

Lineage Plasticity and Spatial Immune Redistribution Patterns in TTF-1/p40 Dual-Positive NSCLC: A Multi-Omics Study

Luning Mao

Fudan University Shanghai Cancer Center

Ziling Huang

Fudan University Shanghai Cancer Center

Mengjie Zhang

Fudan University Shanghai Cancer Center

Leyao Li

Fudan University Shanghai Cancer Center

Yan Jin

Fudan University Shanghai Cancer Center

Yue Wang

Fudan University Shanghai Cancer Center

Qianming Bai

Fudan University Shanghai Cancer Center

Yuan Li

lumoxuan2009@163.com

Fudan University Shanghai Cancer Center

Research Article

Keywords: TTF-1, p40, non-small cell lung cancer, tumor immune microenvironment, multiplex immunohistochemistry

Posted Date: October 4th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11032131/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

In non-small cell lung cancer (NSCLC), lineage plasticity drives therapeutic resistance and phenotypic shifts. How tumors transition between lineages—particularly in rare NSCLCs co-expressing LUAD and LUSC markers (TTF-1 and p40)—remains poorly understood, and static histopathology cannot capture the spatial immune changes accompanying genomic evolution.

Methods

We retrospectively evaluated 110 NSCLCs of adenocarcinomatous origin, stratified by TTF-1/p40 expression into TTF-1 only, focal positive, and diffuse positive groups, integrating targeted next-generation sequencing (NGS) with multiplex immunohistochemistry (mIHC) across 750 regions of interest (ROI) from 30 patients.

Results

The cohort had 11 progression events over a median follow-up of 12.4 months. Progression-free survival (PFS) differed nominally across groups (log-rank $P = 0.0473$), and a post-hoc dichotomized comparison favored the diffuse positive group ($P = 0.0184$). Targeted NGS ($n = 45$) revealed TP53 and EGFR as the most frequently altered genes. Spatial profiling across 750 ROI from 30 patients showed descriptive differences in immune cell distributions across phenotypes. At the region level, diffuse-positive regions showed higher CD8 + and lower FOXP3 + Treg densities compared to TTF-1-only regions (nominal $P < 0.001$). Patient-level mean comparisons showed consistent directional trends but wide inter-patient variability. Complete spatial randomness (CSR)-normalized interaction networks revealed distinct spatial architectures: the TTF-1-only network showed relative separation between PanCK+ tumor cells and effector T cells, whereas the diffuse-positive network showed closer PanCK+-to-T-cell proximity and weaker FOXP3+-mediated interactions. Nearest-neighbor distance metrics and cumulative contact curves further characterized these descriptive spatial interaction patterns.

Conclusions

TTF-1/p40 dual-positive NSCLC may represent a distinct immunogenomic state accompanied by spatial immune remodeling. Given the exploratory endpoints, these observations nominate candidate spatial biomarkers that complement—rather than replace—conventional risk stratification and require prospective, multi-center validation.

1. INTRODUCTION

In non-small cell lung cancer (NSCLC), disruption of lineage fidelity is closely associated with malignant transformation and therapeutic resistance¹. Traditionally, lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) are viewed as distinct entities, each driven by a different transcription factor: thyroid transcription factor-1 (TTF-1) for LUAD and p40 (Δ Np63) for LUSC^{2–4}. However, multi-omics studies have challenged this binary view, showing that lung tumor cells can profoundly shift phenotype—a phenomenon termed lineage plasticity⁴. Well-documented in neuroendocrine transdifferentiation after targeted therapy^{5, 6}, lineage plasticity also includes adenocarcinoma-to-squamous transformation following EGFR blockade^{7, 8}; its significance when tumors co-express both LUAD and LUSC markers remains controversial.

NSCLC with dual-positive TTF-1 and p40 staining poses a major diagnostic and clinical challenge. These lineage-uncommitted tumors fit poorly into current practice guidelines⁹, and new strategies are needed¹⁰. Historically classified as adenosquamous variants, technical artifacts, or dedifferentiated carcinomas, these rare tumors tended to be aggressive with poor outcomes^{6, 11, 12}. Yet this conventional risk stratification relies on static histopathology, which does not reflect tumor evolutionary dynamics.

Two gaps remain. First, the genomic changes driving the transition from adenocarcinoma to the diffuse double-positive phenotype—through a focal intermediate stage—are unknown. Second, although solid histology in these tumors is traditionally equated with immunosuppression, the spatial organization of the tumor immune microenvironment (TIME)¹³ remains largely unexplored. How dual-lineage tumors shape host immunity, and why some phenotypes carry a counterintuitive survival advantage, is unclear.

This study examined how somatic genomic evolution shapes TIME spatial organization and immune activation across TTF-1/p40 phenotypes. To reduce inter-lineage genomic heterogeneity, we built a baseline cohort of adenocarcinomatous origin, excluding pure squamous carcinomas and double-negative variants. Integrating targeted next-generation sequencing (NGS) with spatial multiplex immunohistochemistry (mIHC), we tracked the transformation from classic LUAD homeostasis to a diffuse dual-lineage state.

This study makes three contributions. First, we mapped distinct genomic and functional trajectories of the two double-positive phenotypes: the focal phenotype showed pathway features consistent with a transitional state (including VEGFR2-mediated vascular morphogenesis), whereas the diffuse phenotype shifted away from classic oncogene dependence, with nominally enriched *EPHA5*, *HSPBP1*, and *AC018692.2* mutations. Second, diffuse tumors showed spatial immune remodeling not detectable by cell counting alone: global Treg density was unchanged, but Tregs showed descriptive stromal redistribution patterns compatible with weakened local immunosuppression. Third, this redistribution accompanied closer helper-cell localization and greater cytotoxic effector infiltration, varying by *EGFR* and *EPHA5* status.

2. MATERIALS AND METHODS

2.1 Study Design and Cohort Stratification

To evaluate phenotypic transitions within adenocarcinomatous origins while controlling for inter-lineage genomic heterogeneity, this retrospective study excluded classic pure squamous cell carcinomas and unclassified double-negative tumors. As shown in the study workflow (Fig. 1A), 110 formalin-fixed, paraffin-embedded (FFPE) primary NSCLC samples with clinicopathological data were collected. Routine diagnostic IHC used commercial ready-to-use antibodies: p40 (clone MXR010; Maixin Biotechnologies) on a Roche Ventana platform, and dual-color TTF-1/CK5/6 IHC (clones SPT24/MX040) on a Leica BOND platform. The cohort was stratified into three groups: (1) TTF-1 only ($n = 35$); (2) TTF-1/p40 focal positive ($n = 36$, 5%–50% moderate-to-strong p40 staining); and (3) TTF-1/p40 diffuse positive ($n = 39$, > 50% strong staining for both markers). Targeted NGS was performed in 45 patients (TTF-1 only, $n = 13$; focal, $n = 21$; diffuse, $n = 11$; one contaminated white-blood-cell sample excluded), and mIHC in 30 patients (750 regions of interest; 9 of 30 with molecular testing records). Spatial comparisons were pre-specified at the region level (double-positive versus TTF-1 only regions); the focal group was not analyzed as a separate spatial cohort. Clinical stage followed the 9th edition AJCC/UICC TNM classification. The Institutional Review Board of Fudan University Shanghai Cancer Center approved the protocol, and informed consent was waived given the retrospective design and fully de-identified data.

2.2 NGS and Homologous Recombination Repair (HRR) Scoring

Genomic DNA was extracted from FFPE tumor blocks of 45 cases and sequenced at Burning Rock Biotech (Guangzhou, China) using a validated 68-gene panel on the Illumina NextSeq platform covering core NSCLC actionable mutations¹⁴. One case was profiled with a 520-gene panel with matched peripheral blood as germline control; the remaining 44 were sequenced as tumor-only samples, so germline variants cannot be fully excluded. The pipeline also retains variants from flanking non-coding regions (e.g., the lncRNA *AC018692.2*), fusion partners, and genome-wide backbone SNP probes used for baseline ploidy. Variants in prespecified homologous recombination repair (HRR)-related genes were aggregated with a customized R script as an exploratory HRR-pathway mutation-count descriptor. This score is not based on the established LOH + LST + TAI methodology, has not been benchmarked against whole-genome HRD assays, and should not be interpreted as a validated clinical HRD score^{15–17}.

2.3 Differential Mutation Screening and Somatic Interaction Analysis

Differential mutation frequencies were compared across the three lineage-stratified cohorts using Fisher's exact test with Wilson score 95% confidence intervals. Raw p-values were adjusted by the Benjamini–Hochberg method across all 1,568 genes in the genome-wide call set¹⁸. Driver-versus-passenger annotation cross-referenced established driver lists. Because 44 of 45 samples lacked matched germline controls, variant allele fractions and read depths were inspected to flag likely germline

polymorphisms; flagged variants are reported as exploratory. Gene co-occurrence and mutual exclusivity were analyzed with the somaticInteractions function in the maftools R package (version 4.6.0)¹⁹.

2.4 Functional Pathway Enrichment Analysis

Mutated-gene lists specific to each phenotype were mapped against GO, KEGG, and Reactome. Over-representation analysis used the clusterProfiler R package (version 4.6.0)²⁰. Pathways with BH-adjusted $P < 0.05$ were considered enriched and visualized as a functional evolution matrix heatmap (ComplexHeatmap R package²¹).

2.5 mlHC and Stereological Image Acquisition

We selected 30 patients spanning all three phenotypes (9 TTF-1 only, 4 focal positive, and 17 diffuse positive at the patient level). Diagnostic IHC for stratification used validated TTF-1 and p40 antibodies (Supplementary Table S1). For spatial analysis, mlHC was performed with a 6-color panel (CD8, PanCK, CD4, FOXP3, CD25, and DAPI) through the commercial services of Shanghai KR Pharmtech Co., Ltd. Briefly, 4 μm sections were deparaffinized, rehydrated, and subjected to microwave-assisted antigen retrieval in EDTA buffer (pH 9.0), then blocked and sequentially incubated with primary antibodies. Visualization used a polymer HRP-conjugated secondary antibody with tyramide signal amplification (TSA) fluorophores (1:100); DAPI (1:500) served for nuclear counterstaining. All reagents were from a commercial kit (Cat# RF05525; KR Pharmtech)^{22, 23}. Whole-slide scans were acquired at 20 $\times$ on the KR-HT5 system. Tumor–stroma interface fields were identified by systematic pre-screening, and 25 high-power fields (HPFs) were randomly sampled per sample, balancing per-patient sampling depth against acquisition throughput. This yielded 750 optical fields nested within 30 patients (20–30 per patient, median 25; Supplementary Table S7). Based on local marker profiles, these fields were classified at the region level into 414 TTF-1 only regions (21 cases) and 336 diffuse positive regions (20 cases); eleven mixed cases harbored both region types.

2.6 Spatial Point Pattern and Cellular Interaction Analysis

Multispectral images were unmixed and analyzed in inForm 2.4.0 (Akoya Biosciences). Tissue segmentation partitioned tumor parenchyma and stroma by PanCK mapping; cell segmentation used an AI-driven watershed algorithm on DAPI and membrane patterns. Single and combinatorial phenotypes (e.g., CD8⁺, CD4⁺FOXP3⁺, CD4⁺FOXP3⁺CD25⁺) were defined by mean fluorescence intensity (MFI) thresholds. To mitigate geometric bias from divergent cell densities, we implemented a spatial association metric based on Complete Spatial Randomness (CSR)^{24, 25}. For each target-cell phenotype with density λ (cells/mm²), the CSR expected nearest-neighbor distance was calculated analytically as $E(r) = 500/\sqrt{\lambda}$ (μm) under a homogeneous Poisson process. The interaction weight $W = \log_2[E(r)/r_{\text{obs}}]$, where r_{obs} was the observed nearest-neighbor distance. $W > 0$ indicated spatial attraction, whereas $W < 0$ indicated spatial avoidance. Statistical significance was assessed using a Z score based on the nearest-neighbor distribution; $|Z| > 1.96$ ($P < 0.05$) was considered non-random. Isotropic edge correction (Ripley's circumference method) was applied to boundary calculations. Direct cell–cell apposition was

assessed with a fixed 20- μ m radius, and cumulative interaction zones up to 100 μ m were analyzed separately^{26–29}.

2.7 Statistical Analysis

Analyses used R (version 4.6.0) and GraphPad Prism 10. Skewed continuous variables were compared with Mann–Whitney U or Kruskal–Wallis tests. Progression-free survival (PFS) was measured from surgical resection (wedge $n = 38$, segmentectomy $n = 26$, lobectomy $n = 46$) to documented progression or last follow-up; progression was ascertained from institutional records, and median follow-up by the reverse Kaplan–Meier method. PFS was analyzed with Kaplan–Meier curves and log-rank tests. Given the small event count, multivariable Cox models were fitted with Firth's penalized likelihood³⁰ as primary, alongside conventional estimates; proportional hazards were tested with Schoenfeld residuals, and robustness to unmeasured confounding with the E-value³¹. Spatial immune cell densities were compared between TTF-1-only regions ($n = 414$ ROI) and diffuse-positive regions ($n = 336$ ROI) using independent-sample t tests at the region level. Patient-level mean comparisons were performed by averaging ROI densities within each patient (9 TTF-1-only patients, 17 diffuse-positive patients) and comparing group means with independent-sample t tests. Region-level and patient-level comparisons were treated as exploratory with nominal P -values. Nearest-neighbor distance (NND) metrics and cumulative contact curves were computed descriptively. CSR-normalized chord diagrams visualized spatial interaction networks as descriptive summaries. Two-sided $P < 0.05$ was considered significant.

3. RESULTS

3.1 Clinical and Histopathological Features of NSCLC Cohorts by Lineage Expression

H&E staining revealed distinct histological patterns: the TTF-1 only group showed classic LUAD morphology with predominantly lepidic, acinar, or papillary patterns (Fig. 1B, 1C). Histologic grade and solid components were balanced: IASLC grade did not differ ($P = 0.758$), median solid component was 0% in all groups ($P = 0.176$), and micropapillary components were numerically more frequent in TTF-1 only tumors (11 of 31 evaluable cases; Fig. 1D). Smoking status was balanced (ever-smokers: 11/39 diffuse, 14/36 focal, 13/35 TTF-1 only; $P = 0.578$). Clinically confirmed driver alterations comprised *EGFR* mutations in 74.7% (balanced across groups, $P = 0.362$), *TP53* in 24.2% ($P = 0.044$; enriched in the focal group), and *KRAS* in 8.1% ($P = 0.580$) (Fig. 1E). Per-case morphology, marker extents, and grading are provided in Supplementary Table S4.

Progression events occurred in 5 of 35 TTF-1 only, 5 of 36 focal positive, and 1 of 39 diffuse positive patients (11 total); median follow-up was 12.4 months (IQR 3.4–16.4) by the reverse Kaplan–Meier method (106 patients remained in follow-up; 4 were lost). The three-group comparison reached nominal significance (log-rank $P = 0.0473$, Fig. 1F), and a post-hoc dichotomized comparison (diffuse positive versus the other groups) reached $P = 0.0184$ (Fig. 1G); this dichotomization was not pre-specified and is

hypothesis-generating. Median PFS was not reached in any group; 12-/24-month PFS rates were 84.0%/74.7% (TTF-1 only), 83.5%/75.9% (focal), and 96.7%/96.7% (diffuse). Univariate analyses are shown in Fig. 1H. With an events-per-variable ratio of 2.75, Firth's penalized Cox regression³⁰ served as the primary multivariable analysis (Fig. 1I). After adjustment for clinical stage and perineural invasion, the association between the diffuse phenotype and progression was weakened (adjusted HR 0.16, 95% CI 0.03–1.03, $P = 0.053$). Perineural invasion was retained (HR 53.5, 95% CI 5.8–494, $P < 0.001$); stage was not significant. The conventional Cox model yielded a diffuse-group HR of 0.097 (95% CI 0.011–0.85, $P = 0.035$), illustrating sparse-events bias; both estimates are shown in Fig. 1I. The proportional-hazards assumption was met for all covariates (all $P > 0.4$).

3.2 Spatial Remodeling of the Tumor Immune Microenvironment (TIME)

Quantitative mIHC covered 750 ROI from 30 patients (20–30 per patient, median 25; Supplementary Table S7), classified at the region level as TTF-1 only ($n = 414$, 21 cases) or diffuse positive ($n = 336$, 20 cases) (Fig. 2A); eleven patients contributed both region types. Region-level comparisons showed descriptive differences in immune cell densities between phenotypes. Tumor area fraction did not differ ($P = 0.257$, Fig. 2B), arguing against a purely geometric explanation, though residual architectural confounding cannot be excluded. Region-level comparisons showed higher CD8⁺ and lower FOXP3⁺ Treg densities in diffuse regions (nominal $P < 0.001$, Fig. 2C); at the patient level these differences were not significant (e.g., median stromal FOXP3⁺ density 70.0 versus 50.8, $P = 0.506$). As a descriptive region-level analysis, parenchymal density was plotted against stromal density with per-group linear regression (Fig. 2D); the flatter stromal-to-parenchymal slope in diffuse regions is consistent with lower stromal densities rather than stromal Treg accumulation. These region-level patterns are descriptive and do not account for intra-patient clustering.

Asymmetric nearest neighbor distance (NND) metrics evaluated spatial localization (Fig. 2E). At the patient level, none of the 16 NND metrics differed significantly (minimum nominal $P = 0.121$; Supplementary Table S5). Region-level nominal differences—e.g., shorter CD4⁺-to-PanCK⁺ distances in double-positive regions ($P < 0.05$)—are shown descriptively and should be interpreted as exploratory.

At the interaction-field scale, immune cell contact with tumor cells was quantified by counting cells within 100 μm of PanCK⁺ cells (Fig. 2F), complementing the 20- μm synapse-scale analysis. These cumulative contact curves are descriptive region-level summaries: PanCK⁺-to-CD8⁺ contact was higher, and FOXP3⁺ Treg contact lower, in diffuse regions; patient-level significance was not established.

Finally, CSR-normalized chord diagrams visualized global spatial interaction networks (Fig. 2G): positive weights ($W > 0$, pink) indicate attraction and negative weights ($W < 0$, blue) repulsion/avoidance. The TTF-1 only network showed avoidance weights between PanCK⁺ cells and effector T cells, whereas the diffuse network showed stronger PanCK⁺-to-T-cell attraction and weaker FOXP3⁺-mediated interactions. Consistent with the patient-level results, these networks are descriptive visualizations of redistribution patterns rather than confirmed biological attraction.

3.3 Somatic Mutational Landscapes and Evolutionary Trajectories

NGS profiled 45 tumors (TTF-1 only, $n = 13$; focal, $n = 21$; diffuse, $n = 11$; one contaminated white-blood-cell sample excluded) (Fig. 3A). *TP53* (89%) and *EGFR* (87%) were most frequently altered in this unfiltered call set, which retains germline polymorphisms (clinically confirmed frequencies: 24.2% and 74.7%; Section 3.1). Diffuse positive tumors showed nominally higher mutation frequencies in *EPHA5* (5/11, 45.5%, Wilson 95% CI 21.3–72.0%, versus 0/13 TTF-1 only and 1/21 focal; Fisher $P = 0.0020$), *HSPBP1* (7/11, 63.6%, 95% CI 35.4–84.8%, versus 2/13 and 3/21; $P = 0.0035$), and lncRNA *AC018692.2* (5/11, 45.5%, 95% CI 21.3–72.0%, versus 1/13 and 0/21; $P = 0.0020$) (Fig. 3C, 3E). None survived Benjamini–Hochberg correction across 1,568 genes (all FDR ≈ 0.97)¹⁸. Inspection of variant allele fractions and read depths (BAM files unavailable) indicated these signals were driven by germline polymorphisms and low-depth calls: 4 of 5 *EPHA5* mutations were rs56312931 (p.His959Arg; VAF 0.40–0.67, depth 102–248), all *HSPBP1* calls were the germline indel rs3040014 (depth 1–8), and all *AC018692.2* calls were single-read (depth 1–2). None of these loci constitute drivers; they are reported as exploratory observations only.

GO, KEGG, and Reactome enrichment analyses revealed distinct evolutionary states (Fig. 3B): the TTF-1 only group showed strong enrichment of epithelial-related pathways including alveolar surfactant metabolism; the focal group appeared transitional, marked by VEGFR2-mediated vascular morphogenesis; and the diffuse group diverged from this epithelial program, with enrichment for genomic instability, HRR-pathway alterations, and immune activation. Co-occurrence and mutual exclusivity analysis confirmed these patterns (Fig. 3D).

3.4 Integration of Mutational Genotypes and Spatial Immune Features

Integrative analysis revealed associations between genomic alterations and spatial patterns (Fig. 4A); subgroups are designated by spatial phenotype and mutational status (e.g., TTF-1 only/*EPHA5*-WT, diffuse/*EPHA5*-MUT). These region-level subgroup comparisons are exploratory. The TTF-1 only/*EPHA5*-WT subgroup had significantly lower CD8⁺ T cell density than the other three subgroups (Fig. 4C); the TTF-1 only/*EPHA5*-MUT subgroup showed higher infiltration than TTF-1 only/*EPHA5*-WT but remained lower than diffuse/*EPHA5*-WT. Absolute CD8⁺ density was highest in the diffuse/*EPHA5*-WT subgroup, significantly greater than in all other subsets.

Both TTF-1 only subgroups showed significantly higher FOXP3⁺ Treg infiltration than the diffuse/*EPHA5*-WT subgroup (Fig. 4C), and Treg density in diffuse/*EPHA5*-WT was significantly lower than in its *EPHA5*-MUT counterparts. The NND from CD8⁺ T cells to PanCK⁺ tumor cells was significantly greater in TTF-1 only/*EPHA5*-WT than in the other three subgroups combined (Fig. 4E); similarly, the FOXP3⁺-to-CD8⁺ NND at the TTF-1 only/*EPHA5*-WT baseline was significantly higher than in the TTF-1 only/*EPHA5*-MUT and diffuse/*EPHA5*-MUT subgroups (Fig. 4E).

Stratification by *EGFR* status showed that the diffuse/*EGFR*-WT subgroup had higher CD8⁺ infiltration than the TTF-1 only/*EGFR*-MUT and diffuse/*EGFR*-MUT subgroups (Fig. 4B), and the diffuse/*EGFR*-MUT group had higher CD8⁺ densities than the TTF-1 only/*EGFR*-MUT baseline. The CD8⁺-to-PanCK⁺ NND was markedly higher in TTF-1 only/*EGFR*-MUT than in either diffuse subgroup (Fig. 4D), and the FOXP3⁺-to-CD8⁺ NND corroborated this pattern, being significantly higher in TTF-1 only/*EGFR*-MUT than in both diffuse subsets (Fig. 4D).

4. DISCUSSION

Lineage plasticity is a key driver of therapeutic resistance and histological transformation in NSCLC⁴. Here we characterized the clinical, genomic, and spatial immune features of TTF-1/p40 dual-positive NSCLC. The diffuse positive phenotype showed a trend toward longer PFS, reaching nominal significance in the three-group comparison (log-rank $P = 0.0473$) and a post-hoc dichotomized analysis ($P = 0.0184$). This contrasts with the traditional view of solid adenocarcinoma histology as aggressive³²; in our cohort, solid components and IASLC grade were balanced across groups ($P = 0.176$ and $P = 0.758$), arguing against grade as an explanation. With only 11 events (events-per-variable ratio 2.75), Firth-penalized analysis³⁷ weakened the association, and because treatment data were unavailable, we quantified robustness to unmeasured confounding: the E-value of 11.9 means an unmeasured confounder would need risk-ratio associations of at least this magnitude with both the diffuse phenotype and progression, beyond measured covariates, to explain away the point estimate. However, the confidence interval includes unity and the event count is small, so this signal remains fragile³¹. Within these limitations, our data are consistent with a model in which genomic evolution accompanies—and may contribute to—a more immunologically active spatial state.

Pathway analysis suggested that the three phenotypes reflect distinct states along a putative lineage continuum^{33, 34} rather than discrete random variants, although our cross-sectional design cannot establish temporal ordering. The TTF-1 only group maintains a stable differentiated state reliant on alveolar surfactant metabolism and epithelial cadherin pathways^{35, 36}. The focal group showed enrichment of VEGFR2-mediated vascular morphogenesis and keratinization-related gene sets; whether these signatures correspond to morphological squamous differentiation requires histological confirmation^{37–39}. The diffuse group diverged from this epithelial program, with nominally enriched *EPHA5*, *HSPBP1*, and *AC018692.2* mutations—signals that did not survive FDR correction and largely reflected germline polymorphisms or low-depth calls—plus functional enrichment for genomic instability and senescence escape. Ephrin–Eph signaling regulates cell adhesion, boundary formation, and tissue patterning⁴⁰; whether *EPHA5* alteration compromises lineage fidelity in lung epithelium, permitting co-expression of opposing lineage markers, warrants functional testing. Loss of canonical oncogene dependence and increased genomic instability may accelerate lineage drifting and promote immune interaction.

Within the WHO 2021 classification framework¹, each diffuse dual-positive case was evaluated against the main alternative diagnoses. No diffuse case met the $\geq 10\%$ dual-component threshold for adenosquamous carcinoma¹²; none was solid-predominant LUAD with aberrant p40 expression (median solid component 0% in all groups); none showed the two discrete tumor populations of a collision tumor. Per-case morphology, marker extents, IASLC grade, and molecular annotations for all 110 cases are in Supplementary Table S4. Slides were reviewed by two thoracic pathologists. CK5/6 data from the diagnostic workup will be reported separately; until then, clone-dependent staining cannot be excluded².

Spatial analysis suggested that lineage drifting accompanies differences in immune organization rather than cell number alone. Whereas previous studies focused largely on lymphocyte density^{41–43}, our data emphasize spatial redistribution. Region-level comparisons showed numerically lower stromal FOXP3 + density in diffuse regions (50.8 versus 70.0 cells/mm²), but these are exploratory observations from 750 ROIs that do not account for intra-patient clustering. Tumor area fraction did not differ ($P = 0.257$), arguing against a purely geometric artifact but providing no direct evidence of active redistribution. We therefore report the spatial architecture as descriptive; the term “stromal trapping” is replaced by “spatial redistribution patterns” throughout^{41, 44–46}. NND analysis suggested CD4⁺ helper T cells lie closer to PanCK⁺ tumor cells at the region level^{26, 28, 29}. Whether this configuration weakens local immunosuppression remains unresolved; suppressive function was not directly measured.

Integrative analysis associated somatic genotypes with distinct spatial immunophenotypes, with *EPHA5* and *EGFR* status stratifying immune patterns in opposite directions in these exploratory region-level analyses. In TTF-1 only tumors, *EGFR* mutation or *EPHA5*-WT status marked a highly immunosuppressive state—increased CD8⁺-to-tumor NNDs with stromal Treg clustering—consistent with the immune-excluded phenotype of oncogene-driven lung cancers^{47, 48}. This pattern reversed in the diffuse phenotype: diffuse/*EPHA5*-WT had the highest CD8⁺ density and lowest Treg infiltration. One speculative explanation is that *EPHA5*-WT, in a high-plasticity context, may cooperate with neoantigens potentially arising from lineage infidelity to recruit cytotoxic T cells^{33, 49, 50}, a mechanism requiring experimental validation. Conversely, *EPHA5* or *EGFR* mutations in the diffuse phenotype partially weakened this activation, suggesting that the same mutation may influence the immune microenvironment differently depending on lineage context.

This study has limitations. First, it was a single-center retrospective study of an Asian cohort enriched for never-smoker adenocarcinoma (*EGFR* mutation rate 74.7%), limiting extrapolation; multi-center validation is needed. Second, treatment data were unavailable, and with 11 progression events and 12.4 months of median follow-up, the survival estimates remain fragile. Third, the primary spatial endpoints were not significant at the patient level after accounting for intra-patient clustering, so region-level patterns remain descriptive. Fourth, dual TTF-1/p40 positivity was defined by IHC clones without orthogonal validation, CK5/6 cocktail results were not retrievable, and inter-observer kappa for the 5%/50% thresholds was not computed. Fifth, NGS covered 45 of 110 tumors without matched germline controls; most *EPHA5*, *HSPBP1*, and *AC018692.2* calls reflected germline polymorphisms or low-depth

reads, and the HRR-pathway score is not a validated HRD assay. Finally, mIHC segmentation was not independently benchmarked, the panel focused on T cells, and functional validation of *EPHA5* is pending.

5. CONCLUSION

In conclusion, this study provides the first integrated genomic and spatial immune map of TTF-1/p40 dual-positive NSCLC. The diffuse positive phenotype coupled lineage-plasticity–associated genomic shifts with a more immunologically active spatial state and, in exploratory analyses, favorable prognosis. These hypothesis-generating findings link lineage-plasticity markers to immune-activated phenotypes and warrant prospective validation as prognostic tools.

Abbreviations

AJCC

American Joint Committee on Cancer

CI

confidence interval

CSR

complete spatial randomness

DAPI

4',6-diamidino-2-phenylindole

FDR

false discovery rate

FFPE

formalin-fixed,paraffin-embedded

HR

hazard ratio

HRD

homologous recombination deficiency

HRR

homologous recombination repair

IASLC

International Association for the Study of Lung Cancer

IHC

immunohistochemistry

LOH

loss of heterozygosity

LST

large-scale state transition

LUAD

lung adenocarcinoma

LUSC
lung squamous cell carcinoma
MFI
mean fluorescence intensity
mIHC
multiplex immunohistochemistry
NGS
next-generation sequencing
NND
nearest-neighbor distance
NSCLC
non-small cell lung cancer
PanCK
pan-cytokeratin
PFS
progression-free survival
ROI
region of interest
TAI
telomeric allelic imbalance
TIME
tumor immune microenvironment
TSA
tyramide signal amplification
TTF-1
thyroid transcription factor-1
UICC
Union for International Cancer Control
VAF
variant allele fraction
WHO
World Health Organization.

Declarations

Ethics approval and consent to participate:

This study was conducted in accordance with the Declaration of Helsinki. The protocol was approved by the Ethics Committee of Fudan University Shanghai Cancer Center (approval code 050432-4-2507; date: 7 July 2025). The requirement for informed consent was waived by the Ethics Committee owing to the retrospective study design and the use of fully de-identified data.

Consent for publication:

Not applicable.

Competing interests:

The authors declare that they have no competing interests.

Funding:

This research received no specific grant from any funding agencies in the public, commercial, or not-for-profit sectors.

Author Contribution

Authors' contributions: Conceptualization: Yuan Li, Luning Mao, Ziling Huang; Methodology: Luning Mao, Ziling Huang; Software & Formal Analysis: Luning Mao, Mengjie Zhang; Validation: Luning Mao, Ziling Huang; Investigation: Luning Mao, Ziling Huang, Yan Jin; Resources: Ziling Huang, Leyao Li, Yue Wang, Qianming Bai; Data Curation: Luning Mao, Mengjie Zhang, Leyao Li; Writing – Original Draft Preparation: Luning Mao; Writing – Review & Editing: Yuan Li, Luning Mao; Visualization: Luning Mao; Supervision: Yuan Li. All authors have read and agreed to the published version of the manuscript.

Acknowledgement

The authors express their deep gratitude to the patients and their families. We thank Shanghai KR Pharmtech Co., Ltd. for their technical assistance with mIHC staining and scanning.

Data Availability

The NGS data, mIHC raw images, and de-identified clinical datasets analyzed during the current study are not publicly available at this stage due to institutional regulations and patient privacy concerns, but are available from the corresponding author upon reasonable request. Additional methodological details and baseline tables are available in the Supplementary Materials.

References

1. Nicholson AG, Tsao MS, Beasley MB et al. The 2021 WHO Classification of Lung Tumors: Impact of Advances Since 2015. *J Thorac Oncol* 2022; 17 (3): 362–387.

2. Lin J, Li S, Chen W et al. Diffuse co-expression of TTF-1 and p40 (DeltaNp63) defines a distinct biphenotypic non-small cell lung carcinoma: clinicopathologic and molecular characterization of 19 cases. *Diagn Pathol* 2026; 21 (1).
3. Pelosi G, Fabbri A, Bianchi F et al. DeltaNp63 (p40) and thyroid transcription factor-1 immunoreactivity on small biopsies or cellblocks for typing non-small cell lung cancer: a novel two-hit, sparing-material approach. *J Thorac Oncol* 2012; 7 (2): 281–290.
4. Yatabe Y, Dacic S, Borczuk AC et al. Best Practices Recommendations for Diagnostic Immunohistochemistry in Lung Cancer. *J Thorac Oncol* 2019; 14 (3): 377–407.
5. Offin M, Chan JM, Tenet M et al. Concurrent RB1 and TP53 Alterations Define a Subset of EGFR-Mutant Lung Cancers at risk for Histologic Transformation and Inferior Clinical Outcomes. *J Thorac Oncol* 2019; 14 (10): 1784–1793.
6. Quintanal-Villalonga A, Chan JM, Yu HA et al. Lineage plasticity in cancer: a shared pathway of therapeutic resistance. *Nat Rev Clin Oncol* 2020; 17 (6): 360–371.
7. Park S, Shim JH, Lee B et al. Paired genomic analysis of squamous cell carcinoma transformed from EGFR-mutated lung adenocarcinoma. *Lung Cancer* 2019; 134: 7–15.
8. Schoenfeld AJ, Chan JM, Kubota D et al. Tumor Analyses Reveal Squamous Transformation and Off-Target Alterations As Early Resistance Mechanisms to First-line Osimertinib in EGFR-Mutant Lung Cancer. *Clin Cancer Res* 2020; 26 (11): 2654–2663.
9. Cai J, Chen G, Zhang W. Interpretation of the 2025 NCCN clinical practice guidelines for non-small cell lung cancer (V1-V3). *Journal of Diagnostics Concepts & Practice* 2026; 25 (1): 37–43.
10. Bittner N. Editorial: New diagnostic and therapeutic possibilities in lung cancer. *Pathol Oncol Res* 2025; 31: 1612132.
11. Filosso PL, Ruffini E, Asioli S et al. Adenosquamous lung carcinomas: a histologic subtype with poor prognosis. *Lung Cancer* 2011; 74 (1): 25–29.
12. Zhao R, Xu Y, Chen Y et al. Clonal dynamics and Stereo-seq resolve origin and phenotypic plasticity of adenosquamous carcinoma. *NPJ Precis Oncol* 2023; 7 (1): 80.
13. Luo J, Gao Q, Wang M et al. Single-cell and spatial transcriptome characterize coinhibitory cell-cell communications during histological progression of lung adenocarcinoma. *Front Immunol* 2024; 15: 1430163.
14. Mao X, Zhang Z, Zheng X et al. Capture-Based Targeted Ultradeep Sequencing in Paired Tissue and Plasma Samples Demonstrates Differential Subclonal ctDNA-Releasing Capability in Advanced Lung Cancer. *J Thorac Oncol* 2017; 12 (4): 663–672.
15. Feng J, Lan Y, Liu F et al. Combination of genomic instability score and TP53 status for prognosis prediction in lung adenocarcinoma. *NPJ Precis Oncol* 2023; 7 (1): 110.
16. Pellegrino B, Mateo J, Serra V et al. Controversies in oncology: are genomic tests quantifying homologous recombination repair deficiency (HRD) useful for treatment decision making? *ESMO Open* 2019; 4 (2): e000480.

17. Tsilingiri K, Chalari A, Christopoulou G et al. Genomic scarring score predicts the response to PARP inhibitors in non-small cell lung cancer. *NPJ Precis Oncol* 2024; 8 (1): 291.
18. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 1995; 57 (1): 289–300.
19. Mayakonda A, Lin DC, Assenov Y et al. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* 2018; 28 (11): 1747–1756.
20. Yu G, Wang LG, Han Y et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012; 16 (5): 284–287.
21. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016; 32 (18): 2847–2849.
22. Parra ER. Methods to Determine and Analyze the Cellular Spatial Distribution Extracted From Multiplex Immunofluorescence Data to Understand the Tumor Microenvironment. *Front Mol Biosci* 2021; 8: 668340.
23. Taube JM, Akturk G, Angelo M et al. The Society for Immunotherapy of Cancer statement on best practices for multiplex immunohistochemistry (IHC) and immunofluorescence (IF) staining and validation. *J Immunother Cancer* 2020; 8 (1): e000155.
24. Benimam MM, Meas-Yedid V, Mukherjee S et al. Statistical analysis of spatial patterns in tumor microenvironment images. *Nat Commun* 2025; 16 (1): 3090.
25. Lewis SM, Asselin-Labat ML, Nguyen Q et al. Spatial omics and multiplexed imaging to explore cancer biology. *Nat Methods* 2021; 18 (9): 997–1012.
26. Barua S, Fang P, Sharma A et al. Spatial interaction of tumor cells and regulatory T cells correlates with survival in non-small cell lung cancer. *Lung Cancer* 2018; 117: 73–79.
27. Bindea G, Mlecnik B, Tosolini M et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 2013; 39 (4): 782–795.
28. Cai S, Yang G, Hu M et al. Spatial cell interplay networks of regulatory T cells predict recurrence in patients with operable non-small cell lung cancer. *Cancer Immunol Immunother* 2024; 73 (10): 189.
29. Lopez de Rodas M, Wang Y, Peng G et al. Objective Analysis and Clinical Significance of the Spatial Tumor-Infiltrating Lymphocyte Patterns in Non-Small Cell Lung Cancer. *Clin Cancer Res* 2024; 30 (5): 998–1008.
30. Firth D. Bias Reduction of Maximum Likelihood Estimates. *Biometrika* 1993; 80 (1): 27–38.
31. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med* 2017; 167 (4): 268–274.
32. Travis WD, Brambilla E, Noguchi M et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol* 2011; 6 (2): 244–285.

33. Frankell AM, Dietzen M, Al Bakir M et al. The evolution of lung cancer and impact of subclonal selection in TRACERx. *Nature* 2023; 616 (7957): 525–533.
34. Marjanovic ND, Hofree M, Chan JE et al. Emergence of a High-Plasticity Cell State during Lung Cancer Evolution. *Cancer Cell* 2020; 38 (2): 229–246 e213.
35. Fort G, Arnold H, Camolotto SA et al. Opposing lineage specifiers induce a protumor hybrid identity state in lung adenocarcinoma. *Genes Dev* 2025; 39 (17–18): 1081–1105.
36. Winslow MM, Dayton TL, Verhaak RG et al. Suppression of lung adenocarcinoma progression by Nkx2-1. *Nature* 2011; 473 (7345): 101–104.
37. Han X, Li F, Fang Z et al. Transdifferentiation of lung adenocarcinoma in mice with Lkb1 deficiency to squamous cell carcinoma. *Nat Commun* 2014; 5 (1): 3261.
38. Qin Z, Yue M, Tang S et al. EML4-ALK fusions drive lung adeno-to-squamous transition through JAK-STAT activation. *J Exp Med* 2024; 221 (3): e20232028.
39. Tong X, Patel AS, Kim E et al. Adeno-to-squamous transition drives resistance to KRAS inhibition in LKB1 mutant lung cancer. *Cancer Cell* 2024; 42 (3): 413–428 e417.
40. Pasquale EB. Eph receptors and ephrins in cancer progression. *Nat Rev Cancer* 2024; 24 (1): 5–27.
41. Bagaev A, Kotlov N, Nomie K et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* 2021; 39 (6): 845–865 e847.
42. Kather JN, Suarez-Carmona M, Charoentong P et al. Topography of cancer-associated immune cells in human solid tumors. *Elife* 2018; 7: e36967.
43. Sorin M, Rezanejad M, Karimi E et al. Single-cell spatial landscapes of the lung tumour immune microenvironment. *Nature* 2023; 614 (7948): 548–554.
44. Grout JA, Sirven P, Leader AM et al. Spatial Positioning and Matrix Programs of Cancer-Associated Fibroblasts Promote T-cell Exclusion in Human Lung Tumors. *Cancer Discov* 2022; 12 (11): 2606–2625.
45. Pelka K, Hofree M, Chen JH et al. Spatially organized multicellular immune hubs in human colorectal cancer. *Cell* 2021; 184 (18): 4734–4752 e4720.
46. Salmon H, Franciszkievicz K, Damotte D et al. Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *J Clin Invest* 2012; 122 (3): 899–910.
47. Sugiyama E, Togashi Y, Takeuchi Y et al. Blockade of EGFR improves responsiveness to PD-1 blockade in EGFR-mutated non-small cell lung cancer. *Sci Immunol* 2020; 5 (43): eaav3937.
48. Yang L, He YT, Dong S et al. Single-cell transcriptome analysis revealed a suppressive tumor immune microenvironment in EGFR mutant lung adenocarcinoma. *J Immunother Cancer* 2022; 10 (2): e003534.
49. Al Bakir M, Reading JL, Gamble S et al. Clonal driver neoantigen loss under EGFR TKI and immune selection pressures. *Nature* 2025; 639 (8056): 1052–1059.

Figures



Clinical and histopathological profile of NSCLC stratified by lineage-plastic TTF-1/p40 expression patterns.

(A) Consort diagram detailing the study workflow with numbers at every step: 110 patients analyzed (all with clinical and PFS data); targeted NGS in 45 (TTF-1 only $n = 13$, focal $n = 21$, diffuse $n = 11$; one contaminated white-blood-cell sample excluded); mlHC in 30 patients (9 TTF-1 only, 4 focal, and 17 diffuse at the patient level; 750 ROI, 20–30 per patient, median 25; 9 with molecular testing records). Spatial comparisons were pre-specified at the region level (double-positive vs TTF-1 only regions).

(B) Representative mlHC and H&E microphotographs defining the TTF-1/p40 expression extremes. Scale bars, 200 μm .

(C) H&E staining and p40 expression in different histological subtypes.

(D) Stacked bar charts showing the distribution of histological subtypes across cohorts; solid components were balanced across groups (median 0% in all groups, $P = 0.176$) and micropapillary components were more frequent in TTF-1 only tumors.

(E) Heatmap of clinical, pathological, and mutational features.

(F, G) Kaplan–Meier curves for PFS in independent subgroups (F; three-group log-rank $P = 0.0473$) and combined subgroups (G; post-hoc dichotomized $P = 0.0184$).

(H) Forest plot detailing univariate Cox regression analysis for PFS.

(I) Multivariable Cox regression for progression-free survival, comparing standard and Firth's penalized models ($n = 110$; 11 progression events; events-per-variable ratio = 2.75). Hazard ratios with 95% confidence intervals are shown on a log scale. With Firth's penalty, the adjusted HR for the diffuse-positive phenotype was 0.16 (95% CI 0.03–1.03, $P = 0.053$), versus 0.097 (95% CI 0.01–0.85, $P = 0.035$) with the standard model. The Firth estimate is reported as the primary value.

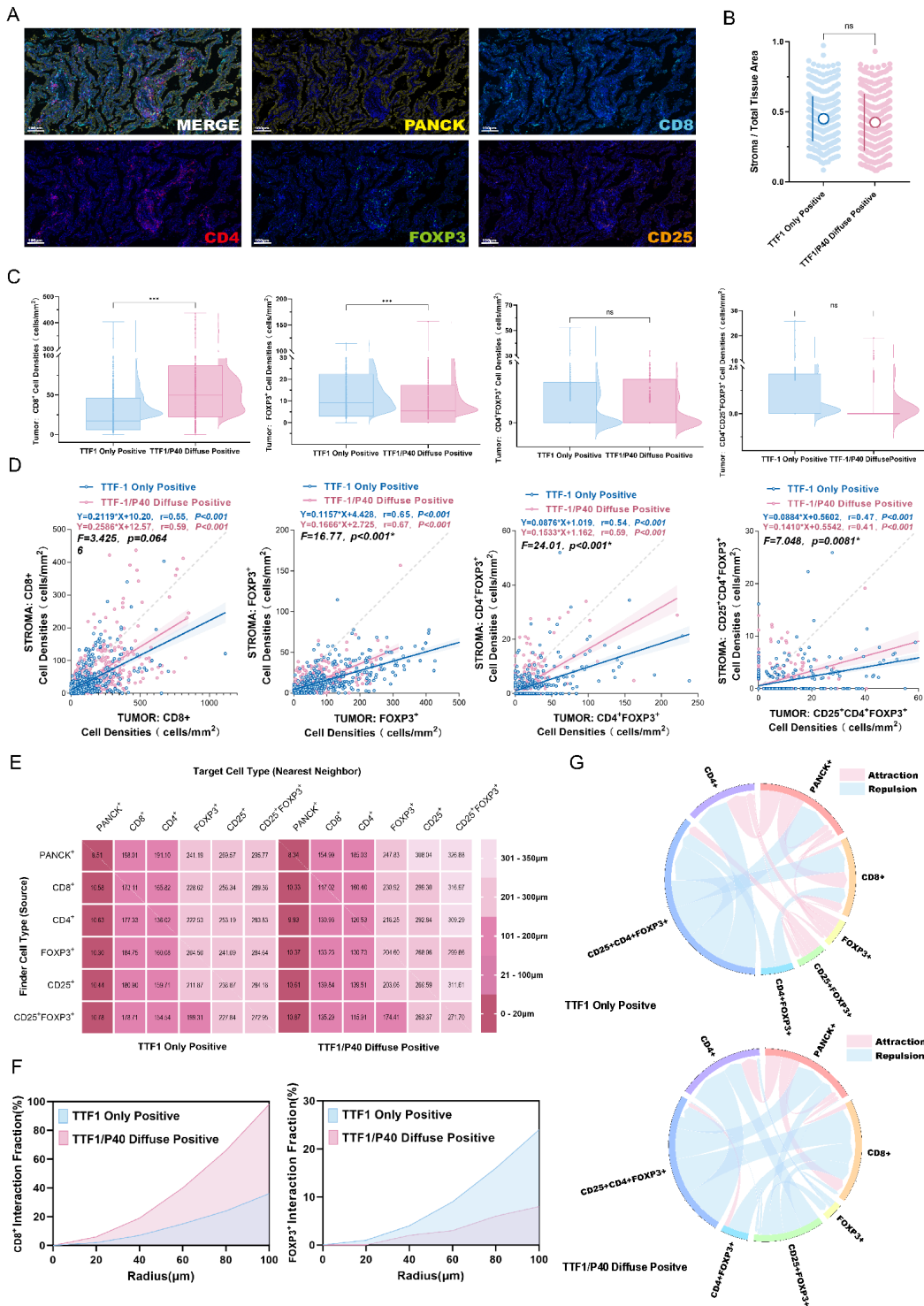


Figure 2

Descriptive spatial redistribution patterns of immune subsets in the TIME.

(A) Representative 6-color mIHC spatial maps (PanCK, CD8, CD4, FOXP3, CD25, and DAPI).

- (B) Stromal area fraction was comparable between TTF-1 only and diffuse positive groups (patient-level $P = 0.257$), arguing against a purely geometric explanation.
- (C) Absolute densities of effector and Treg subpopulations across three spatial phenotypes.
- (D) (D) Region-level linear regression of Treg distribution in stroma versus parenchyma, with per-group fitted lines shown descriptively; patient-level comparisons were non-significant (minimum nominal $P = 0.121$).
- (E) Asymmetric NND matrices showing physical distances between immune and tumor cells.
- (F) Cumulative contact curves showing the fraction of PanCK⁺ tumor cells in contact with effector or suppressive immune cells within 100 μm .
- (G) Finally, to reduce geometric biases from absolute cell densities, we used CSR-normalized chord diagrams to visualize global spatial interaction networks. In these networks, positive interaction weights ($W > 0$, red) indicate attraction, whereas negative weights ($W < 0$, blue) indicate spatial avoidance. The TTF-1 only network showed avoidance weights between PanCK⁺ tumor cells and effector T cells, whereas the diffuse positive network showed stronger PanCK⁺-to-T-cell attraction weights and weaker FOXP3⁺-mediated interactions. Consistent with the patient-level results above, these networks are interpreted as descriptive visualizations of spatial redistribution patterns rather than confirmed biological attraction.

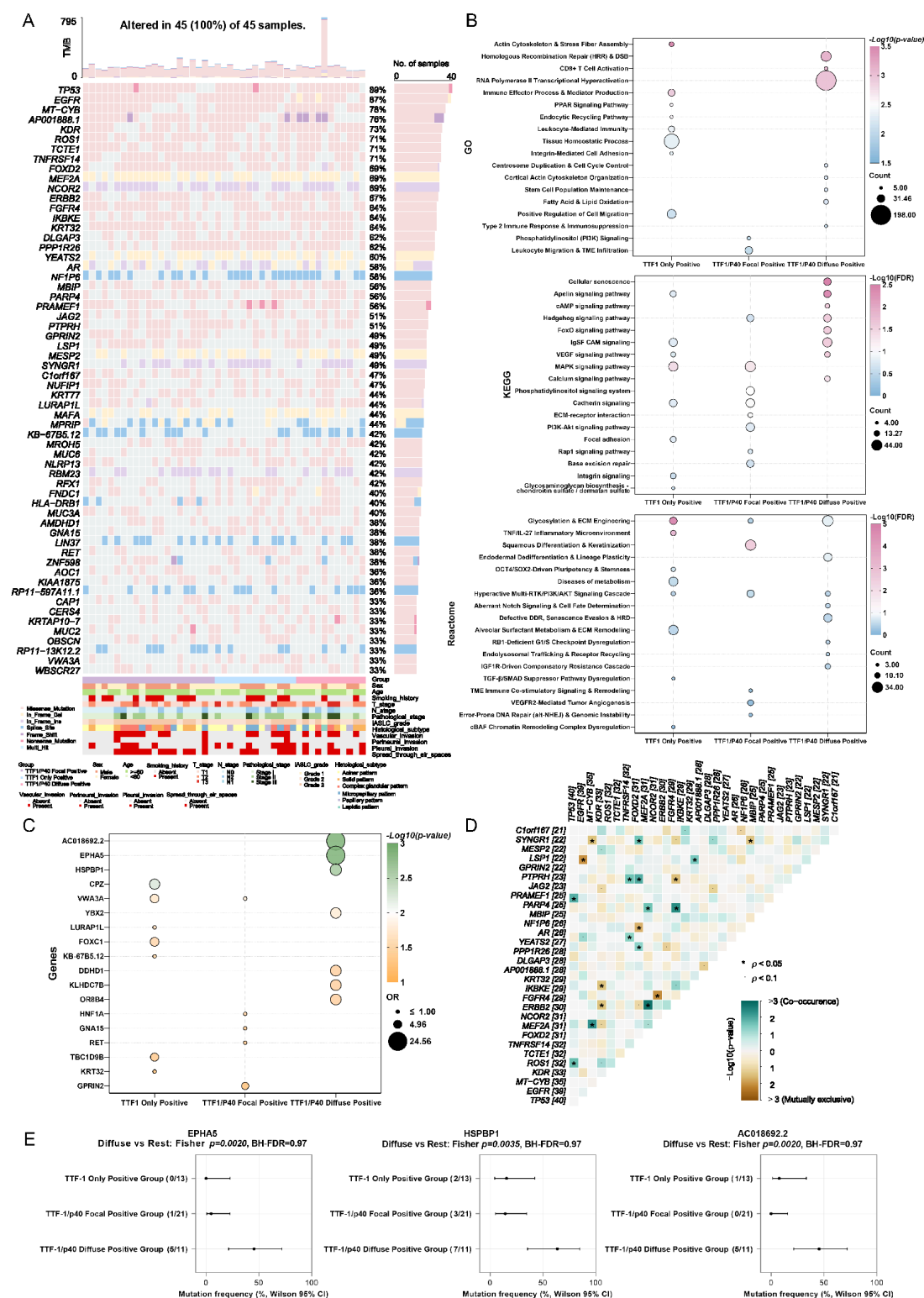


Figure 3

Genomic evolutionary trajectories across the three lineage groups.

(A) Global OncoPrint showing somatic mutations in 45 representative cases, detected by a 68-gene targeted panel.

(B) Functional pathway evolution matrix showing scaled enrichment scores from GO, KEGG, and Reactome.

(C) Bubble plot of nominally enriched genes (*EPHA5*, *HSPBP1*, and *AC018692.2*) in the diffuse dual-lineage group. None survived genome-wide Benjamini–Hochberg FDR correction (all FDR $\approx$ 0.97). Enrichment signals were largely attributable to annotated germline polymorphisms and low-depth calls.

(D) Pairwise somatic interaction matrix showing gene co-occurrence and mutual exclusivity across the three lineage groups.

(E) Mutation frequencies of *EPHA5*, *HSPBP1* and *AC018692.2* across the three lineage groups with Wilson 95% confidence intervals (group sizes: TTF-1 only, n = 13; focal, n = 21; diffuse, n = 11). Mutated samples: *EPHA5* 0/13, 1/21 and 5/11; *HSPBP1* 2/13, 3/21 and 7/11; *AC018692.2* 1/13, 0/21 and 5/11. Nominal Fisher's exact p values for diffuse versus the other groups were 0.002, 0.003 and 0.002, respectively; no gene remained significant after Benjamini–Hochberg correction across 1,568 genes (all FDR $\approx$ 0.97).

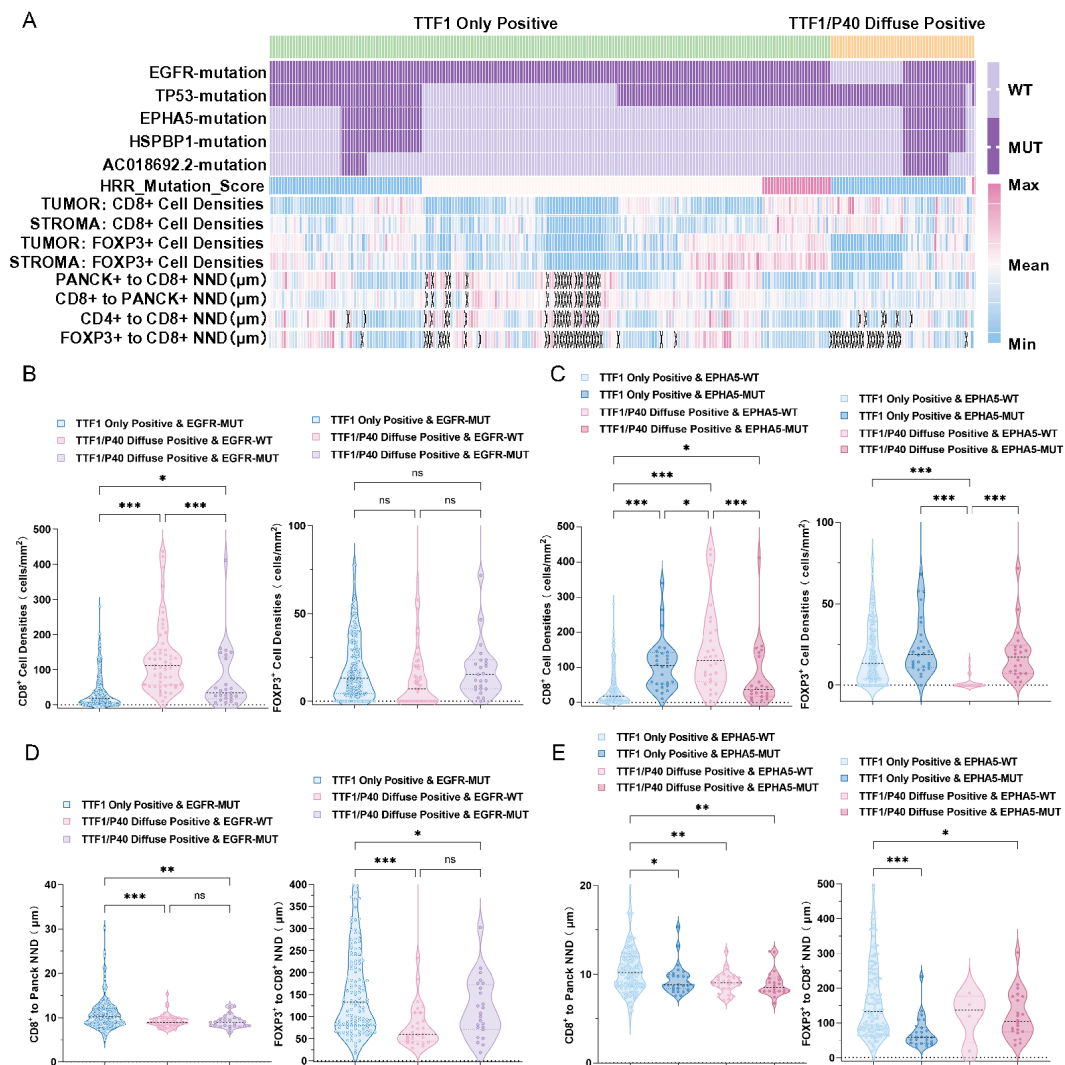


Figure 4

Integration of somatic genotypes and spatial immune phenotypes.

(A) Heatmap of immune cell densities, NND metrics, the exploratory HRR-pathway mutation burden score (mutation-count-based; not a validated LOH+LST+TAI HRD assay), driver gene mutations (*EGFR*

and *TP53*), and candidate loci (*EPHA5*, *HSPBP1*, and *AC018692.2*, hypothesis-generating) by p40 expression status.

(B–C) Violin plots showing the absolute densities of CD8⁺ T cells and FOXP3⁺ Tregs across TTF-1/p40 phenotypes, stratified by *EGFR* (B) and *EPHA5* (C) mutation status.

(D–E) Violin plots of spatial proximity indices (CD8⁺-PanCK⁺ NND; FOXP3⁺-CD8⁺ NND) stratified by *EGFR* (D) and *EPHA5* (E) mutation profiles. Group comparisons were performed using independent-sample t tests and are exploratory with nominal P-values.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [GraphicalAbstract.pdf](#)
- [SupplementaryMaterials.pdf](#)
- [SupplementaryTableS6.xlsx](#)
- [SupplementaryFigureS1.pdf](#)
- [SupplementaryTableS4.xlsx](#)
- [SupplementaryFigureS2.pdf](#)
- [SupplementaryTableS5.xlsx](#)
- [SupplementaryTableS7.xlsx](#)
- [SupplementaryFileList.docx](#)
- [SupplementaryFigureS3.pdf](#)
- [SupplementaryFigureS4.pdf](#)