A systematic pan-cancer study demonstrates the oncogenic function of heterogeneous nuclear ribonucleoprotein C

Chenxi Pan[1,2], Qian Wu[1], Nianjie Feng[1]

[1]Key Laboratory of Fermentation Engineering, Ministry of Education, National "111" Center for Cellular Regulation and Molecular Pharmaceutics, Hubei Key Laboratory of Industrial Microbiology, Hubei Research Center of Food Fermentation Engineering and Technology, Hubei University of Technology, Wuhan 430068, China
[2]The Second Affiliated Hospital of Dalian Medical University, Dalian 116027, China

Correspondence to: Nianjie Feng; **email:** njfeng@hbut.edu.cn
Keywords: HNRNPC, prognosis, phosphorylation, pan-cancer, immune infiltration
Received: December 13, 2021   **Accepted:** March 14, 2022   **Published:** March 28, 2022

ABSTRACT

Although complex links between heterogeneous nuclear ribonucleoprotein C (HNRNPC) and numerous types of cancer have been shown in both cell and animal models, a comprehensive pan-cancer investigation on the features and activities of HNRNPC is still lacking. Based on the Cancer Genome Atlas and Gene Expression Omnibus datasets, we investigated the possible oncogenic effects of HNRNPC in thirty-three cancers. HNRNPC expression was detected in the majority of cancers, and its expression level was shown to be significantly linked with cancer patient prognosis. HNRNPC increased the phosphorylation of S220, which was detected in various cancers, including ovarian cancer and colon cancer. HNRNPC expression was also shown to be related to cancer-associated cell infiltration, most notably in uveal melanoma, testicular germ cell tumors, and thymoma. Additionally, the signaling pathway for vascular endothelial growth factors and RNA transport were implicated in HNRNPC's functioning processes. In short, HNRNPC may further influence cancer progression through gene mutation, protein phosphorylation, cancer associated fibroblasts infiltration and related molecular pathways. This work was intended to provide a relatively thorough knowledge of the oncogenic activities of HNRNPC across a variety of tumor types by performing a systematic pan-cancer investigation.

INTRODUCTION

Malignant tumors are generally acknowledged as a main cause of death and a significant impediment to patients' quality of life in the majority of countries worldwide [1]. Despite recent advancements in neoadjuvant chemo-radiation, targeted therapy, and immunotherapy, the overall survival rates continue to be low [2]. As a result, it is critical to discover novel tumor indicators to accomplish early identification, diagnosis, and therapy.

Heterogeneous nuclear ribonucleoproteins C (HNRNPCs) are members of the HNRNP family which are represented by two subtypes in humans (C1 and C2) [3, 4]. HNRNPC is well-known for its activity in alternative splicing regulation as an RNA binding protein [5]. Alternative

splicing regulators are required for RNA maturation, stabilization, translocation, and posttranscriptional modification [6]. HNRNPC is required for the expression of genes and the proper functioning of biological processes [7–9]. Numerous researches have been conducted in recent years to examine the association between HRNNPC and cancer, which is one of the significant outcomes of aberrant gene expression and biological processes. HNRNPC expression has been reported to be abnormal in a number of malignancies, including prostate cancer [10], breast cancer (BRCA) [11, 12], pancreatic ductal adenocarcinoma [13], lung adenocarcinoma (LUAD) [14], glioblastoma [15], oral squamous cell carcinoma, and hepatocellular carcinoma [16]. Additional research revealed that HNRNPC was involved in the processing of cancer-related genes,

including BRCA1/2, uPAR, MALAT, PDCD4, and cMyc [17–20]. HNRNPC is certainly related to several types of cancer. However, the differential expression of HNRNPC has an uncertain effect on the prognosis of many malignancies, and there is no universally accepted criterion for determining the significance of HNRNPC in distinct cancers. The purpose of this study is to determine the expression and effect of HNRNPC in pan-cancer and to determine whether HNRNPC can be used as a universal cancer marker.

By examining numerous profiles of a large number of human cancers, the pan-cancer project will evaluate the similarities and variations in both genetic and cellular changes across distinct tumor types [21]. On one hand, pan-cancer studies can reflect the cross-tumor similarity reflected by different levels of molecular characteristics. On the other hand, pan-cancer studies have the advantage of having a large sample size. Pan-cancer studies have demonstrated their advantages in finding novel common biomarkers. In this study, gene and protein expression, survival status, genetic alteration, phosphorylation, immune infiltration, and molecular pathways were performed for pan-cancer analysis of HNRNPC. The findings of this study are expected to offer a greater insight on the possible molecular mechanism through which HNRNPC contributes to the development or clinical prognosis of various malignancies.

MATERIALS AND METHODS

Gene expression analysis

We examined the Cancer Genome Atlas (TCGA) database using the TIMER2 (tumor immune estimation resource, version 2) online resource's "Gene DE" module for changes in HNRNPC expression between the tumor and its nearby normal tissues for distinct malignancies or tumor subtypes. To generate box plots illustrating the differential expression among these tumor tissues and their matching normal tissues in the GTEx (Genotype-Tissue Expression) database, we utilized the GEPIA2 (Gene Expression Profiling Interactive Analysis, version 2) web server's "Expression analysis-Box Plots" module [22] with the following settings: P-value threshold = 0.01, log2FC (fold change) = 1. Additionally, we used HEPIA2's "Clinical Stage Plot" module to produce violin plots of HNRNPC expression in each of the TCGA tumors' distinct clinical stages, which we then compared with one another (stage I, stage II, stage III, and stage IV). Furthermore, we used log2 [TPM (Transcripts per million) + 1] transformed expression data from the TPM database as the input for the box and the violin graphs.

In this study, we used the University of Alabama Cancer Database (UALCAN) portal, an interactive online resource for reviewing cancer Omics data [23]. The Clinical proteomic tumor analysis consortium (CPTAC) dataset was also analyzed for protein expression by using the UALCAN portal. Our research team evaluated the levels of HNRNPC (NP 001070910.1) total protein and phosphoprotein expression in primary tumor and normal tissues after entering the term "HNRNPC" into the search box. From the available datasets, BRCA, ovarian cancer, colon cancer, clear cell renal cell carcinoma (ccRCC), uterine corpus endometrial carcinoma (UCEC), LUAD, and pediatric brain cancer were selected for further investigation and analysis.

Staining via immunohistochemistry

Immunohistochemical images of HNRNPC protein expression studies were performed in normal and 10 tumor tissues from the HPA, including liver cancer, lung cancer, BRCA, prostate cancer, and colon cancer, to examine variations in HNRNPC expression at the protein level. HPA51075 was the antibody used for IHC. The number of tumor samples with IHC was 10–12. Antibody staining in cancer types in contemporary human tissue is recorded as not observed, low, medium, or high in the HPA dataset. The intensity of the staining and the proportion of stained cells are used to calculate this score.

Analyses of survival prognoses

Researchers used the "Survival Map" module of the GEPIA2 software to produce the overall survival (OS) and disease-free survival (DFS) significance maps of HNRNPC over all TCGA tumors [22]. Using cutoff values of 50% for high expression and 50% for low expression, respectively, it was possible to distinguish between high and low expression groups. GEPIA2's "Survival Analysis" module was used to create survival plots, which together with the log-rank test were subsequently used to test our hypothesis.

Analysis of genetic variations

When we logged onto the cBioPortal website [24, 25], we were able to quickly obtain information on HNRNPC's genetic alteration features by selecting "TCGA pan-cancer Atlas Studies" from the "Quick select" area. This module summarized the mutation frequency, mutation type, and copy number alteration (CNA) statistics across all TCGA tumors in a single report. The "Mutations" module displays information on HNRNPC mutations in the form of a schematic depiction or as a three-dimensional (3D) representation

of the protein structure. Additionally, we will use the "Comparison" module to investigate whether there are any differences in OS, DFS, Progression-free survival (PFS), and Disease-specific survival between TCGA cancer patients who have and do not have HNRNPC genetic variations. Additionally, log-rank P-values were used to generate Kaplan-Meier plots, which were then graphically represented.

Immune infiltration analysis

We evaluated the connection between HNRNPC expression and immune infiltrates in all of the TCGA tumor samples by using the "Immune-Gene" module of the TIMER2 web server and its "Immune-Gene" module. Fibroblast cells and immune cells that have been linked to cancer were selected for the study. This was determined by the presence of immune infiltration utilizing a variety of algorithms, including the TIMER, CIBERSORT, CIBERSORT-ABS, QUANTISEQ, XCELL, MCPCOUNTER, and EPIC, among others. We were able to calculate the P-values and the partial correlation (Cor) values with the purity-adjusted Spearman's rank correlation test. Information was illustrated using a heatmap and a scatter plot, respectively.

Gene enrichment study for HNRNPC-related genes

Finding matches in the STRING database successfully required us to search for a single protein name ("HNRNPC") and a single organism ("Homo sapiens"). Because of this, we defined the following basic parameters: the lowest required interaction score ["Low confidence (0.150)," the meaning of network edges ("evidence"), the maximum number of interactors to display ("no more than 50 interactors" in the first shell), and active interaction sources ("experiments"). Furthermore, HNRNPC-binding proteins, which had already been found and experimentally extracted, were also identified and isolated.

Based on the datasets of all TCGA pan-cancer Atlas studies, we used the "Coexpression" module of the cBioPortal web to determine the top 200 HNRNPC correlated targeting genes. With the use of GEPIA2's "correlation analysis" module, we were able to conduct pairwise gene Pearson correlation research on HNRNPC and selected genes as well. The log2 TPM was used to generate the dot plot. The correlation coefficient (R) as well as the P-value were shown on the screen. Additionally, we utilized TIMER2's "Gene Corr" module to create heatmap data for the chosen genes, which contained the partial correlation (Cor) and the P-value for the purity-adjusted Spearman's rank correlation test. We also used the TIMER2 "Gene Corr"

module to generate heatmap data for the selected genes. We also used Jvenn, an interactive Venn diagram viewer [26] that allows us to perform an intersection analysis, to compare the HNRNPC-binding and interacting genes. In addition, we combined the two data sets to conduct a KEGG pathway analysis using the combined data (Kyoto encyclopedia of genes and genomes). We submitted the gene lists to KOBAS (KEGG Orthology Based Annotation System) with the required identifier ("Gene list enrichment") and species ("Homo sapiens") settings, and we obtained the functional annotation Visualization as a result. As part of the GO (Gene ontology) enrichment study, we also used Metascape.

RESULTS

HNRNPC expression in pan-cancer

In order to explore the relationship between HNRNPC and various cancers, we first examined whether there were differences in the expression levels of RNA and protein of HNRNPC. We investigated the oncogenic potential of HNRNPC (NC 000014.9 for DNA, NM 004500.4 for mRNA, and NP 004491.2 for protein) by evaluating its expression pattern in a variety of cells and nontumor tissues. According to GTEx, Illumina, BioGPS, and SAGE datasets, HNRNPC is expressed in all observed tissues (all consensus normalized expression levels >1) and has a low tissue selectivity for RNA. Additionally, substantial levels of HNRNPC were detected in the breast, lymph nodes, and ovary (Figure 1A). The expression status of HNRNPC was evaluated using the TIMER2 technique across the TCGA's diverse cancer types. The level of HNRNPC expression in BRCA, Cholangiocarcinoma (CHOL), Colon adenocarcinoma, Esophageal carcinoma (ESCA), Glioblastoma multiforme (GBM), Head and Neck squamous cell carcinoma (HNSC), Kidney chromophobe (KICH), and Liver hepatocellular carcinoma (LIHC) was also assessed (Figure 1B). To further validate the expression of HNRNPC RNA levels in pan-cancer, we checked the results of GTEx database again. The difference in HNRNPC expression between normal and tumor tissues (CHOL and KICH) was further analyzed using the GTEx dataset (Figure 1C, $P < 0.05$). Combined with TCGA database and GTEx database, HNRNPC is expressed significantly differently in CHOL and KICH. The CPTAC dataset revealed that the HNRNPC total protein expression was significantly greater in primary BRCA, ovarian cancer, colon cancer, UCEC, and LUAD tissues than in normal tissues (Figure 1D). Consequently, the IHC results from the HPA were processed and compared to the TCGA-provided HNRNPC gene expression information to evaluate

HNRNPC protein expression (Figure 1E). The results of the data analysis from the two databases were similar, and the HPA database filled in gaps in the TCGA data pertaining to HNRNPC expression in several normal tissues.

At the RNA level and protein level, HNRNPC showed significant differences between tumor tissue and normal tissue. These results support the possibility of HNRNPC as a carcinogen in many aspects. It will lay a foundation for further research on the carcinogenic effect of
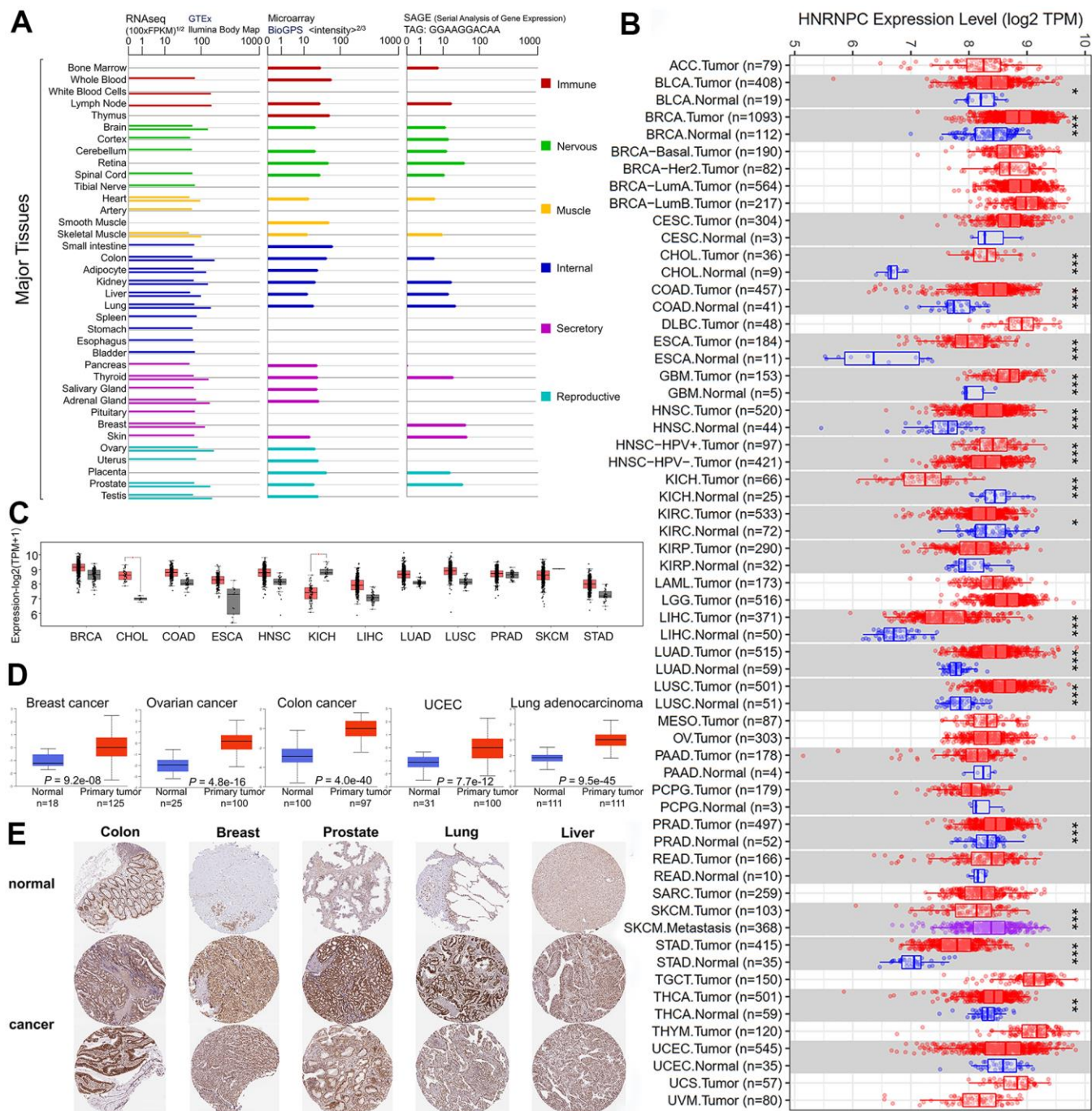

Figure 1. HNRNPC expression in different tumor type. (**A**) The expression levels of the HNRNPC gene varies according to tumor type. in GTEx, Illumina, BioGPS, and SAGE. (**B**) TIMER2 was used to determine the amount of expression of the HNRNPC gene in various malignancies or particular cancer subtypes. (**C**) The TCGA project provided box plot data for pan-cancer, as well as normal tissues from the GTEx database as controls. (**D**) The expression level of HNRNPC total protein was determined in normal and primary tissues using the CPTAC dataset. (**E**) HPA-induced expression of the HNRNPC protein in tumor and tumor-adjacent normal tissues. * P < 0.05; ** P < 0.01; *** P < 0.001.

HNRNPC. The HEPIA2 module "Pathological Stage Plot" was used to examine the link between HNRNPC expression and the pathological stages of various malignancies, such as adrenocortical carcinoma (ACC), CESC, HNSC, KICH, kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), LUAD, skin cutaneous melanoma (SKCM), testicular germ cell tumors (TGCT), and thyroid carcinoma (Figure 2). The above results suggest that HNRNPC may contribute to the development and progression of cancer by influencing the progression of tumor staging.

Relationship between prognosis and HNRNPC expression

To further explore whether the expression of HNRNPC will affect the survival of cancer patients, we analyzed the prognosis of OS and DFS in patients with high and low HNRNPC expression group. According to their levels of HNRNPC expression, cancer patients were distinguished into high- and low expression groups. The connection between HNRNPC expression and prognosis in patients with various types of malignancies was

subsequently examined using the TCGA and GEO datasets. The significantly increased HNRNPC was associated with a poor OS prognosis in the TCGA project for LUAD (P = 0.00011), CESC (P = 0.045), KIPR (P = 0.00072), ACC (P = 0.0033), Sarcoma (SARC) (P = 0.00083), and Pancreatic adenocarcinoma (PAAD) (P = 0.0058) (Figure 3A). The DFS study revealed a connection between high HNRNPC expression and a poor prognosis for LUAD (P = 0.037), ACC (P = 3.6e−05), and BLCA (P = 0.037) patients from the TCGA (Figure 3B). Furthermore, reduced HNRNPC expression was associated with a poor DFS prognosis for GBM (P = 0.014) and a poor OS prognosis for KIRC (P = 0.019) and THYM (P = 0.011).

Relationship between gene mutation characteristics and HNRNPC expression

To investigate whether HNRNPC causes cancer through gene mutation, we analyzed the mutation sites and types of HNRNPC in pan-cancer. The genetic status of HNRNPC was determined in several tumor samples from the TCGA cohorts. The highest frequency of HNRNPC change was 4.28 %, occurring in Ovarian
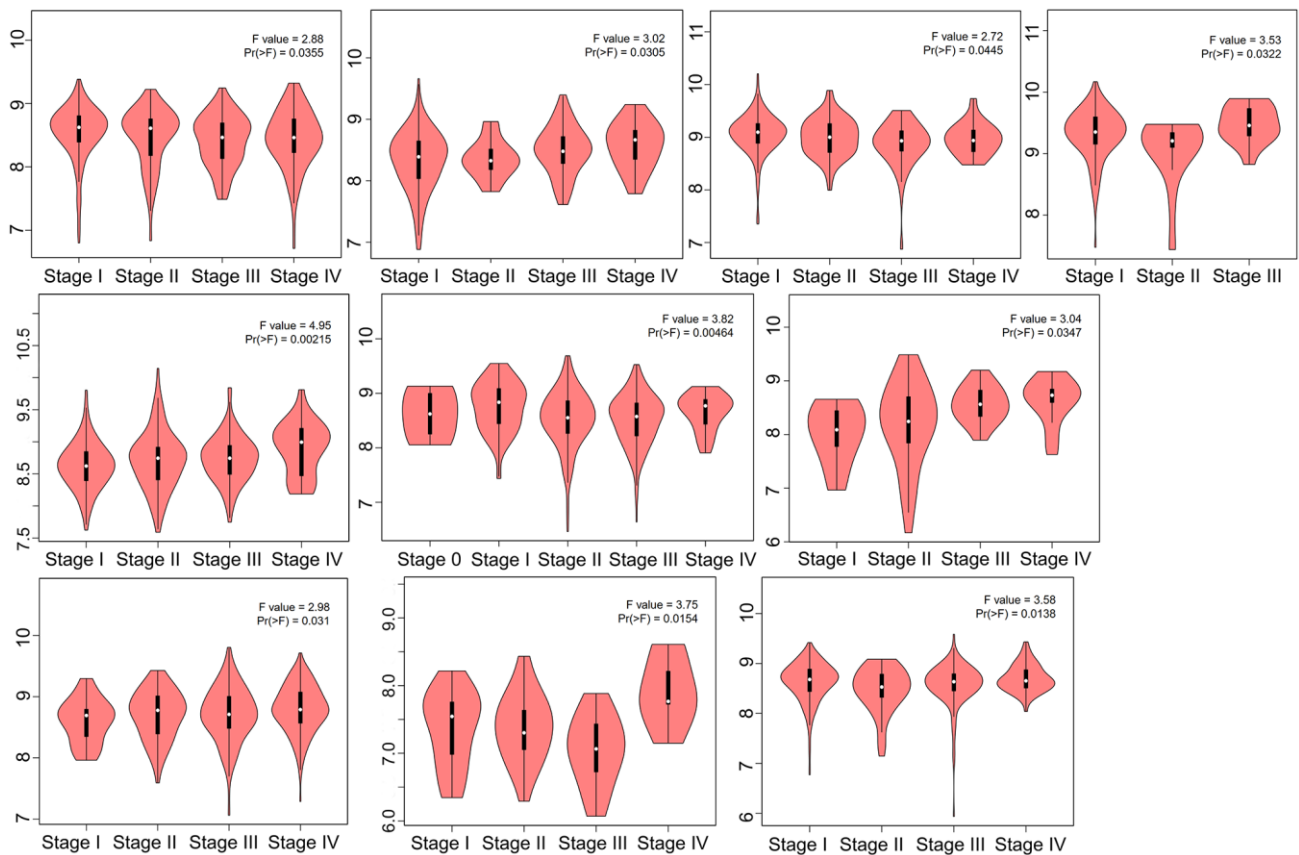

Figure 2. HNRNPC expression in different clinical stage. TCGA data were used to examine the expression level of the HNRNPC gene during the major pathogenic phases. For log-scale, Log2 (TPM+1) was used.

Serous Cystadenocarcinoma with "mutation" as the major type (Figure 4A). Additionally, the "amplification" kind of CNA had a much higher modification frequency of 3.42%. Additionally, it was noteworthy that all SARC, ACC, and ESCA patients with a genetic change (>1% frequency) showed HNRNPC CNA. Figure 4B depicts the types, sites, and case numbers of the HNRNPC genetic mutation. The primary form of the genetic mutation was a missense mutation in HNRNPC. Three cases of UCEC and one case of BRCA, respectively, had

an R167Q/* mutation in the Tudor domain. R167Q/* mutation is capable of causing a missense mutation in the HNRNPC gene, converting R (Arginine) to Q (Glutamine) at the 167 locations on the HNRNPC protein, and resulting in a missense HNRNPC protein. Regrettably, the R167 mutation cannot be mapped onto the 3D structure of HNRNPC (Figure 4C). Then, we further explored the effect of HNRNPC mutations on the prognosis of pan-cancer. The possible relationship between HNRNPC genetic variants and clinical survival
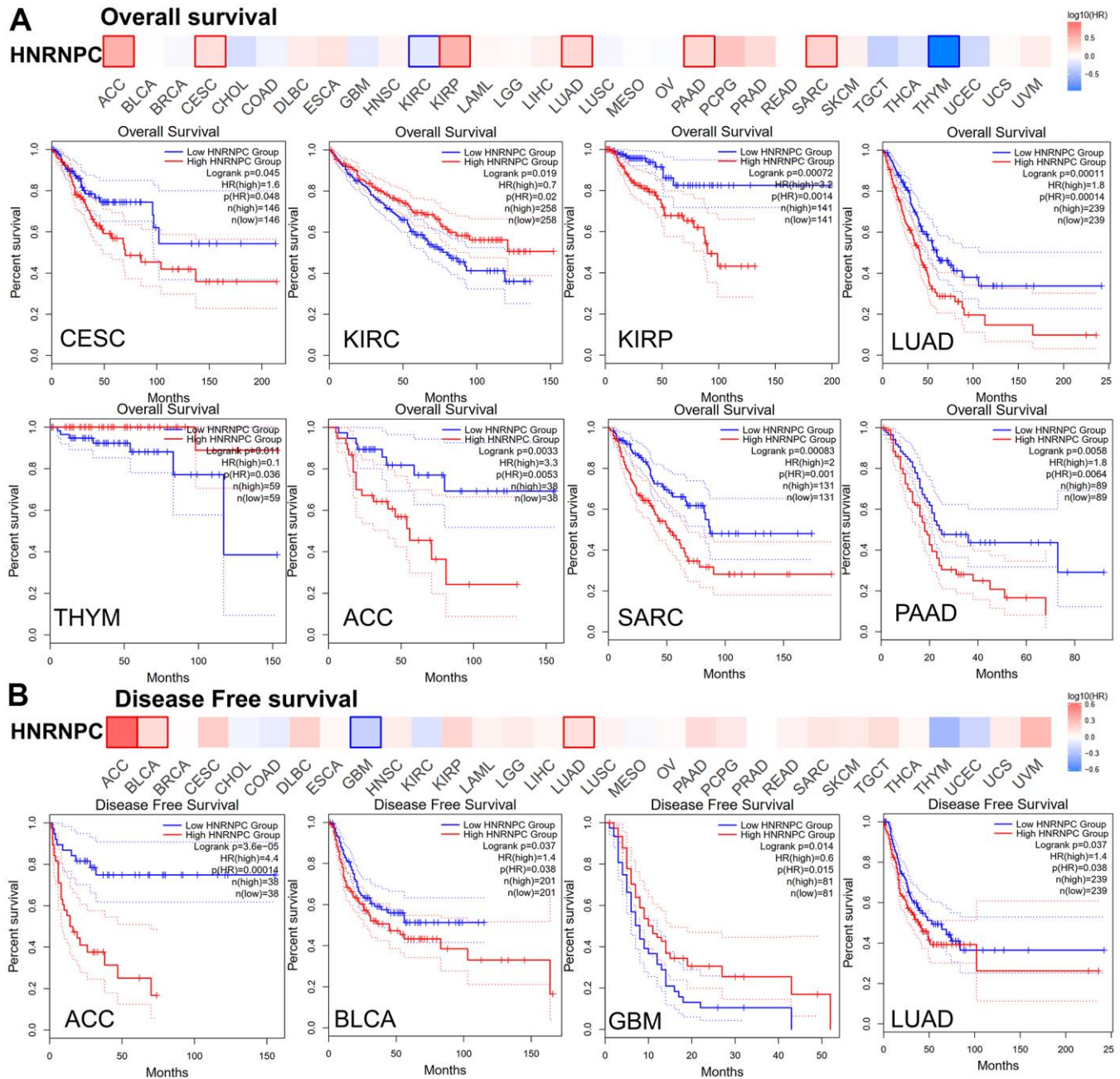
Figure 3. Relationship between HNRNPC gene expression and survival prognosis of cancers in TCGA. (A) overall survival and (B) disease-free survival.
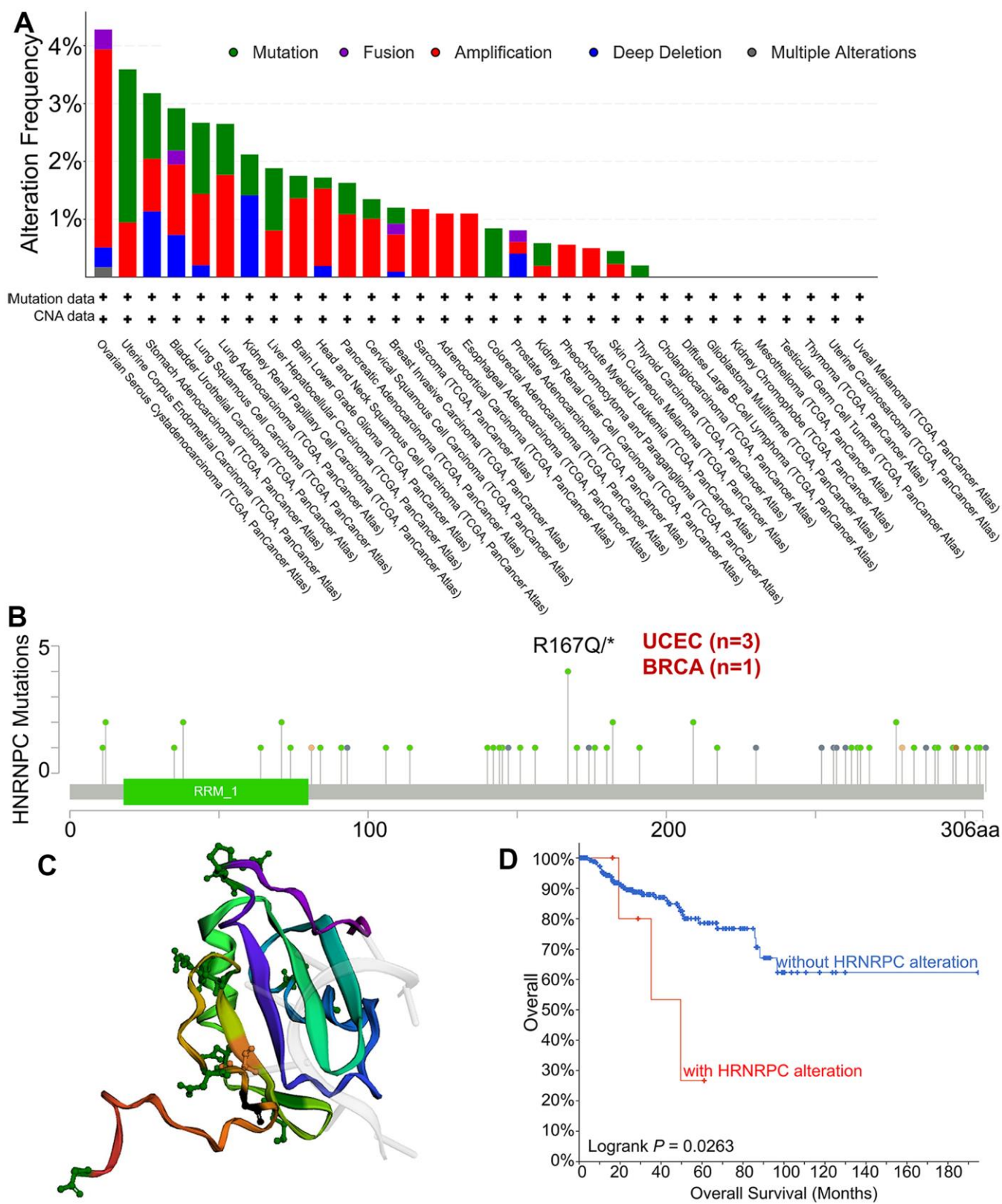
Figure 4. Mutation types and mutation sites of HNRNPC in TCGA tumors. (**A**) Mutation frequency and mutation type of HNRNPC in pan-cancer. (**B**) The mutation site of HNRNPC and the number of related cases at this mutation site. (**C**) The three-dimensional structure of HNRNPC. (**D**) the possible link between mutation status and overall survival in Kidney Renal Papillary Cell Carcinoma (P<0.05).

outcomes in patients with various types of cancer was investigated. In comparison to patients without the HNRNRPC mutation, KIPR demonstrated a worse prognosis (P = 0.0263) (Figure 4D).

Relationship between protein phosphorylation and HNRNPC expression

To clarify whether HNRNPC affects cancer through protein phosphorylation, we analyzed the existing protein phosphorylation database. By analyzing existing protein phosphorylation sites, we wanted to identify the most common mutation sites for HNRNPC. We compared the phosphorylation levels of HNRNPC in both normal and primary tumor tissues. Six distinct types of cancer were investigated using the CPTAC dataset (BRCA, colon cancer, ccRCC, LUAD, ovarian cancer, and UCEC). Figure 5 illustrates the HNRNPC phosphorylation sites and their substantial variations. In primary tumor tissues, the S220 of HNRNPC was more phosphorylated than in normal tissues (P < 0.05). This in turn raised the possibility that S220 phosphorylation plays a role in tumorigenesis.

Relationship between cancer associated fibroblasts and HNRNPC expression

Tumor-infiltrating immune cells have been shown to play an important role in the origin, development, and metastasis of cancer. Here, we explored the effect of HNRNPC on key components of the tumor microenvironment. We employed the TIMER, CIBERSORT, CIBERSORT-ABS, QUANTISEQ, XCELL, MCPCOUNTER, and EPIC algorithms to investigate the relationship between immune cell infiltration and HNRNPC gene expression in a range of TCGA cancer types. It has been shown that cancer-associated fibroblasts in the tumor microenvironment stroma have an influential role on the activity of various tumor-infiltrating immune cells. We checked the effect of HNRNPC on cancer-associated fibroblasts. A statistically significant negative relationship between HNRNPC expression and the estimated infiltration value of cancer-associated fibroblasts was identified in the TCGA tumors LUSC, TGCT, and THYM (Figure 6A). Scatter plots assessing the possible correlations among cancers in the TCGA are shown in Figure 6B using the EPIC algorithm. The degree of HNRNPC expression in LUSC was inversely correlated with the level of cancer-associated fibroblast infiltration (Cor = −0.133, P = 3.6e−03).

Relationship between molecular pathways and HNRNPC expression

To further elucidate the molecular mechanism by which the HNRNPC gene contributes to carcinogenesis,

we searched for HNRNPC-targeting proteins and HNRNPC expression-correlated genes using a variety of pathway enrichment studies. The STRING program was used to identify and validate 100 HNRNPC-binding proteins. The protein interaction network is shown in Figure 7A. The cBioPortal web server was used to retrieve all TCGA tumor expression data and the top 200 genes connected with HNRNPC expression. The associated heatmap data revealed a positive connection between HNRNPC and many genes associated with most types of cancers (Figure 7B). SMARCC1, KPNA4, and YWHAE were identified as common members of the aforementioned two groups by an intersection analysis. As seen in Figure 7C, the amount of HNRNPC expression was strongly linked with the SMARCC1 (R = 0.6), KPNA4 (R = 0.43), and YWHAE (R = 0.55) genes (all P < 0.001) (Figure 7D).

Additionally, the KEGG data indicate that the "VEGF signaling pathway" and the "RNA transport" may be implicated in the way HNRNPC influence tumor development (Figure 8A). Additionally, the GO enrichment analysis revealed that the majority of genes were engaged in metabolic and regulatory biological processes, including mRNA metabolism, nucleo-cytoplasmic transport, and RNA location (Figure 8B, 8C).

DISCUSSION

HNRNPC is a critical protein that plays a role in the cell cycle, especially in transcription and translation [27]. HNRNPC was shown to be involved in the initial steps of spliceosome assembly and pre-mRNA splicing [28]. The TreeFam website (http://www.treefam.org) was used in this study to demonstrate the conservation of HNRNPC structure across many species (Supplementary Figure 1). HNRNPC was shown to have comparable roles in a variety of animals. HNRNPC expression of DNA, RNA, and protein did not vary significantly among the various tissues investigated (Figure 1A). Single cell sequencing study revealed that HNRNPC is expressed in all organ cells (Supplementary Figure 2). As a result, no significant change in HNRNPC expression was observed between tissues and cells. Additionally, it may be suggested that an aberrant HNRNPC expression was a significant factor in abnormal cell behavior. Following that, the survival curve, protein phosphorylation, immunological infiltration, and molecular process will be discussed.

Survival prognosis curve

GEPIA2 was utilized to analyze HNRNPC expression and survival maps in lung adenocarcinoma tissues of the TCGA-LUAD project (Supplementary Figure 3).
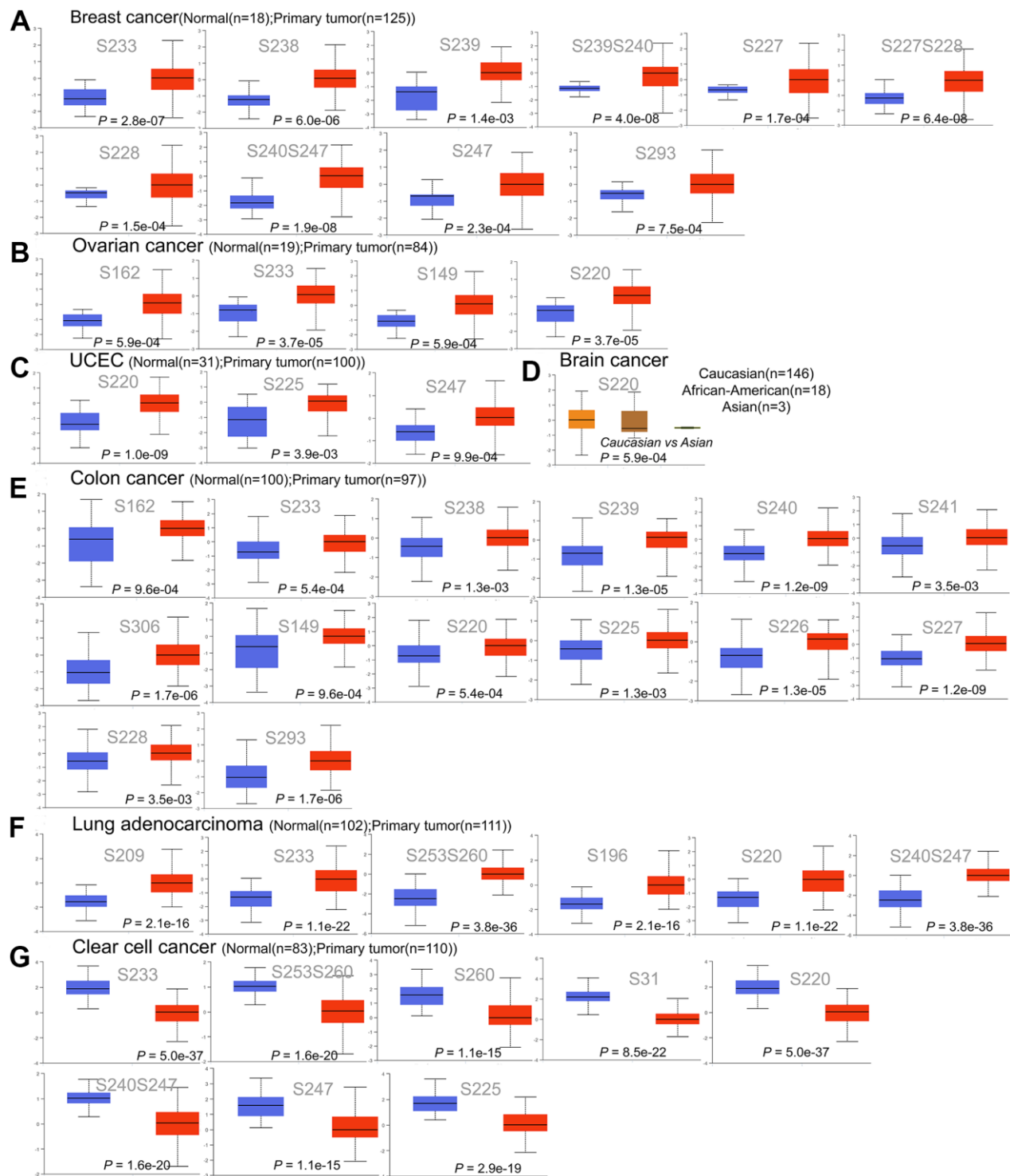
Figure 5. Phosphorylation of the HNRNPC protein in a variety of cancers. The UALCAN was used to compare the expression levels of HNRNPC phosphoprotein in normal and primary tissue using the CPTAC dataset. (**A**) Phosphorylation sites of HNRNPC proteins in breast cancer, including S233, S238, S239, S239S240, S227, S277S228, S228, S240S247, S247 and S293. (**B**) Phosphorylation sites of HNRNPC proteins in ovarian cancer, including S162, S233, S149 and S220. (**C**) Phosphorylation sites of HNRNPC proteins in UCEC, including S220, S225 and S247. (**D**) Phosphorylation sites of HNRNPC proteins in brain cancer, including S220. (**E**) Phosphorylation sites of HNRNPC proteins in colon cancer, including S162, S233, S238, S239, S240 and S241. (**F**) Phosphorylation sites of HNRNPC proteins in Lung adenocarcinoma, including S209, S233, S253S260, S196, S230 and S240S247. (**G**) Phosphorylation sites of HNRNPC proteins in clear cell cancer, including S233, S253S260, S260, S31, S220, S240S247, S247 and S255. Blue box plot: normal tissue; red box plot: tumor tissue; orange box plot: Caucasian; brown box plot: African-American; green box plot: Asian.

There was a strong association between HNRNPC expression in patients with LUAD (OS:P = 0.00011; DFS: P = 0.037). Then, using the Kaplan-Meier plotter, a series of survival analyses were conducted on the GSE200014/ GSE200751/ GSE216302/ GSE227110 cohorts. Our analysis revealed strong correlations between HNRNPC expression and clinical prognosis of OS and first progression-free survival. High HNRNPC

expression has been related to poor outcomes in LUAD [14, 29]. Additionally, surgical therapy for LUAD is beneficial.

Moreover, the Kaplan-Meier plotter was utilized to confirm and investigate gastric cancer (GC) (Supplementary Figure 4A, 4B). In all cohorts, GC prognosis was strongly associated with HNRNPC
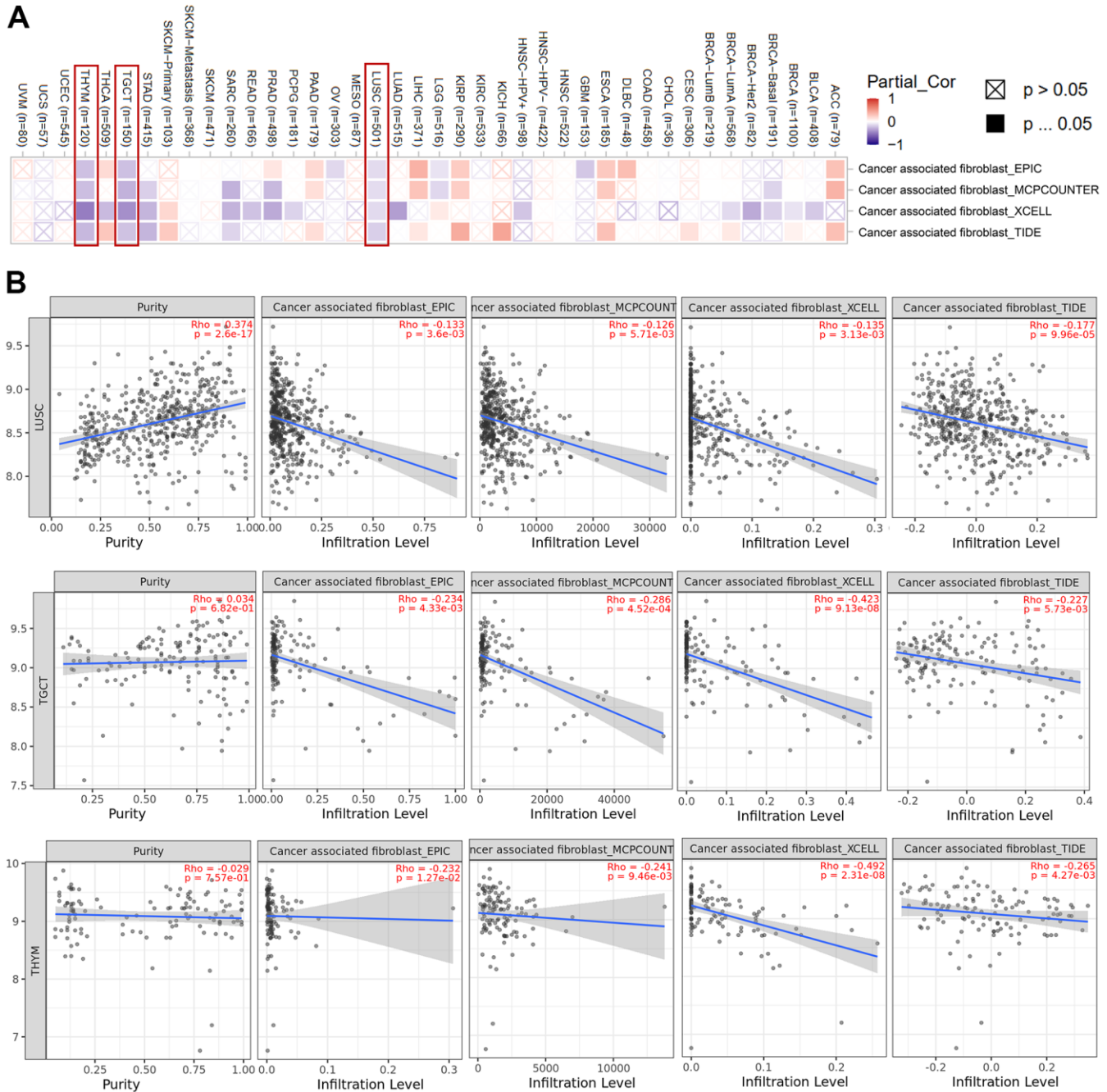
Figure 6. Cancer-associated fibroblasts associated with HNRNPC expression. (**A**) Heat map showing cancer-associated fibroblasts associated with HNRNPC expression in pan-cancer. (**B**) HNRNPC expression and cancer-associated fibroblasts estimation value scatter plots in LUSC, TGCT and THYM.

expression in OS and PFS. High HNRNPC levels are associated with poor OS and PFS. Chemoresistance may be promoted by overexpression of HNRNPC in parental cells [30]. Additionally, there was a link between HNRNPC expression and Her-2(−) status (Supplementary Figure 4C, 4D). According to earlier studies, the link between HNRNPC and Her-2 in GC has not significantly progressed. Patients in the GSE200751 cohort had disparate patterns in terms of ovarian cancer subtypes (Supplementary Figure 5).
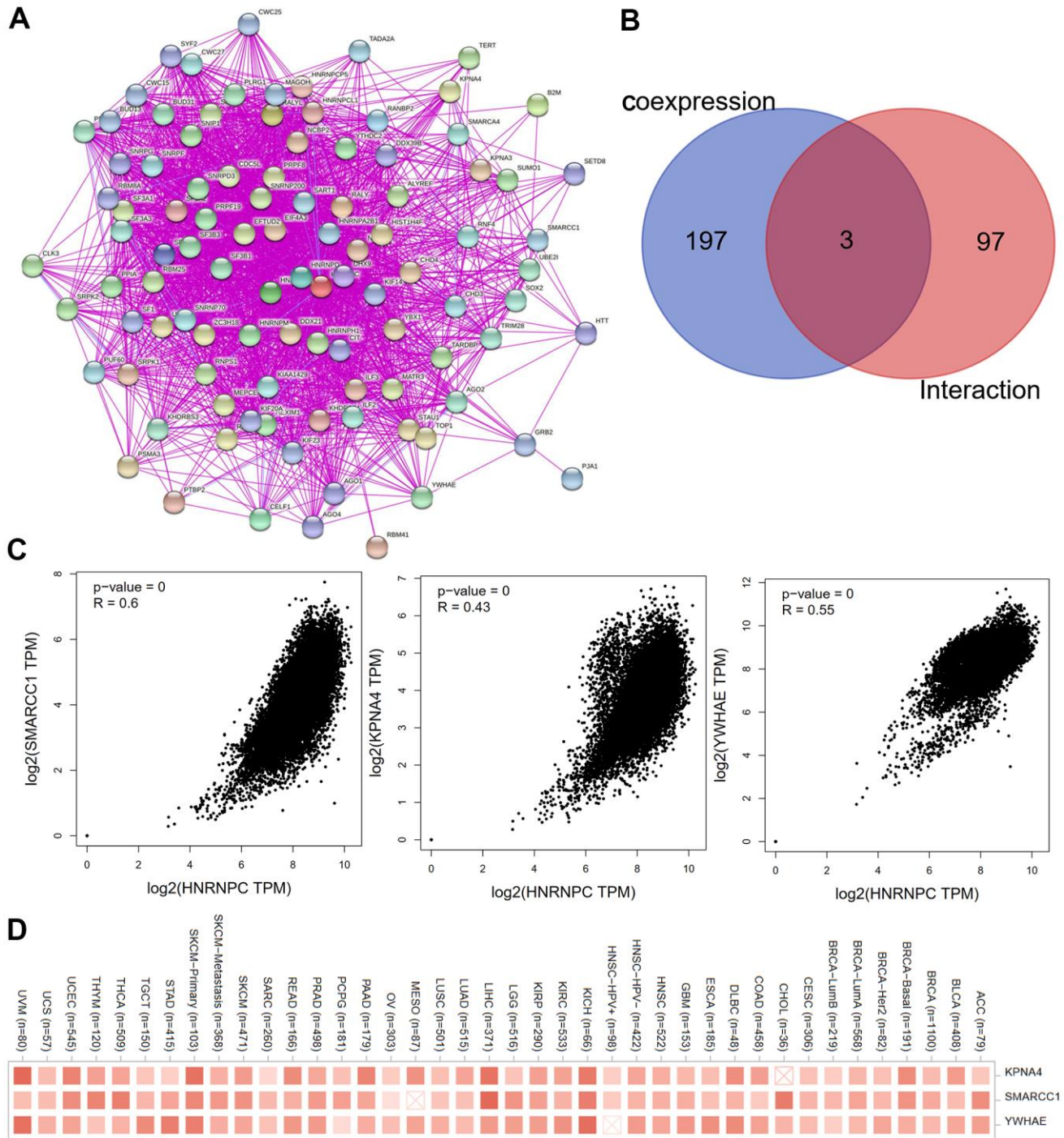

Figure 7. Screening for HNRNPC-related genes. (**A**) The interaction network of the first hundred experimentally determined HNRNPC-binding proteins by STRING. (**B**) Venn diagram was used to look at the intersection genes of the two databases (100 HNRNPC-binding proteins by STRING; 200 HNRNPC-related genes using the cBioPortal online). (**C**) The scatter plot was used to observe the correlation between KPNA4, SMARCC1, YWHAE and HNRNPC. (**D**) Corresponding heatmap data for KPNA4, SMARCC1, YWHAE in pan-cancer.
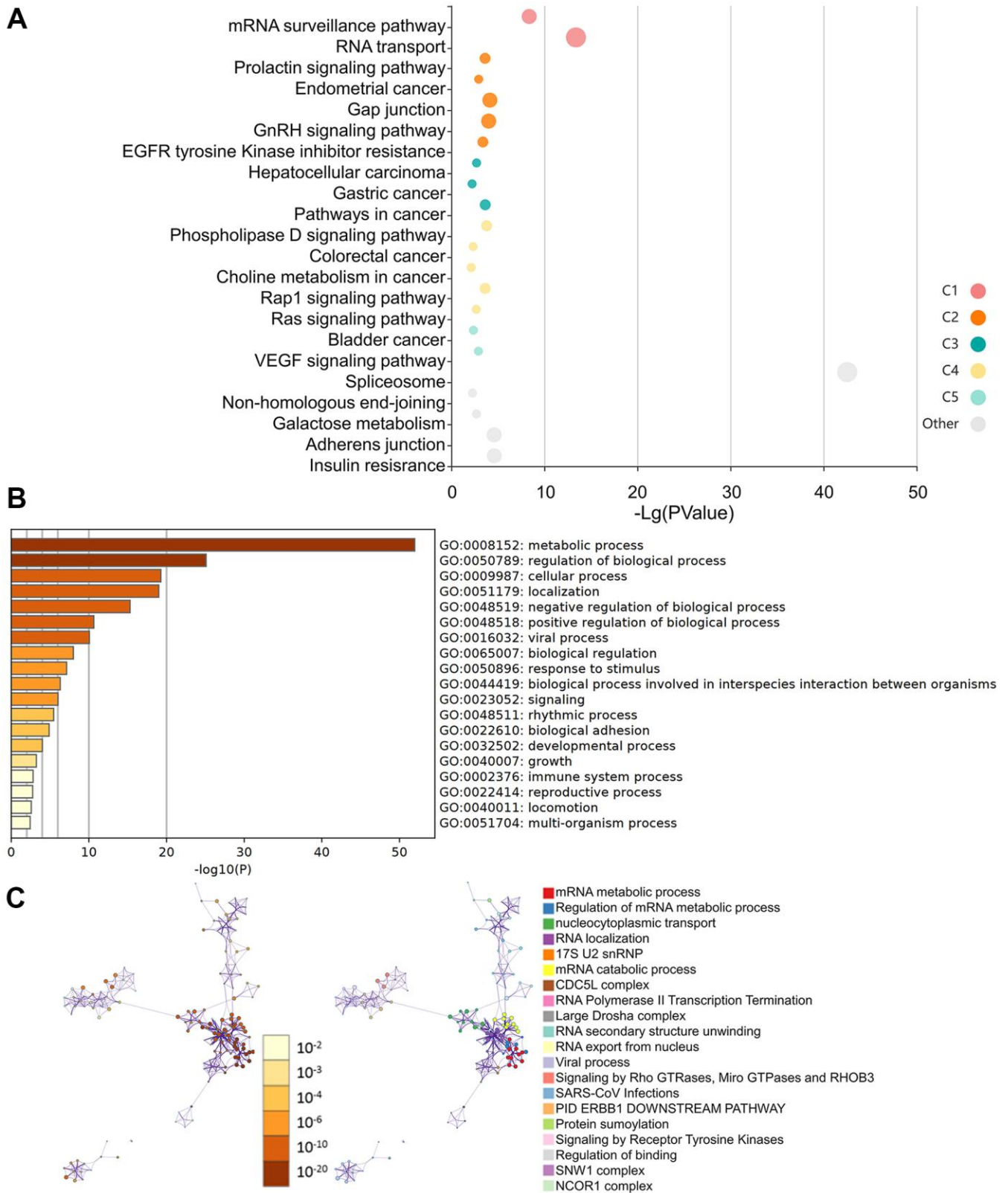
Figure 8. Molecular pathways involved in HNRNPC-related genes. (**A**) Analysis of KEGG pathways using HNRNPC-binding (STING database) and related genes (the cBioPortal online) with a bubble chart. (**B**) Bar chart illustrating the molecular function data for HNRNPC-binding (STING database) and related genes (the cBioPortal online). (**C**) Bar chart illustrating the molecular function data for HNRNPC-binding (STING database) and related genes (the cBioPortal online).

It has been found that enhanced expression of Mir-744-5p directly inhibited mRNA transcription and decreased nuclear factor IX (NFIX) and HNRNPC. The HNRNPC protein was previously reported to induce apoptosis. HNRNPC has been shown to inhibit mir-21 expression and phosphorylation of AKT in previous studies [26]. Additionally, it was discovered that Taxol had a greater effect on ovarian cancers with elevated HNRNPC expression in the survival curve (Supplementary Figure 5). Taxol may act as a selective HNRNPC inhibitor.

Phosphorylation of proteins

Using the CPTAC database, HNRNPC protein expression was found to be greater in cancerous than in normal tissues (Figure 1D). Additionally, HNRNPC contains a large number of phosphorylation sites that are statistically significant in many types of cancer (Figure 5). S220 expression was significantly different in all six malignancies. The S220 locus may operate as a phosphorylation site for HNRNPC activation in cancers. The cell cycle-regulated protein kinase has been shown to be involved in the phosphorylation of the HNRNPC protein. Phosphorylation of HNRNPC may regulate the formation and disassembly of HNRNP complexes, as well as the activity of RNA binding proteins and cell localization [31]. Meanwhile, it has been established that HNRNPC may bind to p53, inhibit its phosphorylation, and block the apoptotic pathway [32, 33]. In summary, S220 may be a key location of phosphorylation for HNRNPC proteins, and it may play a role in malignancy and cancer progression.

Immune infiltration

HNRNPC is intimately associated with the generation of Interferon γ (IFN-γ) in tumor cells, and it may play a significant role in the tumor microenvironment due to its unique immunological activity [9, 10].

HNRNPC expression was observed to be connected with tumor-associated fibroblast expression in LUSC, TGCT, and THYM (Figure 6A, 6B). Additionally, the expression of HNRNPC and CD8+ T cells was almost perfectly associated with uveal melanoma (UVM) (Supplementary Figure 6A). Similarly, HNRNPC expression in HNSC, HNSC-HPV-, SKCM, and SKCM-Metastasis was favorably linked with CD4+ T cells (Supplementary Figure 6B). HNRNPC expression was shown to be negatively correlated with T cell regulatory gene expression in TGCT (Supplementary Figure 6C). HNRNPC expression was strongly linked with B cells in LIHC and with monocytes in SKCM (Supplementary Figure 6E, 6F). Diverse macrophage types and HNRNPC expression patterns were observed (Supplementary Figure 6G). This might be because M0, M1, and M2 cells play distinct roles in the immunological inflammatory response. In summary, HNRNPC expression and immune cell infiltration patterns vary significantly among cancer types. The intricacy and variety of the immunological activity of HNRNPC will be a strategic challenge for future cancer research.

The molecular mechanism

In pan-cancer investigations, it was postulated for the first time that HNRNPC may be connected to the VEGF pathway (Figure 8A). Additionally, GO analysis revealed that HNRNPC's primary activities remained focused on mRNA metabolism control, nuclear cytoplasmic transport, and RNA localization (Figure 8B, 8C). In summary, the foregoing pan-cancer investigations established for the first time a statistically significant correlation between HNRNPC expression and clinical prognosis, protein phosphorylation, immune cell infiltration, and gene mutation. The subsequent examination of clinical tumor samples will provide fresh light on the role of HNRNPC in tumorigenesis.

Abbreviations

HNRNPC: heterogeneous nuclear ribonucleoprotein C; BRCA: breast cancer; LUAD: lung adenocarcinoma; TCGA: the Cancer Genome Atlas; UALCAN: the University of Alabama Cancer Database; CPTAC: The Clinical proteomic tumor analysis consortium; ccRCC: clear cell renal cell carcinoma; UCEC: uterine corpus endometrial carcinoma; OS: overall survival; DFS: disease-free survival; CAN: copy number alteration; PFS: Progression-free survival; 3D: three-dimensional; Cor: the partial correlation; CHOL: Cholangiocarcinoma; ESCA: Esophageal carcinoma; GBM: Glioblastoma multiforme; HNSC: Head and Neck squamous cell carcinoma; KICH: Kidney chromophobe; LICH: Liver hepatocellular carcinoma; ACC: adrenocortical carcinoma; KIRC: kidney renal clear cell carcinoma; KIRP: kidney renal papillary cell carcinoma; SKCM: skin cutaneous melanoma; TGCT: testicular germ cell tumors; SARC: Sarcoma; PAAD: Pancreatic adenocarcinoma; IFN-γ: Interferon γ; UYM: uveal melanoma.

AUTHOR CONTRIBUTIONS

Chenxi Pan: Methodology, Software, Data Curation, Writing - Original Draft. Qian Wu: Writing - Review and Editing, Supervision, Validation. Nianjie Feng: Writing - Review and Editing, Supervision, Project administration, Funding acquisition.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. 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.
https://doi.org/10.3322/caac.21492 PMID:30207593

2. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. CA Cancer J Clin. 2021; 71:7–33.
https://doi.org/10.3322/caac.21654 PMID:33433946

3. Choi YD, Grabowski PJ, Sharp PA, Dreyfuss G. Heterogeneous nuclear ribonucleoproteins: role in RNA splicing. Science. 1986; 231:1534–9.
https://doi.org/10.1126/science.3952495
PMID:3952495

4. Merrill BM, Barnett SF, LeStourgeon WM, Williams KR. Primary structure differences between proteins C1 and C2 of HeLa 40S nuclear ribonucleoprotein particles. Nucleic Acids Res. 1989; 17:8441–9.
https://doi.org/10.1093/nar/17.21.8441
PMID:2587210

5. König J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol. 2010; 17:909–15.
https://doi.org/10.1038/nsmb.1838 PMID:20601959

6. Zarnack K, König J, Tajnik M, Martincorena I, Eustermann S, Stévant I, Reyes A, Anders S, Luscombe NM, Ule J. Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. Cell. 2013; 152:453–66.
https://doi.org/10.1016/j.cell.2012.12.023
PMID:23374342

7. Geuens T, Bouhy D, Timmerman V. The hnRNP family: insights into their role in health and disease. Hum Genet. 2016; 135:851–67.
https://doi.org/10.1007/s00439-016-1683-5
PMID:27215579

8. Han SP, Tang YH, Smith R. Functional diversity of the hnRNPs: past, present and perspectives. Biochem J. 2010; 430:379–92.
https://doi.org/10.1042/BJ20100396
PMID:20795951

9. Chaudhury A, Chander P, Howe PH. Heterogeneous nuclear ribonucleoproteins (hnRNPs) in cellular processes: Focus on hnRNP E1's multifunctional regulatory roles. RNA. 2010; 16:1449–62.
https://doi.org/10.1261/rna.2254110 PMID:20584894

10. Wang S, Xu G, Chao F, Zhang C, Han D, Chen G. HNRNPC Promotes Proliferation, Metastasis and Predicts Prognosis in Prostate Cancer. Cancer Manag Res. 2021; 13:7263–76.
https://doi.org/10.2147/CMAR.S330713
PMID:34584453

11. Wu Y, Zhao W, Liu Y, Tan X, Li X, Zou Q, Xiao Z, Xu H, Wang Y, Yang X. Function of HNRNPC in breast cancer cells by controlling the dsRNA-induced interferon response. EMBO J. 2018; 37:e99017.
https://doi.org/10.15252/embj.201899017
PMID:30158112

12. Sarbanes SL, Le Pen J, Rice CM. Friend and foe, HNRNPC takes on immunostimulatory RNAs in breast cancer cells. EMBO J. 2018; 37:e100923.
https://doi.org/10.15252/embj.2018100923
PMID:30389667

13. Huang XT, Li JH, Zhu XX, Huang CS, Gao ZX, Xu QC, Zhao W, Yin XY. HNRNPC impedes m6A-dependent anti-metastatic alternative splicing events in pancreatic ductal adenocarcinoma. Cancer Lett. 2021; 518:196–206.
https://doi.org/10.1016/j.canlet.2021.07.016
PMID:34271104

14. Guo W, Huai Q, Zhang G, Guo L, Song P, Xue X, Tan F, Xue Q, Gao S, He J. Elevated Heterogeneous Nuclear Ribonucleoprotein C Expression Correlates With Poor Prognosis in Patients With Surgically Resected Lung Adenocarcinoma. Front Oncol. 2021; 10:598437.
https://doi.org/10.3389/fonc.2020.598437
PMID:33569346

15. Park YM, Hwang SJ, Masuda K, Choi KM, Jeong MR, Nam DH, Gorospe M, Kim HH. Heterogeneous nuclear ribonucleoprotein C1/C2 controls the metastatic potential of glioblastoma by regulating PDCD4. Mol Cell Biol. 2012; 32:4237–44.
https://doi.org/10.1128/MCB.00443-12
PMID:22907752

16. Sun DQ, Wang Y, Liu DG. Overexpression of hnRNPC2 induces multinucleation by repression of Aurora B in hepatocellular carcinoma cells. Oncol Lett. 2013; 5:1243–9.

https://doi.org/10.3892/ol.2013.1167
PMID:23599772

17. Anantha RW, Alcivar AL, Ma J, Cai H, Simhadri S, Ule J, König J, Xia B. Requirement of heterogeneous nuclear ribonucleoprotein C for BRCA gene expression and homologous recombination. PLoS One. 2013; 8:e61368.
https://doi.org/10.1371/journal.pone.0061368
PMID:23585894

18. Shetty S. Regulation of urokinase receptor mRNA stability by hnRNP C in lung epithelial cells. Mol Cell Biochem. 2005; 272:107–18.
https://doi.org/10.1007/s11010-005-7644-2
PMID:16010978

19. Yang Y, Shaffer AL 3rd, Emre NC, Ceribelli M, Zhang M, Wright G, Xiao W, Powell J, Platig J, Kohlhammer H, Young RM, Zhao H, Yang Y, et al. Exploiting synthetic lethality for the therapy of ABC diffuse large B cell lymphoma. Cancer Cell. 2012; 21:723–37.
https://doi.org/10.1016/j.ccr.2012.05.024
PMID:22698399

20. Kim JH, Paek KY, Choi K, Kim TD, Hahm B, Kim KT, Jang SK. Heterogeneous nuclear ribonucleoprotein C modulates translation of c-myc mRNA in a cell cycle phase-dependent manner. Mol Cell Biol. 2003; 23:708–20.
https://doi.org/10.1128/MCB.23.2.708-720.2003
PMID:12509468

21. Liu Z, Zhang S. Toward a systematic understanding of cancers: a survey of the pan-cancer study. Front Genet. 2014; 5:194.
https://doi.org/10.3389/fgene.2014.00194
PMID:25071824

22. Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res. 2019; 47:W556–60.
https://doi.org/10.1093/nar/gkz430
PMID:31114875

23. Chen F, Chandrashekar DS, Varambally S, Creighton CJ. Pan-cancer molecular subtypes revealed by mass-spectrometry-based proteomic characterization of more than 500 human cancers. Nat Commun. 2019; 10:5679.
https://doi.org/10.1038/s41467-019-13528-0
PMID:31831737

24. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013; 6:pl1.
https://doi.org/10.1126/scisignal.2004088
PMID:23550210

25. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012; 2:401–4.
https://doi.org/10.1158/2159-8290.CD-12-0095
PMID:22588877

26. Kleemann M, Schneider H, Unger K, Sander P, Schneider EM, Fischer-Posovszky P, Handrick R, Otte K. MiR-744-5p inducing cell death by directly targeting HNRNPC and NFIX in ovarian cancer cells. Sci Rep. 2018; 8:9020.
https://doi.org/10.1038/s41598-018-27438-6
PMID:29899543

27. Li H, Coghlan A, Ruan J, Coin LJ, Hériché JK, Osmotherly L, Li R, Liu T, Zhang Z, Bolund L, Wong GK, Zheng W, Dehal P, et al. TreeFam: a curated database of phylogenetic trees of animal gene families. Nucleic Acids Res. 2006; 34:D572–80.
https://doi.org/10.1093/nar/gkj118 PMID:16381935

28. Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T. N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature. 2015; 518:560–4.
https://doi.org/10.1038/nature14234 PMID:25719671

29. Zhu J, Wang M, Hu D. Deciphering N6-Methyladenosine-Related Genes Signature to Predict Survival in Lung Adenocarcinoma. Biomed Res Int. 2020; 2020:2514230.
https://doi.org/10.1155/2020/2514230
PMID:32258108

30. Huang H, Han Y, Zhang C, Wu J, Feng J, Qu L, Shou C. HNRNPC as a candidate biomarker for chemoresistance in gastric cancer. Tumour Biol. 2016; 37:3527–34.
https://doi.org/10.1007/s13277-015-4144-1
PMID:26453116

31. Piñol-Roma S, Dreyfuss G. Cell cycle-regulated phosphorylation of the pre-mRNA-binding (heterogeneous nuclear ribonucleoprotein) C proteins. Mol Cell Biol. 1993; 13:5762–70.
https://doi.org/10.1128/mcb.13.9.5762-5770.1993
PMID:8395012

32. Shen Y, Liu S, Fan J, Jin Y, Tian B, Zheng X, Fu H. Nuclear retention of the lncRNA SNHG1 by doxorubicin attenuates hnRNPC-p53 protein interactions. EMBO Rep. 2017; 18:536–48.
https://doi.org/10.15252/embr.201643139
PMID:28264987

33. Christian KJ, Lang MA, Raffalli-Mathieu F. Interaction of heterogeneous nuclear ribonucleoprotein C1/C2 with a novel cis-regulatory element within p53 mRNA as a
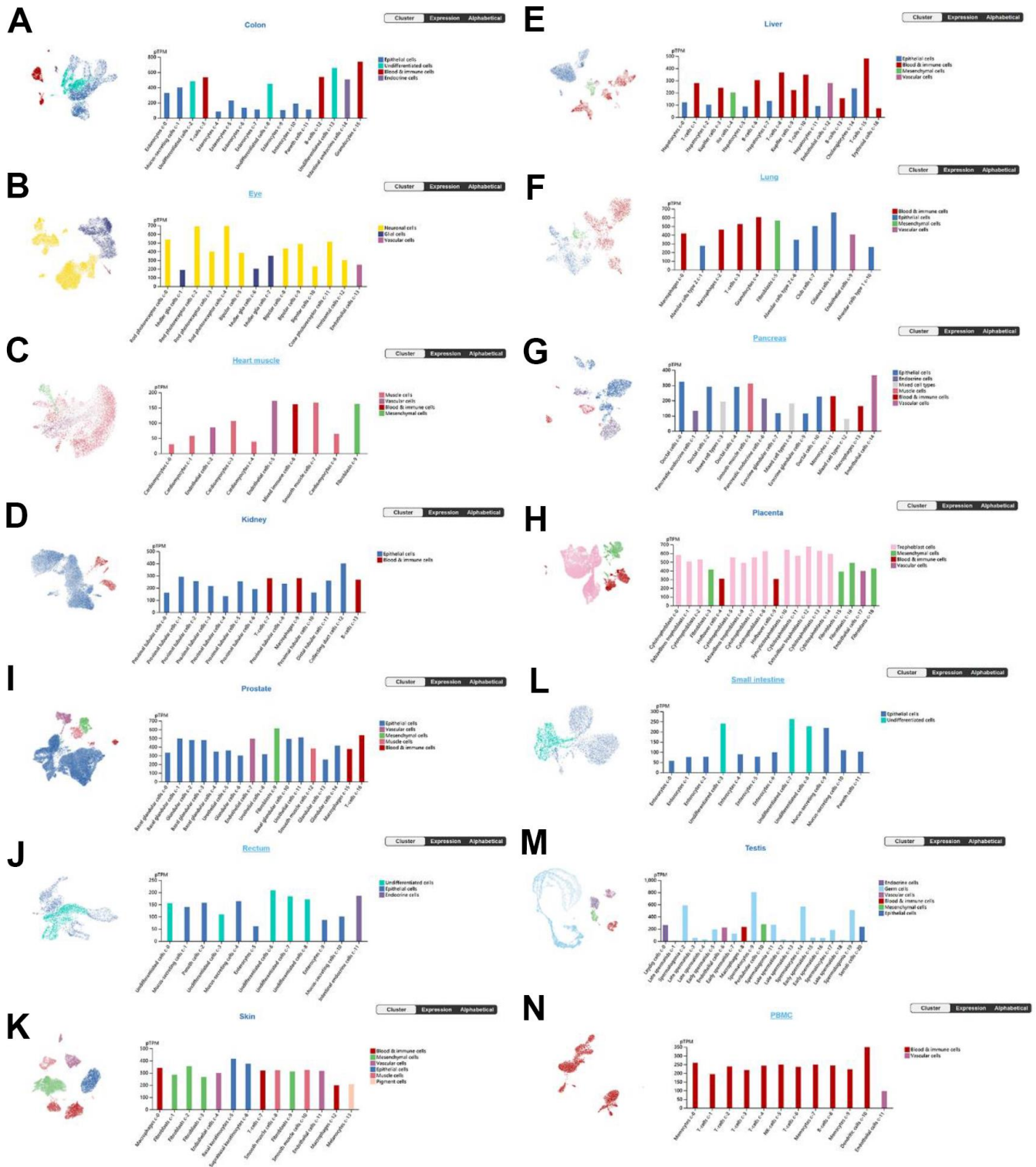
response to cytostatic drug treatment. Mol Pharmacol. 2008; 73:1558–67.
https://doi.org/10.1124/mol.107.042507
PMID:18296503
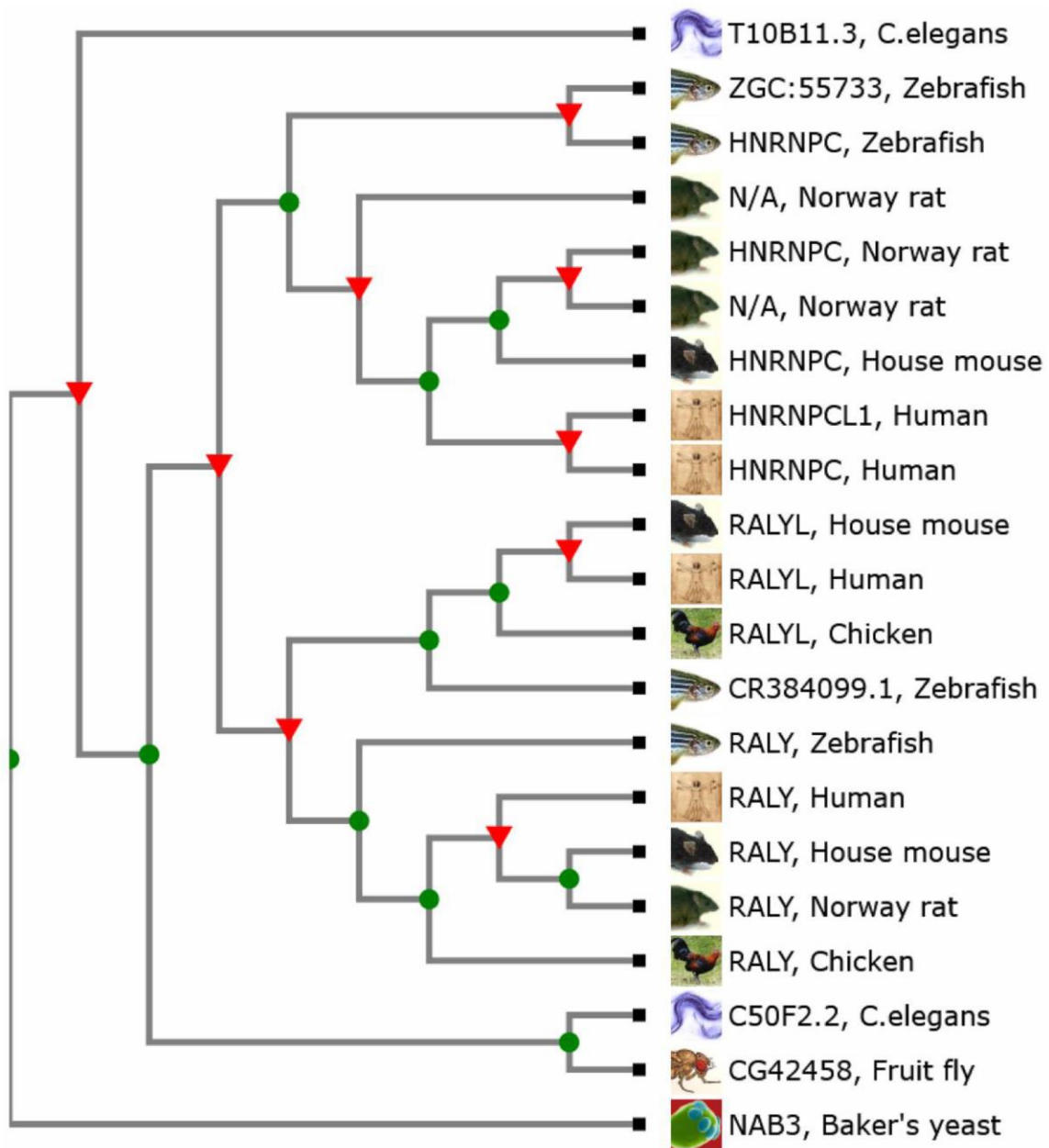
Supplementary Figure 1. Analysis the gene tree about HNRNPC in different races by TreeFam web (http://www.treefam.org).
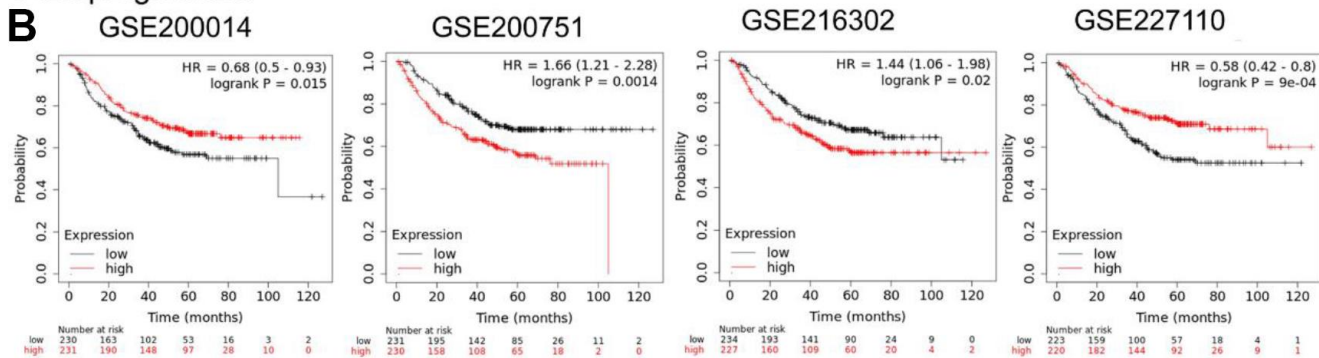
Supplementary Figure 2. Analysis the single cell types atlas in different organs by HPA database. (**A**) Colon, (**B**) Eye, (**C**) Heart muscle, (**D**) Kidney, (**E**) Liver, (**F**) Lung, (**G**) Pancreas, (**H**) Placenta, (**I**) Prostate, (**J**) Rectum, (**K**) Skin, (**L**) Small intestine, (**M**) Testis, (**N**) PBMC.

Lung adenocarcinoma
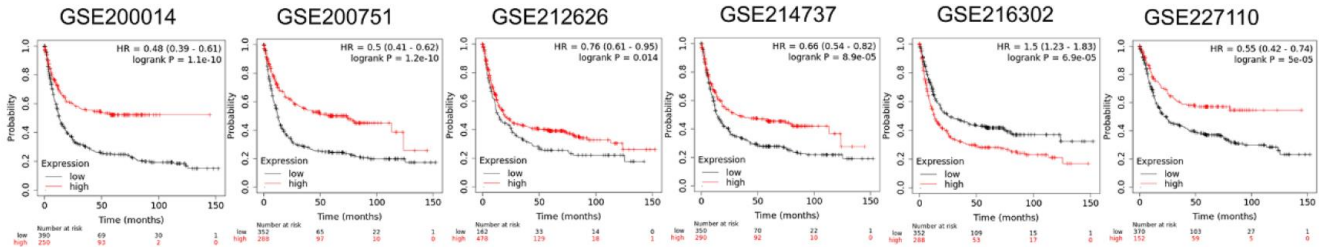## Overall survival
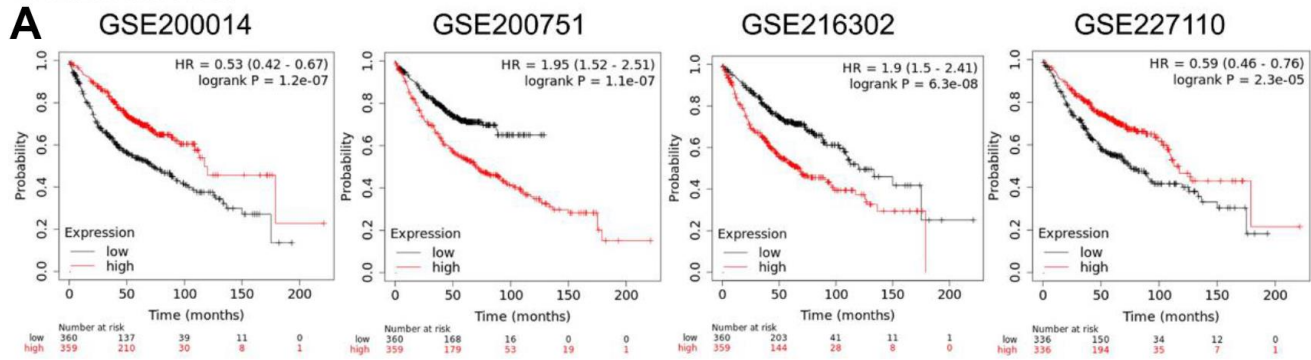
Supplementary Figure 3. Relationship between HNRNPC expression and prognosis of Lung adenocarcinoma, including (**A**) Overall survival progression, (**B**) First progression and (**C**) Only surgical margins negative progression.
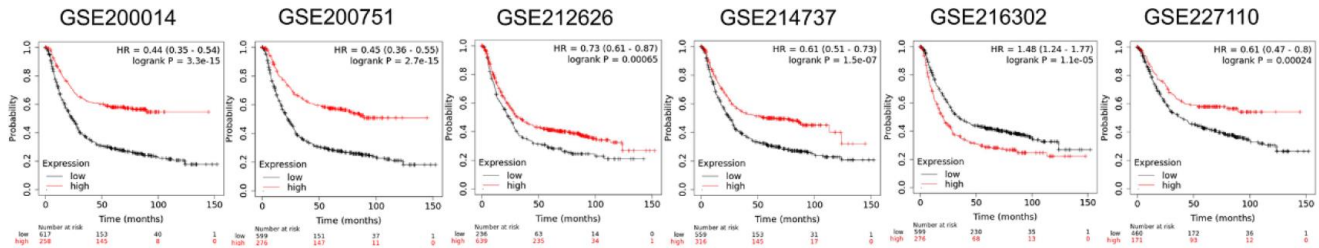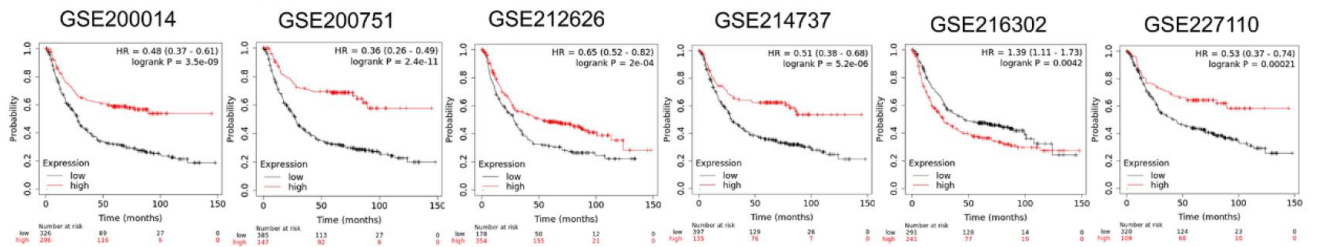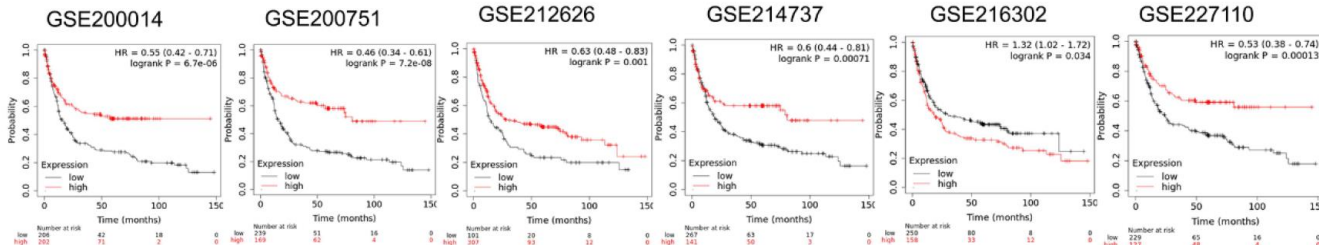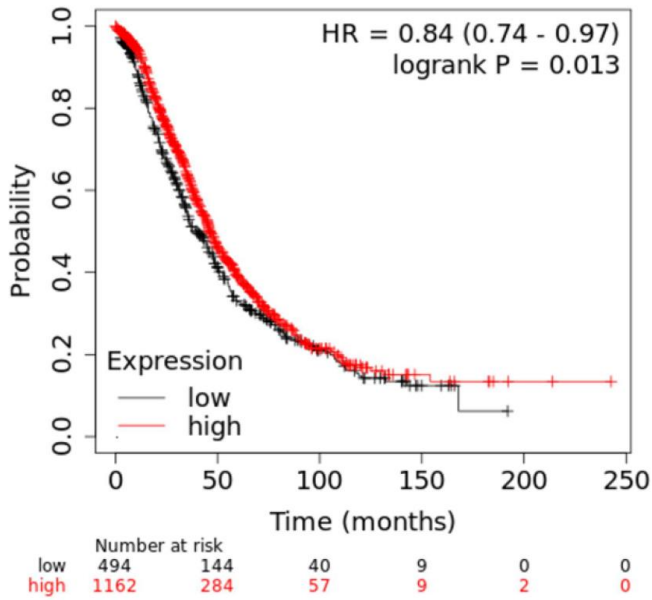
Supplementary Figure 4. Relationship between HNRNPC expression and prognosis of Gastric cancer, including (**A**) Overall survival progression, (**B**) First progression, (**C**) Overall survival progression-Her-2 and (**D**) First progression-Her-2.
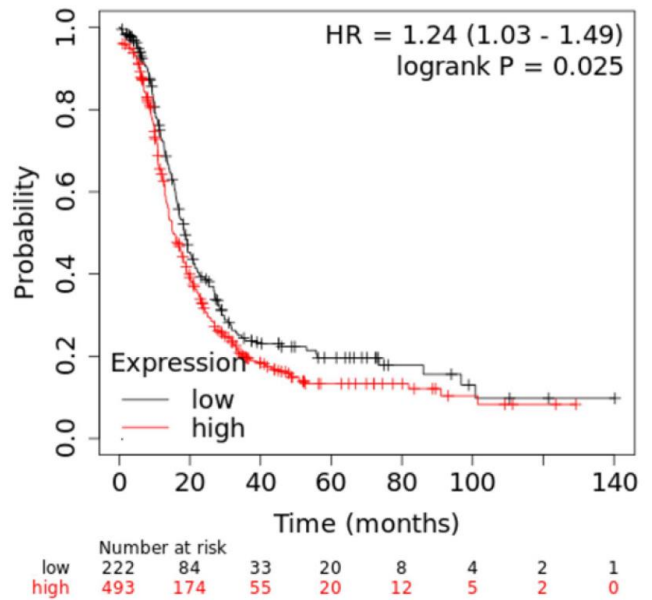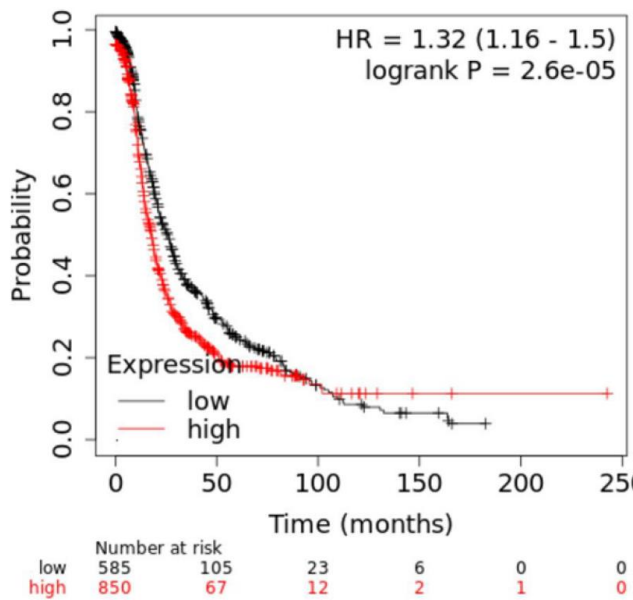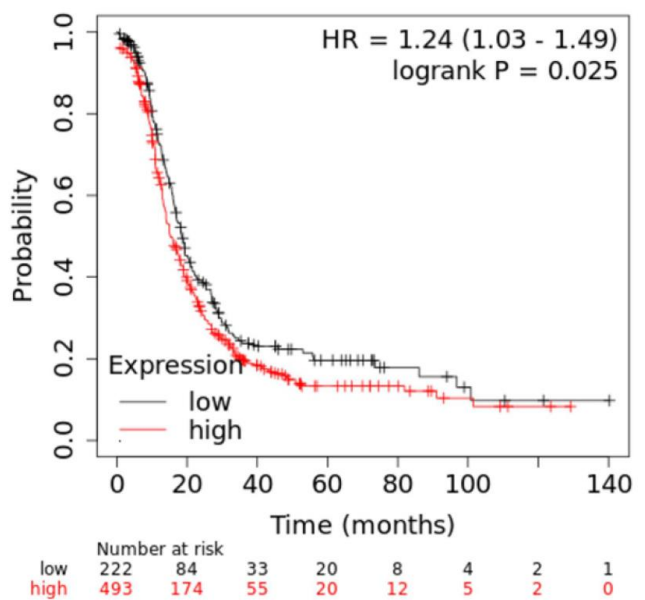
Ovarian cancer

GSE200751

Overall survival

HR = 0.84 (0.74 - 0.97)
logrank P = 0.013

Expression
— low
— high

Number at risk
low	494	144	40	9	0	0
high	1162	284	57	9	2	0

OS-Chemotherapy Taxol

HR = 1.24 (1.03 - 1.49)
logrank P = 0.025

Expression
— low
— high

Number at risk
low	222	84	33	20	8	4	2	1
high	493	174	55	20	12	5	2	0

Progression free survival

HR = 1.32 (1.16 - 1.5)
logrank P = 2.6e-05

Expression
— low
— high

Number at risk
low	585	105	23	6	0	0
high	850	67	12	2	1	0

PFS-Chemotherapy Taxol

HR = 1.24 (1.03 - 1.49)
logrank P = 0.025

Expression
— low
— high

Number at risk
low	222	84	33	20	8	4	2	1
high	493	174	55	20	12	5	2	0

Supplementary Figure 5. Relationship between HNRNPC expression and prognosis of ovarian cancer.
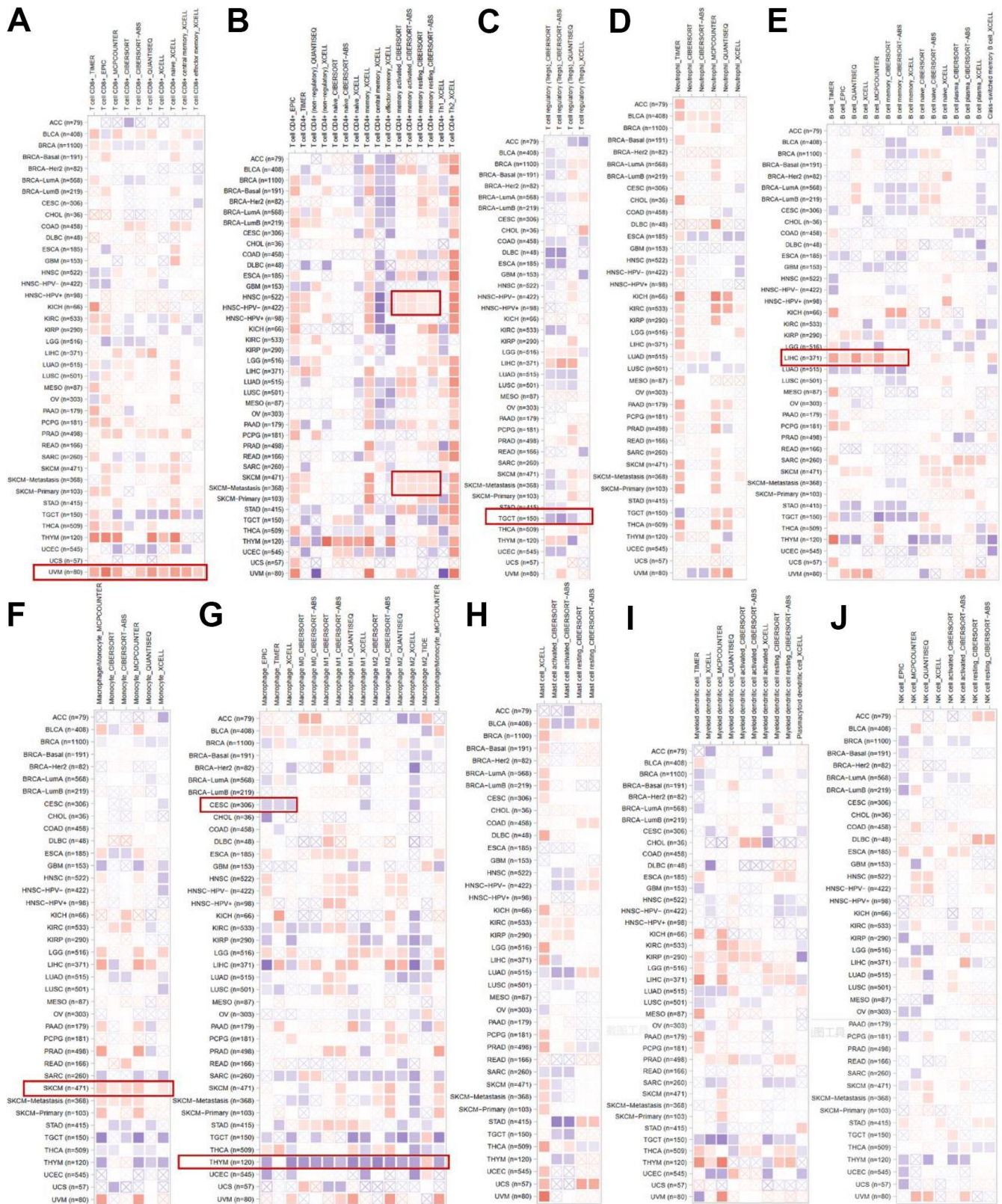
Supplementary Figure 6. Some algorithms were used to explore the potential correlation between immune cells and cancer in TCGA, such as (**A**) CD8+ T cell, (**B**) CD4+ T cell, (**C**) regulatory T cell, (**D**) Neutrophil, (**E**) B cell, (**F**) Monocyte, (**G**) Macrophage, (**H**) Mast cell, (**I**) Myeloid dendritic cell, (**J**) NK cell.