



## OPEN ACCESS

EDITED BY  
Ping-Ching Hsu,  
University of Arkansas for Medical  
Sciences, United States

REVIEWED BY  
Guojun Lu,  
Nanjing Chest Hospital, China  
Erjia Zhu,  
Tongji University, China

\*CORRESPONDENCE  
Yan Yu,  
yuyan@hrbmu.edu.cn

<sup>†</sup>These authors have contributed equally  
to this work and share first authorship

SPECIALTY SECTION  
This article was submitted to Cancer  
Genetics and Oncogenomics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 21 March 2022  
ACCEPTED 19 July 2022  
PUBLISHED 22 August 2022

CITATION  
Zeng Y, Zhou L, Jia D, Pan B, Li X and  
Yu Y (2022), Comprehensive analysis for  
clarifying transcriptomics landscapes of  
spread through air spaces in  
lung adenocarcinoma.  
*Front. Genet.* 13:900864.  
doi: 10.3389/fgene.2022.900864

COPYRIGHT  
© 2022 Zeng, Zhou, Jia, Pan, Li and Yu.  
This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# Comprehensive analysis for clarifying transcriptomics landscapes of spread through air spaces in lung adenocarcinoma

Yuan Zeng<sup>1†</sup>, Lingli Zhou<sup>2†</sup>, Dexin Jia<sup>1</sup>, Bo Pan<sup>1</sup>, Xiaomei Li<sup>3</sup> and Yan Yu<sup>1\*</sup>

<sup>1</sup>Department of Medical Oncology, Harbin Medical University Cancer Hospital, Harbin, China,

<sup>2</sup>Department of Respiratory Medicine, Suizhou Hospital, Hubei University of Medicine, Hubei, China,

<sup>3</sup>Department of Pathology, Harbin Medical University Cancer Hospital, Harbin, China

Patients with spread through air spaces (STAS) have worse postoperative survival and a higher recurrence rate in lung adenocarcinoma, even in the earliest phases of the disease. At present, the molecular pathogenesis of STAS is not well understood. Therefore, to illustrate the underlying pathogenic mechanism of STAS, we accomplished a comprehensive analysis of a microarray dataset of STAS. Differential expression analysis revealed 841 differentially expressed genes (DEGs) between STAS<sub>positive</sub> and STAS<sub>negative</sub> groups. Additionally, we acquired two hub genes associated with survival. Gene set variation analysis (GSVA) confirmed that the main differential signaling pathways between the two groups were hypoxia VHL targets, PKC, and pyrimidine metabolism pathways. Analysis of immune activity showed that the increased expression of MHC-class-I was observed in the STAS<sub>positive</sub> group. These findings provided novel insights for a better knowledge of pathogenic mechanisms and potential therapeutic markers for STAS treatment.

## KEYWORDS

lung adenocarcinoma, spread through air spaces, differential expression analysis, protein–protein interaction, enrichment analysis, RNA-seq

## Introduction

Lung adenocarcinoma is the most frequent histologic subtype of lung cancer and has a high risk of recurrence or metastasis at the early disease stage (Shi et al., 2016). It has been proposed that lung cancer metastasis occurs through hematogenous spread, lymphatic spread, and direct infiltration (Han et al., 2021). In 2015, the World Health Organization (WHO) pointed out that spread through air spaces (STAS) is a recently identified pattern of tumor invasion (metastasis) (Travis et al., 2015). STAS is defined as tumor cells within air spaces in the lung parenchyma beyond the edge of the primary tumor (Eguchi et al., 2019; Qi et al., 2021). Notably, several studies have demonstrated that STAS is an independent factor indicating lung cancer recurrence and poor outcomes in patients (Yang et al., 2018; Liu et al., 2020; Onozato et al., 2021).

TABLE 1 Clinical factors of STAS<sub>positive</sub> and STAS<sub>negative</sub> groups.

Factors		Overall (n = 19)	Positive (n = 11)	Negative (n = 8)	p-value
Age group (years, %)	≥60	11 (57.9)	7 (63.6)	4 (50.0)	0.658
	<60	8 (42.1)	4 (36.4)	4 (50.0)	
Gender (%)	Female	9 (47.4)	4 (36.4)	5 (62.5)	0.37
	Male	10 (52.6)	7 (63.6)	3 (37.5)	
Tumor_size (cm)	Median [IQR]	2.00 [1.35, 3.00]	2.50 [1.50, 3.00]	1.75 [1.15, 2.62]	0.558
N <sub>stage</sub> (%)	N0	17 (89.5)	9 (81.8)	8 (100.0)	0.228
	N2	2 (10.5)	2 (18.2)	0 (0.0)	
p <sub>stage</sub> (%)	IA1	4 (21.1)	2 (18.2)	2 (25.0)	0.454
	IA2	5 (26.3)	2 (18.2)	3 (37.5)	
	IA3	7 (36.8)	5 (45.5)	2 (25.0)	
	IB	1 (5.3)	0 (0.0)	1 (12.5)	
	IIIA	2 (10.5)	2 (18.2)	0 (0.0)	
Smoking (%)	No	11 (57.9)	6 (54.5)	5 (62.5)	1
	Yes	8 (42.1)	5 (45.5)	3 (37.5)	
EGFR (%)	Mutant-type	4 (21.1)	2 (18.2)	2 (25.0)	1
	Wild-type	15 (78.9)	9 (81.8)	6 (75.0)	
ALK (%)	Mutant-type	4 (21.1)	3 (27.3)	1 (12.5)	0.603
	Wild-type	15 (78.9)	8 (72.7)	7 (87.5)	
KRAS (%)	Mutant-type	4 (21.1)	3 (27.3)	1 (12.5)	0.603
	Wild-type	15 (78.9)	8 (72.7)	7 (87.5)	

Additionally, STAS is an insidious pattern of invasion that is invisible to pathologists on gross examination and surgeons on external analysis of tumor specimens at the time of surgery. Unfortunately, there is no current, reliable radiological method to detect STAS (Kadota et al., 2015). Therefore, there is a pressing need to investigate the novel potential and pathogenic mechanisms of STAS.

In lung adenocarcinoma, STAS occurrence is related to the interactions between neutrophils and the tumor. The tumor environment releases interleukin-8 to promote the apoptosis of the local tumor neutrophils and activates neutrophils to promote tumor cell shedding from the primary tumor body, allowing them to migrate along the lung basement membrane to another alveolar cavity, resulting in STAS (Wislez et al., 2004; Wislez et al., 2007). With the developments and broad applications of high-throughput technology in biological and biomedical research fields, these tools can monitor genome-wide gene transcription levels and provide insight into biological processes involved in gene regulation (Gerstner et al., 2020; Hess et al., 2020). Therefore, we predicted the pathogenic mechanisms and candidate markers of STAS using high-throughput sequencing.

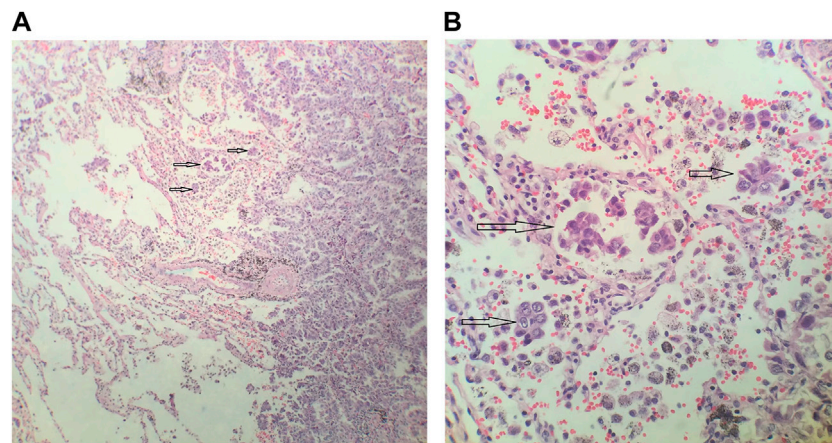
In our study, we performed an integrative analysis of the gene expression dataset of STAS in lung adenocarcinoma. Firstly, differential expression analysis showed that there were 841 differentially expressed genes (DEGs). Among them, two hub genes, *CXCL8* and *CPB2*, were identified as relevant to

survival. Finally, through gene set variation analysis (GSVA), we found that hypoxia VHL targets, PKC, and pyrimidine metabolism pathways were the three main differential signaling pathways. We also observed increased MHC-class-I expression in the STAS<sub>positive</sub> group by immune activity analysis. In conclusion, the results provided novel insights into the potential biomarkers and underlying molecular mechanisms of STAS in lung adenocarcinoma.

## Materials and methods

### Tissue samples

Tumor samples were obtained from March to July 2021 from 19 lung adenocarcinoma patients who underwent surgical resection at Harbin Medical University Cancer Hospital. The inclusion criteria of this study were as follows: 1) histologically confirmed adenocarcinoma; 2) complete clinicopathological information. Exclusion criteria were as follows: 1) preoperative radiotherapy and/or chemotherapy and 2) tumor with other components, including neuroendocrine or squamous differentiation. The clinicopathological characteristics of samples are shown in Table 1. The research was approved by the Ethics Committees of Harbin Medical University Cancer Hospital; all patients in this trial agreed to participate and signed written consent.

**FIGURE 1**

Representative histopathologic images of spread through air spaces (STAS) in lung adenocarcinoma. (A) Magnification  $\times 100$  (B). Magnification  $\times 400$ .

## Data collection

Clinicopathological characteristics were obtained from medical records, such as age, gender, smoking status, tumor size, N stage, TNM stage, EGFR mutation, KRAS mutation, and EGFR mutation.

Two pathologists independently inspected all samples, and disagreements were resolved through discussion and consensus. STAS consists of micropapillary clusters, solid nests, or tumor individual tumor cells at the outside edge of the surrounding alveolar lacuna of lung parenchyma. Notably, micropapillary clusters are defined as papillary structures without central fibrovascular cores in the alveolar space. Additionally, solid nests are defined as solid collections of tumor cells filling air spaces. It is well understood that the tumor cells leave the edge of the lung cancer mass and enter the alveoli and bronchioles of the peripheral pulmonary parenchyma. Representative histopathologic images of STAS in lung adenocarcinoma (arrows) are shown in [Figures 1A,B](#). According to the pathological results, samples were divided into the STAS\_positive and STAS\_negative groups.

## RNA isolation, quality control, library preparation, and sequencing

Total RNA was isolated from paraffin-embedded tissue using the RNeasy Mini kit (Qiagen). The quantity and quality of the extracted total RNA were then determined using a Qubit RNA BR Assay Kit and Agilent 2200 TapeStation. RNA-seq libraries were constructed according to the manufacturer's instructions using the TruSeq Stranded mRNA Prep kit (Illumina). RNA quality control was verified by applying the MultiNA Microchip Electrophoresis System. Total RNA samples were transferred into 96-well plates and diluted to  $1 \mu\text{g}$  of  $50 \mu\text{l}$ . The purification steps during library

preparation were performed using the MinElute PCR Purification kit (Qiagen). After purification and PCR amplification, the final cDNA library was generated according to Illumina's RNA-seq Library Preparation Protocol. We conducted deep sequencing on the Illumina HiSeq 1500 platform using a TruSeq Rapid SBS kit (Illumina) in a 50-base single-end mode.

## Read preprocessing and gene count normalization

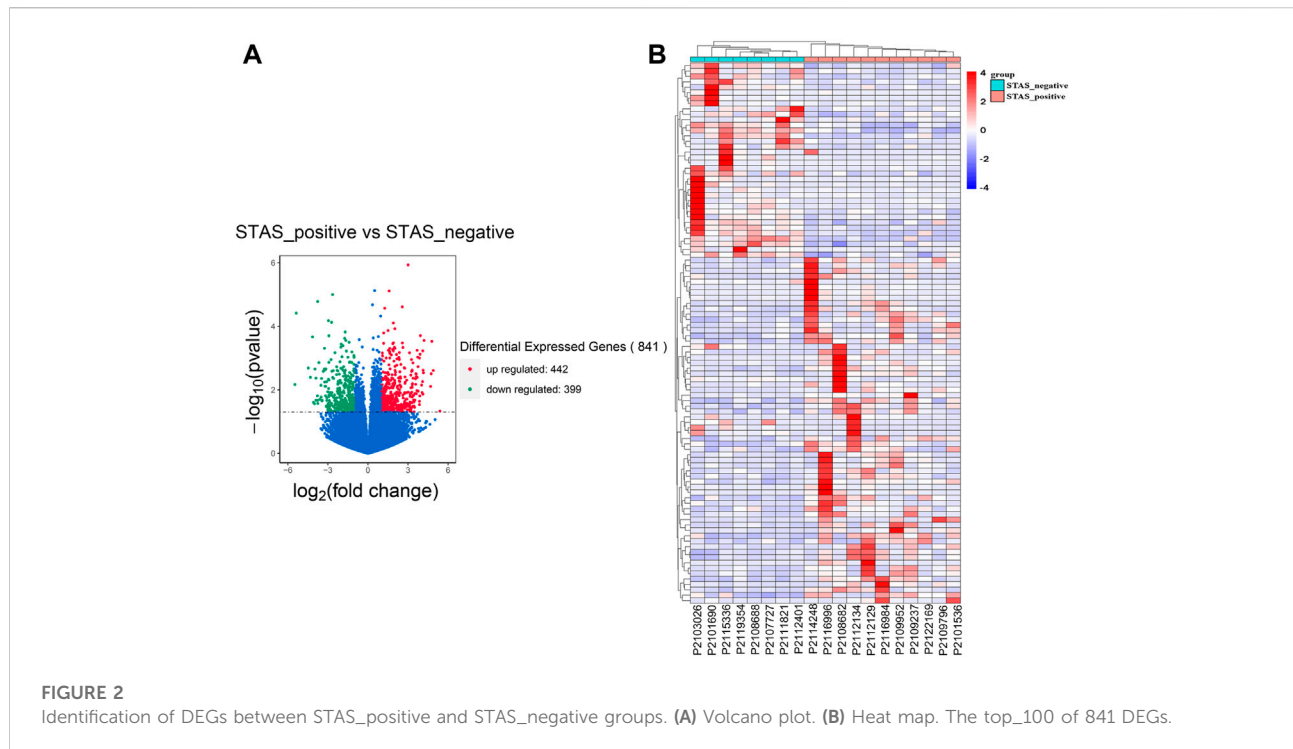
The obtained raw paired-end reads were trimmed using the fastp tool to remove low-quality reads and adapter contaminants. Then, raw read data were aligned to the human reference genome (build 37.2) using the hisat2 (version 2.1.0) tool. DESeq2 and the upper quartile normalization method were used to normalize the raw read counts.

## Identification of differentially expressed genes

FPKM was used to calculate the gene expression level. The DEGs were calculated using the R package DESeq2 ([Zhang et al., 2021](#)).  $p$ . adj  $< 0.05$  and  $|\log_2\text{fold change}| > 2$  were defined as thresholds. The R package ggplot2 was used to construct the Volcano and Heat maps.

## The construction of the protein–protein interaction network

PPI network analysis has emerged as a useful approach to identifying potential new targets and mechanisms from a



systematic perspective (Wu et al., 2020). PPI network was constructed using the STRING database (<https://string-db.org>) (Buttacavoli et al., 2020; Ying et al., 2020). The Cytoscape software was used to analyze the hub genes (Szklarczyk et al., 2017). Log-rank test was used to compare differences in overall survival (OS) and progression-free survival (PFS) between the low- and high-expression groups by median cutoff for each gene. The statistical difference of hub genes between normal and tumor samples was compared through the Wilcox test. All the analysis methods and R package were implemented by R version 4.0.3.

## Gene functional enrichment analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses of DEGs were carried out using the clusterProfiler package (Version 4.0.3) (Xu et al., 2018). In addition, pathway enrichment analysis of DEGs was conducted using the ReactomePA package (Chow et al., 2020).

## Gene set variation analysis

The gene sets of H (hallmark gene sets), C2 (curated gene sets), and C5 (GO gene sets) were downloaded from the MSigDB database (<http://software.broadinstitute.org/gsea/msigdb/index>).

jsp) for gene set variation analysis (GSVA) (Migliavacca et al., 2019; Park et al., 2017). GSVA was used to identify pathways enriched among expressed genes in two groups. The significant difference was set at a  $p$ -value <0.05.

## Analysis of factors related to immune microenvironment

We used single-sample gene set enrichment analysis (ssGSEA) to determine the proportions of 28 types of immune cells in the tumor microenvironment (Thorsson et al., 2018). The CIBERSORT software was applied to evaluate the relative abundance of tumor-infiltrating immune and stromal cells (Conley et al., 2021).

The proportions of 28 types of immune cells, the score of immune activity, and the score of tertiary lymphoid structure (TLS) between the two groups were compared and analyzed using the Wilcoxon ranked-sum test.

## CancerSEA

CancerSEA (<http://biocc.hrbmu.edu.cn/CancerSEA/>) was used to comprehensively investigate the functional status of cancer cells at the single-cell level (Yuan et al., 2019). It could provide information on DEGs in various cancers with multiple functional states.

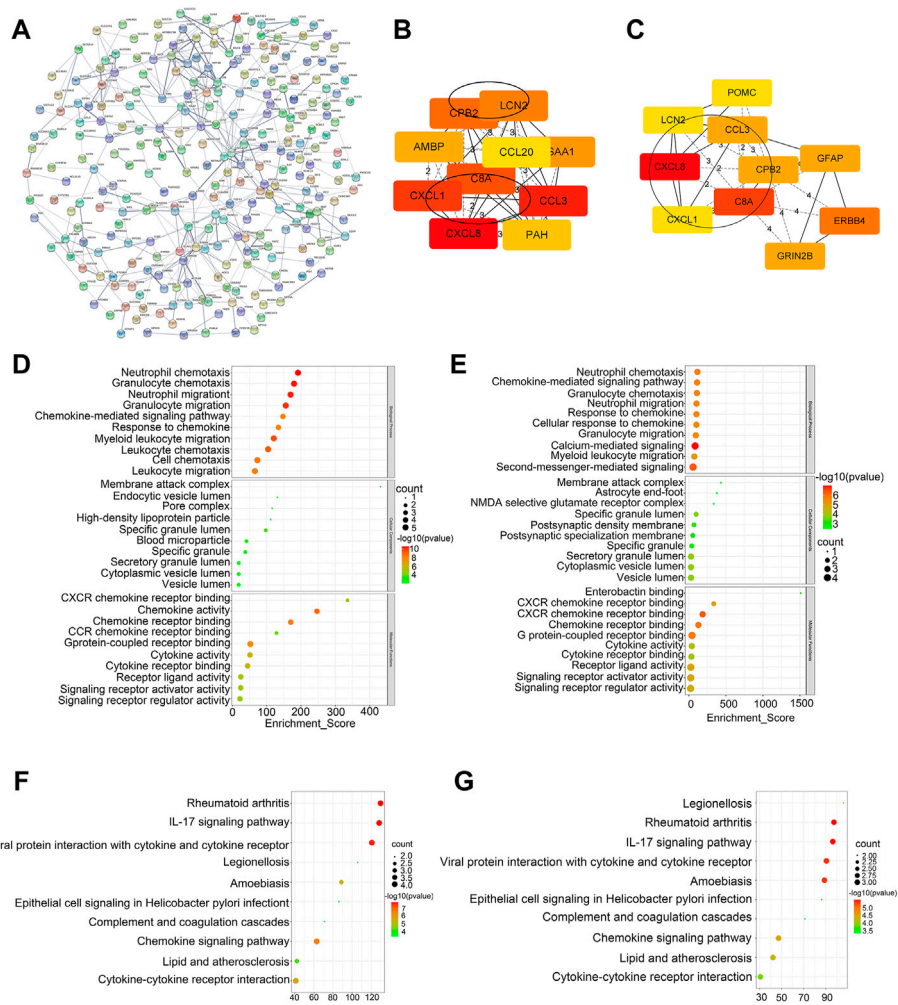


FIGURE 3

Construction of PPI network and screening of hub genes. (A) PPI network of DEGs. (B) Top 10 genes ranked by the MCC method. (C) Top 10 genes ranked by the MNC method. (D–E) GO analysis of top 10 genes ranked by the MCC and MNC method. (F, G) KEGG analysis of top 10 genes ranked by the MCC and MNC method.

## Results

### Identification of DEGs between STAS\_positive and STAS\_negative groups

The dataset was analyzed for DEGs, as shown in Figure 2A. A total of 841 genes were differentially expressed, with 442 upregulated and 399 downregulated. Moreover, the top\_100 DEGs were subjected to cluster analysis. The results showed that the top\_100 DEGs could stratify patients in the STAS\_positive versus STAS\_negative groups (Figure 2B).

### Estimation of DEGs on the protein–protein interaction network

We performed the PPI network analysis to further investigate the interactions among the 841 DEGs (Figure 3A). The results indicated that part of DEGs strongly correlated with other genes. Six hub genes were identified by overlapping the top 10 genes obtained using MCC and MNC ranking methods (Figures 3B,C). The top 10 hub genes ranked by the MCC and MNC methods were functionally annotated through GO terms and KEGG

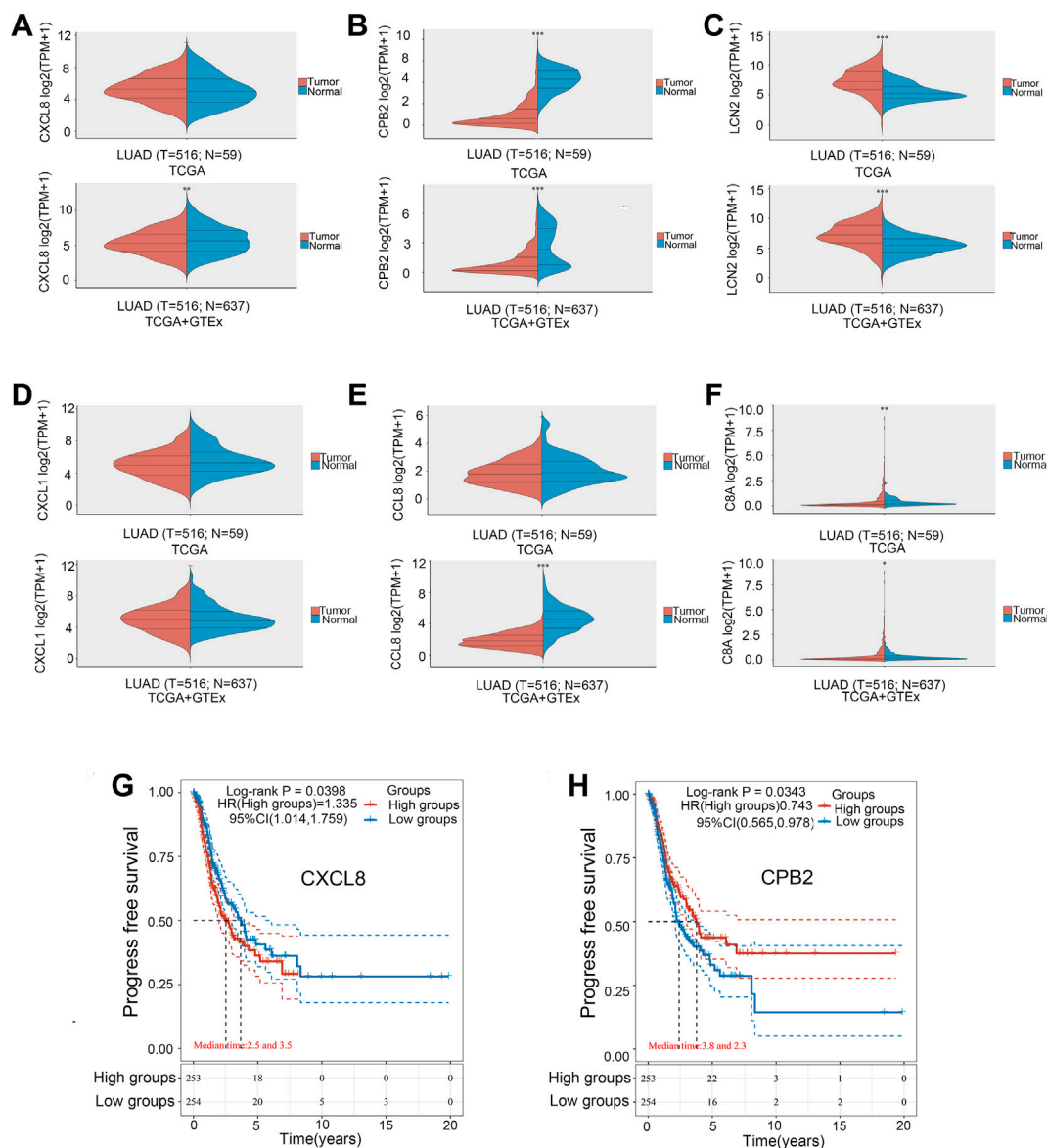


FIGURE 4

Gene expression levels in the tumor versus normal tissue. (A) CXCL8. (B) CPB2. (C) LCN2. (D) CXCL1. (E) CCL3. (F) C8A. Kaplan–Meier survival curve. (G) CXCL8. (H) CPB2. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

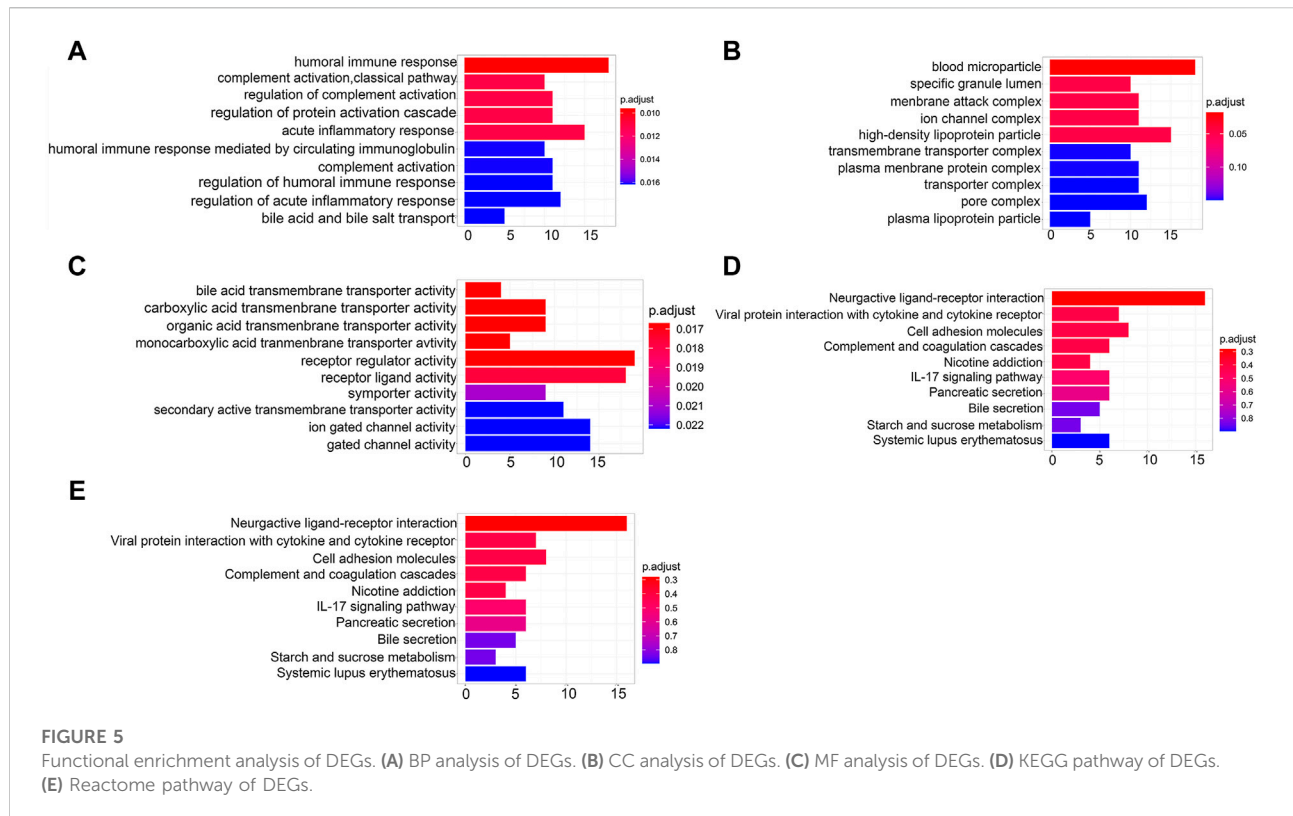
pathways (Figures 3D–G). Based on the GO and KEGG analysis, these hub genes were mainly enriched in cytokine and chemokine-related pathways.

The six hub genes were *CXCL8*, *CXCL1*, *CCL3*, *C8A*, *CPB2*, and *LCN2*. The TCGA and GTE databases revealed that the other five genes significantly differed in gene expression in lung adenocarcinoma tumors versus normal tissues, except for the *CXCL1* gene (Figures 4A–F). Clinical data for lung adenocarcinoma were obtained from TCGA. We then performed the OS and PFS analyses of these six hub genes. Figures 4G,H show that the high-CXCL8 expression group

had significantly worse PFS than the low-CXCL8 expression group ( $p = 0.0398$ ). It is worth noting that patients with high CPB2 expression had significantly better PFS than those in the low-expression group ( $p = 0.0343$ ). However, no significant differences were found in the other four genes.

## Functional enrichment analysis of DEGs

GO analysis covers biological process (BP), cellular component (CC), and molecular function (MF). BP



enrichment revealed that the identified DEGs were involved in humoral immune response, complement activation, and regulation of protein activation cascade (Figure 5A). CC enrichment revealed that DEGs were involved in blood microparticles, specific granule lumen, and membrane attack complexes (Figure 5B). MF enrichment revealed that DEGs were involved in bile acid transmembrane transporter activity, carboxylic acid transmembrane transporter activity, and organic acid transmembrane transporter activity (Figure 5C). Slight enrichment of GO analysis was observed, but the results were not statistically significant ( $p > 0.05$ ).

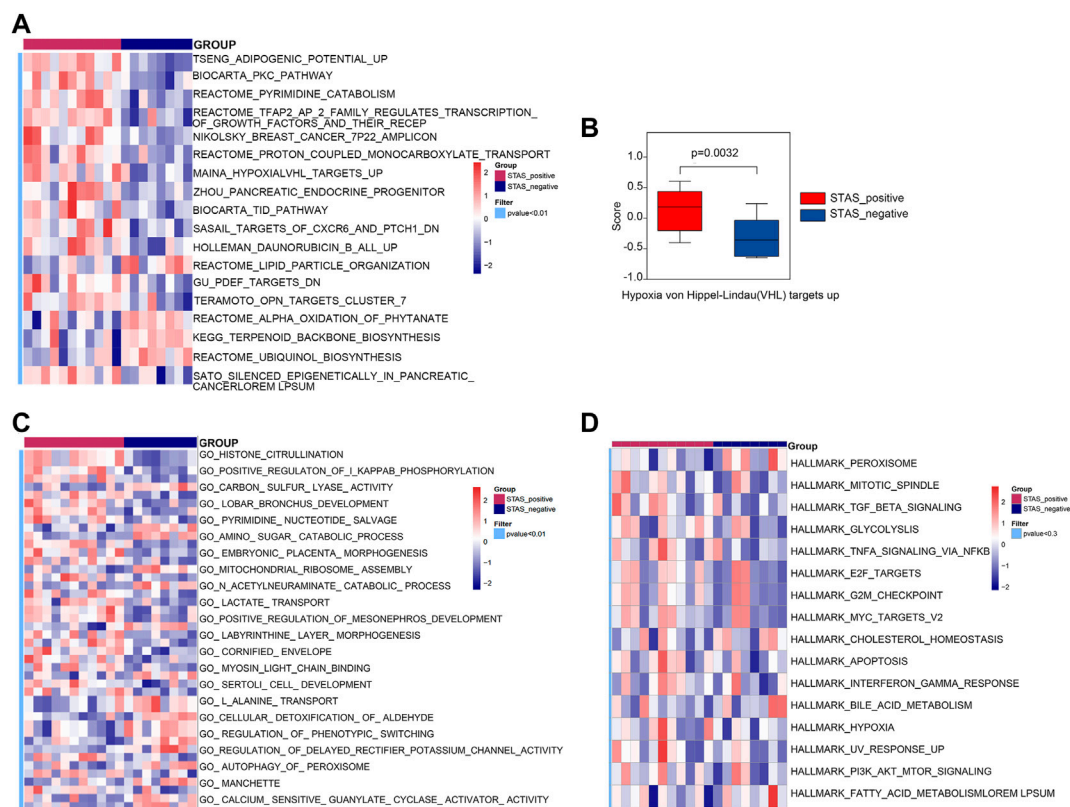
DEGs were also analyzed using the KEGG and Reactome pathway enrichment analyses. KEGG pathway analysis showed that the DEGs were enriched in IL-17 signaling and involved in the interaction of neuroactive ligands with the receptors and cell adhesion molecules (Figure 5D). The Reactome pathway analysis indicated that the DEGs were enriched in the transport of organic anions, bile acids, and bile salt metabolism and recycling of bile acids and bile salts

signaling pathways (Figure 5E). However, the findings had no statistical significance ( $p > 0.05$ ).

## Evaluation of gene set differences between STAS\_positive and STAS\_negative groups

GSVA was performed using the curated gene sets (C2), GO gene sets (C5), and hallmark gene sets (H). The comparative analysis result using the C2 gene set demonstrated that there were 118 significantly different biological processes and signaling pathways between STAS\_positive and STAS\_negative groups. Only those biological processes and pathways with  $p < 0.01$  are shown in Figure 6A. The main differential signaling pathways included hypoxia VHL targets, PKC, and pyrimidine metabolism pathways. Figure 6B shows a boxplot of hypoxia VHL targets.

The analysis result using the C5 gene set indicated that 247 biological processes and signaling pathways



**FIGURE 6** Gene set variation analysis (GSVA) between STAS\_positive and STAS\_negative groups. **(A)** Heat map derived by C2 gene set. **(B)** Box plot of Hypoxia VHL targets up. **(C)** Heat map derived by C5 gene set. **(D)** Heat map derived by Hallmark gene set.

significantly differed between the two groups. **Figure 6C** only considered biological processes and pathways with  $p < 0.01$ . The differential biological processes were mainly involved in histone citrullination, protein arginine deiminase activity, and positive regulation of I-kappaB phosphorylation.

The analysis result using the Hallmark gene set revealed that 50 signaling pathways differed in score values between the two groups, but the differences were not statistically significant ( $p > 0.05$ ). **Figure 6D** shows 16 signaling pathways with  $p < 0.3$ , such as TGF- $\beta$ , MYC\_TARGETS\_V2, IFN- $\gamma$  response, and other signaling pathways.

## Integrative analysis of immune-related factors between STAS\_positive and STAS\_negative groups

The signature score of nine immune features of tumor samples was calculated (**Figure 7A**). We performed a

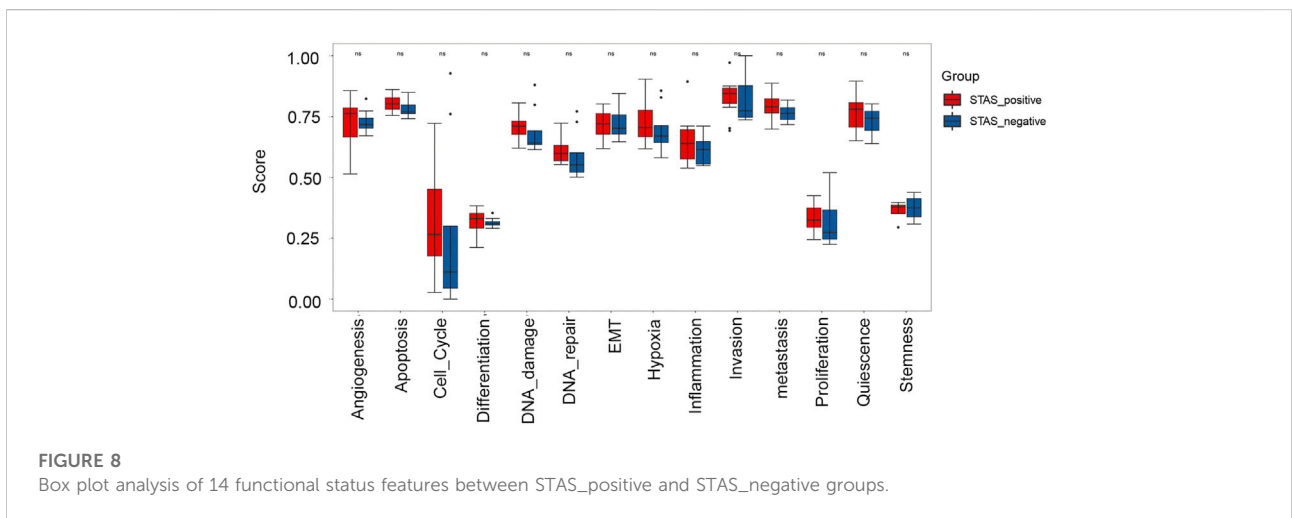
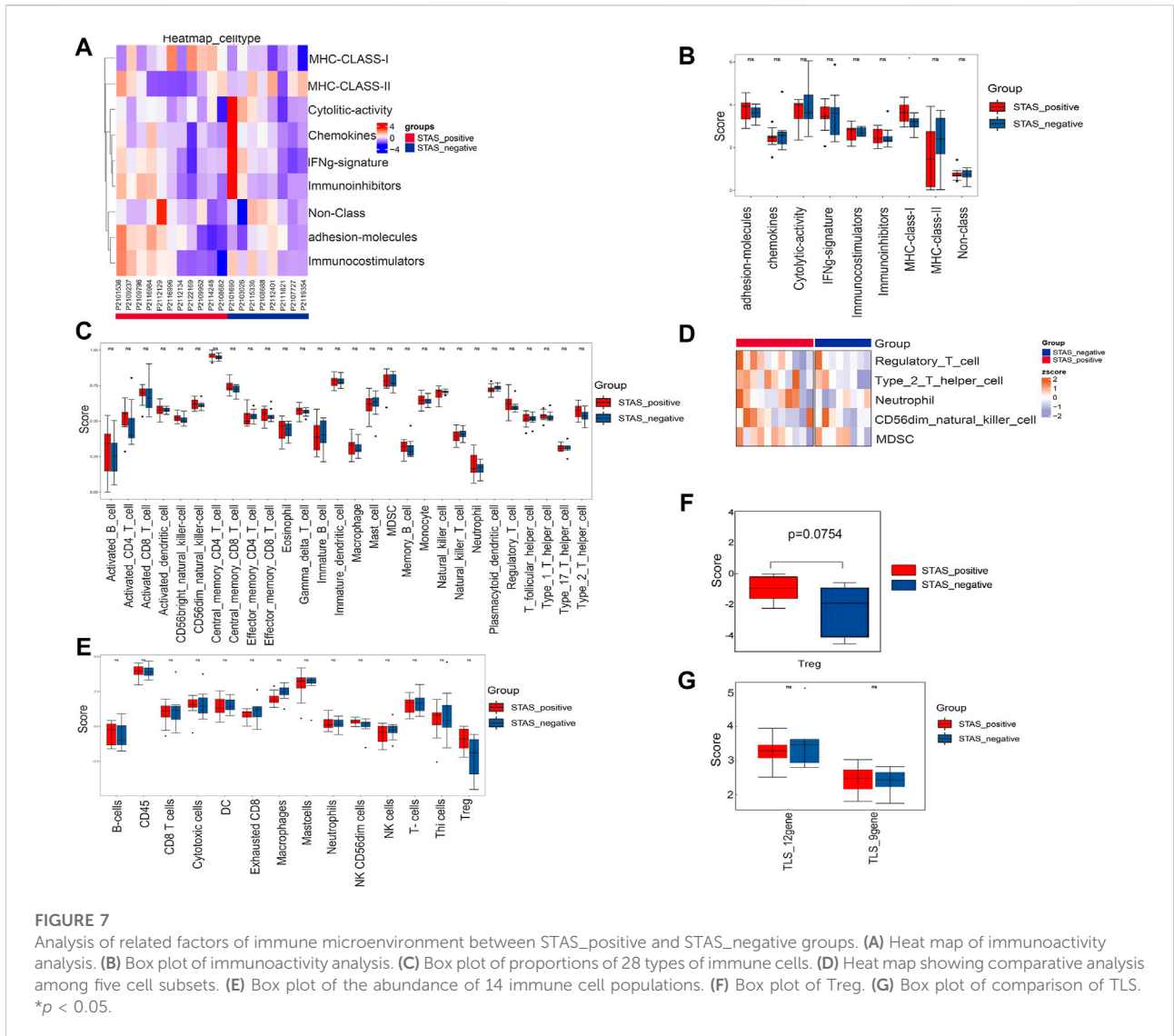
comparative analysis and found that the expression of MHC-Class-I in the STAS\_positive group was significantly higher than that in the STAS\_negative group (**Figure 7B**).

The proportions of 28 types of immune cells in the tumor microenvironment are shown in **Figure 7C**. Although not statistically significant, we did observe more regulatory T cells (Tregs), type 2 T helper cells, neutrophils, CD56<sup>dim</sup> natural killer (NK) cells, and myeloid-derived suppressor cells (MDSC) in the STAS\_positive group (**Figure 7D**).

**Figure 7E** shows the abundance of 14 immune cell populations. Wilcoxon test analysis indicated that the abundance of Tregs tended to increase in the STAS\_positive group ( $p = 0.0754$ ) (**Figure 7F**).

TLS is an aggregate of immune cells (mainly T cells and B cells). The scores of two TLS signatures were calculated according to the TLS signature-related gene expression levels reported in the studies (**Figure 7G**). However, no significant differences were found between the two groups ( $p > 0.05$ ).





## Assessment of functional status features in CancerSEA

Fourteen functional status features collected from the CancerSEA database were analyzed. Results exhibited that metastasis, hypoxia, DNA damage, proliferation, apoptosis, and cell cycle showed a trend to be higher in the STAS\_positive group but failed to reach statistical significance ( $p > 0.05$ ) (Figure 8A).

## Discussion

More than half of stage IA lung adenocarcinoma patients with STAS who receive radical surgery will relapse within 5 years (Onozato et al., 2013). Investigations have demonstrated that STAS occurrence is closely related to clinical outcomes, such as disease recurrence and survival rates. Furthermore, STAS is considered an independent prognostic factor for OS and DFS (Lu et al., 2017; Yanagawa et al., 2018; Aly et al., 2019). However, the current knowledge about markers and molecular pathogenesis of STAS in lung adenocarcinoma is limited. In our study, we performed a comprehensive bioinformatics analysis of gene expression in STAS\_positive and STAS\_negative samples and screened key genes that may be involved in the pathogenesis of STAS. We found that high CXCL8 and low CPB2 expressed in STAS\_positive samples were clinically significant. In addition, high CXCL8 and low CPB2 expression levels were risk factors for unfavorable survival through analysis of TCGA. Consequently, we identified CXCL8 and CPB2 as valuable biomarkers for the diagnosis and prognosis of patients with STAS, and these may be promising therapeutic targets for the treatment of STAS as well.

CXCL8 is a chemokine that severely promotes neutrophilic inflammation through its receptors, CXCR1/CXCR2 (Ha et al., 2017). CXCR1 and CXCR2 are widely expressed on neutrophils, endothelial cells, cancer cells, and tumor-associated macrophages (Gregson et al., 2013). The CXCL8-CXCR1/2 axis plays a crucial part in tumor progression and metastasis by regulating the proliferation and self-renewal of cancer stem cells (Li et al., 2019). CXCL8 is overexpressed in many solid tumors, including lung, esophageal, breast, and colon tumors. Notably, CXCL8 binds CXCR1/2 in the tumor microenvironment, promoting tumor cell proliferation and growth through autocrine and paracrine mechanisms (Xie, 2001; Lin et al., 2004). Additionally, CXCL8 regulates hepatocellular carcinoma (HCC) cell proliferation and migration. The increased expression of CXCL8 in HCC cells accelerates tumor proliferation, migration, and invasion and is strongly correlated with clinical stage and tumor infiltration (Yang et al., 2020).

In addition to its regulatory properties, CXCL8 is a potential biomarker to predict tumor progression and

prognosis in many malignancies. For example, in melanoma and breast cancer, CXCL8 overexpression plays a key role in metastasis and poor patient survival outcomes (Wu et al., 2012; Fang et al., 2017). In addition, higher expression of CXCL8 has been noted in ovarian cancer cell lines with high metastasis compared with the parental cell lines (Milliken et al., 2002). Similarly, CXCL8 promotes tumor growth and metastasis and predicts bad outcomes in colorectal cancer (Xiao et al., 2015). Therefore, these studies demonstrate an important role for CXCL8 in tumorigenesis and metastasis. In the present study, we found that high CXCL8 expression could promote the occurrence of STAS, which is consistent with previous studies noted above.

CPB2, also known as thrombin activatable fibrinolysis inhibitor (TAFI), plays a central role in coagulation and fibrinolysis (Mosnier and Bouma, 2006). It is understood that coagulation and fibrinolytic systems are correlated with physiologic and pathological processes such as tumor growth and invasion (Higuchi et al., 2009). Additionally, plasma CPB2 levels are positively associated with many human diseases (Boffa and Koschinsky, 2007). Similarly, increased expression of CPB2 has been reported in breast, ovarian, lung, gastric, and hepatic cancer cells (Yeung et al., 2013; Miller et al., 2011; Hataji et al., 2004; Fidan et al., 2012). Furthermore, tumors with high CPB2 expression have been related to more advanced tumor stages (Balcik et al., 2011). Conversely, it has been reported that downregulation of CPB2 expression by siRNA reduces breast cancer cell proliferation, migration, and invasion (Yu et al., 2017). Contrary to the findings of the above studies, our results showed that low expression of CPB2 may facilitate the appearance of STAS and be relevant to worse PFS in adenocarcinoma. At the same time, previous research in cell line models has demonstrated that CPB2 results in the suppression of breast cancer cell invasion and migration, which is consistent with our findings (Bazzi et al., 2016). The reason for this difference may be that there are fewer articles about the CPB2 gene. Therefore, the CPB2 gene is worth exploring further to shed light on underlying mechanisms in cancer cells.

There are some noted limitations in our study. Firstly, a major limitation is the heterogeneity of the analyzed samples, so all significant results should be interpreted with caution. Secondly, multiple lines of evidence suggest that these predictions based on epigenetic profiling are largely accurate (Sharp et al., 2011). However, these computer-based predictions do not obligatorily reflect the actual gene pool and need to be verified experimentally in future work.

In conclusion, we revealed the potential markers and underlying molecular mechanisms of STAS in lung adenocarcinoma through a systematic and comprehensive analysis of the high-throughput sequencing RNA-seq dataset.

Our results provided promising clues and laid the groundwork for developing new effective clinical therapies for STAS. However, further study is needed to better understand STAS and develop new treatments by integrating more data.

## Data availability statement

The original contributions presented in the study are publicly available. These data can be found here: GSA database: HRA002248.

## Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committees of The Affiliated Tumour Hospital, Harbin Medical University. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

YY, XM, and BP conceived and designed the experiments. YZ, LL, and DX analyzed the studies and extracted data. All authors participated in the writing, reading, and revising of the manuscript and approved the final version of the manuscript.

## References

- Aly, R. G., Rektman, N., Li, X., Takahashi, Y., Eguchi, T., Tan, K. S., et al. (2019). Spread through Air Spaces (STAS) is prognostic in atypical carcinoid, large cell neuroendocrine carcinoma, and small cell carcinoma of the lung. *J. Thorac. Oncol.* 14 (9), 1583–1593. doi:10.1016/j.jtho.2019.05.009
- Balcik, O. S., Albayrak, M., Uyar, M. E., Dagdas, S., Yokus, O., Ceran, F., et al. (2011). Serum thrombin activatable fibrinolysis inhibitor levels in patients with newly diagnosed multiple myeloma. *Blood Coagul. Fibrinolysis* 22 (4), 260–263. doi:10.1097/MBC.0b013e3283442cf9
- Bazzi, Z. A., Lanoue, D., El-Youssef, M., Romagnuolo, R., Tubman, J., Cavallo-Medved, D., et al. (2016). Activated thrombin-activatable fibrinolysis inhibitor (TAFI) attenuates breast cancer cell metastatic behaviors through inhibition of plasminogen activation and extracellular proteolysis. *BMC Cancer* 16, 328. doi:10.1186/s12885-016-2359-1
- Boffa, M. B., and Koschinsky, M. L. (2007). Curiouser and curiouser: Recent advances in measurement of thrombin-activatable fibrinolysis inhibitor (TAFI) and in understanding its molecular genetics, gene regulation, and biological roles. *Clin. Biochem.* 40 (7), 431–442. doi:10.1016/j.clinbiochem.2006.10.020
- Buttacavoli, M., Albanese, N. N., Roz, E., Pucci-Minafra, I., Feo, S., and Cancemi, P. (2020). Proteomic profiling of colon cancer tissues: Discovery of new candidate biomarkers. *Int. J. Mol. Sci.* 21 (9), 3096. doi:10.3390/ijms21093096
- Chow, J., Hoffend, N. C., Abrams, S. I., Schwaab, T., Singh, A. K., and Muhitch, J. B. (2020). Radiation induces dynamic changes to the T cell repertoire in renal cell carcinoma patients. *Proc. Natl. Acad. Sci. U. S. A.* 117 (38), 23721–23729. doi:10.1073/pnas.2001933117
- Conley, B. A., Staudt, L., Takebe, N., Wheeler, D. A., Wang, L., Cardenas, M. F., et al. (2021). The exceptional responders initiative: Feasibility of a national cancer institute pilot study. *J. Natl. Cancer Inst.* 113 (1), 27–37. doi:10.1093/jnci/djaa061
- Eguchi, T., Kameda, K., Lu, S., Bott, M. J., Tan, K. S., Montecalvo, J., et al. (2019). Lobectomy is associated with better outcomes than sublobar resection

## Acknowledgments

Thanks to Genecast biotechnology company (Jiangsu, China) for helping with the RNA sequencing.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.900864/full#supplementary-material>

in spread through air spaces (STAS)-Positive T1 lung adenocarcinoma: A propensity score-matched analysis. *J. Thorac. Oncol.* 14 (1), 87–98. doi:10.1016/j.jtho.2018.09.005

Fang, Q. I., Wang, X., Luo, G., Yu, M., Zhang, X., and Xu, N. (2017). Increased CXCL8 expression is negatively correlated with the overall survival of patients with ER-negative breast cancer. *Anticancer Res.* 37 (9), 4845–4852. doi:10.21873/anticancer.11892

Fidan, E., Kavgaci, H., Orem, A., Yilmaz, M., Yildiz, B., Fidan, S., et al. (2012). Thrombin activatable fibrinolysis inhibitor and thrombin-antithrombin -III - complex levels in patients with gastric cancer. *Tumour Biol.* 33 (5), 1519–1525. doi:10.1007/s13277-012-0403-6

Gerstner, N., Kehl, T., Lenhof, K., Müller, A., Mayer, C., Eckhart, L., et al. (2020). GeneTrail 3: Advanced high-throughput enrichment analysis. *Nucleic Acids Res.* 48 (W1), W515–W520. doi:10.1093/nar/gkaa306

Gregson, A. L., Wang, X., Weigt, S. S., Palchevskiy, V., Lynch, J. P., 3rd, Ross, D. J., et al. (2013). Interaction between *Pseudomonas* and CXC chemokines increases risk of bronchiolitis obliterans syndrome and death in lung transplantation. *Am. J. Respir. Crit. Care Med.* 187 (5), 518–526. doi:10.1164/rccm.201207-1228OC

Ha, H., Debnath, B., and Neamati, N. (2017). Role of the CXCL8-CXCR1/2 Axis in cancer and inflammatory diseases. *Theranostics* 7 (6), 1543–1588. doi:10.7150/thno.15625

Han, J., Wu, C., Wu, Y., Deng, H., Gao, J., Han, H., et al. (2021). Comparative study of imaging and pathological evaluation of pneumonic mucinous adenocarcinoma. *Oncol. Lett.* 21 (2), 125. doi:10.3892/ol.2020.12386

Hataji, O., Taguchi, O., Gabazza, E. C., Yuda, H., D'Alessandro-Gabazza, C. N., Fujimoto, H., et al. (2004). Increased circulating levels of thrombin-activatable fibrinolysis inhibitor in lung cancer patients. *Am. J. Hematol.* 76 (3), 214–219. doi:10.1002/ajh.20079

- Hess, J. F., Kohl, T. A., Kotrová, M., Rönsch, K., Paprotka, T., Mohr, V., et al. (2020). Library preparation for next generation sequencing: A review of automation strategies. *Biotechnol. Adv.* 41, 107537. doi:10.1016/j.biotechadv.2020.107537
- Higuchi, T., Nakamura, T., Kakutani, H., and Ishii, H. (2009). Thrombomodulin suppresses invasiveness of HT1080 tumor cells by reducing plasminogen activation on the cell surface through activation of thrombin-activatable fibrinolysis inhibitor. *Biol. Pharm. Bull.* 32 (2), 179–185. doi:10.1248/bpb.32.179
- Kadota, K., Nitadori, J. I., Sima, C. S., Ujiie, H., Rizk, N. P., Jones, D. R., et al. (2015). Tumor spread through air spaces is an important pattern of invasion and impacts the frequency and location of recurrences after limited resection for small stage I lung adenocarcinomas. *J. Thorac. Oncol.* 10 (5), 806–814. doi:10.1097/JTO.0000000000000486
- Li, Y., He, Y., Butler, W., Xu, L., Chang, Y., Lei, K., et al. (2019). Targeting cellular heterogeneity with CXCR2 blockade for the treatment of therapy-resistant prostate cancer. *Sci. Transl. Med.* 11 (521), eaax0428. doi:10.1126/scitranslmed.aax0428
- Lin, Y., Huang, R., Chen, L., Li, S., Shi, Q., Jordan, C., et al. (2004). Identification of interleukin-8 as estrogen receptor-regulated factor involved in breast cancer invasion and angiogenesis by protein arrays. *Int. J. Cancer* 109 (4), 507–515. doi:10.1002/ijc.11724
- Liu, A., Sun, X., Xu, J., Xuan, Y., Zhao, Y., Qiu, T., et al. (2020). Relevance and prognostic ability of twist, slug and tumor spread through air spaces in lung adenocarcinoma. *Cancer Med.* 9 (6), 1986–1998. doi:10.1002/cam4.2858
- Lu, S., Tan, K. S., Kadota, K., Eguchi, T., Bains, S., Rekhman, N., et al. (2017). Spread through Air Spaces (STAS) is an independent predictor of recurrence and lung can-cer-specific death in squamous cell carcinoma. *J. Thorac. Oncol.* 12 (2), 223–234. doi:10.1016/j.jtho.2016.09.129
- Migliavacca, E., Tay, S. K. H., Patel, H. P., Sonntag, T., Civiletto, G., McFarlane, C., et al. (2019). Mitochondrial oxidative capacity and NAD<sup>+</sup> biosynthesis are reduced in human sarcopenia across ethnicities. *Nat. Commun.* 10 (1), 5808. doi:10.1038/s41467-019-13694-1
- Miller, T. W., Balko, J. M., Fox, E. M., Ghazoui, Z., Dunbier, A., Anderson, H., et al. (2011). ER $\alpha$ -dependent E2F transcription can mediate resistance to estrogen deprivation in human breast cancer. *Cancer Discov.* 1 (4), 338–351. doi:10.1158/2159-8290.CD-11-0101
- Milliken, D., Scotton, C., Raju, S., Balkwill, F., and Wilson, J. (2002). Analysis of chemokines and chemokine receptor expression in ovarian cancer ascites. *Clin. Cancer Res.* 8 (4), 1108–1114.
- Mosnier, L. O., and Bouma, B. N. (2006). Regulation of fibrinolysis by thrombin activatable fibrinolysis inhibitor, an unstable carboxypeptidase B that unites the pathways of coagulation and fibrinolysis. *Arterioscler. Thromb. Vasc. Biol.* 26 (11), 2445–2453. doi:10.1161/01.ATV.0000244680.14653.9a
- Onozato, M. L., Kovach, A. E., Yeap, B. Y., Morales-Oyarvide, V., Klepeis, V. E., Tammi, S., et al. (2013). Tumor islands in resected early-stage lung adenocarcinomas are associated with unique clinicopathologic and molecular characteristics and worse prognosis. *Am. J. Surg. Pathol.* 37 (2), 287–294. doi:10.1097/PAS.0b013e31826885fb
- Onozato, Y., Nakajima, T., Yokota, H., Morimoto, J., Nishiyama, A., Toyoda, T., et al. (2021). Radiomics is feasible for prediction of spread through air spaces in patients with non small cell lung cancer. *Sci. Rep.* 11 (1), 13526. doi:10.1038/s41598-021-93002-4
- Park, J., Yang, J., Wenzel, A. T., Ramachandran, A., Lee, W. J., Daniels, J. C., et al. (2017). Genomic analysis of 220 CTCLs identifies a novel recurrent gain-of-function alteration in RLTPR (p.Q575E). *Blood* 130 (12), 1430–1440. doi:10.1182/blood-2017-02-768234
- Qi, L., Xue, K., Cai, Y., Lu, J., Li, X., and Li, M. (2021). Predictors of CT morphologic features to identify spread through air spaces preoperatively in small-sized lung adenocarcinoma. *Front. Oncol.* 10, 548430. doi:10.3389/fonc.2020.548430
- Sharp, A. J., Stathaki, E., Migliavacca, E., Brahmachary, M., Montgomery, S. B., Dupre, Y., et al. (2011). DNA methylation profiles of human active and inactive X chromosomes. *Genome Res.* 21 (10), 1592–1600. doi:10.1101/gr.112680.110
- Shi, J., Hua, X., Zhu, B., Ravichandran, S., Wang, M., Nguyen, C., et al. (2016). Somatic genomics and clinical features of lung adenocarcinoma: A retrospective study. *PLoS Med.* 13 (12), e1002162. doi:10.1371/journal.pmed.1002162
- Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., et al. (2017). The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 45 (D1), D362–D368. doi:10.1093/nar/gkw937
- Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The immune landscape of cancer. *Immunity* 48 (4), 411–412. e14. doi:10.1016/j.immuni.2019.08.004
- Travis, W. D., Brambilla, E., Nicholson, A. G., Yatabe, Y., Austin, J. H. M., Beasley, M. B., et al. (2015). The 2015 World Health organization classification of lung tumors: Impact of genetic, clinical and radiologic advances since the 2004 classification. *J. Thorac. Oncol.* 10 (9), 1243–1260. doi:10.1097/JTO.0000000000000630
- Wislez, M., Antoine, M., Rabbe, N., Gounant, V., Poulot, V., Lavolé, A., et al. (2007). Neutrophils promote arogenous spread of lung adenocarcinoma with bronchioloalveolar carcinoma features. *Clin. Cancer Res.* 13 (12), 3518–3527. doi:10.1158/1078-0432.CCR-06-2558
- Wislez, M., Philippe, C., Antoine, M., Rabbe, N., Moreau, J., Bellocq, A., et al. (2004). Upregulation of bronchioloalveolar carcinoma-derived C-X-C chemokines by tumor infiltrating inflammatory cells. *Inflamm. Res.* 53 (1), 4–12. doi:10.1007/s00011-003-1215-3
- Wu, H. T., Chen, W. T., Li, G. W., Shen, J. X., Ye, Q. Q., Zhang, M. L., et al. (2020). Analysis of the differentially expressed genes induced by cisplatin resistance in oral squamous cell carcinomas and their interaction. *Front. Genet.* 10, 1328. doi:10.3389/fgene.2019.01328
- Wu, S., Singh, S., Varney, M. L., Kindle, S., and Singh, R. K. (2012). Modulation of CXCL8 expression in human melanoma cells regulates tumor growth, angiogenesis, invasion, and metastasis. *Cancer Med.* 1 (3), 306–317. doi:10.1002/cam4.28
- Xiao, Y. C., Yang, Z. B., Cheng, X. S., Fang, X. B., Shen, T., Xia, C. F., et al. (2015). CXCL8, overexpressed in colorectal cancer, enhances the resistance of colorectal cancer cells to anoikis. *Cancer Lett.* 361 (1), 22–32. doi:10.1016/j.canlet.2015.02.021
- Xie, K. (2001). Interleukin-8 and human cancer biology. *Cytokine Growth Factor Rev.* 12 (4), 375–391. doi:10.1016/s1359-6101(01)00016-8
- Xu, P., Yang, J., Liu, J., Yang, X., Liao, J., Yuan, F., et al. (2018). Identification of glioblastoma gene prognosis modules based on weighted gene co-expression network analysis. *BMC Med. Genomics* 11 (1), 96. doi:10.1186/s12920-018-0407-1
- Yanagawa, N., Shiono, S., Endo, M., and Ogata, S. Y. (2018). Tumor spread through air spaces is a useful predictor of recurrence and prognosis in stage I lung squamous cell carcinoma, but not in stage II and III. *Lung Cancer* 120, 14–21. doi:10.1016/j.lungcan.2018.03.018
- Yang, L., Yang, Y., Ma, P., Zheng, B., Liu, W., Zhang, Z., et al. (2018). Spread through air spaces predicts a worse survival in patients with stage I adenocarcinomas >2 cm after radical lobectomy. *J. Thorac. Dis.* 10 (9), 5308–5317. doi:10.21037/jtd.2018.09.22
- Yang, S., Wang, H., Qin, C., Sun, H., and Han, Y. (2020). Up-regulation of CXCL8 expression is associated with a poor prognosis and enhances tumor cell malignant behaviors in liver cancer. *Biosci. Rep.* 40 (8), BSR20201169. doi:10.1042/BSR20201169
- Yeung, T. L., Leung, C. S., Wong, K. K., Samimi, G., Thompson, M. S., Liu, J., et al. (2013). TGF- $\beta$  modulates ovarian cancer invasion by upregulating CAF- derived versican in the tumor microenvironment. *Cancer Res.* 73 (16), 5016–5028. doi:10.1158/0008-5472.CAN-13-0023
- Ying, X., Jin, X., Wang, P., He, Y., Zhang, H., Ren, X., et al. (2020). Integrative analysis for elucidating transcriptomics Landscapes of glucocorticoid-induced osteoporosis. *Front. Cell. Dev. Biol.* 8, 252. doi:10.3389/fcell.2020.00252
- Yu, C., Luan, Y., Wang, Z., Zhao, J., and Xu, C. (2017). Suppression of TAFI by siRNA inhibits invasion and migration of breast cancer cells. *Mol. Med. Rep.* 16 (3), 3469–3474. doi:10.3892/mmr.2017.7031
- Yuan, H., Yan, M., Zhang, G., Liu, W., Deng, C., Liao, G., et al. (2019). CancerSEA: A cancer single-cell state atlas. *Nucleic Acids Res.* 47 (D1), D900–D908. doi:10.1093/nar/gky939
- Zhang, H., Wang, Y., Feng, J., Wang, S., Wang, Y., Kong, W., et al. (2021). Integrative analysis for elucidating transcriptomics Landscapes of systemic lupus erythematosus. *Front. Genet.* 12, 782005. doi:10.3389/fgene.2021.782005