



Oct 22, 2025

Comprehensive transcriptomic analysis of brainstem-regionalized organoids and associated diffuse midline glioma organoids

 [Nature Cancer](#)

DOI

<https://dx.doi.org/10.17504/protocols.io.6qpvrwq13lmk/v1>

Rijndert Ariese^{1,2}, Noëlle ommann^{3,4}, Cristian Ruiz Moreno^{1,5}, Christian Mayer⁶, Hendrik G. Stunnenberg¹, Anna Alemany^{3,4}, Anne C. Rios^{1,2,7}

¹Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands;

²Oncode Institute, Utrecht, the Netherlands;

³Department of Anatomy and Embryology, Leiden University Medical Center, Leiden, the Netherlands;

⁴The Novo Nordisk Foundation Center for Stem Cell Medicine (reNEW), Leiden node, the Netherlands;

⁵Department of molecular Biology, Faculty of Science, Radboud University, Nijmegen, the Netherlands;

⁶Max Planck Institute of Neurobiology, Martinsried, Germany;

⁷Department of Biology, Faculty of Science, Utrecht University, Utrecht, the Netherlands



Rijndert Ariese

Create & collaborate more with a free account

Edit and publish protocols, collaborate in communities, share insights through comments, and track progress with run records.

Create free account

OPEN  ACCESS





DOI: <https://dx.doi.org/10.17504/protocols.io.6qpvrwq13lmk/v1>

External link: <https://doi.org/10.1038/s43018-025-01084-0>

Protocol Citation: Rijndert Ariese, Noëlle ommann, Cristian Ruiz Moreno, Christian Mayer, Hendrik G. Stunnenberg, Anna Alemany, Anne C. Rios 2025. Comprehensive transcriptomic analysis of brainstem-regionalized organoids and associated diffuse midline glioma organoids. **protocols.io** <https://dx.doi.org/10.17504/protocols.io.6qpvrwq13lmk/v1>

Manuscript citation:

Bessler, N., Wezenaar, A.K.L., Ariese, H.C.R. *et al.* De novo H3.3K27M-altered diffuse midline glioma in human brainstem organoids to dissect GD2 CAR T cell function. *Nat Cancer* (2026). <https://doi.org/10.1038/s43018-025-01084-0>

License: This is an open access protocol distributed under the terms of the **[Creative Commons Attribution License](#)**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Protocol status: Working

We use this protocol and it's working

Created: October 21, 2025

Last Modified: October 22, 2025

Protocol Integer ID: 230382

Keywords: sequencing, brain organoids, tumor modelling, lineage tracing, midline glioma organoids this protocol detail, midline glioma organoid, corresponding diffuse midline glioma, comprehensive transcriptomic analysis of brainstem, comprehensive transcriptomic analysis, regionalized organoid, transcriptomic analysis, underlying tumor growth, novo dmgl tumor, tumor growth within dmgo, bulk rna, robust transcriptional framework, fgf4, tumor support, comparative effects of fgf4, bearing organoid, rna, pontine regionalization, clonal expansion, fgf2 in hindbrain

Abstract

This protocol details the transcriptomic analyses conducted to characterize a newly developed brainstem-regionalized organoid (BrO) and its corresponding diffuse midline glioma-bearing organoid (DMGO). Bulk RNA sequencing was used to evaluate media optimization and protocol standardization, focusing on the comparative effects of FGF4 versus FGF2 in hindbrain-pontine regionalization and assessing potential batch variability. Single-cell (sc) RNA sequencing provided a comprehensive cellular and molecular characterization of the BrO and the de novo DMG tumors arising within it, complemented by TrackerSeq-based genetic lineage tracing to investigate mechanisms of clonal expansion and tumor support. The protocol is expected to yield a robust transcriptional framework describing BrO patterning and DMG development, validate FGF4 as a key driver of pontine identity, and reveal pontine-glial lineage-specific dynamics underlying tumor growth within DMGO.

Sample preparation

1 Bulk RNA sequencing

- 1.1 Organoids were collected in 2 ml DNA low-binding tubes (Eppendorf, Cat.#0030108078) and mechanically dissociated in Hank's Balanced Salt Solution (HBSS, Gibco, Cat.#14025092), before spinning down at 800 x g for 5 minutes at 4°C. Supernatant was removed, and the cell pellet snap-frozen on dry ice and stored at -80°C.
- 1.2 Total RNA was extracted using the RNeasy Mini Kit (Qiagen, Cat.#74104) according to manufacturer's instructions. RNA concentration and integrity were evaluated by running 1µl of the total RNA sample on an RNA 6000 pico gel (Agilent, Cat.#5067-1513) using a 2100 Bioanalyzer (Agilent, Cat.#G2939BA).
- 1.3 Samples with an RNA Integrity Number above 8 were processed and sequenced on an Illumina NextSeq500 platform.

2 Single-cell RNA sequencing

- 2.1 **For time-course sequencing in BrO:** Brainstem organoids from the same batch were dissociated at selected timepoints to capture various stages of development. Single-cell suspensions were collected from the following stages: embryoid body (day 5), neural induction (day 11), post-Matrigel embedding (day 14), and neuronal specification/maturation (days 20, 30, 60, 90, and 120). To ensure sufficient cell numbers, multiple organoids were pooled, with 24 organoids used for earlier timepoints and as few as 7 for later stages. For timepoints after Matrigel embedding, Cell Recovery Solution (Corning, 354253) was used to dissolve the Matrigel by incubating 15 minutes at 4°C. Organoids were halved and washed in HBSS without Ca^{2+} and Mg^{2+} . Dissociation was performed using a Neural Tissue Dissociation Kit (Miltenyi Biotec, 130-092-628). Organoids were minced, enzyme mix added, incubated on an orbital shaker at 37°C and resuspended in regular intervals with a P1000 until a single cell suspension was reached.
- 2.2 Cells were filtered through 70 µm and 20 µm pre-separation filters to remove debris. The filtrate was centrifuged and washed by resuspension in HBSS without Ca^{2+} and Mg^{2+} . Cells were counted on the automated Countess Cell Counter (Thermo Fisher Scientific). After pelleting, the cells were resuspended in 1mL of mFreSR cryopreservation medium and frozen at -80°C for 24 hours and transferred to liquid nitrogen for long-term storage.
- 2.3 On the sequencing day samples were thawed and transferred to 10mL of pre-warmed DMEM containing 10% FBS and centrifuged. The cells were washed twice with PBS 5% BSA and filtered through a 40 µm Flowmi cell strainer. Viable cell counts were obtained

using Trypan Blue and cells were resuspended in an appropriate volume to capture 30,000 cells.

- 2.4 Single-cell RNA sequencing libraries were generated using the Chromium Single Cell 3' v4 Library & Gel Bead Kit (10x Genomics) following the manufacturer's protocol.
- 2.5 **For DMGO sequencing:** Cell Recovery Solution (Corning, 354253) was used to dissolve the Matrigel by incubating 15 minutes at 4°C. Organoids were halved and washed in HBSS without Ca²⁺ and Mg²⁺. Dissociation was performed using a Neural Tissue Dissociation Kit (Miltenyi Biotec, 130-092-628). Organoids were minced, enzyme mix added, incubated on an orbital shaker at 37°C and resuspended in regular intervals with a P1000 until a single cell suspension was reached.
- 2.6 Cell suspensions were washed twice with PBS +/- (Mg²⁺/Ca²⁺, 3% FBS) and viability assessed by DAPI staining (1:5,000). Suspensions were filtered using a 40 µm Flowmi cell strainer and FACS sorted on a CytoFLEX SRT Benchtop Cell Sorter (Beckman Coulter): tumour cells were enriched for GFP and DAPI events were excluded.
- 2.7 Each DMGO was stained with a different TotalSeq-A anti-human hashtag (Hashtag A0251-A0265, Biolegend). The hashing oligos were designed to recognize most human cells using a combination of two clones against CD298 and β2 microglobulin. The cell pellet was resuspended in staining buffer (50 µl for 500,000 cells) and 5 µl of human Fc blocking reagent added (FcX Human truStain, Biolegend 422301). After 10 min incubation at 4°C, 1 µl of a unique cell hashing antibody was added to each sample and incubated for 20 min at 4°C and then washed 3x using PBS+0.04%BSA. Samples were pooled and single-cell libraries were prepared with the Chromium Single Cell 3' v3 kit (10x Genomics) per manufacturer's instructions.
- 2.8 **For lineage tracing:** three libraries were generated per sample: gene-expression, hashing and a 10x-barcoded TrackerSeq¹ lineage library. Lineage barcodes were further enriched by two nested PCRs (10 cycles each).
 - PCR1: Read1-Forward 5'-CTACACGACGCTCTTCCGATCT-3' and eGFP-Reverse 5'-CTTCTCGTTGGGGTCTTT-3'; Ta 60 °C, 30 s extension.
 - PCR2: Standard P5-Read1-Forward and 10x "Fun series 70x" Reverse; Ta 60 °C, 30 s extension.
- 2.9 All single cell RNA sequencing libraries were sequenced on the Illumina NovaSeq 6000 platform (2×150 bp).

Time-course analysis in BrO

3 Benchmarking integration of time-course data

- 3.1 Snapseed `annotate_hierarchy()` command was used to establish a multi-level initial annotation for label-aware integration, with default settings with a predefined set of marker genes⁵. Employing 3,000 highly variable genes (HVGs) and timepoint as a batch covariate, the `scib-metrics`⁶ package was used to evaluate integration performance across several methods. For semi-supervised integration approaches the coarse level pre-annotation derived from snapseed was used as initial guidance. `scPoli` was selected on its weighed performance in batch correction and preserving biological variance.
- 3.2 For visualization, PAGA was used to generate a coarse-grained graph representation of the data. This graph informed the construction of the final UMAP embedding.

4 HNOCA and HDBCA reference mapping

- 4.1 The integrated organoid time course dataset was aligned to the 3,000 HVGs of the human neural organoid cell atlas (HNOCA)⁵. For genes not present in the organoid dataset, expression values were filled with zeros. The original `scPoli` model was downloaded from the publication's GitHub repository and loaded into the mapper module of the HNOCA-tools package, v0.1.1. The organoid dataset was projected into the HNOCA space using `map_query()`. A weighted k-nearest neighbor (wkNN) graph was computed with $k = 100$. This neighbor graph facilitated the transfer of cell labels from the HNOCA dataset and the generation of a shared UMAP representation that combined the HNOCA and brainstem organoids datasets. Presence scores were calculated using the `get_presence_scores()` function from the HNOCA-tools package.
- 4.2 The CellRanger-processed dataset of the human developing brain cell atlas (HDBCA)⁷ was downloaded and cells with fewer than 300 detected genes were excluded. The dataset was normalized by total counts per cell, log-transformed and scaled. The HDBCA gene expression space was then intersected with the gene set from the HNOCA. A higher-resolution annotation of the HDBCA⁸ was integrated into the dataset. An `scVI` model was trained on the 2,000 HVGs of the HDBCA with 'donor_id' specified as the batch key. The trained `scVI` model was fine-tuned with `scANVI`, using 'CellClass_Mossi' as cell type labels. The organoid dataset was aligned to the 2,000 HVGs used to train the `scANVI` model. Missing gene expressions in the brainstem organoid dataset were padded with zeros.
The aligned brainstem organoid dataset was loaded into the trained `scANVI` model, and a query model was trained. Cell type labels were then transferred to the brainstem organoid dataset using the `scvi-tools predict()` method.

5 HNOCA comparative abundance and gained presence analysis

- 5.1 The scCODA algorithm⁹, as implemented in the pertpy¹⁰ package v0.9.4, was employed to analyse cell type compositional changes between the brainstem organoid model and HNOCA datasets. Timepoints prior to day 30 were excluded, because of progenitors predominance. The 'publication_id' key was used as the covariate_obs, and 'bio_sample' was specified as the sample_identifier. Cell type annotations were grouped by region. For cell types lacking a regional annotation, the cell type transferred from the HNOCA dataset was used. scCODA was executed iteratively, with each cell type selected once as the reference, using default parameters and the No-U-Turn Sampler (NUTS) for Bayesian inference. A majority voting system identified cell types that were credibly differentially abundant, defined as those deemed significant in more than half of the iterations.
- 5.2 To validate the regional identity of the brainstem organoid model in the context of the HNOCA, presence scores were calculated for both datasets using the HDBCA as a reference, as described⁵. The gained presence score was determined by subtracting the HNOCA presence scores from the brainstem organoid presence scores for each cell in the HDBCA. Gained presence scores, based on positive differences, were averaged across each HDBCA cluster and grouped by region.

6 **Glial differential expression analysis**

- 6.1 Cells expressing more than 200 genes and belonging to either OPC or Glioblast lineages in the HDBCA, HNOCA and the brainstem organoid model were aggregated into a pseudobulk object with three pseudoreplicates per dataset using the python implementation of decoupler. EgdeR was used to compute DEGs for each glial lineage using HDBCA_vs_HNOCA and HDBCA_vs_BrO, correcting for cell numbers, as well as median and standard deviation of the number of detected genes per pseudobulk sample. Genes were tested using the Genewise Negative Binomial Generalized Linear Model implemented in edgeR. Genes with an absolute log2FC above 1 and a qvalue below 0.05 were labelled as differentially expressed. For DEG distribution between the HNOCA and the brainstem organoid model, scipy's fisher's exact test was used to calculate the odds ratio and p-value.

7 **Voxhunt analysis**

- 7.1 To compare scRNA-seq data from the brainstem organoid model with mouse spatial gene expression data, the VoxHunt¹¹ R package was utilized. Spatial gene expression samples from embryonic day (E) 13 and E18 were selected. For each region at annotation level 'custom_1', the top 10 genes were identified based on their provided Area under the Curve.

Tumour analysis in DMGO

8 Inferred copy number variation (iCNV) analysis

- 8.1 To confirm tumorigenicity, we inferred the copy number of variation (iCNV) from single-cell gene expression data using SCEVAN¹² (v1.0.1). On a sample-by-sample basis, we used a non-transfected organoid as a healthy reference to estimate iCNVs in the malignant cells.

9 Annotation and reference comparisons

- 9.1 Samples from different sequencing runs were integrated using the scanpy implementation of harmony. On the integrated representation a diffusion map was calculated which served as input for a ForceAtlas representation of the data. Leiden clusters were annotated using published tumour state marker genes¹³. Annotation was subsequently confirmed by projecting the tumour into two different patient datasets^{13,14} using the RunAzimuth function from the Azimuth package (v.0.5.0).
- 9.2 For comparison of the DMGO tumour model to other tumours or models^{13–16} multiple references were merged, normalized, scaled and annotation unified. Anchors were obtained by FindTransferAnchors(), using the first 30 PCs of the shared PCA space and all genes of the reference dataset as features, and used to transfer the reference annotation to the query tumour data. The labels were transferred to the query Seurat as a separate assay and used as input for visualization.

10 Gene set enrichment analysis

- 10.1 The AddModuleScore() function in the Seurat package was used with default settings to compute gene set enrichments from curated lists of marker genes. Published molecular signatures for spatially restricted oligodendrocyte precursor lineages⁷ were used for similarity comparison of cNMF generated Program 1 and 2.

Lineage tracing analysis

11 TrackerSeq barcode recovery

- 11.1 Reads containing a perfect match to the TrackerSeq motif CTGA..CTG..ACT..GAC..TGA..CTG..ACT..GAC..GACT were extracted and tabulated per cell barcode and UMI to generate barcode occurrence counts for both the nested-PCR

library and the gene-expression library. All downstream analyses used the nested-PCR dataset.

- 11.2 For each cell barcode, total sequencing reads, total unique UMIs, and a mean “oversequencing” metric defined as the average number of reads per UMI per TrackerSeq barcode were quantified. Cells were retained if total reads exceeded 100 in experiment 1 or 1,000 in experiment 2, which enriches for cells in which at least one barcode exhibits mean oversequencing of $\sim 25\times$ (experiment 1) or $\sim 10\times$ (experiment 2). For retained cells, we computed for each TrackerSeq barcode both the read fraction (barcode reads/total reads) and UMI fraction (barcode UMIs/total UMIs) and kept barcodes that reached $\geq 10\%$ in both fractions in at least one cell. We then refined assignments by recalculating barcode fractions using the maximum oversequencing value per barcode per cell and removed barcodes with a resulting fraction < 0.2 . When multiple barcodes persisted within a cell, we first merged near-identical barcodes (< 4 edits). If more than one barcode still remained (N barcodes), we assumed a multinomial distribution and removed any barcode with a fraction below $1/N - (1/N)^{21/N} - (1/N)^{21/N} - (1/N)^2$ (mean–variance criterion). The final clonal barcode for each cell was defined as the union of the organoid identifier and the retained TrackerSeq barcode.

12 Clonal family comparison

- 12.1 The fraction of each clonal family was computed for each sample. Big clones were defined as detected in more than 20% of all the cells from each sample. Sample DMG0155 was excluded as it presented as only one large clone ($> 99\%$ of cells with the same barcode). Large versus small clones were compared by using the function `rank_genes_groups` from the scanpy python package and selected DEGs ($\log_2$ fold change > 0.7) to perform METASCAPE analysis.

13 Consensus non-negative matrix factorization (cNMF)

- 13.1 Gene expression programs in tumour scRNA-seq data were inferred via cNMF¹⁷ (v1.4.1), using 100 iterations of NMF with different random seeds for each value of k , the number of components, from 5 to 9. For each value of k , metrics indicating stability, Silhouette score and Frobenius error were calculated and the k maximizing the Silhouette score and minimizing the Frobenius error was selected for further processing. Outlier components from the selected k were filtered by removing components with a higher mean distance to most similar component of 0.02, which resulted in a program activity matrix and a gene scores matrix. To compare the meta programs to our previous annotations we used `scclusteval`¹⁸ (v1.0).

Protocol references

1. Bandler, R. C. *et al.* Single-cell delineation of lineage and genetic identity in the mouse brain. *Nature* **601**, 404–409 (2022).
2. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
3. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
4. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* **8**, 329–337.e4 (2019).
5. He, Z. *et al.* An integrated transcriptomic cell atlas of human neural organoids. *Nature* **635**, 690–698 (2024).
6. Luecken, M. D. *et al.* Benchmarking atlas-level data integration in single-cell genomics. *Nat. Methods* **19**, 41–50 (2022).
7. Braun, E. *et al.* Comprehensive cell atlas of the first-trimester developing human brain. *Science* **382**, eadf1226 (2023).
8. Albiach, A. M. *et al.* Glioblastoma is spatially organized by neurodevelopmental programs and a glial-like wound healing response. *bioRxiv* 2023.09.01.555882 (2023) doi:10.1101/2023.09.01.555882.
9. Büttner, M., Ostner, J., Müller, C. L., Theis, F. J. & Schubert, B. scCODA is a Bayesian model for compositional single-cell data analysis. *Nat. Commun.* **12**, 6876 (2021).
10. Heumos, L. *et al.* Pertpy: an end-to-end framework for perturbation analysis. *bioRxiv* 2024.08.04.606516 (2024) doi:10.1101/2024.08.04.606516.
11. Fleck, J. S. *et al.* Resolving organoid brain region identities by mapping single-cell genomic data to reference atlases. *Cell Stem Cell* **28**, 1148–1159.e8 (2021).
12. Falco, A. D., Caruso, F., Su, X.-D., Iavarone, A. & Ceccarelli, M. A variational algorithm to detect the clonal copy number substructure of tumors from scRNA-seq data. *Nat. Commun.* **14**, 1074 (2023).
13. Liu, I. *et al.* The landscape of tumor cell states and spatial organization in H3-K27M mutant diffuse midline glioma across age and location. *Nat. Genet.* **54**, 1881–1894 (2022).
14. Jessa, S. *et al.* K27M in canonical and noncanonical H3 variants occurs in distinct oligodendroglial cell lineages in brain midline gliomas. *Nat. Genet.* **54**, 1865–1880 (2022).
15. Filbin, M. G. *et al.* Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq. *Science* **360**, 331–335 (2018).
16. Gillen, A. E. *et al.* Single-Cell RNA Sequencing of Childhood Ependymoma Reveals Neoplastic Cell Subpopulations That Impact Molecular Classification and Etiology. *Cell Rep.* **32**, 108023 (2020).
17. Kotliar, D. *et al.* Identifying gene expression programs of cell-type identity and cellular activity with single-cell RNA-Seq. *eLife* **8**, e43803 (2019).
18. Tang, M. *et al.* Evaluating single-cell cluster stability using the Jaccard similarity index. *Bioinformatics* **37**, 2212–2214 (2020).