PIPseq™ Reveals Neuronal and Glial Cell Diversity in Freshly Dissociated Mouse Brain

WHITE PAPER

Introduction

Single cell RNA sequencing (scRNA-seq) is an invaluable technique for elucidating the cellular diversity and function of the central nervous system (CNS). scRNA-seq of the CNS aims to delineate the individual transcriptional profiles of cells that are transcriptionally distinct, yet too close in proximity to be distinguishable by microdissection and traditional sequencing. There are dozens of diverse cell types in the CNS, with hundreds of transcriptionally distinct subtypes. scRNA-seq allows researchers to further understand diverse topics such as molecular taxonomy of the CNS, the molecular mechanisms of neurodegenerative and neurodevelopmental disorders, as well as how the transcriptome changes at distinct time points and disease states. In order to compile the amount of data necessary to reveal these insights, researchers require high-throughput, high-sensitivity, low cost sample preparation tools.

While scRNA-seq has markedly improved our ability to identify novel and rare cell types, the largest individual scRNA-seq datasets represent only about 1% of the over 100 million cells in the adult mouse brain[1]. Pre-templated Instant Partitions (PIPseq™) technology is a novel scRNA-seq technique that is simple, scalable, and accessible due to its simple workflow and competitive cost. Compared to other scRNA-seq technologies, PIPseq provides rapid partitioning of fragile cell types while preserving transcripts that may otherwise be degraded during lengthy collection procedures. This method is highly scalable; inputs ranging from specific brain regions to entire mouse brains can be processed and sequenced in a single reaction, eliminating batch effects and greatly increasing the number of cells that can be captured from individual samples (e.g. animals, patient post mortem tissue, etc.). The scalability of PIPseq also allows the discovery of rare, lowly abundant cell types/subtypes/sub-states that may only be present in fractions of a percent in any particular CNS region. In addition, the scalability and low costs affords the end user the ability to budget for greater biological replication – important for reproducibility of datasets, and again improving power of final datasets with increased cell capture. Here, we present preliminary studies demonstrating the use of the PIPseq T20 3' Single Cell RNA Kit (v2.1) to sequence and characterize the cellular heterogeneity of an adult mouse brain.

Methods

Single cell RNA capture was performed using the PIPseq T20 3' Single Cell Capture and Lysis Kit. A fresh, postnatal day (P) 30 C57BL/6 mouse forebrain (olfactory bulbs and cerebellum removed) was enzymatically and mechanically dissociated[2], then filtered through a 20 µM filter. The cells were counted and a total of 40,000 cells were added to a T20 PIPs reagent tube and processed through the cell capture and lysis workflow. The final PIP-encapsulated cell
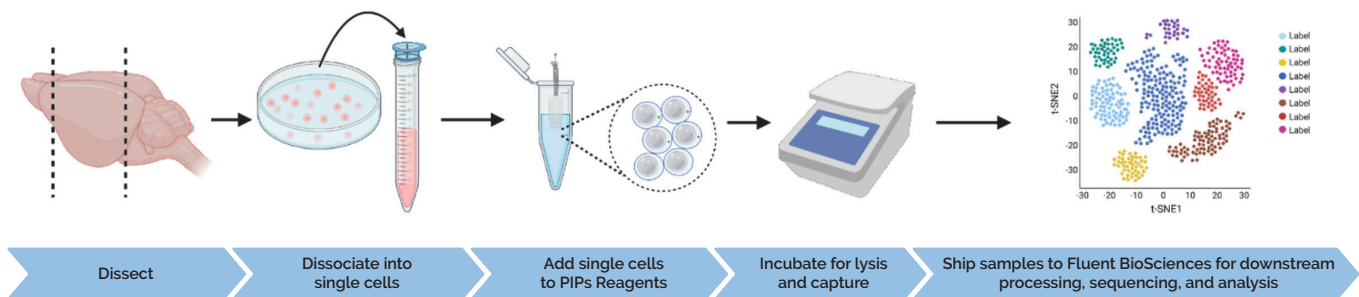
sample was shipped from the point of collection to Fluent BioSciences for further processing, library preparation and next-generation sequencing on an Illumina NextSeq™ 2000, using a P3 reagent cartridge, at a read depth of 14,500 bp paired end reads. For the comparison between PIPseq and 10x Genomics, cells from the same P30 mouse brain were processed via the Chromium Next GEM Single Cell 3' Kit v3.1 workflow, with library production completed following supplier protocol (Illumina) and sequencing on an Illumina NovaSeq 6000 System at a depth of 304 million and 551 million reads, respectively.

Following sequencing, 10x data were processed with Cell Ranger v6.1.0 and PIPseq data were processed with PIPseeker™ v1.0.0, which performs barcode whitelisting and error correction prior to alignment with STARsolo. Because 10x data were sequenced at nearly 2x the read depth per called cell compared to PIPseq, the FASTQs were downsampled to the same depth with PIPseeker to create barcoded FASTQ files and then performed alignment in STARsolo (to mimic Cell Ranger) while holding all variables constant. Cell calling was performed in the same manner for all samples, using Cell Ranger's standard transcripts/cell thresholding approach in combination with EmptyDrops[3]. Filtered matrices were then processed with Seurat 4.0[4] to generate UMAPs for gross visualization. For head-to-head comparisons, downsampled 10x and PIPseq data were integrated using the Harmony batch-correction algorithm[5], and co-clustered to bin related cell types into the same clusters. Differential gene expression analysis was performed separately for each of the merged datasets using the Wilcoxon rank-sum approach without any filtering by p-value or log-fold-change in expression. Automated cell type identification was performed using the ScType tool[6] with a custom mouse brain marker gene database[7].

Results

Here, we demonstrate the performance of PIPseq using fresh mouse brain cells and compare the resulting differentially expressed genes to data generated from the 10x Genomics Chromium instrument and v3.1 reagent kits. Using Fluent BioSciences' Service Lab, our collaborator dissected and dissociated the brain tissue in their own lab and processed the single cells using the PIPseq T20 3' Single Cell RNA Kit (v2.1). After a brief incubation, the sample was shipped overnight to Fluent BioSciences for downstream processing, library preparation, sequencing, and analysis (Figure 1).

Figure 1:

| Dissect | Dissociate into single cells | Add single cells to PIPs Reagents | Incubate for lysis and capture | Ship samples to Fluent BioSciences for downstream processing, sequencing, and analysis |

PIPseq Service Model Workflow. *Fresh, P30 mouse C57BL/6 brains were extracted and the olfactory bulbs and cerebellum removed as shown. The remaining cortex was then dissociated into a single cell suspension, filtered, and added to the PIPs reagent tube. The reagent tube was incubated on the PIPseq dry bath using a pre-programmed lysis and capture profile. After incubation, the tubes were packaged and shipped back to Fluent BioSciences for downstream processing, library preparation, sequencing on a NextSeq 2000, and analysis using Fluent's analytical pipeline, PIPseeker.*

Figure created with BioRender

The analysis revealed 12 distinct clusters, comprehensively representing neurons (clusters 0-2), ependymal cells (3), astrocytes (4), oligodendrocytes (5), D2 medium spiny neurons (6), pericytes (7), choroid plexus epithelial cells (8), microglia (9), and vasculature cells (10-12) (Figure 2A). The corresponding dot plot represents the top 3 differentially expressed genes in each cluster (Figure 2B), which include specific cell marker genes found in the Linnarsson Lab Mouse Brain Atlas (mousebrain.org) for each cluster. These results verify that PIPseq is capable of processing and confidently resolving all expected cell types from freshly dissected mouse forebrain.
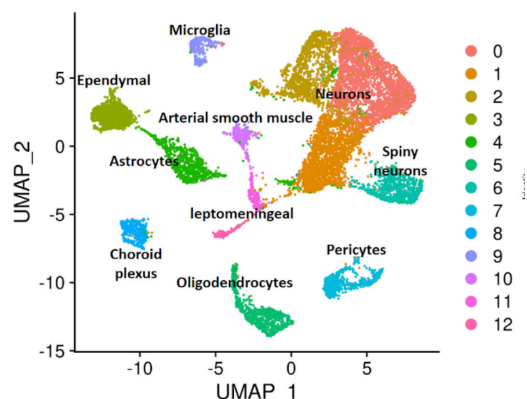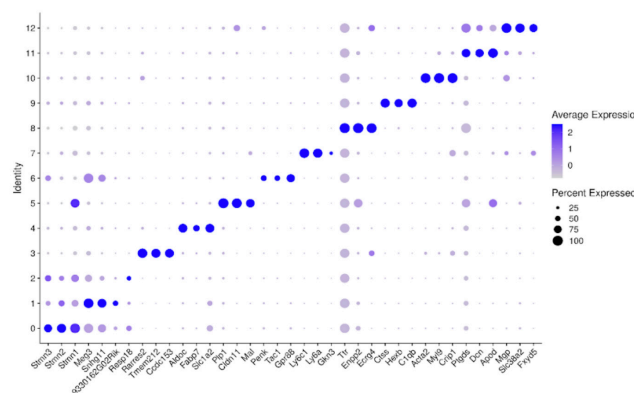
Figure 2:

Figure 2A:

Figure 2B:

PIPseq resolves major cell types from brain-derived single cells. (A) UMAP visualization of scRNA-seq data revealed 12 distinct cell clusters that encompassed all major expected cell types from mouse cerebrum. *(B)* Dot plot of the top differentially expressed genes used to guide the clustering and cell type annotation.

For comparison with 10x scRNA-seq, half of the single cell suspension obtained from P30 C57BL/6 mouse brain was processed through the 10x Chromium Controller instrument and Chromium Next GEM Single Cell 3' Kit v3.1 according to supplier protocol (Figure 3). For a more appropriate comparison between PIPseq and 10x results, we first normalized the number of cell barcodes and reads per cell. This was achieved using the Cell Ranger cell calling strategy on both datasets, and then ensuring that both samples targeted similar cell calls. Because the default cell calling strategy led to too many cell calls for the 10x sample, we set the called cells to match those called for the PIPseq sample. This prevented inclusion of several thousand low-quality barcodes in the 10x sample that interfered with clustering. With identical cell number and read depth, we computed median genes in cells, median transcripts in cells, and sequencing saturation (Table 1). Holding the total number of cells constant between PIPseq and 10x, PIPseq demonstrates better cluster resolution (UMAP silhouette score 0.42) compared to 10x (UMAP silhouette score 0.22).

Table 1:

PLATFORM	MAPPING (%)	CELL BARCODES	READS PER CALLED CELL	MEDIAN GENES IN CELLS	MEDIAN UMIs IN CELLS	SEQUENCING SATURATION
PIPseq	72%	21,203	11,200	778	1106	0.37
10x v3.1	50.4%	21,203	11,200	1254	2081	0.17

Sequencing metrics after normalizing the number of cell barcodes and reads per cell between PIPseq and 10x v3.1.
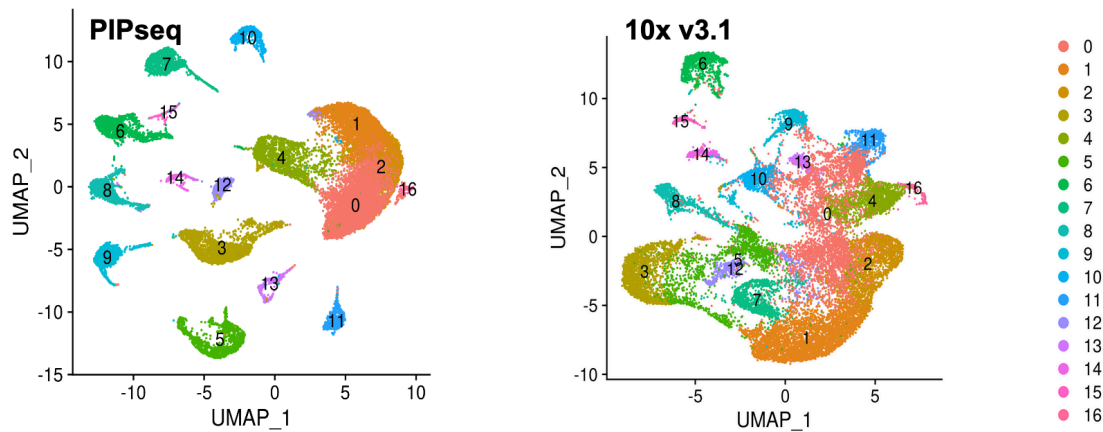
PIPseq demonstrates superior cluster resolution compared to a matched 10x v3.1 sample. *UMAP visualization of the same P30 C57BL/6 mouse brain processed for PIPseq (left) and 10x v3.1 NEXTGems (right). Silhouette scores were higher for PIPseq relative to 10x (0.42 vs. 0.22, respectively), consistent with the improved cluster separation observed in UMAP space.*

In order to compare the differentially expressed genes between the PIPseq and 10x data, we merged the raw matrices and integrated the two datasets using the Harmony batch-correction algorithm, to ensure that matched cell types were binned into the same cluster across both platforms. Differential expression analysis was then performed separately for each dataset without any thresholding. The resulting marker gene tables were then merged so as to only include genes that were detected with at least 1 count per cluster for each platform. Next, we obtained the fold-change, regardless of p-value, and ran a correlation analysis for all markers across clusters, comparing PIPseq and 10x (data not shown). This approach identified a strong correlation between average log2-fold change (log2-FC) of marker genes from 10X and PIPseq (Figure 4, r = 0.861).

We similarly compared the top 10 differentially expressed marker genes for the major identified cell types in heatmap format (Figure 5) and observed that average log2-fold change values were highly similar between platforms (Pearson correlation; r = 0.851, data not shown). Table 2 displays the proportional abundances of cell types identified for each platform along with the top differentially expressed genes for each identified cell type shown in the heatmap. Overall, this demonstrates that the biological data derived from PIPseq is very well aligned with respect to the genes and cell types captured by 10X.
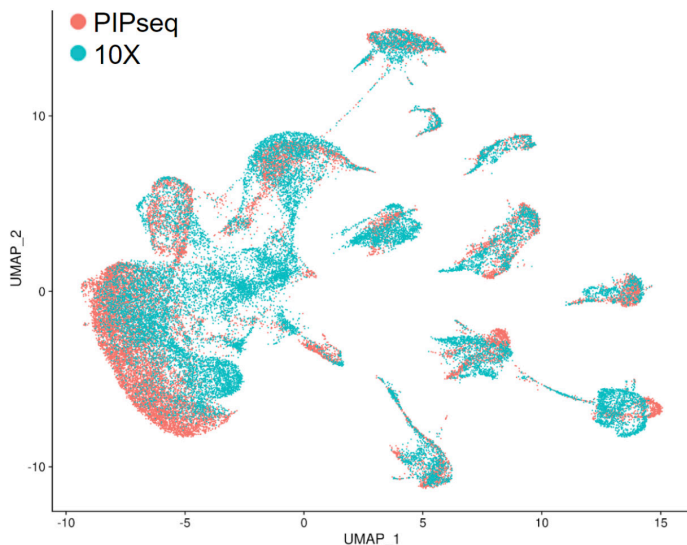

Figure 4:

High concordance of differentially expressed genes between PIPseq and 10X. *PIPseq data was integrated and clustered with matched 10x data generated from the same C57BL/6 mouse brain single cell suspension to assess the similarity of cell types/ subtype frequencies between platforms. There was a strong correlation between the differentially expressed genes derived from both methods (r = 0.861) showing that the gene expression and thus putative biological inferences drawn from each technique are very similar.*

Figure 5:

Cell type marker expression is conserved between PIPseq (left) and 10x (right). *Heatmaps of the top 10 marker genes derived from 10X (right) show consistent gene expression among matched cell types when comparing between platforms. Marker genes displayed here for each cell type are listed in Table 2.*

Table 2:

CELL TYPE	PIPseq (CELL TYPE %)	10X (CELL TYPE %)	MARKER GENE
Excitatory neurons, amygdala	22.4	13.3	Nrgn, C1ql3, Olfm1, Rtn1, Cck, Mef2c, Calm2, Chgb, Ywhah, Ctxn1
Excitatory neurons, cerebral cortex	21.8	11.0	Meg3, Snhg11, Rian, R3hdm1, Kcnq1ot1, Grin2a, Snhg14, Grin2b, Syt7, Ahi1
Astrocytes	9.7	11.4	Slc1a3, Apoe, Aldoc, Plpp3, Atp1a2, Slc1a2, Ntsr2, Gja1, Mt1, Bcan
R-LM border Cck interneurons	2.1	18.7	Tnr, Mbp
Vascular endothelial cells, venous	6.6	6.3	Cldn5, Ly6c1, Flt1, Ly6a, Slco1a4, Itm2a, Cxcl12, Bsg, Pltp, Ptprb
Mature oligodendrocytes	5.9	6.8	Plp1, Mal, Cnp, Mag, Trf, Mobp, Cldn11, Mbp, Cryab, Ermn
D2 medium spiny neurons, striatum	7.0	4.8	Penk, Pde10a, Gpr88, Tac1, Rgs9, Adcy5, Ppp1r1b, Pde1b, Rasd2, Scn4b
Microglia, activated	4.9	4.8	Hexb, C1qa, C1qb, Ctss, C1qc, Cx3cr1, Csf1r, P2ry12, Ctsd, Tyrobp
Choroid plexus epithelial cells	2.6	6.5	Ttr, Enpp2, Ecrg4, Folr1, 2900040C04Rik, Prlr, Igfbp2, Sostdc1, Kcnj13, Calml4
Vascular smooth muscle cells, arterial	4.3	3.8	Acta2, Crip1, Tagln, Tpm2, Myl9, Myh11, Igfbp7, Myl6, Vim, Mylk
Ependymal cells	4.4	2.6	Rarres2, Dbi, Tmem212, Ccdc153, Plac9b, Fam183b, Mia, 1700094D03Rik, Plac9a, Dynlrb2
Red Blood Cells / Other	2.4	3.4	Hba-a1, Hbb-bs, Hbb-bt, Hba-a2, Alas2, Bpgm, Isg20, Ube2l6, Tent5c, Rec114
Pericytes and Vascular leptomeningeal	2.3	3.4	Ptgds, Vtn, Apod, Mgp, Dcn, Ifitm1, Igf2, Col1a2, Rgs5, Nupr1
Neuronal intermediate progenitor cells	2.3	2.0	Hist1h2ap, Hmgb2, Sox11, Ccnd2, Hmgn2, Tubb2b, Igfbpl1, Sox4, Dlx6os1, Ppp1r14b
Perivascular macrophages	1.1	1.0	Lyz2, Cd74, Pf4, Mrc1, Selenop, H2-Eb1, C1qa, H2-Ab1, Fcer1g, Stab1

Cell type percentages and associated marker genes. *Percent abundances for the major cell types observed for PIPseq and 10x are displayed along with the top 10 marker genes derived from the 10x dataset. The cell types shown here match the same order and coloration displayed in the heatmap (Figure 5).*

Conclusion

PIPseq is a powerful, novel single cell sequencing technology that makes scRNA-seq more accessible and affordable for a wide number of researchers. Here, we demonstrate PIPseq's ability to easily partition, process, and characterize 20,000 cells and compare the biological results to another commercial scRNA-seq technology, 10X Genomics. Of note, PIPseq demonstrates superior cluster resolution, is highly compatible with standard tissue dissociation protocols and does not require myelin removal or cell fixation before processing (for such samples from the CNS). In addition, the data presented here is derived from Fluent BioSciences' Service Lab, where tissue is dissociated and processed in one location and then shipped overnight to Fluent for further processing providing additional support for the value of this model. This represents an easy on-ramp to single cell analysis for researchers without previous single cell experience. PIPseq can also be performed entirely in a customer's laboratory with affordable, reagent kits and minimal equipment. By processing 20K cells simultaneously, PIPseq allows unbiased and robust inquiry into the heterogeneity of the cells in the CNS.

References

1. Armand, Ethan J., Junhao Li, Fangming Xie, Chongyuan Luo, and Eran A. Mukamel. 2021. "Single-Cell Sequencing of Brain Cell Transcriptomes and Epigenomes." Neuron 109 (1): 11–26.

2. Hasel, P., I. V. L. Rose, J. S. Sadick, R. D. Kim, and S. A. Liddelow. 2021. "Neuroinflammatory Astrocyte Subtypes in the Mouse Brain." Nature Neuroscience 24 (10).

3. Kaminow B., Yunusov D., and Dobin A. 2022. "STARsolo: accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data." BioRxiv. https://doi.org/10.1101/2021.05.05.442755.

4. Macosko EZ., Basu A., Satija R., et al. 2015. "Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets." Cell 165 (5).

5. Korsunsky, I., Millard, N., Fan, J. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16, 1289–1296 (2019).

6. Ianevski, A., Giri, A.K. & Aittokallio, T. Fully-automated and ultra-fast cell-type identification using specific marker combinations from single-cell transcriptomic data. Nat Commun 13, 1246 (2022).

7. La Manno G, Siletti K, Furlan A, Gyllborg D, Vinsland E, Mossi Albiach A, Mattsson Langseth C, Khven I, Lederer AR, Dratva LM, Johnsson A, Nilsson M, Lönnerberg P, Linnarsson S. Molecular architecture of the developing mouse brain. Nature. 2021 Aug;596(7870):92-96

Visit Fluentbio.com to Learn More