Liu et al. Journal of Translational Medicine (2025) 23:1196
https://doi.org/10.1186/s12967-025-07111-z

Journal of Translational Medicine

RESEARCH Open Access

Genome-wide association, single-cell, and spatial transcriptomics analyses reveal the role of the STK24-expressing positive cells in LUAD progression and the tumor microenvironment, identifying STK24 as a potential therapeutic target

Junzhi Liu[1*†], Huimin Li[2†], Yuheng Jiao[3] and Meng Hou[1*]

Abstract

Background Lung adenocarcinoma (LUAD), a subtype of non-small cell lung cancer (NSCLC), has high incidence and poor prognosis. Although anti-programmed cell death protein 1 (PD1) therapy and epidermal growth factor receptor (EGFR) inhibitors benefit some patients, many remain unresponsive. Genome-wide association studies (GWAS) can identify risk loci, while single-cell transcriptomics enables exploration of genetic variations and mechanisms in LUAD.

Methods We integrated GWAS data from FinnGen, IEU OpenGWAS, and GWAS Catalog with single-cell and spatial transcriptomics from GEO to investigate LUAD genetics and pathology. GWAS analyses identified SNP loci, risk factors, LD Score, and SMR results, revealing a novel apoptosis-related locus, Serine/Threonine Kinase 24 (STK24) (SMR $P < 0.05$). Single-cell analysis identified STK24-positive epithelial cells (STK24posEpi) and their transcription factors (SCENIC), developmental trajectory (Monocle), and tumor microenvironment roles (CellCall). Spatial mapping (RCTD) and trajectory analysis (stlearn) were combined with cell communication (CellChat), spatial signal flow (Commot), and co-localization (Misty). Bayesian deconvolution linked STK24posEpi to TCGA-LUAD bulk transcriptomes, assessing overall survival (OS), immune infiltration, and clinical traits. A prognostic model (STK24SuperPC, mean c-index 0.61) was built using multiple machine learning algorithms to predict survival and immunotherapy response. Laboratory experiments tested STK24's effects on LUAD cell proliferation and apoptosis.

†Junzhi Liu and Huimin Li contributed equally to this work.

*Correspondence:
Junzhi Liu
junzhil97@163.com
Meng Hou
houmengdr@163.com

Full list of author information is available at the end of the article

Results STK24 was identified as a LUAD risk factor (pSMR < 0.05). STK24posEpi communicated with fibroblasts, endothelial cells, and others via MIF and PDGF pathways, and was associated with apoptosis (Importance > 0.5), metastasis (Importance > 2), DNA repair (Importance > 1), and DNA damage (Importance > 2). Located at the origin of the trajectory, STK24posEpi correlated with poor prognosis (HR > 0, $P < 0.05$) and reduced lymphocyte infiltration (Pearson's r < 0, $P < 0.05$). STK24 overexpression enhanced LUAD cell proliferation, migration, and colony formation, while reducing apoptosis, increasing BCL-2 and Caspase-3, and decreasing Cleaved-Caspase-3 expression.

Conclusions STK24, an apoptosis-related gene, promotes LUAD progression by enhancing proliferation and inhibiting apoptosis, with STK24posEpi interacting in the tumor microenvironment through PDGF and MIF pathways. These findings offer potential therapeutic targets and support personalized LUAD treatment strategies.

Keywords Single-cell RNA sequencing, Spatial transcriptomics, Cell apoptosis, STK24, BayesPrism, Biomarker

Introduction

Lung adenocarcinoma (LUAD) is one of the most common types of non-small cell lung cancer (NSCLC). Statistics show that lung cancer is the most common cancer worldwide and the leading cause of cancer-related deaths [1]. Within NSCLC, lung adenocarcinoma occupies a considerable proportion, and its incidence has been rising in recent years [2]. Molecular characteristics and genetic variations in lung adenocarcinoma, such as epidermal growth factor receptor (EGFR) mutations, have been widely studied, with these mutations playing a critical role in the occurrence and development of lung adenocarcinoma [3]. Targeted therapies for these mutations have become the standard treatment for patients with EGFR mutations, significantly improving their survival rates and quality of life [3]. Additionally, the prognosis of lung adenocarcinoma is closely related to its histological characteristics. Studies have shown that non-terminal respiratory unit (non-TRU) types of lung adenocarcinoma are associated with poorer prognosis [2]. In terms of treatment, although options for targeted therapy and chemotherapy continue to expand, therapeutic efficacy remains limited for certain subtypes of lung adenocarcinoma, particularly those exhibiting phenotypic transformation [4]. Therefore, understanding the biological characteristics and clinical manifestations of lung adenocarcinoma is essential for formulating personalized treatment plans [5].

Advancements in human genetics have mitigated the risks of residual confounding and reverse causation, while large-scale genomic analyses offer insights into potential causal mechanisms for complex phenotypes like LUAD [2]. This analytical approach combines techniques such as GWAS and Mendelian randomization, which can reveal genetic variations associated with disease and their biological implications. Through studies with large sample sizes, scientists can identify key proteins and genes related to LUAD, thereby providing crucial clues for future therapeutic targets [6]. Moreover, advances in functional genomics have enabled researchers to gain deeper insights into how these genetic variations affect disease occurrence and development [7]. This approach is not limited to LUAD but can also be applied to studies of other complex diseases, helping to reveal underlying biological mechanisms and therapeutic strategies [8, 9].

During apoptosis, damaged or abnormal cells are eliminated, which maintains tissue homeostasis. Dysregulation of apoptotic pathways is a key factor in LUAD tumorigenesis. Research has shown that apoptosis plays a complex role in cancer development and progression, as it can inhibit tumor formation but, in certain cases, also promote tumor progression [10]. For instance, the loss of apoptosis can lead to the survival of malignant cells, thereby promoting tumor growth and metastasis [11]. Additionally, cells may employ various mechanisms to evade apoptosis, such as anoikis resistance, mitochondrial DNA depletion, and regulation by Cellular FLICE-like inhibitory (c-FLIP) protein [12]. These escape mechanisms allow tumor cells to resist treatment, adding to the complexity of cancer therapy [13].

As a key factor in the apoptosis pathway, STK24 (serine/threonine protein kinase 24) has been found to play an important regulatory role in various cancers. Research shows that STK24 is highly expressed in lung adenocarcinoma (LUAD) and is associated with poor prognosis in patients [14]. By regulating KLF Transcription Factor 5 (KLF5), STK24 promotes the proliferation and migration of lung cancer cells [14]. In gastric cancer, knockdown of STK24 increases cell migration and liver metastasis, and promotes tumor progression by inhibiting the CDH1 gene and enhancing stemness [15]. Additionally, STK24 may be involved in the immune regulation process of LUAD, making it a potential therapeutic target [14]. In tumor immune evasion, STK24 exerts its effect through the AKT-PD-L1 axis [16]. The deletion of STK24 leads to a significant reduction in tumor growth, a process that depends on cytotoxic CD8 + T cells and NK cells. STK24 binds with AKT and directly phosphorylates AKT, promoting the induction of PD-L1. Inhibiting STK24 can significantly enhance the effectiveness of anti-PD-1 blockade therapy [16]. In non-small cell lung cancer (NSCLC), STK24 induces tumorigenesis by

regulating the STAT3/VEGFA signaling pathway [17]. The high expression of STK24 is positively correlated with the proliferation, migration, invasion, and tumor angiogenesis capacity of cancer cells. Inhibiting STK24 can significantly suppress tumor progression and angiogenesis [17]. Moreover, the expression of STK24 is regulated by DNA copy number and methylation [18]. In lung adenocarcinoma, abnormal expression of STK24 is an independent indicator of poor prognosis, with dysregulation associated with changes in DNA copy number and methylation [18]. Finally, the interaction network between STK24 and Cerebral Cavernous Malformations 3 Protein (CCM3) plays an important role in the vesicle exocytosis of neutrophils [19]. STK24 inhibits exocytosis by competing with the lipid-binding C2B domain of UNC13D, while CCM3 reversely regulates STK24-mediated inhibition of exocytosis through its calcium-sensitive interaction with the C2A domain of UNC13D [19]. In conclusion, STK24 plays a crucial role in various cancers through multiple mechanisms. A deeper understanding of its regulatory network may provide new strategies for cancer treatment.

We conducted a meta-analysis of GWAS for LUAD, as well as risk factor identification and new locus analysis for LUAD. We found that STK24, associated with apoptosis, showed significant co-localization with LUAD and was positively correlated with LUAD, suggesting that STK24 is a risk gene for LUAD. Subsequently, we used single-cell transcriptomics to classify tumor cells based on STK24 expression, defining STK24-positive epithelial cells (STK24posEpi), and further analyzed their heterogeneity, functional status, and role in the tumor microenvironment. By integrating spatial transcriptomics with deconvolution techniques, we analyzed the spatial distribution characteristics, cell communication, signaling flows, and spatial evolutionary trajectories of STK24posEpi, validating these findings with single-cell transcriptomic analysis. Additionally, we explored the relationship of STK24posEpi with clinical features such as tumor staging and grading through Bayesian deconvolution counting, and developed a prognostic model based on deep learning algorithms using the characteristic genes of STK24posEpi.

The objective of this study is to determine the role of STK24 as a risk gene in LUAD through GWAS analysis, explore its cellular and molecular mechanisms using single-cell transcriptomics, and analyze the spatial distribution and role of STK24-positive epithelial cells (STK24posEpi) in the tumor microenvironment using spatial transcriptomics. We also aim to validate these findings through laboratory experiments and ultimately develop a prognostic model based on the characteristics of STK24posEpi by analyzing its correlation with clinical features.

Methods and materials

LUAD genome-wide association study

We conducted a fixed-effects inverse-variance weighted meta-analysis on LUAD utilizing data from the published studies GCST004744 (n = 66,756), finn-b-C3_NSCLC_ADENO (n = 218,792), and ieu-a-984 (n = 65,864), encompassing a total sample size of 351,412 individuals. The analysis was performed using the METAL software [20]. Variants with a minor allele frequency (MAF) of less than 0.5% were excluded, yielding a total of 10,227,138 associations. Annotation of our results was conducted using FUMA [21] with its default settings. Following the standard parameters of FUMA, distinct variants were characterized by an R^2 value of less than 0.001. Associations were identified as being more than 500 kilobases away from any previously reported indexed variant in datasets GCST004744, finn-b-C3_NSCLC_ADENO, and ieu-a-984. Prior research [22] led us to identify the gene closest to an indexed variant and the gene identified by POPS as the top gene for GWAS-significant loci. With Ensembl genes within a 500 kb window, we chose the gene with the highest PoP score as the prioritized gene. We then performed enrichment analysis with Metascope on all genes suggested by the nearest gene and PoPs.

Risk factors associated with LUAD GWAS variants

For genetic variants meeting the GWAS threshold for LUAD ($P < 5 \times 10 - 8$), we assessed genetic associations with six LUAD risk factors using data from the IEU OpenGWAS project [23] and the NHGRI-EBI GWAS Catalog. These risk factors include: tuberculosis (TB) [23], smoking [24], Idiopathic pulmonary fibrosis (IPF) [25], COPD (chronic obstructive pulmonary disease) [23], Worked with materials containing asbestos (Asbestos) [26], and Nitrogen oxides air pollution (Airpollution) [23].

SMR analysis was conducted to identify LUAD-related pathogenic genes

The SMR analysis is described in detail by Zhu et al. [27]. In our study, we used eQTL data from lung tissue in the GTEx database as our eQTL dataset. SMR and HEIDI tests assessed cis-eQTL enrichment in LUAD, with a genome-wide significance threshold of PSMR < 0.05 for SMR detection. A PHEIDI > 0.05 indicated no gene heterogeneity. SMR analysis utilized SMR software.

LD (linkage disequilibrium) score regression

Utilizing European LD scores from 1000 Genomes Project Phase 3 for HapMap2 SNPs, we assessed genetic correlations between lung adenocarcinoma and six lung cancer-related traits using LD Score regression [28]. A standardized association statistics program (MungeSumstats) was used [29].

Data acquisition

The scRNA-seq datasets (GSE127465, GSE131907, GSE149655, and GSE148071) and bulk transcriptome data (GSE31210, GSE37745, GSE41271, GSE68465, GSE72094, GSE42127, GSE72094, GSE11969) were obtained from the GEO database. Bulk transcriptome data TCGA-LUAD were obtained from the UCSC Xena database. Spatial transcriptomics data (E-MTAB-13530) were obtained from the ArrayExpress database. GWAS data for GCST004744 were obtained from the GWAS Catalog database. GWAS data for ieu-a-984 were obtained from the OpenGWAS database. GWAS data for finn-b-C3_NSCLC_ADENO were obtained from the FinnGen database (R9).

scRNA-seq analysis

A scRNA-seq dataset was processed using Seurat R [30]. The following quality exclusion demand were carried out: (a) genes detected in less than three cells, (b) cells expressing less than 50 genes in total, and (c) cells having ≥ 5% of mitochondria-expressed gene. The data were normalized using the SCTransform method. Harmony was used to correct batch effects and integrate Seurat objects [31]. The expression matrix dimensionality was reduced using principal component analysis (PCA), focusing on the top 2,000 highly variable genes, which were selected using the FindVariableFeatures function (selection.method = "vst"). The optimal number of dimensions, based on the standard deviations shown in the ElbowPlot function, was used as the input for the FindNeighbors function (dims = 15). Cluster analysis was performed using the FindClusters function with a resolution of 1.5, which computes the nearest neighbors. Unsupervised clustering was conducted and visualized using Uniform Manifold Approximation and Projection (UMAP) [32]. Cluster annotations were determined based on the expression of canonical markers. Meanwhile, cell-cycle scores were meticulously calculated using Seurat's CellCycleScoring function.

Recognition of cell type

Cell cluster-based differential expression analysis was deployed on all genes using the FindAllMarkers function of Seurat to screen the marker genes intrinsic to each cluster [30]. Adjusted P value < 0.05, the absolute value of the $\log_2$ (fold change, FC) > 0.25, and expression percentage > 0.25 were deemed as thresholds. As well as finding and annotating diverse cell subclusters according to composition patterns, the singleR package identified a variety of cell subclusters. All results were manually verified, rectified, and corroborated according to the CellMarker database.

Cell to cell communication analysis

CellCall is a toolkit designed to infer intercellular communication networks and internal regulatory signals by integrating both intracellular and intercellular signals [33]. (1) CellCall collects ligand-receptor-transcription factor (L-R-TF) axis data based on the KEGG pathway dataset. (2) By using prior knowledge of L-R-TF interactions, CellCall infers intercellular communication by combining the expression of ligands/receptors and the activity of downstream transcription factors of certain L-R pairs. (3) CellCall incorporates a pathway activity analysis method to identify key pathways involved in communication between certain cell types. (4) CellCall offers a rich set of visualization options (e.g., Circos plot, Sankey plot, bubble plot) to visually present the results of intercellular communication.

Single-cell trajectory construction

Using the Monocle2 R package (v2.26.0), trajectory analyses were conducted at single-cell level, assuming that a unidimensional 'time' parameter would reveal high-dimensional expression values to explain cell state transitions [34]. To establish an object (expressionFamily = negbinomial.size), epithelial cell subclusters were incorporated into the R computational condition. Based on Monocle2, mean_expression and dispersion_empirical were used to categorize cellular types in a pseudotime sequence based on the dispersion_fit. Indexes were condensed using the reduceDimension() function with parameters reduction_method = "DDRTree" and max_components = 2, and the plot_cell_trajectory visualization function was used to display the minimum spanning tree on cells. Genes with pseudotime-related variations were identified using the "differentialGeneTest" function (q-val < 10 – 5) and displayed on the plot_pseudotime_heatmap. These genes were also categorized into subgroups based on their expression patterns.

Single-cell regulatory network inference and clustering (SCENIC) of transcription factors (TFs)

The SCENIC algorithm [35] is crafted to analyze regulatory frameworks linked to TFs and identify regulons of TFs and their target genes. It uses a log-normalized expression matrix from Seurat, with genes and cells in columns, as input. A motif dataset (hg19-tss-centered-10 kb-7species.mc9nr.feather) is then used to identify suitable regulons for each TF, which are then correlated with potential targets via Spearman's correlation. The "runSCENIC" method was then used to build Gene Regulatory Networks, or regulons. Finally, AUCell software evaluated regulon activity, using a default threshold to represent TF states: '0' for 'off' and '1' for 'on.'

Cellular trajectory reconstruction (CytoTRACE) analysis

The Cellular (Cyto) Trajectory Reconstruction Analysis using gene Counts and Expression (CytoTRACE) tool represents a modern approach to scRNA-Seq data analysis, facilitating the simplification and quantification of gene expression levels that are related to gene counts. It gives each cell a stemness score, effectively predicting differentiation states and surpassing earlier methods in large datasets [36]. Using the CytoTRACE R package version 0.3.3, tumor cell scores were calculated on a scale from 0 to 1, where higher scores suggest increased stemness and lower scores suggest decreased stemness.

Spatial transcriptome analysis

The spatial transcriptomics (ST) data was handled and illustrated using Seurat R package. We deleted the genes related to mitochondria, ribosomes, and red blood cells, as well as genes with expression levels of less than 10 spots. The data was then normalized using the SCTransform technique, followed by dimensionality reduction using Principal Component Analysis (PCA). We used FindClusters and hematoxylin–eosin (HE) staining to select the appropriate number of cell clusters, and the clusters were visualized using RunUMAP. Annotations of cell populations were based on HE staining sections and genes that vary by cluster. The ST data was analyzed using SpatialDimPlot and SpatialFeaturePlot functions.

Analysis of ST data using RCTD

Using RCTD (Robust cell type decomposition), cell types from the reference scRNA-seq dataset were compared with the ST data [37]. The identification of unique marker genes for each cell type was facilitated through the application of Seurat function FindAllMarkers, with the analysis limited to markers demonstrating positive log2-transformed fold changes. In addition, we adhered strictly to the standard RCTD analysis pathway, focusing on both the reference and Visium ST data in full doublet mode.

Spatial data functional information estimation

For individual sites, signaling pathway engagements were estimated using Pathway RespOnsive GENes for activity inference's (PROGENy) [38, 39] model matrix (v1.7-1), based on the top 1000 genes from each transcriptional dataset and the normalized data.

Dependencies between cells on a spatial map

The Multiview Intercellular SpaTial modeling framework (MISTy) [40] code was used in mistyR (v1.14.0) for estimating how the preponderance of primary cell subtypes effects the presence of other cell types. Our model was based on three spatial contexts: (i) Assessing correlations among deconvolution estimates in a specific area, (ii) a juxta view gathering observable deconvolution estimates from nearby neighbors (up to 5 units away), and (iii) a para view assigning weight to deconvolution estimates from more distant neighbors (within 15 units). In contrast, the median standardized importance of each perspective across slides was interpreted as depending on cell types in different spatial contexts, such as colocalization or mutual exclusion, without implying a causal relationship. Predictors with an R^2 below 10% were excluded before aggregation for each slide.

A MISTy model was established to display the distribution of PROGENy pathway activity scores, aiming to link tissue structure with functional traits. The analysis methods were as described in previous literature [41].

Analyzing spatial trajectory

Using stLearn and PAGA (Partition-based graph abstraction) trajectory analysis on SME-normalized data, the pseudotime trajectory was determined to identify links within subclusters [42]. The advancement of malignancy across the sections was represented using a pseudotime spatial trajectory algorithm, thereby distinguishing connections amid the sub-clusters at spatial and transcriptional level.

Deconvolution analysis with BayesPrism

We performed deconvolution analysis using BayesPrism [43], a Bayesian modeling framework designed for cell type composition estimation. BayesPrism leverages single-cell RNA-seq data to generate cell type-specific signatures, allowing for accurate decomposition of bulk RNA-seq samples into distinct cellular components. Count data or TPM data were used for the TCGA-LUAD dataset and GEO dataset, respectively, to ensure consistency.For single-cell reference data, stringent quality control steps were applied, including filtering for low-quality cells, doublets, and other artifacts. A reference matrix was constructed from the single-cell RNA-seq data by aggregating cell type-specific gene expression profiles. Genes with high cell type specificity were prioritized to enhance deconvolution accuracy. Additionally, batch effect corrections were applied as necessary to mitigate any biases introduced by differences across datasets. BayesPrism was employed to perform the deconvolution, using the single-cell reference as a guide. The parameters were optimized to maximize the robustness of cell type proportion estimates across different cohorts. For each sample, BayesPrism produced posterior estimates of cell type proportions. The deconvolution results from TCGA-LUAD and GEO datasets were integrated and analyzed to assess the consistency of cell type distributions across the cohorts. Differences in cell type proportions between LUAD samples and normal controls were

statistically evaluated to identify cell types potentially involved in LUAD pathogenesis.

Machine learning algorithms

The method for establishing the prognostic model was as described in previous literature [44]. In addition, we used timeROC to evaluate the predictive performance of the model and performed survival analysis using the survival package.

Immunotherapy and potential drug prediction

The Tumor Immune Dysfunction and Exclusion (TIDE) algorithm was used to predict immunotherapy outcomes in a large transcriptomic cohort. The datasets CTRP (Cancer Therapeutics Response Portal) and PRISM (profiling relative inhibition simultaneously in mixtures) were used to gather drug sensitivity data for cancer cell lines.

Cell culture

In RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS; HyClone), A549 and PC9 cells were obtained from the Cell Bank of the Chinese Academy of Sciences (Shanghai) and cultured under tightly controlled conditions of 5% CO_2 at 37 °C.

Establishment of STK24-overexpressing cell lines

To overexpress STK24 expression, the full-length cDNA sequence of STK24 was amplified by PCR and then subcloned into pLVX vector (GenePharma), termed pLVX -STK24. The empty pLVX vector was used in control group. A549 and PC9 cells were transfected with pLVX-STK24-Puro and pLVX-Vector-Puro plasmids using Lipofectamine 3000. After 48 h of transfection, cells were selected with 10 μg/mL puromycin at a 1:1000 dilution ratio. Western blotting was then used to assess STK24 expression in various monoclonal cultures.

Cell viability assay

We performed digestion of adherent tumor cells (with 0.25% trypsin (Trypsin, Gibco) for 1 min, terminated by adding complete RPMI 1640 medium (supplemented with 10% fetal bovine serum)), collected the cells into a centrifuge tube, centrifuged, and counted the cells. Finally, we seeded the cells into a 96-well plate at a density of 2000–3000 cells per well. Additionally, according to the manufacturer's instructions, we assess cell viability using the Cell Counting Kit-8 (APExBIO, USA) at 24, 48, 72, and 96 h. For these treated cells, a concentration gradient of 0, 2, 4, 6, 8, 10, and 12 μg/mL of the drug is used for analysis. This systematic approach allows for the accurate monitoring of cell proliferation over time.

Western blotting (WB)

Cells are carefully lysed in an ice-cold buffer, meticulously formulated with both phosphatase and protease inhibitors to preserve the integrity of the proteins. To determine the protein concentration, the BCA assay is employed, providing a reliable measure of protein levels. Next, the proteins are resolved using a 4–12% SDS-PAGE gradient, enabling their separation based on size, and are subsequently transferred to a PVDF membrane for further analysis. Once on the membrane, a blocking step is performed to prevent nonspecific binding, followed by two rounds of incubation with specific antibodies, ensuring the precision of the detection. The immunoreactive proteins are then visualized through a chemiluminescence solution, enabling the clear identification of the target proteins. The antibodies utilized in this experiment include anti-STK24 (Solarbio, K002053P, at a dilution of 1:1000), anti-Caspase-3 (Abmart, M005851, also at 1:1000), anti-Cleaved Caspase-3 (Abmart, MB0711, at 1:1000), anti-BCL-2 (Solarbio, K003505P, diluted 1:1000), and the ubiquitous anti-β-actin (Santa, sc-58673, 1:1000). Each of these antibodies plays a pivotal role in detecting the desired proteins, contributing to the precision of the immunoblotting procedure.

Transwell assays

To assess cellular migration, the Boyden chamber, featuring 8-μm pores, was employed. In this experiment, a suspension of 1×10^5 A549 and PC9 cells was introduced into the upper compartment of the chamber, where they were bathed in 200 μL of serum-free medium. Meanwhile, the lower chamber was filled with a nutrient-rich medium supplemented with 10% fetal bovine serum (FBS), providing a contrasting environment. After a precise incubation period of 24 h, the cells underwent fixation and staining, preparing them for quantification. Migration was then assessed by counting the number of cells that had traversed the porous membrane, with data collected from six randomly selected fields under a microscope.

The 5-ethynyl-2′-deoxyuridine (EdU) assay

After transfection, cells were seeded into 24-well plates (5×10^4 cells per well) and cultured overnight in an incubator at 37 °C. Cell proliferation was assessed using the BeyoClick EdU Cell Proliferation Kit with Alexa Fluor 488 according to the manufacturer's protocol. Proliferating cells (EdU-positive) were identified by green fluorescence through Azide 488 staining, while all cell nuclei were counterstained blue with Hoechst 33,342. Three randomly selected microscopic fields were captured for analysis to ensure representative sampling. The EdU-positive rate was calculated as: (Number of EdU-positive cells / Total number of cells) × 100%.

Immunofluorescence staining

Cells were meticulously cultured on 18-mm coverslips, where they were left undisturbed to acclimate overnight in an incubator at 37 °C. Once sufficiently stabilized, the slides underwent a blocking step using 5% bovine serum albumin, effectively reducing nonspecific binding. Next, the cells were incubated with the primary antibody, which was followed by a subsequent incubation with secondary antibodies, each conjugated with a fluorescent label, to ensure precise and vivid detection. To illuminate the cell nuclei, DAPI (Thermo) was applied, staining them in striking contrast. Images were then meticulously captured with the LSM 880 laser scanning microscope (Zeiss), ensuring the clearest, most detailed visual representation. The primary antibody, chosen for its specificity and reliability, was the anti-BCL-2 (Abmart, T40056, 1:200), playing a critical role in targeting and visualizing the desired protein.

Colony formation assay

To evaluate the colony-forming ability of cells, we carefully seeded 800 to 1,000 cells per well in 6-well plates and cultured them in an incubator at 37 °C for 10–14 days to allow colony formation. After incubation, the colonies were fixed with methanol for 15–30 min to preserve their structures, then stained with 0.1% crystal violet solution for 15 min to enhance visibility. Manual counting of stained colonies was performed to quantitatively assess cell proliferation and colony-forming capacity under specific experimental conditions.

Apoptosis assay

Cells were cultured overnight in complete medium at 37 °C in a 5% CO_2 incubator. The cells, both adherent and suspended, were subsequently harvested. Prior to the analysis, they underwent a washing step with ice-cold PBS to ensure optimal conditions for further processing. The next critical step in the procedure involved staining the cells with Annexin V-FITC and Propidium Iodide (PI), two vital markers for distinguishing apoptotic cells from healthy ones. Annexin V-FITC binds to phosphatidylserine, which translocates to the outer leaflet of the plasma membrane during early apoptosis. PI, on the other hand, enters the cell only when the membrane integrity is compromised, marking late-stage apoptosis or necrosis. To assess the extent of apoptosis, we utilized a FACS flow cytometer, a sophisticated tool capable of detecting the fluorescence signals emitted by both the Annexin V and PI markers. Cells that displayed positivity for both markers were classified as apoptotic, offering a detailed analysis of the cellular demise process. Crucially, to ensure the reliability and reproducibility of the results, all experiments were meticulously repeated three times, guaranteeing that the data collected would be robust and statistically significant.

Statistical analysis

The results are presented as the mean, accompanied by the standard deviation. In order to assess the differences in categorical variables, the chi-squared test was employed, with a significance threshold set at a P value of less than 0.05. To account for the multiple comparisons made, the P values were adjusted using the Benjamini–Hochberg procedure, ensuring greater accuracy. All analyses, as well as graphical representations, were generated using R software, version 4.4.2. All experiments were performed with at least three biological replicates to ensure the reliability and reproducibility of the results and results were analyzed using the t-test to assess statistical significance.

Results

Genome-wide meta-analysis identifies STK24 as a risk gene for LUAD

The workflow of our study design is shown in Fig. 1. First, we conducted a meta-analysis of LUAD genome-wide association results from three studies: ebi-a-GCST004744 (ncases = 11,273; ncontrols = 55,483), finn-b-C3_NSCLC_ADENO (ncases = 571; ncontrols = 218,221), and ieu-a-984 (ncases = 11,245; ncontrols = 54,619) (Fig. 2A). After quality control, we obtained 10,227,138 gene variants associated with LUAD. Supplementary Fig. 1 presents the quantile–quantile (Q-Q) plot of the meta-analysis. We observed 69 genome-wide significant signals associated with LUAD that are distinct from those reported in the three prior studies (Fig. 2B and Supplementary Table 2). Based on a 500 KB region around each index SNP (Supplementary Table 2), we identified the gene closest to the index SNP and the gene with the highest priority score (PoPS). For all genes indicated by proximity and PoPS, we performed enrichment analysis using Metascope (Fig. 2C), and found that GO:0043065: positive regulation of apoptotic process was enriched, along with other relevant pathways such as R-HSA-380108: chemokine receptors bind chemokines, GO:0006954: inflammatory response, and R-HSA-1280215: cytokine signaling in the immune system. Among the 69 novel loci, none were associated with tuberculosis (TB), but 9 loci were related to smoking (rs10846701/UBC, rs11255971/IL2RA, rs114965893/TACC3, rs117355568/RPA3, rs1800057/ATM, rs2023689/PIK3CG, rs3127545/TNFRSF14, rs4757381/IFITM3, rs62155733/IL1R1), and 3 loci were related to air pollution (rs11639926/PLCG2, rs1985385/IRF4, rs78323695/IRF4) (Fig. 2D–E, Supplementary Table 3). We then conducted a genetic correlation analysis of LUAD and its risk factors, finding that air pollution (rg = 0.161, P value = 0.033), smoking (rg = 0.312,

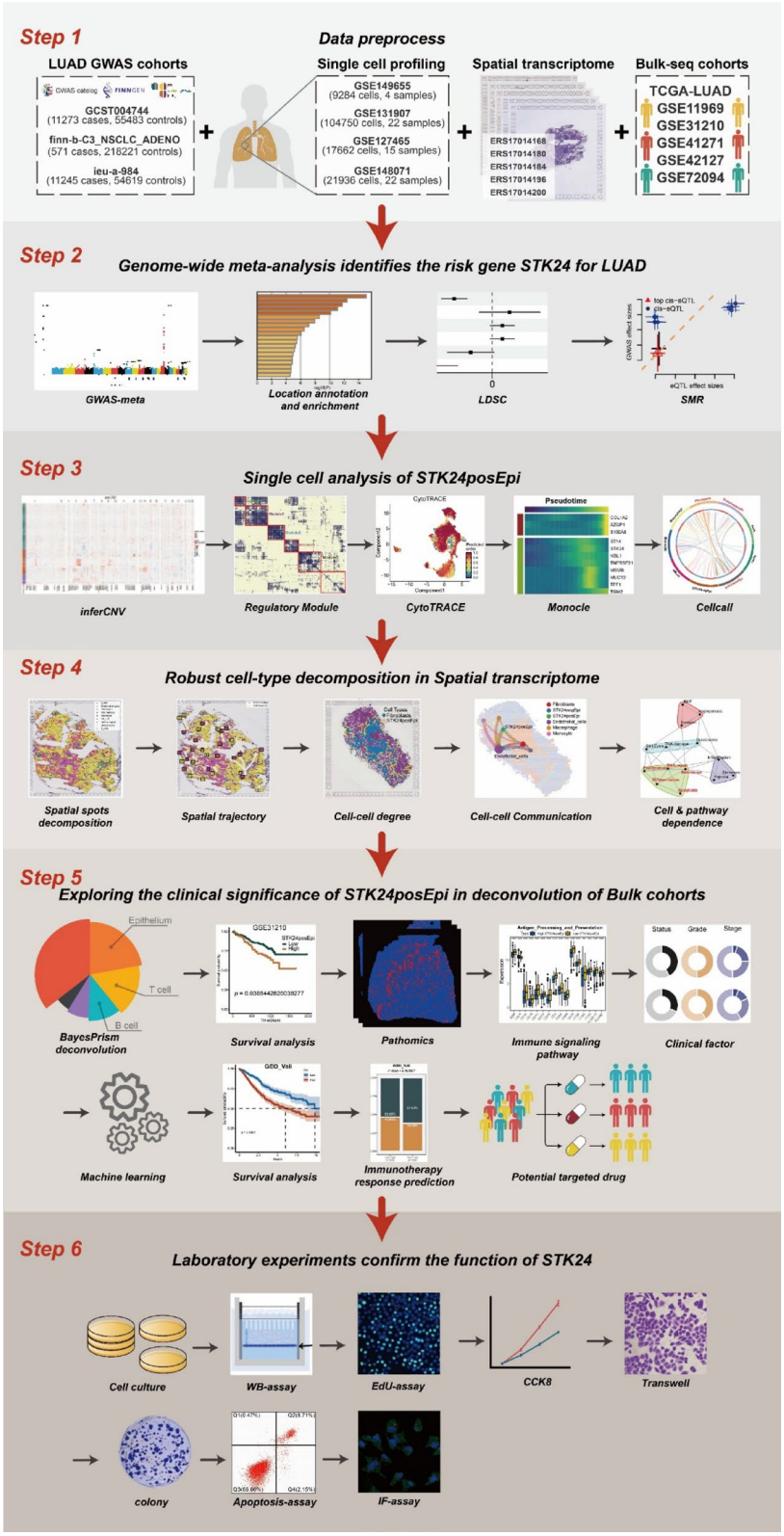

Fig. 1 The workflow of the study

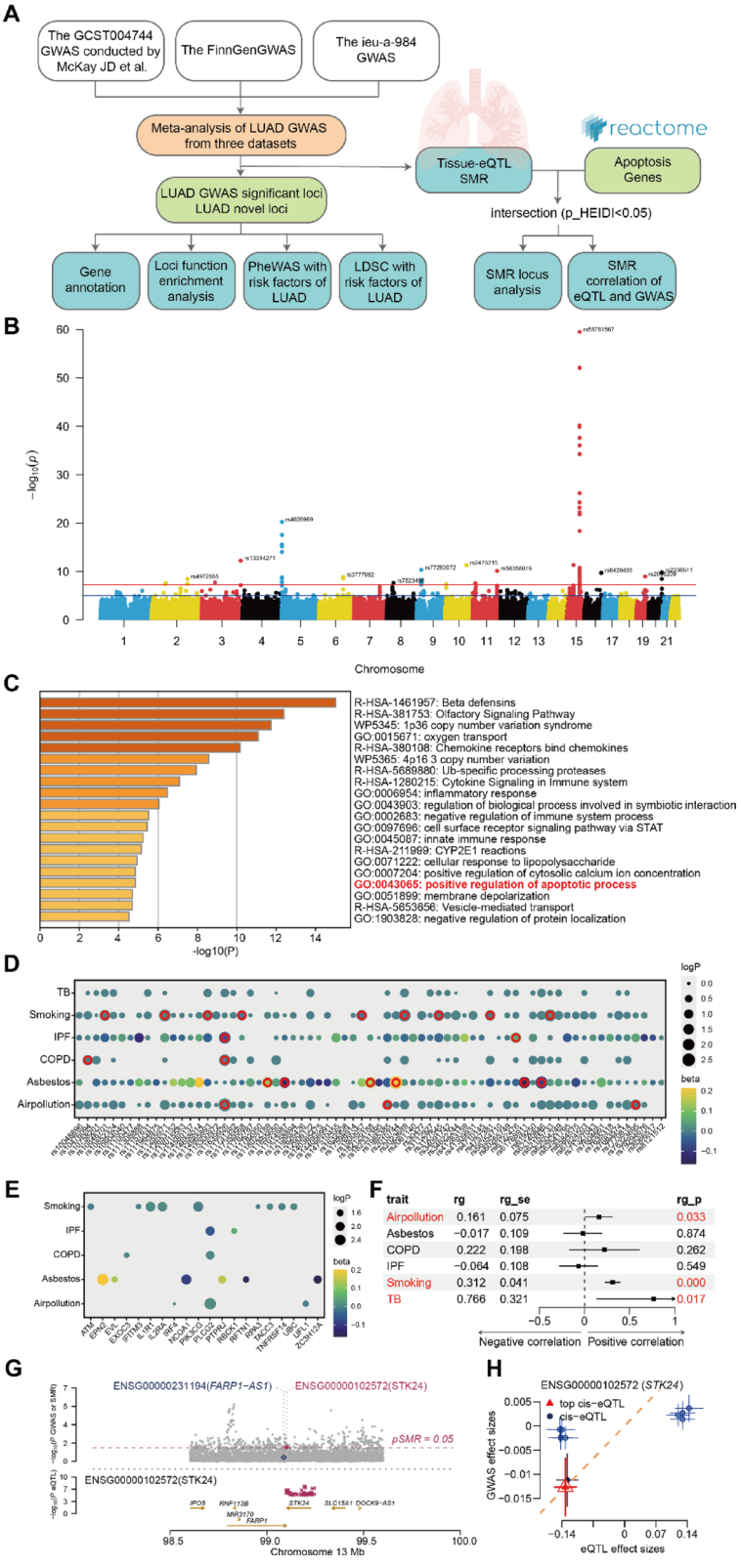

Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 Genome-wide meta-analysis identifies STK24 as a risk gene for LUAD. **A** Dataset and analysis schematic. **B** Manhattan plots displaying LUAD associations from a GWAS meta-analysis of 307,196 individuals. **C** Bar chart displaying the enrichment analysis of pathways related to newly identified loci genes. **D** Plots showing genetic associations of 68 LUAD loci with risk factors for LUAD (red circles indicate $P < 0.05$). **E** Bubble plot annotating genes with $P < 0.05$ against LUAD-related risk factors. **F** Genetic associations between LUAD and various risk factors (n = 456,380 for air pollution, n = 106,915 for asbestos, n = 112,583 for COPD, n = 1,254,748 for IPF, n = 518,633 for smoking, n = 462,933 for tuberculosis (TB)). The error bars' center shows the genetic association's effect estimate (beta coefficient) between loci and trait, with the bars representing the 95% confidence interval. All tests were two-sided and unadjusted for multiple comparisons. **G** SMR analysis co-localization plot for STK24. (H) SMR analysis correlation plot for STK24.

P value = 0.000), and tuberculosis (TB) (rg = 0.766, *P* value = 0.017) were positively correlated with LUAD. No significant associations were found between LUAD and asbestos, chronic obstructive pulmonary disease (COPD), or idiopathic pulmonary fibrosis (IPF). We then performed SMR analysis of the LUAD-meta GWAS cohort using lung tissue-specific expression quantitative trait locus (eQTL) from the Genotype-Tissue Expression (GTEx) database (Supplementary Table 4) and intersected the results with apoptosis-related genes (as shown in the workflow in Fig. 2A). We found that STK24 exhibited significant colocalization with LUAD and was positively associated with LUAD, indicating that STK24 is a risk gene for LUAD (Fig. 2G–H).

Elevated expression of STK24 in LUAD epithelial cells

By integrating four single-cell LUAD cohorts (Supplementary Fig. 2), we explored STK24 expression in LUAD. After removing batch effects, we performed dimensionality reduction, clustering, and cell annotation (Fig. 3A). We found that STK24 was highly expressed in epithelial cells, macrophages, and monocytes, with significantly higher STK24 expression in epithelial cells from tumor tissues compared to normal tissues (Fig. 3B, C). The research concentrated on the function of tumor epithelial cells in LUAD, thus we solely analyzed STK24 expression in these cells. We employed the Cottrazm R package to examine the specificity of STK24 expression in tumor tissues by identifying and characterizing benign and malignant areas in spatial transcriptomics (Supplementary Fig. 3A–E). STK24 expression was elevated in malignant tissues relative to benign tissues (Fig. 3D, E). In both paired and unpaired data from the TCGA-LUAD cohort, STK24 expression was higher in tumor tissues compared to normal tissues (Fig. 3F, Supplementary Fig. 3F). Immunohistochemical staining from the Human Protein Atlas (HPA) database showed stronger STK24 staining in tumor tissues compared to normal tissues (Fig. 3G). These conclusions were further confirmed in validation cohorts (Supplementary Fig. 3G–H). Subsequently, we performed univariate and multivariate Cox analyses to evaluate the clinical significance of STK24. In five different bulk cohorts, univariate survival analysis indicated that STK24, T stage, and stage were associated with LUAD survival and were identified as risk factors (HR > 1). Multivariate Cox survival analysis further demonstrated that STK24 is an independent risk factor for LUAD, independent of traditional clinical variables (Fig. 3H).

Based on expression levels, epithelial cells were divided into two groups (STK24 group): STK24-positive (STK24posEpi) and STK24-negative (STK24negEpi) (Fig. 4A). First, we analyzed the malignancy of the STK24 group using infercnv (Supplementary Fig. 4), dividing epithelial cells into seven clusters. Clusters 2, 4, and 5 had higher copy number variant (CNV) scores, and we found that the CNV score of STK24posEpi was higher than that of STK24negEpi in cluster 2, with STK24posEpi showing copy number variations (CNVs) primarily in chr13q_gain, chr6q_gain, and chr19q_gain (Fig. 4B). Subsequent Gene Ontology—Biological Process (GO-BP) and Reactome pathway enrichment analysis showed upregulation of pathways related to cell differentiation, cell migration, and glycosylation in STK24posEpi, including epidermal cell differentiation, regulation of cell migration, and O-linked glycosylation (Fig. 4C). Further transcriptional regulatory analysis in LUAD epithelial cells was conducted to explore the internal mechanisms driving pathway activation. Through connectivity specificity index evaluation, six transcriptional modules were identified (Fig. 4D). STK24posEpi scored higher in module 2, while STK24negEpi scored higher in module 6 (Fig. 4E, F). Transcription factor activity analysis showed enhanced activity of key genes in module 2, such as KLF9, MAZ, THAP1, KDM5B, ZNF740, ELF1, FOXK2, and ZNF384 in STK24posEpi (Fig. 4G). Research has shown that KLF9 suppresses gastric cancer cell invasion and metastasis [45] and suppresses cell growth and induces apoptosis in prostate cancer [46, 47]. In lung cancer, ELF1 suppresses autophagy to reduce cisplatin resistance [48]. THAP1 has been identified as a nuclear pro-apoptotic factor that links prostate apoptosis response-4 (Par-4) with PML nuclear bodies [49]. KDM5B has also been linked to apoptosis [50, 51].

STK24posEpi: unveiling the origin of epithelial cells in LUAD patients

We used the CytoTRACE algorithm to find that STK24posEpi showed higher tumor stemness scores (Fig. 5A–C), indicating its position at the origin in the developmental hierarchy. The developmental trajectory predicted by the Monocle algorithm revealed a gradual decrease in the proportion of STK24posEpi cells as pseudotime progressed (Fig. 5D, E). We analyzed the

differential genes across different developmental clusters in the trajectory (Fig. 5F). Early development was primarily characterized by the expression of NUPR1, FABP5, LAMP3, and PGC, while later stages showed the expression of STK24, TGM2, and TFF1. We further performed pseudotime analysis on tumor-associated pathways. We found that epithelial-mesenchymal transition (EMT), angiogenesis (ANG), DNA damage repair (DNA), the phosphatidylinositol 3-kinase (PI3K) pathway, apoptosis (APO), and Wnt signaling (WNT) exhibited an upward trend throughout pseudotime (Fig. 5G). Notably, STK24posEpi displayed higher tumor-related features throughout the entire pseudotime continuum when compared to STK24negEpi.

Subsequently, we employed the RCTD deconvolution method and the stlearn spatial trajectory algorithm to analyze the spatial transcriptomics data (Fig. 6A). Using the RCTD deconvolution approach, distinct cell types from the single-cell dataset were mapped onto the spatial transcriptomics data (Fig. 6B–D, Supplementary Fig. 5A, E). Through the stlearn algorithm, we identified an evolutionary trajectory from STK24posEpi to STK24negEpi within the spatial transcriptome. The trajectory tree further elucidated the developmental relationships among individual cell clusters (Fig. 6E–J, Supplementary Fig. 5B–C, F–G). Lastly, a robust correlation emerged between STK24 expression and the enrichment scores of genes involved in the trajectory's evolutionary path (Fig. 6K–M, Supplementary Fig. 5D, H). In conclusion, STK24posEpi is a key origin of epithelial cell development in LUAD tumors and drives associated tumor-related states/pathways along its trajectory.

STK24posEpi interaction with fibroblasts: a complex network unveiled

Our study revealed significant cell–cell communication interactions between STK24posEpi cells and fibroblasts (Fig. 7A). The core pathways mediating these interactions included the Hippo signaling pathway, estrogen signaling pathway, endocrine resistance, focal adhesion, and resistance to EGFR tyrosine kinase inhibitors (Fig. 7B). Furthermore, we identified several key ligand-receptor pairs involved in STK24posEpi-to-fibroblast communication: PDGFA-PDGFRA, PDGFA-PDGFRB, BMP2-BMPR2, TGFA-EGFR, VEGFA-PDGFRA, and VEGFA-PDGFRB (Fig. 7C). Similarly, critical ligand-receptor pairs from fibroblasts to STK24posEpi cells included IL6-IL6ST, INHBA-ACVR2A, INHBA-ACVR2B, and IL6-IL6R (Fig. 7C). These ligands and receptors highlight the complex interplay between STK24posEpi cells and fibroblasts.

To validate these findings, we conducted heterotypic cellular network analysis and cellular co-localization analysis using spatial transcriptomics (Fig. 7D). We identified extensive cellular networks between STK24posEpi and fibroblasts (Fig. 7E–H). Subsequently, we performed spatial cellular co-localization analysis using the mistyR algorithm, revealing significant co-localization between STK24posEpi and fibroblasts in the intra, juxta_5, and para_15 regions (Fig. 7I). These observations were consistent across multiple spatial transcriptomics samples (Supplementary Fig. 5A–C).

Subsequently, we conducted spatial cell–cell communication and signaling flow analyses (Fig. 8A). Cell communication analysis revealed robust interactions between STK24posEpi and fibroblasts (Fig. 8B). Within the PDGF signaling pathway, STK24posEpi predominantly functioned as the sender while fibroblasts acted as receivers (Fig. 8C–E). Co-expression of PDGF pathway-associated ligands and receptors was consistently observed across multiple transcriptomic samples (Fig. 8F–G). In the VEGF signaling pathway, STK24posEpi again served as the sender, with fibroblasts mediating signal transmission and endothelial cells functioning as terminal receivers (Fig. 8H). Spatial interactions among STK24posEpi, fibroblasts, and endothelial cells exhibited consistent patterns across diverse spatial transcriptomic samples (Fig. 8I-J). Similarly, in the MIF signaling pathway, STK24posEpi operated as the sender, fibroblasts as mediators, and macrophages as receivers (Fig. 8K). Notably, this tripartite communication system was validated in multiple spatial transcriptomic samples (Fig. 8L, M).

Further investigation of MIF signaling using the commot algorithm revealed spatial distribution patterns of MIF signals (Fig. 8N). The spatial map demonstrated directional signaling flows (indicated by black arrows) originating from STK24posEpi and fibroblast-enriched regions, consistent with cell type distributions (Fig. 6B–D, Supplementary Fig. 5A, E). Signal intensity maps displayed distinct spatial segregation between MIF senders (light blue) and CD74_CD44 receivers (red), with pronounced signal strength observed in regions of active MIF transmission (Fig. 8N). These findings were further corroborated in an independent spatial transcriptomic sample (Fig. 8O).

The complex interplay of STK24posEpi, apoptosis, and tumor dynamics

We attempted to reveal the underlying mechanisms of related pathways through spatial transcriptomics and performed heterogeneity analysis of 12 tumor cell states or pathways included in the cancer single-cell state atlas (CancerSEA) database using the MistyR algorithm (Fig. 9A). We found that apoptosis scores in malignant regions of spatial transcriptomics were significantly higher than in benign regions (Fig. 9B). Subsequently, we conducted pathway dependency analysis using the MistyR algorithm and discovered that apoptosis is associated with a series of pathways, including proliferation,

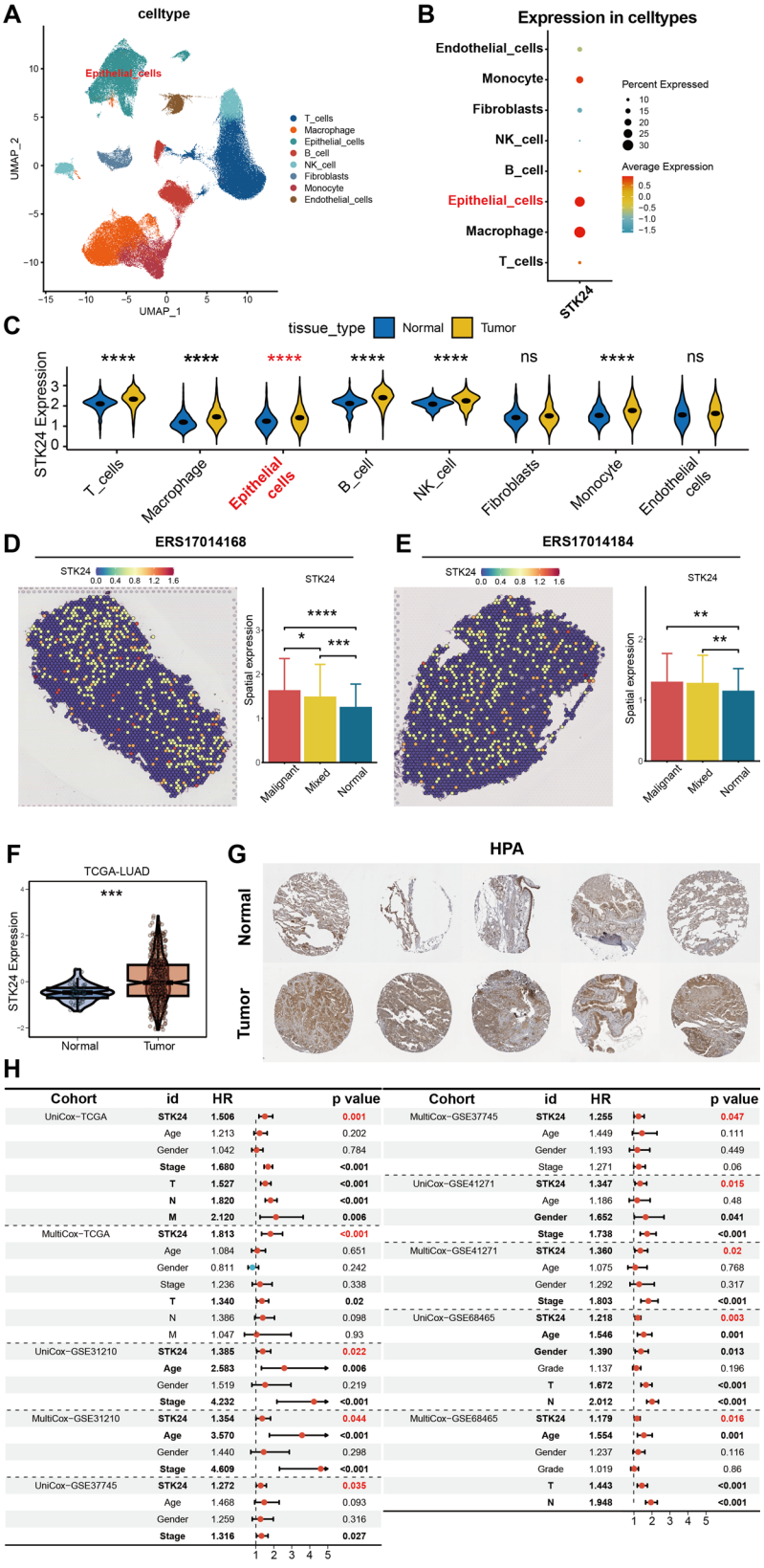

Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 STK24 is elevated in LUAD epithelial cells. **A** UMAP showing cell types after batch correction and dimensionality reduction clustering. **B** Bubble plot showing STK24 expression levels across various cell types. **C** Violin plot showing STK24 expression in normal and tumor cells across various cell types. **D**, **E** STK24 expression levels and regional variation analysis in spatial transcriptomics. **F** Violin plot showing STK24 expression in normal and tumor samples in the TCGA-LUAD cohort. **G** Immunohistochemistry results showing STK24 staining in LUAD and normal tissue samples from the HPA database. **H** Independent prognostic analysis to evaluate whether the association between STK24 and tumor survival is independent of traditional clinical variables. ****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, ns $P > 0.05$

DNA damage, differentiation, and metastasis (Fig. 9C–F). These observations were validated in multiple spatial transcriptomic samples (Supplementary Fig. 7A–D).

In a detailed study of the role of STK24posEpi, we observed that STK24posEpi primarily activates pathways or states related to the cell cycle, DNA damage repair, proliferation, and metastasis. We then evaluated the spatial distribution of tumor state/pathway scores and found that STK24posEpi is highly active in pathways or tumor states associated with proliferation (Fig. 9G), DNA damage (Fig. 9H), the cell cycle (Fig. 9I), and DNA repair (Fig. 9I). Using the MistyR algorithm, we found that STK24posEpi exhibits a strong spatial dependency on tumor pathways/states related to proliferation, the cell cycle, DNA repair, and DNA damage. This spatial relationship is evident in the intra, juxta_5, and para_15 regions (Fig. 9J, K). In conclusion, our analysis highlights the close relationship between STK24posEpi and tumor states/pathways such as cell proliferation, DNA repair, and metastasis.

STK24posEpi is associated with LUAD prognosis, lymphocyte infiltration, and clinical factors

First, we performed an isogenic cell network analysis of STK24posEpi (Fig. 10A). A strong isogenic cell network of STK24posEpi was found across multiple spatial transcriptomics samples (Fig. 10B). We then used Bayesian deconvolution to map STK24posEpi to a bulk transcriptome cohort, finding that the proportion of STK24posEpi was significantly higher in tumor samples than in normal samples (Supplementary Fig. 8A). Similarly, the proportion of STK24posEpi cells was higher in malignant regions than in benign regions (Supplementary Fig. 8B). We further analyzed the clinical factors associated with STK24posEpi, and survival analysis revealed that higher STK24posEpi cell proportions were associated with worse prognosis (Fig. 10C, Supplementary Fig. 9A). The proportion of STK24posEpi cells was significantly higher in Stage 3 tumors compared to Stage 1 (Supplementary Fig. 8C). In TNM staging, the proportions were higher in T2 and T3 compared to T1 (Supplementary Fig. 8D), and in N2 compared to N0 (Supplementary Fig. 8E). In age groups, patients aged ≤ 65 had a higher proportion of STK24posEpi cells (Supplementary Fig. 8F). We then divided the TCGA-LUAD cohort samples into high and low groups based on STK24posEpi cell proportions. Using the TCGA_TilMap database, we found that the High STK24posEpi group had lower tumor-infiltrating lymphocyte (TIL) scores (Fig. 10D), and pathology images showed reduced tumor-infiltrating lymphocytes in the High STK24posEpi group (Fig. 10E). Correlation analysis showed a negative correlation between STK24posEpi cell proportions and B cells across multiple bulk cohorts (Fig. 10F, Supplementary Fig. 9B-C). Survival analysis showed that higher B cell and NK cell proportions were associated with better prognosis (Supplementary Fig. 9D–E), and there was a negative correlation between STK24posEpi cell proportions and NK cells (Supplementary Fig. 9F–J). Additionally, BCR signaling pathway and antigen processing and presentation-related genes were downregulated in the High STK24posEpi group (Fig. 10G, H). Finally, we analyzed the clinical factors in the High STK24posEpi group and found lower survival rates, as well as lower proportions of T1 and N0 tumor grades (Fig. 10I). In summary, we found that STK24posEpi is associated with reduced tumor-infiltrating lymphocytes and is correlated with poor prognosis and worse tumor grading in patients.

Machine deep learning models: predicting patient survival and informing treatment decisions

We identified the top 100 signature genes of STK24posEpi using the starTracer algorithm [52], filtered them with the parameter n = 1, and finally retained 78 signature genes (Supplementary Table 5) for building the prognostic model via machine learning (Supplementary Fig. 10A). We found that the SuperPC algorithm ranked high in terms of C-index across multiple cohorts, so it was selected as the final prognostic model (named STK24SuperPC). Subsequent survival analysis revealed that patients with higher risk scores calculated by the STK24SuperPC model had poorer prognoses (Supplementary Fig. 10B-G). The AUC values for the TCGA-LUAD set were 0.66, 0.65, 0.66, 0.68, and 0.65 at the 1-, 2-, 3-, 4-, and 5-year marks, respectively. The AUC values for the GSE42127 set were 0.71, 0.66, 0.64, 0.63, and 0.61 at the 1-, 2-, 3-, 4-, and 5-year marks, respectively (Supplementary Fig. 10H-L).

Further analysis using the TIDE algorithm to predict immune therapy responses revealed that in the TCGA (Chi-Square.test $P = 9.7e{-}0.5$), GSE72094 (Chi-Square.test $P = 0.001387$), GSE41271 (Chi-Square.test $P = 0.028875$), and GSE42127 (Chi-Square.test $P = 0.029409$) cohorts, patients in the high-risk score

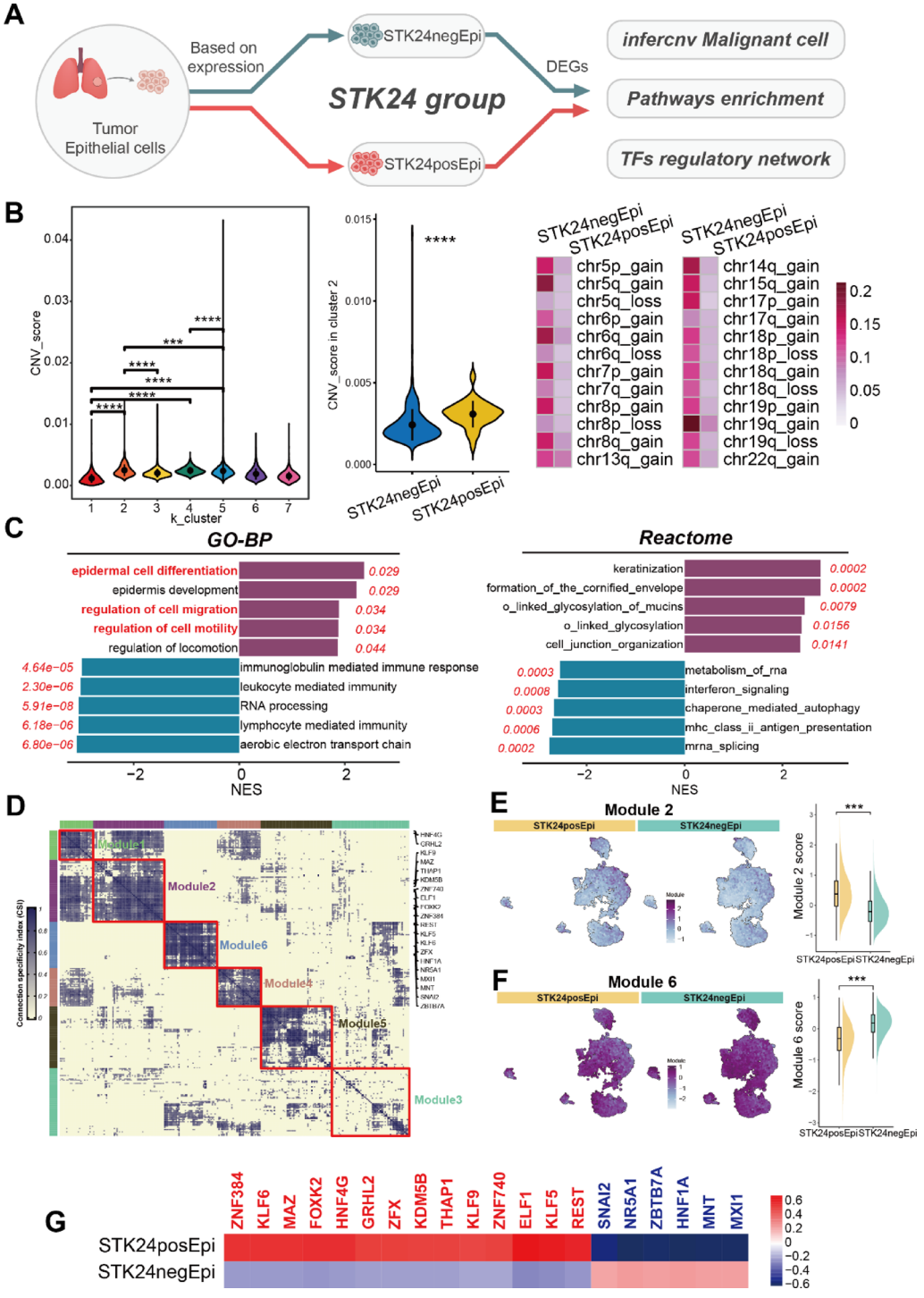

Fig. 4 Identification and functional analysis of the STK24posEpi cell population. **A** Schematic diagram identifying STK24-positive tumor epithelial cells (STK24posEpi). **B** InferCNV analysis showing CNV scores of tumor epithelial cells and the differences in copy number variations between STK24posEpi and STK24negEpi. **C** Bar chart showing enrichment analysis of GO-BP and Reactome pathways. **D** Transcription factor regulatory module analysis in epithelial cells. **E**, **F** Differential comparison of regulatory submodules in the STK24 group. **G** Heatmap showing the activity of different transcription factors in the STK24 group (red text: higher TF activity in STK24posEpi; blue text: higher TF activity in STK24negEpi). ****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, ns $P > 0.05$

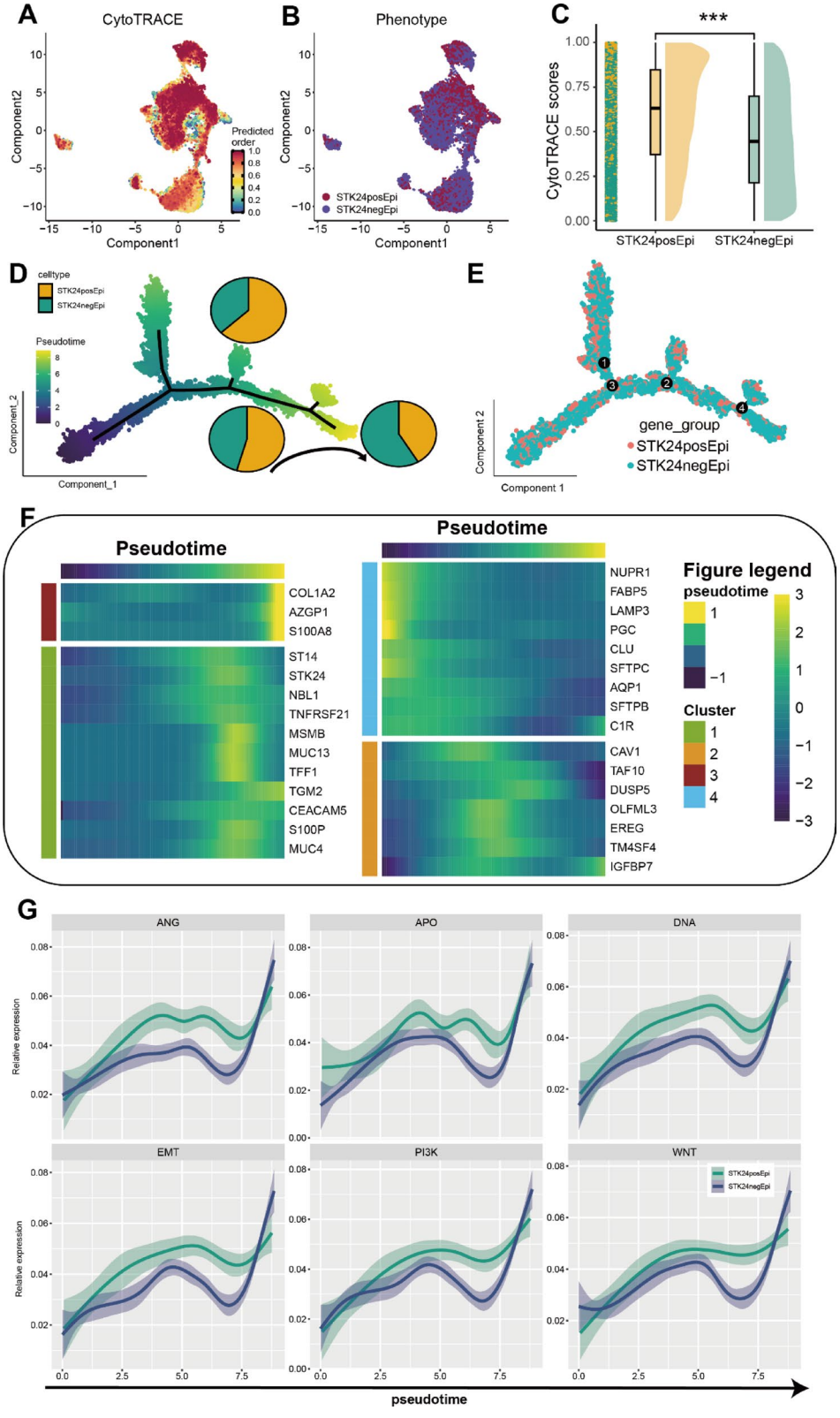

Fig. 5 Exploring the cell origins of the STK24 Group at the single-cell level. **A** CytoTRACE analysis of the differentiation potential of STK24 Group cells. **B** Subgrouping of STK24 Group. **C** Raincloud plot showing a comparison of CytoTRACE scores. **D**, **E** Analysis and display of cell type proportions along the developmental trajectory of STK24 Group cells. **F** Pseudotime analysis of genes related to cell developmental pathways. **G** Comparative pseudotime analysis of tumor activity-related pathways. ****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, ns $P > 0.05$

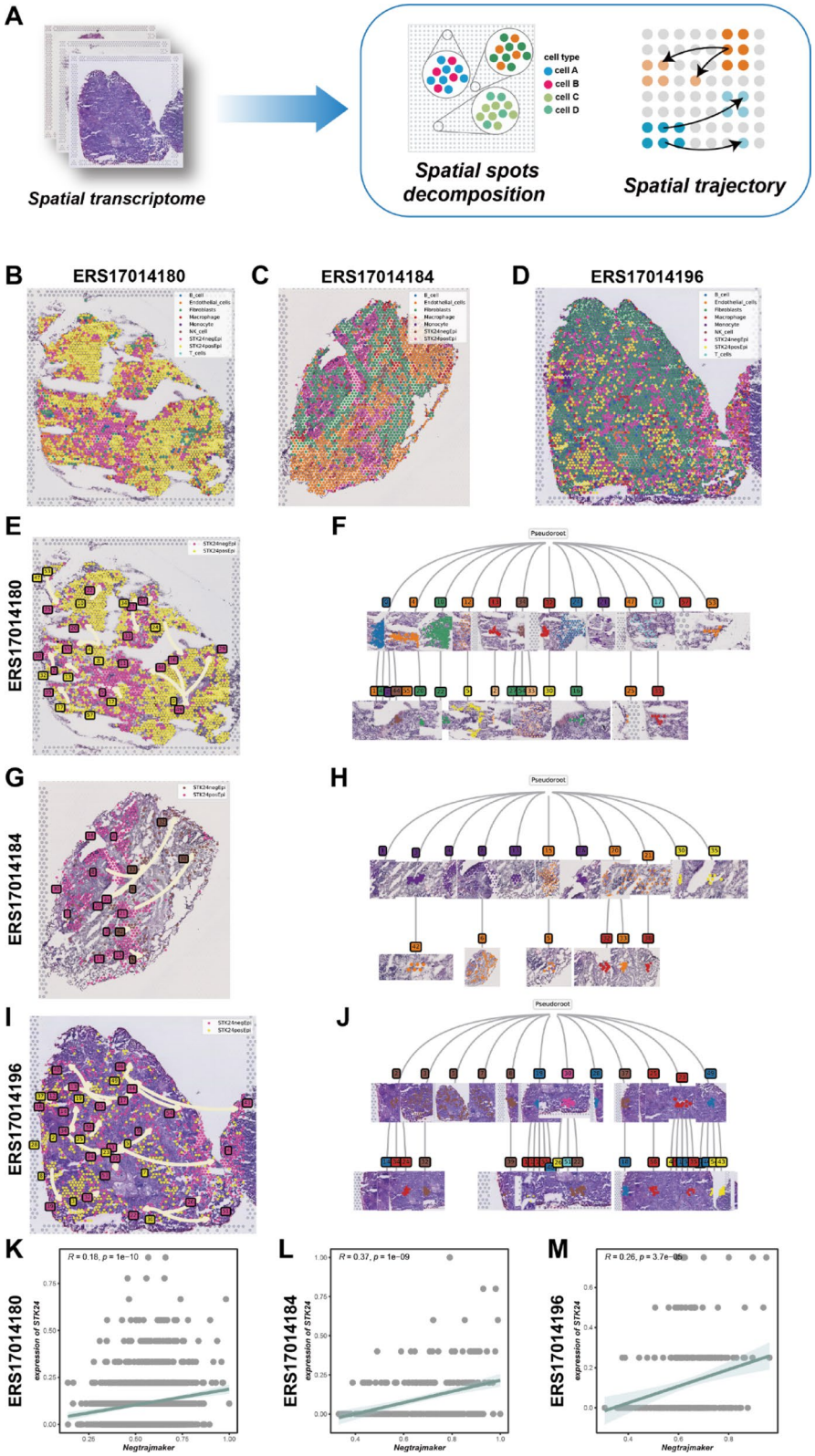

Fig. 6 Exploring the origins of STK24 Group cells through spatial transcriptomics (ST). **A** Schematic diagram of RCTD deconvolution and spatial trajectory analysis of spatial transcriptomics data. **B**–**D** Cell types after ST deconvolution. **E**, **F** Cell developmental trajectory and trajectory tree in ST ERS17014180. **G**, **H** Cell developmental trajectory and trajectory tree in ST ERS17014184. **I**, **J** Cell developmental trajectory and trajectory tree in ST ERS17014196. (K-M) Scatter plots showing the correlation between STK24 gene expression and developmental trajectory genes

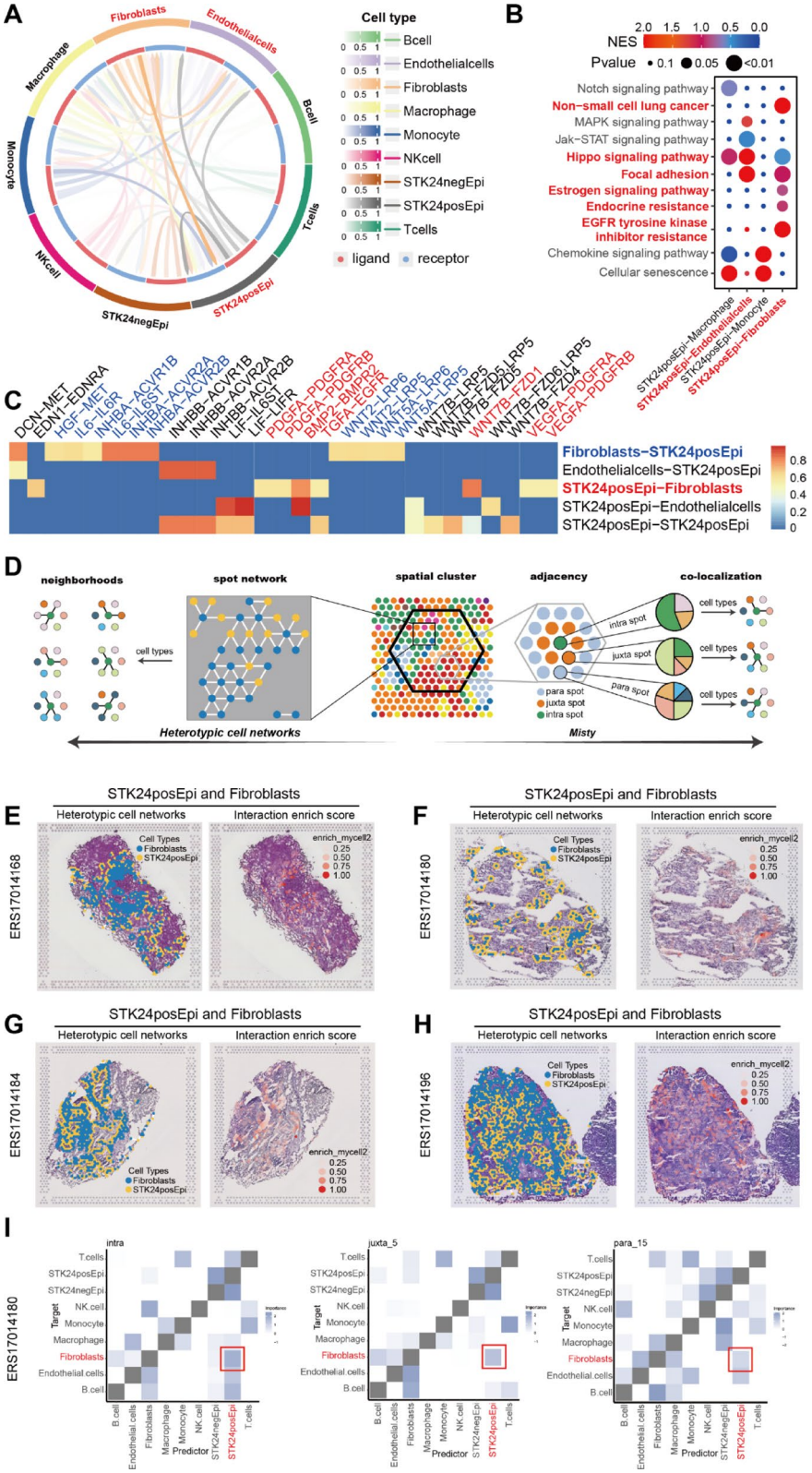

Fig. 7 Interactions between STK24-positive tumor epithelial cells (STK24posEpi) and fibroblasts. **A** Analysis of interaction strength between STK24posEpi and various cell types. **B** Activated pathways in various cell communications. **C** Analysis of activated ligand-receptor pairs. **D** Schematic diagram of Heterotypic cellular network analysis and cell co-localization analysis of spatial tran-scriptomics data. **E**–**H** Spatial transcriptomics heterotypic cell network analysis shows colocalization of STK24posEpi and fibroblasts. **I** Heatmap displaying cell–cell dependency analysis in the colocated, neighboring, and extended neighboring (15-point) regions of the spatial transcriptomics data

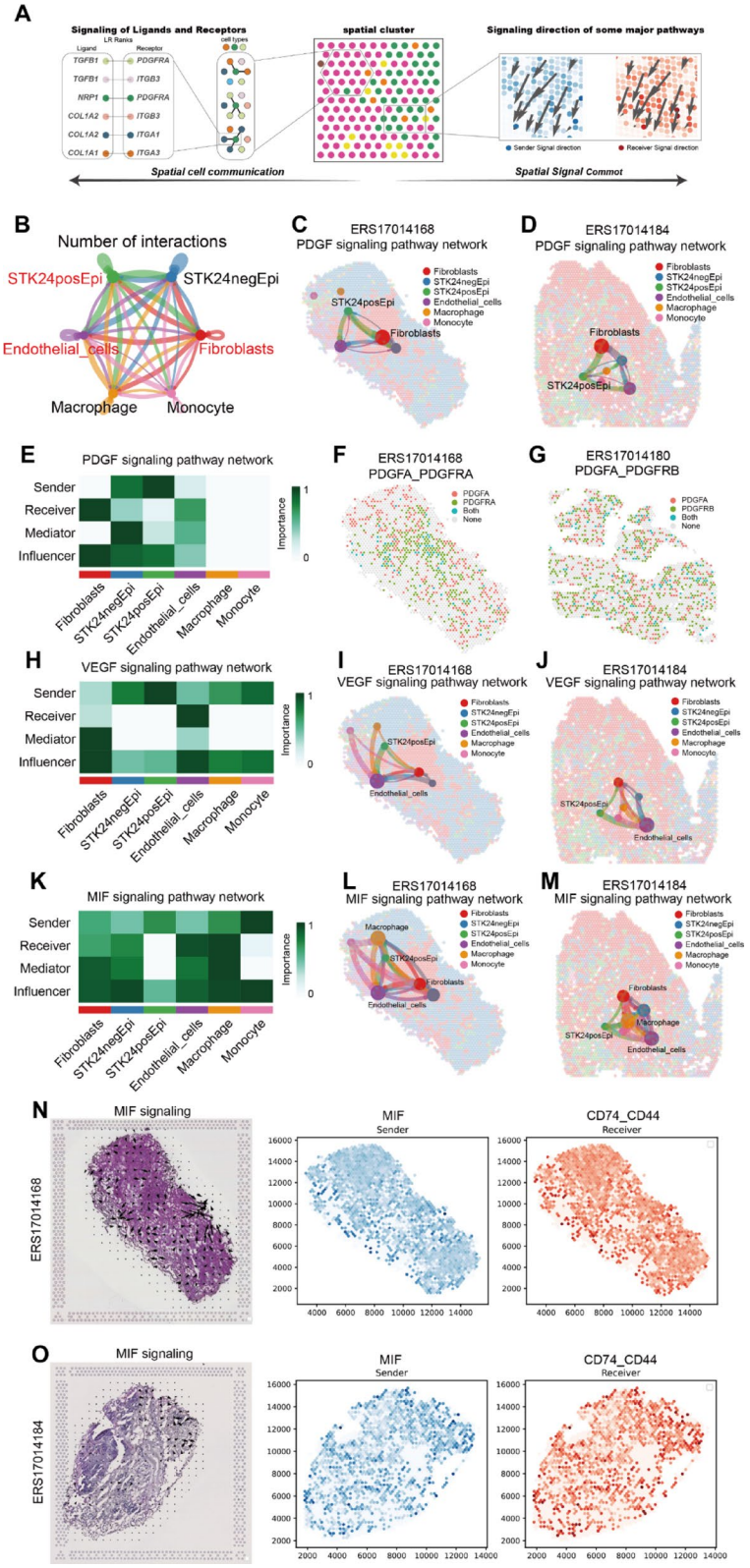

Fig. 8 (See legend on next page.)

(See figure on previous page.)

Fig. 8 Communication and signal flow changes between STK24posEpi and fibroblasts in spatial transcriptomics (ST). **A** Schematic diagram of Cell–cell communication analysis and signal flow direction analysis of spatial transcriptomics data. **B** Analysis of communication intensity between STK24posEpi and fibroblasts by integrating multiple spatial transcriptomics samples. **C**, **D** Communication between STK24posEpi and fibroblasts in the PDGF signaling pathway across different spatial transcriptomics samples. **E** Importance of Sender, Receiver, Mediator, and Influencer in different cell types in the PDGF signaling pathway. **F**, **G** Expression and co-expression of ligand-receptor pairs related to the PDGF signaling pathway in various spatial transcriptomics samples. **H** Importance of Sender, Receiver, Mediator, and Influencer in different cell types in the VEGF signaling pathway. **I**, **J** Communication between STK24posEpi and fibroblasts in the VEGF signaling pathway across different spatial transcriptomics samples. **K** Importance of Sender, Receiver, Mediator, and Influencer in different cell types in the MIF signaling pathway. **L**, **M** Communication between STK24posEpi and fibroblasts in the MIF signaling pathway across different spatial transcriptomics samples. **N**, **O** COMMOT analysis showing the direction of MIF signal flow and expression of Senders and Receivers in various spatial transcriptomics samples

group had fewer favorable responses to immunotherapy compared to the low-risk score group (Fig. 11A–D). Additionally, in the TCGA (wilcox.test $P = 9.7e{-}0.5$), GSE72094 (wilcox.test $P = 0.0014$), GSE41271 (wilcox.test $P = 0.0036$), and GSE42127 (wilcox.test $P = 0.035$) cohorts, the Exclusion score of patients in the high-risk score group was found to be higher than that of the low-risk score group, which was associated with a poorer prognosis for immunotherapy outcomes (Fig. 11E–H). In the clinical treatment sensitivity analysis, we used the CTRP and PRISM databases, finding that the high-risk group exhibited higher sensitivity to certain chemotherapy drugs, including cisplatin, docetaxel, and etoposide (Fig. 11I–K). Moreover, paclitaxel, GSK461364, leptomycin B, SB-743921, epothilone-B, and cabazitaxel were identified as potential treatment candidates for this group (Fig. 11L–M).

STK24 overexpression: a key driver of cell proliferation and apoptotic inhibition

In our study, we used overexpression plasmids to increase the levels of STK24 in A549 and PC9 lung adenocarcinoma cell lines. As the control group, we transfected pLVX-Vector-Puro plasmids, and as the experimental group, we transfected pLVX-STK24-Puro plasmids. After 48 h of transfection, cells were selected with 10 μg/mL puromycin at a 1:1000 dilution ratio. Western blot (WB) analysis showed a significant increase in STK24 expression in the STK24-OE group (T.test A549: $P = 0.0014$; PC9: $P < 0.001$), confirming the successful plasmid transfection and protein expression (Fig. 12A, B). The EdU assay revealed an increase in DNA replication activity in A549 (T.test $P = 0.025$) and PC9 (T.test $P = 0.0077$) cells following STK24 overexpression (Fig. 12C–F). The CCK8 assay indicated enhanced cell proliferation after STK24 overexpression (T.test A549: $P < 0.01$; PC9: $P < 0.01$) (Fig. 12E, F). Invasion assays showed that STK24 overexpression increased the invasion ability of A549 (T.test $P = 0.0017$) and PC9 (T.test $P = 0.0093$) cells (Fig. 12G–I). Two-dimensional colony formation assays demonstrated that STK24 overexpression enhanced the colony formation ability of these cells (T.test A549: $P < 0.001$; PC9: $P = 0.015$) (Fig. 12J–L). Flow cytometry analysis showed a significant reduction in apoptosis rates in STK24-overexpressing cells (T.test A549: $P = 0.0016$; PC9: $P < 0.001$), indicating an anti-apoptotic effect (Fig. 13A, B). Western blot analysis revealed that STK24 overexpression inhibited the expression of Cleaved-Caspase3 (T.test A549: $P < 0.001$; PC9: $P = 0.0014$) while upregulating the expression of BCL-2 (T.test A549: $P = 0.0011$; PC9: $P = 0.0027$) and Caspase3 (T.test A549: $P < 0.001$; PC9: $P < 0.001$) (Fig. 13C–F). Immunofluorescence staining showed increased fluorescence intensity of Caspase3 in the STK24 overexpression group (Fig. 13G–H). In conclusion, STK24 not only promotes the proliferation and migration of LUAD cells but also inhibits apoptosis.

Discussion

Our genetic analysis of 10,227,138 LUAD cases identified 69 distinct LUAD variants through GWAS. Our study expands the understanding of all biological pathways associated with previously discovered LUAD risk loci and identifies STK24 as a risk gene for LUAD. Research indicates that smoking, air pollution, and tuberculosis are risk factors for lung adenocarcinoma (LUAD). Chronic smoking is a recognized risk factor for various lung diseases, including lung cancer, as it contributes to the development of mutations and cellular changes in lung tissue. Additionally, air pollution has become a significant environmental risk factor, with evidence showing that exposure to pollutants such as particulate matter (PM), nitrogen dioxide (NO2), and sulfur dioxide (SO2) exacerbates respiratory conditions and increases the incidence of lung cancer. Furthermore, the relationship between air pollution and tuberculosis (TB) is increasingly recognized, as studies suggest that environmental air pollution can elevate the risk of TB, potentially contributing to the development of lung diseases, including LUAD [53, 54]. Our findings are consistent with these studies, further supporting our conclusions.

Apoptosis plays a critical role in the pathogenesis of lung adenocarcinoma (LUAD). Recent studies underscore the significance of various forms of regulated cell death, including apoptosis, in cancer biology. The B-cell lymphoma 2 (Bcl-2) protein family, comprising pro-apoptotic and anti-apoptotic members, is essential in regulating apoptosis in response to cellular stress and damage. In LUAD, dysregulation of these proteins can lead to

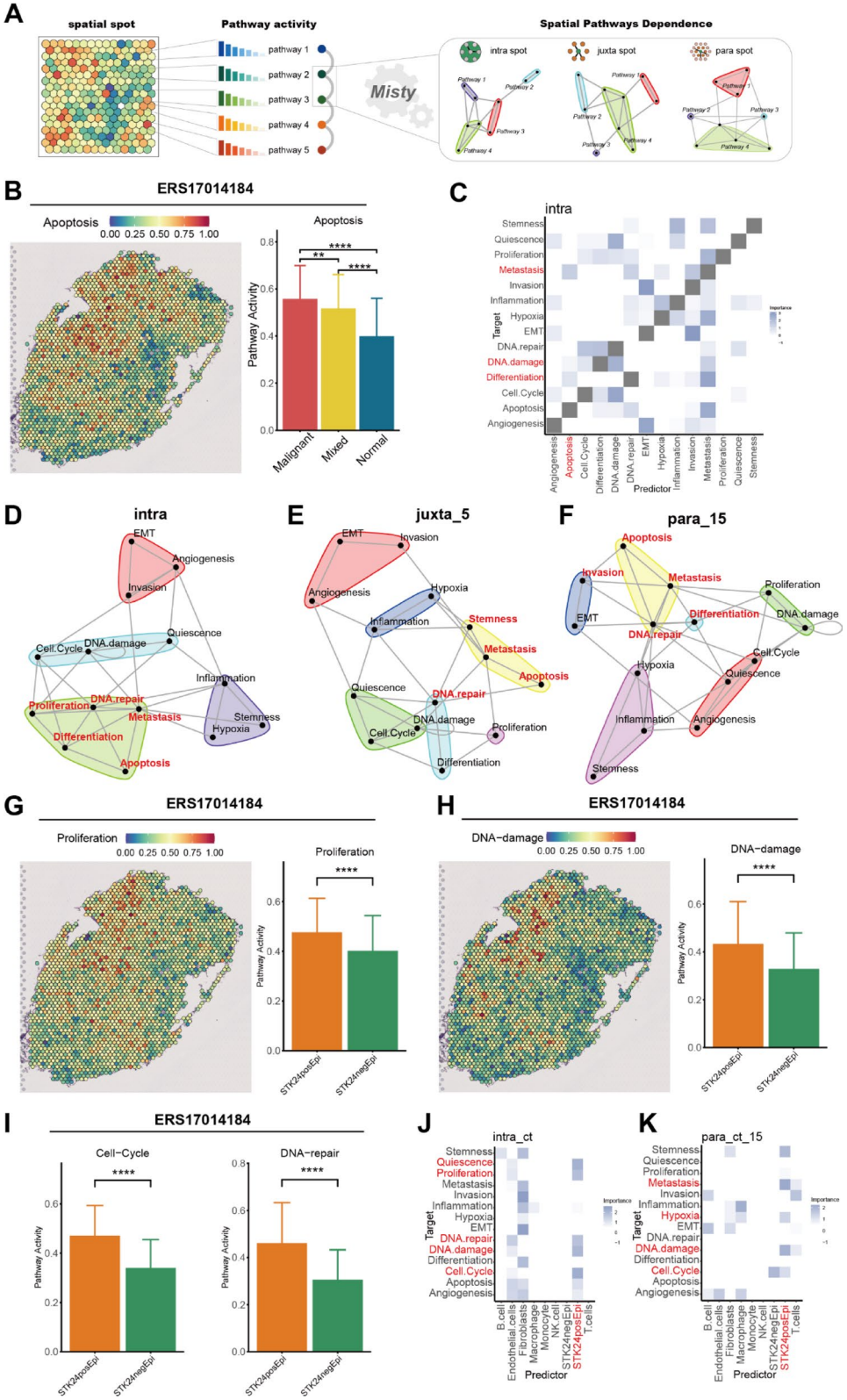

Fig. 9 (See legend on next page.)

(See figure on previous page.)

Fig. 9 Exploration of apoptosis and STK24posEpi-related pathways in spatial transcriptomics (ST). **A** Schematic diagram of Pathway dependency analysis of spatial transcriptomics data. **B** Enrichment results for the ST apoptosis pathway and comparison of differences between regions. **C** Heatmap displaying apoptosis-dependent cell pathways within regions in the spatial context. **D**, **F** Network diagrams showing apoptosis-dependent cell pathways in intra (**D**), juxta_5 (**E**), and para_15 (**F**) regions. **G** Enrichment results for the ST cell proliferation pathway and comparison of differences between the STK24 Group. **H** Enrichment results for the ST cell damage pathway and comparison of differences between the STK24 Group. **I** Comparison of ST cell cycle and DNA repair pathways between the STK24 Groups. **J**, **K** Heatmaps showing cell pathway dependency analysis for different cell types within the intra (**J**) and para_15 (**K**) regions in the spatial context. ****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, ns $P > 0.05$

apoptosis resistance, allowing cancer cells to survive and proliferate despite therapeutic interventions [55]. Additionally, the interplay between apoptosis and other forms of cell death, such as pyroptosis and ferroptosis, has gained increasing recognition in cancer. Pyroptosis, characterized by the activation of inflammatory caspases, can also impact the tumor microenvironment and immune response, potentially influencing tumor progression and response to treatment [56]. Similarly, ferroptosis, a form of iron-dependent cell death, is implicated in the pathogenesis of various cancers, including LUAD, potentially leading to oxidative stress and inflammation [57]. Understanding the mechanisms of interaction between apoptosis and other regulated cell death pathways in LUAD can provide insights into new therapeutic strategies. For instance, targeting apoptotic pathways or enhancing ferroptosis in LUAD cells may improve treatment outcomes and overcome resistance to conventional therapies [55, 57]. In summary, apoptosis is a key factor in LUAD pathogenesis, and its regulation, along with interactions with other cell death mechanisms, presents a promising avenue for further research and therapeutic development. Our study identifies STK24 as an apoptosis-related gene with elevated expression in LUAD epithelial cells.

In the analysis combining single-cell and spatial transcriptomics, we found strong communication between STK24-positive epithelial cells (STK24posEpi) and fibroblasts, with key signaling molecules including PDGF, MIF, and VEGF. Tumor cells and fibroblasts play a critical role in shaping the tumor microenvironment through various signaling pathways and interactions. Platelet-derived growth factor (PDGF) is one of the key factors in this dynamic, known to promote fibroblast activation and proliferation, thus enhancing the supportive stroma around the tumor. This interaction is crucial for tumor progression, as activated fibroblasts, often called cancer-associated fibroblasts (CAFs), secrete a range of growth factors and cytokines, leading to inflammatory signaling in the microenvironment [58]. Additionally, macrophage migration inhibitory factor (MIF) has been identified as another important factor in this context. MIF not only supports tumor cell survival and invasion but also influences fibroblast behavior within the tumor microenvironment. Studies have shown that MIF can enhance the expression of inflammatory cytokines, further promoting a tumor-supportive environment [59]. Furthermore, vascular endothelial growth factor (VEGF) is essential for angiogenesis and is often upregulated in response to inflammatory signals produced by tumor cells and fibroblasts, resulting in increased blood supply that promotes tumor growth and metastasis [60]. The interaction among these factors (tumor cells, fibroblasts, PDGF, MIF, VEGF, and other inflammatory signals) creates a complex microenvironment that not only supports tumor growth but also contributes to therapy resistance. Understanding these interactions is crucial for developing targeted therapies that can disrupt these tumor-promoting signals and potentially improve patient prognosis [61–63].

Our findings show that STK24posEpi is associated with poor prognosis in lung adenocarcinoma (LUAD) patients and a decrease in tumor-infiltrating lymphocytes (TILs). Specifically, in cases of high STK24 expression, both natural killer (NK) cells and B cells are significantly reduced. This observation aligns with other studies highlighting the importance of the tumor microenvironment in LUAD, where immune cell infiltration plays a critical role in patient prognosis [64]. Furthermore, higher tumor grading indicates a lower proportion of STK24posEpi in tumors, suggesting that the presence of this specific epithelial marker decreases as tumor aggressiveness increases. The correlation between tumor grading and immune cell infiltration has been demonstrated in various cancers, underscoring the potential of immune-related biomarkers as prognostic indicators [65]. The relationship between STK24 expression and immune cell dynamics in LUAD emphasizes the need for further research into the mechanisms by which STK24 affects the tumor microenvironment and immune evasion. Understanding these interactions could pave the way for new therapeutic strategies aimed at enhancing the immune response in LUAD patients [66]. In the context of LUAD, the upregulation of apoptosis-related genes such as STK24 may be a compensatory mechanism in response to a tumor microenvironment often characterized by increased inflammatory signaling and altered cellular signaling pathways [67]. Understanding the regulatory mechanisms of these apoptosis-related genes can provide insights into potential therapeutic targets for LUAD, as emphasized by the identification of hub genes that impact patient survival [68]. Overall, the increased expression of STK24 in LUAD epithelial cells highlights the complex relationship between apoptosis and cancer biology, suggesting that further research into its role may

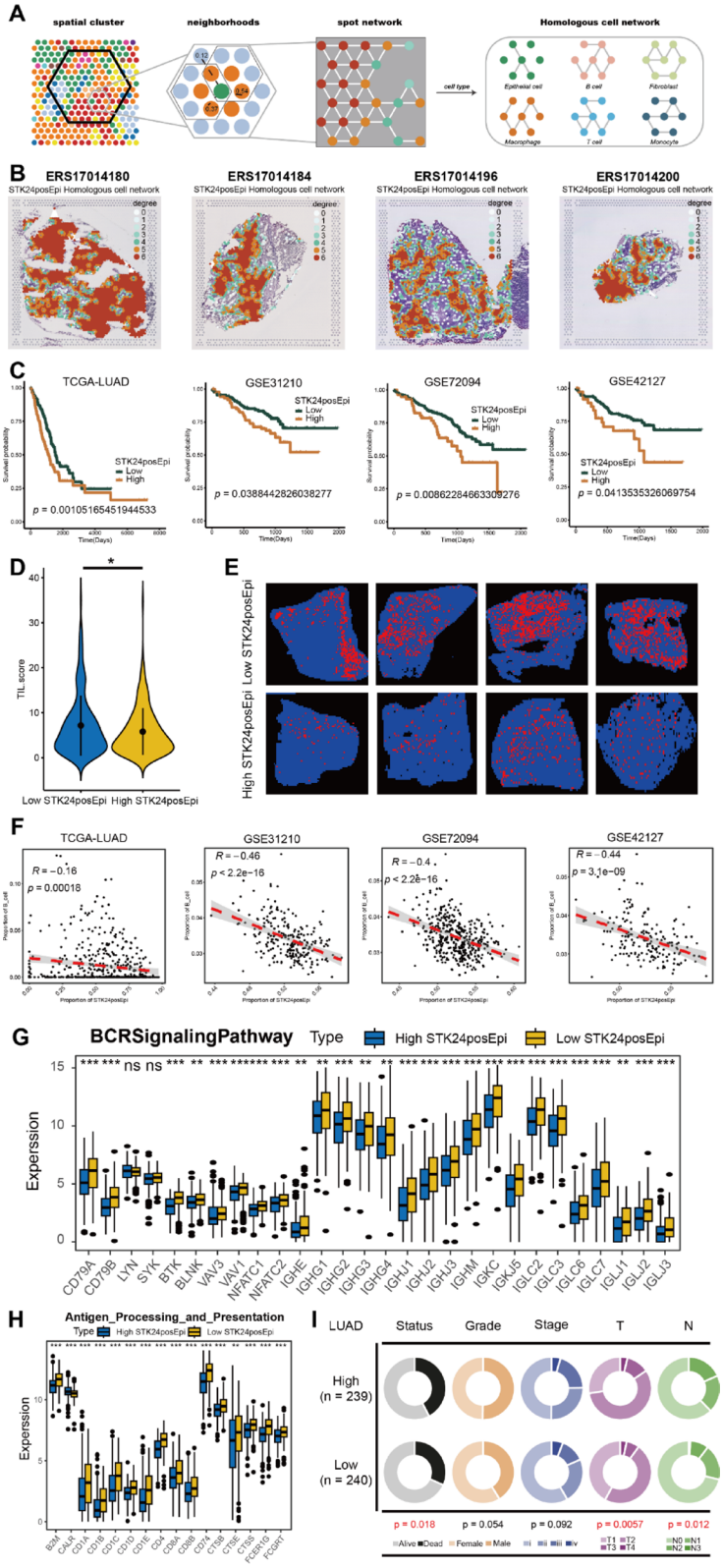

Fig. 10 (See legend on next page.)

(See figure on previous page.)

Fig. 10 Clinical significance of STK24posEpi. **A** Schematic diagram of Homotypic cellular network analysis of spatial transcriptomics data. **B** Homotypic cell network analysis of STK24posEpi in spatial transcriptomics. **C** Survival analysis of STK24posEpi across multiple bulk transcriptome cohorts after Bayesian deconvolution. **D** Comparison of tumor-infiltrating lymphocyte scores between STK24posEpi Groups in the TCGA-LUAD cohort. **E** Histological slides showing differences in tumor-infiltrating lymphocytes between STK24posEpi Groups in the TCGA-LUAD cohort. **F** Correlation analysis of STK24posEpi and B cells in multiple bulk transcriptomes. **G** Differential expression of BCR signaling pathway-related genes between STK24posEpi Groups in the TCGA-LUAD cohort. **H** Differential expression of antigen processing and presentation pathway-related genes between STK24posEpi Groups in the TCGA-LUAD cohort. **I** Comparison of clinical factors between STK24posEpi Groups in the TCGA-LUAD cohort. ****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, ns $P > 0.05$

provide valuable information for the development of targeted therapies for lung adenocarcinoma.

We developed a risk model using Machine learning, named STK24SuperPC, to predict the immunotherapy response in LUAD patients and guide medication. In subsequent experimental validation, we found that STK24 promotes the proliferation of STK24-positive cells and inhibits their apoptosis. However, our study has some limitations, particularly the lack of in vitro validation, so the conclusions still require further confirmation. In summary, our findings suggest that STK24posEpi may serve as a valuable prognostic marker in LUAD, reflecting not only the tumor's biological behavior but also its interaction with the immune landscape, which is important for developing effective immunotherapies.

Conclusion

STK24, an apoptosis-related gene, promotes LUAD progression by enhancing proliferation and inhibiting apoptosis, with STK24posEpi interacting in the tumor microenvironment through PDGF and MIF pathways. These findings offer potential therapeutic targets and support personalized LUAD treatment strategies.

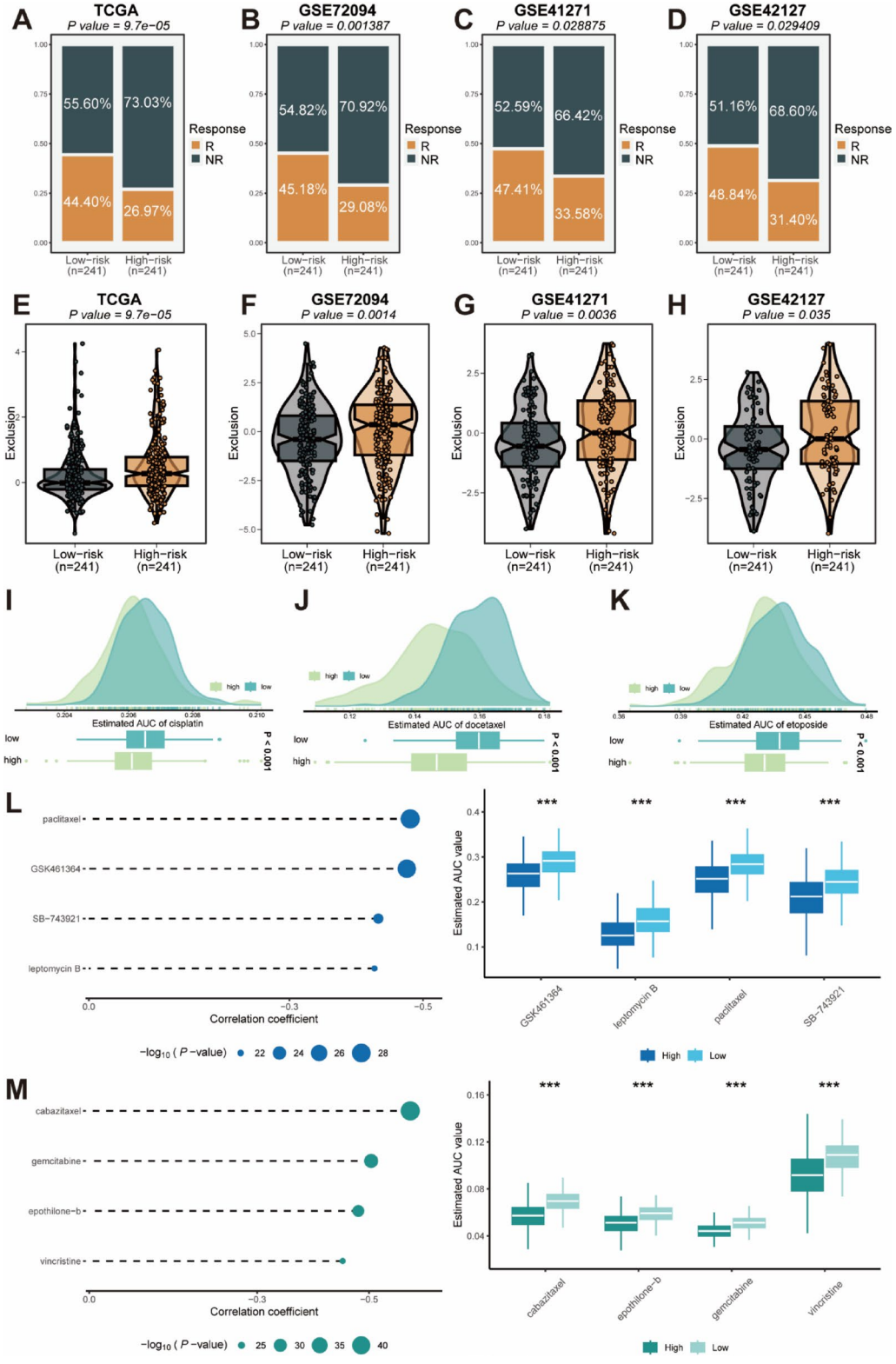

Fig. 11 Machine learning model based on STK24posEpi marker genes for discovering potential therapeutic drugs. **A**–**D** Prediction of immunotherapy response in TCGA and other GEO cohorts via machine learning models. **E**–**H** Differences in Exclusion scores between high-risk and low-risk groups in TCGA and other GEO cohorts. **I**–**K** Drug sensitivity analysis of cisplatin (**I**), docetaxel (**J**), and etoposide (**K**) in risk score groups. **L**–**M** Analysis of potential small molecule drugs based on high-risk and low-risk score groups. ****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, ns $P > 0.05$

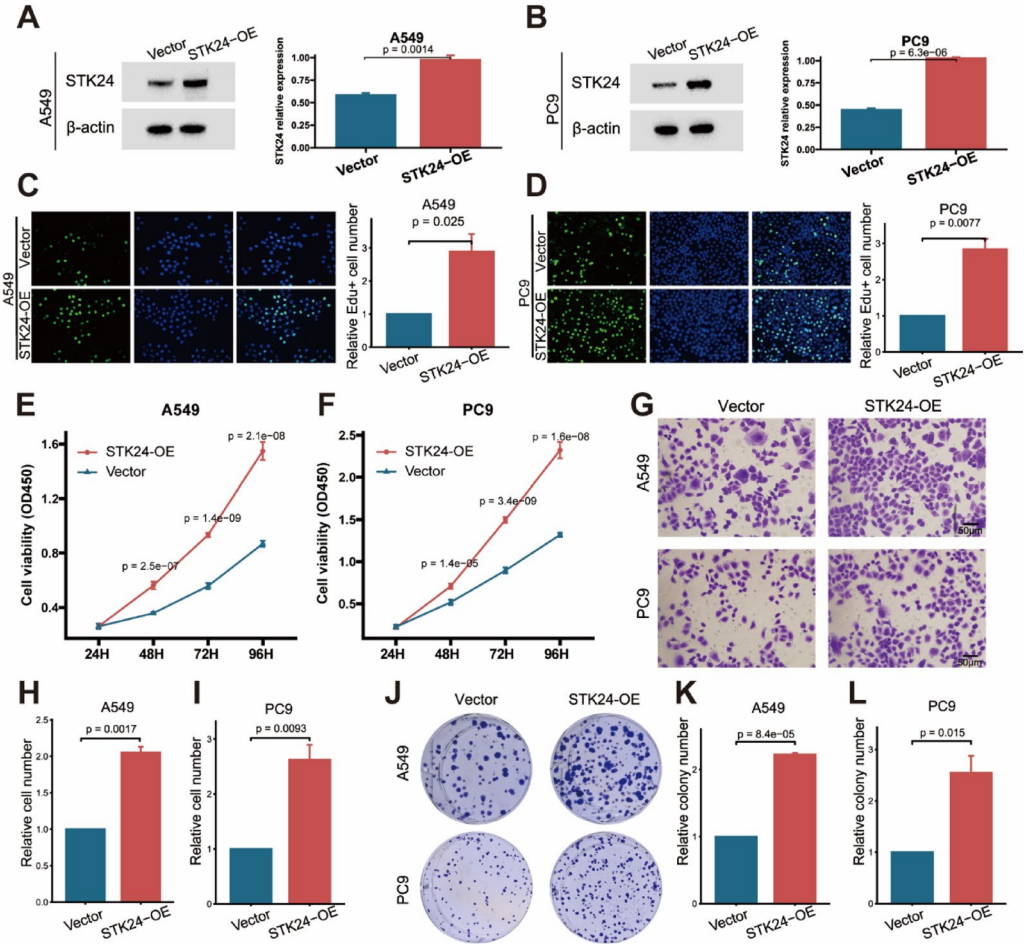

Fig. 12 Overexpression of STK24 promotes proliferation and invasion of LUAD cells. **A**, **B** Western blot analysis showing the efficiency of STK24 overexpression in A549 and PC9 cells. **C**, **D** EdU assay and bar charts showing that STK24 overexpression promotes DNA replication in A549 and PC9 cells. **E**, **F** CCK8 assay demonstrating that STK24 overexpression promotes the growth of A549 and PC9 cells. **G**–**I** Transwell invasion assay and bar charts showing that STK24 promotes the invasion capability of A549 and PC9 cells. **J**–**L** Colony formation assay and bar charts showing that STK24 promotes colony formation in A549 and PC9 cells

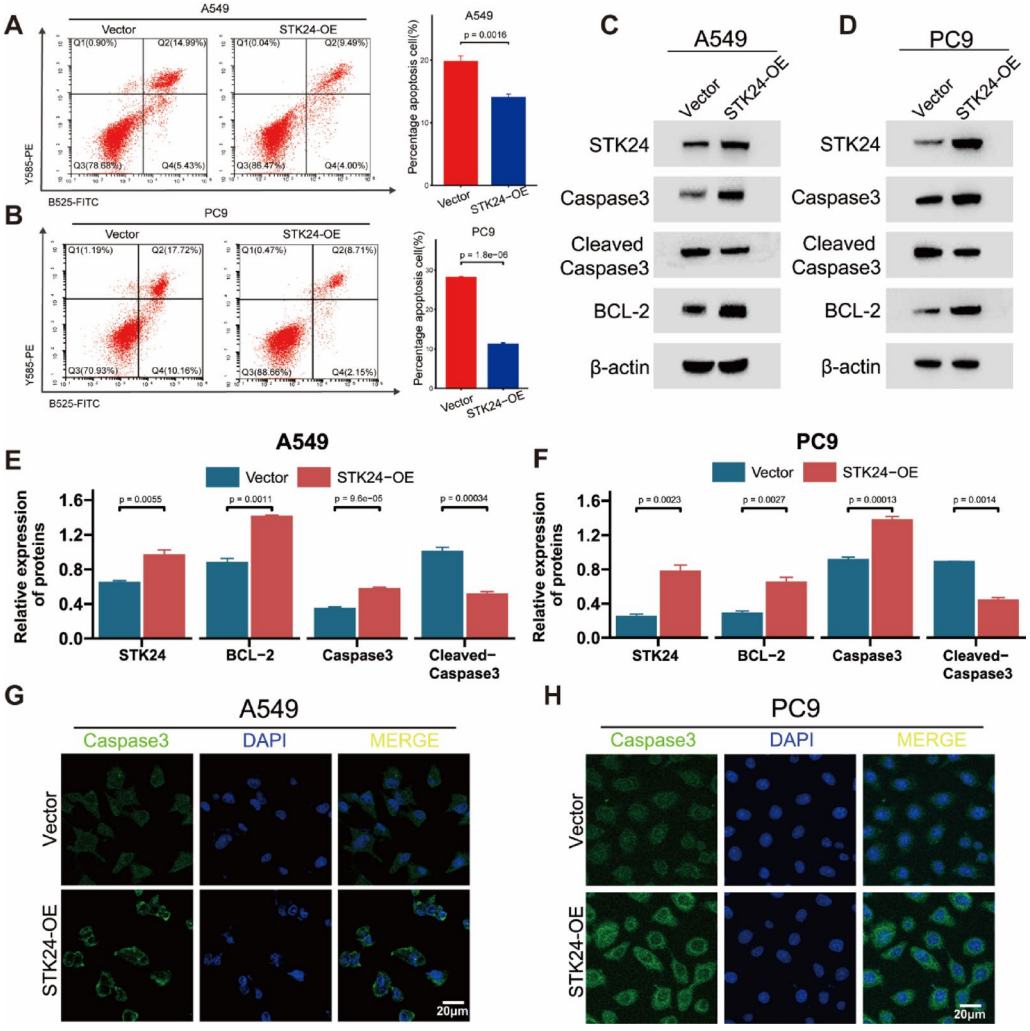

Fig. 13 Overexpression of STK24 inhibits apoptosis in LUAD cells. **A**, **B** Flow cytometry analysis showing that STK24 overexpression inhibits the apoptosis rate in A549 and PC9 cells. **C**–**F** Western blot analysis showing that STK24 overexpression promotes the expression of BCL-2 and Caspase 3, and inhibits the expression of Cleaved-Caspase 3, with the statistical analysis presented in bar charts. **G**, **H** Immunofluorescence assay showing that STK24 overexpression increases the expression of Caspase 3

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-025-07111-z.

Supplementary Material 1

Supplementary Table 1

Supplementary Table 2

Supplementary Table 3

Supplementary Table 4

Supplementary Table 5

Acknowledgements
None.

Author contributions
The Corresponding author is Meng Hou. (I) Conception and design of the experiments: JZL and HML; (II) Experimentation: HML and JZL (III) Administrative support: MH; (IV) Data collection and assembly: JZL, HML, and YHJ; (V) Data analysis and writing: JZL, and HML. All authors have read and approved the final version of the manuscript.

Funding
None.

Data availability
The open-access datasets are available on the following URLs: GSE127465, GSE131907, GSE149655, GSE153935 (https://www.ncbi.nlm.nih.gov/geo), the ST-data (https://www.ebi.ac.uk/biostudies/arrayexpress/studies/), and the Cancer Genome Atlas (TCGA) database (https://xena.ucsc.edu).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
None.

Competing interests
None declared.

Author details
[1]Department of Oncology, Shanghai Pulmonary Hospital, School of Medicine, Tongji University, Shanghai 200433, China
[2]Internal Medicine Residency Training Base, Gongli Hospital of Shanghai Pudong New Area, Shanghai 200135, China
[3]Department of Heart Failure, Shanghai East Hospital, School of Medicine, Tongji University, Shanghai 200120, China

Received: 15 February 2025 / Accepted: 8 September 2025
Published online: 30 October 2025

References

1. Garcia-Campelo R, Bernabe R, Cobo M, Corral J, Coves J, Domine M, et al. SEOM clinical guidelines for the treatment of non-small cell lung cancer (NSCLC) 2015. Clin Transl Oncol. 2015;17(12):1020–9.
2. Sumiyoshi S, Yoshizawa A, Sonobe M, Kobayashi M, Sato M, Fujimoto M, et al. Non-terminal respiratory unit type lung adenocarcinoma has three distinct subtypes and is associated with poor prognosis. Lung Cancer. 2014;84(3):281–8.
3. Guo Y, Cao R, Zhang X, Huang L, Sun L, Zhao J, et al. Recent progress in rare oncogenic drivers and targeted therapy for non-small cell lung cancer. Onco Targets Ther. 2019;12:10343–60.
4. Roca E, Pozzari M, Vermi W, Tovazzi V, Baggi A, Amoroso V, et al. Outcome of EGFR-mutated adenocarcinoma NSCLC patients with changed phenotype to squamous cell carcinoma after tyrosine kinase inhibitors: a pooled analysis with an additional case. Lung Cancer. 2019;127:12–8.
5. Song J, Liang K, Wei T, Li L, Huang Z, Chen G, et al. Expression and predictive significance of FHL1 and SLIT3 in surgically resected lung adenocarcinoma. Comb Chem High Throughput Screen. 2023;26(12):2226–37.
6. Zhang L, Xiong Y, Zhang J, Feng Y, Xu A. Systematic proteome-wide Mendelian randomization using the human plasma proteome to identify therapeutic targets for lung adenocarcinoma. J Transl Med. 2024;22(1):330.
7. Silverman EK. Applying functional genomics to chronic obstructive pulmonary disease. Ann Am Thorac Soc. 2018;15(Suppl 4):S239–42.
8. Inshaw JRJ, Cutler AJ, Burren OS, Stefana MI, Todd JA. Approaches and advances in the genetic causes of autoimmune disease and their implications. Nat Immunol. 2018;19(7):674–84.
9. Zhu Z, Hasegawa K, Camargo CA Jr., Liang L. Investigating asthma heterogeneity through shared and distinct genetics: insights from genome-wide cross-trait analysis. J Allergy Clin Immunol. 2021;147(3):796–807.
10. Wong RS. Apoptosis in cancer: from pathogenesis to treatment. J Exp Clin Cancer Res. 2011;30(1):87.
11. Bai L, Smith DC, Wang S. Small-molecule SMAC mimetics as new cancer therapeutics. Pharmacol Ther. 2014;144(1):82–95.
12. Ergun S, Aslan S, Demir D, Kayaoglu S, Saydam M, Keles Y, et al. Beyond death: unmasking the intricacies of apoptosis escape. Mol Diagn Ther. 2024;28(4):403–23.
13. Ichim G, Tait SW. A fate worse than death: apoptosis as an oncogenic process. Nat Rev Cancer. 2016;16(8):539–48.
14. Li Y, Liu Y, Wang K, Xue D, Huang Y, Tan Z, et al. STK24 promotes progression of LUAD and modulates the immune microenvironment. Mediators Inflamm. 2023;2023:8646088.
15. Chen YL, Wang CY, Fang JH, Hsu HP. Serine/threonine-protein kinase 24 is an inhibitor of gastric cancer metastasis through suppressing CDH1 gene and enhancing stemness. Am J Cancer Res. 2021;11(9):4277–93.
16. Wang N, Jiang Y, Li M, Wang H, Pan J, Tang Y, et al. Protein kinase STK24 promotes tumor immune evasion via the AKT-PD-L1 axis. Adv Sci. 2024;11(12):e2304342.
17. Lai S, Wang D, Sun W, Cao X. Serine/threonine-protein kinase STK24 induces tumorigenesis by regulating the STAT3/VEGFA signaling pathway. J Biol Chem. 2023;299(3):102961.
18. Huang N, Lin W, Shi X, Tao T. STK24 expression is modulated by DNA copy number/methylation in lung adenocarcinoma and predicts poor survival. Future Oncol. 2018;14(22):2253–63.
19. Zhang Y, Tang W, Zhang H, Niu X, Xu Y, Zhang J, et al. A network of interactions enables CCM3 and STK24 to coordinate UNC13D-driven vesicle exocytosis in neutrophils. Dev Cell. 2013;27(2):215–26.
20. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010;26(17):2190–1.
21. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8(1):1826.
22. Rasooly D, Peloso GM, Pereira AC, Dashti H, Giambartolomei C, Wheeler E, et al. Genome-wide association analysis and Mendelian randomization proteomics identify drug targets for heart failure. Nat Commun. 2023;14(1):3826.
23. Lyon MS, Andrews SJ, Elsworth B, Gaunt TR, Hemani G, Marcora E. The variant call format provides efficient and robust storage of GWAS summary statistics. Genome Biol. 2021;22(1):32.
24. Karlsson Linner R, Biroli P, Kong E, Meddens SFW, Wedow R, Fontana MA, et al. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nat Genet. 2019;51(2):245–57.
25. Zhou W, Kanai M, Wu KH, Rasheed H, Tsuo K, Hirbo JB, et al. Global biobank meta-analysis initiative: powering genetic discovery across human disease. Cell Genom. 2022;2(10):100192.
26. Jiang L, Zheng Z, Fang H, Yang J. A generalized linear mixed model association tool for biobank-scale data. Nat Genet. 2021;53(11):1616–21.
27. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016;48(5):481–7.
28. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C, et al. Ld score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015;47(3):291–5.
29. Murphy AE, Schilder BM, Skene NG. Mungesumstats: a Bioconductor package for the standardization and quality control of many GWAS summary statistics. Bioinformatics. 2021;37(23):4593–6.
30. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573–87.
31. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with harmony. Nat Methods. 2019;16(12):1289–96.
32. Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol. 2018. https://doi.org/10.1038/nbt.4314.
33. Zhang Y, Liu T, Hu X, Wang M, Wang J, Zou B, et al. Cell call: integrating paired ligand-receptor and transcription factor activities for cell-cell communication. Nucleic Acids Res. 2021;49(15):8520–34.
34. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol. 2014;32(4):381–6.
35. Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. Nat Methods. 2017;14(11):1083–6.
36. Gulati GS, Sikandar SS, Wesche DJ, Manjunath A, Bharadwaj A, Berger MJ, et al. Single-cell transcriptional diversity is a hallmark of developmental potential. Science. 2020;367(6476):405–11.
37. Cable DM, Murray E, Zou LS, Goeva A, Macosko EZ, Chen F, et al. Robust decomposition of cell type mixtures in spatial transcriptomics. Nat Biotechnol. 2022;40(4):517–26.
38. Schubert M, Klinger B, Klunemann M, Sieber A, Uhlitz F, Sauer S, et al. Perturbation-response genes reveal signaling footprints in cancer gene expression. Nat Commun. 2018;9(1):20.
39. Holland CH, Szalai B, Saez-Rodriguez J. Transfer of regulatory knowledge from human to mouse for functional genomics analysis. Biochem Biophys Acta. 2020;1863(6):194431.
40. Tanevski J, Flores ROR, Gabor A, Schapiro D, Saez-Rodriguez J. Explainable multiview framework for dissecting spatial relationships from highly multiplexed data. Genome Biol. 2022;23(1):97.
41. Kuppe C, Ramirez Flores RO, Li Z, Hayat S, Levinson RT, Liao X, et al. Spatial multi-omic map of human myocardial infarction. Nature. 2022;608(7924):766–77.
42. Pham D, Tan X, Xu J, Grice LF, Lam PY, Raghubar A, Vukovic J, Ruitenberg MJ, Nguyen Q: stLearn: integrating spatial location, tissue morphology and gene expression to find cell types, cell-cell interactions and spatial trajectories within undissociated tissues; 2020.
43. Chu T, Wang Z, Pe'er D, Danko CG. Cell type and gene expression deconvolution with BayesPrism enables Bayesian integrative analysis across bulk and single-cell RNA sequencing in oncology. Nat Cancer. 2022;3(4):505–17.

44. Liu Z, Guo C, Dang Q, Wang L, Liu L, Weng S, et al. Integrative analysis from multi-center studies identities a consensus machine learning-derived lncRNA signature for stage II/III colorectal cancer. EBioMedicine. 2022;75:103750.
45. Li Y, Sun Q, Jiang M, Li S, Zhang J, Xu Z, et al. KLF9 suppresses gastric cancer cell invasion and metastasis through transcriptional inhibition of MMP28. FASEB J. 2019;33(7):7915–28.
46. Shen P, Sun J, Xu G, Zhang L, Yang Z, Xia S, et al. KLF9, a transcription factor induced in flutamide-caused cell apoptosis, inhibits AKT activation and suppresses tumor growth of prostate cancer cells. Prostate. 2014;74(9):946–58.
47. Shen P, Cao X, Sun L, Qian Y, Wu B, Wang X, et al. KLF9 suppresses cell growth and induces apoptosis via the AR pathway in androgen-dependent prostate cancer cells. Biochem Biophys Rep. 2021;28:101151.
48. Zhao L, Wu X, Zhang Z, Fang L, Yang B, Li Y. ELF1 suppresses autophagy to reduce cisplatin resistance via the miR-152-3p/NCAM1/ERK axis in lung cancer cells. Cancer Sci. 2023;114(6):2650–63.
49. Roussigne M, Cayrol C, Clouaire T, Amalric F, Girard JP. THAP1 is a nuclear proapoptotic factor that links prostate-apoptosis-response-4 (Par-4) to PML nuclear bodies. Oncogene. 2003;22(16):2432–42.
50. Cui G, Liu D, Li W, Li Y, Liang Y, Shi W, et al. Original research: miR-194 inhibits proliferation and invasion and promotes apoptosis by targeting KDM5B in esophageal squamous cell carcinoma cells. Exp Biol Med. 2017;242(1):45–52.
51. Huang J, Jin S, Guo R, Wu W, Yang C, Qin Y, et al. Histone lysine demethylase KDM5B facilitates proliferation and suppresses apoptosis in human acute myeloid leukemia cells through the miR-140-3p/BCL2 axis. RNA. 2024;30(4):435–47.
52. Zhang F, Huang K, Chen R, Liu Z, Zhao Q, Hou S, et al. StarTracer is an accelerated approach for precise marker gene identification in single-cell RNA-Seq analysis. Commun Biol. 2024;7(1):1128.
53. Lai TC, Chiang CY, Wu CF, Yang SL, Liu DP, Chan CC, et al. Ambient air pollution and risk of tuberculosis: a cohort study. Occup Environ Med. 2016;73(1):56–61.
54. Lin YJ, Lin HC, Yang YF, Chen CY, Ling MP, Chen SC, et al. Association between ambient air pollution and elevated risk of tuberculosis development. Infect Drug Resist. 2019;12:3835–47.
55. Morsi RZ, Hage-Sleiman R, Kobeissy H, Dbaibo G. Noxa: role in cancer pathogenesis and treatment. Curr Cancer Drug Targets. 2018;18(10):914–28.
56. An S, Hu H, Li Y, Hu Y. Pyroptosis plays a role in osteoarthritis. Aging Dis. 2020;11(5):1146–57.
57. Li S, Li Y, Wu Z, Wu Z, Fang H. Diabetic ferroptosis plays an important role in triggering on inflammation in diabetic wound. Am J Physiol Endocrinol Metab. 2021;321(4):E509–20.
58. Curran CS, Keely PJ. Breast tumor and stromal cell responses to TGF-beta and hypoxia in matrix deposition. Matrix Biol. 2013;32(2):95–105.
59. Garcia-Gerique L, Garcia M, Garrido-Garcia A, Gomez-Gonzalez S, Torrebadell M, Prada E, et al. MIF/CXCR4 signaling axis contributes to survival, invasion, and drug resistance of metastatic neuroblastoma cells in the bone marrow microenvironment. BMC Cancer. 2022;22(1):669.
60. Tjomsland V, Spangeus A, Valila J, Sandstrom P, Borch K, Druid H, et al. Interleukin 1alpha sustains the expression of inflammatory factors in human pancreatic cancer microenvironment by targeting cancer-associated fibroblasts. Neoplasia. 2011;13(8):664–75.
61. Mito I, Takahashi H, Kawabata-Iwakawa R, Horikawa M, Ida S, Tada H, et al. Tumor-derived exosomes elicit cancer-associated fibroblasts shaping inflammatory tumor microenvironment in head and neck squamous cell carcinoma. Oral Oncol. 2023;136:106270.
62. Vera RE, Lamberti MJ, Almada LL, Tolosa EJ, Vrabel AM, Sigafoos AN, et al. GLI1 interaction with p300 modulates SDF1 expression in cancer-associated fibroblasts to promote pancreatic cancer cells migration. Biochem J. 2023;480(3):225–41.
63. Peng Z, Ren Z, Tong Z, Zhu Y, Zhu Y, Hu K. Interactions between MFAP5 + fibroblasts and tumor-infiltrating myeloid cells shape the malignant microenvironment of colorectal cancer. J Transl Med. 2023;21(1):405.
64. Lin J, Wu C, Ma D, Hu Q. Identification of P2RY13 as an immune-related prognostic biomarker in lung adenocarcinoma: a public database-based retrospective study. PeerJ. 2021;9:e11319.
65. Zhao R, Ding D, Yu W, Zhu C, Ding Y. The lung adenocarcinoma microenvironment mining and its prognostic merit. Technol Cancer Res Treat. 2020;19:1533033820977547.
66. Deng C, Li Z, Guo S, Chen P, Chen X, Zhou Q, et al. Tumor PD-L1 expression is correlated with increased TILs and poor prognosis in penile squamous cell carcinoma. Oncoimmunology. 2017;6(2):e1269047.
67. Kao S, Shaterian A, Cauvi DM, Dang X, Chun HB, De Maio A, et al. Pulmonary preconditioning, injury, and inflammation modulate expression of the candidate tumor suppressor gene ECRG4 in lung. Exp Lung Res. 2015;41(3):162–72.
68. Chawhan AP, Dsouza N. Identifying the key hub genes linked with lung squamous cell carcinoma by examining the differentially expressed and survival genes. Mol Genet Genomics. 2024;299(1):76.

Publisher's Note