PSAT1 is a promising tumor microenvironment-related prognostic biomarker in non-small cell lung cancer

Hongmu Li
Sun Yat-sen University Cancer Center
Huikai Miao
Sun Yat-sen University Cancer Center
Mingyue Zeng
Sun Yat-sen University Cancer Center
Youfang Chen
Sun Yat-sen University Cancer Center
Minglei Yang
Sun Yat-sen University
Zhesheng Wen (wenzhsh@sysucc.org.cn)
Sun Yat-sen University Cancer Center

Research Article

Keywords: non-small-cell lung cancer, tumor microenvironment, ESTIMATE algorithm, PSAT1, pan-cancer analysis

Posted Date: May 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1641195/v1

Abstract

Background This study aimed to identify tumor microenvironment-related prognostic genes in non-small cell lung cancer (NSCLC).

Methods Transcriptome profiles and clinical data of NSCLC samples were collected from Gene Expression Omnibus (GEO) Datasets. The ESTIMATE algorithm was used to calculate stromal and immune scores. Heatmap plots were applied for screening differentially expressed genes (DEGs). Venn diagrams were used to assess the intersection of gene groups. DEGs associated with poor prognosis were selected for subsequent analysis. TCGA and Kaplan-Meier Plotter database were used to validate the results of the above analyses. These findings were experimentally validated in NSCLC tissues using quantitative reverse transcription PCR, western blotting, and immunohistochemical staining. Gene Expression Profiling Interactive Analysis 2 and STRING database were used to investigate co-expression and functional networks associated with hub genes. TIMER database was used to assess the correlation between hub genes and the six main types of tumor-infiltrating immune cells. Finally, pan-cancer analysis of hub genes was conducted to investigate the clinical prognosis of diverse cancers.

Results Overall survival analysis of the GSE37745 dataset identified PSAT1 as a risk gene. Kaplan–Meier Plotter data also showed that high expression of PSAT1 results in shorter survival time. Experimental validation confirmed the above results. In addition, the expression level of PSAT1 was correlated with tumor purity and immune cell infiltration in NSCLC. The pan-cancer analysis of PSAT1 offers a relatively comprehensive understanding of the oncogenic roles in different tumors.

Conclusion Our results indicate that PSAT1 is a promising tumor microenvironment-related prognostic biomarker for NSCLC.

Introduction

Lung cancer, the most common cause of cancer death worldwide, leads to 1.6 million deaths each year (Torre, Bray et al. 2015). Approximately 85% of these deaths are caused by non-small-cell lung cancer (NSCLC), the most common subtypes of which are lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) (Molina, Yang et al. 2008). Important advancements in the treatment of NSCLC have been achieved in the past decades. However, the survival rates for NSCLC remain low (Herbst, Morgensztern et al. 2018). Investigating the genomic and host factors that drive the progression of pre-invasive lesions may help to develop better screening strategies and improve patient prognosis.

Nowadays, tumor microenvironment (TME) is considered important in inhibiting tumor cell apoptosis and promoting immune escape, cell proliferation, angiogenesis, invasion, and metastasis (Whiteside 2008). TME comprises a complex population of various cells. Immune cells and stromal cells are the main non-tumor cells in the TME (Hanahan and Coussens 2012, Kim, Gao et al. 2019); accumulated evidence indicates that they may become biomarkers for tumor treatment and prognosis (Junttila and de Sauvage 2013, Chen, Zhang et al. 2017, Kortlever, Sodir et al. 2017). The ESTIMATE algorithm first described by

Yoshihara et al. (Yoshihara, Shahmoradgoli et al. 2013) has helped to estimate the proportions of stromal and immune cells in tumors to predict the associations between these cells and clinical features and outcomes (Yang, Liu et al. 2020, Huang, Zhou et al. 2021). Previous studies have reported that scores calculated by the ESTIMATE algorithm can be utilized as a predictor of nontumor cell infiltration in the TME in breast cancer (Wang, Zhu et al. 2020), bladder cancer (Whitehouse 1989), and gastric cancer (Zeng, Li et al. 2019).

In this study, the ESTIMATE algorithm was used to identify prognostic genes associated with the TME in NSCLC. The expression patterns and prognostic values of differentially expressed genes (DEGs) were validated using the Cancer Genome Atlas (TCGA) and Kaplan-Meier Plotter analysis, and then we identified hub genes. Finally, to investigate the potential molecular mechanisms of target genes in the pathogenesis or clinical prognosis of different cancers, a pan-cancer analysis of target genes was conducted.

Methods

Patients

The gene expression profiles and corresponding clinical information of patients with NSCLC were extracted from Gene Expression Omnibus (GEO) datasets (https://www.ncbi.nlm.nih.gov/). The inclusion criteria of NSCLC samples were as follows: (i) gene expression profiling of NSCLC was available in the dataset; (ii) complete clinical data of patients with NSCLC were available, including gender, age, tumor–node–metastasis stage, and survival time; (iii) the dataset sample size was ≥ 150. Finally, the GSE37745 dataset including 196 patients with NSCLC was included in the study.

Estimation of stromal and immune scores

The stromal and immune scores in the GEO dataset were calculated using the ESTIMATE algorithm for each NSCLC sample. The correlations between clinical features and stromal/immune scores were also analyzed. The prognostic values of stromal and immune scores were estimated using the Kaplan–Meier method and log-rank test.

Screening of prognostic genes among differentially expressed genes

The samples in the GSE37745 dataset were divided into two groups according to stromal scores (high and low) and immune scores (high and low), respectively. DEGs were identified using the “limma” package of R software. Genes that satisfied the following criteria were identified as DEGs between high-score and low-score groups: fold change > 2 and false discovery rate (FDR) < 0.05. To further explore the common characteristics, overlapping DEGs were identified using Venn diagrams. Genes associated with poor prognosis were identified from the overlapping DEGs by analyzing the overall survival (OS) data using Kaplan–Meier plots and log-rank test; $P < 0.05$ was considered statistically significant.

Validation of poor prognosis genes and selected hub genes

To ensure robust results, poor prognosis genes were validated in multiple independent cohorts. The Cancer Genome Atlas (Cancer Genome Atlas Research, Weinstein et al. 2013) and Genotype-Tissue Expression (GTEx) (Consortium 2013) dataset were used to verify the expression of prognosis risk genes using Gene Expression Profiling Interactive Analysis 2 (GEPIA2, https://gepia2.cancer-pku.cn/)(Tang, Kang et al. 2019). Kaplan–Meier Plotter analysis (http://kmplot.com/analysis)(Nagy, Munkacsy et al. 2021) was used to assess the effect of the target gene on the prognosis of patients with NSCLC.

Quantitative real-time PCR (qRT-PCR)

The expression profiles were examined by quantitative *qRT-PCR*. This study was approved by the Sun Yat-sen University Cancer Center ethics committee (YB2018-85). The samples were derived from patients who had undergone lung radical surgery at our center. Total RNA was isolated from ESCC tissue samples and the adjacent normal tissue samples using the TRIzol (TIANGEN, Beijing, China) and reverse-transcribed to complementary DNA using PrimeScript™ RT Master Mix (ES Science, Shanghai, China). Three hub gene was amplified using qRT-PCR by SYBR Green Master Mix (ES Science, Shanghai, China). For each sample, the qRT-PCR assays were conducted in triplicate in 10-mL reaction volumes. *GAPDH* served as the internal control to normalize hub gene expression. The relative expression was calculated using the comparative threshold cycle (2-Ct) method.

Western Blot (WB)

Protein was extracted using RIPA lysis buffer (Beyotime) from NSCLC tissues and normal tissues in our center. Protein concentrations were determined using BCA Protein Assay Kit (Beyotime). Primary antibodies against PSAT1 (1:1000, Proteintech) was incubated overnight at 4°C. The appropriate HRP-conjugated secondary antibodies (goat anti-mouse, 1:5000, Proteinrtech) were used and incubated at room temperature for 1 h followed exposure imaging.

Immunohistochemistry (IHC) staining

Human pathological slides were constructed with formalin-fixed paraffin-embedded NSCLC tissues and normal tissues. IHC staining was performed on the TMA by using the appropriate dilatation of primary antibodies (PSAT1, 1:500, Proteintech) followed by incubation with a secondary antibody conjugated with HRP (mouse anti-rabbit, 1:50, Proteintech). Finally, we evaluated hub genes' levels in a cohort of 99 NSCLC patients with median 5-y follow-up. Annotation parameters include an evaluation of (i) staining intensity (0, 1, 2, 3, 4), (ii) fraction of stained cells (0-100%), the final score is the product of staining intensity and staining area.

STRING and GEPIA2 database analysis

To better understand the biological significance of hub genes, the co-expression genes were identified from the STRING database (STRING, www.string-db.org/). Then, the Gene Ontology (GO) and Kyoto

Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of co-expression genes were performed using the “clusterProfiler”, “org.Hs.eg.db” and “enrichplot” packages of R software. GEPIA2 was used to plot survival heatmaps of the top co-expressed genes in NSCLC.

Transcription factor–miRNA–target regulatory network construction

Numerous studies have revealed that transcription factors (TFs) regulate gene expression by interacting with microRNAs (miRNAs) (Chandra Mangalhara, Manvati et al. 2017). Understanding the cross-talk between these two regulators and their targets is critical to unveiling the complex molecular regulatory mechanisms of cancers. In this study, to explore the potential molecular mechanisms of target genes in oncogenesis and progression of NSCLC, iRegulon plug-in v1.3 of Cytoscape software were used to construct the TF–miRNA–target network (Janky, Verfaillie et al. 2014).

Correlation between target genes and six types of infiltrating immune cells

Tumor-infiltrating immune cells, as prominent components of the TME, are closely associated with initiation, progression, or metastasis of cancer (Fridman, Galon et al. 2011). TIMER (https://cistrome.shinyapps.io/timer/) was used to analyze the correlation between target gene expression and the abundance of immune infiltrates, including B cells, $CD4^+$ T cells, $CD8^+$ T cells, neutrophils, macrophages, and dendritic cells via gene modules (Li, Fan et al. 2017).

Pan-cancer analysis of target genes

The potential molecular mechanisms of target genes in different cancers were investigated using GEPIA2 and Kaplan–Meier Plotter analysis to conduct a pan-cancer analysis of target genes, including gene expression analysis and survival status.

Statistical analysis

Gene expression data from the TCGA and GTEx databases were analyzed using Student’s t-test. The correlation analysis was performed in the TIMER database using Spearman’s correlation. Data processing and plotting in this study was performed by Perl (Version Strawberry-Perl-5.32.0.1; https://www.perl.org/), Cytoscape software (Version x64 3.6.1 https://cytoscape.org/) (Shannon, Markiel et al. 2003), R software (Version x64 3.6.3; https://www.r-project.org/) and the GraphPad Prism Software (San Diego, California USA). Results with $P < 0.05$ were considered statistically significant.

Results

Correlation of stromal and immune scores with different clinical features in patients with non-small-cell lung cancer

A total of 196 patients with NSCLC were included in this study, of which 107 (54.6%) were males, and 89 (45.4%) were females. Of the 196 patients, 106, 66, and 24 had LUAD, LUSC, and large cell neuroendocrine carcinoma, respectively. Regarding the clinical classification stage, 40 cases were in stage IA, 90 cases in stage IB, 6 cases in stage IIA, 29 cases in stage IIB, 21 cases in stage IIIA, 6 cases in stage IIIB, and 4 cases in stage IV (Table 1).

Table 1
Clinical features of patients in GSE37745

Clinical features		Count (%)
Age (y)		
	<=65	102(52%)
	> 65	94(48%)
Gender		
	Male	107(55%)
	Female	89(45%)
Status		
	survive	51(26%)
	died	145(74%)
Histology		
	LUAD	106(54%)
	LUSC	66(34%)
	LCNEC	24(12%)
Stage		
	IA	40(20%)
	IB	90(46%)
	IIA	6(3%)
	IIB	29(15%)
	IIIA	21(11%)
	IIIB	6(3%)
	IV	4(2%)

LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; LCNEC, Large cell neuroendocrine carcinoma

Then, we used the ESTIMATE algorithm to examine the distribution of stromal and immune scores for each patient. The results showed that LUAD and early clinical stages of NSCLC were associated with high immune scores ($P < 0.05$, Fig. 1A, B). Otherwise, patients with high immune scores and high stromal scores showed better survival rate ($P < 0.05$, Fig. 1C, D). Survival analysis showed that the OS in high and low stromal/immune score groups was not significantly different ($P > 0.05$, Fig. 1E, F).

Screening of differentially expressed genes between different score groups

In total, 893 genes were differentially expressed between high and low stromal score groups in the GSE37745 dataset, including 738 upregulated and 155 downregulated genes. Similarly, 787 genes were differentially expressed between high and low immune score groups, including 635 upregulated and 152 downregulated genes. The top 20 DEGs from different score groups were plotted, respectively (Fig. 2A, B). The number of overlapping DEGs between the stromal score groups and the immune score groups was 437, including 380 upregulated and 57 downregulated genes (Fig. 2C, D).

Identification of prognostic genes from differentially expressed genes

The Kaplan–Meier method was used to analyze the relationship between OS and the 437 overlapping DEGs in the GSE37745 dataset, and 43 genes were found to be closely associated with OS (**Supplementary table 1**). Therefore, *PLAUR*, *PSAT1*, and *RAD54B* were identified as prognosis risk genes in NSCLC (Fig. 2E).

Validation of prognostic genes and selected hub genes

The expression of *PSAT1* was verified using TCGA dataset and GEPIA2, which demonstrated that only *PSAT1* was highly expressed in LUAD and LUSC (Fig. 3A) and upregulated with increasing tumor stage (*P* = 1.991e-05, Fig. 3B). Kaplan–Meier Plotter data showed that higher expression of *PSAT1* led to lower survival time in NSCLC (Fig. 3C). The results were further validated by verifying the expression level of *PSAT1* using qRT-PCR. As expected, in comparison with matched adjacent nontumor tissues, *PSAT1* was significantly upregulated in cancer tissues (Fig. 3D and **Supplementary table 2**). In addition, we examined the expression of PSAT1 in NSCLC tissues and normal tissues by WB. A higher level of PSAT1 was observed in the NSCLC tissues and tumor samples, whereas a lower expression of PSAT1 was detected in normal tissues (Fig. 3E). The protein expression of PSAT1 was also investigated by immunohistochemical staining using tissue sections obtained from patients with NSCLC at our cancer center, which showed that PSAT1 was more highly expressed in tumor tissues than in peritumor tissues (Fig. 3F). Given the results of above, we evaluated hub genes' levels in a cohort of 99 NSCLC patients with median 5-y follow-up in which high-score patients displayed poor OS (Fig. 3G and **Supplementary table 3)**.

Enrichment analysis of PSAT1 co-expression genes

To further explore the molecular mechanism of the *PSAT1* gene in tumorigenesis, we identified the top 20 *PSAT1*-binding proteins using the STRING database (Fig. 4A). GO term enrichment showed that *PSAT1* co-expression genes were mainly involved in the alpha-amino acid metabolic and biosynthetic process, cellular amino acid metabolic process, carboxylic acid biosynthetic process, organic acid biosynthetic

process, and cellular amino acid biosynthetic process (Fig. 4B**)**. The KEGG pathway enrichment analysis indicated that biosynthesis of amino acids; glycine, serine and threonine metabolism; and carbon metabolism are involved in the effect of *PSAT1* on tumor pathogenesis (Fig. 4C and **Supplementary table 4**). Remarkably, the top 20 positively co-expressed genes showed a high probability of becoming high-risk markers of NSCLC (Fig. 4D).

Transcription factor–miRNA–target network construction

To understand the regulatory factors of *PSAT1* in NSCLC, we analyzed the miRNAs and transcription factors (TFs) associated with *PSAT1*. This network consisted of 24 nodes and 36 interactions, from which we were able to identify 12 TFs—PAX9, TP53, ETS1, FOXI1, IRF3, NR2F1, NFYA, E2F1, NKX2-3, GATA5, SRF, and EBF1and 12miRNAs—including hsa-miR-1914-5p, hsa-miR-192-5p, hsa-miR-7109-5p, hsa-miR-874-3p, hsa-miR-7113-3p, hsa-miR-1224-5p, hsa-miR-4656, hsa-miR-6886-5p, hsa-miR-3619-3p, hsa-miR-1207-5p, hsa-miR-665 and hsa-miR-659-3p.that potentially interact with *PSAT1* (Fig. 4E and **Supplementary table 5**).

Immune infiltration analysis of PSAT1 in NSCLC

We evaluated the correlation between *PSAT1* expression and immune invasion in NSCLC using the TIMER database. The results revealed that PSAT1 expression levels were correlated with infiltration of B cells (LUAD: r = -0.072, $P = 1.14e^{-01}$; LUSC: r = -0.024, $P = 5.98e^{-01}$), $CD8^+$ T cells (LUAD: r = 0.081, $P = 7.24e^{-02}$; LUSC: r = -0.065, $P = 1.57e^{-01}$), $CD4^+$ T cells (LUAD: r = -0.193, $P = 1.89e^{-05}$; LUSC: r = -0.298, $P = 3.56e^{-11}$), macrophages (LUAD: r = -0.117, $P = 9.94e^{-03}$; LUSC: r = -0.197, $P = 1.54e^{-05}$), neutrophils (LUAD: r = -0.018, $P = 6.95e^{-01}$; LUSC: r = -0.239, $P = 1.30e^{-07}$), and dendritic cells (LUAD: r = -0.115, $P = 1.07e^{-02}$; LUSC: r = -0.209, $P = 4.68e^{-06}$; Fig. 4F).

Pan-cancer analysis of PSAT1

Gene expression analysis indicated a significant difference in *PSAT1* expression between healthy tissues and tumor tissues of breast invasive carcinoma (BRCA), cervical squamous cell carcinoma and endocervical adenocarcinoma, Colon adenocarcinoma, lymphoid neoplasm diffuse large B-cell lymphoma, glioblastoma multiforme, Brain lower grade glioma, ovarian serous cystadenocarcinoma, prostate adenocarcinoma, rectal adenocarcinoma, stomach adenocarcinoma (STAD), thymoma, and uterine corpus endometrial carcinoma (UCEC; Fig. 5A, $P < 0.05$). We also found that *PSAT1* expression is closely associated with the pathological stages of BRCA, kidney chromophobe, kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), STAD, testicular germ cell tumor, thyroid carcinoma (THCA), and UCEC (Fig. 5B, $P < 0.05$). Next, we assessed the prognostic value of *PSAT1* expression across cancers in Kaplan-Meier Plotter, and results showed that *PSAT1* plays a detrimental role in BRCA, KIRC, KIRP, liver hepatocellular carcinoma (LIHC), pancreatic adenocarcinoma (PAAD), sarcoma (SARC), THCA, and UCEC (Fig. 5C, P **< 0.05**).

Discussion

A wealth of evidence shows that the TME plays an important role in carcinoma (Lin, Karakasheva et al. 2016, Skarstein, Jensen et al. 2019). A deeper understanding of tumor cell–TME interactions contribute to developing novel therapeutic strategies for NSCLC. In this study, analysis of the GSE37745 dataset showed that *PSAT1* may be considered as a tumor microenvironment-related prognostic biomarker in NSCLC. Online and experimental validation indicated high expression of *PSAT1* in NSCLC tissues, and survival analysis by Kaplan-Meier Plotter suggested that high expression of *PSAT1* was linked to poor prognosis in NSCLC, which may contribute to the research on the progression of NSCLC and could be regarded as biomarkers and therapeutic targets in NSCLC.

TFs are DNA-binding proteins that play important roles in the regulation of gene expression, cell proliferation, and apoptosis (Lee and Young 2013). miRNAs are short, non-coding RNAs consisting of 18–25 nucleotides that regulate the translation of mRNAs (Doench and Sharp 2004). It is well known that TFs promote or repress transcription at the pre-transcription level, whereas miRNAs play an important regulatory role at the post-transcriptional level. In our study, Cytoscape was used to construct a regulation network of TFs, miRNAs, and their targets, involving 12 TFs and 34 miRNAs. Although miRNAs such as miR-34 (Bommer, Gerin et al. 2007), miR-28 (Davidson, Larsen et al. 2010), miR-155 (Yanaihara, Caplen et al. 2006), miR-21 (Markou, Tsaroucha et al. 2008), and miR-31 (Liu, Sempere et al. 2010) are evidently associated with the development and progression of NSCLC, miRNAs associated with NSCLC need further research.

PSAT1 is an enzyme involved in serine biosynthesis, and excessive activation of the serine/glycine metabolic pathway is considered to promote the occurrence of cancer by promoting cell cycle progression, cell proliferation, and tumorigenesis (Baek, Jun et al. 2003, DeBerardinis 2011, Yang, Wu et al. 2015). A growing amount of evidence indicates that *PSAT1* is an oncogene that affects cancer progression and metastasis (Ojala, Sundstrom et al. 2002, Martens, Nimmrich et al. 2005). Previous studies have demonstrated that PSAT1 may facilitate cell cycle progression via regulation of the GSK3β/β-catenin/cyclin D1 pathway (Yang, Wu et al. 2015).

Through correlation analysis, we reported that PSAT1 expression is associated with several immune infiltrating cells in LUSC and LUAD (Fig. 4F). These results suggest that PSAT1 is involved in the regulation of tumor immune cells. The study of Yuan fan et al. also revealed that PSAT1 may promote the progression of endometrial cancer by affecting immune cell infiltration(Fan, Li et al. 2021).

Finally, the pan-cancer study also suggested that *PSAT1* may serve as a good prognostic biomarker in diverse cancers, including KIRC, KIRP, LIHC, PAAD, SARC, THCA, and UCEC.

Conclusion

In summary, our results indicate that *PSAT1* is a promising prognostic biomarker for NSCLC; it is necessary to further study the potential therapeutic value of *PSAT1*.

Abbreviations

BRCA	Breast invasive carcinoma
DEGs	Differentially expressed genes
DLBC	Lymphoid neoplasm Diffuse large B-cell lymphoma
GEO	Gene expression omnibus
GEPIA2	Gene Expression Profiling Interactive Analysis 2
KEGG	Kyoto Encyclopedia of Genes and Genomes
KIRC	Kidney renal clear cell carcinoma
KIRP	Kidney renal papillary cell carcinoma
LIHC	Liver hepatocellular carcinoma
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
miRNA	microRNA
NSCLC	Non-small-cell lung cancer
PAAD	Pancreatic adenocarcinoma
PSAT1	Phosphoserine aminotransferase 1
READ	Rectal adenocarcinoma
SARC	Sarcoma
STAD	Stomach adenocarcinoma
TCGA	The Cancer Genome Atlas
TF	Transcription factor
THCA	Thyroid carcinoma
TIMER	Tumor Immune Estimation Resource
TME	Tumor microenvironment
UCEC	Uterine corpus endometrial carcinoma

Declarations

Ethics approval and consent to participate

The human tissue acquisition and experiment conduction in this study were approved by the Ethics Committee of Sun Yat-sen University Cancer Center (NO. YB2018-85).

Consent for publication

Not applicable.

Availability of data and materials

Publicly available datasets were analyzed in this study. This data can be found here: Gene Expression Omnibus (GEO) Datasets (https://www.ncbi.nlm.nih.gov/). All data generated or analyzed during this study are included in this published article and its supplementary information files.

Funding

This work was supported by the National Natural Science Foundation of China (NO.81871986).

Author contributions

Zhesheng Wen and Hongmu Li contributed to the study design. Huikai Miao, Mingyue Zeng and Youfang Chen contributed to literature search. Huikai Miao contributed to collecting cancer tissue and experimental validation, Hongmu Li and Minglei Yang wrote the article and performed data analysis. Zhesheng Wen and Minglei Yang contributed to edit, supervision and funding acquisition. All authors gave the final approval of the version to be submitted.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We acknowledge TCGA and GEO database for providing their platforms and contributors for uploading their meaningful datasets. We sincerely thank Ms. Baolian Qiu for helping to edit this article.

References

1. Baek, J. Y., D. Y. Jun, D. Taub and Y. H. Kim (2003). "Characterization of human phosphoserine aminotransferase involved in the phosphorylated pathway of L-serine biosynthesis." Biochem J **373**(Pt 1): 191-200.
2. Bommer, G. T., I. Gerin, Y. Feng, A. J. Kaczorowski, R. Kuick, R. E. Love, Y. Zhai, T. J. Giordano, Z. S. Qin, B. B. Moore, O. A. MacDougald, K. R. Cho and E. R. Fearon (2007). "p53-mediated activation of miRNA34 candidate tumor-suppressor genes." Curr Biol **17**(15): 1298-1307.
3. Cancer Genome Atlas Research, N., J. N. Weinstein, E. A. Collisson, G. B. Mills, K. R. Shaw, B. A. Ozenberger, K. Ellrott, I. Shmulevich, C. Sander and J. M. Stuart (2013). "The Cancer Genome Atlas

Pan-Cancer analysis project." Nat Genet **45**(10): 1113-1120.

4. Chandra Mangalhara, K., S. Manvati, S. K. Saini, K. Ponnusamy, G. Agarwal, S. K. Abraham and R. N. K. Bamezai (2017). "ERK2-ZEB1-miR-101-1 axis contributes to epithelial-mesenchymal transition and cell migration in cancer." Cancer Lett **391**: 59-73.
5. Chen, F., Y. Zhang, E. Parra, J. Rodriguez, C. Behrens, R. Akbani, Y. Lu, J. M. Kurie, D. L. Gibbons, G. B. Mills, Wistuba, II and C. J. Creighton (2017). "Multiplatform-based molecular subtypes of non-small-cell lung cancer." Oncogene **36**(10): 1384-1393.
6. Consortium, G. T. (2013). "The Genotype-Tissue Expression (GTEx) project." Nat Genet **45**(6): 580-585.
7. Davidson, M. R., J. E. Larsen, I. A. Yang, N. K. Hayward, B. E. Clarke, E. E. Duhig, L. H. Passmore, R. V. Bowman and K. M. Fong (2010). "MicroRNA-218 is deleted and downregulated in lung squamous cell carcinoma." PLoS One **5**(9): e12560.
8. DeBerardinis, R. J. (2011). "Serine metabolism: some tumors take the road less traveled." Cell Metab **14**(3): 285-286.
9. Doench, J. G. and P. A. Sharp (2004). "Specificity of microRNA target selection in translational repression." Genes Dev **18**(5): 504-511.
10. Fan, Y., X. Li, L. Tian and J. Wang (2021). "Identification of a Metabolism-Related Signature for the Prediction of Survival in Endometrial Cancer Patients." Front Oncol **11**: 630905.
11. Fridman, W. H., J. Galon, M. C. Dieu-Nosjean, I. Cremer, S. Fisson, D. Damotte, F. Pages, E. Tartour and C. Sautes-Fridman (2011). "Immune infiltration in human cancer: prognostic significance and disease control." Curr Top Microbiol Immunol **344**: 1-24.
12. Hanahan, D. and L. M. Coussens (2012). "Accessories to the crime: functions of cells recruited to the tumor microenvironment." Cancer Cell **21**(3): 309-322.
13. Herbst, R. S., D. Morgensztern and C. Boshoff (2018). "The biology and management of non-small cell lung cancer." Nature **553**(7689): 446-454.
14. Huang, H., W. Zhou, R. Chen, B. Xiang, S. Zhou and L. Lan (2021). "CXCL10 is a Tumor Microenvironment and Immune Infiltration Related Prognostic Biomarker in Pancreatic Adenocarcinoma." Front Mol Biosci **8**: 611508.
15. Janky, R., A. Verfaillie, H. Imrichova, B. Van de Sande, L. Standaert, V. Christiaens, G. Hulselmans, K. Herten, M. Naval Sanchez, D. Potier, D. Svetlichnyy, Z. Kalender Atak, M. Fiers, J. C. Marine and S. Aerts (2014). "iRegulon: from a gene list to a gene regulatory network using large motif and track collections." PLoS Comput Biol **10**(7): e1003731.
16. Junttila, M. R. and F. J. de Sauvage (2013). "Influence of tumour micro-environment heterogeneity on therapeutic response." Nature **501**(7467): 346-354.
17. Kim, I. S., Y. Gao, T. Welte, H. Wang, J. Liu, M. Janghorban, K. Sheng, Y. Niu, A. Goldstein, N. Zhao, I. Bado, H. C. Lo, M. J. Toneff, T. Nguyen, W. Bu, W. Jiang, J. Arnold, F. Gu, J. He, D. Jebakumar, K. Walker, Y. Li, Q. Mo, T. F. Westbrook, C. Zong, A. Rao, A. Sreekumar, J. M. Rosen and X. H. Zhang

(2019). "Immuno-subtyping of breast cancer reveals distinct myeloid cell profiles and immunotherapy resistance mechanisms." Nat Cell Biol **21**(9): 1113-1126.

18. Kortlever, R. M., N. M. Sodir, C. H. Wilson, D. L. Burkhart, L. Pellegrinet, L. Brown Swigart, T. D. Littlewood and G. I. Evan (2017). "Myc Cooperates with Ras by Programming Inflammation and Immune Suppression." Cell **171**(6): 1301-1315 e1314.
19. Lee, T. I. and R. A. Young (2013). "Transcriptional regulation and its misregulation in disease." Cell **152**(6): 1237-1251.
20. Li, T., J. Fan, B. Wang, N. Traugh, Q. Chen, J. S. Liu, B. Li and X. S. Liu (2017). "TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells." Cancer Res **77**(21): e108-e110.
21. Lin, E. W., T. A. Karakasheva, P. D. Hicks, A. J. Bass and A. K. Rustgi (2016). "The tumor microenvironment in esophageal cancer." Oncogene **35**(41): 5337-5349.
22. Liu, X., L. F. Sempere, H. Ouyang, V. A. Memoli, A. S. Andrew, Y. Luo, E. Demidenko, M. Korc, W. Shi, M. Preis, K. H. Dragnev, H. Li, J. Direnzo, M. Bak, S. J. Freemantle, S. Kauppinen and E. Dmitrovsky (2010). "MicroRNA-31 functions as an oncogenic microRNA in mouse and human lung cancer cells by repressing specific tumor suppressors." J Clin Invest **120**(4): 1298-1309.
23. Markou, A., E. G. Tsaroucha, L. Kaklamanis, M. Fotinou, V. Georgoulias and E. S. Lianidou (2008). "Prognostic value of mature microRNA-21 and microRNA-205 overexpression in non-small cell lung cancer by quantitative real-time RT-PCR." Clin Chem **54**(10): 1696-1704.
24. Martens, J. W., I. Nimmrich, T. Koenig, M. P. Look, N. Harbeck, F. Model, A. Kluth, J. Bolt-de Vries, A. M. Sieuwerts, H. Portengen, M. E. Meijer-Van Gelder, C. Piepenbrock, A. Olek, H. Hofler, M. Kiechle, J. G. Klijn, M. Schmitt, S. Maier and J. A. Foekens (2005). "Association of DNA methylation of phosphoserine aminotransferase with response to endocrine therapy in patients with recurrent breast cancer." Cancer Res **65**(10): 4101-4117.
25. Molina, J. R., P. Yang, S. D. Cassivi, S. E. Schild and A. A. Adjei (2008). "Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship." Mayo Clin Proc **83**(5): 584-594.
26. Nagy, A., G. Munkacsy and B. Gyorffy (2021). "Pancancer survival analysis of cancer hallmark genes." Sci Rep **11**(1): 6047.
27. Ojala, P., J. Sundstrom, J. M. Gronroos, E. Virtanen, K. Talvinen and T. J. Nevalainen (2002). "mRNA differential display of gene expression in colonic carcinoma." Electrophoresis **23**(11): 1667-1676.
28. Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker (2003). "Cytoscape: a software environment for integrated models of biomolecular interaction networks." Genome Res **13**(11): 2498-2504.
29. Skarstein, K., J. L. Jensen, H. Galtung, R. Jonsson, K. Brokstad and L. A. Aqrawi (2019). "Autoantigen-specific B cells and plasma cells are prominent in areas of fatty infiltration in salivary glands of patients with primary Sjogren's syndrome." Autoimmunity **52**(7-8): 242-250.
30. Tang, Z., B. Kang, C. Li, T. Chen and Z. Zhang (2019). "GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis." Nucleic Acids Res **47**(W1): W556-W560.

31. Torre, L. A., F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent and A. Jemal (2015). "Global cancer statistics, 2012." CA Cancer J Clin **65**(2): 87-108.
32. Wang, Y., M. Zhu, F. Guo, Y. Song, X. Fan and G. Qin (2020). "Identification of Tumor Microenvironment-Related Prognostic Biomarkers in Luminal Breast Cancer." Front Genet **11**: 555865.
33. Whitehouse, P. J. (1989). "Parkinson's disease and Alzheimer's disease: new neurochemical parallels." Mov Disord **4 Suppl 1**: S57-62.
34. Whiteside, T. L. (2008). "The tumor microenvironment and its role in promoting tumor growth." Oncogene **27**(45): 5904-5912.
35. Yanaihara, N., N. Caplen, E. Bowman, M. Seike, K. Kumamoto, M. Yi, R. M. Stephens, A. Okamoto, J. Yokota, T. Tanaka, G. A. Calin, C. G. Liu, C. M. Croce and C. C. Harris (2006). "Unique microRNA molecular profiles in lung cancer diagnosis and prognosis." Cancer Cell **9**(3): 189-198.
36. Yang, S., T. Liu, H. Nan, Y. Wang, H. Chen, X. Zhang, Y. Zhang, B. Shen, P. Qian, S. Xu, J. Sui and G. Liang (2020). "Comprehensive analysis of prognostic immune-related genes in the tumor microenvironment of cutaneous melanoma." J Cell Physiol **235**(2): 1025-1035.
37. Yang, Y., J. Wu, J. Cai, Z. He, J. Yuan, X. Zhu, Y. Li, M. Li and H. Guan (2015). "PSAT1 regulates cyclin D1 degradation and sustains proliferation of non-small cell lung cancer cells." Int J Cancer **136**(4): E39-50.
38. Yoshihara, K., M. Shahmoradgoli, E. Martinez, R. Vegesna, H. Kim, W. Torres-Garcia, V. Trevino, H. Shen, P. W. Laird, D. A. Levine, S. L. Carter, G. Getz, K. Stemke-Hale, G. B. Mills and R. G. Verhaak (2013). "Inferring tumour purity and stromal and immune cell admixture from expression data." Nat Commun **4**: 2612.
39. Zeng, D., M. Li, R. Zhou, J. Zhang, H. Sun, M. Shi, J. Bin, Y. Liao, J. Rao and W. Liao (2019). "Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures." Cancer Immunol Res **7**(5): 737-750.

Figures

Figure 1

The comparisons of stromal scores and immune scores in different clinical features and clinical outcomes. (**A**) The differences of immune scores of NSCLC in clinical stages. (**B**) The differences of immune scores of NSCLC in varied types. (**C**) The differences of immune scores of NSCLC in overall survival. (**D**) The differences of stromal scores of NSCLC in overall survival. (**E**) The differences of immune scores of NSCLC in overall survival, p = 0.427. (**F**) The differences of stromal scores of NSCLC in overall survival. P = 0.380. There is no statistical difference.

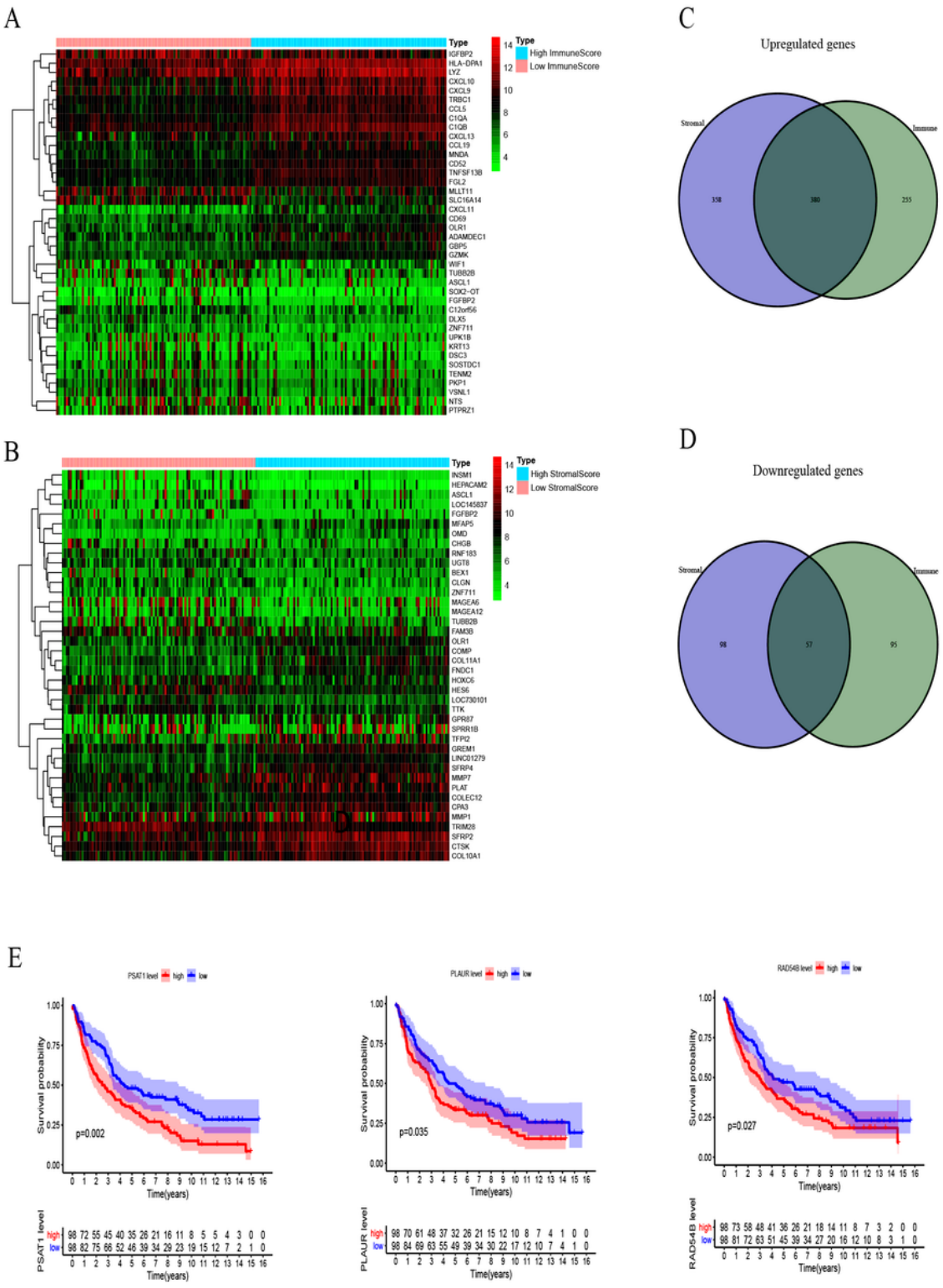

Figure 2

DEGs between high- and low- immune /stromal score group. (**A**) Heatmap of top 20 DEGs of immune scores (low- and high-) in GSE37745 dataset. (**B**) Heatmap of top 20 DEGs of stromal scores (low- and high-). (**C**) The commonly up-regulated DEGs between immune-/ and stromal- score group. (**D**) The commonly down-regulated DEGs between immune-/ and stromal- score group. **(E)** Survival analysis of

the prognosis-related genes in GSE37745 dataset, P<0.05 was considered statistically significant. PLAUR, PSAT1 and RAD54B were considered as risk genes.

Figure 3

Validation of hub genes. (A) Expression of PLAUR based on TCGA dataset through GEPIA website. **(B)** Based on the TCGA data, the relationship between the expression levels of PSAT1 and pathological stages. **(C)** The relationship between PSAT1 and survival rate of patients with NSCLC. **(D)** Expression of PSAT1 in NSCLC, P < 0.05. **(E)** Western blot experiment showed that PSAT1 was overexpressed in cancer tissue. **(F)** Immunohistochemistry confirmed the differential expression of PSAT1 in NSCLC tissues and normal tissues. **(G)** Strong staining for PSAT1 correlates with poor OS in NSCLC patients (n = 99). P < 0.001 by log-rank test.

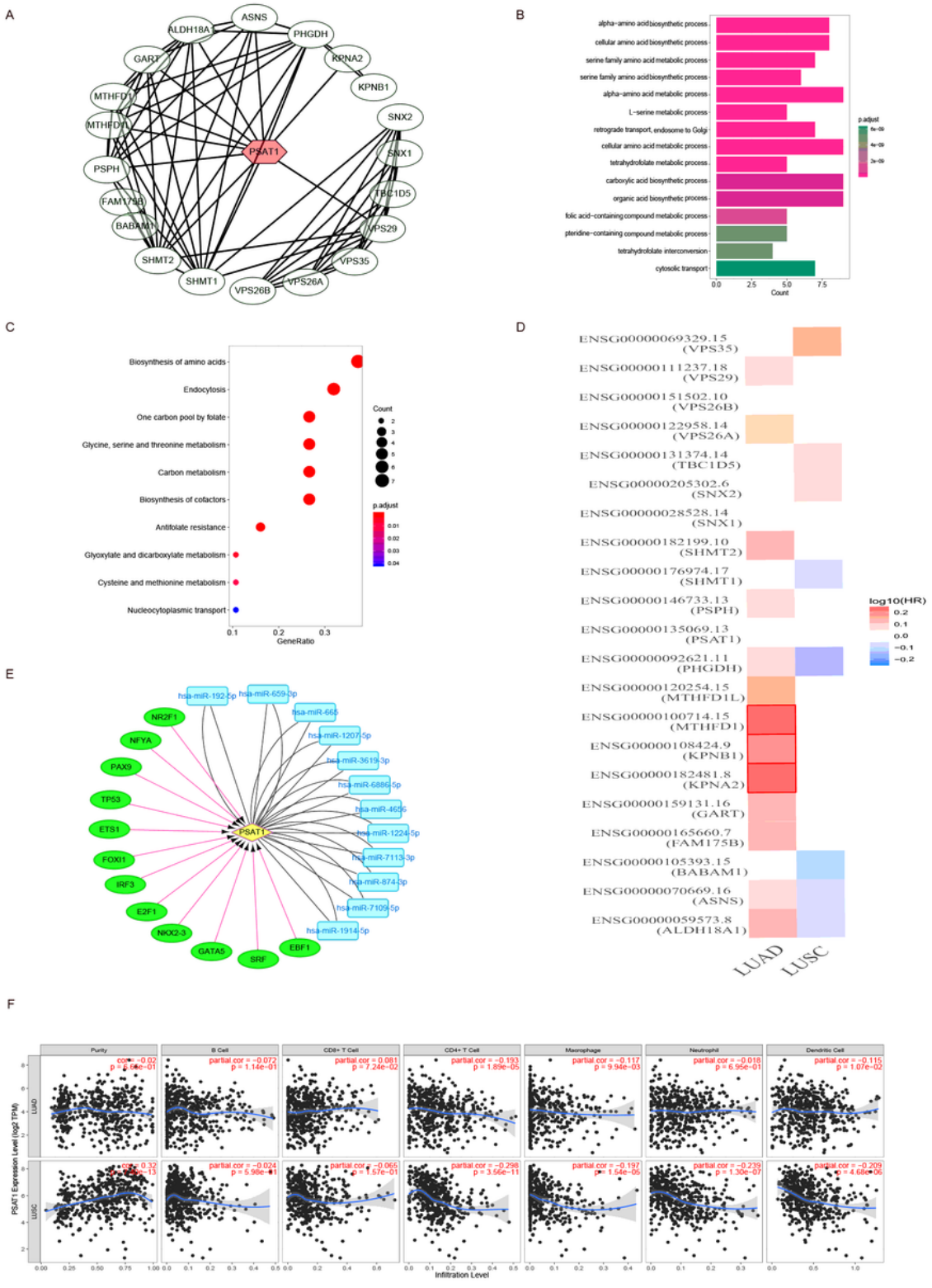

Figure 4

Investigate the potential molecular mechanism of PSAT1. (A) PPI network of top 20 PSAT1-binding proteins based on the STRING database. **(B** and **C)** GO and KEGG pathway of PSAT1 co-expression genes. **(D)** Survival heatmaps of the top 20 genes correlated with PSAT1 in NSCLC. **(E)** TF-miRNA-Target regulatory network. Blue oval represented the miRNA; Green oval represent TFs. **(F)** Correlations between PSAT1 expression level and NSCLC immune infiltration obtained from the TIMER database.

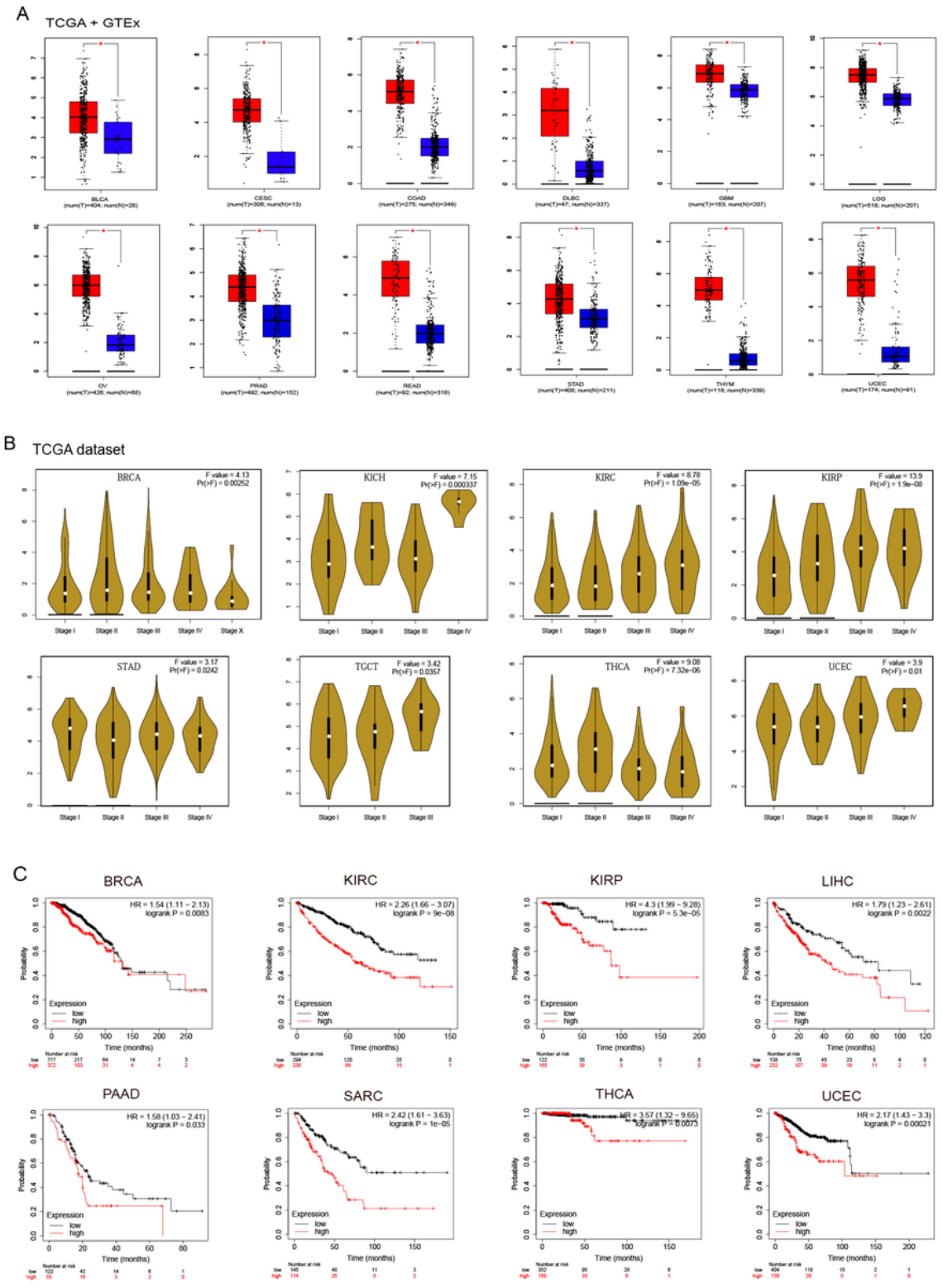

Figure 5

Pan-cancer analysis of PSAT1. (A and B) Expression level of *PSAT1* gene in different tumors and pathological stages were analyzed through GEPIA2. **(C)**The relationship between PSAT1 and survival rate of diverse tumors based on Kaplan-Meier Plotter. BRCA: Breast Invasive Carcinoma; CESC: Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma, DLBC: Lymphoid Neoplasm Diffuse Large B-cell Lymphoma, GBM: Glioblastoma Multiforme, KICH: Kidney Chromophobe, KIRC: Kidney Renal Clear

Cell Carcinoma, KIRP: Kidney Renal Papillary Cell Carcinoma, LGG: Brain Lower Grade Glioma, LIHC: Liver Hepatocellular Carcinoma, OV: Ovarian Serous Cystadenocarcinoma, PAAD: Pancreatic Adenocarcinoma, PRAD: Prostate Adenocarcinoma, READ: Rectum Adenocarcinoma, SARC: Sarcoma, STAD: Stomach Adenocarcinoma, TGCT: Testicular Germ Cell Tumors, THCA: Thyroid , Carcinoma, THYM: Thymoma, UCEC: Uterine Corpus Endometrial Carcinoma.

Supplementary Files

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

- Supplementarytable1.xlsx
- Supplementarytable2.xlsx
- Supplementarytable3.xlsx
- Supplementarytable4.xlsx
- Supplementarytable5.xlsx