

# NUF2 Correlates with Poor Prognosis in Non-Small Cell Lung Cancer: A Systematic Multi-Omics Analysis

Mengqing Chen (✉ [cmq2016@outlook.com](mailto:cmq2016@outlook.com))

the Affiliated Hospital of Southwest Medical University <https://orcid.org/0000-0001-5576-2297>

**Yue Zhang**

the Affiliated Hospital of Southwest Medical University

**Yuling Liang**

the Affiliated Hospital of Southwest Medical University

**Dan Luo**

the Affiliated Hospital of Southwest Medical University

**Wenjun Wang**

the Affiliated Hospital of Southwest Medical University

---

## Research article

**Keywords:** non-small cell lung cancer, NUF2, prognosis, immune infiltration

**Posted Date:** August 7th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-52844/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

**Background:** Lung cancer is one of the most common malignant tumors and the leading causes of cancer-related deaths worldwide. As a component of the nuclear division cycle 80 complex, NUF2 is a part of the conserved protein complex related to the centromere. Although the high expression of NUF2 has been reported in many different types of human cancers, no studies have conducted systematic multi-omics analysis on NUF2 in non-small cell lung cancer.

**Methods:** In this analysis, NUF2 difference analysis in non-small cell lung cancer were evaluated in GEO, TCGA, Oncomine, UALCAN databases, and a prognosis analysis of NUF2 based on Kaplan-Meier was performed. R language was used to analyze the differential expression genes, functional Annotation and protein-protein interaction. GSEA analysis of differential expression genes were also carried out. Mechanism analysis about exploring the characteristic of NUF2, multi-omics and correlation analysis were carried out using UALCAN, cBioportal, GEPIA and TIMER.

**Results:** The expression of NUF2 in tumor tissues was significantly higher than that in normal tissues. The analysis of TCGA and UALCAN database samples proved that NUF2 expression was connected with stage, TNM stage and smoking habits. Meanwhile, the overall survival curve also validated that high expression of NUF2 indicated poor prognosis. Furthermore, immune infiltration analysis showed NUF2 had correlation with immune cells and NUF2 altered group had a poor prognosis than unaltered group in non-small cell lung cancer.

**Conclusion:** Our results demonstrated that data mining efficiently reveals NUF2 expression and potential regulatory mechanism in non-small cell lung cancer, laying a foundation for further study of the role of NUF2 in diagnosis and treatment in non-small cell lung cancer.

## Introduction

Lung cancer is one of the most common malignant tumors and the leading causes of cancer-related deaths worldwide. The five-year survival rate of lung cancer patients mainly depends on the stage of the disease and regional differences, with fluctuations ranging from 4–17%[1]. Due to limited treatment options, lung cancer usually does not require further morphological classification. Therefore, in the past few decades, non-small cell lung cancer (NSCLC) and small cell lung cancer(SCLC) are the most commonly used diagnostic terms for lung cancer[2]. NSCLC are approximately 80–85% of lung cancers, which contains approximately 40–50% cases of lung adenocarcinoma(LUAD) and 20–30% cases of lung squamous cell carcinoma(LUSC). Up to present, the research progress of NSCLC has found many target genes that correlated with tumor metastasis, invasion, prognosis, such as EGFR, KRAS, ALK, ROS1 etc. Thus, the target therapy of EGFR, ALK, ROS1 have been the first-line therapy of NSCLC for a long time. Soon after, medicine of PD-1 and PD-L1 have become new treatment solutions in NSCLC. However, due to the poor prognosis, there is an urgent need to discover new target or biomarker for NSCLC.

The nuclear division cycle 80(NDC80) complex is a heterotetrameric protein complex, and NUF2, as an important part of the NDC80 complex, is essential for kinetochore-microtubule attachment and chromosome separation[3]. Previous evidence demonstrated that NUF2 plays as a prognostic biomarker and therapeutic target in hepatocellular carcinoma, breast cancer and oral cancer[4–6], and NDC80 complex gene might be an early indicator of diagnosis and prognosis of lung adenocarcinoma[7]. These studies have shown that NUF2 might be a candidate biomarker and therapeutic target in cancer with great potential, but the role of the NUF2 gene in NSCLC has not been systematically explored in multiple aspects.

In recent years, more and more platforms, databases and various data sets on the Internet have enabled cancer researchers to use multi-omics data to conduct bioinformatics analysis of cancer. In order to get a better understanding of the roles of the NUF2 gene in NSCLC, we performed expression profiling, prognosis valuation, DNA methylation, gene mutation, immune infiltration and clinic pathologic association significance of the NUF2 gene with the public-accessible cancer genomics database. To lay a foundation for further exploring new markers and understanding of the mechanism of the occurrence and development of non-small cell lung cancer.

## Materials And Methods

### Oncomine Database Analysis

The Oncomine database(<https://www.oncomine.org/resource/main.html>) compiles 65 gene chip data sets, 4700 chips and 480 million gene expression data[8]. It is a large and comprehensive database which was used to facilitate data mining. Here, we employed this database to evaluate the transcription level of NUF2 in multiple tumors, especially lung cancer.

### UALCAN Database Analysis

UALCAN(<http://ualcan.path.uab.edu/index.html>) is a comprehensive interactive web platform for analyzing omics data in cancer[9]. It can help researchers to perform biomarker identification, expression profile analysis, survival analysis of related genes, and can also query information in TCGA databases through related links. All in all, it is a simple, fast and effective TCGA data mining and analysis website tool. In this article, we use the database to evaluate the expression profile of NUF2 genes, the significance of DNA methylation and clinicopathological associations.

### GEO and TCGA Database Analysis

To further verify the analysis in Oncomine and UALCAN, we downloaded data for NSCLC from the GEO and TCGA database. We accessed the TCGA data portal to download mRNA expression quantification profiles(HTSeq-FPKM) and the mRNA expression profiles of GSE77803 and GSE32863 were acquired from GEO database. The TCGA dataset and GSE77803 dataset contained adenocarcinoma and squamous cell carcinoma, while the GSE32863 only

contained adenocarcinoma data in order to verify the analysis from Oncomine and UALCAN. Clinical variables, such as stage, age, AJCC TMN cancer stage[10], were used for the assessment of the correlation between the expression of NUF2 and the parameters.

### Kaplan-Meier plotter

Kaplan Meier plotter (KMplotter) is often used for single-gene or multi-gene prognostic analysis of various malignant tumors[11]. We assessed the effect of NUF2 expression on the survival rate in NSCLC and the subtype of it (adenocarcinoma and squamous cell carcinoma). We also further explore the effect of NUF2 expression on the survival rate due to the enriched and decreased immune cells in NSCLC.

### Functional annotation and Protein-protein interaction

Gene ontology(GO) and Kyoto Encyclopedia of Genes and Genomes(KEGG) pathway enrichment analysis were utilized to evaluate the possible functions of the differential expression genes(DEGs) from GSE77803 which grouped by the expression of NUF2. Protein-protein interaction analysis(PPI, <https://string-db.org/>) for NUF2 were performed by STRING database[12]. Molecular Complex Detection(MCODE) from Cytoscape were used to find clusters (highly interconnected regions) in a network.

### Gene-set Enrichment Analysis(GSEA)

In order to further verify the enrichment analysis of the KEGG pathway, gene set enrichment analysis was performed using the GSEA program (v.4.0.3) [13].The data sets used include the H: hallmark gene set and C2.CPKEGG (KEGG gene set) from the Broad Molecular Signatures Database (MSigDB) set. The number of random sample permutations is set to 1,000, and the significance threshold is P<0.05.

### GEPIA database analysis

GEPIA (<http://gepia.cancer-pku.cn/index.html>) is an online comprehensive database that can be used to standardize the analysis of the RNA-seq data of 9,736 tumor samples and 8,587 normal control samples from the TCGA and the Genotype-Tissue Expression(GTEX) datasets[14]. We first use the GEPIA database to analyze the correlation between NUF2 and some special genes, and also use it to understand the correlation between NUF2 expression and immune cell subset markers.

### c-BioPortal analysis

cBio Cancer Genomics Portal (c-BioPortal) (<http://cbioportal.org>), currently including 225 cancer studies, is an open bioinformatics analysis platform that can be used to interactively explore multidimensional cancer genomics data sets[15]. It provides opportunities to explore, visualize and analyze a variety of cancer genomics data. We used c-BioPortal to analyze NUF2 and the components of NDC80 complex alterations in TCGA NSCLC sample. The OncoPrint displays an overview alterations of genes above per sample in TCGA-NSCLC, while the Mutations dispalys mutational site of every gene in detail. We also used the Comparison/Survival to explore the difference between gene altered group and unaltered group.

### TIMER database analysis

TIMER (<https://cistrome.shinyapps.io/timer/>) is a network platform for comprehensive analysis of tumor infiltrating immune cells[16]. It allows the user to input function-specific parameters and dynamically display the result graph to easily obtain the immunological, clinical and genomic characteristics of the tumor. Therefore, in our study, we use it to explore the association of NUF2 expression with immune cell (B cells, CD4 + T cells, NK cells, CD8 + T cells, Th1 cells, neutrophils, Monocyte, macrophages, Treg and dendritic cells, etc) infiltration and the subgroup markers of them.

## Results

### 1. Differential expression of NUF2 in NSCLC

We developed a flow diagram to show our process(Fig. 1). We initially evaluated NUF2 transcription levels between tumor and normal tissues in pan-cancer from Oncomine and TIMER(Fig. 2A), and further analyzed the expression of NUF2 in NSCLC via UALCAN, TCGA, GEO. As shown in Fig. 2A, data in these databases revealed that mRNA expression of NUF2 were significantly higher in NSCLC. Then, we further studied the transcription levels of NUF2 in LUSC and LUAD ( $P<0.01$ ). NUF2 mRNA was significantly over-expressed in TCGA-FPKM (LUSC) and GSE32863(LUAD)( $P<0.01$ )(Fig. 2B).

### 2. The association of NUF2 expression with prognosis and clinicopathological factors in NSCLC

We next explored the correlation between the transcription level of NUF2 with prognosis and clinicopathological factors. Thus, we utilize the Kaplan-meier plotter in order to evaluate whether NUF2 expression relates to the prognosis of NSCLC, to reveal NUF2 high expression to be significantly associated with a poorer OS in NSCLC(OS HR=1.91, 95%CI=1.38-2.65, logrank  $P=7e-05$ ) and its subgroup, like OS in LUAD(OS HR=1.54, 95%CI=1.21-1.97, logrank  $P=0.00047$ ) (Fig. 3Ba-b). But there was no obviously significant connection between the expression of NUF2 and the OS of LUSC(Fig. 3Bc). We further explored the association between NUF2 expression and clinicopathological features. Subgroup analysis of mutiple clinic pathological features of NSCLC in UALCAN showed that the transcription level of NUF2 was significantly higher in NSCLC(both LUSC and LUAD) than normal group based on stage and smoking habits. We also found the transcription level of NUF2 was significant difference based on T and N stage of AJCC TMN cancer stage in NSCLC via TCGA database (Fig. 3A).

### 3. Functional annotation and protein-protein interaction

The DEGs analysis was performed according to the expression of NUF2 in GSE77803. The volcano plot showed 262 genes(red dots) significant positive and 164 genes(green dots) negative correlations with NUF2, and the top 20 significant genes set positively and negatively correlated with NUF2 as shown in the heat map(Fig. 2Ca). To go a step further, we assessed the association of the top 20 positive genes and negative gene in Fig. 2Cc. We next wanted to determine the functional annotation and protein-protein interaction network(PPI network) of the differential expressed genes above (including NUF2) in NSCLC. As shown in Fig. 4Aa, DEGs were mainly enriched in organelle fission and nuclear division via Biological Process(BP) GO annotations, spindle via Cellular Component(CC) GO annotations and ATPase activity via Molecular Function(MF) GO annotations respectively.

KEGG pathway enrichment of NUF2 interactive genes showed that cell cycle was the most enriched pathways(Fig. 4Ab). Among them, the cell cycle had the smallest P-value( $P=2.45e-10$ ) and the largest number of involved consensus genes(count=18). As shown in Fig. 4Ba, there were 22 nodes and 120 edges in the network of PPI via STRING. The vast majority of the nodes were upregulated DEGs in the network. TTK, RRM2, RAD51AP1, PBK, NDC80, MELK, KIF4A, DLGAP5, CEP55, CCNB1, CDC20 have the largest edges in the network(Fig. 4Bc). In addition, we found a significant module via MCODE in the Cytoscape and the most significant pathway in the module was enriched in cell cycle(Fig. 4Bd).

We further used GSEA enrichment analysis to demonstrate these results in MSigDB database. In curated gene sets(C2.CPKEGG), the genes related with NUF2 enriched in cell cycle, pyrimidine metabolism, purine metabolism and DNA replication; while in hallmark gene sets, the genes related with NUF2 enriched in G2M checkpoint, E2F targets, mitotic spindle, PI3K-AKT-mTOR signaling(Fig. 4C).

### 4. Basic characteristic of NUF2 and the correlation of NUF2 with special genes

To understand the characteristic of NUF2, we explored the basic information and mechanism of NUF2 in NSCLC on Genecards and Compartments. As shown in Fig. 4Aa, NUF2 was a protein-coding RNA which located at q23.3 in Chromosome 1. In compartments, Fig. 4Ab showed the subcellular locations of NUF2 and the highest confidence of subcellular locations were nucleus and cytosol. These results helped us to know the functional location of NUF2 in the cells. In order to know the association of NUF2 with some special genes such as EGFR, KRAS, ROS1 and genes related with cell cycle, we explored it by using GEPIA. We found NUF2 was associated with KRAS, EGFR, ROS1, PIK3CA and also related to CDK1/2/4/6, E2F1 which are enriched in cell cycle. NUF2 had a positive correlation with KRAS, EGFR, PIK3CA, CDK1/2/4/6, E2F1; while had a negative correlation with ROS1(Fig. 5B). By analyzing the relationship of NUF2 with cell cycle and tumor-related genes, it showed that NUF2 was involved in tumorigenesis and development.

### 5. NUF2 DNA methylation status in NSCLC

By using the UALCAN website for differential methylation analysis, we found that the promoter methylation level of NUF2 in LUSC was higher than that of normal tissues. In contrast, the promoter methylation level of NUF2 in LUAD was lower than in normal tissues(Fig. 6A). Then, based on different clinical characteristics, we further discovered whether the promoter methylation level of NUF2 was correlated with clinical characteristics. The subgroup analysis results showed the promotor methylation of NUF2 was possibly impact by stage, smoking status and N stage of AJCC TMN cancer stage(Fig. 6B).

### 6. NUF2 alteration in NSCLC

We then used the cBioPortal to determine the types and frequency of NUF2 alterations based on Whole exome sequencing data from NSCLC(data including 42.3% LUSC and 57.7% LUAD) in TCGA. To highlight the role of NUF2 in NSCLC, we also compared NUF2 with NDC80, SPC24, SPC25, which were the component of NDC80 complex. As shown in Fig. 7A, the NDC80 complex was totally altered in 148 of 1144(12.9%) in NSCLC patients, while NUF2 was altered in 92 of 1144(8%) and most of the cases are amplification. The alteration of NUF2 accounts for 62% of the total alteration. To go a step further, we explored the specific alteration of every genes, and we also found that the number of alteration sites in NUF2 was more than others(Fig. 7Ac-f). We next explored the correlation between alteration and the prognosis of NSCLC in cBioportal. As shown in Fig. 7B, the altered group significantly linked with a poorer prognosis in NUF2( $P=0.0407$ ) while selecting genomic profiles as somatic mutations, and we didn't find obvious difference in other genes.

### 7. The transcription level of NUF2 associated with immune cells infiltration and their type markers in NSCLC

We next assessed whether the transcription level of NUF2 in NSCLC was related to immune cell infiltration. TIMER database was used for correlation analysis. In LUSC, the results have found that NUF2 was positively related with tumor purity( $cor=0.388, P=1.25e-18$ ), negatively correlated with CD4 + T cells(partial.cor=-0.213,  $P=2.96e-06$ ) and macrophages(partial.cor=-0.254,  $P=1.88e-08$ ), and weakly negatively correlated with dendritic cells(partial.cor=-0.132,  $P=4.01e-03$ ) and neutrophils(partial.cor=-0.137,  $P=2.83e-03$ ). We also found that in LUAD, NUF2 was negatively correlated with B cells(partial.cor=-0.19,  $P=2.67e-05$ ), CD4 + T cells(partial.cor=-0.178,  $P=8.61e-05$ ), macrophages(partial.cor=-0.107,  $P=1.89e-02$ ), and dendritic cells(partial.cor=-0.133,  $P=3.23e-03$ )(Fig. 8A).

We next want to find out the relationships between the expression of NUF2 and the type markers of various immune cells in NSCLC. The markers of B cells, CD8+ T cells, neutrophils, macrophages, dendritic cells(DCs), NK cells, Th1 cells, Treg, monocyte were tested via TIMER database. Markers of immune cells were considered to investigate further association of NUF2 expression with immune cells. No matter whether the correlation adjusted or not, NUF2 in LUSC was negatively correlated with markers in several immune cells, such as FCRL2, CD19, MS4A1 in B cells; FCGR3B, CEACAM3, SIGLEC5, FPR1, CSF3R, S100A12 in neutrophils; CD68, CD84, CD163, MS4A4A in macrophages; CD209 in dendritic cells; FOXP3, CCR8 in Treg; C3AR1, CD86, CSF1R in monocyte(Table 1). Meanwhile, NUF2 in LUAD was negatively correlated with MS4A1 in B cells, CD8A; CD8B in CD8+T cells; CSF3R, S100A12 in neutrophils, KIR3DL3, NCR1 in NK cells; CSF1R in monocyte. We next used GEPIA to verify the results. In LUSC, correlation results between NUF2 and

markers of macrophages, neutrophils, dendritic cells in GEPIA are similar to those in TIMER. In LUAD, correlation results between NUF2 and markers of B cells, CD8+T cells, NK cells in GEPIA are similar to those in TIMER(Table 2).

### 8. Prognosis analysis of NUF2 expression in NSCLC based on immune cells infiltration

We have demonstrated that the expression of NUF2 have some relationship with the immune cell infiltration in NSCLC, and the expression of NUF2 was also related to the prognosis of NSCLC. So we speculated whether the expression of NUF2 in NSCLC would have influence on the prognosis partly affected by immune cell infiltration. We analyzed the correlation between NUF2 expression and prognosis based on the enrichment of related immune cell in NSCLC. The results revealed that higher expression of NUF2 of LUAD in enriched B cells, enriched CD4+T cells and enriched macrophages(HR=2.01) had a poor prognosis respectively(Fig. 8B). The results in LUSC described that high expression of NUF2 in enriched macrophages have a better prognosis(Fig. 8B).

## Discussion

With the appearance of modern genetics and molecular biology in the 1980s, researchers started to an gradually clear view of the genetic alterations that underlie uncontrolled proliferation in tumor cells[17]. These disadvantage contain (but are not limited to) the high expression of genes which drive or participate in cell cycle progression, such as cyclin D1(CCND1), cyclin-dependent kinase 4 (CDK4) and Cyclin Dependent Kinase 6(CDK6)[18]. NUF2 is a composition of a molecular linker between the kinetochore attachment site and the tubulin subunits within the lattice of the attached plus ends[19]. NUF2 is also an oncogene that is over-expressed on several cancers[5, 20], including lung adenocarcinoma. Comparing with previous study, the NUF2 expression in multiple databases was significantly higher than in normal tissues, and also significantly higher in LUAD and LUSC[7].

In addition, we selected datasets randomly from TCGA and GEO to assess the expression of NUF2 in LUAD and LUSC, and we also found NUF2 was significantly higher than in normal tissues. To summarize the results from prognosis analysis and correlation analysis with clinicopathological factors, the higher expression of NUF2 indicate the poorer prognosis, poorer clinicopathologic stage. To verify the functions of NUF2 in NSCLC, no matter whether in GO, KEGG, PPI and GSEA, NUF2 was enriched in cell cycle in NSCLC. And this was consistent with the idea that the pathological process of tumor was inseparable with the abnormalities in cell cycle. The cyclin-dependent kinase (CDK)-RB-E2F axis was the key transcription mechanism that drives the cell cycle procession. One or more important parts of this axis (cyclins, CDKs, CDK inhibitors and the RB family of proteins) are changed, which occurs in almost all cancers and leads to oncogenic E2F increased activity and uncontrolled proliferation[21]. In our study, we found the association between NUF2 and these genes related to cell cycle procession or oncogenes in the progresss of NSCLC, such as E2F1, CDK1/2/4/6, KRAS, EGFR, ROS1, PIK3CA. And we also found NUF2 was highly expressed in nuclear via Genecards. Therefore, NUF2 maybe inferred to be a key gene involved in the occurrence, development and prognosis of NSCLC.

The variation of the epigenome are thought to indicate the influence of genetic and environmental risk factors on multiple disease. So far, DNA methylation is still the only epigenetic marker that can be stably detected in various samples[22]. The change of normal DNA methylation patterns contain DNA hypomethylation, which occurs pathologically in normally unmethylated regions of the genome, and DNA hypermethylation, which usually occurs in the CpG islands of gene promoters[23]. In DNA methylation analysis, we found NUF2 methylation level was lower in LUAD than in normal tissues.

Surprisingly, we found NUF2 methylation level was higher in LUSC than in normal tissues. We have found the same trend while exploring the association between NUF2 methylation level and clinicopathological features. The difference of NUF2 methylation levels in different types of NSCLC may suggest that the NUF2 gene has different epigenetic regulation in LUSC and LUAD. Since methylation in the promoter region of a gene can inhibit gene expression at the transcription level, the overexpression of NUF2 transcription level in non-small cell lung cancer has been confirmed by multiple studies including our research. As there is no current research on NUF2 methylation, further research was needed to explore the locus of NUF2 gene methylation, and more researches are needed to explore the mechanism between NUF2 expression and NUF2 methylation. Combing the relationship between NUF2 with cell cycle-related genes and lung cancer oncogenes, future studies need to further explore whether the different levels of NUF2 methylation have different effects on the occurrence, development and prognosis between lung adenocarcinoma and lung squamous cell carcinoma, and compare the difference of NUF2 methylation between LUAD and LUSC.

Although previous study have found missense mutations of NUF2 via large-scale tumor sequencing projects[24], there was seldom study to explore the alteration of NUF2 in NSCLC. Another analysis from cBioportal revealed the alteration frequency of NUF2 in NSCLC. The alteration rate of NUF2(8%) was higher than other components of NDC80 complex. And we also found gene amplification was frequently found in NUF2 than other type of alteration. Genome instability was the molecular genetic marker of tumorigenesis. And gene amplification, as the main form of genome instability, plays an important role in the occurrence and development of many human malignant tumors[25, 26]. This result indicated that NUF2 gene was mainly amplified in non-small cell lung cancer, suggesting that NUF2 gene amplification may be related to the occurrence and development of NSCLC. Due to the lack of research on NUF2 gene amplification, more research need to focus on the manifestations of NUF2 gene amplification in NSCLC, and the relationship between NUF2 gene amplification and NSCLC drug resistance or tumor cell escape growth inhibition[27, 28]. Due to more and more gene mutation found in tumor occurrence and metastasis, we assess whether the alteration of NUF2 affects tumor prognosis. And we found the difference between altered group and unaltered group of NUF2 in NSCLC, but no significance in NDC80 and SPC25. This indicated that the NDC80 complex may affect the prognosis of NSCLC via NUF2 mutation.

Another important aspect of this study was that the expression of NUF2 was correlated with the infiltration level of various immune cells in NSCLC. A large number of studies have explored immune cells in malignant tumor tissues to explore and their relationship with tumor occurrence, development and prognosis[29, 30]. Immune cell infiltration in primary lung squamous cell carcinoma have a better prognosis[31]. It was also found that the presence of

high levels of both CD8 + T cells and CD4 + T cells were a significant marker of a better prognosis for patients in NSCLC[32]. We explored the relationship between NUF2 and immune cell infiltration in NSCLC. Our results demonstrate that there is a negative relationships between NUF2 expression level and infiltration level of multiple immune cells in LUAD and LUSC. Moreover, the correlation between NUF2 expression and the marker genes of immune cells implicate the role of NUF2 in regulating tumor immunology in LUSC and LUAD. Through the analysis above, no matter in LUAD or LUSC, NUF2 gene was negatively correlated with macrophages and DCs. These results reveal the potential regulating role of NUF2 in polarization of tumor-associated macrophages(TAM) and DCs. TAMs infiltration was closely related with tumor specific pathological characteristics, such as immunosuppression, vascular lymphangiogenesis, invasion and metastasis, drug resistance, etc[33]. Although DCs constitute a rare immune cell population in multiple tumors, these cells are central for the initiation of antigen-specific immunity and tolerance[34]. Thus, these results also reveal the potential regulating role of NUF2 in cancer immunology. However, the pathway through which NUF2 regulated immune cell infiltration still needed further research. After that, the results of prognostic valuation based on immune cells showed a poorer survival rate of NUF2 high-expression in enriched macrophages both in LUAD and LUSC. These messages reminded us that NUF2 may affect the survival rate of NSCLC via TAM.

## Conclusions

To sum up, NUF2 transcription levels increased significantly in NSCLC. NUF2 was enriched in cell cycle via functional annotation and PPI analysis. In multi-omics analysis, we found the NDC80 complex may affect the prognosis of NSCLC via NUF2 mutation and NUF2 methylation level was lower in LUAD and higher in LUSC. Elevated NUF2 expression was negatively correlated with immune cells infiltration and prognosis of NSCLC, and NUF2 may affect the survival rate of NSCLC via TAM. However, further studies need to be done to assess the diagnostic and therapeutic role of NUF2 in NSCLC.

## Abbreviations

NSCLC: non-small cell lung cancer; SCLC: small cell lung cancer; LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; NDC80: the mitochondrial cycle 80 complex; KMplotter: Kaplan-Meier plotter; AJCC TNM: American Joint Committee on Cancer TNM stage; GO: Gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; PPI: Protein-protein interaction analysis; GSEA: Gene-set Enrichment Analysis; MSigDB: the Broad Molecular Signatures Database set; GTEx: the Genotype-Tissue Expression; c-BioPortal: cBio Cancer Genomics Portal; BP: Biological Process; CC: Cellular Component; MF: Molecular Function; MCODE: Molecular Complex Detection; DCs: dendritic cells; CCND1: cyclin D1, CDK4: cyclin-dependent kinase 4; CDK6: Cyclin Dependent Kinase 6; CDK: cyclin-dependent kinase; TAM: tumor-associated macrophages.

## Declarations

### Acknowledge

Not applicable.

There is no conflict of interests.

### Authors' contributions

Conception and design of the work: MC and WW. Acquisition, analysis and interpretation of data: MC, YZ , YL and DL. Drafting and revising of the article: MC and WW. Final approval of the manuscript and agreement to be accountable for all aspects of the work: All authors.

### Funding

No funding was obtained for this study.

### Availability of data and materials

The datasets analysed during the current study are available in the TCGA and GEO database, <https://portal.gdc.cancer.gov/> and <https://www.ncbi.nlm.nih.gov/>.

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare they have no competing interests.

### Author details

<sup>1</sup>Department of respiratory and critical care medicine, the Affiliated Hospital of Southwest Medical University, Luzhou 646000, Sichuan Province, China

## References

1. Hirsch FR, Scagliotti GV, Mulshine JL, Kwon R, Curran WJ, Wu YL, Paz-Ares L: **Lung cancer: current therapies and new targeted treatments.** *Lancet (London, England)* 2017, **389**(10066):299-311.
2. Siegel RL, Miller KD, Jemal A: **Cancer statistics, 2020.** *CA: a cancer journal for clinicians* 2020, **70**(1):7-30.
3. Suzuki A, Badger BL, Haase J, Ohashi T, Erickson HP, Salmon ED, Bloom K: **How the kinetochore couples microtubule force and centromere stretch to move chromosomes.** *Nature cell biology* 2016, **18**(4):382-392.
4. Thang PM, Takano A, Yoshitake Y, Shinohara M, Murakami Y, Daigo Y: **Cell division cycle associated 1 as a novel prognostic biomarker and therapeutic target for oral cancer.** *International journal of oncology* 2016, **49**(4):1385-1393.
5. Wang Y, Tan PY, Handoko YA, Sekar K, Shi M, Xie C, Jiang XD, Dong QZ, Goh BKP, Ooi LL et al: **NUF2 is a valuable prognostic biomarker to predict early recurrence of hepatocellular carcinoma after surgical resection.** *International journal of cancer* 2019, **145**(3):662-670.
6. Xu W, Wang Y, Wang Y, Lv S, Xu X, Dong X: **Screening of differentially expressed genes and identification of NUF2 as a prognostic marker in breast cancer.** *International journal of molecular medicine* 2019, **44**(2):390-404.
7. Sun ZY, Wang W, Gao H, Chen QF: **Potential therapeutic targets of the nuclear division cycle 80 (NDC80) complexes genes in lung adenocarcinoma.** *Journal of Cancer* 2020, **11**(10):2921-2934.
8. Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Varambally R, Yu J, Briggs BB, Barrette TR, Anstet MJ, Kincaid-Beal C, Kulkarni P et al: **Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles.** *Neoplasia (New York, NY)* 2007, **9**(2):166-180.
9. Chandrashekhar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, Varambally S: **UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses.** *Neoplasia (New York, NY)* 2017, **19**(8):649-658.
10. Kamarajah SK, Frankel TL, Sonnenday C, Cho CS, Nathan H: **Critical evaluation of the American Joint Commission on Cancer (AJCC) 8th edition staging system for patients with Hepatocellular Carcinoma (HCC): A Surveillance, Epidemiology, End Results (SEER) analysis.** *Journal of surgical oncology* 2018, **117**(4):644-650.
11. Győrffy B, Surowiak P, Budczies J, Lánczky A: **Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer.** *PLoS one* 2013, **8**(12):e82241.
12. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P et al: **STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets.** *Nucleic acids research* 2019, **47**(D1):D607-d613.
13. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES et al: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.** *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102**(43):15545-15550.
14. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z: **GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses.** *Nucleic acids research* 2017, **45**(W1):W98-w102.
15. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E et al: **Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal.** *Science signaling* 2013, **6**(269):pl1.
16. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B, Liu XS: **TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells.** *Cancer research* 2017, **77**(21):e108-e110.
17. Petroni G, Formenti SC, Chen-Kiang S, Galluzzi L: **Immunomodulation by anticancer cell cycle inhibitors.** *Nature reviews Immunology* 2020.
18. Harao M, Hirata S, Irie A, Senju S, Nakatsura T, Komori H, Ikuta Y, Yokomine K, Imai K, Inoue M et al: **HLA-A2-restricted CTL epitopes of a novel lung cancer-associated cancer testis antigen, cell division cycle associated 1, can induce tumor-reactive CTL.** *International journal of cancer* 2008, **123**(11):2616-2625.
19. DeLuca JG, Moree B, Hickey JM, Kilmartin JV, Salmon ED: **hNuf2 inhibition blocks stable kinetochore-microtubule attachment and induces mitotic cell death in HeLa cells.** *The Journal of cell biology* 2002, **159**(4):549-555.
20. Obara W, Sato F, Takeda K, Kato R, Kato Y, Kanehira M, Takata R, Mimata H, Sugai T, Nakamura Y et al: **Phase I clinical trial of cell division associated 1 (CDCA1) peptide vaccination for castration resistant prostate cancer.** *Cancer science* 2017, **108**(7):1452-1457.
21. Kent LN, Leone G: **The broken cycle: E2F dysfunction in cancer.** *Nature reviews Cancer* 2019, **19**(6):326-338.
22. Teschendorff AE, Relton CL: **Statistical and integrative system-level analysis of DNA methylation data.** *Nature reviews Genetics* 2018, **19**(3):129-147.
23. Jung G, Hernández-llán E, Moreira L, Balaguer F, Goel A: **Epigenetics of colorectal cancer: biomarker and therapeutic potential.** *Nature reviews Gastroenterology & hepatology* 2020, **17**(2):111-130.
24. Kamburov A, Lawrence MS, Polak P, Leshchiner I, Lage K, Golub TR, Lander ES, Getz G: **Comprehensive assessment of cancer missense mutation clustering in protein structures.** *Proceedings of the National Academy of Sciences of the United States of America* 2015, **112**(40):E5486-5495.
25. Brison O: **Gene amplification and tumor progression.** *Biochimica et biophysica acta* 1993, **1155**(1):25-41.

26. Knuutila S, Björkqvist AM, Autio K, Tarkkanen M, Wolf M, Monni O, Szymanska J, Larramendy ML, Tapper J, Pere H *et al*: **DNA copy number amplifications in human neoplasms: review of comparative genomic hybridization studies.** *The American journal of pathology* 1998, **152**(5):1107-1123.
27. Turke AB, Zejnullah K, Wu YL, Song Y, Dias-Santagata D, Lifshits E, Toschi L, Rogers A, Mok T, Sequist L *et al*: **Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC.** *Cancer cell* 2010, **17**(1):77-88.
28. Li Y, Zou L, Li Q, Haibe-Kains B, Tian R, Li Y, Desmedt C, Sotiriou C, Szallasi Z, Igglehart JD *et al*: **Amplification of LAPTMB and YWHAZ contributes to chemotherapy resistance and recurrence of breast cancer.** *Nature medicine* 2010, **16**(2):214-218.
29. Tobin JWD, Keane C, Gunawardana J, Mollee P, Birch S, Hoang T, Lee J, Li L, Huang L, Murigneux V *et al*: **Progression of Disease Within 24 Months in Follicular Lymphoma Is Associated With Reduced Intratumoral Immune Infiltration.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2019, **37**(34):3300-3309.
30. Melero I, Rouzaut A, Motz GT, Coukos G: **T-cell and NK-cell infiltration into solid tumors: a key limiting factor for efficacious cancer immunotherapy.** *Cancer discovery* 2014, **4**(5):522-526.
31. Ruffini E, Ascoli S, Filosso PL, Lyberis P, Bruna MC, Macrì L, Daniele L, Oliaro A: **Clinical significance of tumor-infiltrating lymphocytes in lung neoplasms.** *The Annals of thoracic surgery* 2009, **87**(2):365-371; discussion 371-362.
32. Hiraoka K, Miyamoto M, Cho Y, Suzuki M, Oshikiri T, Nakakubo Y, Itoh T, Ohbuchi T, Kondo S, Katoh H: **Concurrent infiltration by CD8+ T cells and CD4+ T cells is a favourable prognostic factor in non-small-cell lung carcinoma.** *British journal of cancer* 2006, **94**(2):275-280.
33. Bohn T, Rapp S, Luther N, Klein M, Bruehl TJ, Kojima N, Aranda Lopez P, Hahlbrock J, Muth S, Endo S *et al*: **Tumor immuno-evasion via acidosis-dependent induction of regulatory tumor-associated macrophages.** *Nature immunology* 2018, **19**(12):1319-1329.
34. Wculek SK, Cueto FJ, Mujal AM, Melero I, Krummel MF, Sancho D: **Dendritic cells in cancer immunology and immunotherapy.** *Nature reviews Immunology* 2020, **20**(1):7-24.

## Tables

Table 1

Correlation results between NUF2 and markers of immune cells via TIMER LUSC, lung squamous cell carcinoma; LUAD, lung adenocarcinoma; NK cells, Natural killer cells; Th 1 cells, type I helper T cells; Treg, regulatory T cells; COR, r value of Spearman's correlation; P, p-value; Purity, correlation adjusted by purity; Age correlation adjusted by age.

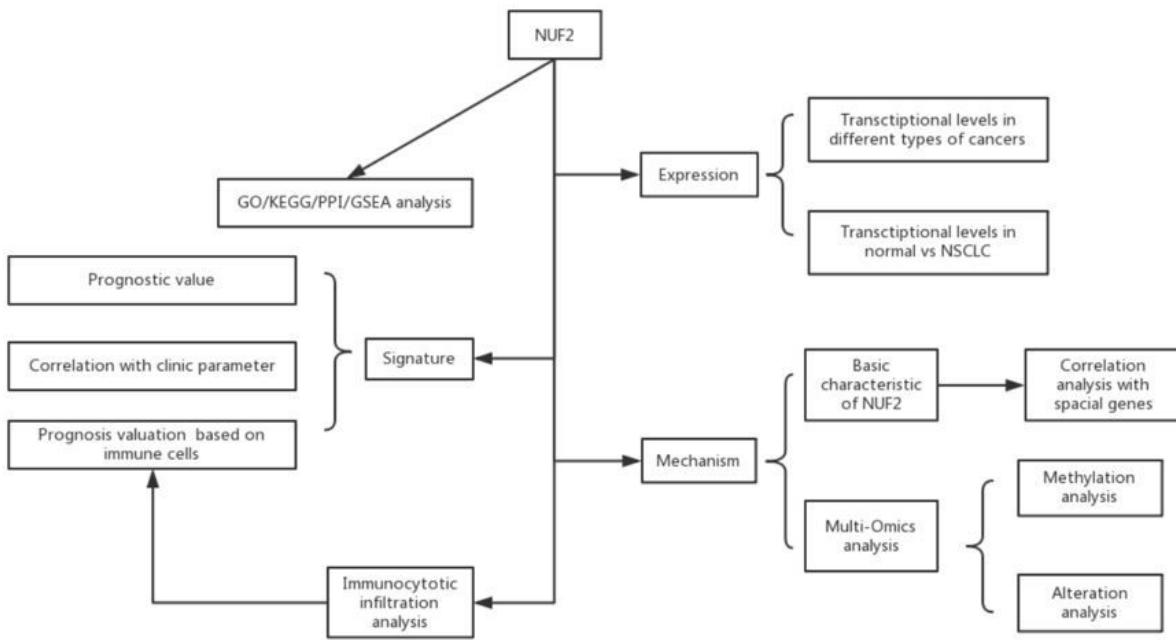
Cell type	Gene	LUAD						LUSC							
		Markers		None		Purity		Age		None		Purity		Age	
		COR	P	COR	P	COR	P	COR	P	COR	P	COR	P	COR	P
B cells	FCRL2	-0.045	0.307	-0.034	0.450	-0.038	0.405	-0.198	8.21E-06	-0.178	7.76E-05	-0.198	9.38E-06		
	CD19	-0.064	0.147	-0.058	0.200	-0.0643	0.158	-0.215	1.25E-06	-0.200	1.02E-05	-0.215	1.58E-06		
	MS4A1	-0.163	0.0002	-0.147	0.001	-0.159	0.0005	-0.164	0.0002	-0.147	0.001	-0.168	0.0002		
CD8 + T cells	CD8A	0.137	0.002	0.140	0.002	0.139	0.002	-0.005	0.910	0.006	0.901	-0.007	0.879		
	CD8B	0.181	3.79E-05	0.176	8.83E-05	0.181	6.12E-05	0.068	0.131	0.080	0.082	0.070	0.1229		
Neutrophils	FCGR3B	0.068	0.125	0.077	0.086	0.068	0.134	-0.244	3.03E-08	-0.234	2.32E-07	-0.242	5.93E-08		
	CEACAM3	-0.082	0.064	-0.080	0.077	-0.083	0.067	-0.322	1.59E-13	-0.316	1.64E-12	-0.326	1.18E-13		
	SIGLEC5	-0.086	0.050	-0.090	0.045	-0.080	0.078	-0.289	4.23E-11	-0.287	1.63E-10	-0.289	6.80E-11		
	FPR1	-0.027	0.541	-0.017	0.699	-0.019	0.689	-0.275	3.59E-10	-0.266	3.31E-09	-0.278	3.69E-10		
	CSF3R	-0.144	0.001	-0.146	0.001	-0.149	0.001	-0.327	5.55E-14	-0.317	1.21E-12	-0.328	7.79E-14		
	S100A12	0.112	0.011	0.123	0.006	0.116	0.011	-0.168	0.0002	-0.185	4.74E-05	-0.167	0.0002		
Macrophages	CD68	-0.044	0.318	-0.041	0.354	-0.038	0.406	-0.282	1.37E-10	-0.281	4.26E-10	-0.274	6.39E-10		
	CD84	-0.072	0.104	-0.063	0.159	-0.070	0.123	-0.155	0.0005	-0.147	0.001	-0.152	0.0007		
	CD163	0.002	0.970	0.003	0.951	0.007	0.872	-0.257	5.04E-09	-0.251	2.85E-08	-0.259	6.06E-09		
	MS4A4A	-0.074	0.093	-0.069	0.126	-0.072	0.110	-0.232	1.54E-07	-0.226	5.64E-07	-0.233	1.81E-07		
Dendritic cells	CD209	0.032	0.465	0.031	0.498	0.029	0.526	-0.242	3.95E-08	-0.239	1.24E-07	-0.239	8.25E-08		
	CD1C	-0.52	4.31e-37	-0.513	2.01e-34	-0.504	1.51e-32	-0.33	3.60e-14	-0.167	2.41e-04	-0.32	4.19e-13		
	CD141	-0.312	4.26e-13	-0.31	1.95e-12	-0.294	4.51e-11	-0.197	9.34e-06	-0.191	2.71e-05	-0.187	3.08e-05		
NK cells	KIR3DL3	0.187	1.97E-05	0.198	9.24E-06	0.187	3.43E-05	0.024	0.595	0.0286	0.533	0.040	0.376		
	NCR1	0.131	0.003	0.137	0.002	0.123	0.007	-0.042	0.346	-0.040	0.389	-0.039	0.386		
Th1 cells	TBX21	-0.014	0.743	-0.007	0.872	-0.012	0.789	-0.087	0.053	-0.076	0.096	-0.084	0.061		
Treg	FOXP3	0.015	0.731	0.017	0.700	-0.001	0.987	-0.183	3.79E-05	-0.183	5.61E-05	-0.181	5.49E-05		
	CCR8	0.018	0.677	0.0256	0.565	0.008	0.861	-0.163	0.0003	-0.165	0.0003	-0.161	0.0003		
Monocyte	C3AR1	-0.063	0.154	-0.055	0.221	-0.060	0.186	-0.227	2.92E-07	-0.228	6.92E-07	-0.229	2.73E-07		
	CD86	0.007	0.871	0.0107	0.813	0.008	0.856	-0.228	2.50E-07	-0.229	4.05E-07	-0.228	3.20E-07		
	CSF1R	-0.160	0.0002	-0.151	0.0008	-0.164	0.0003	-0.316	4.15E-13	-0.320	7.73E-13	-0.317	6.69E-13		

Table 2  
Correlation results between NUF2 and markers of immune cells via GEPIA LUSC, lung squamous cell carcinoma; LUAD, lung adenocarcinoma; NK cells, Natural killer cells; R, r value of Spearman's correlation; P, p-value.

Cell type	Gene	LUSC	
	markers	R	P
Macrophag-es	CD68	-0.25	2.2e-08
	CD84	-0.2	6.6e-06
	CD163	-0.28	5.4e-10
	MS4A4A	-0.25	4.3e-08
Neutrophils	FCGR3B	-0.21	2.5e-06
	CEACAM3	-0.3	2.1e-11
	SIGLEC5	-0.3	1.8e-11
	FPR1	-0.29	1e-11
	CSF3R	-0.3	1.1e-11
	S100A12	-0.14	0.0018
Dendritic cells	CD209	-0.24	1.2e - 07
	CD141	-0.13	0.0029
	CD1C	-0.33	7.2e - 14
Cell type	Gene	LUAD	
	markers	R	P
B cells	FCRL2	-0.07	0.13
	CD19	-0.077	0.09
	MS4A1	-0.18	8.5e-05
Macrophag-es	CD68	-0.073	0.11
	CD84	-0.12	0.0096
	CD163	0.024	0.59
	MS4A4A	-0.099	0.03
Dendritic cellsp	CD209	0.0081	0.86
	CD141	-0.31	4.4e-12
	CD1C	-0.51	1e-33
CD8 + T cells	CD8A	0.11	0.012
	CD8B	0.16	0.00028
NK cells	KIR3DL3	0.2	1.5e-05
	NCR1	0.079	0.082

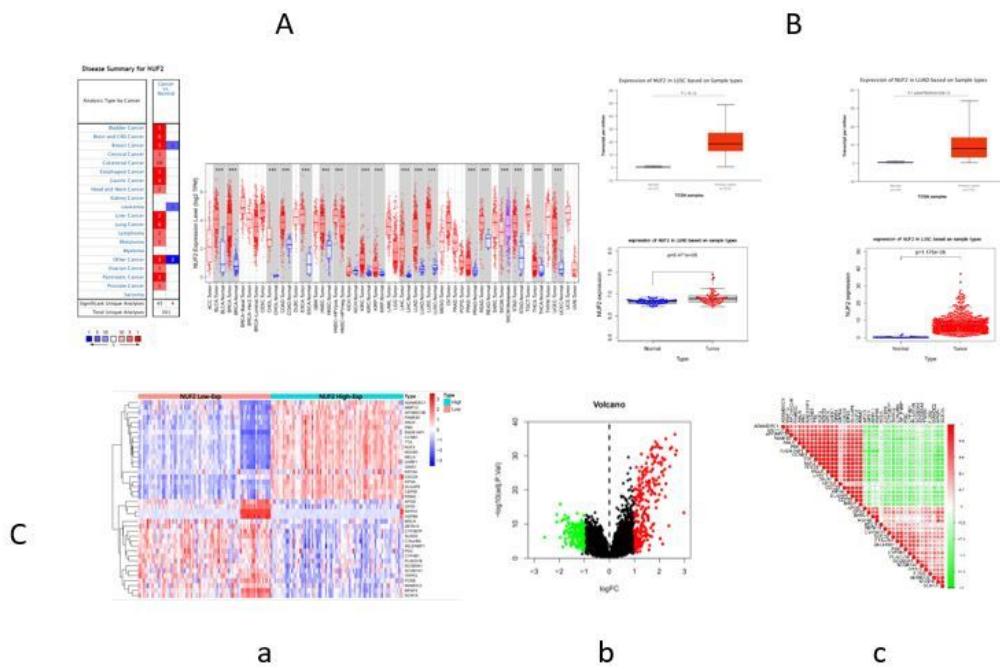
LUSC, lung squamous cell carcinoma; LUAD, lung adenocarcinoma; NK cells, Natural killer cells; R, r value of Spearman's correlation; P, p-value.

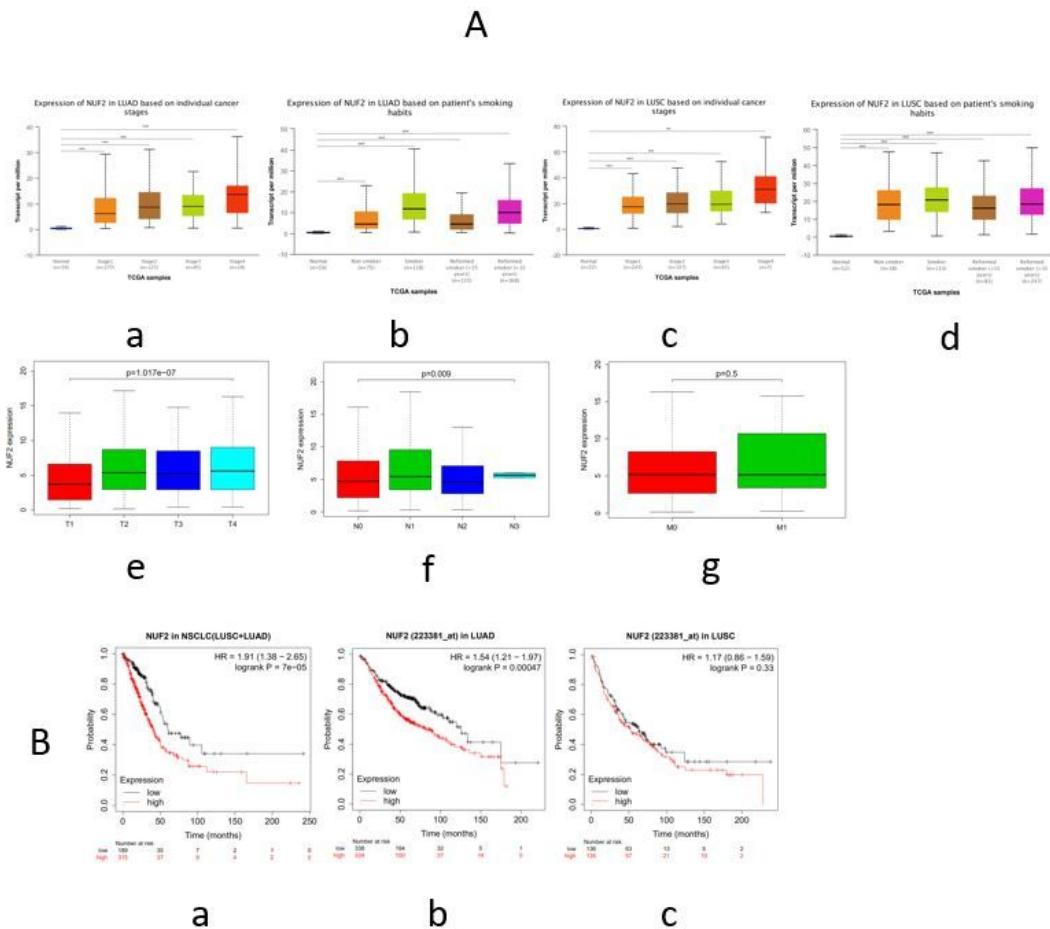
## Figures



**Figure 1**

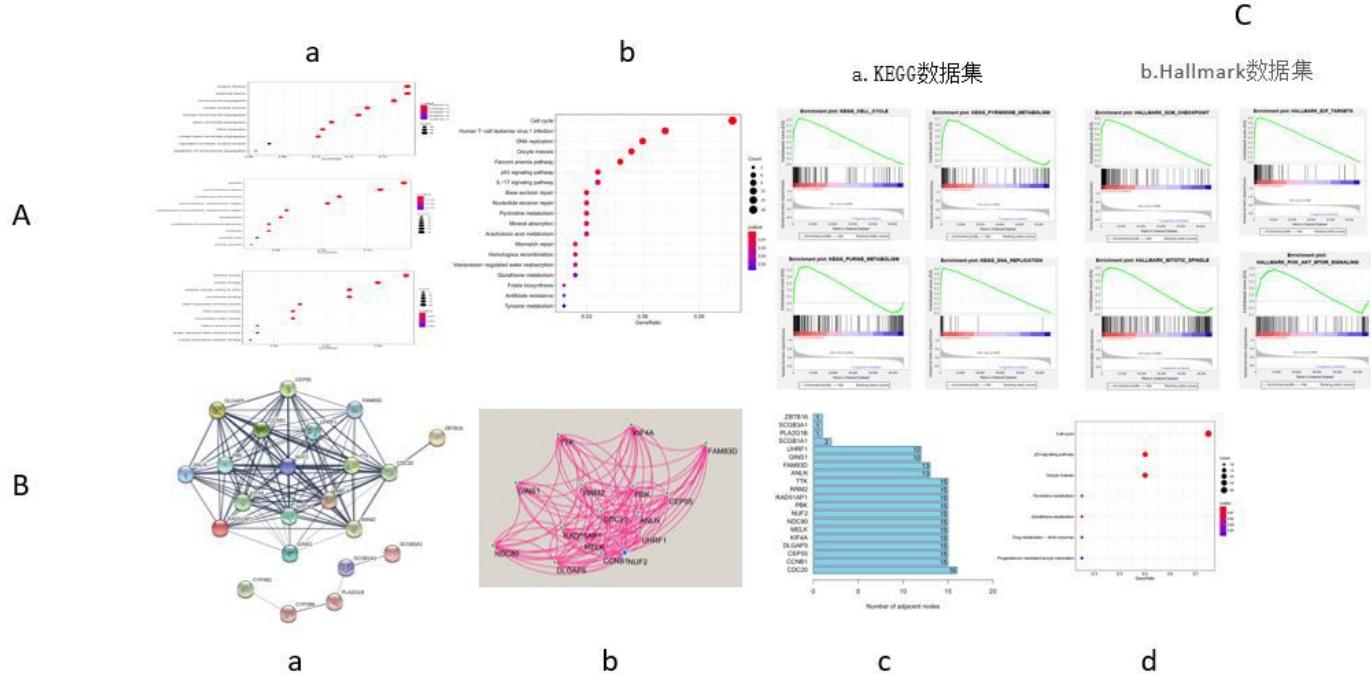
Flow diagram of data collection and method implementation in this work.





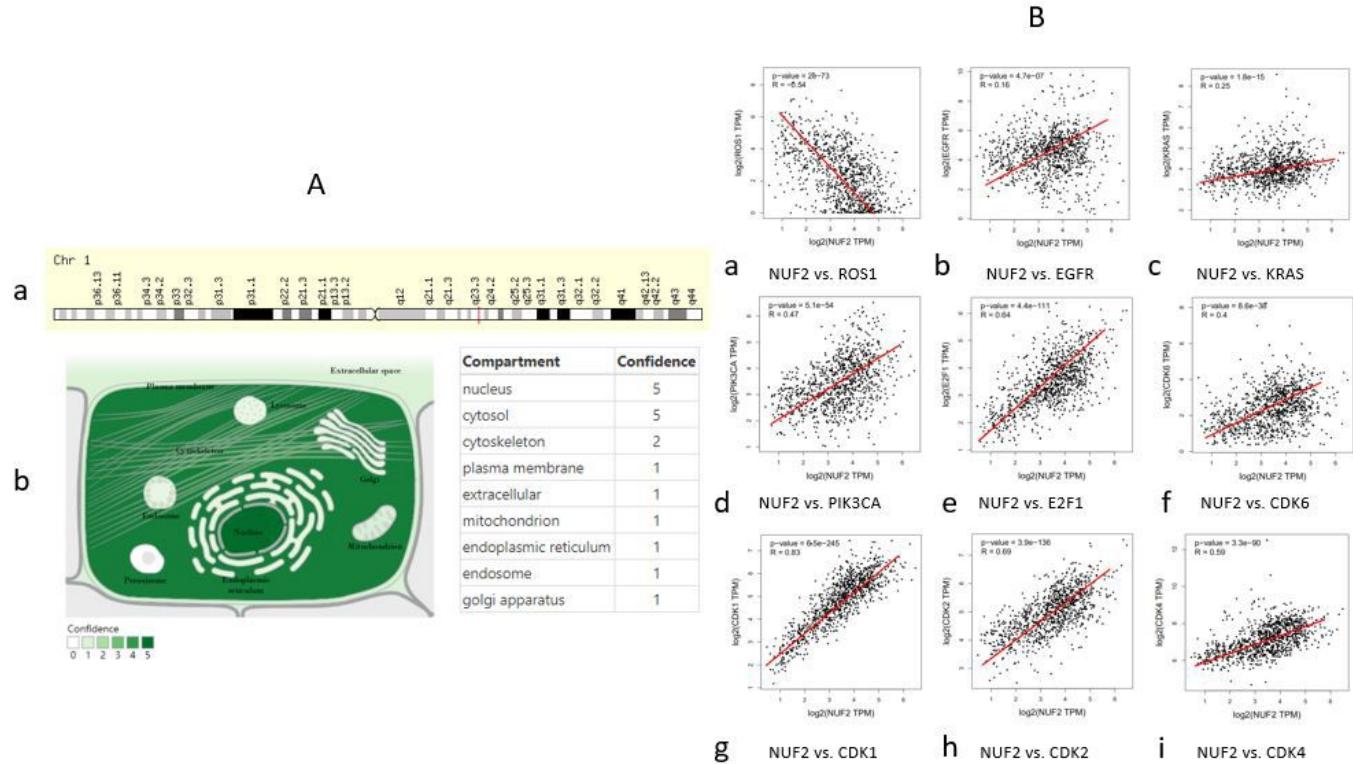
**Figure 3**

prognosis analysis and the association of NUF2 expression with clinicopathological factors in NSCLC.



**Figure 4**

Functional annotation and protein-protein interaction. A. Significantly enriched GO annotations(a) and KEGG pathways(b) of NUF2 in NSCLC B. Protein-protein interaction network of NUF2 and module analysis. b-d show the significant module via MCODE in the Cytoscape and the most significant pathway in the module via pathway enrichment analysis.



**Figure 5**

Basic characteristic of NUF2 and the correlation of NUF2 with special genes.

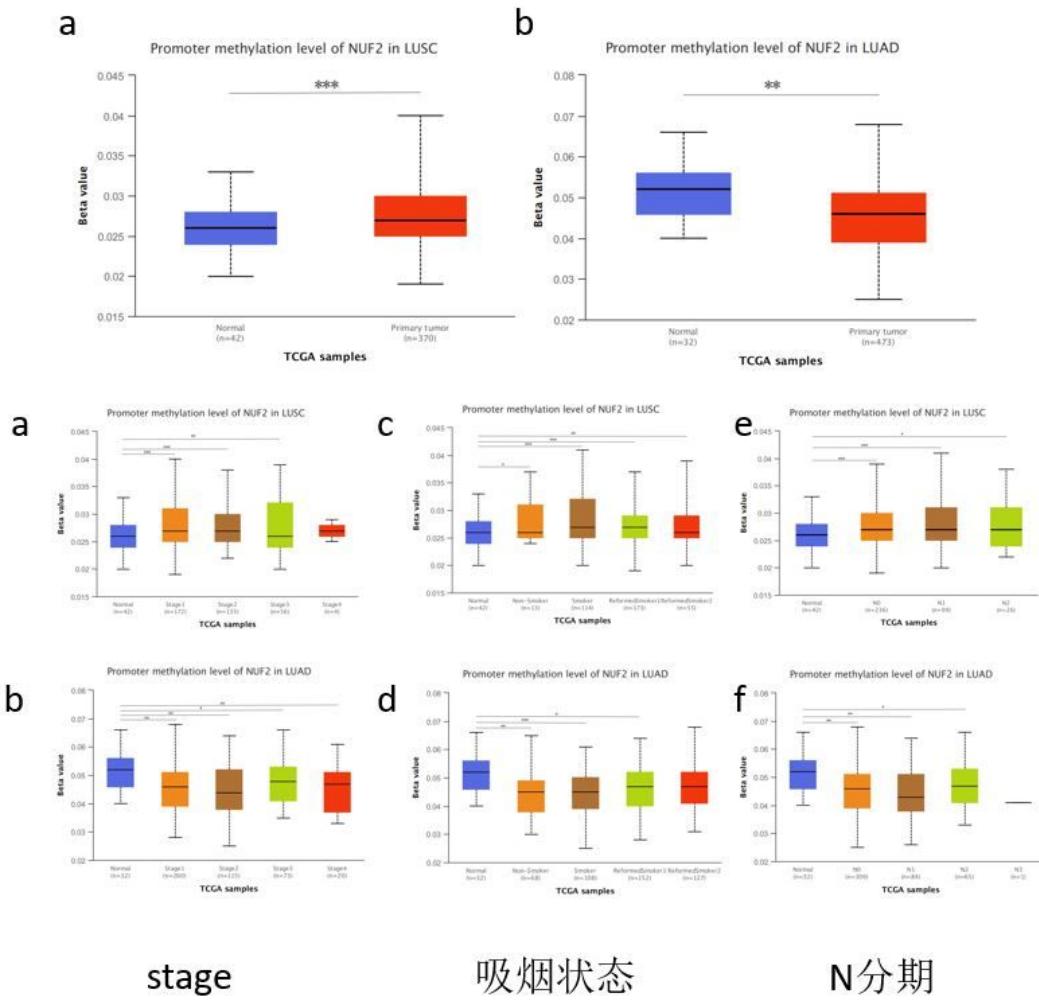
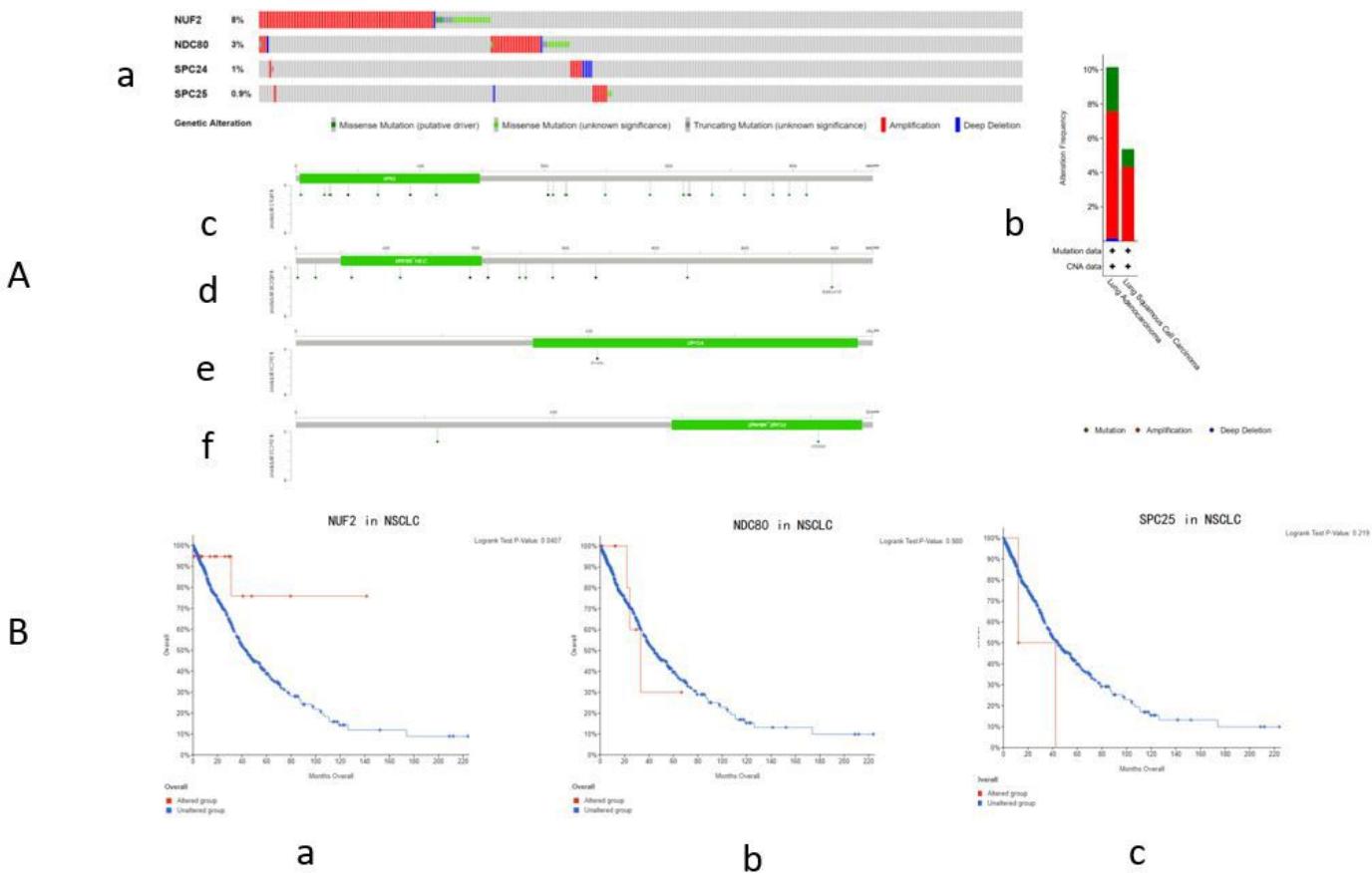


Figure 6

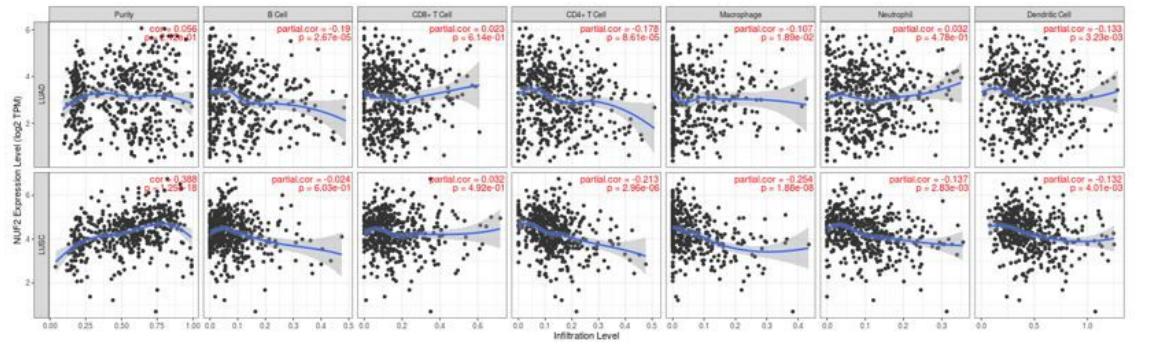
DNA methylation level of NUF2 and the association with clinicopathological factors in NSCLC



**Figure 7**

NUF2 alteration in NSCLC A. the types and frequency of NUF2 alterations B. prognosis analysis of NUF2, NDC80, SPC24, SPC25 in altered group and unaltered group.

A



B

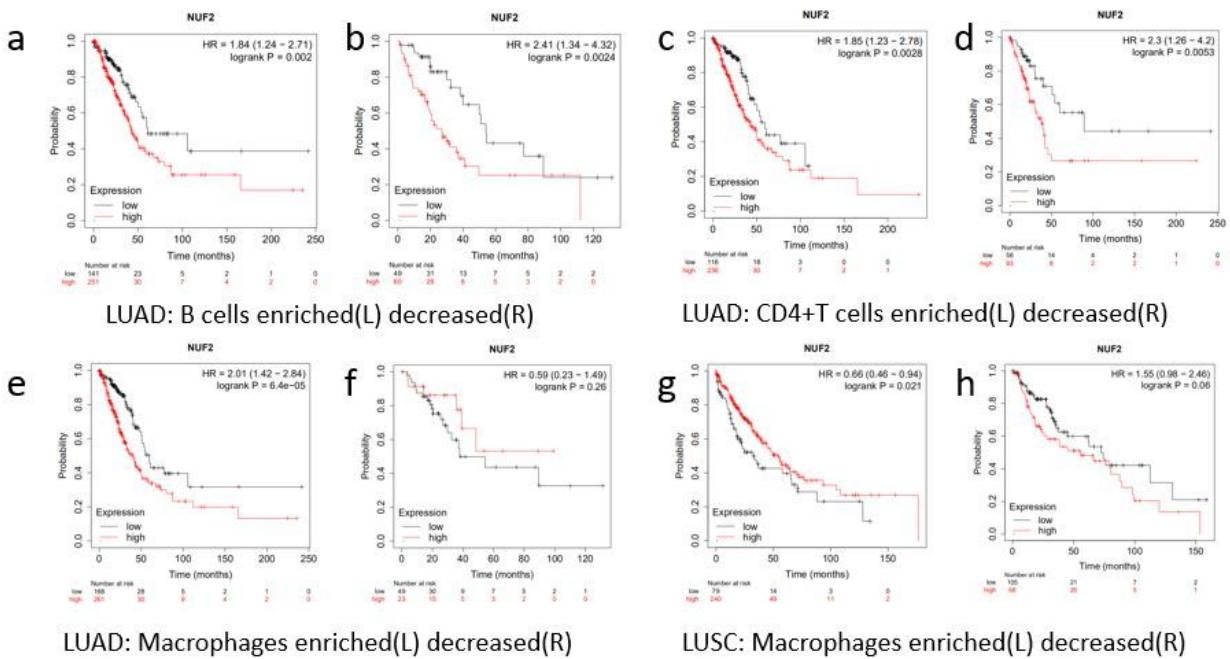


Figure 8

NUF2 expression associated with immune infiltration and prognosis analysis based on immune infiltration