Supplementary information

Pengran Yu, Kaihua Cao, Zhihua Cheng, Yu Zeng, WenChuan Zhang
Department of Neurosurgery, Shanghai Ninth People's Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200011, China
Correspondence: zhangwench88@sjtu.edu.cn

Pengran Yu and Kaihua Cao contributed equally to this work.

Cohort roles and sample reconciliation

The study reanalyzed de-identified public records. GLASS was the longitudinal bulk discovery source; GSE174554 and its related GSE138794 matrices formed the reconciled single-nucleus discovery series. CARE GSE274546 served as the separately deposited external human cohort. Original authors' metadata, clinical timepoints, sample names, gene identifiers, and file checksums were reconciled before comparisons. CARE counts were checked against all 121 author sample files and annotations. Recovery of a missing discovery sample was not counted as an independent cohort. Only primary-to-first-recurrence comparisons entered the main human longitudinal analysis. Patient pairs, not cells, regions, spots, or technical libraries, supplied biological replication.

Bulk and single-nucleus analysis

The strict GLASS analysis retained 84 adult IDH-wildtype grade IV patients with documented intervening radiotherapy or temozolomide. Log2(transcripts per million [TPM] + 1) expression for 17,923 genes was modeled with patient-blocked limma-trend, and the 50 qualifying genes in each direction were frozen. Original inclusion, sensitivity, and external bulk analyses were retained [3,30].

Raw unique molecular identifier (UMI) counts, author tumor labels, and documented QC rules were used for discovery and CARE. Recovery scripts preserved annotation provenance and reused original QC cells when only labels were missing. Discovery quality control used sample-specific lower gene and UMI thresholds with floors of 200 genes and 300 UMIs; the mitochondrial threshold was bounded between 5% and 20%. Scrublet was attempted when at least 200 cells passed initial filters, with expected doublet rate 0.06 and seed 20260913; estimated doublets were excluded. Per-sample thresholds and failures are recorded in the quality-control tables. CARE retained author-QC nuclei. A 21,810-symbol assay intersection standardized log1p(CP10k) module scoring; discovery-derived control genes were fixed before CARE scoring. Four malignant labels were assigned by the maximal harmonized Neftel-like score [9]. Author-defined malignancy and a stricter CNV sensitivity remained distinct. Pseudobulk counts were summed within patient, specimen, and compartment. Main thresholds were 50 malignant cells per sample, or 20 cells for a state/normal compartment. edgeR expression filtering, TMM normalization, robust dispersion estimation, patient-blocked QL testing, and limma-voom sensitivity were applied [30,31]. Patient-level aggregation avoids treating nuclei from the same patient as independent replicates, a problem emphasized in single-cell differential-expression benchmarks [32,33,34,35]. QC and normalization choices were documented following general single-cell analysis principles [36]. Full tested universes and paired denominators accompany the results.

Mean cell scores and composition fractions were reduced to specimen values before patient-paired tests. Two-sided t tests, 95% t intervals, and Wilcoxon sensitivities were retained. Composition denominators were all retained QC nuclei; missing communication cell groups were instead kept unobserved. The CARE 43-pair and 52-pair composition analyses address different selection rules and are not independent replications. Frozen external gene tests used 36 hypotheses together for BH

correction, preserving discovery directions; missing or unestimable genes could not count as successful replication.

Brain-reference construction and deconvolution

Author-labeled CARE pseudobulks from 52 clinical pairs defined malignant, oligodendrocyte, neuronal, astrocyte, myeloid, OPC, vascular (endothelial and pericyte), and lymphocyte compartments. Each reference contributor required at least 20 nuclei of that type, and each type required at least five contributing donors. Reference profiles averaged sample-type CPM with equal donor weight. Within the common measured universe, 60 specific genes per compartment were selected with mean CPM >= 1, excluding mitochondrial and ribosomal genes, the frozen GLASS program, and every four-state signature/control gene. The full reference contained 480 markers. Every donor-held-out fold rebuilt both marker selection and reference profiles from the remaining donors.

The CIBERSORT v1.03 CoreAlg implemented nu-support-vector regression with nu = 0.25, 0.50, or 0.75, coefficient clipping, sum normalization, and minimum-RMSE selection [15]. Its core was used unchanged, while the wrapper explicitly standardized reference and mixture without quantile normalization or automatic antilog transformation. No mixture-permutation fit P values were generated or used for filtering. Weighted NNLS used the same reference with inverse-square-root mean-reference weights. This is an independent estimator sensitivity, not an independent reference cohort. Observed held-out pseudobulks were assessed against captured-UMI RNA fractions; synthetic held-out-donor mixtures provided additional coverage. Compartment-level gates required observed Spearman rho >= 0.50 and MAE <= 0.15. They were fixed in the extension protocol before deconvolution outputs were inspected.

The full CARE reference was applied to 168 GLASS bulk specimens. Two-sided paired t tests, 95% t intervals, Wilcoxon sensitivities, and BH correction across eight compartments within each method were retained. All 84 patient pairs remained included. Correlations between paired weight changes and the frozen program or ten original target-expression changes used Spearman statistics with BH correction across all 176 tests. These same-sample correlations were interpreted as exploratory context.

Expanded malignant programs and continuous axes

Published metaprogram and biological-process genes were read from the published Nomura supplementary tables [14], with source hashes preserved. Cell-level log1p(CP10k) expression was averaged within each malignant specimen; each signature score subtracted frozen expression-matched controls. Controls used 24 discovery expression bins, 30 controls per signature gene, exclusion of all signature genes, and seed 20260920. A minimum 80% coverage rule was applied before paired tests. Biological and technical signature families were corrected separately, with an additional HC3 regression sensitivity for paired UMI change. Expanded CARE state fractions used author annotations and malignant-cell denominators; these labels were not inferred from an argmax over the new signatures.

For continuous axes, discovery means and standard deviations of patient-level four-state changes and PCA loadings were fixed before projecting CARE. After excluding all directly defining score and control genes, expression filtering and TMM normalization preceded logCPM calculation with prior count two. For each gene and axis, a limma model related recurrent-minus-primary expression to axis score and mean-log-UMI change, with robust, trend-aware empirical Bayes inference. Discovery BH correction jointly covered both axes; CARE correction covered all 627 selected hypotheses, assigning untestable results P = 1. Concordance required the same slope direction and external FDR < 0.05. Sensitivities added the initial state projected with fixed loadings, performed leave-one-patient-out

slope refits, and tested partial-rank associations after residualizing ranks on UMI change with n - 3 degrees of freedom. These analyses were specified as a post-validation exploratory extension, not as blinded or preregistered validation.

Representation comparison and attribution

Verified local Geneformer-V1-10M weights used raw counts mapped to model Ensembl identifiers, collapsed duplicates, division by model gene medians, and descending rank truncation to 2,048 tokens. Penultimate gene embeddings were mean-pooled. The AC comparison sampled at most 64 cells per specimen with fixed seed 20260915. Geneformer, four state scores, fixed 100-gene RNA, and training-selected 2,000-gene RNA models used identical cells and 20 repeated patient-grouped five-fold splits. Standardization and variance selection were fitted only within training folds. Logistic regression used fixed C = 1 with liblinear; full-discovery fits were then frozen for CARE. No CARE tuning or orientation reversal was allowed. AUROC intervals used 2,000 patient-pair bootstraps; external label-swap tests used 2,000 within-patient swaps with BH correction across four models.

Linear interventional SHAP values [22] were computed in log-odds units from standardized features and fitted coefficients using each training fold's mean background. Only held-out patients contributed to attribution summaries. A fixed first patient in identifier order supplied the single-fold waterfall example; it was not selected for a favorable prediction. Correlated-feature attribution remains associative.

Real knockout calibration

GSE198522 control and KAT5-knockout matrices were downloaded with source receipts [21]. QC required at least 200 detected genes, 500 UMIs, and at most 20% mitochondrial UMIs. A fixed random 128-cell sample per library established real-KO and control embedding centroids. The initial benchmark required at least ten target-token-positive control cells. After documenting limited coverage, an explicitly labeled sensitivity searched all QC control cells and selected at most 64 eligible cells per model. Three other tokens within 64 rank positions provided matched deletion controls. Contextual effects compared post-deletion mean embeddings with original embeddings of the same remaining genes. Model-mean virtual vectors were aligned with the fixed observed KO-minus-control direction. Model-level exact sign-flip tests and 10,000 model bootstraps assessed the control-adjusted alignment. The sorted 827 model was excluded in a predefined-setting sensitivity; no new target was selected by these results.

Communication pathways and networks

CellChat 2.2.0.9001 [37,38] used protein interactions whose ligand/receptor complex subunits and cofactors were all in the common assay universe. Non-protein interactions were excluded. Each paired-patient object used a minimum of 20 and a maximum of 300 cells per available cell group, fixed seed 20260915, truncated mean with trim 0.1, raw normalized signaling data, no population-size weighting, and 100 cell-label permutations. Only within-stage scores were extracted. All 12,201 eligible pathway-source-target comparisons were corrected together; cell-label permutation P values were not biological recurrence tests.

Ranked gene-set analysis and Hallmark summaries followed established enrichment frameworks [39,40]. KEGG 2021 Human was downloaded from Enrichr with a checksum, and ranked fgsea analyses used gene sets with 15-500 measured genes. Joint correction covered five bulk or malignant analyses. CAMERA sensitivity [41] used patient-blocked voom expression, comparing estimated non-negative inter-gene correlation with fixed correlation 0.01. WGCNA reference parameters were

retained [42], with additional powers 8 and 16 and five fixed random 67-of-84 patient subsamples. Adjusted Rand indices and best-matching module Jaccard indices quantify partition sensitivity. DEG/module overlap remains descriptive because both use the same discovery data.

The archived scTenifold analysis used 22 patients, 32 malignant nuclei per specimen, 1,508 genes, ten resampled networks per stage, and ten fixed virtual targets [13]. It was retained as a limited original computational analysis, with its original patient denominator retained. Per-target figures exclude the deleted gene itself and distinguish unsigned network FC from expression fold change.

Spatial clinical and drug context

Spatial scores retained the original deposited coordinates, coverage rules, UMI-adjusted within-section correlations, and patient aggregation. Program genes were excluded from comparison marker sets. Core and edge regions were not inferred without deposited annotations. Survival models used continuous standardized program scores and clinical covariates, with complete-case counts and proportional-hazards checks retained. DepMap 24Q4, PRISM 24Q2, and GDSC2 release 8.5 used archived verified files and biological-model aggregation; these were not labeled the latest releases. Complete source tables and analysis scripts accompany the manuscript [23,24,43,44,45,46].

Multiple testing and evidential status

Benjamini–Hochberg FDR control [47] was applied within explicitly stated biological test families. Pooled cell counts were not used as patient replication. Frozen tests, exploratory extensions, method sensitivities, and descriptive figures were distinguished throughout. The 36 frozen gene-state tests, 627 later axis-association tests, and 176 deconvolution-context tests answer different questions and were not merged or selectively substituted for one another. Full tables include failed calibration gates and non-significant tests.

No prospective sample-size calculation was performed for this secondary analysis; all eligible pairs meeting the stated criteria were retained.

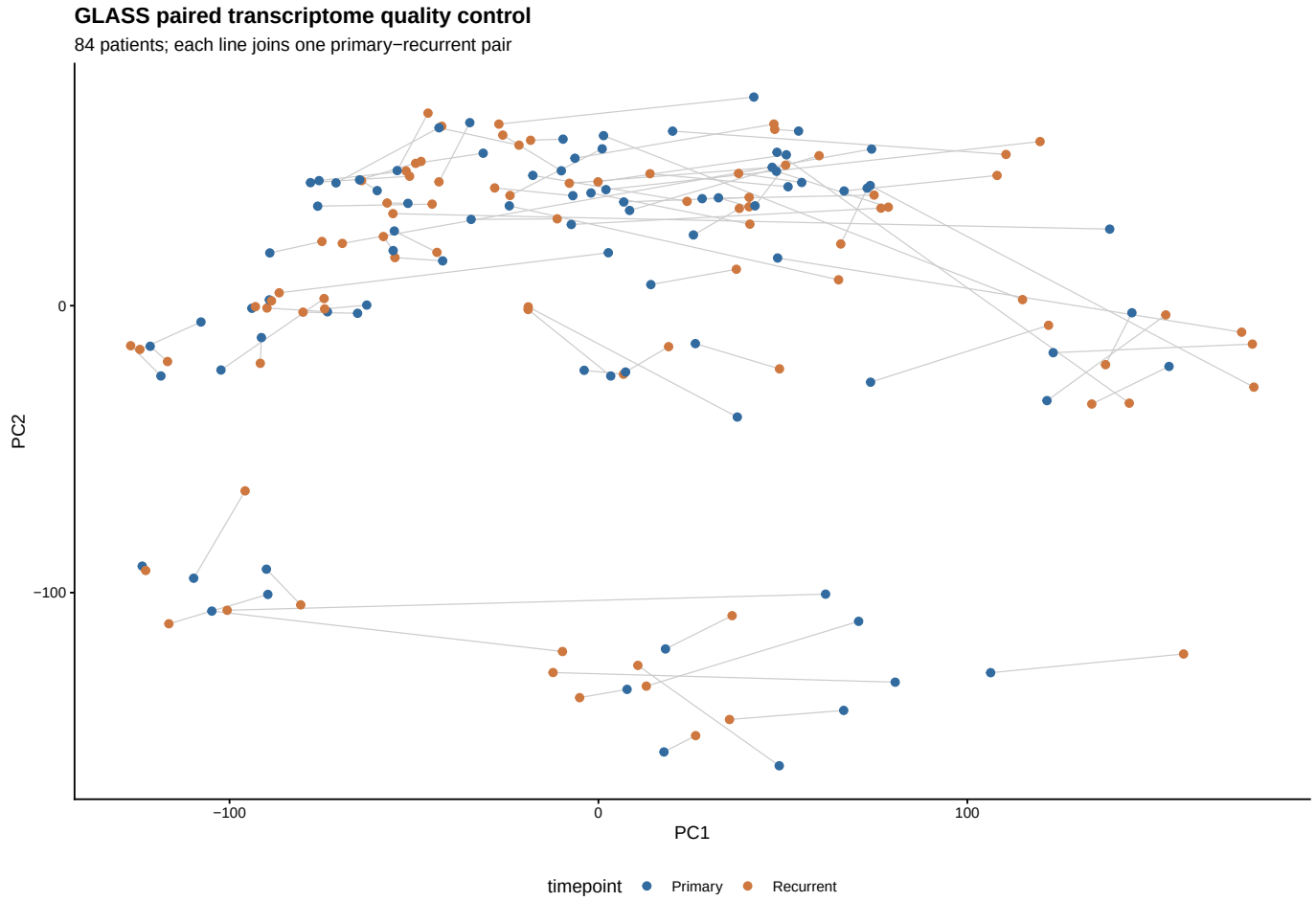
Data sources

Public expression and annotation data were obtained from GLASS and the Gene Expression Omnibus. The principal accessions are GSE174554 and GSE138794 (single-nucleus discovery), GSE274546 (CARE), GSE339484 (external bulk), GSE198522 (KAT5 intervention), GSE237183, GSE318562 and GSE286413 (spatial data). Additional panel/model datasets are GSE254873, GSE254874, GSE254145, GSE271619 and GSE271620. GEO accessions are accessible at https://www.ncbi.nlm.nih.gov/geo/. GLASS provenance is linked to https://doi.org/10.1016/j.cell.2022.04.038. Clinical and functional analyses used TCGA, CGGA325, CGGA693, DepMap 24Q4, PRISM 24Q2 and GDSC2 release 8.5. The supplementary data archive contains derived result tables, sample eligibility records and a source register. Source-level access and reuse conditions continue to apply.

Code availability

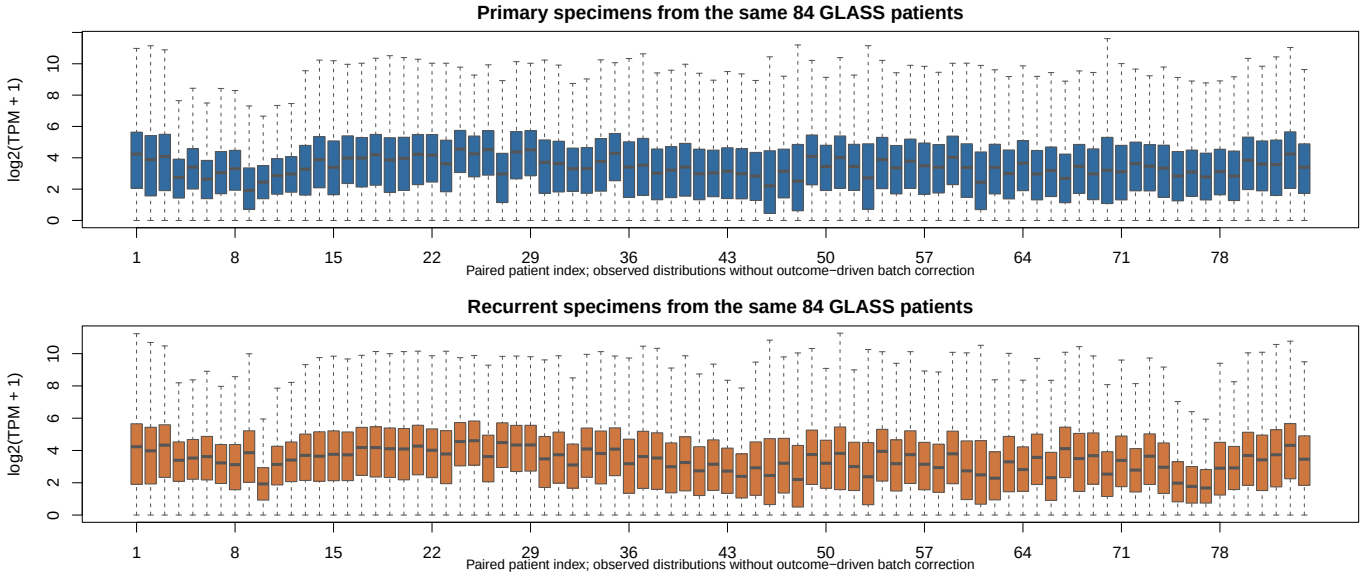
Custom analysis scripts and the figure provenance index are provided in the supplementary code archive. The archive documents software dependencies and data-path configuration. Original public matrices, pretrained Geneformer weights and third-party CIBERSORT software must be obtained separately under their respective terms; they are not redistributed in the code archive.

Supplementary Figure S1 GLASS bulk expression principal components



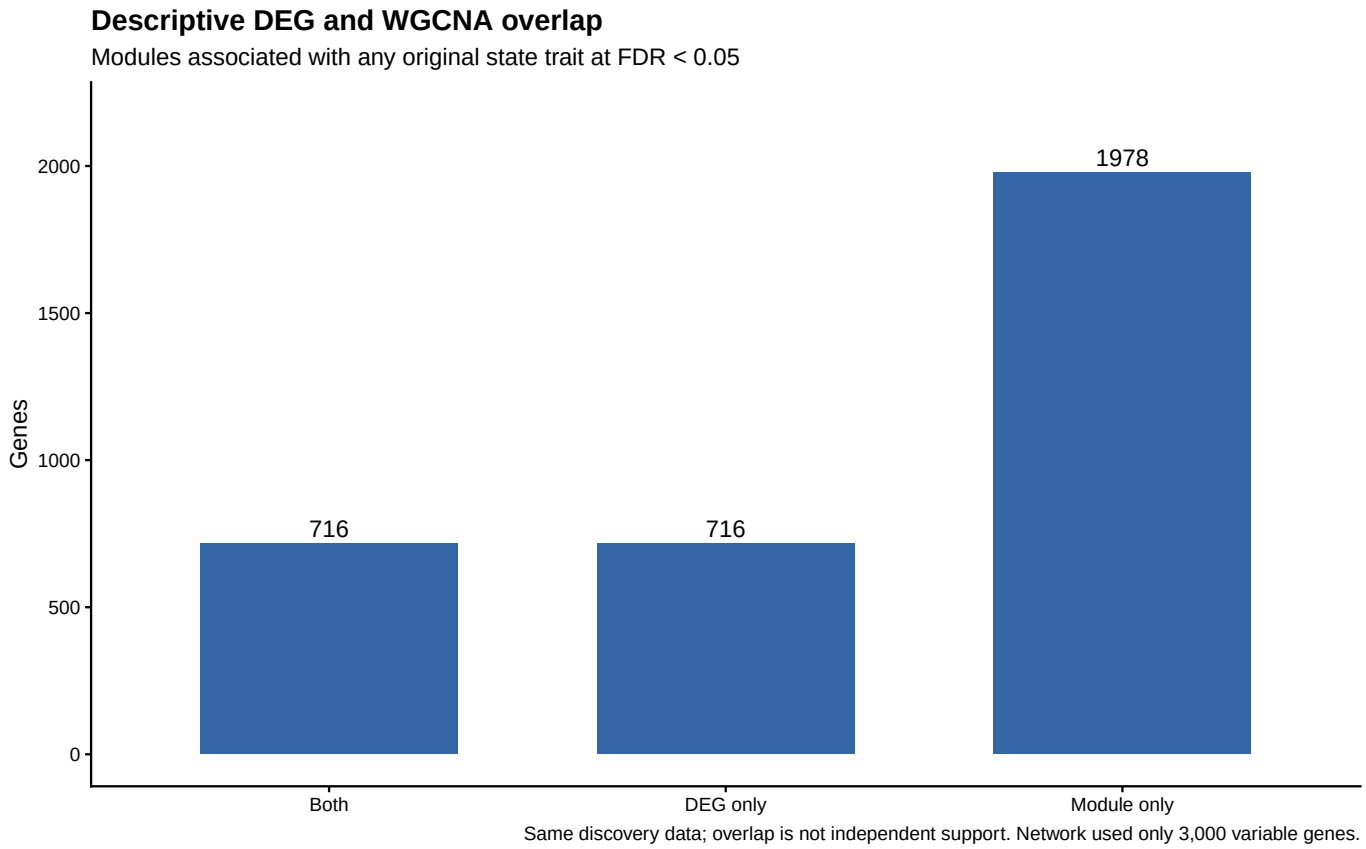
Supplementary Figure S1. Principal components of the retained bulk expression profiles, colored by study metadata. Samples remain paired in downstream inference; visual separation is not a recurrence test and cannot identify a batch-free causal effect.

Supplementary Figure S2 GLASS expression distributions



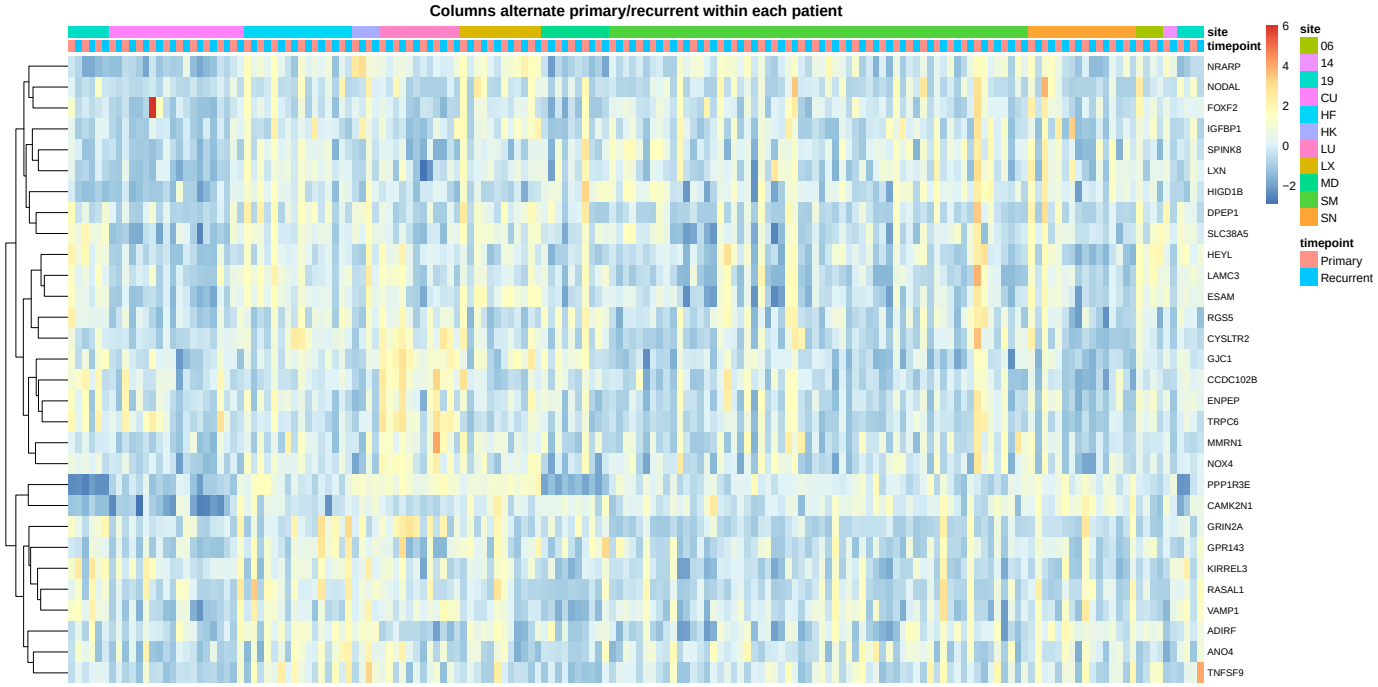
Supplementary Figure S2. Distribution checks for the paired bulk input and retained expression universe. Log transformation and gene selection are documented in the source scripts. Comparable distributions do not rule out collection-site or regional sampling differences.

Supplementary Figure S3 Overlap between bulk differential expression and network modules



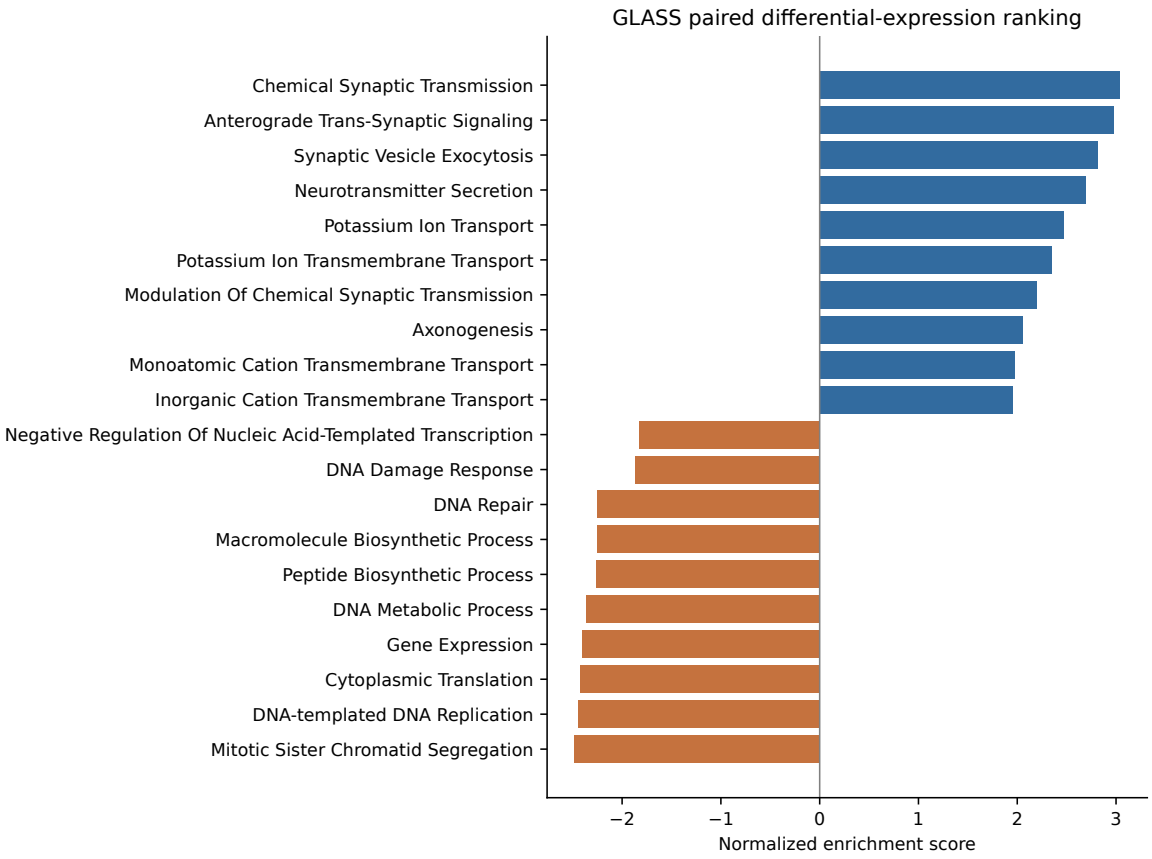
Supplementary Figure S3. Overlap between GLASS DE genes and WGCNA modules from the same paired-patient discovery data. This overlap is descriptive and does not supply an independent replication test or an additional therapeutic-evidence layer.

Supplementary Figure S4 Paired GLASS differential-expression heatmap



Supplementary Figure S4. The thirty highest-ranked genes from the paired GLASS differential-expression table, ordered by limma P value, are shown across 168 specimens from 84 patients. Columns alternate primary and recurrent specimens within patient; annotation bars indicate timepoint and collection site. Rows are standardized expression, clustered for display. This is a selected discovery-data heatmap, not the full frozen 100-gene program or an independent validation.

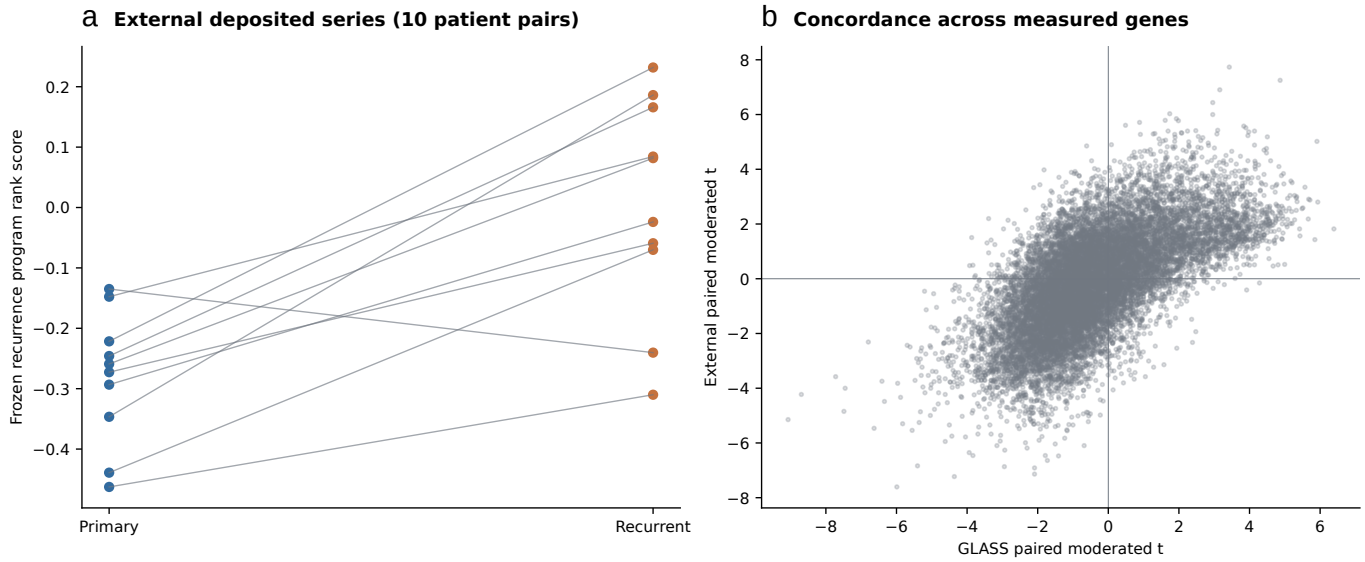
Supplementary Figure S5 GO biological process enrichment



Ten lowest-FDR terms per direction are displayed from the full GO table; terms are overlapping, not independent mechanisms.
This is ranked enrichment of all tested bulk genes, not over-representation of a selected DEG/WGCNA intersection.
The same 84-patient discovery data underlie both the gene ranking and this biological summary.

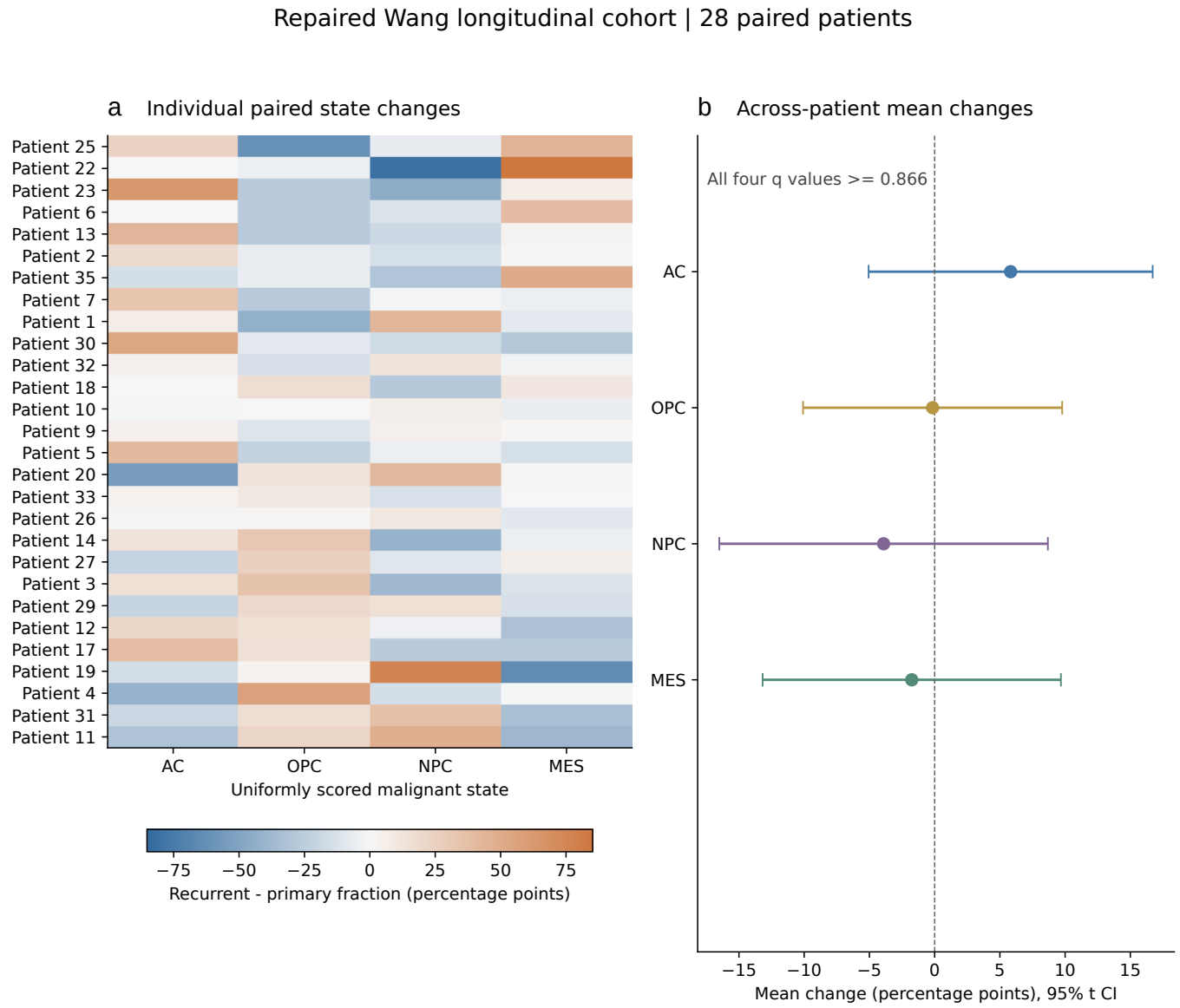
Supplementary Figure S5. The ten lowest-FDR GO terms within each enrichment direction are shown from the archived full ranked analysis of 84 paired GLASS patients. Related terms overlap in genes and cannot be counted as independent mechanisms. Input is the genome-wide paired ranking, not a selected gene intersection. The full result table retains all tests.

Supplementary Figure S6 External bulk replication of the frozen program



Supplementary Figure S6. GSE339484 expression-abundance analysis. The main eligible subset comprises ten grade IV IDH1-wildtype pairs; an additional legacy grade III pair enters sensitivity only. Program mean change is 0.287 (95% CI 0.158–0.416), and 96 measured program genes are directionally concordant with GLASS. Cross-deposit expression matching cannot establish genetic independence of anonymized patients.

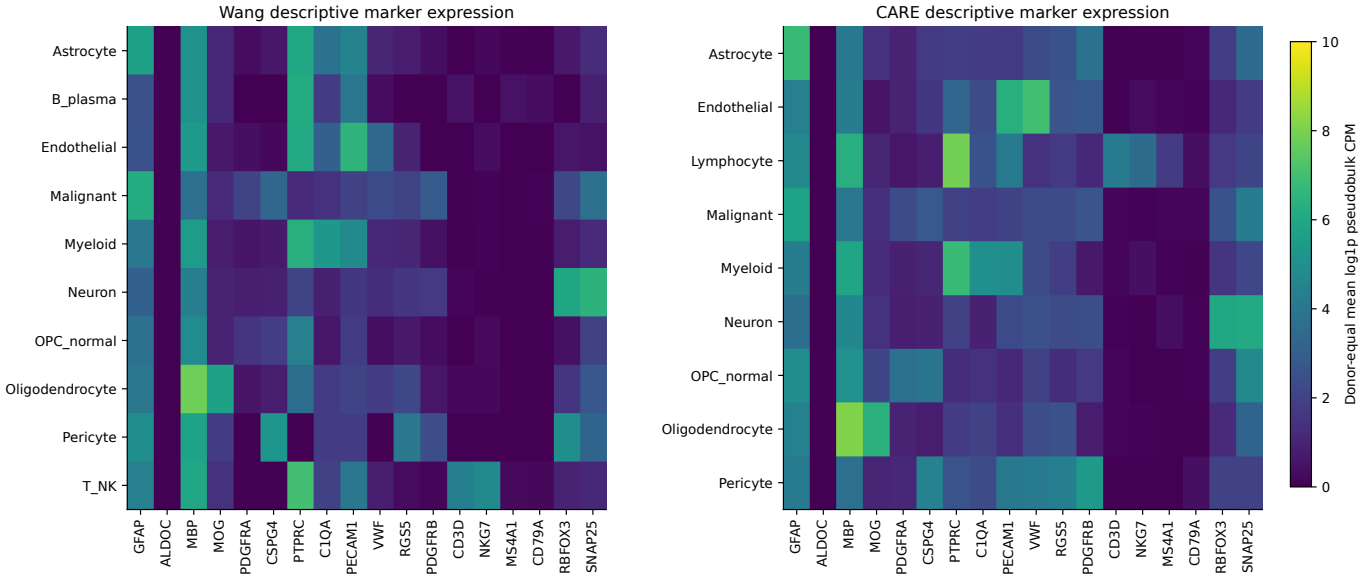
Supplementary Figure S7 Heterogeneous malignant state changes



Rows ordered by a continuous discovery PC1; no trajectory classes were assigned.
Each patient has equal weight. FDR covers the nine prespecified primary score/fraction tests.

Supplementary Figure S7. a, Recurrent-minus-primary fractions of harmonized AC, OPC, NPC and MES malignant states in 28 reconciled discovery pairs. Rows are patients ordered by continuous discovery PC1, without assigning trajectory classes. b, Patient-equal mean changes with two-sided 95% paired t intervals. FDR covers the nine prespecified primary score/fraction tests; none supports a cohort-wide shift. Continuous state-change axes are retrospective and do not predict recurrence from baseline.

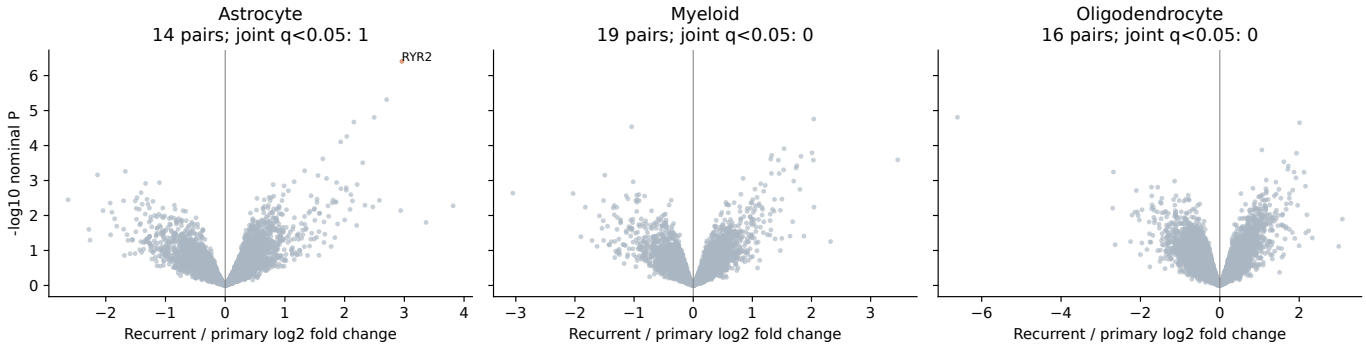
Supplementary Figure S8 Marker evidence for compartment annotation



At least 20 nuclei per contributing sample/type; means first aggregate timepoints within patient.
Wang normal labels were marker-inferred; this is annotation transparency, not an independent validation of those labels.
CARE uses author labels. Malignancy remains author-defined; shared neural markers alone do not diagnose malignant cells.

Supplementary Figure S8. Descriptive expression of 18 marker genes across available compartments in the 28-pair discovery and 52-pair CARE composition cohorts. Contributors require at least 20 nuclei per sample and type. Pseudobulk log1p(CPM) is averaged within donor before donor-equal aggregation. Wang normal labels are themselves marker-inferred, so the heatmap shows annotation consistency rather than independent validation. CARE labels are author supplied. Neural markers shared by malignant and normal cells do not independently establish malignancy.

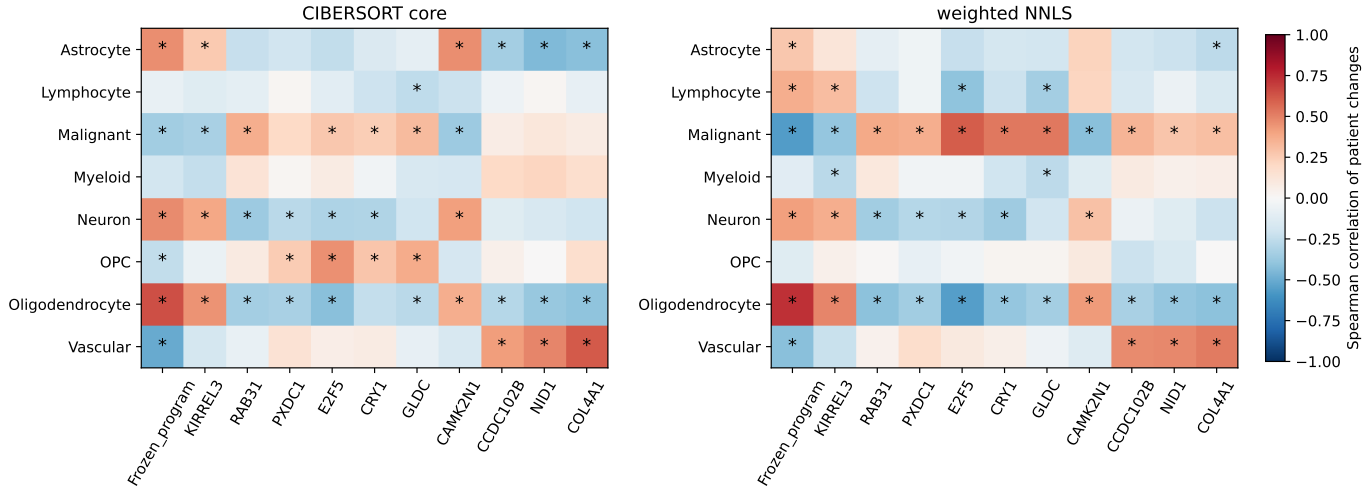
Supplementary Figure S9 Normal-compartment paired differential expression



Patient-blocked edgeR QL; >=20 cells per compartment/sample and >=8 paired patients.
Normal labels are marker-inferred; all tested gene-compartment hypotheses share the indicated FDR family.

Supplementary Figure S9. Discovery pseudobulks with at least 20 nuclei per compartment and timepoint include 14 astrocyte, 19 myeloid and 16 oligodendrocyte pairs. Other compartments fail the eight-pair inferential threshold. RYR2 in astrocytes is the single association surviving joint compartment correction; marker-inferred normal annotations remain a limitation.

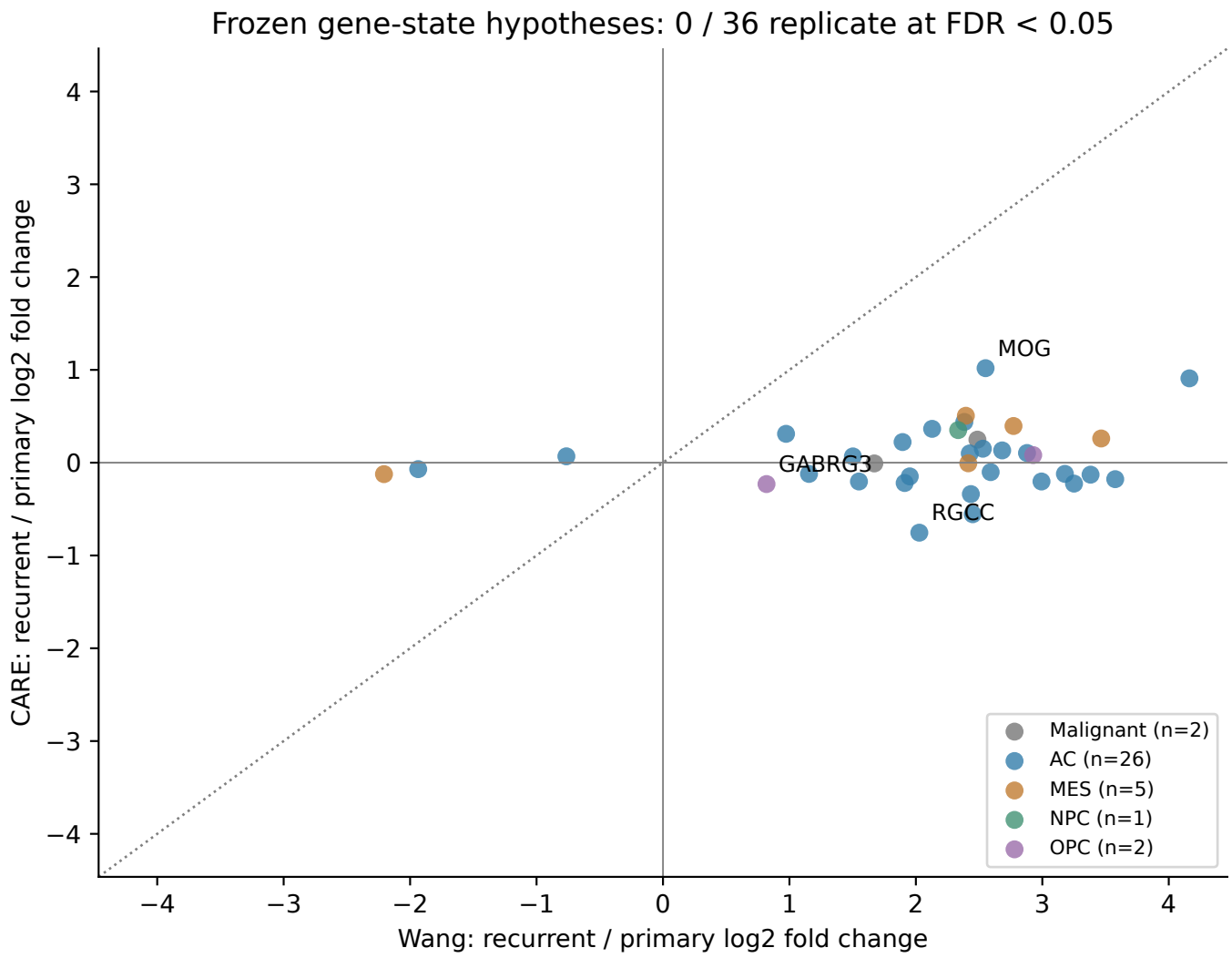
Supplementary Figure S10 Deconvolution context for the frozen program and ten genes



84 patient pairs; * BH FDR < 0.05 across all 176 feature-compartment-method tests.
These associations use the same bulk data on both sides and do not validate the original ten targets.
Reference-specificity errors remain for malignant, OPC and lymphocyte weights; see the held-out benchmark.

Supplementary Figure S10. Spearman correlations between recurrent-minus-primary expression/score changes and estimated RNA-weight changes across all 84 GLASS patients. Stars mark BH FDR < 0.05 across all 176 feature-compartment-method tests. Same-bulk correlations are not independent validation. Interpretation of compartments failing held-out calibration remains restricted; neither a star nor its sign establishes a causal target.

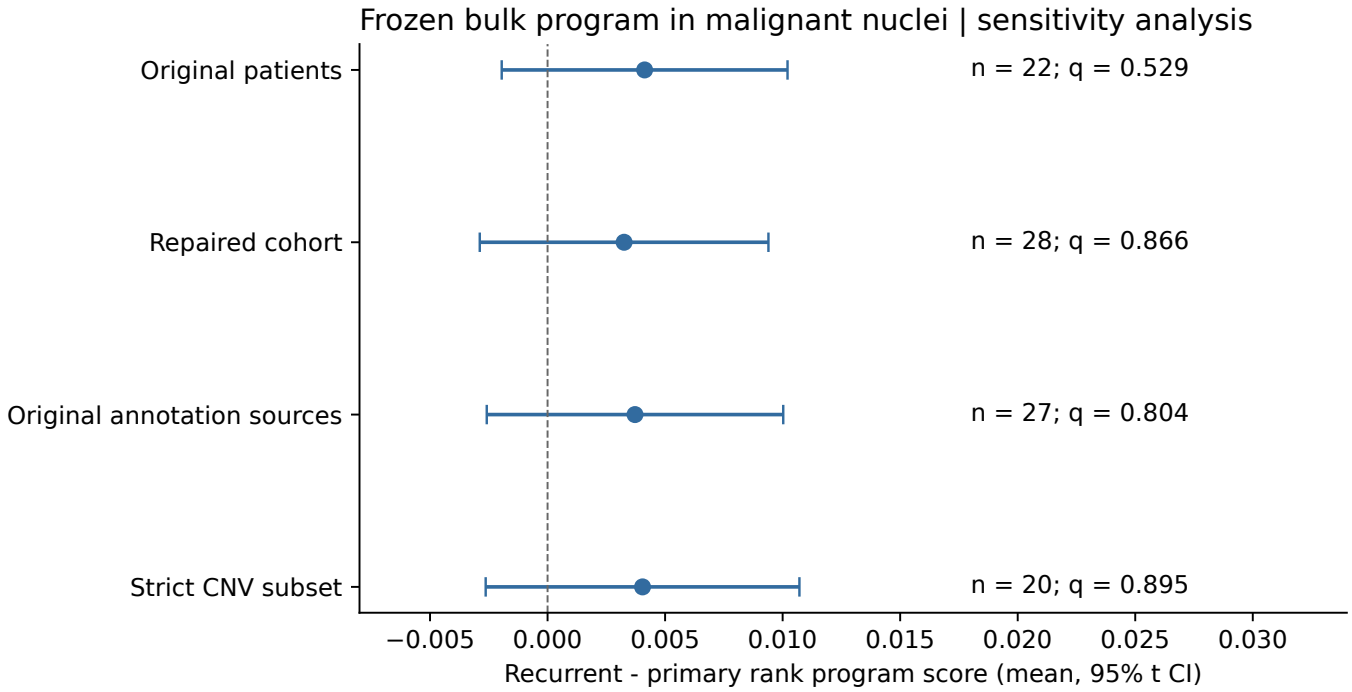
Supplementary Figure S11 Frozen external gene hypotheses



Exploratory discovery gate q < 0.10; external correction covers all 36 hypotheses.
CARE patient pairs vary by compartment (28–43). Best external FDR: MOG, 0.099.

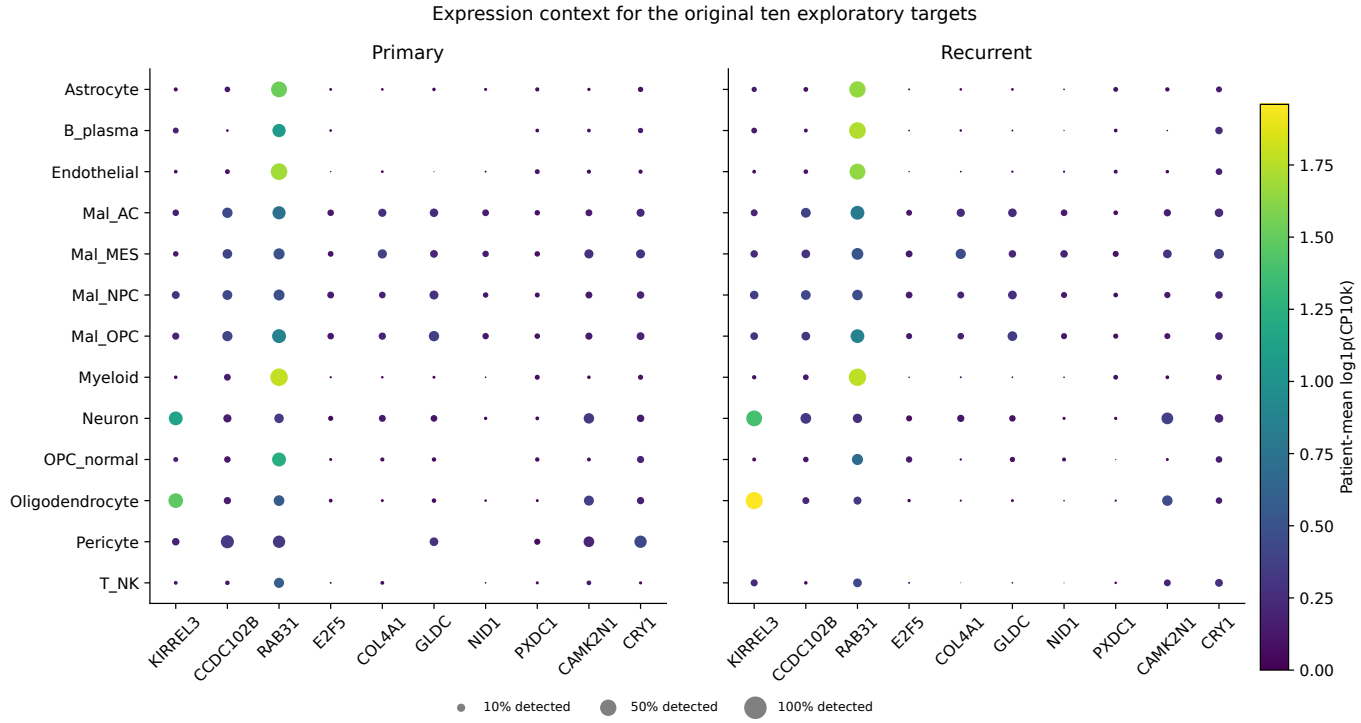
Supplementary Figure S11. All 36 frozen exploratory gene-state hypotheses are plotted using paired discovery and CARE log2 fold changes. Colors indicate malignant compartments; legend n values count hypotheses, not patients. CARE supplies 28–43 paired patients depending on compartment. External BH correction covers all 36 tests, and none passes FDR < 0.05. Discovery eligibility used joint q < 0.10, which was exploratory selection rather than a claim of corrected 0.05 discovery. The line marks equal cohort effects; labeled points illustrate discordance or the best external FDR, not validated targets.

Supplementary Figure S12 Frozen program scoring sensitivity in malignant nuclei



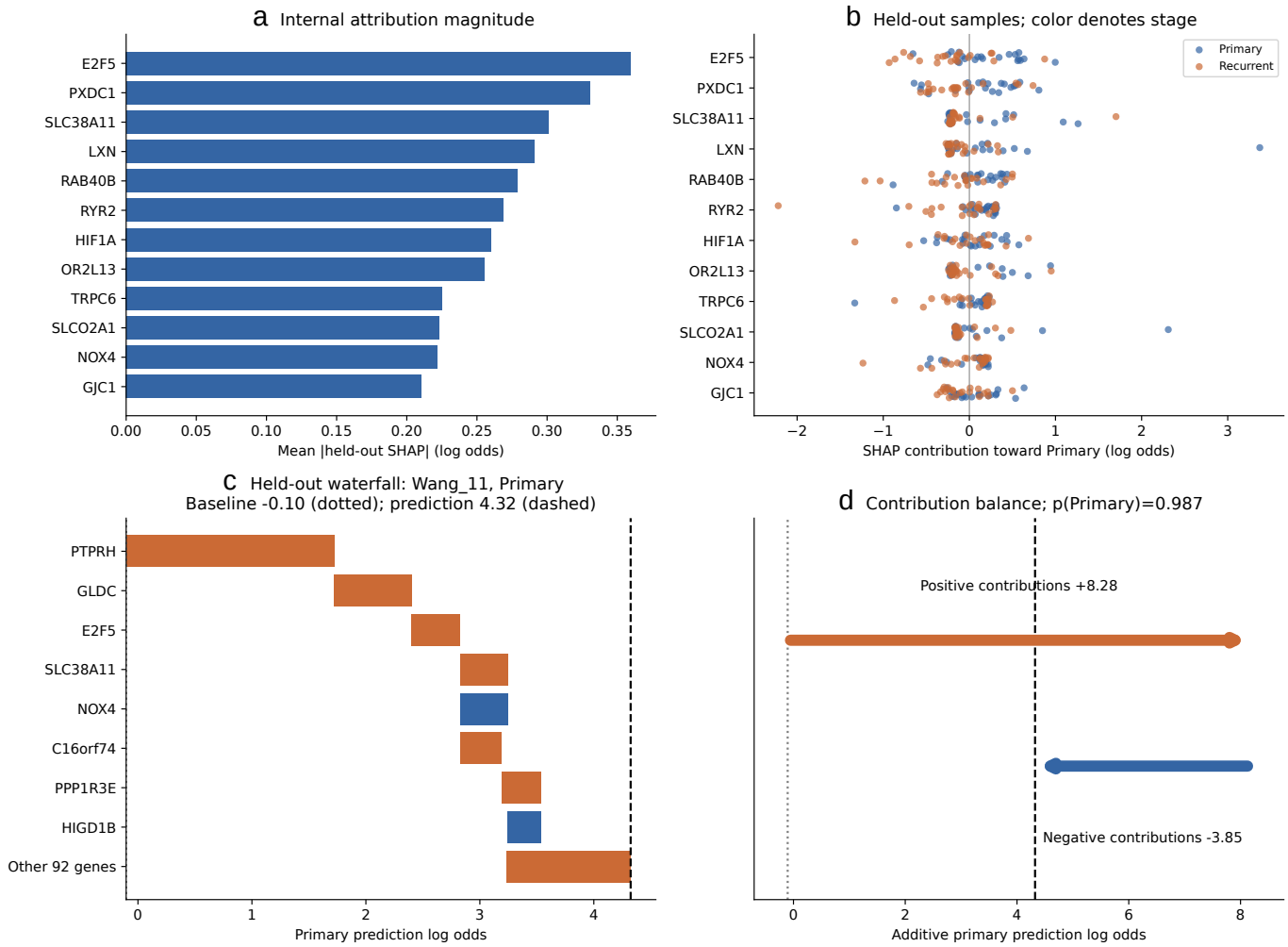
Supplementary Figure S12. Mean recurrent-minus-primary rank-program score changes in discovery malignant nuclei. The original 22-patient cohort, reconciled 28-patient cohort, 27-patient original-annotation subset, and 20-patient strict-CNV subset are compared as overlapping sensitivities, not independent cohorts. Intervals are 95% paired t intervals; the displayed FDRs come from the corresponding nine-test program/score families. No subset supports a corrected average program shift.

Supplementary Figure S13 Expression of the original ten follow-up genes by compartment



Supplementary Figure S13. The original fixed ten genes are shown across the displayed sampled compartments. Bubble size and color describe expression or detection as labeled. These are localization summaries, not statistical evidence of recurrence-specific dependency.

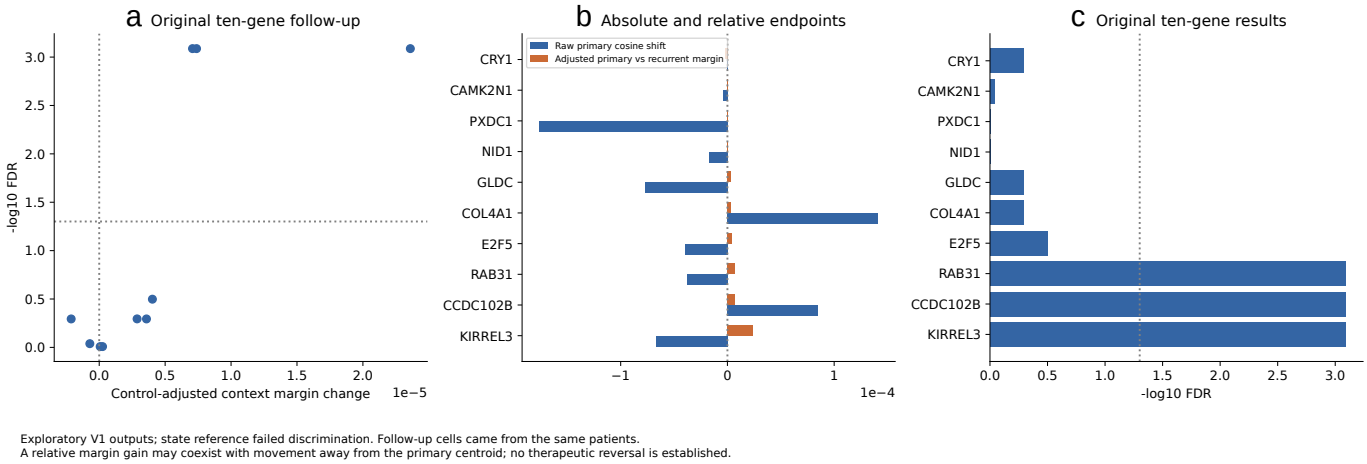
Supplementary Figure S14 Held-out RNA model attribution



Exploratory explanation of an internally fitted model that did not validate in CARE. Contributions are associative, not causal.
A–B average each sample over 20 held-out repeats; C–D explain one fixed held-out fold using its training-mean background.

Supplementary Figure S14. Linear interventional SHAP for held-out predictions using training-fold mean backgrounds. Only held-out discovery patients contribute. The waterfall uses the first patient in identifier order, fixed before inspecting a favorable prediction. Contributions are log-odds decompositions of a classifier, not causal gene effects [22].

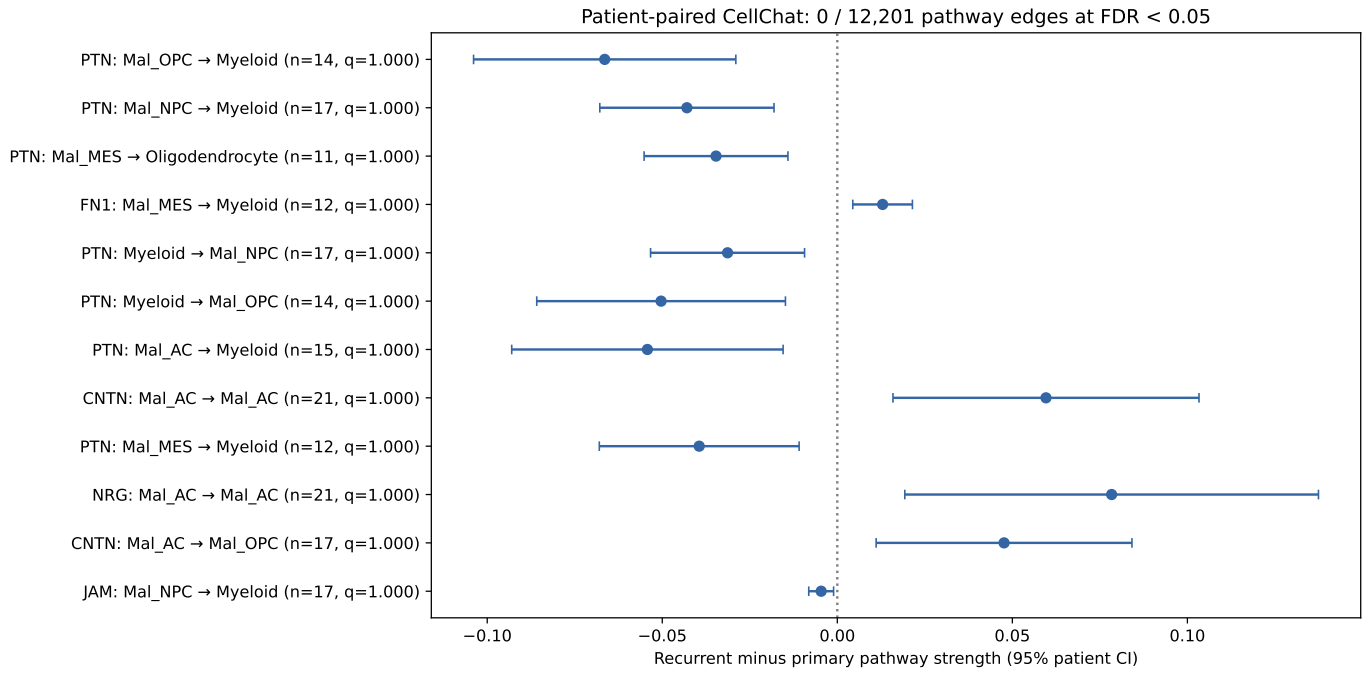
Supplementary Figure S15 Original Geneformer perturbation endpoints



Exploratory V1 outputs; state reference failed discrimination. Follow-up cells came from the same patients.
A relative margin gain may coexist with movement away from the primary centroid; no therapeutic reversal is established.

Supplementary Figure S15. a, Control-adjusted context-margin changes and FDR for the original ten exploratory follow-up genes. b, Absolute movement toward the primary centroid is distinguished from the primary-versus-recurrent margin change. c, Per-gene FDR values for that original endpoint. The original 22-pair whole-malignant reference failed discrimination; held-out cells came from the same patient set. A relative margin gain may coexist with movement away from the primary centroid and does not establish therapeutic reversal. Token-order and patient-label audits are retained in the accompanying tables.

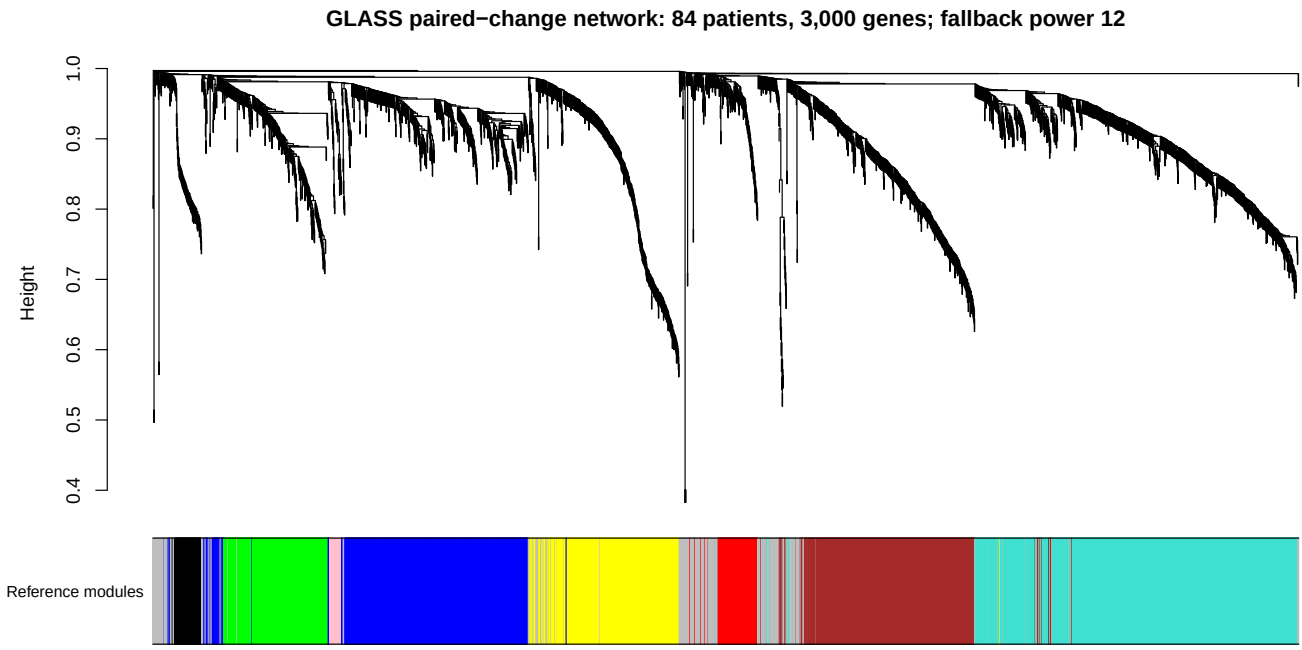
Supplementary Figure S16 Patient-paired CellChat comparisons



Shown: 12 smallest adjusted P values (ties ordered by nominal P). All eligible edges remain in the source table.
Fixed protein LR universe; paired shared scaling; no population-size weighting; missing cell groups not zero-filled.

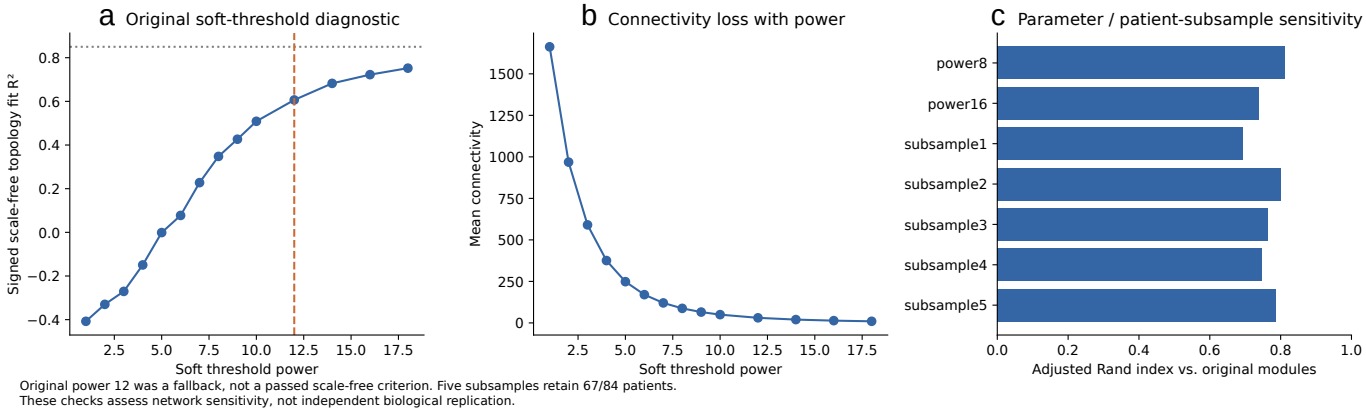
Supplementary Figure S16. Twenty-eight paired objects share within-patient expression scaling. Only within-stage pathway-source-target scores are compared. At least eight complete pairs are required for each edge; missing cell groups are not assumed to have zero signaling. None of 12,201 eligible edges survives joint FDR correction. Cell-label permutation tests and patient recurrence tests address different questions.

Supplementary Figure S17 GLASS reference network dendrogram



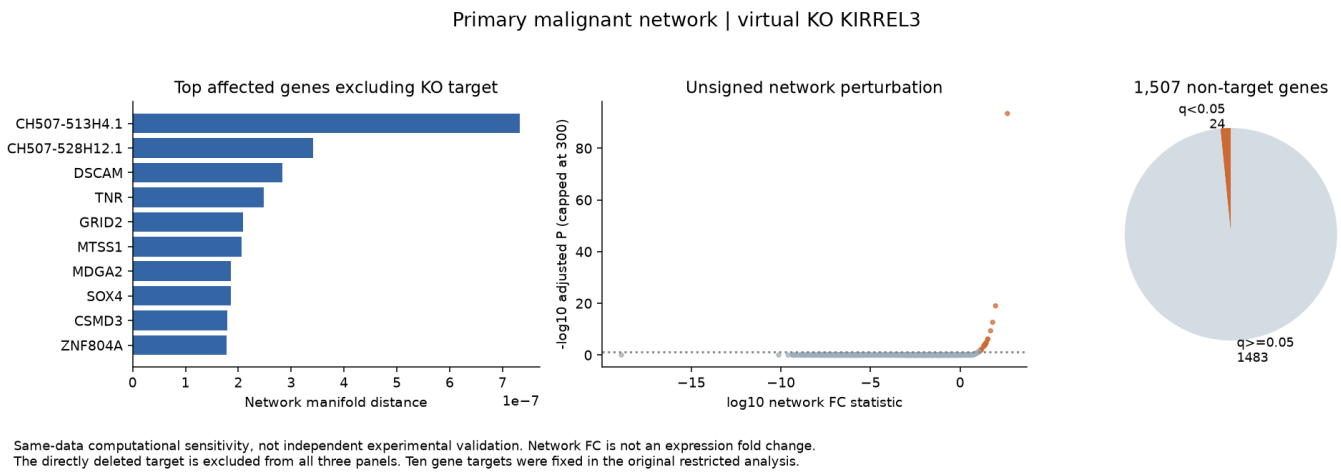
Supplementary Figure S17. WGCNA of 3,000 genes across 84 paired GLASS change profiles. Module colors encode the reference partition. Soft-threshold power 12 was a fallback because the scale-free criterion was not achieved; module membership alone is not validated mechanism.

Supplementary Figure S18 Network parameter and patient-sampling stability

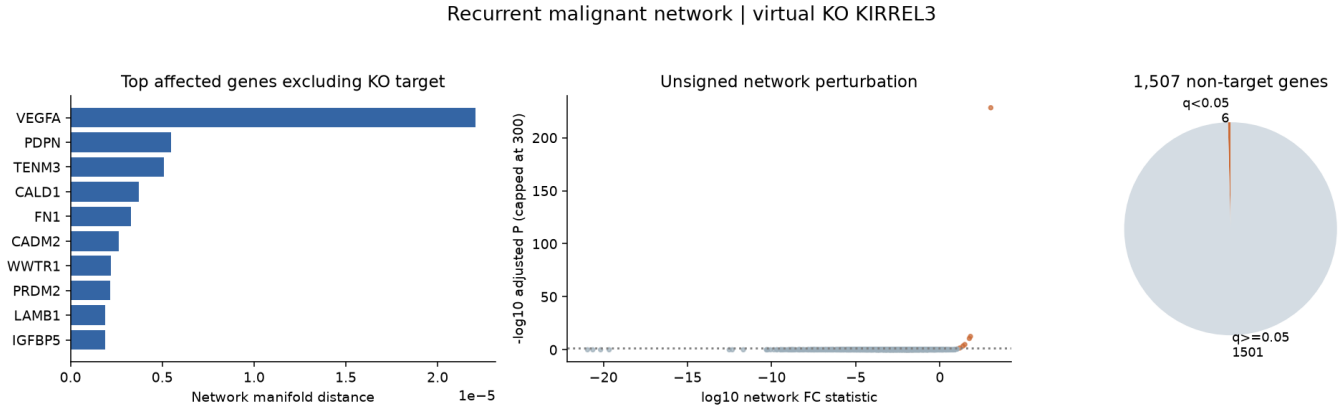


Supplementary Figure S18. a, Scale-free fit across soft-threshold powers; the reference power 12 is a fallback, not a passed fit criterion. b, Mean network connectivity across powers. c, Adjusted Rand indices against the reference partition for powers 8 and 16 and five fixed 67-of-84-patient subsets. Best-matching module overlaps are retained in the source tables. Parameter and sampling stability do not constitute independent biological replication.

Supplementary Figure S19 Individual network knockout of KIRREL3



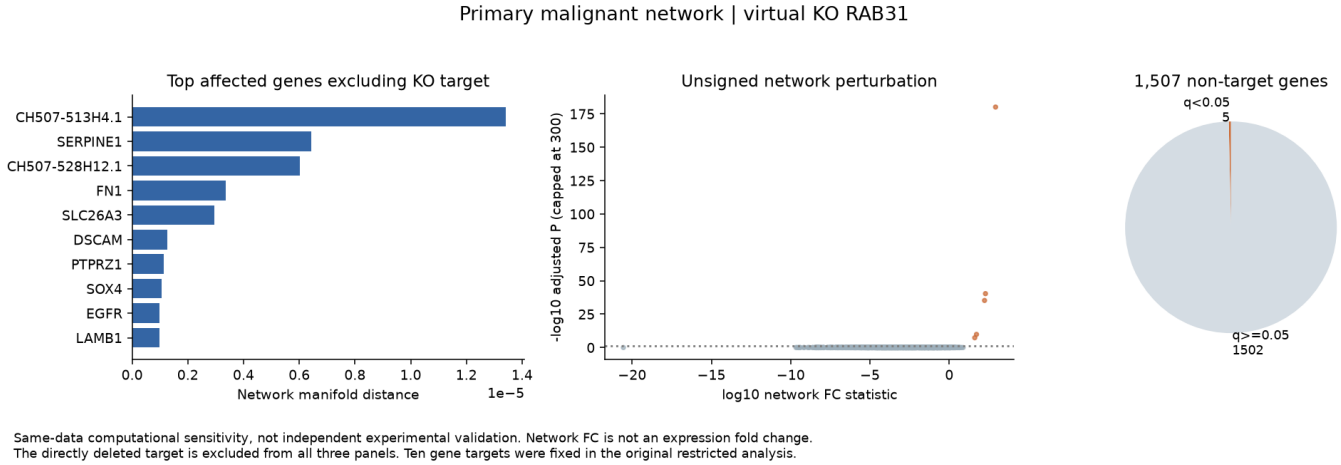
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.



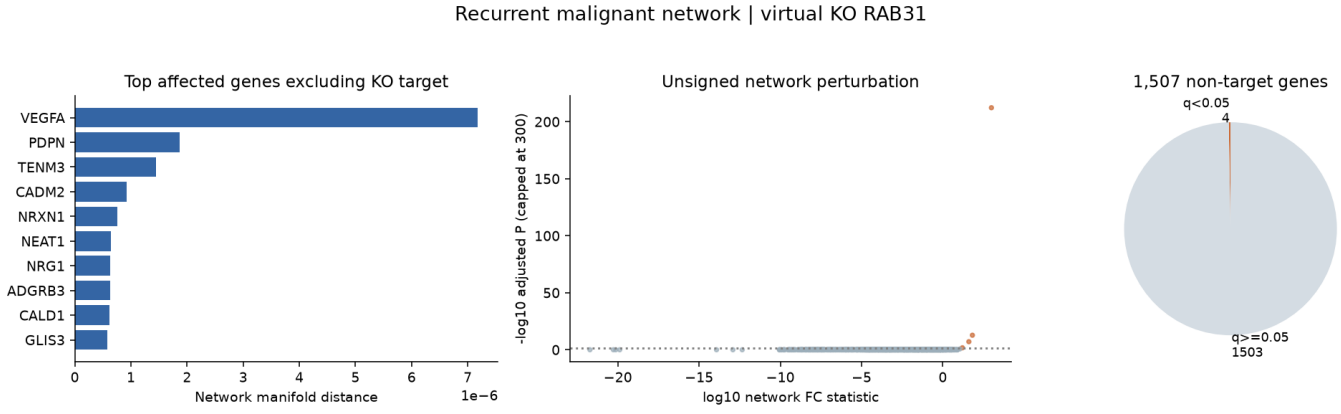
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.

Supplementary Figure S19. Upper, primary-stage network; lower, recurrent-stage network, for the original fixed KIRREL3 virtual knockout. The archived analysis uses 704 sampled malignant nuclei per stage from 22 paired patients and a 1,508-gene network. Ten resampled networks per stage support the displayed computational scores. The deleted gene is omitted from downstream rankings. Network FC is unsigned network dysregulation, not expression fold change; these plots do not establish an experimental knockout response or validate a recurrence-reversal treatment.

Supplementary Figure S20 Individual network knockout of RAB31



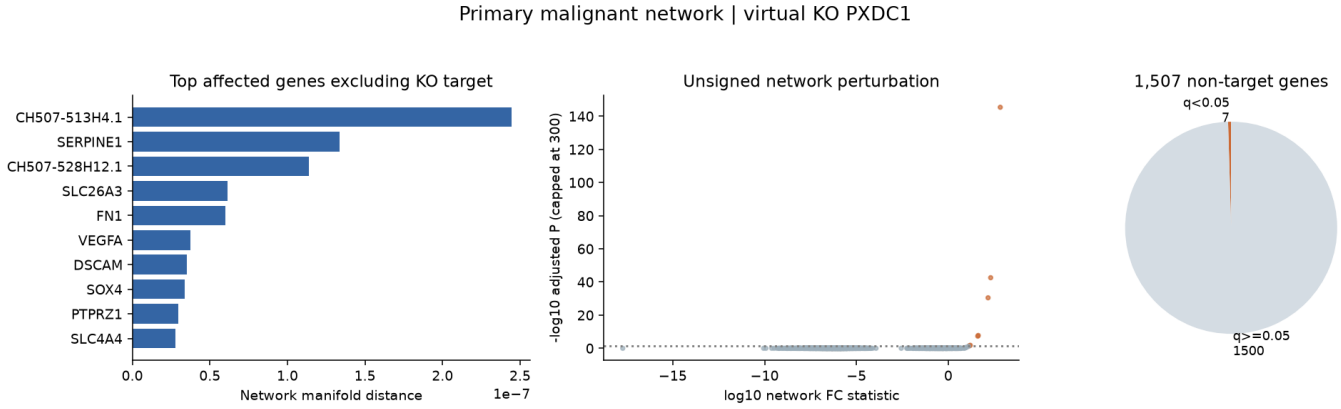
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.



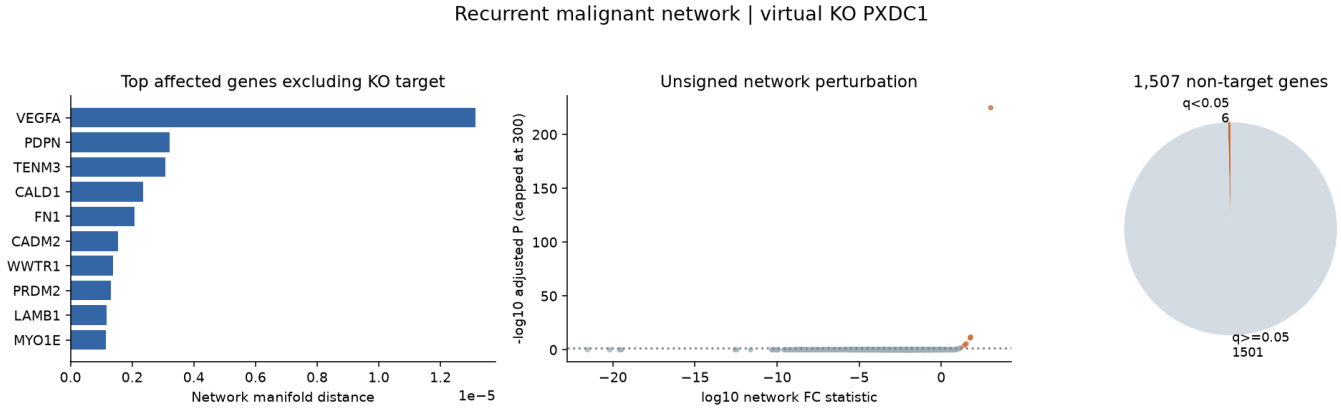
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.

Supplementary Figure S20. Upper, primary-stage network; lower, recurrent-stage network, for the original fixed RAB31 virtual knockout. The archived analysis uses 704 sampled malignant nuclei per stage from 22 paired patients and a 1,508-gene network. Ten resampled networks per stage support the displayed computational scores. The deleted gene is omitted from downstream rankings. Network FC is unsigned network dysregulation, not expression fold change; these plots do not establish an experimental knockout response or validate a recurrence-reversal treatment.

Supplementary Figure S21 Individual network knockout of PXDC1



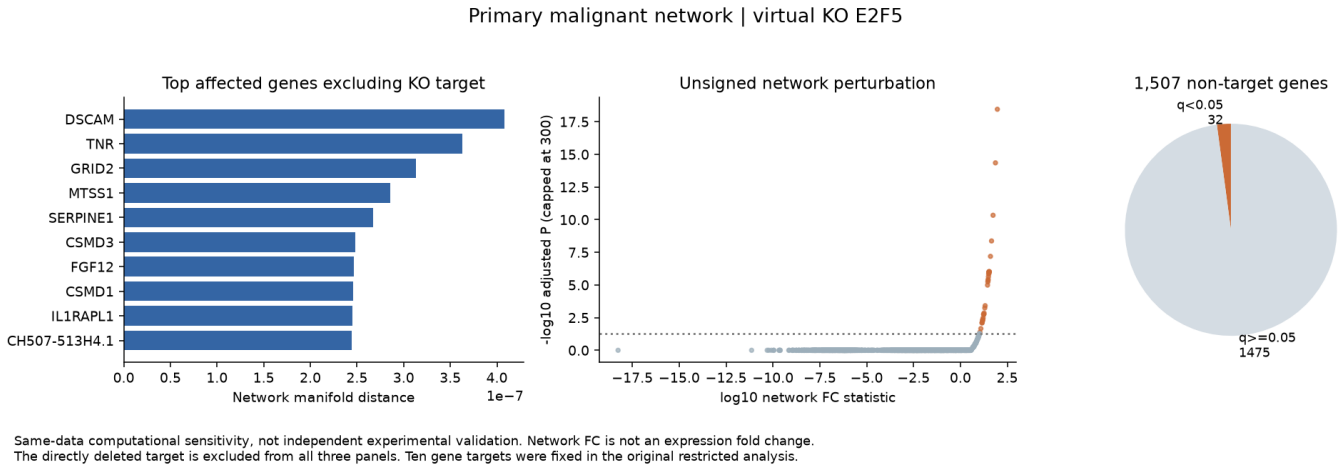
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.



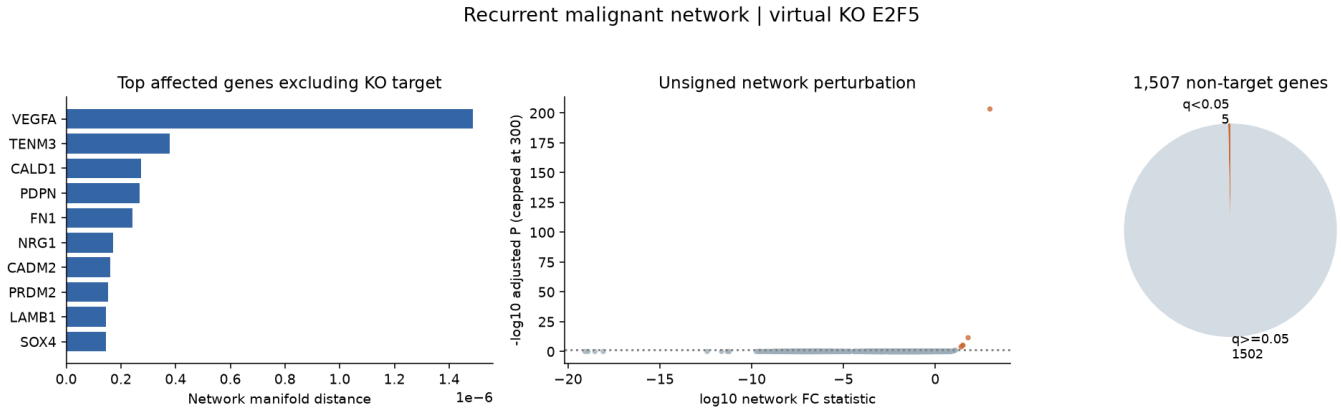
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.

Supplementary Figure S21. Upper, primary-stage network; lower, recurrent-stage network, for the original fixed PXDC1 virtual knockout. The archived analysis uses 704 sampled malignant nuclei per stage from 22 paired patients and a 1,508-gene network. Ten resampled networks per stage support the displayed computational scores. The deleted gene is omitted from downstream rankings. Network FC is unsigned network dysregulation, not expression fold change; these plots do not establish an experimental knockout response or validate a recurrence-reversal treatment.

Supplementary Figure S22 Individual network knockout of E2F5



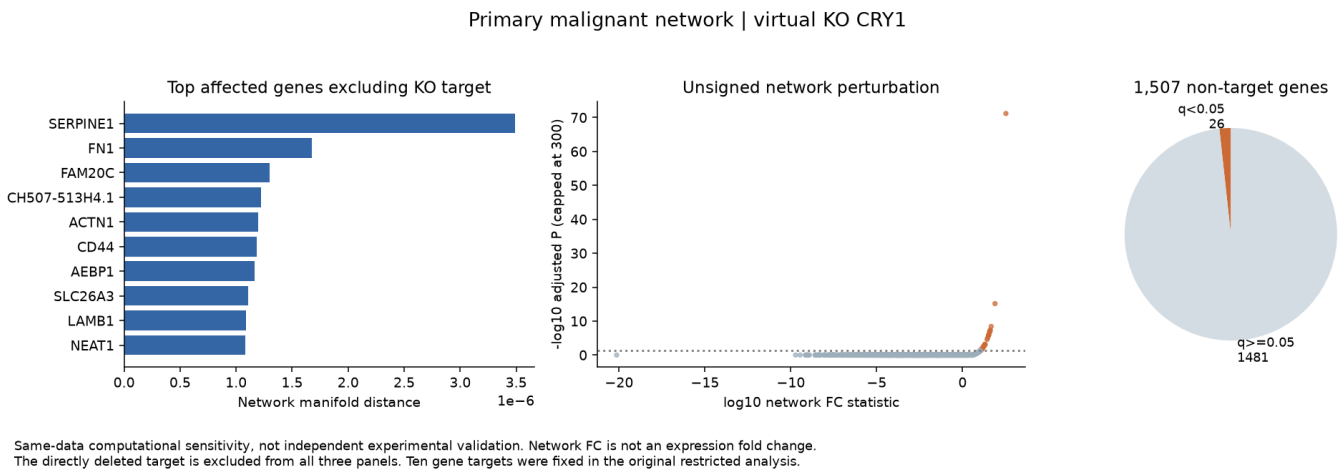
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.



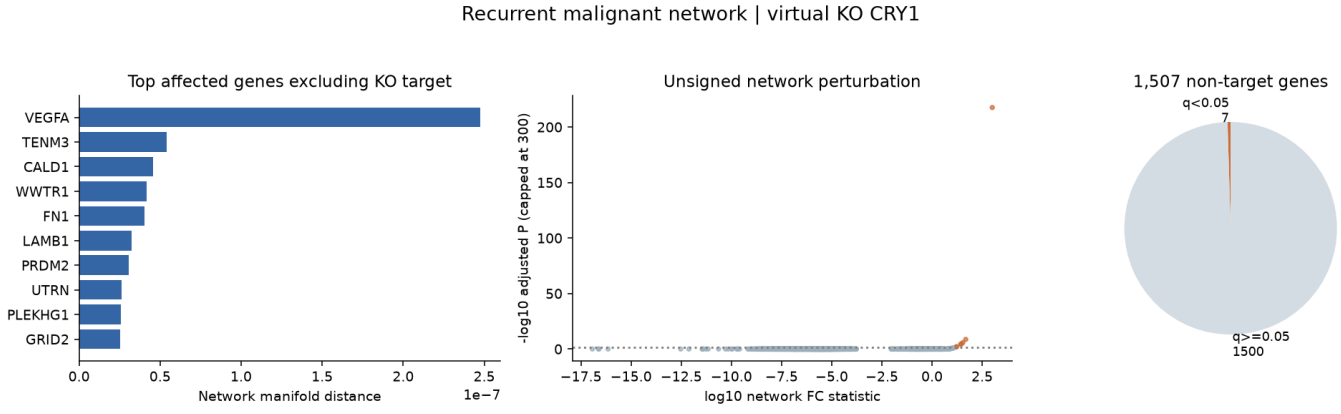
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.

Supplementary Figure S22. Upper, primary-stage network; lower, recurrent-stage network, for the original fixed E2F5 virtual knockout. The archived analysis uses 704 sampled malignant nuclei per stage from 22 paired patients and a 1,508-gene network. Ten resampled networks per stage support the displayed computational scores. The deleted gene is omitted from downstream rankings. Network FC is unsigned network dysregulation, not expression fold change; these plots do not establish an experimental knockout response or validate a recurrence-reversal treatment.

Supplementary Figure S23 Individual network knockout of CRY1



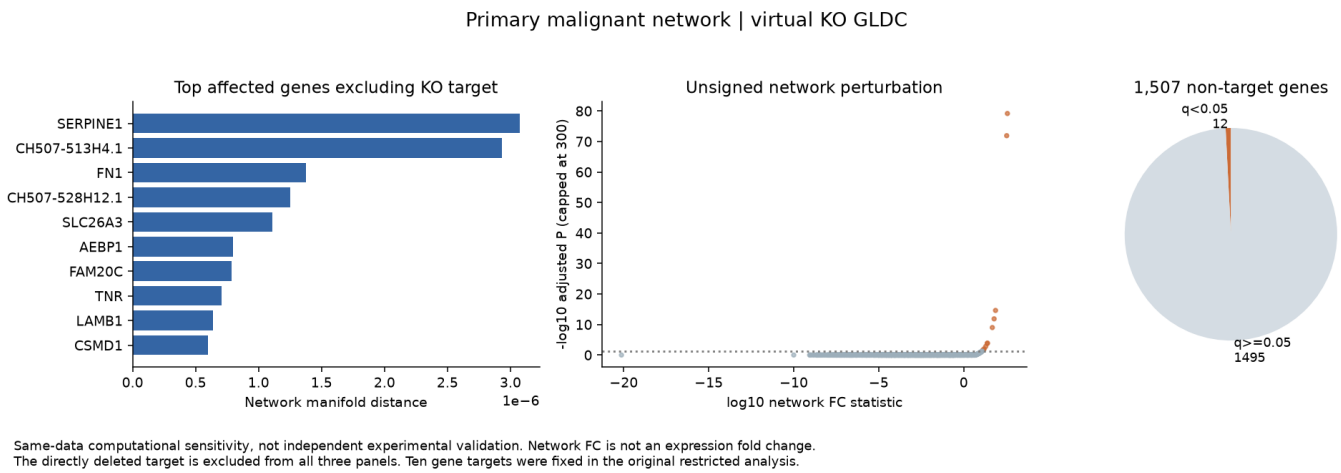
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.



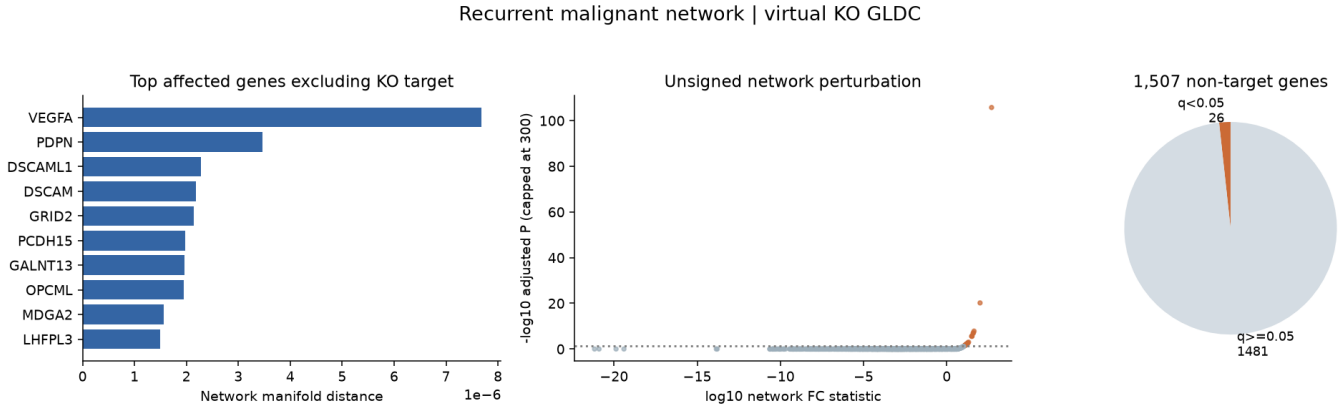
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.

Supplementary Figure S23. Upper, primary-stage network; lower, recurrent-stage network, for the original fixed CRY1 virtual knockout. The archived analysis uses 704 sampled malignant nuclei per stage from 22 paired patients and a 1,508-gene network. Ten resampled networks per stage support the displayed computational scores. The deleted gene is omitted from downstream rankings. Network FC is unsigned network dysregulation, not expression fold change; these plots do not establish an experimental knockout response or validate a recurrence-reversal treatment.

Supplementary Figure S24 Individual network knockout of GLDC



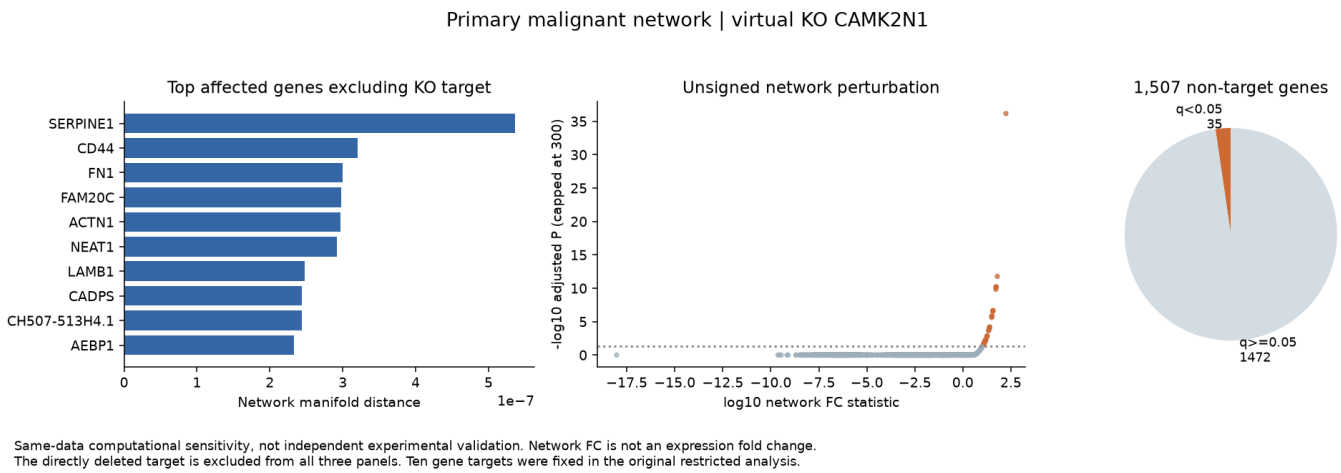
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.



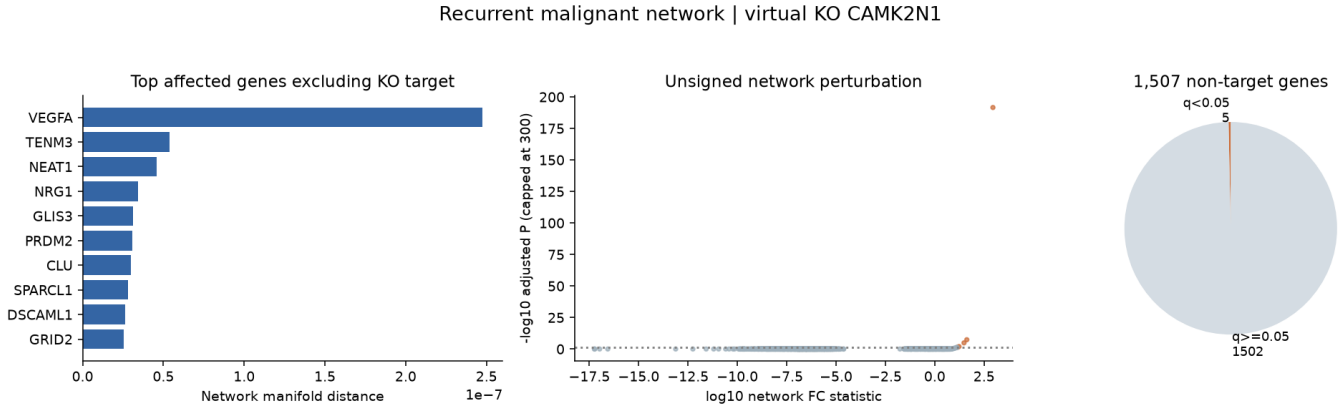
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.

Supplementary Figure S24. Upper, primary-stage network; lower, recurrent-stage network, for the original fixed GLDC virtual knockout. The archived analysis uses 704 sampled malignant nuclei per stage from 22 paired patients and a 1,508-gene network. Ten resampled networks per stage support the displayed computational scores. The deleted gene is omitted from downstream rankings. Network FC is unsigned network dysregulation, not expression fold change; these plots do not establish an experimental knockout response or validate a recurrence-reversal treatment.

Supplementary Figure S25 Individual network knockout of CAMK2N1



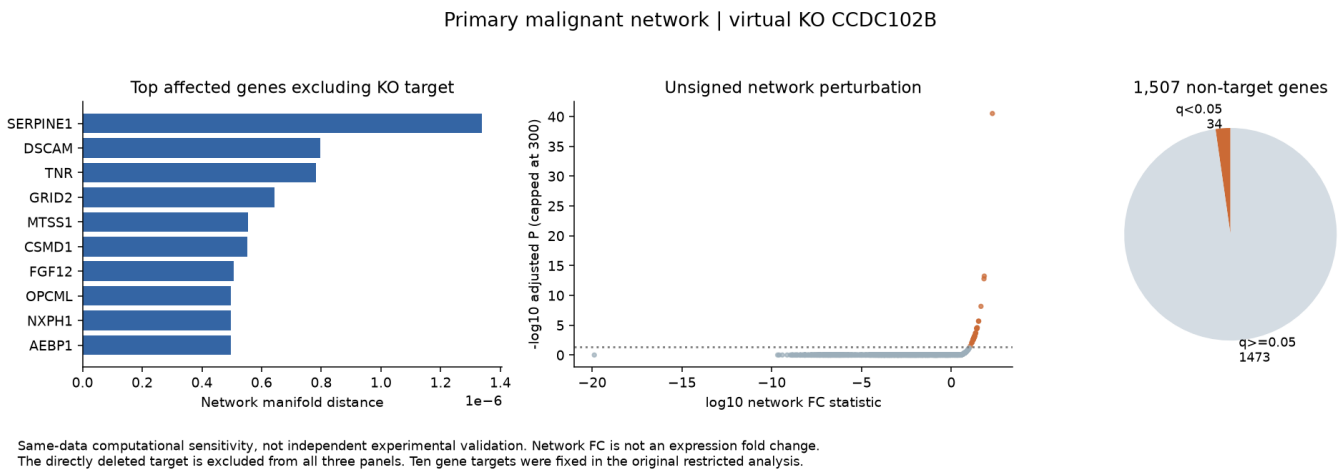
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.



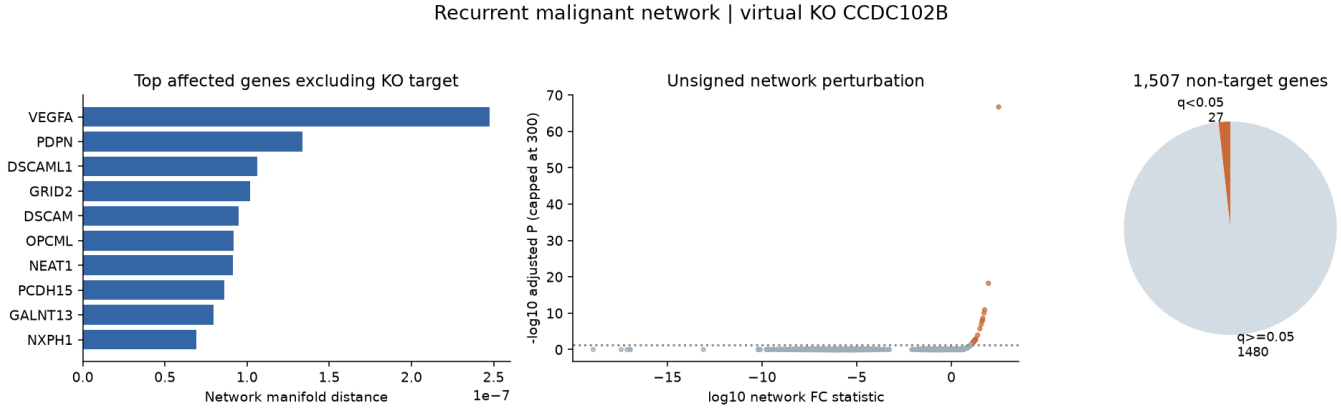
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.

Supplementary Figure S25. Upper, primary-stage network; lower, recurrent-stage network, for the original fixed CAMK2N1 virtual knockout. The archived analysis uses 704 sampled malignant nuclei per stage from 22 paired patients and a 1,508-gene network. Ten resampled networks per stage support the displayed computational scores. The deleted gene is omitted from downstream rankings. Network FC is unsigned network dysregulation, not expression fold change; these plots do not establish an experimental knockout response or validate a recurrence-reversal treatment.

Supplementary Figure S26 Individual network knockout of CCDC102B



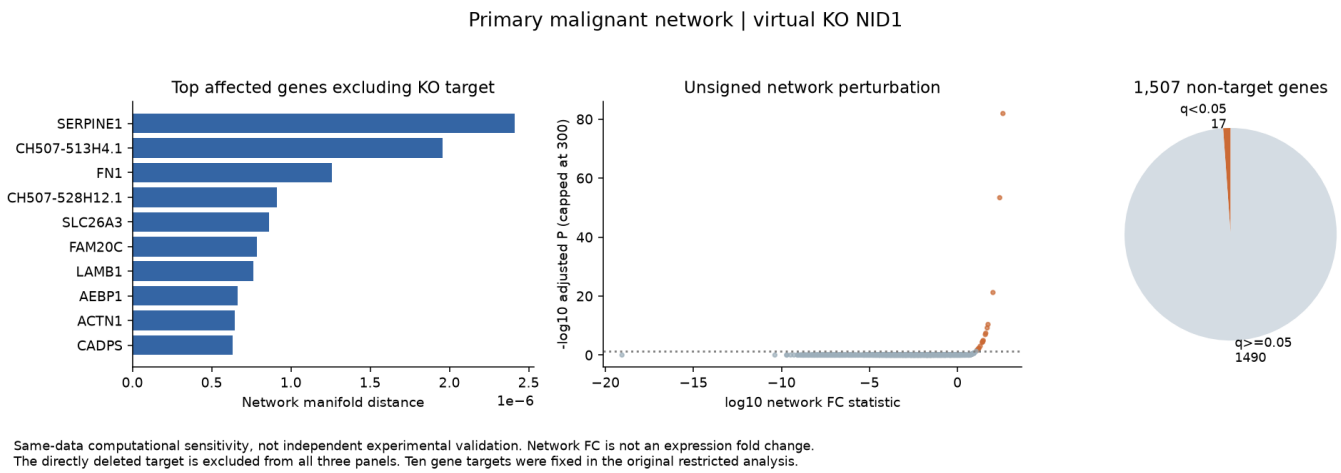
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.



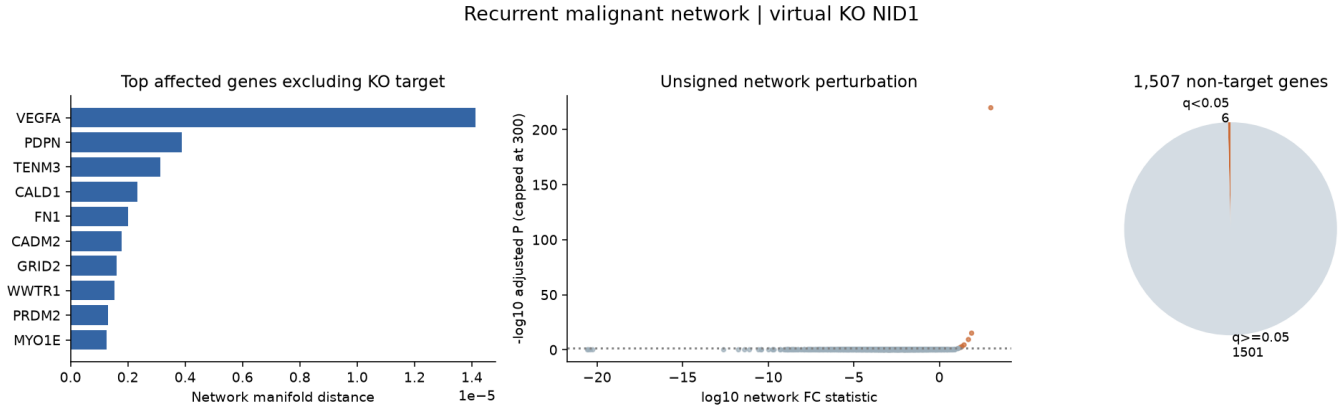
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.

Supplementary Figure S26. Upper, primary-stage network; lower, recurrent-stage network, for the original fixed CCDC102B virtual knockout. The archived analysis uses 704 sampled malignant nuclei per stage from 22 paired patients and a 1,508-gene network. Ten resampled networks per stage support the displayed computational scores. The deleted gene is omitted from downstream rankings. Network FC is unsigned network dysregulation, not expression fold change; these plots do not establish an experimental knockout response or validate a recurrence-reversal treatment.

Supplementary Figure S27 Individual network knockout of NID1



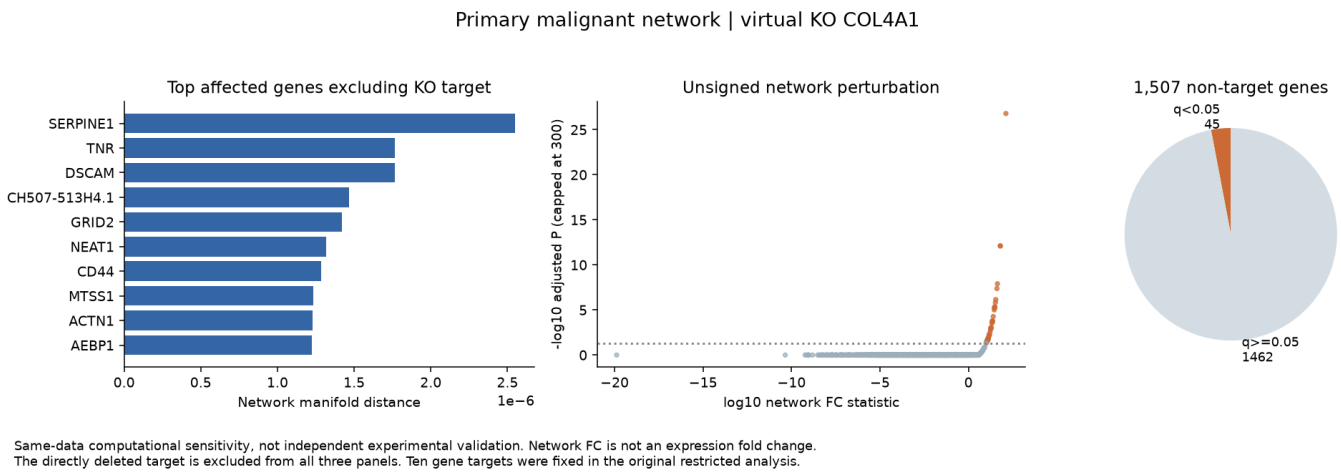
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.



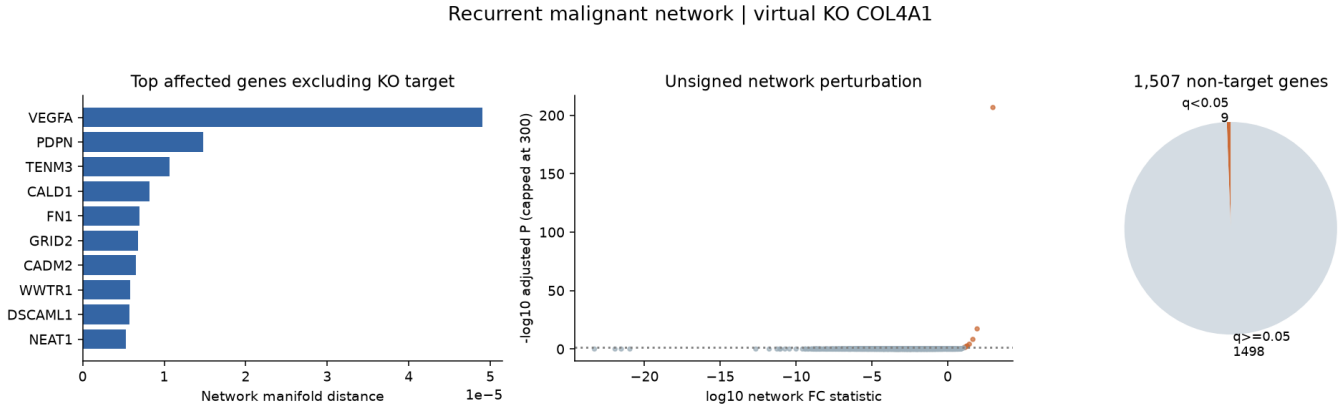
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.

Supplementary Figure S27. Upper, primary-stage network; lower, recurrent-stage network, for the original fixed NID1 virtual knockout. The archived analysis uses 704 sampled malignant nuclei per stage from 22 paired patients and a 1,508-gene network. Ten resampled networks per stage support the displayed computational scores. The deleted gene is omitted from downstream rankings. Network FC is unsigned network dysregulation, not expression fold change; these plots do not establish an experimental knockout response or validate a recurrence-reversal treatment.

Supplementary Figure S28 Individual network knockout of COL4A1



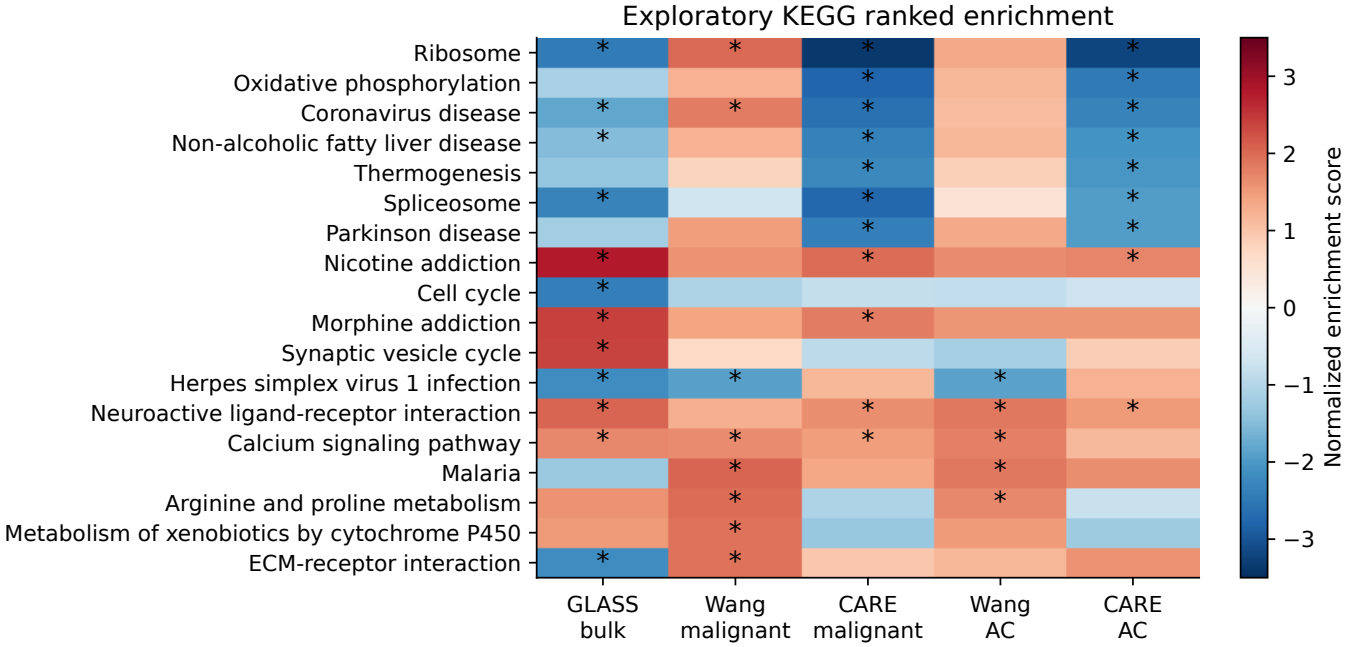
Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.



Same-data computational sensitivity, not independent experimental validation. Network FC is not an expression fold change.
The directly deleted target is excluded from all three panels. Ten gene targets were fixed in the original restricted analysis.

Supplementary Figure S28. Upper, primary-stage network; lower, recurrent-stage network, for the original fixed COL4A1 virtual knockout. The archived analysis uses 704 sampled malignant nuclei per stage from 22 paired patients and a 1,508-gene network. Ten resampled networks per stage support the displayed computational scores. The deleted gene is omitted from downstream rankings. Network FC is unsigned network dysregulation, not expression fold change; these plots do not establish an experimental knockout response or validate a recurrence-reversal treatment.

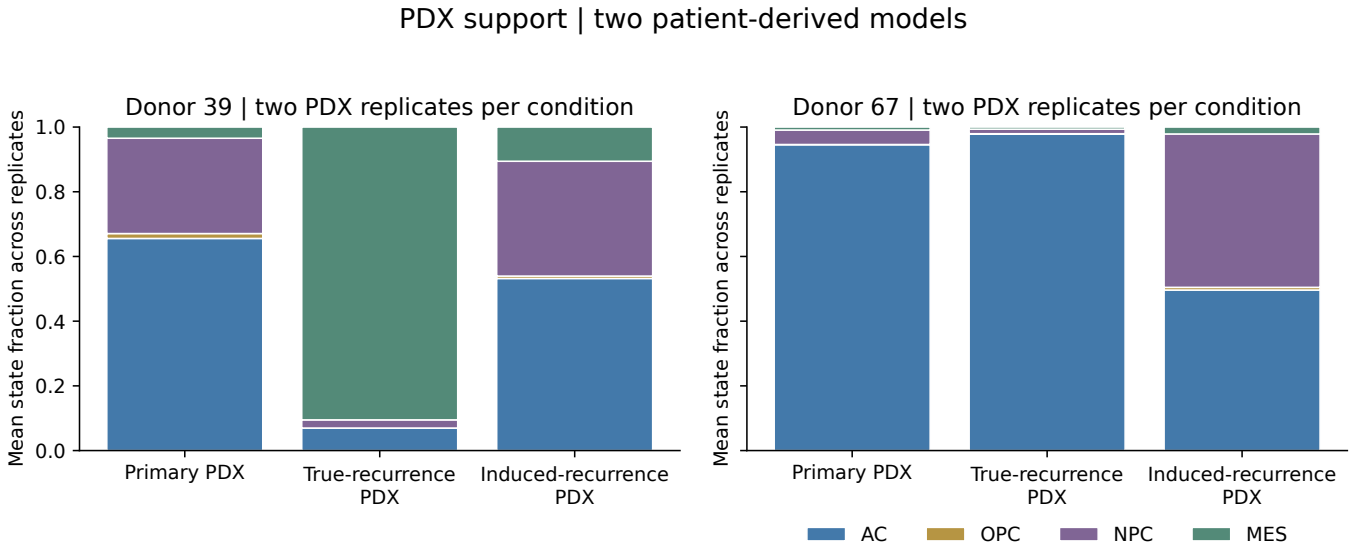
Supplementary Figure S29 KEGG enrichment and correlation-sensitive interpretation



Top 5 |NES| among within-analysis FDR < 0.05 per analysis; * joint FDR < 0.05 across all five analyses. KEGG 2021 Human. No pathway passed FDR < 0.05 in the sensitivity that estimates inter-gene correlation (CAMERA; either single-cell cohort/compartment). Pathway support is method-dependent.

Supplementary Figure S29. Ranked enrichment across bulk and malignant-compartment analyses. BH correction covers the tested analysis family. CAMERA sensitivity with estimated inter-gene correlation yielded no corrected pathway in the single-cell compartments; the displayed enrichment is therefore method-dependent.

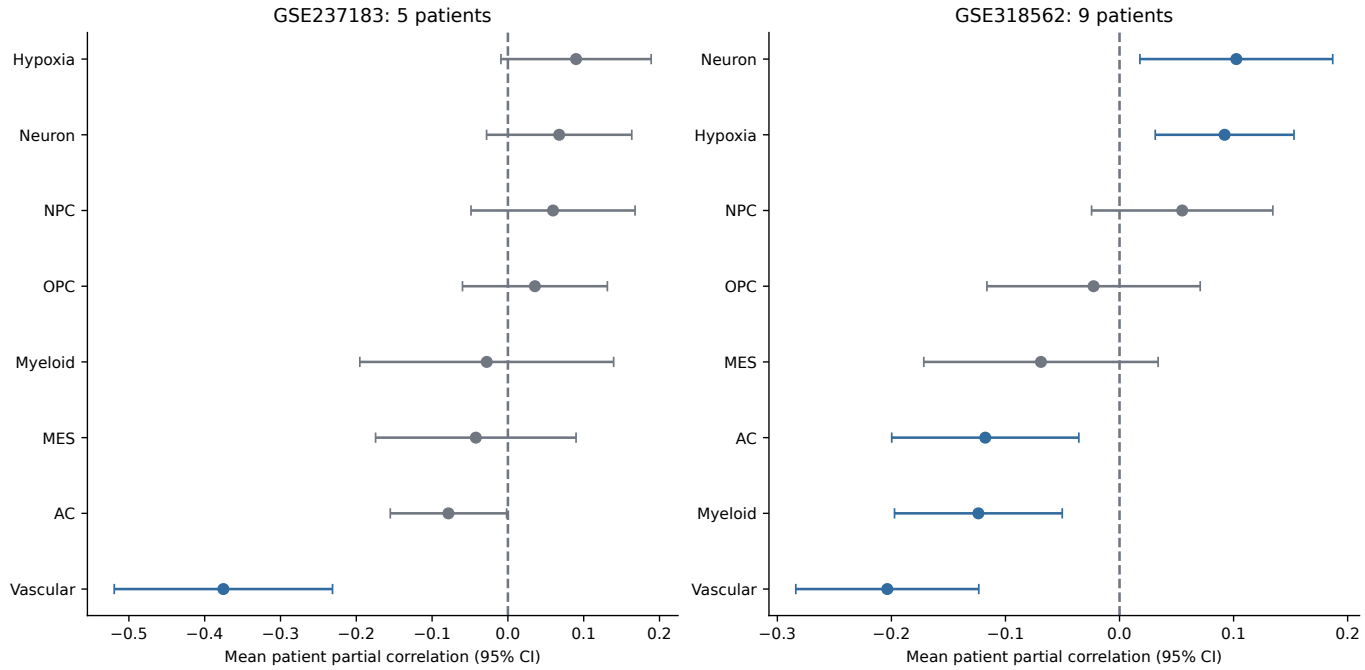
Supplementary Figure S30 Patient-derived model state composition



Human-mapped GIC counts; four-state projection. Replicates are not independent patients.
No cell-level significance tests; these models do not independently validate a new drug target.

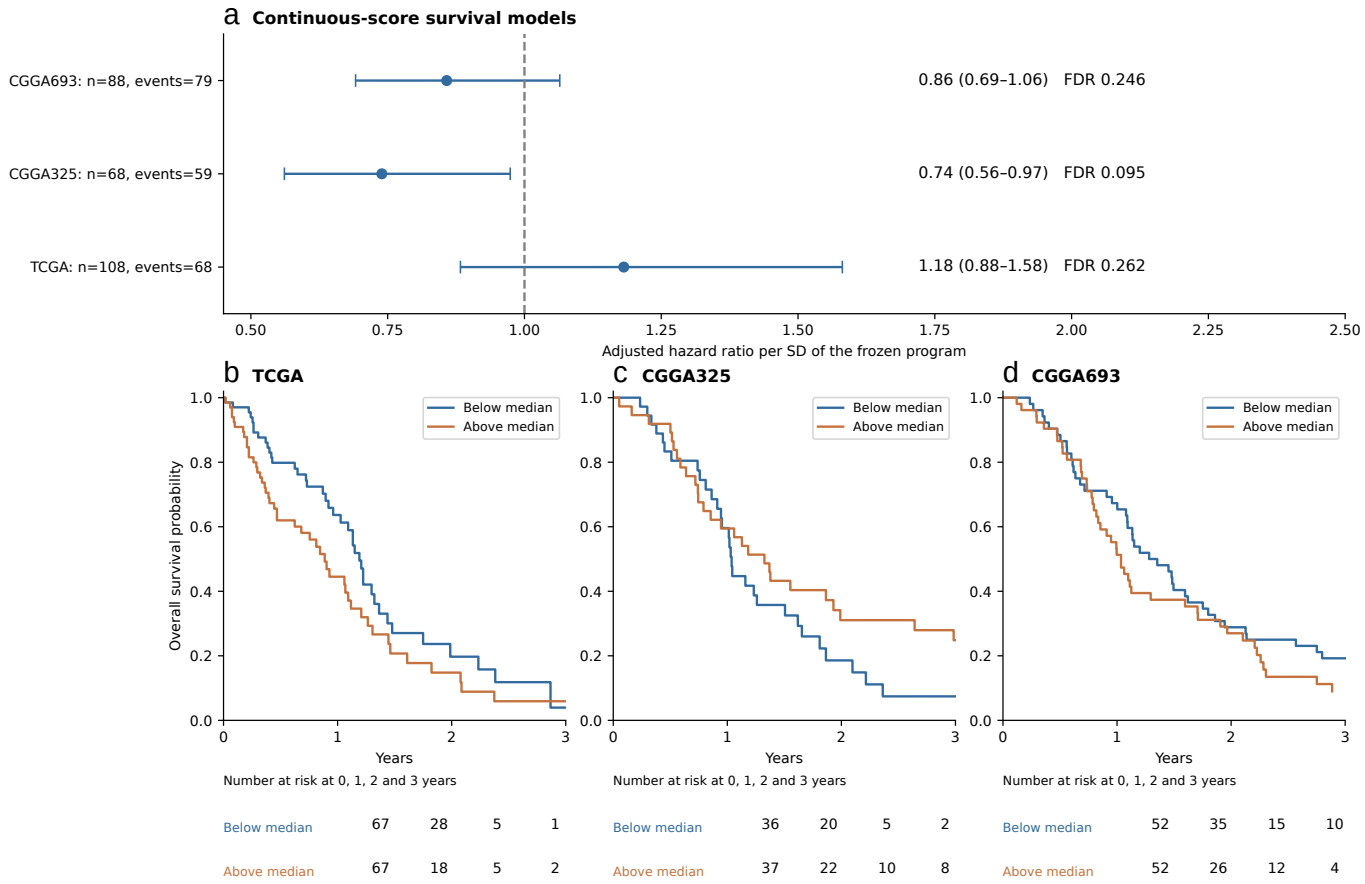
Supplementary Figure S30. GSE271619 contains 12 specimens from only two patient-derived models. Biological-model denominators are two; individual libraries do not supply independent patient replication. Divergent state responses are descriptive and are not a clinical recurrence validation.

Supplementary Figure S31 Spatial inference with patients as replicates



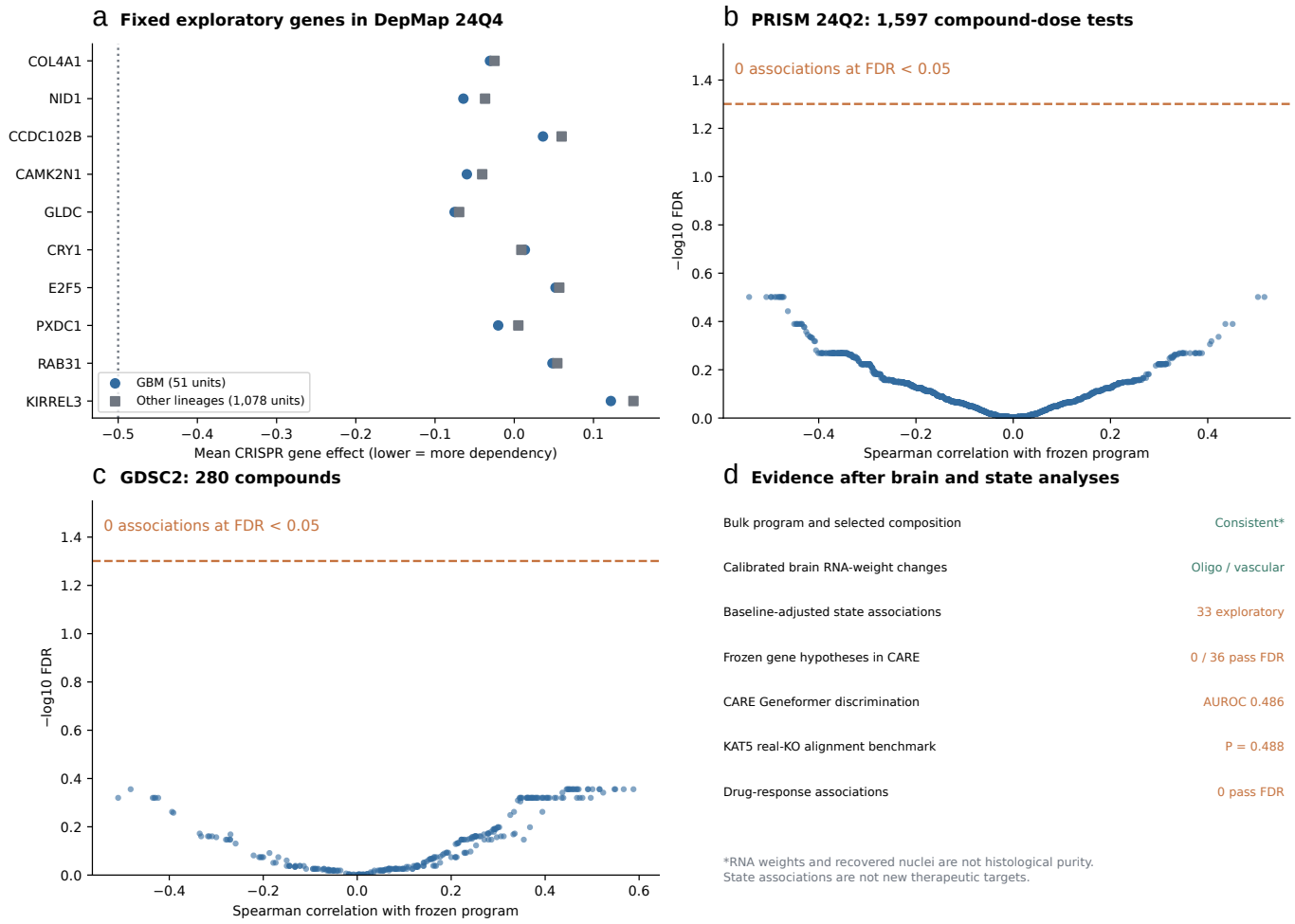
Supplementary Figure S31. GSE237183 and GSE318562 contribute five and nine patients. Section-level partial correlations adjust for UMI, use non-overlapping marker genes, and are aggregated within patient. The two-patient third cohort remains descriptive. Physical spots are not independent clinical replicates or invented core-edge annotations.

Supplementary Figure S32 Clinical outcome associations of the frozen tissue program



Supplementary Figure S32. TCGA, CGGA325 and CGGA693 outcome-eligible cohorts include 134, 73 and 104 patients, with adjusted complete-case counts 108, 68 and 88. Continuous-score hazard ratios have inconsistent directions and none survives correction across three adjusted models. TCGA proportional-hazards violations motivate the reported time-varying sensitivity. A prognostic biomarker is not established.

Supplementary Figure S33 Functional context and current evidence



Supplementary Figure S33. Archived DepMap 24Q4, PRISM 24Q2, and GDSC2 8.5 context, with the current evidence summary. a–c show dependency and response associations for the frozen program and ten original follow-up genes. d distinguishes calibrated composition findings, exploratory baseline-adjusted gene-axis associations, and unsuccessful model or target validation. These cell-line associations do not resolve matched human recurrence or establish therapeutic efficacy.

Supplementary references

Pengran Yu, Kaihua Cao, Zhihua Cheng, Yu Zeng, WenChuan Zhang
Department of Neurosurgery, Shanghai Ninth People's Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200011, China
Correspondence: zhangwench88@sjtu.edu.cn

1. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. New England Journal of Medicine. 2005;352:987-996. https://doi.org/10.1056/NEJMoa043330

2. Hegi ME, Diserens AC, Gorlia T, et al. MGMT Gene Silencing and Benefit from Temozolomide in Glioblastoma. New England Journal of Medicine. 2005;352:997-1003. https://doi.org/10.1056/NEJMoa043331

3. Varn FS, Johnson KC, Martinek J, et al. Glioma progression is shaped by genetic evolution and microenvironment interactions. Cell. 2022;185:2184-2199.e16. https://doi.org/10.1016/j.cell.2022.04.038

4. Wang L, Jung J, Babikir H, et al. A single-cell atlas of glioblastoma evolution under therapy reveals cell-intrinsic and cell-extrinsic therapeutic targets. Nature Cancer. 2022;3:1534-1552. https://doi.org/10.1038/s43018-022-00475-x

5. Verhaak RG, Hoadley KA, Purdom E, et al. Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell. 2010;17:98-110. https://doi.org/10.1016/j.ccr.2009.12.020

6. Wang Q, Hu B, Hu X, et al. Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. Cancer Cell. 2017;32:42-56.e6. https://doi.org/10.1016/j.ccell.2017.06.003

7. Patel AP, Tirosh I, Trombetta JJ, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science. 2014;344:1396-1401. https://doi.org/10.1126/science.1254257

8. Richards LM, Whitley OKN, MacLeod G, et al. Gradient of Developmental and Injury Response transcriptional states defines functional vulnerabilities underpinning glioblastoma heterogeneity. Nature Cancer. 2021;2:157-173. https://doi.org/10.1038/s43018-020-00154-9

9. Neftel C, Laffy J, Filbin MG, et al. An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. Cell. 2019;178:835-849.e21. https://doi.org/10.1016/j.cell.2019.06.024

10. Theodoris CV, Xiao L, Chopra A, et al. Transfer learning enables predictions in network biology. Nature. 2023;618:616-624. https://doi.org/10.1038/s41586-023-06139-9

11. Zhang Y, Venkatesh MS, Theodoris CV. Discovery of candidate therapeutic targets with Geneformer. Nature Protocols. 2026. https://doi.org/10.1038/s41596-026-01364-8

12. Chen H, Venkatesh MS, Gómez Ortega J, et al. Scaling and quantization of large-scale foundation model enables resource-efficient predictions in network biology. Nature Computational Science. 2026;6:450-463. https://doi.org/10.1038/s43588-026-00972-4

13. Osorio D, Zhong Y, Li G, et al. scTenifoldKnk: An efficient virtual knockout tool for gene function predictions via single-cell gene regulatory network perturbation. Patterns. 2022;3:100434. https://doi.org/10.1016/j.patter.2022.100434

14. Nomura M, Spitzer A, Johnson KC, et al. The multilayered transcriptional architecture of glioblastoma ecosystems. Nature Genetics. 2025;57:1155-1167.

https://doi.org/10.1038/s41588-025-02167-5

15. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. Nature Methods. 2015;12:453-457. https://doi.org/10.1038/nmeth.3337

16. Wang X, Park J, Susztak K, et al. Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. Nature Communications. 2019;10:380. https://doi.org/10.1038/s41467-018-08023-x

17. Jew B, Alvarez M, Rahmani E, et al. Accurate estimation of cell composition in bulk expression through robust integration of single-cell information. Nature Communications. 2020;11:1971. https://doi.org/10.1038/s41467-020-15816-6

18. Sutton GJ, Poppe D, Simmons RK, et al. Comprehensive evaluation of deconvolution methods for human brain gene expression. Nature Communications. 2022;13:1358. https://doi.org/10.1038/s41467-022-28655-4

19. Avila Cobos F, Alquicira-Hernandez J, Powell JE, et al. Benchmarking of cell type deconvolution pipelines for transcriptomics data. Nature Communications. 2020;11:5650. https://doi.org/10.1038/s41467-020-19015-1

20. Spitzer A, Johnson KC, Nomura M, et al. Deciphering the longitudinal trajectories of glioblastoma ecosystems by integrative single-cell genomics. Nature Genetics. 2025;57:1168-1178. https://doi.org/10.1038/s41588-025-02168-4

21. Mihalas AB, Arora S, O'Connor SA, et al. KAT5 regulates neurodevelopmental states associated with G0-like populations in glioblastoma. Nature Communications. 2025;16:4327. https://doi.org/10.1038/s41467-025-59503-w

22. Lundberg SM, Lee SI. A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems. 2017;30:4765–4774. https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

23. Al-Dalahmah O, Argenziano MG, Kannan A, et al. Re-convolving the compositional landscape of primary and recurrent glioblastoma reveals prognostic and targetable tissue states. Nature Communications. 2023;14:2586. https://doi.org/10.1038/s41467-023-38186-1

24. Greenwald AC, Darnell NG, Hoefflin R, et al. Integrative spatial analysis reveals a multi-layered organization of glioblastoma. Cell. 2024;187:2485-2501.e26. https://doi.org/10.1016/j.cell.2024.03.029

25. Puchalski RB, Shah N, Miller J, et al. An anatomic transcriptional atlas of human glioblastoma. Science. 2018;360:660-663. https://doi.org/10.1126/science.aaf2666

26. Tirosh I, Venteicher AS, Hebert C, et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. Nature. 2016;539:309-313. https://doi.org/10.1038/nature20123

27. Hara T, Chanoch-Myers R, Mathewson ND, et al. Interactions between cancer cells and immune cells drive transitions to mesenchymal-like states in glioblastoma. Cancer Cell. 2021;39:779-792.e11. https://doi.org/10.1016/j.ccell.2021.05.002

28. Büttner M, Ostner J, Müller CL, et al. scCODA is a Bayesian model for compositional single-cell data analysis. Nature Communications. 2021;12:6876. https://doi.org/10.1038/s41467-021-27150-6

29. Kolde R, Laur S, Adler P, et al. Robust rank aggregation for gene list integration and meta-analysis. Bioinformatics. 2012;28:573-580. https://doi.org/10.1093/bioinformatics/btr709

30. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research. 2015;43:e47.

https://doi.org/10.1093/nar/gkv007

31. Chen Y, Chen L, Lun ATL, et al. edgeR v4: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets. Nucleic Acids Research. 2025;53:gkaf018. https://doi.org/10.1093/nar/gkaf018

32. Squair JW, Gautier M, Kathe C, et al. Confronting false discoveries in single-cell differential expression. Nature Communications. 2021;12:5692. https://doi.org/10.1038/s41467-021-25960-2

33. Zimmerman KD, Espeland MA, Langefeld CD. A practical solution to pseudoreplication bias in single-cell studies. Nature Communications. 2021;12:738. https://doi.org/10.1038/s41467-021-21038-1

34. Crowell HL, Soneson C, Germain PL, et al. muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. Nature Communications. 2020;11:6077. https://doi.org/10.1038/s41467-020-19894-4

35. Soneson C, Robinson MD. Bias, robustness and scalability in single-cell differential expression analysis. Nature Methods. 2018;15:255-261. https://doi.org/10.1038/nmeth.4612

36. Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. Molecular Systems Biology. 2019;15:e8746. https://doi.org/10.15252/msb.20188746

37. Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using CellChat. Nature Communications. 2021;12:1088. https://doi.org/10.1038/s41467-021-21246-9

38. Jin S, Plikus MV, Nie Q. CellChat for systematic analysis of cell–cell communication from single-cell transcriptomics. Nature Protocols. 2025;20:180-219. https://doi.org/10.1038/s41596-024-01045-4

39. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences. 2005;102:15545-15550. https://doi.org/10.1073/pnas.0506580102

40. Liberzon A, Birger C, Thorvaldsdóttir H, et al. The Molecular Signatures Database Hallmark Gene Set Collection. Cell Systems. 2015;1:417-425. https://doi.org/10.1016/j.cels.2015.12.004

41. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. Nucleic Acids Research. 2012;40:e133. https://doi.org/10.1093/nar/gks461

42. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559. https://doi.org/10.1186/1471-2105-9-559

43. Corsello SM, Nagari RT, Spangler RD, et al. Discovering the anticancer potential of non-oncology drugs by systematic viability profiling. Nature Cancer. 2020;1:235-248. https://doi.org/10.1038/s43018-019-0018-6

44. Yang W, Soares J, Greninger P, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Research. 2013;41:D955-D961. https://doi.org/10.1093/nar/gks1111

45. Zhao Z, Zhang KN, Wang Q, et al. Chinese Glioma Genome Atlas (CGGA): A Comprehensive Resource with Functional Genomic Data from Chinese Glioma Patients. Genomics, Proteomics & Bioinformatics. 2021;19:1-12. https://doi.org/10.1016/j.gpb.2020.10.005

46. Brennan CW, Verhaak RG, McKenna A, et al. The Somatic Genomic Landscape of Glioblastoma. Cell. 2013;155:462-477. https://doi.org/10.1016/j.cell.2013.09.034

47. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B: Statistical

Methodology. 1995;57:289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x