Article

Discovery Medicine 2026; 38(209): 1927–1944
https://doi.org/10.24976/Discov.Med.202638209.150

Dynamic Regulatory Mechanisms of the DHX9/ILF3 Complex During Lung Adenocarcinoma Cell Differentiation: A Single-Cell Transcriptomic Analysis

Weishu An[1], Mengting Wang[1], Junyi An[1], Xianyi Cao[2], Ming Xu[1,*]

[1]Department of Oncology, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, 201999 Shanghai, China
[2]Department of Nuclear Medicine, Xin Hua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, 200092 Shanghai, China
*Correspondence: mingxu.msu@gmail.com (Ming Xu)

Submitted: 24 March 2026 Revised: 30 April 2026 Accepted: 8 May 2026 Published: 20 June 2026

Background: The DExH-box helicase 9 (DHX9)/interleukin enhancer-binding factor 3 (ILF3) complex participates in RNA processing and post-transcriptional regulation, but its behavior during lung adenocarcinoma cell differentiation remains poorly defined, particularly at single-cell resolution. This study aimed to characterize the dynamic expression patterns and potential biological associations of the DHX9/ILF3 complex during epithelial-state transitions in lung adenocarcinoma using single-cell transcriptomic analysis.

Methods: We analyzed publicly available single-cell RNA sequencing (scRNA-seq) datasets from the Gene Expression Omnibus (GEO), including the discovery dataset GSE131907 and the independent validation dataset GSE123902, to profile lung adenocarcinoma epithelial cells at single-cell resolution. The sequencing data were processed using a standard single-cell analytical workflow, including quality control, dimensionality reduction, clustering, pseudotime trajectory inference, co-expression analysis, and functional program scoring. To verify key transcriptional patterns identified in the single-cell dataset, we measured the mRNA levels of selected genes in A549 and H1975 cells using quantitative real-time PCR.

Results: Single-cell analysis revealed multiple epithelial states, including normal-like epithelial cells, transitional tumor-associated states, and malignant cells. Across these states, core members of the DHX9/ILF3 complex, including DHX9, ILF3, ILF2, and HNRNPU, displayed coordinated yet dynamic expression patterns along pseudotime trajectories. DHX9 expression rose during the early stages of the trajectory and then declined, whereas ILF3 increased more gradually and remained relatively elevated in intermediate states. Correlation analysis indicated positive associations among key complex members at the single-cell level. Cells with higher DHX9/ILF3 complex activity exhibited stronger malignant transcriptional programs, including proliferation, stemness, and MYC-related signaling. External validation using the GSE123902 dataset reproduced the state-dependent expression patterns and DHX9–ILF3 correlation identified in the discovery dataset. qRT-PCR confirmed coordinated upregulation of DHX9, ILF3, ILF2, and HNRNPU in lung adenocarcinoma cell lines compared with normal epithelial cells, with similar patterns observed in A549 and H1975 cells.

Conclusions: The DHX9/ILF3 complex exhibits coordinated yet dynamic activity during epithelial state transitions in lung adenocarcinoma. Its expression is associated with malignant transcriptional programs and is reproducible across independent single-cell datasets. These findings refine current understanding of RNA-binding protein–mediated regulation of tumor cell states and suggest the DHX9/ILF3 complex as a potential regulatory axis in lung adenocarcinoma.

Keywords: DHX9/ILF3 complex; single-cell RNA sequencing; lung adenocarcinoma; cell differentiation

Introduction

Lung adenocarcinoma is the most common subtype of non-small cell lung cancer (NSCLC), accounting for about 40% of cases and remaining a leading cause of cancer-related death worldwide [1,2]. Targeted and immune-based therapies have improved outcomes in selected patients, but overall survival remains unsatisfactory. A key challenge is the pronounced heterogeneity within tumors, which contributes to inconsistent treatment responses and makes it difficult to trace how cancer cells acquire or alter their differentiation states [3,4]. Clarifying the regulatory factors that guide these transitions may help uncover new therapeutic targets.

RNA-binding proteins (RBPs) serve as pivotal regulators of post-transcriptional processes, including alternative splicing, RNA turnover, intracellular transport, and protein synthesis. The DExH-box helicase 9 (DHX9)/interleukin enhancer-binding factor 3 (ILF3) complex has attracted attention for its involvement in multiple layers of RNA metabolism and its ability to form multi-protein assemblies with diverse partners. DHX9 (RNA helicase A) is

an ATP-dependent helicase involved in transcriptional regulation, RNA processing, and genome surveillance [5,6]. ILF3 (NF90/NF110 isoforms) binds double-stranded RNA and modulates transcript stability and translation [7,8]. Together with ILF2, HNRNPU, SFPQ, NONO, and FET-family proteins (FUS, EWSR1, TAF15), these factors assemble into ribonucleoprotein complexes that respond to developmental and environmental cues [9].

Accumulating evidence suggests that abnormal expression of components of the DHX9/ILF3 complex contributes to tumor progression. DHX9 is upregulated in several cancers, including breast, hepatocellular, and colorectal carcinoma, where it supports proliferation, genome maintenance, and metastasis [6,10,11]. ILF3 is likewise increased in many tumors and has been associated with poor prognosis, probably through stabilization of oncogenic transcripts [12]. However, how this complex functions as an integrated unit during cancer cell differentiation, and how its components change over time at single-cell resolution, remains unclear.

Bulk RNA sequencing averages transcriptional signals across mixed cell populations and therefore masks intermediate or rare cellular states. In contrast, single-cell RNA sequencing (scRNA-seq) resolves transcriptional profiles at the individual-cell level, allowing reconstruction of differentiation continua and identification of heterogeneous subgroups [13,14]. This approach has reshaped our view of tumor evolution, drug resistance, and tumor–microenvironment interactions [15].

Cellular differentiation proceeds through a sequence of molecular events in which transcriptional and post-transcriptional regulation mechanisms gradually establish distinct cellular phenotypes. In tumors, disruption of this process can support phenotypic plasticity, treatment failure, and metastatic spread. Multiprotein assemblies, such as the DHX9/ILF3 complex, may function at key regulatory points, yet how they form, maintain appropriate component ratios, and relate to cell-cycle control in lung adenocarcinoma remains poorly understood.

Recent computational developments, including pseudotime inference, network modeling, and multidimensional mapping, offer analytical tools for interpreting scRNA-seq datasets. Integrating these approaches can reveal how differentiation progresses and clarify the involvement of particular regulatory complexes.

In this work, we analyzed single-cell transcriptomes of lung adenocarcinoma cells and applied several complementary computational approaches to explore the behavior of the DHX9/ILF3 complex during differentiation, profiling thirteen of its components across a large number of individual cells.

We mapped differentiation paths, characterized temporal and cell-cycle-linked expression patterns, and constructed low- and three-dimensional expression landscapes. Key findings were further validated using qRT-PCR in A549 and H1975 cells. Collectively, these data demonstrate how the DHX9/ILF3 complex behaves as a coordinated regulatory unit and support its potential relevance as a therapeutic target.

Materials and Methods

Cell Culture and Sample Preparation

Human lung adenocarcinoma cell lines A549 (ATCC, CCL-185; Manassas, VA, USA) and NCI-H1975 (ATCC, CRL-5908; Manassas, VA, USA) were used as tumor models. Both cell lines were cultured in RPMI-1640 medium (Gibco, Thermo Fisher Scientific, Grand Island, NY, USA; Cat. No. 11875-093) supplemented with 10% fetal bovine serum (FBS; Gibco, Thermo Fisher Scientific; Cat. No. 10099-141) and 1% penicillin–streptomycin solution (Invitrogen, Thermo Fisher Scientific, Carlsbad, CA, USA; Cat. No. 15140-122). Normal human bronchial epithelial (NHBE) cells (Lonza, Basel, Switzerland; Cat. No. CC-2541) served as controls and were maintained in bronchial epithelial growth medium (BEGM BulletKit; Lonza; Cat. No. CC-3170) according to the manufacturer's protocol. All cells were kept at 37 °C in a humidified incubator with 5% CO_2 and passaged at 80–90% confluence, usually every 2–3 days. Only passages 3–10 were used to reduce passage-related drift. Viability was assessed by trypan blue exclusion, and preparations with >95% viable cells were used for downstream experiments. Cell line identity was verified by short tandem repeat (STR) profiling before experimentation. Genomic DNA was analyzed using the GlobalFiler PCR Amplification Kit (Cat. No. 4476135; Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA), followed by capillary electrophoresis on an ABI 3500 Genetic Analyzer (Applied Biosystems, Thermo Fisher Scientific, Foster City, CA, USA). STR profiles were analyzed using GeneMapper ID-X Software v1.6 (Applied Biosystems, Thermo Fisher Scientific, Foster City, CA, USA). The STR patterns matched the reference profiles for A549 and NCI-H1975 cells, and NHBE cells were authenticated according to the supplier's reference information. Mycoplasma contamination was routinely examined using the MycoSEQ Mycoplasma Detection Kit (Cat. No. 4460626; Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA), and all cell cultures were confirmed to be mycoplasma-free before use.

Single-Cell RNA-Seq Data Acquisition

Single-cell RNA-seq data used for the discovery analysis were obtained from the Gene Expression Omnibus (GEO) database under accession number GSE131907, and the independent validation dataset was obtained under accession number GSE123902. The GSE131907 dataset was used to construct the main epithelial atlas, pseudotime trajectory, DHX9/ILF3 complex activity analyses, and downstream transcriptional programs, whereas GSE123902 was

used as an external validation cohort. Both datasets were generated using the 10× Genomics Chromium platform. Raw sequencing data and/or processed count matrices were downloaded and subjected to standard quality-control procedures. Data quality was assessed by examining UMI abundance, gene coverage, and the proportion of mitochondrial transcripts.

RNA Extraction and Quantitative Real-Time PCR

Total RNA was extracted using TRIzol Reagent (Invitrogen, Thermo Fisher Scientific; Cat. No. 15596026) according to the manufacturer's instructions. RNA concentration and purity were measured using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Samples with A260/A280 ratios of 1.8–2.0 were used for cDNA synthesis. For reverse transcription, 1 µg of total RNA was processed using the PrimeScript RT Reagent Kit with gDNA Eraser (Takara Bio, Shiga, Japan; Cat. No. RR047A) in a 20-µL reaction. Quantitative real-time PCR was performed on a QuantStudio 5 Real-Time PCR System (Applied Biosystems, Thermo Fisher Scientific; Model No. A28574) using TB Green Premix Ex Taq II (Takara Bio; Cat. No. RR820A). Each 20-µL reaction contained 10 µL TB Green mix, 0.4 µL of each primer (10 µM), 2 µL diluted cDNA, and nuclease-free water. Cycling conditions were as follows: 95 °C for 30 s, followed by 40 cycles of 95 °C for 5 s and 60 °C for 34 s. Melt-curve analysis was performed to confirm amplification specificity. All reactions were run in triplicate. GAPDH was used as the internal control, and relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method. Primer sequences were: DHX9: F 5′-CAGGAGAGAGAGTTACTGCCT-3′; R 5′-CTCTGCTGCTCGGTCATTCTG-3′; ILF3: F 5′-CATTACGCCCATGAAACGCC-3′; R 5′-TAAAGATGGGGGCATGGACG-3′; ILF2: F 5′-AGGCCCTTTGTACCACATATC-3′; R 5′-ATCCTGTGCTCTTAGGCTTTC-3′; HNRNPU: F 5′-GAGATTGCTGCCCGAAAGAAGC-3′; R 5′-TTCGCTGGAAGCCTGCAAACAG-3′; GAPDH: F 5′-GCACCGTCAAGGCTGAGAAC-3′; R 5′-TGGTGAAGACGCCAGTGGA-3′.

Single-Cell RNA-Seq Data Processing and Quality Control

Raw reads were processed with Cell Ranger v6.1.2 (10× Genomics) for demultiplexing, alignment to GRCh38, barcode assignment, and UMI counting. Gene–barcode matrices were imported into R 4.2.0 and analyzed using Seurat v4.3.0. For the discovery dataset (GSE131907), cells were retained if they expressed 200–6000 genes, contained 500–50,000 UMIs, and had <10% mitochondrial transcripts. For the validation dataset (GSE123902), the same general filtering strategy was applied, with minor dataset-specific adjustments when necessary to preserve high-quality epithelial cells. Putative doublets were removed using DoubletFinder v2.0.3. Expression data were normalized using Seurat's LogNormalize method with a scale factor of 10,000, and 2000 highly variable genes were identified using the "vst" method. The ScaleData function was used to center and scale expression values while regressing out mitochondrial content and cell-cycle scores. Patient- or sample-specific effects were not directly regressed out during this gene-level scaling step, because doing so could remove biologically meaningful inter-sample variation in epithelial tumor states. Instead, the discovery and validation cohorts were processed independently, and sample-level technical heterogeneity within each dataset was addressed during downstream integration using Harmony, with sample identity specified as the batch variable. After batch correction, low-dimensional embeddings were recalculated from the Harmony space for downstream analyses. Batch-correction performance was further evaluated quantitatively using integration local inverse Simpson's index (iLISI) together with batch-adjusted and cell-type silhouette metrics.

Dimensionality Reduction and Cell Clustering

Principal component analysis (PCA) was performed on the variable genes (RunPCA). Thirty principal components were chosen based on elbow plots. Shared nearest-neighbor graphs were built with FindNeighbors, and unsupervised clustering was performed with FindClusters (Louvain algorithm, resolution 0.8). UMAP embeddings were generated by RunUMAP. For the main analysis, epithelial-related cells were extracted from the integrated lung adenocarcinoma (LUAD) single-cell dataset and re-clustered to construct the epithelial-only atlas. Cell identities were assigned by combining canonical marker expression with SingleR v1.10.0, which compares single-cell profiles with reference datasets. Major epithelial states, including alveolar epithelial type I (AT1), alveolar epithelial type II (AT2), Club, Ciliated, tumor-associated transitional states (tS1, tS2, and tS3), and Malignant cells, were annotated based on established marker genes and cluster-specific transcriptional features.

Pseudotime Trajectory Analysis

Seurat objects were converted to Monocle 3 (v1.3.1) cell_data_set objects. Trajectories were learned by learn_graph on the UMAP manifold. Root cells were defined from relatively differentiated epithelial populations located at the early trajectory region, and pseudotime values were then calculated for downstream dynamic analyses. Cell states were subsequently grouped into normal-like, transitional, and malignant stages according to pseudotime position and cluster identity. Dynamic expression changes of DHX9/ILF3 complex members and complex module scores along pseudotime were visualized using smoothed trend curves. To further assess dynamic coordination be-

tween DHX9 and ILF3, rolling local Spearman correlations were calculated across ordered pseudotime windows.

Differential Expression and Functional Enrichment Analysis

Seurat's FindMarkers (Wilcoxon rank-sum test) was used for differential expression analysis. Genes with adjusted $p < 0.05$ and $\log_2$ fold change >0.25 were considered significant. For volcano-plot visualization, a more stringent effect-size threshold ($|\log_2$ fold change$| > 0.5$) was used to highlight genes with stronger expression differences. Gene Ontology (GO) and Hallmark/pathway enrichment analyses were performed with clusterProfiler v4.4.0. Results were visualized with volcano, dot, and bar plots. For malignant-state comparisons, differential expression analysis was performed between malignant epithelial cells and relatively normal-like epithelial populations. For complex-associated analyses, cells were stratified into DHX9-high versus DHX9-low and ILF3-high versus ILF3-low groups based on gene-expression thresholds, or into complex-high versus complex-low groups based on DHX9/ILF3 complex-score thresholds. In the primary analysis, the "high" and "low" groups were defined as cells within the upper and lower tertiles of the corresponding gene-expression value or module score, respectively, whereas the middle tertile was excluded from dichotomized comparisons to increase contrast between groups. To evaluate the robustness of the findings to threshold selection, sensitivity analyses were additionally performed using alternative cutoff strategies, including median-based dichotomization and upper-versus-lower quartile comparisons. The major associations remained directionally consistent across these analyses and are shown in **Supplementary Fig. 1**.

Gene Co-Expression Network Analysis

To explore coordinated expression among DHX9/ILF3 complex components, pairwise Spearman correlations were calculated across epithelial cells for DHX9, ILF3, ILF2, HNRNPU, SFPQ, and NONO. The resulting correlation matrix was used to visualize co-expression relationships among core complex members. Genes showing coordinated expression with the DHX9/ILF3 complex were assembled into co-expression networks using igraph v1.3.5, and hub genes were ranked according to weighted connectivity. Network-associated gene programs were further characterized by differential expression and functional enrichment analyses.

Cell Cycle and Functional Program Scoring

Cell cycle phases were assigned on the basis of S-phase and G2/M-phase scores using Seurat's CellCycleScoring function. In parallel, malignant-associated transcriptional programs, including proliferation, epithelial–mesenchymal transition (EMT), stemness, hypoxia, glycolysis, PI3K-mTOR signaling, MYC targets, and DNA repair, were quantified at the single-cell level using curated gene sets and AddModuleScore. These program scores were compared across epithelial states, pseudotime stages, and complex-high versus complex-low groups.

DHX9/ILF3 Complex Score Construction and Validation Analysis

A DHX9/ILF3 complex score was calculated using AddModuleScore on the basis of representative complex members, including DHX9, ILF3, ILF2, HNRNPU, SFPQ, and NONO. This score was used to evaluate state-specific DHX9/ILF3 complex activity, dynamic trajectory changes, and associations with malignant phenotypes in the discovery dataset. For external validation, the same scoring strategy was applied to epithelial cells in GSE123902, and the reproducibility of DHX9 expression, ILF3 expression, and DHX9/ILF3 complex score across epithelial states was assessed between the discovery and validation datasets.

Statistical Analysis

Statistical analyses were performed in R 4.2.0 and GraphPad Prism 9.0 (GraphPad Software, San Diego, CA, USA). qRT-PCR data were analyzed using one-way ANOVA with Tukey's post hoc test. For single-cell analyses, Wilcoxon rank-sum tests, Kruskal–Wallis tests, or Spearman correlation analyses were applied as appropriate. Data are shown as mean $\pm$ SD from at least three independent experiments. $p < 0.05$ was considered significant ($p < 0.01$ and $p < 0.001$ as indicated). For scRNA-seq, multiple testing was controlled by the Benjamini–Hochberg false discovery rate. Figures were generated with ggplot2, ComplexHeatmap, and custom scripts, and assembled in Adobe Illustrator v26.0 (Adobe Inc., San Jose, CA, USA).

Results

Epithelial Cell Landscape and Subtype Characteristics in LUAD

Single-cell RNA sequencing data from the GSE131907 cohort were used to characterize the epithelial landscape of LUAD. After integration of epithelial-related cells, UMAP visualization revealed multiple clearly separated clusters, indicating pronounced transcriptional heterogeneity within the epithelial compartment (Fig. 1A). Based on canonical marker genes, epithelial cells were further annotated into several subtypes, including malignant epithelial cells, AT1, AT2, ciliated cells, club cells, and several tumor-associated epithelial states (tS1, tS2, and tS3), together with a small population of undetermined cells (Fig. 1B). Cells derived from different samples were broadly distributed across these clusters rather than forming sample-specific groups, suggesting that the integration procedure effectively minimized batch effects (Fig. 1C). Examination of lineage markers showed that canonical epithelial genes such as EPCAM, KRT19, and MUC1 were

widely expressed across epithelial populations, whereas subtype-specific markers displayed distinct enrichment patterns, including SFTPA1, SFTPC, and NAPSA in AT2 cells, AGER in AT1 cells, SCGB1A1 in club cells, and FOXJ1 in ciliated cells (Fig. 1D,E). Analysis of cellular composition across samples further indicated variability in the proportions of epithelial subtypes among LUAD specimens, reflecting inter-sample heterogeneity (Fig. 1F). Quantification of cell numbers showed that malignant epithelial cells represented the largest fraction of the epithelial compartment, followed by tumor-associated epithelial states such as tS1 and tS2 (Fig. 1G). Consistently, the average expression heatmap revealed distinct transcriptional signatures among epithelial subpopulations, supporting the reliability of the cell annotation and highlighting the diverse epithelial states present in LUAD tissues (Fig. 1H). Because epithelial cells were obtained from multiple samples, we further quantified batch-correction performance beyond visual UMAP inspection. In GSE131907, Harmony improved cross-sample mixing, with median iLISI increasing from 1.00 to 3.95 and batch ASW increasing from 0.75 to 0.89, while cell-type ASW remained relatively stable (0.49 before correction versus 0.47 after correction). A similar but more modest improvement was observed in GSE123902, with median iLISI increasing from 1.00 to 1.5939 and batch ASW increasing from 0.85 to 0.86, whereas cell-type ASW changed only slightly from 0.59 to 0.56. Together, these findings support that Harmony effectively reduced sample-level batch effects while preserving the major epithelial-state structure (**Supplementary Figs. 2,3**; **Supplementary Table 1**).

Pseudotime Trajectory Reveals the Progression From Differentiated Epithelial States to Malignant States in LUAD

Using epithelial cells from the GSE131907 dataset, pseudotime trajectory analysis was performed to investigate epithelial state progression during LUAD development. Re-clustering of epithelial cells identified several major populations, including AT1, AT2, Club, Ciliated, tS1, tS2, and Malignant cells, which formed a gradual spatial continuum in the UMAP embedding (Fig. 2A). Trajectory inference indicated that epithelial cell states progressed along a continuous developmental path rather than through discrete transitions (Fig. 2B,C). Cells with canonical differentiation features, including AT1, AT2, Club, and Ciliated cells, were predominantly enriched at the early stages of pseudotime, whereas tS1 and tS2 cells occupied intermediate positions along the trajectory (Fig. 2D). In contrast, malignant epithelial cells were primarily located at the late pseudotime stage, suggesting a terminal abnormal cellular state. Consistent with this pattern, cells derived from normal samples were largely distributed at early pseudotime regions, while tumor-derived cells were enriched toward the middle and late trajectory stages (Fig. 2E). The inferred ordering of epithelial states suggested a progression pattern from differentiated epithelial cells through transitional tumor-associated states and ultimately to malignant cells (Fig. 2F). Moreover, stemness and differentiation scores displayed opposite trends along pseudotime, with stemness progressively increasing while differentiation-associated features decreased during the transition toward malignant states (Fig. 2G). Density distribution analysis further confirmed stage-specific enrichment of epithelial subtypes along the trajectory (Fig. 2H). Collectively, these results support a continuous evolutionary framework from differentiated epithelial states to transitional tumor-associated populations and finally malignant epithelial cells in LUAD.

Transcriptional Program Alterations During Epithelial Malignant Progression in LUAD

To characterize transcriptional reprogramming during LUAD epithelial progression, we first visualized the epithelial cell states included in the downstream malignant-program analysis in the GSE131907 dataset (Fig. 3A). We then compared malignant epithelial cells with relatively normal-like epithelial populations. Differential expression analysis revealed extensive transcriptional alterations between the two groups, with numerous genes significantly upregulated in malignant epithelial cells (Fig. 3B). Pathway activity analysis based on Hallmark gene sets further showed that several oncogenic programs, including MYC targets, PI3K/AKT/mTOR signaling, hypoxia response, and glycolysis, were markedly activated in tumor-associated epithelial states, whereas pathways associated with epithelial homeostasis were relatively reduced (Fig. 3C). Program-level scoring analysis demonstrated substantial heterogeneity among epithelial subtypes with respect to proliferation, EMT, stemness, and cellular stress responses, with overall differences shown in Fig. 3D. Fig. 3E showed distinct state-dependent patterns for malignant and CNV-like scores. Malignant scores were highest in tS2 and malignant epithelial cells, indicating enrichment of malignant transcriptional programs in these populations. In contrast, CNV-like scores did not peak in the malignant group; instead, the highest CNV-like signal was observed in ciliated cells, while malignant and transitional epithelial states showed intermediate levels. Program scores also differed significantly among tS1, tS2, tS3, and malignant epithelial states, and the corresponding post hoc group differences are indicated by letter annotations in Fig. 3F. Malignant-core scores were highest in tS2, whereas hypoxia scores were highest in malignant cells; EMT and stemness patterns also varied across epithelial states. Furthermore, pseudotime analysis revealed that malignant core programs and proliferation-associated signals gradually increased as epithelial cells progressed toward later stages of the trajectory (Fig. 3G). Pathway enrichment analysis of differentially expressed genes further confirmed that glycolysis, hypoxia, MYC signaling, and PI3K/AKT/mTOR

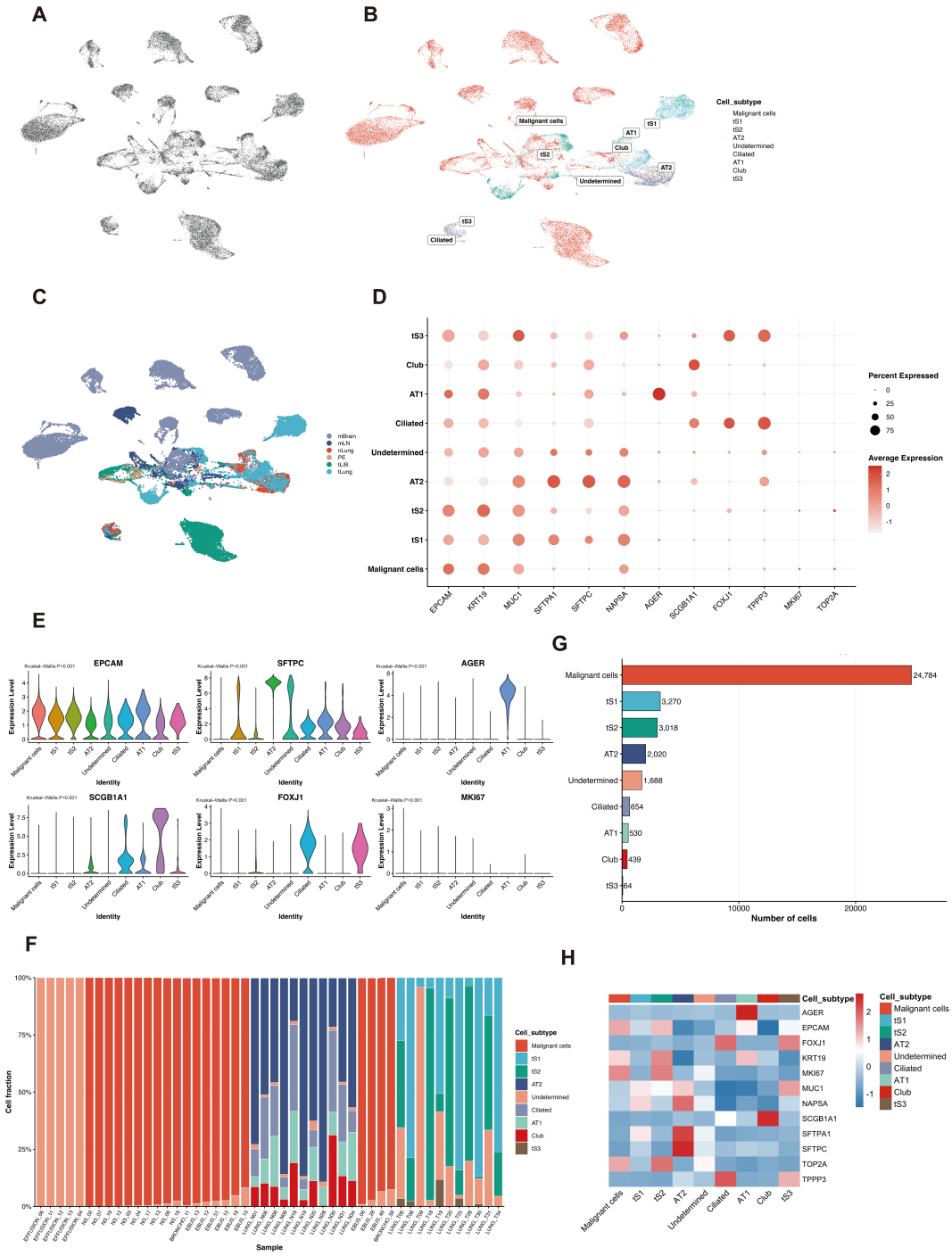

Fig. 1. Epithelial cell atlas and subtype characteristics in LUAD. (A) UMAP projection of integrated epithelial-related cells showing multiple transcriptionally distinct clusters. (B) Cell subtype annotation based on canonical markers identified malignant epithelial cells, AT1, AT2, ciliated, club, tumor-associated epithelial states (tS1–tS3), and undetermined cells. (C) Cells colored by sample origin demonstrate broad distribution across clusters, indicating minimal batch effects after integration. (D) DotPlot showing representative marker gene expression patterns across epithelial subtypes. (E) Violin plots confirming subtype-specific expression of selected markers. (F) Proportional distribution of epithelial subtypes across LUAD samples. (G) Cell number statistics for each epithelial subtype. (H) Heatmap of average gene expression highlighting distinct transcriptional signatures among epithelial populations. LUAD, lung adenocarcinoma; UMAP, Uniform Manifold Approximation and Projection; AT1, alveolar epithelial type I; AT2, alveolar epithelial type II.

signaling were prominently enriched in malignant epithelial cells (Fig. 3H). Together, these findings indicate that LUAD epithelial cells undergo extensive transcriptional remodeling during malignant progression, characterized by

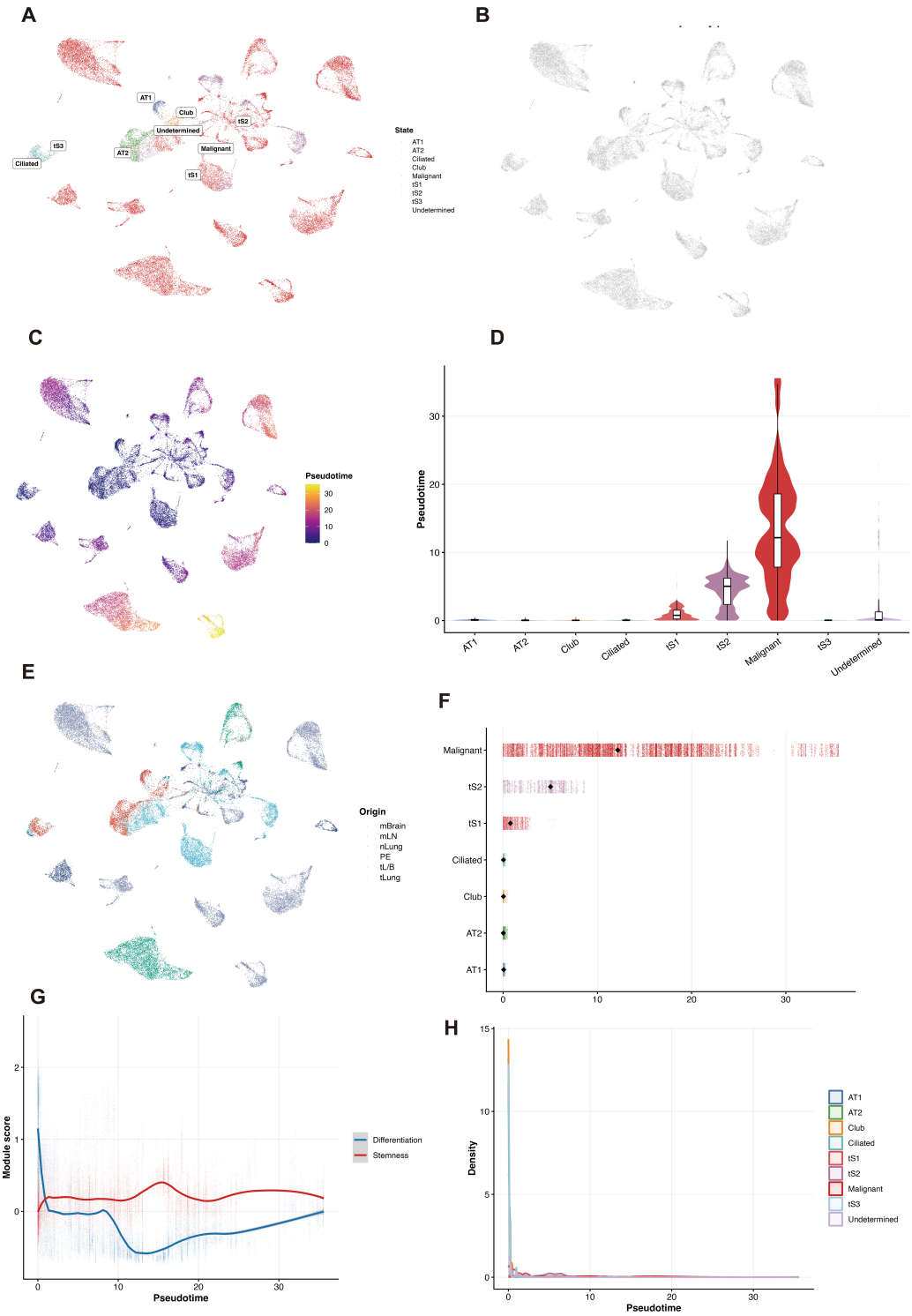

Fig. 2. Functional enrichment and dynamic expression patterns of DHX9/ILF3 complex. (A) UMAP of re-clustered epithelial cells showing AT1, AT2, Club, Ciliated, tS1, tS2, and Malignant populations. (B,C) Pseudotime trajectory inferred from epithelial cells, revealing a continuous developmental path rather than discrete state transitions. (D) Distribution of epithelial subtypes along pseudotime, with AT1, AT2, Club, and Ciliated cells enriched in early stages, transitional states (tS1 and tS2) occupying intermediate positions, and malignant cells located in the late stage. (E) Distribution of cells by sample origin across the trajectory, showing that normal-derived cells mainly appear at early pseudotime whereas tumor-derived cells are enriched in later stages. (F) Trajectory ordering illustrating the inferred progression sequence from differentiated epithelial cells to transitional tumor-associated states and finally malignant cells. (G) Stemness and differentiation scores along pseudotime showing opposite trends during state progression. (H) Density distribution of epithelial subtypes along pseudotime confirming stage-specific enrichment of different epithelial states.

activation of oncogenic signaling pathways, metabolic reprogramming, and acquisition of malignant cellular programs.

Expression Landscape of DHX9/ILF3 Complex Members Across Epithelial States in LUAD

To investigate the expression landscape of the DHX9/ILF3 complex across epithelial states in LUAD, we analyzed the expression of major complex members in epithelial cells from the GSE131907 dataset. Feature plots revealed that DHX9, ILF3, ILF2, and HNRNPU were broadly expressed across epithelial populations but displayed clear differences in signal intensity among cell states (Fig. 4A–D). To characterize the expression pattern of the DHX9/ILF3 complex across epithelial states, we examined key complex members in epithelial cells from the GSE131907 dataset. Heatmap analysis showed that complex members displayed clear state-specific expression patterns rather than uniform distribution across epithelial populations (Fig. 4E). DHX9 and HNRNPU exhibited relatively higher expression in Ciliated, tS1, and tS3 states, whereas ILF2 showed strong enrichment in tS2, tS3, and Malignant cells. ILF3 displayed relatively elevated expression in tS1 and tS3, while SFPQ was particularly increased in Ciliated and tS1 populations, and NONO showed higher expression in tS2 cells. In contrast, several complex components showed comparatively lower expression in AT1, AT2, and Club cells, suggesting reduced complex activity in more differentiated epithelial states. Violin plot analysis further confirmed the heterogeneous expression of individual complex members across epithelial states, with state-dependent distributional differences shown in Fig. 4F. Likewise, DHX9/ILF3 complex scores varied across epithelial states (Fig. 4G), with relatively higher score distributions generally observed in transitional populations than in normal-like epithelial states. Finally, single-cell correlation analysis revealed a moderate positive association between DHX9 and ILF3 expression (Spearman $\rho = 0.370$, $p = 4.62 \times 10^{-223}$, N = 34,779), suggesting coordinated expression of these two core complex members across epithelial cells (Fig. 4H). Collectively, these results indicate that the DHX9/ILF3 complex exhibits clear state-dependent activity patterns and tends to be more active in transitional epithelial states associated with tumor progression.

Trajectory Dynamics of the DHX9/ILF3 Complex During Epithelial State Progression

To explore the dynamic behavior of the DHX9/ILF3 complex during epithelial state transition, we analyzed expression changes along the pseudotime trajectory in the GSE131907 dataset (Fig. 5). The expression trends of major complex members showed distinct but coordinated dynamics along pseudotime (Fig. 5A–D). DHX9 and ILF3 exhibited relatively stable expression during early and intermediate stages, followed by a mild decline toward the late stage, whereas ILF2 displayed a clear mid-trajectory peak and HNRNPU reached its highest level in the intermediate stage before gradually decreasing. Consistently, the DHX9/ILF3 complex module score showed a non-linear pattern along pseudotime, characterized by an initial increase, subsequent decline, and partial recovery at later stages (Fig. 5E). State-stratified trajectory analysis indicated that transitional epithelial states, particularly tS2, displayed relatively higher complex activity in intermediate pseudotime regions (Fig. 5F). Rolling correlation analysis further showed that DHX9 and ILF3 maintained a stable positive correlation throughout the trajectory (Fig. 5G). In addition, expression of several complex members, including ILF2 and HNRNPU, varied across normal-like, transitional, and malignant epithelial stages (Fig. 5H). Together, these results suggest that the DHX9/ILF3 complex undergoes dynamic regulation during epithelial state progression and is preferentially active in transitional tumor-associated states.

Transcriptional Programs Associated With the DHX9/ILF3 Complex

To investigate the transcriptional programs associated with the DHX9/ILF3 complex, we analyzed co-expression relationships and downstream gene signatures in epithelial cells from the GSE131907 dataset (Fig. 6). Correlation analysis showed that complex members (DHX9, ILF3, ILF2, HNRNPU, SFPQ, and NONO) exhibited weak-to-moderate positive correlations with each other, with pairwise correlation coefficients ranging approximately from $\rho = 0.11$ to 0.27 (Fig. 6A). Co-expression network analysis further revealed that these RNA-binding proteins formed a tightly connected module and were linked to additional genes such as STMN1 and SOX4 (Fig. 6B), which ranked among the most highly connected nodes in the network based on weighted connectivity analysis (Fig. 6C). Differential expression analysis comparing DHX9-high and DHX9-low cells identified multiple proliferation-associated genes, including UBE2S, TOP2A, and MYBL2, that were significantly upregulated in the DHX9-high group (Fig. 6D). Similarly, ILF3-high cells showed increased expression of genes such as STAT1, NACC1, and TOP2A, indicating partially overlapping transcriptional programs between DHX9 and ILF3 (Fig. 6E). GO biological process enrichment analysis of DHX9/ILF3-associated genes revealed significant enrichment in pathways related to DNA replication, nuclear transport, nucleocytoplasmic transport, and Wnt signaling regulation (Fig. 6F). Consistent with these findings, a shared downstream gene program activated in both DHX9-high and ILF3-high cells included UBE2S, TOP2A, STAT1, MYBL2, and CDKN2A (Fig. 6G). Finally, complex members and the overall DHX9/ILF3 complex score showed weak-to-moderate positive correlations with proliferation, EMT, and stemness programs ($\rho \approx 0.10$–0.27), indicating modest but directionally consistent asso-

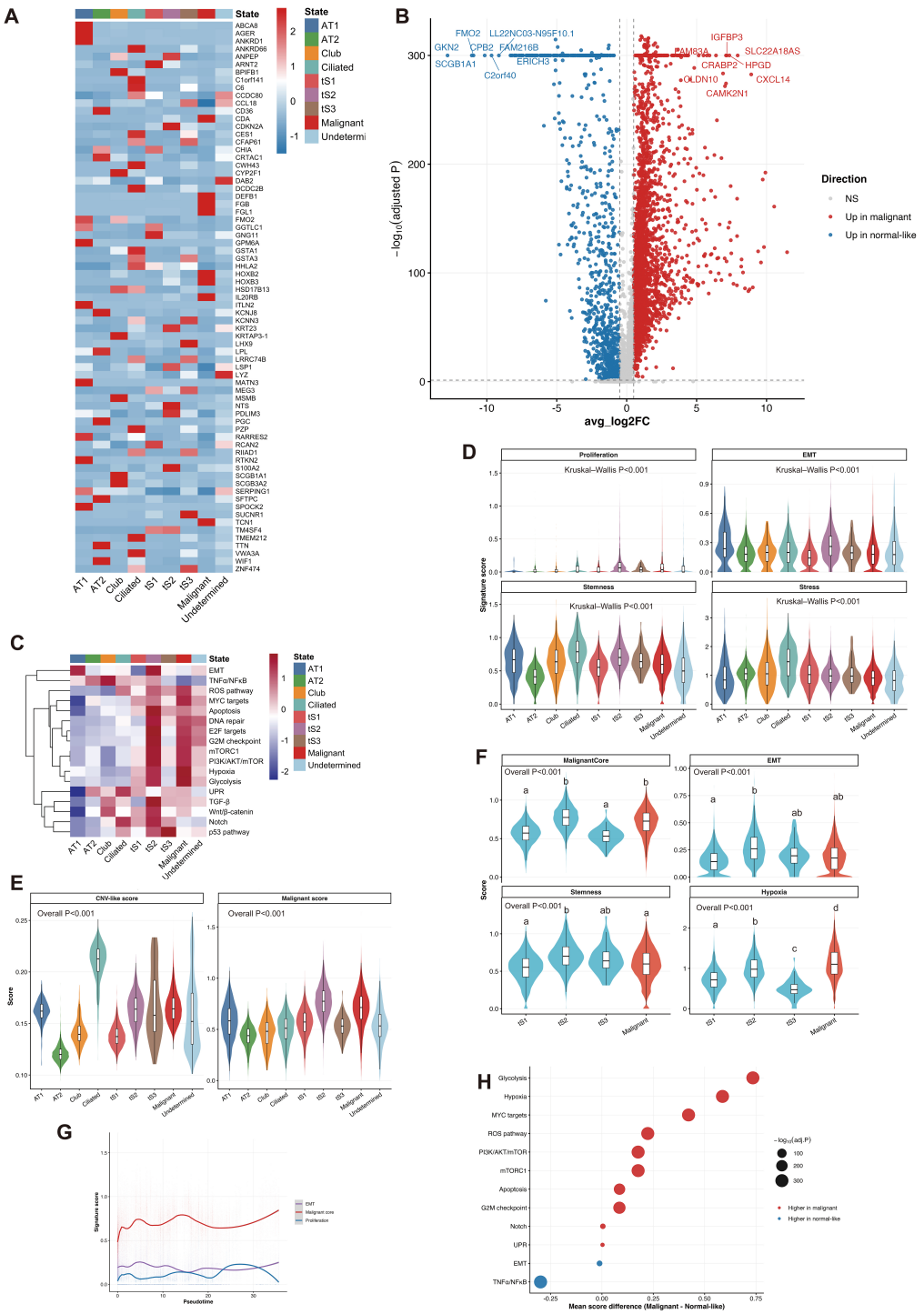

Fig. 3. Temporal dynamics and cell cycle regulation of DHX9/ILF3 complex. (A) UMAP visualization of epithelial cell states used for transcriptional program analysis. (B) Volcano plot showing differentially expressed genes between malignant epithelial cells and relatively normal-like epithelial cells. (C) Heatmap of Hallmark pathway activities across epithelial states. (D) Program score comparison showing heterogeneity in proliferation, EMT, stemness, and stress-response signatures. (E) Comparison of malignant score and CNV-like score across epithelial states. (F) Comparison of malignant-associated transcriptional program scores across tS1, tS2, tS3, and malignant epithelial states; overall *p* values were calculated using the Kruskal–Wallis test, followed by Dunn's multiple-comparison test for pairwise post hoc analysis. Different letters indicate statistically significant differences between groups after multiple-comparison correction. (G) Pseudotime analysis showing gradual activation of malignant core programs along epithelial progression. (H) Bubble plot of enriched pathways in malignant epithelial cells. EMT, epithelial–mesenchymal transition; CNV, copy number variation.

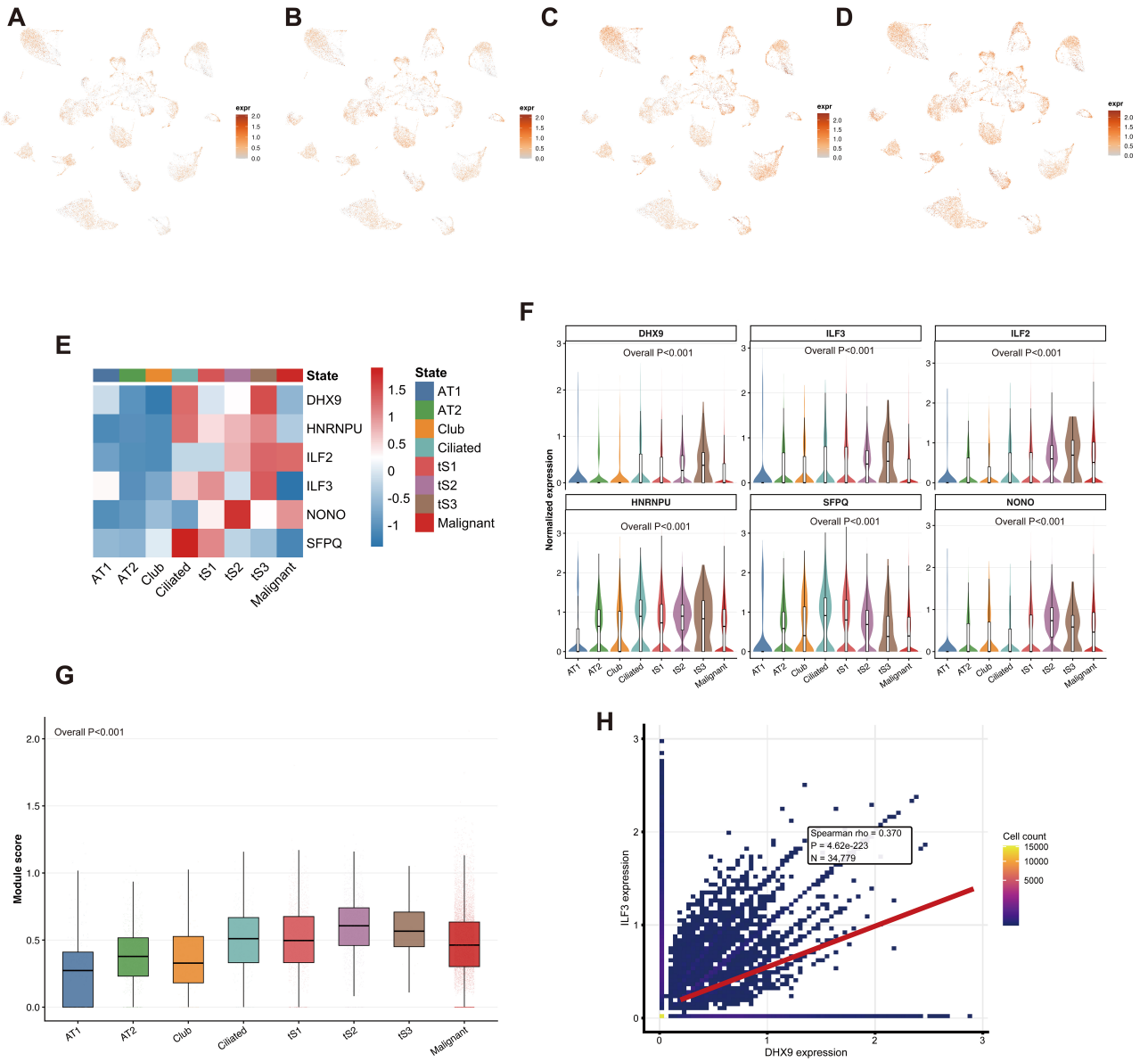

Fig. 4. Expression landscape of DHX9/ILF3 complex members across epithelial states in LUAD. (A–D) Feature plots showing the expression distribution of DHX9, ILF3, ILF2, and HNRNPU across epithelial cell states. (E) Heatmap showing the average expression of major complex members (DHX9, HNRNPU, ILF2, ILF3, NONO, SFPQ) across epithelial states. (F) Violin plots illustrating the expression distribution of representative complex members across epithelial subtypes. (G) Module score analysis showing the overall activity of the DHX9/ILF3 complex across epithelial states. (H) Single-cell correlation analysis between DHX9 and ILF3 expression.

ciations with malignant transcriptional states (Fig. 6H). Together, these results suggest that the DHX9/ILF3 complex is linked to transcriptional programs, promoting tumor cell proliferation and malignant phenotypes in LUAD epithelial cells.

DHX9/ILF3 Complex–High Cells Are Associated With Malignant Epithelial States

To investigate the relationship between DHX9/ILF3 complex activity and malignant epithelial states, we examined the distribution and functional associations of complex-high cells in the dataset (Fig. 7). UMAP visualization showed that DHX9-high (Fig. 7A) and ILF3-high cells (Fig. 7B) were enriched in multiple epithelial clusters rather than uniformly distributed, suggesting state-specific expression patterns. Comparison of cell-state composition revealed that the complex-high group contained a higher fraction of malignant cells and transitional states such as tS2, whereas normal epithelial states such as AT2 were relatively reduced (Fig. 7C). Consistent with this result, density analysis along the pseudotime trajectory indicated that complex-high cells were modestly shifted toward intermediate-to-late trajectory regions, whereas complex-low cells were

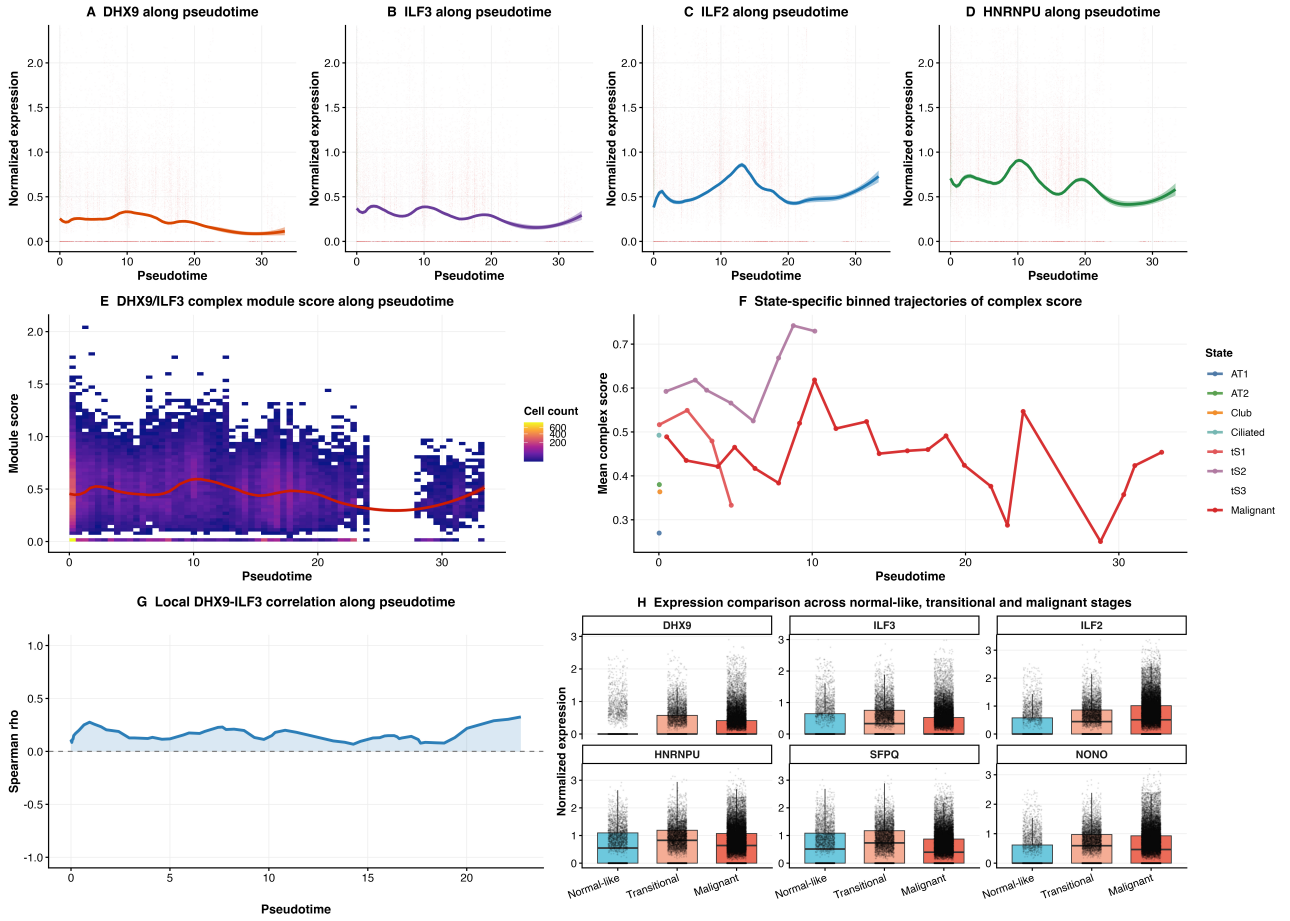

Fig. 5. Trajectory dynamics of the DHX9/ILF3 complex during epithelial state progression. (A–D) Expression trends of major complex members (DHX9, ILF3, ILF2, and HNRNPU) along pseudotime. (E) DHX9/ILF3 complex module score across pseudotime. (F) State-specific binned trajectories of complex module score across epithelial states. (G) Local rolling correlation between DHX9 and ILF3 expression along pseudotime. (H) Expression comparison of complex members across normal-like, transitional, and malignant epithelial stages.

more prevalent in earlier stages (Fig. 7D). Functional program analysis further showed that complex-high cells exhibited higher pathway scores across several tumor-related signatures, including MYC targets, stemness, malignant core programs, proliferation, glycolysis, PI3K-mTOR signaling, EMT, hypoxia, and DNA repair (Fig. 7E). Correlation analysis confirmed that DHX9/ILF3 complex score was positively associated with malignant score (Spearman $\rho = 0.281$, $p < 0.001$; Fig. 7F). Similarly, DHX9/ILF3 complex score showed positive correlations with both stemness ($\rho = 0.265$, $p < 0.001$) and proliferation ($\rho = 0.310$, $p < 0.001$) programs (Fig. 7G), suggesting that higher complex scores are associated with enhanced tumor cell growth potential. Finally, DHX9/ILF3 complex scores showed different overall distributions between tumor-derived and normal lung samples (Fig. 7H). Together, these findings suggest that higher DHX9/ILF3 complex scores are associated with malignant epithelial states and enhanced proliferative and stem-like transcriptional programs in LUAD. To assess whether these findings depended on the choice of dichotomization threshold, we performed sensitivity analyses using alternative cutoff strategies, including median split and upper-versus-lower quartile comparisons. Across these analyses, the overall trends in cell-state composition, pseudotime distribution, and malignant program enrichment remained directionally consistent with the primary tertile-based analysis, supporting the robustness of the complex-high versus complex-low comparisons (**Supplementary Fig. 3**).

Validation of DHX9–ILF3 Complex Activity in an Independent Epithelial Dataset

To validate the findings observed in the discovery dataset, we analyzed epithelial cells from the independent GSE123902 dataset (Fig. 8). UMAP visualization identified several epithelial states, including AT1, AT2, Club, Ciliated, tS1, tS2, and Malignant cells, which formed well-separated clusters in low-dimensional space, indicating substantial epithelial heterogeneity in the validation dataset (Fig. 8A). Quantification of cell-state composition

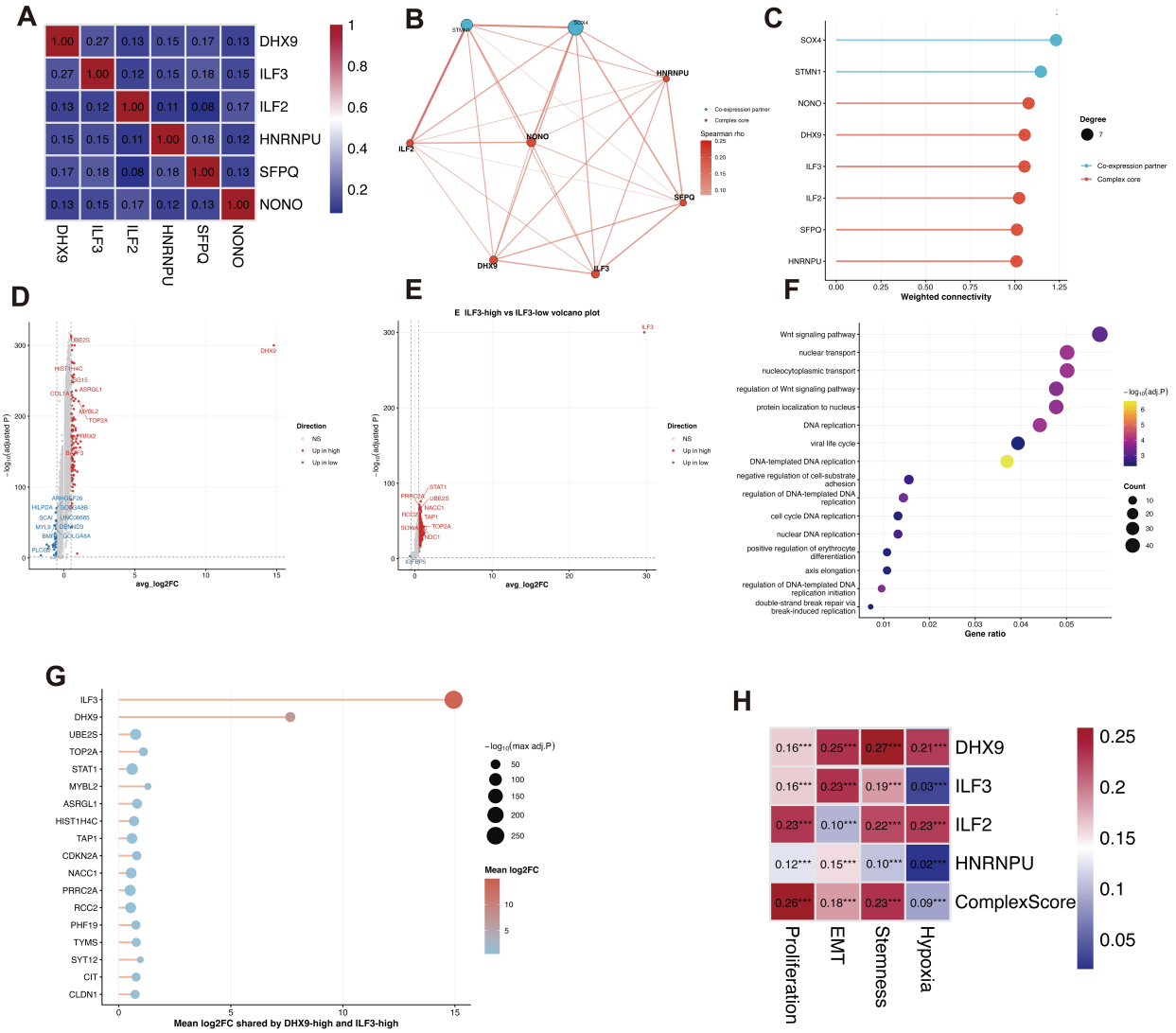

Fig. 6. Transcriptional programs associated with the DHX9/ILF3 complex. (A) Correlation heatmap showing pairwise expression correlations among complex members (DHX9, ILF3, ILF2, HNRNPU, SFPQ, NONO). (B) Co-expression network surrounding the DHX9/ILF3 complex. (C) Hub gene ranking based on weighted connectivity in the co-expression network. (D) Volcano plot of differentially expressed genes between DHX9-high and DHX9-low cells. (E) Volcano plot of differentially expressed genes between ILF3-high and ILF3-low cells. (F) GO biological process enrichment of DHX9/ILF3-associated genes. (G) Top shared-up gene program downstream of DHX9-high and ILF3-high conditions. (H) Correlation matrix between complex activity and malignant transcriptional signatures. ***$p < 0.001$.

showed that Malignant cells represented the largest fraction (28.7%), followed by AT2 (23.5%) and Club cells (22.1%), whereas transitional states such as tS1 (3.3%) and tS2 (12.1%) were less abundant (Fig. 8B). Consistent with the discovery dataset, feature plots further showed that DHX9 and ILF3 remained broadly expressed across epithelial populations in the validation cohort, with relatively stronger signals in several transitional and malignant states (Fig. 8C,D). Analysis of the DHX9/ILF3 complex score revealed that it varied across epithelial states, with relatively higher values observed in several transitional and malignant populations compared with normal-like epithelial cells (Fig. 8E). Pseudotime trajectory inference further showed a continuous state transition pattern across epithelial cells in the validation dataset (Fig. 8F1). Along this trajectory, the DHX9/ILF3 complex score displayed dynamic changes, initially decreasing in early pseudotime and subsequently increasing during intermediate stages before stabilizing at late stages (Fig. 8F2), consistent with the dynamic pattern observed in the discovery dataset. Correlation analysis confirmed that DHX9 and ILF3 expression remained positively correlated in validation epithelial cells (Spearman $\rho = 0.345$, $p < 0.001$) (Fig. 8G1). Rolling correlation analysis further showed that the strength of DHX9–ILF3 asso-

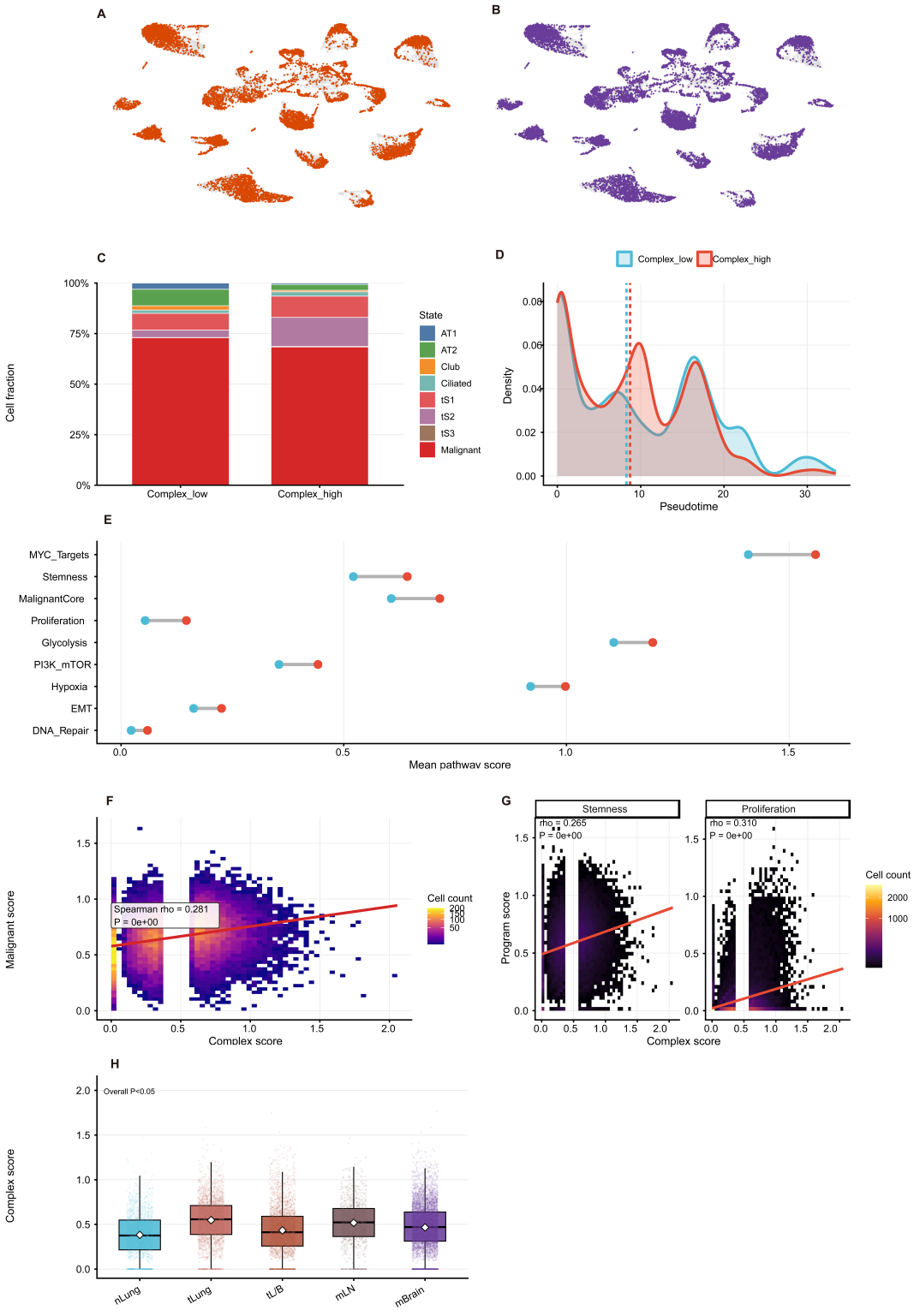

Fig. 7. DHX9/ILF3 complex–high cells are associated with malignant epithelial states. (A) UMAP distribution of DHX9-high cells. (B) UMAP distribution of ILF3-high cells. (C) Cell-state composition comparison between complex-high and complex-low groups. (D) Density distribution of pseudotime between complex-high and complex-low cells. (E) Comparison of tumor-related pathway scores between complex-high and complex-low cells. (F) Correlation between complex score and malignant score. (G) Correlation between complex score and stemness/proliferation programs. (H) Comparison of complex score across different sample origins.

ciation varied along pseudotime, with stronger correlations observed during intermediate trajectory stages and weaker correlations at later stages (Fig. 8G2). Finally, comparison between the discovery and validation datasets demonstrated highly consistent patterns of DHX9 expression, ILF3 expression, and DHX9/ILF3 complex scores across epithelial states, indicating strong reproducibility of the DHX9–ILF3 complex–associated transcriptional features (Fig. 8H). Together, these results confirm that the DHX9–ILF3 complex exhibits conserved activity patterns across independent datasets and is associated with epithelial state transitions and malignant programs in LUAD.

Experimental Validation of DHX9/ILF3 Complex Gene Expression by qRT-PCR

qRT-PCR confirmed the scRNA-seq findings. In both A549 and H1975 cells, DHX9, ILF3, ILF2, and HNRNPU were significantly upregulated relative to NHBE cells. The expression patterns of these genes were consistent between the two cell lines, showing a strong positive correlation ($r = 0.92$, $p < 0.01$). In addition, the ILF3/DHX9 ratio remained relatively stable across both cell lines, suggesting coordinated regulation among complex members. Overall, expression levels of all four genes were significantly higher in lung adenocarcinoma cells compared with normal epithelial cells ($p < 0.001$) (Fig. 9).

Discussion

Using scRNA-seq, we tracked changes in the DHX9/ILF3 complex as lung adenocarcinoma cells transition from stem-like to differentiated states. Based on single-cell RNA sequencing analysis of LUAD epithelial cells from the GSE131907 dataset and external validation using the GSE123902 dataset, our results indicate that the DHX9/ILF3 RNA-binding complex shows coordinated yet dynamic regulation across epithelial differentiation states. The data indicate that these RNA-binding proteins function as a coordinated unit whose activity is linked to differentiation stage, cell-cycle phase, and pseudotime progression.

A key finding is the staggered behavior of DHX9 and ILF3. DHX9 increases early along pseudotime and then declines, whereas ILF3 rises more slowly and remains high in intermediate regions. This dynamic pattern suggests that individual components of the DHX9/ILF3 complex may participate in different stages of epithelial state transitions rather than acting synchronously. This pattern implies that DHX9 predominantly functions during initial stages, where remodeling of RNA or chromatin is likely required, whereas ILF3 remained relatively elevated in intermediate states, a pattern consistent with a possible role in maintaining these transcriptional programs. The non-synchronous timing indicates that individual components operate at different phases rather than undergoing uniform, simultaneous shifts.

Furthermore, strong pairwise associations among DHX9, ILF3, ILF2, and HNRNPU, together with a consistently maintained ILF3/DHX9 ratio of roughly 1.4–1.5, imply that the expression of these genes is under coordinated regulation and subject to quantitative balancing [16]. The maintenance of such proportions mirrors observations made for other multi-protein assemblies in which stable stoichiometry is important for functional output [16]. Our single-cell correlation analysis further supports this coordinated regulation, showing moderate positive correlations among these RNA-binding proteins across epithelial cells. Notably, higher DHX9 levels in G2/M-phase cells, while remaining aligned with the expression of other complex members, suggest that its regulation is influenced by both differentiation status and cell-cycle progression [17–19]. The appearance of two distinct DHX9 peaks versus a single ILF3 peak further implies a temporal rhythm in which different subsets of the complex become predominant at specific time intervals.

Network-based analyses indicated that DHX9 occupies a central position within a group of interacting RNA-binding proteins, including ILF3, SFPQ, NONO, MATR3, FUS, and EWSR1 [20]. Because of this arrangement, altered DHX9 expression may affect several RNA-regulatory pathways simultaneously. In parallel, the broad expression distribution observed for DHX9 and ILF3 across different cell subsets underscores marked heterogeneity within the tumor [21,22], a feature often associated with differences in cellular behavior and drug sensitivity [23]. Our pseudotime trajectory analysis revealed that epithelial cell states form a gradual continuum from differentiated epithelial populations through transitional tumor-associated states and finally to malignant cells, supporting the concept that differentiation advances through a series of transitional steps rather than sharp transitions between fixed identities [24,25].

An additional point that warrants cautious interpretation is the magnitude of the associations between the DHX9/ILF3 complex score and malignant transcriptional programs. Although the observed correlations with proliferation, stemness, EMT, and related signatures were statistically significant, their effect sizes were modest ($\rho \approx 0.10$–0.27), indicating that the complex explains only a limited proportion of transcriptional variation at the single-cell level. These findings therefore should not be interpreted as evidence of strong regulatory dominance. Rather, they suggest that the DHX9/ILF3 axis may represent one contributory component within a broader and multifactorial transcriptional network that shapes LUAD epithelial-state progression. Given the intrinsic biological heterogeneity and technical noise of scRNA-seq data, modest but reproducible associations may still be informative; however, they require further functional validation before stronger mechanistic conclusions can be drawn.

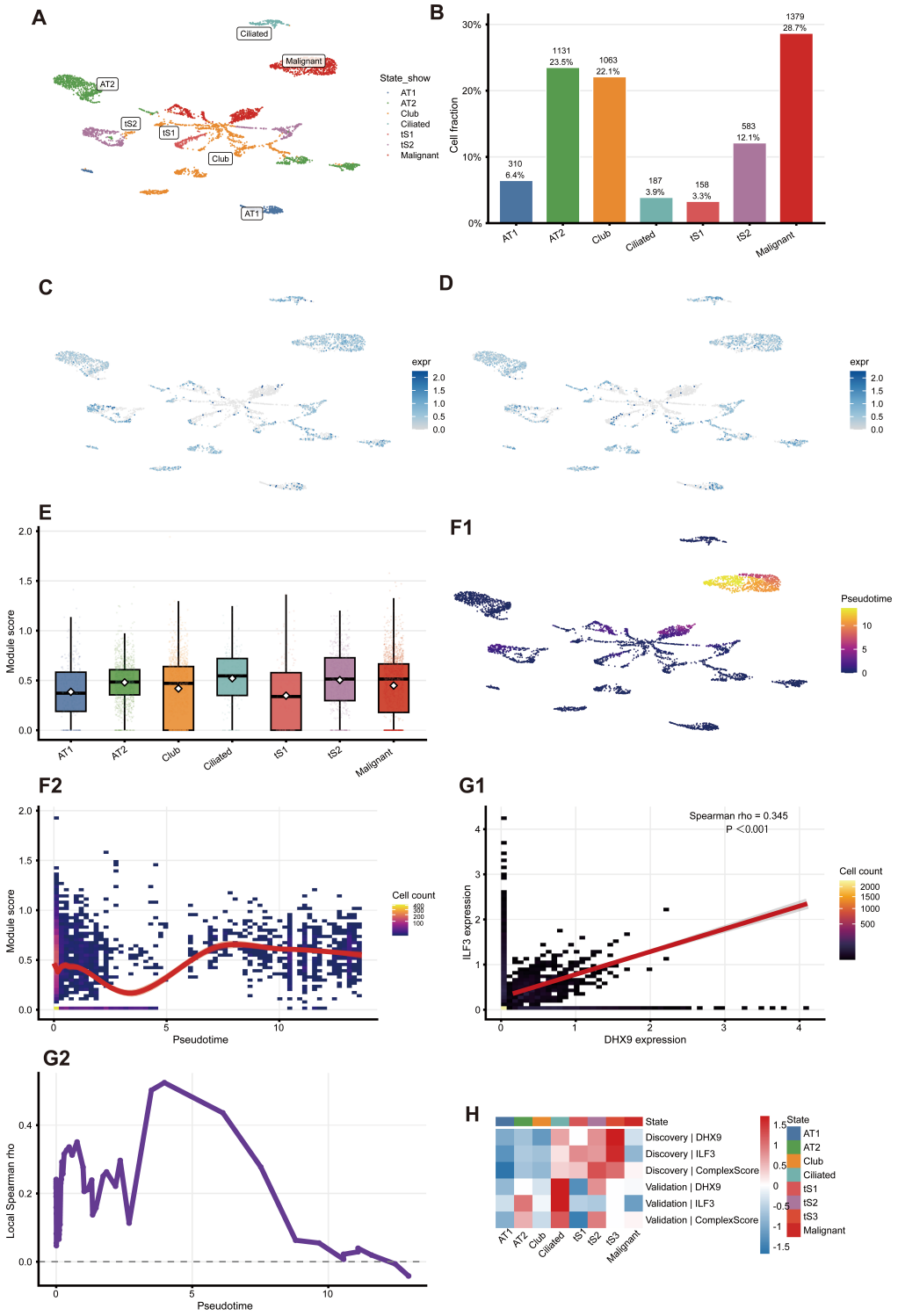

Fig. 8. Validation of DHX9–ILF3 complex activity in an independent epithelial dataset. (A) UMAP visualization of epithelial cells in the validation dataset showing AT1, AT2, Club, Ciliated, tS1, tS2, and Malignant populations. (B) Proportional distribution of epithelial cell states. (C) Feature plot showing DHX9 expression across epithelial states in the validation dataset. (D) Feature plot showing ILF3 expression across epithelial states in the validation dataset. (E) Distribution of DHX9/ILF3 complex module scores across epithelial states. (F1) UMAP colored by pseudotime. (F2) Changes in complex module score along pseudotime. (G1) Correlation between DHX9 and ILF3 expression in validation epithelial cells. (G2) Rolling correlation between DHX9 and ILF3 along pseudotime. (H) Heatmap comparing expression patterns of DHX9, ILF3, and complex module scores between discovery and validation datasets.

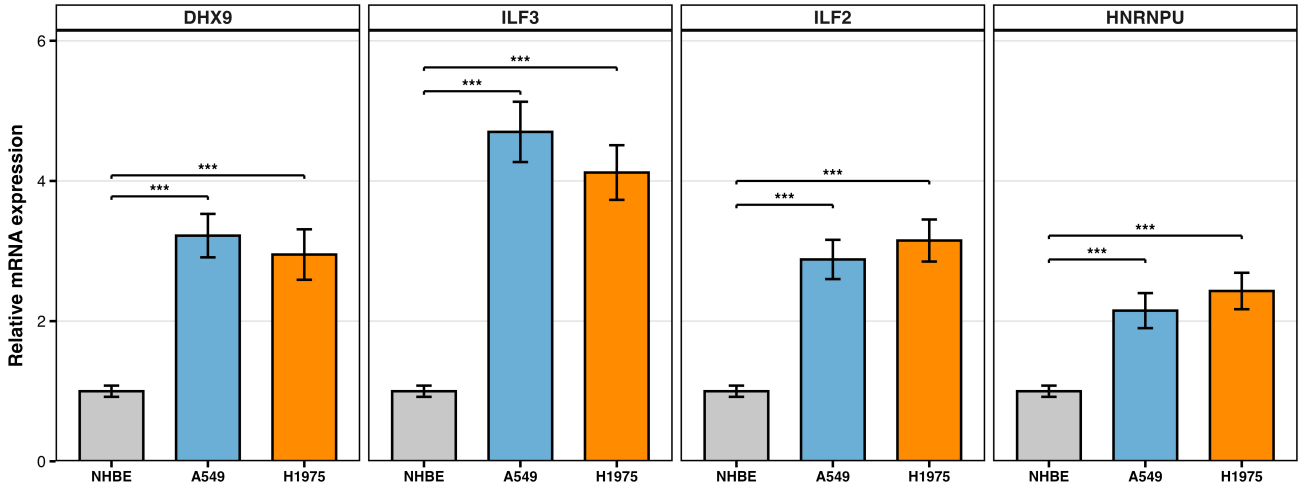

Fig. 9. qRT-PCR validation of DHX9/ILF3 complex gene expression in lung adenocarcinoma cell lines. qRT-PCR validation of DHX9, ILF3, ILF2, and HNRNPU expression in NHBE, A549, and H1975 cells (mean ± SD, n = 3). Statistical significance was assessed using one-way ANOVA with Tukey's post hoc test; ***$p < 0.001$ versus NHBE.

Differences in expression variability among complex members also provide clues. Genes such as DHX9, ILF3, MATR3, and FUS exhibit high variability, characteristic of dynamically regulated factors, whereas more stable components such as HNRNPU may serve as structural scaffolds. Enrichment of functions related to vesicle organization, migration, and adhesion suggests that the complex connects RNA-level regulation to changes in cell morphology and interactions.

Notably, HNRNPU stands out as a condition-responsive component, showing clear shifts across treatments and links to pathways involved in apoptosis and differentiation. Earlier bulk studies reported global upregulation of DHX9 and ILF3 in tumors, but could not resolve the temporal and cell-to-cell variation shown here. Our single-cell analysis indicates that dysregulation in lung adenocarcinoma involves not only elevated expression but also altered timing and coordination with differentiation state and proliferative programs. This finding suggests that therapeutic strategies targeting the DHX9/ILF3 axis might be more effective if directed toward particular stages of vulnerability, rather than through continuous, broad suppression.

Patterns in the expression maps showed both densely populated regions and sparsely occupied areas, suggesting the presence of preferred regulatory states. Cells tended to accumulate within these stable configurations and move between them along restricted trajectories. Such structured transcriptional landscapes suggest that epithelial cells occupy discrete yet interconnected regulatory states during tumor progression. Clarifying which of these states sustain tumor growth or therapeutic resistance could support strategies aimed at shifting cells toward less aggressive or more differentiated outcomes.

Despite these findings, several limitations should be acknowledged. First, this study was based primarily on publicly available scRNA-seq datasets, which may still contain residual dataset-specific biases despite batch correction and external validation. Second, the present analyses are based mainly on transcriptomic association, co-expression, and trajectory inference and therefore do not establish causality or direct mechanistic interaction among DHX9, ILF3, and related complex members. Third, although qRT-PCR validation in LUAD cell lines supported the reproducibility of the observed expression trends, cell-line models cannot fully recapitulate the cellular heterogeneity and microenvironmental complexity of patient tumors. Future studies incorporating additional clinical samples, multicellular systems, and perturbation-based experiments will be required to clarify the biological significance and functional relevance of the DHX9/ILF3-associated patterns identified here.

Conclusions

Single-cell transcriptomic profiling showed that the DHX9/ILF3 complex exhibits coordinated, state-dependent expression patterns during epithelial state transitions in lung adenocarcinoma. Its score was associated with pseudotime progression and malignant transcriptional programs, and this association was reproducible across an independent validation dataset. These findings identify the DHX9/ILF3 complex as a candidate regulatory axis in LUAD that warrants further mechanistic investigation.

Availability of Data and Materials

The single-cell RNA sequencing datasets analyzed in this study are publicly available from the Gene Expression Omnibus (GEO) database. The discovery dataset is available under accession number GSE131907, and the indepen-

dent validation dataset is available under accession number GSE123902. All data used in this study were obtained from public databases, and no newly generated sequencing dataset was deposited. Additional analysis scripts and processed results are available from the corresponding author upon reasonable request.

Author Contributions

WSA and MX designed the research study. MTW, JYA and XYC performed the research. WSA, MTW, JYA and XYC analyzed the data. MX drafted the article. All authors contributed to important editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work to take public responsibility for appropriate portions of the content and agreed to be accountable for all aspects of the work in ensuring that questions related to its accuracy or integrity.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

Not applicable.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.24976/Discov.Med.202638209.150.

References

[1] Riely GJ, Wood DE, Ettinger DS, Aisner DL, Akerley W, Bauman JR, *et al*. Non-Small Cell Lung Cancer, Version 4.2024, NCCN Clinical Practice Guidelines in Oncology. Journal of the National Comprehensive Cancer Network: JNCCN. 2024; 22: 249–274. https://doi.org/10.6004/jnccn.2204.0023.

[2] Hendriks LEL, Remon J, Faivre-Finn C, Garassino MC, Heymach JV, Kerr KM, *et al*. Non-small-cell lung cancer. Nature Reviews Disease Primers. 2024; 10: 71. https://doi.org/10.1038/s41572-024-00551-9.

[3] Kang DH, Lee J, Im S, Chung C. Navigating the Complexity of Resistance in Lung Cancer Therapy: Mechanisms, Organoid Models, and Strategies for Overcoming Treatment Failure. Cancers. 2024; 16: 3996. https://doi.org/10.3390/cancers16233996.

[4] Crucitta S, Cucchiara F, Mathijssen R, Mateo J, Jager A, Joosse A, *et al*. Treatment-driven tumour heterogeneity and drug resistance: Lessons from solid tumours. Cancer Treatment Reviews. 2022; 104: 102340. https://doi.org/10.1016/j.ctrv.2022.102340.

[5] Bohnsack KE, Yi S, Venus S, Jankowsky E, Bohnsack MT. Cellular functions of eukaryotic RNA helicases and their links to human diseases. Nature Reviews Molecular Cell Biology. 2023; 24: 749–769. https://doi.org/10.1038/s41580-023-00628-5.

[6] Murayama T, Nakayama J, Jiang X, Miyata K, Morris AD, Cai KQ, *et al*. Targeting DHX9 Triggers Tumor-Intrinsic Interferon Response and Replication Stress in Small Cell Lung Cancer. Cancer Discovery. 2024; 14: 468–491. https://doi.org/10.1158/2159-8290.cd-23-0486.

[7] Zhang J, Ma Q, Yu Q, Xiao F, Zhang Z, Feng H, *et al*. PSMD3-ILF3 signaling cascade drives lung cancer cell proliferation and migration. Biology Direct. 2023; 18: 33. https://doi.org/10.1186/s13062-023-00389-3.

[8] Chen S, Luo Y, Ruan S, Su G, Huang G. RNA binding protein ILF3 increases CEP55 mRNA stability to enhance malignant potential of breast cancer cells and suppress ferroptosis. Hereditas. 2025; 162: 10. https://doi.org/10.1186/s41065-025-00372-0.

[9] Murphy JJ, Surendranath K, Kanagaraj R. RNA-Binding Proteins and Their Emerging Roles in Cancer: Beyond the Tip of the Iceberg. International Journal of Molecular Sciences. 2023; 24: 9612. https://doi.org/10.3390/ijms24119612.

[10] Chellini L, Pieraccioli M, Sette C, Paronetto MP. The DNA/RNA helicase DHX9 contributes to the transcriptional program of the androgen receptor in prostate cancer. Journal of Experimental & Clinical Cancer Research. 2022; 41: 178. https://doi.org/10.1186/s13046-022-02384-4.

[11] Zhu G, Luo L, He Y, Xiao Y, Cai Z, Tong W, *et al*. AURKB targets DHX9 to promote hepatocellular carcinoma progression via PI3K/AKT/mTOR pathway. Molecular Carcinogenesis. 2024; 63: 1814–1826. https://doi.org/10.1002/mc.23775.

[12] Wang S, Zhu L, Wang Y, Han Y, Wang Q, Yang W, *et al*. ILF3 promotes colorectal cancer cell resistance to ferroptosis by enhancing cysteine uptake and GSH synthesis via stabilizing SLC3A2 mRNA. Cell Death & Disease. 2025; 16: 549. https://doi.org/10.1038/s41419-025-07872-x.

[13] De Zuani M, Xue H, Park JS, Dentro SC, Seferbekova Z, Tessier J, *et al*. Single-cell and spatial transcriptomics analysis of non-small cell lung cancer. Nature Communications. 2024; 15: 4388. https://doi.org/10.1038/s41467-024-48700-8.

[14] Zhang J, Song C, Tian Y, Yang X. Single-Cell RNA Sequencing in Lung Cancer: Revealing Phenotype Shaping of Stromal Cells in the Microenvironment. Frontiers in Immunology. 2022; 12: 802080. https://doi.org/10.3389/fimmu.2021.802080.

[15] Liao P, Huang Q, Zhang J, Su Y, Xiao R, Luo S, *et al*. How single-cell techniques help us look into lung cancer heterogeneity and immunotherapy. Frontiers in Immunology. 2023; 14: 1238454. https://doi.org/10.3389/fimmu.2023.1238454.

[16] Prindle JR, de Cuba OIC, Gahlmann A. Single-molecule tracking to determine the abundances and stoichiometries of freely-diffusing protein complexes in living cells: Past applications and future prospects. The Journal of Chemical Physics. 2023; 159: 071002. https://doi.org/10.1063/5.0155638.

[17] Varzideh F, Gambardella J, Kansakar U, Jankauskas SS, Santulli G. Molecular Mechanisms Underlying Pluripotency and Self-Renewal of Embryonic Stem Cells. International Journal of Molecular Sciences. 2023; 24: 8386. https://doi.org/10.3390/ijms24098386.

[18] Choksi SP, Byrnes LE, Konjikusic MJ, Tsai BWH, Deleon R, Lu Q, *et al*. An alternative cell cycle coordinates multiciliated cell differentiation. Nature. 2024; 630: 214–221. https://doi.org/10.1038/s41586-024-07476-z.

[19] Zhou Y, Nakajima R, Shirasawa M, Fikriyanti M, Zhao L, Iwanaga R, *et al*. Expanding Roles of the E2F-RB-p53 Pathway in Tumor Suppression. Biology. 2023; 12: 1511. https://doi.org/10.3390/biology12121511.

[20] Bertoldo JB, Müller S, Hüttelmaier S. RNA-binding proteins in cancer drug discovery. Drug Discovery Today. 2023; 28: 103580. https://doi.org/10.1016/j.drudis.2023.103580.

[21] Seguin L, Durandy M, Feral CC. Lung Adenocarcinoma Tumor Origin: A Guide for Personalized Medicine. Cancers. 2022; 14: 1759. https://doi.org/10.3390/cancers14071759.

[22] Sinjab A, Rahal Z, Kadara H. Cell-by-Cell: Unlocking Lung Cancer Pathogenesis. Cancers. 2022; 14: 3424. https://doi.org/10.3390/cancers14143424.

[23] Fu YC, Liang SB, Luo M, Wang XP. Intratumoral heterogeneity and drug resistance in cancer. Cancer Cell International. 2025; 25: 103. https://doi.org/10.1186/s12935-025-03734-w.

[24] Pan X, Zhang X. Studying temporal dynamics of single cells: expression, lineage and regulatory networks. Biophysical Reviews. 2023; 16: 57–67. https://doi.org/10.1007/s12551-023-01090-5.

[25] Weiler P, Van den Berge K, Street K, Tiberi S. A Guide to Trajectory Inference and RNA Velocity. Methods in Molecular Biology (Clifton, N.J.). 2023; 2584: 269–292. https://doi.org/10.1007/978-1-0716-2756-3_14.