Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Integrated single-cell profiling reveals subclonal lineage signatures predicting spatial architecture, immune exclusion, and EGFR dependency in head and neck cancer

Rupesh Chaturvedi
rupesh.chaturvedi.jnu@gmail.com
Jawaharlal Nehru University
Anshu Yadav
Jawaharlal Nehru University
Dyumn Dwivedi
Jawaharlal Nehru University
Jyoti Pandey
Jawaharlal Nehru University
Deepali Mishra
Jawaharlal Nehru University
Praveen Kujur
Jawaharlal Nehru University
PRABHJOT KAUR
Jawaharlal Nehru University
RANA SINGH
Gautam Budha University

Article

Keywords:

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11073932/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Integrated single-cell profiling reveals subclonal lineage signatures predicting spatial architecture, immune exclusion, and EGFR dependency in head and neck cancer.

Anshu Yadav[1], Dyumn Dwivedi[1,2,3], Prabhjot kaur[1], Jyoti Pandey[4], Praveen K Kujur[5], Deepali Mishra[5], Rana P. Singh[5,6], Rupesh Chaturvedi[1,7]

1. *School of Biotechnology (SBT), Jawaharlal Nehru University (JNU), New Delhi 110067, India*
2. *Max Institute of Medical Excellence (MAX-MIME), Max Healthcare Institute Limited, New Delhi 110017, India*
3. *Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, Uttar Pradesh 201002, India*
4. *School of Medicine, Washington University, St. Louis, MO 63110, USA*
5. *School of Life science (SLS), Jawaharlal Nehru University (JNU), New Delhi 110067, India*
6. *Gautam Budha University, Greater Noida, Uttar Pradesh , 201312, India*
7. *Nanofludiks, Greater Noida, Gautam Buddha Nagar, Greater Noida, Uttar Pradesh, 201310, India*

***Corresponding author**: Prof. Rupesh Chaturvedi : rupesh.chaturvedi.jnu@gmail.com & Prof. Rana P. Singh: ranaps@hotmail.com*

ABSTRACT

Intratumoral heterogeneity in head and neck squamous cell carcinoma (HNSCC) drives therapeutic resistance, yet the spatial microenvironmental rules governing subclonal survival remain unresolved. Building upon our initial characterization of hypoproliferative (HOPS) and hyperproliferative (HRPS) subclonal lineages, here we perform transcriptomic integration of these subclones to derive consensus lineage signatures and uncover their distinct functional programs. We demonstrate that HRPS clones deploy cell-autonomous, centralized signaling portfolios centered on cell cycle and metabolic expansion within the tumor core. Conversely, slow-cycling HOPS clones execute a decentralized extracellular matrix (ECM) communication program, enabling them to preferentially colonize the invasive leading edge. Spatial transcriptomic deconvolution across independent patient cohorts reveals that leading-edge HOPS enrichment establishes a protective immune-exclusion barrier, marked by diminished CD8+ T-cell infiltration and poor overall survival. Crucially, pharmacogenomic profiling and functional assays demonstrate that HOPS, matrix-remodeling phenotype confers a targetable vulnerability to EGFR inhibition, where baseline HOPS scores predict selective hypersensitivity to the EGFR inhibitor AZD3759, and EGF stimulation dynamically induces the HOPS program. Our findings reveal how subclonal signalling structure the HNSCC tumor microenvironment and provide a biomarker-driven framework to target invasive subclonal niches.

Introduction

Head and neck squamous cell carcinoma (HNSCC) remains one of the most lethal upper aerodigestive malignancies, characterized by poor overall survival and high rates of locoregional recurrence [1]. Despite multimodal therapeutic strategies and the clinical integration of immune checkpoint blockade (ICB), objective response rates remain disappointing [2]. Beyond human papillomavirus (HPV) status in oropharyngeal tumors, robust molecular biomarkers to stratify patient risk and predict targeted therapy response in routine clinical diagnostics remain urgently needed [3]. A primary catalyst of therapeutic failure, locoregional invasion, and treatment resistance in HNSCC is profound intratumoral heterogeneity (ITH) [4][5]. A key driver of this heterogeneity is the emergence of distinct functional subclonal lineages; however, the exact spatial microenvironmental rules governing subclonal survival, invasion, and immune evasion within human tumors remain incompletely understood.

Recently, single-cell RNA sequencing (scRNA-seq) has cataloged malignant cellular diversity in HNSCC [6], yet conventional scRNA-seq analyses often capture transient transcriptomic snapshots without resolving hardwired clonal programs or their physical microenvironmental context [7]. Conversely, clinical bulk transcriptomics and non-spatial multi-omic approaches lack the spatial resolution required to separate distinct subclonal populations operating within the tumor core from those interfacing with the surrounding stroma [8]. To overcome these limitations, we previously established a three-dimensional (3D) single-cell-derived clonal spheroid framework, isolating two co-existing, genetically matched phenotypic lineages: Hyperproliferative Spheroids (HRPS), which drive rapid biomass expansion, and Hypoproliferative Spheroids (HOPS), which exhibit slow-cycling kinetics, stem-like properties, and chemoradiation resistance [9]. However, whether these functionally divergent clonal lineages map onto conserved, physically organized topological niches in patient tumors and how they orchestrate tumor microenvironment (TME) dynamics remains unresolved.

In this study, we integrated single-cell transcriptomic profiles of 3D clonal spheroid lineages (HRPS and HOPS) to establish consensus subclonal lineage signatures and resolve their distinct functional programs. We validated and mapped these subclonal signatures across independent bulk RNA-seq cohorts (TCGA-HNSC) , spatial transcriptomics datasets , and serial patient-derived xenograft (PDX) models to define their spatial topology within intact microenvironments and evaluate their longitudinal stability in vivo. Finally, by linking our subclonal signatures to high-throughput pharmacogenomic databases (DepMap/PRISM) and functional signaling assays, we identified distinct, mutually exclusive therapeutic vulnerabilities for each subclonal state highlighting selective sensitivity of the invasive HOPS lineage to targeted EGFR and MEK pathway inhibition.

Materials and Methods

Single-Cell Dataset Integration, Alignment, and Modularity Clustering

Single-cell RNA sequencing (scRNA-seq) expression matrices derived from hypoproliferative (HOPS) and hyperproliferative (HRPS) 3D clonal spheroid lineages were processed using Seurat (v4.0.0) in R. Quality-control (QC) filtering was performed independently on each sample matrix. Cells were retained for downstream analysis if they expressed between 200 and 2,500 unique genes and exhibited < 15% mitochondrial transcript fraction. Raw count matrices were log-normalized (NormalizeData, scale factor = 10,000). For intra-lineage characterization, individual HOPS and HRPS datasets were analyzed independently: the top 2,000 highly variable features (HVGs) were extracted (FindVariableFeatures, vst method), scaled (ScaleData), and subjected to Principal Component Analysis (PCA). Graph-based Louvain clustering (FindClusters, resolution = 0.5) on individual topologies resolved 4 distinct sub-clusters in the HRPS lineage (Hyper_0–3) and 6 distinct sub-clusters in the HOPS lineage (Hypo_0–5), visualized via Uniform Manifold Approximation and Projection (UMAP).

To construct a unified subclonal landscape across lineages, joint anchor-based dataset integration was conducted using Seurat's Canonical Correlation Analysis (CCA) framework (FindIntegrationAnchors and IntegrateData. Dimensionality reduction was executed via Principal Component Analysis (PCA) on the integrated highly variable genes. Unbiased subclonal topological clustering was conducted using the Louvain community detection algorithm via the *FindClusters* function based on a shared nearest neighbor (SNN) modularity optimization graph at a resolution of 0.5. To visualize the global subclonal landscape, coordinate embedding configurations and the resulting 7 distinct cellular clusters (Clusters 0–6) were projected into a two-dimensional coordinate space via Uniform Manifold Approximation and Projection (UMAP).

Differential Expression Modeling and Signature Generation

In accordance with single-cell integration benchmarking guidelines [10], integrated matrix values were restricted exclusively to low-dimensional visualization (UMAP) and cluster partitioning. All downstream differential gene expression (DEG) modeling, marker identification, and lineage signature derivations were conducted directly on unintegrated, log-normalized UMI counts (DefaultAssay(object) = "RNA") derived from merged count matrices.

To identify robust, non-redundant marker genes defining the HOPS versus HRPS subclonal lineages while controlling for cellular detection rates and pseudoreplication artifacts, a dual-testing methodology was implemented. Differential gene expression across individual cells was conducted using the MAST (Model-based Analysis of Single-cell Transcriptomics) hurdle model within Seurat (FindMarkers, test.use = "MAST") on unintegrated RNA counts. MAST models discrete detection frequency and continuous expression intensity using a two-part generalized linear model adjusted for cellular detection rates. Genes meeting an adjusted P-value (Padj<0.01) and an absolute log2(fold-change)>0.50 were retained.

To eliminate pseudoreplication artifacts, single-cell raw UMI counts were aggregated by biological sample replicate per subclonal lineage using sum-aggregation (AggregateExpression) to construct pseudo-bulk expression profiles. Differential expression testing on pseudo-bulk profiles was executed using DESeq2 (v1.36.0) on raw integer counts. Empirical Bayes shrinkage (apeglm) was applied to adjust log2(fold-change) estimates for low-expression variance. Final consensus subclonal signatures were defined strictly as the intersection of high-confidence single-cell MAST DEGs and pseudo-bulk DESeq2 DEGs (Padj < 0.01, log2FC > 0.50). The HRPS signature was constructed from top-ranked upregulated genes enriched in the hyperproliferative clonal compartment, whereas the HOPS signature comprised top-ranked upregulated genes defining the hypoproliferative clonal compartment.

In Silico Characterization of Intercellular Signaling

Intercellular communication architectures across the subclonal clusters were inferred using CellChat (v1.1.3). Input data comprised the log-normalized single-cell expression matrix containing annotated subclonal states (HRPS, HOPS, and Clusters 0–6). Ligand-receptor interactions were evaluated using the *CellChatDB.human* reference, specifically examining Secreted Signaling, ECM-Receptor, and Cell-Cell Contact pathways. Expression matrices were pre-processed using *identifyOverExpressedGene* and *identifyOverExpressedInteractions* with a minimum threshold of 10 cells per cluster.

Global communication probabilities (interaction strengths) between sender and receiver clusters were computed via *computeCommunicationProbability* using a trimean parameter to account for dropout effects. Pathway-level probabilities were calculated via *computeCommunicationProbabilityPathway*. Network centrality metrics (outdegree, indegree, closeness, betweenness) were derived using *netAnalysis_computeCentrality*. Differential signaling outputs between HOPS and HRPS niches were ranked using rankNet, and dominant axes including Laminin, Collagen, Desmosome, EGF, and Midkine (MK) were visualized via chord and circle diagrams.

Clinical Stratification and Continuous Phenotypic Scoring in TCGA-HNSCC

Raw RNA-seq read counts and clinicopathological data for 528 HNSCC patients were retrieved from the Genomic Data Commons (GDC) portal via TCGAbiolinks (v2.24.0). Counts were normalized and variance-stabilized via vst in DESeq2. To quantify transcriptomic enrichment across the bulk TCGA-HNSCC cohort without introducing classification bias, non-parametric single-sample Gene Set Enrichment Analysis (ssGSEA) was performed via the GSVA package (v1.44.0). ssGSEA derived separate enrichment scores (ES) for the HOPS (ESHOPS) and HRPS (ESHRPS) programs for each patient.

A continuous **Relative Lineage Index** was derived for each patient to capture the relative dominance of the hypoproliferative/invasive program over the hyperproliferative program:

Relative Lineage Index = ESHOPS − ESHRPS

To assess clinical trajectory, the Relative Lineage Index was evaluated as both a **continuous metric** in multivariate Cox proportional hazards models (adjusting for age, sex, AJCC clinical stage, and smoking history) and as a **stratified metric**. For stratified survival visualizations, patients were categorized into **HOPS-dominant**(Relative Lineage Index>upper quartile) and **HRPS-dominant** (Relative Lineage Index<lower quartile) cohorts to assess phenotypic extremes, alongside median-split sensitivity testing across the full distribution. Overall survival (OS) curves were generated using the Kaplan-Meier estimator, and differences were evaluated using the Mantel-Cox log-rank test via the survival (v3.3-1) and survminer (v0.4.9) R packages. Pathological nodal status (pN0 vs. pN+) associations were evaluated using two-sided Wilcoxon rank-sum and Kruskal-Wallis tests.

Deconvolution of Tumor-Infiltrating Immune Cells and Immune Evasion Profiling

The relative proportions of 22 human immune cell subsets were estimated from bulk TCGA-HNSC transcriptomes using the CIBERSORTx digital cytometry platform. The vst-normalized matrix was run against the LM22 signature matrix in absolute mode with batch correction enabled and 100 permutations. Samples with a CIBERSORTx output $P<0.05$ were retained. Overall tumor purity, stromal content, and total immune infiltration scores were calculated using the ESTIMATE algorithm (v1.0.13).

Immune evasion mechanisms were predicted using the TIDE (Tumor Immune Dysfunction and Exclusion) web framework. Normalized expression matrices were centered against the cohort mean. TIDE calculated continuous Exclusion scores, Dysfunction scores, and overall ICB response probabilities. Continuous Relative Lineage Index values were correlated against the expression of canonical checkpoint molecules (*CD274*, *PDCD1*, *CTLA4*, *HAVCR2*, *LAG3*) and cytolytic markers (*GZMA*, *GZMB*, *PRF1*) using Spearman rank correlation, visualized via ComplexHeatmap (v2.12.0) with Benjamini-Hochberg FDR correction ($FDR<0.05$).

Spatial Transcriptomic Mapping and Architectural Analysis

The spatial transcriptomics dataset GSE208253, comprising 12 oral cavity SCC samples profiled using the 10x Genomics Visium platform, was retrieved from the GEO repository . Preprocessing steps and data normalization of the dataset were described earlier [11]. Unsupervised Louvain clustering served to define cluster 1 as “Tumor Core” (TC; *CLDN4* as marker), cluster 3 as “Leading Edge” (LE; *LAMC2* as marker), and cluster 2, exhibiting features of both, as “Transitory region”, as described by Arora et al. [12]. Cluster annotation was used to define spots corresponding to the LE, transitory, and TC regions from each sample.

To evaluate the geographical distribution of the subclonal phenotypes, continuous HOPS and HRPS enrichment scores were calculated for each spot using the AddModuleScore function in Seurat based on population-wide Z-score transformations. A relative lineage score ($Score_{HOPS}-Score_{HRPS}$) was then mapped onto the annotated spatial coordinates. To detect expression differences of these subclonal signatures between all tumor clusters, differential analysis was conducted using a one-way ANOVA test to assess the statistical significance, and a Tukey’s post-hoc test was then applied to identify specific pairs of clusters with significant differences ($*p<0.05$,$****p<0.01$).

Patient-Derived Xenograft (PDX) Generation and Transcriptomic Tracking

All animal procedures were performed in strict accordance with the requirements and guidelines of the Committee for Control and Supervision of Experiments on Animals (CCSEA), Government of India, and were formally approved by the Institutional Animal Ethics Committee (IAEC) of Jawaharlal Nehru University (JNU).

To establish patient-derived xenograft (PDX) models, 6-week-old female BALB/c nude mice were utilized. Primary tumor tissue derived from an HNSCC resection specimen (stage T3N0M0, male, 42 years old) was collected immediately following surgery, washed twice with 1× phosphate-buffered saline (PBS), and sectioned into fragments approximately 1–2 mm in diameter. Tumor fragments were suspended in Matrigel

and subcutaneously implanted into the flanks of recipient mice. A portion of the primary tissue was flash-frozen and stored at −80◦C for bulk RNA sequencing.

When primary xenograft tumors reached a volume of approximately 1,000 mm3 (designated as the F0 generation), tumors were surgically excised and sectioned under sterile conditions. Sectioned fragments were immediately engrafted subcutaneously into secondary recipient BALB/c nude mice to generate the F1 generation, while paired tissue fragments were preserved at −80◦C for transcriptomic profiling.

Total RNA was extracted from harvested tumor tissues across generations and subjected to high-throughput bulk RNA sequencing. To evaluate whether scRNA-seq-derived subclonal lineage signatures were conserved in primary tumor tissue across serial transplantation, single-sample enrichment scores (ESHOPS and ESHRPS) were mapped onto human-specific PDX transcriptomes using GSVA. Pairwise Pearson correlation analyses (r) were performed between single-cell signature profiles and bulk PDX transcriptomes across serial passages (F0→F1) to quantify longitudinal lineage stability and fidelity *in vivo*.

Pharmacogenomic Evaluation and Growth Factor Response Validation

To evaluate the relationship between subclonal lineage dominance and therapeutic vulnerability, pharmacogenomic profiling was conducted utilizing cell line models from the Dependency Map (DepMap) portal. Analysis focused specifically on squamous cell carcinoma models of the head and neck (HNSCC), lung (LUSC), and esophagus (ESCA).

The following primary files were retrieved from DepMap Public 24Q2: [13]

1. Gene Expression Profile Matrix: https://depmap.org/portal/data_page/?tab=allData&releasename=DepMap+Public+24Q2&filename=OmicsExpressionAllGenesTPMLogp1Profile.csv, containing log2(TPM+1)expression counts across human cancer lines.

2. Model and Sample Information Tables**:** https://depmap.org/portal/data_page/?tab=allData&releasename=DepMap+Public+24Q2&filename=Model.csv and https://depmap.org/portal/data_page/?tab=allData&releasename=DepMap%20Public%2024Q2&filename=OmicsProfiles.csv for structural cell annotation.

3. Drug Repurposing Screening Arrays**:** Log-fold change (LFC) viability values from the PRISM https://depmap.org/portal/data_page/?tab=allData&releasename=PRISM%20Repurposing%20Public%2024Q2&filename=Repurposing_Public_24Q2_Extended_Primary_Data_Matrix.csv along with substance annotation records https://depmap.org/portal/data_page/?tab=allData&releasename=PRISM%20Repurposing%20Public%2024Q2&filename=Repurposing_Public_24Q2_Extended_Primary_Compound_List.csv

Continuous lineage dominance tracking scores were derived for each individual cell line by applying the composite signature formula across the standardized transcript arrays. Association distributions between calculated lineage traits and categorical drug LFC vectors were evaluated using the non-parametric **Spearman rank correlation** method, with structural false discovery rate (FDR) control strictly enforced at 5% using Benjamini-Hochberg multi-test modifications.

To independently evaluate phenotypic shifts under growth factor stimulation, transcriptomic profiles from FaDu cells cultured with or without 9 nM recombinant human EGF for 6 and 72 h were obtained from *Schinke et al.* [14]. Independent-sample **Student's t-tests** were used to assess the statistical significance of lineage enrichment trajectory shifts following growth factor exposure, with thresholding defined at a two-tailed $p<0.05$.

Statistical Analysis

Statistical details for each experiment are specified in the figure legends and corresponding methods subsections. All computational analyses were executed in R (v4.3.1) and Python (v3.9.12). For continuous parametric data, two-tailed Student's t-tests or one-way ANOVA were performed; for non-parametric data, Wilcoxon rank-sum or Kruskal-Wallis tests were applied. Survival metrics were evaluated using Kaplan-Meier analyses and log-rank tests. Correlations were computed via Spearman or Pearson coefficients as appropriate. Significance thresholds were defined as $P<0.05$ or $FDR<0.05$ for multi-hypothesis testing.

Results:

Single-cell transcriptomic integration reveals distinct functional states in HNSCC clonal phenotypes

Single-cell RNA sequencing (scRNA-seq) was performed on single-cell suspensions isolated from slow growing hypoproliferative (HOPS) and fast growing hyperproliferative (HRPS) clonal spheroids derived from UMSCC-22b cells, as previously characterized **(Fig. 1A)** [9] . To delineate the transcriptomic programs underlying subclonal heterogeneity, we integrated scRNA-seq datasets derived from HRPS and HOPS subclonal populations. Graph-based Louvain clustering and UMAP projection of the integrated dataset identified seven transcriptionally distinct cellular states (Clusters 0–6) (**Fig. 1B, C**). Cross-mapping the integrated geometry to individual lineage distributions revealed clear compartmentalization across UMAP space: HRPS cells predominantly occupied Clusters 0, 1, and 3, whereas HOPS cells were sequestered within Clusters 1, 2, 4, 5, and 6 (**Fig. 1D**).

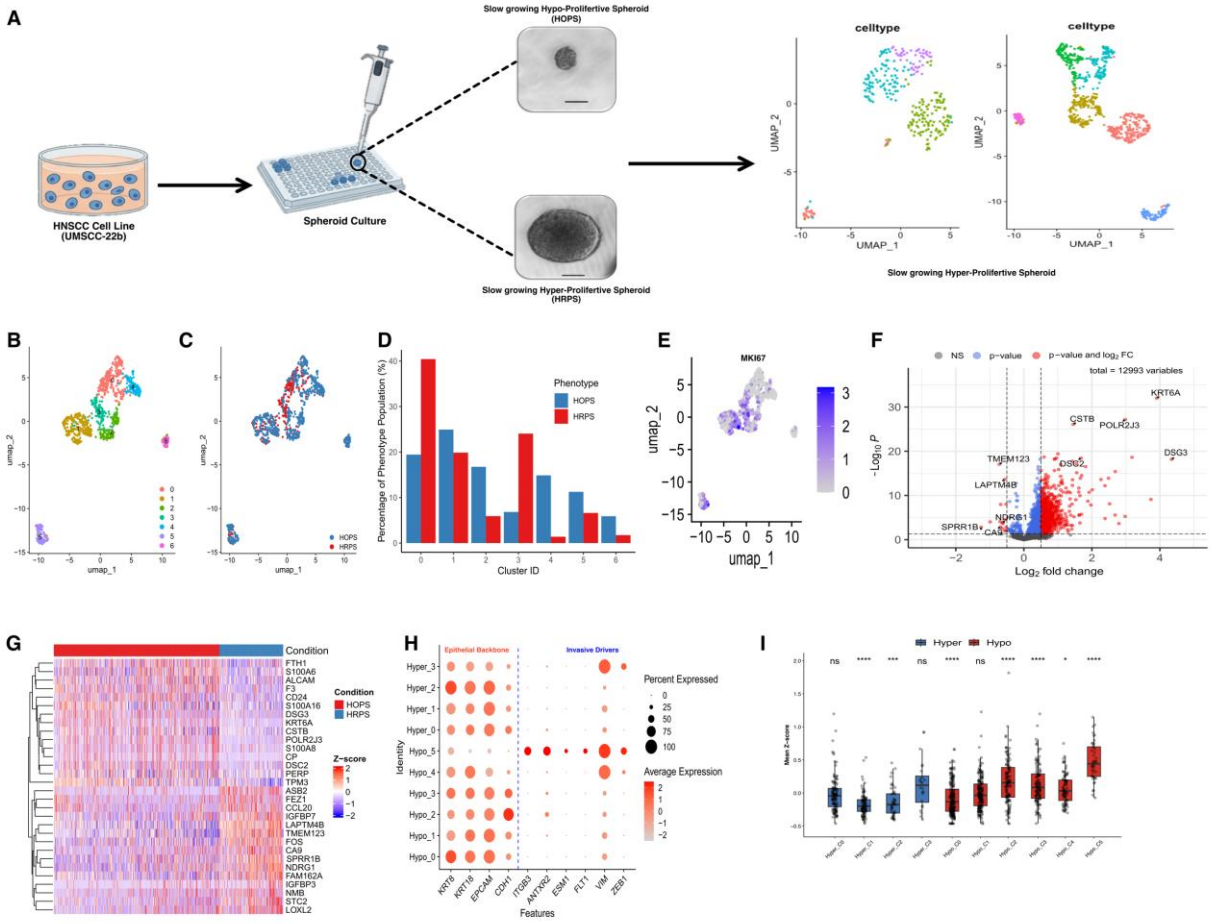


Figure 1 *(**A**) Study Overview. **(B)** UMAP visualization of integrated HRPS and HOPS cells**.** Cells are colored by their clonal phenotype of origin, showing distinct separation between hyperproliferative (HRPS) and hypoproliferative (HOPS) populations. **(C)** Unbiased clustering of the merged dataset**.** Identification of seven transcriptionally distinct clusters (0–6). HRPS cells predominantly occupy Clusters 0, 1, 3 while HOPS cells are sequestered in Clusters 1, 2, 4, 5, and 6.**(D)** Relative distribution of clonal phenotypes across clusters**. (E)** Feature plot for the expression of MKI67 **(F)** Volcano plot displaying differentially expressed genes (DEGs). Significant genes (p-adj < 0.05, |log2FC| > 0.25) are highlighted. **(G)** Heatmap of the top 30 DEGs across condition-specific cells. **(H)** Dot plot showing the expression of an "Epithelial Backbone" alongside "Invasive Drivers" across cluster-specific identities. **(I)** Violin plots showing the distribution of composite p-EMT across individual sub-clusters of hyperproliferative (Hyper_0–3) and hypoproliferative (Hypo_0–5) lineages.*

Expression of proliferation marker (*MKI67*) localized primarily to the HRPS-enriched clusters (Clusters 0, 1, and 3), marking this compartment as the primary proliferative niche (**Fig. 1E**). Pairwise differential expression modeling via MAST and DESeq2 revealed that the HRPS lineage is characterized by a metabolic-stress adaptation program. HRPS cells exhibited significant upregulation of canonical hypoxia-responsive genes (*CA9*, *NDRG1*, *STC2*), metabolic stress transporters (*FAM162A*, *LAPTM4B*), and growth-promoting binding proteins (*IGFBP3*, *IGFBP7*) (**Fig. 1F, G**). Gene Ontology (GO) pathway analysis confirmed enrichment for pathways including "Response to Hypoxia". (**Fig. S1**).

In contrast, the HOPS lineage was enriched for structural architecture and intercellular adhesion machinery over proliferative turnover genes. HOPS clusters were characterized by a coordinated enrichment of desmosomal junction markers (*DSC2*, *DSG3*), cell adhesion molecules (*ALCAM*, *CD24*), and calcium-binding S100 family members (*S100A6*, *S100A8*, *S100A16*) (**Fig. 1G**). GO analysis corroborated this structural specialization, enriching for "Epidermis Development", "Focal Adhesion", and "Ameboidal-type Cell Migration" programs (**Fig. S1**).

To evaluate phenotypic plasticity across subclonal niches, we evaluated canonical epithelial and mesenchymal drivers across normalized individual condition mappings (Hyper_0–3 and Hypo_0–5). While both lineages maintained a core epithelial backbone (*KRT8*, *KRT18*, *EPCAM*, *CDH1*), a hybrid epithelial-mesenchymal (p-

EMT) signature emerged specifically within sub-cluster Hypo_5. This sub-population co-expressed mesenchymal drivers (*VIM*, *ZEB1*) alongside specialized vascular and ECM interaction mediators (*ITGB3*, *ANTXR2*, *ESM1*, *FLT1*) (**Fig. 1I**).

Crucially, to test whether this p-EMT state represented a generalized lineage feature or a compartmentalized sub-state, we quantified the composite p-EMT Z-score distribution across all sub-clusters **(Fig 1I**). Quality control filtering confirmed that score variations were independent of sequencing depth or mitochondrial gene fractions. Rather than a uniform lineage-wide elevation, the p-EMT program was sharply compartmentalized. Within HOPS, Hypo_5 exhibited a distinct expression peak (median Z-score = 0.48), significantly exceeding all other HOPS sub-clusters (Hypo_0, Hypo_1, and Hypo_4; Benjamini-Hochberg FDR-adjusted $P<0.0001$, pairwise Wilcoxon rank-sum tests). Within HRPS, Hyper_3 displayed a secondary, modest elevation relative to Hyper_0–2 (FDR-adjusted $P<0.01$), but remained significantly below the magnitude observed in Hypo_5.

Importantly, while Hypo_5 shares features with the partial-EMT (p-EMT) program reported in primary human HNSCC (*Puram et al., Cell 2017*), our high-resolution analysis demonstrates that this program is not a pervasive HOPS trait, but rather a compartmentalized sub-state preferentially expanded within the HOPS landscape. Together, these transcriptomic data establish that HNSCC subclonal divergence follows a functional division of labor: an HRPS metabolic-stress engine optimized for biomass expansion, and a structural HOPS lineage harboring localized Hypo_5 niches enriched for plastic machinery .

Single-cell communication profiling reveals a shift from metabolic signaling in HRPS to matrix-remodeling networks in HOPS

Having identified localized partial-EMT (p-EMT) plasticity within the HOPS lineage (Hypo_5), we next sought to determine whether subclonal phenotypes operate through distinct microenvironmental cell-cell communication (CCC) networks. To preserve lineage-specific signaling fidelity and avoid cross-dataset alignment distortions, ligand-receptor dynamics were modeled using CellChat directly on the unintegrated individual lineage topologies: HRPS (Hyper_0 – Hyper_3) and HOPS (Hypo_0 – Hypo_5) (**Fig. 2A–D**).

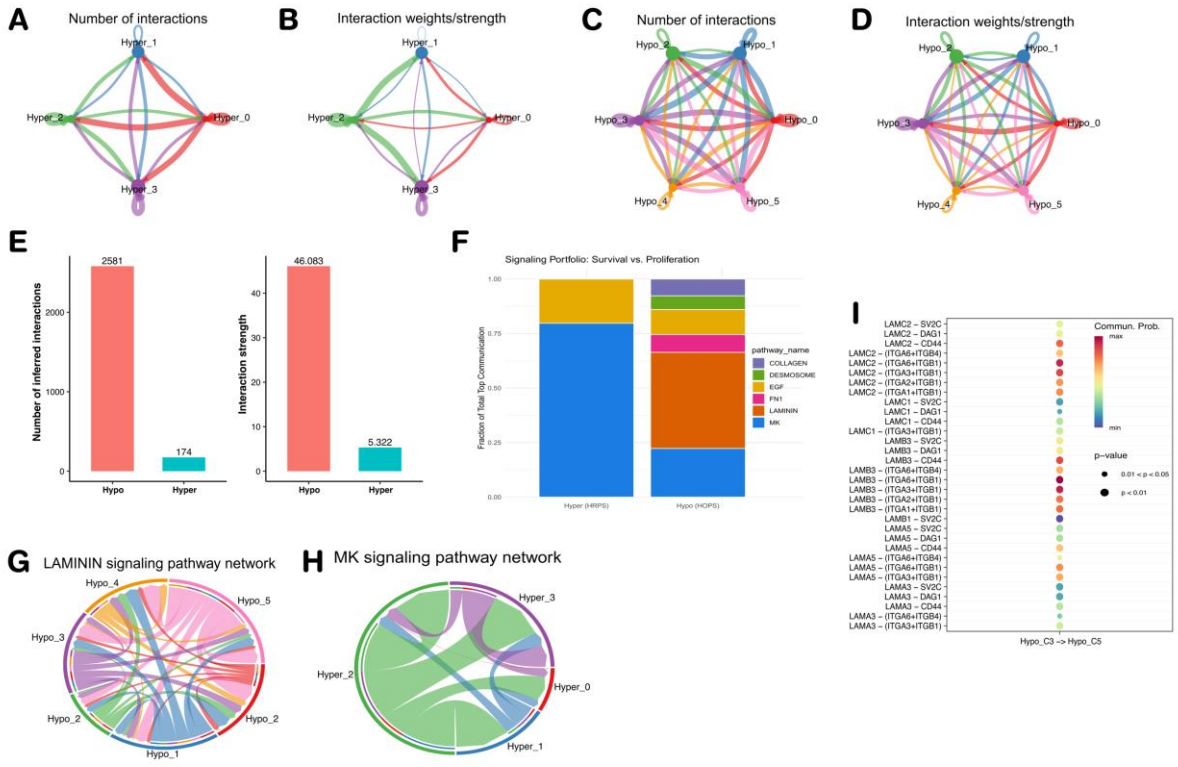


Figure 2. Integrated ligand-receptor analysis reveals a coordinated structural signaling axis in HOPS lineages (A-D) *Global cellular connectivity. Hierarchical circle plots displaying inferred interaction numbers and interaction strengths across individual clusters in hyperproliferative (HRPS; A, B) and hypoproliferative (HOPS; C, D) subclonal topologies. HOPS constructs a dense, decentralized network compared to the sparse signaling topology of HRPS.* ***(E)****Quantification of communication volume. Bar plots comparing total inferred intercellular interactions (left) and overall interaction strength (right), demonstrating the high signaling capacity of the HOPS lineage relative to HRPS.* ***(F)*** *Signaling portfolio diversification. Relative contribution stacked bar chart of top inferred pathways. HRPS signaling is centralized and dominated by Midkine (MK, ~75%), whereas HOPS signaling is decentralized across structural extracellular matrix (ECM) and adhesion machinery (LAMININ, COLLAGEN, DESMOSOME, FN1, EGF).* ***(G, H)*** *Directed pathway chord networks. Chord diagrams illustrating communication flux for the decentralized LAMININ network across HOPS sub-clusters (G) and the centralized MK pathway across HRPS sub-clusters (H).* ***(I)*** *Molecular interaction probabilities targeting the Hypo_5 p-EMT niche. Dot plot displaying specific ligand-receptor communication probabilities from sender structural cluster Hypo_C3 to receiver cluster Hypo_C5. Note the enrichment of Hypo_C3-derived laminin subunit chains (LAMC2, LAMA5, LAMB3) binding integrin heterodimers (ITGA6+ITGB4, ITGA3+ITGB1) and non-integrin ECM receptors (CD44, DAG1) on Hypo_C5 cells (P<0.05, P<0.01).*

Quantitative network comparison revealed a dramatic disparity in intercellular signaling capacity between the two subclonal states. The HOPS lineage exhibited an interconnected, **decentralized** signaling network, displaying a >14-fold higher total number of inferred interactions (2,581 vs. 174) and an ~8.6-fold higher overall interaction strength (46.08 vs. 5.32) relative to HRPS (**Fig. 2E**). While HRPS communication was overwhelmingly **centralized** toward cell-autonomous growth drivers dominated by Midkine signaling (**MK**, accounting for ~75% of total HRPS communication) (**Fig. 2F**) the HOPS lineage constructed a highly diversified signaling portfolio. HOPS communication was enriched for structural matrix components, cell-adhesion machinery, and survival pathways, including **LAMININ**, **COLLAGEN**, **DESMOSOME**, **FN1**, and **EGF** signaling networks (**Fig. 2F**).

To determine how this dense structural network supports the specialized Hypo_5 p-EMT niche, we dissected directional ligand-receptor pair probabilities targeting Hypo_5 (Hypo_C5) as a receiver population. We identified the adjacent structural cluster (Hypo_C3) as a primary signaling hub driving Hypo_C5 activation via robust Laminin-Integrin interaction axes (**Fig. 2I**). Specifically, Hypo_C3 derived laminin chains (*LAMC2*, *LAMA5*, *LAMB3*) strongly engaged canonical integrin complexes

(*ITGA6+ITGB4*, *ITGA3+ITGB1*, *ITGA1+ITGB1*) alongside non-integrin ECM receptors (*CD44*, *DAG1*) on Hypo_C5 cells (P<0.01) (**Fig. 2I**).

Together, these network analyses demonstrate that hyperproliferative HRPS clones act as self-contained units reliant on autonomous growth cues (MK), whereas hypoproliferative HOPS clones construct a dense extracellular matrix niche that actively fuels localized p-EMT transformation through paracrine Laminin-Integrin crosstalk.

Clonal lineage signatures independently predict overall survival and correlate with nodal progression in primary HNSCC

To evaluate the clinical relevance of our subclonal lineage programs in primary patient tumors, we translated the HRPS and HOPS single-cell transcriptomic signatures to the HPV-negative HNSCC cohort from The Cancer Genome Atlas (TCGA-HNSC; n=277) (**Fig. 3**).

We first assessed whether these clonal states hold independent prognostic value for patient survival. Stratifying patients based on the relative lineage index (High HOPS / Low HRPS vs. High HRPS / Low HOPS) using threshold optimization via CutoffFinder yielded significant overall survival separation on Kaplan-Meier analysis (HR=1.48,95% CI: 1.04–2.12,P=0.035) (**Fig. 3A**). To establish whether this survival separation holds independent risk stratification value, we constructed Cox proportional hazards models (**Fig. 3B, C**). In univariable analysis, the continuous lineage index showed a trend toward increased mortality hazard (HR=1.29,95% CI: 0.82–2.02,P=0.27), while ESTIMATE-derived tumor purity was non-significant (HR=0.85,95% CI: 0.56–1.29,P=0.45).

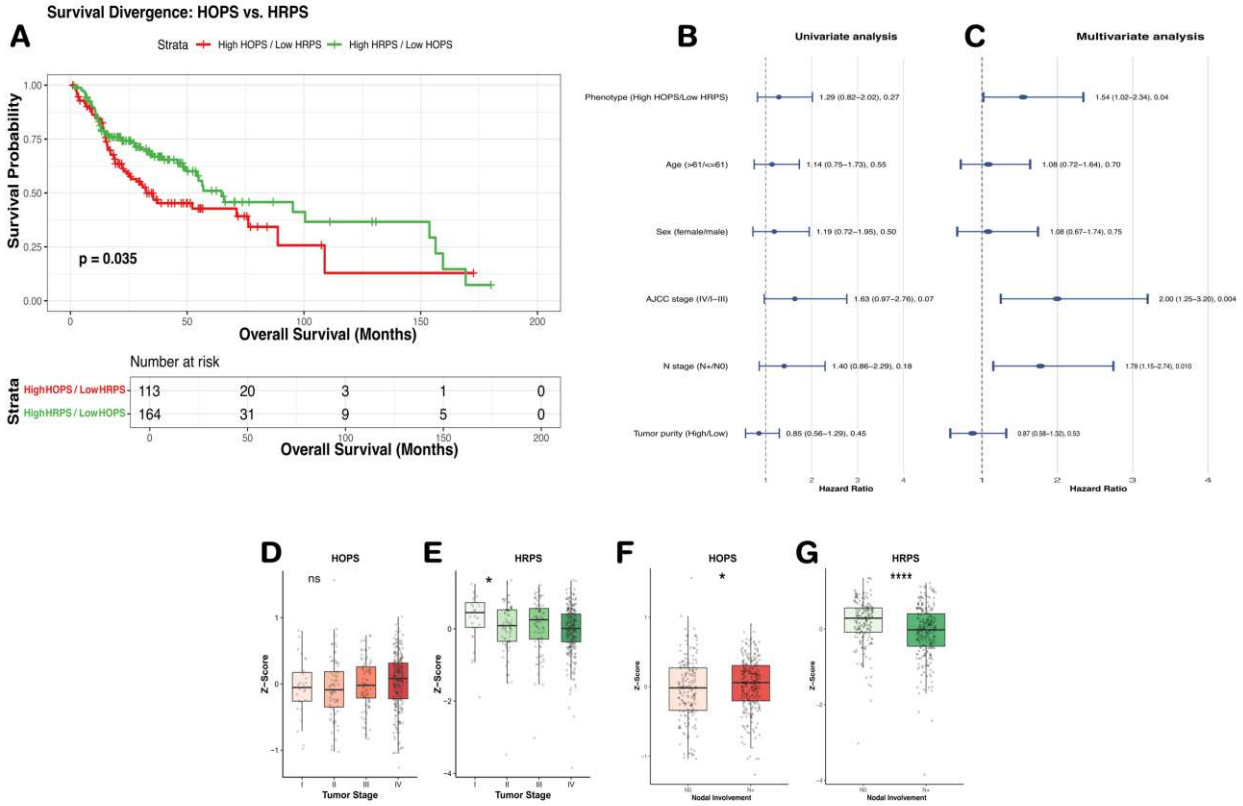


Figure 3. The HOPS lineage signature independently predicts adverse overall survival and correlates with nodal dissemination in primary HNSCC. (a) *Kaplan-Meier survival analysis of the HPV-negative TCGA-HNSC cohort (n=277) stratified by relative lineage index (High HOPS / Low HRPS vs. High HRPS / Low HOPS). Threshold optimization via CutoffFinder confirmed significant overall survival separation (HR=1.48,95% CI: 1.04–2.12,P=0.035, log-rank test).* ***(b, c)*** *Clonal signature dynamics across primary AJCC tumor stages (Stage I, n=27; Stage II, n=74; Stage III, n=81; Stage IV, n=276).* ***(b)*** *HOPS score distributions across stages (P=ns).* ***(c)*** *HRPS score distributions across stages, showing a statistically significant reduction with advancing clinical stage (P<0.05, Kruskal-Wallis test).* ***(d, e)*** *Lineage signature scoring in relation to regional lymph node metastasis (N0 vs. N+ cohorts). Comparison of* ***(d)*** *HOPS scores (P=0.0452) and* ***(e)*** *HRPS scores (P=9.25×10−6) between node-negative and node-positive primary tumors (two-tailed Wilcoxon rank-sum tests).* ***(f, g)*** *Cox proportional hazards regression models for overall survival. Forest plots displaying* ***(f)*** *univariable hazard ratios and* ***(g)*** *multivariable hazard ratios adjusted for age, sex, AJCC clinical stage, N stage, and ESTIMATE tumor purity (HR=1.54,95% CI: 1.02–2.34,P=0.040).*

Upon controlling for clinical covariates including age, sex, AJCC stage, nodal stage, and tumor purity in a multivariable model, the High HOPS / Low HRPS phenotype emerged as a statistically significant, independent predictor of adverse overall survival (HR=1.54,95% CI: 1.02–2.34,P=0.040; overall model Likelihood Ratio P=0.0308). Crucially, tumor purity remained non-significant in the multivariable setting (HR=0.87,95% CI: 0.58–1.32,P=0.53), confirming that the adverse prognosis associated with the HOPS program reflects intrinsic cancer cell-state shifts rather than non-neoplastic stromal infiltration.

We next evaluated how these lineage states distribute across primary AJCC tumor stages (Stages I–IV) (**Fig. 3D,E**). Evaluating HOPS expression across disease stages revealed a nominal upward trend from Stage I to Stage IV that did not achieve overall statistical significance (**Fig. 3D**, P=ns). In contrast, the structural HRPS signature exhibited a statistically significant decrease with advancing clinical stage (**Fig. 3E**, P<0.05, Kruskal-Wallis test). This shift reflects a biological transition wherein early-stage primary tumors maintain a proliferative baseline, whereas advanced Stage IV tumors show up-regulated structural adhesion.

Crucially, this lineage division demonstrated strong regional compartmentalization during metastasis (**Fig. 3F, G**). Comparison of signature Z-scores between node-negative (N0) and node-positive (N+) tumors revealed that the HOPS signature was significantly enriched in patients with regional lymph node metastases (**Fig. 3F**, P=9.25×10−6, two-tailed Wilcoxon rank-sum test). Conversely, the HRPS signature was markedly suppressed in the N+ cohort relative to N0 controls (**Fig. 3G**, P=0.0452).

Together, these clinical findings confirm that the hypoproliferative HOPS state is not a passive bystander, but an invasive phenotype that enriched in node-positive disease and independently predicts adverse overall survival in HNSCC patients.

Clonal lineages correlate with divergent immune infiltration profiles and maps to conserved spatial landmarks.

To investigate whether subclonal lineage identity dictates the tumor microenvironment (TME) architecture in primary HNSCC, we performed deconvolution and TME scoring across the TCGA-HNSC cohort (**Fig. 4**).

Deconvolution of 22 leukocyte subsets using CIBERSORTx demonstrated that HRPS and HOPS tumors occupy distinct immunological states (**Fig. 4A**). HOPS-dominant tumors (High_Hypo) exhibited a pronounced depletion of anti-tumor effector populations including cytotoxic CD8+ T cells and activated Natural Killer (NK) cells compared to HRPS-dominant tumors (High_Hyper) (P<0.05) (**Fig. 4A**). Concurrently, HOPS tumors were significantly enriched for M0-polarized macrophages. Hierarchical clustering of actionable immune checkpoint molecules (*CD274*/PD-L1, *CTLA4*, *HAVCR2*/TIM-3, *LAG3*, *IDO1*) further demonstrated global upregulation of exhaustion targets in the HRPS lineage, reflecting an immune-active ("hot") phenotype (**Fig 4B**). In contrast, the HOPS lineage displayed coordinate reduction across all evaluated checkpoint molecules, establishing an immune-cold phenotype. To quantify global microenvironmental properties, we applied the ESTIMATE algorithm (**Fig. 4C**). Consistent with the deconvolution profiles, High_Hypo tumors exhibited a significantly lower composite ImmuneScore relative to High_Hyper controls (P<0.0001), along with a reduced overall ESTIMATEScore (P<0.05) (**Fig. 4C**). Predictive modeling using the Tumor Immune Dysfunction and Exclusion (TIDE) framework demonstrated that the HOPS phenotype was significantly enriched for clinical non-responders to immune checkpoint blockade (77.8%, n=175/225) compared to the HRPS phenotype (39.6%, n=91/230) (**Fig 4D**). Granular evaluation of individual TIDE immune evasion signatures revealed that the refractoriness of the HOPS phenotype is mechanistically driven by severe T-cell Exclusion (P<0.0001, Wilcoxon rank-sum test; **S2**).

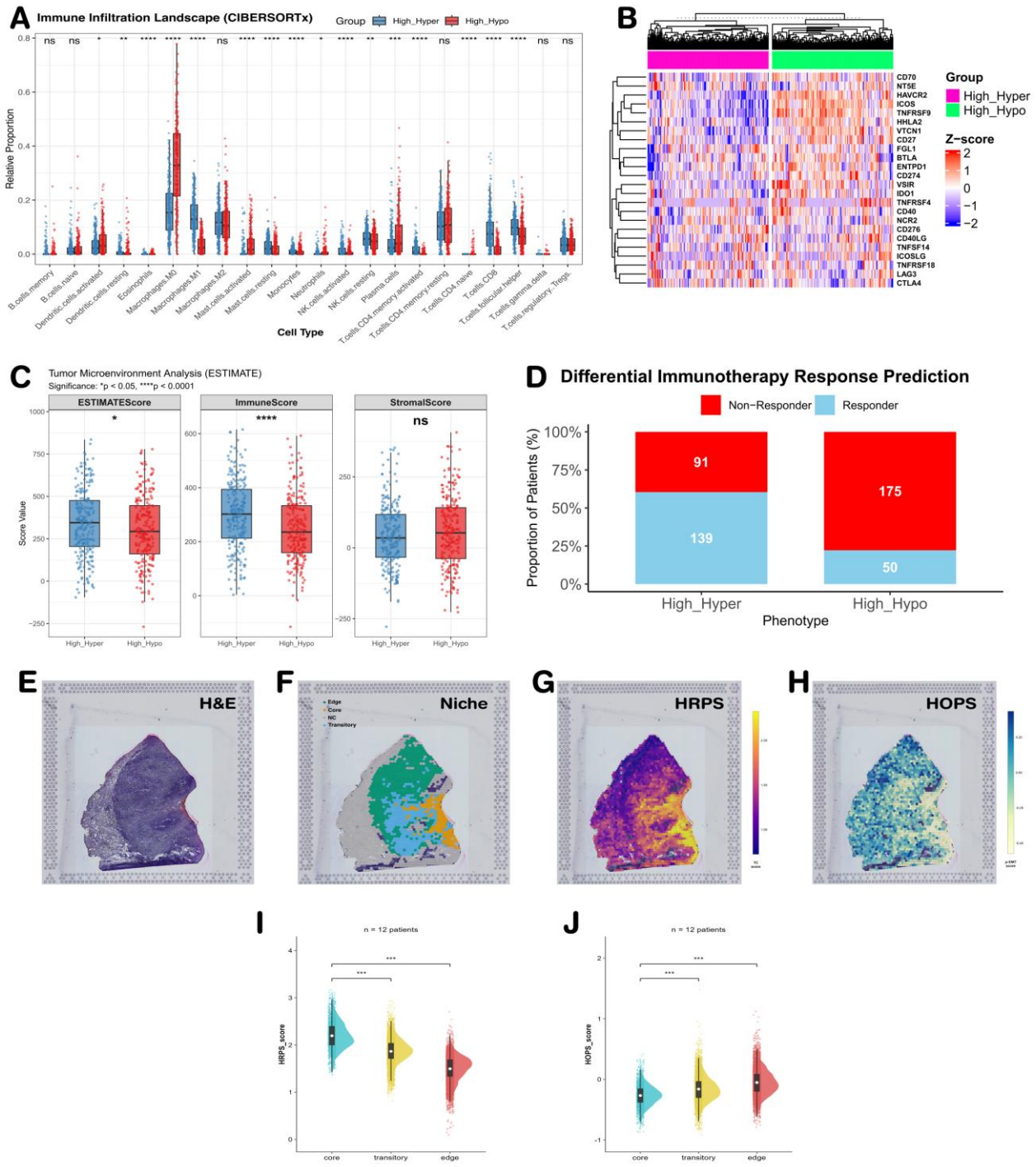


Figure 4. Clonal lineage signatures define an immune-cold microenvironment refractory to immune checkpoint blockade. ***(A)*** *Relative abundance of 22 immune cell subsets estimated using CIBERSORTx deconvolution across High_Hyper (HRPS) and High_Hypo (HOPS) cohorts. (* Wilcoxon rank-sum test; ∗P<0.05,P<0.01, ∗∗∗P<0.001, ∗∗∗∗P<0.0001, ns=non-significant).* ***(B)*** *Hierarchical clustering and Z-score expression heatmap of actionable immune checkpoint.* ***(C)*** *Global tumor microenvironment quantification using the ESTIMATE algorithm. Boxplots display comparative ESTIMATEScore (∗P<0.05), ImmuneScore (∗∗∗∗P<0.0001), and StromalScore (ns) across lineage phenotypes.* ***(D)*** *Predicted clinical response to immune checkpoint blockade (ICB) modeled via the Tumor Immune Dysfunction and Exclusion (TIDE) framework.* ***(E-H)*** *Spatial transcriptomic profiling of a representative HNSCC tissue section (patient 32, GSE208253).* ***(E)*** *H&E histological section.* ***(F)*** *Spot annotation map*

delineating defined microenvironmental niches: Tumor Core (TC, green), Leading Edge (LE/Edge, orange), Transitory (light blue), and Non-Cancerous stroma (NC, purple). ***(G)*** *Spatial transcriptomic heatmaps displaying the topographical distribution of the HRPS score and* ***(H)*** *the HOPS score.* ***(I, J)*** *Centered, SD-normalized score distributions across microenvironmental zones depicted as raincloud plots.* ***(I)*** *HRPS score displays significant step-wise attenuation from the core to the leading edge (∗∗∗P<0.001).* ***(J)*** *HOPS score demonstrates significant step-wise enrichment from the tumor core to the invasive leading edge (∗∗∗P<0.001, linear mixed-effects model / post-hoc ANOVA).*

Next, to resolve the physical architecture giving rise to this immune-cold microenvironment, we projected single-cell subclonal signatures onto high-resolution spatial transcriptomics data (n=12 HNSCC tissue sections, GSE208253) across defined microenvironmental domains (**Fig 4E-H**). Histological mapping verified three distinct spatial zones: the cohesive Tumor Core (TC), the invasive Leading Edge (LE), and an intermediate Transitory domain (**Fig 4F**). Projecting continuous signature scores onto spatial coordinates demonstrated distinct topographical compartmentalization (**Fig 4E-H, S4**). The HRPS signature was tightly sequestered within interior tumor nests (**Fig 4G**), whereas the HOPS signature localized along the outer boundary interfacing with surrounding stroma (**Fig 4H**). Across the full spatial cohort (n=12), linear mixed-effects modeling confirmed a step-wise attenuation of the HRPS score from the core to the leading edge ($P<0.001$; **Fig 4I**), alongside progressive step-wise HOPS enrichment toward the invasive margin ($P<0.001$; **Fig 4J**).

Differential expression analysis between HRPS and HOPS states across spatial landmark genes revealed a clear molecular duality. The HRPS lineage selectively overexpressed canonical tumor core marker involved in squamous differentiation and stress response (*SPRR1B*) (**Fig S3**). Conversely, the HOPS lineage was enriched for leading-edge markers governing partial epithelial-mesenchymal transition (p-EMT), cell-matrix adhesion, and basement membrane remodeling (*SERPINE1*, *VIM*, *ITGA5*, *SNAI2*, *LAMC2*) (**Fig S3)**.

Together, these data demonstrate that subclonal lineage identity fundamentally shapes the TME landscape: hyperproliferative HRPS clones drive interior expansion within the tumor core inside an inflamed, checkpoint-rich niche, whereas hypoproliferative HOPS clones selectively colonize the invasive leading edge, coordinating matrix remodeling and establishing a structural barrier that mediates T-cell exclusion and primary immunotherapy resistance.

Patient-derived xenograft models preserve subclonal lineage fidelity across serial *in vivo* passaging

To determine whether the single-cell transcriptomic signatures of the HRPS and HOPS lineages reflect stable, intrinsic tumor phenotypes rather than transient *in vitro* adaptations, we established patient-derived xenograft (PDX) models across serial *in vivo* generations (**Fig. 5**).

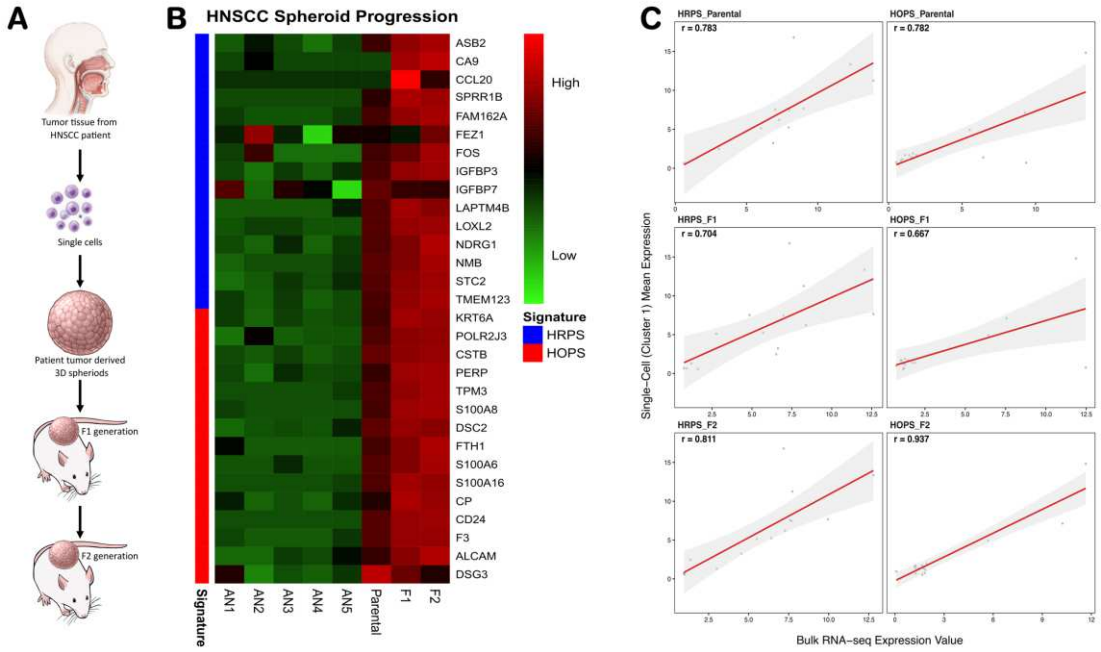


Figure 5. Patient-derived xenograft models recapitulate subclonal lineage identities across serial *in vivo* passaging. ***(A)*** *Schematic representation of the experimental workflow for patient-derived xenograft (PDX) generation. Primary tumor tissue from an HNSCC patient (T3N0M0/M/42) was dissociated into single cells, established as 3D spheroids, and subcutaneously engrafted with Matrigel into female BALB/c nude mice to yield the parental (F0) generation, followed by serial transplantation to generate F1 and F2 xenograft cohorts.* ***(B)*** *Expression heatmap of canonical HRPS (blue bar) and HOPS (red bar) marker genes across adjacent normal tissues (AN1–AN5; n=5), primary patient tumor (Parental), and derived PDX models (F1 and F2 generations). Both lineage signatures show robust, coordinated upregulation in parental tumors and serial PDX generations relative to adjacent normal controls.* ***(C)*** *Pearson correlation scatter plots comparing single-cell transcriptomic profiles (Cluster 1 mean expression) with bulk RNA-sequencing data across parental tumor and PDX generations (HRPS, left; HOPS, right). Transcriptomic fidelity is highly preserved across generations, exhibiting strong correlation in parental (HRPS r=0.783; HOPS r=0.782), F1 (HRPS r=0.704; HOPS r=0.667), and F2 generations (HRPS r=0.811; HOPS r=0.937).*

Primary HNSCC tissue (T3N0M0/M/42) was processed into 3D clonal spheroids and subcutaneously engrafted into female BALB/c nude mice to generate parental (F0) tumors, followed by serial transplantation to establish F1 and F2 xenograft generations (**Fig. 5A**). Bulk RNA-sequencing of matched primary tumors (PT), serial PDX tissues (F0–F2), and adjacent normal mucosal controls (AN1–AN5; n=5) was performed to track subclonal signature stability across passaging (**Fig. 5B**).

Heatmap profiling confirmed that canonical gene sets defining both the metabolic HRPS state (*CA9*, *NDRG1*, *STC2*, *IGFBP3*, *FAM162A*) and the structural HOPS state (*KRT6A*, *S100A8*, *DSC2*, *CD24*, *DSG3*) were markedly suppressed or absent in adjacent normal tissues (AN1–AN5) (**Fig. 5B**). Conversely, both lineages exhibited coordinated, robust expression across the primary patient tumor (Parental) and were faithfully maintained across F1 and F2 PDX generations (**Fig. 5B**).

To quantitatively evaluate transcriptomic retention, we correlated single-cell cluster expression profiles against bulk RNA-sequencing data from matched xenograft generations (**Fig. 5C**). Both lineages demonstrated high correlation with primary patient tumors (HRPS Parental r=0.783; HOPS Parental r=0.782) . Crucially, serial *in vivo* passaging preserved these lineage identities with increasing fidelity, reaching strong correlation values in the F2 generation for both HRPS (r=0.811) and HOPS (r=0.937).

These findings demonstrate that the subclonal HRPS and HOPS lineage programs represent stable, biologically hardwired cellular identities that persist across serial patient-derived xenograft passaging, providing a robust preclinical platform for lineage-targeted therapeutic interventions.

Pharmacogenomic profiling identifies EGF-driven therapeutic vulnerabilities selective for the HOPS subclonal lineage

To identify actionable therapeutic strategies capable of selectively targeting distinct subclonal populations, we mined high-throughput pharmacogenomic data from the DepMap and PRISM databases. Correlating our single-cell lineage signatures with small-molecule viability screens across head and neck squamous cell carcinoma lines mapped distinct, mutually exclusive drug sensitivities for each subclonal state (**Fig. 6**).

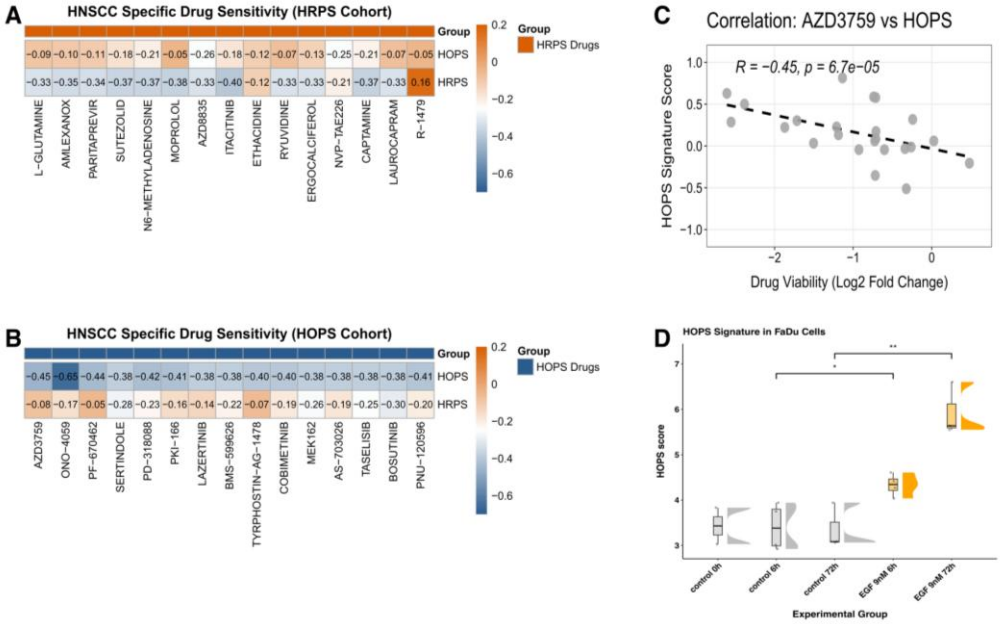


Figure 6. Pharmacogenomic profiling and functional validation identify lineage-specific therapeutic vulnerabilities in HNSCC. (A, B) *Heatmaps displaying Spearman correlation coefficients (R) between single-cell lineage scores and small-molecule drug viability values (log2-fold change) across HNSCC cell lines from DepMap/PRISM databases.* ***(A)****HRPS-selective vulnerabilities enriched for metabolic and anti-proliferative inhibitors (L-Glutamine, R=−0.33; Amlexanox, R=−0.35; Paritaprevir, R=−0.34).* ***(B)*** *HOPS-selective vulnerabilities enriched for EGFR/RTK and downstream MAPK pathway inhibitors (AZD3759, R=−0.45; ONO-4059, R=−0.65; PF-670462, R=−0.44; Cobimetinib, R=−0.40).* ***(C)*** *Scatter plot demonstrating a significant inverse correlation between baseline HOPS signature score (y-axis) and cellular viability upon treatment with the EGFR inhibitor AZD3759 (x-axis, log2-fold change viability) across HNSCC cell lines (R=−0.45, P=6.7×10−5).* ***(D)*** *Raincloud and boxplot distribution of HOPS signature scores in FaDu cells under serum-free control conditions (0h, 6h, 72h) versus recombinant EGF stimulation (9 nM; 6h and 72h). EGF treatment induces a time-dependent, significant upregulation of the HOPS program (∗P<0.05at 6h,P<0.01 at 72h, Student's t-test).*

Cell lines displaying a prominent hyperproliferative (HRPS) signature exhibited selective sensitivity to metabolic perturbations and anti-proliferative compounds (**Fig. 6A**). Top HRPS hits included L-Glutamine deprivation (R=−0.33), the TBK1/IKK\varepsilon inhibitor amlexanox (R=−0.35), and paritaprevir (R=−0.34). These sensitivities reflect the energetic reliance of HRPS tumor cores on rapid cell-cycle progression and biomass accumulation.

Conversely, cells displaying a high HOPS signature were uniquely sensitive to receptor tyrosine kinase (RTK) and downstream MAPK cascade inhibitors (**Fig. 6B**). Among top HOPS candidates, the blood-brain barrier-penetrating EGFR inhibitor AZD3759 emerged as a primary hit ($R=-0.45$), alongside the BTK inhibitor ONO-4059 ($R=-0.65$) and MEK1/2 inhibitors including cobimetinib ($R=-0.40$) and MEK162 ($R=-0.38$).

To confirm the selectivity of EGFR inhibition in HOPS-dominant models, we analyzed the direct correlation between baseline HOPS signature scores and AZD3759 viability response across HNSCC lines (**Fig. 6C**). Baseline HOPS scores exhibited a strong, statistically significant inverse linear correlation with drug viability ($R=-0.45$, $p<0.0001$), confirming that elevated HOPS expression predicts heightened sensitivity to EGFR targeted inhibition (**Fig. 6C**).

Because HOPS-high lines exhibited pronounced sensitivity to EGFR and MEK inhibitors, we analysed whether extracellular EGFR activation directly regulates the HOPS program. We evaluated HOPS signature dynamics in FaDu cells under serum-free control conditions versus recombinant EGF stimulation (9 nM) across time (**Fig. 6D**). While serum-starved control cells maintained low baseline HOPS expression across 72 hours, EGF stimulation triggered a significant, time-dependent induction of the HOPS signature at 6 hours ($p<0.05$) that peaked at 72 hours ($p<0.01$).

Together, these pharmacogenomic and functional findings demonstrate that subclonal lineage identities dictate distinct therapeutic vulnerabilities: HRPS clones rely on metabolic pathways supporting rapid proliferation, whereas HOPS clones are driven by active EGFR/MAPK signaling, rendering them uniquely vulnerable to targeted RTK and MEK pathway inhibition.

Discussion

Intratumoral heterogeneity (ITH) has emerged as a critical driver of therapeutic resistance, disease recurrence, and poor clinical prognosis across solid malignancies, including HNSCC [15] . In this study, we integrated single-cell transcriptomics, high-resolution spatial transcriptomics, serial patient-derived xenograft (PDX) passaging, computational immune profiling, and high-throughput pharmacogenomics to comprehensively characterize the subclonal architecture and functional divergence of subclonal lineages in HNSCC.

Spatial transcriptomics analysis revealed that subclonal lineage identity dictates physical organization within intact tumor architecture [16]. The observed spatial segregation wherein HRPS clones sequester within interior tumor core nests while HOPS clones selectively colonize the invasive tumor-stroma interface demonstrates that subclonal divergence operates within conserved spatial landmarks . The strong overexpression of basement membrane and matrix remodeling factors (*LAMC2*, *VIM*, *SERPINE1*, *ITGA5*, *SNAI2*) along the HOPS-dominated leading edge further supports the role of cell-matrix integrin interactions and partial epithelial-mesenchymal transition (p-EMT) in driving local tissue invasion. This structural duality aligns with prior reports by Puram et al. and related studies demonstrating that p-EMT programs and laminin-332 deposition at the invasive front facilitate localized extracellular matrix degradation and lymph node metastasis in HNSCC [17] . Conversely, the selective upregulation of squamous differentiation and stress-response markers (*KRT8*, *KRT18*, *SPRR1B*) in interior HRPS tumor nests underscores an evolutionary division of labor, mirroring previous observations where central tumor regions adapt to hypoxic and metabolic stress through rapid proliferation and cell-cycle progression [18].

Extending beyond cellular mapping, our multi-omic integration revealed that subclonal spatial positioning directly dictates the tumor immune microenvironment (TME) [19]. Deconvolution of 22 leukocyte subsets demonstrated that while HRPS core regions maintain an inflamed, immune-active state characterized by high ImmuneScores and elevated expression of actionable checkpoints (*CD274*/PD-L1, *CTLA4*, *HAVCR2*, *LAG3*, *IDO1*), HOPS-dominant tumors establish a profoundly immune-cold microenvironment. This refractoriness to ICB was mechanistically linked via TIDE modeling to severe structural T-cell exclusion ($P<0.0001$) rather than intrinsic T-cell exhaustion. The concurrent depletion of cytolytic effector cells (CD8+ T cells and activated NK cells) alongside selective enrichment of unpolarized

M0-macrophage precursors at the HOPS-dominated margin suggests that margin-bound HOPS cells construct a physical and microenvironmental barrier restricting effector lymphocyte entry into the tumor core. These observations complement recent findings showing that dense stromal matrix barriers and non-activated myeloid precursors at the tumor margin physically impede T-cell infiltration, driving primary resistance to anti-PD-1/PD-L1 therapies [20].

Tracking transcriptomic stability across serial patient-derived xenograft (PDX) generations confirmed that these subclonal programs represent hardwired biological identities rather than transient tissue-culture artifacts. Both HRPS and HOPS signatures were markedly suppressed or absent in adjacent normal mucosa (AN1–AN5) while exhibiting robust, coordinated expression across parental primary tumors and derived xenografts. Crucially, cross-correlation analysis demonstrated maximum transcriptomic retention in F2 xenografts (HRPS $r=0.811$; HOPS $r=0.937$), confirming that these lineage signatures remain durable across serial *in vivo* passaging. This longitudinal stability validates our PDX model as a reliable platform for lineage-targeted drug discovery, consistent with established guidelines emphasizing the necessity of serial xenograft stability for biomarker translation [21].

Our data comprehensively characterized the therapeutic vulnerabilities governing each subclonal state. Pharmacogenomic mining across DepMap/PRISM SCC lines revealed mutually exclusive drug sensitivities: HRPS clones exhibited reliance on metabolic expansion and anti-proliferative pathways (L-Glutamine deprivation, amlexanox), whereas HOPS cells were uniquely sensitive to receptor tyrosine kinase (RTK) and downstream MAPK cascade inhibitors. Among top candidates, the EGFR inhibitor AZD3759 displayed a strong inverse linear correlation with HOPS signature expression ($R=-0.45$, $P=6.7\times10-5$). Functionally, recombinant EGF stimulation triggered a time-dependent induction of the HOPS program in FaDu cells ($P<0.01$ at 72 h), establishing that extracellular EGFR signaling dynamically regulates the leading-edge HOPS state. These findings align with prior functional screens in squamous cancers demonstrating that hyperactivation of the EGFR-MEK-ERK signaling axis directly drives p-EMT programs, invadopodia formation, and local motility [22] .

Furthermore, recent clinical trials evaluating MAPK/MEK inhibition such as the FRAME trial combining avutometinib with FAK inhibitors have shown significant clinical efficacy in driven solid tumors [23], supporting our hypothesis that targeted EGFR and MEK pathway inhibition represents a potent strategy to eliminate invasive leading-edge subclonal populations.

Lastly, while benefiting from cross-platform spatial and single-cell sequencing integration, our study has potential limitations. Computational immune deconvolution applied to bulk cohorts relies on predefined reference single-cell matrices, which may mask rare transitional myeloid phenotypes. Additionally, while our PDX models confirm transcriptomic stability across serial passaging *in vivo*, immunocompromised mice lack functional T-cell signaling, necessitating future humanized mouse models or prospective clinical trial cohorts to validate subclone-targeted combination strategies with ICB. Nevertheless, cross-platform validation across bulk RNA-seq (TCGA-HNSC), single-cell transcriptomics, high-resolution spatial transcriptomics (n=12), and experimental PDX/cell-line validations reinforces the validity and translational robustness of our findings.

Author Contribution declaration

Anshu Yadav: Conceptualization, data curation, software and formal analysis, interpretation of data, including statistics, biostatistics, computational analysis, investigation, visualization, methodology, writing−original draft, writing-review, and editing. **Dyumn Dwivedi:** Conceptualization, analysis, and interpretation of data, including statistics, biostatistics, computational analysis, visualization, writing review, and editing.. **Prabhjot Kaur:** Writing – review & editing, Writing – original draft. **Jyoti Pandey**: Conceptualization , Data Curation. **Deepali Mishra**: Formal analysis, Data Curation. **Praveen K Kujur:** Formal analysis, Data Curation. **Rana P Singh**: Conceptualization, formal analysis, funding acquisition, study supervision, project administration. **Rupesh Chaturvedi:** Conceptualization, formal analysis, funding acquisition, study supervision, project administration, writing, and editing.

Acknowledgements

This work was supported by the Department of Science & Technology (VI-D&P/546/2016-17/TDT (C)). Anshu Yadav and Dyumn Dwivedi were supported by CSIR-UGC-JRF, CCRAS, respectively, in India.

Use of Generative AI and AI-Assisted Technologies in the Writing Process

The writing of this paper involved the use of generative AI and AI- assisted technologies only to enhance the clarity, coherence, and overall quality of the manuscript. The authors acknowledge the contributions of AI in the writing process while ensuring that the final content reflects the authors' own insights and interpretations.

FUNDING

This research was funded by the Department of Science & Technology (VI-D&P/546/2016-17/TDT (C)), India.

Data availability

Data will be made available on request.

References

[1] F. Bray *et al.*, 'Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries', *CA. Cancer J. Clin.*, vol. 74, no. 3, pp. 229–263, May 2024, doi: 10.3322/caac.21834.

[2] J. V. Poulose and C. T. Kainickal, 'Immune checkpoint inhibitors in head and neck squamous cell carcinoma: A systematic review of phase-3 clinical trials', *World J. Clin. Oncol.*, vol. 13, no. 5, pp. 388–411, May 2022, doi: 10.5306/wjco.v13.i5.388.

[3] B. Burtness *et al.*, 'Pembrolizumab alone or with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-048): a randomised, open-label, phase 3 study', *The Lancet*, vol. 394, no. 10212, pp. 1915–1928, Nov. 2019, doi: 10.1016/S0140-6736(19)32591-7.

[4] I. Dagogo-Jack and A. T. Shaw, 'Tumour heterogeneity and resistance to cancer therapies', *Nat. Rev. Clin. Oncol.*, vol. 15, no. 2, pp. 81–94, Feb. 2018, doi: 10.1038/nrclinonc.2017.166.

[5] E. A. Mroz, A. D. Tward, C. R. Pickering, J. N. Myers, R. L. Ferris, and J. W. Rocco, 'High intratumor genetic heterogeneity is related to worse outcome in patients with head and neck squamous cell carcinoma', *Cancer*, vol. 119, no. 16, pp. 3034–3042, Aug. 2013, doi: 10.1002/cncr.28150.

[6] S. V. Puram *et al.*, 'Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer', *Cell*, vol. 171, no. 7, pp. 1611-1624.e24, Dec. 2017, doi: 10.1016/j.cell.2017.10.044.

[7] V. K. A. Sreenivasan, J. Henck, and M. Spielmann, 'Single-cell sequencing: promises and challenges for human genetics', *Med. Genet. Mitteilungsblatt Berufsverb. Med. Genet. EV*, vol. 34, no. 4, pp. 261–273, Dec. 2022, doi: 10.1515/medgen-2022-2156.

[8] D. Gong, J. M. Arbesfeld-Qiu, E. Perrault, J. W. Bae, and W. L. Hwang, 'Spatial oncology: Translating contextual biology to the clinic', *Cancer Cell*, vol. 42, no. 10, pp. 1653–1675, Oct. 2024, doi: 10.1016/j.ccell.2024.09.001.

[9] J. Pandey *et al.*, 'Uncovering functional divergence and cellular clusters with specific gene signatures in HNSCC clonal spheroids', *Cell Commun. Signal.*, vol. 23, no. 1, p. 527, Dec. 2025, doi: 10.1186/s12964-025-02360-3.
[10] M. D. Luecken *et al.*, 'Benchmarking atlas-level data integration in single-cell genomics', *Nat. Methods*, vol. 19, no. 1, pp. 41–50, Jan. 2022, doi: 10.1038/s41592-021-01336-8.
[11] J. Zhou *et al.*, 'EGFR-mediated local invasiveness and response to Cetuximab in head and neck cancer', *Mol. Cancer*, vol. 24, no. 1, p. 94, Mar. 2025, doi: 10.1186/s12943-025-02290-1.
[12] R. Arora *et al.*, 'Spatial transcriptomics reveals distinct and conserved tumor core and edge architectures that predict survival and targeted therapy response', *Nat. Commun.*, vol. 14, no. 1, p. 5029, Aug. 2023, doi: 10.1038/s41467-023-40271-4.
[13] B. DepMap, 'DepMap 24Q2 Public'. Figshare+, p. 25031191226 Bytes, 2024. doi: 10.25452/FIGSHARE.PLUS.25880521.
[14] H. Schinke *et al.*, 'A transcriptomic map of EGFR-induced epithelial-to-mesenchymal transition identifies prognostic and therapeutic targets for head and neck cancer', *Mol. Cancer*, vol. 21, no. 1, p. 178, Sep. 2022, doi: 10.1186/s12943-022-01646-1.
[15] Y.-C. Fu, S.-B. Liang, M. Luo, and X.-P. Wang, 'Intratumoral heterogeneity and drug resistance in cancer', *Cancer Cell Int.*, vol. 25, no. 1, p. 103, Mar. 2025, doi: 10.1186/s12935-025-03734-w.
[16] M. John *et al.*, 'Spatial transcriptomics reveals profound subclonal heterogeneity and T-cell dysfunction in extramedullary myeloma', *Blood*, vol. 144, no. 20, pp. 2121–2135, Nov. 2024, doi: 10.1182/blood.2024024590.
[17] J. D. Aleman, C. D. Young, S. D. Karam, and X. Wang, 'Revisiting laminin and extracellular matrix remodeling in metastatic squamous cell carcinoma: What have we learned after more than four decades of research?', *Mol. Carcinog.*, vol. 62, no. 1, pp. 5–23, Jan. 2023, doi: 10.1002/mc.23417.
[18] N. Sureka *et al.*, 'Hypoxic microenvironment in cancer: role in metabolic reprogramming', *Front. Oncol.*, vol. 16, p. 1771365, 2026, doi: 10.3389/fonc.2026.1771365.
[19] J.-E. See, Y. Liu, and J. Plummer, 'Spatial omics of immunity: Mapping cellular landscapes in tissue microenvironments', *Curr. Opin. Immunol.*, vol. 101, p. 102790, Aug. 2026, doi: 10.1016/j.coi.2026.102790.
[20] Z. Zhang, Y. Gu, S. Qiao, L. Sang, L. Zhang, and F. Zhou, 'Navigating the landscape of CD8+ T cell tumor infiltration: Barriers, mechanisms, and therapeutic opportunities', *Cell Press Blue*, vol. 1, no. 4, p. 100056, Jul. 2026, doi: 10.1016/j.cpblue.2026.100056.
[21] Y. Liu, W. Wu, C. Cai, H. Zhang, H. Shen, and Y. Han, 'Patient-derived xenograft models in cancer therapy: technologies and applications', *Signal Transduct. Target. Ther.*, vol. 8, no. 1, p. 160, Apr. 2023, doi: 10.1038/s41392-023-01419-2.
[22] A. Fejza, L. Camicia, E. Poletto, G. Carobolante, M. Mongiat, and E. Andreuzzi, 'ECM Remodeling in Squamous Cell Carcinoma of the Aerodigestive Tract: Pathways for Cancer Dissemination and Emerging Biomarkers', *Cancers*, vol. 13, no. 11, p. 2759, Jun. 2021, doi: 10.3390/cancers13112759.
[23] S. Banerjee *et al.*, 'Defactinib with avutometinib in patients with solid tumors: the phase 1 FRAME trial', *Nat. Med.*, vol. 31, no. 9, pp. 3074–3080, Sep. 2025, doi: 10.1038/s41591-025-03763-y.

Supplementary Files

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

- SupplementaryFigureS1.docx
- dataandcodeaccess.docx
- graphicalabstract.pdf