pathway, which is worth further research.

Declarations

Competing interests:

The authors have no conflicts of interest to declare.

Authors' contributions:

Dr. Zhigang Luo collected the tissue samples, performed pathological diagnosis, and conducted LC-ESI-MS/MS analysis and immunohistochemistry staining. Dr. Jian He performed protein preparation and mRNA analysis. Prof. Shuhai Lin performed LC-ESI-MS/MS analysis and data analysis. Dr. Feng Shou and Dr. Jingjun Zhang collected patients' information and data analysis including the survival analysis. Dr. Lei Chu and Prof. Xingcai Zhang were contributing on critical review, commentary and revision on the paper.

Funding:

National Natural Science Foundation of China (NSFC-21377106)

Data availability statement:

The data that support the findings of this findings of this study are available on request from the corresponding author. The data are not publicly available due to publication privacy for the moment.

Ethic approval: The study was filed and approved by the ethics committee of the People's Hospital of Jianyang City.

Consent to Publish

The Author confirms:

• that the work described has not been published before;

• that it is not under consideration for publication elsewhere;

• that its publication has been approved by all co-authors;

•that its publication has been approved explicitly by the responsible authorities at the institution where the work is carried out.

Acknowledgements:

This work was supported by the National Natural Science Foundation of China (NSFC-21377106).

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021. 71(3): 209–249.
2. Xiao W, Zheng S, Yang A, et al. Breast cancer subtypes and the risk of distant metastasis at initial diagnosis: a population-based study. Cancer Manag Res. 2018. 10: 5329–5338.
3. Xiao W, Zheng S, Yang A, et al. Incidence and Survival Outcomes of Breast Cancer with Synchronous Hepatic Metastases: A Population-Based Study. J Cancer. 2018. 9(23): 4306–4313.
4. Ye W, Siwko S, Tsai R. Sex and Race-Related DNA Methylation Changes in Hepatocellular Carcinoma. Int J Mol Sci. 2021. 22(8).
5. Zhang T, Xu J, Ye L, et al. Age, Gender and Geographic Differences in Global Health Burden of Cirrhosis and Liver Cancer due to Nonalcoholic Steatohepatitis. J Cancer. 2021. 12(10): 2855–2865.
6. Ajayi F, Jan J, Singal AG, Rich NE. Racial and Sex Disparities in Hepatocellular Carcinoma in the USA. Curr Hepatol Rep. 2020. 19(4): 462–469.
7. Kim SY, Song HK, Lee SK, et al. Sex-Biased Molecular Signature for Overall Survival of Liver Cancer Patients. Biomol Ther (Seoul). 2020. 28(6): 491–502.

8. Zheng B, Zhu YJ, Wang HY, Chen L. Gender disparity in hepatocellular carcinoma (HCC): multiple underlying mechanisms. Sci China Life Sci. 2017. 60(6): 575–584.
9. Li Y, Xu A, Jia S, Huang J. Recent advances in the molecular mechanism of sex disparity in hepatocellular carcinoma. Oncol Lett. 2019. 17(5): 4222–4228.
10. Wu MH, Ma WL, Hsu CL, et al. Androgen receptor promotes hepatitis B virus-induced hepatocarcinogenesis through modulation of hepatitis B virus RNA transcription. Sci Transl Med. 2010. 2(32): 32ra35.
11. Lee JS. Exploring cancer genomic data from the cancer genome atlas project. BMB Rep. 2016. 49(11): 607–611.
12. Gao D, Chen T, Chen S, et al. Targeting Hypoxic Tumors with Hybrid Nanobullets for Oxygen-Independent Synergistic Photothermal and Thermodynamic Therapy. Nanomicro Lett. 2021, 13(1), 99.
13. Yang Z, Gao D, Guo X, et al. Fighting immune cold and reprogramming immunosuppressive tumor microenvironment with red blood cell membrane-camouflaged nanobullets. ACS Nano. 2020, 14 (12), 17442.
14. Wei D, Yu Y, Zhang X, et al. Breaking the Intracellular Redox Balance with Diselenium Nanoparticles for Maximizing Chemotherapy Efficacy on Patient-Derived Xenograft Models. ACS Nano. 2020, 14 (12), 16984.
15. Parekh G, Shi Y, Zheng J, Zhang X, Leporatti S. Nano-carriers for targeted delivery and biomedical imaging enhancement. Ther Deliv. 2018. 9(6): 451–468.
16. Liu C, Sun S, Feng Q, et al. Arsenene Nanodots with Selective Killing Effects and their Low-Dose Combination with ß-Elemene for Cancer Therapy. Adv Mater. 2021. 33(37): e2102054.
17. Ji X, Ge L, Liu C, et al. Capturing functional two-dimensional nanosheets from sandwich-structure vermiculite for cancer theranostics. Nat Commun. 2021. 12(1): 1124.
18. Kong N, Zhang H, Feng C, et al. Arsenene-mediated multiple independently targeted reactive oxygen species burst for cancer therapy. Nat Commun. 2021. 12(1): 4777.
19. Wang Y, Gao D, Liu Y, et al. Immunogenic-cell-killing and immunosuppression-inhibiting nanomedicine. Bioact Mater. 2021. 6(6): 1513–1527.
20. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976. 72: 248–54.
21. Nia A, Dhanasekaran R. Genomic Landscape of HCC. Curr Hepatol Rep. 2020. 19(4): 448–461.
22. Meier T, Timm M, Montani M, Wilkens L. Gene networks and transcriptional regulators associated with liver cancer development and progression. BMC Med Genomics. 2021. 14(1): 41.
23. Yim SY, Lee JS. Genomic Perspective on Mouse Liver Cancer Models. Cancers (Basel). 2019. 11(11).
24. Cerapio JP, Marchio A, Cano L, et al. Global DNA hypermethylation pattern and unique gene expression signature in liver cancer from patients with Indigenous American ancestry. Oncotarget. 2021. 12(5): 475–492.

25. Toh TB, Lim JJ, Chow EK. Epigenetics of hepatocellular carcinoma. Clin Transl Med. 2019. 8(1): 13.
26. Li H, Rong Z, Wang H, et al. Proteomic analysis revealed common, unique and systemic signatures in gender-dependent hepatocarcinogenesis. Biol Sex Differ. 2020. 11(1): 46.
27. Rong Z, Fan T, Li H, et al. Differential Proteomic Analysis of Gender- dependent Hepatic Tumorigenesis in Hras12V Transgenic Mice. Mol Cell Proteomics. 2017. 16(8): 1475–1490.
28. Li S, Xue F, Zheng Y, et al. GSTM1 and GSTT1 null genotype increase the risk of hepatocellular carcinoma: evidence based on 46 studies. Cancer Cell Int. 2019. 19: 76.
29. Chen J, Ma L, Peng NF, Wang SJ, Li LQ. A meta-analysis of the relationship between glutathione S-transferases gene polymorphism and hepatocellular carcinoma in Asian population. Mol Biol Rep. 2012. 39(12): 10383–93.
30. Liu K, Zhang L, Lin X, et al. Association of GST genetic polymorphisms with the susceptibility to hepatocellular carcinoma (HCC) in Chinese population evaluated by an updated systematic meta-analysis. PLoS One. 2013. 8(2): e57043.
31. Caramujo-Balseiro S, Faro C, Carvalho L. Metabolic pathways in sporadic colorectal carcinogenesis: A new proposal. Med Hypotheses. 2021. 148: 110512.

Figures

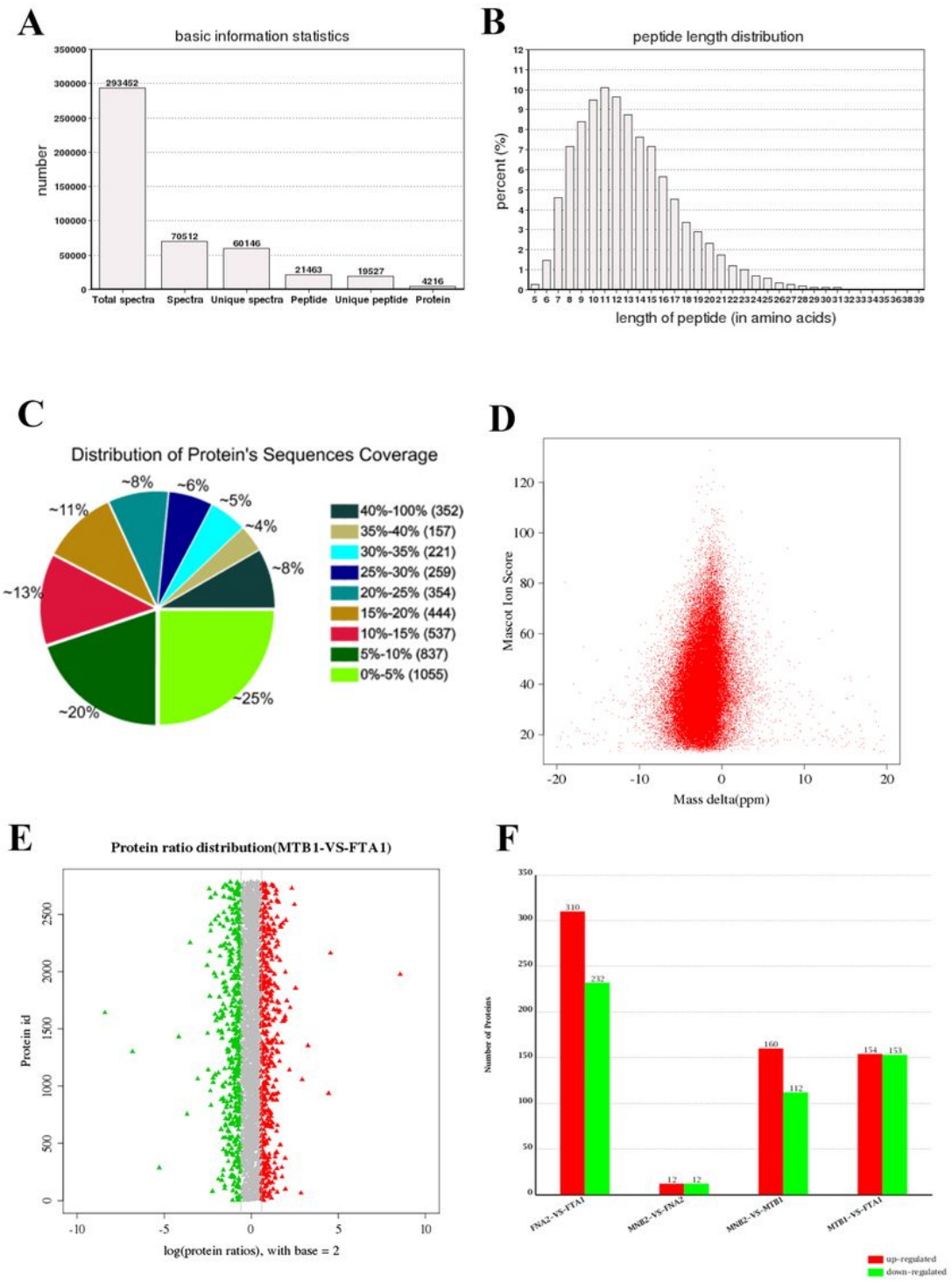

Figure 1

Identification Basic Information and Ratio Distribution and Differential Protein Quantity Information.

Legend: Peptide length distribution(B). Distribution of Protein's Sequence Coverage(C). The error distribution between the real value and the theoretical value of the relative molecular weight of all matched peptides(D). The distribution of protein abundance ratios between samples. The abscissa is the

protein abundance ratio information processed by the logarithm of the base 2, that is, the difference multiple of the protein. The red dot in the figure indicates an up-regulated protein with a fold difference greater than 1.5 folds. the green dot indicates a down-regulated protein with a fold difference greater than 1.5 folds (E). There were 154 up-regulated proteins and 153 down-regulated proteins between male and female tumor samples, 232 up-regulated and 310 down-regulated proteins between male tumor and non-tumor tissues, 112 up-regulated and 160 down-regulated proteins between female tumor and non-tumor (F).

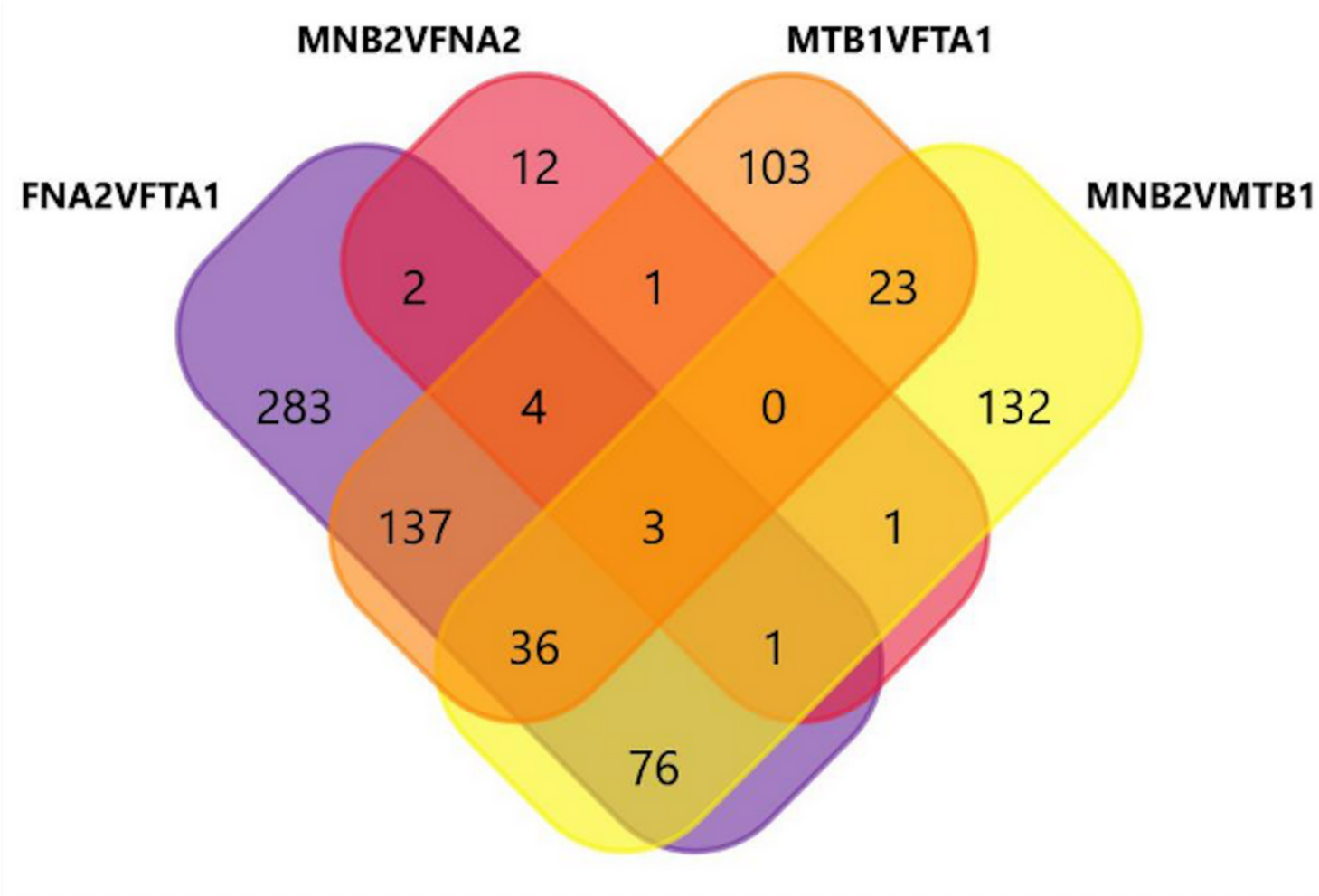

Figure 2

Venn Diagram on the Common Differentially Expressed Genes among Samples

Legend: The common differentially expressed genes (DEGs) from TCGA LIHC RNA-seq data for proofing GSTM1 expression differences and effects on hepatocellular carcinoma. Venn diagram shows the overlap of differential proteins between four groups of samples, including female tumor tissue (FTA1), female non-tumor tissue (FNA2), male tumor tissues (MTB1), and male non-tumor tissues (MNB2): MTB1VFTA1, MNB2VMTB1, MNB2VFNA2, FNA2VFTA1.

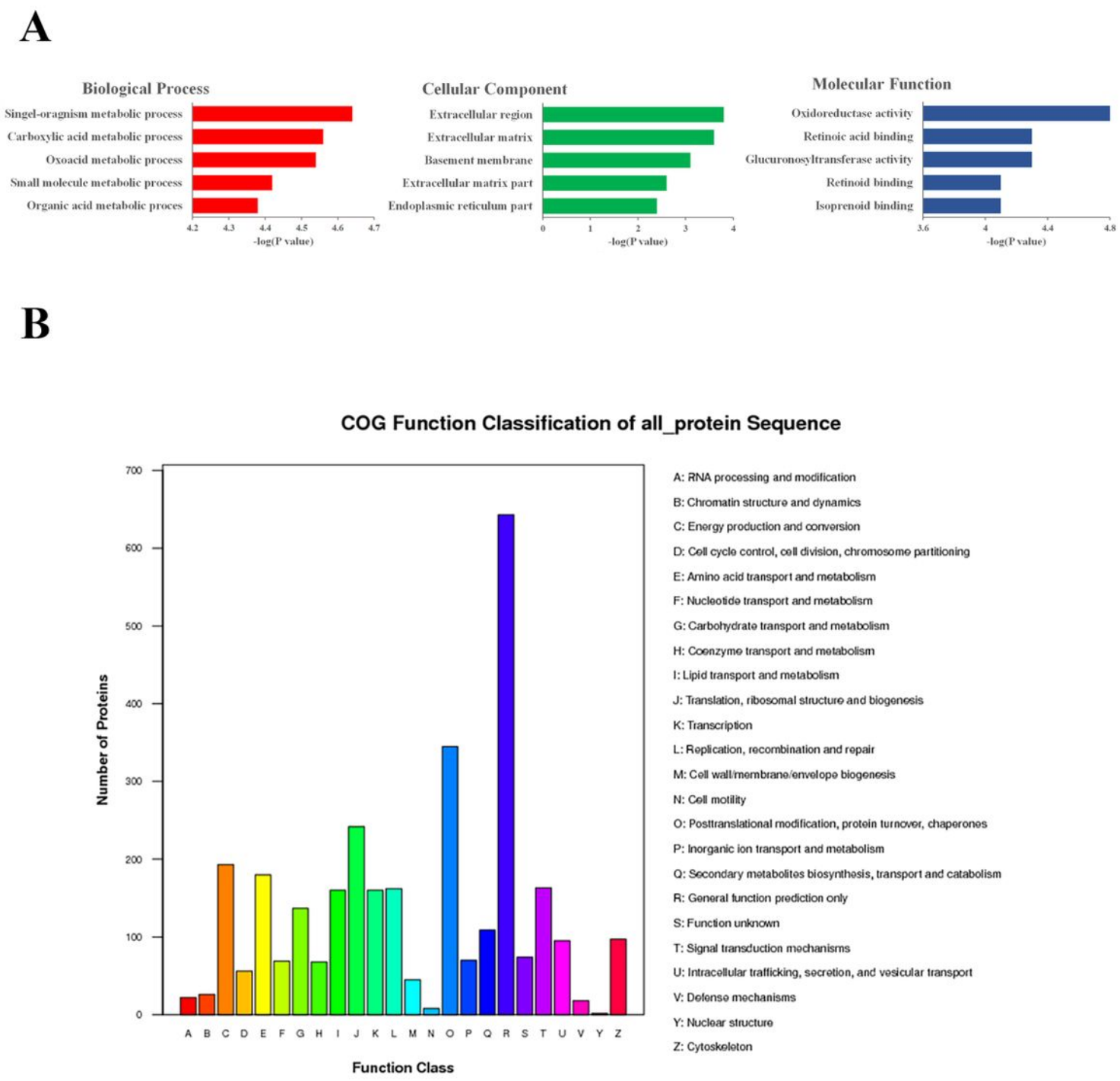

Figure 3

The Results of GO Enrichment Analysis

Legend: The main enrichment biological processes were single-organism metabolic process, carboxylic acid metabolic process, oxoacid metabolic process, small molecule metabolic process and organic acid metabolic process(A). COG Function Classification of all-protein Sequence(B).

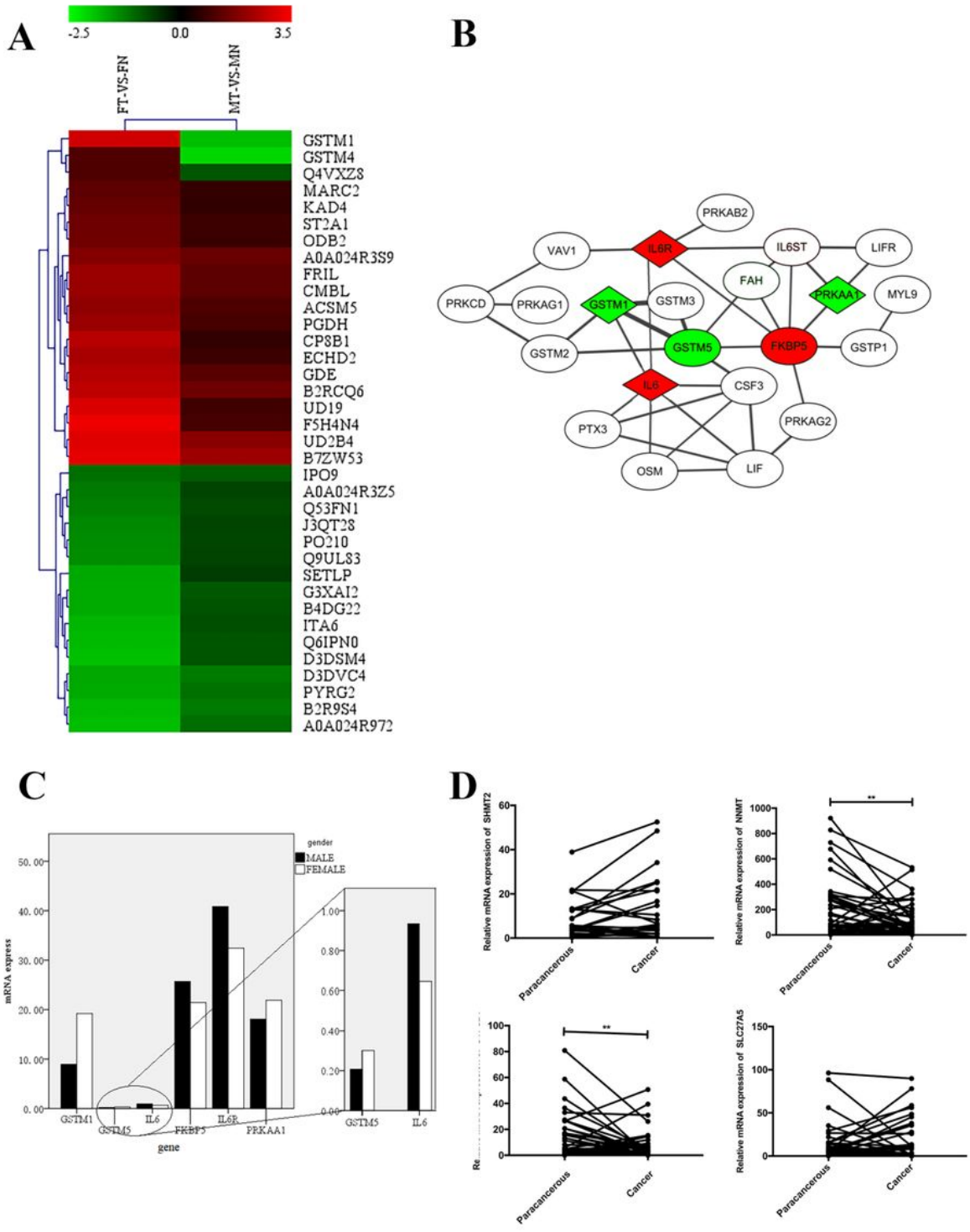

Figure 4

Heat Map of 36 proteins and Co-expression predicted by GeneMANIA.

Legend: Heat Map of 36 proteins (A) and Gene expression in relation to GSTM1 in public database data(B). GSTM1 is highly expressed in female HCC which is consistent with our iTRAQ data. IL6 and IL6R are highly expressed in male HCC and PRKAA1 is highly expressed in female HCC(C). The RNA

expressions on CYP8B1and NNMT in cancer are lower than paracancerous (p=0.0034 and p=0.0023), while the expressions SHMT2, SLC27A5 and MTHFD2 are not (D).

Figure 5

Kaplan-Meier survival analysis of GSTM1 expression levels.

Legend: The OS (overall survival) of the low expression group was lower than that of the high expression group, which suggests that low GSTM1 expression level has a worse prognosis.

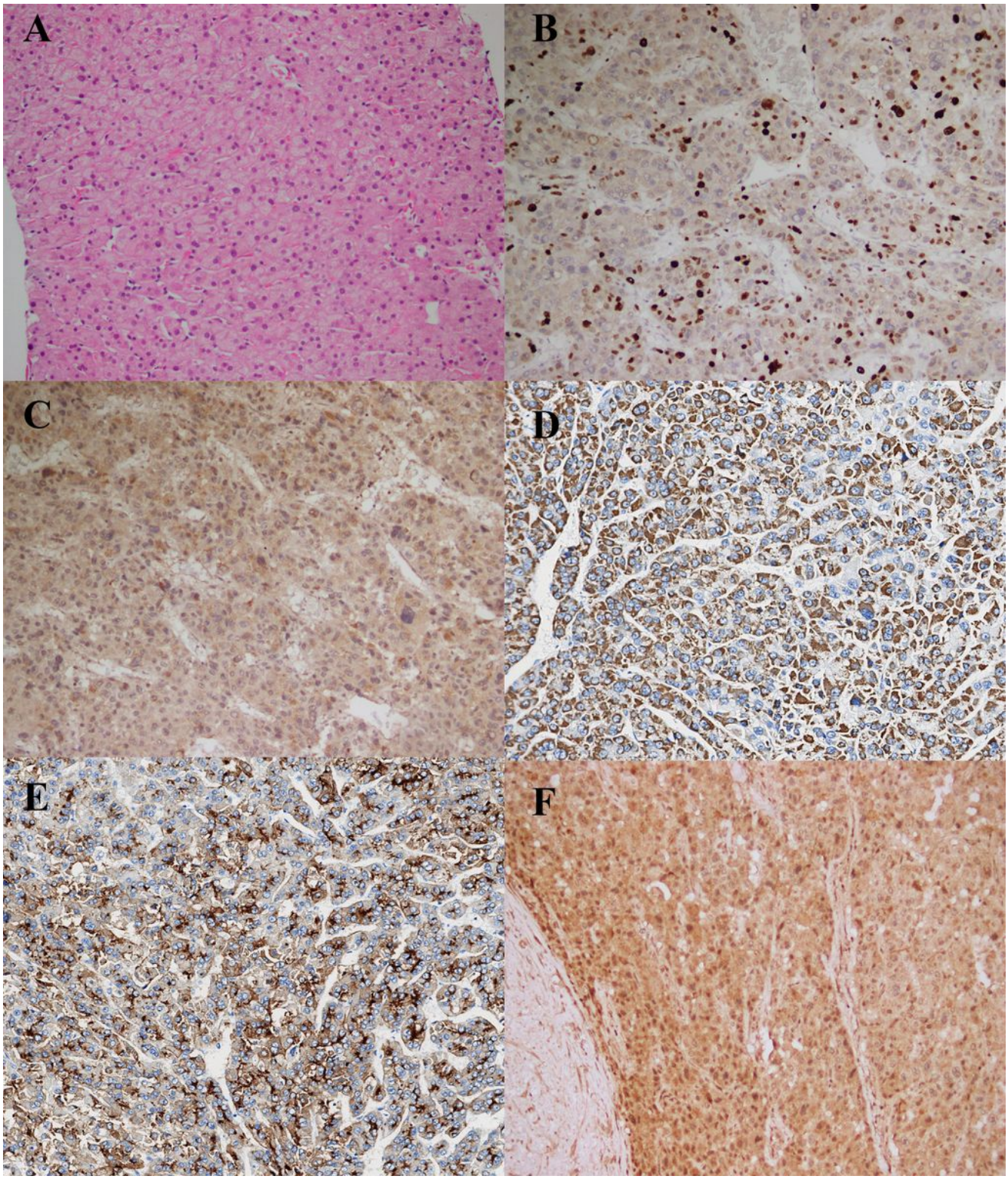

Figure 6

The Proteins Expressions Related to Cytochrome P450.

Legend: The proteins expression of CYP8B1 (B), MTHFD2(C), NNMT(D), SHMT2(E), SLC27A5(F) are different. MTHFD2, NNMT, SHMT2 are higher in tumors than in paracancerous (A) while CYP8B1,

SLC27A5 are contrary, which are higher in paracancerous than in tumor. The proteins expression on CYP8B1 are consisted with the RNA (Table 2).