# CYSTM1: A Novel Biomarker for Hepatocellular Carcinoma Prognosis

**Jun Du**
Wannan Medical College
**Jinguo Wang** ( shwangjg@sina.com )
Yijishan hospital, wannan medical college https://orcid.org/0000-0002-6100-6022

**Research**

**Keywords:** Hepatocellular cancer, CYSTM1, TCGA, prognosis evaluation, bioinformatics analysis

**Posted Date:** January 21st, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-150866/v1

# Abstract

**Background**: The expression and molecular mechanism of cysteine rich transmembrane module containing 1 (CYSTM1) in human tumor cells remains unclear. The aim of this study was to determine whether CYSTM1 could be used as a potential prognostic biomarker for hepatocellular carcinoma (HCC).

**Methods**: We first demonstrated the relationship between CYSTM1 expression and HCC in various public databases. Secondly, Kaplan–Meier analysis and Cox proportional hazard regression model were performed to evaluate the relationship between the expression of CYSTM1 and the survival of HCC patients which data was downloaded in the cancer genome atlas (TCGA) database. Finally, we used the expression data of CYSTM1 in TCGA database to predict CYSTM1-related signaling pathways through bioinformatics analysis.

**Results**: The expression level of CYSTM1 in HCC tissues was significantly correlated with T stage ($p = 0.039$). In addition, Kaplan–Meier analysis showed that the expression of CYSTM1 was significantly associated with poor prognosis in patients with early-stage HCC ($p = 0.003$). Multivariate analysis indicated that CYSTM1 is a potential predictor of poor prognosis in HCC patients ($p = 0.036$). The results of biosynthesis analysis demonstrated that the data set of CYSTM1 high expression was mainly enriched in neurodegeneration and oxidative phosphorylation pathways.

**Conclusion**: CYSTM1 is an effective biomarker for the prognosis of patients with early-stage HCC and may play a key role in the occurrence and progression of HCC.


# Introduction

HCC is a one of malignant tumor with high degree of malignancy, easy metastasis and recurrence, which is harmful to human health. Its mortality rate ranks the fourth among all kinds of cancer mortality[1]. In recent years, more and more studies have focused on the molecular mechanism of HCC, but it is still unclear. Because the early symptoms and clinical signs of HCC are not obvious, most patients are in advanced stage of cancer at the time of diagnosis due to its occult onset and rapid progress. Therefore, HCC patients generally have poor prognosis and low survival rate[2]. Research on the molecular mechanisms of early diagnosis and treatment of HCC is of great significance for reducing HCC mortality. CYSTM1 is a highly conserved cysteine-rich transmembrane protein in all eukaryotes. It may play a role of stress resistance in eukaryotes, including human beings[3]. The gene is located on chromosome 5q31.3 (C5orf32), consisting of 97 amino acids. However, the expression of CYSTM1 in human tumor tissues and its role in tumor development have not been reported. Therefore, we analyzed the expression patterns of CYSTM1 in HCC patients in this study. The potential mechanism of CYSTM1 in HCC was elaborated by bioinformatics analysis, which laid the foundation for the follow-up experiments.

# Materials And Methods

# Public data extraction and processing

The gene expression quantification data and clinical data of 374 HCC samples and 50 normal samples were downloaded from the GDC data portal (https://portal.gdc.cancer.gov/). We used Perl (https://www.perl.org/) scripts to decompress the downloaded compressed files in batches, which contain gene expression and clinical data of GC samples. Then, all sample IDs and RNA-seq data were integrated into a matrix file. Next, according to the Ensembl database (http://asia.ensembl.org/index.html), the Ensembl ID was converted to the gene symbol. Finally, we added the gene attribute (protein coding or lincRNA) after the gene symbol for subsequent operations. R (https://www.r-project.org/) language scripts and various packages were used to make the images and process the data. The expression difference of CYSTM1 in normal gastric and GC tissues was verified in the Oncomine database (https://www.oncomine.org). The threshold was set to the following parameters: *p*-value of 0.001, fold change of 1.5, and gene rank of all. The correlation between CYSTM1 expression and prognosis in HCC patients was evaluated by Kaplan–Meier plotter (http://kmplot.com). Gene Expression Profiling Interactive Analysis (GEPIA) databases was used to analyse the mRNA expression of CYSTM in HCC samples (http://gepia.cancer-pku.cn/index.html). The human protein atlas (HPA) database was used to analyze the protein expression of CYSTM1 between normal and HCC tissues (www.proteinatlas.org) .

# Gene Set Enrichment Analysis (gsea)

GSEA was performed using GSEA v3.0 (https://www.gsea-msigdb.org) and JAVA 8 (https://www.java.com) to identify CYSTM1 associated with gene sets. First, two files were prepared: (1) a .gct file containing CYSTM1 expression data of 374 HCC patients, and (2) a .cls file that divides the expression data of CYSTM1 into high expression and low expression. Then, the number of permutations was set to 1000 to test CYSTM1 correlations with the phenotypes using the c2.cp.kegg.v6.2 gene set database. Ultimately, with the normalized (NOM) *p*-value < 0.05 and false discovery rate (FDR) < 0.05, the gene sets with significant enrichment of CYSTM1 high expression-related genes were considered the enrichment gene sets.

### Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis

To find the co-expressed genes of CYSTM1, we first merged the same gene expression data in the processed matrix file and took the average. Secondly, the gene expression data of all the coding proteins in the matrix file was screened out, and the correlation was tested with the expression data of CYSTM1 one by one. Finally, we screened out co-expressed genes of CYSTM1 using Pearson's correlation coefficient > 0.4 and p-value < 0.05 as thresholds. Then, we used R language scripts and R packages to perform GO and KEGG enrichment analyses. We considered the co-expressed genes of CYSTM1 significantly enriched in the GO term and the KEGG pathway at p-value < 0.05 and FDR q-value < 0.05.

# Statistical analysis

R 4.0.2 was used for statistical analysis. The Kaplan–Meier survival curve and log-rank test were used to compare the survival rate between the CYSTM1 high expression and low expression groups. We performed univariate Cox-regression analysis of all clinicopathological parameters, and then integrated these

parameters into multivariate Cox-regression analysis to determine independent predictors of survival in HCC patients. *p*-values < 0.05 were considered statistically significant.

## Results

# Expression of CYSTM1 in HCC patients

According to the results of TCGA analysis, the expression of CYSTM1 mRNA in normal tissues and HCC tissues was significant different ($p < 0.001$), and its expression in cancer tissues was 2.3 times higher than that in normal tissues (Fig. 1A). In GEPIA database, the expression of CYSTM1 was up-regulated in HCC tissues ($p < 0.001$)(Fig. 1B). Similarly, The mRNA expression of CYSTM1 in HCC tissues in the Oncomine database respectively increased by 1.5 and 1.7 times in the two chips (Fig. 1C). The results of immunohistochemical staining in HPA database showed that the protein expression of CYSTM1 was strongly positive in HCC tissues, but negative in normal tissues. Moreover, CYSTM1 was mainly localized in the cell membrane and cytoplasm of HCC tissues (Fig. 1D).

# Relationship between the expression of CYSTM1 and clinicopathological features

In this study, we downloaded the clinical data of 374 HCC patients from the TCGA database, including survival time, survival status, age, gender, histological grade, T stage, N stage, M stage and TNM stage. HCC samples with unknown survival time were deleted, and 342 samples were left for subsequent analysis. We analyzed the relationship between the expression of CYSTM1 and the clinicopathological characteristics of HCC patients, which showed that the expression of CYSTM1 was significantly different from the histological grade ($p = 0.004$), T stage ($p = 0.046$) and M stage ($p = 0.048$)(Fig. 2). According to the median value of CYSTM1 expression, HCC patients were divided into low expression groups ($n = 170$) and high expression groups ($n = 172$). The results showed that the expression of CYSTM1 was statistically significant with T stage ($p = 0.039$) (Table 1).

Table 1
Relationship between CYSTM1 expression and clinicopathological features in patients with HCC.

| Variables | n (%) | CYSTM1 expression in HCC samples | | | |
|---|---|---|---|---|---|
| | 342 (100) | Low | High | $c^2$ | *p*-value |
| **Age (years)** | | | | 0.939 | 0.333 |
| < 60 | 162 (47.4) | 85 | 77 | | |
| ³ 60 | 180 (52.6) | 85 | 95 | | |
| **Gender** | | | | 1.251 | 0.263 |
| Male | 233 (68.1) | 111 | 122 | | |
| Female | 109 (31.9) | 59 | 50 | | |
| **Grade** | | | | 4.645 | 0.302 |
| G1 | 45 (13.2) | 25 | 20 | | |
| G2 | 168 (49.1) | 85 | 83 | | |
| G3 | 117 (34.2) | 57 | 60 | | |
| G4 | 12 (3.5) | 3 | 9 | | |
| **T stage** | | | | 8.347 | **0.039*** |
| T1 | 171 (50) | 95 | 76 | | |
| T2 | 85 (24.9) | 35 | 50 | | |
| T3 | 76 (22.2) | 38 | 38 | | |
| T4 | 10 (2.9) | 2 | 8 | | |
| **N stage** | | | | 1.747 | 0.418 |
| N0 | 251 (47.4) | 129 | 122 | | |
| N1 | 4 (47.4) | 1 | 3 | | |
| Nx | 87 (47.4) | 40 | 47 | | |
| **M stage** | | | | 1.504 | 0.471 |
| M0 | 261 (76.3) | 125 | 136 | | |
| M1 | 4 (1.2) | 2 | 2 | | |
| Mx | 77 (22.5) | 43 | 34 | | |
| **AJCC stage** | | | | 6.058 | 0.108 |

*$p$-value < 0.05 was considered statistically significant.

| | | | |
|---|---|---|---|
| Stage I | 169 (49.4) | 95 | 74 |
| Stage II | 83 (24.3) | 34 | 49 |
| Stage III | 85 (24.9) | 39 | 46 |
| Stage IV | 5 (1.4) | 2 | 3 |

*$p$-value < 0.05 was considered statistically significant.

# Relationship between expression of CYSTM1 and survival

The survival time of HCC patients between the CYSTM1 low expression group and the CYSTM1 high expression group was statistically significant (p = 0.003). This result indicates that the prognosis of the CYSTM1 high expression group is poor. After that, we performed subgroup analysis of each clinicopathological features which showed that the expression of CYSTM1 and survival time were statistically significant in the ≥ 60 years old group (p = 0.002), stage I group (p = 0.006) and T1 stage group (p = 0.003)(Fig. 3). The results of these subgroup analyses indicated that the expression of CYSTM1 is more predictive of prognosis in early-stage HCC patients ≥ 60 years old. Then we verified its in the Kaplan-Meier database and derived similar results (Fig. 4). Univariate analysis demonstrated that T stage (p < 0.001), AJCC stage (p < 0.001) and CPXM1 expression (p = 0.005) are related to the overall survival of HCC patients (Fig. 5A). In addition, multivariate analysis manifested that the expression of CPXM1 is an independent prognostic factor for HCC patients (p = 0.036)(Fig. 5B).

# Gsea Analysis Of Cystm1-related Signaling Pathways

We performed GSEA analysis on the CPXM1 high and low expression datasets in TCGA to determine the various pathways that could take part in HCC. With normal p-value < 0.05 and false discovery rate q-value < 0.05 as the threshold, we listed the top nine pathways related to the CPXM1 high expression dataset (Table 2), including huntingtons disease, alzheimers disease, oxidative phosphorylation, parkinsons disease, proteasome, vibrio cholerae infection, lysosome, snare interactions in vesicular transport and glutathione metabolism (Fig. 6).

Table 2
GSEA terms that are significantly enriched in the high CYSTM1 expression.

| | GSEA team | Size | ES | NES | NOM p-val | FDR q-val |
|---|---|---|---|---|---|---|
| 1 | KEGG_HUNTINGTONS_DISEASE | 181 | 0.70 | 2.34 | 0 | 0 |
| 2 | KEGG_ALZHEIMERS_DISEASE | 166 | 0.68 | 2.29 | 0 | 0 |
| 3 | KEGG_OXIDATIVE_PHOSPHORYLATION | 132 | 0.75 | 2.17 | 0 | 0 |
| 4 | KEGG_PARKINSONS_DISEASE | 128 | 0.72 | 2.15 | 0 | 0 |
| 5 | KEGG_PROTEASOME | 46 | 0.81 | 2.09 | 0 | 0 |
| 6 | KEGG_VIBRIO_CHOLERAE_INFECTION | 54 | 0.67 | 2.07 | 0 | 0 |
| 7 | KEGG_LYSOSOME | 121 | 0.62 | 1.99 | 0 | 0.004 |
| 8 | KEGG_SNARE_INTERACTIONS_IN_VESICULAR_TRANSPORT | 38 | 0.68 | 1.97 | 0 | 0.005 |
| 9 | KEGG_GLUTATHIONE_METABOLISM | 49 | 0.58 | 1.94 | 0 | 0.011 |

ES, enrichment score; NES, normalized enrichment score; NOM, nominal; FDR, false discovery rate. NOM p-val < 0.05 and FDR q-val < 0.05 are considered as significant.

In order to further elucidate the molecular mechanism of cystm1 in gastric cancer, we screened out cystm1 co expression genes from TCGA database and visualized their correlation (Fig. 7A). We then analyzed the CYSTM1 co-expressed gene expression correlation by GO and KEGG analyses. GO analysis showed that the CYSTM1 co-expressed genes were mainly enriched in unfolded protein binding, NADH dehydrogenase activity, NADH dehydrogenase (ubiquinone) activity, NADH dehydrogenase (quinone) activity, ubiquitin binding, etc (Fig. 7B). In addition, KEGG analysis showed that CYSTM1 co-expressed genes were mainly enriched in signaling pathways related to huntington disease, prion disease, proteasome, amyotrophic lateral sclerosis, parkinson disease, etc (Fig. 7C). Chord plot displays of the relationship between CYSTM1 co-expressed genes and GO analysis are shown in Fig. 8A. The CYSTM1 co-expressed gene profiles are displayed in each GO analysis by hierarchical clustering (Fig. 8B).

# Discussion

HCC is one of the most fatal cancers in the world which occurrence and development are a multi-gene, multi-step and multi-stage process[3, 4]. With the development of molecular biology, more and more biomarkers have been found. However, the prognosis of HCC is still poor due to the lack of effective biomarkers to predict early-stage HCC[5, 6]. Therefore, it is of great significance to find a specific and sensitive tumor marker to assist in the diagnosis and treatment of early-stage HCC. This study is the first time to propose that CYSTM1 could be used as a biomarker for the prognosis of HCC, and the overexpression of CYSTM1 is significantly related to the clinicopathological characteristics and prognosis of HCC patients. In addition, multivariate Cox proportional hazard regression model showed that CYSTM1 was an independent risk factor for survival of HCC patients.

Nearly 10 years ago, CYSTM1 was first proposed and characterized, when it was proved to be a cysteine-rich transmembrane module[3]. It is noteworthy that CYSTM1 is also expressed and located on the cell membrane of human tissues. In the past researches, CYSTM1 has only a few reports in some eukaryotes, and it has been proved that CYSTM family proteins play an important role in resistance to drug, resistance to metal ions, and resistance to viruses[8, 9]. These functions may be related to cysteines, the acid residues and the cytoplasmic polar disordered head on CYSTM1, and these structures are highly conserved in different species. The C-terminal transmembrane helix of CYSTM1 contains 5–6 cysteines, among which 3–4 continuous cysteines constitute the cysteine patch. This may change the redox potential or radical quenching of the membrane, thus playing an antioxidant role[10]. CDT1, a member of CYSTM1 subfamily which rich in cysteine polypeptide, is heterologously expressed in yeast to prevent cadmium from entering cells[8]. Vallee and Margoshe reported for the first time that metallothioneins (MTs) are cadmium binding proteins in horse kidney cortex[11], and it contained a high proportion of cysteines[12]. These results suggest that CYSTM1 may act as a metallothionein-like protein in cell membrane.

Through KEGG enrichment analysis, we found that CYSTM1 co-expression genes are mainly concentrated in pathway of neurodegeneration - multiple diseases, especially huntington's disease, which is consistent with the results of transcriptome analysis by Mastrokolias using next-generation sequencing to predict biomarkers of huntington's disease[13]. Copper could increase the aggregation of poly-glutamine (polyQ) in vivo and in vitro, but MTs could protect huntington model cells from the toxic effects of polyQ[14]. Then through GO enrichment analysis, it is concluded that CYSTM1 co-expressed genes are mainly enriched in unfolded protein binding and NADH dehydrogenase activity. The former function is related to the results of KEGG enrichment analysis, because neuronal cells are highly sensitive to unfolded protein. Long-term accumulation of unfolded protein will cause endoplasmic reticulum stress, which may lead to cell apoptosis and necrosis if stress exists permanently[15]. The function of the latter may be related to the resistance to cellular oxidative stress. The tumor microenvironment also contains a large amount of reactive oxygen species (ROS). These ROS could be produced by tumor-related fibroblasts, inflammatory cells, vascular endothelial cells, hypoxic internal environment through a variety of ways[16]. At the same time, the effect of ROS on tumor is bidirectional. On the one hand, it could promote tumor growth and progression by stimulating tumor cells migration and invasion[17]. On the other hand, high levels of ROS could cause cell apoptosis or necrosis, which is detrimental to the progress of tumor[]. However, MTs could protect DNA from damage by exchanging various toxic metal ions and oxygen free radicals[18]. Therefore, the up-regulation of CYSTM1 expression may be to protect tumor cells from apoptosis by resisting high ROS levels. Of course, these results need a large number of scientific experiments to verify, and it is also one of our follow-up research topics.

# Conclusion

CYSTM1 is present on the cell membrane of human HCC cells. Moreover, overexpression of CYSTM1 may be a potential tumor biomarker for poor prognosis of HCC. The stress resistance function of CYSTM1 in eukaryotes and the bioinformatics analysis results make us predict that oxidative stress resistance may be one of the most important functions of CYSTM1 expression in HCC. However, it is necessary to further explore the biological function of CYSTM1 in human tumor cells.

# Abbreviations

CYSTM1: cysteine rich transmembrane module containing 1; HCC: hepatocellular carcinoma; TCGA: the cancer genome atlas; GO: gene ontology; KEGG: kyoto encyclopedia of genes and genomes.

# Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Availability of data and materials

The datasets used in the present study are available from pubilc databases.

### Competing interests

The authors declare that they have no competing interests.


### Funding

Not applicable.


### Authors' contributors

DJ planned the study and wrote the manuscript. WJG resolved the dispute and supervised the study. All authors screened and approved the final manuscript.


### Acknowledgments

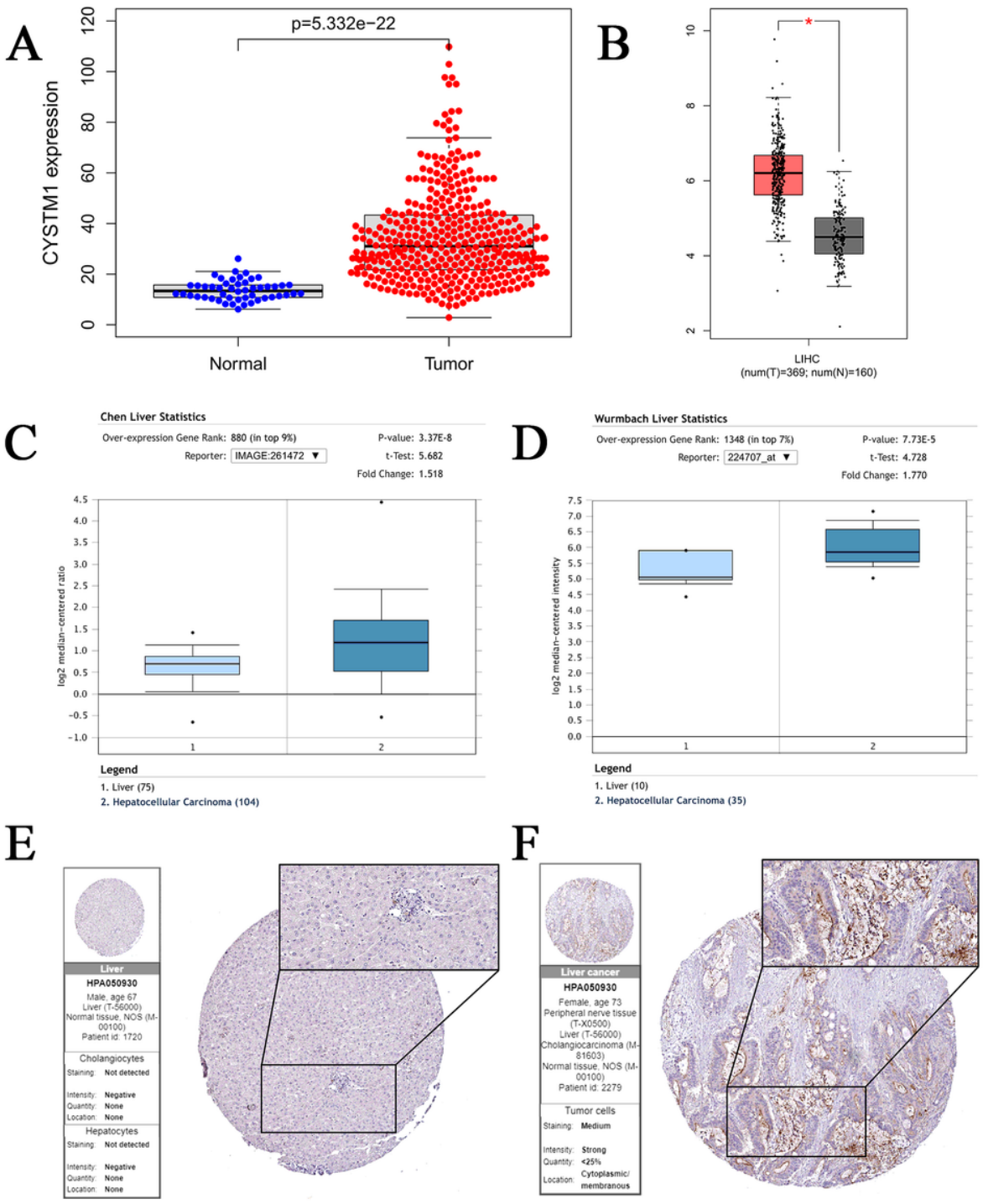
The authors thank the teacherss of Central Laboratory, The First Affiliated Hospital of Wannan Medical College for providing us with valuable advice.


# References


1. Bray F, Ferlay J, Soerjomataram I. **et al**. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68:394–424.
2. Longo DL, Villanueva A. Hepatocellular Carcinoma[J]. N Engl J Med. 2019;380(15):1450–62.
3. Aravind L. CYSTM, a novel cysteine-rich transmembrane module with a role in stress tolerance across eukaryotes. Bioinformatics. 2010;26:149 – 52.
4. Khan FS, Ali I, Afridi UK. **et al**. Epigenetic mechanisms regulating the development of hepatocellular carcinoma and their promise for therapeutics[J]. Hepatol Int. 2017;11:45–53.

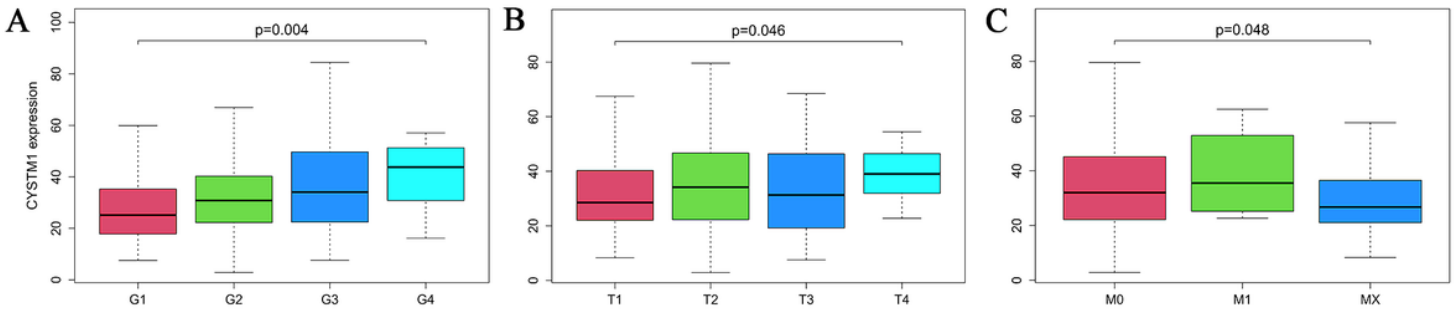
5. Yates LR, Campbell PJ. Evolution of the cancer genome.[J]. Nat Rev Genet. 2012;13:795–806.
6. Chen JG, Zhang SW. Liver cancer epidemic in China: past, present and future.[J]. Semin Cancer Biol. 2011;21:59–69.
7. Chen Y, Liu W, Shang Y. **et al**. Folic acid-nanoscale gadolinium-porphyrin metal-organic frameworks: fluorescence and magnetic resonance dual-modality imaging and photodynamic therapy in hepatocellular carcinoma[J]. Int J Nanomedicine. 2018;14:57–74.
8. Masato K, Shuichi M, Yoshihiro T. **et al**. Novel Cysteine-Rich Peptides from Digitaria ciliaris and Oryza sativa Enhance Tolerance to Cadmium by Limiting its Cellular Accumulation. Plant Cell Physiol. 2009;50:106–17.
9. Lee JK, Kim M, Choe J. **et al**. Characterization of uvi15+, a stress-inducible gene from Schizosaccharomyces pombe. Mol Gen Genet. 1995;246:663–70.
10. Deneke SM. Thiol-based antioxidants. Curr Top Cell Regul. 2000;36:151–80.
11. Margoshes M, Vallee BL. Cadmium protein from equine kidney cortex. J Am Chem Soc. 1986;79:4813–4.
12. Sutherland DEK, Stillman MJ. The "magic numbers" of metallothionein. Metallomics. 2011;3:444–63.
13. Mastrokolias A, Ariyurek Y, Goeman JJ. **et al**. Huntington's disease biomarker progression profile identified by transcriptome sequencing in peripheral blood. Eur J Hum Genet. 2015;23:1349–56.
14. Hands SL, Mason R, Sajjad MU. **et al**. Metallothioneins and copper metabolism are candidate therapeutic targets in Huntington's disease. Biochem Soc Trans. 2010;38:552–8.
15. Ghemrawi R, Khair M. Endoplasmic Reticulum Stress and Unfolded Protein Response in Neurodegenerative Diseases. Int J Mol Sci. 2020;21:6127.
16. Policastro LL, Ibañez IL, Notcovich C. **et al**. The tumor microenvironment: characterization, redox considerations, and novel approaches for reactive oxygen species-targeted gene therapy. Antioxid Redox Signal. 2013;19:854–95.
17. Ko E, Seo HW, Jung G. Telomere length and ROS levels are positively associated with a high risk of mortality and recurrence in hepatocellular carcinoma. Hepatol. 2017;67:1378–91.
18. Chandel NS, Tuveson DA. The promise and perils of antioxidants for cancer patients. N Engl J Med. 2014;371:177–8.

# Figures



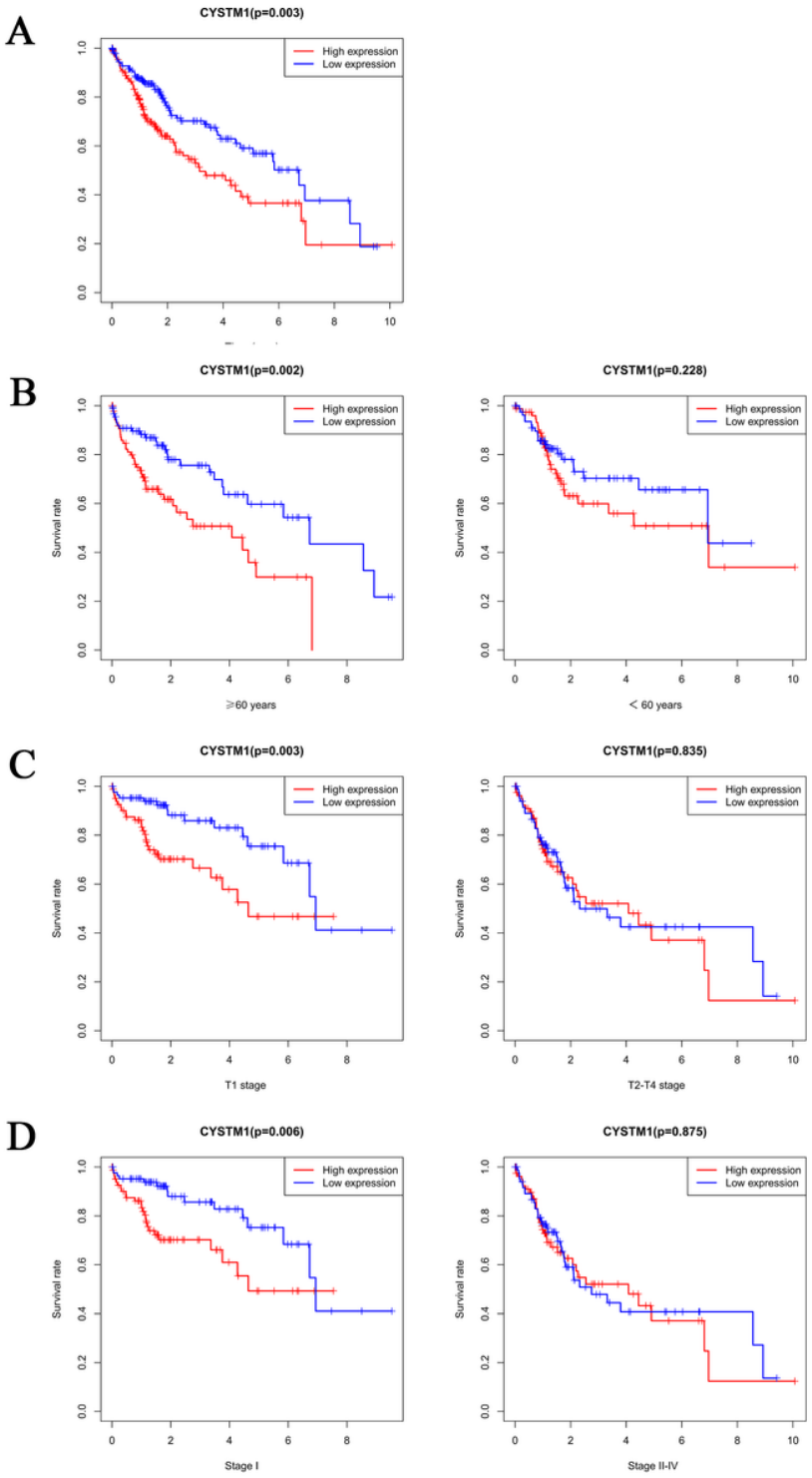

## Figure 1

CYSTM1 is upregulated in HCC tissues. (A) Dot plots represent CYSTM1 expression levels in HCC tissues (n = 374) and normal tissues (n = 50) according to the data from the TCGA-LIHC cohort. (B) Boxpolt represent CYSTM1 expression levels in gastric cancer tissues (n = 369) and normal gastric tissues (n = 160) according to the data from GEPIA-LIHC cohort. (C-D) The Chen and Wurmbach liver analyses in the Oncomine database are presented, illustrating the overexpression of CYSTM1 in HCC tissues. (E-F) The expression of CYSTM1 protein in normal tissue and HCC tissue were visualized using immunohistochemistry via HPA.



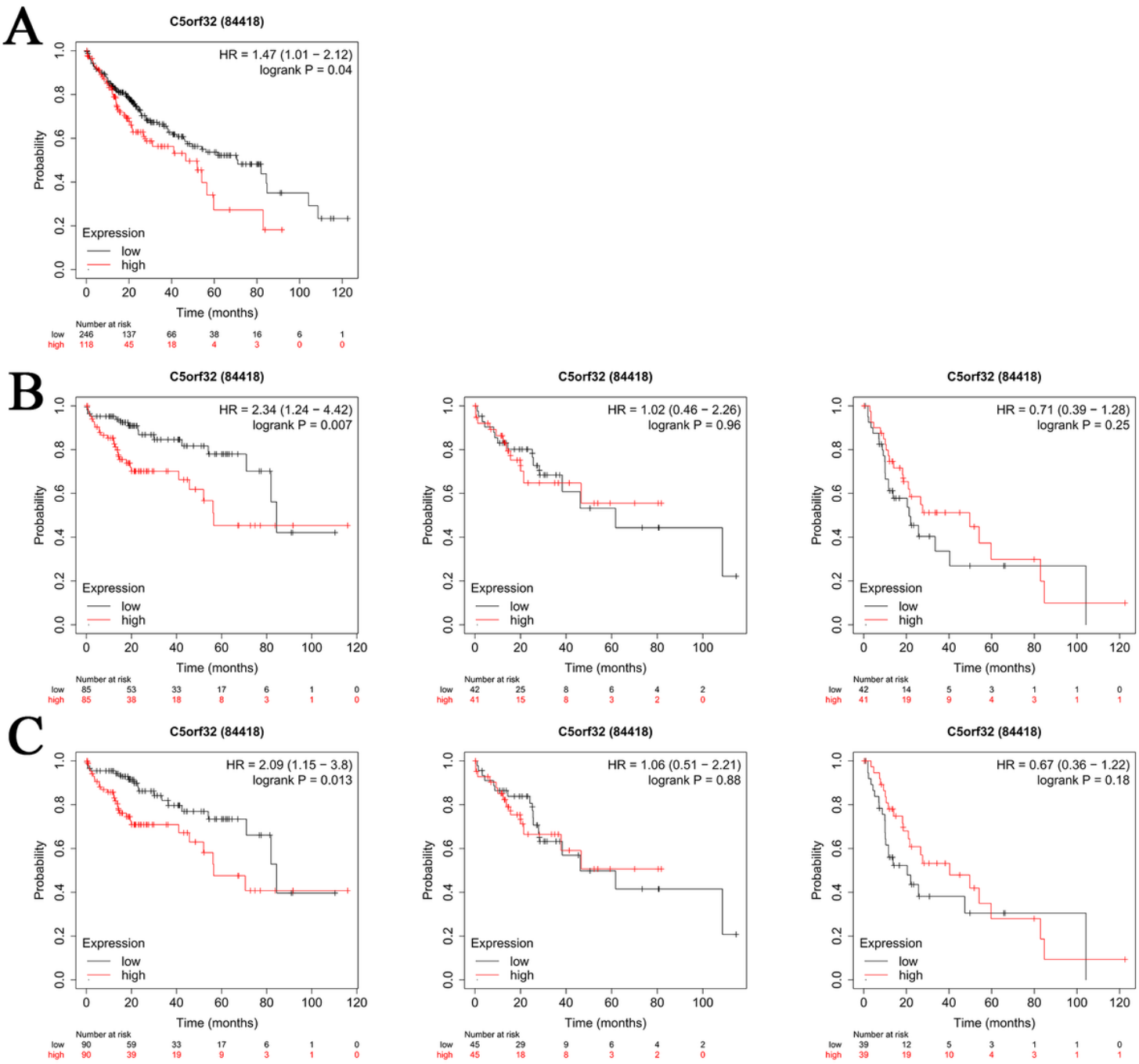

## Figure 2

Association between CYSTM1 expression and clinicopathological features, including (A) histological grade, (B) T stage, and (C) M stage.



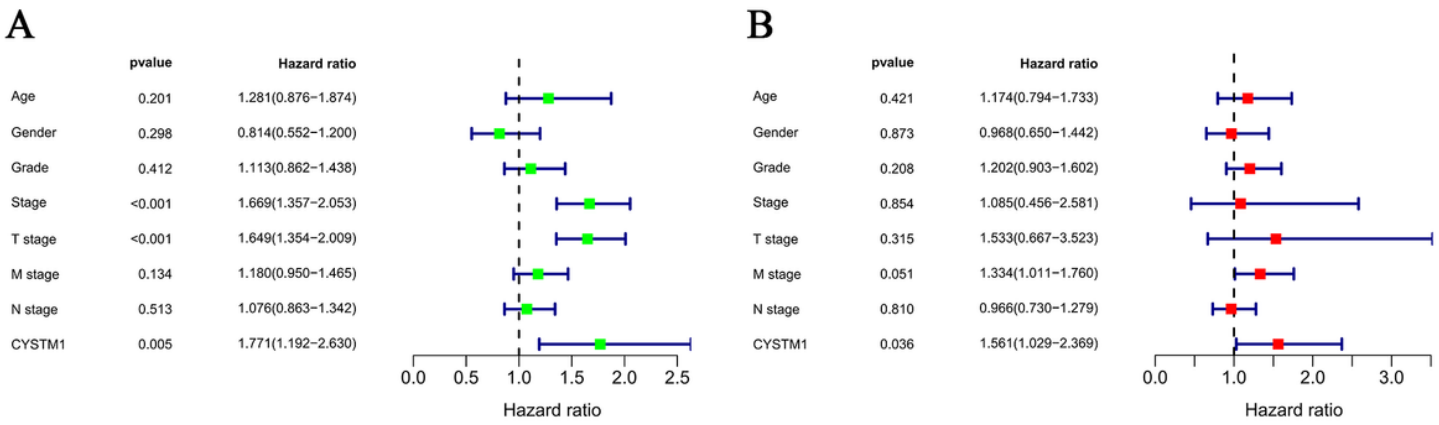

## Figure 3

High level CYSTM1 expression was associated with poor prognosis in HCC patients which data was downloaded from TCGA-LIHC. Kaplan-Meier curves of overall survival were stratified by (A) CYSTM1 expression (low vs. High), (B) patients with ≥60 years old and <60 years old, (C) patients with T1 stage and T2-4 stage, (D) patients with stage 1 and stage 2-4.



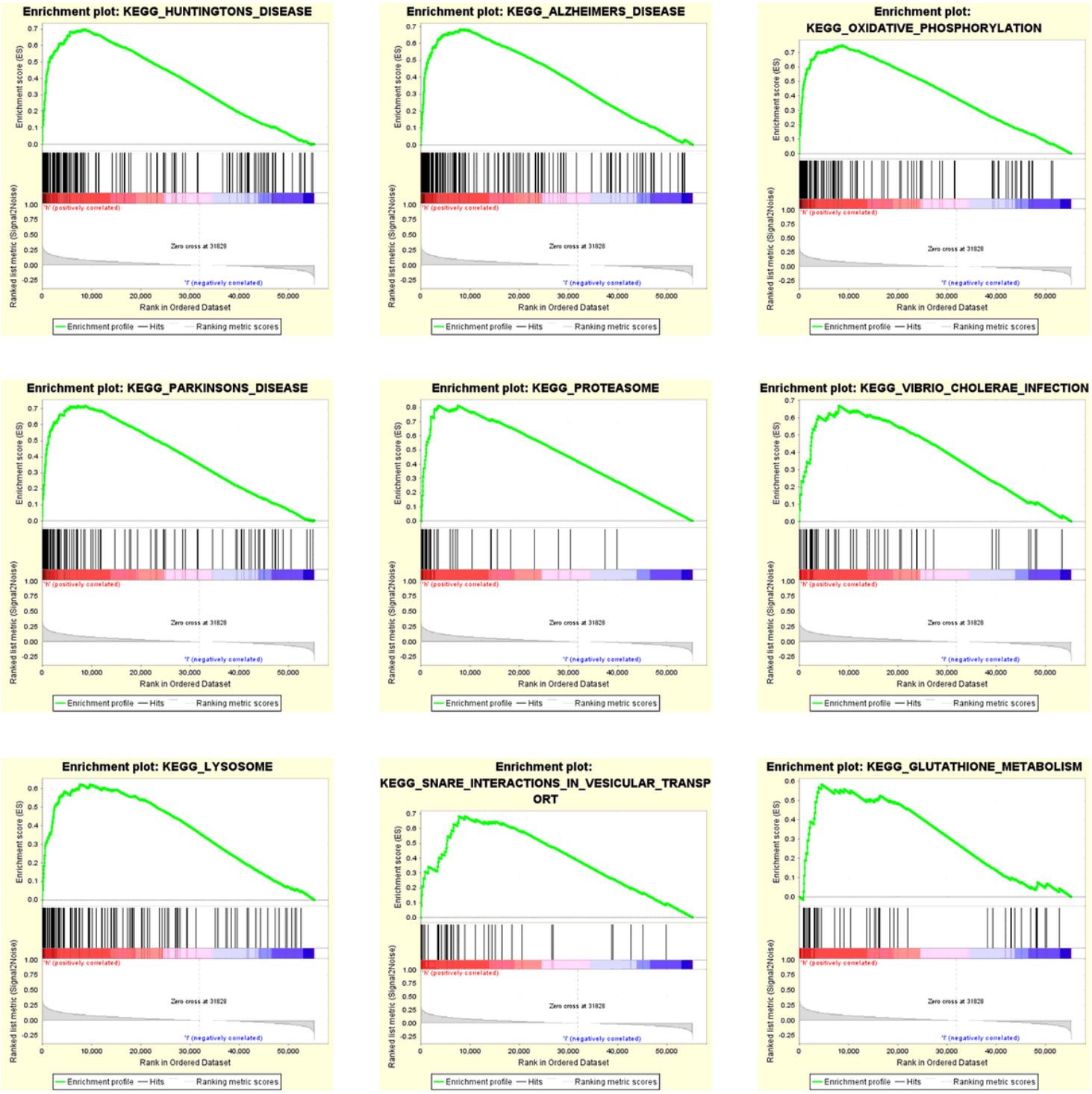

## Figure 4

High level CYSTM1 expression was associated with poor prognosis in HCC patients which data was downloaded from Kaplan-Meier database. Kaplan-Meier curves of overall survival were stratified by (A) CYSTM1 expression (low vs. High), (B) patients with stage 1, stage 2 and stage 3, (C) patients with T1 stage, T2 stage and T3 stage. Because of the small number of patients with T4 and stage 4, the survival curves could not be displayed.



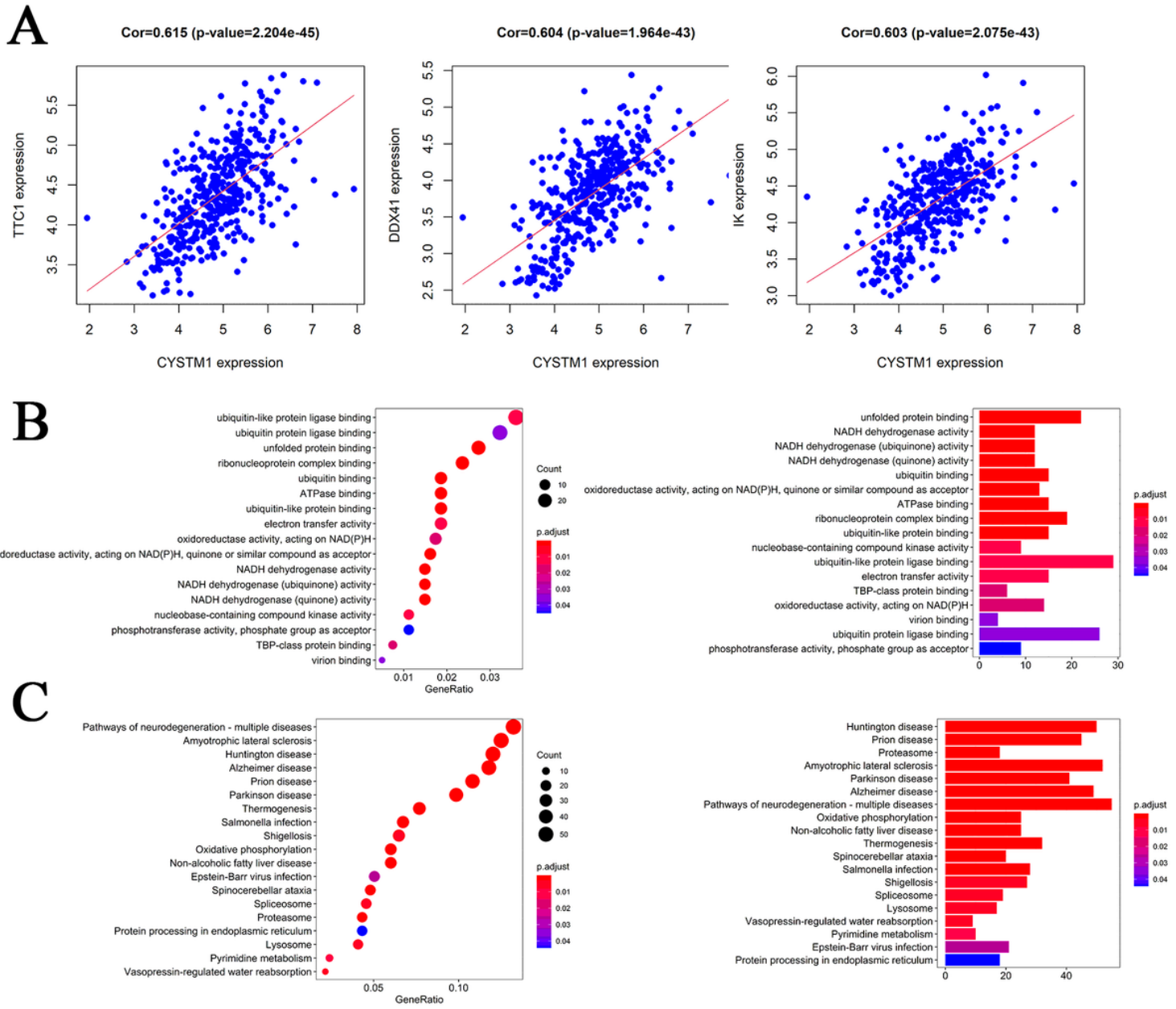

## Figure 5

Univariate and multivariate OS Analysis in TCGA-LIHC cohort.



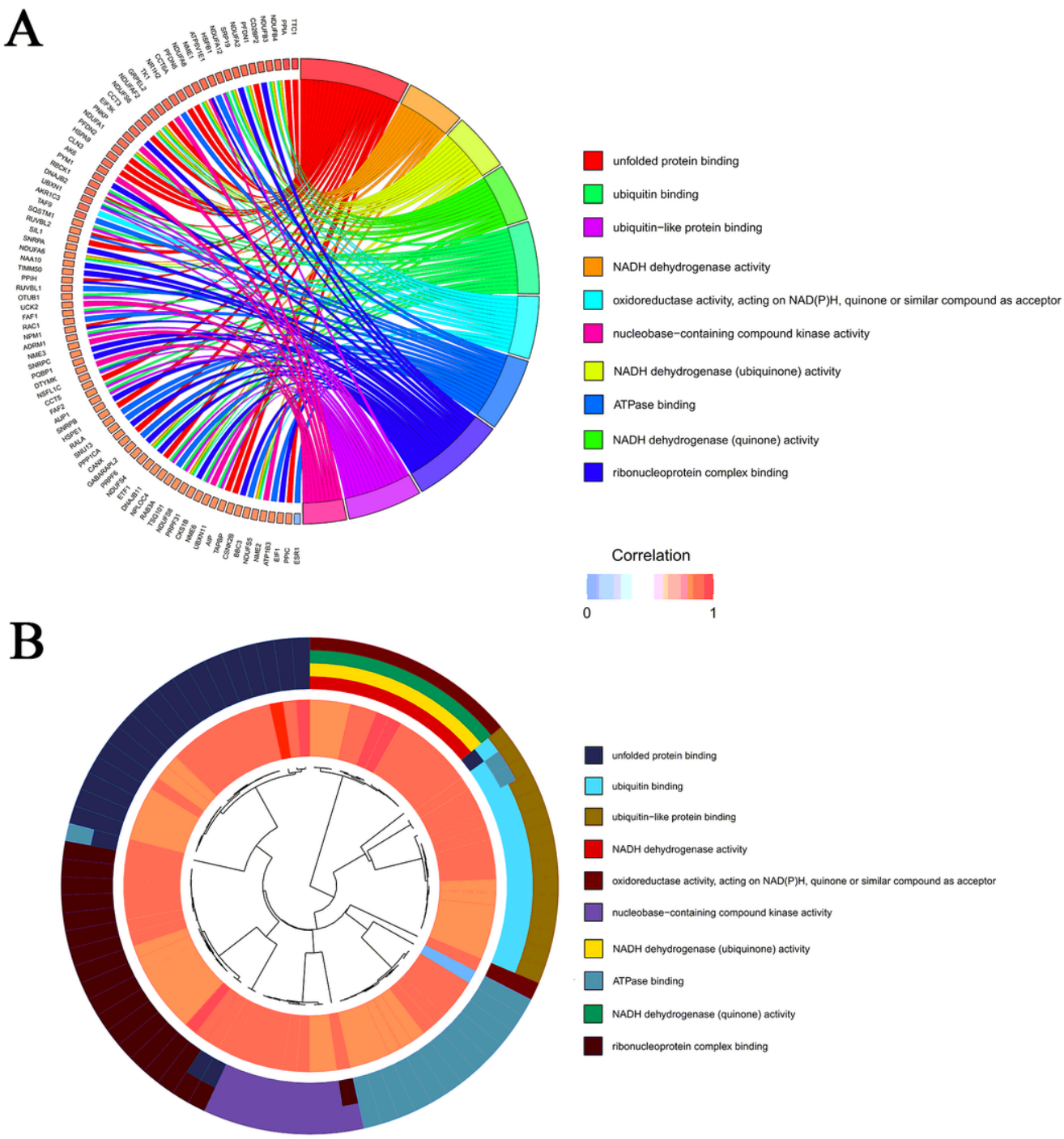

## Figure 6

GSEA result of CYSTM1 in TCGA dataset. KEGG pathways, named “KEGG_HUNTINGTONS_DISEASE”, “KEGG_ALZHEIMERS_DISEASE”, “KEGG_OXIDATIVE_PHOSPHORYLATION”, “KEGG_PARKINSONS_DISEASE”, “KEGG_PROTEASOME”, “KEGG_VIBRIO_CHOLERAE_INFECTION”, “KEGG_LYSOSOME”, “KEGG_SNARE_INTERACTIONS_IN_VESICULAR_TRANSPORT ” and “KEGG_GLUTATHIONE_METABOLISM”.




## Figure 7

GO and KEGG analyses of CYSTM1 co-expressed genes in TCGA database. (A) The most relevant co-expressed genes of CPXM1, including TTC1, DDX41 and IK (Cor > 0.6). (B) CYSTM1 co-expressed genes in TCGA were mainly enriched in GO pathways related to unfolded protein binding and NADH dehydrogenase activity(left picture is dotplot; right picture is barplot). (C) CPXM1 co-expressed genes in TCGA were mainly enriched in KEGG pathways related to pathways of neurodegeneration and Oxidative phosphorylation (left picture is dotplot; right picture is barplot).




## Figure 8

(A) Hierarchical clustering of the CYSTM1 co-expressed gene profiles in each KEGG pathway. (B) Chord plot displays of the relationship between the CYSTM1 co-expressed genes and KEGG pathways.