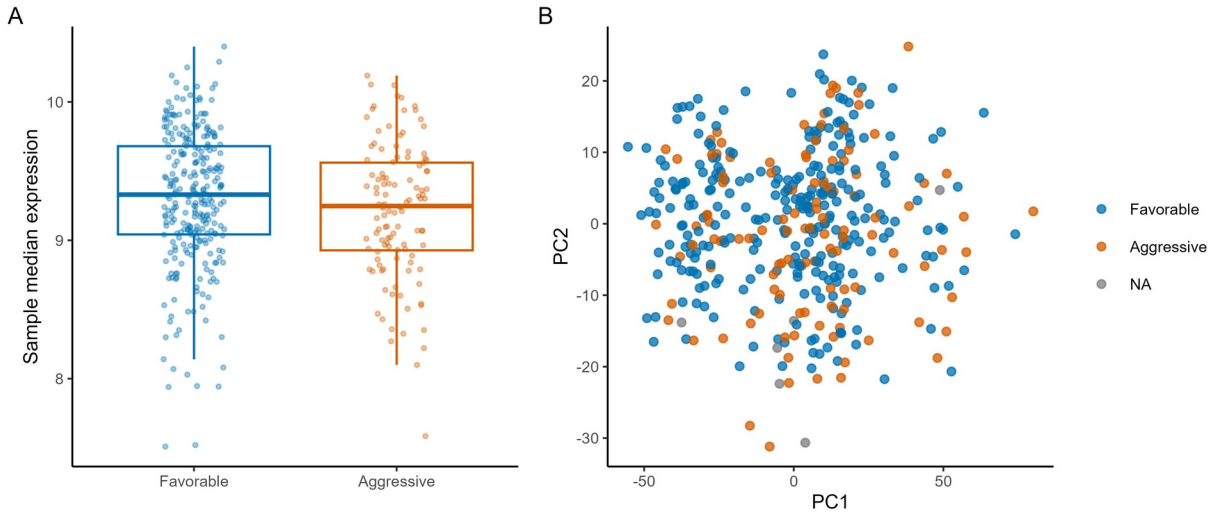


Supplementary Materials

Differentiation–Proliferation Imbalance and Therapy-Associated State Remodeling in MYCN-Non-Amplified Neuroblastoma: Multi-Cohort and Paired Single-Cell Analyses

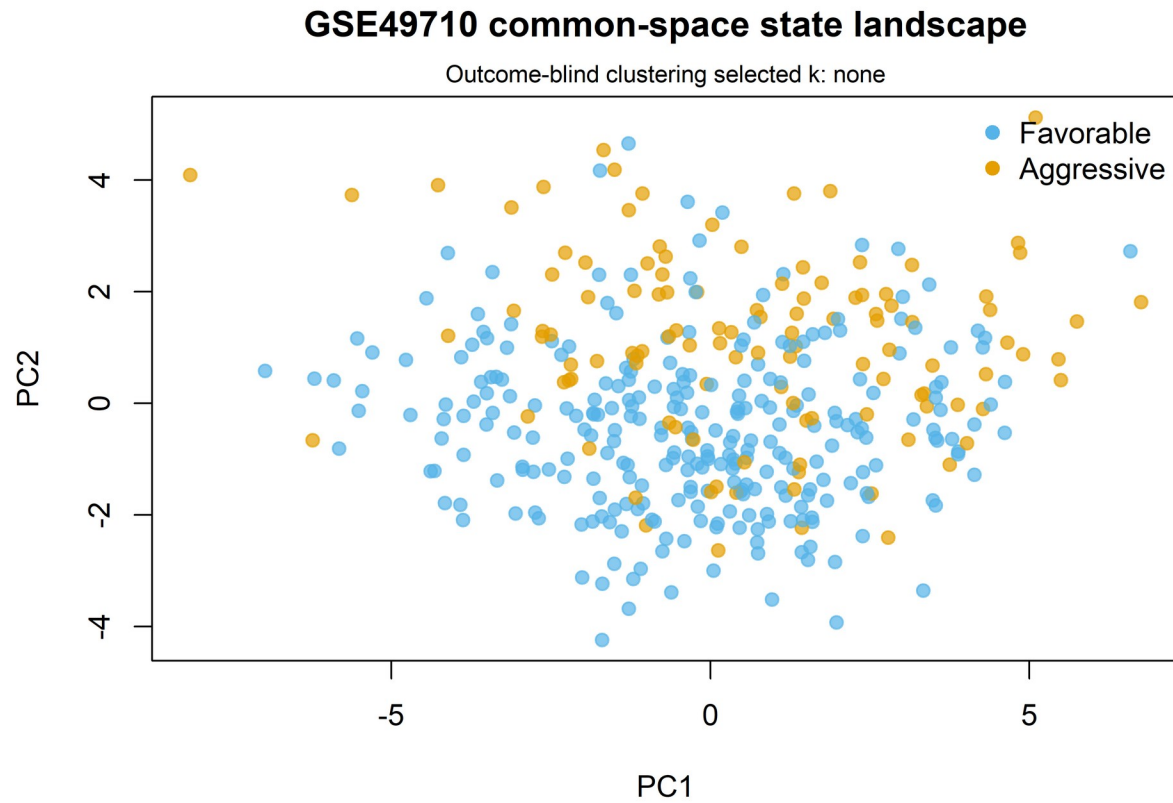
Ruiting Yang, Fei Li, Jing Zhang, Shihao Xu, Bin Wang, Tao Huang, Hairui Hu, Panpan Lv, and Zhuowei Luo

Supplementary Figure S1. Expression quality control in GSE49710.



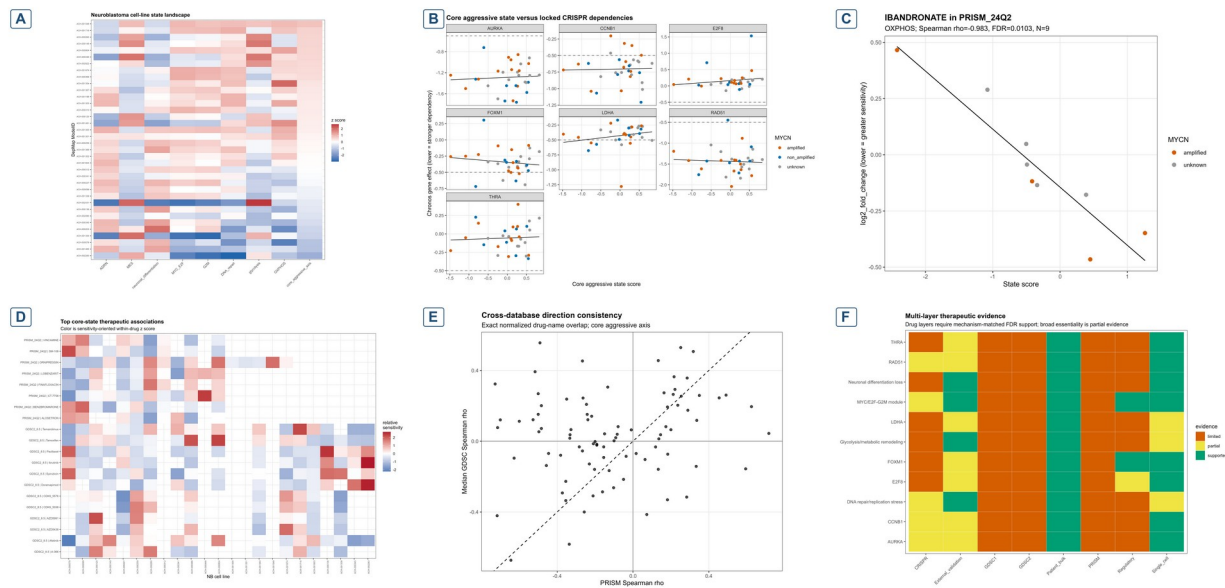
(A) Distribution of sample-level median expression in the favorable and aggressive groups. Boxes show the median and interquartile range; points represent samples. (B) Principal-component analysis of the expression matrix, colored by the prespecified phenotype. These plots were used to evaluate expression distributions and sample-level structure before the locked analyses.

Supplementary Figure S2. Outcome-blind state landscape in the discovery cohort.



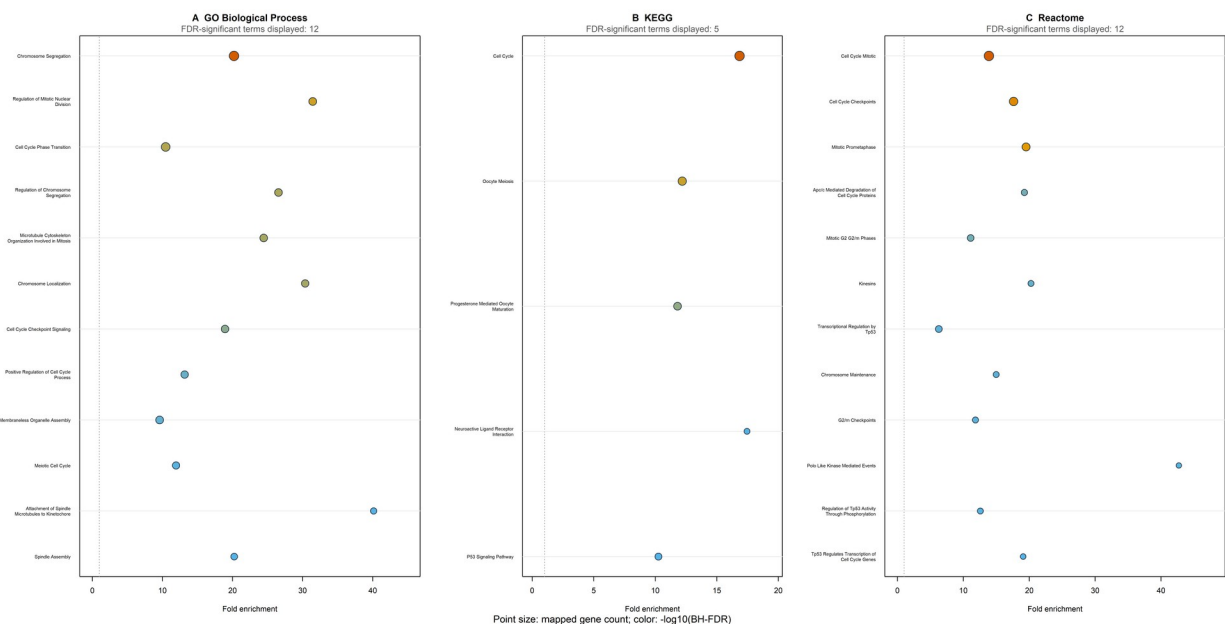
Principal-component representation of the 14 prespecified state and pathway scores in the common-gene space. Points are colored by the clinical phenotype only for display. The outcome-blind clustering diagnostics did not support selection of a stable discrete cluster number; the main analyses therefore retained continuous program scores.

Supplementary Figure S3. Exploratory functional-genomic and pharmacogenomic context.



(A) Neuroblastoma cell-line state landscape. (B) Associations between the core aggressive-state score and locked CRISPR dependencies. (C) Representative within-database drug-state association. (D) Top core-state therapeutic associations. (E) Cross-database direction comparison. (F) Multilayer evidence grading. These analyses are exploratory; broadly essential genes and database-specific associations were not interpreted as validated MYCN-non-amplified therapeutic vulnerabilities.

Supplementary Figure S4. Functional enrichment of differentially expressed genes by GO, KEGG, and Reactome.



(A) Gene Ontology Biological Process, (B) KEGG, and (C) Reactome. Only terms with Benjamini-Hochberg FDR < 0.05 in the frozen analyses are shown. GO and Reactome terms were ranked by FDR and de-redundified using a gene-set Jaccard-overlap threshold of 0.65, retaining 12 terms per database; all five significant KEGG terms are displayed. The x-axis shows fold enrichment, point size indicates the mapped-gene count, and color indicates $-\log_{10}(\text{FDR})$. Complete results and selection rules are provided with the Supplementary Data.

Supplementary Tables and Data Files

Item	Description	Packaged location
Supplementary Table S1	Complete differential-expression results and prespecified quality-control sensitivity branches.	Supplementary_Data/S1/
Supplementary Table S2	Complete Hallmark GSEA results.	Supplementary_Data/S2/
Supplementary Table S3	Complete GO Biological Process over-representation results.	Supplementary_Data/S3/
Supplementary Table S4	Complete KEGG over-representation results.	Supplementary_Data/S4/
Supplementary Table S5	Complete Reactome over-representation results.	Supplementary_Data/S5/
Supplementary Table S6	Program-gene coverage, group comparisons, adjusted clinical models, and score correlations.	Supplementary_Data/S6/
Supplementary Table S7	Cross-cohort effect directions and replication classifications.	Supplementary_Data/S7/
Supplementary Table S8	Prespecified state models and outcome-blind clustering diagnostics.	Supplementary_Data/S8/
Supplementary Table S9	Paired single-cell program, candidate-gene, and state-composition tests.	Supplementary_Data/S9/
Supplementary Table S10	Prespecified and global regulatory analyses, PROGENy results, and candidate-gene localization.	Supplementary_Data/S10/
Supplementary Table S11	Neuroblastoma cell-line availability, MYCN annotation, and CRISPR dependency audit.	Supplementary_Data/S11/
Supplementary Table S12	Drug-state associations, cross-database direction checks, and final therapeutic-evidence grading.	Supplementary_Data/S12/

Large tabular results are supplied as separate CSV/XLSX data files in Supplementary_Data_S1-S12.zip. Each logical Supplementary Table folder contains the complete frozen result files and preserves the original row-level outputs.