

Supplemental Materials:

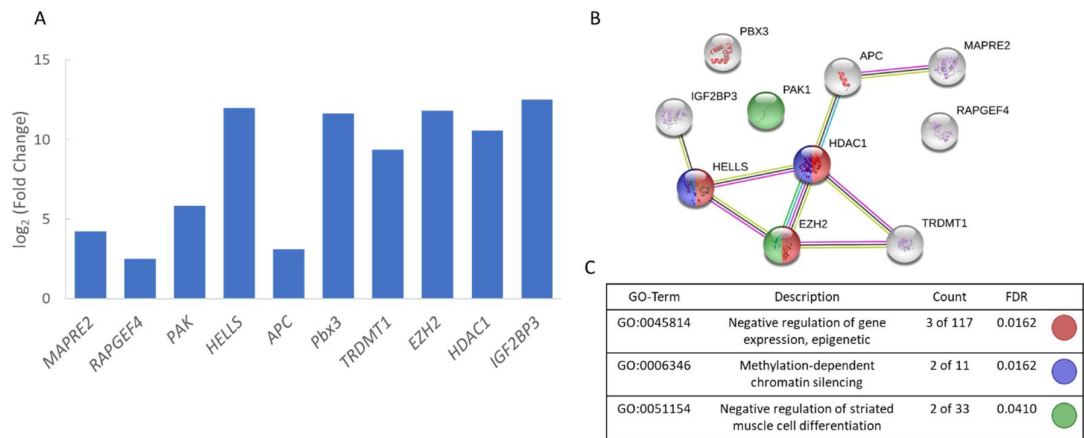


Figure S1. Epigenetic regulators are overexpressed in ATRT relative to normal brain. (A) Log₂ fold changes of the ten most differentially expressed genes (lowest FDR) in frozen ATRT tissue (*n* = 17) relative to normal brain tissue (*n* = 7). Gene expression is measured using NanoString technology with a previously published custom gene panel [1]. (B) STRING 11.0 protein-protein interaction network of ten genes shows high connectivity relative to a random protein set (*p* = 0.000175) (C) All significant results from gene-set enrichment analysis in STRING (FDR < 0.05). HDAC1, EZH2, and HELLS are implicated in GO terms associated with epigenetic regulation [2].

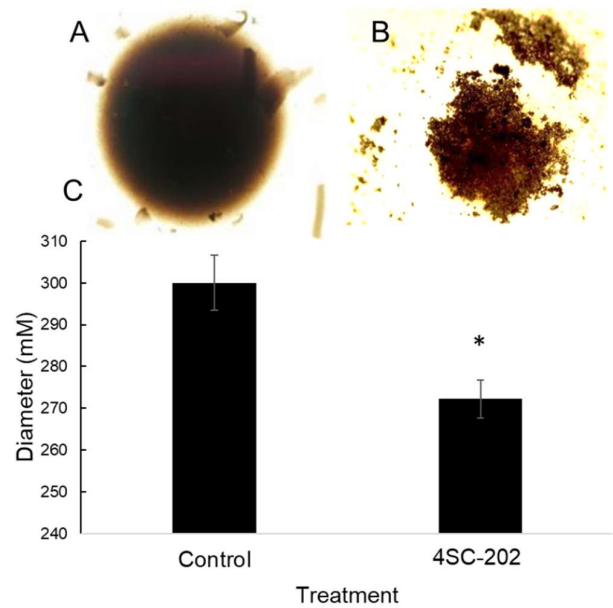


Figure S2. 4SC-202 reduces growth of 3-dimensional ATRT-06 spheroids. (A) Spheroids treated with DMSO control (0.02%) are more defined and have significantly increased diameter compared to (B) 4SC-202-treated (56 nM) spheroids. (C) Diameters of control and treated spheroids. Spheroids were imaged using the Olympus (BX51) and diameters quantified using cellSens (Standard).

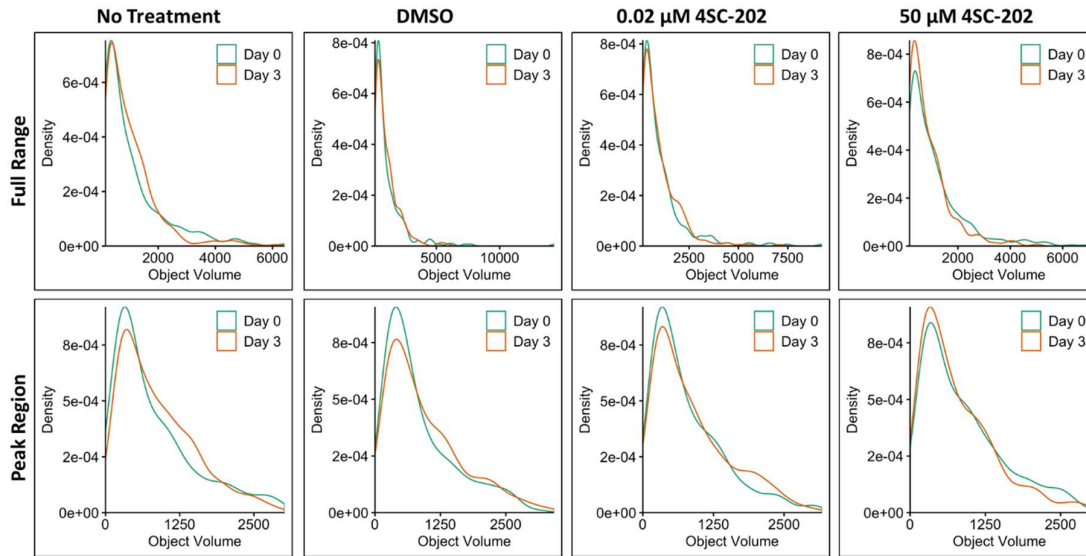


Figure S3. Density plots of cell size from segmented confocal images of ATRT scaffold model. In each of the experimental conditions, the most common segmented volume is between 400–500 μm^3 . Based on the approximate cell volume, it is likely that the peak for the absolute maxima represents the population of non-clustered cells. Experimental conditions with high populations of non-clustered cells would also be more likely to have fewer, well defined clusters, which may be represented by the clear substructures in the density plot. In contrast, experimental conditions with a lower population of non-clustered cells and broader peaks are likely to represent models that have higher levels of clustering. The control experimental conditions show a shortening of the central peak from day 0 (blue green) to day 3 (orange) and an increase in the size of the peaks at larger volumes, suggesting that there is an increase in the number and size of cell clusters. The density plot for the 0.02 μM 4SC-202-treated scaffold likewise has a shorter peak on day 3 relative to day 0, but the substructures are less pronounced excepting the peak at around 2000 μm^3 . In contrast, for the 50 μM treated scaffold, the central peak is higher in day 3 relative to day 0, suggesting that there is a higher percentage of cells that are isolated or in very small clusters on day 3 relative to the control groups.

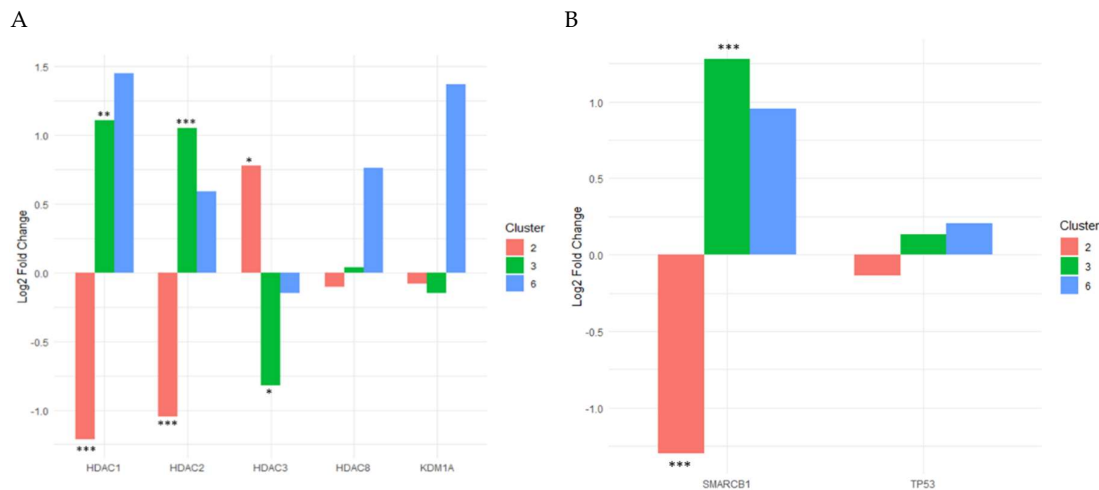


Figure S4. 4SC-202 significantly reduces *HDAC1*, *HDAC2*, and *SMARCB1* gene expression. (A) Log₂ fold change of the average expression and significance of Class I HDACs and *KDM1A* (LSD1). (B) Log₂ fold change of the average expression and significance of *SMARCB1* (SNF5/INI1/BAF47) and *TP53* (P53). Log fold changes and significance were exported from Loupe Cell Browser and are based off of a locally distinguishing significant feature comparison of clusters 2, 3, and 6. p value < 0.001 (***), p value < 0.01 (**), p value < 0.1 (*).

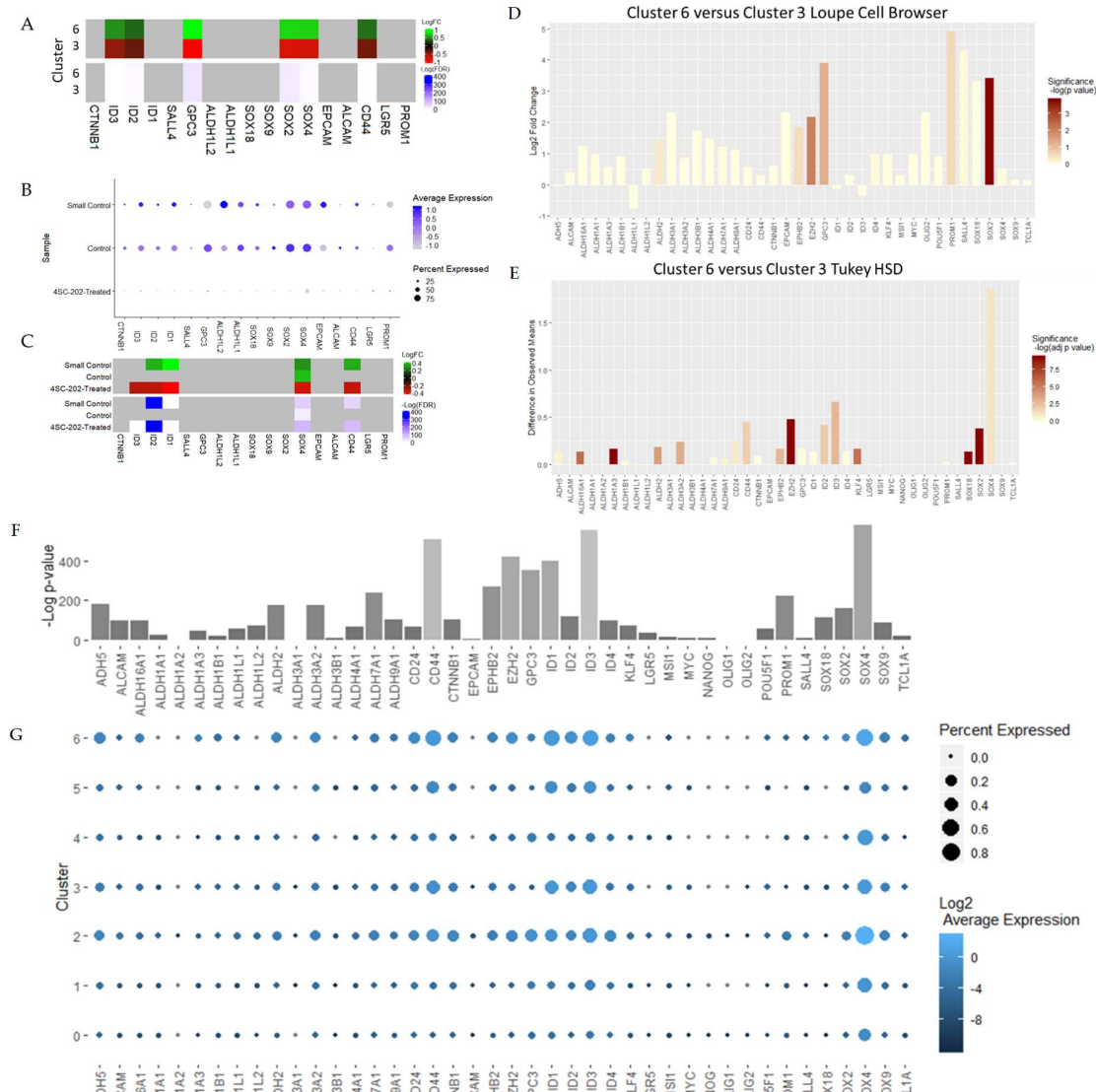


Figure S5. Stem cell markers are underexpressed in 4SC-202-treated spheroids and overexpressed in cluster 6. (A) Log₂ fold change (LFC, top heatmap) and significance of stem cell marker expression (bottom heatmap) between cluster 6 and cluster 3. Both clusters are principally composed of control cells that have a low percentage of mtRNA. Significance is calculated using the Wilcoxon Rank Sum test in Seurat ($p < 0.01$, $|LFC| > 0.25$) [3,4]. Green indicates positive log fold expression changes and red indicates negative log fold expression changes. Non-significant changes are grey. Significance is shown as the negative log of the FDR. No significance is grey, low significance is white, and high significance is blue. Heatmaps are generated with the ComplexHeatmap package in R [5]. (B) Dot plot visualization of the average expression and percentage of cells expressing stem cell markers or stem cell-related genes across samples. Color of the dot indicates the average expression of the feature and the size indicates what percentage of the cells express the feature. (C) Log₂ fold change and significance of stem cell marker expression between experimental conditions. Heatmaps were generated as described above. (D) Log₂ fold change and significance of stem cell-related genes in cluster 6 relative to cluster 3 as calculated in Loupe Cell Browser using locally distinguishing significant feature comparison tool. Color indicates the significance where high significance is red and low significance is tan. (E) Difference in observed means and significance of stem cell-related genes based on Tukey's Honest Significant Difference Test between clusters 6 and 3 following an ANOVA of all clusters (see F-G). Color indicates the significance where high significance is red and low

significance is tan. (F) Barplot of the negative log of the significance of an ANOVA test for each stem cell-related gene. Data were scaled by a size factor that was calculated as the total UMI count per cell divided by the median UMI sum across all cells prior to testing. (G) Dotplot of average scaled expression and percentage of cells expressing a gene across all clusters. Cells with a gene expression above 0 were considered to express a gene. Color indicates the log₂ transformed average expression, where light blue indicates a higher average expression.

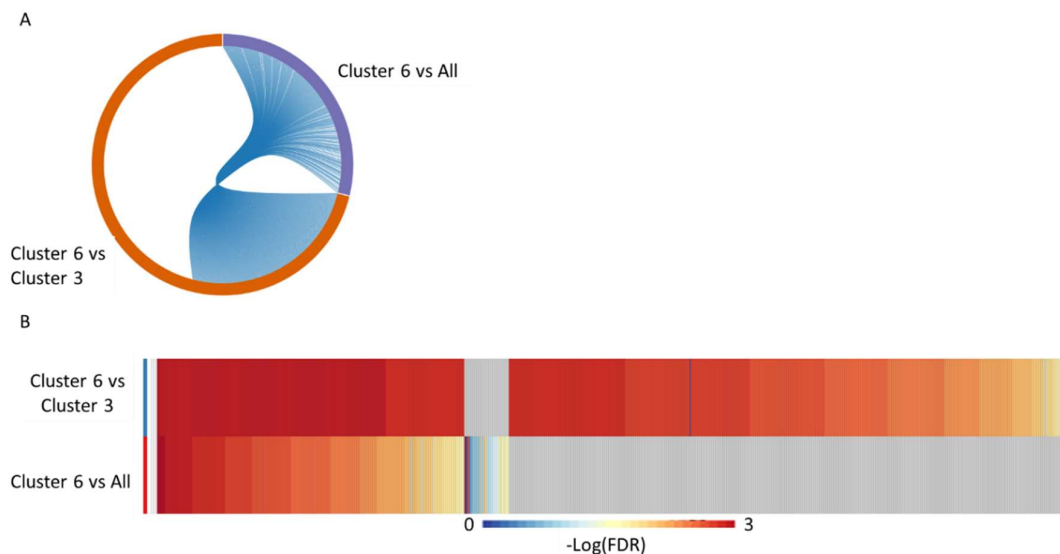


Figure S6. Highly differentially expressed genes are stable across comparisons. (A) Circos diagram of the overlap in the differentially expressed genes when comparing Cluster 6 to all cells (control and treated) or to just Cluster 3 (principally control cells). Differentially expressed genes are considered genes that have an FDR < 0.05. (B) Heatmap of the $-\log(\text{FDR})$. Heatmap and Circos diagram were generated from differentially expressed genes using NetworkAnalyst [6]. Because Cluster 3 and Cluster 6 are both principally composed of control spheroids, the comparison of these two clusters is close to a comparison of the two populations within the healthy control. Cluster 6 versus all takes into consideration the effect of the drug and includes a comparison with the drug-treated population.

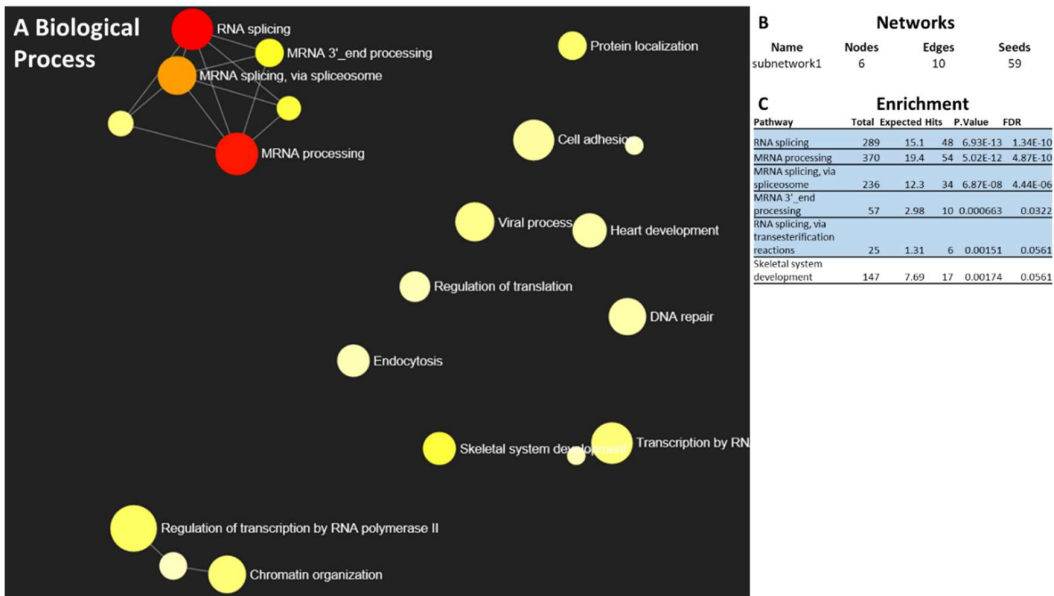


Figure S7. Enriched biological process network. (A) Network of enriched biological process terms in the differentially expressed gene list for the cluster 2 (4SC-202-treated) versus cluster 3 (control) comparison of the single-cell RNA-seq dataset. Color is related to the p-value from the over-representation analysis (ORA) with red being more significant and tan being less significant ($p < 0.05$). Size is directly related to the number of differentially expressed genes mapped to each term. Connections between terms indicate overlap in the gene list mapped to each term. (B) List of subnetworks in the biological process network. (C) Enriched terms and their significance ($FDR < 0.05$). Terms related to mRNA processing are highlighted blue.

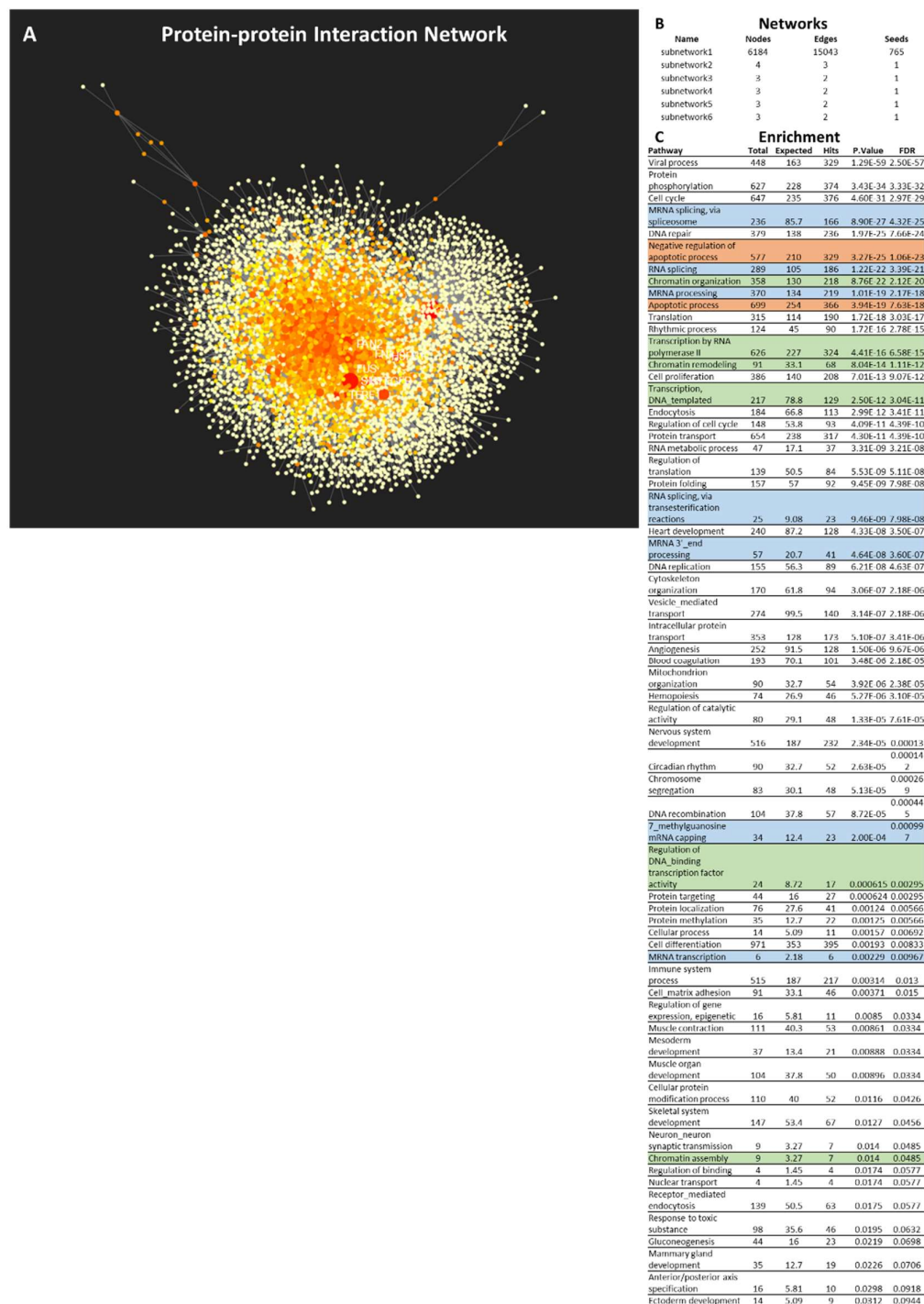


Figure S8. Cerebellum protein-protein interaction (ppi) network. (A) Network of protein-protein interactions initiated from differentially expressed genes (DEGs) in cluster 2 (4SC-202-treated) relative to cluster 3 (control) of the single-cell RNA-seq dataset. Only DEGs and proteins that directly interact with DEG seeds are included. The size of the node indicates the degree value and the color indicates the betweenness centrality values with red indicating higher centrality and tan indicating lower centrality. (B) List of subnetworks visualized in the ppi network. (C) Enriched terms and their significance ($FDR < 0.05$). Terms related to mRNA processing are highlighted blue, terms related to transcription regulation are highlighted green, and terms related to apoptosis are highlighted orange.

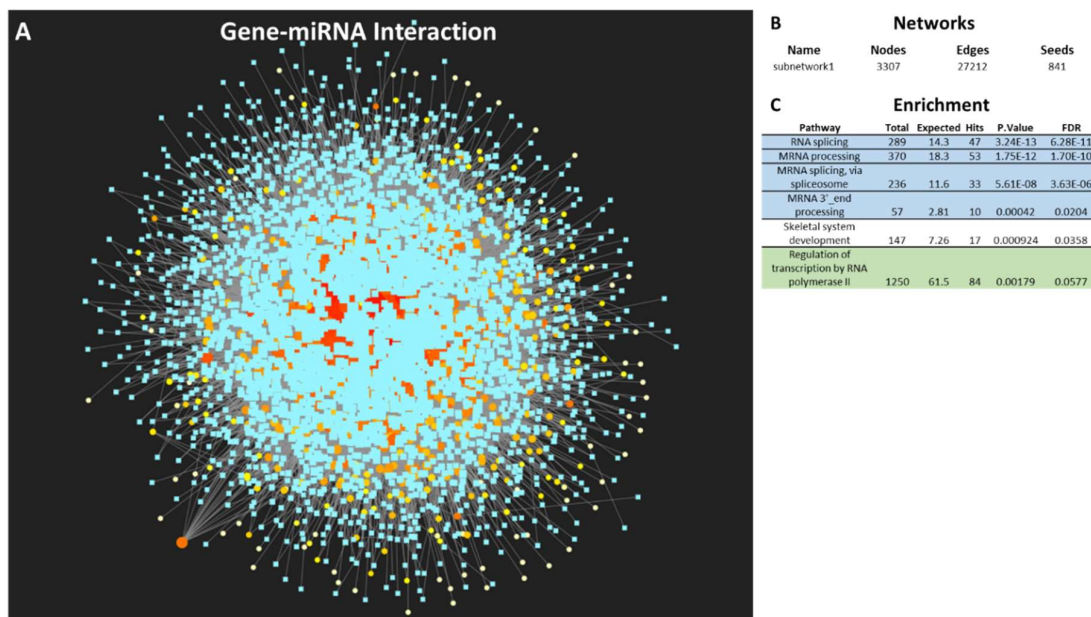


Figure S9. Gene-miRNA interaction network. **(A)** Network of gene-miRNA interactions seeded from differentially expressed genes in cluster 2 (4SC-202-treated) relative to cluster 3 (control) of the single-cell RNA-seq dataset (DEGs). Only miRNAs that are directly related to DEGs are included (blue squares). The size of DEGs (circles) indicate the degree value and the color indicates the betweenness centrality values with red indicating higher centrality and tan indicating lower centrality. **(B)** Subnetwork details for only subnetwork in the gene-miRNA interaction network. **(C)** Enriched terms and their significance (FDR < 0.05). Terms related to mRNA processing are highlighted blue and terms related to transcription regulation are highlighted green.

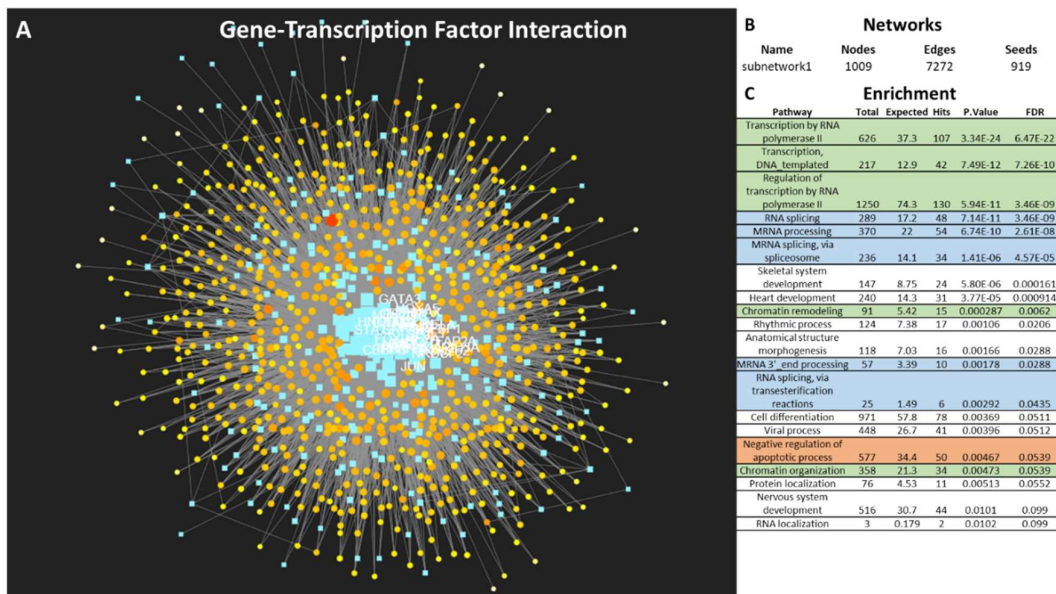


Figure S10. Gene-transcription factor interaction network. (A) Network of gene-transcription factor interactions seeded from differentially expressed genes (DEGs) in cluster 2 (4SC-202-treated) relative to cluster 3 (control) of the single cell RNA-seq dataset. Only DEGs (color circles) and transcription factors (blue squares) that directly interact with them are included in the network. Circle size is determined by the degree value and circle color indicates the betweenness centrality values with red indicating higher centrality and tan lower centrality. (B) Subnetwork details for only subnetwork in the gene-transcription factor interaction network. (C) Enriched terms and their significance (FDR < 0.05). Terms related to mRNA processing are highlighted blue, terms related to transcription regulation are highlighted green, and terms related to apoptosis are highlighted orange.

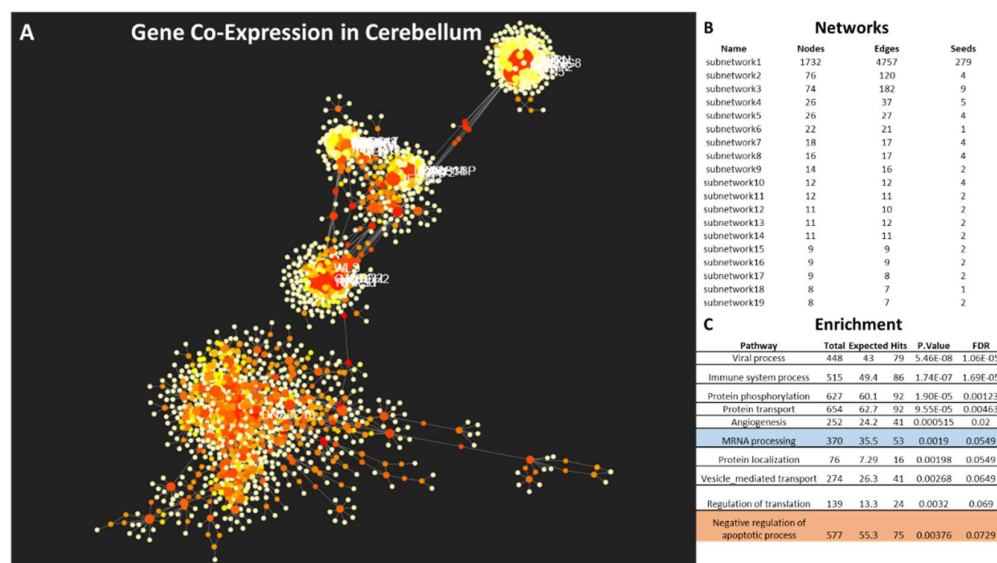


Figure S11. Cerebellum gene co-expression network. (A) Network of co-expressed genes seeded Figure 2. (4SC-202-treated) relative to cluster 3 (control) in the single-cell RNA-seq dataset. Node size is determined by degree value and node color by betweenness centrality. Red nodes are more central than are lighter colored nodes. (B) Details of subnetworks in the gene co-expression network. (C) Enriched terms and their significance (FDR < 0.05). Terms related to mRNA processing are highlighted blue and terms related to apoptosis are highlighted orange.

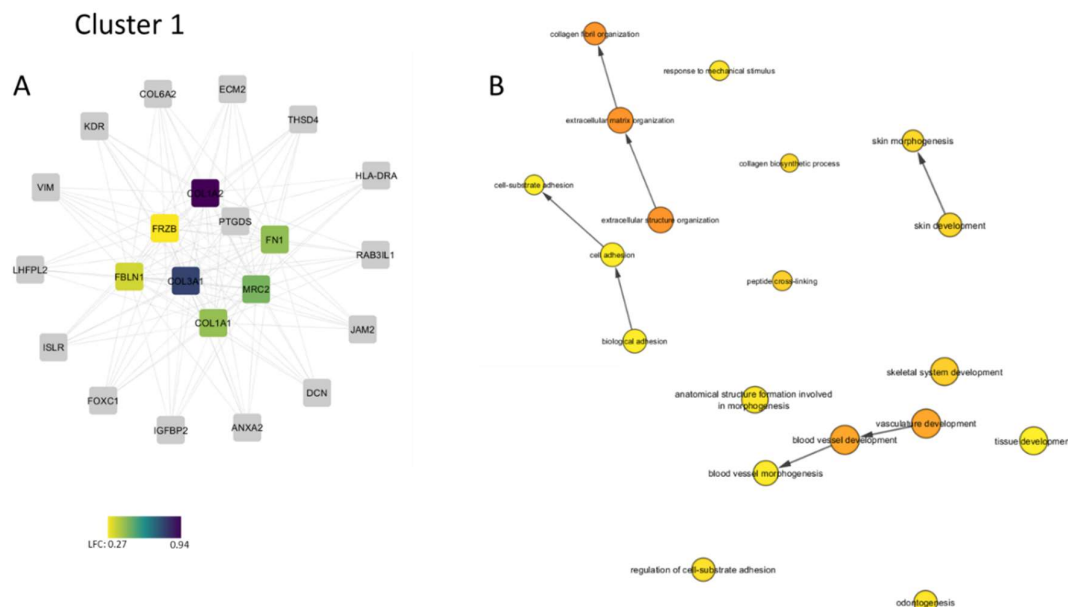


Figure S12. Densely connected subregion (cluster) 1 and biological process enrichment from integrated systems biology analysis **(A)** Network of bioelements and their interactions. The MCODE score for this cluster is 11.905 based on the size and density of the subgraph. Lines are unweighted and indicate interactions in the integrated systems biology network (Figure 7). Color is based on the log fold change of the gene expression between cluster 2 and cluster 3 of the single cell dataset. Grey squares indicate genes that were not differentially expressed. **(B)** Overrepresented GO terms for the densely connected subgraph ($p < 0.05$) and their ontological hierarchy (arrows). Coloring of GO nodes indicates significance where red is high and yellow is low and the size of the circles indicate how many differentially expressed genes were mapped to each term.

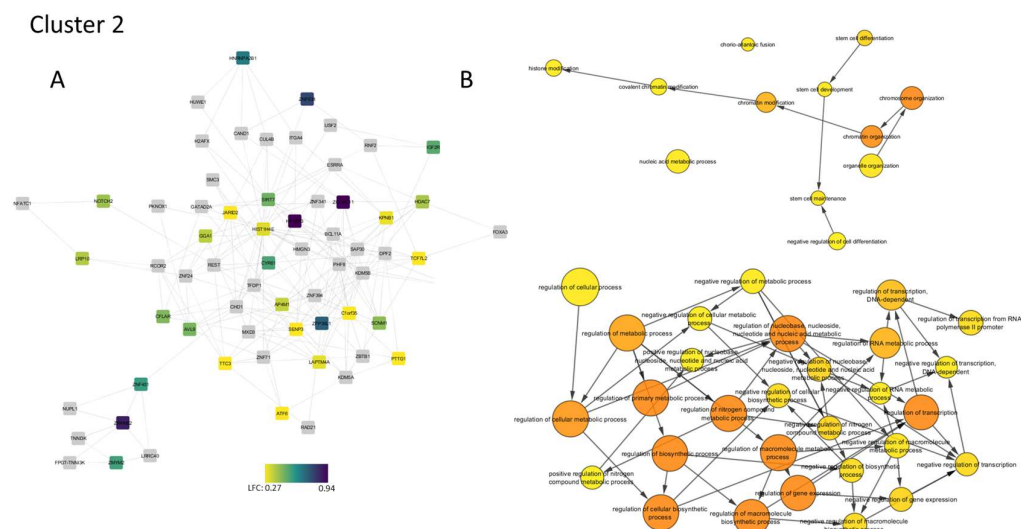


Figure S13. Densely connected subregion (cluster) 2 and biological process enrichment from integrated systems biology analysis **(A)** Network of bioelements and their interactions. The MCODE score for this cluster is 6.469 based on the size and density of the subgraph. Lines are unweighted and indicate interactions in the integrated systems biology network (Figure 7). Color is based on the log fold change of the gene expression between cluster 2 and cluster 3 of the single cell dataset. Grey squares indicate genes that were not differentially expressed. **(B)** Overrepresented GO terms for the densely connected subgraph ($p < 0.05$) and their ontological hierarchy (arrows). Coloring of GO nodes indicates significance where red is high and yellow is low and the size of the circles indicate how many differentially expressed genes were mapped to each term.

[illegible]

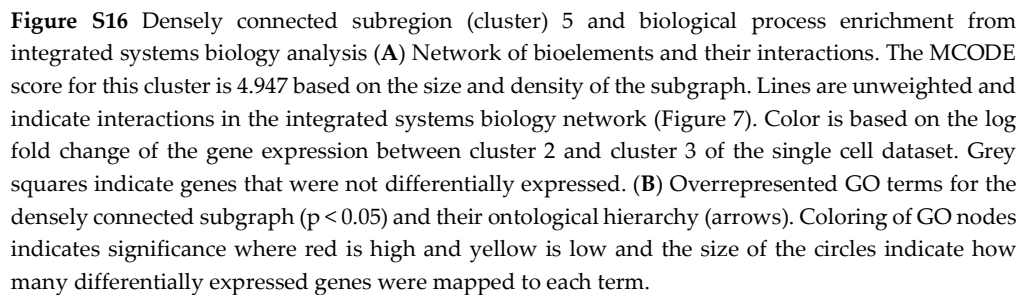
Cluster 4

A

LFC: 0.27 0.94

B

Figure S15. Densely connected subregion (cluster) 4 and biological process enrichment from integrated systems biology analysis **(A)** Network of bioelements and their interactions. The MCODE score for this cluster is 5.625 based on the size and density of the subgraph. Lines are unweighted and indicate interactions in the integrated systems biology network (Figure 7). Color is based on the log fold change of the gene expression between cluster 2 and cluster 3 of the single cell dataset. Grey squares indicate genes that were not differentially expressed. **(B)** Overrepresented GO terms for the



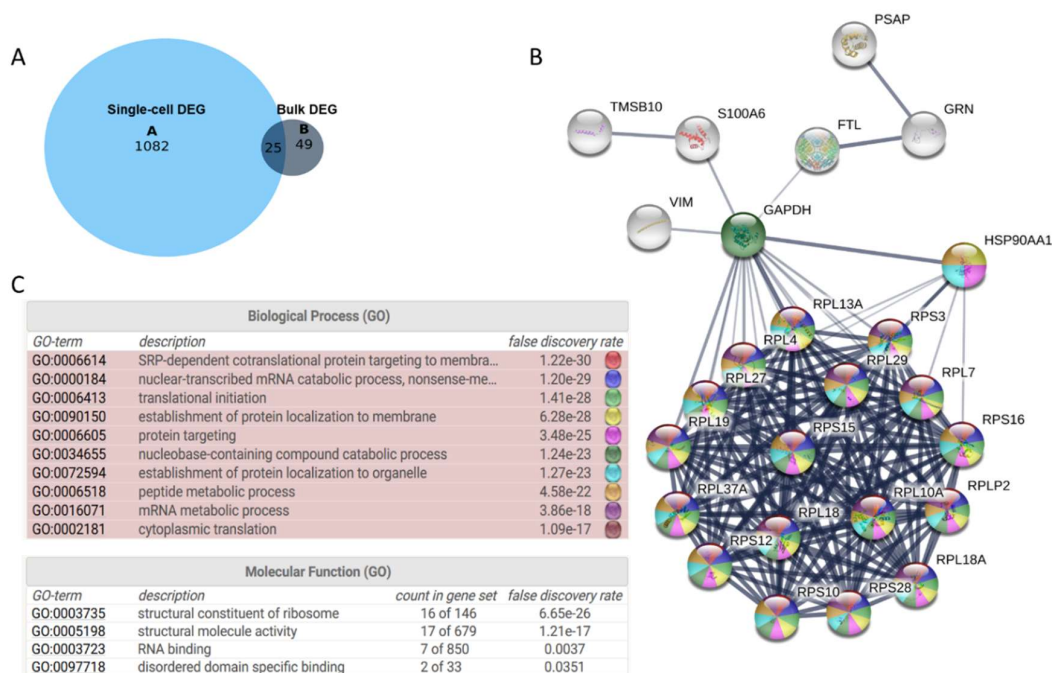


Figure S17. Genes differentially expressed between 4SC-202-treated and control spheroids in both single-cell and bulk RNA-sequencing datasets are enriched for ribosome-related GO terms. **(A)** Venn diagram of differentially expressed genes for bulk and single cell sequencing results (FDR < 0.05, |LFC| > 0.25, LFC restriction imposed for single cell dataset only). **(B)** STRING 11.0 diagram of 25 genes differentially expressed in both the single cell RNA-sequencing and the bulk RNA-sequencing datasets [2]. Protein nodes are colored by the enriched GO biological process terms they are involved in and edge thickness indicates the strength of the evidence of the interaction between proteins. **(C)** Enriched GO biological process terms and GO molecular function terms. Color of biological process terms corresponds to color of protein nodes. FDR < 0.05. Only the first 10 terms are shown of the enriched GO biological process terms. 16 of the 25 terms are mapped to structural constituent of ribosome, and enriched biological process terms include terms related to translation, translocation of proteins, and mRNA catabolic processes.

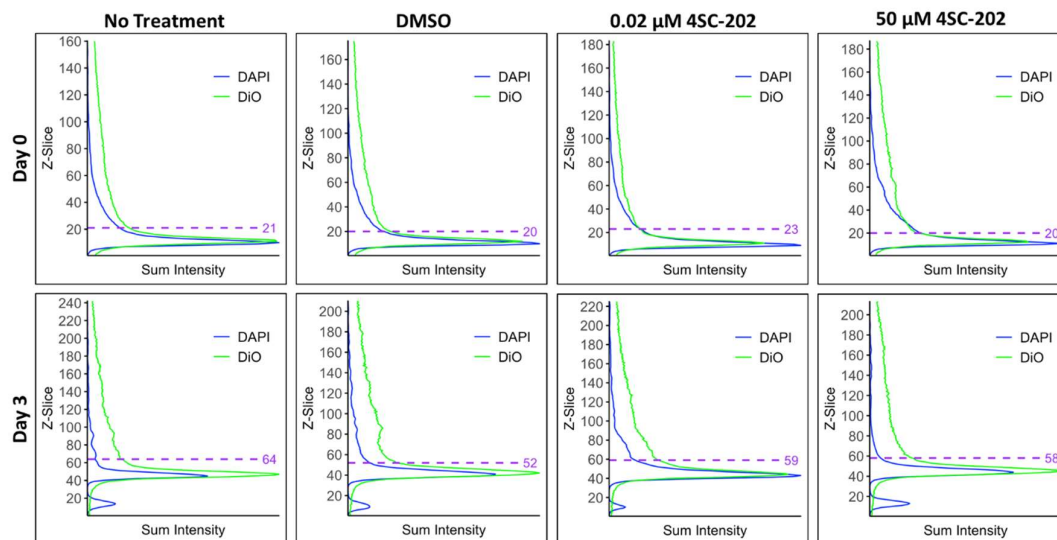


Figure S18. Cropping location relative to the total intensity along the z-axis. Each slice's pixel intensities were summed into two values per slice: one for sum DAPI intensity, and one for sum DiO intensity. The blue line indicates DAPI channel slice intensities, and the green line indicates DiO channel slice intensities. Dashed purple lines indicate crop points along the z-stacks. Slice numbers are indicated to the right of each dashed line. For image pre-processing, slices less than the indicated value were removed.

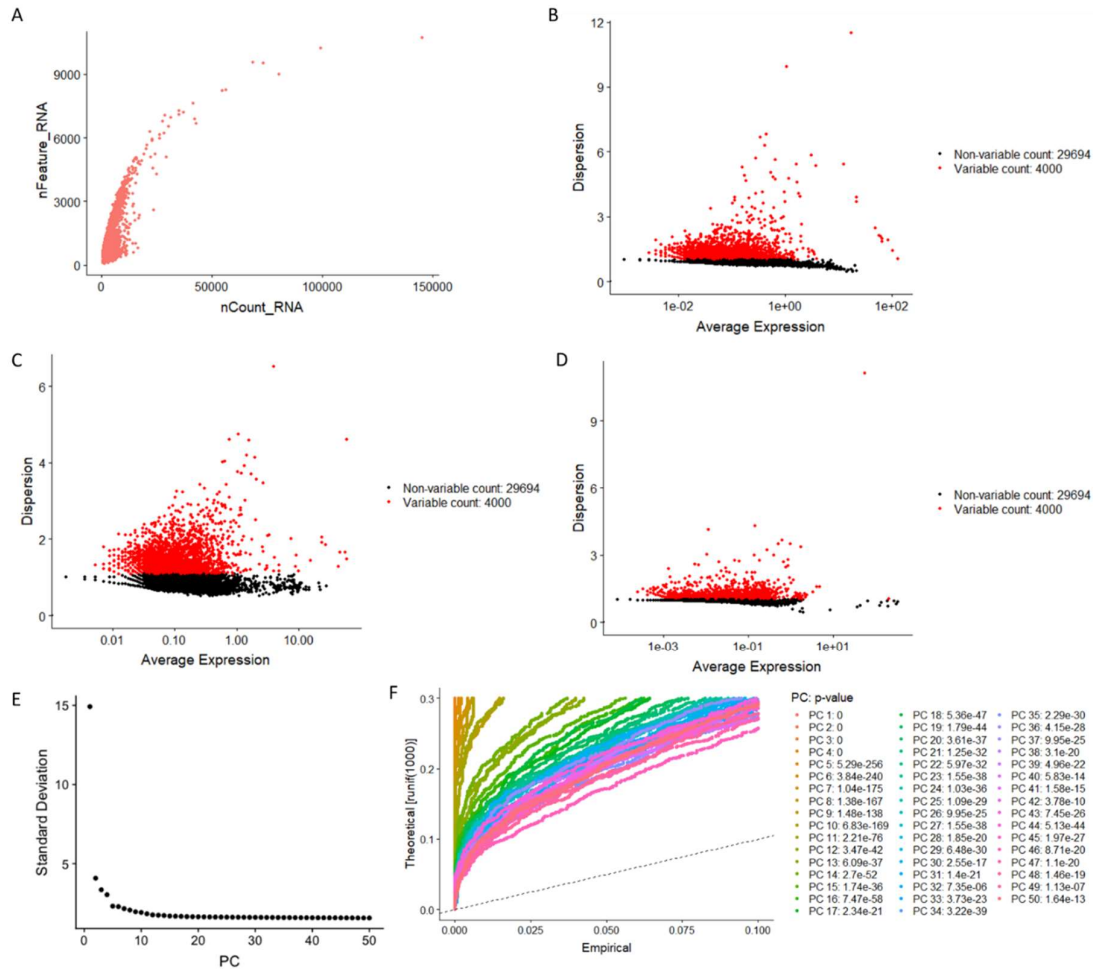


Figure S19. Parameter selection plots for single cell analysis in Seurat [3]. (A) Number of features versus number of RNA molecule plot for outlier identification. Cells with less than 40,000 counts of RNA were kept for analysis. (B–D) Variable feature plots of 4000 retained features for (B) small control spheroid, (C) control spheroid, and (D) 4SC-202-treated spheroid. (E) Elbow plot showing the standard deviations of the first 50 principle components (PCs). (F) Jack Straw Analysis and p values for the first 50 PCs. Jack Straw Q-Q plot shows distribution of p values for each PC relative to a theoretical uniform distribution (dashed line). P values are based on a proportion test of the number of significant p values in the PC compared to the theoretical uniform distribution.

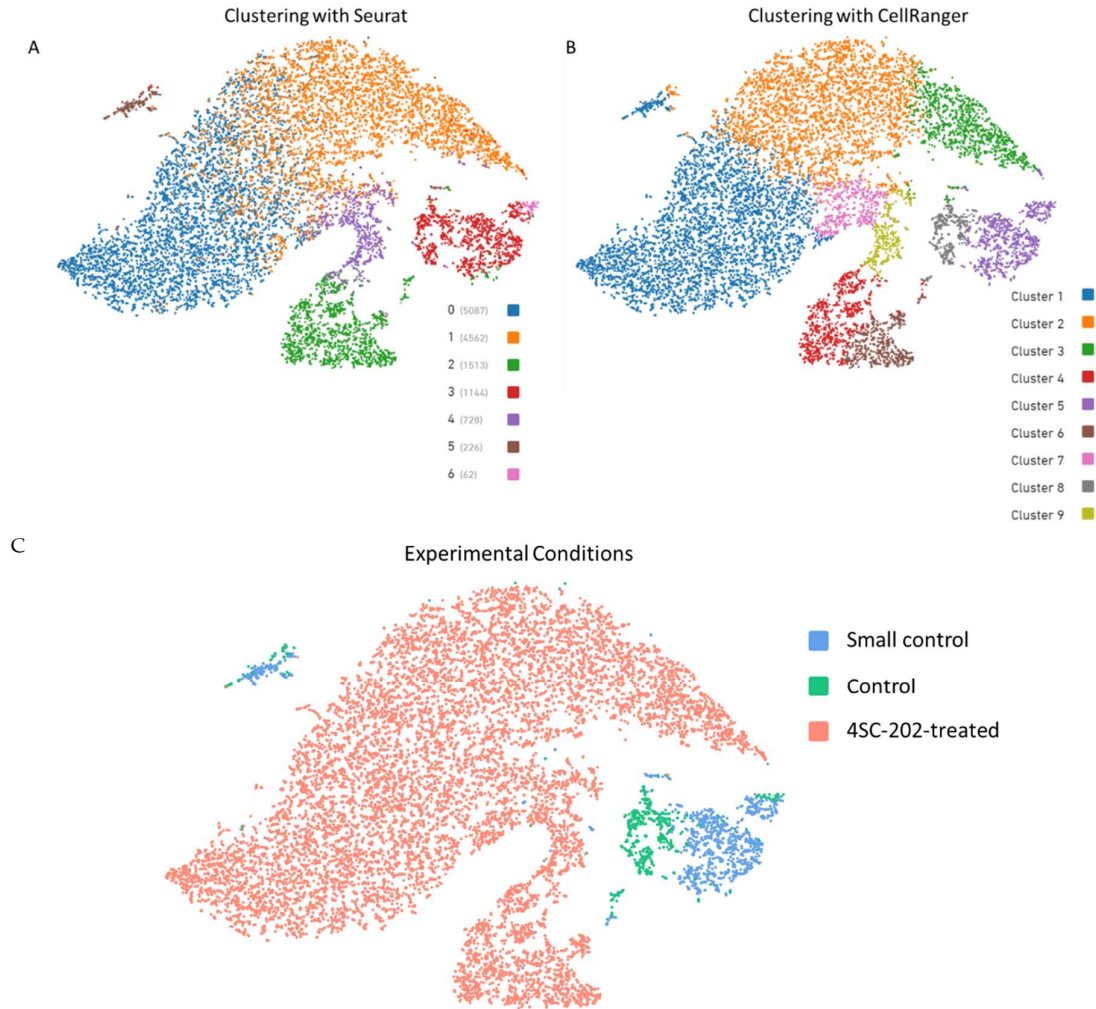


Figure S20. Seurat graph-based clusters correlate with clusters identified using the CellRanger pipeline. **(A)** 10× Genomics Loupe Cell Browser TSNE plot of scRNA-seq data with colors indicating the clusters calculated using the Seurat algorithm **(B)** Loupe Cell Browser TSNE plot of scRNA-seq data with colors indicating the clusters identified using the CellRanger pipeline. While Seurat and CellRanger clusters identify different numbers of clusters, the CellRanger clusters tend to be sub-clusters of the Seurat clusters. One notable exception is Cluster 6, which appears in the Seurat clustering but not in the CellRanger clustering. Cluster 6 does occupy a distinct area in the TSNE plot calculated by the CellRanger pipeline. **(C)** Loupe Cell Browser TSNE plot of experimental conditions.

References:

1. Messerli, S.M.; Hoffman, M.M.; Gnimpieba, E.Z.; Kohlhof, H.; Bhardwaj, R.D. 4SC-202 as a potential treatment for the pediatric brain tumor medulloblastoma. *Brain Sci.* **2017**, doi:10.3390/brainsci7110147
2. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613.
3. Stuart, T.; Butler, A.; Hoffman, P.; Hafemeister, C.; Papalexi, E.; Mauck, W.M.; Hao, Y.; Stoeckius, M.; Smibert, P.; Satija, R. Comprehensive Integration of Single-Cell Data. *Cell* **2019**, *177*, 1888–1902.e21.
4. Butler, A.; Hoffman, P.; Smibert, P.; Papalexi, E.; Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **2018**, *36*, 411–420.
5. Gu, Z.; Eils, R.; Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **2016**, *32*, 2847–2849.
6. Zhou, G.; Soufan, O.; Ewald, J.; Hancock, R.E.W.; Basu, N.; Xia, J. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res.* **2019**, *47*, W234–W241.