

Bioinformatics Analysis of SPP1 Expression, Prognostic Value, and Molecular Mechanisms in Lung Adenocarcinoma

Dong Yao^{1*}

¹ Beijing Anzhen Nanchong Hospital, Capital Medical University & Nanchong Central Hospital, Nanchong, China

Email Address

1763876092@qq.com (Dong Yao)

*Correspondence: 1763876092@qq.com

Received: 2 December 2025; **Accepted:** 23 December 2025; **Published:** 13 January 2026

Abstract:

Objective: To investigate the expression, clinical prognostic value, and potential molecular mechanisms of SPP1 in lung adenocarcinoma (LUAD) to provide a target for the precise diagnosis and treatment of LUAD. **Methods:** Multiple public databases, including TIMER, GEPIA, and CPTAC, were integrated to perform bioinformatics analysis of SPP1 mRNA/protein expression, its correlation with clinicopathological features and prognosis, to construct a co-expression gene network, and to conduct KEGG pathway enrichment analysis. **Results:** SPP1 was highly expressed at both mRNA and protein levels in LUAD. The overall survival (OS) and progression-free survival (PFS) of patients with increased expression level were worse ($P < 0.001$), and this expression level was related to the patient's age and clinical stage. The co-expressed genes were related to the integrin and matrix metalloproteinase families and were enriched in tumor-related pathways such as ECM-receptor interaction and PI3K-Akt. **Conclusion:** The increase of SPP1 level is related to the bad results and the deterioration of LUAD and may promote tumorigenesis and development by regulating related pathways, suggesting that this may become an important predictor and intervention direction of lung adenocarcinoma.

Keywords:

Lung Adenocarcinoma, SPP1, Bioinformatics Analysis, Prognostic Biomarker, Molecular Mechanism

1. Introduction

Lung cancer is one of the most prevalent and fatal malignancies globally, representing a significant threat to human health [1]. The International Agency for Research on Cancer (IARC), a cancer research institution under the World Health Organization, has published some figures, showing that in 2020, about 2.2 million people in the world were diagnosed with lung cancer, and 1.8 million people died of the disease [2]. Lung adenocarcinoma (LUAD) is the most common non-small cell lung cancer (NSCLC), accounting for 40% to 50% of all such cases [3]. While recent advancements in precision medicine, including targeted and immunotherapies, have

improved the management of LUAD, the disease's asymptomatic nature at early stages means that patients are often diagnosed in the mid-to-late stages[4]. Therefore, combined with a high propensity for recurrence and distant metastasis, results in a 5-year survival rate that remains under 20%. Consequently, a deeper understanding of the key molecular drivers of LUAD pathogenesis, along with their expression patterns and regulatory networks, is essential to identify reliable biomarkers for early diagnosis and effective therapeutic targets, thereby enhancing patient outcomes.

Osteopontin (OPN), also called Secreted Phosphoprotein 1 (SPP1), is an acidic glycoprotein with 284 amino acids, which has many functions in our body. This protein exists in many body tissues and organs, such as bone, kidney, breast and liver [5]. By binding to integrin receptors (such as $\alpha v\beta 3$, $\alpha 5\beta 1$) and CD44 molecules, SPP1 participates in regulating a variety of physiological and pathological processes, including cell adhesion, migration, proliferation, apoptosis, angiogenesis, and immune regulation [6]. Recent studies have shown that SPP1 gene is extremely active in many cancers, and it is closely related to the occurrence, development, spread and metastasis of these cancers and the final results of patients [7].

However, research on the expression characteristics, clinical significance, and regulatory mechanisms of SPP1 in LUAD is not yet systematic or in-depth. Some existing studies have reported that SPP1 is highly expressed in LUAD tissues, but these studies have small sample sizes and did not perform cross-validation using multiple databases [8]. Meanwhile, there is still controversy regarding the association of SPP1 with clinicopathological features and prognostic value in LUAD patients, and its specific molecular mechanism in the occurrence and development of LUAD has not been fully elucidated. Based on this, this study adopted bioinformatics analysis methods, integrating multiple public database resources such as TIMER, GEPIA, CPTAC, HPA, KM Plotter, and GeneMANIA, to systematically analyze the mRNA and protein expression levels of SPP1 in LUAD, to explore its relationship with patients' clinicopathological features and prognosis, to construct a co-expression gene network of SPP1, and to mine its potential molecular regulatory mechanisms through KEGG pathway enrichment analysis, in order to provide a theoretical basis for further elucidating the role of SPP1 in LUAD, and at the same time, to provide new ideas and targets for the precise diagnosis and treatment of LUAD.

2. Materials and Methods

2.1. Data Sources and Research Tools

This study only uses publicly available data, so there is no need for ethical review. Databases and tools used: (1) TIMER 2.0: Analyzed the mRNA expression differences of SPP1 in pan-cancer and LUAD through the Gene_DE module [9]. (2) GEPIA: Verified the expression differences of SPP1 in LUAD and normal lung tissues and its association with clinical stage ($P < 0.05$, $|\log_2FC| > 1$) [10]. (3) CPTAC: Analyzed the protein expression of SPP1 in LUAD and its association with clinicopathological features [11]. (4) The Human Protein Atlas (HPA): Observed the protein localization and expression intensity of SPP1 through immunohistochemical sections [12]. (5) KM Plotter: Analyzed the differences in OS and PFS between high/low expression groups of SPP1 in LUAD patients (log-rank test) [13]. (6) GeneMANIA: Selected 50 SPP1-related genes and constructed a co-expression network [14]. (7) KEGG: Performed enrichment analysis on the core pathways of SPP1 and co-expressed genes ($P < 0.05$) [15].

2.2. Research Methods

2.2.1. mRNA Expression Analysis

Utilizing TIMER 2.0 to analyze mRNA expression differences of SPP1 in pan-cancer and LUAD; GEPIA verified expression differences between LUAD tumor and normal tissues and their correlation with clinical stages (Stage I-IV).

2.2.2. Protein Expression and Clinical Association Analysis

Utilizing the CPTAC LUAD dataset (tumor/normal, n=111 each), t-test compared SPP1 protein expression differences; combined with its clinical information, ANOVA or t-test compared expression differences across different stratifications (age, gender, etc.). HPA immunohistochemistry verified protein localization and expression intensity.

2.2.3. Prognostic Association Analysis

Utilizing the KM Plotter LUAD dataset, analyzed OS and PFS differences between high/low SPP1 expression groups, with log-rank test assessing survival significance.

2.2.4. Co-expression Network Construction

Acquired 50 SPP1-related genes through GeneMANIA, constructed and analyzed co-expression networks.

2.2.5. KEGG Enrichment Analysis

Collated SPP1 and co-expressed genes, KEGG enrichment analysis screened core pathways ($P < 0.05$).

2.3. Statistical Analysis

Employed database-specific tools for statistical analysis: t-test/ANOVA compared inter-group expression differences, log-rank test analyzed survival curves, Fisher's exact test calculated enrichment P-values.

3. Results

3.1. mRNA expression characteristics of SPP1 in LUAD

Analysis using TIMER 2.0 revealed that SPP1 is highly expressed in multiple tumors, including LUAD ($P < 0.05$, Figure 1A). Verification via GEPIA confirmed that SPP1 mRNA expression levels in LUAD tumor tissues were significantly elevated compared to normal lung tissues ($P < 0.05$, Figure 1B). Furthermore, expression levels increased with advancing clinical stage, with Stage III-IV showing significantly higher expression than Stage I-II ($P < 0.05$).

3.2 Protein expression characteristics of SPP1 in LUAD

Protein group analysis of CPTAC showed that the level of SPP1 protein z-score in LUAD tumor samples was significantly higher than that in adjacent normal tissues ($P < 0.05$, Figure 1C). Immunohistochemical validation by HPA indicated weak to negative SPP1 staining in normal lung tissues (Figure 1D), whereas LUAD tissues exhibited moderate to strong positive staining localized to the cytoplasm and cell membrane (Figure 1E).

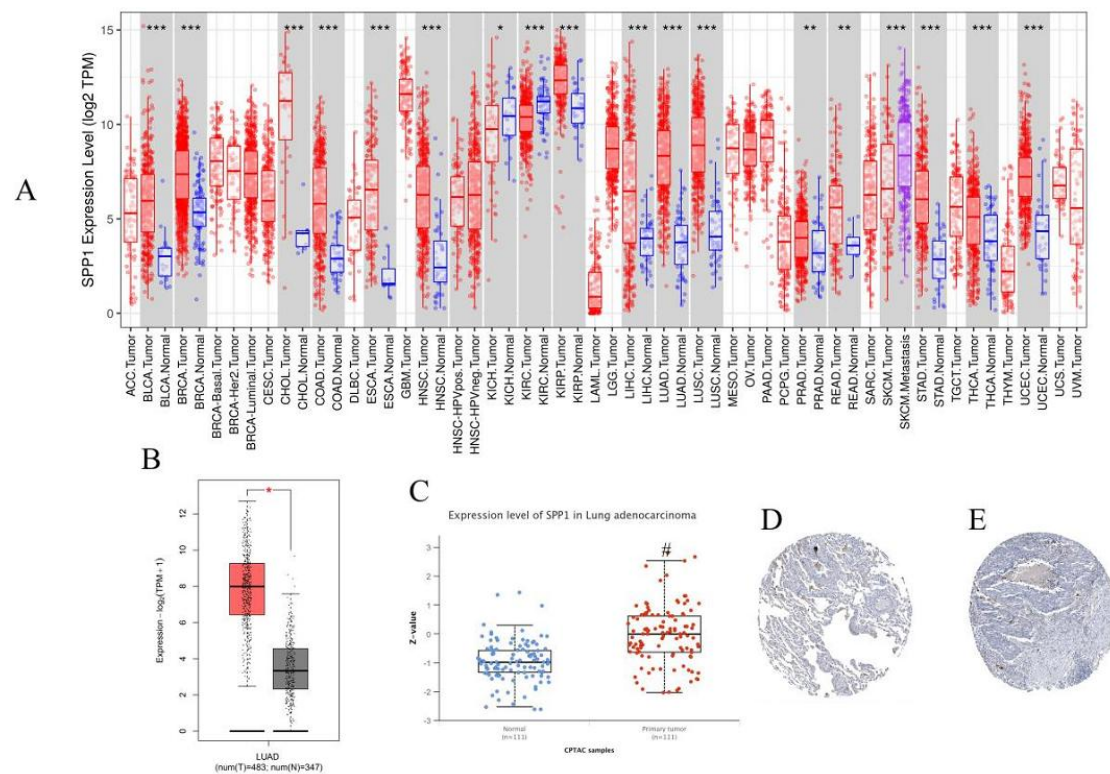
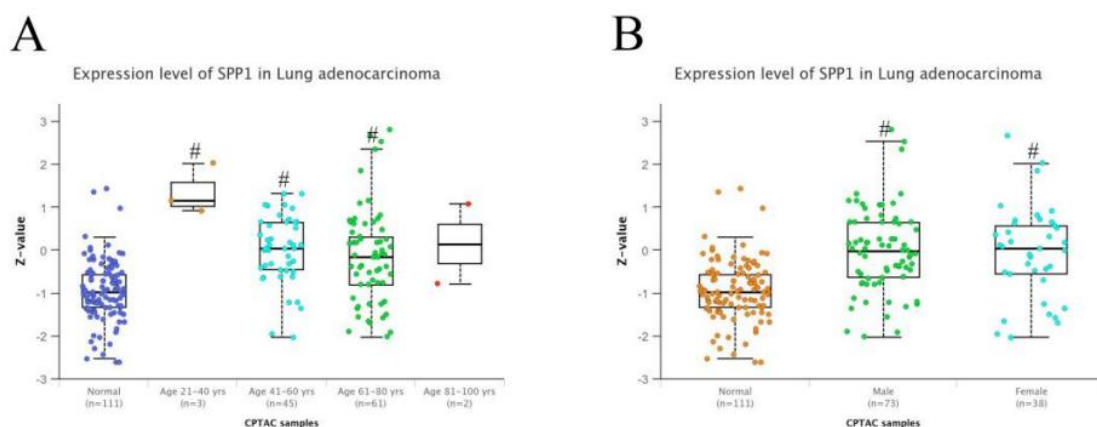


Figure 1. Expression of SPP1 in LUAD. *A*: TIMER database analyzed mRNA expression differences of SPP1 in pan-cancer and LUAD. *B*: GEPIA database verified mRNA expression differences between LUAD tumor and normal tissues. *C*: CPTAC database analyzed protein expression differences of SPP1. *D-E*: Immunohistochemistry results from HPA database (*D*: Weakly positive/negative in normal lung tissue; *E*: Strongly positive in LUAD tissue, localized in cytoplasm and cell membrane). *, # indicate $P < 0.05$.

3.3. Correlation between SPP1 expression and clinical and pathological features of LUAD

CPTAC analysis (Figure 2) revealed that the 61-80 age group exhibited the highest SPP1 expression, with both the 41-60 and 61-80 age groups showing significantly higher expression than the ≤ 40 and > 80 age groups ($P < 0.05$). Stage III expression was significantly greater than Stage I-II ($P < 0.05$), with no significant increase observed in Stage IV. It was found that there was no significant difference between different sexes or races ($P > 0.05$), which indicated that the expression of SPP1 might be more closely related to age and clinical stage.



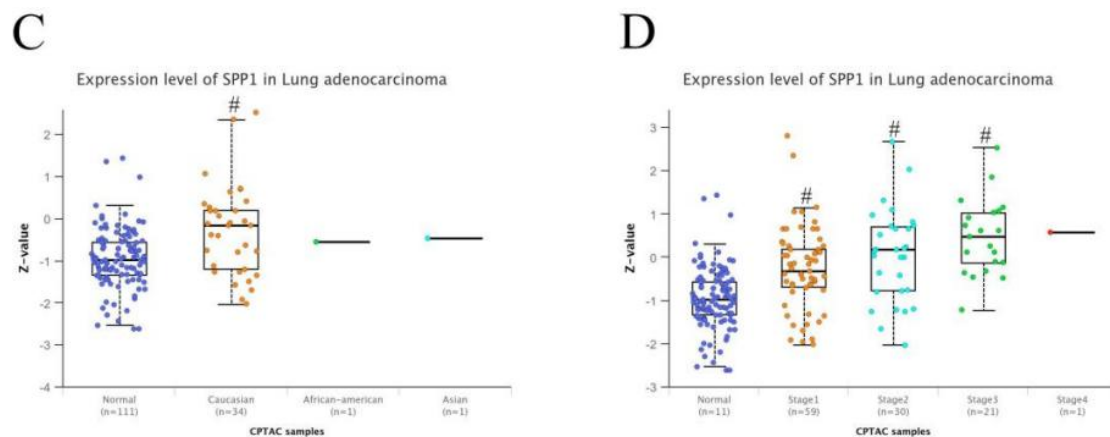


Figure 2. Association between SPP1 expression and clinicopathological features of LUAD. A: Differences in SPP1 expression between lung adenocarcinoma and normal tissues across different age groups (21-40 yrs, 41-60 yrs, 61-100 yrs). B: Differences in SPP1 expression between lung adenocarcinoma and normal tissues across different genders (Male, Female). C: Differences in SPP1 expression between lung adenocarcinoma and normal tissues across different ethnic groups (Caucasian, African-American, Asian). D: Differences in SPP1 expression between lung adenocarcinoma and normal tissues across different TNM stages (Stage1, Stage2, Stage3). * indicates $P < 0.05$.

3.4. The expression of SPP1 gene is related to the treatment outcome of those who have been diagnosed with LUAD

KM Plotter analysis (Figure 3) showed that the median overall survival (OS) and progression-free survival (PFS) of patients with high SPP1 expression were 35.2 months and 18.3 months, respectively, which were much shorter than those with low spp1 expression (62.5 months and 32.6 months). The hazard ratios (HRs) were 1.34 (95%CI: 1.19-1.51, $P = 1.1 \times 10^{-6}$) for OS and 1.57 (95%CI: 1.32-1.86, $P = 2 \times 10^{-7}$) for PFS, suggesting that high SPP1 expression is a significant risk factor for poor prognosis in LUAD patients.

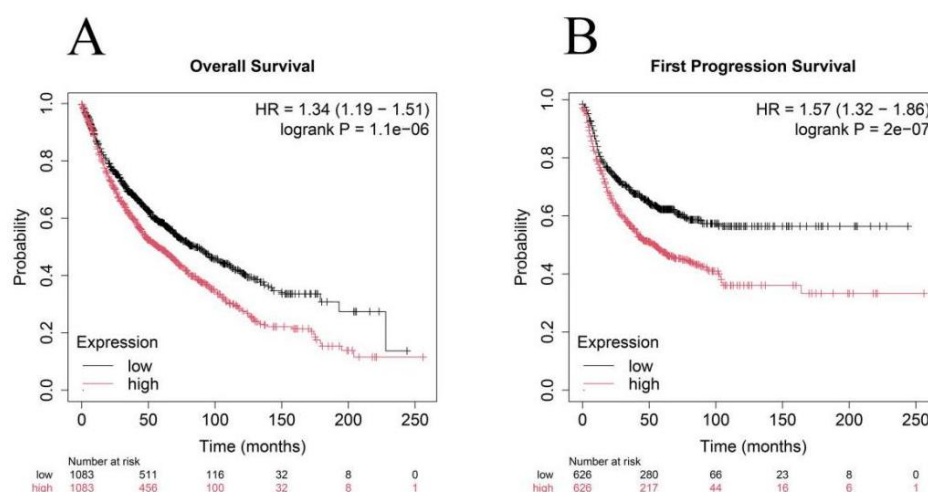


Figure 3. The relationship between the expression of SPP1 and the clinical outcome of lung adenocarcinoma. A: OS survival curve. B: PFS survival curve. HR = hazard ratio, 95% CI = 95% confidence interval.

3.5. SPP1 co-expression gene network analysis

The network constructed by GeneMANIA (Figure 4) identified 50 genes that form close interactions with SPP1, including members of the integrin family (e.g., ITGA4,

ITGA5), the matrix metalloproteinase family (e.g., MMP3, MMP7), and various signal transduction genes (e.g., PIK3CB, AKT1).

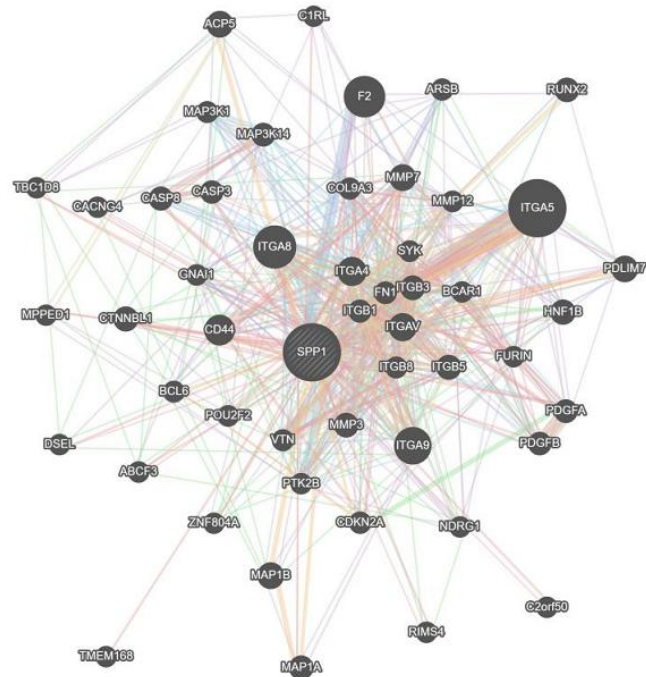


Figure 4. *SPP1* co-expression gene network constructed by GeneMANIA, nodes represent genes, and lines indicate gene interactions.

3.6. KEGG pathway enrichment analysis of *SPP1* and its co-expressed genes

KEGG pathway enrichment analysis (Figure 5) demonstrated significant enrichment of the associated genes in six core pathways ($P<0.05$), including ECM-receptor interaction, integrin signaling, and the PI3K-Akt pathway. Among these, the ECM-receptor interaction and integrin signaling pathways showed the highest degree of enrichment.

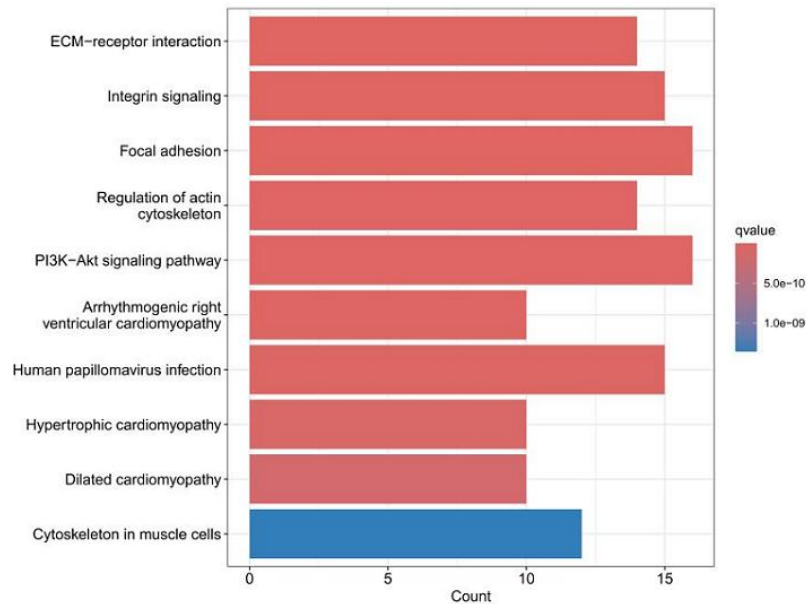


Figure 5. KEGG pathway enrichment map of *SPP1* and co-expressed genes. Drafted based on *SPP1* and 50 co-expressed genes, showing significantly enriched core pathways ($P<0.05$). Bubble

size represents the number of gene enrichments, and the color intensity represents the significance of the enrichment.

4. Discussion

Lung adenocarcinoma is characterized by its high malignancy and poor prognosis, necessitating a major research focus on identifying key molecular targets for clinical intervention [16]. Through a multi-database bioinformatics analysis, this study systematically explores the role and mechanistic pathways of SPP1 in LUAD, aiming to contribute a theoretical foundation for precision diagnostics and therapeutics.

Our findings demonstrate that SPP1 is overexpressed in LUAD. This overexpression is validated across the TIMER, GEPIA, CPTAC, and HPA databases, revealing significantly higher mRNA and protein levels compared to normal tissues [17,18,19]. Furthermore, this expression is positively correlated with clinical advancement, increasing progressively with higher disease stages. This pattern aligns with the established tumor-promoting function of SPP1 in other malignancies, including breast and colorectal cancers, suggesting its potential as a pan-oncogene [18].

Clinical correlation analyses identified a strong association between SPP1 expression and patient age and clinical stage, with the highest expression levels found in patients aged 61-80 and those with Stage III disease. Survival analysis further shows that the increased expression of SPP1 is an important signal, indicating that the prognosis of patients may be worse, which will lead to significantly shortened overall survival (OS) and progression-free survival (PFS) [20]. These results are consistent with the findings of Li et al., and the reliability of our conclusions is strengthened by cross-validation across multiple independent databases[18].

From a mechanistic standpoint, SPP1 exhibits close interactions with genes from the integrin family, the matrix metalloproteinase (MMP) family, and components of the PI3K-Akt signaling pathway [21]. Gene set enrichment analysis reveals that these SPP1-associated genes are predominantly enriched in critical pathways such as ECM-receptor interaction and integrin signaling [22]. We propose a model wherein SPP1 binds to integrin receptors to activate the downstream PI3K-Akt pathway. This activation subsequently upregulates MMP expression, leading to extracellular matrix (ECM) degradation, enhanced tumor cell invasion and metastasis, and suppression of apoptosis. This proposed mechanism demonstrates remarkable conservation with the tumor-promoting roles of SPP1 observed in other cancer types [22].

This study also has certain limitations: (1) This study is a retrospective bioinformatics analysis, and all data are sourced from public databases, which may lead to data bias and lacks actual detection and verification with clinical samples; (2) The specific functions and regulatory mechanisms of SPP1 were not validated through in vitro and in vivo experiments. Subsequent studies need to further confirm the effects of SPP1 on LUAD cell proliferation, invasion, and metastasis through cell experiments (such as siRNA interference or SPP1 overexpression) and animal experiments (such as a nude mouse xenograft model); (3) The role of SPP1 in the LUAD tumor microenvironment was not analyzed, for example, the correlation between SPP1 and immune cell infiltration, which is also an important direction for future research.

5. Conclusions

In summary, this study establishes that SPP1 is overexpressed in LUAD and correlates with advanced patient age and clinical stage, as well as a poor prognosis. The upregulation of SPP1 appears to drive tumor progression, potentially through its regulation of pathways like ECM-receptor interaction. This study found that SPP1 can predict the patient's condition well and provide a new way for the treatment of lung adenocarcinoma. Future work should focus on validating these discoveries in clinical samples and conducting in-depth mechanistic studies using both in vitro and in vivo models to facilitate the eventual clinical translation of these findings.

Conflicts of Interest

The author declares that there is no conflict of interest regarding the publication of this article.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

- [1] Bray, F. et al., Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, 2024. 74(3), 229-263.
- [2] Sung, H. et al., Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*, 2021. 71(3), 209-249.
- [3] Travis, W.D. et al., The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. *J Thorac Oncol*, 2015, 10(9), 1243-1260.
- [4] Chen, W. et al., Cancer statistics in China, 2015. *CA Cancer J Clin*, 2016, 66(2), 115-32.
- [5] Scatena, M.; L. Liaw.; C.M. Giachelli, Osteopontin: a multifunctional molecule regulating chronic inflammation and vascular disease. *Arterioscler Thromb Vasc Biol*, 2007, 27(11), 2302-9.
- [6] O'Regan, A.; J.S. Berman, Osteopontin: a key cytokine in cell-mediated and granulomatous inflammation. *Int J Exp Pathol*, 2000. 81(6), 373-90.
- [7] Kumari, A., D. Kashyap, and V.K. Garg, Osteopontin in cancer. *Adv Clin Chem*, 2024. 118, 87-110.
- [8] Shen, X.Y. et al., Genome-wide analysis reveals alcohol dehydrogenase 1C and secreted phosphoprotein 1 for prognostic biomarkers in lung adenocarcinoma. *J Cell Physiol*, 2019, 234(12), 22311-22320.
- [9] Li, T. et al., TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res*, 2017, 77(21), e108-e110.
- [10] Tang, Z., et al., GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*, 2017, 45(W1), W98-w102.

- [11]Savage, S.R. et al., Pan-cancer proteogenomics expands the landscape of therapeutic targets. *Cell*, 2024. 187(16), 4389-4407.e15.
- [12]Digre, A.; C. Lindskog, The human protein atlas-Integrated omics for single cell mapping of the human proteome. *Protein Sci*, 2023, 32(2), e4562.
- [13]Györfy, B. et al., An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res Treat*, 2010. 123(3), 725-31.
- [14]Franz, M., et al. GeneMANIA update 2018. *Nucleic Acids Res*, 2018. 46(W1), W60-w64.
- [15]Kanehisa, M.; S. Goto, KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*, 2000, 28(1), 27-30.
- [16]Wei, X., et al., Regulation of Ferroptosis in Lung Adenocarcinoma. *Int J Mol Sci*, 2023, 24(19).
- [17]Yi, X. et al., SPP1 facilitates cell migration and invasion by targeting COL11A1 in lung adenocarcinoma. *Cancer Cell Int*, 2022, 22(1), 324.
- [18]Liu, Y. et al., A pan-cancer analysis of the oncogenic role of secreted phosphoprotein 1 (SPP1) in human cancers. *Ann Transl Med*, 2022. 10(6), 279.
- [19]Li, B. et al., Overexpression of SPP1 is a prognostic indicator of immune infiltration in lung adenocarcinoma. *Aging (Albany NY)*, 2024. 16(3), 2953-2977.
- [20]Tang, H. et al., Upregulation of SPP1 Is a Marker for Poor Lung Cancer Prognosis and Contributes to Cancer Progression and Cisplatin Resistance. *Front Cell Dev Biol*, 2021, 9, 646390.
- [21]Zeng, B. et al., SPP1 promotes ovarian cancer progression via Integrin β 1/FAK/AKT signaling pathway. *Onco Targets Ther*, 2018, 11, 1333-1343.
- [22]Dong, B. et al., Macrophage-Related SPP1 as a Potential Biomarker for Early Lymph Node Metastasis in Lung Adenocarcinoma. *Front Cell Dev Biol*, 2021, 9, 739358.



© 2026 by the author(s); licensee International Technology and Science Publications (ITS), this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (<http://creativecommons.org/licenses/by/4.0/>)