



## Research article

# Identification of the H3K36me3 reader LEDGF/p75 in the pancancer landscape and functional exploration in clear cell renal cell carcinoma

Yuwei Zhang<sup>a,b,1</sup>, Wei Guo<sup>b,1</sup>, Yangkun Feng<sup>a,1</sup>, Longfei Yang<sup>a,b</sup>, Hao Lin<sup>b</sup>, Pengcheng Zhou<sup>e</sup>,  
Kejie Zhao<sup>e</sup>, Lin Jiang<sup>d,\*</sup>, Bing Yao<sup>e,\*\*</sup>, Ninghan Feng<sup>a,b,c,\*\*\*</sup>

<sup>a</sup> Nantong University Medical School, Nantong, China

<sup>b</sup> Department of Urology, Jiangnan University Medical Center, Wuxi, China

<sup>c</sup> Wuxi School of Medicine, Jiangnan University, Wuxi, China

<sup>d</sup> Department of Endocrinology and Metabolism, First Affiliated Hospital of Nanjing Medical University, Nanjing, China

<sup>e</sup> Department of Medical Genetics, Nanjing Medical University, Nanjing, China



## ARTICLE INFO

## Keywords:

LEDGF/p75

H3K36me3

Pancancer

SETD2

Clear cell renal cell carcinoma

## ABSTRACT

Lens epithelium-derived growth factor (LEDGF/p75) is a reader of epigenetic marks and a potential target for therapeutic intervention. Its involvement in human immunodeficiency virus (HIV) integration and the development of leukemia driven by MLL (also known as KMT2A) gene fusion make it an attractive candidate for drug development. However, exploration of LEDGF/p75 as an epigenetic reader of H3K36me3 in tumors is limited. Here, for the first time, we analyze the role of LEDGF/p75 in multiple cancers via multiple online databases and in vitro experiments. We used pancancer bulk sequencing data and online tools to analyze correlations of LEDGF/p75 with prognosis, genomic instability, DNA damage repair, prognostic alternative splicing, protein interactions, and tumor immunity. In summary, the present study identified that LEDGF/p75 may serve as a prognostic predictor for tumors such as adrenocortical carcinoma, kidney chromophobe, liver hepatocellular carcinoma, pancreatic adenocarcinoma, skin cutaneous melanoma, and clear cell renal cell carcinoma (ccRCC). In addition, in vitro experiments and gene microarray sequencing were performed to explore the function of LEDGF/p75 in ccRCC, providing new insights into the pathogenesis of the nonmutated SETD2 ccRCC subtype.

## 1. Introduction

LEDGF/p75, encoded by *PSIP1*, was originally identified as a protein copurifying with the general transcriptional coactivator PC4 and described as a transcriptional coactivator related to stress and autoimmune responses[1]. *PSIP1* also codes for an alternative splicing isoform referred to as p52. Compared with p52, LEDGF/p75 has an integrase binding domain (IBD) in addition to the common PWWP domain[2]. LEDGF/p75 has been reported to play a key role in the development of human immunodeficiency virus (HIV) and MLL leukemia[3–5]. HIV integrase can recognize and bind to the IBD of LEDGF/p75 and hijack it to the transcriptionally active region of the genome, allowing the virus to replicate in large numbers[6,7]. Similarly, MLL/MENIN complex, an

important player in the development and progression of MLL leukemia, can bind to the IBD of LEDGF/p75 to promote development and progression of the disease[8,9]. In fact, as a chromatin-binding protein, LEDGF/p75 mediates chromatin localization of several nuclear proteins [10–13].

Posttranslational modification of histones is an important branch of epigenetic inheritance that has been widely reported in various diseases, especially in cancer. H3K36me3 is a 3-methylated modification at the 36th K of histone H3 and mediates several key tumor processes, such as transcriptional elongation, DNA methylation, and DNA damage repair [14]. Studies have reported that LEDGF/p75 reads the H3K36me3 mark via its PWWP domain to recruit functional proteins to the region of actively transcribed genes[15,16]. However, its role in tumors is poorly

\* Correspondence to: Department of Endocrinology and Metabolism, First Affiliated Hospital of Nanjing Medical University, 300 Guangzhou Road, Nanjing 210029, China.

\*\* Correspondence to: Department of Medical Genetics, Nanjing Medical University, 101 Longmian Road, Nanjing 211166, China.

\*\*\* Correspondence to: Nantong University Medical School, 9 Qiangyuan Road, Nantong 226019, China.

E-mail addresses: [jlina0000@163.com](mailto:jlina0000@163.com) (L. Jiang), [byao@njmu.edu.cn](mailto:byao@njmu.edu.cn) (B. Yao), [n.feng@njmu.edu.cn](mailto:n.feng@njmu.edu.cn) (N. Feng).

<sup>1</sup> These authors contributed equally to this work.

<https://doi.org/10.1016/j.csbj.2023.08.023>

Received 22 May 2023; Received in revised form 18 August 2023; Accepted 25 August 2023

Available online 26 August 2023

2001-0370/© 2023 Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

understood. Therefore, the present study was performed to characterize the landscape of LEDGF/p75 across cancers for the first time.

Renal cell carcinoma includes more than 10 histological and molecular subtypes, of which clear cell renal cell carcinoma (ccRCC) is the most common and accounts for the majority of deaths associated with kidney cancer [17]. Genetically, ccRCC results from high-frequency mutations or even deletion of multiple tumor-suppressor genes (VHL, 80%; PBRM1, 29–46%; BAP1, 6–19%; and SETD2, 8–30%), which leads to genomic instability and promotes defects in DNA repair pathways [18]. ccRCC can be classified into clinically and therapeutically relevant subtypes based on the molecular characteristics caused by these defects [19]. SETD2 is an RNA polymerase II-associated histone methyltransferase that catalyzes H3K36me3, which is a transcriptional activity marker. Previous studies indicated that H3K36me3 is only added by SETD2. Although SMYD5 was recently reported to play a role in methylation, SETD2 is the most dominant specific methylase [20].

Research on the presence or absence of SETD2 is of key clinical significance for personalized treatment. Therefore, the present study was performed as preliminary functional exploration of LEDGF/p75, the main reader of H3K36me3, in ccRCC.

## 2. Materials and Methods

### 2.1. Acquisition of basic information about LEDGF/p75

The genomic view for the LEDGF/p75 gene was obtained from the GeneCards database [21]. The features for the domains and regions of LEDGF/p75 were obtained from the UniProt database [22]. A three-dimensional structure for LEDGF/p75 was constructed from AlphaFold [23]. The immunofluorescence graphs of the intracellular location of LEDGF/p75 in U-251MG cells (HPA019697) were obtained from the ATLAS database [24].

### 2.2. Interaction network of LEDGF/P75 and functional enrichment analyses

The LEDGF/p75 protein–protein interaction network with physical interactions was predicted via the GeneMANIA database [25], and the potential pathways are marked in colors. The LEDGF/p75 protein–protein interaction network with known experimental validations was also explored via the STRING database [26]. We further predicted scores in cancers and other diseases based on the LEDGF/p75 protein–protein interaction network via the canSAR database [27].

Gene Ontology (GO) analyses, including biological process, cellular component analyses and molecular function analyses, along with reactome pathways were predicted via the TISIDB database [28]. Gene set enrichment analysis (GSEA) was carried out via R software with canonical pathway gene sets derived from the KEGG pathway database [29]. All R programs used in the present study were uploaded to GitHub (<https://github.com/melondocor/LEDGF/tree/master>).

### 2.3. Analyses of LEDGF/p75's correlation with histone modification, DNA mismatch repair, tumor environment, immune cell infiltration, and immune-related genes

We obtained pancancer expression profiles from the UCSC Xena database [30]. Associations between 5 histone modification genes, 5 DNA mismatch repair genes, tumor environment, immune cell infiltration, immune-related genes and LEDGF/p75 expression in pancancer were visualized using R software.

### 2.4. Differential expression of LEDGF/p75 in normal and tumor groups

We processed expression data from the UCSC Xena database, deleted data with less than three normal samples, and analyzed the remaining data for 21 tumor types via R software. For the remaining 12 tumor

types, we added data from GTEx through GEPIA [31] for further analysis and identified differences in LEDGF/p75 expression between three tumor groups and corresponding normal groups. We further explored differential expression of LEDGF/p75 in different kinds of cells and obtained results from the ATLAS database.

For the LEDGF/p75 protein expression of tumor and normal groups, we searched the UALCAN database [32] and found 7 tumor types with different expression. Subsequently, we searched the ATLAS database and found corresponding immunohistochemical diagrams to show LEDGF/p75 protein expression.

### 2.5. Analyses between LEDGF/p75 expression and patient prognosis

Data on expression and survival in pancancer were obtained from the UCSC Xena database and analyzed via R software. Overall survival (OS), disease-specific survival (DSS), disease-free interval (DFI), and progression-free interval (PFI) were analyzed to indicate patient prognosis. Cox regression analysis and Kaplan–Meier (K–M) analysis were used for forest plots and K–M plots, respectively. In addition, we analyzed LEDGF/p75 expression and clinical stages of patients across cancers.

### 2.6. Genomic alterations of LEDGF/p75 across cancers

We searched the cBioPortal database [33] for information about genomic alterations in LEDGF/p75 across cancers. We first identified the landscape of alteration frequency across cancers, including mutations, structural variants, amplifications, deep deletions and multiple alterations. Then, we searched the ratio of the alteration group in pancancer. We further obtained detailed information on copy-number alterations and mutations in LEDGF/p75. Finally, we analyzed the correlation between genomic alterations of LEDGF/p75 and patient survival.

We processed the mutation data of pancancer from the UCSC Xena database and analyzed them via R software to identify information about the tumor mutational burden (TMB) and microsatellite instability (MSI).

### 2.7. Clinically relevant alternative splicing analyses of LEDGF/p75

To identify clinically relevant alternative splicing (AS) events, the OncoSplicing database [34] was searched for AS events for LEDGF/p75. We chose project 247053 for subsequent analyses, which was the only known splice type in SplAdder methodology according to the OncoSplicing database. Pan plots indicate the reads in, reads out and percent spliced-in (PSI) values in pancancer and normal tissues. PanDiff plots compared the PSI differences of queried AS events (detected in more than 3 cancers) between cancers and adjacent or GTEx normal tissues. Finally, we explored the prognostic significance of LEDGF/p75 AS events across cancers via K–M plots.

### 2.8. Exploration of the immunological roles of LEDGF/p75 across cancers

We first explored the association between LEDGF/p75 expression and immune subtypes across cancers via the TISIDB database. Subsequently, we analyzed detailed subtypes, including C1 (wound healing), C2 (IFN-gamma dominant), C3 (inflammatory), C4 (lymphocyte depleted), C5 (immunologically quiet), and C6 (TGF- $\beta$  dominant). We visualized the distribution across subtypes for the top six tumor types with the highest LEDGF/p75 expression. Heatmaps showed Spearman correlations between immunoinhibitors, chemokines, tumor infiltrating lymphocytes (TILs), and LEDGF/p75 expression across cancers.

We further searched the most confident results using gene expression data in the TIDE database [35], including parts of cancer, subtype, Pearson correlation with cytotoxic T lymphocyte level (CTL Cor), T-cell dysfunction score (T Dysfunction), survival risk score (Risk), survival risk score adjusted for the effect of cytotoxic T lymphocyte (Risk. adj),

and sample count in the dataset (count).

We then compared LEDGF/p75 expression levels across cell lines between pre- and postcytokine-treated samples with the TISMO database[36]; IFN $\gamma$ , IFN $\beta$ , TNF $\alpha$ , and TGF $\beta$ 1 are included as cytokine treatments in the module.

Finally, we searched ROC Plotter[37] to explore the correlation between LEDGF/p75 expression and immunotherapy. Receiver operating characteristic (ROC) curves indicated the high diagnostic value of LEDGF/p75 for assessing immunotherapy outcomes.

## 2.9. Cell culture

Human kidney cancer cell lines (i.e., Caki-1, 786-O, A498) and normal human kidney epithelial cells (HK-2) were acquired from Procell Life Science & Technology Company (Wuhan, China). All cell lines used in this study were tested and authenticated by DNA sequencing using the STR method (ABI 3730XL Genetic Analyzer) and tested for the absence of mycoplasma contamination (MycAlert). All cell lines were cultured in commercial cell culture medium at 37 °C in a 5% CO $_2$  atmosphere.

## 2.10. RNA isolation and quantitative real-time PCR (qRT–PCR)

TRIzol Reagent (Invitrogen, USA) was used to isolate total RNA. HiScript III SuperMix (Vazyme, China) was used to perform reverse transcription. qRT–PCR was used to analyze the expression level of mRNAs, as performed using a SYBR Green Kit (Yeasen, China) with the LightCycler® 96 SW 1.1 system (Roche, Switzerland).

## 2.11. Western blotting

RIPA buffer (Beyotime, China) mixed with protease inhibitor (Beyotime) was used to extract total cell protein. The proteins were separated and then transferred to a polyvinylidene difluoride membrane, which was incubated in 10% milk for 2 h at room temperature (RT). Subsequently, the membrane was incubated in the primary antibody (1:1000, anti-LEDGF/p75: Abcam#ab177159; anti-H3K36me3: Cell Signaling Technology#4909 s; anti-H3: Proteintech#17168–1-AP; anti- $\beta$ -Actin: Proteintech#81115–1-RR) for 12 h at 4 °C and then treated with a matched secondary antibody (1:5000, Proteintech#SA00001–2) at RT for 2 h. Enhanced chemiluminescence (Tanon, China) was used for detection.  $\beta$ -Actin and histone H3 were used as endogenous controls.

## 2.12. siRNA transfection

siRNA sequences, which were designed and synthesized by RiboBio (Guangzhou, China), used to target LEDGF/p75 were GGAAGA-TACCGACCATGAA (5′–3′, KD1) and GCAGCAACTAAACAATCAA (5′–3′, KD2). The siRNAs were transfected with Lipo3000 (Invitrogen) according to the instruction manual.

## 2.13. Cell counting kit-8 (CCK-8) and clone formation

A total of 2000 cells were seeded into 96-well plates and cultured for the indicated times. At each time point, 10  $\mu$ l of CCK-8 reagent (Yeasen) was mixed with the cells for 1 h. Optical density (OD) values were measured at 450 nm.

For the clone formation experiment, 1000/well of the indicated cells were cultured in a 6-well plate for 10 days. Methanol was used to fix the cells for 30 min, followed by crystal violet staining for 30 min.

## 2.14. Transwell assay

Cell migration assays were performed with 24-well no-Matrigel Transwell chambers (Corning, USA). A total of 3  $\times$  10 $^4$  cells were cultured in the upper chamber suspended in 200  $\mu$ l of medium without fetal bovine serum (FBS), and 600  $\mu$ l of medium containing 10% FBS was

added to the bottom chamber. After overnight incubation, crystal violet was used to stain the cells on the lower surface of the chamber for 30 min. Images of three random fields were acquired using a fluorescence microscope, and the cells were counted.

## 2.15. Gene microarray analysis

786-O cells were selected for LEDGF/p75 knockdown treatment, and three biological replicates were prepared for gene microarray detection. Agilent SurePrint G3 Human Gene Expression v3 8 $\times$ 60K Microarray (DesignID:072363) chip experiments and data analysis of 6 samples were performed at Shanghai Ouyi Biomedical Technology Co., Ltd. China.

Feature Extraction software version 10.7.1.1 (Agilent Technologies) was used to process original images and extract original data. The original data were then standardized. Differential genes were screened according to fold change > 1.5 and P value < 0.05. Then, GO and KEGG enrichment analyses of differentially expressed genes were performed to determine biological functions and pathways.

## 2.16. Online databases

Information on all online databases used in the present study can be found in Table 2.

## 2.17. Statistical analyses

All bioinformatics analyses were conducted via R software (version 4.2.2), except for the results obtained from the online databases mentioned in this study. Independent t tests were used to compare normally distributed continuous variables and Mann–Whitney U tests to compare skewed continuous variables. All statistical tests were two-sided. P values less than 0.05 (\*  $P$  < 0.05) were considered significant.

## 3. Results

### 3.1. Biological information of LEDGF/p75

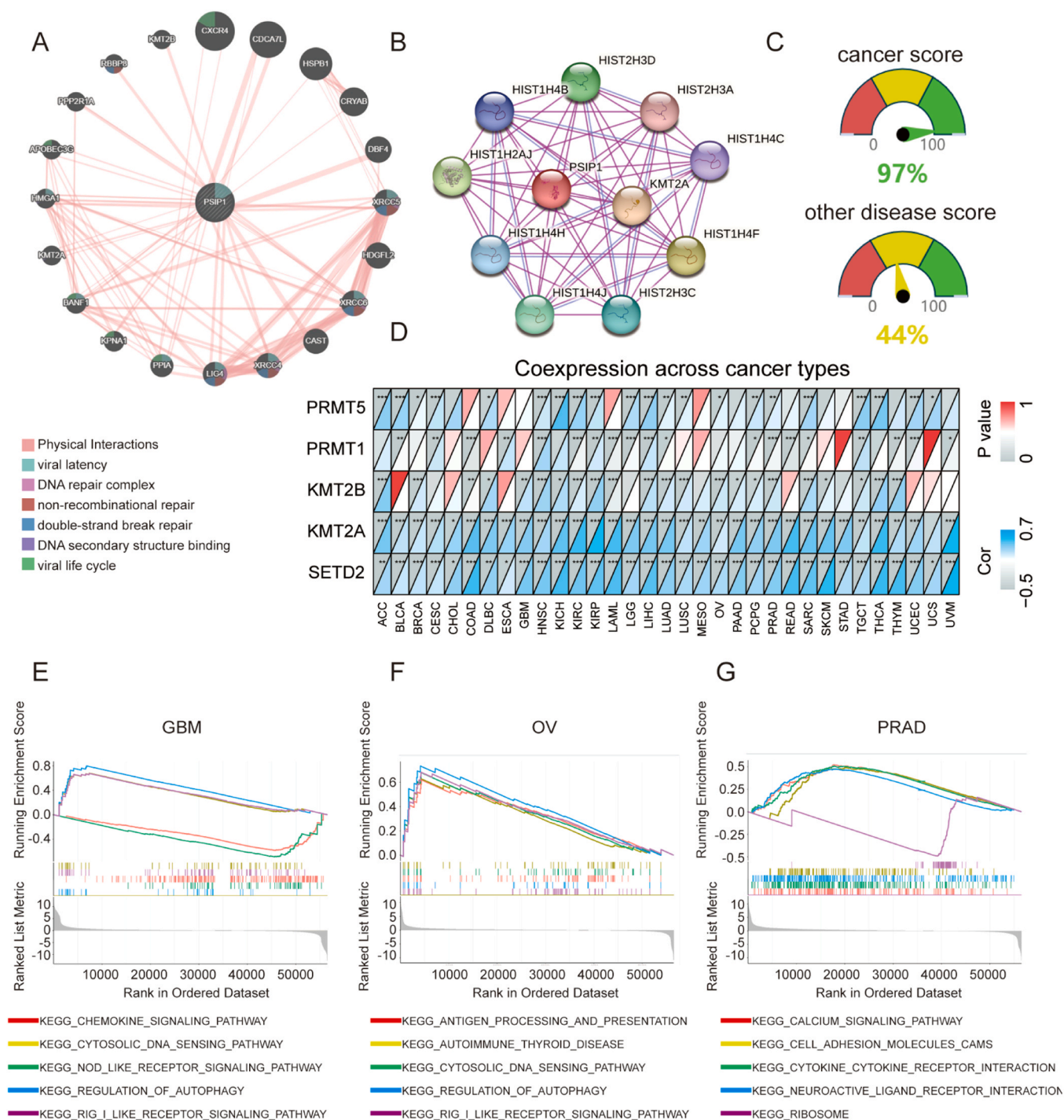
*PSIP1* is a protein-coding gene located in the short arm of the ninth chromosome (Supplementary Fig. 1 A). LEDGF/p75, encoded by *PSIP1*, has two functional domains: the PWWP domain (aa 1–91) and the IBD (aa 347–454) (Supplementary Fig. 1B–C)[38]. The PWWP domain reads the H3K36me3 mark, which is an important regulatory mode in epigenetics[39]. For the IBD, a protein binding hub, previous studies have reported several interacting proteins, such as HIV integrase 1 and 2, MENIN–MLL, CDCA7L, PogZ, CDC7–DBF4, and IWS1 [38].

We further explored localization of LEDGF/p75 in U-251MG cells and found that almost all of the protein in the nucleus, which was consistent with its chromatin identification function (Supplementary Fig. 1D).

### 3.2. LEDGF/p75 is involved in several diseases, especially cancer

We first focused on protein interactions of LEDGF/p75 and predicted proteins that physically bind to it via online tools. As displayed in Fig. 1A, LEDGF/p75 is mainly involved in viral infections and DNA repair, such as the viral life cycle, DNA repair complex and DNA binding. In addition, previous experiments confirmed that LEDGF/p75 binds to histones and histone-modification proteins (Fig. 1B). Finally, we predicted the scores of LEDGF/p75 in cancer and other diseases based on protein–protein interactions, which highlighted the importance of LEDGF/p75 in cancer (Fig. 1C). As LEDGF/p75 was reported as a reader of histone modification marks such as H3K36me3, we further explored the correlation between LEDGF/p75 expression and several common histone modification-related genes. As shown in Fig. 1D, SETD2, a classic writer of H3K36me3, correlated highly positively with LEDGF/





**Fig. 1. LEDGF/p75 plays a vital role in tumors and other diseases.** (A) Physical interactions of LEDGF/p75 and other key proteins indicate potential functions of LEDGF/p75. (B) LEDGF/p75 binds to histones and several histone modification proteins. The purple lines represent experimentally determined interactions, and the blue lines show the results from curated databases. (C) Predicted scores of LEDGF/p75 in cancers and other diseases based on protein–protein interactions. Different colors represent different rating levels of the canSAR database. (D) Correlations between LEDGF/p75 expression and 5 histone modification genes in pancancer. (E–G) Results of GSEA-KEGG analyses. \*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ .

p75 expression in all 33 tumor types. Interestingly, apart from the previously reported lysine methylation factors (SETD2, KMT2A, KMT2B), our results showed that LEDGF/p75 expression also correlated with arginine methylation factors (PRMT1, PRMT5) in some tumor types, suggesting a potential biological role.

GO analyses, including biological process, cellular component, and molecular function, indicated vital roles for LEDGF/p75 in both HIV infection and cancers (Supplementary Table 1). For instance, LEDGF/

p75 functions in viral latency, response to oxidative stress, and chromatin binding. The reactome pathway of LEDGF/p75 illustrates the details of LEDGF/p75 in HIV integration and the viral life cycle (Supplementary Table 2). Then, we performed GSEA-KEGG analyses to further explore LEDGF/p75 function. LEDGF/p75 expression correlated with regulation of autophagy in GBM and OV and with cell adhesion molecules in PRAD, which indicates the potential function of LEDGF/p75 in tumor metastasis and development (Fig. 1E–G). All TCGA



abbreviations are provided in Table 1.

### 3.3. LEDGF/p75 is differentially expressed across cancers

We analyzed LEDGF/p75 expression data from the UCSC Xena database and found LEDGF/p75 mRNA to be significantly differentially expressed in 15 of 21 tumor types. Among the 15 types, LEDGF/p75 was downregulated in 11 types (BLCA, BRCA, KICH, KIRC, KIRP, LUAD, LUSC, PRAD, READ, THCA, and UCEC) and highly expressed in another 4 (CHOL, HNSC, LIHC, and PCPG) (Fig. 2A). We supplemented data from GTEx normal tissues and performed the analyses for another 12 tumor types. We found LEDGF/p75 to be highly expressed in DLBC and THYM but reduced in OV (Fig. 2B–D).

Next, we explored LEDGF/p75 expression in various cell lines and found LEDGF/p75 to be differentially expressed in different cell lines (Fig. 2E). The top two cell lines with the highest LEDGF/p75 expression were HEL and NTERA-2, and the expression level of HEL cells was more than twice that of NTERA-2 cells. HEL, a cancer cell line derived from myeloid cells, is an erythroleukemia cell line (AML M6 in relapse after treatment for Hodgkin's disease). Such high expression of LEDGF/p75 in HEL cells suggests that it may play a role in erythroleukemia. Previous studies have reported that LEDGF/p75 is essential for MLL-rearranged leukemogenesis [40]. Whether there is a deeper connection between the two diseases other than both being blood cancers and whether this connection is related to LEDGF/p75 remains unclear.

Then, we analyzed LEDGF/p75 protein expression across cancers via online tools. LEDGF/p75 was significantly reduced in 5 tumor types (BRCA, COAD, HNSC, LUAD, and UCEC) but markedly increased in another 2 (LIHC and OV) (Fig. 2F–L). As expected, immunohistochemical results from the ATLAS database showed trends similar to the above results.

These findings indicate significant differences in LEDGF/p75 expression across cancers. After integrated analysis of differences in

**Table 1**  
Abbreviations of 33 tumors.

| Abbreviation | Full name                             |
|--------------|---------------------------------------|
| ACC          | Adrenocortical Carcinoma              |
| KIRC         | Kidney Renal Clear Cell Carcinoma     |
| PRAD         | Prostate Adenocarcinoma               |
| BLCA         | Bladder Urothelial Carcinoma          |
| KIRP         | Kidney Renal Papillary Cell Carcinoma |
| READ         | Rectum Adenocarcinoma                 |
| BRCA         | Breast Invasive Carcinoma             |
| LAML         | Acute Myeloid Leukemia                |
| SARC         | Sarcoma                               |
| CESC         | Cervical Squamous Cell Carcinoma      |
| LGG          | Lower Grade Glioma                    |
| SKCM         | Skin Cutaneous Melanoma               |
| CHOL         | Cholangiocarcinoma                    |
| LIHC         | Liver Hepatocellular Carcinoma        |
| STAD         | Stomach Adenocarcinoma                |
| COAD         | Colon Adenocarcinoma                  |
| LUAD         | Lung Adenocarcinoma                   |
| TGCT         | Testicular Germ Cell Tumors           |
| DLBC         | Diffuse Large B-cell Lymphoma         |
| LUSC         | Lung Squamous Cell Carcinoma          |
| THCA         | Thyroid Carcinoma                     |
| ESCA         | Esophageal Carcinoma                  |
| MESO         | Mesothelioma                          |
| THYM         | Thymoma                               |
| GBM          | Glioblastoma Multiforme               |
| OV           | Ovarian Serous Cystadenocarcinoma     |
| UCEC         | Uterine Corpus Endometrial Carcinoma  |
| HNSC         | Head and Neck Squamous Cell Carcinoma |
| PAAD         | Pancreatic Adenocarcinoma             |
| UCS          | Uterine Carcinosarcoma                |
| KICH         | Kidney Chromophobe                    |
| PCPG         | Pheochromocytoma and Paraganglioma    |
| UVM          | Uveal Melanoma                        |

**Table 2**  
Information of databases used in the present study.

| Database      | Online link   |
|---------------|---|
| GeneCards     | <a href="https://www.genecards.org/">https://www.genecards.org/</a>   |
| Uniprot       | <a href="https://www.uniprot.org/">https://www.uniprot.org/</a>   |
| AlphaFold     | <a href="https://alphafold.ebi.ac.uk/">https://alphafold.ebi.ac.uk/</a>   |
| ATLAS         | <a href="https://www.proteinatlas.org/">https://www.proteinatlas.org/</a>   |
| GeneMANIA     | <a href="http://genemania.org/">http://genemania.org/</a>   |
| STRING        | <a href="https://cn.string-db.org/">https://cn.string-db.org/</a>   |
| canSAR        | <a href="https://cansar.ai/">https://cansar.ai/</a>   |
| TISIDB        | <a href="http://cis.hku.hk/TISIDB/index.php">http://cis.hku.hk/TISIDB/index.php</a>   |
| KEGG pathway  | <a href="http://www.gsea-msigdb.org/gsea/msigdb/human/collections.jsp">http://www.gsea-msigdb.org/gsea/msigdb/human/collections.jsp</a> |
| UCSC Xena     | <a href="https://xena.ucsc.edu/">https://xena.ucsc.edu/</a>   |
| GEPIA         | <a href="http://gepia.cancer-pku.cn/index.html">http://gepia.cancer-pku.cn/index.html</a>   |
| UALCAN        | <a href="http://ualcan.path.uab.edu/index.html">http://ualcan.path.uab.edu/index.html</a>   |
| cBioPortal    | <a href="https://www.cbioportal.org/">https://www.cbioportal.org/</a>   |
| OncoSplicing  | <a href="http://www.oncosplicing.com/">http://www.oncosplicing.com/</a>   |
| TIDE          | <a href="http://tide.dfci.harvard.edu/">http://tide.dfci.harvard.edu/</a>   |
| TISMO         | <a href="http://tismo.cistrome.org/">http://tismo.cistrome.org/</a>   |
| ROC Plotter   | <a href="https://www.rocplot.org/">https://www.rocplot.org/</a>   |
| DepMap portal | <a href="https://depmap.org/portal/">https://depmap.org/portal/</a>   |

LEDGF/p75 mRNA and protein expression, we found the same trend for LIHC, BRCA, LUAD and UCEC. Notably, LEDGF/p75 expression was opposite at the mRNA and protein levels in HNSC, which requires experimental verification.

### 3.4. LEDGF/p75 expression correlates with patient prognosis

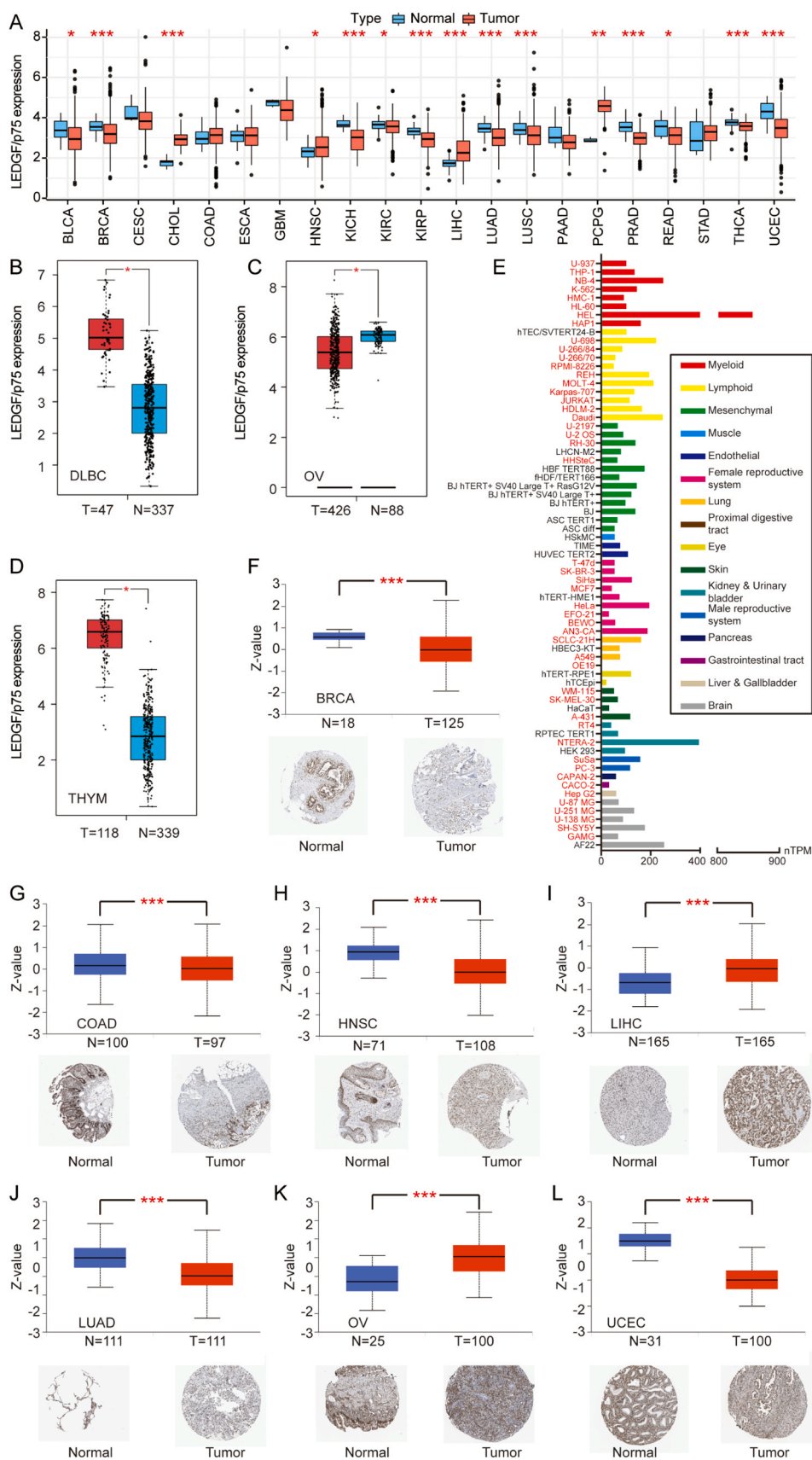
To investigate the correlation between LEDGF/p75 expression and patient prognosis, we performed comprehensive analysis of expression and clinical data from the UCSC Xena database. Cox regression and K–M analyses were used to explore OS, DSS, DFI and PFI. As shown in the forest plot, LEDGF/p75 expression correlated negatively with OS in 4 tumor types (ACC, KICH, LIHC, and UCEC) but positively in 5 (CESC, KIRC, LGG, PAAD, and SKCM) (Fig. 3A). Regarding K–M plots, there was a negative correlation between LEDGF/p75 expression and OS in ACC but a positive correlation in 3 other tumor types (KIRC, LGG, and OV) (Fig. 3B–E). The results for DSS, DFI and PFI are shown in Supplementary Fig. 2 and Supplementary Fig. 3.

Furthermore, we analyzed the correlation between LEDGF/p75 expression and the clinical stage of patients. The results indicated that patients with advanced ACC generally expressed high levels of LEDGF/p75; this phenomenon was reversed in patients with PAAD, SKCM, THCA, or KIRC (Fig. 3F–L). All these analyses indicate that LEDGF/p75 expression is closely associated with patient prognosis across cancers. Specifically, LEDGF/p75 is a potential oncogene in ACC, KICH and LIHC but a potential protective gene in PAAD and SKCM.

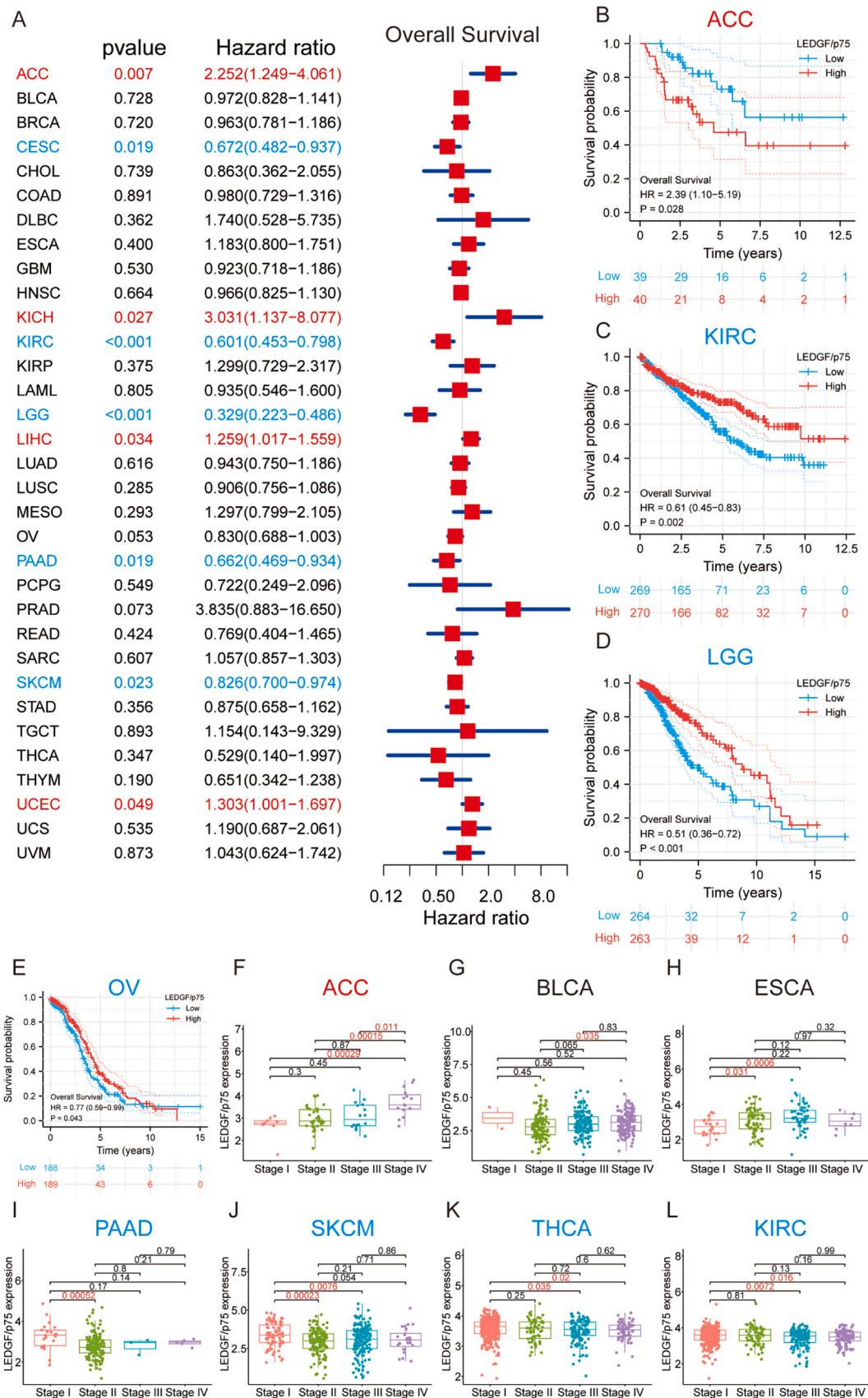
### 3.5. LEDGF/p75 genomic alterations in pancancer and correlation with patient prognosis

As genomic alterations may cause tumorigenesis, we explored LEDGF/p75 gene alterations in the cBioPortal database, which is a multidimensional cancer genomics dataset. We illustrate the landscape of alteration frequency in pancancer in Supplementary Fig. 4A, including mutation, structural variant, amplification, deep deletion, and multiple alterations. We then studied the percentage of altered groups in pancancer. Notably, the LEDGF/p75 gene was altered in nearly 50% of patients with 3 tumor types (COAD\_POLE, ESCA\_POLE, and UCEC\_POLE) (Supplementary Fig. 4B). Putative copy number alterations from GISTIC showed that diploid and shallow deletions were the top two most common types (Fig. 4A). LEDGF/p75 mutation occurred at a low frequency (Fig. 4B), and specific mutation sites are shown in Fig. 4C.

Subsequently, we analyzed the TMB and MSI of LEDGF/p75 across cancers. The TMB results showed that LEDGF/p75 correlated negatively with 8 tumor types (KIRC, KIRP, LGG, PAAD, PRAD, THCA, UVM, and

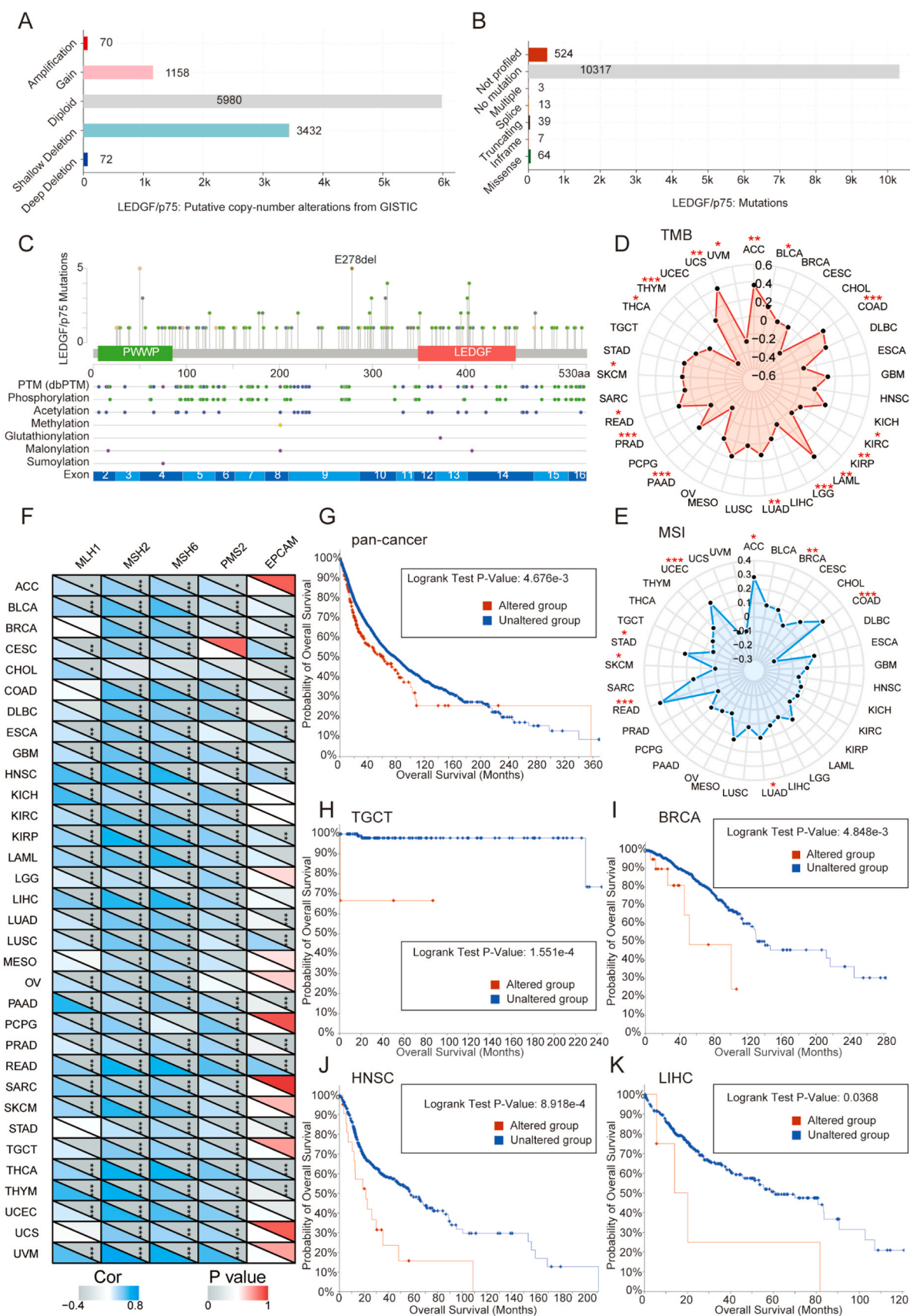


**Fig. 2. Differential expression of LEDGF/p75 in normal and tumor groups.** (A) Differential expression of LEDGF/p75 across cancers from the TCGA database. (B-D) Box plots of LEDGF/p75 expression in tumor and normal (match TCGA normal and GTEx data) groups from the GEPIA database. (E) Differential expression of LEDGF/p75 in different cells (classified according to the source of the organization). The cells of normal tissues are marked in black, and the tumor cells are marked in red. (F-L) Box plots and immunohistochemical diagrams show the differential expression of LEDGF/p75 in 7 tumor types. \*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ .



**Fig. 3. Differential expression of LEDGF/p75 correlates with patient prognosis.** (A) The forest plot shows correlations between LEDGF/p75 expression and the overall survival of patients. (B-E) K–M plots show correlations between LEDGF/p75 expression and the overall survival of patients. (F-L) Box plots show correlations between LEDGF/p75 expression and different stages of patients. Red indicates LEDGF/p75 as a potential oncogene, and blue indicates LEDGF/p75 as a potential protective gene.





**Fig. 4. Genomic instability of LEDGF/p75 is associated with patient prognosis.** (A) Putative copy-number alterations of LEDGF/p75. (B) Mutations in LEDGF/p75. (C) Schematic diagram of LEDGF/p75 mutation sites. (D) The tumor mutation burden (TMB) of LEDGF/p75 across cancers. (E) Microsatellite instability (MSI) of LEDGF/p75 across cancers. (F) The heatmap shows correlations between LEDGF/p75 expression and 5 MMR genes in pancancer. (G–K) K–M plots show that alterations in LEDGF/p75 are associated with the overall survival of patients. \*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ .

THYM) but positively with 8 types (ACC, BLCA, LOAD, LAML, LUAD, READ, SKCM, and UCS) (Fig. 4D). The MSI results showed a positive correlation between 7 tumor types (ACC, BRCA, COAD, LUAD, READ, STAD, and UCEC) and LEDGF/p75 but a negative correlation only between SKCM and LEDGF/p75 (Fig. 4E). Next, we investigated the potential function of LEDGF/p75 in DNA mismatch repair (MMR). As displayed in Fig. 4F, LEDGF/p75 expression correlated highly with MMR genes across cancers, which was consistent with a previously reported result that LEDGF/p75 is involved in DNA damage repair[41].

To further explore the clinical value of LEDGF/p75 gene alterations across cancers, we analyzed their association with patient OS. Our comprehensive analysis showed that patients in the altered group had shorter OS, which was significantly different (Fig. 4G). We then performed a separate analysis of 33 tumor types and found that LEDGF/p75 gene alterations predicted poor patient prognosis in TGCT, BRCA, HNSC, and LIHC (Fig. 4H-K). In conclusion, LEDGF/p75 genomic instability is widespread across cancers and suggests poor prognosis.

### 3.6. Pancancer view of LEDGF/p75 alternative splicing and correlation with patient survival

Alternative splicing (AS) regulates the generation of multiple mRNA and protein products from a single gene. AS plays a crucial role in cancer progression, and cancer cells have general as well as cancer-type-specific and subtype-specific changes during splicing that may have prognostic value and contribute to cancer development and progression[42]. We chose the item (PSIP1\_alt\_3prime\_247053) for subsequent analyses on the OncoSplicing database because it is the only known splice type in SplAdder, a bioinformatics tool for the analysis and quantification of alternative splicing events in RNA sequencing data. The read-in, read-out and PSI values are shown in Fig. 5A, and there were significant differences in LEDGF/p75 AS between tumor and normal tissues. We then visualized the PSI difference in tumor and adjacent normal tissues (Fig. 5B), and the result showed LUSC as the top result. However, the top three changed to LGG, GBM and TGCT when we compared the PSI difference between tumor and GTEx normal tissues (Fig. 5C).

We performed K–M analysis to explore the clinical value of LEDGF/p75 AS. Consistent with our predictions, LEDGF/p75 AS suggested differences in patient prognosis in several tumor types. K–M curves of OS, DSS and PFI showed that a high LEDGF/p75 PSI indicated good prognosis in patients with SKCM (Fig. 5D-F). The same trend was also observed in patients with THCA and CESC (Fig. 5H-I). However, the opposite trend was observed in patients with COAD and LUAD (Fig. 5G, J-K). All of the above results imply the biological importance of LEDGF/p75 AS events across cancers.

### 3.7. LEDGF/p75 is involved in cancer immune infiltration and immunotherapy

To further explore the immunomodulatory effects of LEDGF/p75, we analyzed correlations between immune cells, stromal cells, 22 immune cells, immune checkpoints and LEDGF/p75 expression across cancers (Supplementary Fig. 6); the results showed a tight association between LEDGF/p75 and tumor immunity. Subsequently, we explored whether LEDGF/p75 is differentially expressed in diverse cancer immune subtypes via the TISIDB database. Fig. 6A shows that LEDGF/p75 was significantly associated with immune subtypes in several tumors, and the top six are presented in Supplementary Fig. 5D-I. The detailed subtypes include C1 (wound healing), C2 (IFN-gamma dominant), C3 (inflammatory), C4 (lymphocyte depleted), C5 (immunologically quiet), and C6 (TGF- $\beta$  dominant). In addition, we analyzed the association between immunoinhibitors, chemokines, TILs and LEDGF/p75 expression (Supplementary Fig. 5A-C). As visualized in heatmaps, LEDGF/p75 expression correlated with several immunoinhibitors (CTLA4, IL10, PDCD1, etc.), chemokines (CXCL9, 10, 11, etc.) and TILs (activated CD4, Th2, CD56, etc.) in pancancer.

We searched the TIDE database to evaluate multiple published transcriptomic biomarkers to predict patient response. We list the most confident results about the correlation between LEDGF/p75 expression and CTLs, T dysfunction, and risks in Fig. 6C. The results showed a positive correlation between LEDGF/p75 expression and CTLs in breast cancer but a negative correlation in brain cancer. Moreover, high LEDGF/p75 expression indicated short overall survival for endometrial cancer patients.

We then explored the correlation between LEDGF/p75 and immunotherapy. We compared LEDGF/p75 expression levels across cell lines between pre- and postcytokine-treated samples via the TISMO database, and the box plot is presented in Fig. 6B. Finally, we searched the ROC Plotter database to investigate the correlation between LEDGF/p75 expression and immunotherapy efficiency. Interestingly, high LEDGF/p75 expression indicated effective immunotherapy results in ESCA PD-L1, STAD PD-1, SKCM CTLA-4, and SKCM PD-1 (Fig. 6D-G). Notably, the AUC was higher than 0.65 for ESCA PD-L1 and STAD PD-1, illustrating the high value of the prediction model. In summary, LEDGF/p75 is involved in cancer immune infiltration and immunotherapy, which might guide personalized treatment of tumor patients.

### 3.8. LEDGF/p75 is highly expressed in ccRCC cells and significantly promotes proliferation and metastatic ability

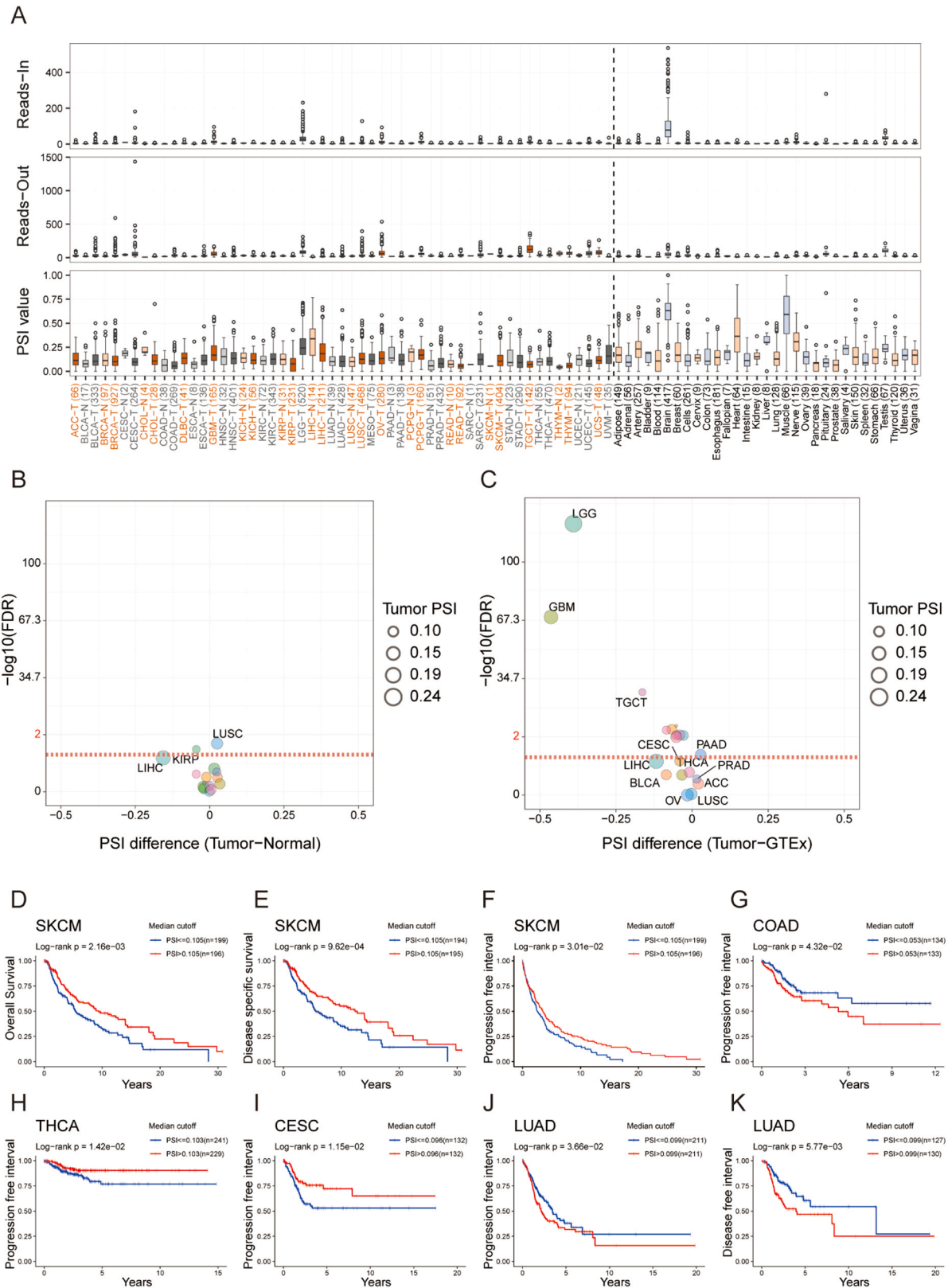
Renal cell carcinoma includes more than 10 histological and molecular subtypes, of which ccRCC is the most common and accounts for the majority of cancer-related deaths[17], and reduction or even deletion of H3K36me3 occurs due to the high mutation rate of SETD2 in patients with advanced ccRCC[43]. This phenomenon naturally stratifies ccRCC patients into subgroups; thus, the study of LEDGF/p75, the reader of H3K36me3, is of great significance. Although studies have reported the vital function of LEDGF/p75 in prostate cancer, leukemia and other kinds of tumors[44–46], there has been no study related to kidney cancer. Therefore, we are the first to conduct preliminary functional exploration of LEDGF/p75 in ccRCC.

Considering the presence of SETD2 mutations in ccRCC cell lines, we searched the DepMap portal database for relevant information. The results revealed SETD2 mutations in A498 and Caki-1 cells: p.V2536fs and p.R400R, respectively. To determine expression of H3K36me3 in ccRCC cell lines, 786-O, A498 and Caki-1 cells were selected for western blotting. As shown in Fig. 7A, H3K36me3 was highly expressed in 786-O and Caki-1 cells but absent in A498 cells, as predicted by the database. For LEDGF/p75, all three ccRCC cell lines expressed higher levels than HK-2 cells. Therefore, we chose 786-O and Caki-1 cells for further studies.

To detect the impact of LEDGF/p75 on ccRCC cell characteristics, we first attempted to knock down LEDGF/p75. Both siRNAs tested achieved > 50% knockdown of the expression level of LEDGF/p75 (Fig. 7B-C). When LEDGF/p75 was knocked down, proliferation and migration ability of ccRCC cells were significantly reduced (Fig. 7D-G), which indicated that LEDGF/p75 is a potential oncogene in ccRCC.

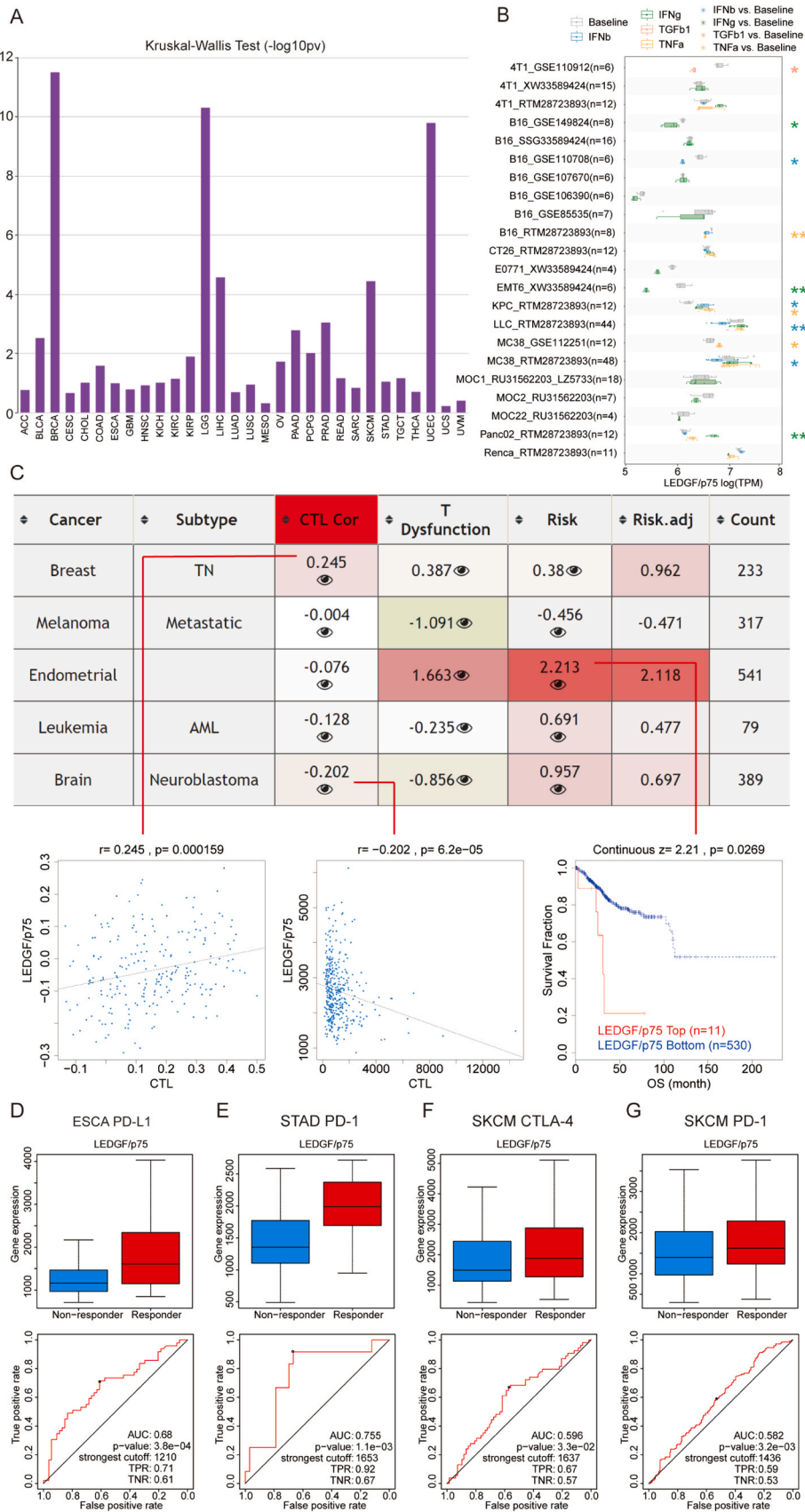
### 3.9. LEDGF/p75 knockdown in ccRCC causes changes to cancer-related genes and pathways

To further explore the role of LEDGF/p75 in ccRCC, we knocked down LEDGF/p75 in 786-O cells and performed gene microarray analysis (Fig. 8A). After LEDGF/p75 was knocked down, 655 genes were upregulated and 512 downregulated (Fig. 8B, Supplementary Table 3). Among them, the most down regulated protein-coding gene was ERO1L. We performed GO and KEGG analyses for all regulated genes. Consistent with our expectations, the transcriptional activity of cells was significantly changed after LEDGF/p75 knockdown (Figs. 8C, 8E). As H3K36me3 is a marker of active transcription[14], knockdown of LEDGF/p75, a reader of H3K36me3, is likely to cause transcriptional inhibition of some downstream genes. Therefore, we reperformed the GO and KEGG analyses of all downregulated genes.

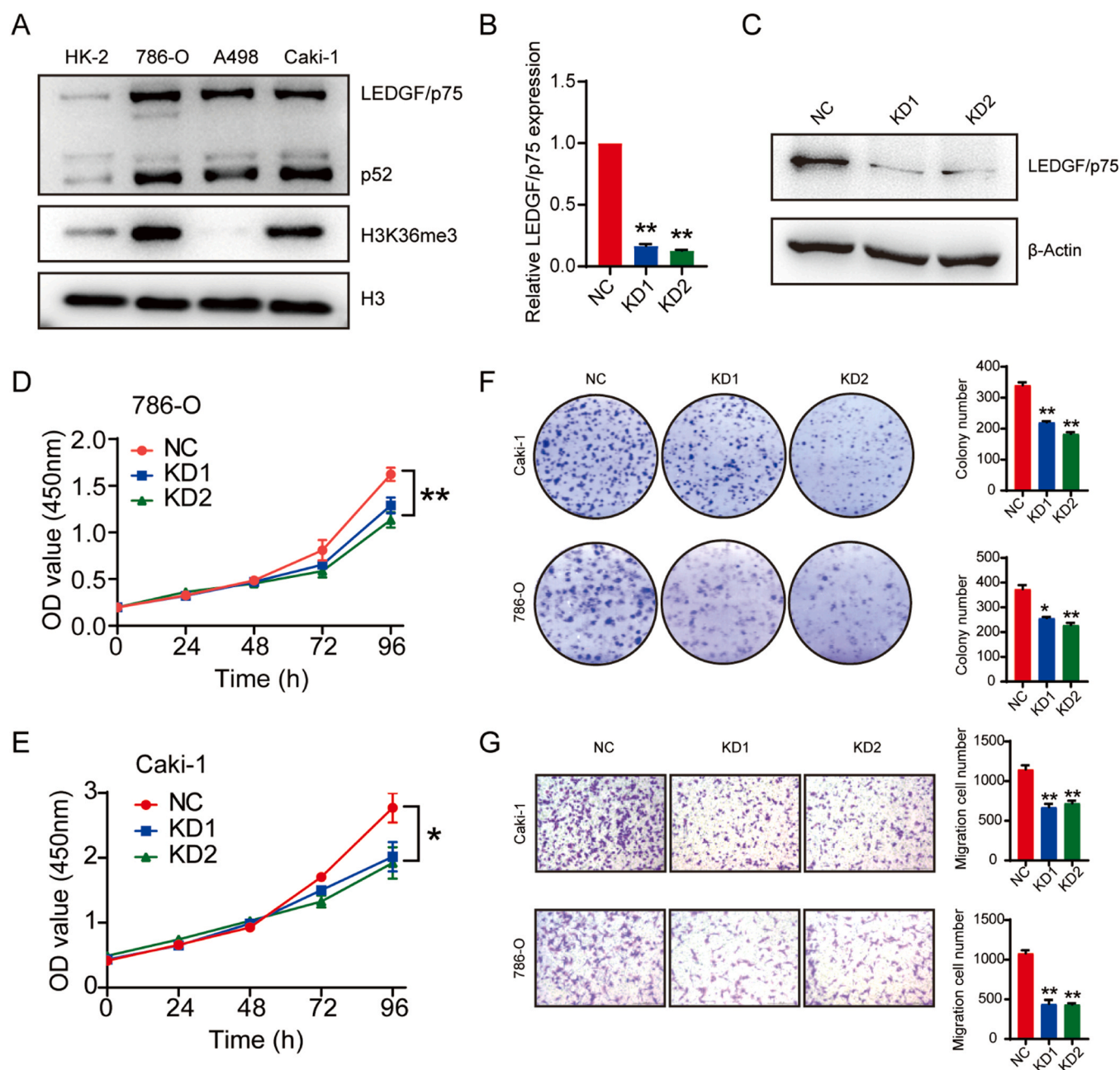


**Fig. 5. LEDGF/p75 alternative splicing correlates with patient prognosis.** (A) Read-in, read-out, and percent spliced in (PSI) values of LEDGF/p75 in pancancer and normal tissues. The red and gray labels represent cancers and adjacent normal tissues, respectively; black labels represent normal tissues. The parts labeled with "Reads-In" and "Reads-Out" on the Y-axis represent read count values, indicating exon splicing in or splicing out, respectively. (B-C) PSI differences between tumor and adjacent normal tissues and between tumor and GTEx normal tissues. The red line refers to 0.05, the dot size represents the tumor PSI value, and different cancers are marked in different colors. (D-K) The PSI value of LEDGF/p75 correlates with patient prognosis in several kinds of cancer.





**Fig. 6. LEDGF/p75 is involved in cancer immunity.** (A) Correlations between LEDGF/p75 and immune subtypes were obtained from the TSIDB database. (B) The box plot retrieved from the TISMO database shows LEDGF/p75 expression levels across cell lines between pre- and postcytokine-treated samples. (C) The table shows the correlation between LEDGF/p75 expression and CTLs, T dysfunction, and risks. The bottom graphs show detailed information on corresponding data in the table. (D–G) Box plots show differential expression of LEDGF/p75 in the indicated groups. ROC curves illustrate the feasibility of LEDGF/p75 expression as an indicator of the effectiveness of immunotherapy. \*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ .



**Fig. 7.** LEDGF/p75 significantly promotes the proliferation and metastatic ability of kidney cancer cells. (A) Expression of LEDGF/p75 and H3K36me3 in kidney cancer cell lines. (B–C) qRT–PCR and western blot experiments show the high efficiency of LEDGF/p75 knockdown. (D–F) CCK-8 and colony formation assays show that knockdown of LEDGF/p75 significantly inhibits proliferation of kidney cancer cells. (G) Transwell experiments show that knockdown of LEDGF/p75 significantly inhibits metastasis of kidney cancer cells.

Interestingly, the results of GO analysis showed that after LEDGF/p75 knockdown, the activity of several functional proteins, such as the Wnt protein, were changed (Fig. 8D). The p53 signaling pathway was also affected by changes in LEDGF/p75 (Fig. 8F).

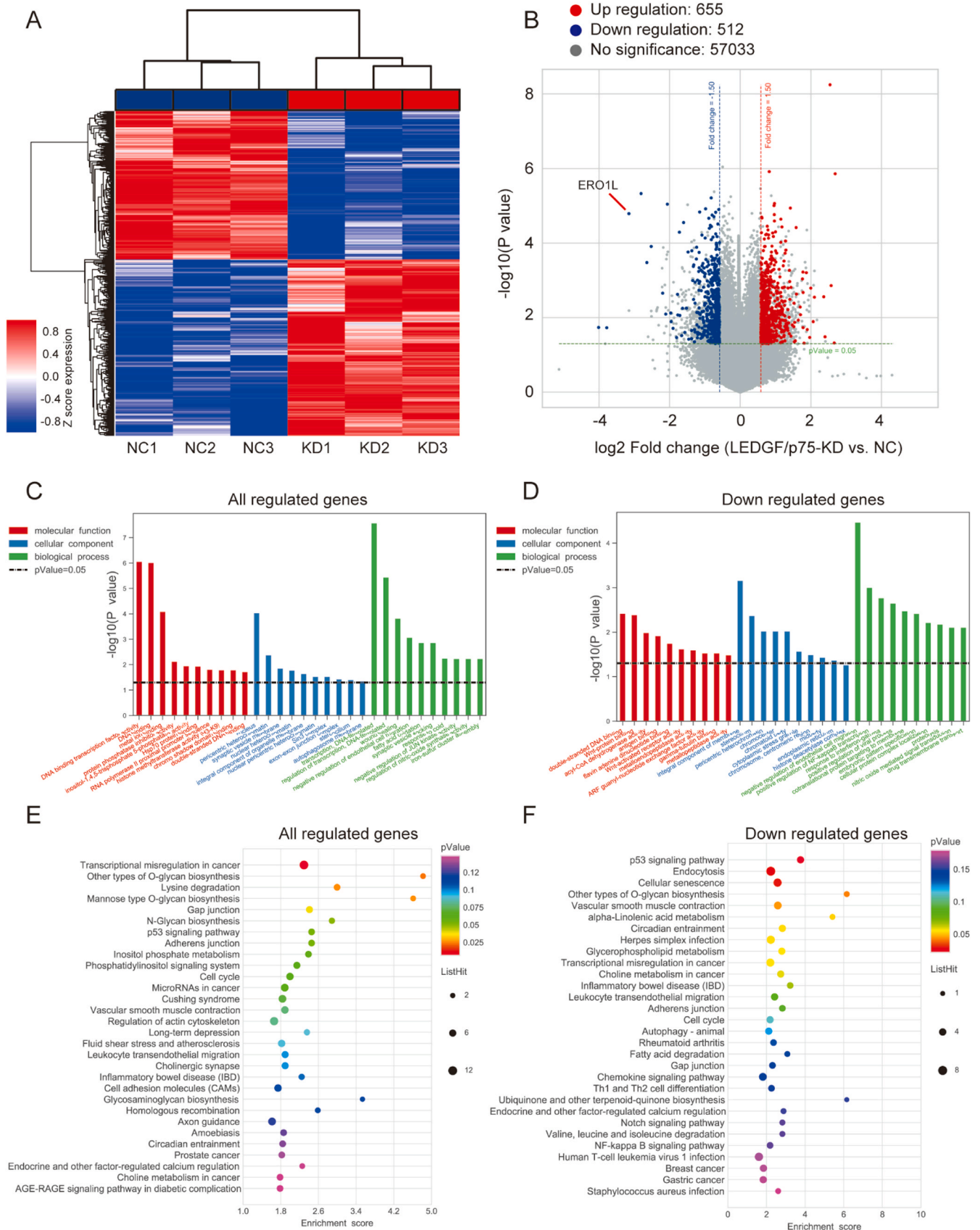
#### 4. Discussion

Studies to date on LEDGF/p75 have mainly focused on HIV and MLL diseases, but there are few studies on its role in tumors. In fact, as a reader of histone modification marks, LEDGF/p75 mediates chromatin binding of many nuclear proteins, thus playing an important biological function in tumors[14]. As a main reader of H3K36me3, its potential role in cancer is worthy of further study. Here, for the first time, we comprehensively analyze the LEDGF/p75 landscape across cancers

using multiple online databases, in vitro experiments and gene microarray sequencing. The present study detailed the basic information, clinical significance, genomic instability, alternative splicing, and cancer immunity related to LEDGF/p75.

The PWWP and IBD of LEDGF/p75 perform the functions of chromatin recognition and protein binding, respectively; specifically, processes of chromatin and DNA binding, transcription regulation, protein–protein interactions, epitope recognition, HIV integration, and stress survival are involved, and homology to the hepatoma-derived growth factor protein family is notable[38]. Accordingly, LEDGF/p75 plays an important role in a variety of biological processes, consistent with the results of our GO analysis, protein interaction prediction, and GSEA across cancers.

LEDGF/p75 expression varied widely among normal groups, tumor



**Fig. 8.** LEDGF/p75 knockdown in ccRCC causes changes in cancer-related genes and pathways. (A) Heatmap of differentially expressed genes in different groups. (B) Volcano plot of differentially expressed genes after knockdown of LEDGF/p75. (C) Gene Ontology (GO) analysis of all regulated genes. (D) GO analysis of downregulated genes. (E) Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of all regulated genes. (F) KEGG analysis of all downregulated genes.



groups and cell lines, suggesting a disease-specific nature of LEDGF/p75. It is worth noting that LEDGF/p75 expression in HEL cells was much higher than that in dozens of other cells. As a cancer cell line derived from myeloid cells, HEL is an erythroleukemia cell line (AML M6 in relapse after treatment for Hodgkin's disease). Whether there is a deeper connection between erythroleukemia and MLL (known to be dependent on LEDGF/p75) other than both being blood cancers and whether that connection is related to LEDGF/p75 is a question that has not yet been answered.

Compared with normal tissues, LEDGF/p75's transcription and protein expression levels were lower in BRCA, LUAD, and UCEC and higher levels in LIHC. However, LEDGF/p75 transcriptional and protein levels in COAD, HNSC, OV, and KIRC did not seem to be uniform. There are many reasons for this. TCGA is a database based on tumor data, and the number of normal tissues is far less than the number of tumor tissues, which may cause a degree of error. In addition, posttranscriptional regulation and posttranslational modification have a great impact on transcription and translation level. Therefore, experiments need to be carried out to verify the findings.

Furthermore, we comprehensively analyzed the OS, DSS, DFI, PFI and clinical stages of patients. High expression of LEDGF/p75 in ACC, KICH, and LIHC suggested poor prognosis, with the opposite in PAAD and SKCM.

Genomic instability leads to the development of tumors, including mutation, structural variant, amplification, deep deletion, and multiple alterations. Our study indicates that genomic alterations in LEDGF/p75 occur in multiple cancers and are associated with poor prognosis in patients with TGCT, BRCA, HNSC, and LIHC. We also analyzed the relationship between LEDGF/p75 and tumor immunity. Our study reports the immune subtypes of LEDGF/p75, the relationship between LEDGF/p75 and immunoinhibitors, chemokines and TILs, and the efficacy of immunotherapy. Our analysis may provide additional treatment options for patients with cancer.

Previous studies have reported that LEDGF/p75 plays different key roles in cancers such as cervical cancer[47], breast cancer[48], ovarian cancer[49], and prostate cancer[50]. However, there is still no report about the function of LEDGF/p75 as a reader of H3K36me3 in kidney cancer.

A high proportion of SETD2 mutations in patients with ccRCC resulted in substantial reduction or even deletion of H3K36me3, naturally grouping patients into clinically and therapeutically relevant subtypes. Therefore, studying LEDGF/p75, a key reader of H3K36me3, is of great significance for clinical diagnosis and treatment of ccRCC patients. We selected 786-O and Caki-1 cells with high H3K36me3 expression to perform functional experiments after LEDGF/p75 was knocked down. The results suggested that LEDGF/p75 is a potential oncogene in ccRCC. Interestingly, LEDGF/p75 is significantly protected from mutations in the ccRCC cohort, which is consistent with its role as an oncogene. Therefore, targeting LEDGF/p75 interference may be a feasible personalized therapy for ccRCC patients without SETD2 mutation. However, the current experimental results are only a preliminary exploration of this hypothesis, and experiments are still needed for rigorous verification in the future.

To further explore the role of LEDGF/p75 in ccRCC, we knocked down LEDGF/p75 and performed gene microarray analysis. Considering that H3K36me3 is a transcriptional activation mark, we focused on the genetic functions that were lowered after LEDGF/p75 knockdown. Interestingly, we found that the p53 signaling pathway changed, which is worthy of further research. After LEDGF/p75 was knocked down, ERO1L was the most significantly decreased protein-coding gene among 512 downregulated genes. Overexpression of ERO1L, an endoplasmic reticulum oxidase, is related to the development and progression of many cancers, such as lung adenocarcinoma, glioblastoma and low-grade glioma, pancreatic ductal adenocarcinoma, and kidney renal papillary cell carcinoma[51]. Whether there is a regulatory axis of LEDGF/p75-ERO1L in SETD2 nonmutant ccRCC is worth exploring.

In summary, we performed multidirectional analysis of LEDGF/p75 across cancers and identified it as a prognostic biomarker. The present study preliminarily explored its potential function in kidney cancer, especially for personalized treatment of patients with SETD2 nonmutant ccRCC.

## Funding

This study was funded by Wuxi Taihu Lake Talent Plan, Leading Talents in Medical and Health Profession Project: Research and application of early screening and accurate diagnosis and treatment of prostate cancer (THRCJH20200104).

## CRedit authorship contribution statement

Ninghan Feng, Bing Yao, and Lin Jiang designed this study and provided clinical guidance as well as data interpretation. Yuwei Zhang, Wei Guo, and Yangkun Feng performed the analyses and experiments. Yuwei Zhang and Wei Guo prepared the figures for this study. Longfei Yang, Hao Lin, Pengcheng Zhou and Kejie Zhao checked the data. Yuwei Zhang drafted the article. All authors reviewed the manuscript, provided comments and approved the final version.

## Declaration of Competing Interest

The authors declare that they have no conflicts of interest.

## Data Availability

The data of this study are available from the corresponding author on reasonable request.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.csbj.2023.08.023](https://doi.org/10.1016/j.csbj.2023.08.023).

## References

- Ge H, Si Y, Roeder RG. Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. *EMBO J* 1998;17:6723–9.
- Nishizawa Y, Usukura J, Singh DP, Chylack Jr LT, Shinohara T. Spatial and temporal dynamics of two alternatively spliced regulatory factors, lens epithelium-derived growth factor (ledgf/p75) and p52, in the nucleus. *Cell Tissue Res* 2001;305:107–14.
- Llano M, Saenz DT, Meehan A, Wongthida P, Peretz M, Walker WH, et al. An essential role for LEDGF/p75 in HIV integration. *Science* 2006;314:461–4.
- Okuda H, Kanai A, Ito S, Matsui H, Yokoyama A. AF4 uses the SL1 components of RNAP1 machinery to initiate MLL fusion- and AEP-dependent transcription. *Nat Commun* 2015;6:8869.
- Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* 2002;110:521–9.
- Ciuffi A, Bushman FD. Retroviral DNA integration: HIV and the role of LEDGF/p75. *Trends Genet* 2006;22:388–95.
- Wang GP, Ciuffi A, Leipzig J, Berry CC, Bushman FD. HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Res* 2007;17:1186–94.
- Canella A, Van Belle S, Brouns T, Nigita G, Carlon MS, Christ F, et al. LEDGF/p75-mediated chemoresistance of mixed-lineage leukemia involves cell survival pathways and super enhancer activators. *Cancer Gene Ther* 2022;29:133–40.
- Yokoyama A, Cleary ML. Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell* 2008;14:36–46.
- De Rijck J, Bartholomeeusen K, Ceulemans H, Debyser Z, Gijssbers R. High-resolution profiling of the LEDGF/p75 chromatin interaction in the ENCODE region. *Nucleic Acids Res* 2010;38:6135–47.
- Sharma S, Cermakova K, De Rijck J, Demeulemeester J, Fabry M, El Ashkar S, et al. Affinity switching of the LEDGF/p75 IBD interactome is governed by kinase-dependent phosphorylation. *Proc Natl Acad Sci USA* 2018;115:E7053–62.
- Huang J, Gurung B, Wan B, Matkar S, Veniaminova NA, Wan K, et al. The same pocket in menin binds both MLL and JUND but has opposite effects on transcription. *Nature* 2012;482:542–6.

- [13] Cermakova K, Demeulemeester J, Lux V, Nedomova M, Goldman SR, Smith EA, et al. A ubiquitous disordered protein interaction module orchestrates transcription elongation. *Science* 2021;374:1113–21.
- [14] Xiao C, Fan T, Tian H, Zheng Y, Zhou Z, Li S, et al. H3K36 trimethylation-mediated biological functions in cancer. *Clin Epigenetics* 2021;13:199.
- [15] Pradeepa MM, Sutherland HG, Ule J, Grimes GR, Bickmore WA. Psp1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLoS Genet* 2012;8:e1002717.
- [16] Turlure F, Maertens G, Rahman S, Cherepanov P, Engelman A. A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo. *Nucleic Acids Res* 2006;34:1653–65.
- [17] Hsieh JJ, Purdue MP, Signoretti S, Swanton C, Albiges L, Schmidinger M, et al. Renal cell carcinoma. *Nat Rev Dis Prim* 2017;3:17009.
- [18] Hsieh JJ, Le VH, Oyama T, Ricketts CJ, Ho TH, Cheng EH. Chromosome 3p Loss-Orchestrated VHL, HIF, and Epigenetic Deregulation in Clear Cell Renal Cell Carcinoma. *J Clin Oncol* 2018;36. JCO2018792549.
- [19] Jonasch E, Walker CL, Rathmell WK. Clear cell renal cell carcinoma ontogeny and mechanisms of lethality. *Nat Rev Nephrol* 2021;17:245–61.
- [20] Zhang Y, Fang Y, Tang Y, Han S, Jia J, Wan X, et al. SMYD5 catalyzes histone H3 lysine 36 trimethylation at promoters. *Nat Commun* 2022;13:3190.
- [21] Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr Protoc Bioinforma* 2016;54. 1 30 31-31 30 33.
- [22] C. UniProt, UniProt: the Universal Protein Knowledgebase in 2023, *Nucleic Acids Res*, (2022).
- [23] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;596:583–9.
- [24] Uhlen M, Fagerberg L, Hallstrom BM, Lindskog P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. *Science* 2015;347:1260419.
- [25] Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* 2010;38. W214-220.
- [26] Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* 2021; 49:D605–12.
- [27] Mitsopoulos C, Di Micco P, Fernandez EV, Dolciami D, Holt E, Mica IL, et al. canSAR: update to the cancer translational research and drug discovery knowledgebase. *Nucleic Acids Res* 2021;49:D1074–82.
- [28] Ru B, Wong CN, Tong Y, Zhong JY, Zhong SSW, Wu WC, et al. TISIDB: an integrated repository portal for tumor-immune system interactions. *Bioinformatics* 2019;35:4200–2.
- [29] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005;102:15545–50.
- [30] Goldman MJ, Craft B, Hastie M, Repecka K, McDade F, Kamath A, et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol* 2020; 38:675–8.
- [31] 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 Res* 2017; 45:W98–102.
- [32] Chandrashekar DS, Karthikeyan SK, Korla PK, Patel H, Shovon AR, Athar M, et al. UALCAN: an update to the integrated cancer data analysis platform. *Neoplasia* 2022;25:18–27.
- [33] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012;2:401–4.
- [34] Zhang Y, Yao X, Zhou H, Wu X, Tian J, Zeng J, et al. OncoSplicing: an updated database for clinically relevant alternative splicing in 33 human cancers. *Nucleic Acids Res* 2022;50:D1340–7.
- [35] Fu J, Li K, Zhang W, Wan C, Zhang J, Jiang P, et al. Large-scale public data reuse to model immunotherapy response and resistance. *Genome Med* 2020;12:21.
- [36] Zeng Z, Wong CJ, Yang L, Ouaraoui N, Li D, Zhang W, et al. TISMO: syngeneic mouse tumor database to model tumor immunity and immunotherapy response. *Nucleic Acids Res* 2022;50:D1391–7.
- [37] Fekete JT, Gyorfy B. ROCplot.org: Validating predictive biomarkers of chemotherapy/hormonal therapy/anti-HER2 therapy using transcriptomic data of 3,104 breast cancer patients. *Int J Cancer* 2019;145:3140–51.
- [38] Cermakova K, Weydert C, Christ F, De Rijck J, Debysers Z. Lessons Learned: HIV Points the Way Towards Precision Treatment of Mixed-Lineage Leukemia. *Trends Pharm Sci* 2016;37:660–71.
- [39] Wang H, Farnung L, Dienemann C, Cramer P. Structure of H3K36-methylated nucleosome-PWWP complex reveals multivalent cross-gyre binding. *Nat Struct Mol Biol* 2020;27:8–13.
- [40] El Ashkar S, Schwaller J, Pieters T, Goossens S, Demeulemeester J, Christ F, et al. LEDGF/p75 is dispensable for hematopoiesis but essential for MLL-rearranged leukemogenesis. *Blood* 2018;131:95–107.
- [41] Ui A, Chiba N, Yasui A. Relationship among DNA double-strand break (DSB), DSB repair, and transcription prevents genome instability and cancer. *Cancer Sci* 2020; 111:1443–51.
- [42] Bonnal SC, Lopez-Oreja I, Valcarcel J. Roles and mechanisms of alternative splicing in cancer - implications for care. *Nat Rev Clin Oncol* 2020;17:457–74.
- [43] Xie Y, Sahin M, Sinha S, Wang Y, Nargund AM, Lyu Y, et al. SETD2 loss perturbs the kidney cancer epigenetic landscape to promote metastasis and engenders actionable dependencies on histone chaperone complexes. *Nat Cancer* 2022;3: 188–202.
- [44] Ortiz-Hernandez GL, Sanchez-Hernandez ES, Ochoa PT, Elix CC, Alkashgari HR, McMullen JRW, et al. The LEDGF/p75 Integrase Binding Domain Interactome Contributes to the Survival, Clonogenicity, and Tumorsphere Formation of Docetaxel-Resistant Prostate Cancer Cells. *Cells* 2021;10.
- [45] Roudaia L, Speck NA. A MENage a Trois in leukemia. *Cancer Cell* 2008;14:3–5.
- [46] Dagaard M, Kirkegaard-Sorensen T, Ostenfeld MS, Aaboe M, Hoyer-Hansen M, Orntoft TF, et al. Lens epithelium-derived growth factor is an Hsp70-2 regulated guardian of lysosomal stability in human cancer. *Cancer Res* 2007;67:2559–67.
- [47] Leitz J, Reuschenbach M, Lohrey C, Honegger A, Accardi R, Tommasino M, et al. Oncogenic human papillomaviruses activate the tumor-associated lens epithelial-derived growth factor (LEDGF) gene. *PLoS Pathog* 2014;10:e1003957.
- [48] Singh DK, Gholamalamdari O, Jadaliha M, Ling X, Lin Li, YC, Zhang Y, et al. PSIP1/p75 promotes tumorigenicity in breast cancer cells by promoting the transcription of cell cycle genes. *Carcinogenesis* 2017;38:966–75.
- [49] Sapoznik S, Cohen B, Tzuman Y, Meir G, Ben-Dor S, Harmelin A, et al. Gonadotropin-regulated lymphangiogenesis in ovarian cancer is mediated by LEDGF-induced expression of VEGF-C. *Cancer Res* 2009;69:9306–14.
- [50] Liedtke V, Rose L, Hiemann R, Nasser A, Rodiger S, Bonaventura A, et al. Over-expression of LEDGF/p75 in HEP-2 cells enhances quitoimmune IgG response in patients with benign prostatic hyperplasia-a novel diagnostic approach with therapeutic consequence? *Int J Mol Sci* 2023;24.
- [51] Shergalis AG, Hu S, Bankhead 3rd A, Neamati N. Role of the ERO1-PDI interaction in oxidative protein folding and disease. *Pharm Ther* 2020;210:107525.