Journal Pre-proof

Comparing the impact of sample multiplexing approaches for single-cell RNA-sequencing on downstream analysis using cerebellar organoids

Kseniia Sarieva, Theresa Kagermeier, Vladislav Lysenkov, Francesco Castagnetti, Zeynep Yentuer, Katharina Becker, Julia Matilainen, Nicolas Casadei, Simone Mayer

PII: S2589-0042(26)00155-0

DOI: https://doi.org/10.1016/j.isci.2026.114780

Reference: ISCI 114780

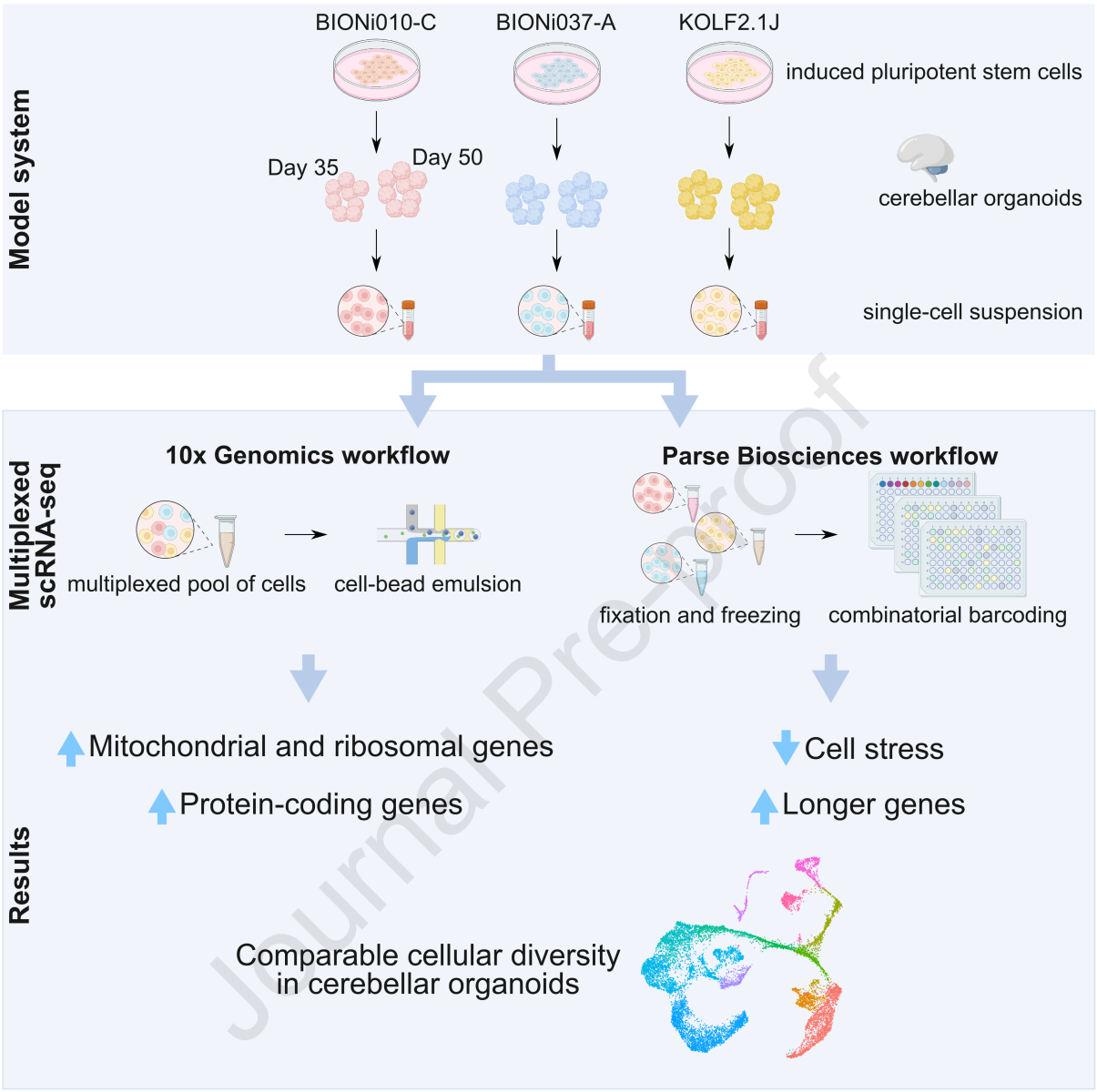
To appear in: *iScience*

Received Date: 18 June 2025

Revised Date: 10 November 2025

Accepted Date: 20 January 2026

Please cite this article as: Sarieva, K., Kagermeier, T., Lysenkov, V., Castagnetti, F., Yentuer, Z., Becker, K., Matilainen, J., Casadei, N., Mayer, S., Comparing the impact of sample multiplexing approaches for single-cell RNA-sequencing on downstream analysis using cerebellar organoids, *iScience* (2026), doi: https://doi.org/10.1016/j.isci.2026.114780.



Comparing the impact of sample multiplexing approaches for single-cell RNA-sequencing on downstream analysis using cerebellar organoids

Kseniia Sarieva[1]*, Theresa Kagermeier[1,2,8]*, Vladislav Lysenkov[3,4], Francesco Castagnetti[8], Zeynep Yentuer[1,2,5,6,8], Katharina Becker[1], Julia Matilainen[7], Nicolas Casadei[3,4], Simone Mayer[1,5,6,8,9,10#]

1, Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, 72076 Germany

2, International Max Planck Research School, Graduate Training Centre of Neuroscience, University of Tübingen, Tübingen, 72076 Germany

3, Institute of Medical Genetics and Applied Genomics, University Hospital Tübingen, 72076 Germany

4, NGS Competence Center Tübingen, Tübingen, 72076 Germany

5, Heidelberger Akademie der Wissenschaften, Heidelberg, 69117 Germany

6, IMPRS, The Mechanisms of Mental Function and Dysfunction, University of Tübingen, 72076 Germany

7, German PCH patient network (PCH-Familie e.V.), Böblingen, 71034 Germany

8, Zoological Institute, Karlsruhe Institute of Technology (KIT), 76131 Germany

9, Institute of Biological and Chemical Systems - Functional Molecular Systems, Karlsruhe Institute of Technology (KIT), 76131 Germany

10, Lead contact

*, equal contribution authors

#, corresponding author: Prof. Dr. Simone Mayer
Address: Wolfgang-Gaede-Straße 1a, 76131 Karlsruhe, Germany
Email: simone.mayer@kit.edu

Summary

Multiplexing overcomes limited throughput in single-cell RNA sequencing (scRNA-seq). Commercial strategies include Parse Biosciences combinatorial barcoding (Parse) and 10x Genomics CellPlex with microfluidic capture (10x). It is currently unknown how these techniques differ when characterizing complex tissues. Cerebellar organoids are a highly relevant model for studying cerebellar evolution, development, and disease. Yet, their extensive characterization through scRNA-seq is ongoing. Therefore, we compared the two multiplexing techniques using cerebellar organoids. While both strategies demonstrated technical reproducibility and revealed comparable cellular diversity, we found more stressed cells in 10x than in Parse. Additionally, Parse covered a higher gene biotype diversity and showed lower mitochondrial and ribosomal protein-coding transcript fractions. In summary, we demonstrate that both techniques provide similar insight into cerebellar organoid biology, but the flexibility of experimental design, capture of long transcripts, and the level of cell stress caused by the two workflows differ.

Keywords

transcriptomics, neurodevelopment, neural organoids, stem cells, multiplexing, bioinformatics

Introduction

Single-cell RNA-sequencing (scRNA-seq) has revolutionized our approach to characterize cell types, states, and lineages in various biological systems and is increasingly used in drug screening. While biological replica are essential for robust statistical analysis and the detection of even subtle changes between experimental conditions, replication has often been limited by technically challenging workflows and high costs[1,2]. Additionally, effective cell sampling maximizes the capture of cellular heterogeneity including rare cell populations[3]. Recent advances in commercialized kits now allow sample multiplexing, increasing both the number of cells assayed and the number of possible biological replicates. While combinatorial barcoding (as provided commercially by Parse Biosciences, hereafter Parse) is inherently multiplexed, microfluidic approaches (as provided commercially by 10x Genomics, hereafter 10x) require an additional labeling step for barcoding, mediated by antibodies or lipids[4]. However, increasing the number of samples remains technically challenging when working with fresh tissue because dissociation, a highly manual process, needs to be parallelized[5]. Fixation of the dissociated cells before capture (as performed in the Parse workflow) overcomes this obstacle, and different samples, for instance from different experimental time points, can be sequenced together, thereby avoiding batch effects of the capture. The kits allow multiplexing of up to 12 (10x) or 96 samples (Parse). The higher the number of multiplexed samples, the lower are the per-sample costs of cell capture with both strategies.

Since scRNA-seq multiplexing is widely used and datasets from different studies and experimental approaches are increasingly compared and integrated, it is important to consider the effects of the chosen multiplexing approach on the results. A recent study comparing both technologies using peripheral blood mononuclear cells (PBMCs) demonstrated that Parse had a higher sensitivity for detecting rare cell types[6]. Furthermore, it was shown that Parse covered a wider range of gene lengths, and that 10x was biased towards more GC-rich transcripts[6]. However, it remains unclear, to what extent these differences affect downstream analysis and highly complex 3D samples that require dissociation such as neural organoids.

Regionalized neural organoids recapitulate the development of specific brain regions with their specialized neural cell populations, making them a particularly powerful tool to study human neurodevelopment[7], to model neurological disorders[8,9], and to test on- and off-target effects of

pharmaceuticals[10,11]. The human cerebellum has long been thought to mainly be involved in motor learning and coordination[12], however, more recent insights into cerebellar function, describe its major contribution to cognitive functions such as attention, task execution, working memory, language and social behavior[13], and a contribution to neurodevelopmental disorders (NDD) such as autism spectrum disorder (ASD)[14,15]. Two cerebellar progenitor zones, the ventricular zone (VZ) and the rhombic lip (RL), arise from the rhombencephalon[16,17]. The VZ gives rise to all inhibitory neurons of the future cerebellum, including Purkinje cells (PC) and inhibitory neurons of the deep cerebellar nuclei. The RL generates all excitatory neurons, including granule cells (GC) and excitatory neurons of the deep cerebellar nuclei[18]. Progenitors and neurons from both progenitor zones can now be generated in human cerebellar organoids, placing them in a unique position to model cerebellar disorders such as cerebellar hypoplasias, Dandy-Walker Syndrome, ataxias, and medulloblastoma as pioneered in several recent studies[9,19,20]. However, the protocols underlying their generation are still being improved[21-23], and few single-cell RNA datasets of selected cell lines are available[21,23,24].

Here, we addressed two important gaps in knowledge related to multiplexing in scRNAseq and cerebellar organoid generation by comparing the technical features between the two multiplexing strategies, Parse and 10x, in complex tissue-like samples, cerebellar organoids derived from three control iPSC lines at two time points.

Results

Experimental design and quality assessment

To assess the reproducibility of cerebellar organoid differentiation and comparability of two multiplexed scRNA-seq methods, we differentiated three iPSC lines (BIONi010-C, BIONi037-A, and KOLF2.1J) into cerebellar organoids (Fig. 1A). Cell lines were handled in parallel throughout the experimental period. On day 28 (D28) and day 42 (D42) of differentiation, organoids were collected for quality control assessment by immunohistochemistry. We observed the expression of Purkinje cell marker SKOR2 at both time points (Fig. S1A), the granule cell precursor markers BARHL1 and ATOH1 were expressed at D42 (Fig. S2C), and general neuroectodermal commitment was indicated by the presence of neural precursor marker SOX2 as well as the early pan-neuronal markers Tuj1 (D28 and D42, Fig. S1C, Fig. S2B) and MAP2 (D42, Fig. S2B). Further, all cell lines demonstrated cell division at both timepoints, indicated by the expression of Ki-67 (Fig. S1B, Fig. S2A). Samples for scRNA-seq were harvested on day 35 (D35) and day 50 (D50) of differentiation. Pools of 24 organoids per cell line and time point were dissociated. One aliquot of each cell suspension was used for 10x, the other for Parse scRNA-seq workflow. This experimental design minimized the effect of biological variability and focused on technical differences between 10x and Parse.

Libraries were sequenced to achieve over 50,000 reads per cell (Table S1), and raw FASTQ files were downsampled to 50,000 reads per cell to allow a direct comparison of gene detection sensitivity (Supplementary Table 1). Alignment of reads was performed through technology-specific pipelines: cellranger v7.2.0 *multi* pipeline (10x) and split-pipe v1.1.2 (Parse).

In both technologies, most reads were mapped to the genome (93.2% for 10x, 91.8% for Parse, Fig. S3A, Table S2), with exonic reads constituting 56.3% of all reads in 10x, and 30.1% in Parse (Fig. S3A, Table S2). Valid barcodes were identified for 97.2% of reads for 10x and 79.9% for Parse (Fig. S3A, Table S2). The cell recovery rate was 42.7% for 10x and 16.5% for Parse (Fig. S3B, Table S2).

For further comparisons, technology-specific cell-by-gene matrices were merged. We found that 32,408 genes had a non-zero expression in both technologies, while 2,159 and 12,098 genes were uniquely expressed in 10x and Parse, respectively (Fig. S3C). After merging count matrices from both technologies, we only retained genes that had a non-zero expression in more than 8 cells, resulting in the count matrix with 38,580 genes (Fig. S3D).

For further analysis, we used the following combination of metadata parameters to assign cells to samples unless stated otherwise: (1) technology (10x vs Parse); (2) day of differentiation (D35 vs D50) of cerebellar organoids; and (3) sequencing library (L1 and L2). Day of differentiation was used as a covariate to acknowledge both biological differences in the stage of organoid differentiation and technical differences arising from harvesting D35 and D50 samples on different days. The sequencing library was used as a covariate to show the reproducibility of the workflow within each technology.

After cell-level quality control (QC), we recovered 87.2% of cells from 10x and 95.6% of cells from Parse datasets (10x, 29,505 out of 33,951 cells; Parse, 14,542 out of 15,226 cells, Fig. S3E). Interestingly, the number of genes per cell was higher in Parse both before and after QC ($p < 0.001$, Fig. 1B). While protein-coding genes were the most abundant in both technologies (Fig. 1B), Parse recovered a higher proportion of non-coding RNAs (ncRNA) reads, including long non-coding RNA (lncRNA) (Fig. S3F). Additionally, the percentage of mitochondrial and ribosomal protein-coding transcripts was lower in Parse than in 10x. In contrast, the percentage of reads originating from transcription factors (TF) among protein-coding genes was higher in Parse than in 10x (Fig. 1b, Supplementary Fig. 3e). In line with previous findings[6], the correlation of gene expression between the two technologies across cells was only moderate (Pearson's $r = 0.6$) (Fig. 1C), indicating differential gene detection between the two technologies.

Different RNA-seq technologies are known to have biases in gene detection based on gene properties such as GC content and gene length[6,25]. To characterize these biases, we analyzed the correspondence between gene abundance and gene length or GC content (Fig. 1D, Fig. S3F). While using all expressed genes per technology revealed small but statistically significant differences in these parameters ($p < 0.001$, Fig. S3F), gene length and GC content of differentially expressed genes (DEG) per technology (10x, 2,737 DEGs; Parse, 4,055 DEGs) differed to a higher extent (Fig. 1D), reminiscent of previously published results[6]. We observed a bias towards detecting longer genes in Parse, both for protein-coding genes and lncRNA (Fig. S3G). Finally, we performed an extensive analysis of gene detection sensitivity and biases (Table S2) largely corroborating results from the previous benchmarking study[6]. We therefore suggest that the observed differences are characteristic features of 10x and Parse technologies independent of sample type.

Technical and biological differences between technologies

Next, data normalization revealed highly variable genes for Principal Component Analysis (PCA) as well as Uniform Manifold Approximation and Projection (UMAP) on unintegrated data (Fig. 2A). As expected from previous results[6] and our QC, both PCA and UMAP revealed major differences between the technologies (Fig. 2A). We hypothesized that these differences arise from sample preparation where cells for Parse were immediately fixed and frozen after dissociation, while cells undergoing 10x capture were depleted of nutrients and passed through microfluidic channels of the instrument before lysis.

Hence, we hypothesized that cellular stress may contribute to differences between samples. We analyzed the expression of gene ontology (GO) modules involved in different modalities of cellular stress and its

downstream effects, such as integrated stress response (ISR) (Supplementary Fig. 4a). In hierarchical clustering of average GO module expression scores, samples from the two technologies clustered apart. The major differences came from three terms: response to oxidative stress, glycolytic process, and ISR signaling (Fig. S4A). Using only these three modules and the random set for hierarchical clustering led to the same results (Fig. 2B, Fig. S4A).

We further determined the number of stressed cells using Gruffi [26] using the top cell stress terms from the module expression analysis: glycolytic process (GO:0006096) and ISR signaling (GO:0140467). We found that the percentage of stressed cells varied between technologies but also between days of organoid differentiation (Fig. 2C,D, Fig. S4B). There were more stressed cells in the 10x data and both technologies captured more stressed cells in D50 cerebellar organoids (Fig. 2D). This finding can be explained by the diffusion-based distribution of nutrients in organoids leading to an increasing nutrient deficiency as organoids grow (D50 vs. D35)[27,28]. We therefore removed cells that were classified as stressed by Gruffi (6,595 out of 44,047 cells that passed QC) from further analysis, integrated normalized counts by sample using reciprocal PCA, and repeated PCA and UMAP. This analysis revealed that the data from the two technologies can be easily integrated (Fig. 2E).

To analyze the biological reproducibility of the cerebellar organoid protocol between different iPSC lines, we characterized the cellular diversity. We first aimed to understand whether organoids had neural identity. We, therefore, performed reference-query mapping of our dataset onto the human developmental transcriptome using Azimuth[29,30]. This reference dataset includes cell types from various tissues, including the nervous system and the cerebellum. We first assigned our cells with cell types from the reference dataset featuring cells from 15 human organs between 72 and 129 days post-conception[29,30] (Fig. S4C). High prediction scores were assigned to the cells annotated as skeletal muscle, bronchiolar and alveolar epithelial cells, enteric nervous system glia, astrocytes, and some neuronal cells (Fig. S4D). However, prediction scores varied between cells (0.59 ± 0.26, mean ± SD, Fig. S4D), with most cells not reaching a high-confidence prediction score of 0.75[30]. Therefore, we did not rely on the annotation of certain cell types but grouped the cells into two categories – neural and non-neural (Fig. 2F, Table S3). We found a considerable portion of cells having non-neural identity (Fig. 2F) with subsets of cells expressing muscular markers (e.g., *MYOD1* and *MYOG*[31]) and endo-/mesodermal markers (e.g., *FGF10*[32]) (Fig. 2G). Accordingly, cells expressing muscular markers were annotated as muscular cells with high confidence (Fig. S4D). In contrast, most cells classified as neural expressed the pan-neuronal marker *STMN2* (Fig. 2G). Among those cells, there were cells annotated as granule neurons and Purkinje neurons, albeit with lower prediction scores (Fig. S4C,D). Overall, the proportion of neural cells ranged from 46.0% to 60.7% per sample (Fig. 2H). Importantly, considerable differences were observed between the three iPSC lines that the organoids were generated from with BIONi010-C cell line having the highest number of neural cells (Fig. 2I).

To cross-validate this assignment we adapted Gruffi[26] for detecting neural and non-neural transcriptomic signatures. We used GO terms for endoderm (GO:0001706) and mesoderm formation (GO:0001707) for selecting non-neural cells and GO terms for nervous system development (GO:0007399) and neurogenesis (GO:0022008) for selecting neural cells (Fig. S4E). The results between reference-query mapping and Gruffi were mostly coherent (Fig. S4F). Inconsistent annotations were observed for putatively muscular cells (positive for *MYOG* and *MYOD1*), which were incorrectly classified as neural by Gruffi. We suggest that this discrepancy may be due to the shared excitability between neural and muscular cells.

Characterization of neural cell diversity

Utilizing the results of reference-query mapping with the human developmental transcriptome[29], we subset neural cells (19,526 neural cells out of 37,452 cells) and downsampled 10x and Parse datasets to an equal number of cells (resulting in 7,212 cells per technology) before repeating integration and dimensionality reduction. Next, we aimed to reveal the brain regional identity of the neural cells within the cerebellar organoids[23] by correlating regional marker gene expression (inferred from E15 mouse brain, Table S4) with our dataset and human brain transcriptomic data from postconceptional week (PCW) 12-13 from Brain Span[33,34]. All samples had the highest correlation with the cerebellum (Fig. S5A). However, when similarity scores were not scaled, they were higher for 10x than for Parse samples (Fig. 3A). Next, we assigned cell identities to the neural cells by combining cerebellar canonical marker gene [18,35-37] with differential gene expression (DGE). We identified both RL-derived cellular lineages (RL, granule precursor cells (GPC), and GC) and VZ-derived newborn PCs (Fig. 3B,C). A subset of neuronal cells was characterized as hindbrain neurons (Fig. 3B). While overall proportions of cells captured by the two technologies were similar (Fig. 3D, Fig. S5B), dividing progenitors, PAX6-positive RL and dividing RL cell populations were significantly enriched in Parse (Fig. 3D, Fig. S5B). We then visualized the distribution of cell types in organoids originating from different cell lines (Fig. S5C). This analysis revealed differences in proportions of different neural cell types between cell lines (Fig. S5C). This highlights the necessity to use multiple cell lines and batches of differentiation when characterizing the reproducibility of new neural organoid protocols.

To our knowledge, currently scRNA-seq cerebellar organoid datasets are available for D60[21] or D90[22] of differentiation. We hypothesized that our scRNA-seq analysis at D35 and D50 provides insights into the establishment of cell type diversity during differentiation. Indeed, when we visualized the distribution of cell types between the two sampling time points, we found that RL-derived populations had higher proportions in D35 than in D50 of differentiation while several neuronal populations, including newborn PCs, demonstrated the opposite trend (Fig. S5D,E). Therefore, cerebellar organoids recapitulated the temporal progression of cell type proportions characteristic of the developing cerebellum[18,36]. To characterize the similarity of our cerebellar organoids with the developing human cerebellum, we performed reference-query mapping with a primary cerebellar transcriptomic dataset, subset to only include prenatal samples[37]. While finding general agreement in cell type annotations, we noticed differences in both assigned cell type identities (Fig. 3E) and prediction scores, which were higher in Parse than in 10x data (Fig. S5F). We further compared our data with a recent scRNA-seq cerebellar organoids dataset (Fig. S5G,H)[21]. The prediction scores were higher than for the comparison with the human cerebellar developmental transcriptome (Fig. S5F). This time, however, prediction scores were higher for 10x than for Parse cells (Fig. S5I). Interestingly, both reference datasets were generated using 10x. Therefore, expectedly, our organoid data aligns more with organoid data obtained from a different protocol than with primary tissue.

In summary, we found that the cerebellar organoids indeed acquired a mid-gestational human cerebellar regional identity. We also found robust differentiation into both major cerebellar lineages, RL- and VZ-derived cells. Small variances in the different parameters were found between 10x and Parse technologies.

Secondary analysis between techniques reveals differences in cell stress signatures and neurodevelopment-related gene regulatory networks activity

During QC, we found differences in the percentage of reads originating from ribosomal and mitochondrial protein-coding genes between the two technologies (Fig. 1B). We also found a subset of cells expressing cell stress-related genes, and this proportion was higher for 10x (Fig. 2D). Therefore, we analyzed whether the neural cells preserved these transcriptomic features and performed DGE analysis between the different technologies within individual cell types. For that, we split the dataset by cell type, technology, cell line, and day of differentiation and pseudobulked cells for DESeq2. DEGs were spread across all cell types (Fig. 4A, Fig. S6A). Especially mitochondrial and ribosomal protein-coding genes were upregulated in 10x compared to Parse (Table S5), including GPCs (Fig. 4B). More genes were upregulated in 10x compared to Parse across all cell types (Fig. S6A). Interestingly, there were a few genes with large fold change and relatively large p-values upregulated in either of the two technologies (Fig. S6B). To functionally characterize the differences in gene expression between the techniques, we performed gene set enrichment analysis (GSEA) and clustered the output in a semantic similarity matrix (Fig. 4C). Here, we describe findings for GSEA in GPCs, as a representative cell type with relatively high cell numbers and a medium number of DEGs. In GPCs, the normalized expression score for all statistically significant GO terms was less than 0, indicating their upregulation in 10x compared with Parse (Table S6). One cluster of enriched GO terms was related to nucleotide processing, another to mitochondrial respiration. These two clusters of GO terms included not only mitochondria-encoded protein-coding genes (Fig. 1B), but also nuclear-encoded genes involved in mitochondrial function (e.g., NDUF, Table S6). Another group of enriched GO terms in GPCs was described as related to neuron projection assembly (Fig. 4C).

To reveal the upstream mechanisms leading to the transcriptional changes across cell types, we performed upstream regulator analysis (URA). That predicted a variety of TFs to be differentially active in either of the technologies, and that these transcriptional changes were coordinated across cell types (Fig. 4D). For example, we found ER stress-induced TFs *XBP1, ATF4* and *ATF6*, and *NFE2L2* and *NRF1,* which mediate oxidative stress response and are involved in maintaining mitochondria redox homeostasis[38-40] to be upregulated in 10x. These predictions are in line with our previous findings (Fig. 2B, Fig. S4A), demonstrating a higher proportion of stressed cells in 10x than Parse. Since we found that the Parse dataset had a larger proportion of reads originating from TFs (Fig. 1B), we decided to extend our analysis to gene regulatory network (GRN) analysis using SCENIC[41]. Average area under the curve (AUC) scores per cell type and technology were z-score normalized and subjected to k-means clustering (Fig. 4E). We found that, albeit overall small z-scores, the two technologies clustered apart (column clusters 1 and 3 for 10x, and 2 and 4 for Parse) but also cell types divided into two meta groups based on the activity of GRNs (column clusters 1 and 2 were enriched in neurons, while column clusters 3 and 4 contained predominantly progenitor cell types, Fig. 4E, Figure S6C). Additional examples of cell type-associated regulon activity can be found in Fig. S6C. In summary, transcriptional differences between technologies did not mask transcriptional differences between cell types.

Discussion

In this study, we compared two broadly used and commercialized approaches for sample multiplexing of scRNA-seq: 10x and Parse using cerebellar organoids, as an example of a complex 3D sample that requires dissociation. Regionalized neural organoids, such as cerebellar organoids, are commonly used in neuroscience research but can be challenging due to heterogeneity between samples, batches, and iPSC lines and require in-depth characterization [7,42]. To compare scRNA-seq datasets across experiments and

to differentiate technical and biological causes of variance, it is essential to understand artefacts and biases introduced by experimental pipelines of cell capture techniques. We generated cerebellar organoids[23] from three iPSC lines, dissociated the samples at D35 and D50, subjected them to 10x and Parse cell capture, and sequenced the resulting libraries. We compared the methods based on library efficiency, differential transcript capture, cell type enrichment, and secondary analysis insights.

Sample preparation differs considerably between the two technologies: Parse samples are fixed after dissociation, whereas cells are kept alive until lysis in 10x. Consequently, Parse provides more flexibility in sample processing and allows handling higher sample numbers in one sequencing run, which is advantageous for larger experiments. Further, we observed differences in the cell recovery rates—42.7% for 10x and 16.5% for Parse (Fig. S3B), and for scarce samples, higher recovery is beneficial to maximize data output. The downside of fixation may be decreased RNA quality, reflected in a lower number of recovered genes[43]. Next, there are substantial differences in library preparation protocols. Namely, one important difference is that Parse, in addition to oligo(dT) primers, uses random hexamer primers for reverse transcription, thus allowing multipriming. Since both technologies rely on reverse transcription of mRNA up to its 5' end, multipriming may allow for more robust recovery of longer transcripts. Random hexamer primers also allow capturing transcripts devoid of polyA tails, non-coding RNAs, and nascent transcripts. Other steps of library preparation, albeit different, rely on reactions that should be either immune to transcript-specific biases or whose effect is difficult to predict.

For alignment and generation of count tables, we used technology-specific pipelines, thus leveraging optimal settings for both library construction protocols. Both approaches use the same tools for data processing with minor changes, thus potentially introducing minor variations into the resulting count tables.

Consistent with previous findings[6] and suggested effects of different sample processing and library generation protocols, we observed differences in the number of detected genes and their properties. Namely, 10x resulted in a higher number of genes, a higher number of protein-coding genes, including mitochondrial and ribosomal protein-coding genes compared to Parse (Fig. 1B, Fig. S3E). Furthermore, 10x captured transcripts with higher GC content, while Parse captured longer transcripts (Fig. 1B, Fig. S3E,H). Previous studies showed a connection between gene length and neurodevelopment and NDDs[44,45]. Interestingly, BCL11b (CTIP2) (102,911 bps), a TF crucial for neuronal maturation and differentiation, is predicted to be upregulated in Parse in DAB1/CALB1/CALB2 HindN in our data (Fig. 4D). The clinical features of BCL11b-associated NDDs include ASD, intellectual disability, and cerebellar hypoplasia[45], which have been previously modeled in organoids[8,9]. These findings highlight transcript length as an important factor, suggesting Parse may be better suited for studying long transcripts upon experimental manipulations.

Further, Parse covered more transcripts encoding TFs among protein-coding genes (Fig. 1B, Fig. S3E). To investigate if this bias had effects on GRN activity, we employed SCENIC analysis. Interestingly, Parse showed higher z-scores for neurodevelopment and maturation-related regulons (Fig. 4E), in contrast to the upregulation of neuron processes assembly-related terms in 10x in GSEA (Fig. 4C). Additionally, we identified cell type- and technique-specific differences in regulon activity. For instance, *NFIA* regulon had higher z-scores in RL derivates in Parse (Fig. 4E), a TF relevant for GC maturation and linked to NDDs and gliomas[46,47]. Taken together, the GRN analysis reveals not only cell type but also technique-driven differences in regulon activity of identical biological samples.

During QC, we found that the percentage of mitochondrial and ribosomal protein-coding genes was higher in 10x samples (Fig. 1B), corroborating previous findings[6]. DGE analysis revealed the upregulation of mitochondrial protein-coding genes and other genes involved in mitochondrial function (Fig. 4B). Hence, the differences in mitochondrial transcripts might be partially explained by higher cell stress in the 10x data and mitochondrial involvement in stress response pathways[48].

We analyzed stress-specific modulators and identified three modules (oxidative stress, glycolysis, and ISR) that separated the two technologies. 10x showed stronger stress module expression overall, particularly at D50. Regionalized neural organoids have been reported to show high expression of stress pathway-related transcripts due to *in vitro* culturing conditions and insufficient oxygen supply[26-28,49]. The cascade of events unfolding upon persistent hypoxia may explain the elevated stress response-associated transcriptional signature at D50 compared to D35 of differentiation (Fig. 2B)[50]. Additionally, tissue dissociation for single-cell capture can induce stress response[51]. Since stressed cells are common in scRNA-seq organoid datasets, a bioinformatic approach called Gruffi was developed to remove these cells[26]. Using Gruffi, we found a noticeably higher percentage of stressed cells in the 10x compared to the Parse dataset at both time points (Fig. 2D). Immediate fixation of cells after the dissociation in Parse may limit the induction of stress-related genes, in contrast to live cells processing in 10x. These findings suggest that identical samples of cerebellar organoids show a technology- and time point-specific stress response reflected in transcriptional signature and striking differences in the number of cells identified as stressed (Fig. 2D).

To assess the biological reproducibility of organoid differentiation, we assessed the percentage of neural cells[29], revealing neural commitment of 52.1% of all cells, suggesting the initial tissue specification could be improved. Different neural organoid protocols[52,53] and a recently published protocol for cerebellar organoids[21] use dual SMAD inhibition during initiation of differentiation to prevent meso- and endodermal fates thus promoting neural induction[54]. In contrast, the cerebellar differentiation protocol used in this study employs only one SMAD pathway inhibitor[23], and dual SMAD inhibition could improve neuroectodermal commitment. Furthermore, we noticed substantial differences in differentiation efficiencies between the iPSC lines, with the KOLF2.1J-derived cerebellar organoids demonstrating the lowest proportion of neural cells. This suggests that iPSC line-inherent mechanisms influence the differentiation efficiency[55], underscoring the importance of using isogenic control iPSCs when analyzing pathogenic variants[56]. A recent study suggests adjusting small molecule and growth factor concentrations in cortical organoids for individual iPSC lines to reduce off-target tissue[40]. This approach could reduce line-to-line variability. Further, a recent preprint demonstrates structural variants in neurodevelopmental genes in KOLF2.1 line that could affect neural differentiation[57]. Despite the differences between the three iPSC lines used in this study, our cerebellar organoids generated cerebellar cells of both RL and VZ lineage. Comparing our data set with a recently published cerebellar organoid transcriptomic dataset[21] revealed similar cell populations.

In conclusion, our comparison of Parse and 10x encompassed library efficiency, differential transcript capture, cell type preferences, and secondary analysis outcomes, showing distinct strengths of each method. While 10x provided higher cell recovery and gene detection rates, Parse captured longer transcripts and a wider range of transcript lengths and resulted in lower cell stress—important for regionalized neural organoids, in which cell stress may be a key artifact[27]. These technical differences have relevant biological implications, making it essential to choose the appropriate method based on specific research goals. Future studies should consider these factors to improve the accuracy and biological relevance of single-cell transcriptomic analyses. Finally, we demonstrated cerebellar organoid

differentiation and in-depth characterization on three iPSC lines, highlighting the importance of using diverse cell lines to capture line-to-line variability.

Limitations of the study

In the current study, we used cerebellar organoids as a model system to showcase technical differences arising from two single-cell capturing and multiplexing techniques. Our experimental design has several limitations. First, the study relied on one differentiation per line, reducing our ability to separate protocol effects from batch- and line-specific variability. Next, although the organoids produced relevant cerebellar lineages, neural commitment was incomplete, and a substantial fraction of cells adopted non-neural fates, indicating that early tissue specification was not fully optimized. More recent cerebellar organoid differentiation protocols[21,58,59] may better suppress meso-/endodermal trajectories and improve neural induction through dual SMAD inhibition, and should be evaluated in future work. In addition, differentiation efficiency differed markedly across cell lines, with the KOLF2.1J line showing the lowest neural yield, suggesting that line-inherent properties may confound direct comparisons[57,60]. Finally, some technical features of the single-cell workflows may introduce differences in count tables due to bioinformatic processing. In our analysis, we generated count tables using technology-specific pipelines, ensuring optimal settings. While most of the bioinformatic tools are shared between pipelines, some minor differences exist. For example, while both Cell Ranger and split-pipe perform UMI demultiplexing, only Cell Ranger reports further UMI correction and exclusion of low-quality UMIs.

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Simone Mayer (simone.mayer@kit.edu).

Materials availability

This study did not generate new unique reagents or new iPSC cell lines.

Data and code availability

- Single-cell RNA sequencing data are available at CellXGene (link: https://cellxgene.cziscience.com/collections/0dd101f7-9829-44b3-a323-18b113eabeb4).
- Code: This study did not generate novel code, and the required functions for the analysis and data visualization are described in the STAR Methods section.
- All other items: Further requests for additional information should be directed to the lead contact, Simone Mayer (simone.mayer@kit.edumailto:).

Acknowledgements

We thank Antje Schulze-Selting, Elisabeth Gustafsson, Christina Kulka, and Ezgi Atay for technical support. We thank Christopher Sifuentes, Yogesh Singh, and Vincent Hammer for strategic and technical discussions.

We are grateful for financial support from the Hertie Foundation (Gemeinnützige Hertie-Stiftung), the Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg state postgraduate fellowship (to TK), Add-on Fellowship of the Joachim Herz Foundation (to KS), and the Heidelberger Akademie der Wissenschaften (WIN Kolleg). NGS sequencing for the Parse libraries was performed with the support of

the DFG-funded NGS Competence Center Tübingen (INST 37/1049-1). This project has been made possible in part by grant number 2022-316727 from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation. This research has been partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy via the Excellence Cluster 3D Matter Made to Order (EXC-2082/1 – 390761711).

Authors' contributions

KS: Conceptualization, Methodology, Software, Formal analysis, Writing – Original draft, Writing – Review & Editing, Visualization, Project administration; **TK**: Conceptualization, Methodology, Investigation, Writing – Original draft, Writing – Review & Editing, Visualization, Project administration; **VL**: Methodology, Software, Formal analysis, Writing – Original draft; Writing – Review & Editing; **FC**: Investigation, Writing – Original draft; Writing – Review & Editing; **ZY**: Investigation, Writing – Original draft; Writing – Review & Editing; **KB**: Investigation; Writing – Review & Editing; **JM**: Funding acquisition; Writing – Review & Editing; **NC**: Conceptualization, Methodology, Formal analysis, Writing – Original draft, Writing – Review & Editing, Resources, Supervision; **SM**: Conceptualization, Methodology, Writing – Review & Editing, Resources, Supervision, Funding acquisition.

Declaration of interests

The authors declare no competing interests. Kseniia Sarieva is currently affiliated with Oncode Institute, Hubrecht Institute, KNAW, and University Medical Center Utrecht, Utrecht, the Netherlands. Vladislav Lysenkov is currently affiliated with the Technical University of Munich, Munich, Germany.

Declaration of generative AI and AI-assisted technologies

During the preparation of this work, the authors used ChatGPT in order to improve the language and readability. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Figure titles and legends

Fig. 1. Study design, quality control, and potential biases in the data. **A**, Three iPSC lines (BIONi010-C, BIONi037-A, and KOLF2.1J) were differentiated to cerebellar organoids until days 35 and 50. The organoids generated from the same cell line were pooled and dissociated into single cells when each single-cell suspension was split into two portions. One set of single-cell suspensions was immediately subjected to sample multiplexing with CellPlex and processed in 10x Genomics 3'GEX+FB pipeline. The second set of single-cell suspensions was frozen until all samples were available. The samples were further processed though Parse Biosciences Evercode v2 pipeline. Libraries were sequenced, and the resulting FASTQ files were processed with technology-specific computational pipelines. Count matrices were further analyzed. Graphic was created with BioRender.com. **B**, Quality statistics after quality control. Color represents sample identity with respect to technology (10x or Parse), day of differentiation (D35 or D50), and library (L1 or L2). 10x, n = 29,505, Parse, n = 14,542 cells. Three-way ANOVA, p-values represent differences between technologies, *** $p < 0.001$. **C**, Left, density scatter plot showing correlation of average gene expression between the two technologies. Right, scatter plot showing correlation of average gene expression between the two technologies. Color represents gene group. **D**, Distributions of gene GC content and gene length for differentially expressed genes between technologies. Two-sided t-test, *** $p < 0.001$. See also Figure S1-3 and Table S1 and S2.

Fig. 2. Assessment of neural lineage identity. **A**, PCA and UMAP plots for globally normalized and unintegrated data. **B**, Heatmap representing mean module expression scores of gene ontology terms related to aspects of cell stress. **C**, UMAP plot representing cell stress status of cells based on Gruffi assessment. **D**, Percentage of stressed cells based on Gruffi assessment. **E**, RPCA and UMAP plots for globally normalized and RPCA-integrated data originating from non-stressed cells. **F**, UMAP plot representing neural lineage status of cells based on reference-query integration with human developmental transcriptome[29]. **G**, Feature plots showing expression of selected genes to highlight developmental lineages. **H**, Percentage of neuroectodermal cells based on reference-query integration with human developmental transcriptome. **I**, Percentage of neuroectodermal cells per cell line based on reference-query integration with human developmental transcriptome. For **A, D, E, H, I**, color represents sample identity with respect to technology (10x or Parse), day of differentiation (D35 or D50), and library (L1 or L2). See also Figure S3 and Table S3.

Fig. 3. Assessment of regional identity and cell type annotation. **A**, Heatmap of similarity metric of VoxHunt algorithm comparing samples with human neocortical RNA-seq data from BrainSpan using brain regional markers obtained from Mouse Brain Atlas at E13. **B**, UMAP plots for globally normalized and RPCA-integrated neural data with manually annotated clusters. **C**, Violin plots for expression of canonical markers of hindbrain development. **D**, Stacked bar plot representing average proportion of individual cell types between technologies. **E**, UMAP plot representing cell type identity as assigned based on reference-query integration with human cerebellar transcriptome[37]. **F**, Feature plots showing prediction score based on reference-query integration with human cerebellar transcriptome. See also Figure S5 and Table S4.

Fig. 4. Differential gene expression between technologies. **A**, Strip plot displaying DEGs between technologies per cell type. Genes represented in grey are not differentially expressed. Color represents log10 adjusted p-value for differentially expressed genes (absolute log2 fold change > 1, FDR < 10^{-4}). **B**, Volcano plot representing differential gene expression in GPC cluster. **C**, Heatmap representing semantic similarity between GO terms identified as significantly deregulated in GPC cluster by GSEA analysis. **D**,

459 Heatmap representing z-scores for SCENIC regulon activity calculated based on AUC scores. See also
460 Figure S6 and Table S5 and S6.

461

462 STAR★METHODS

463 Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
CORL2 (SKOR2)	Atlas Antibodies	Cat# HPA046206; RRID:AB_2679588
Sox2	R&D Systems	Cat# AF2018; RRID:AB_355110
BARHL1	Atlas Antibodies	Cat# HPA004809; RRID:AB_1078266
ATOH1	Sigma-Aldrich	Cat# WH0000474M1; RRID:AB_1839957
Ki67	Merck	Cat# AB9260; RRID: AB_2142366
Tuj1 (TUBB3)	Atlas Antibodies	Cat# AMAb91394; RRID: AB_2716670
Map2	Abcam	Cat# ab32454; RRID:AB_776174
Bacterial and virus strains		
N/A	N/A	N/A
Biological samples		
N/A	N/A	N/A
Chemicals, peptides, and recombinant proteins		
N/A	N/A	N/A
Critical commercial assays		
Chromium Next GEM Single Cell 3' kit v3.1	10x Genomics	Cat# 1000268
Evercode WT Mini v2	Parse Biosciences	Cat# ECW02110
Deposited data		
scRNA-seq data of hiPSC-derived cerebellar organoids	This paper	https://cellxgene.cziscience.com/collections/0dd101f7-9829-44b3-a323-18b113eabeb4
Human fetal development scRNA-seq	Cao et al.[29]	DOI: 10.1126/science.aba7721
BrainSpan human developmental transcriptome	Miller et al.[34]	DOI: 10.1038/nature13185
Human cerebellar development scRNA-seq	Sepp et al.[37]	DOI: 10.1038/s41586-023-06884-x
hiPSC-dericed cerebellar organoids scRNA-seq	Atamian et al.[21]	DOI: 10.1016/j.stem.2023.11.013
Experimental models: Cell lines		
BIONi010-C (male)	EBiSC	hiPSC
BIONi037-A (female)	EBiSC	hiPSC
KOLF2.1J (male)	Jackson Laboratory	hiPSC
Experimental models: Organisms/strains		
N/A	N/A	N/A
Oligonucleotides		
N/A	N/A	N/A
Recombinant DNA		
N/A	N/A	N/A
Software and algorithms		
Cell Ranger v.7.2.0	10x Genomics	https://www.10xgenomics.com/support/software/cell-ranger/downloads

Split-pipe v.1.1.2	Parse Biosciences	https://www.parsebiosciences.com/
Other		
N/A	N/A	N/A

Experimental model and study participant details

iPSC culture

Commercially available iPSC lines BIONi010-C (EBiSC), BIONi037-A (EBiSC) and KOLF2.1J (Jackson Laboratory) were cultured under standard conditions (37°C, 5% CO2, and 100% humidity) in E8 Flex medium (BIONi010-C and BIONi037-A, Gibco, Cat. no. A2858501) and mTeSR plus (KOLF2.1J, STEMCELL Technologies, Cat. no 100-0276) on hESC-qualified growth factor-reduced Matrigel-coated (Corning, Cat. no. 354277) cell culture dishes (Greiner, Cat. no. 657160). Passaging was performed using Gentle Dissociation Reagent (STEMCELL Technologies, Cat. no. 07174) once cells reached 80%-90% confluency. Medium was supplemented with Thiazovivin (Sigma-Aldrich, Cat. no. 420220) upon passaging for one day. All cell lines were kept under passage 20 and tested for mycoplasma using PCR Mycoplasma Detection Set (TaKaRa, Cat. no. 6601) and pluripotency by immunocytochemistry against OCT4 (rabbit, 1:500, Abcam, Cat. no. ab19857).

Generation of cerebellar organoids

Cerebellar organoids were generated as previously described[23,61] with some modifications: 80-90% confluent iPSCs were dissociated using Accutase (Merck, Cat. no. A6964), and 4,500 cells per well were seeded into 96 well plates (S-bio, Cat. no. MS-9096VZ) in culture medium (Gibco, Cat. no. A2858501), supplemented with 10 µM Y-27632 (Cayman Chemical, Cat. no. 10005583). Once aggregates reached 250 µm in diameter, medium was changed to growth factor-free chemically defined medium (gfCDM) supplemented with 50 ng/ml FGF2 (PeproTech, Cat. no. 100-18B) and 10 µM SB-431542 (Tocris, Cat. No. 1614). At D7, FGF2 and SB-431542 were reduced to 33.3 ng/ml and 6.67 µM, respectively. At D14, media was supplemented with 100 ng/ml FGF19 (PeproTech, Cat. No. 100-32). The medium was changed to Neurobasal Medium at D21, supplemented with 300 ng/ml SDF-1 from D28 to D34. From D35 onwards, media was changed to complete BrainPhys (StemCell Technologies, Cat. no. 5793), supplemented with 10 µg/ml BDNF (PeproTech, Cat. no. 450-02), 100 µg/ml GDNF (PeproTech, Cat. no. 450-10), 100 mg/ml dbcAMP (PeproTech, Cat. no. 1698950) and 250 mM ascorbic acid (Tocris, Cat. no. 4055). All three cell lines were processed in parallel throughout the experiments.

Method details

Immunohistochemistry on organoids

Organoids were fixed at the respective time points in 4% PFA in PBS for 45-60 min at room temperature[62]. The organoids were washed three times for 15 min with 1× PBS and then incubated in 30% sucrose (Sigma-Aldrich, S7903) in PBS solution at 4°C until they sunk to the bottom of the dish. The organoids were embedded in a 1:1 v/v mixture of 30% sucrose in PBS and optimal cutting temperature (OCT) compound (Sakura, 4583) and sectioned on Superfrost Plus slides (R. Langenbrinck GmbH, 03-0060) with a cryostat at 20 µm (Leica). The slides were stored at −80°C.

For immunohistochemistry, slides were thawed for 15 min at room temperature, and the embedding solution was rinsed off with PBS. A hydrophobic pen (PAP pen, Abcam, ab2601) was used to circle the sections to prevent the blocking solution from spilling during incubation. Permeabilization and blocking

were performed with 1% Triton X-100, 0.2% gelatin (Sigma-Aldrich, G1890) and 10% normal donkey serum in PBS for 1 h at room temperature. Primary antibodies were diluted in permeabilization and blocking solution and applied to the sections overnight at 4°C. Subsequently, the slides were rinsed with PBS three times for 15 min, then secondary antibodies were diluted in permeabilization and blocking solution and applied for 3 h at room temperature. The sections were rinsed in PBS three times for 15 min and nuclei were stained with DAPI (1:5000) diluted in PBS for 4 min. The sections were then rinsed in PBS and mounted using ProLong Gold (Invitrogen, P36930). Image acquisition was performed at 20X magnification using ECHO Revolution Hybrid Automated Microscope (DISCOVER ECHO INC.).

Single-cell dissociation of cerebellar organoids, library preparation, and sequencing

On D35 and D50, 24 organoids per cell line were pooled and dissociated using the Papain dissociation kit (Worthington, Cat.No. LK003150) following a published protocol with minor modifications[52]. Cells were counted and divided into aliquots for further processing.

For the 10x Genomics pipeline cells were labeled with cell multiplexing oligos (CMO, 10x Genomics, Cat. no. 1000261) and subsequently pooled at an equal ratio. Cells were counted and loaded onto two lanes of a Chromium Next Gen Chip G (10x Genomics, Cat. no. 1000120) with a targeted cell recovery of 12,000 (D35) and 14,000 (D50) cells per lane. Library preparation was performed with the Chromium Next GEM Single Cell 3' kit v3.1 (10x Genomics, Cat. no. 1000268), and sequencing was performed on NovaSeq 6000 with S1 flow cell kit and 100 cycles (Illumina, Cat. no. 20028319).

Samples for Parse Bioscience workflow were fixed using the Evercode fixation kit for cells (Parse Bioscience, Cat. No. WF300). Fixed samples were stored at –80°C. Samples were characterized by day of differentiation (D35 or D50) and cell line (BIONi010-C, BIONi037-A, or KOLF2.1J). Every sample was loaded as a technical duplicate into 2 independent wells, with all samples spanning wells 1-12. Sequencing was performed using a molarity of 62.4 nM and 3% PhiX spike-in on the Nova Seq 6000 with SP flow cell kit and 200 cycles (Illumina).

Quantification and statistical analysis

Data downsampling, preprocessing, and quality control

Initially, the datasets from 10x and Parse pipelines had different sequencing depths (Supplementary Table 1). To ensure fair comparison, we downsampled both datasets to an average of 50,000 reads per cell. The FASTQ files were downsampled with the *seqtk sample* tool using the same seed for forward and reverse reads. Parse FASTQ files from each of the 2 sub-libraries were demultiplexed into 6 samples and processed *split-pipe* (v1.1.2), resulting in a count matrix. The 2 sub-libraries were merged with *combine* mode of *split-pipe*. For 10x data, read downsampling was performed for individual libraries. Afterwards, downsampled FASTQ files were processed with *cellranger* (v.7.2.0) *multi* pipeline, assigning their cell line of origin based on CMO.

Gene names in count matrices between the two technologies were harmonized as follows: First, ENSEMBL gene identifiers were used to merge expression matrices. Secondly, ENSEMBL identifiers were replaced by HGCN identifiers wherever possible (41,980 genes), and ENSEMBL identifiers were used in other cases (20,930 genes). The merged count matrix was converted into a Seurat object (Seurat v.5.1.0). Gene biotypes were retrieved from bioMart using ENSEMBL identifiers. Ribosomal and mitochondrial protein-coding genes were identified by HGCN names starting with RPS/RPL and MT-, respectively. The percentage of gene expression for ribosomal and mitochondrial protein-coding genes as well as for individual gene

biotypes were calculated using *PercentageFeatureSet()*. For transcription factors (TF) among protein-coding genes, the count matrix was first subset to protein-coding genes, and *PercentageFeatureSet()* was applied using the human TFs list[63].

Next, QC was performed on cell and gene levels. Cells were excluded if they met any of the following criteria: (1) number of genes per cell ≤ 2,000 or ≥ 13,000; (2) number of genes per UMI ≤ 0.8; or (3) percentage of mitochondrial genes ≥ 8%. Genes were excluded if their cumulative expression across all cells was ≤ 8.

Data normalization, clustering, integration, and dimensionality reduction

After QC, data were normalized using Seurat's *NormalizeData()* function with default parameters. Normalized data were then scaled, and principal component analysis (PCA) was performed on the z-scaled expression of the 2,000 most variable features (*FindVariableFeatures()*). Additionally, normalized counts were integrated using *IntegrateData()* function with reciprocal PCA (RPCA). Dimensionality reduction and clustering were performed using both un- and integrated data. *RunUMAP()* function was used for dimensionality reduction with 30 neighbors and 30 principal components (PC). Louvain clustering was performed using *FindClusters()* function.

Technology-specific analyses: correlation analysis, transcript length, and GC content

To analyze the correlation of gene expression between technologies, we used cells that passed QC, averaged the gene expression for each technology, and calculated Pearson's correlation coefficient. DEGs between technologies were identified using the MAST algorithm in *FindMarkers()* function as previously described[6] with the following cutoffs: absolute log2 fold change (log2FC) > 1, adjusted p-value < 0.01. Gene length and GC content were retrieved from bioMart.

Cellular stress assessment

Normalized unintegrated counts were used to analyze the expression of cell stress-related GO terms using *AddModuleScore()* function. A random set of genes with mean GO term size was used as an internal control for module expression analysis. Hierarchical clustering was performed on mean module expression of cell stress-related GO terms across samples. Gruffi cell stress analysis was performed using normalized unintegrated counts following the authors' instructions[26]. Two GO terms were chosen for negative selection: glycolytic process (GO:0006096) and integrated stress response signaling (GO:0140467); and one for positive selection: neurogenesis (GO:0022008).

Germ layer assessment

Normalized integrated counts were used to perform Azimuth reference-query mapping[30] of our dataset with human fetal development transcriptome[29]. Cells were further classified as "neural" and "non-neural" based on cell type assignment from Azimuth (Supplementary Table 3). Gruffi differentiation lineage analysis was performed using normalized integrated counts. Two GO terms were chosen for negative selection: endoderm (GO:0001706) and mesoderm (GO:0001707) formation; and two for positive selection: nervous system development (GO:0007399) and neurogenesis (GO:0022008).

Neural data processing and cell type annotation

After germ layer assessment, the dataset was subset to neural cells by labels originating from Azimuth reference-query mapping and downsampled to retain the equal number of cells in 10x and Parse datasets (7,212 cells per technology). Data normalization, clustering, integration, and dimensionality reduction workflow steps were repeated as previously described.

VoxHunt[33] was used to analyze the brain region identity of cells. 10 genes with the highest AUC scores per brain region in the developing mouse brain at E15 were retrieved, resulting in 186 unique regional marker genes. These marker genes were used to assess the similarity of gene expression profiles between our samples and BrainSpan human developmental transcriptome[34] at postconceptional weeks 12 and 13.

Cell type annotation was performed for clusters at resolution 0.9 by a combination of approaches: (1) retrieving cluster marker genes by *FindAllMarkers()* with MAST (normalized counts provided as input) and ROC (raw counts provided as input) algorithms; (2) visualizing canonical marker gene expression for cell types in the developing mouse and human cerebellum.

Reference-query mapping with published primary cerebellar development and cerebellar organoids transcriptomic datasets

For reference-query mapping of neural-classified cells, we first used human cerebellar development transcriptomic dataset[37] as a reference, downsampling it to 1,000 cells per cell type as defined by the metadata (author_cell_type column). Both the reference and query datasets were normalized, variable features identified, scaled, and PCA was performed using Seurat's default parameters. Integration was performed using *FindTransferAnchors()* function with the "pcaproject" option and 30 PCs. Predicted cell types and prediction scores were obtained from *TransferData(),* wrapped into *MapQuery()*, with default parameters and "author_cell_type" as the reference label. For integration with the cerebellar organoids transcriptomic dataset[21], the same method was used with two adjustments: (1) the complete reference dataset was used for mapping; (2) the reference label was "final.clusters".

Differential gene expression analysis and functional enrichment analysis

For DGE analysis, raw counts from neural cells were used. Cells were grouped by cell type, technology, cell line, and day of differentiation, excluding groups with fewer than 20 cells. Gene counts were aggregated by technology, cell line, and day of differentiation using *AggregateExpression()* function with default settings to sum raw counts per group. No further downsampling was applied to equalize cell group sizes. The aggregated counts were used for *DESeq2* (v.1.42.1) DGE analysis between technologies within individual cell types[64]. Log2FC were shrunk using *apeglm* shrinkageestimator[65]. Volcano plots were generated using *EnhancedVolcano* library (v.1.20.0).

GSEA with GO terms was performed by *clusterProfiler* (v.4.10.1)[66] using Biological Processes gene ontology, gene set size of 50 to 500 genes, false discovery rate (FDR) for p-value adjustment with a q-value threshold of 0.05. For significantly deregulated GO terms, similarity matrices were calculated and simplified using the *binary cut* approach implemented in *simplifyEnrichment* (v.1.12.0) package[67].

Upstream regulator analysis

Upstream regulator analysis was conducted using IPA software (Qiagen). Cell type-specific DESeq2 output matrices were used for IPA core analysis with the following cutoffs: (1) absolute log2FC > 1; (2) q-value < 0.0001. For visualizations, molecule type was restricted to transcription regulators, and bias-corrected z-scores across cell types were used for hierarchical clustering using the *ComplexHeatmap* package (v.2.18.0). When z-scores were not available, they were set to 0.

Gene regulatory network activity analysis

We performed GRN analysis closely following the official pySCENIC protocol[41,68]. The annotated raw count matrix from Seurat and the list of human TFs were processed, inferring importance values of regulatory interactions between TFs and their target genes. The inferred interactions ("adjacencies") were searched

in the cisTarget databases to identify enriched binding motifs. TFs and target genes indicated by the enriched motifs were grouped into regulons, and their enrichment was assessed in each cell. Cells were assigned AUC scores representing activity levels of regulons. Z-scores were calculated based on AUC scores, and k-means clustering of z-scores was performed to reveal groups of co-regulated regulons. Regulon target genes and GO BP were used for gene set overrepresentation analysis (ORA) by clusterProfiler (v.4.10.1) with gene set size of 5 to 500 genes, FDR for p-value adjustment method, and a q-value threshold of 0.1.

Statistics

R v.4.3.2 was used for statistical analysis. Statistical tests are described in text and figure legends, and results are documented in Table S7. Two-sided unpaired t-tests were used to compare two groups. For comparisons with more than two groups, we used three-way ANOVA. Within a set of comparisons (e.g., for quality control metrics), the Benjamini-Hochberg method of p-value adjustments was used.

Supplemental information

Document S1. Figure S1-S6.

Table S1. Sequencing statistics. Related to Figure 1.

Table S2. Summary of sequencing statistics. Related to Figure 1.

Table S3. Assignment of neural vs non-neural cell fate to the human developmental transcriptome dataset. Related to Figure 2.

Table S4. Cerebellar regional marker genes. Related to Figure 3.

Table S5. Differentially expressed genes between 10x and Parse. Related to Figure 4.

Table S6. GSEA results for GPC cluster. Related to Figure 4.

Table S7. Detailed statistical results related to Figures 1, 3, S3, and S5.

References

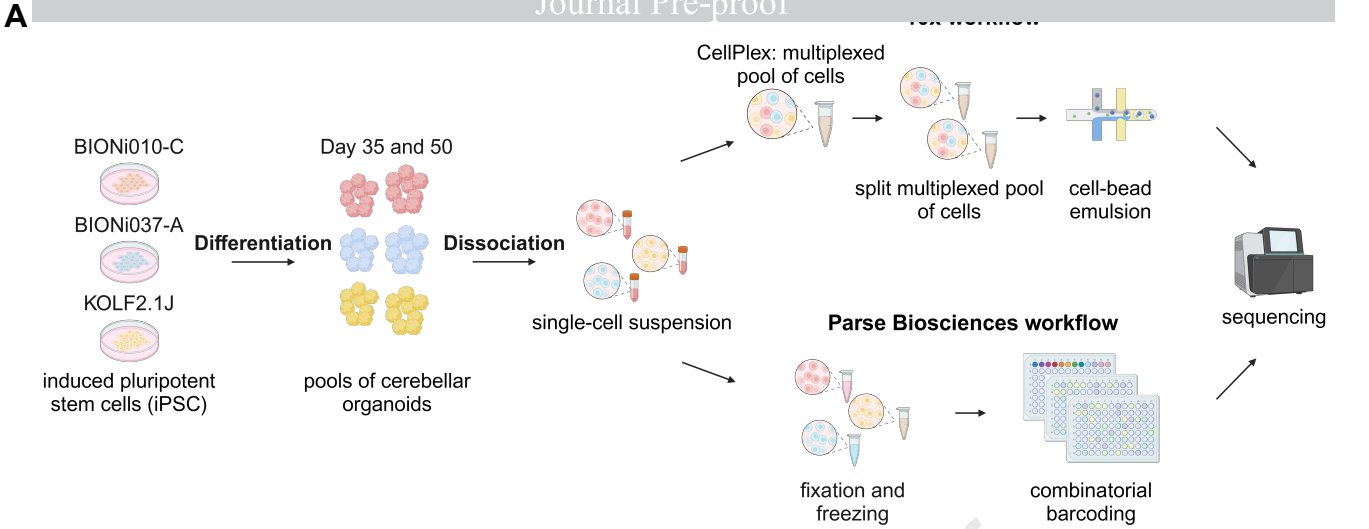
1. Evrony, G.D., Hinch, A.G., and Luo, C. (2021). Applications of Single-Cell DNA Sequencing. Annu Rev Genomics Hum Genet *22*, 171–197. 10.1146/annurev-genom-111320-090436.
2. Yasen, A., Aini, A., Wang, H., Li, W., Zhang, C., Ran, B., Tuxun, T., Maimaitinijiati, Y., Shao, Y., Aji, T., and Wen, H. (2020). Progress and applications of single-cell sequencing techniques. Infect Genet Evol *80*, 104198. 10.1016/j.meegid.2020.104198.
3. Luecken, M.D., and Theis, F.J. (2019). Current best practices in single-cell RNA-seq analysis: a tutorial. Mol Syst Biol *15*, e8746. 10.15252/msb.20188746.
4. Rosenberg, A.B., Roco, C.M., Muscat, R.A., Kuchina, A., Sample, P., Yao, Z., Graybuck, L.T., Peeler, D.J., Mukherjee, S., Chen, W., et al. (2018). Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science *360*, 176–182. 10.1126/science.aam8999.
5. Cheng, J., Liao, J., Shao, X., Lu, X., and Fan, X. (2021). Multiplexing Methods for Simultaneous Large-Scale Transcriptomic Profiling of Samples at Single-Cell Resolution. Adv Sci (Weinh) *8*, e2101229. 10.1002/advs.202101229.
6. Xie, Y., Chen, H., Chellamuthu, V.R., Lajam, A.B.M., Albani, S., Low, A.H.L., Petretto, E., and Behmoaras, J. (2024). Comparative Analysis of Single-Cell RNA Sequencing Methods with and without Sample Multiplexing. International Journal of Molecular Sciences *25*, 3828. 10.3390/ijms25073828.
7. Eichmüller, O.L., and Knoblich, J.A. (2022). Human cerebral organoids — a new tool for clinical neurology research. Nature Reviews Neurology *18*, 661–680. 10.1038/s41582-022-00723-9.
8. Sarieva, K., Kagermeier, T., Khakipoor, S., Atay, E., Yentur, Z., Becker, K., and Mayer, S. (2023). Human brain organoid model of maternal immune activation identifies radial glia cells as selectively vulnerable. Mol Psychiatry *28*, 5077–5089. 10.1038/s41380-023-01997-1.
9. Kagermeier, T., Hauser, S., Sarieva, K., Laugwitz, L., Groeschel, S., Janzarik, W.G., Yentür, Z., Becker, K., Schöls, L., Krägeloh-Mann, I., and Mayer, S. (2024). Human organoid model of pontocerebellar hypoplasia 2a recapitulates brain region-specific size differences. Disease Models & Mechanisms *17*, dmm050740. 10.1242/dmm.050740.
10. Giorgi, C., Lombardozzi, G., Ammannito, F., Scenna, M.S., Maceroni, E., Quintiliani, M., d'Angelo, M., Cimini, A., and Castelli, V. (2024). Brain Organoids: A Game-Changer for Drug Testing. Pharmaceutics *16*. 10.3390/pharmaceutics16040443.
11. Corsini, N.S., and Knoblich, J.A. (2022). Human organoids: New strategies and methods for analyzing human development and disease. Cell *185*, 2756–2769. 10.1016/j.cell.2022.06.051.
12. Schmahmann, J.D. (2019). The cerebellum and cognition. Neurosci Lett *688*, 62–75. 10.1016/j.neulet.2018.07.005.
13. Zhang, P., Duan, L., Ou, Y., Ling, Q., Cao, L., Qian, H., Zhang, J., Wang, J., and Yuan, X. (2023). The cerebellum and cognitive neural networks. Front Hum Neurosci *17*, 1197459. 10.3389/fnhum.2023.1197459.
14. Sathyanesan, A., Zhou, J., Scafidi, J., Heck, D.H., Sillitoe, R.V., and Gallo, V. (2019). Emerging connections between cerebellar development, behaviour and complex brain disorders. Nat Rev Neurosci *20*, 298–313. 10.1038/s41583-019-0152-2.
15. Mapelli, L., Soda, T., D'Angelo, E., and Prestori, F. (2022). The Cerebellar Involvement in Autism Spectrum Disorders: From the Social Brain to Mouse Models. Int J Mol Sci *23*, 3894. 10.3390/ijms23073894.
16. Harada, H., Sato, T., and Nakamura, H. (2016). Fgf8 signaling for development of the midbrain and hindbrain. Dev Growth Differ *58*, 437–445. 10.1111/dgd.12293.

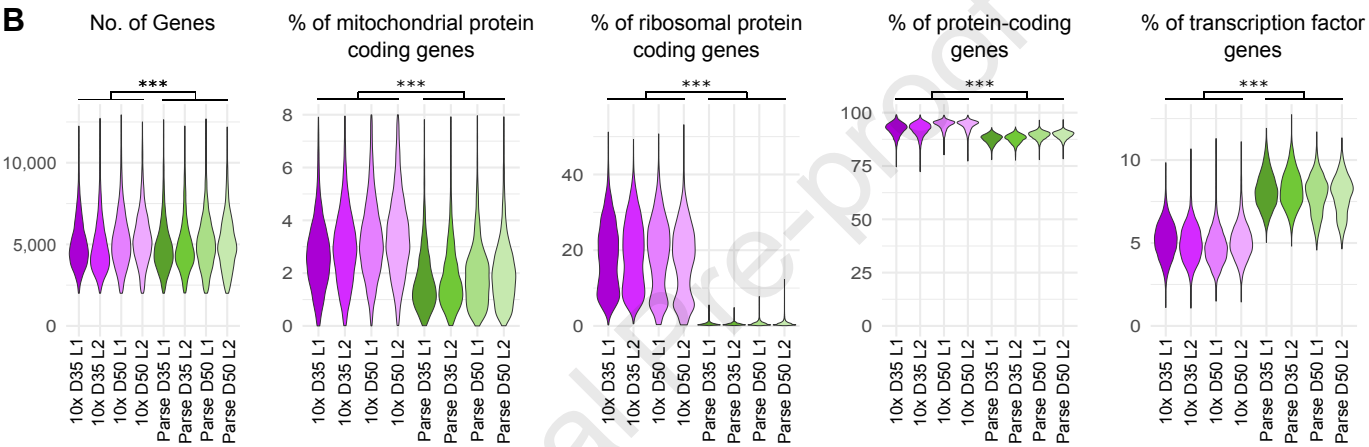
17. Lowenstein, E.D., Cui, K., and Hernandez-Miranda, L.R. (2023). Regulation of early cerebellar development. FEBS J *290*, 2786–2804. 10.1111/febs.16426.
18. Leto, K., Arancillo, M., Becker, E.B.E., Buffo, A., Chiang, C., Ding, B., Dobyns, W.B., Dusart, I., Haldipur, P., Hatten, M.E., et al. (2016). Consensus Paper: Cerebellar Development. The Cerebellum *15*, 789–828. 10.1007/s12311-015-0724-2.
19. van Essen, M.J., Apsley, E.J., Riepsaame, J., Xu, R., Northcott, P.A., Cowley, S.A., Jacob, J., and Becker, E.B.E. (2024). PTCH1-mutant human cerebellar organoids exhibit altered neural development and recapitulate early medulloblastoma tumorigenesis. Dis Model Mech *17*, dmm050323. 10.1242/dmm.050323.
20. Ballabio, C., Anderle, M., Gianesello, M., Lago, C., Miele, E., Cardano, M., Aiello, G., Piazza, S., Caron, D., Gianno, F., et al. (2020). Modeling medulloblastoma in vivo and with human cerebellar organoids. Nat Commun *11*, 583. 10.1038/s41467-019-13989-3.
21. Atamian, A., Birtele, M., Hosseini, N., Nguyen, T., Seth, A., Del Dosso, A., Paul, S., Tedeschi, N., Taylor, R., Coba, M.P., et al. (2024). Human cerebellar organoids with functional Purkinje cells. Cell Stem Cell *31*, 39–51 e36. 10.1016/j.stem.2023.11.013.
22. Nayler, S., Agarwal, D., Curion, F., Bowden, R., and Becker, E.B.E. (2021). High-resolution transcriptional landscape of xeno-free human induced pluripotent stem cell-derived cerebellar organoids. Sci Rep *11*, 12959. 10.1038/s41598-021-91846-4.
23. Silva, T.P., Fernandes, T.G., Nogueira, D.E.S., Rodrigues, C.A.V., Bekman, E.P., Hashimura, Y., Jung, S., Lee, B., Carmo-Fonseca, M., and Cabral, J.M.S. (2020). Scalable Generation of Mature Cerebellar Organoids from Human Pluripotent Stem Cells and Characterization by Immunostaining. J Vis Exp, e61143. 10.3791/61143.
24. Muguruma, K., Nishiyama, A., Kawakami, H., Hashimoto, K., and Sasai, Y. (2015). Self-organization of polarized cerebellar tissue in 3D culture of human pluripotent stem cells. Cell Rep *10*, 537–550. 10.1016/j.celrep.2014.12.051.
25. Zheng, W., Chung, L.M., and Zhao, H. (2011). Bias detection and correction in RNA-Sequencing data. BMC Bioinformatics *12*, 290. 10.1186/1471-2105-12-290.
26. Vértesy, Á., Eichmüller, O.L., Naas, J., Novatchkova, M., Esk, C., Balmaña, M., Ladstaetter, S., Bock, C., Von Haeseler, A., and Knoblich, J.A. (2022). Gruffi: an algorithm for computational removal of stressed cells from brain organoid transcriptomic datasets. The EMBO Journal *41*, e111118. 10.15252/embj.2022111118.
27. Bhaduri, A., Andrews, M.G., Mancia Leon, W., Jung, D., Shin, D., Allen, D., Jung, D., Schmunk, G., Haeussler, M., Salma, J., et al. (2020). Cell stress in cortical organoids impairs molecular subtype specification. Nature *578*, 142–148. 10.1038/s41586-020-1962-0.
28. Qian, X., Song, H., and Ming, G.L. (2019). Brain organoids: advances, applications and challenges. Development *146*, dev166074. 10.1242/dev.166074.
29. Cao, J., O'Day, D.R., Pliner, H.A., Kingsley, P.D., Deng, M., Daza, R.M., Zager, M.A., Aldinger, K.A., Blecher-Gonen, R., Zhang, F., et al. (2020). A human cell atlas of fetal gene expression. Science *370*, eaba7721. 10.1126/science.aba7721.
30. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e3529. 10.1016/j.cell.2021.04.048.
31. Olson, E.N. (1990). MyoD family: a paradigm for development. Genes Dev *4*, 1454–1461. 10.1101/gad.4.9.1454.
32. Itoh, N. (2016). FGF10: A multifunctional mesenchymal–epithelial signaling growth factor in development, health, and disease. Cytokine & Growth Factor Reviews *28*, 63–69. https://doi.org/10.1016/j.cytogfr.2015.10.001.

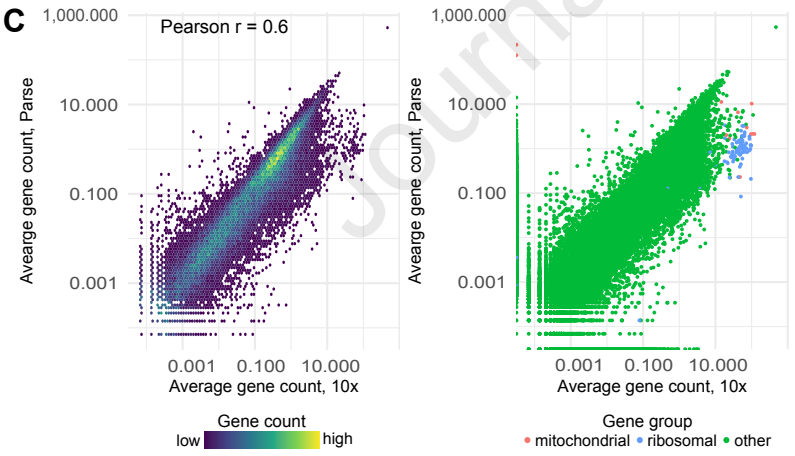
33. Fleck, J.S., Sanchís-Calleja, F., He, Z., Santel, M., Boyle, M.J., Camp, J.G., and Treutlein, B. (2021). Resolving organoid brain region identities by mapping single-cell genomic data to reference atlases. Cell Stem Cell *28*, 1148–1159.e1148. 10.1016/j.stem.2021.02.015.
34. Miller, J.A., Ding, S.-L., Sunkin, S.M., Smith, K.A., Ng, L., Szafer, A., Ebbert, A., Riley, Z.L., Royall, J.J., Aiona, K., et al. (2014). Transcriptional landscape of the prenatal human brain. Nature *508*, 199–206. 10.1038/nature13185.
35. Aldinger, K.A., Thomson, Z., Phelps, I.G., Haldipur, P., Deng, M., Timms, A.E., Hirano, M., Santpere, G., Roco, C., Rosenberg, A.B., et al. (2021). Spatial and cell type transcriptional landscape of human cerebellar development. Nat Neurosci *24*, 1163–1175. 10.1038/s41593-021-00872-y.
36. Haldipur, P., Aldinger, K.A., Bernardo, S., Deng, M., Timms, A.E., Overman, L.M., Winter, C., Lisgo, S.N., Razavi, F., Silvestri, E., et al. (2019). Spatiotemporal expansion of primary progenitor zones in the developing human cerebellum. Science *366*, 454–460. 10.1126/science.aax7526.
37. Sepp, M., Leiss, K., Murat, F., Okonechnikov, K., Joshi, P., Leushkin, E., Spanig, L., Mbengue, N., Schneider, C., Schmidt, J., et al. (2024). Cellular development and evolution of the mammalian cerebellum. Nature *625*, 788–796. 10.1038/s41586-023-06884-x.
38. Chen, X., Shi, C., He, M., Xiong, S., and Xia, X. (2023). Endoplasmic reticulum stress: molecular mechanism and therapeutic targets. Signal Transduct Target Ther *8*, 352. 10.1038/s41392-023-01570-w.
39. Jager, R., Bertrand, M.J., Gorman, A.M., Vandenabeele, P., and Samali, A. (2012). The unfolded protein response at the crossroads of cellular life and death during endoplasmic reticulum stress. Biol Cell *104*, 259–270. 10.1111/boc.201100055.
40. Bertucci, T., Bowles, K.R., Lotz, S., Qi, L., Stevens, K., Goderie, S.K., Borden, S., Oja, L.M., Lane, K., Lotz, R., et al. (2023). Improved Protocol for Reproducible Human Cortical Organoids Reveals Early Alterations in Metabolism with MAPT Mutations. bioRxiv. 10.1101/2023.07.11.548571.
41. Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nature Methods *14*, 1083–1086. 10.1038/nmeth.4463.
42. Camp, J.G., and Treutlein, B. (2017). Human organomics: a fresh approach to understanding human development using single-cell transcriptomics. Development *144*, 1584–1587. 10.1242/dev.150458.
43. Phan, H.V., Van Gent, M., Drayman, N., Basu, A., Gack, M.U., and Tay, S. (2021). High-throughput RNA sequencing of paraformaldehyde-fixed single cells. Nature Communications *12*, 5636. 10.1038/s41467-021-25871-2.
44. Zhao, Y.T., Kwon, D.Y., Johnson, B.S., Fasolino, M., Lamonica, J.M., Kim, Y.J., Zhao, B.S., He, C., Vahedi, G., Kim, T.H., and Zhou, Z. (2018). Long genes linked to autism spectrum disorders harbor broad enhancer-like chromatin domains. Genome Res *28*, 933–942. 10.1101/gr.233775.117.
45. Seigfried, F.A., and Britsch, S. (2024). The Role of Bcl11 Transcription Factors in Neurodevelopmental Disorders. Biology (Basel) *13*. 10.3390/biology13020126.
46. Dini, G., Verrotti, A., Gorello, P., Soliani, L., Cordelli, D.M., Antona, V., Mencarelli, A., Colavito, D., and Prontera, P. (2023). NFIA haploinsufficiency: case series and literature review. Front Pediatr *11*, 1292654. 10.3389/fped.2023.1292654.
47. Chen, K.S., Lim, J.W.C., Richards, L.J., and Bunt, J. (2017). The convergent roles of the nuclear factor I transcription factors in development and cancer. Cancer Lett *410*, 124–138. 10.1016/j.canlet.2017.09.015.

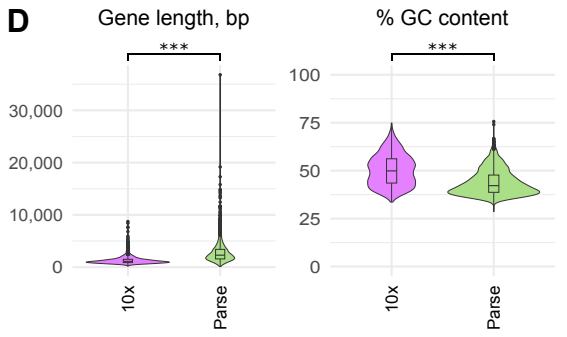
48. Manoli, I., Alesci, S., Blackman, M.R., Su, Y.A., Rennert, O.M., and Chrousos, G.P. (2007). Mitochondria as key components of the stress response. Trends Endocrinol Metab *18*, 190–198. 10.1016/j.tem.2007.04.004.
49. Pasca, A.M., Park, J.Y., Shin, H.W., Qi, Q., Revah, O., Krasnoff, R., O'Hara, R., Willsey, A.J., Palmer, T.D., and Pasca, S.P. (2019). Human 3D cellular model of hypoxic brain injury of prematurity. Nat Med *25*, 784–791. 10.1038/s41591-019-0436-0.
50. Cao, S.S., and Kaufman, R.J. (2014). Endoplasmic reticulum stress and oxidative stress in cell fate decision and human disease. Antioxid Redox Signal *21*, 396–413. 10.1089/ars.2014.5851.
51. Neuschulz, A., Bakina, O., Badillo-Lisakowski, V., Olivares-Chauvet, P., Conrad, T., Gotthardt, M., Kettenmann, H., and Junker, J.P. (2023). A single-cell RNA labeling strategy for measuring stress response upon tissue dissociation. Mol Syst Biol *19*, e11147. 10.15252/msb.202211147.
52. Velasco, S., Kedaigle, A.J., Simmons, S.K., Nash, A., Rocha, M., Quadrato, G., Paulsen, B., Nguyen, L., Adiconis, X., Regev, A., et al. (2019). Individual brain organoids reproducibly form cell diversity of the human cerebral cortex. Nature *570*, 523–527. 10.1038/s41586-019-1289-x.
53. Pasca, A.M., Sloan, S.A., Clarke, L.E., Tian, Y., Makinson, C.D., Huber, N., Kim, C.H., Park, J.Y., O'Rourke, N.A., Nguyen, K.D., et al. (2015). Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture. Nat Methods *12*, 671–678. 10.1038/nmeth.3415.
54. Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nature Biotechnology *27*, 275–280. 10.1038/nbt.1529.
55. Volpato, V., and Webber, C. (2020). Addressing variability in iPSC-derived models of human disease: guidelines to promote reproducibility. Dis Model Mech *13*, dmm042317. 10.1242/dmm.042317.
56. Pantazis, C.B., Yang, A., Lara, E., McDonough, J.A., Blauwendraat, C., Peng, L., Oguro, H., Kanaujiya, J., Zou, J., Sebesta, D., et al. (2022). A reference human induced pluripotent stem cell line for large-scale collaborative studies. Cell Stem Cell *29*, 1685–1702 e1622. 10.1016/j.stem.2022.11.004.
57. Yang, M.J., Sedov, K., Chen, M.Y., Zafar, F., and Schüle, B. (2024). Optical Genome Mapping of the human reference iPSC line KOLF2.1J reveals new smaller structural variants in neurodevelopmental genes. bioRxiv. 10.1101/2024.10.17.618968.
58. Azbukina, N., He, Z., Lin, H.-C., Santel, M., Kashanian, B., Maynard, A., Török, T., Okamoto, R., Nikolova, M., Kanton, S., et al. (2025). Multi-omic human neural organoid cell atlas of the posterior brain. 10.1101/2025.03.20.644368.
59. Amin, N.D., Kelley, K.W., Kaganovsky, K., Onesto, M., Hao, J., Miura, Y., McQueen, J.P., Reis, N., Narazaki, G., Li, T., et al. (2024). Generating human neural diversity with a multiplexed morphogen screen in organoids. Cell Stem Cell *31*, 1831–1846 e1839. 10.1016/j.stem.2024.10.016.
60. Gracia-Diaz, C., Perdomo, J.E., Khan, M.E., Roule, T., Disanza, B.L., Cajka, G.G., Lei, S., Gagne, A.L., Maguire, J.A., Shalem, O., et al. (2024). KOLF2.1J iPSCs carry CNVs associated with neurodevelopmental disorders. Cell Stem Cell *31*, 288–289. 10.1016/j.stem.2024.02.007.
61. Silva, T.P., Bekman, E.P., Fernandes, T.G., Vaz, S.H., Rodrigues, C.A.V., Diogo, M.M., Cabral, J.M.S., and Carmo-Fonseca, M. (2020). Maturation of Human Pluripotent Stem Cell-Derived Cerebellar Neurons in the Absence of Co-culture. Front Bioeng Biotechnol *8*, 70. 10.3389/fbioe.2020.00070.
62. Lancaster, M.A., and Knoblich, J.A. (2014). Generation of cerebral organoids from human pluripotent stem cells. Nat Protoc *9*, 2329–2340. 10.1038/nprot.2014.158.

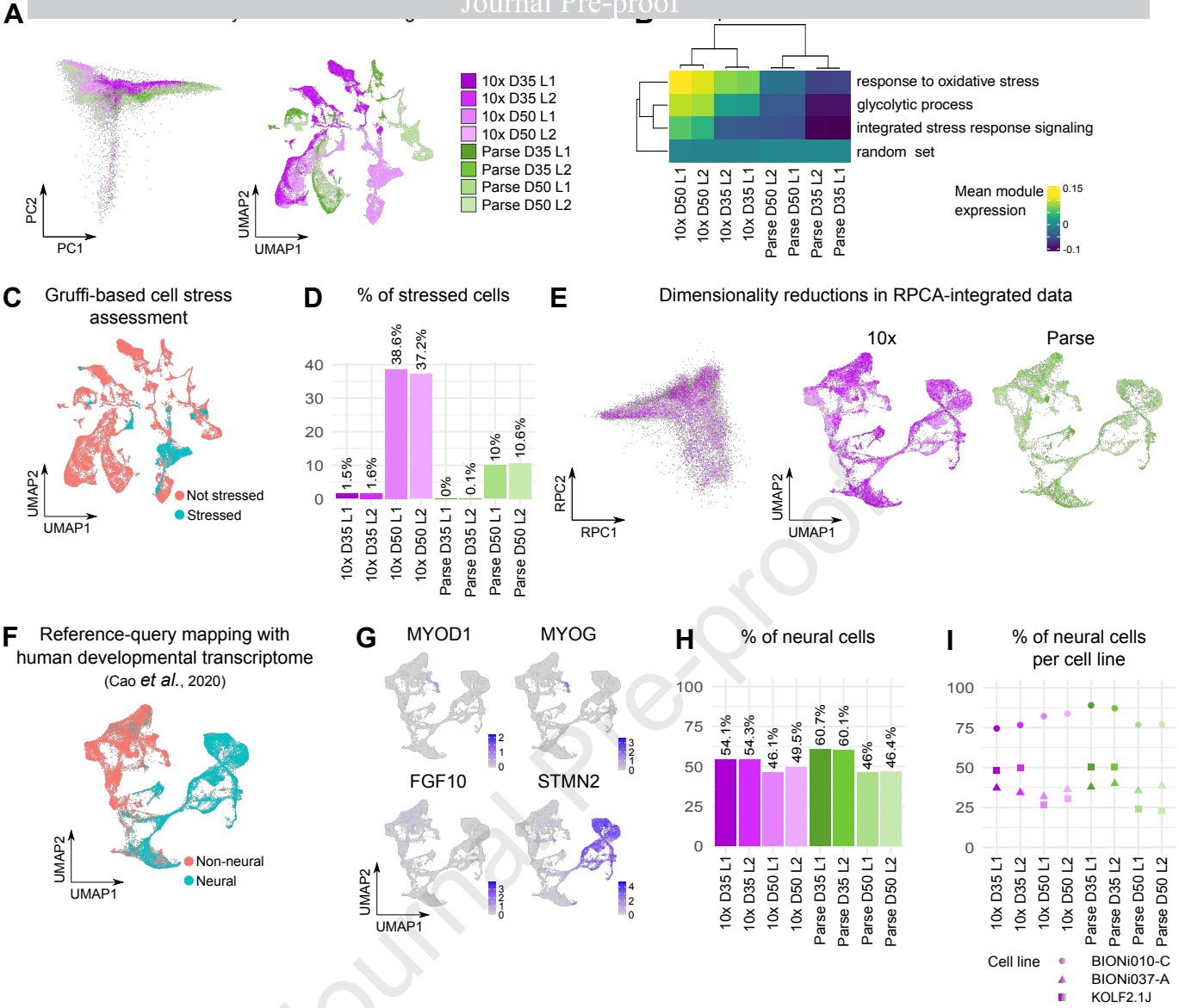
63. Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and Weirauch, M.T. (2018). The Human Transcription Factors. Cell *172*, 650–665. 10.1016/j.cell.2018.01.029.
64. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology *15*, 550. 10.1186/s13059-014-0550-8.
65. Zhu, A., Ibrahim, J.G., and Love, M.I. (2019). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinformatics *35*, 2084–2092. 10.1093/bioinformatics/bty895.
66. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation *2*. 10.1016/j.xinn.2021.100141.
67. Gu, Z., and Hübschmann, D. (2023). SimplifyEnrichment: A Bioconductor Package for Clustering and Visualizing Functional Enrichment Results. Genomics, Proteomics & Bioinformatics *21*, 190–202. 10.1016/j.gpb.2022.04.008.
68. Van De Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., Seurinck, R., Saelens, W., Cannoodt, R., Rouchon, Q., et al. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nature Protocols *15*, 2247–2276. 10.1038/s41596-020-0336-2.

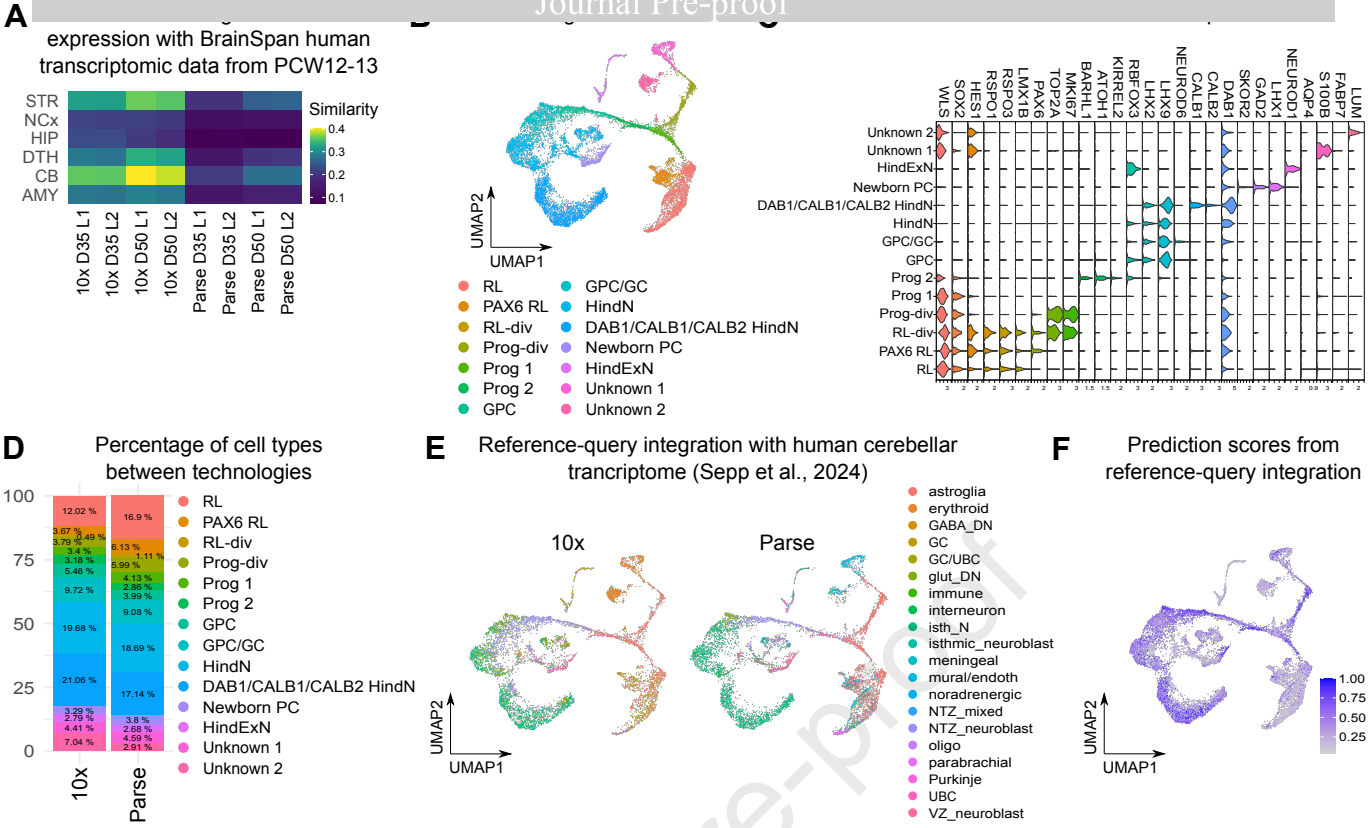


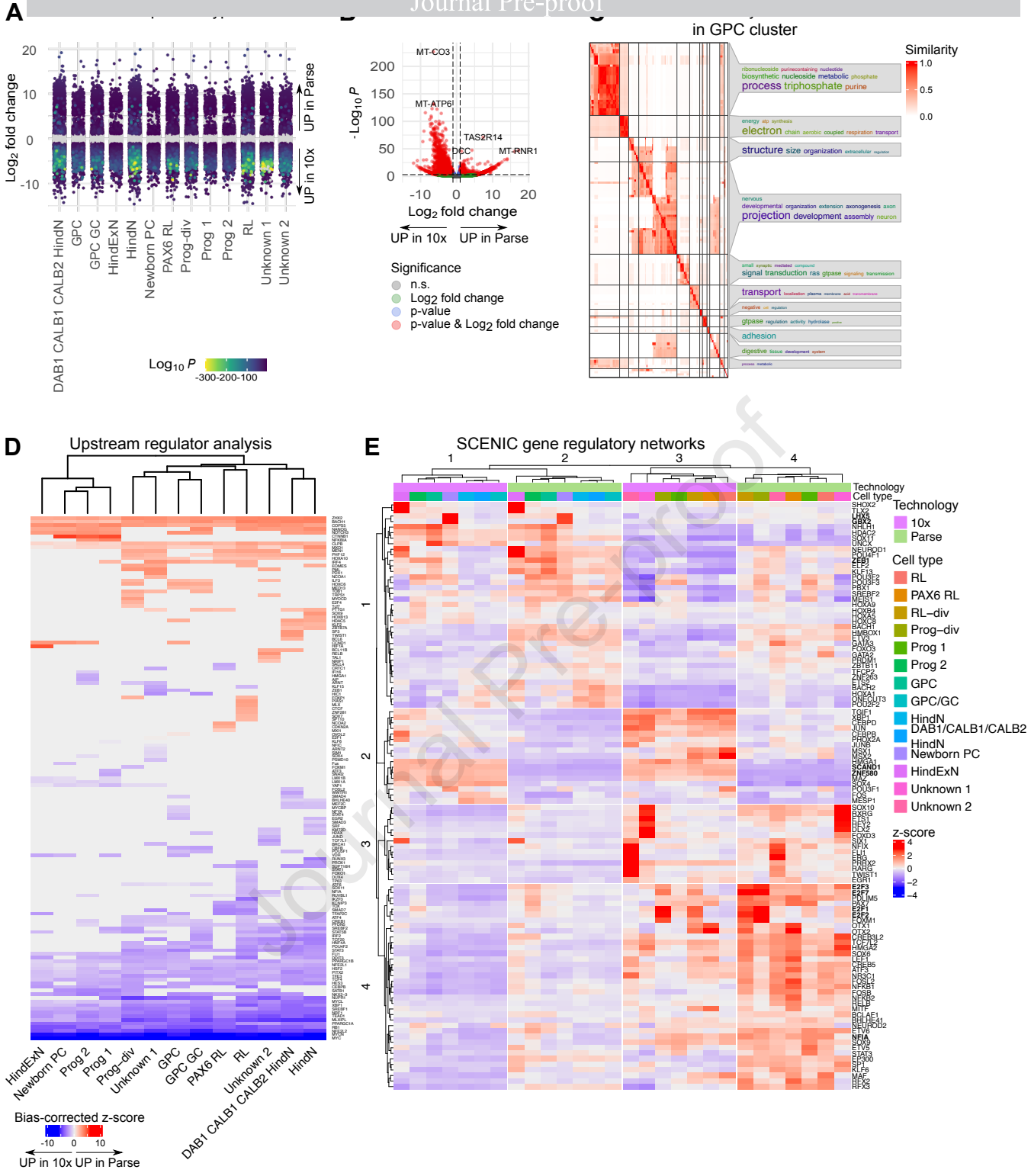












Highlights

10x yields more ribosomal, mitochondrial, and nuclear gene reads than Parse

The Parse workflow results in lower cell stress than 10x.

Data from both technologies can be successfully integrated.

10x and Parse capture the same cellular diversity in human cerebellar organoids.