Comprehensive pan-cancer analysis of heat shock protein 110, 90, 70 and 60 families

Li-rong Yan
The First Affiliated Hospital of China Medical University

Shi-xuan Shen
The First Affiliated Hospital of China Medical University

Ang Wang
The First Affiliated Hospital of China Medical University

Yuan Yuan
The First Affiliated Hospital of China Medical University

Qian Xu ( qxu@cmu.edu.cn )
The First Hospital of China Medical University  https://orcid.org/0000-0001-8073-6656

Research Article

Keywords: heat shock proteins, expression, prognosis, variation, immune cell infiltration, pan-cancer

Posted Date: March 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-304845/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

**Background:** Heat shock proteins (HSPs) are a class of molecular chaperones that function in protein folding and maintain protein structure and function. No study has performed a comprehensive pan-cancer analysis of HSPs. Here we carried out a panoramic analysis of the expression and prognosis of HSP110, HSP90, HSP70 and HSP60 families in 33 tumors, with the aim of deepening the systematic understanding of HSPs in cancer.

**Materials and Methods:** Next-generation sequencing data of multiple tumors were downloaded from TCGA, CCLE and Oncomine databases. RStudio 3.6.1 was used to analyze the expressions, mutations, copy number variations (CNVs), cancer-related signaling pathways, immune cell infiltration and prognosis profiles of HSP110, HSP90, HSP70, and HSP60 families in 33 tumors.

**Results:** HSPA6 and HSPA7 were generally highly expressed while HSPA12A, HSPA12B, HSPA2 and HSPA4L were mainly expressed at low levels in different cancer tissues. The results revealed mainly positive correlations among the expressions of HSPs in different cancers. Expressions of HSP family members were generally associated with poor prognosis in respiratory, digestive, urinary, and reproductive system tumors and associated with good prognosis in cholangiocarcinoma, pheochromocytoma and paraganglioma. TCGA mutation analysis showed that HSP gene mutation rate in cancers was 0%–23%. CCLE mutation analysis indicated that HSP gene mutation rate in 828 cell lines from 15 tumors was 0%–17%. CNV analysis revealed that HSPs have different degrees of gene amplifications and deletions in cancers. Gene mutations of 15 HSPs influenced their protein expressions in different cancers. Copy number amplifications and deletions of 22 HSPs also impacted protein expression levels in pan-cancer. HSP gene mutation was generally a poor prognosis factor in cancers, except for uterine corpus endometrial carcinoma. CNVs in 14 HSPs showed varying influences on survival status in different cancers. HSPs may be involved in the activation and inhibition of multiple cancer-related pathways. HSP expressions were closely correlated with 22 immune cell infiltrations in different cancers.

**Conclusion:** Our results show that HSP families play an important role in the occurrence and development of various tumors and are potential tumor diagnostic and prognostic biomarkers as well as anti-cancer therapeutic targets.

Introduction

The global cancer incidence and mortality rates have risen rapidly in the 21st century, and cancer has become the main cause of death in many countries around the world(1). The main causes of the high mortality rate of cancer lie in the unknown pathogenic mechanisms and the lack of effective treatment protocols. Therefore, exploring the pathogenic mechanism of cancer is critical to develop strategies for cancer treatment.

Heat shock proteins (HSPs) are a class of molecular chaperones that help protein folding and maintain the normal structures and functions of proteins(2). HSPs protect cells from physical and chemical stimuli and stress, such as ischemia, hypoxia, high temperature, metabolic factors, alcohol and drugs, thus maintaining cell homeostasis(3, 4). On the basis of their molecular weight, HSPs are divided into different families, including the HSP110(HSPH) family, HSP90 family, HSP70 (HSPA) family, HSP60 family, HSP40 (DNAJ) family and small heat shock proteins(2, 5). Some research has suggested that different families not only independently regulate the structures and functions of proteins, but also collaborate with each other. For instance, protein folding and degradation are synergistically regulated by the HSP70 and HSP90 families(6, 7). The HSP60 precursor is transformed into mature HSP60 complexes under the cooperating of HSP70 and HSP10(8). HSP family functions and jointly role was critical to sustain homeostasis.

Multiple studies have indicated that HSPs are involved in the occurrence and development of tumors. HSPs modulate cell proliferation, angiogenesis, and the migration, invasion and metastasis of tumor cells, as well as the resistance of tumors to anti-cancer drugs(9-14). Several studies have also suggested HSPs as potential diagnostic and prognostic biomarkers as well as therapeutic targets of tumors(9-14). In addition, mutations in HSPs may have an influence on cancer risk and prognosis(15-
In conclusion, HSPs play an important role in tumor progression when the body is in an unbalanced state of tumorigenesis.

At present, no study has performed a panoramic analysis of HSPs in 33 tumors. In this study, we used The Cancer Genome Atlas (TCGA), Cancer Cell Line Encyclopedia (CCLE), Oncomine and The Human Protein Atlas (THPA) databases to analyze the expressions, mutations, copy number variations (CNVs) and prognosis profiles of HSP110, HSP90, HSP70 and HSP60 families in cancer. We investigated the potential correlation of HSP expressions with cancer-related pathways, immune cell infiltration, and prognosis in pan-cancer, as well as the relationship between HSP mutations and CNVs with HSP expressions. Our study aims include analyzing HSP profiles in different cancers, deepening HSP systematic recognitions, finding potential diagnostic and prognostic biomarkers as well as anti-cancer therapeutic targets and providing important clues for further mechanism research.

Materials And Methods

1. Data collection at tissue, protein and cell levels

The HSP expressions, mutations and CNV data in 33 tumors were downloaded from TCGA database (http://cancergenome.nih.gov/). Clinical information, such as survival status and survival period, was downloaded from UCSC XENA (https://xenabrowser.net/). The details of 33 tumors collected from TCGA were as follows: ACC, Adrenocortical carcinoma; BLCA, Bladder Urothelial Carcinoma; BRCA, Breast invasive carcinoma; CESC, Cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, Cholangiocarcinoma; COAD, Colon adenocarcinoma; DLBC, Lymphoid Neoplasm Diffuse Large B-cell Lymphoma; ESCA, Esophageal carcinoma; GBM, Glioblastoma multiforme; HNSC, Head and Neck squamous cell carcinoma; KICH, Kidney Chromophobe; KIRC, Kidney renal clear cell carcinoma; KIRP, Kidney renal papillary cell carcinoma; LAML, Acute Myeloid Leukemia; LGG, Brain Lower Grade Glioma; LIHC, Liver hepatocellular carcinoma; LUAD, Lung adenocarcinoma; LUSC, Lung squamous cell carcinoma; MESO, Mesothelioma; OV, Ovarian serous cystadenocarcinoma; PAAD, Pancreatic adenocarcinoma; PCPG, Pheochromocytoma, and Paraganglioma; PRAD, Prostate adenocarcinoma; READ, Rectum adenocarcinoma; SARC, Sarcoma; SKCM, Skin Cutaneous Melanoma; STAD, Stomach adenocarcinoma; TGCT, Testicular Germ Cell Tumors; THCA, Thyroid carcinoma; THYM, Thymoma; UCEC, Uterine Corpus Endometrial Carcinoma; UCS, Uterine Carcinosarcoma; UVM, Uveal Melanoma. In addition, expression data of HSP mRNA levels in different cancers were downloaded from the Oncomine database(18). Immunohistochemistry data of HSPs in 16 tumors were collected from THPA database (https://www.proteinatlas.org/). The HSP expressions and mutations data of 828 cell lines from 15 tumors were downloaded from the CCLE database (https://portals.broadinstitute.org/ccle)(19).

2. Comprehensive analysis of HSP expressions in pan-cancer

2.1 Analysis of HSP mRNA levels in pan-cancer tissues and cells

Deseq2 R package was used to analyze expression data in TCGA to identify differentially expressed HSP family members in pan-cancer tissues. \( p < 0.05 \) was considered statistically significant. The Oncomine database was applied to verify the differentially expressed HSP mRNAs in different cancer tissues. The cut-off criteria were \( p < 0.05 \), \(|\log_2 \text{fold change}| \geq 2\) and top 10% gene rank.

We further examined the expression profiles of HSPs in different tumor cell lines using the CCLE database. Kruskal–Wallis rank test was used to compare expressions of HSPs in pan-cancer cell lines. \( p < 0.05 \) was considered statistically significant.

2.2 Analysis of HSP protein levels in pan-cancer tissues

Immunohistochemistry data of HSPs were collected from THPA database to evaluate their expression profiles at protein level in pan-cancer. The cancer types were as follows: glioma, lung cancer, colorectal cancer, testis cancer, renal cancer, head and neck cancer, stomach cancer, pancreatic cancer, lymphoma, ovarian cancer, skin cancer, breast cancer, liver cancer, endometrial cancer, melanoma, thyroid cancer.
2.3 Correlation analysis between HSP expressions and cancer-related pathways

Gene expression datasets were used to perform gene set variation analysis to identify potential signal transduction pathways correlated with the HSP family(20, 21). Pearson's correlation coefficient between HSP expressions and cancer-related pathways was calculated to evaluate the association of HSP expression with pathway activity. \( |r| \geq 0.2 \) and \( p < 0.05 \) were regarded as cut-off criteria.

2.4 Correlation analysis between HSP expressions and immune cell infiltration in pan-cancer

Spearman correlation analysis was performed to evaluate the relationships of HSP family expression with infiltration of 22 immune cell types. The cut-off value was \( |r| \geq 0.3 \) and \( p < 0.05 \). RStudio 3.6.1 was used for data analysis and Cytoscape 3.7.1 was applied for visualization.

2.5 Correlation and interaction of HSP expressions in pan-cancer

HSP family members often cooperate to exert their cellular functions(8, 22). We determined Pearson's correlation coefficient of HSP expressions in different cancers using TCGA data. Ggcorrplot and ggthemes R package were used to perform detailed analysis and visualization. Search Tool for the Retrieval of Interacting Genes (STRING, https://string-db.org/) was applied to predict the network to explore the potential interaction relationships of HSP family members. Cytoscape 3.7.2 was used to perform analysis and visualization. Degree score was determined to evaluate interaction strength among HSPs.

2.6 Correlation analysis between HSP family expression and prognosis in pan-cancer

We analyzed the correlation of HSPs expressions with prognosis using gene expression and clinical information in TCGA. Patients were divided into two groups according to the median expression of HSPs and log-rank test was performed. \( p < 0.05 \) was considered to be statistically significant.

3. Analysis of HSP gene mutations and CNVs in pan-cancer tissues and cells

3.1 HSP gene mutations in pan-cancer tissues and cells

Mutation data were downloaded from TCGA and CCLE databases. RStudio 3.6.1 was used to calculate the gene mutation frequency of HSP molecules in tissues and cell lines. The mutation frequency was defined as mutation proportion in each type of cancer.

3.2 HSP gene CNVs in pan-cancer tissues

CNV data were downloaded from TCGA database. RStudio 3.6.1 was used to calculate the copy number amplification and deletion frequency in tissues. The CNV frequency was defined as CNV proportion in each type of cancer.

3.3 Correlation analysis between HSP gene mutations and CNVs with expressions in pan-cancer

Mann–Whitney U test was applied to statistically identify the influence of gene mutations and CNVs on HSP expressions in pan-cancer. \( p < 0.05 \) was considered to be statistically significant. All analyses were performed by RStudio 3.6.1.

3.4 Correlation analysis between HSP gene mutations and CNVs with prognosis in pan-cancer

Patients were divided into two groups according to the median variation frequency of HSPs to identify the influence of mutations and CNVs on overall survival in pan-cancer. Log-rank test was performed by RStudio 3.6.1. \( p < 0.05 \) was considered to be statistically significant.

Results

1. HSP expression profiles in pan-cancer
1.1 HSP mRNA expressions in pan-cancer

On the basis of a literature review, 22 HSPs among HSP110, HSP90, HSP70 and HSP60 families were selected for analysis (Table 1) (5). We examined the mRNA expression of these 22 HSPs in cancer and non-tumor tissues in TCGA and found that mRNA levels of 10 HSPs were differentially expressed in 11 tumors (p < 0.05) (Figure 1A). HSPA2 mRNA was down-regulated in stomach adenocarcinoma, colon adenocarcinoma, bladder urothelial carcinoma, kidney renal clear cell carcinoma, kidney renal papillary cell carcinoma and kidney chromophobe (Figure 1B). The mRNA expression level of HSPA7 was decreased in colon adenocarcinoma but increased in lung adenocarcinoma, kidney renal clear cell carcinoma and kidney renal papillary cell carcinoma. HSPA1A mRNA was down-regulated in colon adenocarcinoma. HSPA6 mRNA was up-regulated in breast cancer, kidney renal clear cell carcinoma and kidney renal papillary cell carcinoma. HSPA4L mRNA expression was increased in lung squamous cell carcinoma and breast cancer and decreased in prostate adenocarcinoma. HSPA1L mRNA was down-regulated in uterine corpus endometrial carcinoma. HSPA12A mRNA was down-regulated in bladder urothelial carcinoma, breast cancer, uterine corpus endometrial carcinoma and kidney renal papillary cell carcinoma. The mRNA expression levels of TRAP1 and HSPD1 were increased in lung squamous cell carcinoma.

We used the Oncomine database to validate the above findings. The results suggested that HSPs show varying degrees of differential expression in multiple tumors (Figure 1C).

1.2 HSP protein expressions in pan-cancer

We next used the THPA site to examine the expression profiles of HSP proteins in 16 tumors. Immunohistochemistry results indicated that HSP proteins showed different expression intensities in 16 tumors. The expression intensities of HSPs in 10 common tumors including lung cancer, liver cancer, colorectal cancer, pancreatic cancer, renal cancer, prostate cancer, breast cancer, endometrial cancer, ovarian cancer and melanoma were shown in Figure 2A. For example, HSP90B1, HSPA9, TRAP1, HSPH1 and HSPD1 showed high levels of immunostaining in 10 common tumors. HSPA6 showed moderate expression in lung cancer, breast cancer, endometrial cancer, ovarian cancer and melanoma and negative expression in endometrial cancer, renal cancer and liver cancer. HSPA8 showed moderate expression in endometrial cancer and breast cancer and negative expression in other cancers. The immunostaining of HSP9 is shown in Figure 2B.

1.3 HSP expressions in pan-cancer cell lines

CCLE data revealed different expression levels of HSPs in 425 cell lines of 8 tumors, including breast cancer, endometrium cancer, kidney cancer, large intestine cancer, lung cancer, prostate cancer, stomach cancer and urinary tract cancer (Figure 2C). The following HSPs were expressed in cell lines in all eight tumors: HSP90AA1, HSP90AB1, HSPA8, HSP90B1, HSPD1, HSPA5, HSPA9, HSPA1B, HSPA4, HSPH1, HYOU1 and TRAP1. HSPA2 and HSPA1A were mainly expressed in lung cancer and breast cancer cell lines. HSPA12A was mainly expressed in lung cancer and kidney cancer cell lines. HSPA13 was mainly expressed in lung cancer, stomach cancer and endometrium cancer cell lines. HSPA14 and HSPA4L were mainly expressed in lung cancer and large intestine cancer cell lines. HSPA7, HSPA1L, HSPA12B and HSPA6 were rarely expressed in the cell lines examined.

2. Relationships of HSP expressions and interactions in pan-cancer

Previous studies showed that HSPs cooperate with each other to exert critical cellular functions. Therefore, we used TCGA data to analyze the correlations of HSPs expressions in 9 tumors including stomach adenocarcinoma, breast cancer, lung adenocarcinoma, lung squamous cell carcinoma, bladder urothelial carcinoma, colon adenocarcinoma, uterine corpus endometrial carcinoma, kidney renal clear cell carcinoma and kidney renal papillary cell carcinoma. The results revealed mostly positive correlations among the expressions of HSP families in different cancers (Figure 3A). Further analysis results demonstrated that the strongest correlation was observed between HSPA4 and HSPA9 in colon adenocarcinoma (r = 0.94, p < 0.05).
We next used the STRING site to predict the potential PPI network of HSP families. The results showed that there were interactions among HSPs. HSPA4, HSPA5, HSPH1, HSPA12A and HSPD1 showed the highest degree score (Figure 3B).

3. Relationships between HSP expressions and cancer-related pathways

The association of HSP expressions with cancer-related pathways was analyzed. The results demonstrated that HSP family proteins mainly participate in the fatty acid metabolism pathway, oxidative phosphorylation pathway, G2M checkpoint pathway, MTORC1 signaling pathway, mitotic spindle pathway, unfolded protein response pathway, protein secretion, reactive oxygen species pathway, E2F target pathway, MYC target pathway, UV response pathway and xenobiotic metabolism pathway (Figure 4A). We counted the numbers of pathways affected by each HSP; the results suggested that HSPA1L, HSPA12B, HSPA6, HSPA7 and HSPA12A may influence the activity of most cancer-related pathways and thus potently affect tumor development (Supplementary Table 1). In addition, the numbers of activated pathways were more than inactivated pathways affected by HSPs, indicating a cancer promotion role of these genes (Figure 4B).

4. Relationships between HSP expressions and immune cell infiltration in pan-cancer

Previous studies reported that HSPs participate in immune cell reactions (23, 24). We next used Spearman correlation analysis to calculate the association of HSP expressions with infiltration of 22 immune cell types. HSP expressions were closely related to immune cell infiltration in the examined cancers (| r | ≥ 0.3, p < 0.05), except for uterine corpus endometrial carcinoma, cervical squamous cell carcinoma and endocervical adenocarcinoma, bladder urothelial carcinoma, lung squamous cell carcinoma and ovarian serous cystadenocarcinoma. The immune cell types included M1 macrophages, resting mast cells, M0 macrophages, M2 macrophages and CD4 memory activated T cells (Table 2 and Supplementary Table 1). For example, HSPD1 was positively associated with CD4 memory activated T cells and follicular helper T cells but negatively associated with monocytes in stomach adenocarcinoma, while HSPD1 was positively associated with CD4 memory activated T cells and negatively associated with resting mast cells in lung adenocarcinoma (Figure 4C).

5. Relationships between HSP expressions and prognosis in pan-cancer

We analyzed the correlation of HSP expressions with prognosis using gene expression and clinical information in TCGA. All specimens were divided into two groups according to the median expression of HSPs, and log-rank test was performed to evaluate the correlation of expression with prognosis. The expressions of HSP families showed different effects on prognosis in 25 tumors, except for rectum adenocarcinoma, stomach adenocarcinoma, thymoma, prostate adenocarcinoma, pancreatic adenocarcinoma, ovarian serous cystadenocarcinoma and lymphoid neoplasm diffuse large B-cell lymphoma (Figure 5). HSP expressions were related to poor prognosis in sarcoma, acute myeloid leukemia, lung squamous cell carcinoma, uterine corpus endometrial carcinoma, uterine carcinosarcoma, esophageal carcinoma, colon adenocarcinoma, lung adenocarcinoma, bladder urothelial carcinoma, glioblastoma multiforme and lung adenocarcinoma but associated with good prognosis in cholangiocarcinoma, pheochromocytoma and paraganglioma. HSP expressions have different influences on prognosis in the remaining tumors, such as liver hepatocellular carcinoma, adrenocortical carcinoma, kidney renal clear cell carcinoma, head and neck squamous carcinoma, lower grade glioma, mesothelioma and skin cutaneous melanoma.

6. Mutation profiles of HSPs in pan-cancer tissues and cell lines

We next examined the mutation profiles of HSPs using TCGA data. The results showed that mutations of HSPs were mainly present in uterine corpus endometrial carcinoma, colon adenocarcinoma, stomach adenocarcinoma, rectum adenocarcinoma, lung squamous cell carcinoma and lung adenocarcinoma, with a mutation frequency of 0%–23% (Figure 6A). A waterfall plot for the mutation details of HSPs in uterine corpus endometrial carcinoma is shown in Figure 6B. Furthermore, CCLE analysis indicated that the mutation frequency of HSPs in 828 cell lines of 15 tumors was 0%–17% (Figure 6C).

Given that the HSPs often interact synergistically in regulating cellular functions, the mutation correlation among HSP genes were analyzed in uterine corpus endometrial carcinoma of the most common mutations. The results suggested that the mutation could be co-occurred or mutually exclusive among HSPs, but the co-occurrence mutation was more obvious (Figure
The co-occurrence of mutations was frequently observed in HSPA1L-HSP90B1, HSPA1L-HSPA5, HSPA4L-HSPA4 and HSPA4L-HSPA6 pairs in uterine corpus endometrial carcinoma.

CNV analysis demonstrated that HSP family members showed different degrees of gene amplification and deletion in pan-cancer (Figure 7A). For instance, the gene amplification frequency of HSPA6 was 24% in bladder urothelial carcinoma and the frequency of HSP90AA1 gene deletion was 22% in cholangiocarcinoma.

7. Relationships between HSP gene variations and expression in pan-cancer

We then examined the correlations of HSP gene variations with expression in cancer. Our results showed that the mutation of 15 HSPs, not including HSPA7, HSPD1, HSPA4, HSPA1B, HYOU1, HSP90B1 and HSPA12B, affected their protein expression in different tumors (p < 0.05, Figure 7B). Notably, the mutation of six HSPs, including HSPA4L, TRAP1, HSPH1, HSP90AA1, HSPA8 and HSPA9, was associated with their protein expression in uterine corpus endometrial carcinoma (p < 0.05).

We further found that CNVs of HSP families impacted their expression levels in different tumors (p < 0.05) (Table 3 and Supplementary Table 2). For example, copy number amplification of HSPA2 was associated with its increased expression in breast cancer, glioblastoma multiforme, head and neck squamous carcinoma, lung squamous cell carcinoma, uterine corpus endometrial carcinoma and ovarian serous cystadenocarcinoma (Table 3). Increased copy number of HSPD1 was associated with increased expression in stomach adenocarcinoma and breast cancer, while its up-regulation was observed in pancreatic adenocarcinoma regardless of whether HSPD1 gene was amplified or deleted (Table 3).

8. Relationships between HSP gene variations and prognosis in pan-cancer

We next examined the relationships of HSP gene variations with prognosis in pan-cancer using TCGA data. We found that mutations of HSP family genes were generally associated with good prognosis in uterine corpus endometrial carcinoma, while mutations in these genes were a poor prognostic factor in other tumors (Figure 7C). Furthermore, CNVs of HSPs, except for HSPA6, HSPA7, HSPD1, HSPA4L, HSPA14, HSPA12A, HSP90B1 and TRAP1, correlated with survival period in different cancers. HSPA1A CNVs were correlated with shorter survival time in stomach adenocarcinoma (HR=1.317, OR=1.044–1.661, p=0.02) and esophageal carcinoma (HR=1.386, OR=1.007–1.908, p=0.045). HSP90AA1 CNVs was correlated with shorter survival rate in prostate adenocarcinoma (HR=3.391, OR=1.07–10.75, p=0.038) (Table 4).

Discussion

HSPs play important roles in various biological processes in tumor cells including cell proliferation, invasion and migration and are potential clinical biomarkers and therapeutic targets(25, 26). Systematic understanding of HSP profiles in pan-cancer is helpful for us to explore its pathogenic mechanism. In this study, we downloaded the expression, survival, mutation and CNVs data in various tumors from TCGA, Oncomine, CCLE and THPA databases, which were initially used to perform a joint analysis of different levels with HSP110, HSP90, HSP70 and HSP60 families. We comprehensively combed the expression profiles of HSPs at mRNA, protein and cell levels and investigated the associations among their expressions and the correlations of expressions with cancer-related pathways, immune cell infiltration and prognosis.

The expression profile analysis revealed that 11 HSPs were differentially expressed in 11 tumors at the mRNA level. Furthermore, 21 HSPs showed different expression patterns at the protein level in pan-cancer. Sixteen HSPs showed high expression levels in 425 cell lines from 8 tumors. Previous studies have shown that HSPs are differentially expressed in different cancers(27-29). Our analysis indicates that HSP family members have the potential to be diagnostic biomarkers of cancer.

We further analyzed the correlation and interaction relationship among HSP family members. The results revealed a mainly positive correlation among HSP expressions in different cancers. The strongest correlation was observed between HSPA4 and HSPA9 in colon adenocarcinoma (r=0.94, p<0.05). PPI network analysis indicated that there were generally interactions among HSPs. HSPH1 interacts with HSPA5 to promote endoplasmic reticulum stress–induced caspase-3 activation and subsequent
apoptosis(30). The N-terminus and C-terminus of HSPA9 interact with HSP90B1 to regulate tumor cell functions(31). Other studies showed that HSPA9 and HSPA4 influence the biological behavior of tumor cells to promote cancer progression(32, 33). Our study found that HSPA4 and HSPA9 were closely correlated with each other. Therefore, we speculate that HSPA4 and HSPA9 may affect tumor development by cooperating with each other. Further study is needed to explore this possibility.

Our results indicate that HSPs may be involved in the activation and inhibition of different cancer-related pathways, such as the unfolded protein response pathway, mitotic spindle pathway and reactive oxygen species pathway, demonstrating that HSPs play different functions in tumor progression. A previous study reported that HSP90 is involved in activating the IL6/JAK/STAT3 signaling pathway to affect tumor progression(34, 35). HSP70 inhibits oxidative phosphorylation and compensates ATP balance through enhanced glycolytic activity in HeLa cells(36). HSP90 were related to the activation of the unfolded protein response pathway in myeloma plasma cells(37). HSPA12B secreted by tumor-associated endothelial cells induces M2 polarization of macrophages by activating the PI3K/Akt/mTOR signaling pathway, thus promoting the formation of the immunosuppressive microenvironment within tumors(38). Our results suggested that HSPA1L, HSPA12B, HSPA12A, HSPA7, HSPH1, HYOU1, HSPA14 and HSPA13 may inhibit the oxidative phosphorylation pathway. In addition, HSPA1L, HSPA12B, HSPA12A, HSP90AB1, HSPA2, HSPA14, HSPA13 and HSPH1 may inhibit the fatty acid metabolism pathway; HSPA6, HSPH1, HSP90AA1, HSPA14, HSPA5 and HSPA8 may activate the MTORC1 signaling pathway, while HSPA1L, HSPA12B and HSPA12A may inactivate the MTORC1 signaling pathway. Previous studies have demonstrated the close correlation of tumor progression with the above pathways(39-41). Our study provides the first evidence for the association of HSPs with these cancer-related pathways. These findings may provide new clues for expanding our understanding of cancer-related pathways and their specific functions in tumor cells.

Based on previous studies demonstrating an involvement of HSPs in the immune response, we examined the correlation of HSP with immune cell infiltration. The results showed that the HSPs was related to immune cell infiltration in 28 tumors. One study reported that HSPD1 regulates the antibacterial function of neutrophils(42). HSP70 released by heat-stressed tumor cells induced the production of tumor cell chemokines and activation of dendritic cells via the TLR4 pathway, thus initiating anti-cancer immunity(43). M2 macrophages could elevate HSPA5 expression to trigger an inflammatory response, and thus facilitating tumor metastasis. (44). Our results showed that HSPs were closely related to the infiltration of immune cells, such as M1 macrophages, resting mast cells, M0 macrophages, M2 macrophages and CD4 memory activated T cells. These findings may provide new directions for researching immune-targeted therapy for cancer.

Our findings showed that HSPs expressions had various effects on prognosis in 25 tumors. HSPs expressions were related to poor prognosis in several cancers, such as sarcoma, lung squamous cell carcinoma and uterine corpus endometrial carcinoma, and were associated with good prognosis in cholangiocarcinoma, pheochromocytoma and paraganglioma. HSPs expressions were correlated with poor and good prognosis in other tumors, such as liver hepatocellular carcinoma and adrenocortical carcinoma. HSPD1 promotes cell invasion and migration, which contributes to the poor prognosis in oral squamous cell carcinoma(45). The expression of HSP90AB1 also caused the worse prognosis in lung adenocarcinoma(46). HSPA9 expression was positive associated with histological grade, pathological stage and lymphatic metastasis in breast cancer, and its expression was negative associated with shorter disease-free survival and overall survival(47). Our study demonstrated that HSP expressions showed different effects on the prognosis of tumors in the digestive, respiratory, urinary and reproductive system. These results indicate that HSPs have the potential to be prognostic biomarkers of different cancers. These findings may also help establish an important foundation for further mechanism studies on how HSPs affect prognosis.

In this study, HSP expression profiles, gene mutations and CNVs as well as the correlation of gene mutations and CNVs with expression and the impact of expression on prognosis were analyzed in pan-cancer. We found that HSP mutations were frequent in uterine corpus endometrial carcinoma, colon adenocarcinoma, stomach adenocarcinoma, rectum adenocarcinoma, lung squamous cell carcinoma and lung adenocarcinoma, with a mutation frequency of 0%–23%. HSP mutations were the most common in endometrial cancer, while mutations rarely occurred in testicular germ cell tumors, uveal melanoma, thymoma and thymoma. The mutation frequency of HSP genes in different cell lines of 15 tumors was 0%–17%. Previous
analysis illustrated that frequent mutations of HSPH1, HSPD1, HSPA4 and HSP90AA1 were detected in head and neck cancer in head and neck squamous carcinoma(48). Somatic mutations and deletions of HSPA8 were observed in sporadic breast carcinoma(49). The mutation frequency of HSP90AA1 in gestational trophoblastic neoplasia was 18.2%(50). Mutations in HSP genes may play an important role in the occurrence and development of tumors and might be targets of anti-cancer treatments. The mutation correlation among HSPs was analyzed in uterine corpus endometrial carcinoma. The results showed that the co-occurrence of mutations in HSPA1L-HSP90B1, HSPA1L-HSPA5, HSPA4L-HSPA4 and HSPA4L-HSPA6 pairs were frequently observed in uterine corpus endometrial carcinoma. Therefore, we predicted that the mutations of these HSPs were closely related and might affect tumor progression.

In this study, the mutation of 15 HSPs affected protein expressions in different cancers. HSP gene mutations were generally associated with good prognosis in uterine corpus endometrial carcinoma, while mutations were poor prognostic factors in other cancers. Previous studies reported frameshift mutations of HSPA4 in gastric and colorectal cancers with microsatellite instability(51). The HSPA4 frameshift mutation seems to reduce the survival activity of tumor cells, which may partially explain why patients of gastric and colorectal cancers with microsatellite instability have better prognosis than those with microsatellite stable cancer(51). One study demonstrated that mutant HSPH1 showed disrupted cellular localization and interaction with other HSPs, thus abolishing the chaperone activity and anti-apoptotic function of HSPH1 by a dominant-negative manner(52). In addition, HSPH1 mutation enhanced chemosensitivity to drugs, such as oxaliplatin and 5-fluorouracil, and improved prognosis in colorectal cancers with microsatellite instability(52).

CNV analysis indicated that amplification and deletion of HSP genes in pan-cancer. Our results also showed that CNVs impacted the expression level of HSPs in different tumors. CNVs of 14 HSPs were associated with good and poor prognosis in various tumors. Somatic CNVs in general can impact protein expression level and thus influence the occurrence and development of different tumors; these CNVs have potential to be diagnostic and prognostic biomarkers as well as anti-cancer therapeutic targets(53-55). Mutations and CNVs are crucial for tumor progression and are closely related to the pathogenic mechanisms of tumors. In our study, mutations and CNVs of HSP genes were correlated with protein expression as well as prognosis in various cancers. These variations are potential immunotherapeutic targets, and further exploration of pathogenic mechanisms is required.

In conclusion, our study analyzed the expression profiles of HSP110, HSP90, HSP70 and HSP60 families in pan-cancer, the relationships among their expression and the correlations of expression with cancer-related signal transduction pathways, immune cell infiltration and prognosis. We also examined the mutations and CNV profiles of HSPs as well as the association of expression with mutations and CNVs. We found that HSP family members are differently expressed in pan-cancer and are closely related to prognosis. Mutations and CNVs in HSPs exert various effects on expression and prognosis in different tumors. HSPs were closely related to immune cell infiltrations in different cancers. In this study, we comprehensively analyzed HSP110, HSP90, HSP70, and HSP60 families in cancer. These results expand our understanding of these proteins and clarified the potential roles for HSPs as diagnostic and prognostic biomarkers as well as anti-cancer therapeutic targets. Our findings provide promising clues for further research of the pathogenic mechanisms of tumors.

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Authors’ contributions**
Qian Xu and Yuan Yuan conceived and designed this study. Li-rong Yan, Shi-xuan Shen and Ang Wang were responsible for the data analysis and performed data interpretation. Li-rong Yan wrote the paper. Qian Xu and Yuan Yuan revised the manuscript.

**Availability of data and materials**

The data that support the results of this manuscript are available from the corresponding author upon reasonable request.

**Competing interests**

All authors disclose no conflicts of interest that might bias their work.

**Funding**

This work was supported by the National Key R&D Program of China (2017YFC0907402).

**Acknowledgements**

We thank Gabrielle White Wolf, PhD, from Liwen Bianji, Edanz Editing China (www.liwenbianji.cn/ac), for editing the English text of a draft of this manuscript.

**Abbreviations**

HSPs, Heat shock proteins; CNVs, copy number variations; TCGA, The Cancer Genome Atlas; CCLE, Cancer Cell Line Encyclopedia; THPA, The Human Protein Atlas; PPI, protein-protein interaction; ACC, Adrenocortical carcinoma; BLCA, Bladder Urothelial Carcinoma; BRCA, Breast invasive carcinoma; CESC, Cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, Cholangiocarcinoma; COAD, Colon adenocarcinoma; DLBC, Lymphoid Neoplasm Diffuse Large B-cell Lymphoma; ESCA, Esophageal carcinoma; GBM, Glioblastoma multiforme; HNSC, Head and Neck squamous cell carcinoma; KICH, Kidney Chromophobe; KIRC, Kidney renal clear cell carcinoma; KIRP, Kidney renal papillary cell carcinoma; LAML, Acute Myeloid Leukemia; LGG, Brain Lower Grade Glioma; LIHC, Liver hepatocellular carcinoma; LUAD, Lung adenocarcinoma; LUSC, Lung squamous cell carcinoma; MESO, Mesothelioma; OV, Ovarian serous cystadenocarcinoma; PAAD, Pancreatic adenocarcinoma; PCPG, Pheochromocytoma, and Paraganglioma; PRAD, Prostate adenocarcinoma; READ, Rectum adenocarcinoma; SARC, Sarcoma; SKCM, Skin Cutaneous Melanoma; STAD, Stomach adenocarcinoma; TGCT, Testicular Germ Cell Tumors; THCA, Thyroid carcinoma; THYM, Thymoma; UCEC, Uterine Corpus Endometrial Carcinoma; UCS, Uterine Carcinosarcoma; UVM, Uveal Melanoma.

**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: a cancer journal for clinicians. 2018;68(6):394-424.

2. Wu J, Liu T, Rios Z, Mei Q, Lin X, Cao S. Heat Shock Proteins and Cancer. Trends in pharmacological sciences. 2017;38(3):226-56.

3. Macario AJ, Conway de Macario E. Molecular chaperones: multiple functions, pathologies, and potential applications. Frontiers in bioscience : a journal and virtual library. 2007;12:2588-600.

4. Tsukahara F, Yoshioka T, Muraki T. Molecular and functional characterization of HSC54, a novel variant of human heat-shock cognate protein 70. Molecular pharmacology. 2000;58(6):1257-63.

5. Kampinga HH, Hageman J, Vos MJ, Kubota H, Tanguay RM, Bruford EA, et al. Guidelines for the nomenclature of the human heat shock proteins. Cell Stress Chaperones. 2009;14(1):105-11.

6. Bhattacharya K, Weidenauer L, Luengo TM, Pieters EC, Echeverria PC, Bemasconi L, et al. The Hsp70-Hsp90 co-chaperone Hop/Stip1 shifts the proteostatic balance from folding towards degradation. Nature communications. 2020;11(1):5975.
7. Kravats AN, Hoskins JR, Reidy M, Johnson JL, Doyle SM, Genest O, et al. Functional and physical interaction between yeast Hsp90 and Hsp70. Proceedings of the National Academy of Sciences of the United States of America. 2018;115(10):E2210-e9.

8. Böttinger L, Oeljeklaus S, Guiard B, Rospert S, Warscheid B, Becker T. Mitochondrial heat shock protein (Hsp) 70 and Hsp10 cooperate in the formation of Hsp60 complexes. The Journal of biological chemistry. 2015;290(18):11611-22.

9. Gibert B, Simon S, Dimitrova V, Díaz-Latoud C, Arrigo AP. Peptide aptamers: tools to negatively or positively modulate HSPB1(27) function. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2013;368(1617):20120075.

10. Xu L, Chen S, Bergan RC. MAPKAPK2 and HSP27 are downstream effectors of p38 MAP kinase-mediated matrix metalloproteinase type 2 activation and cell invasion in human prostate cancer. Oncogene. 2006;25(21):2987-98.

11. Lin CS, He PJ, Hsu WT, Wu MS, Wu CJ, Shen HW, et al. Helicobacter pylori-derived Heat shock protein 60 enhances angiogenesis via a CXCR2-mediated signaling pathway. Biochemical and biophysical research communications. 2010;397(2):283-9.

12. Tsai YP, Yang MH, Huang CH, Chang SY, Chen PM, Liu CJ, et al. Interaction between HSP60 and beta-catenin promotes metastasis. Carcinogenesis. 2009;30(6):1049-57.

13. Wong CS, Wong VW, Chan CM, Ma BB, Hui EP, Wong MC, et al. Identification of 5-fluorouracil response proteins in colorectal carcinoma cell line SW480 by two-dimensional electrophoresis and MALDI-TOF mass spectrometry. Oncology reports. 2008;20(1):89-98.

14. Yun CW, Kim HJ, Lim JH, Lee SH. Heat Shock Proteins: Agents of Cancer Development and Therapeutic Targets in Anti-Cancer Therapy. Cells. 2019;9(1).

15. He L, Deng T, Luo HS. Heat shock protein 70 gene polymorphisms and cancer risk: a meta-analysis. TheScientificWorldJournal. 2014;2014:540309.

16. Zou T, Liu JY, She L, Yin JY, Li X, Li XP, et al. The Association Between Heat-Shock Protein Polymorphisms and Prognosis in Lung Cancer Patients Treated With Platinum-Based Chemotherapy. Frontiers in pharmacology. 2020;11:1029.

17. Medhi S, Sarma MP, Asim M, Kar P. Genetic variants of heat shock protein A1L2437 and A1B1267 as possible risk factors for hepatocellular carcinoma in India. Journal of viral hepatitis. 2013;20(4):e141-7.

18. Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Varambally R, Yu J, Briggs BB, et al. Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. Neoplasia (New York, NY). 2007;9(2):166-80.

19. Ghandi M, Huang FW, Jané-Valbuena J, Kryukov GV, Lo CC, McDonald ER, 3rd, et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. Nature. 2019;569(7757):503-8.

20. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(43):15545-50.

21. Liberzon A, Subramanian A, Pinchback R, Thorvaldssdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics (Oxford, England). 2011;27(12):1739-40.

22. Genest O, Wickner S, Doyle SM. Hsp90 and Hsp70 chaperones: Collaborators in protein remodeling. The Journal of biological chemistry. 2019;294(6):2109-20.

23. Todryk S, Melcher AA, Hardwick N, Linardakis E, Bateman A, Colombo MP, et al. Heat shock protein 70 induced during tumor cell killing induces Th1 cytokines and targets immature dendritic cell precursors to enhance antigen uptake. Journal of immunology (Baltimore, Md : 1950). 1999;163(3):1398-408.

24. Yamazaki K, Ohsawa Y, Tabela K, Ito H, Ueki K, Oda T, et al. Accumulation of human heat shock protein 60-reactive T cells in the gingival tissues of periodontitis patients. Infection and immunity. 2002;70(5):2492-501.

25. Saini J, Sharma PK. Clinical, Prognostic and Therapeutic Significance of Heat Shock Proteins in Cancer. Current drug targets. 2018;19(13):1478-90.
26. Kumar S, Stokes J, 3rd, Singh UP, Scissum Gunn K, Acharya A, Manne U, et al. Targeting Hsp70: A possible therapy for cancer. Cancer letters. 2016;374(1):156-66.

27. Melendez K, Wallen ES, Edwards BS, Mobarak CD, Bear DG, Moseley PL. Heat shock protein 70 and glycoprotein 96 are differentially expressed on the surface of malignant and nonmalignant breast cells. Cell Stress Chaperones. 2006;11(4):334-42.

28. Lee HW, Kim KM. Clinical significance of heat shock protein 90α expression as a biomarker of prognosis in patients with gastric cancer. Nigerian journal of clinical practice. 2019;22(12):1698-705.

29. Li XS, Xu Q, Fu XY, Luo WS. Heat shock protein 60 overexpression is associated with the progression and prognosis in gastric cancer. PloS one. 2014;9(9):e107507.

30. Meares GP, Zmijewska AA, Jope RS. HSP105 interacts with GRP78 and GSK3 and promotes ER stress-induced caspase-3 activation. Cellular signalling. 2008;20(2):347-58.

31. Takano S, Wadhwa R, Mitsui Y, Kaul SC. Identification and characterization of molecular interactions between glucose-regulated proteins (GRPs) mortalin/GRP75/peptide-binding protein 74 (PBP74) and GRP94. The Biochemical journal. 2001;357(Pt 2):393-8.

32. Starenki D, Sosonkina N, Hong SK, Lloyd RV, Park JI. Mortalin (GRP75/HSPA9) Promotes Survival and Proliferation of Thyroid Carcinoma Cells. International journal of molecular sciences. 2019;20(9).

33. Gu Y, Liu Y, Fu L, Zhai L, Zhu J, Han Y, et al. Tumor-educated B cells selectively promote breast cancer lymph node metastasis by HSPA4-targeting IgG. Nature medicine. 2019;25(2):312-22.

34. Kolosenko I, Grander D, Tamm KP. IL-6 activated JAK/STAT3 pathway and sensitivity to Hsp90 inhibitors in multiple myeloma. Current medicinal chemistry. 2014;21(26):3042-7.

35. Lin H, Kolosenko I, Björklund AC, Protsyuk D, Österborg A, Grandér D, et al. An activated JAK/STAT3 pathway and CD45 expression are associated with sensitivity to Hsp90 inhibitors in multiple myeloma. Experimental cell research. 2013;319(5):600-11.

36. Wang L, Schumann U, Liu Y, Prokopchuk O, Steinacker JM. Heat shock protein 70 (Hsp70) inhibits oxidative phosphorylation and compensates ATP balance through enhanced glycolytic activity. Journal of applied physiology (Bethesda, Md : 1985). 2012;113(11):1669-76.

37. Davenport EL, Moore HE, Dunlop AS, Sharp SY, Workman P, Morgan GJ, et al. Heat shock protein inhibition is associated with activation of the unfolded protein response pathway in myeloma plasma cells. Blood. 2007;110(7):2641-9.

38. Zhou J, Zhang A, Fan L. HSPA12B Secreted by Tumor-Associated Endothelial Cells Might Induce M2 Polarization of Macrophages via Activating PI3K/Akt/mTOR Signaling. OncoTargets and therapy. 2020;13:9103-11.

39. Yu L, Lu M, Jia D, Ma J, Ben-Jacob E, Levine H, et al. Modeling the Genetic Regulation of Cancer Metabolism: Interplay between Glycolysis and Oxidative Phosphorylation. Cancer research. 2017;77(7):1564-74.

40. Li Z, Zhang H. Reprogramming of glucose, fatty acid and amino acid metabolism for cancer progression. Cellular and molecular life sciences : CMLS. 2016;73(2):377-92.

41. Kim LC, Cook RS, Chen J. mTORC1 and mTORC2 in cancer and the tumor microenvironment. Oncogene. 2017;36(16):2191-201.

42. Osterloh A, Geisinger F, Piédavent M, Fleischer B, Brattig N, Breloer M. Heat shock protein 60 (HSP60) stimulates neutrophil effector functions. Journal of leukocyte biology. 2009;86(2):423-34.

43. Chen T, Guo J, Han C, Yang M, Cao X. Heat shock protein 70, released from heat-stressed tumor cells, initiates antitumor immunity by inducing tumor cell chemokine production and activating dendritic cells via TLR4 pathway. Journal of immunology (Baltimore, Md : 1950). 2009;182(3):1449-59.

44. Zhang L, Li Z, Ding G, La X, Yang P, Li Z. GRP78 plays an integral role in tumor cell inflammation-related migration induced by M2 macrophages. Cellular signalling. 2017;37:136-48.

45. Kang BH, Shu CW, Chao JK, Lee CH, Fu TY, Liou HH, et al. HSPD1 repressed E-cadherin expression to promote cell invasion and migration for poor prognosis in oral squamous cell carcinoma. Scientific reports. 2019;9(1):8932.
46. Wang M, Feng L, Li P, Han N, Gao Y, Xiao T. [Hsp90AB1 Protein is Overexpressed in Non-small Cell Lung Cancer Tissues and Associated with Poor Prognosis in Lung Adenocarcinoma Patients]. Zhongguo fei ai za zhi = Chinese journal of lung cancer. 2016;19(2):64-9.

47. Jin H, Ji M, Chen L, Liu Q, Che S, Xu M, et al. The clinicopathological significance of Mortalin overexpression in invasive ductal carcinoma of breast. Journal of experimental & clinical cancer research : CR. 2016;35:42.

48. Fan G, Tu Y, Wu N, Xiao H. The expression profiles and prognostic values of HSPs family members in Head and neck cancer. Cancer cell international. 2020;20:220.

49. Bakkenist CJ, Koreth J, Williams CS, Hunt NC, McGee JO. Heat shock cognate 70 mutations in sporadic breast carcinoma. Cancer research. 1999;59(17):4219-21.

50. Luo L, Lin L, Zhang X, Cai Q, Zhao H, Xu C, et al. Next-Generation Sequencing Panel Analysis of Clinically Relevant Mutations in Circulating Cell-Free DNA from Patients with Gestational Trophoblatic Neoplasia: A Pilot Study. BioMed research international. 2020;2020:1314967.

51. Jo YS, Choi MR, Song SY, Kim MS, Yoo NJ, Lee SH. Frameshift Mutations of HSPA4 and MED13 in Gastric and Colorectal Cancers. Pathology oncology research : POR. 2016;22(4):769-72.

52. Dorard C, de Thonel A, Collura A, Marisa L, Svrcek M, Lagrange A, et al. Expression of a mutant HSP110 sensitizes colorectal cancer cells to chemotherapy and improves disease prognosis. Nature medicine. 2011;17(10):1283-9.

53. Shao X, Lv N, Liao J, Long J, Xue R, Ai N, et al. Copy number variation is highly correlated with differential gene expression: a pan-cancer study. BMC Medical Genetics. 2019;20(1):175.

54. Hieronymus H, Murali R, Tin A, Yadav K, Abida W, Moller H, et al. Tumor copy number alteration burden is a pan-cancer prognostic factor associated with recurrence and death. eLife. 2018;7.

55. Cheau-Feng Lin F, Jeng YC, Huang TY, Chi CS, Chou MC, Chin-Shaw Tsai S. Mitochondrial DNA copy number is associated with diagnosis and prognosis of head and neck cancer. Biomarkers : biochemical indicators of exposure, response, and susceptibility to chemicals. 2014;19(4):269-74.

Tables
| Gene family | Gene name | Old names | chrom | chromStart | chromEnd | strand |
|-------------|-----------|-----------|-------|------------|----------|--------|
| HSP110      | HSPH1     | HSP105    | chr13 | 31134974   | 31162388 | -      |
|             | HSPH2     | HSPA4; APG-2; HSP110 | chr5   | 133051962  | 133106449 | +      |
|             | HSPH3     | HSPA4L; APG-1 | chr4   | 127781821  | 127840733 | +      |
|             | HSPH4     | HYOU1/Grp170; ORP150; HSP12A | chr11  | 119044189  | 119057202 | -      |
| HSP90       | HSPC1     | HSP90AA1; HSPN; LAP2; HSP86; HSPC1; HSPCA; HSP89; HSP90; HSP90A; HSP90N; HSPCAL1; HSPCAL4; FLJ31884 | chr14  | 102080738  | 102139699 | -      |
|             | HSPC3     | HSP90AB1; HSPC2; HSPCB; D6S182; HSP90B; FLJ26984; HSP90-BETA | chr6   | 44246166   | 44253888 | +      |
|             | HSPC4     | HSP90B1; ECGP; GP96; TRA1; GRP94; endoplasmin | chr12  | 103930107  | 103953645 | +      |
|             | HSPC5     | TRAP1; HSP75; HSP90L | chr16  | 3651639    | 3717597 | -      |
| HSP70       | HSPA1A    | HSP70-1; HSP72; HSPA1 | chr6   | 31815464   | 31817946 | +      |
|             | HSPA1B    | HSP70-2   | chr6   | 31827735   | 31830255 | +      |
|             | HSPA1L    | hum70t; hum70t; Hsp-hom | chr6   | 31809619   | 31815065 | -      |
|             | HSPA2     | Heat-shock 70kD protein-2 | chr14  | 64535905   | 64546173 | +      |
|             | HSPA5     | BIP; GRP78; MIF2 | chr9   | 125234853  | 125241330 | -      |
|             | HSPA6     | Heat shock 70kD protein 6 (HSP70B°) | chr1   | 161524540  | 161526910 | +      |
|             | HSPA7     | Heat shock 70kD protein 7 | chr1   | 161606291  | 161608217 | +      |
|             | HSPA8     | HSC70; HSC71; HSP71; HSP73 | chr11  | 123057489  | 123063230 | -      |
|             | HSPA9     | GRP75; HSPA9B; MOT; MOT2; PBP74; mot-2 | chr5   | 138554882  | 138575444 | -      |
|             | HSPA12A   | FLJ13874; KIAA0417 | chr10  | 116671192  | 116742574 | -      |
|             | HSPA12B   | RP23-32L15.1; 2700081N06Rik | chr20  | 3732667    | 3753111 | +      |
|             | HSPA13    | Stch | chr21  | 14371115   | 14383484 | -      |
|             | HSPA14    | HSP70-4; HSP70L1; MGC131990 | chr10  | 14838164   | 14871741 | +      |
| HSP60       | HSPD1     | HSP60; GroEL | chr2   | 197486581  | 197516737 | -      |
| Gene  | CancerType | CellType           | Correlation Coefficient | P value  | Gene  | CancerType | CellType           | Correlation Coefficient | P value  |
|-------|------------|--------------------|-------------------------|----------|-------|------------|--------------------|-------------------------|----------|
| TRAP1 | KIRP       | Macrophages M0     | 0.35                    | <0.001   | HSPA1A| LUAD       | Macrophages M0     | 0.32                    | <0.001   |
| TRAP1 | THYM       | Macrophages M1     | 0.34                    | <0.001   | HSPA13| LUAD       | T cells CD4 memory activated | 0.36          | <0.001   |
| TRAP1 | GBM        | Macrophages M2     | 0.33                    | <0.001   | HSPA14| LUAD       | T cells CD4 memory activated | 0.31          | <0.001   |
| HYOU1 | LAML       | Mast cells resting | 0.36                    | <0.001   | HSPA14| BRCA      | Mast cells resting | -0.3                 | <0.001   |
| HYOU1 | THYM       | Macrophages M1     | 0.34                    | <0.001   | HSPA14| KIRP      | Mast cells resting | -0.3                 | <0.001   |
| HYOU1 | TGCT       | Macrophages M1     | 0.31                    | <0.001   | HSPA14| BRCA      | Mast cells resting | -0.3                 | <0.001   |
| HSPH1 | STAD       | Macrophages M0     | 0.34                    | <0.001   | HSPA14| THYM      | Macrophages M2     | -0.4                 | <0.001   |
| HSPH1 | SARC       | Mast cells resting | -0.35                   | <0.001   | HSPA14| THYM      | Mast cells resting | -0.4                 | <0.001   |
| HSPH1 | THYM       | Mast cells resting | -0.53                   | <0.001   | HSPA14| THYM      | Mast cells resting | -0.4                 | <0.001   |
| HSPH1 | TGCT       | Macrophages M2     | -0.57                   | <0.001   | HSPA14| ACC       | Mast cells resting | -0.53                | <0.001   |
| HSPA9 | KIRP       | Macrophages M1     | 0.38                    | <0.001   | HSPA12B| PCPG      | Mast cells resting | 0.42                 | <0.001   |
| HSPA9 | THYM       | Macrophages M1     | 0.37                    | <0.001   | HSPA12B| TGCT      | Macrophages M2     | 0.39                 | <0.001   |
| HSPA9 | PRAD       | Macrophages M1     | 0.31                    | <0.001   | HSPA12B| ACC       | Macrophages M2     | 0.35                 | <0.001   |
| HSPA9 | STAD       | Mast cells resting | -0.35                   | <0.001   | HSPA12B| KIRC      | Mast cells resting | 0.35                 | <0.001   |
| HSPA8 | PAAD       | T cells CD4 memory activated | 0.3         | <0.001   | HSPA12B| STAD      | T cells CD4 memory activated | 0.34          | <0.001   |
| HSPA8 | THYM       | Mast cells resting | -0.4                    | <0.001   | HSPA12B| SARC      | Macrophages M2     | 0.33                 | <0.001   |
| HSPA7 | THYM       | Macrophages M1     | 0.41                    | <0.001   | HSPA12B| ESCA      | Mast cells resting | 0.32                 | <0.001   |
| HSPA7 | LGG        | Macrophages M1     | 0.32                    | <0.001   | HSPA12B| ACC       | Mast cells resting | 0.31                 | <0.001   |
| HSPA7 | THCA       | Macrophages M1     | 0.3                     | <0.001   | HSPA12B| TGCT      | T cells CD4 memory activated | -0.32       | <0.001   |
| HSPA6 | THYM       | Macrophages M2     | 0.45                    | <0.001   | HSPA12B| PRAD      | Macrophages M1     | -0.34                | <0.001   |
| HSPA6 | THYM       | Macrophages M1     | 0.38                    | <0.001   | HSPA12A| TGCT      | Macrophages M2     | 0.45                 | <0.001   |
| HSPA6 | GBM        | Macrophages M0     | 0.35                    | <0.001   | HSPA12A| KIRP      | Macrophages M2     | -0.32                | <0.001   |
| HSPA6 | SARC       | T cells CD4 memory activated | 0.3         | <0.001   | HSPA12A| KIRP      | Macrophages M2     | -0.32                | <0.001   |
| HSPA6 | LGG        | Macrophages M0     | 0.3                     | <0.001   | HSPA12A| KIRP      | Macrophages M0     | -0.32                | <0.001   |
| HSPA6 | KIRC       | Mast cells resting | -0.32                   | <0.001   | HSPA12A| KIRP      | Macrophages M2     | -0.32                | <0.001   |
| HSPA6 | LGG        | Macrophages M2     | -0.33                   | <0.001   | HSPA12A| KIRP      | Macrophages M0     | -0.32                | <0.001   |
| HSPA6 | SARC       | Mast cells resting | -0.36                   | <0.001   | HSPA12A| KIRP      | Macrophages M2     | -0.32                | <0.001   |
| HSPA5 | LGG        | Macrophages M1     | 0.36                    | <0.001   | HSPA12A| TGCT      | T cells CD4 memory activated | -0.45       | <0.001   |
| HSPA5 | LAML       | Mast cells resting | 0.36                    | <0.001   | HSP90B1| TGCT      | Macrophages M2     | 0.49                 | <0.001   |
| HSPA5 | KIRP       | Macrophages M1     | 0.35                    | <0.001   | HSP90B1| LGG       | Macrophages M0     | 0.31                 | <0.001   |
| HSPA5 | GBM        | Macrophages M1     | 0.31                    | <0.001   | HSP90B1| LGG       | Macrophages M1     | 0.31                 | <0.001   |
| HSPA5 | LGG        | Macrophages M0     | 0.3                     | <0.001   | HSP90B1| SARC      | Mast cells resting | -0.3                 | <0.001   |
| HSPA5 | STAD       | Mast cells resting | -0.3                    | <0.001   | HSP90B1| STAD      | Mast cells resting | -0.32                | <0.001   |
| HSPA5 | SARC       | Mast cells resting | -0.39                   | <0.001   | HSP90AB1| KIRP      | Macrophages M1     | 0.4                   | <0.001   |
| HSPA4L | THYM   | Macrophages M1     | 0.51                    | <0.001   | HSP90AB1| STAD      | Macrophages M0     | 0.35                 | <0.001   |
| Gene   | Tissue | Cell Type               | Correlation | P Value | Gene   | Tissue | Cell Type               | Correlation | P Value |
|--------|--------|-------------------------|-------------|---------|--------|--------|-------------------------|-------------|---------|
| HSPA4L | THYM   | Macrophages M2          | 0.4         | <0.001  | HSP90AB1 | THYM   | Macrophages M1          | 0.33        | <0.001  |
| HSPA4L | THYM   | Macrophages M0          | 0.32        | <0.001  | HSP90AB1 | TGCT   | Macrophages M2          | -0.33       | <0.001  |
| HSPA4L | LGG    | Macrophages M1          | 0.3         | <0.001  | HSP90AB1 | STAD   | Mast cells resting      | -0.36       | <0.001  |
| HSPA4  | TGCT   | T cells CD4 memory activated | 0.32       | <0.001  | HSP90AA1 | STAD   | Mast cells resting      | -0.3        | <0.001  |
| HSPA4  | KIRP   | Macrophages M1          | 0.3         | <0.001  | HSP90AA1 | THYM   | Mast cells resting      | -0.34       | <0.001  |
| HSPA2  | STAD   | Mast cells resting      | 0.31        | <0.001  | HSP90AA1 | TGCT   | Macrophages M2          | -0.37       | <0.001  |
| HSPA2  | READ   | T cells CD4 memory activated | -0.34      | <0.001  | HSPD1   | THYM   | Mast cells resting      | -0.33       | <0.001  |
| HSPA1L | LUAD   | Macrophages M0          | 0.32        | <0.001  | HSPD1   | LUAD   | Mast cells resting      | -0.35       | <0.001  |
| HSPA1B | THYM   | Macrophages M2          | 0.38        | <0.001  | HSPD1   | STAD   | T cells CD4 memory activated | 0.32       | <0.001  |
| HSPA1B | THYM   | Macrophages M0          | 0.36        | <0.001  | HSPD1   | READ   | T cells CD4 memory activated | 0.33       | <0.001  |
| HSPA1B | THYM   | Macrophages M1          | 0.32        | <0.001  | HSPD1   | LUAD   | T cells CD4 memory activated | 0.33       | <0.001  |
| HSPA1B | LGG    | Macrophages M2          | -0.32       | <0.001  | HSPD1   | COAD   | T cells CD4 memory activated | 0.32       | <0.001  |
| Gene name | CNVCat | N   | summarise | p value   | CancerType |
|-----------|--------|-----|-----------|-----------|------------|
| HSPA2     | DEL    | 62  | 10.303(9.167-11.858) | 0.011     | BRCA       |
| GAIN      | 29     |     | 12.051(9.649-12.666) |           |            |
| No Change | 999    |     | 10.844(9.735-12.004) |           |            |
| DEL       | 3      |     | 9.011(8.298-9.706)   | 0.043     | GBM        |
| GAIN      | 2      |     | 12.016(11.549-12.483)|           |            |
| No Change | 159    |     | 11.152(10.204-11.958)|           |            |
| DEL       | 9      |     | 11.291(9.502-12.274) | 0.024     | HNSC       |
| GAIN      | 25     |     | 13.069(11.437-13.797)|           |            |
| No Change | 462    |     | 12.096(10.908-13.071)|           |            |
| DEL       | 43     |     | 11.47(10.509-12.65)  | 0.003     | LGG        |
| GAIN      | 29     |     | 12.356(11.421-13.173)|           |            |
| No Change | 485    |     | 8.607(7.803-9.212)   | 0.006     | LUSC       |
| DEL       | 23     |     | 10.212(9.423-10.997) |           |            |
| GAIN      | 13     |     | 10.18(8.972-11.133)  |           |            |
| No Change | 464    |     | 12.056(10.908-13.071)|           |            |
| GAIN      | 60     |     | 15.065(14.404-15.626)|           |            |
| No Change | 1011   |     | 14.287(13.784-14.794)|           |            |
| DEL       | 8      |     | 13.737(13.554-13.977)| 0.013     | CESC       |
| GAIN      | 21     |     | 14.891(14.629-15.34) |           |            |
| No Change | 265    |     | 13.988(13.442-14.525)|           |            |
| DEL       | 12     |     | 14.339(13.909-14.924)| 0.013     | ESCA       |
| GAIN      | 10     |     | 15.273(15.023-15.438)|           |            |
| No Change | 139    |     | 14.785(14.316-15.189)|           |            |
| DEL       | 27     |     | 13.661(13.244-14.158)| 0.029     | LIHC       |
| GAIN      | 26     |     | 14.896(14.208-15.524)|           |            |
| No Change | 443    |     | 14.036(13.506-14.512)|           |            |
| DEL       | 11     |     | 13.422(12.664-13.688)| 0.002     | KIRC       |
| GAIN      | 7      |     | 14.302(14.146-14.692)|           |            |
| No Change | 512    |     | 14.146(13.715-14.602)|           |            |
| DEL       | 1      |     | 13.507(13.507-13.507)| 0.001     | KIRP       |
| GAIN      | 7      |     | 15.782(15.079-16.231)|           |            |
| No Change | 279    |     | 14.106(13.647-14.754)|           |            |
| DEL       | 3      |     | 13.662(13.968-14.761)| 0.029     | PAAD       |
| GAIN      | 14     |     | 15.097(14.695-15.415)|           |            |
| No Change | 355    |     | 14.489(13.979-14.991)|           |            |
| DEL       | 25     |     | 13.866(13.364-14.265)| 0.001     | LUSC       |
| GAIN      | 34     |     | 15.343(15.184-15.612)|           |            |
| No Change | 441    |     | 14.33(13.791-14.855) |           |            |
| DEL       | 27     |     | 13.798(13.424-14.083)| 0.001     | OV         |
| GAIN      | 68     |     | 14.658(14.153-15.103)|           |            |
| No Change | 282    |     | 14.205(13.669-14.636)|           |            |
| GAIN      | 6      |     | 14.156(13.693-14.566)| 0.032     | PRAD       |
| No Change | 171    |     | 13.399(12.9-13.861)  |           |            |
| DEL       | 4      |     | 14.354(14.184-14.737)| 0.01      |            |
| GAIN      | 6      |     | 15.644(15.17-15.92)  |           |            |
| No Change | 486    |     | 14.09(13.666-14.491) |           |            |
| DEL       | 7      |     | 13.876(13.827-14.053)| 0.005     | SKCM       |
| GAIN      | 8      |     | 15.292(14.784-15.408)|           |            |
| No Change | 455    |     | 14.426(13.823-14.92) |           |            |
| DEL       | 11     |     | 14.022(13.528-14.122)| 0.007     | STAD       |
| GAIN      | 14     |     | 15.382(14.203-15.79) |           |            |
| No Change | 348    |     | 14.44(13.778-15.003) |           |            |
| DEL       | 6      |     | 14.111(13.133-14.18) | 0.003     | UCEC       |
| Gene name | Variations | HR   | 95%CI     | p value | CancerType |
|-----------|------------|------|-----------|---------|------------|
| P90AB1    | ENSG00000096384_Del8 | 2.488 | 1.337 - 4.631 | 0.004  | UCS        |
| PA2       | ENSG00000_Del_Gain6803.9 | 2.924 | 1.374 - 6.223 | 0.005  | PAAD       |
| PA8       | ENSG00000_Del09971.1_Gain | 29.131 | 2.64 - 321.42 | 0.006  | KICH       |
| PA12B     | ENSG00000_Del3_Gain622.9 | 4.453 | 1.545 - 12.833 | 0.006  | LAML       |
| PA12B     | ENSG00000_Del3_Gain622.9.1 | 4.453 | 1.545 - 12.833 | 0.006  | LAML       |
| OU1       | ENSG00000_Del494_Gain8.17 | 29.131 | 2.64 - 321.42 | 0.006  | KICH       |
| PA5       | ENSG00000044574.7 | 29.131 | 2.64 - 321.42 | 0.006  | KICH       |
| PA9       | ENSG00000_Del13013.11 | 2.795 | 1.315 - 5.939 | 0.008  | LAML       |
| PA9       | ENSG00000_Del13013.11 | 3.46  | 1.373 - 8.718 | 0.008  | PRAD       |
| PA9       | ENSG00000_Del13013.11 | 2.795 | 1.315 - 5.939 | 0.008  | LAML       |
| PA9       | ENSG00000_Del13013.11 | 1.927 | 1.154 - 3.217 | 0.012  | ESCA       |
| PA4       | ENSG00000_Del70606.1_Gain | 3.494 | 1.323 - 9.226 | 0.012  | PRAD       |
| PA1B      | ENSG00000_Gain04388.6 | 1.317 | 1.044 - 1.661 | 0.02   | STAD       |
| PA1A      | ENSG00000_Gain04389.9 | 1.317 | 1.044 - 1.661 | 0.02   | STAD       |
| PA1L      | ENSG00000_Gain04390.9 | 1.317 | 1.044 - 1.661 | 0.02   | STAD       |
| PA8       | ENSG00000_Del09971.1_Gain | 0.587 | 0.371 - 0.927 | 0.022  | BLCA       |
| PA2       | ENSG00000_Del_Gain6803.9 | 1.534 | 1.05 - 2.243  | 0.027  | BRCA       |
| PH1       | ENSG00000_Del_Gain0694.18 | 0.671 | 0.469 - 0.959 | 0.028  | LUSC       |
| P90AB1    | ENSG00000096384_Del8 | 3.645 | 1.097 - 12.108 | 0.035  | MESO       |
| P90AA1    | ENSG00000808_Gain4_Del7 | 3.391 | 1.07 - 10.75  | 0.038  | PRAD       |
| PA4       | ENSG00000_Del70606.1_Gain | 1.699 | 1.023 - 2.821 | 0.041  | CESC       |
| PA5       | ENSG00000044574.7 | 1.363 | 1.013 - 1.834 | 0.041  | BRCA       |
| PA8       | ENSG00000_Del09971.1_Gain | 1.944 | 1.019 - 3.709 | 0.044  | MESO       |
| OU1       | ENSG00000_Del494_Gain8.17 | 1.986 | 1.02 - 3.867  | 0.044  | MESO       |
| PA1B      | ENSG00000_Gain04388.6 | 1.386 | 1.007 - 1.908 | 0.045  | ESCA       |
| PA1A      | ENSG00000_Gain04389.9 | 1.386 | 1.007 - 1.908 | 0.045  | ESCA       |
| PA1L      | ENSG00000_Gain04390.9 | 1.386 | 1.007 - 1.908 | 0.045  | ESCA       |
| PH1       | ENSG00000_Del_Gain0694.18 | 4.279 | 1.015 - 18.044 | 0.048  | KIRP       |
| PA13      | ENSG00000_Del55304.5 | 1.806 | 1.004 - 3.249 | 0.048  | CESC       |

N represented sample size.

The table above shows the correlations between copy number variations and prognosis of heat shock proteins in pan-cancer. The table includes the name of the gene, the type of variation (Gain or Del), the hazard ratio (HR), the 95% Confidence Interval (95%CI), and the p-value. The CancerType column indicates the type of cancer associated with each variation.

**Figures**
Figure 1

HSP expressions at mRNA level. (A) The differentially expressed of heat shock proteins in pan-cancer. The color in heat map represents the log2 fold change value between cancer and normal. Red color represents up-regulated and blue color represents down-regulated. * p<0.05, ** p<0.01, *** p<0.001. (B) HSPA2 expression in 16 types of cancers between cancer and normal tissues. (C) Expression profile of HSPs in different human cancer from Oncomine database. Red color represents up-regulated and blue color represents down-regulated. Cell color is determined by the best gene rank percentile for the analyses within the cell. HSP, heat shock protein.
Figure 2

HSP expressions at protein and cell level. (A) HSP family members protein expression in 10 types of cancers. Each gene expression in one cancer was divided into four groups of high expression (red color), medium expression (orange color), low expression (yellow color) and not detected (grey color). (B) The immunohischemistry staining results of HSPA9 protein in 16 types of cancers in The Human Protein Atlas database. (C) Expression profile of HSPs family in cell lines in the CCLE database. Expression increased gradually from blue color to red color. CCLE, Cancer Cell Line Encyclopedia; LUCA, lung cancer; COCA, colorectal cancer; TECA, testis cancer; RACA, renal cancer; HNSC, head and neck cancer; STCA, stomach cancer; PACA; pancreatic cancer; OV, ovarian cancer; SKCA, skin cancer; BRCA, breast cancer; LIHC, liver cancer; ENCA, endometrial cancer; THCA, thyroid cancer.
Figure 3

The correlation analysis and protein-protein interactions among HSPs. (A) The correlations among HSP expressions. Red color represents positive correlation and blue color represents negative color (p<0.05). CCLE, Cancer Cell Line Encyclopedia. HSP, heat shock protein. (B) Protein-protein interaction network among HSP family. The higher the degree score, the larger the size.
Figure 4

The association of HSP expressions with cancer-related pathway and immune cell infiltration. (A) Network displaying the correlation between HSP expressions and cancer-related pathways. Pink color represents HSPs and blue color represents pathways. The red edges represent activated pathways and grey edges represent inhibited pathways. (B) Number of activated and inhibited pathways of each heat shock protein. Pink color represents activated pathways activity and blue color represents inhibited pathways activity. (C) Association of HSPD1 with different immune cell infiltration across different cancer types. HSP, heat shock protein.
Figure 5

Heatmap of the relationship between HSP expressions and overall survival. Red represents a worse prognosis, and blue represents a better prognosis. HSP, heat shock protein.
Figure 6

HSP mutations at tissue and cell level across different cancer types. (A) Mutation frequency of HSPs in tissues across different cancer types. (B) Oncoplot for HSPs in UCEC. HSPH1 has the most frequent mutation in UCEC. UCEC, uterine corpus endometrial carcinoma. (C) Mutation status of HSPs in different cancer cell lines from CCLE database. (D) Correlation among HSP mutations in UCEC. Blue color represents co-occurrence mutation and brown color displayed mutually exclusive. The number represents sample sizes of mutation. UCEC, uterine corpus endometrial carcinoma. CCLE, Cancer Cell Line Encyclopedia. HSP, heat shock protein. UCEC, uterine corpus endometrial carcinoma.
Figure 7

Copy number variation of HSPs and the correlation of their variation with expression and prognosis. (A) Copy number variation frequency of HSPs in across different cancer types. (B) Association of HSP mutations with expression. Orange color represents $p \geq 0.05$, blue color represents $p < 0.05$. (C) Correlation between HSP mutations and prognosis in different cancer types. Red color represents a worse prognosis, blue color represents a better prognosis and brown color represents $p \geq 0.05$. HSPs, heat shock proteins.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementarymaterials.docx