# IDENTIFICATION OF DISEASE-SPECIFIC VULNERABILITY STATES AT THE SINGLE-CELL LEVEL

Matthew D'Antuono[1], Madison Sharp[1], Rishika Chowdary[1], Michael E. Ivan[2,3], Ricardo J. Komotar[2,3], Robert K. Suter[1,*], and Nagi G. Ayad[1,*]

[1]*Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC, USA*
[2]*Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Miami, FL, USA*
[3]*Department of Neurological Surgery, University of Miami Miller School of Medicine, Miami, FL, USA*
**Corresponding Authors: Nagi G. Ayad (Lead Contact)*, Robert K. Suter

**Please address all correspondence to:**

Nagi G. Ayad, Ph.D. (Lead Contact)
Georgetown University, Department of Oncology
3970 Reservoir Rd NW
Washington, D.C. 20057
Email: na853@georgetown.edu

Robert K. Suter, Ph.D.
Georgetown University, Department of Oncology
3970 Reservoir Rd NW
Washington, D.C. 20057
Email: rks82@georgetown.edu

**SUMMARY**

Extensive intratumor heterogeneity in glioblastoma (GBM) impedes successful treatment and complicates drug discovery as it is not obvious which cells a tumor is most dependent on. Here, we posit that single-cell-resolution transcriptomic data can be integrated with loss-of-function screens to identify the most critical cells to target within a tumor. We parsed CRISPR screen data from the Dependency Map (DepMap) Consortium and identified a GBM Dependency Signature (GDS) − 168 genes that are essential for GBM cell viability *in vitro*. Through similarity scoring of GDS transcriptomic profiles in single-cell RNA-sequencing (scRNA-seq) data and iterative hierarchical clustering, we identify and report here 3 single-cell vulnerability states (VS) characterized in 49 GBM tumors using both scRNA-seq and spatial transcriptomic data. These VS reflect single-cell gene dependencies and differ significantly in enrichment profiles and spatial distributions. Importantly, the proportion of VS in each GBM tumor is variable, suggesting a means of stratifying patients in clinical trials. Collectively, we have developed a novel computational pipeline to identify unique vulnerability states in GBM and other cancers, which can be used to identify existing or novel drugs for incurable diseases.

**KEYWORDS**: glioblastoma, single-cell RNA-sequencing analysis, spatial transcriptomics analysis, essential genes, transcriptional cell states

## INTRODUCTION

Glioblastoma (GBM) is the most common primary malignant brain tumor in adults. Development of effective treatments for GBM is challenging, and despite extensive research efforts no new therapies have been approved for GBM since the establishment of the current standard-of-care nearly two decades ago.[1] Without advancements in treatment for GBM, the prognosis is poor with a median overall survival after diagnosis of only 15 months and a 5-year survival rate less than 5%.[1–3] Therefore, new GBM treatments are desperately needed.

One of the driving forces behind treatment and drug development challenges is the extensive inter- and intra-tumor heterogeneity innate to GBM. Intratumor heterogeneity (ITH) in particular presents a large hurdle. Each patient tumor contains cells with high degrees of genomic, transcriptomic, proteomic, and metabolic variability, and this can lead to innate or rapid development of drug resistance and subsequent treatment failure.[4–9] To date, much of the heterogeneity observed in GBM and across all tumor lineages is shown to be the result of recurrent transcriptional modules expressed during cellular differentiation or in response to stressful stimuli such as hypoxia.[7,10–16] In GBM, tumor cells exist along a spectrum of transcriptional cell states that mimic canonical neurodevelopmental cell types.[7] These states are designated as astrocyte-like (AC), mesenchymal-like (MES), neural progenitor-like (NPC), and oligodendrocyte-like (OPC).[7] Multiple studies have demonstrated that these states are highly plastic and the relative proportions of each in a tumor will change under treatment pressure, contributing to drug resistance and tumor recurrence.[7,13,17,18]

As GBM heterogeneity continues to impede drug discovery, a novel method is needed to identify the distinct transcriptional cell states to target to achieve therapeutic efficacy. This framework involves prioritization of those states identified as most essential to a tumor with simultaneous targeting of the remaining cancer cells. Toward that end, single-cell transcriptomic data can be integrated with loss-of-function screens to identify the most essential cells within a GBM tumor. Genome-wide loss-of-function screens such as those conducted by the Dependency Map (DepMap) Consortium offer insight into genetic vulnerabilities of cancer cell lines and make for promising resources to identify candidate genes for targeted therapy development. DepMap performed systematic CRISPR knockout screens in more than 1,000 cell lines across nearly every type of cancer lineage, creating an unparalleled collection of genetic profiles that quantify the essentiality of nearly every gene to each cell lines' viability.[19–23] Importantly, DepMap data has already been successfully leveraged to discover genetic vulnerabilities in cancer cells with translation to *in vivo* models and patient tumors.[24–27]

We sought to leverage CRISPR gene effect scores (ES) from DepMap to identify exploitable "vulnerability states" in GBM – single-cell states with shared, therapeutically-relevant transcriptional profiles that would be useful toward informing drug development strategies. We identified a list of GBM lineage-specific essential genes and used this gene set to inform transcriptomic single-cell state assignments via similarity scoring and iterative

hierarchical clustering. We uncovered three (3) novel vulnerability states within and across GBM tumors from different datasets that express GBM dependency genes to varying degrees. We found these states to have distinct enrichment profiles in both single-cell RNA-seq and spatial transcriptomic data, and they are each spatially distinct among patient GBM tumor sections. Critically, these states are distinct from existing GBM transcriptional classifiers and are informed by CRISPR screen-derived, lineage-specific essential genes, paving the way for drug discovery and development efforts using these single-cell states.

## RESULTS

### Analysis of DepMap genome-wide knockout screen uncovers 168-gene glioblastoma dependency signature (GDS).

We acquired DepMap's Project Achilles open-access genome-wide CRISPR knockout screen post-Chronos effect score (ES) data from their online portal. Lower ES for a gene indicates a cell line is more dependent on that gene for survival. In general, an ES of -1 is commensurate with the median of all common essential genes, while an ES of 0 implies that a gene is not essential to a cell line. The DepMap ES dataset (23Q2 release) contains ES for 17,453 genes across 1,078 cell lines (Figure 1A). We interrogated these data to identify genes whose knockdown negatively impacts viability in GBM cell lines specifically. Our approach strived to select for genes uniquely essential to GBM over other cell line lineages with the aim of identifying genes that, if targeted for knockdown/out, would have minimal off-target/tumor effects and thereby minimize potential toxicities. Moreover, through these methods we sought to minimize identification of pan-cancer essential genes (e.g. CDK1, DNMT1, HDAC3) whose inhibitors have suffered from clinical trial failures due to on- and off-target toxicities.

With these goals in mind, we compared ES for every gene in GBM cell lines to the corresponding ES in all other cell line lineages using a generalized linear model (GLM). Of 1,078 cell lines in the dataset, 49 are GBM-lineage specific (Supplementary Table 1). Genes with significantly lower ES in GBM lines were defined as those with negative effect sizes and a p-value $< 0.05$ by GLM (Figure 1B). Using these criteria, we identified 377 genes whose knockout more significantly impacted GBM cell line viability than other cell lineages (Supplementary Table 2). Importantly, a strict ES cutoff was then set to identify which of the 377 genes actually kill GBM cells *in vitro* – removing genes with mean ES $> 0$, for example. We imposed a cutoff of mean ES $< -0.5$ in GBM cell lines as this quantitative value has been used in previous studies and purported as representing depletion in most cell lines.[20,26] We thereby identified 168 genes uniquely essential to GBM cell lines, a list we term a GBM Dependency Signature (GDS) (Figure 1C, Supplementary Table 3). Some of the top GDS genes by p-value include PRKAR1A, LMNA, JUN, VRK1, MTOR, RAC1, CDC42, and CDK2.

To determine GDS gene expression in patient GBM tumors, we used a single-cell RNA-seq (scRNA-seq) dataset first published by Neftel et al. in 2019. This dataset contains 5,689 single cells (4,865 neoplastic) across 20 adult GBM tumors. Neoplastic cells per tumor ranged from 121 to 435 cells. We applied a Module Score (MS) calculation provided by the scRNA-seq analysis *R* package, *Seurat*, to simultaneously score the entire GDS as a gene set in each single-cell in the Neftel dataset.[28] MSs provide a scaled method of evaluating single-cell level expression of groups of genes.[28,29] We found that neoplastic GBM cells more highly express GDS genes than

non-neoplastic tumor microenvironment (TME) cells in the Neftel dataset (Student's t test, $p < 0.0001$, Figure 1D; non-neoplastic TME cells in this dataset include oligodendrocytes, T-cells, and macrophages).

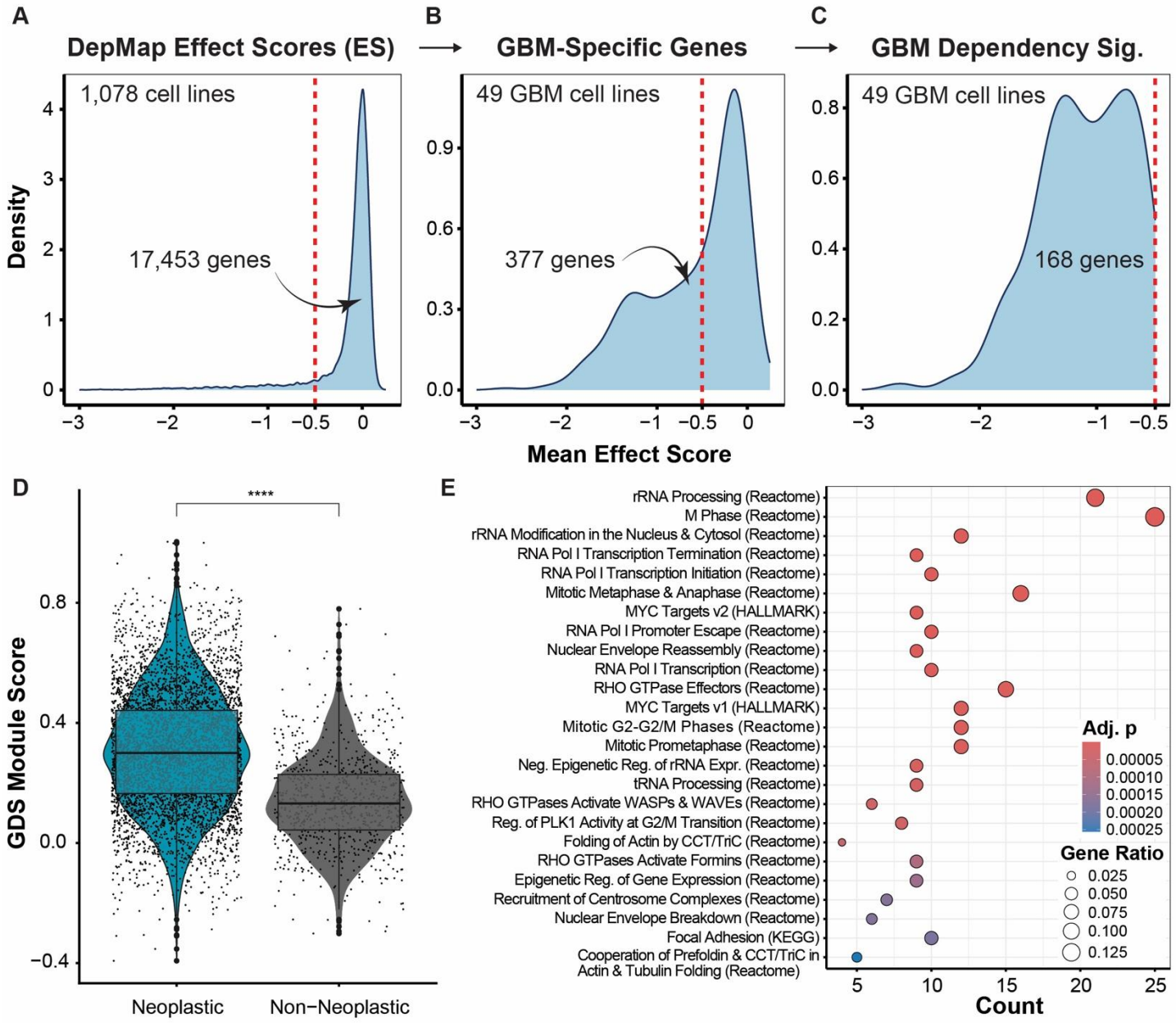



**Figure 1: Analysis of DepMap genome-wide knockout screen uncovers 168-gene glioblastoma dependency signature.**

(**A-C**) Density plots depicting the selection pipeline for the Glioblastoma (GBM) Dependency Signature (GDS). The red dashed line in each represents a mean effect score of -0.5. (**A**) shows the mean effect scores (ES) of all genes (n = 17,453) in all cell lines (n = 1,078) present in the DepMap dataset. After applying a linear model comparing ES in GBM-lineage cell lines (n = 49) to all other cell lines, (**B**) depicts the mean ES of 377 GBM-

specific genes. (**C**) then represents the mean ES of the 168-gene GDS, identified by filtering the 377 genes from (**B**) for those with mean ES < -0.5. (**D**) Violin and overlapping box plot depicting the module scores (MS) of the GDS in each single-cell (black dots) in the Neftel et al. scRNA-seq dataset (5,689 cells, 20 adult GBM tumors), grouped by neoplastic status. Neoplastic cells (n = 4,865) are compared to non-neoplastic cells (n = 824, consisting of T-cells, macrophages, and oligodendrocytes) by a Student's t test (**** = p < 0.0001). (**E**) Dot plot representing Molecular Signatures Database (MSigDB) over-representation analysis performed on the 168-gene GDS (using collections C2 and H). The top 25 over-represented terms here have Benjamini-Hochberg (BH)-adjusted p-value < 0.05 by Fisher's exact test. Terms (pathways) are sorted by descending adjusted p-value (bar color). The gene ratio is the number of GDS genes that are enriched in a pathway over the total number of genes that define it.

To validate this finding, we performed the same MS calculations and comparisons on two additional scRNA-seq datasets of patient-derived GBM tumors – one published by Johnson et al. in 2020 (28,130 cells, 5 adult IDH-wt GBM tumors)[30] and one that we generated ourselves, in-house (42,839 cells, 6 adult GBM tumors). In each independent dataset, we similarly found that GDS genes were significantly more highly expressed in neoplastic compared to non-neoplastic single cells (Supplementary Figures 3-4).

We then analyzed the pathways best represented in the 168-gene GDS. We performed pathway enrichment analysis using collections H, C2, and C3 of the Molecular Signatures Database (MSigDB) and found the GDS to be significantly enriched for pathways related to transcription and translation, adhesion, motility, and invasion (Figure 1E).[31–39] Of the top 25 enriched terms among these collections, eight (8) were related to transcription and the translation machinery (e.g. "rRNA Processing (Reactome)", "RNA Pol I Transcription Initiation [Termination] (Reactome)"). Additionally, 7 of 25 are related to mitosis and cell division (e.g. "Mitotic Metaphase & Anaphase (Reactome)", "Nuclear Envelope Reassembly (Reactome)", "Recruitment of Centrosome Complexes (Reactome)"), and 5 out of 25 are related to actin filaments and migration/adhesion (e.g. "Focal Adhesion (KEGG)", "Folding of Actin by CCT/TriC (Reactome)", and "RHO GTPases Activate WASPs & WAVEs (Reactome)").

Therefore, the 168-gene GDS is not only specific to GBM cell lines, but is more highly expressed in patient GBM tumor cells compared to non-neoplastic TME cells in those same tumors. Moreover, these GBM dependency genes are largely related to cellular mitosis, protein translation, and actin polymerization – key pathways in tumor development and progression.

**Glioblastoma dependency signature is required for identification of vulnerability states.**

We next sought to use the GDS to group single-cells by their patterns of expression of these genes. We hypothesized that cells clustered in such a way would share actionable vulnerabilities. For this analysis, we chose to again use the Neftel scRNA-seq dataset as it contains robust transcriptomic data on the largest number of patient tumors. The pipeline used for single-cell state identification is outlined in Figure 2A. For each of the 20 adult GBM tumors, we calculated the Jaccard similarity index (JS) for the binarized scaled expression of each GDS gene in every cell. This initial step was performed on each tumor separately to minimize tumor-level effects and account for intertumor heterogeneity. Following Jaccard scoring, we performed hierarchical clustering on the resulting JS matrix for every cell in the tumor for all k from 2 through 30. A representative heatmap of this first step from tumor MGH106 is shown in Figure 2B. The remaining 19 JS heatmaps for each tumor are shown in Supplementary Figure 5.

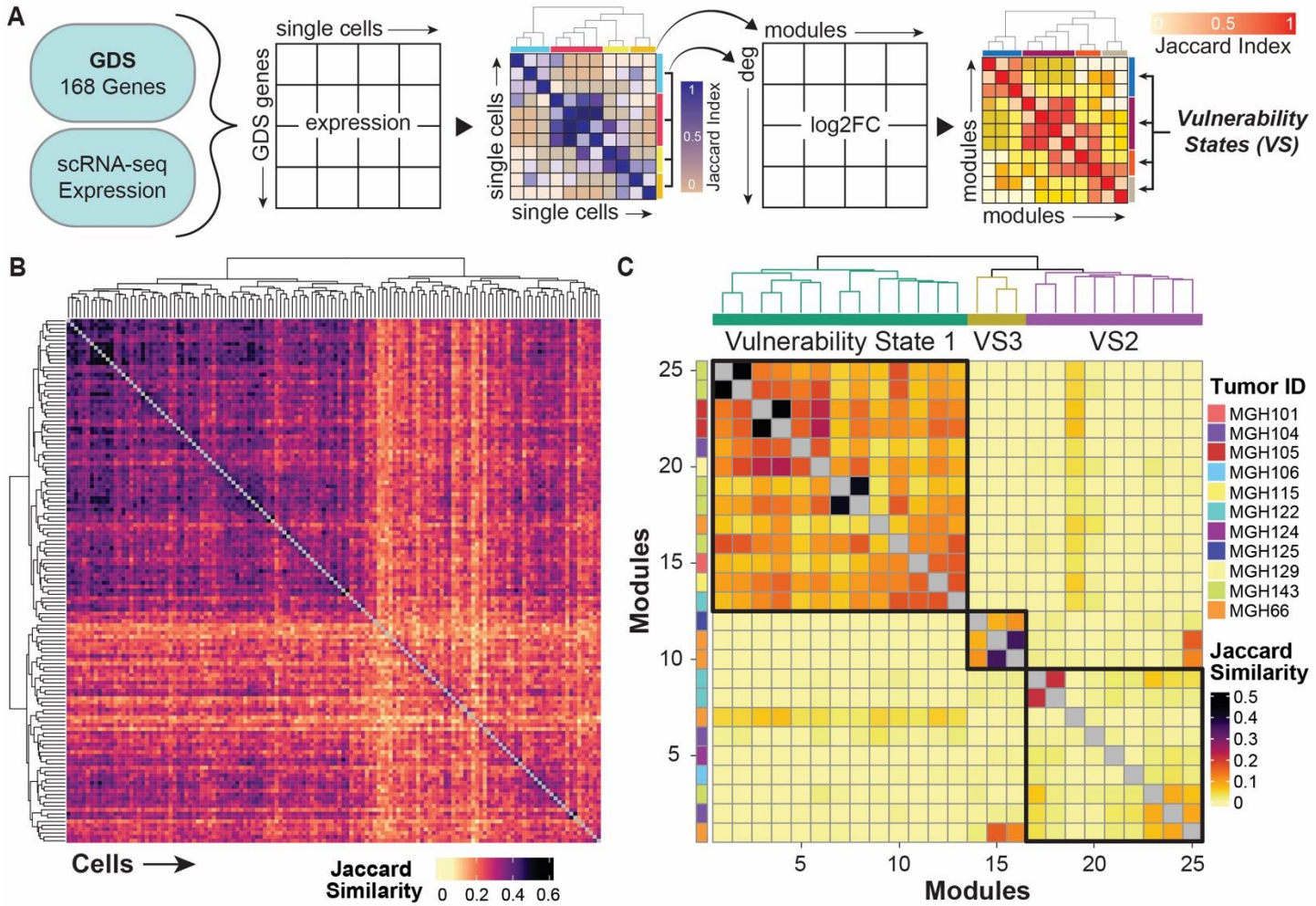



**Figure 2: Glioblastoma dependency signature is required for identification of vulnerability states.** (**A**) Schematic outlining the process of vulnerability state (VS) identification from single-cell expression data of glioblastoma dependency signature (GDS) genes (DEG = differentially expressed genes). (**B**) Representative heatmap of expression-derived Jaccard Index scores (Jaccard Similarity) for every neoplastic single-cell in one

tumor (MGH106) out of 20 total adult GBM tumors in the Neftel et al. scRNA-seq dataset. Clustering of Jaccard scores was performed by Ward D2 hierarchical clustering at all k between 2 and 30 for this and all other tumors (Supplementary Figure 5). Each row and column of the heatmap represents a single neoplastic cell in MGH106. (**C**) Heatmap representing Jaccard scores between pass-filter GDS-derived clusters (modules) across all 20 tumors. Pass-filter clusters included 25 modules from 11 distinct tumors. Each row and column of the heatmap represents one GDS-derived module from one tumor. The tumor identifier (Tumor ID) is used as row annotations to highlight which tumor each module is derived from. Clustering of Jaccard similarities between modules was performed by Ward D2 hierarchical clustering, identifying 3 final signatures (vulnerability states, VS).

We identified 274 clusters of single-cells across all 20 tumors with sufficient numbers of cells for downstream analysis. We then sought to collapse clusters based on shared transcriptomic profiles since there may be overlap across k-values. To do so, we performed differential expression analysis between each cluster within a given tumor, using the top differentially expressed genes (DEG) with $p < 0.05$ and log2FC > 0 as a transcriptomic signature (module) for each cluster (FindAllMarkers from Seurat, test used was MAST).[28,40] At minimum, a given cluster had to have 50 DEG that met these criteria, in addition to at least 1 gene with $p < 0.001$ to be used as a module. We then binarized modules based on whether a DEG was present or absent to then perform JS calculation to score each cluster for similarity to each other. We then removed any clusters with JS > 0.75. This resulted in 25 unique modules across 11 (out of 20) tumors with unique GDS-informed transcriptomic signatures. Following this exclusion, we performed hierarchical clustering to collapse the remaining 25 modules into meta-clusters, which we termed vulnerability states (VS) (Figure 2C). We found 3 VS by this method, with optimal k-value determined by both elbow plot and gap statistic (Supplementary Figure 5). Importantly, rather than clustering by tumor identity, inter-tumor clusters group according to transcriptomic profiles.

We next sought to identify marker genes for these cell states to allow for further characterization as well as cell state assignment and validation in additional data sets. The top 100 genes by log2FC across all modules within a given VS were selected as VS signature genes. Figure 3A shows the pseudo-bulk normalized scaled expression of the 292 unique identified signature genes across 24 of 25 modules from which they were derived. Differential enrichment analysis was then performed on these signature genes using the Molecular Signatures Database (MSigDB) collections H, C2, and C3 (Figure 3B).[31–39] The results of this analysis accurately reflected the signature genes for each VS, with the VS1 signature enriching strongly for neuronal and developmental transcription factor (TF) gene sets, VS2 for transcription and translational machinery, and VS3 for cellular proliferation. Key VS signature genes that corroborate these enrichment results are highlighted in Figure 3A.

To identify DEG between each VS and further characterize these novel cell states, we then performed differential expression analysis. Differential expression profiles between each VS were distinct, with more GDS genes identified as DEG in VS2 than either VS1 or VS3 others (Figure 3C). In total, 32 of 168 GDS genes were identified

among the top 1,500 significant DEG for each VS ($p < 0.05$, log2FC > 0, genes annotated in Figure 3C). VS1 differentially expressed 9 GDS genes: DCTN2, WDHD1, CDK2, VRK1, NOP2, CCT6A, RRN3, CEBPZ, and CDC42. VS2 differentially expressed 21 GDS genes including LMNA and RCC1. VS3 differentially expressed 2 GDS genes: ITGAV and TRMT5. Notably, VS2 had the largest number of GDS genes represented among its significant DEG.

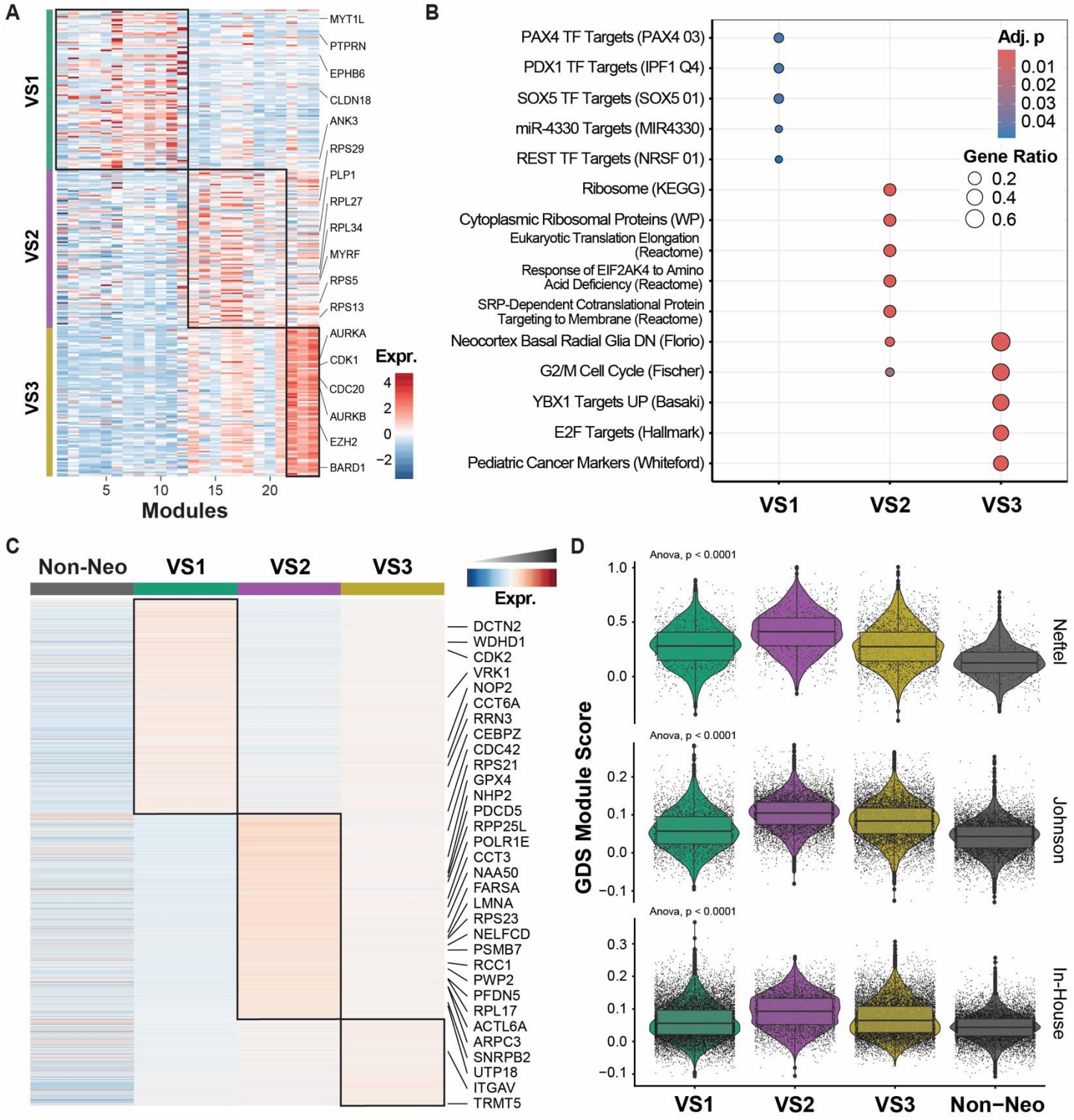


**Figure 3: Glioblastoma vulnerability states are distinct.**
(**A**) Heatmap representing average scaled RNA expression of the n = 100 marker genes identified for each VS. Genes are ranked by descending log2FC for each VS. Labeled genes are those of interest with particular relationships to over-enriched terms identified in (B). (**B**) Dot plot displaying MSigDB differential enrichment analysis results (collections H, C2, and C3) performed on the 100 marker genes per VS as highlighted in (A) (n = 292 unique genes). The database each term is derived from within each collection is written in parentheses following the term name. Color indicates the BH-adjusted p-value and circle size represents the gene ratio. (**C**) The top 1,500 differentially expressed genes by log2 fold change (log2FC) between VS are visualized using a heatmap of scaled, normalized pseudo-bulk gene expression. Total DEG identified for VS3 was 634. Labeled genes are GDS genes (32 out of 168) that are also significantly differentially expressed. DEG were identified using Seurat and the MAST statistical test with a Bonferroni-adjusted p-value < 0.05 and log2FC > 0.0. (**D**) Violin plots displaying MS of the 168-gene GDS in single-cells in each of the datasets used in the study (Neftel, Johnson, and Suter, respectively), with cells grouped by their assigned VS. Each black dot represents one single-cell in the dataset. Statistical testing was performed by ANOVA with Tukey's post-hoc test for individual comparisons and presented as a Bonferroni-adjusted p-value.

Moreover, Gene Ontology (MSigDB collections H and C5) differential enrichment analysis performed on the top 100 significant DEG ($p < 0.05$) with log2FC > 0.5 (Supplementary Figure 5), revealed that each VS had distinctly enriched pathways that mirrored VS signature gene enrichment (Figure 3B). VS2 was more highly enriched for pathways related to neuron myelination. By contrast, VS1 DEGs are enriched for apoptotic and trans-synaptic signaling regulation in addition to transcriptional activator activity, while VS3 DEGs are related to exocytotic pathways (e.g. synaptic exocytosis).

We are therefore able to use GDS genes to inform identification of GBM VS. Moreover, GDS genes are both present and reflected in both the VS signature genes and the significant DEG between VS, especially when comparing our analyses demonstrated in Figure 1D with those in Figure 3B-C and Supplementary Figure 7B. Taken together, VS have distinctive signatures and transcriptional profiles. VS1 largely reflects a stem-like transcriptional signature, VS2 is related to translational machinery in oligodendrocytes and largely recapitulates the GDS (Figure 3D), and VS3 is related to the cell cycle and exocytosis (e.g. platelet alpha granule GOCC).

**Vulnerability states are present across multiple scRNA-seq datasets.**

We next sought to validate the presence of VS in additional GBM tumors and to further study their single-cell transcriptomic profiles. To that end we downloaded, processed, and analyzed a GBM scRNA-seq dataset from Johnson et al. to use in conjunction with our own in-house dataset. The Johnson dataset contains 28,130 single

cells (19,530 neoplastic) from 5 adult GBM tumors (Supplementary Figure 3).[30] Our in-house dataset contains 42,839 single cells (32,159 neoplastic) from 6 adult GBM tumors (Supplementary Figure 4).

We used VS signature genes to score single-cells in each dataset for enrichment of those genes using *singscore*, an independent-sample gene-set scoring method (Supplementary Table 4).[41,42] Singscore enrichment scores were then used for cell state assignment, with cells being assigned to whichever VS scored most highly in that cell (Supplementary Figure 6). VS were identified in both datasets, and the resulting VS proportions by tumor in both the Johnson and our in-house dataset largely recapitulated the proportions observed in the Neftel dataset. Individual tumors in all three datasets differed slightly in the relative proportions of their VS content, with some more highly or more lowly containing some VS than others. As a whole, VS3 makes up the majority of the neoplastic cell populations, followed by VS1, and then by VS2 (Figure 4A).

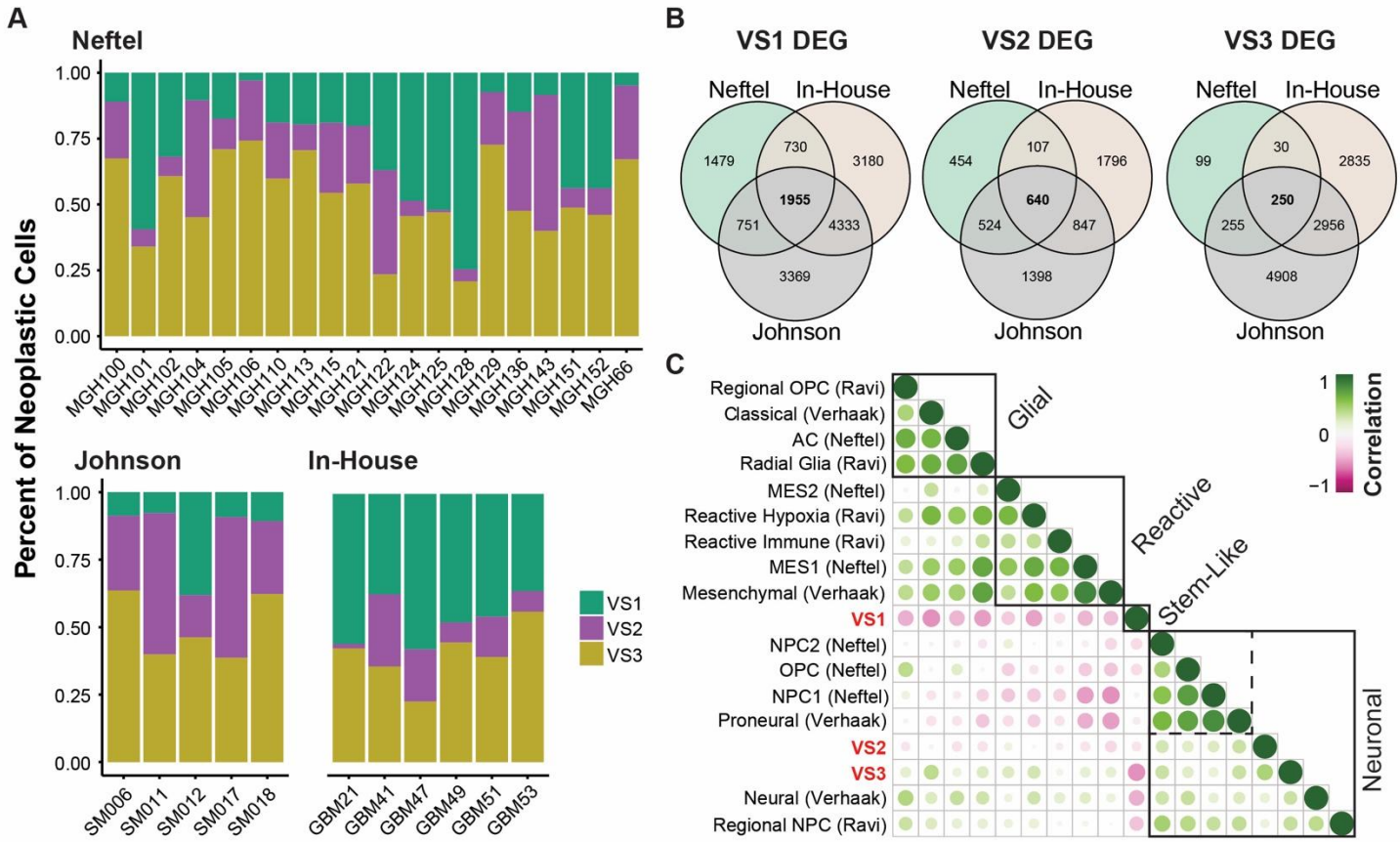



**Figure 4: Vulnerability states are present across multiple scRNA-seq datasets.**
(**A**) Stacked bar plots displaying the proportions of neoplastic cells that are categorized as each VS for each tumor in the Neftel et al., Johnson et al., and our in-house scRNA-seq datasets, respectively. Color corresponds to VS. VS assignments in external datasets were made using singscore enrichment scoring. (**B**) Venn diagrams representing overlapping DEG among each of the 3 scRNA-seq represented in A. For each dataset, DEG were

identified between VS using Seurat's FindMarkers function with log2FC cutoff of 0 and significant p-value < 0.05 as determined by MAST statistical testing. The number of DEG in each category are represented here with the percentage of total DEG per VS among all 3 datasets shown in parentheses. (**C**) Spearman correlation analysis of singscore enrichment scores between existing GBM transcriptional signatures (from Verhaak et al., 2010; Neftel et al., 2019; and Ravi et al., 2022) and our own vulnerability state signatures (red). Clustering of all signatures produces 4 major groups (with 1 sub-group): Glial-like, Reactive-like, Stem-like (VS1), and Neuronal-like (including VS2 and VS3 within their own subcluster). Enrichment scores calculated from our In-House scRNA-seq dataset.

Given the consistency of VS scoring and assignment in each dataset, we wanted to validate the similarities between assigned VS across datasets, independent of the 100-gene VS signatures. To do so, we performed differential expression analysis between VS in the Johnson and in-house datasets, respectively, with the aim of comparing resulting DEG with those DEG identified in the Neftel dataset. DEG were calculated using Seurat's FindAllMarkers function and determined by MAST-adjusted p-value < 0.05 and log2FC > 0. For VS1 across all three scRNA-seq datasets, a total of 1,955 (12.4% of all VS1 DEG) of DEG overlapped. Similarly, for VS2 a total of 640 (11.1%) of DEG overlapped. For VS3, 250 genes overlapped though this only represented 2.21% of VS3 DEG across all three datasets. However, the VS3 DEG identified in the Johnson and in-house dataset had high concordance, with 3,206 (28.3%) overlapping genes. These findings emphasize that the VS-assigned cells identified in the Johnson and our in-house datasets are similar to each other and to VS-assigned cells in the Neftel dataset, allowing us to conclude that VS are not a phenomenon isolated to the dataset from which they were derived.

Once confident in our validation of VS assignments, we finally sought to return to the GDS; since the premise of our investigation was to identify cell states with shared GDS expression patterns, we wanted to identify whether VS express GDS genes in all datasets. We used the GDS MS calculated previously for each single cell in all three datasets and grouped cells by neoplastic identity, comparing GDS expression between assigned VS. As highlighted in Figure 3C and quantified in Figure 3D, VS differently express GDS genes, with VS2 more highly expressing GDS genes compared to both VS1 and VS3 in all 3 datasets (ANOVA $p < 0.0001$ for all three datasets; VS2 MS > VS1 and VS3 MS by Tukey's post-hoc test, $p < 0.0001$ for all three datasets; Figure 3D).

Transcriptional signatures representing bulk, single-cell, and spatial transcriptomic classification systems for GBM from Verhaak et al., Neftel et al, and Ravi et al., respectively, were acquired with the goal of comparing these to our VS signatures.[7,14,43] Each signature was scored for enrichment in every neoplastic single-cell in our in-house dataset using singscore. Enrichment scores were correlated using Spearman correlation to identify similarities between signatures (Figure 4C). Our results largely recapitulated the results observed in Ravi et al.[14] However, the localization and placement of our VS signatures in this plot are critical. First, we observe similar

grouping of glial and reactive (immune response, hypoxia) signatures. While Ravi et al. observed clustering of one neuronal group, we see a larger neuronal cluster with a sub-cluster (dotted lines) owing to our identification and use of VS2 and VS3 signatures. While neuronal-like signatures by this analysis, VS2 and VS3 are unique enough to be distinct from Neftel NPC and OPC states and Verhaak Proneural tumor type, yet similar to the regional neural progenitor-like state observed in Ravi et al. spatial transcriptomic analysis. Critically, VS1 was not similar to any of the existing signatures, as it clustered independently with low Spearman correlations to all other signatures. Importantly, the overlapping DEG from Figure 4B enriched for pathways that corroborate these data (Supplementary Figure 6). Based on VS signature gene enrichment analysis (Figures 3 A-B), VS DEG enrichment analysis (Figure 3C and Supplementary Figures 5-6), and our findings here (Figure 4C), VS1 most likely represents a novel neuronal stem-like signature in GBM.

These findings emphasize that VS are found across GBM tumors, but precise proportions of these cell states differ between patients. Despite this proportional variance, their concordance across tumors is high. Moreover, VS2 shows the highest degree of similarity to the GDS signature. By comparing transcriptional classification signatures from previously-published literature, we show that VS1 is a unique stem-like classification signature for cell states in GBM. We also find that VS2 and VS3 represent neuronal-like signatures related to transcription/translation and exocytosis and the cell cycle, respectively.

**Glioblastoma vulnerability states are spatially distinct.**

Given our findings using the transcriptomic signatures from the literature, particularly those from Ravi et al., and a growing body of research emphasizing the importance of cellular spatial distribution and heterogeneity in GBM,[11,14,44,45] we sought to investigate how our identified VS were distributed spatially in patient tumors. Corroborating our data in spatial samples is essential given continuously emerging evidence that spatial cellular organization plays a role in driving clonal evolution, tumor progression, and treatment resistance.

We obtained spatial transcriptomic (stRNA-seq) data from Ravi et al. (Dryad, https://doi.org/10.5061/dryad.h70rxwdmj) and processed and analyzed raw counts according to our in-house protocols. These data contained 18 stRNA-seq GBM tumor sections. Our in-house scRNA-seq dataset was then used to deconvolute the bulk RNA-seq spots of the 10X Visium stRNA-seq data using SCDC to glean insight into the spatial distribution of vulnerability states in these tumor slices.[46] Our deconvolution results highlight that spatial distribution of VS varies from tumor-to-tumor, with most tumors having lower levels of VS1 present while being either VS2-high or VS3-high, relatively. These findings are quantified in Figure 5A with deconvoluted proportions of non-neoplastic cells and VS presented across all 18 tumor sections.

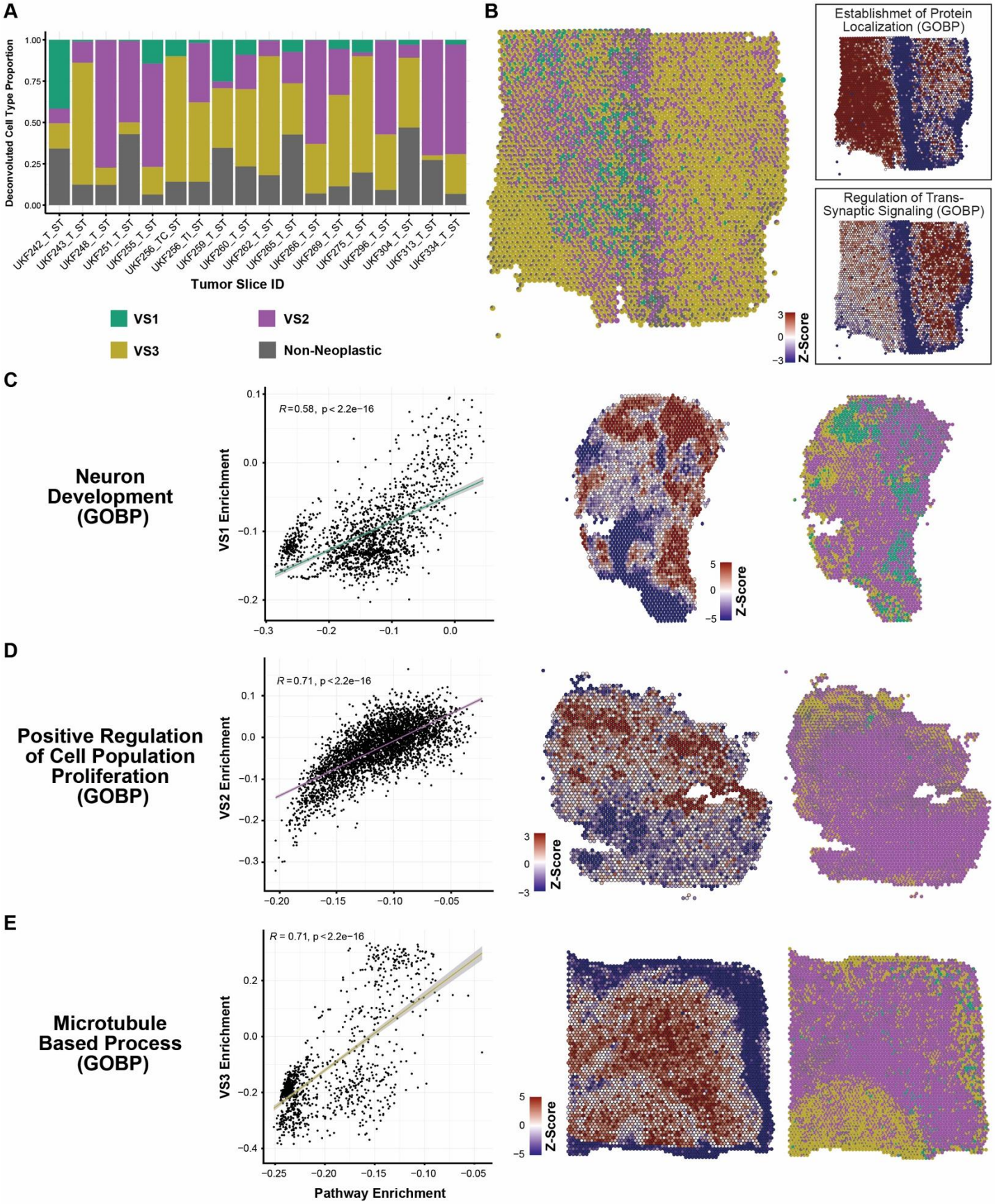

A
Deconvoluted Cell Type Proportion
1.00
0.75
0.50
0.25
0.00
UKF242_T_ST
UKF243_T_ST
UKF248_T_ST
UKF251_T_ST
UKF255_T_ST
UKF256_TC_ST
UKF256_TI_ST
UKF259_T_ST
UKF260_T_ST
UKF262_T_ST
UKF265_T_ST
UKF266_T_ST
UKF269_T_ST
UKF275_T_ST
UKF296_T_ST
UKF304_T_ST
UKF313_T_ST
UKF334_T_ST
Tumor Slice ID
VS1
VS2
VS3
Non-Neoplastic
B
Establishmet of Protein Localization (GOBP)
Regulation of Trans-Synaptic Signaling (GOBP)
Z-Score
3
0
−3
C
Neuron Development (GOBP)
R=0.58, p<2.2e−16
VS1 Enrichment
0.1
0.0
−0.1
−0.2
−0.3
−0.2
−0.1
0.0
Z-Score
5
0
−5
D
Positive Regulation of Cell Population Proliferation (GOBP)
R=0.71, p<2.2e−16
VS2 Enrichment
0.1
0.0
−0.1
−0.2
−0.3
−0.20
−0.15
−0.10
−0.05
Z-Score
3
0
−3
E
Microtubule Based Process (GOBP)
R = 0.71, p<2.2e−16
VS3 Enrichment
0.2
0.0
−0.2
−0.4
−0.25
−0.20
−0.15
−0.10
−0.05
Pathway Enrichment
Z-Score
5
0
−5

**Figure 5: Glioblastoma vulnerability states are spatially exclusive.**
(**A**) 10X Visium stRNA-seq data were acquired from Ravi et al. and deconvoluted using SCDC to identify single-cell type proportions present in each section, with our in-house scRNA-seq dataset as a reference. This stacked bar plot shows the deconvoluted cell type proportions present within each of 18 GBM tumor sections. Cell types included as "Non-Neoplastic" designation include: astrocytes, endothelial cells, fibroblasts, myeloid cells, oligodendrocytes, and T-cells. (**B**) Shows a spatial scatter plot of UKF269 where each spot is represented as a pie chart showing the proportion of cell types identified by deconvolution at that precise location. To the right of the spatial scatter-pie image are two representative images of gene set co-regulation analysis (GESECA), with respective enrichment scores at each spot represented as z-scores. (**C-E**) Display, in succession, select top-upregulated pathways for each of VS1 (C), VS2 (D), and VS3 (E), visualized by spot-level enrichment score correlations, GESECA z-score spatial scatter plots, and spatial scatter-pie images, from left to right. (C) specifically shows VS1 enrichment for Neuron Development, (D) shows VS2 enrichment for Regulation of Cell Population Proliferation, and (E) shows VS3 enrichment for Microtubule Based Processes. All enrichments were performed using fgsea using Gene Ontology Biological Process (GOBP) enrichment set, with correlation analysis performed using Pearson correlation with a level of significance set to < 0.05.

In addition to tumor sections varying by proportion of VS present, spatial distributions of Gene Ontology Biological Processes (GOBP) enriched pathways were unique and overlapped with VS distributions. We performed gene set co-regulation analysis (GESECA) implemented through the fgsea R package independently on each tumor section using the GOBP database.[47] The goal was to identify spatial patterns of gene set expression. For example, we found that one section (UKF269) showed a stark contrast between cells involved in protein localization (Establishment of Protein Localization) compared to those involved in trans-synaptic signaling (Regulation of Trans-Synaptic Signaling) (Figure 5B). Such contrasts are also interesting when examining relative VS proportions in each of these respectively enriched regions – VS1 and VS2 are more present in tissue spaces involved in protein localization while VS3 appears more involved in trans-synaptic signaling unless specifically sharing space with non-neoplastic cell types.

We sought to correlate spot-level VS enrichment to pathway enrichment scores in a spatial landscape. We took the top regulated pathways in each section from our GESECA analysis, scored every spot for enrichment of those pathways, and concurrently scored each spot for VS enrichment using singscore.[41,42] We then ran correlation analyses using Pearson's R to compare VS enrichment with GOBP pathway enrichment directly, across all spots among all sections. Among these comparisons, those pathways with significant correlations to VS enrichment were strikingly similar to our previous scRNA-seq-based enrichment results. VS1 strongly correlated with GOBP terms related to cellular and neuronal development and differentiation ("Neuron Development", $R = 0.58$, $p < 0.0001$, Figure 5C; "Neuron Differentiation", $R = 0.54$, $p < 0.0001$, Supplementary Figure 7). Additional enrichments were for synaptic transmission and synaptic plasticity ("Inhibitory Synapse

Assembly", R = 0.83, p < 0.0001, Supplementary Figure 7; Supplementary File 3). VS2 strongly correlated with terms related to cellular proliferation, response to cellular stress, and regulation of cell death, reinforcing our findings that VS2 most highly expressed GDS genes ("Positive Regulation of Cell Population Proliferation", R = 0.71, p < 0.0001, Figure 5D; "Cellular Response to Stress", R = 0.73, p < 0.0001; "Regulation of Cell Death", R = 0.8, p < 0.0001, Supplementary Figure 7). Interestingly, VS2 enrichment also correlated with immune and inflammatory response-related pathways (Supplementary File 3). Lastly, VS3 enrichment correlated with microtubule-related pathways ("Microtubule Based Process", R = 0.71, p < 0.0001, Figure 5E; "Transport Along Microtubule", R = 0.69, p < 0.0001, Supplementary Figure 7) and pathways related to oxidative phosphorylation including cellular aerobic respiration and ATP production ("Aerobic Respiration", R = 0.79, p < 0.0001, Supplementary Figure 7; Supplementary File 3). These results strongly corroborate those posited here from analysis of our scRNA-seq data – that VS distinctly express transcriptional programs important to GBM – while revealing their spatial correlations with critical tumorigenic pathways.

## DISCUSSION

Transcriptional cell states and state transitions have been described as viable therapeutic targets due to their roles in drug resistance, tumor progression, immune evasion, and recurrence.[7,8,11,13,14,30,48–52] However, it remains unclear which cells states to target and prioritize. Here, we present a novel framework for cell state identification that is informed by intrinsic GBM lineage-specific vulnerabilities. Analysis of genome-wide cell line CRISPR screen data from the DepMap Consortium yielded a 168-gene dependency signature for GBM that we leveraged together with scRNA-seq data to identify shared GDS-related transcriptional modules in cells across patient tumors using similarity scoring and iterative hierarchical clustering. We demonstrate that the GDS is GBM-specific, relevant to patient tumors in both scRRNA-seq and stRNA-seq datasets, and has some degree of shared transcriptional regulation across tumors. The GDS highlights that large-scale *in vitro* screens provide critical insights into tumor vulnerabilities. Moreover, we show that the GDS-derived VS are distinct from existing GBM tumor and cell type classification signatures in both scRNA-seq and stRNA-seq data and share spatial colocalization and enrichment with critical GBM tumorigenic pathways. Taken together, our results uncover three single-cell states derived from critical GBM vulnerabilities, each of which provide evidence that they may make good targets for drug development and discovery.

DepMap and Project Achilles' CRISPR screen data have been used successfully to identify candidate genetic targets for multiple cancer types.[24–27] Gene essentiality is a measure used to quantify the effect of gene loss on a model organism or system. We hypothesized that we could use gene essentiality and single-cell expression in GBM to identify cells to prioritize for drug development. Using DepMap's CRISPR data, we identified 168 genes specifically essential to GBM cell lines. Our GDS complements existing literature showing that some of our top GDS hits such as LMNA, JUN, RAC1, and CDC42 have key roles in GBM aggression and tumorigenicity.[53–56] Translating *in vitro* results to patient data is challenging, yet our investigation highlighted GDS specificity to GBM neoplastic single-cells. Moreover, we identified minimal overlap between the GDS and dependency signatures from other tumor lineages (breast cancer, pancreatic ductal adenocarcinoma, and multiple myeloma; Supplementary Figure 1), emphasizing the specificity of this signature to GBM. Even with the inclusion of common essential genes, we would have expected greater overlaps between dependency signatures if lineage-specific gene essentiality were merely an *in vitro* phenomenon.

As the GDS was relevant for GBM, we used it to uncover three transcriptional states that reflect dependency gene vulnerabilities among neoplastic GBM single-cells. Transcriptional ITH in GBM has been explored extensively, with researchers characterizing cell states in high resolution using both scRNA- and stRNA-seq technologies.[7,11,12,14,30,48–50] We show that our identified VS correlate with – yet are distinct from – some of the leading GBM transcriptomic classification signatures. Our stem-like signature, VS1, critically, does not positively correlate with any modules from Neftel et al., nor those from Ravi et al.'s spatial transcriptomic analysis or

Verhaak et al.'s bulk RNA-seq-based GBH subtype classification scheme.[7,14,43] It does, however, exhibit strong neuronal and synaptic enrichment compounded with upregulation of canonical stem-cell pathways and embryonic TFs (PAX4, SOX5, REST), including the WNT/beta-catenin and Hedgehog signaling pathways. Therefore, VS1 most likely represents a unique population of neoplastic, neural stem cell-like cells despite it not clustering with the NPC1 or NPC2 signatures from Neftel et al. by our analysis. Drexler et al. recently described a high-neural epigenetic signature in GBM that was correlated with poorer patient outcome, higher glioma-neural synaptic integration, and high levels of neural stem cell-like cells, which may be a starting point toward exploring the functional significance of our VS1 cell signature for exploitation toward drug development.[57]

In contrast to VS1, our VS2 and VS3 signatures do correlate with cell state classifications from Ravi et al. (Regional NPC) and Neftel et al. (NPC1, NPC2, OPC) in addition to Verhaak et al.'s Proneural and Neural GBM transcriptional subtypes. In their publication Neftel et al. largely attribute NPC-like states to CDK4 amplification and OPC-like states to PDGFRA amplification, and subsequent research largely corroborates this.[12,13] Both VS2 and VS3 signatures are associated with cell cycle regulation and proliferation, and VS3 is associated with platelet-derived enrichments due to upregulation of exocytotic pathways. We additionally show, however, that VS2 is strongly and uniquely GDS-like, expressing the most GDS-derived genes and sharing overlapping enrichment profiles with the GDS. Given its additional spatial-level correlations with anti-apoptotic and immunoregulatory biological processes, VS2, while sharing qualities of NPC-like cells, is distinct from them. Moreover, VS3 shares characteristics with OPC-like cells, and its enrichment for cellular respiration, metabolic processes, and microtubule processes are in line with findings from Liu et al. who demonstrate that OPC-like cells reside in areas closer to vasculature, implying increased oxygen requirements for VS3 cells.[48]

Recently, Greenwald et al. put forth spatially defined malignant meta-programs identified across spatial GBM sections. In addition to the four Neftel et al. cell states, they identified a hypoxic MES program, a chromatin regulatory program, a proliferation and metabolism-related program, and an astrocyte-like MES program.[11] Similarly, Zheng et al. identified a hypoxic MES module and reactive astrocyte module (in lieu of AC-like) in addition to MES-like, NPC-like, and OPC-like cells in their independent spatial analysis, which they used to predict patient prognosis and tumor aggressiveness using machine learning methods.[45] The VS we identified share some similarities with these newer transcriptional profiles, our correlation analysis shows concretely that they are distinct. Moreover, researchers have shown – and we have observed here – that cells in any given state tend to be spatially clustered,[11,48] yet our VS also overlay with high degrees of correlation with specific pathway enrichment profiles highlighting the functional significance of these states. While the relevance of this phenomenon for treatment remains to be fully elucidated, we posit that the aforementioned developmental- and spatially-regulated transcriptional states provide great insight into the biology of tumors and that our VS capture differences between cells most relevant for drug discovery.

When attempting to identify novel therapeutic targets or strategies, we desire highly specific targets that would minimize off-target effects and toxicities. Our decision to include common essential genes in our GDS contradicts this logic since common essential genes have notable impacts on cell viability across all lineages. However, the essential genes that were included a) still met our criteria of having significantly lower ES in GBM than other lineages and b) being more inclusive of genes at this step allows for more robust identification of VS downstream. On a related note, our decision to focus exclusively on the CRISPR screen data rather than the RNAi data published by DepMap was to minimize off-target effects and to ensure confidence in complete gene silencing across thousands of cell lines of varying phenotypes. Another limitation of the GDS is that the genes in it are not necessarily transcriptionally co-regulated. This is important because we use the GDS as if co-expression of these genes as a whole has biological significance, and this may be true only for some genes in the GDS but not others. Nonetheless, expression of individual GDS genes at all is biologically important as measured by the DepMap screen, and our approach toward VS identification using Jaccard scoring allowed for those co-regulated GDS genes to have a larger impact on scores. Moreover, given the robustness of our results across both scRNA- and stRNA-seq datasets, our approach and methodology demonstrate that the GDS does inform biological significance through VS.

Importantly, our VS had high concordance across all three scRNA-seq datasets used in this study. Though only 250 DEG (2.21% of all VS3 DEG) overlap between significant DEG from the Neftel, Johnson and in-house datasets, this is likely due to the relatively high number of VS3 DEG identified in the Johnson and in-house datasets (8,367 and 6,071, respectively) compared to only 634 in the Neftel dataset.

Though deconvolution methods and technologies continue to improve, none are perfect, and spatial data provides unique challenges when data are based on spots rather than single cells. Our SCDC-based deconvolution of stRNA-seq data produced relatively low proportions of VS1 detected across 18 GBM sections, leading us to use lower proportional cutoffs when analyzing correlations between spot-level VS and pathway enrichment. Despite this, our spatial enrichment and correlation results were highly congruent with our scRNA-seq data.

Transcriptional cell states are informative for modeling and predicting treatment response and resistance on the single-cell level.[13,16,16,30,50,57–63] Devising effective therapies requires identification of those states and transcriptional modules specifically, over cell types, are either the most vulnerable to drug treatment or the most essential to tumor integrity. This comprehensive characterization of GBM vulnerabilities across both scRNA- and stRNA-seq data highlights the importance of multimodal approaches to drug discovery. We demonstrate that the application of essential gene signatures to expression profiling can uncover shared transcriptional regulatory modules across tumors that have actionable vulnerabilities based on extensive enrichment and correlation analyses. We thereby present a novel framework for targeted therapy development, small molecule discovery

and design, and drug repurposing. Individual patient tumors can be stratified by their relative VS transcriptional profiles, allowing for the potential for tumor-specific therapeutic targeting. Gene essentiality thereby provides a framework from which we can understand and exploit tumor transcriptional profiles on the single-cell level, yielding more options for pharmacological development in the treatment of GBM.

Critically, our pipeline as presented here need not be specific to GBM, and its applicability to other cancer types holds great promise. Such applications would not only validate the robustness and versatility of our methodology, but enhance the clinical relevance and translational applicability by uncovering novel therapeutic opportunities across different malignancies. By demonstrating the efficacy of our approach and the importance of essential genes on a cell-state level across multiple cancer types, we can enhance personalized medicine and improve patient outcomes at a broad scale for other difficult-to-treat cancers.

**Acknowledgements:**

We would like to thank Siddharth Jain for their helpful comments and advice. This work has been supported by BellRinger at the Georgetown University Lombardi Comprehensive Cancer Center (LCCC). We would like to thank LCCC and the Georgetown University Tumor Biology Program. We additionally thank all other members of the Ayad laboratory for invaluable discussion of this research.

**Author Contributions**:

**Conceptualization**: N.G.A., R.K.S., M.S., M.D. **Methodology, investigation, visualizations**: M.D. **Single-cell RNA sequencing samples and data**: R.K.S, M.E.I, R.J.K. **Computational design**: M.D., R.K.S., M.S., R.C. **Data analysis**: M.D. **Writing and revision**: M.D., R.K.S., N.G.A. **Supervision**: N.G.A., R.K.S.

**Competing Interest Statement**: N.G.A. is a founder of Decode Therapeutics. The authors declare no financial or personal conflicts of interest.

## STAR METHODS

### Characterization of glioblastoma dependency signature

CRISPR loss of function screen gene effect scores in addition to cell line metadata were downloaded directly from the DepMap portal. Downloaded data were from the 23Q2 release (May 2023). The post-Chronos effect scores dataset contained effect scores for 17,453 genes knocked out in 1,078 cell lines.[20] Cell line metadata were subset to include only those lines profiled by CRISPR screen. A linear regression model was then used to compare effect scores of all genes in the 49 GBM cell lines to the effect scores in all other cell line lineages.[64] Genes with significantly lower effect scores in GBM cell lines (negative effect size by linear regression and Bonferroni-adjusted p-value < 0.05) and with a mean effect score < -0.5 were selected as GDS genes. All analyses performed in this study were done using R and RStudio unless otherwise specified.[65,66]

### External single-cell RNA sequencing data set acquisition and processing

Previously published data from Neftel et al. was downloaded from GEO (GSE131928) as processed, log-normalized counts.[7] These data were then subject to our in-house pipeline for scRNA-seq processing using R and Seurat.[28] The log-normalized counts matrix containing expression data for 26 tumors was converted to a Seurat object and downloaded metadata from the authors added. Six (6) pediatric GBM tumors were excluded from downstream analysis at this point, leaving a total of 20 GBM tumors. Individual cells were confirmed to have < 20% mitochondrial transcripts and contain data on more than 100 features. Subsequently, cells were split by individual tumor and subject to normalization and variance stabilization using a regularized negative binomial regression implemented within the scTransform R package.[28,67] Individual tumor scRNA-seq data were subsequently integrated using an anchor-based technique utilizing the Pearson residuals.[28,68] Normalized, integrated values were then used for shared-nearest-neighbor (SNN) clustering, dimensional reduction, cell type verification, and cell type identification. Downstream analysis including differential expression and enrichment scoring was performed using log-normalized counts (see below).

Previously published data from Johnson et al. was downloaded from Synapse (https://synapse.org/singlecellglioma) and processed using our in-house scRNA-seq pipeline as described above.[30]

### Single-cell RNA sequencing of patient glioblastoma tumors

We performed scRNA-seq of six GBM tumors resected at the University of Miami. Three of these tumors were from patients with newly diagnosed GBM, and three from patients diagnosed with recurrent GBM. The newly

diagnosed GBM samples were resected and sequenced prior to patients receiving treatment, while the three recurrent tumors were resected from patients who had previously received standard of care treatment with TMZ chemotherapy and radiotherapy. Tissue collection was performed under approved IRB study 20170887 (MODCR00002196) at the University of Miami. University of Miami Hospital pathologists separated resected tissue into three spatially distinct groups, namely the necrotic core, enhancing edge, and infiltrating tumor. The tissue was transferred directly from the operating suite to the laboratory on ice. Tissue from the enhancing edge was used for dissociation and scRNA-seq.

Tumor samples were dissociated using the Worthington Biochemical Papain Dissociation System.[69] Tumor tissue was finely minced with a sterile surgical blade and incubated in sterile Earl's Balanced Salt Solution (EBSS) containing 20 units of papain per mL in 1mM L-cysteine with 0.5 mM EDTA and DNase for 45 minutes at 37 C while shaking. Following papain incubation, the tissue was gently triturated with a 5 mL serological pipette until it was adequately dissociated. The resulting suspensions were pelleted and resuspended in EBSS with ovomucoid protease inhibitor, bovine serum albumin (BSA), and DNase. This suspension was layered on top of ovomucoid protease inhibitor solution with BSA to form a gradient, which was centrifuged at 70xg for 6 minutes. The supernatant containing cellular debris was then removed. The pellets were subsequently resuspended, and the gradient centrifugation was repeated. The resulting pellet was then washed 2 times with ice-cold PBS with 0.1% BSA, passed through 0.45 uM cell strainer, and resuspended in ice-cold PBS with 0.1% BSA. This final suspension was then checked for cell viability by staining an aliquot with acridine orange and propidium iodide (AO/PI) and counted on a Nexcelom K2 Cellometer. All samples had cell viability greater than 90% prior to loading on the single-cell chips.

Single cells were captured using a 10X Genomics Chromium Controller, and patient single-cell gene expression libraries were prepared using 10X Genomics 5' Gene Expression chemistry. All libraries were sequenced on an Illumina NextSeq500. Alignment and demultiplexing was performed using CellRanger, and filtered unique molecular identifier (UMI) matrices were used for downstream analysis using the Seurat package in R.[28] Within Seurat, poor quality cells were further filtered by removing outliers based on the number of detected UMIs and unique transcripts, as well as percent expression of mitochondrial transcripts. Pass-filter cells were those with greater than 200 and less than 4,500 detected features, and with less than 12.5% mitochondrial transcript expression. Filtered data were normalized and variance stabilized using a regularized negative binomial regression implemented within the scTransform R package, and individual tumor scRNA-seq datasets were subsequently integrated with an anchor-based technique utilizing the Pearson residuals.[28,68] Normalized, integrated values were then used for initial shared-nearest-neighbor (SNN) clustering, dimensional reduction, and cell type identification. Downstream analysis including differential expression and enrichment scoring was performed using log-normalized counts.

**Discovery of glioblastoma vulnerability states**

The processed Neftel scRNA-seq dataset was split by patient tumor (n = 20 adult GBM) and the scaled, log-normalized expression of the 168-gene GDS genes was binarized. For each independent tumor, scaled expression of a gene > 0 was assigned "1" and all other values were assigned "0". The Jaccard Similarity Index (Jaccard Index, Jaccard Similarity, JS) was then computed between pairs of single cells based on these binarized expression values to create a pairwise cell-cell JS matrix for each tumor. The agglomerative coefficient was then computed for clustering methods including ward, average, single, and complete clustering. For each tumor, the ward clustering method had the highest agglomerative coefficient and was chosen for all subsequent clustering.

The JS matrix for each tumor was then clustered by Ward D2 hierarchical clustering for all k between 2 and 30. For each value of k, clusters were excluded if it included fewer than 5 cells or greater than 80% of the total number of neoplastic cells in the tumor to ensure high quality differential expression analysis (downstream). After exclusions, differential expression was performed between all clusters for each k using Seurat's FindMarkers with a log2FC threshold set to > 0 using MAST statistical testing.[40] Clusters at each k were retained if they contained greater than 50 genes ($N_{50}$) with an adjusted p-value < 0.05 and at least 1 gene ($N_1$) with an adjusted p-value < 0.001.

After clustering on each independent tumor, JS were calculated between retained clusters using the defined $N_{50}$ genes (modules) for each cluster. For pairs of modules with > 75% similarity (JS > 0.75), we retained the larger module (cluster with the highest number of genes in $N_{50}$). This resulted in 25 modules identified across 11 tumors, which were subsequently subjected to JS and Ward D2 hierarchical clustering with k = 3, uncovering 3 vulnerability states (VS). All heatmaps were visualized using the *pheatmap* R package.[70]

**Single-cell vulnerability state signature identification and cell state assignment**

Signature genes for each identified VS were identified by aggregating the top 100 genes across all modules comprising the VS, ranked by log2FC. Genes with more than one instance within a single signature were aggregated by taking the mean log2FC prior to this, resulting in 100-gene signatures per VS, consisting of 292 unique genes across all VS.

VS signature genes were used for VS assignment within the Neftel, Johnson, and Suter datasets. This was done using singscore, an independent-sample gene-set scoring method implemented in R, to score each neoplastic single-cell for expression of the 300 marker genes (Supplementary Table 4), binned according to VS (100 genes per VS).[41,42] Neoplastic cells were then assigned to a VS based on which VS marker gene-set had the highest singscore (Supplementary Figure 6).

**Differential expression and differential enrichment analysis**

All differential expression analyses were performed using Seurat's FindMarkers function with a baseline log2FC cutoff > 0 and statistical testing performed with MAST.[28,40]

Differential enrichment or over-representation analyses were performed using the Molecular Signatures Database with the collections named in each instance. R packages used for these analyses include enrichR, msigdbr, and ClusterProfiler.[31,71–73]

Comparisons between DEG profiles between datasets were performed using the VennDiagram R package to highlight overlapping DEG for each VS.[74]

**Cell type classification correlation analysis**

Transcriptional signatures from previously-published GBM classification systems were acquired from Verhaak et al. (2010), Neftel et al. (2019), and Ravi et al. (2022).[7,14,43] Each single-cell in our in-house scRNA-seq dataset was scored for enrichment of these signatures, in addition to our VS signatures, using singscore as mentioned previously. We then calculated Spearman correlations on the resulting single-cell enrichment scores to calculate and subsequently visualize with *corrplot* similarities between transcriptional classifiers.[75]

**Spatial transcriptomic data acquisition, processing, deconvolution, and analysis**

10X Visium spatial transcriptomic RNA sequencing (stRNA-seq) data was acquired from Ravi et al. (2022) as h5 counts with corresponding H&E-stained section images from Dryad.[76] These data were processed and analyzed using Seurat as described above for scRNA-seq data. Of 28 total sections, 18 are GBM tumor sections; 8 of 28 are surgical entry site normal cortex, and the 2 remaining samples are IDH-mutant and thereby not classified a GBM and not analyzed here. In short, poor quality spots without counts data were removed, and filtered data were normalized and variance stabilized image-by-image using scTransform.[67] Individual images were subsequently merged and integrated using a canonical correlation analysis (CCA)-based method implemented through Seurat. Normalized, integrated values for all spots across all sections were then used for SNN clustering and dimensionality reduction. Downstream analyses including differential expression and enrichment were performed on scaled data regressed for cell cycle genes and mitochondrial RNA.

For spot-based deconvolution of stRNA-seq data, we used SCDC, an RNA-seq deconvolution method that can leverage cell-type specific gene expression information from multiple scRNA-seq references to address batch-

effect confounding.[46] Raw counts of the Ravi et al. stRNA-seq data were used for spot-based deconvolution with our in-house scRNA-seq dataset used as the reference set to provide single-cell type annotations. Visualization of deconvolution results from SCDC was performed with *SPOTlight* to represent each spot as a pie chart with each cell type predicted at that spot colored as a pie section (see Figure 5).[77]

**Gene set co-regulation analysis and spot-level correlation analysis**

The 18 GBM tumor spatial sections were subject to gene set co-regulation analysis (GESECA) using the fgsea R package[47] and the Gene Ontology Biological Processes (GOBP).[78,79] Input data were the results of reverse principal component analysis performed on scaled RNA-seq data from the Spatial assay with cell cycle genes and mitochondrial RNA regressed out. Minimum gene set size was set to 2, maximum size was set to 500. Enriched pathways were filtered for those with BH-adjusted p-values less than 0.05, and then the top 50 by adjusted p-value were visualized using the z-score of the GESECA enrichment score.

In addition to spot-level proportion data on VS from SCDC deconvolution, we used singscore to calculate an enrichment score for each VS at each spot across all 18 tumor sections. Similarly, we used singscore to calculate enrichment scores for significantly enriched GOBP terms to allow for accurate correlation comparisons between VS and GOBP pathways. We then correlated VS and GOBP enrichment scores for all spots across all 18 sections. For each VS, we correlated those spots with > 0% (VS1) or > 50% (VS2 and VS3) content in any given spot with GOBP enriched terms using singscore-based enrichment scores. Different proportional cutoff values were used for each VS given the relatively low levels of VS1 detected across all GBM sections in the stRNA-seq data to include similar numbers of spots in each independent correlation analysis. Correlations greater than 0.5 (VS1 and VS2) or 0.3 (VS3) are available in Supplementary File 3.

**Data Availability**: All data and code needed to generate and reproduce the materials and conclusions presented in this manuscript are available in the body of the text, the Supplementary Materials, and/or the Ayad Laboratory GitHub repository (https://github.com/AyadLab/scZeroDay). Data, code, and materials are made available to any researcher. Our generated single-cell data of patient GBM tumors, including processed Seurat objects, are available via the Gene Expression Omnibus (GEO) with accession ID GSE229779 (Reviewer Token: yzghmqawxbodtmf). Neftel et al. scRNA-seq data were obtained from GEO with ID GSE131928. Johnson et al. scRNA-seq data were obtained from Synapse at https://synapse.org/singlecellglioma. Spatial transcriptomic data published by Ravi et al. were obtained from Dryad at https://doi.org/10.5061/dryad.h70rxwdmj.

## SUPPLEMENTARY MATERIALS

**Supplementary Table 1:** DepMap metadata of 49 glioblastoma cell lines
**Supplementary Table 2:** Statistics from 377 genes with significantly lower effect scores in glioblastoma compared to all other cell line lineages
**Supplementary Table 3:** Statistics from 168-gene glioblastoma dependency signature
**Supplementary Table 4:** Vulnerability state signature genes (n = 292)
**Supplementary Table 5:** Top 100 differentially expressed genes between vulnerability states
**Supplementary Table 6:** List of top 1,500 differentially expressed genes between vulnerability states

**Supplementary File 1:** Zip file containing 19 heatmaps displaying Jaccard scores for individual tumor samples from Neftel et al. (2019) not showing in Figure 2
**Supplementary File 2:** Enrichment results from GESECA enrichment analysis on each of 18 tumor sections submitted as zipped R data files
**Supplementary File 3:** Spot-level correlation analysis results between VS and significantly enriched terms from Gene Ontology Biological Processes

**Supplementary Figure 1: Further characterization of glioblastoma dependency signature**
(**A**) Scatter plot depicting the differences in gene ES in GBM-lineage cell lines (n = 49) compared to all other cell lines (n = 1,029) by a linear model using all 17,453 genes profiled by DepMap's CRISPR screens. Negative effect size indicates genes with lower effect scores in GBM cell lines. The negative log of the adjusted p-value is represented on the y-axis, with a p-value cutoff of 0.05 represented by the horizontal dashed line. Blue coloring (both shades) represents genes with a significant p-value (< 0.05). Candidate GBM-specific essential genes are those with p < 0.05 and negative effect size (light blue). (**B**) Scatter plot depicting the second level of selection for the GBM dependency signature (GDS). All genes with adjusted p-value < 0.05 from (A) were plotted as their mean ES in all GBM-lineage cell lines (y-axis) versus their effect size. Identified GDS genes are those with a negative effect size and a mean effect score < -0.5 (white). (**C**) Additional enrichment analysis of the 168 genes that comprise the GDS, conducted using MSigDB's Gene Ontology Molecular Function (GOMF) gene sets. Color indicates BH-adjusted p-value. Gene ratio (x-axis) is the ratio of genes present in the GDS to the genes present in the named GOMF gene set (fractional representation to the right of each bar). (**D**) Four-way Venn diagram depicting the number of overlapping dependency genes identified for multiple cancer cell line lineages using methods identical to those used for GDS identification. (BC = breast cancer, PDAC = pancreatic ductal adenocarcinoma, MM = multiple myeloma).

**Supplementary Figure 2: Characterization of 20 adult glioblastoma tumors subset from Neftel et al. scRNA-seq dataset**

(**A**) Stacked bar plot showing the proportion of cell types in the n = 20 adult GBM tumors present in the Neftel et al. (2019) scRNA-seq dataset. (**B**) UMAP plot of the Neftel et al. (2019) scRNA-seq dataset with cell type labels. (**C-G**) UMAP plots depicting scaled expression of named cell type lineage markers. An expression value of 0 includes cells with negative scaled expression. (**H**) Boxplot with underlying violin plot that shows the MS of the GDS in each single cell (black dots) in the Neftel et al. scRNA-seq dataset, grouped by cell type. Cell type MS are compared by an ANOVA with Tukey's post-hoc test for multiple comparisons. (**I**) Boxplot with underlying violin plot that shows the MS of the GDS in each single cell (black dots) in the Neftel et al. scRNA-seq dataset, grouped by Neftel cell states. Cell state GDS MS are compared by an ANOVA with Tukey's post-hoc test (**** = $p < 0.0001$, *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$, n.s. = not significant).

**Supplementary Figure 3: Characterization of 5 adult glioblastoma tumors subset from Johnson et al. scRNA-seq dataset**

(**A**) Stacked bar plot showing the proportion of cell types present in the Johnson et al. (2021) scRNA-seq dataset. (**B-G**) UMAP plots depicting scaled expression of named cell type lineage markers. An expression value of 0 includes cells with negative scaled expression. (**H**) UMAP plot of the Johnson et al. (2021) scRNA-seq dataset with cell type labels. (**I**) Boxplot with underlying violin plot that shows the MS of the GDS in each single cell (black dots) in the Johnson et al. scRNA-seq dataset, grouped by cell type. Cell type MS are compared by an ANOVA with Tukey's post-hoc test. (**I**) Boxplot with underlying violin plot that shows the MS of the GDS in each single cell (black dots) in the Johnson et al. scRNA-seq dataset, grouped by VS assignment. Cell state MS are compared by an ANOVA with Tukey's post-hoc test (**** = $p < 0.0001$, *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$, n.s. = not significant).

**Supplementary Figure 4: Characterization of 6 adult glioblastoma tumors by scRNA-seq**

(**A**) Stacked bar plot showing the proportion of cell types identified in our in-house scRNA-seq dataset. (**B-G**) UMAP plots depicting scaled expression of named cell type lineage markers. An expression value of 0 includes cells with negative scaled expression. (**H**) UMAP plot of our in-house scRNA-seq dataset with identified cell type labels. (**I**) Boxplot with underlying violin plot that shows the MS of the GDS in each single cell (black dots) in our in-house dataset, grouped by cell type. Cell type MS are compared by an ANOVA with Tukey's post-hoc test. (**I**) Boxplot with underlying violin plot that shows the MS of the GDS in each single cell (black dots) grouped by VS assignment. Cell state MS are compared by an ANOVA with Tukey's post-hoc test (**** = $p < 0.0001$, *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$, n.s. = not significant).

**Supplementary Figure 5: Further characterization of glioblastoma vulnerability states**

(**A-B**) Elbow a gap statistic plots that were used to choose the optimal k-cutoff for identification of the 3 VS. (**C**) Heatmap showing the pseudo-bulk scaled expression of the top 100 DEG between VS in the Neftel et al. (2019) dataset. (**D**) Heatmap showing the single-cell scaled expression of the 1,500 DEG between each VS in the Neftel

dataset. (**E**) The top 5 differentially enriched terms between the top 100 DEG per VS using MSigDB's C5 and H collections. (**F**) To top 5 differentially enriched terms between the top 100 DEG per VS using MSigDB's C2 and H collections.

**Supplementary Figure 6: Vulnerability state differentially expressed genes are conserved across datasets**

(**A-C**) Bar plots depicting the top enriched pathways for overlapping VS DEG between all three scRNA-seq datasets (Neftel, Johnson, In-House) with (**A**) corresponding to VS1 enrichments, (**B**) corresponding to VS2 enrichments, and (**C**) corresponding to VS3 enrichments. Enrichment analysis was performed using enrichR and the MSigDB Hallmarks gene set. Color represents the BH-adjusted p-value, and the gene ratio represents the number of overlapping DEG compared to the number of genes for a given term. (**D**) Correlation analysis between VS signatures and Kandoth et al. (2013) pan-cancer significantly mutated genes (SMG) mutational signatures.

**Supplementary Figure 7**: **Additional spatial transcriptomic spot-level correlation and enrichment analysis findings**

(**A, D, G, J, M, P**) Spot-level enrichment score scatter plot with Pearson's correlation analysis comparing VS singscore enrichment to the named GOBP pathway singscore enrichment. Each black dot represents a single spot with VS proportion > 50% by SCDC deconvolution (> 0% for VS1). Spots were also filtered to contain < 20% non-neoplastic cells to ensure the majority of cells in spots used for correlation analysis contain tumor cells. Colored line represents the line of best fit with error shaded in gray. (**B, E, H, K, N, Q**) Spatial scatter plot showing scaled GESECA enrichment scores for the named term (by row in the figure) at each spot. (**C, F, I, L, O, R**) Spatial scatter pie plots showing the proportion of each VS or non-neoplastic cells predicted to be present at each spot by SCDC deconvolution analysis.

## REFERENCES


1. Stupp Roger, Mason Warren P., van den Bent Martin J., et al. Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. *N Engl J Med*. 2005;352(10):987-996. doi:10.1056/NEJMoa043330
2. Stupp R, Hegi ME, Mason WP, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol*. 2009;10(5):459-466. doi:10.1016/S1470-2045(09)70025-7
3. Tamimi AF, Juweid M. Epidemiology and Outcome of Glioblastoma. In: De Vleeschouwer S, ed. *Glioblastoma*. Codon Publications; 2017. Accessed June 28, 2024. http://www.ncbi.nlm.nih.gov/books/NBK470003/
4. Qazi MA, Bakhshinyan D, Singh SK. Deciphering brain tumor heterogeneity, one cell at a time. *Nat Med*. 2019;25(10):1474-1476. doi:10.1038/s41591-019-0605-1
5. Le Rhun E, Preusser M, Roth P, et al. Molecular targeted therapy of glioblastoma. *Cancer Treat Rev*. 2019;80:101896. doi:10.1016/j.ctrv.2019.101896
6. Meyer M, Reimand J, Lan X, et al. Single cell-derived clonal analysis of human glioblastoma links functional and genomic heterogeneity. *Proc Natl Acad Sci*. 2015;112(3):851-856. doi:10.1073/pnas.1320611111
7. Neftel C, Laffy J, Filbin MG, et al. An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell*. 2019;178(4):835-849.e21. doi:10.1016/j.cell.2019.06.024
8. Patel AP, Tirosh I, Trombetta JJ, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science*. 2014;344(6190):1396-1401. doi:10.1126/science.1254257
9. Suvà ML, Tirosh I. The Glioma Stem Cell Model in the Era of Single-Cell Genomics. *Cancer Cell*. 2020;37(5):630-636. doi:10.1016/j.ccell.2020.04.001
10. Barkley D, Moncada R, Pour M, et al. Cancer cell states recur across tumor types and form specific interactions with the tumor microenvironment. *Nat Genet*. 2022;54(8):1192-1201. doi:10.1038/s41588-022-01141-9
11. Greenwald AC, Darnell NG, Hoefflin R, et al. Integrative spatial analysis reveals a multi-layered organization of glioblastoma. *Cell*. 2024;187(10):2485-2501.e26. doi:10.1016/j.cell.2024.03.029
12. Couturier CP, Ayyadhury S, Le PU, et al. Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nat Commun*. 2020;11:3406. doi:10.1038/s41467-020-17186-5
13. Couturier CP, Nadaf J, Li Z, et al. Glioblastoma scRNA-seq shows treatment-induced, immune-dependent increase in mesenchymal cancer cells and structural variants in distal neural stem cells. *Neuro-Oncol*. 2022;24(9):1494-1508. doi:10.1093/neuonc/noac085
14. Ravi VM, Will P, Kueckelhaus J, et al. Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma. *Cancer Cell*. 2022;40(6):639-655.e13. doi:10.1016/j.ccell.2022.05.009

15. Weinberg RA. *The Biology of Cancer*. W.W. Norton & Company; 2007. doi:10.1201/9780203852569
16. Richards LM, Whitley OKN, MacLeod G, et al. Gradient of Developmental and Injury Response transcriptional states defines functional vulnerabilities underpinning glioblastoma heterogeneity. *Nat Cancer*. 2021;2(2):157-173. doi:10.1038/s43018-020-00154-9
17. Yabo YA, Heiland DH. Understanding glioblastoma at the single-cell level: Recent advances and future challenges. *PLoS Biol*. 2024;22(5):e3002640. doi:10.1371/journal.pbio.3002640
18. Saini A, Ballesta A, Gallo JM. Cell state-directed therapy – epigenetic modulation of gene transcription demonstrated with a quantitative systems pharmacology model of temozolomide. *CPT Pharmacomet Syst Pharmacol*. 2023;12(3):360-374. doi:10.1002/psp4.12916
19. DepMap B (2023). This DepMap release contains data from CRISPR knockout screens from project Achilles, as well as genomic characterization data from the CCLE project. For more information, please see README.txt.. figshare. doi:https://doi.org/10.6084/m9.figshare.22765112.v2
20. Dempster JM, Boyle I, Vazquez F, et al. Chronos: a CRISPR cell population dynamics model. Published online February 25, 2021:2021.02.25.432728. doi:10.1101/2021.02.25.432728
21. Dempster JM, Rossen J, Kazachkova M, et al. Extracting Biological Insights from the Project Achilles Genome-Scale CRISPR Screens in Cancer Cell Lines. Published online July 31, 2019:720243. doi:10.1101/720243
22. Pacini C, Dempster JM, Boyle I, et al. Integrated cross-study datasets of genetic dependencies in cancer. *Nat Commun*. 2021;12(1):1661. doi:10.1038/s41467-021-21898-7
23. Meyers RM, Bryan JG, McFarland JM, et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat Genet*. 2017;49(12):1779-1784. doi:10.1038/ng.3984
24. Pacini C, Duncan E, Gonçalves E, et al. A comprehensive clinically informed map of dependencies in cancer cells and framework for target prioritization. *Cancer Cell*. 2024;42(2):301-316.e9. doi:10.1016/j.ccell.2023.12.016
25. Sandoval GJ, Pulice JL, Pakula H, et al. Binding of TMPRSS2-ERG to BAF Chromatin Remodeling Complexes Mediates Prostate Oncogenesis. *Mol Cell*. 2018;71(4):554-566.e7. doi:10.1016/j.molcel.2018.06.040
26. Shimada K, Bachman JA, Muhlich JL, Mitchison TJ. shinyDepMap, a tool to identify targetable cancer genes and their functional connections from Cancer Dependency Map data. *eLife*. 2021;10:e57116. doi:10.7554/eLife.57116
27. Wang X, Wang S, Troisi EC, et al. BRD9 defines a SWI/SNF sub-complex and constitutes a specific vulnerability in malignant rhabdoid tumors. *Nat Commun*. 2019;10(1):1881. doi:10.1038/s41467-019-09891-7
28. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573-3587.e29. doi:10.1016/j.cell.2021.04.048

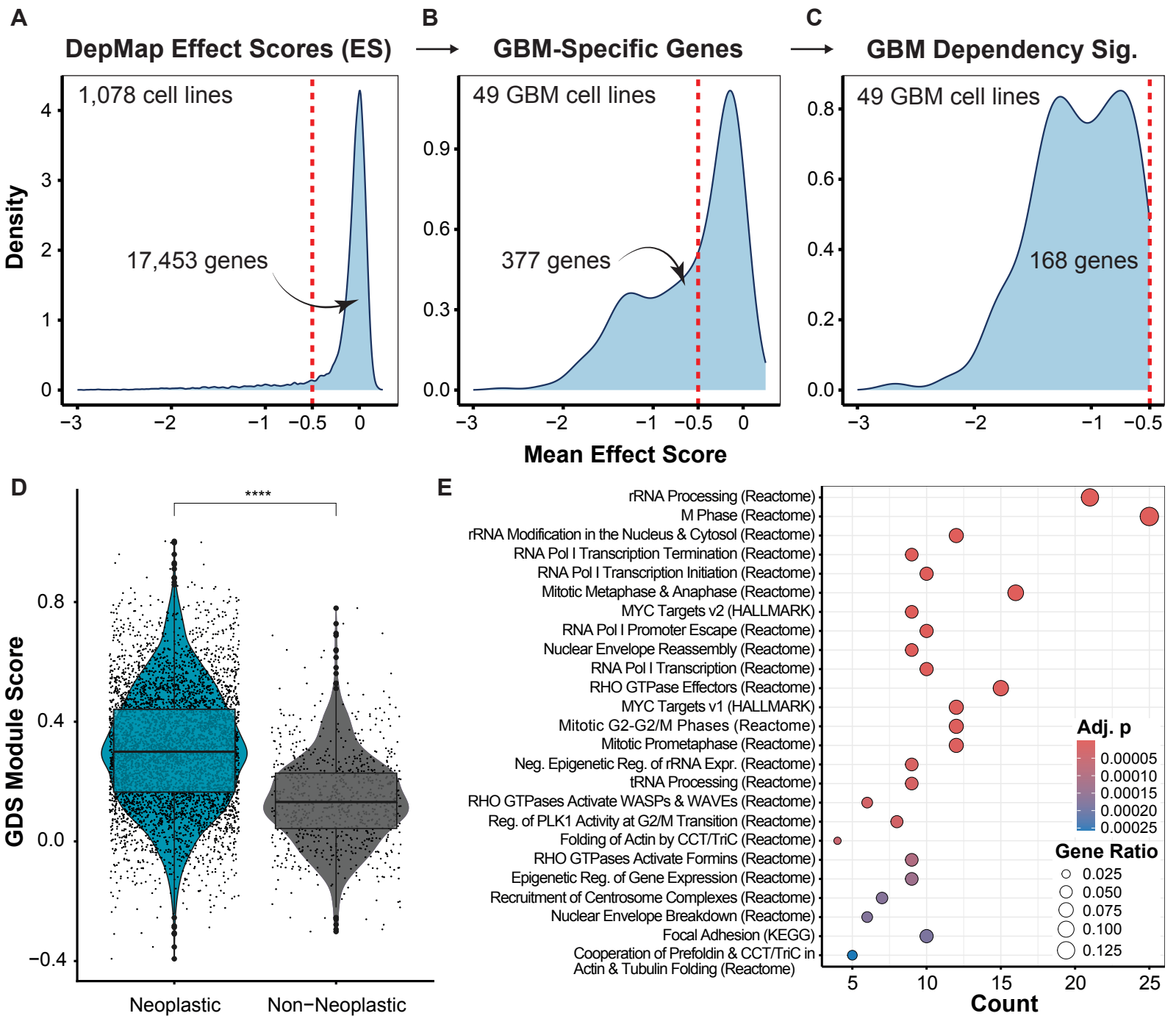
29. Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*. 2016;352(6282):189-196. doi:10.1126/science.aad0501
30. Johnson KC, Anderson KJ, Courtois ET, et al. Single-cell multimodal glioma analyses identify epigenetic regulators of cellular plasticity and environmental stress response. *Nat Genet*. 2021;53(10):1456-1468. doi:10.1038/s41588-021-00926-8
31. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27(12):1739-1740. doi:10.1093/bioinformatics/btr260
32. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417-425. doi:10.1016/j.cels.2015.12.004
33. Milacic M, Beavers D, Conley P, et al. The Reactome Pathway Knowledgebase 2024. *Nucleic Acids Res*. 2024;52(D1):D672-D678. doi:10.1093/nar/gkad1025
34. Kolmykov S, Yevshin I, Kulyashov M, et al. GTRD: an integrated view of transcription regulation. *Nucleic Acids Res*. 2021;49(D1):D104-D111. doi:10.1093/nar/gkaa1057
35. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res*. 2017;45(D1):D353-D361. doi:10.1093/nar/gkw1092
36. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res*. 2016;44(D1):D457-462. doi:10.1093/nar/gkv1070
37. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res*. 2010;38(Database issue):D355-360. doi:10.1093/nar/gkp896
38. Jassal B, Matthews L, Viteri G, et al. The reactome pathway knowledgebase. *Nucleic Acids Res*. 2020;48(D1):D498-D503. doi:10.1093/nar/gkz1031
39. Griss J, Viteri G, Sidiropoulos K, Nguyen V, Fabregat A, Hermjakob H. ReactomeGSA - Efficient Multi-Omics Comparative Pathway Analysis. *Mol Cell Proteomics MCP*. 2020;19(12):2115-2125. doi:10.1074/mcp.TIR120.002155
40. Finak G, McDavid A, Yajima M, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol*. 2015;16:278. doi:10.1186/s13059-015-0844-5
41. Foroutan M, Bhuva DD, Lyu R, Horan K, Cursons J, Davis MJ. Single sample scoring of molecular phenotypes. *BMC Bioinformatics*. 2018;19(1):404. doi:10.1186/s12859-018-2435-4
42. Bhuva DD, Cursons J, Davis MJ. Stable gene expression for normalisation and single-sample scoring. *Nucleic Acids Res*. 2020;48(19):e113. doi:10.1093/nar/gkaa802
43. Verhaak RGW, 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(1):98-110. doi:10.1016/j.ccr.2009.12.020

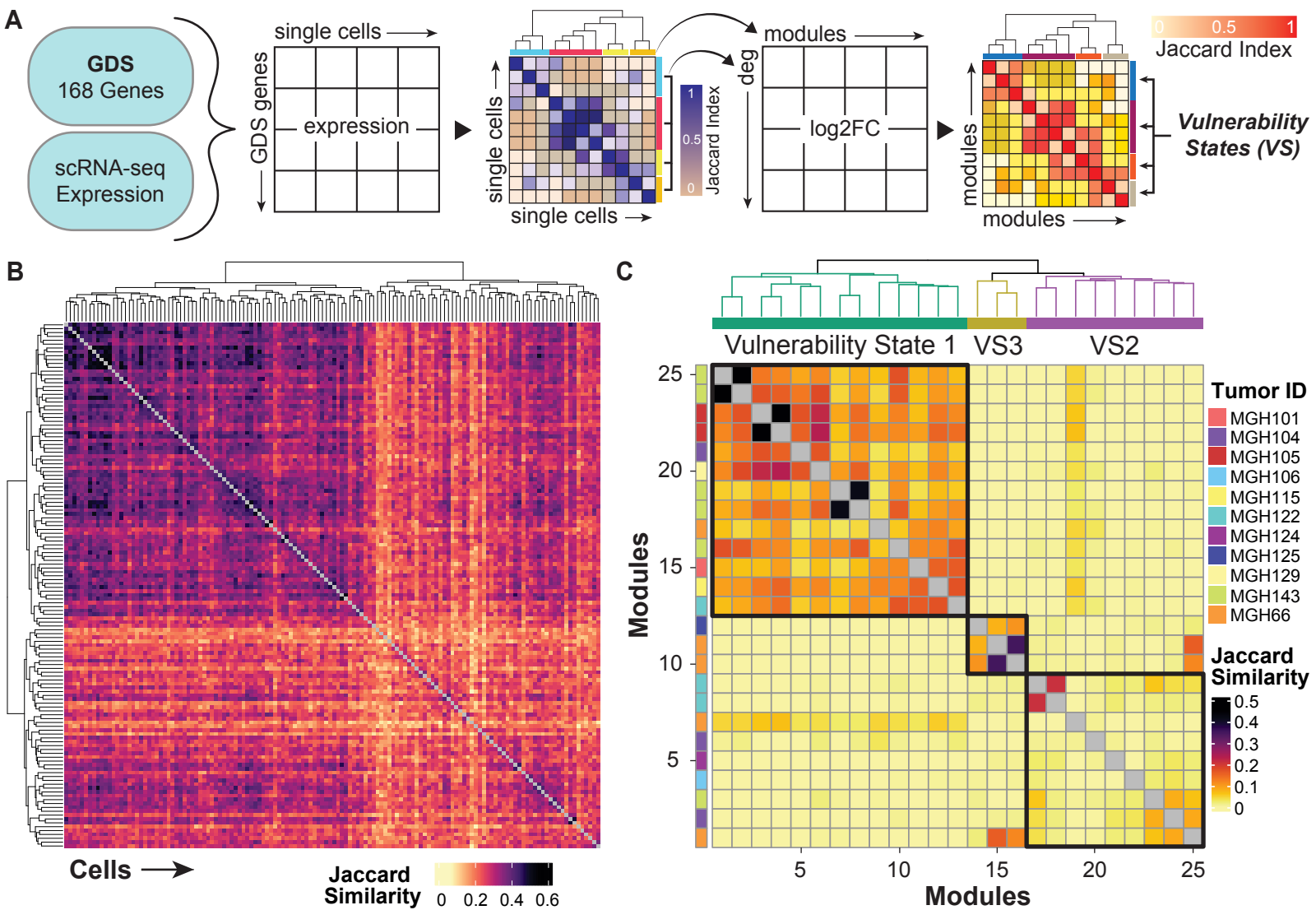
44. Ren L, Li J, Wang C, et al. Single cell RNA sequencing for breast cancer: present and future. *Cell Death Discov*. 2021;7:104-104. doi:10.1038/s41420-021-00485-1
45. Zheng Y, Carrillo-Perez F, Pizurica M, Heiland DH, Gevaert O. Spatial cellular architecture predicts prognosis in glioblastoma. *Nat Commun*. 2023;14(1):4122. doi:10.1038/s41467-023-39933-0
46. Dong M, Thennavan A, Urrutia E, et al. SCDC: bulk gene expression deconvolution by multiple single-cell RNA sequencing references. *Brief Bioinform*. 2021;22(1):416-427. doi:10.1093/bib/bbz166
47. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. Published online February 1, 2021:060012. doi:10.1101/060012
48. Liu M, Ji Z, Jain V, et al. Spatial transcriptomics reveals segregation of tumor cell states in glioblastoma and marked immunosuppression within the perinecrotic niche. *Acta Neuropathol Commun*. 2024;12:64. doi:10.1186/s40478-024-01769-0
49. Ren Y, Huang Z, Zhou L, et al. Spatial transcriptomics reveals niche-specific enrichment and vulnerabilities of radial glial stem-like cells in malignant gliomas. *Nat Commun*. 2023;14(1):1028. doi:10.1038/s41467-023-36707-6
50. Gangoso E, Southgate B, Bradley L, et al. Glioblastomas acquire myeloid-affiliated transcriptional programs via epigenetic immunoediting to elicit immune evasion. *Cell*. 2021;184(9):2454-2470.e26. doi:10.1016/j.cell.2021.03.023
51. 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(6):779-792.e11. doi:10.1016/j.ccell.2021.05.002
52. Bhaduri A, Di Lullo E, Jung D, et al. Outer Radial Glia-like Cancer Stem Cells Contribute to Heterogeneity of Glioblastoma. *Cell Stem Cell*. 2020;26(1):48-63.e6. doi:10.1016/j.stem.2019.11.015
53. Gatti G, Vilardo L, Musa C, et al. Role of Lamin A/C as Candidate Biomarker of Aggressiveness and Tumorigenicity in Glioblastoma Multiforme. *Biomedicines*. 2021;9(10):1343. doi:10.3390/biomedicines9101343
54. Zhou S, Niu R, Sun H, Kim SH, Jin X, Yin J. The MAP3K1/c-JUN signaling axis regulates glioblastoma stem cell invasion and tumor progression. *Biochem Biophys Res Commun*. 2022;612:188-195. doi:10.1016/j.bbrc.2022.04.057
55. Glogowska A, Thanasupawat T, Beiko J, Pitz M, Hombach-Klonisch S, Klonisch T. Novel CTRP8-RXFP1-JAK3-STAT3 axis promotes Cdc42-dependent actin remodeling for enhanced filopodia formation and motility in human glioblastoma cells. *Mol Oncol*. 2022;16(2):368-387. doi:10.1002/1878-0261.12981
56. Hao J, Han X, Huang H, et al. Sema3C signaling is an alternative activator of the canonical WNT pathway in glioblastoma. *Nat Commun*. 2023;14(1):2262. doi:10.1038/s41467-023-37397-w
57. Drexler R, Khatri R, Sauvigny T, et al. A prognostic neural epigenetic signature in high-grade glioma. *Nat Med*. 2024;30(6):1622-1635. doi:10.1038/s41591-024-02969-w
58. Cheng YL, Banu MA, Zhao W, Rosenfeld SS, Canoll P, Sims PA. Multiplexed single-cell lineage tracing of

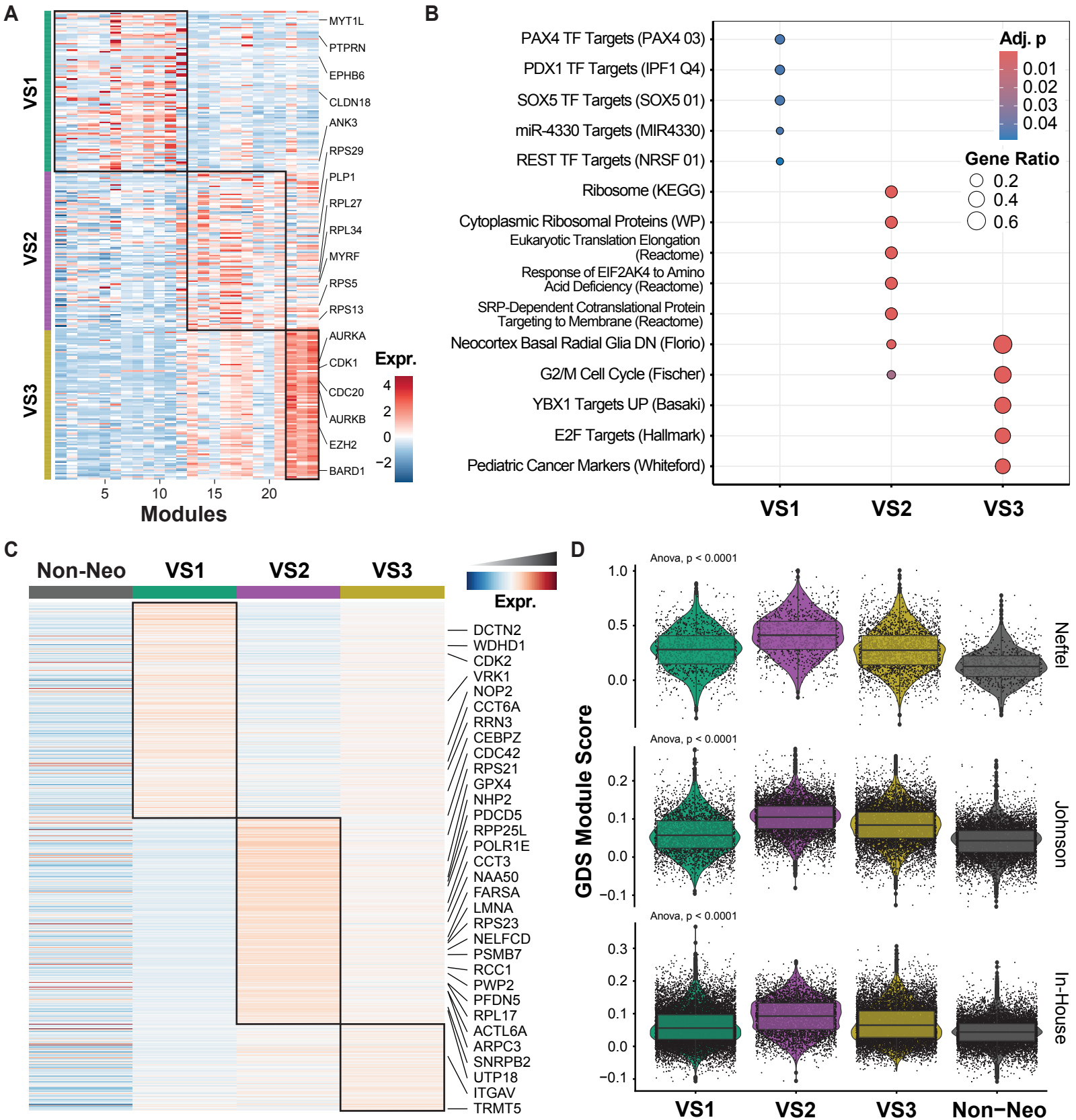
mitotic kinesin inhibitor resistance in glioblastoma. *Cell Rep*. 2024;43(5). doi:10.1016/j.celrep.2024.114139
59. Eyler CE, Matsunaga H, Hovestadt V, Vantine SJ, van Galen P, Bernstein BE. Single-cell lineage analysis reveals genetic and epigenetic interplay in glioblastoma drug resistance. *Genome Biol*. 2020;21(1):174. doi:10.1186/s13059-020-02085-1
60. Larsson I, Dalmo E, Elgendy R, et al. Modeling glioblastoma heterogeneity as a dynamic network of cell states. *Mol Syst Biol*. 2021;17(9):e10105. doi:10.15252/msb.202010105
61. Rabé M, Dumont S, Álvarez-Arenas A, et al. Identification of a transient state during the acquisition of temozolomide resistance in glioblastoma. *Cell Death Dis*. 2020;11(1):19. doi:10.1038/s41419-019-2200-2
62. Rukhlenko OS, Halasz M, Rauch N, et al. Control of cell state transitions. *Nature*. 2022;609(7929):975-985. doi:10.1038/s41586-022-05194-y
63. Yabo YA, Niclou SP, Golebiewska A. Cancer cell heterogeneity and plasticity: A paradigm shift in glioblastoma. *Neuro-Oncol*. 2022;24(5):669-682. doi:10.1093/neuonc/noab269
64. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47. doi:10.1093/nar/gkv007
65. R CT. A Language and Environment for Statistical Computing. Published online 2024. https://www.R-project.org
66. Posit team. RStudio: Integrated Development Environment for R. Published online 2023. http://www.posit.co/
67. Choudhary S, Satija R. Comparison and evaluation of statistical error models for scRNA-seq. *Genome Biol*. 2022;23(1):27. doi:10.1186/s13059-021-02584-9
68. Hao Y, Stuart T, Kowalski MH, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol*. 2024;42(2):293-304. doi:10.1038/s41587-023-01767-y
69. Huettner JE, Baughman RW. Primary culture of identified neurons from the visual cortex of postnatal rats. *J Neurosci*. 1986;6(10):3044-3060. doi:10.1523/JNEUROSCI.06-10-03044.1986
70. Kolde R. pheatmap: Pretty Heatmaps. Published online January 4, 2019. Accessed July 1, 2024. https://cran.r-project.org/web/packages/pheatmap/index.html
71. Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*. 2021;2(3). doi:10.1016/j.xinn.2021.100141
72. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci*. 2005;102(43):15545-15550. doi:10.1073/pnas.0506580102
73. Dolgalev I. msigdbr: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format. Published online March 30, 2022. Accessed July 1, 2024. https://cran.r-project.org/web/packages/msigdbr/index.html
74. Chen H. VennDiagram: Generate High-Resolution Venn and Euler Plots. Published online April 12, 2022. Accessed July 1, 2024. https://cran.r-project.org/web/packages/VennDiagram/index.html
75. Wei T, Simko V, Levy M, et al. corrplot: Visualization of a Correlation Matrix. Published online November

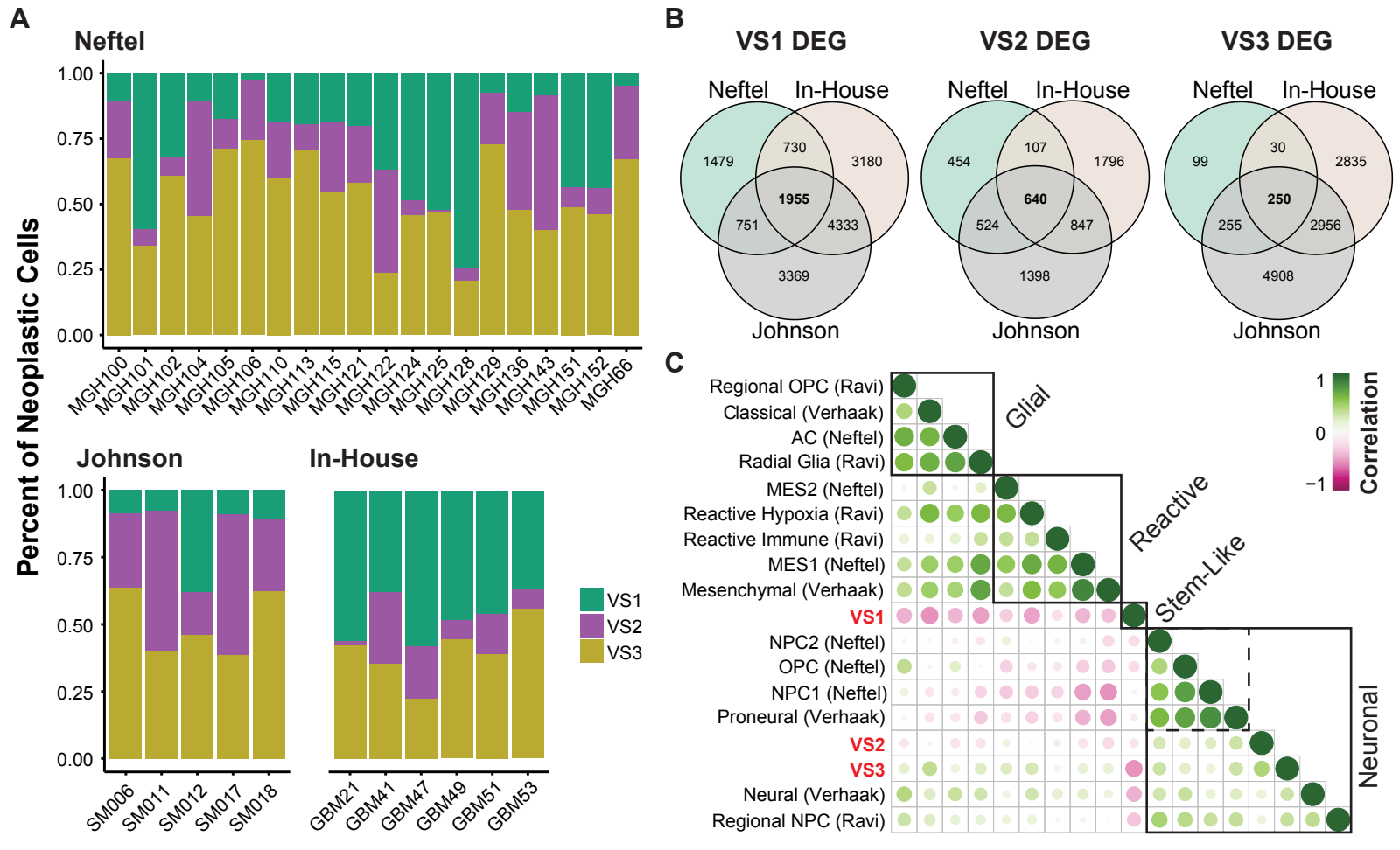
18, 2021. Accessed July 1, 2024. https://cran.r-project.org/web/packages/corrplot/index.html

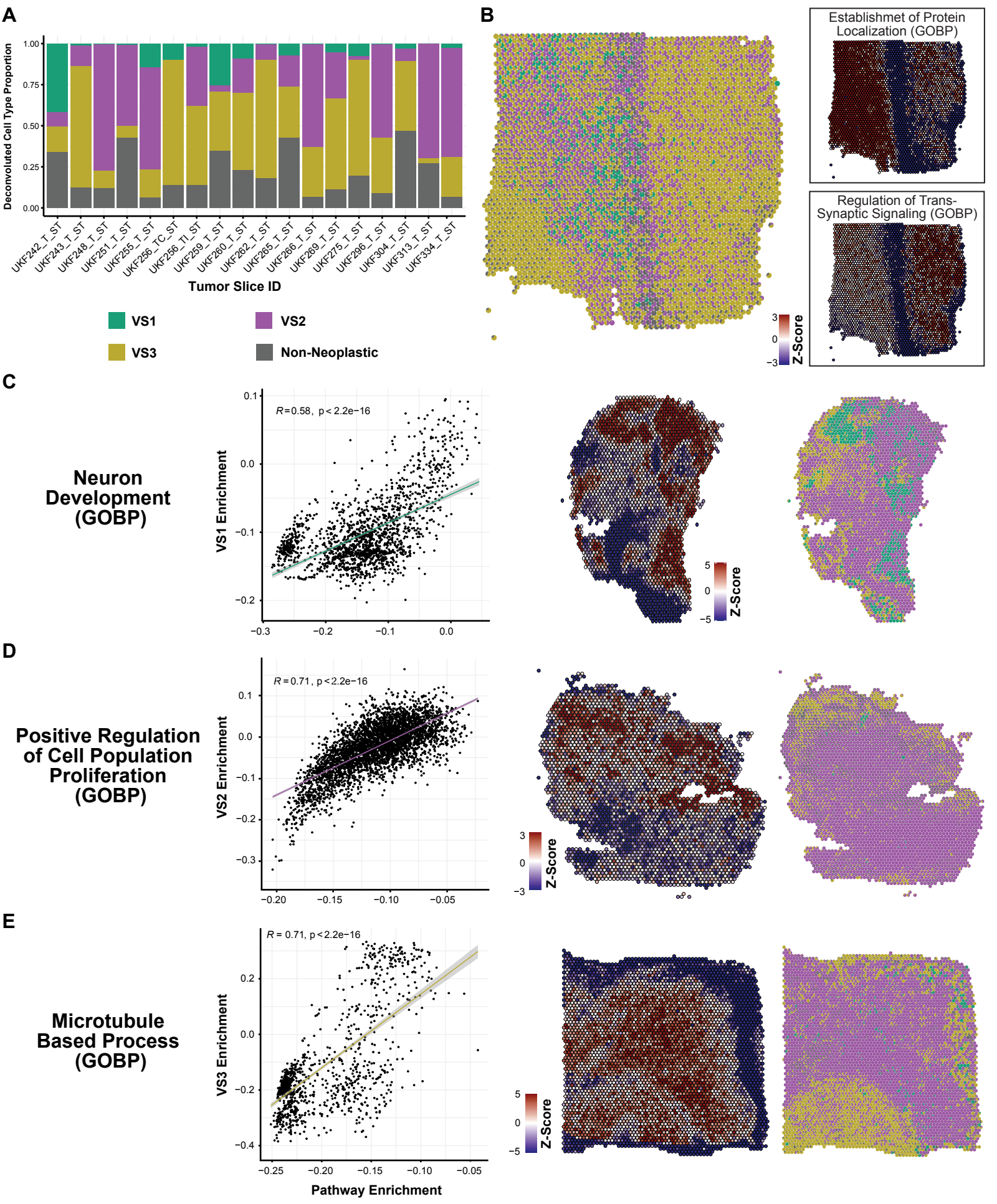
76. Ravi VM, Will P, Kueckelhaus J, et al. Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma. Published online June 2022. doi:10.5061/dryad.h70rxwdmj
77. Elosua-Bayes M, Crowell H. SPOTlight: Spatial Transcriptomics Deconvolution. Published online 2024. Accessed July 1, 2024. http://bioconductor.org/packages/SPOTlight/
78. The Gene Ontology Consortium, Aleksander SA, Balhoff J, et al. The Gene Ontology knowledgebase in 2023. *Genetics*. 2023;224(1):iyad031. doi:10.1093/genetics/iyad031
79. Ashburner M, Ball CA, Blake JA, et al. Gene Ontology: tool for the unification of biology. *Nat Genet*. 2000;25(1):25-29. doi:10.1038/75556


A
DepMap Effect Scores (ES)
1,078 cell lines
17,453 genes
Density
B
GBM-Specific Genes
49 GBM cell lines
377 genes
C
GBM Dependency Sig.
49 GBM cell lines
168 genes
Mean Effect Score
D
****
GDS Module Score
Neoplastic
Non−Neoplastic
E
rRNA Processing (Reactome)
M Phase (Reactome)
rRNA Modification in the Nucleus & Cytosol (Reactome)
RNA Pol I Transcription Termination (Reactome)
RNA Pol I Transcription Initiation (Reactome)
Mitotic Metaphase & Anaphase (Reactome)
MYC Targets v2 (HALLMARK)
RNA Pol I Promoter Escape (Reactome)
Nuclear Envelope Reassembly (Reactome)
RNA Pol I Transcription (Reactome)
RHO GTPase Effectors (Reactome)
MYC Targets v1 (HALLMARK)
Mitotic G2-G2/M Phases (Reactome)
Mitotic Prometaphase (Reactome)
Neg. Epigenetic Reg. of rRNA Expr. (Reactome)
tRNA Processing (Reactome)
RHO GTPases Activate WASPs & WAVEs (Reactome)
Reg. of PLK1 Activity at G2/M Transition (Reactome)
Folding of Actin by CCT/TriC (Reactome)
RHO GTPases Activate Formins (Reactome)
Epigenetic Reg. of Gene Expression (Reactome)
Recruitment of Centrosome Complexes (Reactome)
Nuclear Envelope Breakdown (Reactome)
Focal Adhesion (KEGG)
Cooperation of Prefoldin & CCT/TriC in Actin & Tubulin Folding (Reactome)
Adj. p
0.00005
0.00010
0.00015
0.00020
0.00025
Gene Ratio
0.025
0.050
0.075
0.100
0.125
Count


A
GDS
168 Genes
scRNA-seq
Expression
single cells
GDS genes
expression
single cells
single cells
1
0.5
0
Jaccard Index
modules
deg
log2FC
modules
modules
0
0.5
1
Jaccard Index
Vulnerability
States (VS)
B
Cells
Jaccard
Similarity
0
0.2
0.4
0.6
C
Vulnerability State 1
VS3
VS2
Modules
Modules
5
10
15
20
25
Tumor ID
MGH101
MGH104
MGH105
MGH106
MGH115
MGH122
MGH124
MGH125
MGH129
MGH143
MGH66
Jaccard
Similarity
0.5
0.4
0.3
0.2
0.1
0


A
VS1
VS2
VS3
MYT1L
PTPRN
EPHB6
CLDN18
ANK3
RPS29
PLP1
RPL27
RPL34
MYRF
RPS5
RPS13
AURKA
CDK1
CDC20
AURKB
EZH2
BARD1
Expr.
4
2
0
−2
5
10
15
20
Modules
B
PAX4 TF Targets (PAX4 03)
PDX1 TF Targets (IPF1 Q4)
SOX5 TF Targets (SOX5 01)
miR-4330 Targets (MIR4330)
REST TF Targets (NRSF 01)
Ribosome (KEGG)
Cytoplasmic Ribosomal Proteins (WP)
Eukaryotic Translation Elongation (Reactome)
Response of EIF2AK4 to Amino Acid Deficiency (Reactome)
SRP-Dependent Cotranslational Protein Targeting to Membrane (Reactome)
Neocortex Basal Radial Glia DN (Florio)
G2/M Cell Cycle (Fischer)
YBX1 Targets UP (Basaki)
E2F Targets (Hallmark)
Pediatric Cancer Markers (Whiteford)
Adj. p
0.01
0.02
0.03
0.04
Gene Ratio
0.2
0.4
0.6
VS1
VS2
VS3
C
Non-Neo
VS1
VS2
VS3
Expr.
DCTN2
WDHD1
CDK2
VRK1
NOP2
CCT6A
RRN3
CEBPZ
CDC42
RPS21
GPX4
NHP2
PDCD5
RPP25L
POLR1E
CCT3
NAA50
FARSA
LMNA
RPS23
NELFCD
PSMB7
RCC1
PWP2
PFDN5
RPL17
ACTL6A
ARPC3
SNRPB2
UTP18
ITGAV
TRMT5
D
GDS Module Score
Anova, p < 0.0001
Neftel
Johnson
In-House
VS1
VS2
VS3
Non−Neo


A
Neftel
Percent of Neoplastic Cells
1.00
0.75
0.50
0.25
0.00
MGH100
MGH101
MGH102
MGH104
MGH105
MGH106
MGH110
MGH113
MGH115
MGH121
MGH122
MGH124
MGH125
MGH128
MGH129
MGH136
MGH143
MGH151
MGH152
MGH66
Johnson
In-House
SM006
SM011
SM012
SM017
SM018
GBM21
GBM41
GBM47
GBM49
GBM51
GBM53
VS1
VS2
VS3
B
VS1 DEG
VS2 DEG
VS3 DEG
Neftel
In-House
Johnson
1479
730
3180
1955
751
4333
3369
454
107
1796
640
524
847
1398
99
30
2835
250
255
2956
4908
C
Regional OPC (Ravi)
Classical (Verhaak)
AC (Neftel)
Radial Glia (Ravi)
MES2 (Neftel)
Reactive Hypoxia (Ravi)
Reactive Immune (Ravi)
MES1 (Neftel)
Mesenchymal (Verhaak)
VS1
NPC2 (Neftel)
OPC (Neftel)
NPC1 (Neftel)
Proneural (Verhaak)
VS2
VS3
Neural (Verhaak)
Regional NPC (Ravi)
Glial
Reactive
Stem-Like
Neuronal
Correlation
1
0
−1


A
Deconvoluted Cell Type Proportion
1.00
0.75
0.50
0.25
0.00
UKF242_T_ST
UKF243_T_ST
UKF248_T_ST
UKF251_T_ST
UKF255_T_ST
UKF256_TC_ST
UKF256_TI_ST
UKF259_T_ST
UKF260_T_ST
UKF262_T_ST
UKF265_T_ST
UKF266_T_ST
UKF269_T_ST
UKF275_T_ST
UKF296_T_ST
UKF304_T_ST
UKF313_T_ST
UKF334_T_ST
Tumor Slice ID
VS1
VS2
VS3
Non-Neoplastic
B
Establishmet of Protein Localization (GOBP)
Regulation of Trans-Synaptic Signaling (GOBP)
3
0
−3
Z-Score
C
Neuron Development (GOBP)
R = 0.58, p < 2.2e−16
VS1 Enrichment
0.1
0.0
−0.1
−0.2
−0.3
−0.2
−0.1
0.0
5
0
−5
Z-Score
D
Positive Regulation of Cell Population Proliferation (GOBP)
R = 0.71, p < 2.2e−16
VS2 Enrichment
0.1
0.0
−0.1
−0.2
−0.3
−0.20
−0.15
−0.10
−0.05
3
0
−3
Z-Score
E
Microtubule Based Process (GOBP)
R = 0.71, p < 2.2e−16
VS3 Enrichment
0.2
0.0
−0.2
−0.4
−0.25
−0.20
−0.15
−0.10
−0.05
Pathway Enrichment
5
0
−5
Z-Score