Lopez *et al. Cell Communication and Signaling* (2025) 23:386
https://doi.org/10.1186/s12964-025-02405-7

Cell Communication and Signaling

RESEARCH Open Access

Intercellular signaling reinforces single-cell level phenotypic transitions and facilitates robust re-equilibrium of heterogeneous cancer cell populations

Daniel Lopez[1], Darren R. Tyson[2] and Tian Hong[3*]

Abstract

Background Cancer cells within tumors exhibit a wide range of phenotypic states driven by non-genetic mechanisms, such as epithelial-to-mesenchymal transition (EMT), in addition to extensively studied genetic alterations. Conversions among cancer cell states can result in intratumoral heterogeneity which contributes to metastasis and development of drug resistance. However, mechanisms underlying the initiation and/or maintenance of such phenotypic plasticity are poorly understood. In particular, the role of intercellular communications in phenotypic plasticity remains elusive.

Methods In this study, we employ a multiscale inference-based approach that integrates single-cell transcriptomic data to predict phenotypic changes and tumor population dynamics. Our computational framework combines ligand-receptor interaction inference (*CellChat*), transcription factor activity estimation (*decoupleR*), and causal signaling network reconstruction (*CORNETO*) to analyze single-cell RNA sequencing (scRNA-seq) data and investigate how intercellular interactions influence cancer cell phenotypes, with a particular focus on EMT-related gene programs. We further use mathematical models based on ordinary differential equations, informed by network inferences, to examine how intercellular communication shapes phenotypic dynamics at the population level from a dynamical systems perspective.

Results Our inference approach reveals that signaling interactions between cancerous cells in small cell lung cancer (SCLC) result in the reinforcement of the phenotypic transition in single cells and the maintenance of population-level intratumoral heterogeneity. Additionally, we find a recurring signaling pattern across multiple types of cancer in which the mesenchymal-like subtypes utilize signals from other subtypes to support its new phenotype, further promoting the intratumoral heterogeneity. Our models show that inter-subtype communication both accelerates the development of heterogeneous tumor populations and confers robustness to their steady state phenotypic compositions.

*Correspondence:
Tian Hong
hong@utdallas.edu

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

Conclusions Our work highlights the critical role of intercellular signaling in sustaining intratumoral heterogeneity, and our approach of computational analysis of scRNA-seq data can infer inter- and intra-cellular signaling networks in a holistic manner.

Background

Most tumors develop and evolve as complex ecosystems under strong environmental selective pressures, leading to a unique collection of cancer cells that exhibit a wide range of genotypic and phenotypic characteristics [1, 2] . This intratumoral heterogeneity promotes aggressive disease progression, increased resistance to therapeutic interventions, and poor overall survival [1, 3–5] . While genetic diversity is a well-known driver of intratumoral heterogeneity [6] there is increasing evidence that non-genetic mechanisms, such as epigenetic, transcriptional, and/or translational changes, also significantly contribute to the intratumoral heterogeneity and disease progression [4, 5, 7, 8] . These non-genetic mechanisms can create distinct cancer cell states through a process called phenotypic plasticity, where cells are dynamic, reversible, and responsive to regulatory changes [3, 9, 10] .

Phenotypic plasticity has recently been recognized as a hallmark of cancer and a key driver of tumor aggressiveness [3] . It influences diverse and often opposing cellular behaviors in cancer, including stemness and differentiation, drug-sensitive and drug-resistant states, and transitions between epithelial and mesenchymal cell-states [11] . Increasing efforts are being made to characterize the intrinsic cellular factors that drive phenotypic plasticity [12–14] . However, despite extensive molecular characterization, the dynamics of phenotypic plasticity at the single-cell and population levels remain largely unclear. This is particularly true regarding how non-cell-autonomous effects regulate intratumoral heterogeneity and whether the intercellular communication between different cell states stabilizes or destabilizes these phenotypes.

One such cancer where phenotypic plasticity is particularly evident is small cell lung cancer (SCLC) [15, 16] . SCLC is a neuroendocrine (NE) carcinoma that constitutes approximately 15% of all lung cancer cases and has a dismal five-year survival rate of less than 7% [17] . Despite the high similarity to pulmonary NE cells and having highly consistent morphological characteristics, SCLC presents substantial inter- and intratumoral heterogeneity, featuring distinct molecular subtypes with varied biological behaviors [18–22] . These subtypes are categorized based on the enriched expression of one of four transcription factors (TFs): *ASCL1*, *NEUROD1*, *POU2F3*, *YAP1* [23] . The subtype naming conventions: SCLC-A, -N, -P and -Y, correspond to the enrichment of these four TFs, respectively. These molecular subtypes delineate into two overarching categories, NE (SCLC-A2, -A, and -N) and non-NE (SCLC-P and -Y), with the NE subtypes typically exhibiting some level of ASCL1 expression while non-NE counterparts do not. Recent studies have demonstrated that SCLC tumors will often comprise multiple cell types, with the different subtypes cooperating to drive tumorgenicity [21, 22] . The dynamic regulation of TFs regulates intratumoral compositions, and this diversity is essential as different subtypes play distinct biological roles, impacting therapeutic response [21, 22] .

Further highlighting the phenotypic plasticity evident within SCLC, previous work by us [24] and others [25, 26] has linked the different SCLC subtypes to the epithelial-mesenchymal transition (EMT) program, a cellular process in which cell-cell interactions are remodeled, resulting in cells losing their epithelial properties and assuming a more mesenchymal phenotype [27] . Within SCLC, the NE subtype SCLC-A2 demonstrates a strong epithelial-like phenotype whereas the other NE subtypes, SCLC-A and SCLC-N, display a partial EMT state (Fig. 1A). non-NE subtypes, SCLC-P and SCLC-Y, also display a partial EMT state, albeit with mesenchymal gene expression signatures that differ from that of the NE subtypes. This correspondence with EMT further demonstrates the plastic nature of this cancer.

Most research on SCLC has focused on describing its differing cell states, with very few studies attempting to define these states at the single-cell level. Consequently, there is still much to learn about the transitions between different cell states and whether intercellular communication influences the intratumoral heterogeneity. Given the plasticity present within SCLC, we utilized single-cell RNA-sequencing (scRNA-seq) SCLC data to investigate whether intercellular communications influence cell-fate transitions, and if so, whether these extracellular signals reinforce the current phenotype of a cell or push it towards a different phenotype.

To investigate whether intercellular communications play a role in driving intratumoral heterogeneity within SCLC, we adapted a single-cell multiscale inference-based approach that integrates both intercellular and intracellular signaling information from scRNA-seq data [28] . This multiscale framework links cell-cell communication networks (e.g., ligand-receptor signaling) to downstream transcriptional programs within individual cells (e.g., transcription factor activity), enabling us to assess how external signals influence regulatory states and to subsequently model cell dynamics at the population level. By connecting these layers of signaling, we aim to determine whether interactions between cancer cells reinforce

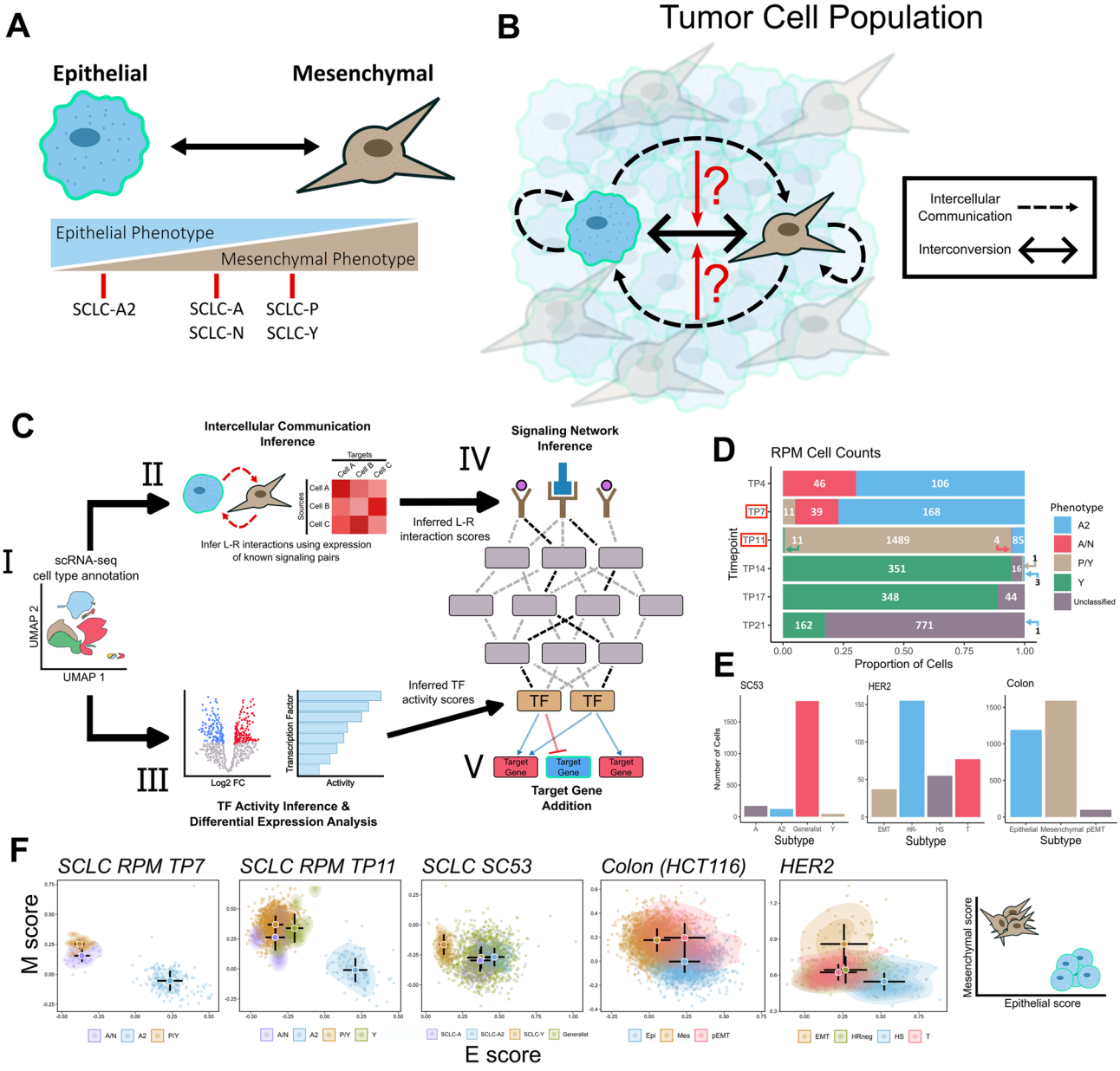

Fig. 1 Multiscale Inference Approach to Investigate Role of Intercellular Communication on Cellular Plasticity. **A** Schematic depicting EMT and correspondence to SCLC subtypes. Subtypes are characterized based on transcriptomic expression. **B** Investigating whether intercellular communications affect the cell state transitions between subtypes and the overall intratumoral heterogeneity. Black dashed arrows represent inferred cell-cell communication and the black, double-arrow represents the interconversion between subtypes. **C** I: Pre-processed and annotated scRNA-seq data. Pipeline is applied to four different datasets across three different cancers II: Inference of cell-cell communications using *CellChat*. *CellChat* models communication using a law of mass action model (see reference [32]), which generates a matrix of L-R interaction scores. These interaction scores are then used as one of the inputs for the *CORNETO* algorithm. III: Differential expression and TF activity analysis. Differentially expressed genes were identified using *Seurat* and TF activity scores were computed via *decoupleR*. IV: *CORNETO* performs signaling network inference by using the L-R interaction scores, TF activities and a prior knowledge graph consisting of protein-protein interactions. The nodes within the prior knowledge graph are assigned weights based on their log-fold values, allowing the *CORNETO* algorithm to find the optimal path connecting the receptors to TFs (dashed black lines). V: Differentially expressed SCLC and EMT gene markers are added to the network based on filtering criteria (see Methods). **D** Filled bar chart showing the cell counts for the SCLC RPM dataset at TP7. X-axis represents the proportion of cells and y-axis is the timepoint. The cell numbers for each cell type are shown. TP7 and TP11 (red boxes) are the focus of our analysis within the RPM dataset due to their prominent cell heterogeneity. **E** Cell counts for SCLC SC53, HER2 and colon cancer datasets, respectively. The cell type annotations are from the original authors. **F** Scatter plots of single-sample gene set enrichment analysis (ssGSEA). Each dot represents a cell in the dataset. X- and y-axis represent enrichment scores of epithelial and mesenchymal genes, respectively. Summary of the results depicting the arrangement of epithelial and mesenchymal cells on the far right

existing phenotypes or drive transitions towards a different phenotype (Fig. 1B). To explore the effect of cell–cell communications on phenotypic heterogeneity, we utilized scRNA-seq data from ex vivo cultured cells obtained from a genetically engineered mouse model that incorporates a constitutively active form of *MYC*, coincident with deletions of *Rb* and *p53* (RPM) [29] . The cells undergo a transition over time in culture from a NE to non-NE state and these time series data enable the exploration of intercellular communications between the NE and non-NE populations at individual time points. Through this approach, we identified the activation of several well-established EMT pathways, revealing a consistent pattern of convergence towards activating mesenchymal genes within the mesenchymal phenotype. Additionally, our analysis revealed that the epithelial phenotype employs both paracrine and autocrine signaling mechanisms to maintain its epithelial state. To see if these mechanisms are consistent across epithelial cancers that involve EMT, we applied this method to colon and breast cancer datasets. The convergence of EMT pathways towards the mesenchymal phenotype is present across all 3 cancers. However, SCLC appears to be unique in its utilization of both paracrine and autocrine signals to sustain its phenotypic state. Overall, our results show recurring roles of intercellular communications in maintaining newly formed cell states, and they shed light into non-cell-autonomous mechanisms of intratumoral heterogeneity.

Results

A multiscale inference approach to explore the signaling mechanisms maintaining phenotypic heterogeneity in cancer cell populations

To assess whether cell-cell communications contribute to intratumoral heterogeneity and phenotypic plasticity it is essential to connect intercellular signaling with downstream intracellular processes and determine their overall impact on maintaining phenotypic diversity. To accomplish this, we adapted a methodology that integrates various approaches to explore both intercellular and intracellular signaling events (see Methods, Fig. 1C, Supplementary Fig. S1, and reference [28]). As a first step, we used subtype labels provided by the original data producers [21, 30, 31] to classify cells along epithelial and mesenchymal axes. These preassigned annotations formed the foundation for evaluating subtype-specific signaling dynamics across datasets. We then infer the active signaling pathways and ligand–receptor (L–R) interactions to assess intracellular signaling from processed scRNA-seq data. For the inference of cell-cell communication, we opted for *CellChat* [32] due to its robustness to noise and its capacity to incorporate heteromeric complexes, which has been shown to reduce false-positive predictions [33] . Additionally, *CellChat* is one of the fastest cell communication inference tools and has been demonstrated to achieve higher specificity compared to other tools, meaning that it is less likely to infer spurious interactions not supported by data [34] .

The inferred active signaling pathways are then used to predict transcriptional activity and identify the relative strength of active signaling pathways in each cell state. We then used an integer linear programming tool for contextualizing causal networks [35] to integrate the transcription factor (TF) activity inference scores, differential expression analysis, and L–R interactions. This approach finds the smallest-sign consistent network that explains the measured inputs and outputs, connecting the receptors to the downstream TFs. To aid in assessing the relationships between EMT states cellular differentiation, we incorporate downstream SCLC [21] and EMT [36] target genes in the network—adding specific target genes based on the congruence of the inferred regulatory modes between TF and target gene (i.e. activation or inhibition) and experimental values of gene expression for the target gene. This multiscale integration allows us to capture dynamic changes in signaling networks across different cell states and enables a detailed exploration of the potential signaling mechanisms involved in maintaining the phenotypic heterogeneity within the cancerous population.

We are interested in studying the interactions between and transitions among the epithelial and mesenchymal states in the SCLC RPM dataset, focusing primarily on timepoints 7 and 11 (TP7 & TP11) when both NE and non-NE subtypes are present in relatively high abundance (Fig. 1D). These time points also coincide with a transitional phase in the tumor's evolution, as the population begins shifting from a predominance of NE to non-NE phenotypes. Six subtypes had previously been defined via archetype analysis (SCLC-A2, -A/N, -P/Y, -Y, Generalists, and Unclassified [21]), with the SCLC-A2 subtype displaying more epithelial-like properties, while the SCLC-P/Y and SCLC-Y subtypes exhibiting more mesenchymal-like features [24] . Archetype analysis is an approach based on Pareto optimality theory applied to cell specification that identifies phenotypic states at the extremes of high dimensional feature space [37] and applying stringent thresholds enriches phenotypes with distinct features and excludes cells with intermediate levels of features from two or more archetypes. Unclassified cells were those that were in close proximity to an archetype vertex but could not be confidently assigned to any SCLC subtype based on enrichment testing with known marker genes (see Methods), whereas Generalists were cells with no clear proximity to any archetype vertex, and they were excluded from downstream subtype-specific analyses. For Generalists, we performed additional

robustness tests with various proximity thresholds and the results will be described later in this subsection.

We then extended the multiscale approach to three additional datasets involving cancers undergoing EMT: SC53 (Human SCLC circulating tumor cells-derived xenograft sample [25]), HCT116 colon cancer cell line [30] and a HER2 Crainbow mouse [31] (Fig. 1E & Supplementary Fig. S2). The cell type classifications in these datasets were determined by the authors who generated the data. In SC53 there are four identified subtypes: SCLC-A, -A2, -Y and Generalists. Notably, the SCLC-Y subtype exhibited gene expression profiles consistent with a more mesenchymal-like state. In the colon cancer dataset, three EMT-associated states were characterized: epithelial (Epi), mesenchymal (Mes), and partial-EMT (pEMT) [30] . The HER2 dataset has four cell-states that were inferred using trajectory analysis: hormone-sensitive (HS), hormone-receptor negative (HR-), EMT and a transitional (T) state [31] .

Given the different strategies of cell type annotation across the datasets, we assessed whether each dataset contained distinguishable epithelial- and mesenchymal-like populations across the different datasets. Through gene set enrichment analysis methods, we revealed a clear separation of epithelial and mesenchymal populations across the different datasets (Fig. 1F). To further quantify this separation, we computed silhouette scores and Euclidean distances between centroids of the most epithelial- and most mesenchymal-like annotated cell types in each dataset. These metrics confirmed the presence of distinct epithelial- and mesenchymal-like clusters (Supplementary Fig. S3). Additionally, given the presence of a substantial number of Generalist cells in the SCLC RPM dataset, we explored how adjusting the archetype assignment threshold affected cell annotation. Lowering the threshold from 0.80 to 0.70 enabled reclassification of some Generalist cells into more defined subtypes, without substantially altering the overall clustering structure (Supplementary Fig. S4). We also examined lower thresholds, but these were not used in downstream analyses, as they resulted in substantial shifts in cell annotations, such as cells initially labeled as Unclassified were reassigned to the SCLC-P subtype (Supplementary Fig. S4).

Diverse signaling pathways converge on activating key mesenchymal pathways/genes

We first analyzed the RPM dataset and identified 48 active signaling pathways across the six different timepoints (Supplementary File S1). Many of these pathways are well-established contributors to EMT in cancer [27, 38–59] (Fig. 2A). Notably, we observed a recurring pattern among these EMT pathways wherein they converge towards the mesenchymal non-NE cell types (SCLC-P/Y in Fig. 2B). This convergence is mediated by paracrine (NOTCH) signaling from NE (E-like) to non-NE (M-like) cells and autocrine signaling within the non-NE cell population (WNT and SPP1). To assess whether these inferred signaling patterns reflect true biological structure rather than arising from cell-type proportions or noise, we constructed a null model by randomly shuffling subtype labels. Across 25 iterations, only a single significant L-R interaction was identified, whereas the real model identified 383 interactions. This supports that the observed communication network is specifically structured by the true subtype organization (Supplementary Fig. S5). In addition, we observed that autocrine and paracrine signals are also received by NE cells according to our data analysis pipeline (Fig. 2B). In this subsection, we will focus on the signaling effects on the non-NE or M-like cells in SCLC and other cancer types. The analysis of the signals that non-NE cells receive will be presented in detail in a later subsection.

To elucidate how intercellular communication affects intracellular signaling, we constructed a signaling network for the Day 7 SCLC-P/Y cell type. This network identified TGFβ and NOTCH as primary ligands and captured key lineage-supporting target genes (Fig. 2C). Within the network, we observed the activation of numerous mesenchymal markers (like vimentin, PDGF-C and Axl) and the inhibition of epithelial markers, especially Epcam. Consistent with previous experimental findings [16] we noted that the activation of *Myc* and Notch signaling promotes the non-NE SCLC fate. Our network further highlights that the inhibition of *Ascl1*—a NE and SCLC-A & -A2 subtype marker downstream of the *Notch2* receptor—and the activation of *Myc* facilitates the upregulation of mesenchymal markers. Notably, the ligand–receptor interactions captured in the network are well known EMT modulators. Since the ligands involved in the activation of mesenchymal markers within the SCLC-P/Y network originate from SCLC-A2, -A/N and -P/Y cells, our network underscores the contributions of both NE and non-NE cells toward mesenchymal transitions. This approach yielded similar results in the TP11 SCLC-P/Y cell type by capturing the activation of lineage-supporting genes, with both NE and non-NE cells playing a role in activating those genes (Supplementary Fig. S6). Additionally, to further assess the robustness of our network results, we decreased the assignment threshold in the archetype analysis to assign more Generalist cells into one of the SCLC subtype archetypes. While this increased the total network size for the SCLC-P/Y group at TP7, the expanded network still captured a majority of the same nodes as the original network. This indicates that the inclusion of more Generalist cells does not substantially alter the overall network structure (Supplementary Fig. S7).

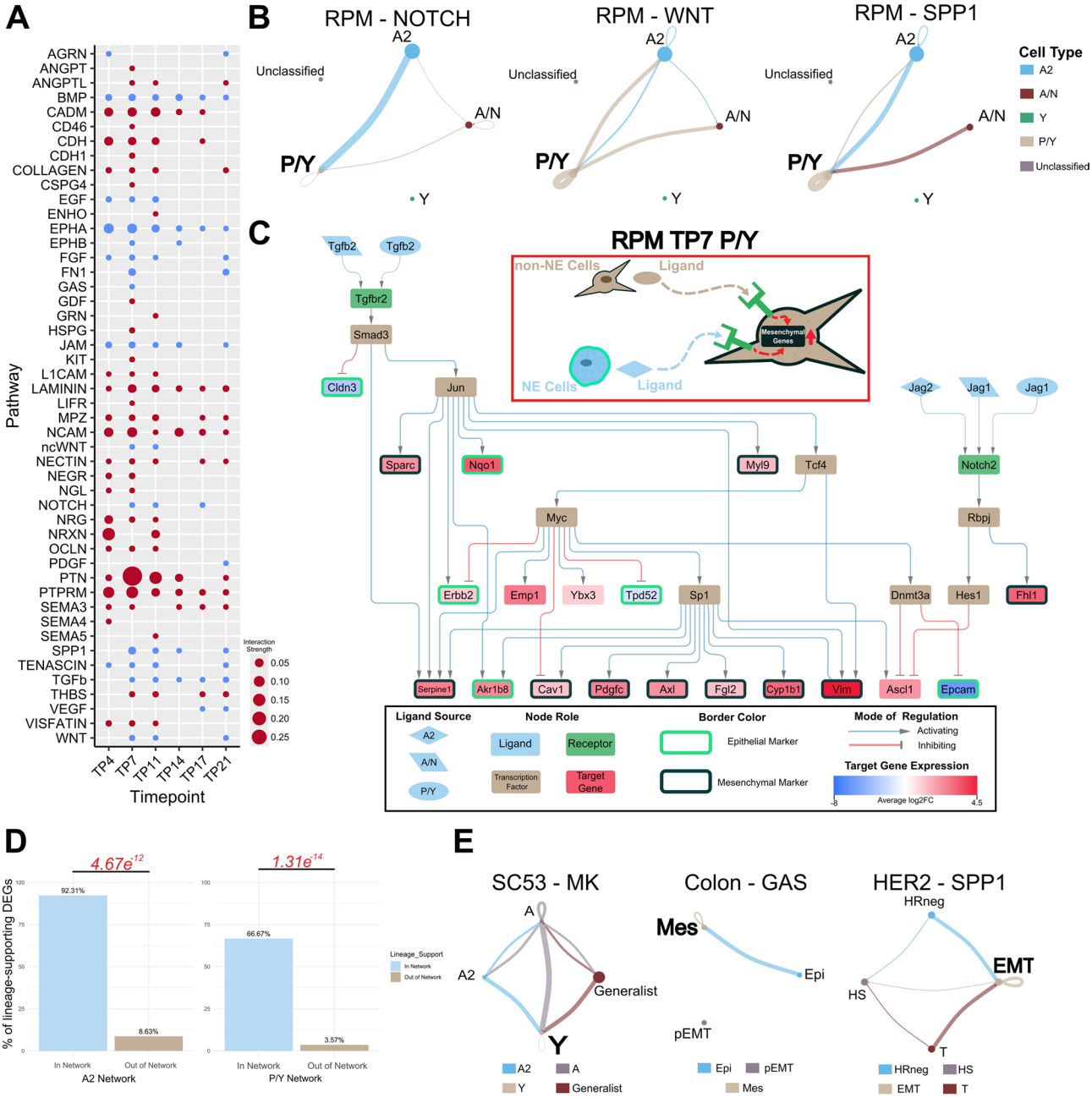

Fig. 2 Mesenchymal Phenotypes Utilize Autocrine and Paracrine Signaling to Reinforce SCLC Subtypes. **A** Inferred active pathways in the SCLC RPM dataset. Dot size is representative of the relative interaction strength for a given timepoint. Blue circles are pathways known to be involved in EMT signaling. **B** Inferred *CellChat* signaling of pathways known to be involved in EMT at TP7. Dot size is proportional to cell number. The legend on the right contains the cell type each dot represents. Line color represents the source of signal/ligand. Line width represents interaction strength. **C** RPM TP7 SCLC-P/Y cell type inferred signaling network connecting both intercellular and intracellular signaling. The red box contains an illustrative summary of the network. The black box is a legend for the network. Node color represents the role in the signaling pathway. The shape of the ligand nodes represents the subtype source of the ligand. The edges connecting a ligand to a receptor vary in width corresponding to interaction strength. Average log2 fold-change of target gene expression is colored as indicated by the scale bar. Epithelial target genes are indicated by a green border around the node and mesenchymal genes have a black border. **D** Percentage of lineage-supporting DEGs among DEGs that are either included in the inferred network (In Network) or excluded (Out of Network) for each SCLC subtype. Fisher's exact test was used to calculate significance and P-values are labeled in red. See Supplementary Table S1 for full statistical analysis. **E** Inferred *CellChat* signaling of EMT pathways from other datasets. Left: Human SCLC CDX. Subtypes are characterized through archetype analysis. Middle: HCT116 colon cancer. Subtype labels are Epi (Epithelial), Mes (Mesenchymal), and pEMT (Partial EMT) and they are from the original publication. Left: HER2 breast cancer mouse. Subtype labels are HS (Hormone-sensitive), HR-negative (Hormone-receptor negative), T (Transitional state), and EMT. The more mesenchymal states in these datasets are in bold text and are Y, Mes, and EMT, respectively

We then assessed the importance of lineage supporting genes captured in the networks by creating a 2×2 contingency table and determining the ratios of lineage supporting genes included in the inferred signaling networks versus those excluded from the network. We found a significantly higher proportion of lineage supporting genes within the network compared to outside the network (Fig. 2D, Supplementary Table S1), suggesting that the inferred intercellular communication driving the mesenchymal (M) state transcriptional program was not simply due to the random selection of broadly upregulated M genes.

To assess the generalizability of these findings across other cancers undergoing EMT, we performed a similar analysis on three additional scRNA-seq datasets. We observed the activation of intercellular communication-driven mesenchymal pathways converging towards a more mesenchymal phenotype in all three datasets (Fig. 2E, Supplementary Fig. S8A and Supplementary Table S2), whereas the signaling effects on epithelial-like cells are less clear (discussed in a later subsection 'SCLC utilizes autocrine and paracrine signaling to maintain epithelial state'). Nonetheless, application of the signaling network pipeline to the colon and HER2 datasets yielded lineage-stabilizing networks, albeit less comprehensive compared to those observed in SCLC (Supplementary Fig. S8A and Supplementary Table S2). In addition, in the HER2 dataset, we observed that fibroblasts were dominant contributors to many signaling pathways (Supplementary Fig. S8B), highlighting the potentially important role of stromal components in shaping the signaling landscape. These findings suggest that in some tumor contexts, phenotypic heterogeneity may be more strongly influenced by signals originating from the tumor microenvironment than by direct communication between cancer cells. In contrast, within SCLC, direct intercellular communication between NE and non-NE phenotypes appears to play a prominent role in reinforcing the non-NE/mesenchymal (SCLC-P/Y or -Y) state.

A mathematical model predicts the roles of inter-subtype feedback in stabilizing phenotypic compositions of heterogeneous tumor cell populations

Analysis of published data showed that the co-existence of epithelial-(E-)like cells and mesenchymal-(M-)like cancer cell subpopulations is a recurring pattern in multiple cancer types (Fig. 1F). Through our inference pipeline in this work, we also found a type of intercellular communication shared by multiple cancer types and datasets: cells at the state enriched with M genes receive signals from cells at a more E-like state, and the signals were used to maintain the recently acquired M state (Fig. 2). Previous work has shown the functions of autocrine signaling in maintaining cell states [60, 61] but it is unclear how the inter-subtype signals that we found in multiple contexts can influence intratumoral population dynamics. We therefore built a simple mathematical model based on the following dimensionless ordinary differential equations (ODEs)

$$\begin{aligned} dX_1/dt &= r_1 X_1 - k_1 X_1 + 1/\left(1 + (w_1/K_2)^{n_2}\right) k_2 X_2 \\ dX_2/dt &= r_2 X_2 + k_1 X_1 - 1/\left(1 + (w_1/K_2)^{n_2}\right) k_2 X_2 \end{aligned} \quad (1)$$

which depict two subpopulations (X_1 and X_2) with proliferation rate constants r_1 and r_2 respectively, and basal interconversion rate constants (k_1 and k_2 for X_1-to-X_2 conversion and X_2-to-X_1 conversion, respectively) (Fig. 3A). The signal that originates from X_1 and received by X_2 is modeled by an inhibitory Hill function that influences the overall conversion rate from X_2 to X_1, as the inferred inter- and intra-cellular activities common to all cancer types. Parameter K_2 determines the threshold of the inactivation and n_2 determines the nonlinearity of the response. w_1 is the abundance of X_1 relative to the total size of the population (i.e. $w_1 = X_1/(X_1 + X_2)$. This Hill function effectively serves as a feedback mechanism that may influence the dynamics of the cancer cell population. Simply, the model describes the dynamics of a cell population with two discrete states (of any type) with interactions between them. Note that this model focuses on cell population dynamics in the timescale of hours and days rather than single-cell behaviors subject to significant stochasticity. The time evolution of cell fractions was previously shown to be gradual and can be appropriately described by deterministic models similar to our framework [62] .

We simulated the model with an initial population containing X_1 but not X_2 with a parameter set adjusted such that the steady state ratio of the two populations is approximately 1:1 (Fig. 3B, red), a ratio consistent with experimentally observed NE to non-NE transitions of SCLC cells and the average non-NE fraction of human SCLC tumors [15, 20] (See Methods). We found that removing the feedback by increasing threshold parameter K_2 to a very large value resulted in both a longer time to an equilibrium of two subpopulations and a lower steady state fraction of X_2 due to the increased number of cells converting from X_2 to X_1 (Fig. 3B, black). The shorter time to reach equilibrium with the population-level feedback may be related to a previously established function of negative feedback loop in accelerating responses [63] . While lowered X_2 fraction might suggest a role of the feedback signaling (Fig. 3A, red) in facilitating the formation of mesenchymal-like X_2 cell population, one can argue that this advantage can be simply achieved by adjusting the basal conversion rate constant k_2 or other parameters. We therefore tested another feedback-free model in which there was a compensatory reduction

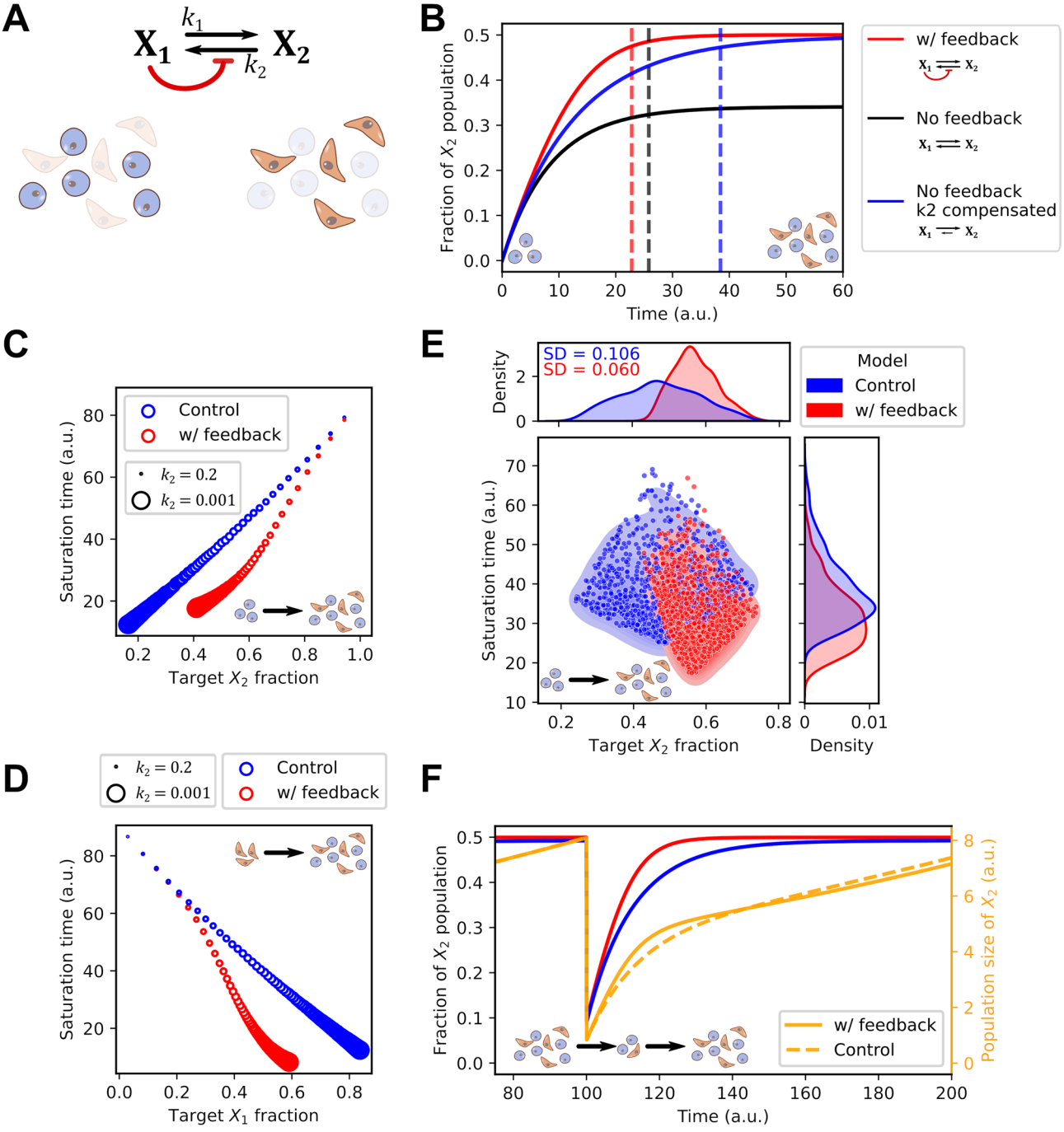

Fig. 3 Simulations with an ODE model for inter-subtype communication. **A** Network diagram and illustration of two cell types (Epithelial-like and Mesenchymal-like). X_1 and X_2 are two state variables representing the sizes of the two cell populations, respectively. Black arrows show transitions. Red arrow shows the intercellular communication responsible for maintaining the M-like cell state. **B** Time course trajectories (solid curves) for the main model and two perturbed models. Y-axis shows the quantity $X_2/(X_1+X_2)$. Dashed vertical lines show positions of saturation times, defined as the time at which 95% of the steady state level of the fraction is reached. Initial conditions: $X_1=10$, and $X_2=0$. See Methods for parameter values. **C** Saturation time as a function of steady state fraction of X_2. Color code is the same as Panel B. 100 evenly spaced k_2 values in the indicated range were chosen to achieve various steady state X_2 fractions for each model. Circle sizes indicate k_2 values. Initial conditions are parameter values are the same as B. **D** Same simulations as in C except for initial conditions. Here, $X_1=0$, and $X_2=10$. **E** Distributions of saturation time and steady state X_2 fraction of two models with 1000 sets of randomly chosen values for all parameters with 50% decrease/increase from the basal values as in B. **F** Time course trajectories of a fractional killing scenario in which the fractions of the two populations reached steady state and 90% of X_2 cells were removed at Time 100. The orange lines are the total population size and the red and blue lines represent X_2 fraction from the same models shown in *B*. The control model (blue and dashed orange lines) is the no-feedback model with compensated k_2

of k_2 (Fig. 3B, blue) that produced the same 1:1 steady state ratio for the two types of cells. We found that the model with the inter-subtype signaling needed a significantly shorter time to achieve the equilibrium of the two subpopulations compared to the perturbed model that achieved the same level of heterogeneity (approximately three days shorter than control. See Methods.). This confirms the importance of the cell communication-mediated feedback, rather than merely kinetic rate constants that influence cell fractions, in controlling re-equilibrium time. We found that this acceleration was consistent in a range of steady state (target) ratios of the two subpopulations, but it was more prominent when the two subpopulations were comparable in size (Fig. 3C). Gopal et al. observed that a re-equilibrium of multiple SCLC subtypes with comparable fractions can be achieved within 7 days, a time window significantly smaller than that for doubling tumor cell populations [22, 64] suggesting an importance of timely equilibrium. Lim et al. made a similar observation [15] . The relatively short time to diverse populations may contribute to SCLC's aggressive and recalcitrant properties [21] . In addition to the acceleration of equilibrium, with the same range of the scanned k_2 values, the model with the feedback had a narrower range of steady state compositions of the two populations (Fig. 3C). This suggests that the inter-subtype feedback can be helpful to support desired phenotypic compositions in facing variations of kinetic rate constants. Interestingly, both acceleration and stabilization properties of the feedback were observed when we simulated a heterogeneity restoration from a purely M-like population that reflects a biologically plausible mesenchymal-to-epithelial transition (or a non-NE to NE transition in SCLC) (Fig. 3D). Next, we asked whether cell-composition stabilization is a general behavior when all kinetic rate constants are perturbed simultaneously. We generated 1000 parameter sets, each with randomly chosen parameter values from a uniform distribution bounded by ± 50% deviations from the basal ones (Fig. 3E). We found that the model with inter-subtype feedback had a smaller variation of phenotypic composition compared with the control model with respect to the global perturbation of parameters (control standard deviation (SD): 0.106, SD with feedback: 0.06) (Fig. 3E, top). The feedback model also had shorter saturation time for heterogeneity restoration, as expected ($p < 10^{-15}$, Mann-Whitney U test) (Fig. 3E, right).

We hypothesized that the acceleration function of the feedback loop may help the tumor cell population to re-establish a heterogeneous population under a subtype-specific "fractional killing" condition, i.e., a depletion of one of the subtypes. Indeed, after we perturbed a population already at equilibrium with equal numbers of each state (X_1:X_2 ratio of 1:1) by removing 90% of X_2 cells from the system, the model with feedback recovered more rapidly compared to the feedback-free model (Fig. 3F). Taken together, our mathematical model suggests two population-level functions of inter-subtype communication that we inferred from single-cell data: it accelerates the acquisition of heterogeneous tumor cell population from either a relatively homogeneous initial or out-of-equilibrium population, and it renders lower sensitivity of the steady state phenotypic compositions with respect to the perturbations of kinetic rate constants.

SCLC utilizes autocrine and paracrine signaling to maintain epithelial state

In our analysis of the SCLC RPM dataset, we found that epithelial-like (NE) cells receive some paracrine and autocrine signals (e.g. Figure 2B), but it was not clear whether related intracellular signaling programs may be affecting the phenotype. We therefore asked how the cross-talk between the different phenotypes affects the epithelial state. In the RPM dataset, our analysis revealed that the epithelial state is sustained through a combination of paracrine and autocrine signaling mechanisms. We applied our analytical pipeline to construct a network for the Day 7 epithelial, SCLC-A2 subtype (Fig. 4A). The network captured the activation of several epithelial marker genes. We observed the activation of the epithelial marker *Cdh1* in both the Day 7 and Day 11 network (Supplementary Fig. S9A). Additionally, the Day 7 network showed the activation of the SCLC-A2 marker, *Ascl1*. Notably, we identified *Sp1* as a key transcription factor involved in the activation of several epithelial markers. Interestingly, *Sp1* was also present in the SCLC-P/Y network, suggesting its potential to influence both NE and non-NE cell fate determination.

Similar to observations in the SCLC-P/Y network, the SCLC-A2 network revealed the participation of both NE and non-NE cells in maintaining the epithelial state within SCLC-A2 cells. This interplay is also evident in the other inferred signaling pathways (Fig. 4B). Specifically, the inferred interaction involving *Jam3* ligands within the JAM signaling pathway is consistent with prior work demonstrating *Jam3*'s role in establishing the epithelial phenotype [65] . Furthermore, we identified the CDH1 pathway as being activated by autocrine signaling within SCLC-A2 cells. The presence of CDH1 and JAM signaling pathways was also observed in human SCLC SC53 samples, operating in a manner similar to our findings in the SCLC-A2 network (Fig. 4C and Supplementary Fig. S9B). This suggests that the maintenance of the SCLC epithelial state involves signaling interactions between epithelial and mesenchymal cells.

When applying the signaling network pipeline to the epithelial states of the other cancers, a network could only be generated for the colon cancer epithelial state but

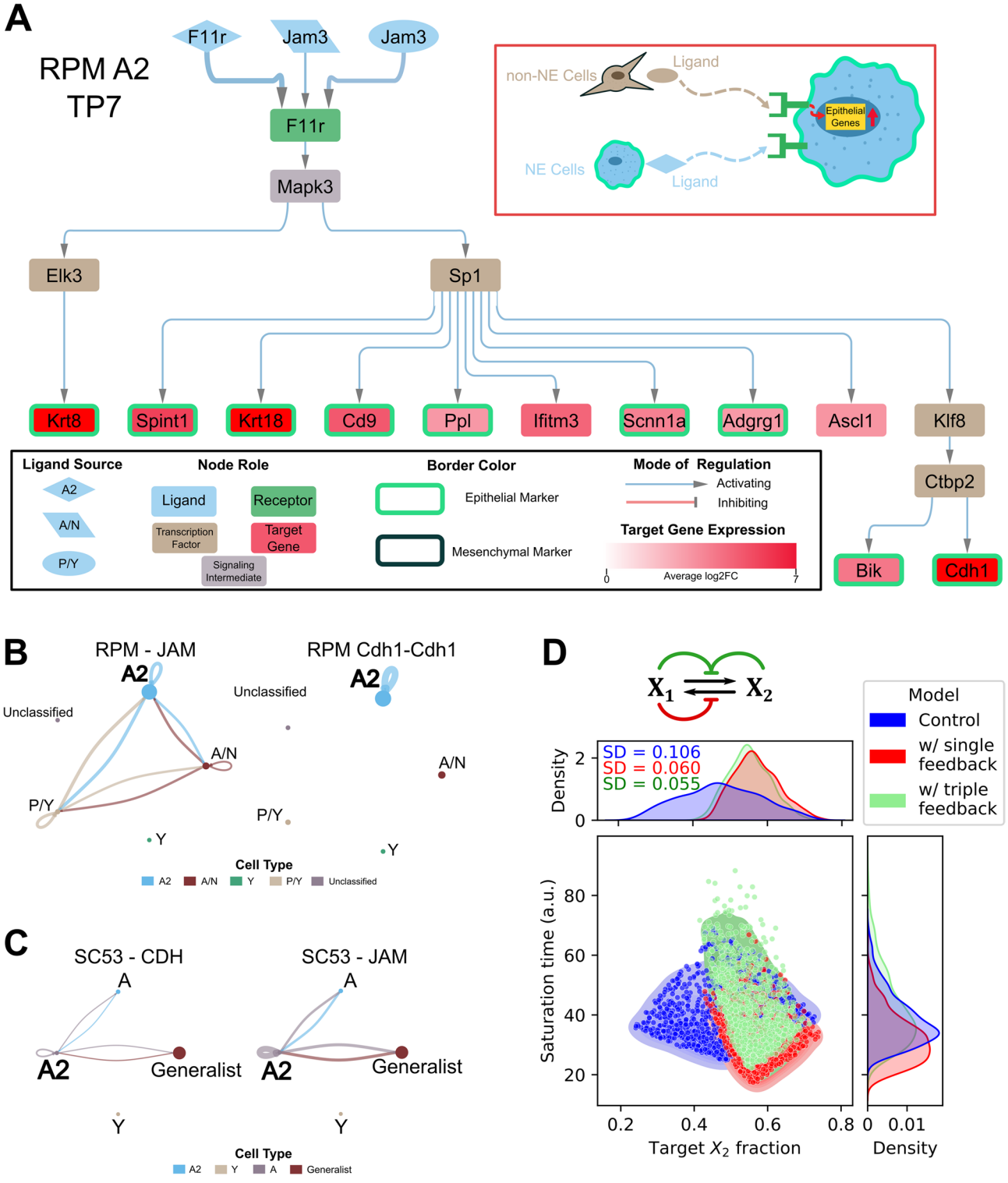

Fig. 4 Epithelial Phenotype is Maintained by Autocrine and Paracrine Signaling. **A** RPM TP7 SCLC-A2 cell type inferred signaling network. The inset indicated with a red rectangle is an illustrative summary of the network. The legend of the network is shown in the inset black rectangle. **B** From the RPM data, the inferred JAM signaling pathway is shown on the left and the inferred L-R interaction of CDH1 is shown on the right. The Cell Type legend on the bottom contains the cell type each dot represents. **C** Inferred CDH signaling pathway (left) and JAM pathway (right) from the human SCLC SC53 dataset. **D** Distributions of saturation time and steady state X_2 fraction of models with 3 types of feedback (red and green arrows) in terms of the saturation time and phenotypic composition (measured as fractions of X_2 cells). Results from the triple-feedback model are shown in green. A single-feedback model containing the E-to-M signal only (red) and a feedback-free model (blue) were used for comparison. 1000 random parameter sets were used for each model (see Methods for basal parameter values and ranges). Initial conditions: $X_1 = 10$, and $X_2 = 0$

not the HER2 epithelial states (Supplementary Fig. S8A). The colon cancer network does not capture the activation of any of the overexpressed epithelial genes present within this cell type. Similarly, for the HER2 epithelial states, a causal inference network could not be generated from the receptors to transcription factors. These suggest that cell–cell communication does not contribute significantly to the development or maintenance of the epithelial states in these systems. It is possible, however, that additional cell types not present in the data are responsible for regulating the epithelial state, which could explain the less comprehensive nature of the colon cancer network and the inability to generate a network for the HER2 epithelial state by our approach. This highlights the different levels of influence that cell–cell communication may play in maintaining the intratumoral heterogeneity, with the epithelial SCLC state being more sensitive to these signals. Nonetheless, we examined the potential roles of the SCLC-specific autocrine and paracrine signaling at the cell population level with mathematical modeling. We built a triple-feedback model that incorporates three types of cell communications (Fig. 4D, top). With 1000 sets of randomly generated parameter sets each of the triple-feedback and benchmark models (i.e. the feedback-free model and the single-feedback model described in Fig. 3), we found that signals that the NE cells (E-like cells) received produced the narrowest distribution of the phenotypic composition among the three models when parameters were perturbed globally (SD with single-feedback: 0.06. SD with triple-feedback: 0.055) (Fig. 4D, top). This indicates that these forms of communication can further enhance the robustness of phenotypic composition. However, the feedback on NE cells dampened the acceleration of the re-equilibrium compared to the single-feedback model ($p < 10^{-13}$, Mann-Whitney U test) (Fig. 4D. right). This suggests that the system has a tradeoff between the speed and the accuracy (subpopulation fractions) of (re-)establishment of cellular heterogeneity. Overall, we found that all types of cell communication-based feedback inferred from single-cell data had significant effects on the equilibrium of the heterogeneous cancer cell populations.

To test the robustness of our conclusions on the roles of feedback in adjusting the saturation time and the variations of the phenotypic compositions, we performed sensitivity analysis by systematically scanning each parameter over a wide range (See Supplementary Text and Supplementary Table S3). We found that our observations are consistent in a large parameter region (Supplementary Fig. S13): the feedback control of the non-NE-to-NE transition (red arrow in Fig. 4D) reduces both saturation time and variability of phenotypic compositions, whereas the feedback on the NE-to-non-NE transition (green arrows in Fig. 4D) reduces variability of phenotypic compositions at the expense of slowing down the re-equilibrium. Additionally, our parameter scanning with the strength of each feedback loop showed that the tradeoff between accelerating re-equilibrium and stabilizing phenotypic compositions was manifested when we varied the strength of the paracrine feedback driven by non-NE cells, whereas the autocrine-driven feedback played a negative effect on both functions (Supplementary Fig. S14). Our results suggest that the three types of feedback have distinct roles in controlling the dynamics of the re-equilibrium of heterogeneous cancer cell populations.

Discussion

Elucidating the dynamics and mechanisms that govern phenotypic plasticity within cancer tumors is essential for developing therapeutic strategies and tackling two major unresolved clinical challenges: cancer metastasis and therapeutic resistance [10, 11] . While considerable progress has been made in characterizing phenotypic plasticity at the gene expression level [12, 60, 61, 66] many aspects remain poorly understood. Identifying the mechanisms that drive intratumoral heterogeneity and regulate phenotypic plasticity is a critical step for effective cancer treatment, as different cell types within a tumor can respond differently to therapies [22, 67–71] . In this study, we investigated whether intercellular communications play a role in controlling cell fate transitions and whether these interactions stabilize or destabilize cellular phenotypes. We applied a multiscale inference-based approach to different solid tumor scRNA-seq datasets to investigate the crosstalk between cell states and how they influence one another. Within SCLC, we found the pivotal role of intercellular signaling in maintaining the phenotypic diversity among the cancer cell population, particularly in the context of EMT. The inferred SCLC-P/Y signaling network captures the activation of both *Myc and* Notch signaling, which is consistent with recent observations as the activation of these two components is seen within the non-NE subtypes [15, 16] . Additionally, the networks capture the mesenchymal nature of SCLC-P/Y subtype [24] as the activation of many differentially overexpressed mesenchymal markers are present within the network. These findings support the view that Notch acts as a mesenchymal-promoting signal, reinforcing the non-NE phenotype. In contrast, the SCLC-A2 subtype displayed signaling patterns suggestive of epithelial state maintenance, characterized by *Jam* family signaling. This is consistent with literature showing that *JAM3* promotes epithelial integrity by facilitating tight junction formation and suppressing migratory phenotypes [65] . In our network models, both autocrine and paracrine *Jam* signaling appear to reinforce epithelial identity in SCLC-A2 cells, while SCLC-P/Y cells received reinforcing

Notch-mediated mesenchymal cues from both SCLC-P/Y and -A2 sources. Among several cancer types that we analyzed, this epithelial state maintenance mechanism appears to be unique to SCLC and specific to the SCLC-A2 subtype. These inferred signaling pathways may also represent therapeutic targets to modulate plasticity. A recent study showed that treatment with an FGFR inhibitor induced a transition toward NE-like states in SCLC [72] . This finding illustrates how targeting specific signaling pathways can influence lineage plasticity in SCLC and perhaps disrupting autocrine or paracrine signals, such as Notch of Jam3, could offer therapeutic strategies to reshape phenotypic equilibria within tumors.

Applying this multiscale methodology to the colon and HER2 cancer datasets yielded less-comprehensive networks. One possible explanation for this difference is that the tumor microenvironment (apart from tumor cell heterogeneity) may play a larger role in influencing phenotypic plasticity in colon and HER2 breast cancer. In all three datasets, our primary analysis focused on intercellular communications within the cancerous population. In SCLC, the cell–cell communication between the cancerous cells appears sufficient enough to influence the cellular phenotypes. However, this does not appear to be the case with colon and HER2 cancers, where other factors in the tumor microenvironment may have a greater impact on tumor cell plasticity [73, 74] . In the HER2 dataset, for example, fibroblasts emerged as dominant contributors to key signaling pathways, suggesting a stronger influence of stromal-derived cues (Supplementary Fig. S8B). This dataset was derived from an in vivo setting, where interactions with immune, stromal, and endothelial cells may likely shape tumor behavior more substantially. Together, these findings highlight how the relative contribution of tumor cell-intrinsic versus microenvironmental signaling can vary across tumor types and biological contexts. Future work incorporating systematic modeling of both malignant and non-malignant compartments will be crucial to fully elucidate the signaling dynamics governing phenotypic plasticity.

Due to the feasibility of identifying transcriptional programs for E-like and M-like cell states, we focused on a 2-subtype (E and M) framework for both our intracellular transcriptional program inference and our mathematical models. In many cancer types, the heterogeneous population containing three or more states are observed, and the potential role of their crosstalks in lineage stabilization can be analyzed with our approach. One possible extension of our work is a model with E, M and hybrid E/M states (in SCLC, they correspond to A2, P/Y and A/N respectively). However, a comprehensive list of genes involved in the transcriptional program of the hybrid state is not very well defined. A systematic analysis of the influence on lineage dynamics and modeling are therefore difficult with our current framework. Nonetheless, we built a 3-subtype, scRNA-seq data-informed model that incorporates the feedback that originates from the hybrid state (Supplementary Fig. S15). This model confirmed the role of M-to-E signaling in accelerating the re-equilibrium revealed by the 2-subtype model, and it further suggests similar roles of the feedback driven by the hybrid state (Supplementary Fig. S15 and Table S4). Recent research has started to reveal some factors that (de-)stabilize the hybrid E/M state [75] . Future models with more molecular details of the hybrid state will help to dissect the effects of individual feedback loops in a more comprehensive manner. Another potential future development of the model is the inclusion of signaling and transcriptional networks within cells; this would enable modeling complex features such as EMT as a continuum and provide a bridge between our multiscale data analysis pipeline and the model output. We expect that other modeling strategies—such as stochastic differential equations, Boolean networks, agent-based modeling or combinations of these—will uncover new insights regarding the cell communication-driven dynamics of cancer cell subpopulations. Finally, the predictions made by our models about the role of cell communication–mediated feedback in tumor cell population dynamics should be tested in future experiments. However, analyzing tumor cell population dynamics at the single-cell level is especially challenging, especially when considering cell state transitions and cell death and division events. Performing these experiments ex vivo in rudimentary ways may be possible, although quantitative assessment of more than a few molecular species over time in live cells remains a significant challenge and would limit any broad molecular analyses (e.g., single-cell RNA-seq) to fixed end points.

It is possible that technical differences in how tumor cells are classified may contribute to the weaker signaling networks observed in the colon and HER2 datasets. Our pipeline is robust to different cell type annotation strategies within a dataset, as demonstrated in SCLC where both archetype-based and Leiden-clustering based annotations yielded consistent signaling patterns (Supplementary Fig. S10). However, the initial structure and diversity of a dataset may still influence the depth and resolution of the inferred networks. In the colon cancer dataset, $EpCAM^{high}$ and $EpCAM^{low}$ populations were FACS-sorted and merged, and cell type annotation was performed based on EpCAM expression and epithelial/mesenchymal scores derived from gene set enrichment. While this approach captures a broad epithelial spectrum, reliance on a limited set of sorting markers and downstream scores may constrain the granularity of cell state definitions. Conversely, in the HER2 dataset, cell states were defined based on terminal branches from

pseudotime trajectory analysis, which may obscure intermediate states. In addition to annotation-based limitations, technical noise inherent to scRNA-seq can further reduce the detectability of lowly expressed genes, including key ligands, receptors and TFs [76] . Such constraints in both data acquisition and expression resolution can compress biological variance and limit the granularity of detectable signaling. Therefore, while our pipeline is flexible with respect to subtype definition, it still relies on sufficient expression structure and transcriptomic depth, which may be more limited in the colon and HER2 datasets.

Phenotypic diversity in cell populations can be supported by both autonomous and non-autonomous mechanisms. Autonomous mechanisms include intrinsic transcriptional fluctuations, which can stochastically initiate phenotypic transitions [60] as well as multi-stable regulatory networks, where cells can switch between stable phenotypic states based on underlying network architectures [77] . Additionally, post-transcriptional mechanisms provide another layer of intrinsic regulation [78] . However, while these autonomous processes can trigger phenotypic transition, it may be difficult to maintain this phenotypic diversity over time. Therefore, non-autonomous mechanisms may be required, such as paracrine/autocrine signaling which can play a critical role in reinforcing phenotypic heterogeneity through intercellular communication [79–81] . Our findings highlight the importance of these non-autonomous signaling mechanisms, demonstrating that cell–cell interactions can be essential for sustaining the intratumoral phenotypic heterogeneity. While our population-level models provide a quantitative framework for describing these roles of intercellular communications based on ligand-receptor interactions, other forms of intercellular interactions, such as resource competition and density-dependent effects on growth rates of subpopulations, can also influence dynamics of phenotypic heterogeneity, as described in recent work [82] . Future modeling work is therefore needed to dissect the roles of multiple types of interactions in heterogeneity recovery of cancer cell populations.

Cell–cell interactions are fundamental to shaping and maintaining multicellular structures and tissue integrity. Among these, cell–cell communications play a pivotal role in coordinating cellular behavior across short and long distances within tissues. Here, we inferred cell–cell communications from scRNA-seq data using *CellChat*. While these methods lack spatial context, the limitation arises from the input data rather than the inference framework itself. *CellChat* includes a comprehensive reference database with cytokines and long-range signaling molecules, enabling recovery of both short- and long-range interactions. Although spatial resolution is absent in standard scRNA-seq, previous studies have shown that transcriptomic inference methods can still capture biologically meaningful paracrine signaling. For instance, WNT and TGF-β signaling predicted from scRNA-seq were validated in colonic stem cells [74] and IL6-IL6R-mediated signaling was confirmed in non-small cell lung cancer [83] . These findings suggest that when expression signals are sufficiently strong, long-range communication can be inferred even without spatial data. Nonetheless, transcript-based predictions rely on mRNA levels as proxies for protein activity and are thus susceptible to false positives [84, 85] . CellChat mitigates this by incorporating multimeric L-R complexes and cofactors to improve specificity and reduce false positives [33] but validation remains essential. Similarly, *CORNETO* integrates these inferred interactions with prior knowledge to reconstruct causal signaling pathways, but its accuracy depends on upstream inference quality and completeness of the prior network. These limitations emphasize the importance of integrating spatial transcriptomics and experimental validation. Integrating spatial transcriptomics with scRNA-seq offers a promising avenue to overcome these limitations, as it preserves the spatial arrangement of cells, allowing for more accurate inference of both short- and long-range communications [86, 87] . Future work will benefit from leveraging such integrative datasets and complementary validation strategies as it will refine our understanding of the role of cell–cell communications in maintaining intratumoral heterogeneity within tumors. In addition, this work focused on signaling with ligand-receptor interactions, which can in turn contribute to nonhomogeneous spatial patterns at the tissue scale [88] . It will be of interest to study the reciprocity of cell-cell communications and tissue patterning in future modeling work.

Intercellular signaling is known to contribute to the intratumoral heterogeneity within SCLC, with the activation of Notch signaling resulting in NE to non-NE cell fate switching in 10–50% of tumor cells [15] . Additionally, the non-NE subtype exhibits a reduced proliferative rate but relatively greater chemoresistance, and these cells support the growth and survival of the NE subtype within admixed tumors [18] . This dynamic interplay between NE and non-NE subtypes highlights the role of intercellular signaling in maintaining a functional heterogeneity that benefits tumor survival and progression. A recent study suggests that SCLC subtypes not only coexist but may actively cooperate to optimize essential tumor functions, with NE and non-NE cells interacting in mutually beneficial ways to foster tumor growth and adapt to changing external conditions, such as treatment [21] . Additionally, it has been suggested that non-genetic mechanisms, such as cell-cell interactions between SCLC cell types, provides the capability for some tumors to

reemerge once therapy is withdrawn through commensal niche-like interactions, where one cell type fosters the growth or survival of another [22] . These cooperative interactions are critical for maintaining the phenotypic diversity needed for tumor adaptability. Importantly, the rate at which phenotypic heterogeneity reaches equilibrium is likely driven by such cooperative mechanisms, enabling tumors to rapidly adapt by leveraging the distinct but complementary functions of different cell types. Disrupting these signaling networks or undermining the cooperative interactions between subtypes could impair the tumor's ability to maintain this adaptability. thereby enhancing therapeutic efficacy.

We are only beginning to uncover the role in cancer of non-cell-autonomous signaling on phenotypic plasticity—a key driver of tumor progression and therapeutic resistance. Prior studies have defined transcriptional subtypes of SCLC and linked subtype programs to key hallmarks like increased cellular proliferation and evading immune destruction [21] . Our work builds on these foundations by advancing from static subtype classification to dynamic modeling of how intercellular communication contributes to cell state regulation. Our results indicate that SCLC tumor cells have significant responses to extracellular signals emanating from different tumor cell subtypes and that transitions to mesenchymal phenotypes (especially SCLC-P/Y) are enhanced by ligands released from both NE and non-NE sources. Additionally, and somewhat surprisingly, the more epithelial SCLC-A2 state is likewise stabilized by signals from both NE and non-NE sources. Our mathematical model of the dynamics of cell state heterogeneity equilibration suggests that feedback mechanisms dependent on intercellular signaling are important modulators of how quickly and how accurately equilibrium can be reestablished. Altogether, our results support an important role for intercellular communication in controlling the dynamics of establishing equilibrium of intrinsic tumor cell state heterogeneity.

Conclusions

Single cell-based inference approach reveals a key role of intercellular signaling in maintaining intratumoral heterogeneity. Cell-cell communications reinforce cell state transitions in multiple cancers at the single-cell level, and they support phenotypic heterogeneity at the population level by accelerating re-equilibrium and conferring robustness of population compositions.

Methods

Single-cell RNA-sequencing data

Single-cell RNA sequencing data were downloaded from Gene Expression Omnibus (GEO) at GSE149180 (RPM mouse tumor time course) [16] GSE138474 (Human SCLC CDX) [25] GSE154930 (HCT116 colon cancer) [30] and GSE152422 (HER2 breast cancer mouse isoform). RPM mouse tumor dataset was preprocessed as described by Groves et al. [21] Python package *Scanpy* (version 1.8.0) was used for filtering and normalization of total counts. Log-transformation was performed using the *log1p* function from the *Numpy* (version 1.17) package and scaling was done using *Scanpy*.

Human CDX data were preprocessed as described by Gay et al. [89] Cells were filtered to remove non-tumor cells. Only the SC53 tumors were used in this analysis. *Scanpy* was used to normalize the total counts by cell and the data was then log-transformed.

Cell type annotation for RPM and SC53 datasets was performed as described by Groves, et al. [21] Briefly, archetype analysis was applied to gene expression data, allowing for flexible characterization of the transcriptional landscape based on functional phenotypic features. This method approximates the cell phenotype space as a low-dimensional polytope that encapsulates the gene expression data [90] . The vertices of the multidimensional space represent archetypes, which are transcriptional programs optimized for specific biological tasks and correspond to the major SCLC subtypes (e.g., SCLC-A, -N, -P, -Y) [21] . Each cell is assigned an "archetype score", which quantifies how closely the cell's expression profile aligns with each archetype vertex [90] . Cells with an archetype score greater than 0.80 were labeled as members of that archetype. To assess the statistical significance of archetype assignments, we applied a permutation enrichment test (as implemented in the original codebase) to compute p-values for each cell-archetype pair. Cells not significantly associated with any archetype (after multiple testing correction, *Holm-Bonferroni*) were labeled as Unclassified. In the SCLC RPM dataset, some subtypes labels represent hybrid states (SCLC-A/N and SCLC-P/Y), indicating cells that exhibit transcriptional characteristics of two neighboring archetypes. There are also Generalist cells present in this dataset, which are cells that are not strongly associated with any of the defined archetypes, instead they occupy a more central position within the archetype space. In our analysis, we used the original 0.80 archetype score threshold, as defined by Groves et al., for all primary analyses to maintain consistency and ensure conservative, high-confidence classification of cell states. To evaluate the robustness of our multiscale inference pipeline, we also applied our full analysis to cell assignments generated using a slightly more permissive threshold (0.70). We were able to assign more Generalist cells into one of the archetypes by lowering the archetype assignment score threshold.

As an additional validation strategy, we identified representative cell clusters corresponding to different SCLC states using Leiden clustering (Supplementary Fig. S10).

Cluster identities were inferred using marker genes reported in a previous study [20] . We selected clusters that showed transcriptional profiles consistent with the SCLC-A2 and -P/Y archetypes and re-applied the *CellChat* analysis (See Supplementary Section: Cluster-based validation of archetype assignments).

Note that archetype analysis is suitable for cancer cell populations with complex heterogeneity (e.g. cells in SCLC datasets) but it does not replace other annotation methods that identify cell types with distinct lineages. We therefore did not use archetype analysis for processing other datasets where non-cancer cells (e.g. fibroblast) are present. The deposited colon cancer data was already preprocessed as described by Sacchetti, et al. [30] $EpCAM^{high}$ and $EpCAM^{low}$ raw count matrices were merged together in R and processed for downstream analysis using the *Seurat* package. Dimension reduction was performed using PCA, tSNE, and UMAP. Cell type annotation was performed based on EpCAM expression. Cells were also assigned epithelial and mesenchymal scores which were computed using gene set enrichment scoring packages.

HER2 breast cancer data was preprocessed as described by Ginzel, et al. [31] Raw count matrices were processed using *Seurat* (version 4.0.0). Scores for S-phase and G_2-M cell-cycle using the *CellCycleScoring* function. Data were log-normalized and scaled after regressing out total UMI counts, percent mitochondrial gene expression, and cell-cycle phase. Dimension reduction was done using PCA and UMAP. Individual clusters were annotated based on known marker gene expression. Trajectory analysis was performed with *Monocle 2* (version 2.12.0) on a subset of the data which contains only the epithelial compartment, and cellular identities were inferred in the terminal branches using a list of curated gene sets obtained from literature search and gene enrichment analysis [31] .

Epithelial-mesenchymal enrichment analysis

Two methods were used to compute epithelial and mesenchymal scores: non-negative principal component analysis (nnPCA) [24, 91] and single-sample gene set enrichment analysis (ssGSEA) [92] . An EMT gene set was used for the enrichment analysis [36] . Since some of the datasets originated from mouse models, human gene symbols from the gene set were converted to their mouse orthologs using the function *convert_human_to_mouse_symbols* from *nichenetr* [93] .

Quantification of epithelial and mesenchymal cell type separation

To assess how well epithelial- and mesenchymal-like populations were separated within each dataset, we calculated silhouette scores and Euclidean distances between cell type centroids using E- and M-scores derived from ssGSEA. For each dataset, we selected the two most epithelial- and mesenchymal-like annotated cell types and computed silhouette scores using the *silhouette* function from the R package *cluster* (version 2.1.4). Cluster assignments were based on cell type annotations. Euclidean distances between E- and M-like populations were calculated as the distance between their mean coordinates (centroids) in the two-dimensional E/M-score space.

Connecting cell-cell communication to TFs and downstream target genes

To infer subtype-specific intracellular signaling pathways, we applied a multiscale pipeline integrating L-R interactions, TF activity, and prior biological knowledge (Fig. 1C, Supplementary Fig. S1). *CellChat* (version 1.6.1) was used to identify intercellular communication events from scRNA-seq data based on differentially expressed ligands and receptors. For each cell type, inferred L-R interactions where the cell is the receiver were used as signaling inputs for downstream network inference.

To infer TF activities, we used the *decoupleR* package (version 2.6.0) with the *CollecTRI* signed gene regulatory network [94] applying a univariate linear model to predict gene expression based on TF-target interaction weights [95] . Activity scores were computed using log-fold changes and filtered for significance (Benjamini-Hochberg correction, p-value < 0.05) and a positive activity score. Differential expression was computed using *Seurat's FindMarkers* function (Bonferroni correction, p-value < 0.05, Supplementary File S2) and used both to support TF activity inference and to assign node weights in the prior knowledge graph. Causal signaling networks were then inferred using the LIANA + framework and *CORNETO* package [28, 35] where we integrate *CellChat*-derived L-R scores, inferred TF activities, and a prior knowledge graph derived from *OmniPath* [96] . *CORNETO* solves an integer linear programming (ILP) problem to reconstruct the smallest-sign consistent, directed network linking receptors to TFs, constrained by data-derived node weights. Ligands and downstream target genes were subsequently added based on inferred L-R pairs and TF-target interactions, with directional consistency and expression filters. Additional implementation details are provided in the Supplementary Material (Supplementary Sect. 1).

While the inferred signaling and regulatory edges are supported by prior knowledge and data-driven scores, we acknowledge that they remain hypothetical in the absence of experimental validation (e.g., ChIP-seq). As such, the reconstructed networks should be interpreted as testable mechanistic hypotheses, rather than definitive causal maps. Additionally, while methods such as NicheNet are well suited for prioritizing ligands based on their ability to regulate observed transcriptional

responses in receiver cells, they do not reconstruct the intracellular signaling intermediates that mediate ligand effects. In contrast, the pipeline used here infers full causal signaling chains from ligands to target genes via receptors, TFs, and signaling intermediates, offering a more mechanistic and interpretable view of subtype-specific regulatory programs.

The code of our multiscale inference pipeline is available at https://github.com/DanielL543/scRNA_seq_multiscale_inference.

Statistical analysis of networks

To assess whether the inferred signaling networks reinforce or alter the phenotypic identity of each cell state, we defined lineage-supporting genes based on the expected expression profile of each subtype. These are genes whose expression alighsn with the known phenotype of the cell type, such as epithelial markers and SCLC-A2 specific genes in the SCLC-A2 subtype, or mesenchymal markers in more mesenchymal-like subtypes (e.g., SCLC-P/Y). These gene sets were curated from published literature [21, 36] .

We then used cell-type specific differential expression to compare the distribution of lineage-supporting genes inside vs. outside the network. Specifically, we constructed 2×2 contingency tables where the two dimensions are (1) the number of differentially expressed genes supporting or not supporting the cell lineage and (2) the number of differentially expressed genes in the network or out of the network. Fisher's exact test was used to evaluate the association between network inclusion and lineage support (See Supplementary Tables 1 and 2 for contingency tables and odds ratios).

Mathematical modeling

For modeling the inter-subtype feedback that we identified as a conserved mechanism across cancer types, the two-variable ODE system shown in Eq. 1 was used to simulate a cell population containing two subtypes of cancer cells (X_1: NE, or E-like cells; X_2: non-NE, or M-like cells). We estimated parameter values from experimental data (See Supplementary Text and Supplementary Fig. S11 for detail). In particular, a target steady state percentage of X_2 (50%) was used during parameter estimation and model analysis, and this percentage was estimated based on the average fraction of non-NE cells from 81 SCLC tumors with inferred compositions [20, 97] . Another study with ex vivo NE-to-non-NE transition showed a similar fraction [15] . Parameter scanning was performed to determine the ability of the parameter set to achieve the target X_2 fraction and to examine the robustness of the conclusions regarding saturation time and variability of the steady state X_2 fraction (see description below, Supplementary Fig. S12-14, and Supplementary Text for sensitivity analysis). Parameter values and ranges with related experimental observations are listed in Supplementary Table S3. A representative parameter set was used for visualizing simulation trajectories: $r_1 = 0.01$, $r_2 = 0.001$, $k_1 = k_2 = 0.04$, $K_2 = 0.5$, and $n_2 = 6$. For feedback-free models (control), K_2 was set to 1000, which effectively changed the multiplier $1/(1+(w_1/K_2)^{n_2})$ to 1, thereby removing the feedback. k_2 was adjusted to allow a feedback-free model (k_2 compensated model) to achieve the target fraction of X_2 cells. The model is dimensionless. One time unit in the model corresponds to 5 h approximately. The initial conditions for the scenario of a purely E-like cell population are $X_1 = 10$, $X_2 = 0$, whereas initial conditions $X_1 = 0$, $X_2 = 10$ were used to simulate the opposite scenario. To examine the behavior of the model with random parameter values around the basal parameter set, 1000 parameter sets were generated with a random sampling from uniform distributions bounded by $(0.5p,\ 1.5p)$ where p is the basal value for each parameter described above. We assumed that the timescale for signal processing is significantly shorter than that for cell state transitions, so the delay in transmitting signals is not considered in our models.

To incorporate additional feedback regulations that we inferred from the time course data for SCLC, we used the following ODEs to describe a "triple-feedback" model:

$$\begin{aligned}\frac{dX_1}{dt} &= r_1X_1 - \frac{1}{1+\left(\frac{w_1}{K_3}\right)^{n_3}+\left(\frac{w_2}{K_4}\right)^{n_4}k_1X_1} + \frac{1}{1+\left(\frac{w_1}{K_2}\right)^{n_2}k_2X_2}\\ \frac{dX_2}{dt} &= r_2X_2 + \frac{1}{1+\left(\frac{w_1}{K_3}\right)^{n_3}+\left(\frac{w_2}{K_4}\right)^{n_4}k_1X_1} - \frac{1}{1+\left(\frac{w_1}{K_2}\right)^{n_2}k_2X_2}\end{aligned} \quad (2)$$

In addition to the variables and parameters in the single-feedback model (Eq. 1), w_2 is the fraction of X_2 cells ($w_2 = X_2/(X_1+X_2)$); K_3 is the threshold of the autocrine effect on cell state transition from E-like cells to M-like cells; K_4 is the threshold of the M-to-E signal effect on cell state transition from E-like cells to M-like cells; n_3 and n_4 describe the nonlinearity of the two effects, respectively. To test the robustness of the main conclusions regarding the saturation time and variability of the X_2 fraction, we performed sensitivity analysis by scanning each model parameters, including those feedback related ones, with a range covering up to two orders of magnitude (See Supplementary Text, Supplementary Fig. S13 and Supplementary Fig. S14). For random parameter sampling, however, feedback related parameters (K and n) were held constant (See values in Supplementary Table S3).

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12964-025-02405-7.

Supplementary File S1 Ligand-receptor interactions across five datasets

Supplementary File S2 Subtype-specific differential expression results

Supplementary Material file containing Supplementary Figures S1-S15, Tables S1-S4, and Supplementary Text

Acknowledgements
The authors thank the members of Tian Hong's laboratory for helpful comments.

Author contributions
Developed computational methods: D.L. and T.H. Analyzed data: D.L., D.R.T. and T.H. Wrote the manuscript: D.L., D.R.T. and T.H. All authors read and approved the final manuscript.

Funding
This work is supported by grants from National Institutes of Health (R35GM149531 awarded to T.H. and R50CA243783 awarded to D.R.T.) and from National Science Foundation (2243562 awarded to T.H.).

Data availability
No datasets were generated during the current study. Previously published datasets GSE138474. GSE149180, GSE154930 and GSE152422 were analyzed in this study. Code for data analysis and modeling is at the GitHub repository https://github.com/DanielL543/scRNA_seq_multiscale_inference.

Declarations

Ethics approval and consent to participate
This work does not involve human participants or animals.

Consent for publication
This manuscript does not contain any individual person's data.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Biochemistry & Cellular and Molecular Biology, The University of Tennessee, Knoxville, Knoxville, TN 37916, USA
[2]Department of Pharmacology and Cancer Biology, Duke University School of Medicine, Durham, NC 27710, USA
[3]Department of Biological Sciences, The University of Texas at Dallas, Richardson, TX 75080, USA

Received: 20 March 2025 / Accepted: 16 August 2025
Published online: 28 August 2025

References

1. Vitale I, Shema E, Loi S, Galluzzi L. Intratumoral heterogeneity in cancer progression and response to immunotherapy. Nat Med. 2021;27:212–24.
2. Burkhardt DB, Juan S, Lock BP, Krishnaswamy JG, S., Chaffer CL. Mapping phenotypic plasticity upon the cancer cell state landscape using manifold learning. Cancer Discov. 2022;12:1847–59.
3. Hanahan D. Hallmarks of cancer: new dimensions. Cancer Discov. 2022;12:31–46.
4. Marine J-C, Dawson S-J, Dawson MA. Non-genetic mechanisms of therapeutic resistance in cancer. Nat Rev Cancer. 2020;20:743–56.
5. Marusyk A, Janiszewska M, Polyak K. Intratumor heterogeneity: the Rosetta stone of therapy resistance. Cancer Cell. 2020;37:471–84.
6. McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. Cell. 2017;168:613–28.
7. Sun X, Yu Q. Intra-tumor heterogeneity of cancer cells and its implications for cancer treatment. Acta Pharmacol Sin. 2015;36:1219–27.
8. Flavahan WA, Gaskell E, Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. Science. 2017;357:eaal2380.
9. Welch DR, Hurst DR. Defining the hallmarks of metastasis. Cancer Res. 2019;79:3011–27.
10. Gupta PB, Pastushenko I, Skibinski A, Blanpain C, Kuperwasser C. Phenotypic plasticity: driver of cancer initiation, progression, and therapy resistance. Cell Stem Cell. 2019;24:65–78.
11. Jain P, et al. Dynamical hallmarks of cancer: phenotypic switching in melanoma and epithelial-mesenchymal plasticity. Semin Cancer Biol. 2023;96:48–63.
12. Pillai M, Jolly MK. Systems-level network modeling deciphers the master regulators of phenotypic plasticity and heterogeneity in melanoma. iScience. 2021;24(10):103111.
13. Stylianou N, et al. A molecular portrait of epithelial–mesenchymal plasticity in prostate cancer associated with clinical outcome. Oncogene. 2019;38:913–34.
14. Watanabe K, Panchy N, Noguchi S, Suzuki H, Hong T. Combinatorial perturbation analysis reveals divergent regulations of mesenchymal genes during epithelial-to-mesenchymal transition. Npj Syst Biol Appl. 2019;5:1–15.
15. Lim JS, et al. Intratumoural heterogeneity generated by Notch signalling promotes small-cell lung cancer. Nature. 2017;545:360–4.
16. Ireland AS, et al. MYC drives Temporal evolution of small cell lung cancer subtypes by reprogramming neuroendocrine fate. Cancer Cell. 2020;38:60–e7812.
17. Gazdar AF, Bunn PA, Minna JD. Small-cell lung cancer: what we know, what we need to know and the path forward. Nat Rev Cancer. 2017;17:725–37.
18. Calbo J, et al. A functional role for tumor cell heterogeneity in a mouse model of small cell lung cancer. Cancer Cell. 2011;19:244–56.
19. Shue YT, et al. A conserved yap/notch/rest network controls the neuroendocrine cell fate in the lungs. Nat Commun. 2022;13:2690.
20. Wooten DJ, et al. Systems-level network modeling of small cell lung cancer subtypes identifies master regulators and destabilizers. PLoS Comput Biol. 2019;15:e1007343.
21. Groves SM, et al. Archetype tasks link intratumoral heterogeneity to plasticity and cancer hallmarks in small cell lung cancer. Cell Syst. 2022;13:690–e71017.
22. Gopal P, et al. Multivalent state transitions shape the intratumoral composition of small cell lung carcinoma. Sci Adv. 2022;8:eabp8674.
23. Rudin CM, et al. Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. Nat Rev Cancer. 2019;19:289–97.
24. Groves SM, et al. Involvement of Epithelial-Mesenchymal transition genes in small cell lung cancer phenotypic plasticity. Cancers (Basel). 2023;15:1477.
25. Stewart CA, et al. Single-cell analyses reveal increased intratumoral heterogeneity after the onset of therapy resistance in small-cell lung cancer. Nat Cancer. 2020;1:423–36.
26. Allison Stewart C, et al. Dynamic variations in epithelial-to-mesenchymal transition (EMT), ATM, and SLFN11 govern response to PARP inhibitors and cisplatin in small cell lung cancer. Oncotarget. 2017;8:28575–87.
27. Dongre A, Weinberg RA. New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. Nat Rev Mol Cell Biol. 2019;20:69–84.
28. Dimitrov D, et al. LIANA + provides an all-in-one framework for cell–cell communication inference. Nat Cell Biol. 2024;26:1613–22.
29. Mollaoglu G, et al. MYC drives progression of small cell lung cancer to a variant neuroendocrine subtype with vulnerability to Aurora kinase Inhibition. Cancer Cell. 2017;31:270–85.
30. Sacchetti A, et al. Phenotypic plasticity underlies local invasion and distant metastasis in colon cancer. Elife. 2021;10:e61461.
31. Ginzel JD, et al. HER2 isoforms uniquely program intratumor heterogeneity and predetermine breast cancer trajectories during the occult tumorigenic phase. Mol Cancer Res. 2021;19:1699–711.
32. Jin S, et al. Inference and analysis of cell-cell communication using cellchat. Nat Commun. 2021;12:1088.
33. Dimitrov D, et al. Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. Nat Commun. 2022;13:3224.
34. Xie Z, Li X, Mora A. A comparison of Cell-Cell interaction prediction tools based on scRNA-seq data. Biomolecules. 2023;13:1211.
35. Rodriguez-Mier P, Garrido-Rodriguez M, Gabor A, Saez-Rodriguez J. Unified knowledge-driven network inference from omics data. 2024.10.26.620390 Preprint at https://doi.org/10.1101/2024.10.26.620390 (2024).
36. Panchy N, Watanabe K, Takahashi M, Willems A, Hong T. Comparative single-cell transcriptomes of dose and time dependent epithelial-mesenchymal spectrums. NAR Genom Bioinform. 2022;4:lqac072.
37. Shoval O, et al. Evolutionary Trade-Offs, Pareto optimality, and the geometry of phenotype space. Science. 2012;336:1157–60.

38. Zhang H, et al. AGRN promotes lung adenocarcinoma progression by activating Notch signaling pathway and acts as a therapeutic target. Pharmacol Res. 2023;194:106819.
39. Gjerdrum C, et al. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. Proc Natl Acad Sci. 2010;107:1124–9.
40. Gordon KJ, Kirkbride KC, How T, Blobe GC. Bone morphogenetic proteins induce pancreatic cancer cell invasiveness through a Smad1-dependent mechanism that involves matrix metalloproteinase-2. Carcinogenesis. 2009;30:238–48.
41. Pickup MW, et al. Deletion of the BMP receptor BMPR1a impairs mammary tumor formation and metastasis. Oncotarget. 2015;6:22890–904.
42. Lu Z, Ghosh S, Wang Z, Hunter T. Downregulation of caveolin-1 function by EGF leads to the loss of E-cadherin, increased transcriptional activity of. Cancer Cell. 2003; 4:499-515
43. Shibue T, Weinberg RA, EMT. CSCs, and drug resistance: the mechanistic link and clinical implications. Nat Rev Clin Oncol. 2017;14:611–29.
44. Lo H-W, et al. Epidermal growth factor receptor cooperates with signal transducer and activator of transcription 3 to induce Epithelial-Mesenchymal transition in cancer cells via Up-regulation of TWIST gene expression. Cancer Res. 2007;67:9066–76.
45. Pirozzi G, et al. Epithelial to mesenchymal transition by TGFβ-1 induction increases stemness characteristics in primary Non small cell lung cancer cell line. PLoS ONE. 2011;6:e21548.
46. Tomlinson DC, Baxter EW, Loadman PM, Hull MA, Knowles MA. FGFR1-Induced epithelial to mesenchymal transition through MAPK/PLCγ/COX-2-Mediated mechanisms. PLoS ONE. 2012;7:e38972.
47. Tian Y, et al. Junctional adhesion molecule-A, an epithelial–mesenchymal transition inducer, correlates with metastasis and poor prognosis in human nasopharyngeal cancer. Carcinogenesis. 2015;36:41–8.
48. Communal L, et al. Low junctional adhesion molecule-A expression is associated with an epithelial to mesenchymal transition and poorer outcomes in high-grade serous carcinoma of uterine adnexa. Mod Pathol. 2020;33:2361–77.
49. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. Nat Rev Mol Cell Biol. 2014;15:178–96.
50. Abedini A, Sayed C, Carter LE, Boerboom D, Vanderhyden BC. Non-canonical WNT5a regulates Epithelial-to-Mesenchymal transition in the mouse ovarian surface epithelium. Sci Rep. 2020;10:9695.
51. Yuan X, et al. Notch signaling and EMT in non-small cell lung cancer: biological significance and therapeutic application. J Hematol Oncol. 2014;7:87.
52. Yin J, Guo Y, Li Z. Platelet-derived growth factor-B signalling might promote epithelial-mesenchymal transition in gastric carcinoma cells through activation of the MAPK/ERK pathway. Contemp Oncol. 2021;25:1.
53. Craene BD, Berx G. Regulatory networks defining EMT during cancer initiation and progression. Nat Rev Cancer. 2013;13:97–110.
54. Xu C, et al. SPP1, analyzed by bioinformatics methods, promotes the metastasis in colorectal cancer by activating EMT pathway. Biomed Pharmacother. 2017;91:1167–77.
55. Nagaharu K, et al. Tenascin C induces Epithelial-Mesenchymal Transition–Like change accompanied by SRC activation and focal adhesion kinase phosphorylation in human breast cancer cells. Am J Pathol. 2011;178:754–63.
56. Lin L-C, Hsu S-L, Wu C-L, Hsueh C-M. TGFβ can stimulate the p38/β-catenin/PPARγ signaling pathway to promote the EMT, invasion and migration of non-small cell lung cancer (H460 cells). Clin Exp Metastasis. 2014;31:881–95.
57. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. J Clin Invest. 2009;119:1420–8.
58. Qi L, et al. Wnt3a expression is associated with epithelial-mesenchymal transition and promotes colon cancer progression. J Experimental Clin Cancer Res. 2014;33:107.
59. Yang D, et al. WNT4 secreted by tumor tissues promotes tumor progression in colorectal cancer by activation of the Wnt/β-catenin signalling pathway. J Experimental Clin Cancer Res. 2020;39:251.
60. Schuh L, et al. Gene networks with transcriptional bursting recapitulate rare transient coordinated high expression States in cancer. Cell Syst. 2020;10:363–e37812.
61. Akbar MW, et al. A stemness and EMT based gene expression signature identifies phenotypic plasticity and is A predictive but not prognostic biomarker for breast cancer. J Cancer. 2020;11:949–61.
62. Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature. 2008;453:544–7.
63. Rosenfeld N, Elowitz MB, Alon U. Negative autoregulation speeds the response times of transcription networks. J Mol Biol. 2002;323:785–93.
64. Harris K, et al. Small cell lung cancer doubling time and its effect on clinical presentation: A concise review. Clin Med Insights Oncol. 2012;6:199–203.
65. Mandicourt G, Iden S, Ebnet K, Aurrand-Lions M, Imhof BA. JAM-C regulates tight junctions and Integrin-mediated cell adhesion and Migration *. J Biol Chem. 2007;282:1830–7.
66. Torre EA, et al. Genetic screening for single-cell variability modulators driving therapy resistance. Nat Genet. 2021;53:76–85.
67. Housman G, et al. Drug resistance in cancer: an overview. Cancers. 2014;6:1769–92.
68. Nami B, Ghanaeian A, Black C, Wang Z. Epigenetic Silencing of HER2 expression during Epithelial-Mesenchymal transition leads to trastuzumab resistance in breast cancer. Life. 2021;11:868.
69. Jordan NV, et al. HER2 expression identifies dynamic functional States within Circulating breast cancer cells. Nature. 2016;537:102–6.
70. Kim C, et al. Chemoresistance evolution in Triple-Negative breast cancer delineated by Single-Cell sequencing. Cell. 2018;173:879–e89313.
71. Tirosh I, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science. 2016;352:189–96.
72. Desai P et al. Microenvironment shapes small-cell lung cancer neuroendocrine States and presents therapeutic opportunities. CR Med. 2024;5(6):101610.
73. Zapatero MR, et al. Trellis tree-based analysis reveals stromal regulation of patient-derived organoid drug responses. Cell. 2023;186:5606–e561924.
74. Qin X, et al. An oncogenic phenoscape of colonic stem cell polarization. Cell. 2023;186:5554–e556818.
75. Boareto M, et al. Notch-Jagged signalling can give rise to clusters of cells exhibiting a hybrid epithelial/mesenchymal phenotype. J Royal Soc Interface. 2016;13:20151106.
76. Brennecke P, et al. Accounting for technical noise in single-cell RNA-seq experiments. Nat Methods. 2013;10:1093–5.
77. Ye Y, Kang X, Bailey J, Li C, Hong T. An enriched network motif family regulates multistep cell fate transitions with restricted reversibility. PLoS Comput Biol. 2019;15:e1006855.
78. Nordick B, Yu PY, Liao G, Hong T. Nonmodular oscillator and switch based on RNA decay drive regeneration of multimodal gene expression. Nucleic Acids Res. 2022;50:3693–708.
79. Gregory PA, et al. An autocrine TGF-beta/ZEB/miR-200 signaling network regulates establishment and maintenance of epithelial-mesenchymal transition. Mol Biol Cell. 2011;22:1686–98.
80. Neelakantan D, et al. EMT cells increase breast cancer metastasis via paracrine GLI activation in neighbouring tumour cells. Nat Commun. 2017;8:15773.
81. Yamamoto M, et al. NF-κB non-cell-autonomously regulates cancer stem cell populations in the basal-like breast cancer subtype. Nat Commun. 2013;4:2299.
82. Jain P et al. Cell-state transitions and density-dependent interactions together explain the dynamics of spontaneous epithelial-mesenchymal heterogeneity. iScience. 2024;27(7):110310.
83. Choi H, et al. Transcriptome analysis of individual stromal cell populations identifies Stroma-Tumor crosstalk in mouse lung cancer model. Cell Rep. 2015;10:1187–201.
84. Almet AA, Cang Z, Jin S, Nie Q. The landscape of cell–cell communication through single-cell transcriptomics. Curr Opin Syst Biology. 2021;26:12–23.
85. Armingol E, Officer A, Harismendy O. Lewis, N. E. Deciphering cell–cell interactions and communication from gene expression. Nat Rev Genet. 2021;22:71–88.
86. Cang Z, et al. Screening cell–cell communication in Spatial transcriptomics via collective optimal transport. Nat Methods. 2023;20:218–28.
87. Garcia-Alonso L, et al. Mapping the Temporal and Spatial dynamics of the human endometrium in vivo and in vitro. Nat Genet. 2021;53:1698–711.
88. Paul S, Adetunji J, Hong T. Widespread biochemical reaction networks enable turing patterns without imposed feedback. Nat Commun. 2024;15:8380.
89. Gay CM, et al. Patterns of transcription factor programs and immune pathway activation define four major subtypes of SCLC with distinct therapeutic vulnerabilities. Cancer Cell. 2021;39:346–e3607.
90. Hart Y, et al. Inferring biological tasks using Pareto analysis of high-dimensional data. Nat Methods. 2015;12:233–5.
91. Panchy N, Watanabe K, Hong T, Interpretable. Scalable, and transferrable functional projection of Large-Scale transcriptome data using constrained matrix decomposition. Front Genet. 2021;12:719099.

92. Barbie DA, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature. 2009;462:108–12.
93. Browaeys R, Saelens W, Saeys Y. NicheNet: modeling intercellular communication by linking ligands to target genes. Nat Methods. 2020;17:159–62.
94. Müller-Dott S, et al. Expanding the coverage of Regulons from high-confidence prior knowledge for accurate Estimation of transcription factor activities. Nucleic Acids Res. 2023;51:10934–49.
95. Badia-i-Mompel P, et al. DecoupleR: ensemble of computational methods to infer biological activities from omics data. Bioinform Adv. 2022;2:vbac016.
96. Türei D, et al. Integrated intra- and intercellular signaling knowledge for multicellular omics analysis. Mol Syst Biol. 2021;17:e9923.
97. George J, et al. Comprehensive genomic profiles of small cell lung cancer. Nature. 2015;524:47–53.

Publisher's Note