

Co-expression Network and Prognosis Analyses of Pyroptosis-related Genes in Non-small Cell Lung Cancer

Zixiao Liu

Xinyang Normal University

Xudong Liu

Xinyang Normal University

Yu Zhang

Xinyang Normal University

Yongjie Zhou

Xinyang Normal University

Shuaibin Lian (✉ shuai_lian@xynu.edu.cn)

Xinyang Normal University

Lei Wang

Xinyang Normal University

Research Article

Keywords: Pyrolysis, Gasdermin family, NSCLC, Co-expression network analysis, Small molecules

Posted Date: June 24th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-627404/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Lung cancer is very difficult to diagnose in its early stages because of its initial asymptomatic characteristics. In recent years, pyrolysis has been shown identified as a novel type of programmed cell death with inflammation mediated by the gasdermin family. In this study, 33 differentially-expressed pyroptosis-related genes were commonly identified in both lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) patients. Tumor-related gasdermin family genes that were significantly differentially expressed in non-small cell lung cancer (NSCLC) tissues were identified by our co-expression network analysis. Among them, the mRNA level of *GSDMB* gene had significant impacts on tumor staging and survival rates of NSCLC patients. Therefore, this gene is a potential new therapeutic target for the treatment of NSCLC. In addition, the high expression levels of *GSDMC/D* were significantly correlated with the low overall survival (OS), progression-free survival (FP) and post-progression survival (PPS) of NSCLC patients. Therefore, this gene is a potential oncogene for NSCLC. Furthermore, four small molecules (erastin, cefotiam, metanephrine, and vorinostat) that could most significantly reverse the NSCLC gene expression were identified. They interacted with GSDMB proteins mainly through H-bonds and hydrophobic interactions. This study provides new therapeutic targets and prognostic makers for NSCLC patients.

Introduction

Lung cancers are highly fatal diseases. Due to their initial asymptomatic characteristics, it is difficult to detect and diagnose lung cancers during their early stages, which is also the main cause of the high death rate of lung cancer patients^{1,2}. In the past few decades, the survival rate of lung cancer patients has not been improved significantly. There are about 1.8 million new lung cancer cases and 1.6 million lung cancer-related deaths annually, yielding a mortality rate of more than 80%³. In the United States, lung cancer is the second most common disease that causes deaths, and the 10-year survival rate for all stages of lung cancer is less than 7%^{4,5}. Current treatments for lung cancers include surgical resection, high-dose stereotactic body radiation therapy, and molecular targeted therapies. Despite the continuous development of molecular targeted drugs and immune checkpoint inhibitors, as well as the improvements in the treatment methods for lung cancers, the 5-year survival rates for lung cancers with different locations of lesions and stages remain poor^{6,7}. In addition to the limited efficacies of current treatment options for lung cancers, the existing molecular pathological studies on lung cancers only benefited the development of molecular targeted therapies for LUAD, while the research on LUSC and small cell lung cancer is still in progress⁸. Early diagnosis and new therapeutic targets will benefit the clinical management and targeted therapies of lung cancers.

Pyrolysis is a type of programmed cell death mediated by the gasdermin proteins and accompanied by inflammatory and immune responses. Pyrolysis activates caspase-1 through promoting inflammasome assembly, which in turn facilitates the secretion of bioactive interleukin-1 β /18, thereby achieving inflammatory cell death^{9,10}. As a multi-protein signaling complex mainly expressed in myeloid cells,

inflammasomes play important roles in immunology, molecular pathology, and disease treatment¹¹. Inflammasomes can protect cells from pathogen invasion and prevent cancer-related diseases by activating the human immune system. Previous studies have shown that the NLRP3 inflammasome in macrophages can optimally activate cells by combining purinergic signaling with the activation of the complement cascade, a finding that can potentially promote the research and development of anti-cancer drugs^{12,13}. As an important component of the inflammasome, NEK7 can mediate the activation of NLRP3 inflammasome through facilitating the interaction between the NF- κ B signaling and NLRP3, thereby affecting the pathogenesis of inflammatory bowel diseases. In addition, unlike apoptosis, pyrolysis may induce cytolytic death to cause diseases, during which it releases immunogenic factors (such as interleukin-1 β) and triggers inflammation¹⁴. Some studies have shown that pyrolysis can inhibit the growth of tumors and has the potential to be utilized in the treatments for cancers. For example, liver cells treated with chemical reagents can overexpress *TXNIP* gene, which in turn activates the NLRP3 inflammasome and caspase-1-mediated hepatocyte lysis¹⁵. Inflammasomes can also function as signal-transmitting complexes in immune cells in brain microglia, thereby regulating brain functions and the occurrence of diseases¹⁶.

The gasdermin family proteins, namely Gasdermin A (GSDMA), Gasdermin B (GSDMB), Gasdermin C (GSDMC), Gasdermin D (GSDMD), Gasdermin E (GSDME, also known as DNFA5), and Pejvakin (PJVK, also known as DFNB59), are the main executors of pyroptosis¹⁷. Under an atomic force microscope, it can be observed that gasdermins are pore-forming effector proteins containing N-terminal and C-terminal domains. The gasdermin-N domain can bind to membrane lipids, phosphoinositides and cardiolipins, thereby generating pores in the cell membrane^{18,19}. Through these cell membrane pores (10-20 nm in diameter), cellular contents can be slowly released from cells, thereby leading to the formation of larger oligomeric pores and activating a strong inflammatory response, which ultimately leads to imbalance of ion homeostasis and induces cell death^{20,21}. Among the gasdermins, GSDMA/B have been associated with the susceptibility of inflammatory bowel diseases²²; *GSDMC* overexpression can indicate a poor prognosis in LUAD patients²³; the expression of *GSDMD* promotes the response of CD8⁺ T cells to lung cancer cells²⁴; and *GSDME* has been regarded as a tumor suppressor gene²⁵.

In this research, we systematically studied the expression of pyroptosis-related genes in normal and tumorous (N&T) tissues from NSCLC patients, identified gene modules closely related to the N&T tissues through weighted correlation network analysis, and analyzed the expression levels, prognostic significances and mutation rates of gasdermin genes, as well as the correlations between their expression levels with cancer stages in the NSCLC patients. Next, the functions of 150 genes whose expression levels were closely associated with gasdermin gene mutations were predicted. Finally, the prediction and molecular docking analysis of small molecule drugs that can target the gasdermin family proteins were carried out (see **Fig. 1** for the workflow of this study). Our findings indicate that pyroptosis has important impacts on tumor development, diagnosis, and treatment of NSCLC.

Materials And Methods

Data preparation

The original data [the fragments per kilobase of transcript per million mapped reads (FPKM) values and clinical information] of the NSCLC patients were downloaded from The Cancer Genome Atlas database by using the SangerBox tool (SangerBox, <http://sangerbox.com/>). The “Deseq2”³⁹ and “edger” packages⁴⁰ were used to identify differentially expressed genes (DEGs) (thresholds were $padj < 0.05$, $|\log_2FC| > 1$, false discovery rate < 0.05) between the N&T tissues. We identified 4,851 and 6,513 DEGs in the LUAD and LUSC cohorts, respectively (**Table S1**). Subsequently, 40 and 45 pyroptosis-related genes were identified from the LUAD and LUSC cohorts, respectively, based on a previous review⁴¹⁻⁴⁴ (**Table S2**).

Weighted co-expression network and immunohistochemical analyses

The “module detection” function of WGCNA and R packages (version 1.68) were utilized to identify the gene co-expression modules between the N&T tissues of the NSCLC patients. The networks were constructed by adjusting the parameters for the ‘block-wiseModules’ function of WGCNA for the LUAD and LUSC DEG datasets, so that the input data were allowed for automatic and unsupervised network constructions. We obtained the nodes and edges of the WGCNA network by calculating the correlations among gene expression values. Each node in the WGCNA network corresponds to a gene, and the edges were constructed by identifying the correlations between the expression levels of the DEGs. The parameter settings for co-expression network construction were as follows: “powers” = 10, “minKMEtoStay” = 0.7, “minModuleSize” = 30, “mergeCutHeight” = 0.25, and “minCoreKME” = 0.8, other parameters were in their default settings. Next, nodes that had edges with a correlation value of $r < 0.5$ and a weighted threshold of $p < 0.5$ were deleted. Finally, the nodes and edges of the DEGs shared by the LUAD and LUSC cohorts were visualized using the Cytoscape (<https://cytoscape.org/>) software⁴⁵.

The Human Protein Atlas (HPA)⁴⁶ is a public database containing more than 26,000 proteins and more than 17,000 target human genes (<https://www.proteinatlas.org/>). The molecular profiles of gasdermin family proteins in NSCLC patients were determined by using the HPA database

Differential expression analyses using UALCAN and GEPLA

UALCAN (<http://ualcan.path.uab.edu>) is an online platform for comprehensive, user-friendly and interactive analysis of omics data based on TCGA database and the clinical data of 31 cancer types. In our study, UALCAN was used to determine the differences in expression levels of gasdermin family genes between the N&T tissues of the NSCLC patients ($p\text{-value} < 0.05$).

GEPLA (<http://gepia.cancer-pku.cn/>)⁴⁷ is an interactive website for online analysis and mining of cancer data deposited in TCGA and Genotype-Tissue Expression (GTEx) databases. Utilizing a standard data processing pipeline, GEPLA can be used to perform customizable analyses, such as identification of gene biomarkers, N&T differential expression analyses, gene profiling based on cancer types or pathological stages, patient survival analyses, and identification of genes with similar functions. A multi-gene comparison method was adopted to investigate the relationships between mRNA expression levels of the

five gasdermin family genes and the staging of NSCLC tumors by using the “Expression DIY” function of GEPLA. *P* values < 0.05 in Student’s *t*-tests were indicative of significant differences.

Analysis of prognostic significances of the gasdermin genes using Kaplan Meier plotter

The Kaplan Meier Plotter database (<https://kmplot.com/analysis/index.php>) has comprehensive and sufficient chromatin immunoprecipitation sequencing (ChIP-seq) and RNA-seq data for analyses of gene expression at the mRNA level. The database contains data from GEO, EGA and TCGA databases and can accurately assess the effects of 54,000 genes on the survival of patients with 21 types of cancer. Breast cancer (6234 cases), ovarian cancer (2190 cases), lung cancer (3452 cases), and stomach cancer (1440 cases) are the most studied cancer datasets⁴⁸. In this study, we used the PostgreSQL server of the Kaplan Meier Plotter database, which integrates both gene expression and clinical data of cancer patients, to evaluate the prognostic significances of the gasdermin genes. Based on the mRNA levels of gasdermin family genes, the cancer patients were divided into high- and low-expression groups. Comparisons between the two groups of patients were performed by using the Kaplan-Meier survival chart, and the hazard ratios (HRs) and 95% confidence intervals (CIs) as well as the log-rank *p* values were calculated. *P* value < 0.05 is considered a significant difference.

cBioPortal, Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses

cBioPortal (<http://www.cbioportal.org/>) for Cancer Genomics is an open source for interactive exploration of multiple cancer genomics datasets⁴⁹. By using cBioPortal, we analyzed the genetic data of NSCLC and objectively analyzed the influences of the five gasdermin genes on the development and prognosis of NSCLC. In this study, we selected the LUAD (TCGA, Firehose Legacy) dataset, which includes 515 cases with pathology reports, and analyzed the genomic profiles of the five gasdermin genes, including their mutations, their putative copy-number alteration data from Genomic Identification of Significant Targets in Cancer (GISTIC), and their mRNA expression z-scores (RNASeq V2 RSEM, z-score threshold = ± 1.8)⁵⁰. Mutations in the gasdermin family genes and their relationships with the OS of NSCLC patients are shown as Kaplan-Meier plots, and the log-rank test was performed to determine the diversity between survival curves. A *p*-value < 0.05 was considered statistically significant. By using the “co-expression” module of cBioPortal, genes whose expression levels were closely associated with gasdermin gene mutations were identified. Afterwards, the top 30 co-expressed genes with the largest Pearson’s correlation coefficients with each gasdermin family gene were identified (**Table S3**). Using the GO and KEGG modules in the DAVID database (<https://david.ncifcrf.gov/>)⁵¹, the functions and pathways of the 150 co-expressed genes associated with gasdermin gene mutations were analyzed (*P*-value < 0.05).

Prediction and molecular docking analyses of small molecule drugs that can target the gasdermin family proteins

Connectivity Map (CMAP) (<http://www.broadinstitute.org>) is an online database that links small molecules to potential drug⁵². In this study, the differentially expressed gasdermin genes were imported into the CMAP database to obtain enrichment values (-1~1). An enrichment value close to 1 represents a positive correlation between the expression of a gene and a drug, while an enrichment value close to -1 represents a negative correlation. The top four most significantly enriched small molecules were selected according to the *P* value (the smaller the *P* value, the more significant the enrichment), and their 3D conformations were analyzed in PubChem (<http://www.pubchem.ncbi.nlm.gov>), a database containing structure information of small molecules. Using the four most significantly enriched small molecule drugs as ligands and the gasdermin proteins as receptors, a molecular docking analysis was performed using the docking method of “Auto Dock Vina” in the YASARA software⁵³.

Results

Differential expression analyses of pyroptosis-related genes in the NSCLC patients

We identified 4851 and 6513 differentially expressed genes (DEGs) in the LUAD and LUSC cohorts, respectively, of which the down-regulated genes accounted for 60.17% and 55.86%, respectively (**Fig. 2A**). Moreover, 40 and 45 pyroptosis-related genes were identified in the LUAD and LUSC cohorts, respectively, and 33 pyroptosis-related genes were found in both cohorts (**Figs. 2B-C**). Interestingly, these common pyroptosis-related genes exhibited similar differential expression patterns in the LUAD and LUSC cohorts, with 57.58% of the genes being down-regulated in tumor tissues. Previous studies have shown that the expression of *GSDMC* gene can be used as a prognostic indicator in LUAD patients, which can be utilized in the future treatment of lung cancers²³. In addition, downregulation of *GSDMD* gene was found to inhibit lung cancer tumor proliferation²⁶. Therefore, the expression of this gene is an effective prognostic indicator for NSCLC. Functional enrichment analyses showed that these genes were principally enriched in cellular production of functional proteins and regulation of transcription activity (**Fig. 2D**), such as positive regulation of cytokine production, regulation of interleukin-1 β production, regulation of DNA-binding activities of transcription factors, and regulation of interleukin-1 production, among which interleukin-1 β plays an essential role in pyroptosis.

Co-expression network analyses of pyroptosis-related genes and the DEGs identified in the LUAD and LUSC cohorts

Subsequently, we tried to identify correlated modules between the identified DEGs and the N&T tissues of the LUAD and LUSC cohorts and constructed a co-expression network using the pyroptosis-related genes and the identified DEGs with WGCNA²⁷ tools.

Our results indicated that the N&T tissues of the LUAD and LUSC cohorts were closely correlated with the DEGs, with a Pearson correlation coefficient *r* of > 0.5 (**Fig. 3**). For the LUAD cohort, we identified three gene modules (turquoise, red, and purple modules) associated with the tumor tissues, and the corresponding Pearson correlation coefficients were *r* = -0.71 (*P*-value < 5e-06), *r* = 0.57 (*P*-value < 0.002)

and $r = 0.52$ (P -value $< 3e-08$), respectively. Besides, a module (red) associated with the normal tissues (**Fig. 3A**) was identified for the LUAD cohort. These four modules contained eleven differentially expressed pyroptosis-related genes. For the LUSC cohort, we identified two modules (turquoise and green modules) related to the tumor tissues, and the corresponding Pearson correlation coefficients were $r = -0.73$ (P -value $< 5e-07$) and $r = 0.56$ (P -value $< 1e-02$), respectively. A blue module relevant to the normal tissues was identified for the LUSC cohort (**Fig. 3B**). Ten differentially expressed pyroptosis-related genes were involved in the three modules for the LUSC cohort.

The findings of our co-expression network analysis for the LUAD cohort are summarized in **Fig. 3C**. Four gene modules were associated with the N&T tissues in the LUAD cohort. Notably, there were three key pyroptosis-related genes (namely *GSDMB*, *GSDME* and *CARD16*) in the turquoise module, four key pyroptosis-related genes (namely *GSDMA*, *GSDMD*, *NLRP3* and *CASP3*) in the red module, and four key pyroptosis-related genes in the magenta module (namely *GSDMC*, *IL6*, *NLRP14* and *PJVK*). Most DEGs in the co-expression modules were downregulated. The proportions of downregulated genes in the red, magenta and turquoise modules were 66.7%, 58.3% and 55.6%, respectively. Notably, 63.6% of the pyroptosis-related genes were downregulated. Our co-expression network analysis of the LUSC cohort (**Fig. 3D**) revealed that there were three gene modules associated with the N&T tissues. There were three pyroptosis-related genes (namely *TLR4*, *TXNIP* and *TLR3*) in the blue module, four pyroptosis-related genes in the turquoise module (namely *GSDMB*, *GSDMD*, *GSDME* and *PLCG1*), and three pyroptosis-related genes in the green module (namely *IFI16*, *CASP5* and *GSDMC*). In addition, the proportions of downregulated genes in the blue, turquoise and green modules were 60%, 64% and 55.6%, respectively.

These results indicated that downregulation of the pyroptosis-related genes and the DEGs played an essential role in the development of LUAD and LUSC. In each module, one pyroptosis-related gene interacted with almost all the DEGs. This finding indicated that the pyroptosis-related genes were key regulators of the DEGs during the development of LUAD and LUSC. Furthermore, we noted that five gasdermin genes (*GSDMA/B/C/D/E*) and four gasdermin genes (*GSDMB/C/D/E*) were expressed in the tumor tissues of LUAD and LUSC patients, respectively. This indicated that these gasdermin genes might play important roles in the development of LUAD and LUSC. These results may provide a new perspective for identification of novel prognostic biomarkers and therapeutic targets for patients with LUAD and LUSC.

Verification of expression levels of the gasdermin proteins

In order to further verify the correlations between the gasdermin family proteins and NSCLC, we analyzed the protein levels of gasdermin family proteins in the N&T tissues of LUAD and LUSC patients based on the HPA database (**Fig. 4**). The results showed moderate *GSDMB/C/E* expression, negative *GSDMA* expression and strong *GSDMD* expression in normal tissues. The numbers of NSCLC patients with gasdermin proteins being expressed in their tumor tissues were *GSDMA*: 11 cases (negative: 10 cases and weak: 1 case); *GSDMB*: 8 cases (negative: 6 cases and weak: 2 cases); *GSDMC*: 11 cases (negative: 7 cases, strong: 2 cases, and moderate: 2 cases); *GSDMD*:

10 cases (negative: 3 cases, weak:3 cases, strong: 2 cases, and moderate: 2 cases); and GSDME: 12 cases (negative: 8 cases, weak: 3 cases, and negative: 1 case).

The above results showed that GSDMC was significantly upregulated in tumor tissues compared with the corresponding normal tissues, indicating a positive correlation between the expression of GSDMC and lung cancer development. Compared with GSDMC, GSDMA/B/D/E were expressed at much lower levels in tumor tissues, which is consistent with the WGCMA results. This also proved the credibility of our identification of the module genes.

Association between the aberrant expression of gasdermin genes and tumor stages in the NSCLC patients

Next, we used the UALCAN database to compare the expression levels of gasdermin genes between 515 primary tumor and 59 normal tissue samples from the LUAD cohort, and between 503 primary tumor and 52 normal tissue samples from the LUSC cohort. The expression levels of *GSDMA/B/C/E* genes were significantly higher in primary NSCLC tumor tissues than in the corresponding normal tissues, especially those of *GSDMB/C* genes (**Figs. 5A-E**). Furthermore, the expression level of *GSDMD* was higher in normal tissues than in primary tumor tissues in the LUSC cohort, and there was no significant difference in *GSDMD* expression between primary tumor tissues and normal tissues in the LUAD cohort. Using the GEPIA database, we also analyzed the correlations between the expression levels of gasdermin genes and the NSCLC patients' tumor stages (**Fig. S1**). The expression level of *GSDMB* was significantly correlated with the NSCLC patients' tumor stages, and NSCLC patients with earlier tumor stages tended to express a higher level of *GSDMB*. In order to comprehensively investigate the aberrant expression of the gasdermin family genes in the NSCLC patients, we also compared the relative expression levels of gasdermin family genes in the LUAD and LUSC tissues (**Fig. 5F**). Interestingly, *GSDMD* had the highest expression levels among gasdermin family genes in the LUAD and LUSC tissues.

These results indicated that the gasdermin family genes had significantly different expression levels in the LUAD and LUSC tissues. Meanwhile, the high expression of *GSDMB* may affect the tumor stages of the NSCLC patients.

Survival analyses of the gasdermin genes in the NSCLC patients

By using data mining methods in the Kaplan–Meier Plotter database, we investigated the correlations between the expression of gasdermin family genes and the OS, FP, and PPS of the NSCLC patients. To do this, the NSCLC patients were divided into high-expression and low-expression groups based on the cutoff values in each cohort (**Fig. 6**).

Our results indicated that the expression levels of *GSDMB/C/D* genes were significantly correlated with the OS, FP and PPS of the NSCLC patients. In contrast, there was no significant correlation between *GSDME* expression and either the OS or the FP of the NSCLC patients, whereas and increased *GSDME* expression was associated with the PPS of the NSCLC patients. For *GSDMB*, the survival rates of

patients in the high-expression group were higher than those of patients in the low-expression group, and the OS, FP and PPS of the NSCLC patients were positively correlated with *GSDMB* expression. These results suggest that *GSDMB* is a potential tumor suppressor gene. Unlike *GSDMB*, the expression levels of *GSDMC/D* showed negative correlations with the OS, FP and PPS of the NSCLC patients. Unfortunately, we were not able to establish a correlation between the expression level of *GSDMA* and the NSCLC patients' survival based on the Kaplan–Meier Plotter database.

Analysis of the prognostic significances of gasdermin gene mutations

Heritable mutations are crucial targets of research on clinical medicine and cancer development^{28,29}. In order to investigate the prognostic significances of gasdermin gene mutations, we performed mutation analyses on the five genes of *GSDMA/B/C/D/E* based on the cBioPortal online tool.

Gasdermin gene mutations were identified in 237 out of 515 patients with LUAD (46%) (**Fig. 7A**). *GSDMD*, *GSDMC* and *GSDME* had the highest mutation rates of 20%, 18%, and 14%, respectively. Besides, based on the expression levels of gasdermin genes (RNA-Seq V2 RSEM), we also calculated the Pearson's correlation coefficients between gasdermin family via the cBioPortal online tool for the LUAD cohort (TCGA, Firehose Legacy). The results indicated an evident positive correlation between *GSDMB*, *GSDMC*, *GSDMD* and *GSDME* genes (**Fig. 7B**). Finally, we analyzed the relationships between gasdermin gene mutations and the OS of the NSCLC patients. Kaplan Meier plot analysis and log-rank test results revealed that gasdermin gene mutations were associated with a shorter OS (**Fig. 7C**, $p = 5.668E-3$) in the NSCLC patients.

By analyzing the prognostic significances of gasdermin gene mutations in the NSCLC patients, we obtained 150 genes whose expression levels were significantly associated with gasdermin gene mutations by using the "co-expression" module of cBioPortal. Subsequently, we used the GO and KEGG modules in DAVID to analyze the potential functions of these 180 genes, and the results are shown in **Fig. 7D**. We found that gasdermin gene mutations in NSCLC were significantly associated with biological processes (BP) terms, including GO: 0006954 (inflammatory response), GO: 0006955 (immune response), GO: 0045087 (innate immune response), GO: 0002250 (adaptive immune response) and GO: 0050900 (leukocyte migration). Furthermore, we also found that cellular components (CC) terms such as GO: 0016021 (integral component of membrane), GO: 0005886 (plasma membrane), GO: 0070062 (extracellular exosome), GO: 0005887 (integral component of plasma membrane) and GO: 0005615 (extracellular space) were significantly related to gasdermin gene mutations. Among molecular functions (MF) terms, however, only GO: 0044389 (ubiquitin-like protein ligase binding) was associated with gasdermin gene mutations. The KEGG analysis revealed that hsa05150 (*Staphylococcus aureus* infection) and hsa04610 (Complement and coagulation cascades) were closely related to the gasdermin gene mutations in NSCLC.

The above results indicated that gasdermin gene mutations may play potentially important roles in the clinical treatment of NSCLC patients and in the development of NSCLC models. In particular, *GSDMC/D/E*,

which had the highest mutation rates among the gasdermin family genes and an apparent positive correlation among them, may be important for the prognosis of NSCLC patients. Furthermore, functional enrichment analyses suggest that genes whose expression levels were significantly associated with gasdermin gene mutations may play important roles in inflammatory and immune responses.

Prediction and molecular docking analyses of small molecule drugs that can potentially target gasdermin family proteins

In order to identify novel therapeutic drugs for NSCLC, we used the CMAP database to match gasdermin family proteins with therapeutic small molecules. The top four most significantly enriched small molecule drugs and their corresponding enrichment values and 3D conformations are shown in **Table 1** and **Fig. 8A** respectively. Among these small molecule drugs, cefotiam (enrichment value: -0.826), metanephrine (enrichment value: -0.739), and vorinostat (enrichment value: -0.525) were associated with significant negative fractions. These candidate small molecule drugs have the potential to reverse gene expression profiles in NSCLC, providing a guidance for the development of targeted drugs for NSCLC.

Our molecular docking analysis revealed that the four small molecules mainly interacted with GSDMB via H-bonds and hydrophobic interactions (**Fig. 8B**). For example, metanephrine can potentially form a H-bonds (2.0Å) with the O atom of Asp65 of GSDMB. In addition, metanephrine can potentially form hydrophobic interactions with Trp230, Trp340, Glu153, and Tyr155 of GSDMB.

Table 1: The four most significantly enriched small molecule drugs.

Rank	CMAP name	Mean	N	Enrichment value	P	CID
1	Erastin	0.701	4	0.838	0.00107	11214940
2	Vorinostat	-0.508	12	-0.525	0.00136	5311
3	Cefotiam	-0.701	4	-0.826	0.00177	43708
4	Metanephrine	-0.691	5	-0.739	0.00244	21100

Discussion

Pyrolysis plays important roles in tumor proliferation, invasion and metastasis. Therefore, exploring the regulatory mechanisms of pyrolysis in tumor cells and identification of pyrolysis-related diagnostic markers for cancers may benefit the clinical treatment and prevention of cancers^{30,31}. In this study, we analyzed the expression profiles of differentially expressed pyroptosis-related genes in the N&T tissues of LUAD and LUSC patients, predicted N&T tissues-related modules through WGCNA, and identified the associations between gasdermin genes with tumor tissues of the two cancers through co-expression network analyses of pyrolysis-related genes and DEGs. Next, we used the HPA database to further verify the correlations between the gasdermin genes and NSCLC. To our knowledge, this should be the first systematic study of the gasdermin family genes in N&T tissues of NSCLC patients.

GSDMA has a regulatory effect on the expression of neighboring genes in the blood, which can reduce the incidence of systemic sclerosis and contribute to the treatment of childhood acute lymphoblastic leukemia³²⁻³⁴. As a functional gene implicated in the development of asthma, *GSDMB* is involved in multiple pathways that can regulate the pathogenesis of asthma and slow down the deterioration of pneumonia³⁵. According to a previous study on the expression level of *GSDMC*, *GSDMC* may play important roles in the treatment of lumbar spinal stenosis and LUAD^{23,36}. In addition, it was observed that downregulation of *GSDMD* could slow down the development of LUAD by regulating the EGFR/Akt signaling pathway and had an impact on the treatment of acute kidney injury and diabetes, indicating that *GSDMD* is an important prognostic biomarker for these diseases^{26,37}. *GSDME* is important for inducing pyroptosis in lung cancers and plays important roles in immune responses, clinical diagnosis and treatments of tumors³⁸. In our study, we found that the expression levels of *GSDMA/B/C/E* were higher in tumor tissues than in normal tissues. *GSDMD* had the highest expression level among the gasdermin family genes and its expression was significantly higher in normal tissues than in tumor tissues. These results were consistent with those of differential expression and co-expression network analyses, verifying the results of these analyses and demonstrating the significant roles of gasdermin family genes in NSCLC. In addition, we found that the expression level of *GSDMB* was significantly associated with tumor staging and survival in the NSCLC patients. Therefore, *GSDMB* may become a new therapeutic target in the future. *GSDMC/D/E* had the highest gene mutation rates among the gasdermin family genes. Notably, high expression levels of *GSDMC/D* were significantly correlated with poor OS, FP, and PPS in the NSCLC patients, indicating that *GSDMC/D* are potential oncogenes for NSCLC.

To develop new targeted drug therapies for NSCLC, we used the CMAP database to identify small molecule drugs that can potentially target the gasdermin family proteins. The higher the enrichment value of a small molecule drug, the more likely it can target the gasdermin family proteins. Among the enriched small molecule drugs, metanephrine (enrichment value: -0.739) is an epinephrine metabolite widely used in the production of enzyme-linked immunosorbent assay kits, but has not been used in the treatment of NSCLC so far; cefotiam (enrichment value: -0.826), a second-generation cephalosporin widely used in the production of antibiotics, has good antibacterial effects and has been widely used in clinical practice. Our molecular docking analyses revealed that the four small molecules mainly interacted with *GSDMB* via H-bonds and hydrophobic interactions. Our results provide a valuable guidance for the development of novel small molecule drugs for NSCLC.

In conclusion, in this comprehensive multi-dimensional comparison study, we explored the prognostic significances of gasdermin family genes for NSCLC by using multiple online databases and bioinformatics technologies. Our results provide new insights into the prognosis and treatment of NSCLC.

Declarations

Acknowledgments

This publication was financially supported by the Nanhu Scholars Program for Young Scholars of XYNU.

Author contributions

Shuinbin Lian and Lei Wang conceived the work. Yu Zhang and Yongjie Zhou collected the data. Zixiao Liu and Xudong Liu analyzed the data. Zixiao Liu. and Shuinbin Lian. wrote the paper with input from all the co-authors.

Competing interests

The authors declare no competing interests.

References

1. Nasim, F., Sabath, B. F. & Eapen, G. A. Lung Cancer. *The Medical clinics of North America* **103**, 463-473, doi:10.1016/j.mcna.2018.12.006 (2019).
2. Jacobsen, M. M. *et al.* Timeliness of access to lung cancer diagnosis and treatment: A scoping literature review. *Lung cancer (Amsterdam, Netherlands)* **112**, 156-164, doi:10.1016/j.lungcan.2017.08.011 (2017).
3. Hirsch, F. R. *et al.* Lung cancer: current therapies and new targeted treatments. *Lancet (London, England)* **389**, 299-311, doi:10.1016/s0140-6736(16)30958-8 (2017).
4. Cheung, C. H. Y. & Juan, H. F. Quantitative proteomics in lung cancer. *Journal of biomedical science* **24**, 37, doi:10.1186/s12929-017-0343-y (2017).
5. Schabath, M. B. & Cote, M. L. Cancer Progress and Priorities: Lung Cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **28**, 1563-1579, doi:10.1158/1055-9965.Epi-19-0221 (2019).
6. Hamann, H. A., Ver Hoeve, E. S., Carter-Harris, L., Studts, J. L. & Ostroff, J. S. Multilevel Opportunities to Address Lung Cancer Stigma across the Cancer Control Continuum. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer* **13**, 1062-1075, doi:10.1016/j.jtho.2018.05.014 (2018).
7. Oudkerk, M., Liu, S., Heuvelmans, M. A., Walter, J. E. & Field, J. K. Lung cancer LDCT screening and mortality reduction - evidence, pitfalls and future perspectives. *Nature reviews. Clinical oncology* **18**, 135-151, doi:10.1038/s41571-020-00432-6 (2021).
8. Oberndorfer, F. & Müllauer, L. Molecular pathology of lung cancer: current status and perspectives. *Current opinion in oncology* **30**, 69-76, doi:10.1097/cco.0000000000000429 (2018).
9. Karki, R. & Kanneganti, T. D. Diverging inflammasome signals in tumorigenesis and potential targeting. *Nature reviews. Cancer* **19**, 197-214, doi:10.1038/s41568-019-0123-y (2019).
10. Xia, X. *et al.* The role of pyroptosis in cancer: pro-cancer or pro-"host"? *Cell death & disease* **10**, 650, doi:10.1038/s41419-019-1883-8 (2019).

11. Awad, F. *et al.* Inflammasome biology, molecular pathology and therapeutic implications. *Pharmacology & therapeutics* **187**, 133-149, doi:10.1016/j.pharmthera.2018.02.011 (2018).
12. Man, S. M. & Kanneganti, T. D. Regulation of inflammasome activation. *Immunological reviews* **265**, 6-21, doi:10.1111/imr.12296 (2015).
13. Ratajczak, M. Z. *et al.* NLRP3 inflammasome couples purinergic signaling with activation of the complement cascade for the optimal release of cells from bone marrow. *Leukemia* **33**, 815-825, doi:10.1038/s41375-019-0436-6 (2019).
14. Frank, D. & Vince, J. E. Pyroptosis versus necroptosis: similarities, differences, and crosstalk. *Cell death and differentiation* **26**, 99-114, doi:10.1038/s41418-018-0212-6 (2019).
15. Heo, M. J. *et al.* Alcohol dysregulates miR-148a in hepatocytes through FoxO1, facilitating pyroptosis via TXNIP overexpression. *Gut* **68**, 708-720, doi:10.1136/gutjnl-2017-315123 (2019).
16. Heneka, M. T., McManus, R. M. & Latz, E. Inflammasome signalling in brain function and neurodegenerative disease. *Nature reviews. Neuroscience* **19**, 610-621, doi:10.1038/s41583-018-0055-7 (2018).
17. Feng, S., Fox, D. & Man, S. M. Mechanisms of Gasdermin Family Members in Inflammasome Signaling and Cell Death. *Journal of molecular biology* **430**, 3068-3080, doi:10.1016/j.jmb.2018.07.002 (2018).
18. Liu, Y. *et al.* Visualization of perforin/gasdermin/complement-formed pores in real cell membranes using atomic force microscopy. *Cellular & molecular immunology* **16**, 611-620, doi:10.1038/s41423-018-0165-1 (2019).
19. Ding, J. *et al.* Pore-forming activity and structural autoinhibition of the gasdermin family. *Nature* **535**, 111-116, doi:10.1038/nature18590 (2016).
20. Broz, P., Pelegrín, P. & Shao, F. The gasdermins, a protein family executing cell death and inflammation. *Nature reviews. Immunology* **20**, 143-157, doi:10.1038/s41577-019-0228-2 (2020).
21. Zhang, Y., Chen, X., Gueydan, C. & Han, J. Plasma membrane changes during programmed cell deaths. *Cell research* **28**, 9-21, doi:10.1038/cr.2017.133 (2018).
22. Söderman, J., Berglind, L. & Almer, S. Gene Expression-Genotype Analysis Implicates GSDMA, GSDMB, and LRRC3C as Contributors to Inflammatory Bowel Disease Susceptibility. *BioMed research international* **2015**, 834805, doi:10.1155/2015/834805 (2015).
23. Wei, J. *et al.* Overexpression of GSDMC is a prognostic factor for predicting a poor outcome in lung adenocarcinoma. *Molecular medicine reports* **21**, 360-370, doi:10.3892/mmr.2019.10837 (2020).
24. Xi, G. *et al.* GSDMD is required for effector CD8(+) T cell responses to lung cancer cells. *International immunopharmacology* **74**, 105713, doi:10.1016/j.intimp.2019.105713 (2019).
25. Zhang, Z. *et al.* Gasdermin E suppresses tumour growth by activating anti-tumour immunity. *Nature* **579**, 415-420, doi:10.1038/s41586-020-2071-9 (2020).
26. Gao, J. *et al.* Downregulation of GSDMD attenuates tumor proliferation via the intrinsic mitochondrial apoptotic pathway and inhibition of EGFR/Akt signaling and predicts a good

- prognosis in non-small cell lung cancer. *Oncology reports* **40**, 1971-1984, doi:10.3892/or.2018.6634 (2018).
27. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics* **9**, 559, doi:10.1186/1471-2105-9-559 (2008).
 28. Cheng, H. H., Sokolova, A. O., Schaeffer, E. M., Small, E. J. & Higano, C. S. Germline and Somatic Mutations in Prostate Cancer for the Clinician. *Journal of the National Comprehensive Cancer Network : JNCCN* **17**, 515-521, doi:10.6004/jnccn.2019.7307 (2019).
 29. Coelho, M. C., Pinto, R. M. & Murray, A. W. Heterozygous mutations cause genetic instability in a yeast model of cancer evolution. *Nature* **566**, 275-278, doi:10.1038/s41586-019-0887-y (2019).
 30. Fang, Y. *et al.* Pyroptosis: A new frontier in cancer. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie* **121**, 109595, doi:10.1016/j.biopha.2019.109595 (2020).
 31. Zhaolin, Z., Guohua, L., Shiyuan, W. & Zuo, W. Role of pyroptosis in cardiovascular disease. *Cell proliferation* **52**, e12563, doi:10.1111/cpr.12563 (2019).
 32. Moreno-Moral, A. *et al.* Changes in macrophage transcriptome associate with systemic sclerosis and mediate GSDMA contribution to disease risk. *Annals of the rheumatic diseases* **77**, 596-601, doi:10.1136/annrheumdis-2017-212454 (2018).
 33. Gatineau-Sailliant, S. *et al.* Impact of DARC, GSDMA and CXCL2 polymorphisms on induction toxicity in children with acute lymphoblastic leukemia: A complementary study. *Leukemia research* **86**, 106228, doi:10.1016/j.leukres.2019.106228 (2019).
 34. Das, S. *et al.* GSDMB induces an asthma phenotype characterized by increased airway responsiveness and remodeling without lung inflammation. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 13132-13137, doi:10.1073/pnas.1610433113 (2016).
 35. Das, S., Miller, M. & Broide, D. H. Chromosome 17q21 Genes ORMDL3 and GSDMB in Asthma and Immune Diseases. *Advances in immunology* **135**, 1-52, doi:10.1016/bs.ai.2017.06.001 (2017).
 36. Jiang, H. *et al.* Two GWAS-identified variants are associated with lumbar spinal stenosis and Gasdermin-C expression in Chinese population. *Scientific reports* **10**, 21069, doi:10.1038/s41598-020-78249-7 (2020).
 37. Wang, Y. *et al.* TLR4/NF- κ B Signaling Induces GSDMD-Related Pyroptosis in Tubular Cells in Diabetic Kidney Disease. *Frontiers in endocrinology* **10**, 603, doi:10.3389/fendo.2019.00603 (2019).
 38. Zhang, C. C. *et al.* Chemotherapeutic paclitaxel and cisplatin differentially induce pyroptosis in A549 lung cancer cells via caspase-3/GSDME activation. *Apoptosis : an international journal on programmed cell death* **24**, 312-325, doi:10.1007/s10495-019-01515-1 (2019).
 39. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
 40. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).

41. Ye, Y., Dai, Q. & Qi, H. A novel defined pyroptosis-related gene signature for predicting the prognosis of ovarian cancer. *Cell death discovery* **7**, 71, doi:10.1038/s41420-021-00451-x (2021).
42. Wang, B. & Yin, Q. AIM2 inflammasome activation and regulation: A structural perspective. *Journal of structural biology* **200**, 279-282, doi:10.1016/j.jsb.2017.08.001 (2017).
43. Wang, Y. *et al.* Chemotherapy drugs induce pyroptosis through caspase-3 cleavage of a gasdermin. *Nature* **547**, 99-103, doi:10.1038/nature22393 (2017).
44. Liu, L. & Sun, B. Neutrophil pyroptosis: new perspectives on sepsis. *Cellular and molecular life sciences : CMLS* **76**, 2031-2042, doi:10.1007/s00018-019-03060-1 (2019).
45. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* **13**, 2498-2504, doi:10.1101/gr.1239303 (2003).
46. Uhlen, M. *et al.* A pathology atlas of the human cancer transcriptome. *Science (New York, N.Y.)* **357**, doi:10.1126/science.aan2507 (2017).
47. Tang, Z. *et al.* GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic acids research* **45**, W98-w102, doi:10.1093/nar/gkx247 (2017).
48. Nagy, Á., Lániczky, A., Menyhárt, O. & Gyórfy, B. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Scientific reports* **8**, 9227, doi:10.1038/s41598-018-27521-y (2018).
49. Wu, P. *et al.* Integration and Analysis of CPTAC Proteomics Data in the Context of Cancer Genomics in the cBioPortal. *Molecular & cellular proteomics : MCP* **18**, 1893-1898, doi:10.1074/mcp.TIR119.001673 (2019).
50. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science signaling* **6**, pl1, doi:10.1126/scisignal.2004088 (2013).
51. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**, 44-57, doi:10.1038/nprot.2008.211 (2009).
52. Lamb, J. *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science (New York, N.Y.)* **313**, 1929-1935, doi:10.1126/science.1132939 (2006).
53. Land, H. & Humble, M. S. YASARA: A Tool to Obtain Structural Guidance in Biocatalytic Investigations. *Methods in molecular biology (Clifton, N.J.)* **1685**, 43-67, doi:10.1007/978-1-4939-7366-8_4 (2018).

Figures

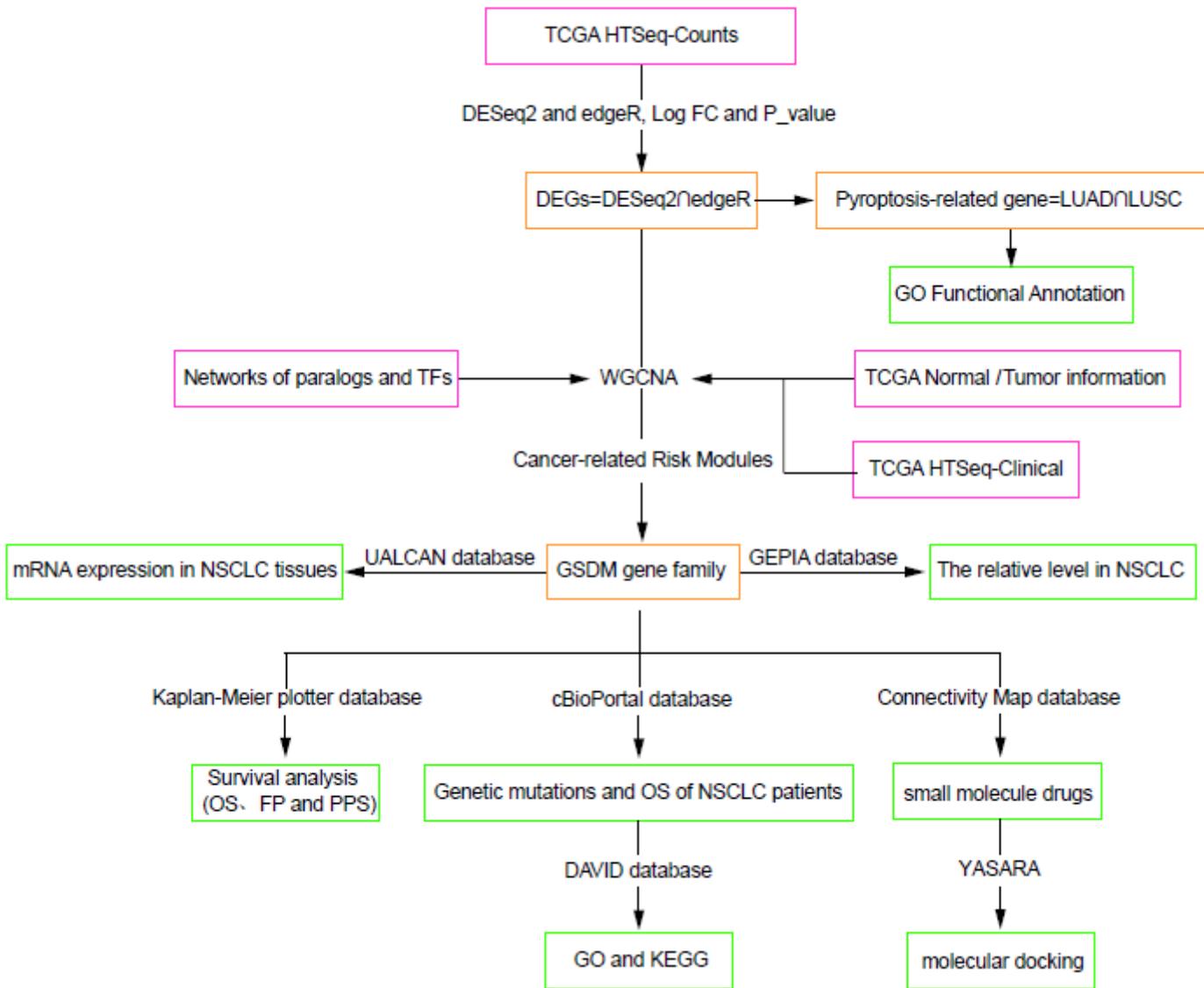


Figure 1

A diagram showing the workflow of this study. The workflow includes data collection and processing (the pink box), differential expression analysis and identification of key genes (the light yellow box), multi-database analyses of gene functions, gene expression levels, survival rates and potential targeted small molecule drugs (the pale green box).

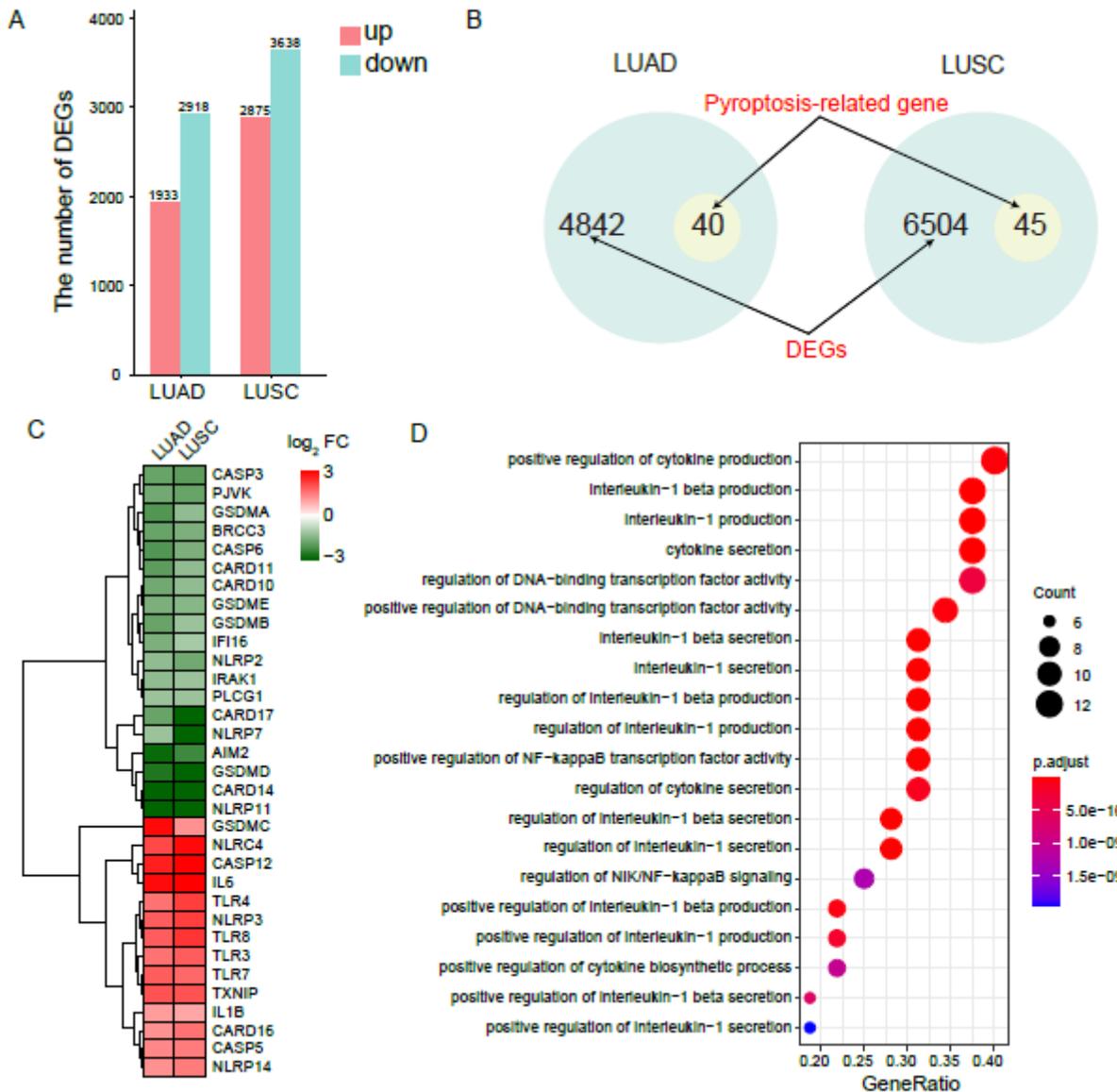


Figure 2

(A) Histograms of upregulated and downregulated DEGs in the LUAD and LUSC cohorts; (B) Venn diagrams of differentially expressed pyroptosis-related genes in the LUAD and LUSC cohorts; (C) Heat maps of the expression profiles of the 33 differentially expressed pyroptosis-related genes in NSCLC; (D) GO functional annotations of the top 20 differentially expressed pyroptosis-related genes.

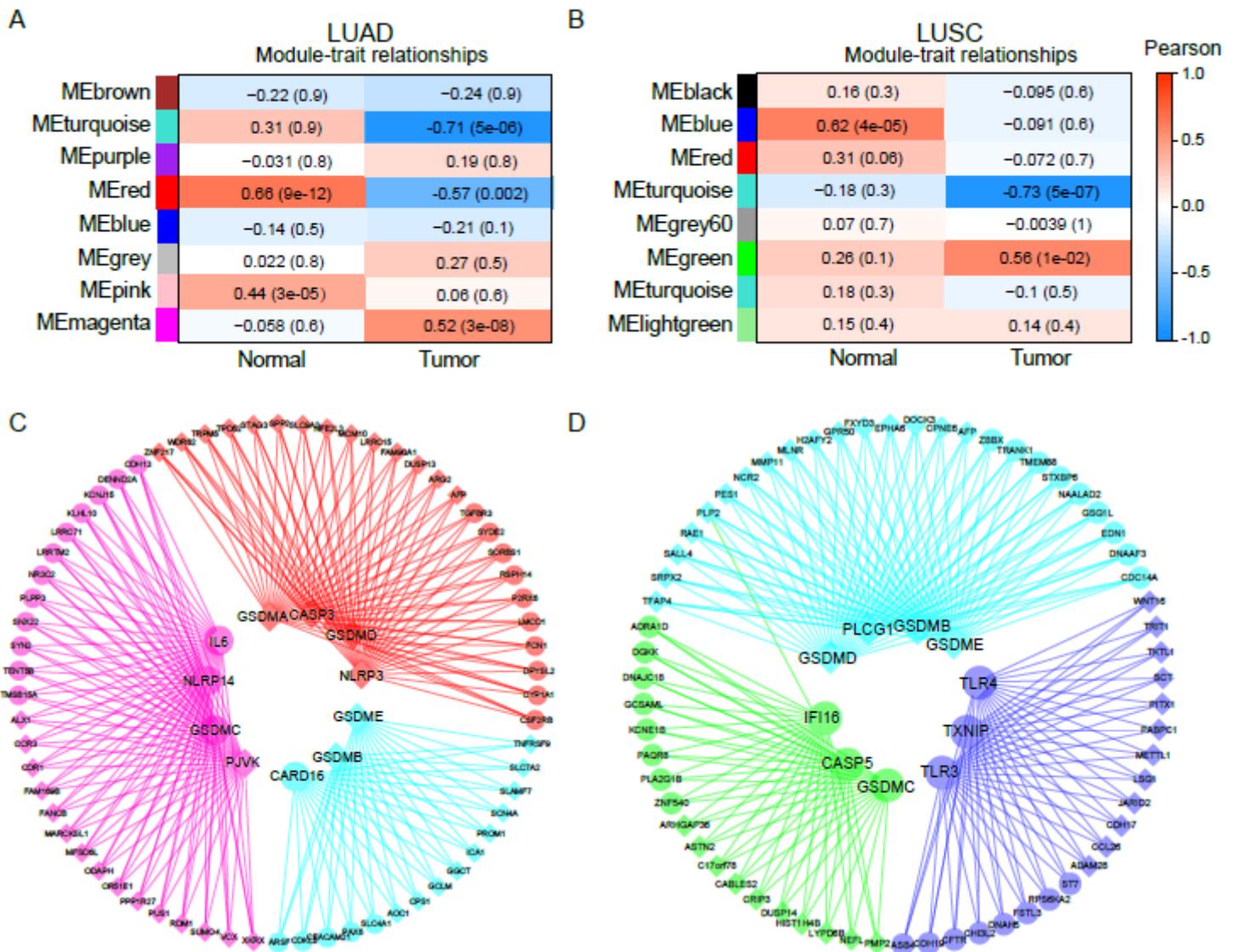


Figure 3

(A-B) Gene co-expression modules associated with the N&T tissues of the LUAD and LUSC patients. The colors indicate the gene co-expression modules. The Pearson's correlation coefficients and the corresponding P-values are presented in the form of Pearson's correlation coefficient (P-value). (C-D) Co-expression networks between pyroptosis-related genes and the gene modules related to the N&T tissues of the LUAD and LUSC patients. The color represents the corresponding gene modules related to the N&T tissues; the lines represent the correlations in expression between the pyroptosis-related genes and the DEGs. The inner circles are the key pyroptosis-related genes, while the outer circles are the DEGs. Circles represent upregulated genes, while quadrilaterals represent down-regulated genes.

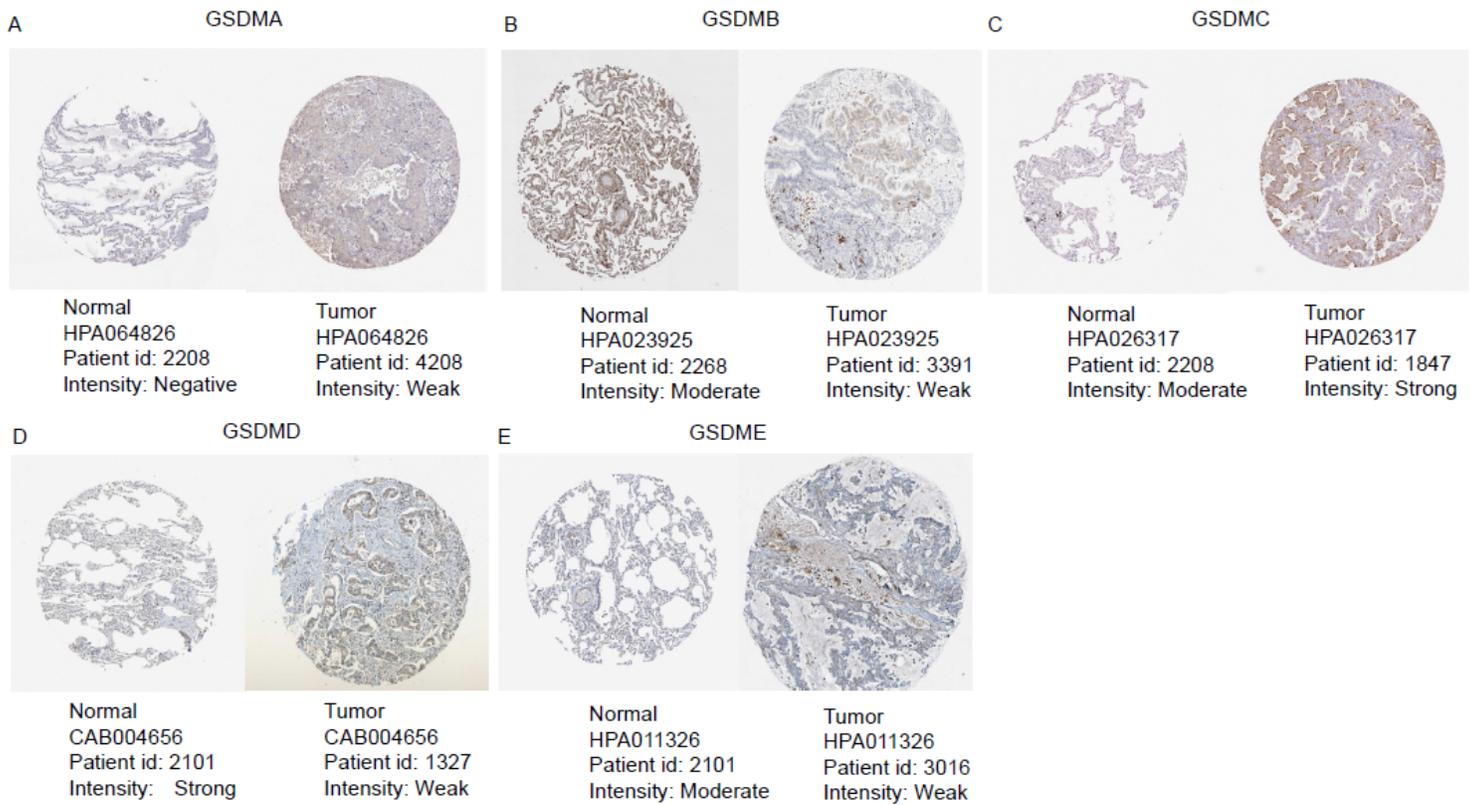


Figure 4

Levels of gasdermin family proteins in the N&T tissues derived from the HPA database. (A) The protein level of GSDMA; (B) the protein level of GSDMB; (C) the protein level of GSDMC; (D) the protein level of GSDMD; and (E) the protein level of GSDME.

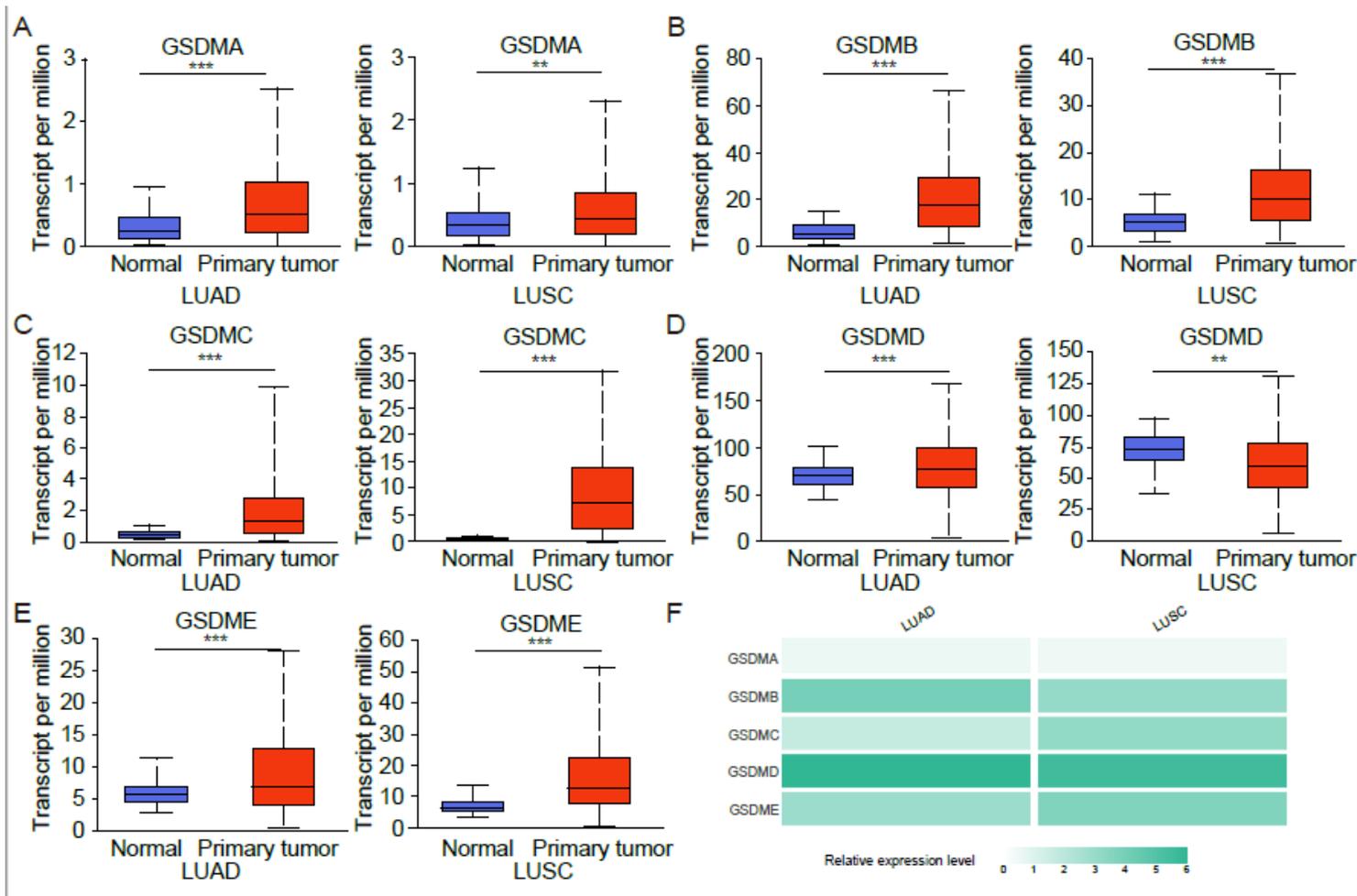


Figure 5

(A-E) The expression levels of gasdermin family genes in the N&T tissues of the LUAD and LUSC patients (UALCAN, ***p-value < 0.001, **p-value < 0.01). (F) The relative expression levels of gasdermin family genes in the NSCLC patients.

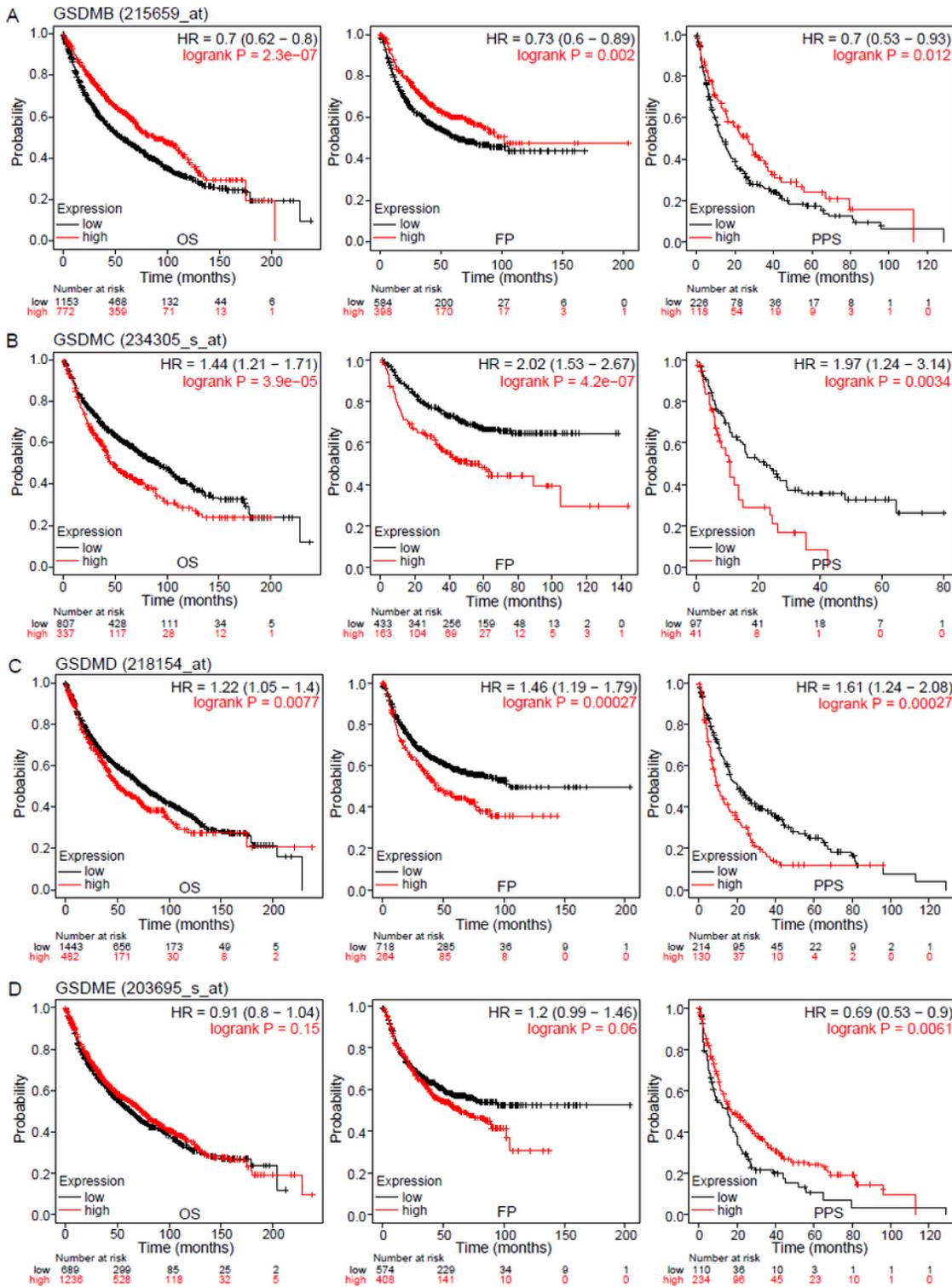


Figure 6

Survival analyses of (A) GSDMB; (B) GSDMC; (C) GSDMD; and (D) GSDME genes in the NSCLC patients. In the comparative analysis of OS, FP, and FPS survival curves, the red curve represents high expression of these gasdermin family genes and the black curve represents low expression of these gasdermin family genes (p-value < 0.05).

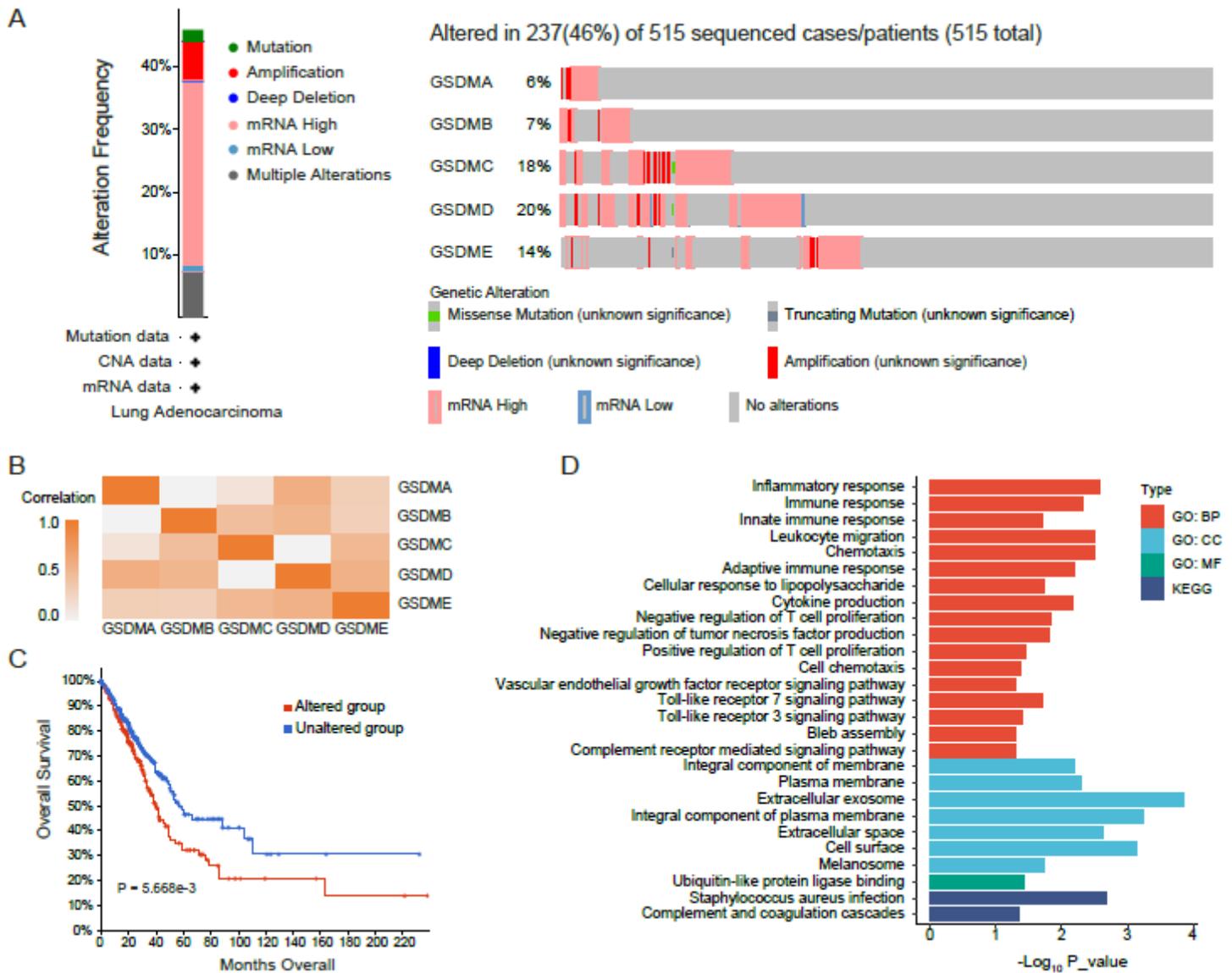


Figure 7

(A) Summary of gasdermin gene mutations in the LUAD cohort; (B) Correlations among the gasdermin family genes in the LUAD cohort; (C) the OS of the NSCLC patients; (D) GO and KEGG enrichment analyses of the gasdermin family genes and the 150 genes whose expression levels were closed associated with gasdermin gene mutations in the NSCLC patients.

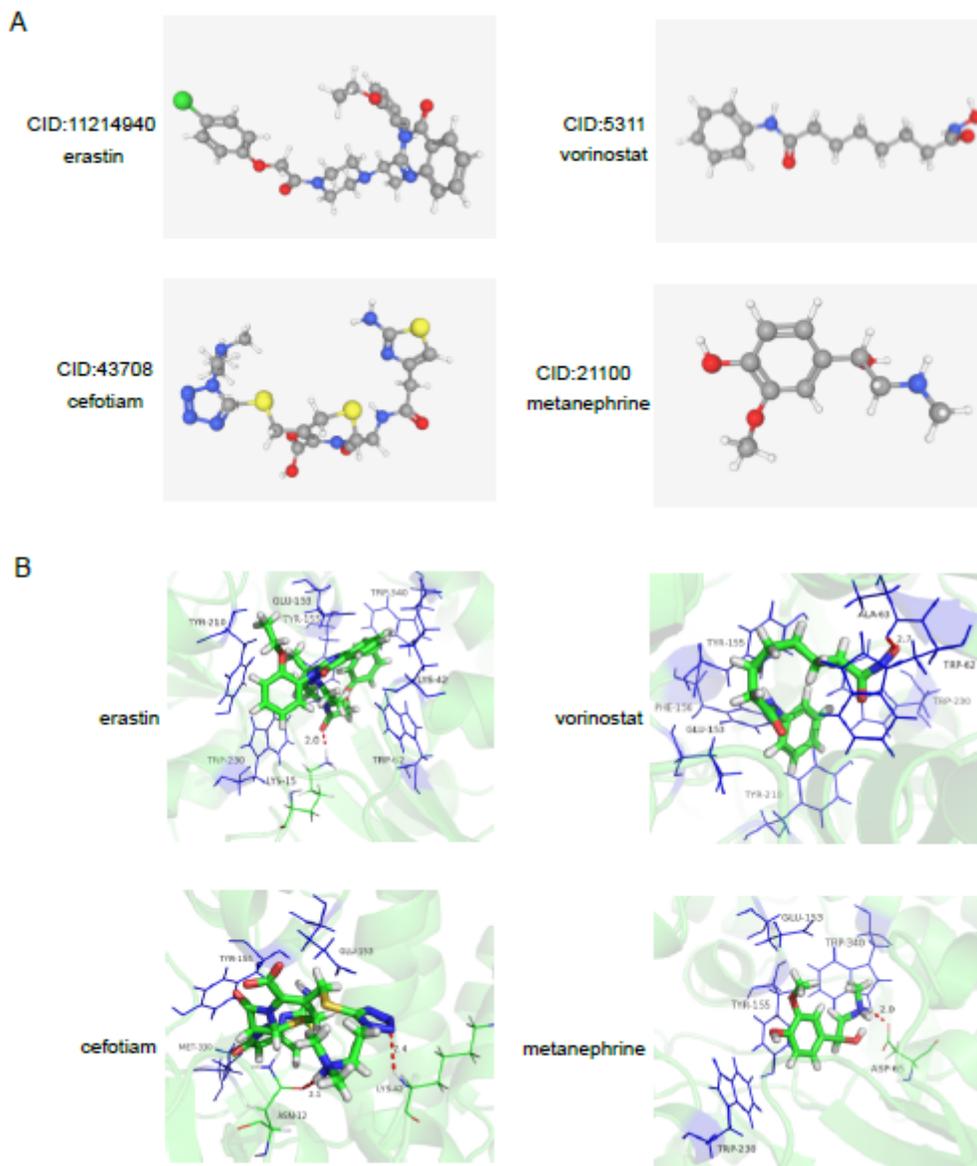


Figure 8

(A) The 3D conformations of the four most significantly enriched small molecule drugs by the gasdermin family proteins; (B) Molecular docking models of the four small molecules binding to GSDMB. The red dotted lines represent H-bonds and the blue thin lines represent hydrophobic interactions.

Supplementary Files

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

- [Fig.S1.pdf](#)
- [TableS1.xlsx](#)
- [TableS2.xlsx](#)
- [TableS3.xlsx](#)