Isolated nuclei from frozen tissue are the superior source for single cell RNA-seq compared with whole cells.

Andrew Jiang[1]*, Klaus Lehnert[1], Suzanne J. Reid[1], Renee R. Handley[1], Jessie C. Jacobsen[1], Syke R. Rudiger[2], Clive J. McLaughlan[2], Paul J. Verma[2], C. Simon Bawden[2] and Russell G. Snell[1]

[1] Applied Translational Genetics Group, School of Biological Sciences, The University of Auckland, Auckland, New Zealand

[2] Molecular Biology and Reproductive Technology Laboratories, South Australian Research and Development Institute, Adelaide, SA, Australia

* To whom correspondence should be addressed. Email: ajia169@aucklanduni.ac.nz

Abstract

The isolation of intact single cells from frozen tissue is a challenge due to the mechanical and physical stress inflicted upon the cell during the freeze-thaw process. Ruptured cells release ambient RNA into the cell suspension, which can become encapsulated into droplets during droplet based single cell RNA-seq library preparation methods. The presence of ambient RNA in droplets has been suggested to impact data quality, however there have been limited reports on single cell RNA-seq data from frozen tissue. Here, we compare the results of single cell RNA-seq derived from disaggregated cells from frozen brain tissue with single nuclei RNA-seq derived from purified nuclei of identical tissue using the 10X Genomics Chromium 3' gene expression assay. Our results indicated that presence of ambient RNA in the cell suspension resulted in single cell RNA-seq data with a 25-fold lower gene count, a 5-fold lower UMI count per cell and a 2-fold lower fraction of reads per cell compared with single nuclei RNA-seq data.

Cell clustering with the single cell RNA-seq data was unable to resolve the heterogeneity of brain cell types. Our conclusion is that nuclei from frozen tissue are the superior substrate for single cell transcriptome analysis.

Introduction

Obtaining fresh tissue is often not practical when conducting single cell transcriptome experiments. As a result, frozen tissue is typically the starting resource for single cell RNA-seq (scRNA-seq). Recovery of individual intact cells from frozen tissue is difficult due to the mechanical and physical stress inflicted upon the cell during the freeze-thaw process. The freeze-thaw process alone can cause cell membrane rupture and result in a release of ambient RNA into cell suspensions which can be partitioned into microfluidic droplets in the library making process. The co-encapsulation of ambient RNA with native cellular RNA can result in cross contamination of transcripts from multiple different cells and cell populations. This will have downstream consequences in clustering and can confound data analysis [1]. Therefore, scRNA-seq library preparation protocols recommend performing single nuclei RNA-seq (snRNA-seq) as an alternative to scRNA-seq for frozen samples [2, 3]. Nuclei are more resistant to damage incurred in the freeze-thaw process. Nevertheless, snRNA-seq presents other problems. First, lower amounts of mRNA are recovered from nuclei than whole cells and this results in RNA libraries with lower numbers of total genes per nuclei and unique molecular identifiers (UMIs) per nuclei which affects clustering [3]. Second, techniques that include quantification of cell surface markers (i.e., CITE-seq, 10X multiome scRNA-seq + ATAC-seq) cannot be performed on nuclei. Lastly, mitochondrial transcripts are completely absent in nuclei.

There have been limited reports of scRNA-seq experiments performed on frozen tissue and therefore our aim was to explore the performance of scRNA-seq on this tissue source. We compare the results of single cell RNA libraries prepared from post-mortem striatal (brain) tissue of four sheep against single nuclei RNA libraries prepared from identical tissue using the 10X Genomics Chromium 3' gene expression assay.

Methods

Frozen brain tissue collection

The tissues used here were isolated from the striatum of four 5-year-old sheep. To obtain the sheep brains a lethal dose of pentobarbitone sodium solution (Lethabarb, 1 mL/2 kg body weight) was administrated intravenously, the brains were removed from the skulls, dissected into samples representing various brain regions and snap frozen initially on dry ice and then in liquid nitrogen. Animal brain weight, body weight and post-mortem delay is detailed in Table 1. Samples were wrapped in tinfoil and stored at -80°C until further use. Tissue collection was performed in accordance with the SARDI/PIRSA Animal Ethics approvals. Preparation of cells and nuclei for scRNA-seq and snRNA-seq respectively was performed on subsamples dissected from the same striatal tissue.

Table 1 Sheep cohort sampling details

Sheep ID	Sex	Age at death (years)	Brain weight (g)	Body weight (kg)	Post-mortem delay (min)[1]
1	F	5	132.46	86.6	41
2	M	5	123.67	74.2	50
3	F	5	128.19	75.6	45
4	M	5	129.24	91.4	62

[1]Post-mortem delay described as the time from death until striatal tissue was dissected and frozen.

Preparation of cells for single cell RNA-seq

Dissociation of neural tissue for scRNA-seq was performed using a modified protocol from 10X Genomics (10X Demonstrated Protocol, CG00055). In brief, 2 mL of 2 mg/mL Papain (Sigma) was added to 30 mg of tissue and incubated at 37 ºC for 20 minutes. The papain solution was then removed without disrupting the tissue, leaving enough solution to cover the tissue. The tissue was dissociated using a Pasteur pipette with 2 mL of Neurobasal media (ThermoFisher Scientific), leaving cell debris to settle on the bottom. The supernatant containing the cells were transferred to a new 15 mL centrifuge tube and centrifuged at 200 xg for 2 minutes. The supernatant was removed, and the cell pellet resuspended in 1 mL Neurobasal media. 5 μL of resuspended cells was stained with 5 μL trypan blue to assess the cellular membrane intactness and number of cells was assessed using the Countess II FL Automated Cell Counter (ThermoFisher Scientific). The cell suspension was diluted by an appropriate volume of Neurobasal media to a final titer of 1000 cells/μL. Cell suspensions from the tissues isolated from animals 1+2 and animals 2+3 were pooled resulting in a total of two cell suspensions.

Preparation of nuclei for single nuclei RNA-seq

Nuclei were extracted from frozen ovine brain tissue utilising an adapted protocol from Krishnaswami et al., 2016 [2]. Briefly, approximately 50-100 mg of brain tissue was transferred to a dounce homogenizer containing 1mL homogenization buffer, HB (250mM sucrose, 25mM KCl, 5mM $MgCl_2$, 10mM Tris buffer pH 8.0, 1µM DTT, 1X protease inhibitor, 0.4 U/μL RNaseIn (ThermoFisher Scientific), 0.2 U/μL Superasein (ThermoFisher Scientific), 0.1 % Triton X-100). Tissue was homogenised using 5 strokes with the loose dounce pestle A, followed by 10-15 strokes of tight dounce pestle B. The homogenate was filtered through a 40 μm strainer into 5 mL Eppendorf tubes and centrifuged at 1000 xg (4 ºC) for 8 minutes. The

supernatant was removed and the pellet was resuspended in 250 μL of HB. A 50% - 29% iodixanol gradient (OptiPrep™ Density Gradient Medium, Sigma) was prepared to allow removal of the myelin layer. 250 μL of 50 % iodixanol was added to the nuclei-HB mixture and slowly layered on top of 500 μL of 29 % iodixanol in a new Eppendorf tube. The resultant gradient was centrifuged at 13,000 g (4 ^{o}C) for 40 minutes. The supernatant and myelin was removed and the purified nuclei pellet was resuspended in 1 mL PBS, 1 % BSA, 0.2 U/μl RNAse inhibitor. 5 μL of resuspended nuclei was stained with 5 μL trypan blue and the quality and number of nuclei was assessed using the Countess II FL Automated Cell Counter (ThermoFisher Scientific). Nuclei suspensions from animals 1+2 and animals 3+4 were pooled resulting in a total of two nuclei suspensions.

Library preparation, sequencing, and quality assessment

Single cell and single nuclei RNA libraries were prepared using the Chromium Next GEM Single Cell 3' Reagent Kit v3.1 (10x Genomics) as per manufacturer's instructions. Single cell RNA and single nuclei RNA libraries were prepared with a targeted cell or nuclei recovery of 8,000. Concentration and quality of the synthesised cDNA were assessed using the Bioanalyzer High-Sensitivity DNA Kit (Agilent, 5067-4627). cDNA was successfully amplified from the suspension containing 8,000 nuclei however cDNA prepared from the suspension containing 8,000 cells was undetectable and therefore library preparation was discontinued at this step. Due to an estimated 10 % of cells in the suspension containing an intact cellular membrane, a second pair of cell suspensions was prepared with a targeted recovery of 80,000 cells to ensure sufficient cDNA could be recovered. cDNA was successfully amplified from these cell suspensions and thus adapter ligation was continued. Sequencing ready libraries were assessed using the Bioanalyzer and sequenced on HiSeq Xten platform to generate on average 384,175,406 (range: 377,815,731 – 392,107,280) 150 bp paired end reads. Alignment of reads was performed using the CellRanger v6.0.1 pipeline with STAR v 2.7.2a to the sheep

Oar_rambouillet_v1.0 reference genome and annotation (Ensembl release 107). Library quality was assessed using CellRanger's web summary report and the ICARUS single cell RNA-seq analysis web server [4]. ICARUS was used to plot unique gene counts per barcode and UMIs mapped per barcode for quality analysis.

Dimensionality reduction, clustering and cell type annotation

Gene count matrices were normalised and scaled by a factor 10,000. Dimensionality reduction with principal component analysis was performed using a set of 2,000 top variable genes (Seurat::FindVariableFeatures) [5]. Clustering was performed with 15 principal components, k-nearest neighbour value of 20 and the Louvain clustering algorithm. Cell type annotation was performed by comparison of cluster marker genes against known striatal cell type markers identified in published scRNA-seq datasets [6-8].

Results

Single cell and single nuclei RNA libraries were prepared using isolated cells or nuclei from frozen sheep striatal tissue. cDNA recovered from the two nuclei suspensions containing 8,000 nuclei showed a cDNA length distribution with maxima at approximately 1200 bp (Figure 2A). The shape of the pherogram for nuclei samples were consistent with 10X Genomics in-house testing for brain samples [9]. The initial attempt to construct a single cell RNA library targeted 8,000 cells. Microscopic examination of intact cellular membranes as screened by trypan blue viability exclusion assay showed 80% and 90% ruptured cells in the two cell suspensions. This fraction of ruptured/unruptured cells is unsurprising given the difficult nature of isolating intact cells from frozen tissue. The freeze-thaw and extraction process induces mechanical and physical stress to the cellular membranes leading to widespread cell lysis and release of mRNA [10, 11]. The cell pools containing both unruptured and ruptured cells were carried forward to cDNA preparation, but no cDNA was detectable in the bioanalyzer pherograms (Figure 1A);

therefore library preparation from these samples was discontinued at this stage. A second pair of cell suspensions was prepared with a targeted recovery of 80,000 cells to ensure sufficient cDNA for the library making process. The percentage of ruptured cells in the second pair of suspensions was estimated at approximately 85% and 93% respectively. cDNA was detected in the Bioanalyzer pherograms (Figure 1A). Although cDNA yield was within the desired range, higher amounts of cDNA in the 350-600 bp range was observed that differed from the size distribution of cDNA from nuclei preparations. Following SPRIselect size selection, sequencing-ready libraries for both nuclei and cell samples had typical cDNA length distribution at 300-1000bp (Figure 1B & 2B).

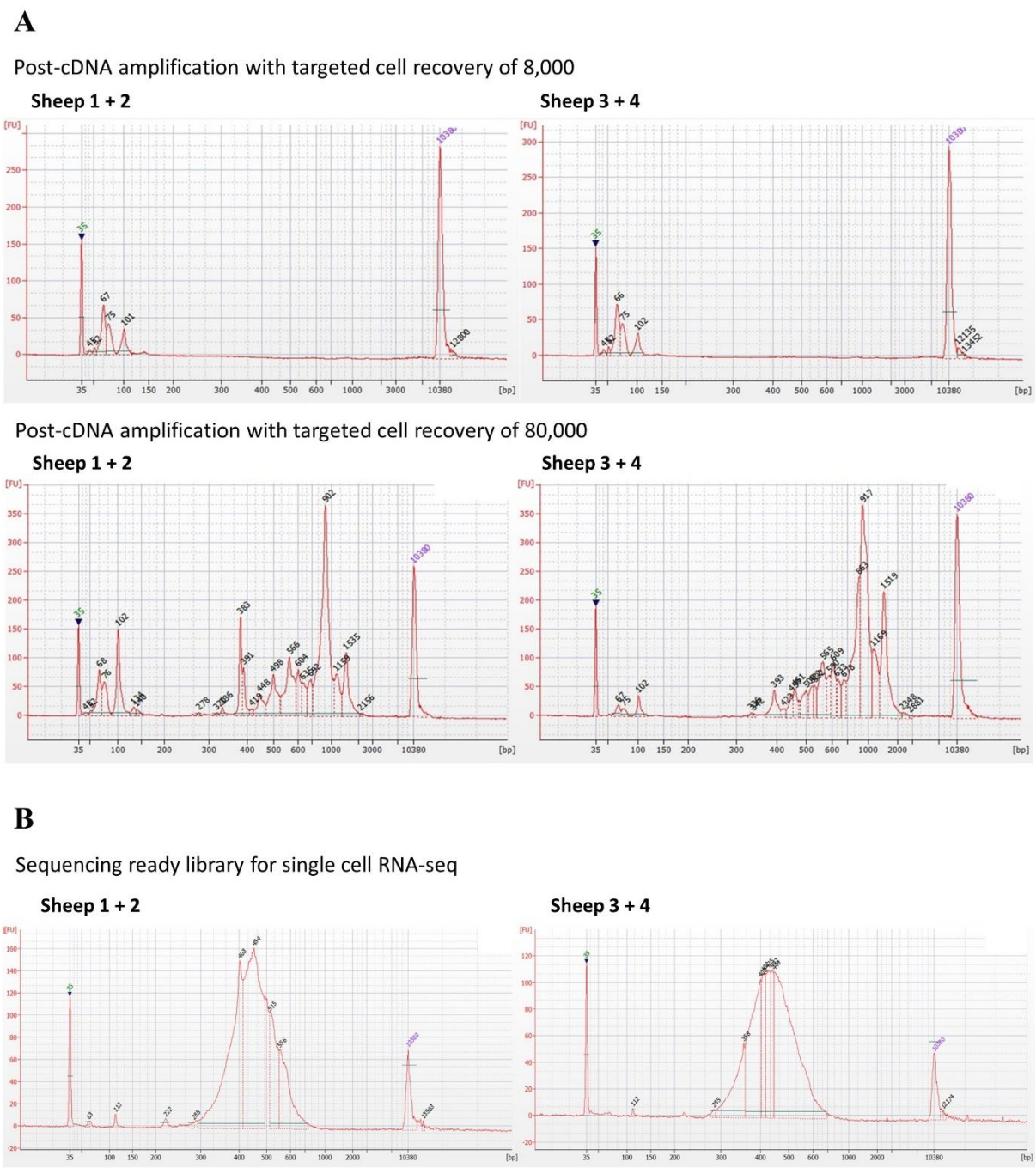


Figure 1 Bioanalyzer pherograms of single cell cDNA and sequencing ready libraries. (A) cDNA libraries with targeted cell recovery of 8,000 and 80,000 were prepared from frozen sheep striatal tissue using the 10X Chromium Next GEM Single Cell 3' v3.1 Kit. No cDNA was detected for the cell suspensions containing 8,000 cells. Higher amounts of cDNA in the 350-600 bp range (indications of ambient RNA) were detected for the cell suspensions containing 80,000 cells. (B) Sequencing ready libraries exhibited a typical cDNA length distribution at 300-1000bp due to SPRIselect size selection.

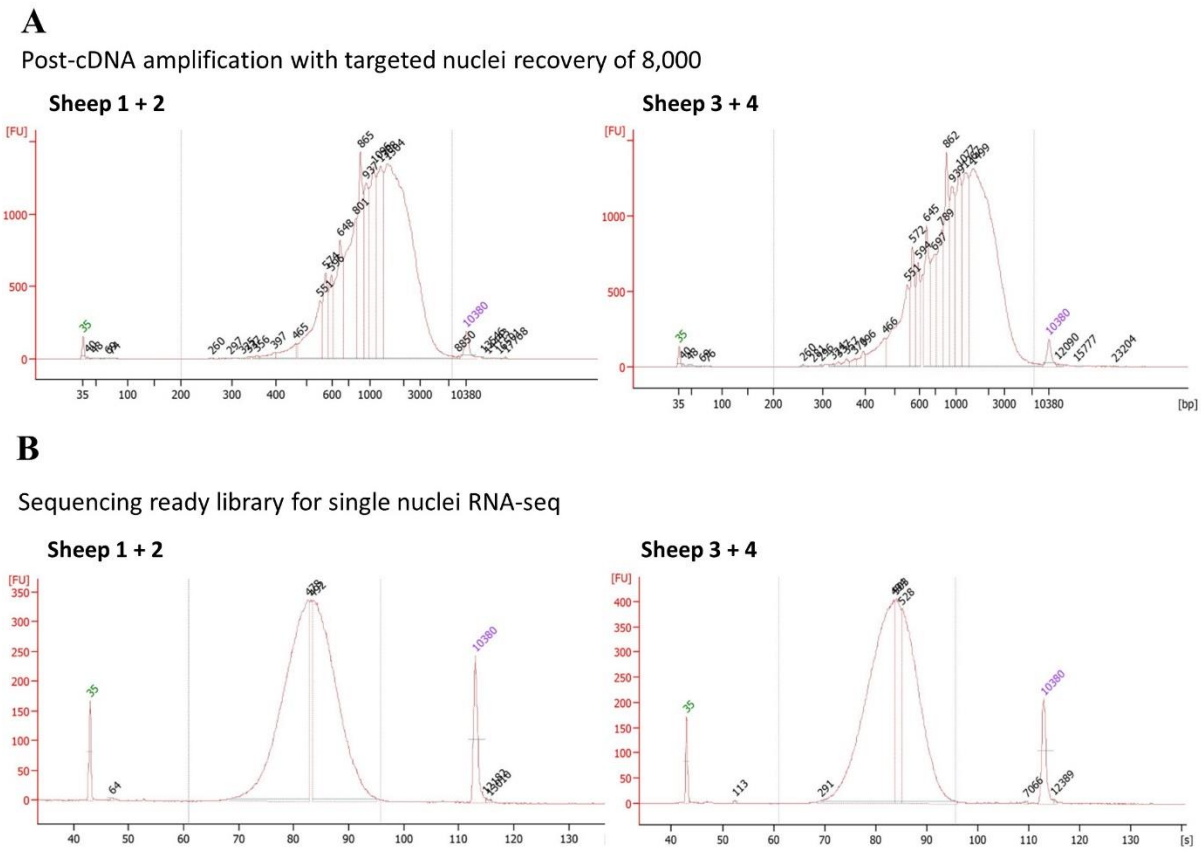


Figure 2 Bioanalyzer pherograms of single nuclei cDNA and sequencing ready libraries. (A) cDNA libraries and (B) sequence libraries were prepared from frozen sheep striatal tissue using the 10x Chromium Next GEM Single Cell 3' v3.1 Kit. The shape of the pherogram was consistent with 10X Genomics in-house testing for brain tissue [8]. (B) Sequencing ready libraries exhibited a typical cDNA length distribution at 300-1000 bp due to SPRIselect size selection.

Reads were aligned to the Ovis aries reference genome with the 10X Genomics CellRanger software. CellRanger summaries indicated excellent read quality with >90 % of barcodes, reads and UMIs containing a Phred score of Q30 or above for single cell and single nuclei RNA libraries (Table 2). 89.6 % and 85.7 % of reads from the single cell RNA libraries, and 96 % and 98.6 % of reads from the single nuclei RNA libraries, were successfully mapped to the reference genome. A ranked plot of every detected 10X barcode in decreasing order of number of UMIs associated with that barcode (barcode rank plot) is used to estimate the number of cells/nuclei in each library. The number of cells/nuclei is estimated by observing a distinctive steep drop-off at two inflection points, known as the cliff and knee (blue and red arrows respectively, Figure 5) that indicates a clear separation between barcodes associated with high UMI count (cell/nuclei associated barcodes) and barcodes associated with low/absent UMI

counts (empty droplets). A characteristic cliff and knee were observed in single nuclei RNA libraries, whereas the steep cliff was not detectable in the data from the single cell RNA libraries. This has resulted in a substantially different number of estimated cells/nuclei between single cell and single nuclei RNA libraries. For the single nuclei RNA libraries, 8,739 and 7,165 barcodes were recovered. In contrast, for single cell RNA libraries, 32,500 barcodes were detected in one library and 2,162 barcodes for the other library. The fraction of reads in barcodes metric is the percentage of reads that are associated with a valid barcode and are mapped to the transcriptome. This fraction was low in single cell RNA libraries (60.8 % and 17 %) compared to single nuclei RNA libraries (90.3 % and 84.6 %). The fraction of reads in barcodes is usually >70 % for a typical single cell RNA and single nuclei RNA libraries [12]. The single cell RNA libraries also contained very low UMI per cell (median UMI counts 504 and 348), and consequently low numbers of genes detected per cell (median gene count 72 and 40). Both metrics were well below single nuclei RNA libraries (median UMI counts 2,612 and 2,386; genes/cell 1,331 and 1,202). Median UMI counts per cell or nucleus usually range from 1,000 – 6,000, while the typical number of genes identified per cell exceeds 1,000 in most cell types and samples [13]. The percentage of reads mapped to the mitochondrial genome was much higher in single cell RNA libraries (42.07 % and 43.77 %) than in single nuclei RNA libraries (1.3 % and 1.0 %). For single nuclei RNA libraries, unsupervised clustering produced distinct clusters that captures the expected heterogeneity of striatal cell types. A total of 13 cell types were identified including oligodendrocytes (expressing marker genes MOG and PLP1), medium spiny neurons (RGS9, PDE10A), oligodendrocyte precursor cells (PCDH15, PDGFRA), microglia (CSF1R, CX3CR1), astrocytes (AQP4, GFAP), neuroblasts (DCX) and interneurons (SST, NPY) (Supplementary 1). In contrast, for single cell RNA libraries, the same clustering procedure resulted in a single large cluster that failed to resolve the cell types in the striatum (Figure 5).

Table 2 Summary statistics for single cell RNA and single nuclei RNA libraries

	Cells		Nuclei	
Sheep IDs	**1+2**	**3+4**	**1+2**	**3+4**
Cells/nuclei targeted	80,000	80,000	8,000	8,000
Estimated Cells/Nuclei	31,974	2,243	8,739	7,165
% Sequencing Saturation[1]	78	93	47	57
Total Reads	382,384,343	384,394,271	377,815,731	392,107,280
Total Genes detected	15,548	8,945	21,520	20,651
Median UMI/Barcode	554	356	2,612	2,210
Mean Reads/Barcode	11,959	17,1375	42,861	49,353
Median Genes/Barcode	64	35	1,331	1,202
Fraction Reads in Barcode	60.9	18.6	90.3	84.6
% Reads Mapped to Genome	89.6	85.7	96.0	98.6
% Reads Mapped to Transcriptome	41.2	27.3	40.0	32.6
% Reads Mapped to Exonic Regions	43.4	29.7	6.5	8.6
% Reads Mapped to Intronic Regions	13.4	15.8	53.4	55.7%
% Reads Mapped to Intergenic Regions	21.5	27.6	30.2	28.5
% Reads to the Mitochondrial Genome	42.1	43.8	1.3	1.0
% Reads Mapped Antisense to Gene	1.0	1.0	19.8	31.1
% Q30 Bases in Barcode	97.2	97.2	97.0	96.6
% Q30 Bases in RNA Read	90.5	90.6	91.3	91.0
% Q30 Bases in UMI	97.3	97.3	96.9	96.6

[1]Sequencing saturation defined as the fraction of reads originating from an already-observed UMI.

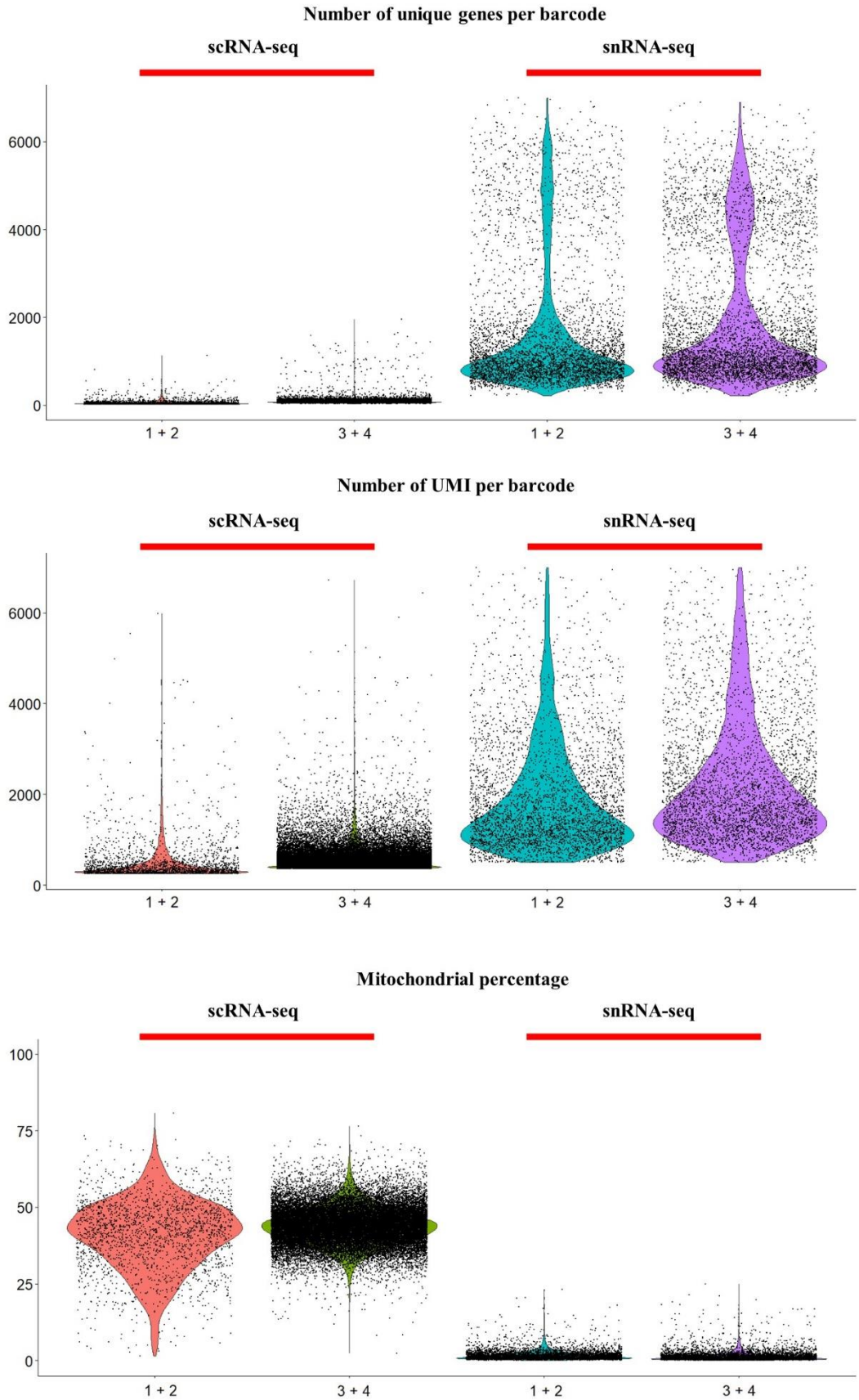


Figure 3 Quality control metrics for single cell RNA and single nuclei RNA libraries. Violin plots indicated single cell RNA libraries on average exhibited a 25-fold lower number of unique genes per barcode and a 5-fold lower UMI per barcode compared to single nuclei RNA libraries. The percentage of reads aligned to the mitochondrial genome (mitochondrial percentage) was also drastically higher in single cell RNA libraries.

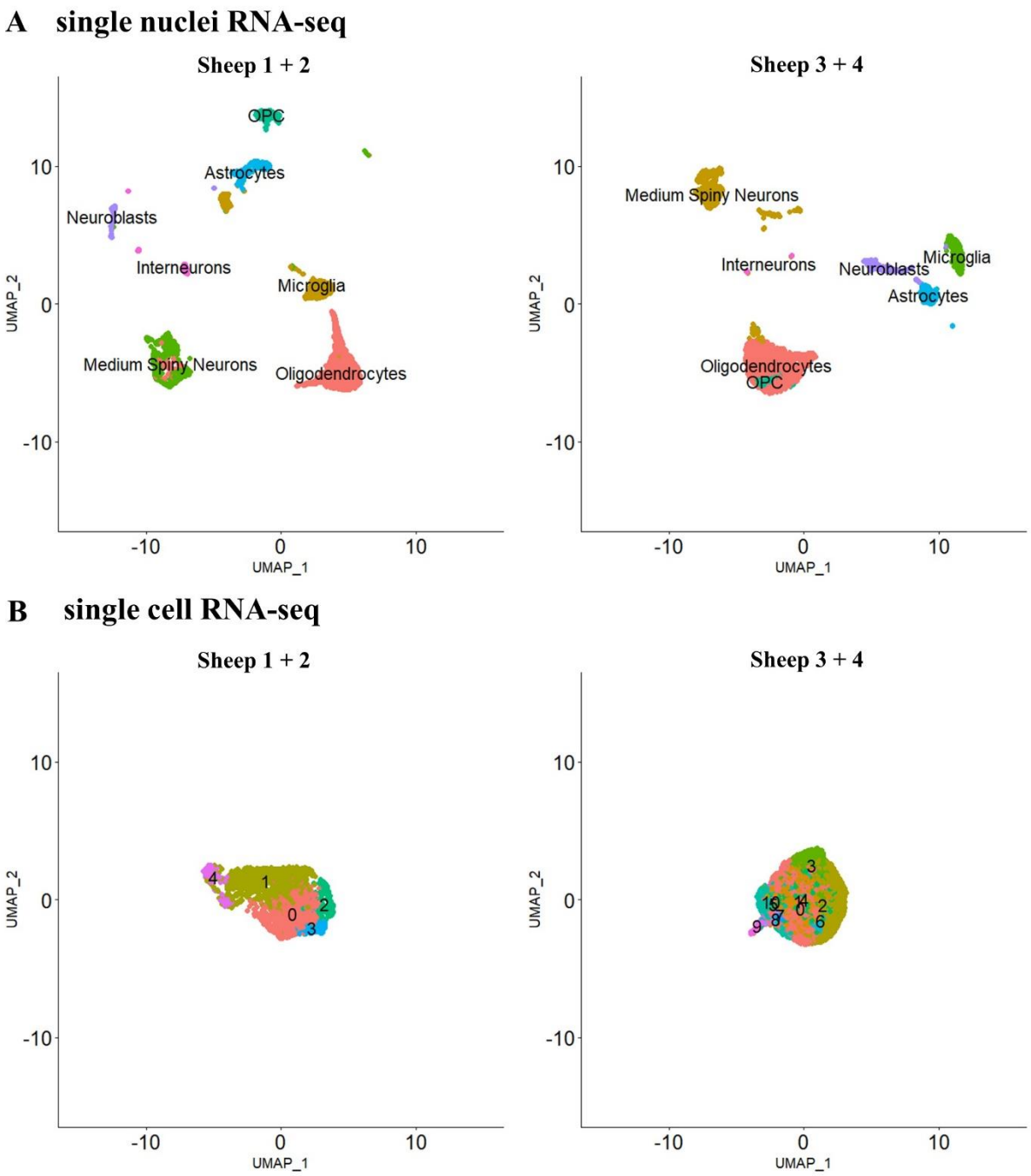


Figure 4 Clustering outcome for single cell RNA and single nuclei RNA libraries. (A) Single nuclei RNA-seq resulted in formation of distinct clusters representing the heterogeneity of gene expression for various cell types present within striatal brain tissue. (B) Single cell RNA-seq was unable to resolve the cellular heterogeneity of striatal cell types.

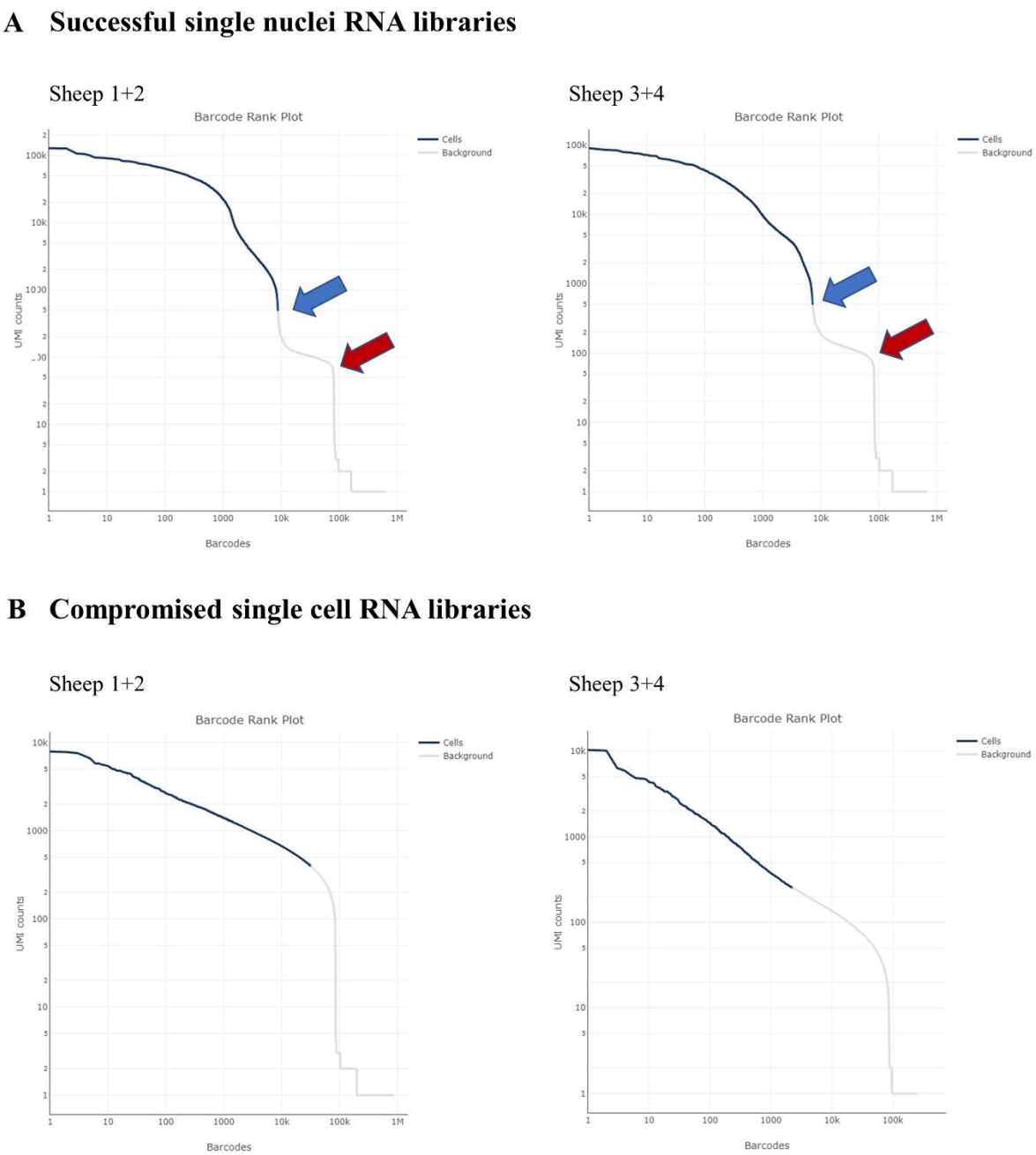


Figure 5 Comparison of barcode rank-plots from a successful single nuclei RNA-seq experiment versus compromised single cell RNA-seq experiments. (A) A barcode rank-plot from a successful experiment should exhibit a cliff (blue arrow) and knee (red arrow) shape indicating a clear separation between cell-associated barcodes and barcodes associated with low quality or empty GEMs. (B) A population of ambient RNA resulted in compromised barcode rank plots with a lack of steep cliff and no clear separation between cell-associated barcodes and empty GEMs.

Discussion

This manuscript assessed the applicability of single cell RNA-seq for frozen brain tissue and compared the outcome of single cell RNA and single nuclei RNA libraries. We observed a very large proportion of cells to have ruptured membranes due to the freeze-thaw dissociation process (average ruptured cells estimated at 87 %). These ruptured cells release RNA into the suspension ('ambient RNA') that can be partitioned with or without intact cells into a droplet gel bead (Gel Bead in Emulsion, GEMs). The ambient RNA is reverse transcribed to cDNA, which was detectable in bioanalyzer pherograms as increased amounts of short cDNA molecules in the 350-600 bp range. This was not evident in the nuclei derived samples. Further, libraries from single cell samples exhibited a compromised barcode rank-plot with a lack of steep cliff and no clear separation between cell-associated barcodes (high UMI counts) and empty GEMs. This lack of the anticipated graphical steep cliff is indicative of GEM encapsulation of cDNA from ambient RNA thereby diluting the number of empty GEMs. Consequently, a large disparity between the number of cell-associated barcodes detected in the single cell RNA libraries (32,500 and 2,162) was observed. The nuclei derived libraries in contrast revealed a consistent detection of nuclei associated barcodes (8,739 and 7,165) that conforms more closely to the initial targeted nuclei recovery of 8,000 per library. A higher proportion of reads that aligned to the mitochondrial genome was also observed in single cell RNA libraries. High mitochondrial content is often an indication of disrupted plasma membrane integrity. Mitochondrial transcripts are enclosed within a double membrane and are more resistant to membrane rupture. Hence, ruptured cells with intact mitochondrial DNA may be partitioned into GEMs, increasing the percentage of mitochondrial transcripts detected.

To address whether the presence of ambient RNA has affected data quality, we assessed common quality control metrics including the number of UMIs per barcode, number of unique

genes per barcode and the fraction of reads in barcodes. Library quality of single cell RNA libraries was severely impacted shown by a 25-fold lower number of genes detected per barcode, a 5-fold lower number of UMIs attributed per barcode and a 2-fold lower fraction of reads attributed to a barcode compared to single nuclei RNA libraries. This has impacted cell clustering with clustering unable to resolve the cellular heterogeneity of striatal cell types. We conclude that maintaining plasma membrane integrity prior to library preparation is essential. Due to the difficulties in maintaining intact cells from frozen tissue, we recommend performing single nuclei RNA-seq as an alternative to scRNA-seq for frozen tissue. Nuclei are more resilient to the mechanical and physical stress induced by the freeze-thaw process and therefore allow the recovery of a high population of intact nuclei that represent the initial composition of the tissue. Our single nuclei RNA-seq datasets showcased a similar cell type composition and diversity to other single nuclei RNA-seq striatal datasets (Table 3) [8, 14]. Single nuclei RNA-seq has been widely applied to characterise transcriptomes from frozen brain tissue [15-22]. In summary, we have demonstrated successful application of single nuclei RNA-seq to frozen brain tissue that generates high quality transcriptomic data reflecting the transcriptome heterogeneity among cell types of the brain.

Table 3 Cell type composition of the striatum as determined by single nuclei RNA-seq datasets

	Sheep 1+2	**Sheep 3+4**	**Malaiya et al., 2021 [14]**	**Lee et al., 2020 [8]**	**Lee et al., 2020 [8]**	**Lee et al., 2020 [8]**
Species	Sheep	Sheep	Mouse	Mouse	Mouse	Human
Cell types						
Medium Spiny neurons	1,334 (15.3%)	1,549 (21.6%)	3,003 (66.4%)	67,144 (61.9%)	35,338 (59.3%)	25,176 (20.1%)
Interneurons	256 (2.9%)	59 (0.8%)	481 (10.6%)	2,943 (2.7%)	1,523 (2.5%)	2,231 (1.8%)
Neuroblasts	195 (2.2%)	485 (6.8%)	NA[1]	3,412 (3.1%)	852 (1.4%)	NA[1]
Oligodendrocytes	4,114 (47.1%)	2,816 (39.3%)	468 (10.3%)	17,333 (16.0%)	13,964 (22.9%)	64,650 (51.6%)
OPC	678 (7.8%)	755 (10.5%)	78 (1.7%)	1,673 (1.5%)	965 (1.6%)	5,069 (4.0%)
Astrocytes	802 (9.2%)	617 (8.6%)	300 (6.6%)	11,696 (10.8%)	5,185 (8.5%)	14,819 (11.8%)
Microglia	1,360 (15.6%)	884 (12.3%)	82 (1.8%)	2,288 (2.1%)	1,415 (2.3%)	8,588 (6.9%)
Endothelial Cells	NA[1]	NA[1]	112 (2.4%)	343 (0.3%)	186 (0.3%)	902 (0.7%)
Mural Cells	NA[1]	NA[1]	NA[1]	396 (0.4%)	198 (0.3%)	1,352 (1.0%)
Ependymal Cells	NA[1]	NA[1]	NA[1]	1,239 (1.1%)	470 (0.7%)	2,384 (1.9%)
Total	8,739	7,165	4,524	108,467	60,924	125,171

[1]Not described in dataset

References

1. Yang, S., et al., *Decontamination of ambient RNA in single-cell RNA-seq with DecontX.* Genome Biology, 2020. **21**(1): p. 57.
2. Krishnaswami, S.R., et al., *Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons.* Nature Protocols, 2016. **11**(3): p. 499-524.
3. Bakken, T.E., et al., *Single-nucleus and single-cell transcriptomes compared in matched cortical cell types.* PLOS ONE, 2018. **13**(12): p. e0209648.
4. Jiang, A., et al., *ICARUS, an interactive web server for single cell RNA-seq analysis.* Nucleic Acids Research, 2022. **50**(W1): p. W427-W433.
5. Satija, R., et al., *Spatial reconstruction of single-cell gene expression data.* Nat Biotechnol, 2015. **33**(5): p. 495-502.
6. Malaiya, S., et al., *Single-Nucleus RNA-Seq Reveals Dysregulation of Striatal Cell Identity Due to Huntington's Disease Mutations.* The Journal of Neuroscience, 2021. **41**(25): p. 5534-5552.
7. Saunders, A., et al., *Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain.* Cell, 2018. **174**(4): p. 1015-1030 e16.
8. Lee, H., et al., *Cell Type-Specific Transcriptomics Reveals that Mutant Huntingtin Leads to Mitochondrial RNA Release and Neuronal Innate Immune Activation.* Neuron, 2020. **107**(5): p. 891-908 e8.
9. *What are the additional peaks in my post-cDNA amplification QC?* ; Available from: https://kb.10xgenomics.com/hc/en-us/articles/217233386.
10. Farrant, J., *Mechanism of Cell Damage During Freezing and Thawing and its Prevention.* Nature, 1965. **205**(4978): p. 1284-1287.
11. McGann, L.E., H.Y. Yang, and M. Walterson, *Manifestations of cell damage after freezing and thawing.* Cryobiology, 1988. **25**(3): p. 178-85.
12. *Interpreting Cell Ranger Web Summary Files for Single Cell Gene Expression Assays*. Available from: https://cdn.10xgenomics.com/image/upload/v1660261286/support-documents/CG000329_TechnicalNote_InterpretingCellRangerWebSummaryFiles_RevA.pdf.
13. Kim, T.H., X. Zhou, and M. Chen, *Demystifying "drop-outs" in single-cell UMI data.* Genome Biology, 2020. **21**(1): p. 196.
14. Malaiya, S., et al., *Single-Nucleus RNA-Seq Reveals Dysregulation of Striatal Cell Identity Due to Huntington's Disease Mutations.* J Neurosci, 2021. **41**(25): p. 5534-5552.
15. Al-Dalahmah, O., et al., *Single-nucleus RNA-seq identifies Huntington disease astrocyte states.* Acta Neuropathol Commun, 2020. **8**(1): p. 19.
16. Mathys, H., et al., *Single-cell transcriptomic analysis of Alzheimer's disease.* Nature, 2019. **570**(7761): p. 332-337.
17. Del-Aguila, J.L., et al., *A single-nuclei RNA sequencing study of Mendelian and sporadic AD in the human brain.* Alzheimer's Research & Therapy, 2019. **11**(1): p. 71.
18. Nguyen, A.T., et al., *APOE and TREM2 regulate amyloid-responsive microglia in Alzheimer's disease.* Acta Neuropathologica, 2020. **140**(4): p. 477-493.
19. Sorrells, S.F., et al., *Immature excitatory neurons develop during adolescence in the human amygdala.* Nature Communications, 2019. **10**(1): p. 2748.
20. Brenner, E., et al., *Single cell transcriptome profiling of the human alcohol-dependent brain.* Human Molecular Genetics, 2020. **29**(7): p. 1144-1153.

21. Grubman, A., et al., *A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation.* Nat Neurosci, 2019. **22**(12): p. 2087-2097.
22. Nagy, C., et al., *Single-nucleus transcriptomics of the prefrontal cortex in major depressive disorder implicates oligodendrocyte precursor cells and excitatory neurons.* Nature Neuroscience, 2020. **23**(6): p. 771-781.

Acknowledgements

The authors wish to thank the incredible on-farm SARDI team for all animal management.

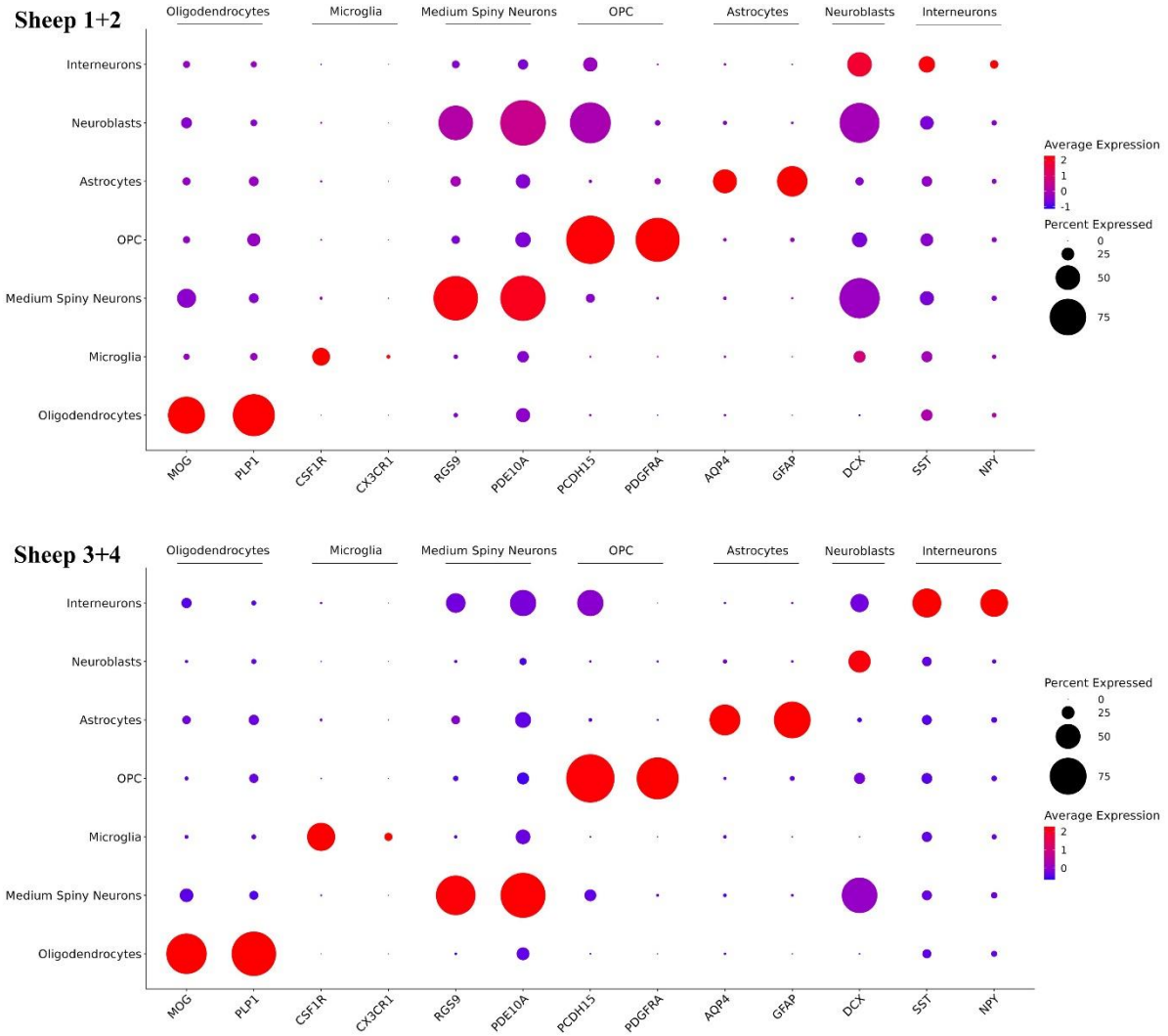
Funding

This work was supported by the CHDI Foundation and New Zealand Ministry of Business Innovation and Employment funding for New Zealand–China Non Communicable Diseases Research [UOOX1601].

Author Contributions

A.J, K.L and R.G.S considered the experiment. S.J.R., R.R.H., S.R.R., C.J.M., collected the samples. A.J performed the single cell RNA-seq experiment and analysed the data. A.J., wrote the manuscript. K.L and R.G.S revised the manuscript. All authors reviewed and approved the final manuscript.

Supplementary



Supplementary 1 Dot plot of enriched cell-type specific markers for annotated cell types in the single cell RNA-seq datasets.