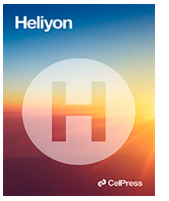
Heliyon 9 (2023) e21328

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



# Thymidine kinase 1 as a target is regulated by the hsa-let-7b-5p/LINC00665 axis and affects NSCLC prognosis

Xu-Dong Zhu [a], Yong-Fei Fan [a], Yi Zhao [a], Xue-Yu Song [a], Xiang-Sen Liu [a], Zhao-Jia Gao [a,b], Kai Yuan [a,b,*]

[a] *Department of Thoracic Surgery, The Affiliated Changzhou No. 2 People's Hospital of Nanjing Medical University, Changzhou, 213003, China*
[b] *Heart and Lung Disease Laboratory, The Affiliated Changzhou No. 2 People's Hospital of Nanjing Medical University, Changzhou, 213003, China*

ARTICLE INFO

*Keywords:*
LINC00665
let-7b-5p
*TK1*
ceRNA
NSCLC

ABSTRACT

*Background:* In the past, multiple studies have offered incremental evidence that indicates that competitive endogenous RNA (ceRNA) regulatory networks are involved in tumor growth and present novel therapeutic targets. Herein, we investigated the impact of thymidine kinase 1 (*TK1*)-related ceRNA networks on the prognosis of non-small cell lung cancer (NSCLC).
*Methods: TK1* expression data in NSCLC and normal tissue samples were retrieved from the Cancer Genome Atlas (TCGA) database and were then compared. Thereafter, the findings of the immunohistochemical staining experiments and clinical follow-up data derived from patients with NSCLC were used for conducting prognostic analysis. The starBase database was searched to determine *TK1*-targeted microRNAs and long non-coding RNAs, and clinical data from TCGA were used for survival analysis to construct a ceRNA network associated with *TK1* expression and prognosis. Finally, the roles played by methylation and immunity in the prognosis and treatment of NSCLC were analyzed.
*Results:* Our findings revealed that the cancer tissues expressed significantly higher *TK1* levels than normal tissues, and the follow-up clinical data revealed that the prognosis was generally worse in the high-expression patients than in the low-expression patients. In addition, clinical data collected from the starBase and TCGA databases showed that the LINC00665/has-let-7b-5p/*TK1* network could influence the growth and prognosis of NSCLC. It was also noted that the *TK1* methylation site was correlated with the prognosis of NSCLC, and immunoprognostic analysis further indicated that patients with higher *TK1* expression levels displayed a worse prognosis.
*Conclusion:* When the regulatory network of LINC00665/has-let-7b-5p/*TK1* was assessed, it was observed that elevated *TK1* levels may affect the prognosis of NSCLC. Therefore, it could be considered a prognostic biomarker and a probable therapeutic target for predicting NSCLC prognosis.

## 1. Introduction

Lung cancer is a deadly cancer that affects people across the globe. Non-small cell lung cancer (NSCLC), including lung squamous

* Corresponding author. The Affiliated Changzhou Second People's Hospital of Nanjing Medical University, Changzhou Second People's Hospital, Changzhou Medical Center, Nanjing Medical University, No. 29 Xinglong Lane, Changzhou, 213003, Jiangsu Province, China.
*E-mail address:* yuankai1978@163.com (K. Yuan).

https://doi.org/10.1016/j.heliyon.2023.e21328
Received 24 May 2023; Received in revised form 14 October 2023; Accepted 19 October 2023
Available online 28 October 2023

cell carcinoma and lung adenocarcinoma, is a major subtype of lung cancer, with a diagnosis rate of around 80–85 % and an extremely poor prognosis [1]. The development of the diagnosis and multimodal treatment strategies has moderately improved the prognosis of patients with NSCLC [2,3]. However, the existing treatment modalities have failed to achieve better outcomes among patients affected by recurring and metastatic NSCLC [4]. Therefore, it has become imperative to investigate the pathological mechanisms of NSCLC from a genetic perspective and explore new treatment modalities.

Long non-coding RNAs (lncRNAs) are described as non-coding RNA molecules that are >200 nucleotides [5]. They can regulate gene expression through modification, transcription, and translation. Advancements in the research strategies indicated that the lncRNAs played a vital role in malignant pathological processes such as tumorigenesis, metastasis, and drug resistance [6]. For example, Lin et al. discovered that the regulatory network of SNHG16/miR-520a-3p/EphA2 could promote the development of NSCLC [7]. Unlike mRNAs, which are large molecules with hundreds of nucleotides, microRNAs (miRNAs) are small non-coding RNAs that assist in the degradation and translation of mRNAs [8].

Currently, several studies have demonstrated that lncRNAs can target the binding miRNAs to regulate mRNAs and control tumor development. This class of transcripts can be regulated at the transcriptional level as they compete for shared miRNAs, and are defined as competitive endogenous RNA (ceRNA) [9,10]. The findings of these studies showed that this network links the functions of protein-coding mRNA and non-coding RNA to regulate the transcription process *via* different reaction elements. The ceRNA activity is highly susceptible to certain factors, such as subcellular localization, binding affinity of miRNA to its sponges, secondary structure of RNA, etc. After these factors become anomalous, they can further lead to human diseases, including cancer.

Thymidine kinase 1 (*TK1*) is a cell cycle-dependent enzyme found in the cytoplasm. It catalyzes the production of thymidine monophosphate [11], which is an essential precursor for DNA synthesis in cancer cells. Many studies have shown that *TK1* has a promoting effect on cancers such as bladder cancer and hepatocellular carcinoma [12]. However, there is a paucity of research that highlights the effect of *TK1* in NSCLC. Therefore, further analysis needs to be conducted to determine the role of *TK1* in NSCLC and its related ceRNAs.

This research aimed to explore the *TK1*-related ceRNA regulatory network in NSCLC and preliminarily determine the mechanisms associated with its effect on the progression and prognosis of NSCLC [13].

## 2. Materials & methods

### 2.1. Data collection and processing

In this study, we examined the differences in the *TK1* expression levels between tumor and healthy tissues across 33 cancers (samples = 11093) using data from the Xiantao Academic website (http://www.xiantao.love/). We then obtained mRNA (mRNA-seq) sequencing data from the Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/) for 1145 NSCLC samples (tumor = 1037, normal = 108) and miRNA (mRNA-seq) sequencing data for 1090 NSCLC samples (tumor = 999, normal = 91). All miRNA sequences in the raw data were transformed into common miRNA names using the starBase 2.0 database (http://starbase.sysu.edu.cn). Thereafter, *TK1* expression in NSCLC was first validated using the HPA database (http://www.proteinatlas.org/). Furthermore, the data presented in the cBioPortal database (http://www.cbioportal.org/) indicated that the genetic mutation and *TK1* expression levels were correlated among patients with NSCLC.

### 2.2. Fabrication of tissue microarrays

We fabricated two tissue microarrays (TMAs): TMA1 and TMA2. TMA1 contained 57 pairs of cancerous and paracancerous tissues (39 males and 18 females) obtained from Superbiotek (Shanghai, China). Inclusion criteria were T1aN0M0 to T4N0M1c tumor stage of NSCLC as defined by the WHO in 2004. TMA2 contained tissue samples collected from 140 patients with NSCLC and 10 normal patients (38 females and 112 males). All patients with NSCLC underwent surgical resection of the tumor between January and December 2005 at Dept. of Thoracic Surgery, Zhongshan Hospital, Fudan University (stage: Ia–IIIa; American Joint Committee on Cancer and Union for International Cancer Control criteria), and were followed-up until July 2013.

### 2.3. Immunohistochemical staining and quantification

Standard immunohistochemical staining was performed on TMA1 and TMA2 tumor tissue samples and normal tissue specimens. For this purpose, the tissue samples were embedded in paraffin blocks, and the solidified paraffinized specimens were sliced into 4-μm sections and mounted onto slides for baking, deparaffinizing, and hydrating [14]. After inactivating endogenous peroxidase activity with 1 mL of $NaN_3$ and 200 mL of 3 % $H_2O_2$, the antigen was recovered with sodium citrate buffer (10 mM; pH 6.0). These samples were incubated with 10-mM tris-buffered saline supplemented with 4 % normal rabbit serum (ProteinTech Group, Inc., China) for 60 min at room temperature. Then, these slides were subjected to overnight incubation at 4 °C, with primary antibodies against *TK1* (1:50, 20299-1-AP, ProteinTech Group, Inc.). Then, secondary antibodies (1:200, K5007, DAKO, China) were developed at 37 °C for 35 min [15]. In the final step, these samples were weakly re-stained using hematoxylin at 37 °C and dehydrated. Then, a quantitative analysis was carried out using Image J22 [16]. We transformed the selected stained areas of the experimental images into average optical density (AOD) values and plotted the AOD values in TMA1 as paired analysis of variance and the AOD values in TMA2 as unpaired analysis of variance with the GraphPad Prism software ($P<0.05$; log-rank test).

### 2.4. Construction of a prognostic model and analysis of overall survival

First, the obtained AOD values of *TK1* immunohistochemistry with the prognostic survival times of the collected patients were compared and the survival curves were plotted. Next, additional data regarding the survival time and survival patterns of patients with NSCLC were collected from the TCGA database. Correlation analysis was performed using the obtained data. *TK1* expression was analyzed using factors such as TNM stage, pathological type, gender, age, and smoking habit ($P<0.05$). The association between *TK1* and overall survival in the ceRNA network was determined using univariate and multivariate Cox regression analysis, and *P*-values $<0.05$ were deemed as statistically significant.

### 2.5. Constructing the ceRNA network in patients with NSCLC

Depending on the ceRNA network settings, the next step involves the construction of a ceRNA regulatory network. Firstly, the starBase database (http://starbase.sysu.edu.cn/) was examined, the "miRNA-mRNA" module was selected, and the programNum$\geq$2 condition was used for acquiring the miRNAs that bind to *TK1*. Secondly, miRNAs that were negatively correlated with mRNA and differed between tumor and normal tissues were filtered by cor $< -0.2$, $P<0.001$, logFC$<0$, and diffP-value$<0.05$. Thirdly, the "miRNA-mRNA" co-expression network was constructed with the Cytoscape software. Similarly, the "miRNA-lncRNA" module of the starBase database was used for filtering the lncRNAs (programNum $\geq$ 2) that bind to miRNAs, while the lncRNAs that were negatively correlated with miRNAs and differed between the tumor and normal tissues were filtered by cor $< -0.2$, *P*-value$<0.001$, logFC$>0$, and diffP-value$<0.05$. Finally, the lncRNA-miRNA-mRNA co-expression network was constructed with the Cytoscape software.

### 2.6. Exploring the ceRNA network with clinical significance

The correlation between lncRNAs in the regulatory network and *TK1* was first investigated with the aid of the "limma" package in the "R" software (cor$>0.2$, $P<0.001$). The eligible lncRNAs were further analyzed to determine the differences (diffP-value$<0.05$, logFC$>0$; *t*-test). The survival differences in miRNA samples with high and low *TK1* expression levels ($P<0.05$) were determined using log-rank tests.

Next, the LNCipedia database (https://lncipedia.org/) was accessed to identify the expressed sequences of all relevant lncRNAs, and these sequences were copied and pasted into the lncLocator website to determine the cellular distribution of relevant lncRNAs. Finally, the starBase database was employed to anticipate base pairing within the relevant ceRNAs.

### 2.7. Methylation and expression analysis of TK1

Considering the effect of methylation on tumor growth, we further analyzed the *TK1*-related methylation levels. Firstly, the UALCAN database (http://ualcan.path.uab.edu/) was investigated to assess the *TK1* methylation levels in NSCLC and normal tissues. The MEXPRESS (https://mexpress.be) database was used to find the possible *TK1* and DNA methylation sites (***$P<0.001$, **$P<0.01$, *$P<0.05$).

### 2.8. Immune infiltration levels of TK1 in NSCLC

Herein, the TIMER database (https://cistrome. shinyapps.io/timer/) was employed to further investigate the relationship between the *TK1* expression and immune infiltration levels. The tumor-infiltrating immune cells (B cells, neutrophils, $CD8^+$ T cells, $CD4^+$ T cells, macrophages, and dendritic cells) were analyzed, and immune infiltrating cells that exhibited a correlation with *TK1* were then identified. Further, survival analysis was used to analyze the immune-infiltrating cells that were related to the NSCLC prognosis ($P<0.05$).

### 2.9. Ethical statement

This study was ethically reviewed and implemented by the Research Ethics Committee of The Affiliated Changzhou No.2 People's Hospital of Nanjing Medical University. During clinical tissue collection, the participants were asked to voluntarily and knowingly sign the informed consent form. All tissue samples from patients were anonymously processed in compliance with ethical and legal requirements.

### 2.10. Statistical analysis

We analyzed all obtained data using GraphPad Prism (ver. 9.1) and R (ver. 4.1) software. Various statistical analyses were completed using the *t*-test, univariate Cox analysis, multivariate Cox analysis, and log-rank test. *P*-values $<0.05$ were considered statistically significant.

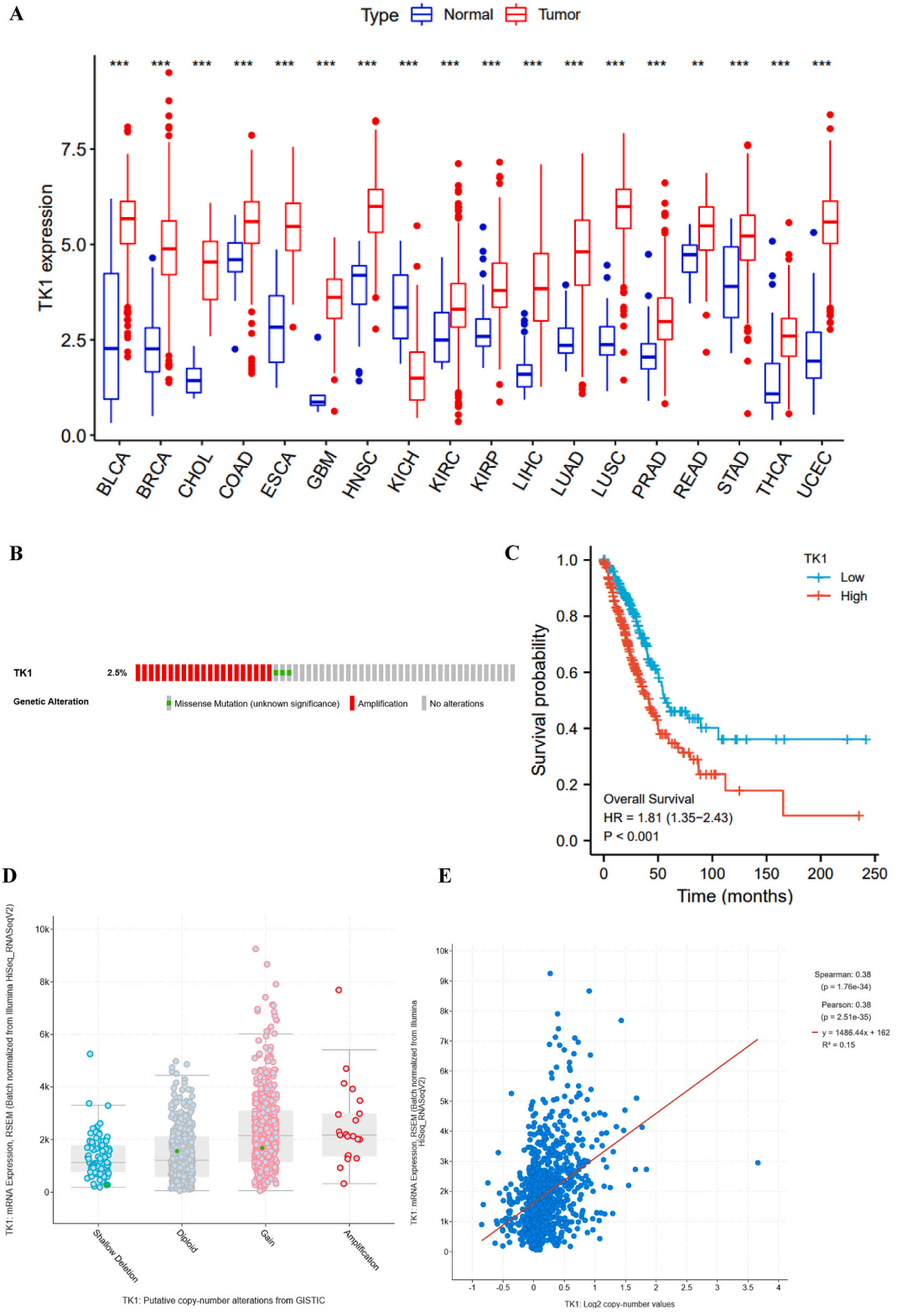
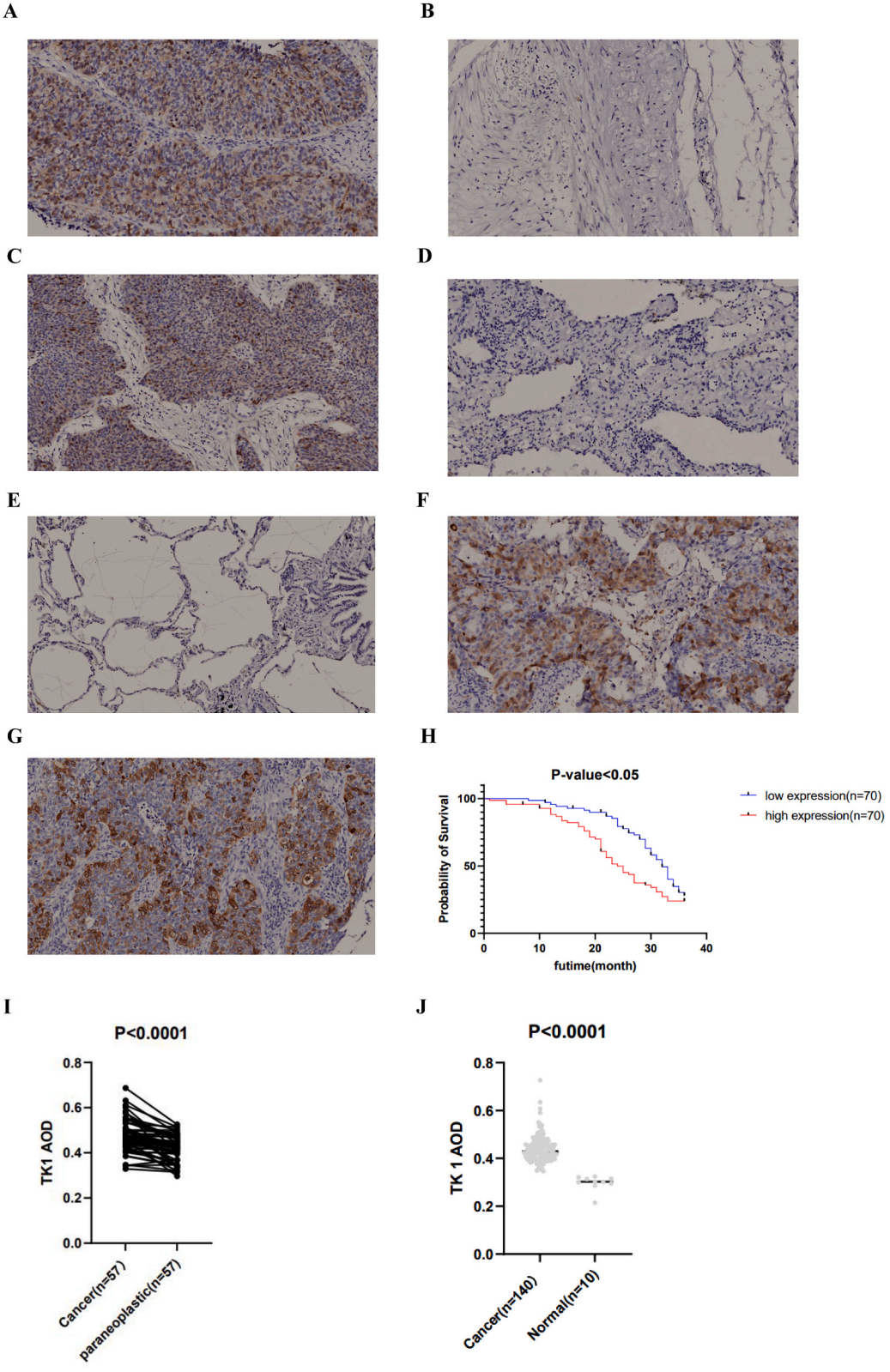



**Fig. 1.** *TK1* expression in human NSCLC. **(A)** Distribution of *TK1* expression in pan-cancerous tissues (*: *P*-value<0.05; **: *P*-value<0.01; ***: *P*-value<0.001). **(B)** Distribution of *TK1* gene changes in TCGA NSCLC shown on biogenetic plots. **(C)** Comparison of Kaplan-Meier survival curves for *TK1* low expression (n = 108) and high expression (n = 1037). **(D)** Dot plot of *TK1* copy number versus mRNA expression. **(E)** Correlation plot between *TK1* copy number and mRNA expression.




(caption on next page)

**Fig. 2.** Immunohistochemical staining and quantitative analysis of *TK1* in NSCLC. (A) Lung adenocarcinoma tissue (×400). (B) Lung adenocarcinoma paraneoplastic tissue (×400). (C) Squamous lung cancer tissue (×400). (D) Squamous lung cancer paraneoplastic tissue (×400). (E) Normal lung tissue (×400). (F) Lung adenocarcinoma tissue (×400). (G) Squamous lung cancer tissue (×400). (H) Survival analysis of *TK1* in TMA2. (I) Paired differential analysis of *TK1* in TM1. (J) Unpaired differential analysis of *TK1* in TMA2.

## 3. Results

### 3.1. The role and prognostic value of TK1 overexpression in NSCLC

The results of bioinformatic analyses we performed indicated that *TK1* levels differed between the different malignant and normal tissues, while the variability was specifically pronounced in lung squamous cell carcinoma and lung adenocarcinoma (Fig. 1A). To understand the mechanism responsible for high *TK1* expression levels in patients with NSCLC, we further investigated *TK1* amplification (Fig. 1B) levels, determined its correlation with the prognosis of patients with NSCLC, and analyzed its genome and copy number. It was observed that patients with NSCLC showing higher *TK1* expression levels exhibited a worse overall survival (Fig. 1C). Meanwhile, when the copy number of the sample was further amplified, the mRNA expression levels in the sample would saturate (Fig. 1D), so it was concluded that the copy number was also positively correlated with mRNA expression (Fig. 1E).

Based on the immunohistochemical analysis of the TMA1 and TMA2 datasets, *TK1* expression was recorded to be significantly higher in tumor tissue samples compared to paracancerous tissues (Fig. 2A–G). The immunostaining results were converted into AOD values, and paired difference analysis was carried out using the TMA1 findings, while the TMA2 results were used for conducting unpaired difference analysis. The results indicated that the tumor tissue samples showed a significantly higher *TK1* expression level compared to the paracancerous (Fig. 2I; $P<0.05$) and normal tissues (Fig. 2J; $P<0.05$). The median AOD values were used for further categorizing the tumor tissues into low- and high-expression groups and the follow-up data were combined for analyzing the survival differences. The results implied that the high-expression group presented a significantly worse prognosis than the low-expression group (Fig. 2H, $P<0.05$). Finally, we predicted *TK1*-related prognostic risk factors. According to the results of pathological characteristics analysis, both univariate and multivariate analyses indicated that T stage, N stage, pathological stage, and age could serve as probable prognosis-related risk factors ($P<0.05$, Table 1; $P<0.05$, Table 2). Thus, we tentatively concluded that *TK1* may serve as a detection target for NSCLC.

### 3.2. Preliminary construction of the triple regulatory network

Herein, data from the starBase database were used for selecting the most suitable *TK1*-related 16 miRNAs and mapping the co-expression networks using the Cytoscape software (Fig. 3A). Out of the 16 miRNAs that were analyzed, we noted that the has-let-7b-5p expression level was significantly lower in NSCLC tissue samples in comparison to the normal tissues (Fig. 3B; $P<0.05$), and it showed a correlation with the *TK1* expression level (cor $< -0.2$; $P<0.05$; Fig. 3C), suggesting that it could be used as a prognostic biomarker. Next, 49 lncRNAs associated with has-let-7b-5p were identified using the starBase database, and a co-expression network

**Table 1**
Correlation of TK1 expression with clinicopathological characteristics of NSCLC.

| Characteristic | Low expression of TK1 | High expression of TK1 | p |
|---|---|---|---|
| T stage, n (%) | | | <0.001 |
| T1 | 182 (17.6 %) | 107 (10.3 %) | |
| T2 | 264 (25.5 %) | 319 (30.9 %) | |
| T3 | 52 (5 %) | 68 (6.6 %) | |
| T4 | 18 (1.7 %) | 24 (2.3 %) | |
| N stage, n (%) | | | <0.001 |
| N0 | 368 (36.3 %) | 300 (29.6 %) | |
| N1 | 75 (7.4 %) | 151 (14.9 %) | |
| N2 | 55 (5.4 %) | 59 (5.8 %) | |
| N3 | 1 (0.1 %) | 6 (0.6 %) | |
| M stage, n (%) | | | 0.348 |
| M0 | 357 (44.3 %) | 416 (51.7 %) | |
| M1 | 18 (2.2 %) | 14 (1.7 %) | |
| Pathologic stage, n (%) | | | <0.001 |
| Stage I& II | 419 (40.9 %) | 405 (39.5 %) | |
| Stage III& IV | 91 (8.9 %) | 110 (10.7 %) | |
| Gender, n (%) | | | <0.001 |
| Female | 254 (24.5 %) | 163 (15.7 %) | |
| Male | 264 (25.5 %) | 356 (34.3 %) | |
| Age, n (%) | | | 0.287 |
| ≤65 | 213 (21.1 %) | 233 (23.1 %) | |
| >65 | 289 (28.6 %) | 274 (27.2 %) | |
| Smoker, n (%) | | | <0.001 |
| No | 63 (6.2 %) | 30 (3 %) | |
| Yes | 439 (43.4 %) | 479 (47.4 %) | |

**Table 2**
Univariate and multifactorial COX analysis of prognostic factors in NSCLC.

| Characteristics | Total(N) | Univariate analysis | | Multivariate analysis | |
|---|---|---|---|---|---|
| | | Hazard ratio (95 % CI) | P value | Hazard ratio (95 % CI) | P value |
| T stage | 1019 | | | | |
| T1&T2 | 860 | | | | |
| T3&T4 | 159 | 1.889 (1.480–2.412) | <**0.001** | 1.366 (1.006–1.854) | **0.045** |
| N stage | 1000 | | | | |
| N0&N1 | 882 | | | | |
| N2&N3 | 118 | 1.799 (1.372–2.357) | <**0.001** | 1.618 (1.172–2.233) | **0.003** |
| M stage | 792 | | | | |
| M0 | 760 | | | | |
| M1 | 32 | 2.269 (1.439–3.577) | <**0.001** | 1.733 (1.046–2.870) | **0.033** |
| Pathologic stage | 1010 | | | | |
| Stage I | 533 | | | | |
| Stage II&Stage III&Stage IV | 477 | 1.913 (1.566–2.337) | <**0.001** | 1.364 (1.034–1.799) | **0.028** |
| Gender | 1022 | | | | |
| Female | 410 | | | | |
| Male | 612 | 1.164 (0.949–1.428) | 0.145 | | |
| Age | 1006 | | | | |
| ≤65 | 445 | | | | |
| >65 | 561 | 1.265 (1.034–1.548) | **0.022** | 1.353 (1.075–1.702) | **0.010** |
| Smoker | 996 | | | | |
| No | 90 | | | | |
| Yes | 906 | 0.883 (0.617–1.263) | 0.496 | | |

was mapped (Fig. 3D). Hence, we initially selected LINC00665, SLC9A3-AS1, LINC01001, HELLPAR, SNHG4, and TMEM147-AS1 using some filtering conditions (cor < −0.2, $P<0.001$). The initial regulatory network was constructed in this manner (Fig. 3E).

### 3.3. Further analysis of the prognosis-related ceRNA regulatory network

Further analysis was carried out for the preliminary construction of ceRNAs. The results indicated that six lncRNAs were differentially expressed in normal and NSCLC tissues ($P<0.001$, Fig. 4A–F). After excluding the less-expressed HELLPAR, we retained five lncRNAs that were associated with patient survival prognosis (Fig. 5A–E), and the prognosis of hsa-let-7b-5p was also strongly correlated with the expression (Fig. 5F). Furthermore, a ceRNA network with LINC00665/has-let-7b-5p/*TK1* as the regulatory axis (Fig. 6A) was also identified. It was noted that LINC00665 was an NSCLC prognosis-associated biomarker and the cellular localization results showed that LINC00665 levels were higher in the cytoplasm (Fig. 6B). Based on this, we tentatively predicted binding sites among LINC00665, has-let-7b-5p, and *TK1* (Fig. 6C and D). The results of the correlation analyses indicated that *TK1* was negatively related to has-let-7b-5p and was positively associated with LINC00665, while LINC00665 also demonstrated a negative correlation with has-let-7b-5p (Fig. 6E–G, $P<0.05$).

### 3.4. Methylation analysis associated with TK1 in NSCLC

Considering the effect of methylation on tumor growth, we further analyzed the methylation status associated with *TK1* using the UALCAN and MEXPRESS databases. After analyzing the UALCAN database, it was noted that the tumor tissue samples showed a significantly lower *TK1* methylation level than that recorded in normal tissues (Fig. 7A and B). Meanwhile, the analysis of the MEXPRESS database showed that a few of the *TK1*-related methylation sites (Fig. S1 and Fig. S2, $P<0.05$) included cg10456035, cg21519872, cg03291825, cg19066150, cg25069807, cg08115732, cg03291825, and cg15227574. All of these could serve as potential sites affecting NSCLC prognosis.

### 3.5. Analysis of immune infiltration associated with TK1 in NSCLC

The TIMER database was further analyzed to better understand the immune infiltration of *TK1*. A significant correlation was observed between the number of *TK1* copies in immune cell infiltrating regions, such as B cells, $CD8^+$ T cells, macrophages, and dendritic cells (Fig. S3 A–B). Further studies showed that *TK1* expression correlated not only with tumor purity but also with the infiltration of $CD8^+$ T cells and neutrophils (Fig. S3 C–D, $P<0.05$). Finally, the findings of survival analysis implied that the infiltration of dendritic cells and neutrophils might be correlated with the prognosis of patients with NSCLC (Fig. S4 E–F, $P<0.05$). Therefore, though the correlation between *TK1* expression and immune response is not clearly defined, additional studies are still needed to explore the immune microenvironment associated with *TK1*.

## 4. Discussion

Although the prognosis of NSCLC has improved with the development of diagnostic and multi-treatment modalities [17,18], it still

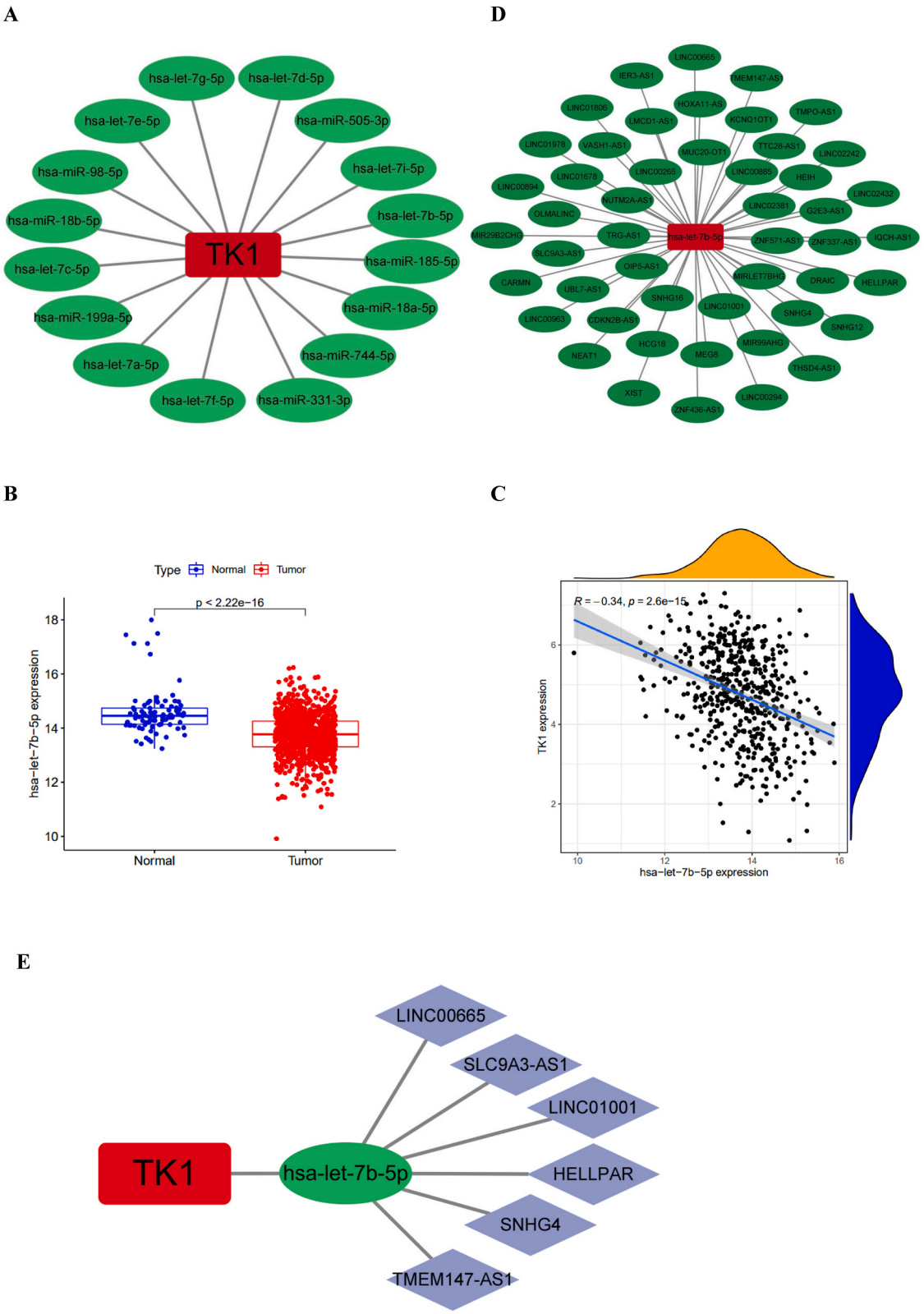



**Fig. 3.** Preliminary construction of lncRNA-miRNA-mRNA triple regulatory network in NSCLC. **(A)** MiRNA network bound to *TK1* from Starbase database. **(B) (C)** Expression and correlation analysis of selected target miRNAs (cor < −0.3, *P*-value<0.001; logFC<0, diffP-value<0.05). **(D)** The lncRNA network bound to has-let-7b-5p in the tarbase database. **(E)** Preliminary constructs of ceRNA regulatory network in NSCLC.

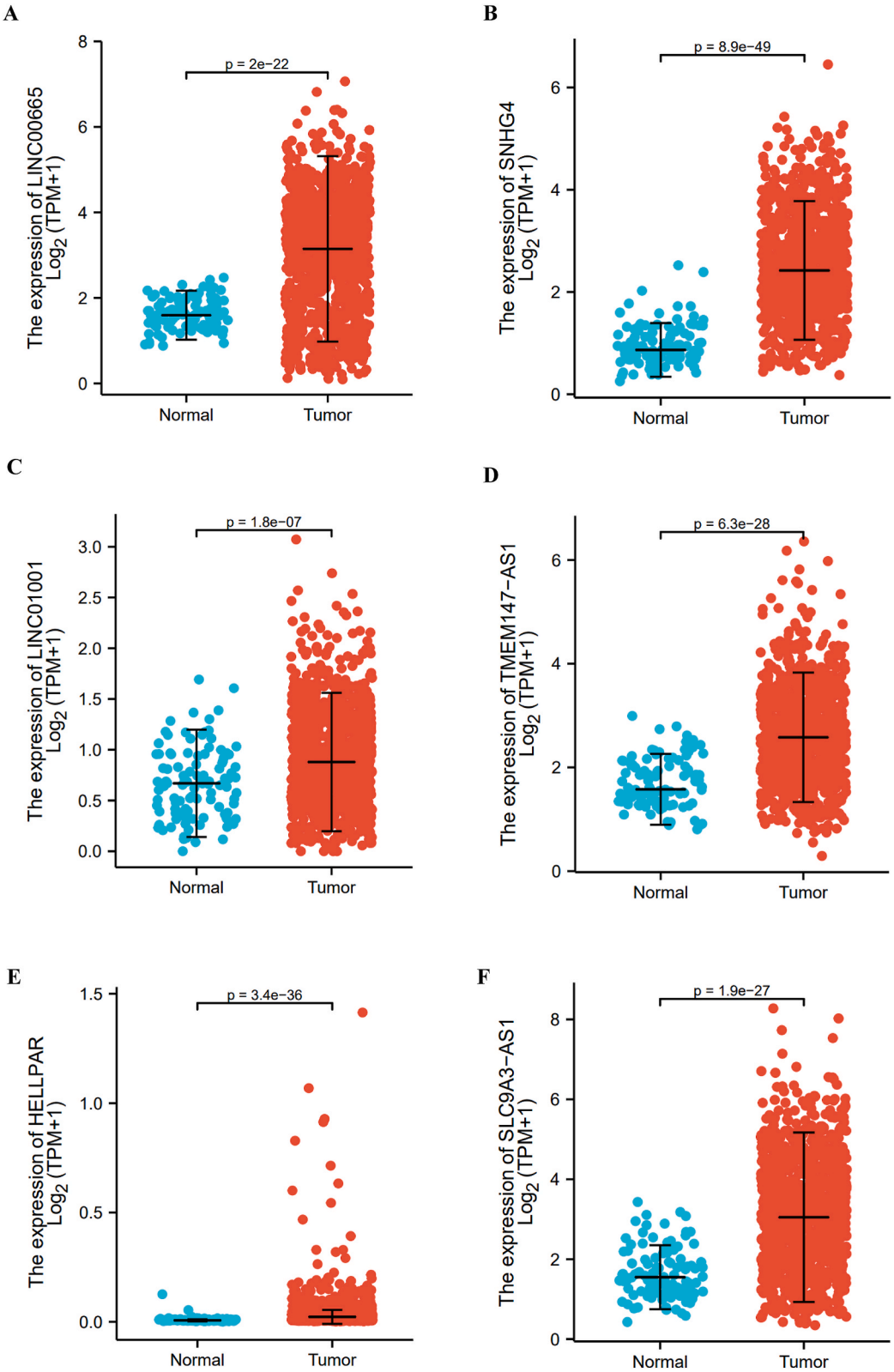



**Fig. 4.** Differences in the expression patterns of six lncRNAs selected from the TCGA NSCLC dataset. **(A–F)** Comparison of the expression patterns of six lncRNAs in normal and NCSCLC tissues.

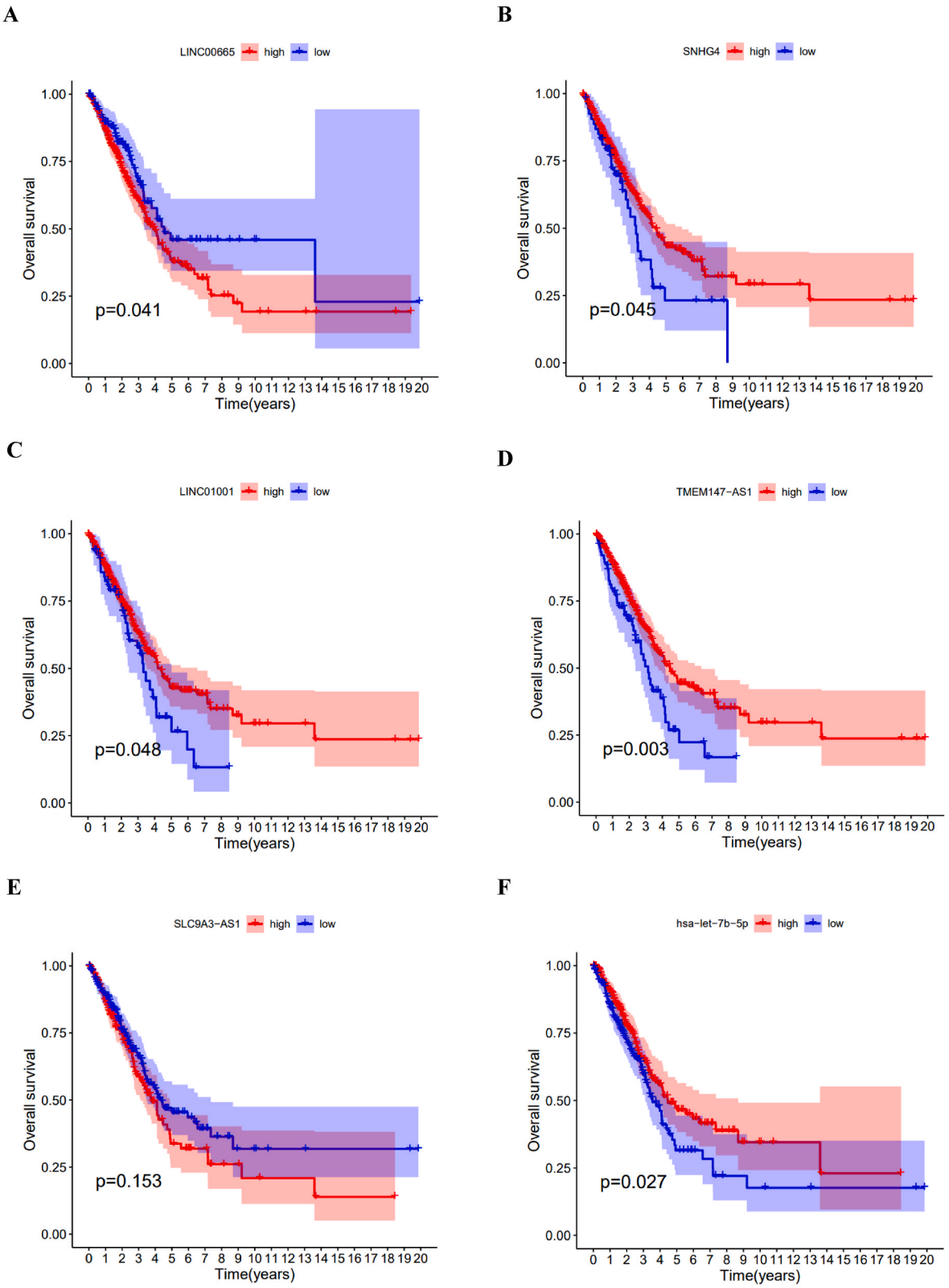



**Fig. 5.** Survival analysis of the five lncRNAs and the targeted miRNA after further screening. **(A–E)** Kaplan-Meier survival curves of screened lncRNAs and **(F)** the targeted miRNA.

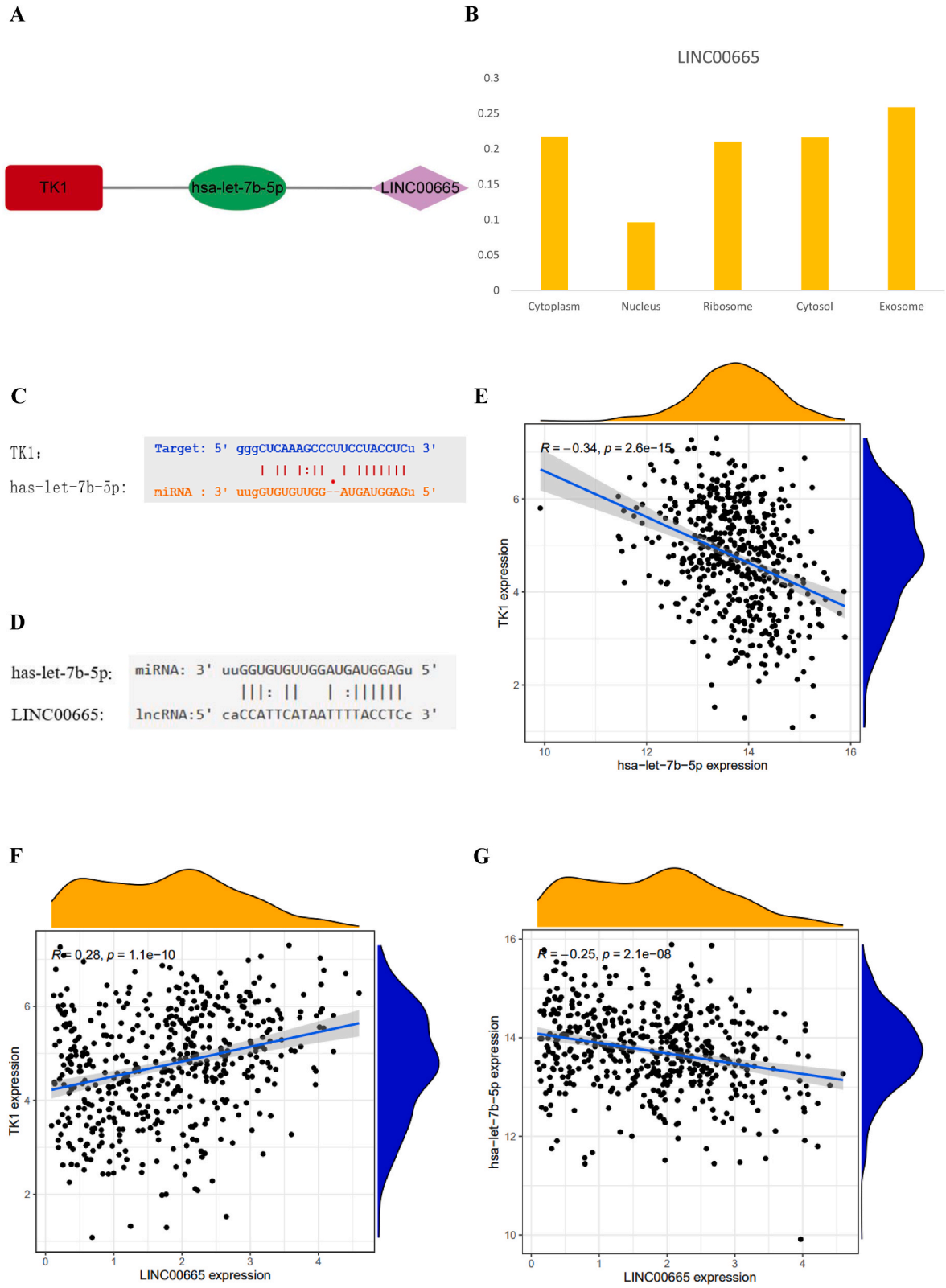



**Fig. 6.** *TK1*-associated ceRNA regulatory network in NSCLC. **(A)** The final constructed ceRNA regulatory network. **(B)** Localization of LINC00665 in NSCLC cells predicted by the LncLocator database. **(C–D)** Base pairing between LINC00665/has-let-7b-5p and *TK1*/has-let-7b-5p predicted by the TargetScan database. **(E–G)** Correlation analysis between three genes in the ceRNA network.

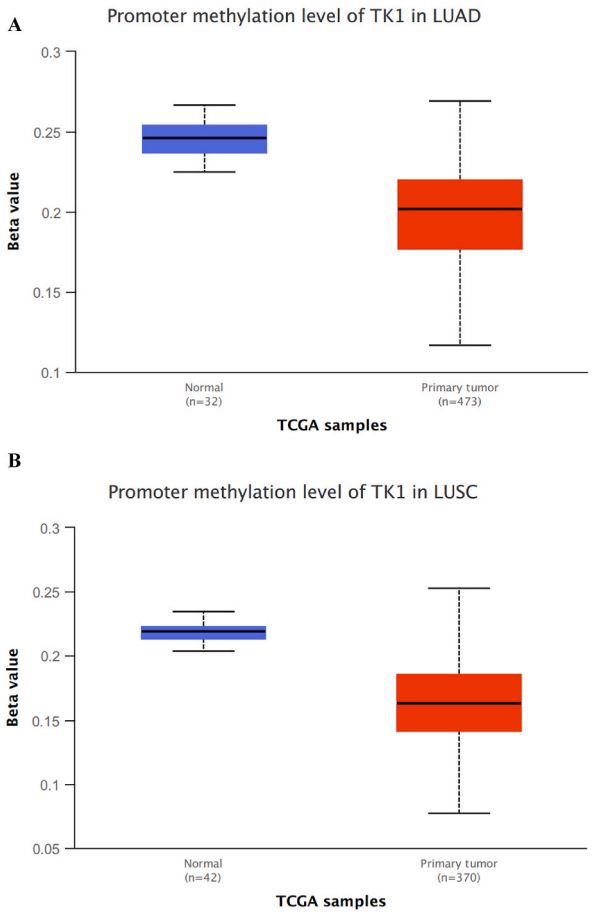



**Fig. 7.** Methylation analysis of *TK1* in NSCLC. **(A)** Comparison of *TK1* methylation expression in LUAD and normal tissues. **(B)** Comparison of *TK1* methylation expression in LUSC and normal tissues. (*: *P*-value<0.05; **: *P*-value<0.01; ***: *P*-value<0.001).

presents the highest mortality rate worldwide [19]. It is particularly important to identify promising biomarkers for improving the prognosis of patients and identifying new therapeutic targets.

In recent years, ceRNA regulatory networks have gained a lot of attention. Several researchers are investigating the association between ceRNA and cancer onset and progression. Pan et al. reported that LINC00339 regulates the miR-148a3p/ROCK1 axis in ovarian cancer tissues, which promotes the proliferation, migration, and invasion of ovarian cancer cells [20]. However, the ceRNA regulatory network has not been studied adequately in NSCLC. Recently, many researchers have started investigating the role of *TK1* in the past few years, and the results have shown that it acts like a regulatory gene in a variety of cancers. On the other hand, its role in NSCLC has not been studied in depth [21]. Therefore, in this study, we strived to explore the role of *TK1* expression in NSCLC for constructing *TK1*-related regulatory networks and determining their correlation with the NSCLC prognosis. Thus, it was observed that *TK1* presented a novel direction for diagnosis and treatment [22].

This study demonstrated that *TK1* expression was significantly upregulated in NSCLC tissues in comparison to normal lung tissues and that it was associated with the poor prognosis of patients with NSCLC. When the results of bioinformatics analysis that were determined in the preliminary stage were combined with the above findings, we hypothesized that *TK1* showed a stimulatory effect on NSCLC. We further acquired the miRNAs and lncRNAs associated with *TK1* and determined their action mechanisms by assessing the databases. Significance analysis and survival analysis allowed us to exclude some interfering genes. Thus, based on our initial results, we constructed an LINC00665/let-7b-5p/*TK1* targeting regulatory axis [23,24]. The statistical data indicated that none of the relevant studies have proposed the relevance of this regulatory network in NSCLC.

The published studies indicated that methylation and immune responses were also involved in gene regulation of tumor growth. The changes in DNA methylation and histone post-translational modifications could help in regulating the specific target genes that control different metastatic phenotypes and are involved in the tumor activity process [25]. In turn, alterations in the immune microenvironment can eventually lead to the development of cancer cells, which antagonizes the cytotoxic function of immune cells and induces immune evasion [26]. Based on database predictions, we noted that the *TK1* gene was hypermethylated in NSCLC tissues in comparison to paracancerous tissues [27]. Our study successfully predicted the correlation between various methylation sites and the prognosis of patients with NSCLC [28]. We also discovered that infiltration of a few immune cells is highly correlated with *TK1*

expression and the prognosis of patients with NSCLC [29]. These findings suggest that *TK1*-induced changes in methylation levels and in the immune environment could affect the prognosis of patients with NSCLC.

The existing studies are in the initial bioinformatics analysis and simple experimental phase. To determine the reliability of the LINC00665/let-7b-5p/*TK*1 regulatory axis, and *TK*1-related alterations in methylation levels and the immune microenvironment, we need to conduct further experiments and clinical trials. Additional studies are needed to present evidence for investigating other related fields.

In conclusion, *TK1* can be used as a prognostic biomarker as it significantly affects the prognosis of NSCLC [30,31]. The related LINC00665/hsa-let-7b-5p/*TK1* regulatory network may also pave the way for future treatment of NSCLC.

## Ethical approval

The Ethics Committee of Changzhou Medical Center, Nanjing Medical University gave its approval to this study (Approval Number: [2022] KY323-01)*.

## Funding

1. Changzhou Sic & Tech Program (Grant number: CZ20220025)
2. Changzhou High-Level Medical Talents Training Project (Grant number: 2022CZBJ069)
3. "333 Project" of Jiangsu Province (Grant number: BRA2020157)
4. 333 High-Level Talent Training Project (Grant number: 2022-2)
5. "Six One Project," Research Projects of High-level Medical Personnel of Jiangsu Province (Grant number: LGY2019025)
6. High-level Talent Selection and Training Project of the 16th Batch of "Six Talent Peak" in Jiangsu Province (Grant number: WSN-245)
7. Medical Scientific Research Foundation of Jiangsu Commission of Health (Grant number: H2018083)

## Data availability statement

The clinical data involved in this study were obtained from publicly available online datasets. The online sites include the following:

Xiantao Academic: https://www.xiantao.love/
TCGA: https://www.cancer.gov/
cBioPortal: https//www.cbioportal.org/
starBase: http://starbase.sysu.edu.cn/
lncLocator: http://www.csbio.sjtu.edu.cn/bioinf/lncLocator/
LNCipedia: https://lncipedia.org/
UALCAN: http://ualcan.path.uab.edu/
TIMER: https://cistrome.shinyapps.io/timer/
MEXPRESS: https://mexpress.be.

Please note that the research data associated with this article is not deposited in the publicly available repository. Please contact us for further research data upon request.

## CRediT authorship contribution statement

**Xu-Dong Zhu:** Data curation, Writing – original draft, Writing – review & editing. **Yong-Fei Fan:** Data curation, Supervision, Validation. **Yi Zhao:** Data curation, Project administration, Validation. **Xue-Yu Song:** Conceptualization, Data curation, Formal analysis. **Xiang-Sen Liu:** Investigation, Methodology, Software. **Zhao-Jia Gao:** Data curation, Investigation, Supervision, Writing – review & editing. **Kai Yuan:** Funding acquisition, Validation, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

We would like to thank Xiantao Academic, TCGA, cBioPortal, starBase, lncLocator, LNCipedia, UALCAN, MEXPRESS, and TIMER open databases. We also thank Bullet Edits Limited for the linguistic editing and proofreading of the manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e21328.

## References


[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, Ca - Cancer J. Clin. 71 (3) (2021) 209–249.

[2] Y. Fan, Y. Zhou, X. Li, M. Lou, Z. Gao, J. Tong, K. Yuan, Long non-coding RNA AL513318.2 as ceRNA binding to hsa-miR-26a-5p upregulates *SLC6A8* expression and predicts poor prognosis in non-small lung cancer, Front. Oncol. 12 (2022), 781903.

[3] M. Jia, Y. Shi, Y. Xie, W. Li, J. Deng, D. Fu, J. Bai, Y. Ma, Z. Zuberi, J. Li, Z. Li, WT1-AS/IGF2BP2 Axis is a potential diagnostic and prognostic biomarker for lung adenocarcinoma according to ceRNA network comprehensive analysis combined with experiments, Cells 11 (1) (2021) 25.

[4] National Lung Screening Trial Research Team, D.R. Aberle, A.M. Adams, C.D. Berg, W.C. Black, J.D. Clapp, R.M. Fagerstrom, I.F. Gareen, C. Gatsonis, P. M. Marcus, J.D. Sicks, Reduced lung-cancer mortality with low-dose computed tomographic screening, N. Engl. J. Med. 365 (5) (2011) 395–409.

[5] S. Yang, Z. Zhang, Q. Wang, Emerging therapies for small cell lung cancer, J. Hematol. Oncol. 12 (1) (2019) 47.

[6] Y. Shi, D.D. Zhang, J.B. Liu, X.L. Yang, R. Xin, C.Y. Jia, H.M. Wang, G.X. Lu, P.Y. Wang, Y. Liu, Z.J. Li, J. Deng, Q.L. Lin, L. Ma, S.S. Feng, X.Q. Chen, X.M. Zheng, Y.F. Zhou, Y.J. Hu, H.Q. Yin, L.L. Tian, L.P. Gu, Z.W. Lv, F. Yu, W. Li, Y.S. Ma, F. Da, Comprehensive analysis to identify DLEU2L/TAOK1 axis as a prognostic biomarker in hepatocellular carcinoma, Mol. Ther. Nucleic Acids 23 (2021) 702–718.

[7] L. Yu, D. Chen, J. Song, LncRNA SNHG16 promotes non-small cell lung cancer development through regulating EphA2 expression by sponging miR-520a-3p, Thorac Cancer 11 (3) (2020) 603–611.

[8] R. Rupaimoole, F.J. Slack, MicroRNA therapeutics: towards a new era for the management of cancer and other diseases, Nat. Rev. Drug Discov. 16 (3) (2017) 203–222.

[9] Y. Li, Z. Yin, J. Fan, S. Zhang, W. Yang, The roles of exosomal miRNAs and lncRNAs in lung diseases, Signal Transduct. Targeted Ther. 4 (2019) 47.

[10] E.A. Braga, M.V. Fridman, A.A. Moscovtsev, E.A. Filippova, A.A. Dmitriev, N.E. Kushlinskii, LncRNAs in ovarian cancer progression, metastasis, and main pathways: ceRNA and alternative mechanisms, Int. J. Mol. Sci. 21 (22) (2020) 8855.

[11] Q. Hua, M. Jin, B. Mi, F. Xu, T. Li, L. Zhao, J. Liu, G. Huang, LINC01123, a c-Myc-activated long non-coding RNA, promotes proliferation and aerobic glycolysis of non-small cell lung cancer through miR-199a-5p/c-Myc axis, J. Hematol. Oncol. 12 (1) (2019) 91.

[12] X. Qi, D.H. Zhang, N. Wu, J.H. Xiao, X. Wang, W. Ma, ceRNA in cancer: possible functions and clinical implications, J. Med. Genet. 52 (10) (2015) 710–718, https://doi.org/10.1136/jmedgenet-2015-103334. Epub 2015 Sep 10. PMID: 26358722.

[13] B. Nisman, H. Nechushtan, H. Biran, H. Gantz-Sorotsky, N. Peled, S. Gronowitz, T. Peretz, Serum thymidine kinase 1 activity in the prognosis and monitoring of chemotherapy in lung cancer patients: a brief report, J. Thorac. Oncol. 9 (10) (2014) 1568–1572.

[14] N.J. Morriss, G.M. Conley, S.M. Ospina, W.P. Meehan Iii, J. Qiu, R. Mannix, Automated quantification of immunohistochemical staining of large animal brain tissue using QuPath software, Neuroscience 429 (2020) 235–244.

[15] Z. Li, J. Zheng, W. Lin, J. Weng, W. Hong, J. Zou, T. Zhang, C. Ye, Y. Chen, Circular RNA hsa_circ_0001785 inhibits the proliferation, migration and invasion of breast cancer cells in vitro and in vivo by sponging miR-942 to upregulate SOCS3, Cell Cycle 19 (21) (2020) 2811–2825.

[16] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis, Nat. Methods 9 (7) (2012) 671–675.

[17] F. Azimi, R.A. Scolyer, P. Rumcheva, M. Moncrieff, R. Murali, S.W. McCarthy, R.P. Saw, J.F. Thompson, Tumor-infiltrating lymphocyte grade is an independent predictor of sentinel lymph node status and survival in patients with cutaneous melanoma, J. Clin. Oncol. 30 (21) (2012) 2678–2683.

[18] N.O. Siemers, J.L. Holloway, H. Chang, S.D. Chasalow, P.B. Ross-MacDonald, C.F. Voliva, J.D. Szustakowski, Genome-wide association analysis identifies genetic correlates of immune infiltrates in solid tumors, PLoS One 12 (7) (2017), e0179726.

[19] M. Evison, AstraZeneca Uk Limited, The current treatment landscape in the UK for stage III NSCLC, Br. J. Cancer 123 (Suppl 1) (2020) 3–9.

[20] L. Pan, Q. Meng, H. Li, et al., LINC00339 promotes cell proliferation, migration, and invasion of ovarian cancer cells via miR-148a-3p/ROCK1 axes, Biomed. Pharmacother. 120 (2019), 109423. TK1.

[21] Q. Shen, Z. Xu, G. Sun, H. Wang, L. Zhang, TFAP4 activates IGF2BP1 and promotes progression of non-small cell lung cancer by stabilizing TK1 expression through m6A modification, Mol. Cancer Res. 20 (12) (2022) 1763–1775, https://doi.org/10.1158/1541-7786.MCR-22-0231. PMID: 36074102.

[22] X. He, M. Wang, Application value of serum *TK1* and PCDGF, CYFRA21-1, NSE, and CEA plus enhanced CT scan in the diagnosis of nonsmall cell lung cancer and chemotherapy monitoring, J. Oncol. 2022 (2022), 8800787.

[23] C. Yang, C. Sun, X. Liang, S. Xie, J. Huang, D. Li, Integrative analysis of microRNA and mRNA expression profiles in non-small-cell lung cancer, Cancer Gene Ther. 23 (4) (2016) 90–97.

[24] K. Wang, W. Huang, R. Chen, P. Lin, T. Zhang, Y.F. Ni, H. Li, J. Wu, X.X. Sun, J.J. Geng, Y.M. Zhu, G. Nan, W. Zhang, X. Chen, P. Zhu, H. Bian, Z.N. Chen, Di-methylation of CD147-K234 promotes the progression of NSCLC by enhancing lactate export, Cell Metab 33 (1) (2021) 160–173.e6.

[25] A. Cock Rada, J.B. Weitzman, The methylation landscape of tumour metastasis, Biol Cell 105 (2) (2013) 73–90, https://doi.org/10.1111/boc.201200029. Epub 2013 Jan 16. PMID: 23198959.

[26] X. Lei, Y. Lei, J.K. Li, W.X. Du, R.G. Li, J. Yang, J. Li, F. Li, H.B. Tan, Immune cells within the tumor microenvironment: biological functions and roles in cancer immunotherapy, Cancer Lett. 470 (2020) 126–133, https://doi.org/10.1016/j.canlet.2019.11.009. Epub 2019 Nov 12. PMID: 31730903.

[27] K. Hyun, J. Jeon, K. Park, J. Kim, Writing, erasing and reading histone lysine methylations, Exp. Mol. Med. 49 (4) (2017) e324.

[28] W.J. Shu, H.N. Du, The methyltransferase SETD3-mediated histidine methylation: biological functions and potential implications in cancers, Biochim. Biophys. Acta Rev. Canc 1875 (1) (2021), 188465.

[29] C. Zhang, S. Ma, X. Hao, Z. Wang, Z. Sun, Methylation status of *TK1* correlated with immune infiltrates in prostate cancer, Front. Genet. 13 (2022), 899384.

[30] Y. Jin, D.L. Chen, F. Wang, C.P. Yang, X.X. Chen, J.Q. You, J.S. Huang, Y. Shao, D.Q. Zhu, Y.M. Ouyang, H.Y. Luo, Z.Q. Wang, F.H. Wang, Y.H. Li, R.H. Xu, D. S. Zhang, The predicting role of circulating tumor DNA landscape in gastric cancer patients treated with immune checkpoint inhibitors, Mol. Cancer 19 (1) (2020) 154.

[31] X. Guo, Y. Zhang, L. Zheng, C. Zheng, J. Song, Q. Zhang, B. Kang, Z. Liu, L. Jin, R. Xing, R. Gao, L. Zhang, M. Dong, X. Hu, X. Ren, D. Kirchhoff, H.G. Roider, T. Yan, Z. Zhang, Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing, Nat. Med. 24 (7) (2018) 978–985.