

Intratumor Heterogeneity Through the Lens of Gene Regulatory Networks

Narein Rao

Rice University

Nicolae Sapoval

Rice University

Luay Nakhleh

Rice University

Hamim Zafar

hamim@iitk.ac.in

Indian Institute of Technology Kanpur

Article

Keywords: intratumor heterogeneity, gene regulatory networks, single-cell RNA-seq, clonal evolution, cancer genomics, Wasserstein distance, regulon analysis, transcription factors, tumor microenvironment

Posted Date: October 23rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10627293/v1>

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

[Read Full License](#)

Additional Declarations: No competing interests reported.

Intratumor Heterogeneity Through the Lens of Gene Regulatory Networks

Narein Rao¹, Nicolae Sapoval¹, Luay Nakhleh^{1,2,*} and Hamim Zafar^{3,4,5,*}

¹Department of Computer Science, Rice University, , 77005, TX, USA

²Department of BioSciences, Rice University, , 77005, TX, USA

³Department of Computer Science and Engineering, Indian Institute of Technology Kanpur, 208016, Uttar Pradesh, India

⁴Department of Biological Sciences and Bioengineering, Indian Institute of Technology Kanpur, , 208016, Uttar Pradesh, India

⁵Mehta Family Center for Engineering in Medicine, Indian Institute of Technology Kanpur, , 208016, Uttar Pradesh, India

*Corresponding authors. nakhleh@rice.edu, hamim@iitk.ac.in

Abstract

Intratumor heterogeneity (ITH), the presence of heterogeneous genomic and transcriptomic populations in tumor and its implications for cancer therapies, has been one of the focal points of computational cancer biology in the past years. While ITH is well characterized at the genomic level, the regulatory wiring connecting clonal genotypes to transcriptional states remains largely unmapped. To understand whether evidence for ITH at the gene regulatory level can be inferred from scRNA-seq data, we analyzed scRNA-seq datasets spanning triple-negative breast cancer, colorectal cancer, and glioblastoma as well as an extended cohort spanning 253 patients across 12 cancer types, and two longitudinal treatment-response cohorts: a stage IV breast cancer patient profiled across treatment, and an 11-patient ovarian cancer cohort profiled before and after neoadjuvant chemotherapy, inferring clone-specific gene regulatory networks (GRNs) and developing quantitative tools to compare them: a Wasserstein distance-based measure of regulatory divergence, a k^{th} order neighborhood similarity framework distinguishing structural from functional rewiring, bootstrap-based edge confidence, and an integrated regulon classification pipeline. Our analyses demonstrate that across cancer types, tumor clones exhibit reproducibly distinct network architectures, and dominant clones show significantly greater regulatory divergence from normal cells than smaller clones, linking regulatory rewiring to clonal fitness. Treatment-induced remodeling exceeded clonal variation, with rewiring exhibiting structural divergence but functional convergence. In the ovarian cancer cohort, cross-phase (pre- versus post-neoadjuvant-chemotherapy) clone pairs showed greater regulatory (Wasserstein) divergence than same-phase clone pairs (a suggestive result given the small number of clone pairs), accompanied by exploratory shifts in the regulatory activity of select transcription factors following treatment. Regulon analysis further resolved clone-variable, sample-specific, and broadly conserved programs, some of which containing transcription factors whose target genes were independently associated with patient survival.

Keywords intratumor heterogeneity, gene regulatory networks, single-cell RNA-seq, clonal evolution, cancer genomics, Wasserstein distance, regulon analysis, transcription factors, tumor microenvironment

Introduction

Cancer is a complex disease characterized by extensive genomic and transcriptomic diversity [10] occurring both across different cancer types and patients, as well as within individual tumors. The latter form of diversity manifested by the presence of multiple genotypically distinct cellular populations within the same tumor tissue is known as intratumor heterogeneity (ITH) and presents a major obstacle in design and delivery of efficient and personalized cancer therapies [53, 12]. A common approach to characterize ITH is based on the clonal cancer evolution hypothesis [61], which involves identifying the clonal subpopulations of tumor cells typically defined by their somatic single-nucleotide variants (SNVs) [90, 19], copy-number aberration (CNA) profiles [13] or a combination of both [74, 11]. However, while the inference of clonal subpopulations

from single-cell DNA data (scDNA) [68, 91, 75, 89] and single-cell RNA data (scRNA) [1, 21] has become increasingly common in modern studies of tumor evolution, the assessment of clone-specific alterations in gene regulatory networks (GRNs) remains limited. In many single-cell RNA-seq cancer studies, CNA inference from epithelial or tumor cells is routinely used to distinguish malignant from non-malignant cells, leveraging large-scale transcriptional shifts associated with chromosomal alterations [79, 69]. However, the tumor subclones defined by these inferred copy-number profiles are rarely further characterized in terms of their GRN activity or regulatory programs. A limited number of studies have performed the general characterization of GRNs [34], explored network-based approaches to quantify ITH [66], and developed methods that integrate protein-protein interaction networks with scRNA-seq expression data to infer context-specific regulatory interactions [93]. Other studies

have focused on regulatory changes associated with specific cancer hallmark pathways [33, 81] or have compared tumor-derived GRNs with regulatory networks inferred from normal tissues to identify cancer-specific rewiring of transcriptional programs [16]. However, these approaches largely treat tumors as relatively homogeneous systems and do not explicitly account for the clonal structure of tumors revealed by single-cell data. Consequently, the heterogeneity of transcriptional regulatory networks across tumor clones within the same tumor or their evolution across different temporal or treatment-associated states remains largely unexplored, limiting our understanding of how regulatory rewiring contributes to clonal diversification and tumor evolution.

To address these, here we test the hypothesis that clone-specific changes in GRNs can be inferred and quantified from scRNA-seq datasets. In particular, we analyzed scRNA-seq datasets from a pilot cohort of colorectal cancer (CRC, $n = 6$), triple-negative breast cancer (TNBC, $n = 12$), and glioblastoma ($n = 11$) patients, an extended cohort of 253 patients spanning 12 cancer types (which includes the pilot datasets), and two longitudinal treatment-response cohorts: one stage IV breast cancer cohort comprising three longitudinal samples (pre-treatment, post-treatment 1, and post-treatment 2), and an 11-patient ovarian cancer cohort profiled before and after neoadjuvant chemotherapy, quantifying the extent of clonal GRN divergence from normal cell subpopulations; the extent of GRN divergence across clones and its association with clonal fitness; therapy-induced GRN divergence and finally the variation in GRN activity and architecture across clones and tumor samples, their association with pathways and established gene meta-programs. Specifically, we used a Wasserstein distance-based framework to quantify regulatory divergence between clones, measuring the minimal redistribution required to transform one clone's regulatory profile into another. We also employed k^{th} order neighborhood similarity analysis to differentiate between structural and functional rewiring, revealing patterns of regulatory convergence and divergence that are not apparent from conventional gene expression analyses. Using these approaches, we sought to answer three key questions: (1) to what extent do tumor clones exhibit distinct regulatory architectures compared to normal cell populations? (2) how does clonal GRN divergence relate to clonal dominance and therapy-induced changes? (3) how does clone-specific regulatory activity shape functional pathways and gene meta-programs across tumors?

Our analyses across cancer types revealed consistent differences in regulatory interactions between tumor clones and normal cells, as well as among tumor clones, with the differential regulatory interactions supported by high bootstrap confidence. Importantly, clonal GRNs deviated significantly from randomized null models, indicating non-random regulatory structure. We further found that the dominant clone exhibited greater divergence from the normal cell population compared to smaller clones, suggesting a potential association between regulatory rewiring and clonal fitness. In a longitudinal scRNA-seq dataset from a stage IV breast cancer patient undergoing treatment, we observed variability in gene activity (measured by betweenness centrality) across both clones and treatment timepoints. Regulatory divergence was more pronounced across treatment timepoints than between clones, with GRN rewiring exhibiting substantial structural changes but relatively limited functional shifts over time. For the ovarian cancer cohort specifically, we additionally asked whether clonal regulatory networks diverge more across treatment phases than within a phase,

and which transcription factors' regulatory activity shifts most consistently following neoadjuvant chemotherapy (Methods). For TNBC and Glioblastoma, we captured the activity of biologically relevant transcription factor–target (TF–TG) interactions (regulons), which displayed both clone- and sample-specific variation, as well as consistent activity across multiple samples. These regulons were further associated with enriched immune or cell state associated pathways and gene programs, enabling characterization of subnetwork-level differences across samples and clones. Notably, variation in TF–TG activity within gene programs was driven either by particular clones or by broader, sample-level effects. By addressing these questions, our study provides a comprehensive framework for linking regulatory network architecture to functional heterogeneity in tumors, uncovering mechanistic insights into tumor evolution and potential therapeutic vulnerabilities.

Results

Tumor clones exhibit distinct gene regulatory network architectures compared to normal cell populations

To assess whether tumor clonal GRNs differ from normal cell populations across cancer types, we analyzed 9 TNBC [86, 63] and 6 CRC [42] scRNA-seq datasets, reconstructed the clonal GRNs and compared GRNs inferred from tumor clones and normal cell populations using multiple complementary network similarity metrics. We observed most of the inferred GRN edges to have bootstrap support > 60 , allowing us to work with reliable regulatory interactions (Fig. 1c-d). We then compared the edge sets of GRNs inferred from clonal populations and those inferred from normal cell populations (Fig. 1a,b). For each dataset, we evaluated edge-level differences using three complementary strategies: UpSet plots were used to visualize overlaps among GRN edge sets, Jaccard similarity was used to quantify normalized overlap while accounting for differences in network size, and Hamming distance was computed using binarized edge-presence vectors constructed from the union of edges across the compared populations (see Methods). Across all three approaches, clonal GRNs consistently exhibited greater similarity to one another than to networks inferred from normal cells (Fig. 1a,b; Supplementary Fig. S1a). We calculated the average number of shared edges across clone–clone and clone–normal subpopulation pairs, observing consistently higher overlap among clone–clone subpopulations (Fig. 1a). On average clone–clone GRN difference was 6076 edges ($\sim 12\%$).

Furthermore, clone–clone pairs showed higher Jaccard similarity than clone–normal pairs across the datasets (Fig. 1b; this difference was significant for TNBC, $p = 0.038$, and a non-significant trend for CRC, $p = 0.089$, which comprises only 6 datasets). From our observations we note that clone–clone GRN similarity remained consistently higher than the clone–normal values (Fig. 1b, Supplementary Fig. S1a). Hamming distance analysis further supported these differences (Supplementary Fig. S1a). Larger clone–clone versus clone–normal differences were associated with disparities in the sizes of the underlying cell populations. For example, samples CID4495, CID4513, CID44971, and TN0131 were outliers among TNBC datasets, while SMC16 and SMC07 were outliers among CRC datasets (Supplementary Fig. S1a). These samples exhibited substantial differences in the number of cells between normal and clonal populations, without a consistent bias toward

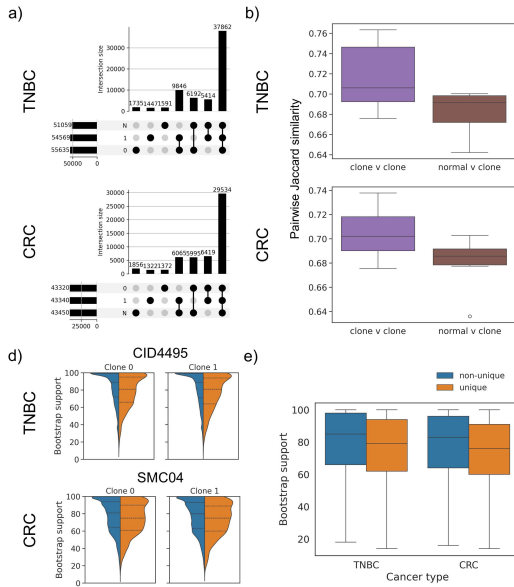


Figure 1 Clonal GRNs exhibit greater structural similarity to each other than to normal GRNs. (a) UpSet plots depicting shared edges between tumor clones and normal cell GRNs for a single TNBC (CID4495) and CRC (SMC04) dataset, highlighting greater edge sharing among tumor clones. (b) Distribution of pairwise Jaccard similarities between edge sets of every pair of clones (purple) and between clones and normal subpopulations (brown). Mann-Whitney U test p -values of 0.038 and 0.089 for TNBC and CRC, respectively. (c) Distribution of bootstrap support values for GRN edges not unique to a clone (blue) and edges unique to clone (orange) inferred for TNBC sample CID4495 (top) and CRC sample SMC04 (bottom) for each clone. (d) Distribution of bootstrap support values for GRN edges not unique to a clone (blue) and edges unique to clone (orange) inferred across all TNBC and CRC cancer datasets.

either population being larger (Supplementary Fig. S2a). Finally, edges unique to clone-specific GRNs tended to have lower bootstrap support compared to the overall set of clone-specific edges (Fig. 1c-d, Supplementary Fig. S1b-c). Together, these analyses across both cancer types highlight consistent structural differences between GRNs inferred from tumor clones and those from normal cell populations, supporting the existence of distinct regulatory architectures in tumor clones.

To test whether these clone-versus-normal regulatory differences generalize beyond this 15-dataset pilot cohort, we repeated the bootstrap-support, Jaccard-similarity, and Hamming-distance analyses in a substantially larger extended cohort of 253 patients spanning 12 cancer types from the same 3CA atlas, which includes the TNBC and CRC datasets analyzed above (Methods; datasets and patients are listed in Supplementary Tables S1–S3, and per-patient results are shown in Supplementary Figs. S11–S18). All three findings were replicated: edges unique to a clone showed lower bootstrap support compared to edges shared with other populations (medians ≈ 71 vs. ≈ 79 out of 100; Mann-Whitney $p \approx 0$; Fig. 2a), clone-clone pairs showed higher Jaccard similarity than clone-normal pairs (medians ≈ 0.64 vs. ≈ 0.59 ; $p = 1.2 \times 10^{-23}$; Fig. 2b), and clone-clone Hamming distance was again smaller than clone-normal Hamming distance ($n = 252$ patients; $p = 3.0 \times 10^{-147}$; Fig. 2c). Together with the pilot cohort, these results indicate that the greater GRN

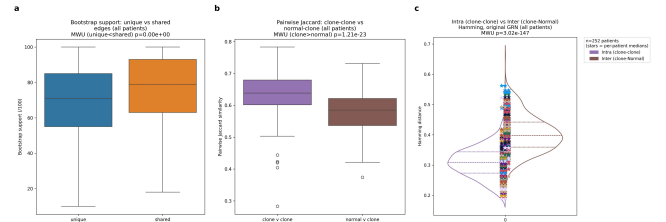


Figure 2 Extended-cohort analysis (253 patients, 12 cancer types, including the pilot TNBC/CRC datasets; exact n varies slightly by panel, Methods) of the clone-versus-normal regulatory differences in Fig. 1. (a) Bootstrap support (/100) for edges unique to a clone versus edges shared with other populations, pooled across all patients (Mann-Whitney $p \approx 0$). (b) Pairwise Jaccard similarity between clone-clone and normal-clone edge sets, pooled across all patients ($p = 1.2 \times 10^{-23}$). (c) Hamming distance between clone-clone and clone-normal edge-presence vectors computed from the original (non-bootstrap) GRN, pooled across all patients ($n = 252$; $p = 3.0 \times 10^{-147}$); stars mark each patient's per-category median. This panel uses the original (non-bootstrap) GRNs, unlike the bootstrap-subsample intra-/inter-clone analysis of Fig. 4a-c.

similarity among tumor clones than between tumor clones and normal cells is a consistent pattern across diverse cancer types.

Regulatory network divergence distinguishes tumor clones and associates with clonal dominance

In order to establish whether robust GRN differences exist between tumor clones, we compared GRNs inferred from subsampled populations across different clones (inter-clone) with those inferred from subsampled populations within the same clone (intra-clone). To quantify these differences, we computed Hamming distances between binarized edge vectors (edge presence/absence) derived from bootstrap subsamples (subsampling fraction 90%) drawn from the same clones and from different clones (Fig. 4a-c). For both TNBC and CRC datasets, distances between subsamples originating from the same clone were significantly smaller than those between GRNs inferred from subsamples belonging to different clones, indicating consistent regulatory differences between clonal populations. To further confirm that these differences reflect clone-specific regulatory architectures rather than artifacts of sampling variability, we performed a shuffle experiment. Specifically, we compared the similarity between inferred clone-specific GRNs measured using Jaccard similarity of edge sets with GRNs inferred from datasets in which clone labels were randomly shuffled within each sample, i.e. a null model (see Methods). Across both cancer types, the Jaccard similarity observed for the original clone-specific GRNs deviated substantially from the distribution obtained from shuffled datasets, with a median deviation of approximately six standard deviations from the shuffled mean (Fig. 4d,e). Although 2 (out of 15) samples (CID44971, SMC16) showed smaller deviations (~ 2 standard deviations), these corresponded to datasets in which the underlying clones differed substantially in cell counts (e.g., clone1 $\gg$ clone2) (Supplementary Fig. S2a). In such cases, shuffled datasets may still be dominated by cells from one clone, limiting the ability of the shuffle procedure to fully randomize clonal structure.

To assess whether GRN divergence is associated with clone size, a proxy for clonal fitness, we quantified regulatory divergence using the 1-Wasserstein distance (Methods) between GRNs inferred for the largest and smallest clonal subpopulations relative to the GRN of

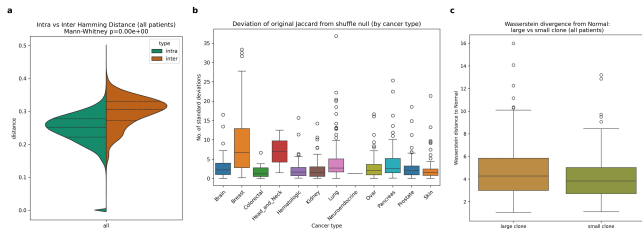


Figure 3 Extended-cohort analysis (253 patients, 12 cancer types, including the pilot TNBC/CRC datasets; exact n varies slightly by panel, Methods) of the intra-/inter-clone divergence and clonal-dominance findings in Fig. 4. (a) Bootstrap intra-clone (same-clone replicate pairs) versus inter-clone (different-clone replicate pairs, Normal excluded) Hamming distance, pooled across all patients (Mann-Whitney $p \approx 0$). (b) Deviation of the original clone-pair Jaccard similarity from its shuffle-null distribution, in standard deviations, by cancer type; note the substantial cancer-type dependence of the effect size (see main text). (c) Wasserstein divergence from Normal for the largest versus smallest non-Normal clone by cell count, pooled across all patients ($n = 253$; binomial $p = 2.4 \times 10^{-23}$; Wilcoxon $p = 2.9 \times 10^{-26}$).

the corresponding normal cell population across all 15 datasets. On performing a binomial and Wilcoxon rank test, we observed that the divergence of the GRNs of the largest clonal subpopulations from the normal cell subpopulation were significantly more than the smaller clones (Fig. 4f). We verified that this effect is not driven by differences in absolute cell counts. The sizes of the large clonal, small clonal, and normal subpopulations varied across datasets, ensuring that the observed pattern was consistent across diverse sampling scenarios (Supplementary Fig. S2b). These results provide evidence that tumor clones exhibit reproducible and biologically meaningful differences in their gene regulatory architectures.

To test whether the intra-/inter-clone divergence, shuffle-null, and clone-size-divergence findings above generalize beyond the original 15-dataset cohort, we repeated all three analyses on the same 253-patient, 12-cancer-type extended cohort (Methods; per-patient and per-cancer-type results in Supplementary Figs. S11–S18). The intra-clone-smaller-than-inter-clone pattern was again observed in the pooled extended cohort (medians ≈ 0.26 vs. ≈ 0.31 ; Mann-Whitney $p \approx 0$; Fig. 3a), and held in the same direction within every individual cancer type we examined. The shuffle-null result also replicated: the original clone-specific Jaccard values deviated significantly from the shuffle-null distribution in all 12 cancer types (Fig. 3b), with median deviations ranging from ≈ 1.2 to ≈ 7.0 standard deviations depending on cancer type (breast and head-and-neck cancers showed the largest median deviations, ≈ 6.7 and ≈ 7.0 SD; the remaining ten cancer types ranged ≈ 1.2 – 2.7 SD), indicating that the effect size is cancer-type-dependent. Finally, the clone-size-associated Wasserstein divergence from normal also replicated ($n = 253$ patients; binomial $p = 2.4 \times 10^{-23}$; Wilcoxon $p = 2.9 \times 10^{-26}$; Fig. 3c), and divergence was weakly but significantly correlated with clone size pooled across large and small clones (Spearman $\rho = 0.20$, $p = 9.1 \times 10^{-6}$), further supporting an association between clonal dominance and regulatory divergence from normal across a broad range of cancer types.

Treatment-driven divergence of gene regulatory networks across longitudinal tumor samples

For the longitudinal data from the stage 4 breast cancer patient, we focused on GRN differences between the pretreatment tumor sample (Pre) and two post-treatment tumor samples: one originating from

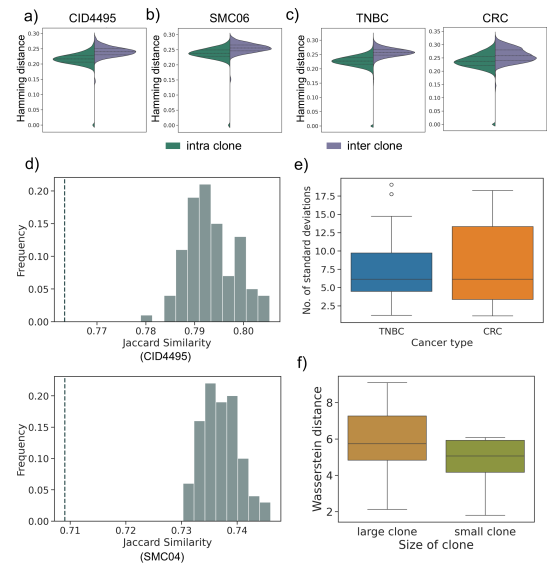


Figure 4 Bootstrap GRNs exhibit stable intra-clone structure, significant inter-clone divergence, with dominant clones exhibiting greater regulatory divergence from normal GRNs. Distribution of Hamming distances between edge vectors of the bootstrap GRNs inferred from sample. Intra-clone (green) distances are measured between all pairs of bootstrap subsamples derived from the same clone, inter-clone (purple) distances are measured between all pairs of bootstrap subsamples derived from different clones. (a) Data for TNBC sample (CID4495) (b) Data for CRC sample (SMC04) (c) Data for all TNBC and CRC samples. (d) Distribution of the Jaccard similarities between edge sets of two tumor subsets obtained by shuffling the cells. Dashed line represent the Jaccard similarity between the edge sets of the clone-specific GRNs. Top: TNBC sample (CID4495). Bottom: CRC sample (SMC04). (e) Distribution of distances between Jaccard similarity of actual clonal GRNs and mean shuffle Jaccard similarity in terms of number of standard deviations across all TNBC and CRC samples. (f) Distribution of Wasserstein distances computed between GRNs inferred from the large clone subpopulation and normal subpopulation (blue); and small clone subpopulations and normal subpopulation (orange) across all 15 datasets (TNBC and CRC). Statistical significance was confirmed using both binomial ($p = 4.8 \times 10^{-4}$) and Wilcoxon rank tests ($p = 6.1 \times 10^{-5}$).

the vicinity of central necrotic tissue (Post 1), and another containing viable cancer cells collected near the tumor margins (Post 2) [62]. Out of 23,928 edges inferred for the Pre tumor GRN, only 10.7% (4,765) were shared with Post 1, and 15.07% (8,528) were shared with Post 2. In total, 3,328 edges were common to all three tumor-specific GRNs, while 7,691 edges were shared between Post 1 and Post 2. Due to these large differences in the inferred GRNs, many nodes are unique to their respective GRN while still exhibiting high degree (Supplementary Fig. S3). Among these nodes, *ETS1*, *KLF6*, and *HEY1* showed some of the highest degrees in the Post 1 and Post 2 GRNs but were absent in the Pre GRN. Additionally, these nodes were also strongly supported by subsampling bootstrap values, with average support across Pre:Post 1:Post 2 subsamples of 0 : 81.9 : 79.9, 0 : 78.1 : 83.1, and 0 : 59.9 : 84.9, respectively. In contrast, *ZEB1* and *MEF2A* were detected only in the Pre GRN with bootstrap support of 76.7 : 0 : 0 and 62.9 : 0 : 0, respectively.

Given the large discrepancies between the inferred GRNs, we further evaluated nodes with high betweenness centrality to identify potential drivers of key differences between the three samples (Supplementary Fig. S4). The transcription factor *EGR1* exhibited

among the highest betweenness centrality in both the Pre and Post 2 networks. In contrast, *GLI2* showed high centrality in Post 1 but was absent from the GRNs in both Pre and Post 2. Similarly, *PLAGL1* [2, 30] and *ELF1* displayed high centrality in Post 2 but did not appear as nodes in the GRNs inferred from the Pre and Post 1 tumor samples. Unlike the clonal subpopulations within the same sample, the differences in edges between the inferred GRNs for the three samples were substantial. Although the differences across time-point-specific GRNs were apparent, we also sought to examine how clonal structure influences these centrality measures. To do so, we analyzed differences in degree and betweenness centrality across each clone within each time point. We did not observe large differences in degree centrality, as all clones displayed similar degree distributions. However, we identified notable patterns in betweenness centrality (Fig. 5a). Specifically, the betweenness centrality of several outlier genes including *FOXF2*, *GLI2*, *CREM*, and *EGR1* appeared to be driven by specific clones to varying degrees at each time point. In some cases, only a small subset of clones contributed substantially to the betweenness property of a gene at a given time point (e.g., *FOXF2*).

To investigate this more systematically, we adopted a more robust framework to study the trajectory divergence of GRNs across time points. To quantify GRN divergence, we used our 1-Wasserstein distance formulation and calculated (1) the divergence of GRNs at each time point from the GRN inferred from the normal cell subpopulation, denoted as $W_{\text{timepoint}}^{\text{normal}}$ and (2) the divergence between GRNs inferred from different time points, denoted as $W_{\text{timepoint1 timepoint2}}^{\text{timepoint1 timepoint2}}$ (see Methods). We observed that the Pre time point exhibited the highest divergence from the normal cells, as expected ($W_{\text{Pre}}^{\text{normal}} = 7.80$), followed by a progressive decrease in divergence from Post 1 to Post 2 ($W_{\text{Post1}}^{\text{normal}} = 5.78$ and $W_{\text{Post2}}^{\text{normal}} = 2.09$, respectively). Next, when examining divergence across time points, we observed higher divergence between Pre-Post 1 ($W_{\text{Post1}}^{\text{Pre}} = 7.39$) and Pre-Post 2 ($W_{\text{Post2}}^{\text{Pre}} = 5.20$) compared to the divergence between Post 1 and Post 2 ($W_{\text{Post2}}^{\text{Post1}} = 3.78$). Notably, the magnitude of these time-point divergences (Pre-Post 1 and Pre-Post 2) exceeded the divergence observed among clonal GRNs within the same time point (i.e., GRN divergence between Pre clone GRNs and Post 1 and Post 2 clone GRNs, respectively) (Fig. 5b). Additionally, we also computed the k^{th} -hop jaccard similarity across genes present in the pre and Post 1 time point GRNs; and pre and Post 2 time point GRNs. On calculating the difference between the 0^{th} -hop jaccard similarity and k^{th} -hop jaccard similarity, we noticed the distribution of differences to incline towards having negative values (Fig. 5c) in both pre vs Post 1 and pre vs Post 2 comparisons. This implies that a large number of interactions across both pre and post time point comparisons showcase structural divergence with functional convergence. Overall, these results suggest that treatment-induced regulatory rewiring may exceed the degree of network rewiring attributable solely to somatic clonal evolution.

Neoadjuvant chemotherapy is associated with shifts in clonal gene regulatory networks in ovarian cancer

The longitudinal stage IV breast cancer dataset above provides a single-patient, proof-of-concept view of treatment-induced regulatory rewiring. To test this at cohort scale, we turned to an independent, longitudinal high-grade serous ovarian cancer (HGSOC) cohort of 11 patients profiled before and after neoadjuvant chemotherapy

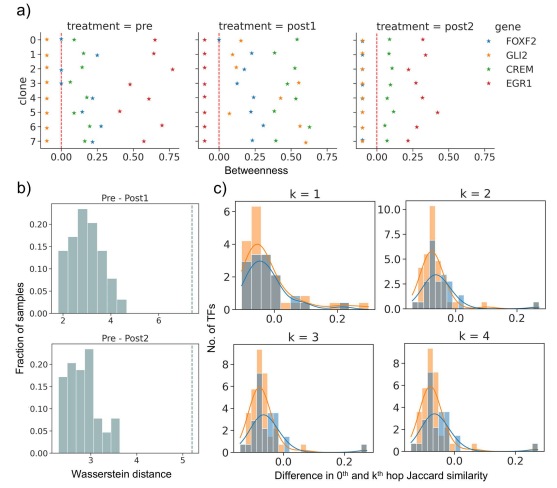


Figure 5 Temporal GRN analysis reveals dynamic, clone-specific centrality shifts, regulatory rewiring across treatment timepoints and structurally divergent network patterns. (a) Betweenness centrality measure for *FOXF2*, *GLI2*, *CREM*, and *EGR1* across clones at different time points (pre, Post 1, Post 2). Red dashed line represents 0 betweenness. Genes located to the left of the red dashed line represent isolated nodes within the GRN. (b) Distribution of Wasserstein distance between clone specific GRNs at time points Pre and Post 1 (top); and Pre and Post 2 (bottom). Grey line represents the Wasserstein distance between whole GRNs inferred at time points Pre and Post 1; and Pre and Post 2. (c) Distribution of difference between the 0^{th} -hop and k^{th} -hop jaccard similarities across common genes in the Pre, Post 1 and Post 2 GRNs. Blue curve represents the comparison between Pre and Post 1 time points and orange curve represents the comparison between Pre and Post 2 time points.

(NACT) [92] (Methods). Because clonal annotation was performed on each patient's pooled pre- and post-treatment cells, individual clones were generally phase-segregated i.e., a given clone's cells came almost entirely from one timepoint consistent with chemotherapy-driven clonal selection, as reported in the original study. We therefore compared GRNs inferred from phase-segregated clones rather than tracking a single clone across both timepoints, which was generally not possible in this cohort (Methods).

Using an absolute cell-count floor to assign each clone to a treatment phase (excluding clones with ambiguous, mixed-phase membership or too few cells in either phase; Methods), 5 of 11 patients (EOC1005, EOC136, EOC153, EOC443, EOC733) contributed at least one usable clone pair, yielding 7 within-phase (pre-pre or post-post) and 7 cross-phase (pre-post) clone-pair comparisons pooled across patients. As with the longitudinal stage IV breast cancer analysis above, we used the 1-Wasserstein distance as our regulatory divergence measure. Cross-phase clone pairs showed greater Wasserstein regulatory divergence than within-phase pairs (medians ≈ 2.79 vs. ≈ 2.16 ; Mann-Whitney $p = 1.75 \times 10^{-2}$, uncorrected; Fig. 6a), consistent with treatment-associated regulatory rewiring exceeding the divergence between independently-sampled clones of the same phase. Given the modest sample size (7 pairs per group, from 5 patients), we treat this as suggestive evidence of treatment-associated divergence rather than a fully powered cohort-level finding.

To identify which regulators are most associated with this divergence, we compared each gene's regulatory out-degree (number of downstream targets) between the combined pre-treatment and

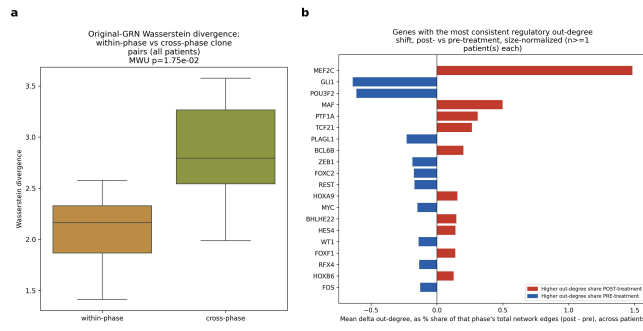


Figure 6 Treatment-phase (pre- vs. post-neoadjuvant chemotherapy) regulatory network comparison in an 11-patient longitudinal ovarian cancer cohort [92]. (a) Wasserstein divergence of original-GRN edge weights, within-phase versus cross-phase clone pairs, pooled across 5 patients contributing usable clone pairs ($p = 1.75 \times 10^{-2}$). (b) Genes with the largest, most consistent pre-to-post shift in regulatory out-degree, expressed as a share of each phase's total network edges, across the 3 patients with a usable clone in both phases; roughly half of the genes shown are supported by only a single patient (see main text) and should be read accordingly.

post-treatment clone networks, for the 3 of 11 patients with at least one usable clone in each phase (EOC1005, EOC136, EOC153). Because a phase's combined network size depends on how many clones survived the phase-assignment filter in that phase independent of any underlying biology we ranked genes by their pre-to-post change in out-degree *share* of that phase's total network edges, rather than by raw degree change (Methods; Fig. 6b). Most of the largest shifts were supported by only a single contributing patient; we treat those as exploratory and instead highlight two genes whose out-degree share shifted in the same direction in both patients where they were detected, and for which that direction is consistent with existing ovarian cancer literature. *MAF* gained out-degree share post-NACT in both contributing patients, including a complete de novo gain in one (zero edges pre-treatment to 278 post-treatment). Large- and small-*MAF* transcription factors, including the closely related *MafB*, have been reported to promote ovarian cancer cell proliferation and invasion and to reduce sensitivity to cisplatin and olaparib, with *Maf* overexpression linked to chemoresistance more broadly across cancer types [18], consistent with *MAF*-driven regulation becoming more prominent among the clones that persist after treatment. *ZEB1*, conversely, lost out-degree share post-NACT in both contributing patients and illustrates why the size-normalized share, rather than raw degree, was necessary: its raw out-degree *increased* in both patients, yet its share of the (larger) post-treatment network *decreased* in both. This is consistent with reports showing that, although *ZEB1* marks a more aggressive, EMT-associated phenotype and predicts poorer overall survival in ovarian cancer broadly, it is not prognostic among platinum-treated patients, suggesting that it is not a specific marker of chemoresistance [71]. This is consistent with our observation that *ZEB1* regulatory activity was not preferentially retained in treatment-surviving clones.

Regulon activity reveals functional heterogeneity across clones and tumor samples

Regulons, sets of genes regulated by a common transcription factor (TF), play an important role in understanding disease-associated regulatory mechanisms [57, 5, 25]. Since the expression levels of

target genes reflect upstream TF activity, regulon analysis provides a useful framework for inferring regulatory activity and characterizing localized GRN dynamics within a given dataset. Having observed clear structural differences across clonal GRNs in multiple cancer types, we next examined regulon activity across clones and samples within each cancer type, followed by pathway enrichment analysis of regulon-associated genes. To ensure reliable regulon definitions, we relied on curated TF-TG interactions from CollecTRI [56]. We first analyzed 12 TNBC samples [86, 63, 21] (see Methods), computed the activity of regulons in each clone of each sample and also identified regulons exhibiting differential activity across clones and samples. Specifically, based on the clonal and sample-level activity, we categorized regulons into three groups: (1) regulons that exhibit differential activity across clones in multiple samples, (2) regulons that do not vary substantially across clones but show broad activity across multiple samples, and (3) regulons showing sample-specific activity (Fig. 7a). Importantly, regulons GRHL2 and SIX1, which are known to be associated with EMT [26, 14] and HEY1 and HEY2, downstream targets of Notch signaling [20] exhibited differential activity across tumor clones in multiple samples. On the other hand, regulons BCL11A and TBX3 (stemness drivers), TRPS1 (marker for TNBC, known to promote epithelial proliferation [4]) and POU2F3 (defines a rare, highly aggressive “Tuft cell-like” subtype of TNBC with a distinct chemosensory gene signature [87]) were found to have activity across all clones in multiple samples.

Notably, some of these regulons contained TFs or TGs that exhibited alterations in copy number profiles (CNP) such as *PBRM1*, suggesting potential links between genomic variation and regulatory activity. To further characterize the functional roles of these regulons, we performed pathway enrichment analysis on the TFs and TGs associated with regulons in each category. While several biologically relevant pathways were shared across samples and clones (e.g. PI3K-Akt signaling pathway [51], FoxO signaling [65]) (Supplementary Fig. S5a), we also identified pathways uniquely enriched within each regulon category (Fig. 7b) such as EGF EGFR signaling [52] and Notch signaling pathways [22]. These results highlight how distinct sets of regulons may contribute to functional differences across clones or samples. One notable observation was the enrichment of Interferon/MHC-II-related gene meta-program within the category of regulons showing broad activity across multiple samples. This raised the question of how GRN structures vary across well-characterized functional meta-programs [80]. To examine the structural impact of clonal variation on these gene meta-programs, we first constructed SCENIC-derived regulatory networks using the inferred CellOracle networks for each clone within every TNBC dataset. We then restricted each network to genes belonging to specific meta-programs associated to the samples considered (see Methods). Fig. 7c illustrates the edges present across clones within each dataset, along with the associated bootstrap support values for each edge for meta-program ‘Interferon/MHC-II’. We observed that some regulatory interactions were present in certain clones but absent in others, or appeared in specific samples but not across the entire cohort. Interactions that varied across clones also included previously known interactions (e.g., *KLF4:SOCS3* [48]). Similarly, we also observed interactions involving *JUND* (a key regulator in TNBC [15, 47]) in the PDAC-related meta-program (Supplementary Fig. S7b); and *DDIT3*-mediated interactions in the stress meta-program (Supplementary Fig. S7d). Several of these metaprogram-associated TFs exhibited outlier degree

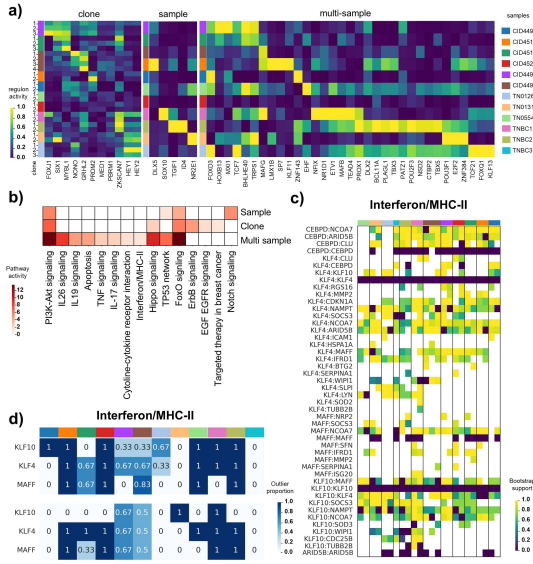


Figure 7 Regulon categories reveal distinct patterns of clone- and sample-specific activity, pathway enrichment, and meta-program-dependent GRN variability across TNBC samples. (a) Activity of three Regulon categories across all TNBC samples: (left to right) regulons that exhibit differential activity across clones in multiple samples, regulons displaying sample-specific activity and regulons that do not vary substantially across clones but show broad activity across multiple samples. x-axis denotes the Regulons, y-axis denotes the samples. (b) Enriched pathways and corresponding pathway enrichment scores for the three regulon categories. Genesets used for pathway enrichment include the TFs and TGs of every regulon within a specific category. (c) Distribution of SCENIC edges across all TNBC samples for meta-program Interferon/MHC-II. Each cell represents the corresponding bootstrap value. Sections within each row (sample) represents clones within each sample. (d) Distribution of degree (top) and betweenness (bottom) outliers, across all TNBC samples for the Interferon/MHC-II meta-program associated TFs. Each cell represents the fraction of clone pairings a gene has been found to be an outlier in.

or betweenness centrality across samples and clones (Fig. 7d, Supplementary Fig. S7). Importantly, some of the TFs (*KLF10* within the Interferon/MHC-II meta-program) or TGs (*PSAT1*, a target of *DDIT3* in the stress-response meta-program) associated with the metaprograms were further found to have significant effects on patient survival in TNBC patients from TCGA cohort [59] (Supplementary Fig. S9a).

Clonal regulatory architecture across glioblastoma meta-programs

We next performed the regulon analysis on 11 glioblastoma samples [58, 82] and again identified three categories of regulons consistent with our observation in TNBC (Fig. 8a). Regulons that showed differential activity across clones in multiple samples included oncogenic drivers *HOXB7* (promotes tumor proliferation and associated with Wnt/ β -catenin pathway) [94], and *BCL11B* (maintains stemness and helps bypass senescence) [43], as well as the angiogenic factor *NR2F2* [70]. Different tumor suppressors such as *EBF1* [44] and *PEG3* [36] showed sample-specific activities. Regulons which were broadly active across clones and patients included neurodevelopmental TFs (*DLX1*, *DLX2*, *SOX11*, *NEUROG1*) [58], TFs associated with mesenchymal and stemness program (*FOXP1*,

WWTR1) [8] and TFs associated with proliferation (*HMGB2*) and chromatin-remodeling (*SMARCC1*) [78]. Interestingly, *SOX11* also exhibited copy-number aberrations in the same datasets where its regulon activity was found to be differentially regulated. Different pathways were enriched for the three classes of regulons (Fig. 8b). TROP2 regulatory signaling [32], Ras signaling [31], Wnt signaling [41] and EGFR tyrosine kinase inhibitor resistance [64] were enriched in the regulons showing differential clonal activity. On the other hand, regulons active across clones and samples enriched for TGF-beta signaling [28], IL26 signaling [88], Notch signaling [7] and Neural Progenitor-like cell state associated pathways [46].

Glioblastoma-related meta-programs [58] were also enriched within the category of multi-sample regulons. In particular, we identified the NPC-1 and NPC-2 meta-programs and examined the structural impact of clonal variation on these gene meta-programs across samples. We observed regulatory edges that were selectively active across clones and samples, with varying bootstrap support values (Fig. 8c). TFs (*TCF12*, *OLIG1*) involved in these edges are known to play vital roles in gliomagenesis [55]. *TCF12* and *OLIG1* have been reported to be co-expressed within oligodendrocyte progenitor-like lineages in GBM, and *Tcf3/Tcf12* double knockout driven by *Olig1*-Cre directly impairs OPC generation and oligodendrocyte differentiation, supporting a functional role for *TCF12* as an E-protein partner within the *OLIG1*-associated transcriptional program [84]. Several of these TFs also exhibited outlier degree or betweenness centrality across clones and samples, suggesting clone-specific regulatory influence within these programs (Fig. 8d). Similar trends were also observed in sub-networks corresponding to other meta-programs in GBM (Supplementary Fig. S8). For example, the interaction *DDIT3:ATF3* was active in multiple samples for the MES2 meta-program, both factors are known to be co-induced in the pseudopalisading necrotic zones of GBM where ER stress and hypoxia converge to drive the MES-like cellular state [35]. Upon evaluating gene importance in the TCGA GBM dataset, [9] we observed that some interactions involved genes with significant associations with patient survival, including *FTL* (MES-2 meta-program), *SOX11* (NPC-1 and NPC-2 metaprograms) (Supplementary Fig. S9b).

Discussion

In this study, we systematically investigated the impact of ITH on GRN architecture across multiple cancer types. By inferring clone-specific GRNs from scRNA-seq data across a pilot cohort of triple-negative breast cancer, colorectal cancer, and glioblastoma datasets, an extended cohort of 253 patients spanning 12 cancer types, and two longitudinal treatment-response cohorts (a single stage IV breast cancer patient and an 11-patient ovarian cancer cohort profiled before and after neoadjuvant chemotherapy), we demonstrate that tumor clones harbor distinct regulatory architectures that differ from both normal cell populations and from one another, suggesting that regulatory rewiring is an important component of tumor evolution. A central finding of our study is that tumor clones exhibit GRN architectures more similar to each other than to normal cells. These differences were consistent across tumor types, supported by multiple metrics and high bootstrap support, indicating robust and reproducible regulatory interactions. The same pattern held in the 253-patient extended cohort: clone-clone pairs were again more similar than clone-normal pairs by both Jaccard and Hamming

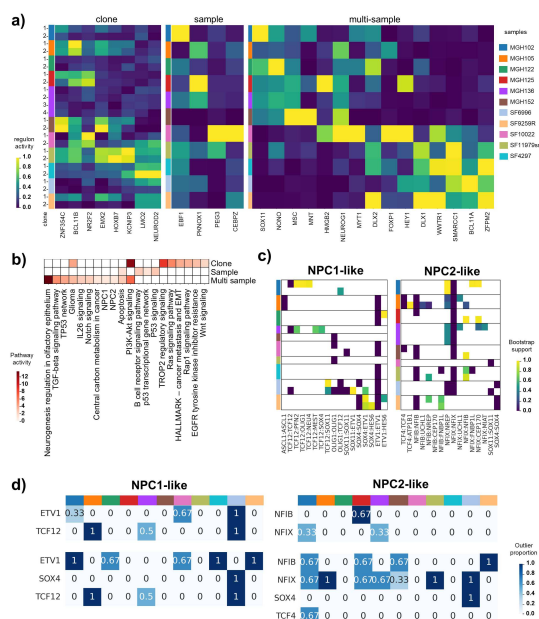


Figure 8 Regulon categories reveal distinct patterns of clone- and sample-specific activity, pathway enrichment, and meta-program-dependent GRN variability across Glioblastoma samples. (a) Activity of three Regulon categories across all Glioblastoma samples: (left to right) regulons that exhibit differential activity across clones in multiple samples, regulons displaying sample-specific activity and regulons that do not vary substantially across clones but show broad activity across multiple samples. x-axis denotes the Regulons, y-axis denotes the samples. (b) Enriched pathways and corresponding pathway enrichment scores for the three regulon categories. (c) Distribution of SCENIC edges across all Glioblastoma samples for meta-programs NPC-1 and NPC-2 respectively. Each cell represents the corresponding bootstrap value. Sections within each row (sample) represents clones within each sample. (d) (left to right) Distribution of degree and betweenness outliers, respectively, across all Glioblastoma samples for NPC-1 TFs and NPC-2 TFs. Each cell represents the fraction of clone pairings a gene has been found to be an outlier in.

measures, and the intra-clone-smaller-than-inter-clone divergence and the deviation from shuffled-label null models were observed across all 12 cancer types, although the magnitude of the null deviation varied by cancer type. Edges unique to clone-specific GRNs tended to have lower bootstrap support than shared edges, consistent with the expectation that clone-specific regulatory interactions are more labile or context-dependent than widely shared ones.

Another key finding of this study is the association between regulatory divergence and clonal dominance. We developed a novel Wasserstein distance framework that captures graded differences in regulatory strength by comparing distributions of edge weights across transcription factors, thereby providing a global measure of weighted regulatory divergence that complements binary metrics such as Jaccard and Hamming distances. Using this Wasserstein distance framework, we found that the dominant clonal subpopulation within each tumor exhibited significantly greater regulatory divergence from normal cells compared to smaller clones, supported by both binomial and Wilcoxon tests, indicating a mechanistic connection between regulatory rewiring and clonal fitness. We interpret this as clones undergoing greater regulatory reprogramming acquiring transcriptional states that confer selective advantages, potentially through activation of proliferative, anti-apoptotic, or immune-evasive

programs [29], consistent with known roles of phenotypic plasticity in cancer [50]. However, regulatory divergence may be either a driver or a consequence of clonal expansion. Longitudinal studies with matched genomic and transcriptomic profiling at the clonal level will be needed to fully disentangle these possibilities [83]. The association replicated in the extended cohort (253 patients; binomial and Wilcoxon $p < 10^{-22}$), although the correlation between clone size and divergence from normal was weak (Spearman $\rho = 0.20$), indicating that clone size explains only a small part of the variation in regulatory divergence.

Our analysis of the longitudinal stage IV breast cancer dataset suggests that treatment can induce large-scale regulatory rewiring that exceeds the variation attributable solely to somatic clonal evolution. Several transcription factors with high betweenness centrality emerged as potential drivers of these treatment-associated changes. The clone-specific betweenness analysis further showed that therapeutic pressure acts differentially on the regulatory programs of distinct clonal populations. The k^{th} order neighborhood similarity analysis further indicated that many interactions exhibited structural divergence with functional convergence [17] suggesting adaptive preservation of key transcriptional outputs under therapeutic pressure [50]. However, we acknowledge that inferences from a single longitudinal patient should be interpreted as proof-of-concept, and extension to larger longitudinal cohorts will be essential to establish generalizability.

The longitudinal ovarian cancer cohort extends this observation to several patients, with important caveats. Cross-phase clone pairs (pre- versus post-neoadjuvant chemotherapy) showed greater Wasserstein regulatory divergence than within-phase pairs, but this rests on 7 within-phase and 7 cross-phase pairs drawn from 5 of 11 patients, and we regard it as suggestive rather than definitive. Because clones were largely phase-segregated, the comparison is between different clones sampled at different times, not the same clone followed across treatment; the data therefore cannot distinguish regulatory rewiring within a clone from the preferential persistence of clones that already differed in their regulatory networks before treatment. At the gene level, most of the largest shifts in regulatory out-degree share were supported by a single patient and are exploratory. The two genes whose shift was in the same direction in both contributing patients, *MAF* (gain) and *ZEB1* (loss), are consistent with the literature on Maf-family factors and reduced platinum/PARP-inhibitor sensitivity [18] and on *ZEB1* as an EMT marker that is not a specific marker of platinum resistance [71], respectively.

The regulon-level analyses across TNBC and glioblastoma datasets revealed three distinct categories of regulatory programs: those exhibiting differential activity across clones in multiple samples, those active specifically within individual samples, and those broadly conserved across clones and samples. The pathway enrichment analysis of these regulon categories further revealed that regulatory heterogeneity across clones corresponds to functional differences in tumor biology. The enrichment of immune-related pathways among conserved regulons is particularly notable, as it suggests that the tumor immune microenvironment may exert a consistent selective pressure on regulatory programs across clones and patients, in contrast to the more stochastic activation of specific oncogenic pathways within individual clones [73]. In both TNBC and glioblastoma, regulatory edges within meta-program subnetworks displayed clone- and sample-specific presence with varying bootstrap

support, indicating that even within a defined functional program, the regulatory wiring is not static across clonal populations. The identification of transcription factors with outlier degree or betweenness centrality within meta-program subnetworks adds a network-topological layer to functional heterogeneity, highlighting clone-specific regulatory hubs as potential therapeutic vulnerabilities. This complements differential expression analysis by capturing the regulatory logic underlying transcriptional changes. Genes within meta-program subnetworks were associated with patient survival in TCGA cohorts for both TNBC and GBM, indicating that the inferred regulatory interactions reflected biologically and clinically meaningful aspects of disease progression.

A limitation we do acknowledge is that GRN inference from single-cell transcriptomic data remains sensitive to noise and modeling assumptions, and the inferred regulatory interactions will require experimental validation. While our analysis supports regulatory modeling as a feature of tumor evolution, the identification of novel TF-TG interactions is inherently constrained by the capabilities and biases of current state-of-the-art GRN inference methods. In addition, clonal populations were computationally inferred rather than determined through direct lineage tracing, which may introduce uncertainty in clone assignments. Importantly, for a subset of datasets, clonal structure was inferred solely from scRNA-seq data rather than from paired DNA-RNA measurements. This reflects a fundamental limitation: inferring clonal identity from transcriptomic data, the same modality used to derive GRNs, raising a concern of inflated associations between clonal identity and regulatory variation. To mitigate these concerns, we employed multiple safeguards. First, we focused on robust, bootstrap-supported regulatory interactions to reduce sensitivity to stochastic noise. Second, we confirmed key observations across multiple samples and cancer types, including a broader cohort that overlaps the pilot cohort by design, to check that identified patterns were reproducible rather than dataset-specific artifacts. Third, comparisons against randomized or shuffled controls were used to establish that observed structural differences exceeded those expected by chance. Another limitation is that our analysis focused on transcription factor-target gene interactions and did not account for epigenomic regulation, post-transcriptional mechanisms, or protein-level regulatory layers. Multi-omic approaches integrating single-cell ATAC-seq, proteomics, and spatial transcriptomics with scRNA-seq would enable a more comprehensive characterization of the regulatory landscape underlying intratumor heterogeneity. Spatial transcriptomics in particular could reveal how tumor microenvironment modulates clone-specific regulatory programs [72]; a dimension that is lost in dissociated single-cell assays. We selected CellOracle after comparing it with GRNBoost2 and Inferelator on a subset of datasets (see Methods; Supplementary Fig. S10), and used the copy-number profiles provided by 3CA where available, so that clone calling rests on curated data, and inferCNV, a well-established tool for CNA inference from scRNA-seq, only where no such profiles existed; no CNA-inference tools were therefore compared. A more thorough benchmarking of alternative approaches, and of how the choice of upstream method affects inferred clonal regulatory divergence, was beyond the scope of the present study and is a valuable direction for future work.

In summary, our work provides a comprehensive framework for characterizing intratumor heterogeneity at the level of gene regulatory networks. As single-cell multi-omic technologies continue to mature and computational methods for GRN inference improve,

the integration of regulatory network analysis into the standard toolkit for studying intratumor heterogeneity promises to yield deeper insights into the mechanistic basis of tumor evolution and, ultimately, to inform the design of more effective and precisely targeted therapeutic strategies.

Methods

Datasets and preprocessing

The pilot cohort comprised single-cell RNA-seq (scRNA-seq) data from 12 TNBC patients [86, 21, 63], 6 CRC patients [42], 11 glioblastoma patients [58, 82], and one stage IV breast cancer cohort comprising three longitudinal samples (pre-treatment, post-treatment 1, and post-treatment 2) [62]. The clone-versus-normal and intra-/inter-clone analyses of the pilot cohort used 15 TNBC and CRC datasets (9 TNBC and 6 CRC). Two further cohorts are described below: an extended cohort of 253 patients spanning 12 cancer types from 3CA, and a longitudinal ovarian cancer cohort of 11 patients profiled before and after neoadjuvant chemotherapy.

The TNBC, CRC and Glioblastoma datasets, along with inferred copy number profiles (CNPs), tumor/normal cell annotations, and cancer-type-specific gene meta-programs, were obtained from the Curated Cancer Cell Atlas (3CA) [80]. Clonal annotations were inferred using Leiden clustering applied to the existing copy number profiles, as suggested in inferCNV [1], with the resolution chosen per patient by visual inspection so as to obtain distinct clones. CNPs for the stage IV breast cancer cohort were inferred by applying inferCNV with settings as described in the original study [62]. Clusters containing fewer than 50 cells were excluded to ensure robustness of downstream analyses.

All datasets were preprocessed using standard workflows in SCANPY [85], including quality control, normalization, and filtering.

For the extended cohort and the longitudinal ovarian cohort introduced below, clones were called with an automated procedure applied to each patient's malignant-cell copy number profile, run with default settings throughout: cells were over-clustered with Leiden (resolution 1.0), and the resulting clusters were then iteratively merged, starting each round with the most similar remaining pair by centroid distance, unless a permutation test (200 permutations, a multivariate separation statistic between the pair's centroids, Bonferroni-adjusted $\alpha = 0.05$ across all pairs still under consideration) indicated that pair was already significantly distinct, in which case merging stopped for that pair and the next-closest remaining pair was tested. Patients with fewer than 50 malignant cells were not clustered at all; any resulting cluster retaining fewer than 50 cells after merging was dropped; and patients retaining fewer than 2 clusters after these filters were excluded from the corresponding analysis entirely, since no within-patient clone comparison is possible from a single clone.

To test the generalizability of the clone-versus-normal and intra-/inter-clone findings described above beyond the 15-dataset pilot analyses (9 TNBC and 6 CRC datasets), we additionally drew an extended cohort of 253 patients spanning 12 cancer types (brain, breast, colorectal, head and neck, hematologic, kidney, lung, neuroendocrine, ovarian, pancreatic, prostate, and skin), also from the Curated Cancer Cell Atlas (3CA) [80] (datasets and patients listed in Supplementary Tables S1–S3); per-patient clone-size

distributions and per-cancer-type bootstrap-support and Hamming-distance results are shown in Supplementary Figs. S11–S18: the same 3CA-provided copy number profiles as the pilot cohort, with clones called via the automated procedure above, followed by the identical GRN-inference, bootstrap, and shuffle-null pipeline described below (the separately-analyzed longitudinal EOC/NACT treatment-phase cohort was excluded from this pooled extended cohort). The extended cohort deliberately includes the pilot TNBC, CRC, and glioblastoma datasets: we first characterized TNBC and CRC in detail and then extended the same analyses to all datasets available in 3CA with sufficient clone structure. The cohort comprises 253 patients; the exact number of patients contributing varies slightly by analysis (e.g., 252 for the clone-clone versus clone-normal Hamming comparison, Fig. 2c, and 253 for the Wasserstein clone-size-divergence analysis, Fig. 3c, which requires a Normal population and ≥ 2 differently-sized non-Normal clones).

To examine treatment-associated regulatory change, we separately obtained the longitudinal ovarian cancer scRNA-seq cohort of Zhang et al. [92] (Gene Expression Omnibus accession GSE165897), comprising 11 patients (EOC1005, EOC136, EOC153, EOC227, EOC349, EOC372, EOC3, EOC443, EOC540, EOC733, EOC87) each profiled at two timepoints: a treatment-naive sample from primary debulking surgery (“pre”) and a post-neoadjuvant-chemotherapy sample from interval debulking surgery (“post”). This cohort is not part of 3CA, so copy number profiles were inferred directly with inferCNV, using parameters matching the original study’s own Methods [92]; clones were then called with the automated Leiden overcluster-and-merge procedure described above. Because clonal annotation was performed on the pooled pre- and post-treatment cells for each patient, individual clones were generally found to be phase-segregated (i.e., composed almost entirely of cells from one timepoint) consistent with chemotherapy-driven clonal selection as reported in the original study. This cohort was analyzed separately using the clone-to-phase assignment and treatment-phase comparison procedure described below (Methods), and was excluded from the pooled extended cohort described above.

Inference of gene regulatory networks

GRNs were inferred for each clone using CellOracle [37], a framework specifically designed for single-cell transcriptomic data that integrates gene expression with prior regulatory information. We initially evaluated three widely used GRN inference methods: CellOracle, GRNBoost2 (as part of the SCENIC pipeline) [5, 54], and Inferelator [76]; all of which have been benchmarked extensively and shown to perform competitively across diverse settings in single-cell and bulk transcriptomic data [67, 76, 77, 60]. All methods were executed using default parameters. For Inferelator, GRNBoost2-derived networks were used as prior information due to the absence of curated priors. Across multiple TNBC and CRC datasets, GRNs inferred using GRNBoost2 and Inferelator exhibited limited overlap between tumor-wide and clone-specific networks (Supplementary Fig. S10), indicating reduced consistency at the clonal level. This observation is consistent with prior benchmarking studies demonstrating that Inferelator performance can vary substantially across model configurations and is particularly sensitive to the availability and quality of prior information. In particular, Inferelator has been shown to exhibit high variability across different model selection strategies (e.g., BBSR, StARS,

AMuSR), with performance depending strongly on preprocessing and prior specification [77]. More broadly, GRN inference methods that incorporate prior biological knowledge have been shown to yield more stable and biologically meaningful networks compared to expression-only approaches such as GRNBoost2 [24]. Consistent with these observations, CellOracle’s integration of prior regulatory information and ensemble regression framework provides improved robustness in single-cell settings. Consequently, all downstream analyses were restricted to CellOracle-derived GRNs. GRNs were generated using CellOracle’s workflow and only edges with p -value < 0.05 were considered.

Each GRN is represented as a directed weighted graph $G = (V, E, W)$, where V denotes the set of genes, E denotes the set of regulatory interactions, and W denotes edge weights reflecting regulatory strength.

Bootstrap and null model analysis

To assess the robustness of inferred regulatory interactions, we performed bootstrap resampling for each clone. Specifically, 100 subsamples were generated without replacement using fractions $f \in \{0.75, 0.9, 0.95\}$ of cells within each clone. GRNs were inferred for each subsample using the same pipeline as the full dataset. Edge support was quantified as the proportion of bootstrap replicates in which a given edge was recovered. The 90% subsamples were used for the intra-/inter-clone comparisons.

To construct a null model, clone labels were randomly permuted 100 times within each dataset, and GRNs were inferred for each randomized partition. This procedure provides a baseline for assessing whether observed GRN differences arise from structured biological variation rather than arbitrary partitioning.

Graph similarity and regulatory divergence measures

To quantify similarity between GRNs, we employed both global and local measures of network similarity.

Global similarity

Global similarity between two GRNs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ was computed using Jaccard similarity of edges:

$$J(G_1, G_2) = \frac{|E_1 \cap E_2|}{|E_1 \cup E_2|}.$$

The corresponding Hamming distance was computed on binarized edge-presence vectors defined over the union of edges across all compared populations of a patient, as the proportion of positions at which the two vectors differ. It is therefore related to, but not identical to, $1 - J$, since J normalizes by the union of the two compared networks only.

Local similarity

To capture node-level differences, we computed Jaccard similarity over local neighborhoods of genes. For a node $v \in V_1 \cap V_2$, let $N_i^{\text{succ}}(v)$ and $N_i^{\text{pred}}(v)$ denote its outgoing and incoming neighbors in G_i . Local similarities were defined as:

$$J^{\text{succ}}(v) = \frac{|N_1^{\text{succ}}(v) \cap N_2^{\text{succ}}(v)|}{|N_1^{\text{succ}}(v) \cup N_2^{\text{succ}}(v)|}, \quad J^{\text{pred}}(v) = \frac{|N_1^{\text{pred}}(v) \cap N_2^{\text{pred}}(v)|}{|N_1^{\text{pred}}(v) \cup N_2^{\text{pred}}(v)|}.$$

Wasserstein distance-based regulatory divergence

To capture graded differences in regulatory strength beyond binary edge overlap, we introduce a weighted divergence measure based on the 1-Wasserstein distance.

For each transcription factor $v \in V_1 \cap V_2$, let $\mathbf{w}_i(v) \in \mathbb{R}^k$ be the vector of outgoing edge weights from v in clone i , defined over the union of targets $k = |N_1^{\text{succ}}(v) \cup N_2^{\text{succ}}(v)|$, with zero entries for edges absent in a given clone. Let $F_i(\cdot)$ denote the empirical cumulative distribution function (CDF) of $\mathbf{w}_i(v)$ after ℓ_1 -normalization. The regulatory divergence of TF v between the two clones is then defined as the 1-Wasserstein distance between these distributions:

$$W_1(v) = \int_0^1 |F_1^{-1}(t) - F_2^{-1}(t)| dt,$$

where F_i^{-1} denotes the quantile function of F_i . Geometrically, $W_1(v)$ is the minimum total “work” required to transform the outgoing regulatory weight profile of TF v in clone 1 into that of clone 2. A global summary of weighted regulatory divergence across the network is given by

$$\bar{W} = \frac{1}{|V_1 \cap V_2|} \sum_{v \in V_1 \cap V_2} W_1(v).$$

k -th order neighborhood similarity

For a node v and clone i , define the k -th order successor neighborhood recursively:

$$N_i^{(1)}(v) = N_i^{\text{succ}}(v), \quad N_i^{(k)}(v) = \bigcup_{u \in N_i^{(k-1)}(v)} N_i^{\text{succ}}(u) \setminus \{v\}, \quad k \geq 2.$$

The k -th order Jaccard similarity of node v between the two clones is then:

$$J^{(k)}(v) = \frac{|N_1^{(k)}(v) \cap N_2^{(k)}(v)|}{|N_1^{(k)}(v) \cup N_2^{(k)}(v)|}.$$

We define the similarity profile of TF v up to order K as the vector

$$\mathbf{J}(v) = (J^{(1)}(v), J^{(2)}(v), \dots, J^{(K)}(v)) \in [0, 1]^K.$$

This profile supports a biologically meaningful taxonomy of TF divergence patterns:

- Structurally divergent, functionally convergent: $J^{(1)}(v)$ low, $J^{(K)}(v)$ high. The TF drives different direct targets in each clone, but these targets converge on shared downstream effectors.
- Structurally convergent, functionally divergent: $J^{(1)}(v)$ high, $J^{(K)}(v)$ low. Direct wiring is similar across clones, but downstream regulatory consequences diverge substantially.
- Globally divergent: $J^{(k)}(v)$ low for all k , indicating pervasive rewiring of the regulatory program of TF v .

Centrality analysis

To identify key regulatory genes within the inferred GRNs, we computed degree centrality and betweenness centrality using the NetworkX Python package (v2.8.8) [27].

Degree Centrality.

The degree centrality of a node $v \in V$ in a network $G = (V, E)$ is defined as the number of edges adjacent to v , normalized by the

maximum possible degree:

$$C_{\text{deg}}(v) = \frac{\deg(v)}{|V| - 1},$$

where $\deg(v)$ is the number of edges adjacent to node v . This measure captures the local connectivity of a gene within the regulatory network.

Betweenness Centrality.

The betweenness centrality of a node $v \in V$ quantifies the fraction of shortest paths between all pairs of nodes that pass through v :

$$C_{\text{bet}}(v) = \sum_{\substack{s, t \in V \\ s \neq t \neq v}} \frac{\sigma_{st}(v)}{\sigma_{st}},$$

where σ_{st} is the total number of shortest paths between nodes s and t , and $\sigma_{st}(v)$ is the number of those paths that pass through node v . Betweenness centrality captures the importance of a gene in mediating regulatory interactions across the network.

Regulon activity and pathway enrichment

Regulon Activity Scoring.

Regulons defined as TF–TG sets derived from either clone-specific GRNs inferred using SCENIC [5] over CellOracle GRN or curated TF–TG interactions from the CollecTRI database [56]. This integration ensures that only biologically supported regulatory interactions are included in each regulon. For each cell in the single-cell RNA-seq dataset, regulon activity scores were computed using the AUCCell framework [5] as implemented in the decoupler package (v1.9.2) [6].

Let G_c denote the ranked gene expression vector for cell c , and $R = \{g_1, g_2, \dots, g_m\}$ be the set of target genes for a regulon. AUCCell computes the area under the recovery curve (AUC) as:

$$\text{AUCCell}_c(R) = \frac{1}{|R|} \sum_{i=1}^{|G_c|} \mathbf{1}(G_c[i] \in R) \cdot \text{rank}(G_c[i])$$

where $\mathbf{1}(\cdot)$ is the indicator function, and $\text{rank}(G_c[i])$ is the rank of gene i in cell c (higher rank = higher expression). Higher AUCCell scores indicate stronger regulatory activity of the TF in that cell.

Regulon Activity Classification.

A regulon was considered differentially active across clones if its AUCCell scores were significantly different between at least two clones in the same sample (Wilcoxon rank-sum test, $p < 0.05$ after Benjamini-Hochberg correction). Regulons were classified as sample-specific if they showed high activity within a sample but not in other samples. In both clone and sample-specific differentially active regulons, this was enforced by applying restrictions on the number of cells the regulons were active in each group (clone/sample). Specifically, the regulon must be active in at least $\mathcal{H}\%$ of cells in the group (clone/sample) that the regulon is highly active in and should be active in at most $\mathcal{L}\%$ of cells in other groups. Depending on the cancer type, $\mathcal{H} \in [40, 50]$ and $\mathcal{L} \in [20, 35]$. Based on these comparisons, three categories of regulons were defined:

1. **Clone-specific regulons:** Highly active in a subset of clones within a sample across multiple samples, driving clonal functional heterogeneity.

2. **Sample-specific regulons:** Highly active in a sample but not in other samples, reflecting sample-level regulatory programs.
3. **Multi-sample regulons:** Observed consistently across multiple samples and clones of a given cancer type, indicating broadly conserved regulatory programs.

Functional Enrichment Analysis.

To identify biological processes, pathways or meta-programs overrepresented among the genes in each regulon or differentially active regulons, functional enrichment analysis was performed using g:Profiler [40]. Specifically, regulon target genes were tested for enrichment against KEGG [38], HALLMARK [45], WIKIPATHWAY [3]. This approach allowed us to link TF activity patterns to functional programs and interpret the potential phenotypic consequences of clonal regulatory heterogeneity.

Meta-program-specific network analysis

To investigate regulatory structures within biologically relevant gene programs, we leveraged cancer-type-specific meta-programs corresponding to the tumor types analyzed (TNBC and glioblastoma). For TNBC, we utilized predefined meta-programs from the Curated Cancer Cell Atlas (3CA) [80], each comprising approximately 50 genes and associated with malignant cell states, including epithelial-mesenchymal transition (EMT1-3), respiration, translation, cell cycle, stress response, and lineage-specific differentiation amongst others. We first restricted our selection of meta-programs to those prevalent in the TNBC samples considered. We then further filtered the meta-programs to those unique to the samples considered. This left us with 11 meta-programs namely: ‘Cell Cycle - G1/S’, ‘Cell Cycle HMG-rich’, ‘Stress (in vitro)’, ‘Unfolded protein response 1’, ‘Protein maturation’, ‘Translation initiation’, ‘EMT-II’, ‘Interferon/MHC-II (II)’, ‘Epithelial Senescence (HNSCC)’, ‘PDAC-related 5’ and ‘Adherens’. For glioblastoma, meta-programs (39–50 genes each) were obtained from 3CA and the original study by [58], representing developmental cell states such as astrocyte-like (AC-like), oligodendrocyte progenitor-like (OPC-like), neural progenitor-like (NPC-like), and mesenchymal-like (MES-like) programs. Following the same filtering procedure as with TNBC, we were left with 12 meta-programs namely: ‘MES2’, ‘MES1’, ‘AC’, ‘OPC’, ‘NPC1’, ‘NPC2’, ‘G1/S’, ‘G2/M’, ‘Stress (in vitro)’, ‘Interferon/MHC-II (I)’, ‘Interferon/MHC-II (II)’, ‘Glioma single-nucleus’.

For each sample, clone-specific GRNs were inferred using SCENIC in conjunction with CellOracle. For each meta-program, the inferred clone-specific GRNs across all samples were restricted to the subset of genes associated with that program, yielding program-specific subnetworks. Edges within these subnetworks were annotated with bootstrap support values derived from the bootstrap experiment, representing the fraction of bootstrap subsamples in which each edge was observed. This enabled a quantitative assessment of regulatory variability both across clones within a sample and across samples.

To further characterize the role of transcription factors (TFs) within these meta-programs, we evaluated the extent to which they exhibited outlier behavior in degree and betweenness centrality. For each sample, pairwise comparisons were performed across all clone combinations, and a linear model was fit between the normalized centrality values of the corresponding clones. A TF was classified as an outlier if its centrality deviated by at least three standard

deviations from the expected value under this model in either clone. An outlier proportion was then computed for each TF as the ratio of the number of clone pairings in which the TF was identified as an outlier to the total number of possible clone pairings within the sample.

This framework enables systematic evaluation of how GRN architecture and transcriptional regulatory activity vary within and across clones and samples, while anchoring these differences to biologically meaningful gene programs.

Cancer patient survival analysis

Patient-level survival analyses were performed using gene expression and clinical data obtained from The Cancer Genome Atlas (TCGA) via the UCSC Xena platform [23]. Normalized gene expression data (RSEM, Toil recompute) and corresponding clinical annotations were retrieved for glioblastoma (GBM) and breast cancer (BRCA) cohorts. Samples were matched across datasets using harmonized TCGA barcodes (first 15 characters), ensuring consistent alignment between expression and clinical records.

For breast cancer, triple-negative breast cancer (TNBC) samples were defined by filtering clinical annotations for estrogen receptor-negative, progesterone receptor-negative, and HER2-negative status. Survival time T_i for each patient i was defined as:

$$T_i = \begin{cases} \text{days.to.death}, & \text{if available} \\ \text{days.to.last.followup}, & \text{otherwise} \end{cases} \quad (1)$$

Event status δ_i was defined as a binary indicator of death:

$$\delta_i = \begin{cases} 1, & \text{if death occurred} \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

Gene expressions were obtained from matched TCGA bulk samples. Patients were stratified into high and low groups based on gene program enrichment using predefined thresholds (median or quantile-based cutoffs).

Survival distributions for the two groups were estimated using the Kaplan–Meier estimator [39]:

$$\hat{S}(t) = \prod_{t_j \leq t} \left(1 - \frac{d_j}{n_j}\right), \quad (3)$$

where d_j denotes the number of events and n_j the number of individuals at risk at time t_j .

Differences between survival curves were assessed using the log-rank test [49]. For two groups, the test statistic is given by:

$$Z = \frac{\sum_j (O_{1j} - E_{1j})}{\sqrt{\sum_j V_{1j}}}, \quad (4)$$

where O_{1j} and E_{1j} denote the observed and expected number of events in group 1 at time t_j , and V_{1j} is the corresponding variance. The statistic asymptotically follows a standard normal distribution, from which the corresponding p -value was computed. Statistical significance was determined based on the log-rank p -value.

Treatment-phase (pre-/post-NACT) regulatory network comparison

Clone-to-phase assignment.

For each of the 11 Zhang et al. [92] patients, every non-Normal clone's cells were cross-tabulated against the sample-level treatment-phase annotation (treatment-naive vs. post-NACT). A clone was assigned to the "pre" or "post" phase if it contained at least 50 cells (an absolute count, not a purity percentage) from that phase and fewer than 50 from the other; a clone with ≥ 50 cells in *both* phases was labeled "mixed" and excluded (its phase could not be cleanly attributed), and a clone with < 50 cells in both was labeled "insufficient_cells" and likewise excluded. This absolute-count rule was chosen over a purity threshold so that a clone large enough to trust in one phase was not excluded merely because the other phase also cleared an arbitrary percentage; all per-clone assignments are reported in a patient-level audit manifest.

Within-phase vs. cross-phase clone-pair divergence.

For each patient with ≥ 2 usable (non-mixed, non-insufficient) clones, every pairwise combination of clones was classified as "within-phase" (both pre, or both post) or "cross-phase" (one pre, one post), and compared using the same original-GRN 1-Wasserstein divergence measure defined above (Sections 4.2–4.6); consistent with its use as the primary GRN-divergence measure for the single-patient longitudinal analysis above, pooled across all contributing patients and compared via the Mann-Whitney U test.

Gene-level regulatory out-degree shift.

For each patient with at least one usable pre-clone *and* at least one usable post-clone (a weaker requirement than the pairwise comparison above, since it needs only one clone per phase), all pre-clones' weighted edges were combined into a single "pre" network and all post-clones' into a single "post" network (an edge recurring across multiple contributing clones had its weight averaged, not summed, across those clones). For each gene present in either network, its out-degree (number of regulatory targets) was computed in both phases. Because a phase network's total edge count depends on how many clones were combined into it (a set-theoretic effect of taking a union over more vs. fewer clones, independent of any underlying biology), each gene's out-degree was additionally expressed as a *share* of that phase network's total edge count, and genes were ranked by the pre-to-post change in this share (Δ out-degree share, percentage points) rather than by raw degree change. A gene with zero out-degree in one phase and nonzero out-degree in the other was additionally flagged as a "de novo emergence" event (gained or lost), since this is a qualitatively different claim (a regulator appearing or disappearing entirely) from an incremental reweighting of an already-active regulator.

Conclusion

Our work provides a comprehensive framework for characterizing intratumor heterogeneity at the level of gene regulatory networks. As single-cell multi-omic technologies continue to mature and computational methods for GRN inference improve, the integration of regulatory network analysis into the standard toolkit for studying intratumor heterogeneity promises to yield deeper insights into the mechanistic basis of tumor evolution and, ultimately, to inform the design of more effective and precisely targeted therapeutic strategies.

The consistency of the clone-versus-normal and intra-/inter-clone findings across 253 patients and 12 cancer types, together with suggestive evidence of treatment-associated regulatory divergence in ovarian cancer, supports the use of clone-resolved regulatory networks as a lens on tumor evolution, while the limitations above indicate where larger and better-controlled studies are needed.

Conflicts of interest

The authors declare no competing interests.

Funding

This research was supported in part by the National Science Foundation, grants IIS-1812822 and IIS-2106837 (L.N.) and the DBT/Wellcome Trust India Alliance grant IA/E/21/1/506298 (H.Z.).

Data availability

All data used in this study are publicly available. Single-cell transcriptomic datasets comprising 12 TNBC patients [86, 21, 63], 6 CRC patients [42], 11 glioblastoma patients [58, 82]; along with inferred CNPs, tumor/normal cell annotations, and cancer-type-specific gene meta-programs, were obtained from the Curated Cancer Cell Atlas (3CA) [80] and one stage IV breast cancer cohort comprising three longitudinal samples (pre-treatment, post-treatment 1, and post-treatment 2) [62]. The extended cohort of 253 patients spanning 12 cancer types was likewise obtained from 3CA [80]. The longitudinal ovarian cancer treatment-phase cohort of 11 patients is publicly available from the Gene Expression Omnibus under accession GSE165897 [92].

All code used to reproduce the analyses, including preprocessing, gene regulatory network inference, and downstream statistical analyses, is available on GitHub: <https://github.com/narein97/Intratumor-Heterogeneity-Through-the-Lens-of-Gene-Regulatory-Networks>

Author contributions statement

H.Z. and L.N. conceived the study. N.R., N.S., H.Z., and L.N. developed the methodology. N.R. implemented the software, performed the experiments and ran baseline methods under the supervision of N.S., H.Z., and L.N. N.R., N.S., and H.Z. analyzed the results. N.R., N.S., and H.Z. drafted the manuscript. All authors reviewed, edited, and approved the final version of the manuscript.

Acknowledgments

Not applicable.

References

1. inferCNV of the Trinity CTAT Project. <https://github.com/broadinstitute/inferCNV>. URL <https://github.com/broadinstitute/inferCNV>.
2. A. Abdollahi, D. Pisarcik, D. Roberts, J. Weinstein, P. Cairns, and T. C. Hamilton. Lot1 (plagl1/zac1), the candidate tumor suppressor gene at chromosome 6q24–25,

- is epigenetically regulated in cancer. *Journal of Biological Chemistry*, 278(8):6041–6049, 2003.
3. A. Agrawal, H. Balci, K. Hanspers, S. L. Coort, M. Martens, D. N. Slenter, F. Ehrhart, D. Digles, A. Waagmeester, I. Wassink, et al. Wikipathways 2024: next generation pathway database. *Nucleic acids research*, 52(D1):D679–D689, 2024.
 4. D. Ai, J. Yao, F. Yang, L. Huo, H. Chen, W. Lu, L. M. S. Soto, M. Jiang, M. G. Raso, S. Wang, D. Bell, J. Liu, H. Wang, D. Tan, C. Torres-Cabala, Q. Gan, Y. Wu, C. Albarracin, M.-C. Hung, F. Meric-Bernstam, I. I. Wistuba, V. G. Prieto, A. A. Sahin, and Q. Ding. TRPS1: a highly sensitive and specific marker for breast carcinoma, especially for triple-negative breast cancer. *Mod. Pathol.*, 34(4):710–719, Apr. 2021.
 5. S. Aibar, C. B. González-Blas, T. Moerman, V. A. Huynh-Thu, H. Imrichova, G. Hulselmans, F. Rambow, J.-C. Marine, P. Geurts, J. Aerts, et al. Scenic: single-cell regulatory network inference and clustering. *Nature Methods*, 14(11):1083–1086, 2017.
 6. P. Badia-i Mompel, J. Vélez Santiago, J. Braunger, C. Geiss, D. Dimitrov, S. Müller-Dott, P. Taus, A. Dugourd, C. H. Holland, R. O. Ramirez Flores, and J. Saez-Rodriguez. decoupler: ensemble of computational methods to infer biological activities from omics data. *Bioinformatics Advances*, 2(1):vbac016, 03 2022. ISSN 2635-0041. doi: 10.1093/bioadv/vbac016. URL <https://doi.org/10.1093/bioadv/vbac016>.
 7. R. Bazzoni and A. Bentivegna. Role of notch signaling pathway in glioblastoma pathogenesis. *Cancers*, 11(3):292, 2019.
 8. K. P. L. Bhat, V. Balasubramaniyan, B. Vaillant, R. Ezhilarasan, K. Hummelink, F. Hollingsworth, K. Wani, L. Heathcock, J. D. James, L. D. Goodman, S. Conroy, L. Long, N. Lelic, S. Wang, J. Gumin, D. Raj, Y. Kodama, A. Raghunathan, A. Olar, K. Joshi, C. E. Pelloski, A. Heimberger, S. H. Kim, D. P. Cahill, G. Rao, W. F. A. Den Dunnen, H. W. G. M. Boddeke, H. S. Phillips, I. Nakano, F. F. Lang, H. Colman, E. P. Sulman, and K. Aldape. Mesenchymal differentiation mediated by NF- κ B promotes radiation resistance in glioblastoma. *Cancer Cell*, 24(3):331–346, Sept. 2013.
 9. C. W. Brennan, R. G. Verhaak, A. McKenna, B. Campos, H. Nounshmehr, S. R. Salama, S. Zheng, D. Chakravarty, J. Z. Sanborn, S. H. Berman, et al. The somatic genomic landscape of glioblastoma. *cell*, 155(2):462–477, 2013.
 10. R. A. Burrell, N. McGranahan, J. Bartek, and C. Swanton. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature*, 501(7467):338–345, 2013.
 11. Z. Chen, F. Gong, L. Wan, and L. Ma. Bitsc 2: Bayesian inference of tumor clonal tree by joint analysis of single-cell snv and cna data. *Briefings in Bioinformatics*, 23(3):bbac092, 2022.
 12. F. Cheng, H. Hong, S. Yang, and Y. Wei. Individualized network-based drug repositioning infrastructure for precision oncology in the panomics era. *Briefings in bioinformatics*, 18(4):682–697, 2017.
 13. S. A. Chowdhury, E. M. Gertz, D. Wangsa, K. Heselmeyer-Haddad, T. Ried, A. A. Schäffer, and R. Schwartz. Inferring models of multiscale copy number evolution for single-tumor phylogenetics. *Bioinformatics*, 31(12):i258–i267, 2015.
 14. B. Coban, Z. Wang, C. yi Liao, K. Beslmüller, M. A. Timmermans, J. W. Martens, J. H. Hundscheid, B. Slutter, A. J. Zweemer, E. Neubert, and E. H. Danen. Grhl2 suppression of nt5e/cd73 in breast cancer cells modulates cd73-mediated adenosine production and t-cell recruitment. *iScience*, 27(5):109738, 2024. ISSN 2589-0042. doi: <https://doi.org/10.1016/j.isci.2024.109738>. URL <https://www.sciencedirect.com/science/article/pii/S258900422400960X>.
 15. Z. Dan, H. Xiujiang, L. Ting, Z. Xiaorong, Z. Hong, Y. Jiqiao, L. Yanchu, and J. Jing. Long non-coding rna btg3-7: 1 and jund co-regulate c21orf91 to promote triple-negative breast cancer progress. *Frontiers in Molecular Biosciences*, 7: 605623, 2021.
 16. M. Dautle, S. Zhang, and Y. Chen. sctiger: A deep-learning method for inferring gene regulatory networks from case versus control scrna-seq datasets. *International Journal of Molecular Sciences*, 24(17), 2023. ISSN 1422-0067. doi: 10.3390/ijms241713339. URL <https://www.mdpi.com/1422-0067/24/17/13339>.
 17. A. de la Fuente. From 'differential expression' to 'differential networking' –identification of dysfunctional regulatory networks in diseases. *Trends in genetics*, 26(7): 326–333, 2010.
 18. Y. Deng, L. Lu, H. Zhang, Y. Fu, T. Liu, and Y. Chen. The role and regulation of maf proteins in cancer. *Biomarker Research*, 11(1):17, 2023. doi: 10.1186/s40364-023-00457-w.
 19. M. Edrisi, M. V. Valecha, S. B. Chowdary, S. Robledo, H. A. Ogilvie, D. Posada, H. Zafar, and L. Nakhleh. Phylovar: toward scalable phylogeny-aware inference of single-nucleotide variations from single-cell dna sequencing data. *Bioinformatics*, 38(Supplement_1):i195–i202, 2022.
 20. M. Fournier, J. Javary, V. Roh, N. Fournier, and F. Radtke. Reciprocal inhibition of notch and sox2 shapes tumor cell plasticity and therapeutic escape in triple-negative breast cancer. *EMBO Molecular Medicine*, 16(12):3184–3217, 12 2024. ISSN 1757-4684. doi: 10.1038/s44321-024-00161-8. URL <https://doi.org/10.1038/s44321-024-00161-8>.
 21. R. Gao, S. Bai, Y. C. Henderson, Y. Lin, A. Schalck, Y. Yan, T. Kumar, M. Hu, E. Sei, A. Davis, et al. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. *Nature Biotechnology*, 39(5): 599–608, 2021.
 22. M. V. Giuli, E. Giuliani, I. Screpanti, D. Bellavia, and S. Checquolo. Notch signaling activation as a hallmark for triple-negative breast cancer subtype. *Journal of Oncology*, 2019(1):8707053, 2019. doi: <https://doi.org/10.1155/2019/8707053>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1155/2019/8707053>.
 23. M. J. Goldman, B. Craft, M. Hastie, K. Repecka, F. McDade, A. Kamath, A. Banerjee, Y. Luo, D. Rogers, A. N. Brooks, et al. The ucsc xena platform for public and private cancer genomics data visualization and interpretation. *Nature Biotechnology*, 38(6):675–678, 2020.
 24. A. Greenfield, C. Hafemeister, and R. Bonneau. Robust data-driven incorporation of prior knowledge into the inference of dynamic regulatory networks. *Bioinformatics*, 29(8):1060–1067, 2013.
 25. X. Guan, F. Polesso, C. Wang, A. Sehwat, R. M. Hawkins, S. E. Murray, G. V. Thomas, B. Caruso, R. F. Thompson, M. A. Wood, et al. Androgen receptor activity in t cells

- limits checkpoint blockade efficacy. *Nature*, 606(7915):791–796, 2022.
26. L. Guo, Y. Rao, Y. Song, J. Hu, Z. Luo, S. Sun, C. Chen, and D. Kong. Six1 transmits signals for breast cancer progression via the zeb1/il6/stat3 signaling axis. *Oncogene*, 44(37):3493–3511, 09 2025. ISSN 1476-5594. doi: 10.1038/s41388-025-03499-6. URL <https://doi.org/10.1038/s41388-025-03499-6>.
 27. A. Hagberg, P. J. Swart, and D. A. Schult. Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Laboratory (LANL), Los Alamos, NM (United States), 2008.
 28. J. Han, C. A. Alvarez-Breckenridge, Q.-E. Wang, and J. Yu. Tgf- β signaling and its targeting for glioma treatment. *American journal of cancer research*, 5(3):945, 2015.
 29. D. Hanahan. Hallmarks of cancer: new dimensions. *Cancer discovery*, 12(1):31–46, 2022.
 30. L. G. Harris, L. K. Pannell, S. Singh, R. S. Samant, and L. A. Shevde. Increased vascularity and spontaneous metastasis of breast cancer by hedgehog signaling mediated upregulation of cyr61. *Oncogene*, 31(28):3370–3380, 2012.
 31. S. L. Holmen and B. O. Williams. Essential role for ras signaling in glioblastoma maintenance. *Cancer research*, 65(18):8250–8255, 2005.
 32. J. Hou, A. Lv, Q. Deng, G. Zhang, X. Hu, and H. Cui. Trop2 promotes the proliferation and metastasis of glioblastoma cells by activating the jak2/stat3 signaling pathway. *Oncology reports*, 41(2):753–764, 2019.
 33. J.-H. Hung, T. W. Whitfield, T.-H. Yang, Z. Hu, Z. Weng, and C. DeLisi. Identification of functional modules that correlate with phenotypic difference: the influence of network topology. *Genome biology*, 11:1–16, 2010.
 34. G. Iacono, R. Massoni-Badosa, and H. Heyn. Single-cell transcriptomics unveils gene regulatory network plasticity. *Genome Biology*, 20(1):110, 06 2019. ISSN 1474-760X. doi: 10.1186/s13059-019-1713-4. URL <https://doi.org/10.1186/s13059-019-1713-4>.
 35. A. Jauhiainen, C. Thomsen, L. Strömbom, P. Grundevik, C. Andersson, A. Danielsson, M. K. Andersson, O. Nerman, L. Rökvist, A. Ståhlberg, et al. Distinct cytoplasmic and nuclear functions of the stress induced protein ddit3/chop/gadd153. *PloS one*, 7(4):e33208, 2012.
 36. X. Jiang, Y. Yu, H. W. Yang, N. Y. Agar, L. Frado, and M. D. Johnson. The imprinted gene *jem1peg3i/em1* inhibits wnt signaling and regulates glioma growth *. *Journal of Biological Chemistry*, 285(11):8472–8480, 03 2010. ISSN 0021-9258. doi: 10.1074/jbc.M109.069450. URL <https://doi.org/10.1074/jbc.M109.069450>.
 37. K. Kamimoto, B. Stringa, C. M. Hoffmann, K. Jindal, L. Solnica-Krezel, and S. A. Morris. Dissecting cell identity via network inference and in silico gene perturbation. *Nature*, 614(7949):742–751, 2023.
 38. M. Kanehisa and S. Goto. Kegg: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1):27–30, 2000.
 39. E. L. Kaplan and P. Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481, 1958.
 40. L. Kolberg, U. Raudvere, I. Kuzmin, P. Adler, J. Vilo, and H. Peterson. g:profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Research*, 51(W1):W207–W212, 05 2023. ISSN 0305-1048. doi: 10.1093/nar/gkad347. URL <https://doi.org/10.1093/nar/gkad347>.
 41. M. Latour, N.-G. Her, S. Kesari, and E. Nurmemmedov. Wnt signaling as a therapeutic target for glioblastoma. *International Journal of Molecular Sciences*, 22(16):8428, 2021.
 42. H.-O. Lee, Y. Hong, H. E. Etlioglu, Y. B. Cho, V. Pomella, B. Van den Bosch, J. Vanhecke, S. Verbandt, H. Hong, J.-W. Min, N. Kim, H. H. Eum, J. Qian, B. Boeckx, D. Lambrechts, P. Tsantoulis, G. De Hertogh, W. Chung, T. Lee, M. An, H.-T. Shin, J.-G. Joung, M.-H. Jung, G. Ko, P. Wirapati, S. H. Kim, H. C. Kim, S. H. Yun, I. B. H. Tan, B. Ranjan, W. Y. Lee, T.-Y. Kim, J. K. Choi, Y.-J. Kim, S. Prabhakar, S. Tejpar, and W.-Y. Park. Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer. *Nature Genetics*, 52(6):594–603, 06 2020. ISSN 1546-1718. doi: 10.1038/s41588-020-0636-z. URL <https://doi.org/10.1038/s41588-020-0636-z>.
 43. C.-K. Liao, K.-M. Fang, K. Chai, C.-H. Wu, C.-H. Ho, C.-S. Yang, and S.-F. Tzeng. Depletion of B cell CLL/lymphoma 11B gene expression represses glioma cell growth. *Mol. Neurobiol.*, 53(6):3528–3539, Aug. 2016.
 44. D. Liao. Emerging roles of the EBF family of transcription factors in tumor suppression. *Mol. Cancer Res.*, 7(12):1893–1901, Dec. 2009.
 45. A. Liberzon, C. Birger, H. Thorvaldsdóttir, M. Ghandi, J. P. Mesirov, and P. Tamayo. The molecular signatures database hallmark gene set collection. *Cell systems*, 1(6):417–425, 2015.
 46. A. Loras, L. G. Gonzalez-Bonet, J. L. Gutierrez-Arroyo, C. Martinez-Cadenas, and M. A. Marques-Torres. Neural stem cells as potential glioblastoma cells of origin. *Life*, 13(4):905, 2023.
 47. M. J. Lukey, K. S. Greene, J. W. Erickson, K. F. Wilson, and R. A. Cerione. The oncogenic transcription factor c-jun regulates glutaminase expression and sensitizes cells to glutaminase-targeted therapy. *Nature Communications*, 7(1):11321, 2016.
 48. S. Madonna, C. Scarponi, R. Sestito, S. Pallotta, A. Cavani, and C. Albanesi. The ifn- γ -dependent suppressor of cytokine signaling 1 promoter activity is positively regulated by ifn regulatory factor-1 and spl but repressed by growth factor independence-1b and krüppel-like factor-4, and it is dysregulated in psoriatic keratinocytes. *The Journal of Immunology*, 185(4):2467–2481, 2010.
 49. N. Mantel. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports*, 50(3):163–170, 1966.
 50. J.-C. Marine, S.-J. Dawson, and M. A. Dawson. Non-genetic mechanisms of therapeutic resistance in cancer. *Nature Reviews Cancer*, 20(12):743–756, 2020.
 51. D. Massihnia, A. Galvano, D. Fanale, A. Perez, M. Castiglia, L. Incorvaia, A. Listì, S. Rizzo, G. Cicero, V. Bazan, et al. Triple negative breast cancer: shedding light onto the role of pi3k/akt/mtor pathway. *Oncotarget*, 7(37):60712, 2016.
 52. H. Masuda, D. Zhang, C. Bartholomeusz, H. Doihara, G. N. Hortobagyi, and N. T. Ueno. Role of epidermal growth factor receptor in breast cancer. *Breast Cancer Research and Treatment*, 136(2):331–345, 11 2012. ISSN 1573-7217. doi: 10.1007/s10549-012-2289-9. URL <https://doi.org/10.1007/s10549-012-2289-9>.

- 1007/s10549-012-2289-9.
53. N. McGranahan and C. Swanton. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer cell*, 27(1):15–26, 2015.
 54. T. Moerman, S. Aibar Santos, C. Bravo González-Blas, J. Simm, Y. Moreau, J. Aerts, and S. Aerts. Grnboost2 and arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics*, 35(12):2159–2161, 2019.
 55. B. L. Myers, K. J. Brayer, L. E. Paez-Beltran, E. Villicana, M. S. Keith, H. Suzuki, J. Newville, R. H. Anderson, Y. Lo, C. M. Mertz, R. K. Kollipara, M. D. Borromeo, Q. R. Lu, R. M. Bachoo, J. E. Johnson, and T. Y. Yue. Transcription factors ascl1 and olig2 drive glioblastoma initiation and co-regulate tumor cell types and migration. *Nature Communications*, 15(1):10363, 11 2024. ISSN 2041-1723. doi: 10.1038/s41467-024-54750-9. URL <https://doi.org/10.1038/s41467-024-54750-9>.
 56. S. Müller-Dott, E. Tsirvouli, M. Vazquez, R. RamirezFlores, P. Badia-i Mompel, R. Fallegger, D. Türei, A. Lægreid, and J. Saez-Rodriguez. Expanding the coverage of regulons from high-confidence prior knowledge for accurate estimation of transcription factor activities. *Nucleic Acids Research*, 51(20):10934–10949, 10 2023. ISSN 0305-1048. doi: 10.1093/nar/gkad841. URL <https://doi.org/10.1093/nar/gkad841>.
 57. U. Nasrullah, K. Stanke, V. Recknagel, S. Bozkurt, P. Wurzel, S. Gauer, G. Imre, C. Münch, J. Pfeilschifter, and W. Eberhardt. The e3 ligase trim25 impairs apoptotic cell death in colon carcinoma cells via destabilization of caspase-7 mrna: A possible role of hnrnp1. *Cells*, 12(1), 2023. ISSN 2073-4409. doi: 10.3390/cells12010201. URL <https://www.mdpi.com/2073-4409/12/1/201>.
 58. C. Neftel, J. Laffy, M. G. Filbin, T. Hara, M. E. Shore, G. J. Rahme, A. R. Richman, D. Silverbush, M. L. Shaw, C. M. Hebert, J. Dewitt, S. Gritsch, E. M. Perez, L. N. Gonzalez Castro, X. Lan, N. Druck, C. Rodman, D. Dionne, A. Kaplan, M. S. Bertalan, J. Small, K. Pelton, S. Becker, D. Bonal, Q.-D. Nguyen, R. L. Servis, J. M. Fung, R. Mylvaganam, L. Mayr, J. Gojo, C. Haberler, R. Geyeregger, T. Czech, I. Slavc, B. V. Nahed, W. T. Curry, B. S. Carter, H. Wakimoto, P. K. Brastianos, T. T. Batchelor, A. Stemmer-Rachamimov, M. Martinez-Lage, M. P. Frosch, I. Stamenkovic, N. Riggi, E. Rheinbay, M. Monje, O. Rozenblatt-Rosen, D. P. Cahill, A. P. Patel, T. Hunter, I. M. Verma, K. L. Ligon, D. N. Louis, A. Regev, B. E. Bernstein, I. Tirosh, and M. L. Suvà. An integrative model of cellular states, plasticity, and genetics for glioblastoma. *Cell*, 178(4):835–849.e21, 2019. ISSN 0092-8674. doi: <https://doi.org/10.1016/j.cell.2019.06.024>. URL <https://www.sciencedirect.com/science/article/pii/S0092867419306877>.
 59. C. G. A. Network et al. Comprehensive molecular portraits of human breast tumours. *Nature*, 490(7418):61–70, 2012.
 60. J. Nourisa, A. Passemiers, M. Stock, B. Zeller-Plumhoff, R. Cannoodt, C. Arnold, A. Tong, J. Hartford, A. Scialdone, Y. Moreau, et al. genernib: a living benchmark for gene regulatory network inference. *bioRxiv*, pages 2025–02, 2025.
 61. P. C. Nowell. The clonal evolution of tumor cell populations: Acquired genetic lability permits stepwise selection of variant sublines and underlies tumor progression. *Science*, 194(4260):23–28, 1976.
 62. K. Otsuji, Y. Takahashi, T. Osako, T. Kobayashi, T. Takano, S. Saeki, L. Yang, S. Baba, K. Kumegawa, H. Suzuki, et al. Serial single-cell rna sequencing unveils drug resistance and metastatic traits in stage iv breast cancer. *NPJ Precision Oncology*, 8(1):222, 2024.
 63. B. Pal, Y. Chen, F. Vaillant, B. D. Capaldo, R. Joyce, X. Song, V. L. Bryant, J. S. Penington, L. Di Stefano, N. Tubau Ribera, S. Wilcox, G. B. Mann, A. T. Papenfuss, G. J. Lindeman, G. K. Smyth, J. E. Visvader, and kConFab. A single-cell rna expression atlas of normal, preneoplastic and tumorigenic states in the human breast. *The EMBO Journal*, 40(11):EMBJ2020107333, 05 2021. ISSN 1460-2075. doi: 10.15252/embj.2020107333. URL <https://doi.org/10.15252/embj.2020107333>.
 64. P. C. Pan and R. S. Magge. Mechanisms of egfr resistance in glioblastoma. *International journal of molecular sciences*, 21(22):8471, 2020.
 65. S.-H. Park, Y. M. Chung, J. Ma, Q. Yang, J. S. Berek, and M. C. Hu. Pharmacological activation of foxo3 suppresses triple-negative breast cancer in vitro and in vivo. *Oncotarget*, 7(27):42110, 2016.
 66. Y. Park, S. Lim, J.-W. Nam, and S. Kim. Measuring intratumor heterogeneity by network entropy using rna-seq data. *Scientific reports*, 6(1):37767, 2016.
 67. A. Pratapa, A. P. Jalihal, J. N. Law, A. Bharadwaj, and a. M. Murali. Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. *Nature methods*, 17(2):147–154, 2020.
 68. Priya, S. B. Chowdary, and H. Zafar. Phalcon: phylogeny-aware variant calling from large-scale single-cell panel sequencing datasets. *bioRxiv*, pages 2024–12, 2024.
 69. S. V. Puram, I. Tirosh, A. S. Parikh, A. P. Patel, K. Yizhak, S. Gillespie, C. Rodman, C. L. Luo, E. A. Mroz, K. S. Emerick, et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell*, 171(7):1611–1624, 2017.
 70. J. Qin, X. Chen, X. Xie, M.-J. Tsai, and S. Y. Tsai. COUP-TFII regulates tumor growth and metastasis by modulating tumor angiogenesis. *Proc. Natl. Acad. Sci. U. S. A.*, 107(8):3687–3692, Feb. 2010.
 71. S. Rae, C. Spillane, G. Blackshields, S. F. Madden, J. Keenan, and B. Stordal. The emt-activator zeb1 is unrelated to platinum drug resistance in ovarian cancer but is predictive of survival. *Human Cell*, 35(5):1547–1559, 2022. doi: 10.1007/s13577-022-00744-y.
 72. N. Rao, T. Kumar, R. Pai, A. Mishra, F. Ginhoux, J. Chan, A. Sharma, and H. Zafar. Charting spatial ligand-target activity using renoir. *bioRxiv*, 2024. doi: 10.1101/2023.04.14.536833. URL <https://www.biorxiv.org/content/early/2024/11/01/2023.04.14.536833>.
 73. R. Rosenthal, E. L. Cadieux, R. Salgado, M. A. Bakir, D. A. Moore, C. T. Hiley, T. Lund, M. Tanić, J. L. Reading, K. Joshi, et al. Neoantigen-directed immune escape in lung cancer evolution. *Nature*, 567(7749):479–485, 2019.
 74. G. Satas, S. Zaccaria, G. Mon, and B. J. Raphael. Scarlet: single-cell tumor phylogeny inference with copy-number constrained mutation losses. *Cell systems*, 10(4):323–332, 2020.
 75. J. Singer, J. Kuipers, K. Jahn, and N. Beerenwinkel. Single-cell mutation identification via phylogenetic inference. *Nature communications*, 9(1):5144, 2018.

76. C. Skok Gibbs, C. A. Jackson, G.-A. Saldi, A. Tjärnberg, A. Shah, A. Watters, N. De Veaux, K. Tchourine, R. Yi, T. Hamamsy, et al. High-performance single-cell gene regulatory network inference at scale: the inferelator 3.0. *Bioinformatics*, 38(9):2519–2528, 2022.
77. C. Skok Gibbs, O. Mahmood, R. Bonneau, and K. Cho. Pmf-grn: a variational inference approach to single-cell gene regulatory network inference using probabilistic matrix factorization. *Genome biology*, 25(1):88, 2024.
78. M. Suvà, E. Rheinbay, S. Gillespie, A. Patel, H. Wakimoto, S. Rabkin, N. Riggi, A. Chi, D. Cahill, B. Nahed, W. Curry, R. Martuza, M. Rivera, N. Rossetti, S. Kasif, S. Beik, S. Kadri, I. Tirosh, I. Wortman, A. K. Shalek, O. Rozenblatt-Rosen, A. Regev, D. Louis, and B. Bernstein. Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. *Cell*, 157(3):580–594, 04 2014. ISSN 0092-8674. doi: 10.1016/j.cell.2014.02.030. URL <https://doi.org/10.1016/j.cell.2014.02.030>.
79. I. Tirosh, B. Izar, S. M. Prakadan, M. H. Wadsworth, D. Treacy, J. J. Trombetta, A. Rotem, C. Rodman, C. Lian, G. Murphy, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell rna-seq. *Science*, 352(6282):189–196, 2016.
80. M. Tyler, A. Gavish, C. Barbolin, R. Tschernichovsky, R. Hoefflin, M. Mints, S. V. Puram, and I. Tirosh. The curated cancer cell atlas provides a comprehensive characterization of tumors at single-cell resolution. *Nature Cancer*, 6(6):1088–1101, 06 2025. ISSN 2662-1347. doi: 10.1038/s43018-025-00957-8. URL <https://doi.org/10.1038/s43018-025-00957-8>.
81. E. Wang, N. Zaman, S. Mcgee, J.-S. Milanese, A. Masoudi-Nejad, and M. O'Connor-McCourt. Predictive genomics: a cancer hallmark network framework for predicting tumor clinical phenotypes using genome sequencing data. In *Seminars in cancer biology*, volume 30, pages 4–12. Elsevier, 2015.
82. L. Wang, H. Babikir, S. Müller, G. Yagnik, K. Shamardani, F. Catalan, G. Kohanbash, B. Alvarado, E. Di Lullo, A. Kriegstein, S. Shah, H. Wadhwa, S. M. Chang, J. J. Phillips, M. K. Aghi, and A. A. Diaz. The phenotypes of proliferating glioblastoma cells reside on a single axis of variation. *Cancer Discovery*, 9(12):1708–1719, 12 2019. ISSN 2159-8274. doi: 10.1158/2159-8290.CD-19-0329. URL <https://doi.org/10.1158/2159-8290.CD-19-0329>.
83. L. Wang, J. Jung, H. Babikir, K. Shamardani, S. Jain, X. Feng, N. Gupta, S. Rosi, S. Chang, D. Raleigh, et al. A single-cell atlas of glioblastoma evolution under therapy reveals cell-intrinsic and cell-extrinsic therapeutic targets. *Nature cancer*, 3(12):1534–1552, 2022.
84. M. Wedel, F. Fröb, O. Elsesser, M.-T. Wittmann, D. C. Lie, A. Reis, and M. Wegner. Transcription factor tcf4 is the preferred heterodimerization partner for olig2 in oligodendrocytes and required for differentiation. *Nucleic acids research*, 48(9):4839–4857, 2020.
85. F. A. Wolf, P. Angerer, and F. J. Theis. Scanpy: large-scale single-cell gene expression data analysis. *Genome Biology*, 19(1):15, 02 2018. ISSN 1474-760X. doi: 10.1186/s13059-017-1382-0. URL <https://doi.org/10.1186/s13059-017-1382-0>.
86. S. Z. Wu, G. Al-Eryani, D. L. Roden, S. Junankar, K. Harvey, A. Andersson, A. Thennavan, C. Wang, J. R. Torpy, N. Bartonicek, T. Wang, L. Larsson, D. Kaczorowski, N. I. Weisenfeld, C. R. Uyttingco, J. G. Chew, Z. W. Bent, C.-L. Chan, V. Gnanasambandapillai, C.-A. Dutertre, L. Gluch, M. N. Hui, J. Beith, A. Parker, E. Robbins, D. Segara, C. Cooper, C. Mak, B. Chan, S. Warrier, F. Ginhoux, E. Millar, J. E. Powell, S. R. Williams, X. S. Liu, S. O'Toole, E. Lim, J. Lundeberg, C. M. Perou, and A. Swarbrick. A single-cell and spatially resolved atlas of human breast cancers. *Nature Genetics*, 53(9):1334–1347, 09 2021. ISSN 1546-1718. doi: 10.1038/s41588-021-00911-1. URL <https://doi.org/10.1038/s41588-021-00911-1>.
87. Y. Yamada, R. Simon, K. Iwane, Y. Nakanishi, Y. Takeuchi, A. Yoshizawa, M. Takada, M. Toi, H. Haga, A. Marx, and G. Sauter. An exploratory study for tuft cells in the breast and their relevance in triple-negative breast cancer: the possible relationship of SOX9. *BMC Cancer*, 23(1):438, May 2023.
88. Y. Yeung, K. McDonald, T. Grewal, and L. Munoz. Interleukins in glioblastoma pathophysiology: implications for therapy. *British journal of pharmacology*, 168(3):591–606, 2013.
89. H. Zafar, Y. Wang, L. Nakhleh, N. Navin, and K. Chen. Monovar: single-nucleotide variant detection in single cells. *Nature methods*, 13(6):505–507, 2016.
90. H. Zafar, A. Tzen, N. Navin, K. Chen, and L. Nakhleh. Sift: inferring tumor trees from single-cell sequencing data under finite-sites models. *Genome biology*, 18:1–20, 2017.
91. H. Zafar, N. Navin, K. Chen, and L. Nakhleh. Siclonet: Bayesian inference of population structure, genotype, and phylogeny of tumor clones from single-cell genome sequencing data. *Genome research*, 29(11):1847–1859, 2019.
92. K. Zhang, E. P. Erkan, S. Jamalzadeh, J. Dai, N. Andersson, K. Kaipio, T. Lamminen, N. Mansuri, K. Huhtinen, O. Carpén, S. Hietanen, J. Oikonen, J. Hynninen, A. Virtanen, A. Häkkinen, S. Hautaniemi, and A. Vähärautio. Longitudinal single-cell rna-seq analysis reveals stress-promoted chemoresistance in metastatic ovarian cancer. *Science Advances*, 8(8):eabm1831, 2022. doi: 10.1126/sciadv.abm1831.
93. J. X. Zhou, R. Taramelli, E. Pedrini, T. Knijnenburg, and S. Huang. Extracting intercellular signaling network of cancer tissues using ligand-receptor expression patterns from whole-tumor and single-cell transcriptomes. *Scientific Reports*, 7(1):8815, 2017.
94. X. Zhou, T. Liang, J. Deng, K. Ng, M. Li, C. Lv, J. Chen, K. Yang, Z. Ma, W. Ma, and P. Wang. Differential and prognostic significance of HOXB7 in gliomas. *Front. Cell Dev. Biol.*, 9:697086, Aug. 2021.

Supplementary Files

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

- [ITHGRNsupplementary.pdf](#)