Extraction of nuclei from archived post-mortem tissues for single-nucleus sequencing applications

Malosree Maitra[1,2], Corina Nagy[1,2], Yu Chang Wang[4], Camila Nascimento[5], Mathew Suderman[6,7], Jean-Francois Théroux[2], Naguib Mechawar[2,3], Jiannis Ragoussis[4,8], Gustavo Turecki[2,3,4]*

[1]These authors contributed equally

[2]McGill Group for Suicide Studies, Douglas Mental Health University Institute, [3]Department of Psychiatry, McGill University, [4]McGill University and Genome Quebec Innovation Centre, Department of Human Genetics, McGill University, [5]Department of Psychiatry, University of Sao Paolo, [6]MRC Integrative Epidemiology Unit, University of Bristol, UK, [7]Population Health Sciences, Bristol Medical School, Bristol, UK, [8]Department of Bioengineering, McGill University

***Correspondence to**: Dr. Gustavo Turecki (gustavo.turecki@mcgill.ca)
6875 LaSalle Blvd, Verdun, Québec, Canada H4H 1R3
Tel.: (514) 761-6131 (ext. 2370)

Abstract

Single-cell and single-nucleus sequencing techniques are a burgeoning field with various biological, biomedical, and clinical applications. Numerous high and low-throughput methods have been developed for sequencing the RNA and DNA content of single cells. However, for all these methods the key requirement is high quality input of a single-cell or single-nucleus suspension. Preparing such a suspension is the limiting step when working with fragile, archived tissues of variable quality. This hurdle can prevent such tissues from being extensively investigated with single-cell technologies. We describe a protocol for preparing single-nucleus suspensions within the span of a few hours that reliably works for multiple post-mortem and archived tissues types using standard lab equipment. Moreover, these preparations are compatible with single-nucleus RNA-seq and ATAC-seq using the 10X Genomics' Chromium system.

Keywords: nuclei extraction, single nucleus suspension, snRNA-seq, snATAC-seq

Introduction

Advancements in technology have allowed researchers to preform large-scale transcriptomic studies at the level of a single-cell. Droplet-based cell isolation has become a favorite in the field for its scalability and simplicity of use with either in-house fluidic set-ups[1,2], or commercially available equipment (10x Genomics)[3]. This technique is particularly interesting for tissues with highly heterogeneous cellular compositions like intestine[4], lung[5], spinal cord[6], and brain[7,8]. There has been particular interest in deconvoluting brain architecture and function, which at its base, starts by accurately identifying all the cells types present[9-11]. However, truly harnessing the power of individual cellular transcriptomes comes with assessing differences between those transcriptomes in different physiological states. This is of particular value for complex diseases where multiple genes contribute with additive effects, making it difficult to identify changes in tissues homogenates[12]. The chemical dissociation of tightly interconnected brain cells and other cell-types has been found to alter transcription profiles[13,14]. Given that nuclear transcriptomes closely reflect the cell's cytosolic profile[15,16], isolating the nuclei from brain tissue has proven to be an excellent strategy for single-cell level studies. Likewise, other tissues that have either been frozen for long term storage or that are formed by syncytium, such as in skeletal muscles[17,18], could benefit from this approach.

Development of the Protocol

Numerous protocols for isolating nuclei from brain cells have been published; some rely on the additional purification by fluorescence assisted cell sorting (FACS)[19], which is costly and not readily available for all researchers, while others have made adjustments to the microfluidics component[10], which can also be a limiting to many labs. Our protocol has been developed for use

with the commercially available Chromium™ Single Cell Controller. This is a highly optimized system that allows scalable single-cell capture. We have adapted our protocol to allow the system to efficiently capture nuclei from archived tissue. Our preparation produces stable and easily quantifiable nuclear suspension even when using archived tissue. We have used this approach to successfully compare the gene expression differences in the post-mortem prefrontal cortex of depressed patients who died by suicide compared to psychiatrically healthy controls[20]. The protocol has also been successfully applied to collect single-nucleus transcriptomic data from surgical samples of glioblastoma[21].

Overview of the procedure

The experimental workflow (**Fig. 1**) begins with dounce-homogenization in low concentration detergent for cellular lysis. Integral to the protocol are numerous wash steps to reduce ambient nucleic acid contamination. Determination of the nuclei numbers and concentration are assessed by a cell counter or hemocytometer. The concentration of the nuclei suspension is important to reduce aggregation, particularly in tissues that have undergone long-term storage and are thus more likely to be damaged, fragile and inclined to aggregate.

Comparison with other methods

As previously mentioned, single-nucleus RNA sequencing (snRNA-seq) protocols exist using FACs which is harder to scale, as well as droplet based protocols which use in house set-ups. Early protocols used relatively fresh frozen tissue, which is not available from most tissue banks where samples are likely to have undergone long term storage[9-11]. Moreover, when studying specific phenotypes for which it is hard to obtain tissues, it is not always possible to select for short post-mortem intervals (PMIs) and archival times. Early protocols were also

limitedly applied to the high-quality tissue which may not be an option for answering certain types of research questions.

As with several more recently published potocols[22-24], we have been able to adapt our nuclear prep to be compatible with a 10X Chromium system which is more likely to be available as a service platform. Furthermore, the wet-lab aspect of the protocol will produce nuclei suitable for multiple post-nuclei capture application such as, whole genome sequencing for the study of somatic mutations or Assay for Transposase-Accessible Chromatin (ATAC-seq) as supported by preliminary results from our lab.

Each of the more recently developed protocols have their own strengths and weaknesses and, in some cases, adaptations for specific tissue types such as macro-dissections for white matter regions[24]. The strength of our protocol is that it is effective and mostly unaffected by variations in PMI or archival times of the samples **(Fig 2)**. Some of these protocols also incorporate ultra-centrifugation[25], which is time consuming and requires specialized equipment. We are able to circumvent the additional challenges that arise with archived tissue such as, the fragility of the cells and organelles upon thawing which typically results in large amounts of debris and ambient RNA than can either, interfere with droplet formation, or be integrated into droplet increasing background sequencing noise. Here, we show that brain tissue which has been stored at -80^{o}C for as long as 25 years, can produce high quality single-nuclei suspensions.

Directly applying either the cell preparation protocol or the demonstration protocol for nuclei developed from 10x Genomics did not produce useable results (**Fig. 3**) , but is effective for samples from cell culture. The modifications made here are primarily for use with post-mortem tissue that has been archived for long periods of time but can also be applied to any post

mortem frozen sample. Similar to previous studies[10,11], we did apply a few modifications to the standard bioinformatic analysis to addresses a number of issues which arise with droplet-bases single-nucleus sequencing. First, we assembled a pre-mRNA reference to account for unprocessed transcripts found in the nucleus[26]. Second, given that previous studies have consistently shown fewer identifiable transcripts in glial cells[10,11] we performed customized barcode filtering to include cells with a wider range of UMIs but removed noise. With these minor modifications to the analysis our wet lab approach produced much improved data compared to the demonstration protocol from 10X Genomics.

Experimental design

The most important factors to take into consideration while designing single-cell or single-nucleus RNA-seq experiments is the potential batch effects. Given that the Chromium system only allows for the capture of 8 samples at a time and that for many experiments that total number of samples to be analyzed may be greater than eight, it may be preferable to create a balanced experimental design if possible. This will help limit the effects of batch to batch variability. For example, if two phenotypic or treatment groups are to be compared, it would be ideal to include equal numbers of samples from each group in every batch. Moreover, other potential co-variates to take into consideration include age, PMI, and sex. It may be possible to account for the effects of these variables by matching samples by these parameters within each batch.

In cases where cell-type specific gene-expression data has been previously published, or single-cell or nucleus gene expression datasets are available, these data can be used for comparison to help determine whether that cell-types identified, and single-nucleus transcriptomic profiles

detected reflect, previously published literature. In cases where such datasets are not available it may be informative to prepare bulk samples in parallel or to perform sequencing for fluorescence assisted nuclei sorting (FANS) purified populations of expected cell-types based on known markers to validate the cell-type identification from the single-nucleus transcriptomic data[27]. High-throughput *in situ* hybridization (ISH)[28] and ISH based nuclei sorting[27] have also been used to confirm experimentally determined cell-types from snRNA-seq. In the case of complex tissues, it can be useful to perform careful dissection and even to cryosection the tissue before preparing nuclei to ensure that the approximate cell-type composition for each sample will be comparable[11,24].

Another strategy which has been recently applied to increase cost-effectiveness as well as to aid in batch effect correction is combining male and female samples in a single capture followed by using the expression of sex-specific, X-chromosome genes such as *XIST* and Y-chromosome genes such as *SRY*[29,30], or the chromosome accessibility ratios for sex-chromosome versus autosomes[31] to separate the cells from each sample. Since both samples are captured on the same lane of the microfluidic chip, it may be possible to account for lane to lane variability using this approach. Moreover, the use of cryosections of histological grade tissue blocks may be a good strategy for accounting for uniform input from a micro-anatomicallly heterogeneous region such as the cerebral cortex[11,23].

Expertise needed to implement the protocol

This protocol requires access to a 10X Genomics' Chromium system and corresponding reagents or an in-house droplet based single-nucleus sequencing system. A hemocytometer or cell-counting microscope will be required for determining proper loading concentration. Wet-lab

work will require familiarity with standard molecular biology approaches such as cDNA synthesis, and sequencing library preparation.

Advantages and Limitations

We are unable to get information about cytoplasmic transcription. Some tissue types may require additional processing such as FACS. Representation of all cell types may not be uniform as different cell types are differentially susceptible to lysis during the isolation procedure.

Materials

REAGENTS:

- Tissue samples: This protocol was successfully applied for processing frozen archived post-mortem prefrontal cortex tissue from the Douglas Bell Canada Brain Bank, post-mortem intestinal tissue (with modifications), and surgical samples of tumor tissue[21]. Use of post-mortem tissues was approved by the Institutional Review Board of the Douglas Hospital.
- NP-40 detergent at 10% concentration (Abcam, cat. no. ab142227)
- Bovine Serum Albumin Fraction V (BioShop, cat. no. ALB001.25)
- Tris HCl (BioShop, cat. no. TSS003.5)
- NaCl (BioShop, cat. no. SOD001.1)
- $MgCl_2.6H_2O$ (BioShop, cat. no. MAG510)
- HCl (BioShop, cat. no. HCL333)
  CAUTION: Concentrated HCl is highly corrosive and should be handled wearing PPE inside a fume hood.
- KCl (BioShop, cat. no. POC308)
- KOH (BioShop, cat. no. PHY202)
  CAUTION: Concentrated KOH is highly corrosive and should be handed wearing PPE inside a fume hood.
- Tricine (BioShop, cat. no. TRI001)
- Glycerol (BioShop, cat. no. GLY001)
- Protector RNAse Inhibitor (Millipore Sigma, 3335399001)

CRITICAL: Other RNAse inhibitors may not be compatible with the protocol and may result in low yield of nuclei.

- Optiprep™ Density Gradient Medium (Millipore Sigma, D1556-250)
- Gibco™ PBS, pH 7.4 (1X), (Thermo Fisher Scientific, cat. no. 10010203)
- Deionized water
- Ethanol 100% (Sigma, cat. no. 459836-500ML)

- Trypan Blue Stain (0.4%) (Thermo Fisher Scientific, cat. no. T10282)
- Hoechst stain (Invitrogen, cat. no. H1399)
- Chromium Single Cell 3' Library & Gel Bead Kit v2 (10X Genomics Inc, cat. no. 120237)
- Chromium Single Cell A Chip Kit (10X Genomics Inc, cat. no. 120236)
- SPRIselect Reagent Kit (Beckman Coulter, cat. no. B23318)
- Tween 20 (Bio-Rad, cat. no. 1610781)
- Buffer EB (250mL) (Qiagen, cat. no. 19086)
- Glycerin (glycerol), 50% (v/v) Aqueous Solution (Ricca Chemical Company (or other), cat. no. 3290-32)
- DynaBeads MyOne™ Silane Beads (5mL) (Thermo Fisher, cat. no. 37002D)
- Low TE Buffer (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA, 100mL) (Thermo Fisher, cat. no. 12090-015)
- Nuclease-Free Water (Thermo Fisher, cat. no. AM9937)
- TapeStation High Sensitivity D1000 Sample Buffer (Agilent, cat. no. 5067-5603)
- TapeStation High Sensitivity D1000 ScreenTape (Agilent, cat. no. 5067-5584)

Software for sequence alignment and gene-barcode counting

- CellRanger version 2.1.0
- bcl2fastq2, version 2.19

Software for secondary analysis in R

- Seurat, version 2.3.0 or higher
- mixtools (1.1.0)
- DeconRNASeq (1.18.0)
- R, version 3.4 or higher s

EQUIPMENT:

- Scalpel
- Spatula
- Weighing boat
- Weighing scale
- Refrigerated bench-top centrifuge for 5mL tubes (Eppendorf, model 5430R)
- Refrigerated bench-top centrifuge for 15 mL tubes (Beckman Coulter, model Allegra X-14R)
- Countess® II FL Automated Cell Counter (Thermo Fisher Scientific, cat. no. AMAQAF1000)
- Countess® II FL Automated Cell Counter Chamber Slides (Thermo Fisher Scientific, cat. no. C10228)
- Flowmi™ Cell Strainer, 40 μm (Bel-Art, cat. no. H13680-0040)
- MACS® SmartStrainers, 30 μm (Miltenyi Biotec, cat. no. 130-098-458)
- 7 ml Tissue Grinder, Dounce (Wheaton, cat. no. 357542)
- 15 ml centrifuge tubes (Corning, cat.no. 430791)

- Centrifuge tube, 50 mL screw cap (Sarstedt, SAR62547205)
- DNA LoBind Microcentrifuge Tubes (Eppendorf™, cat. no. 022431021)
- 250 mL glass bottles
- DNA LoBind Microcentrifuge Tubes 5.0 ml (Eppendorf, cat. no. 30108310)
- INCYTO C-Chip Disposable Hemocytometers (SKC Films Inc., cat. no. DHCN012 or DHCN015)
- PCR Tubes 0.2 ml 8-tube strips (Eppendorf, cat. no. 0030124286)
- 10mL serological pipette
- Invitrogen EVOS FL Auto Cell Imaging System (Thermo Fisher Scientific)
- Chromium Controller (10X Genomics)
- Microseal 'B' Adhesive Seals (Bio-Rad, cat. no. MSB1001)
- twin.tec 96-Well PCR Plate Semi-skirted LoBind (25) (Eppendorf, cat. no. 30129504)
- Divided Polystyrene Reservoirs (25mL, 50) (VWR, cat. no. 41428-960)
- 200UL Filter Tips (Rainin, cat. no. 17007961)
- Pipet-Lite Multi Pipette L8-200XLS+ (Rainin, cat. no. 17013805)
- Flowmi 40µm Cell Strainers (Sigma-Aldrich, cat. no. BAH136800040-50EA)
- TapeStation 2200 (Agilent) or equivalent equipment

REAGENT SETUP

The following reagents should be prepared ahead of time:

- Prepare 1M NaCl, 100mM $MgCl_2$, 1M $MgCl_2$, 1M KOH solutions in ddH_2O in separate 50 mL centrifuge tubes and store at room temperature (21-22^oC for our laboratory).
- Prepare 250 mL of 10% (weight/volume) BSA solution in a glass bottle by dissolving crystalline BSA in ddH_2O at room temperature. Store at 4^oC for up to 1 month. Keep crystalline BSA at -20^oC.
- Prepare 250 mL of 1M Tris HCl buffer in a glass bottle by dissolving in ddH_2O. Adjust pH to 7.4 by adding 1M HCl dropwise. Store at room temperature. This is a time-consuming step.
- Prepare 250 mL of 0.5M Tricine KOH buffer in a glass bottle by dissolving Tricine in ddH_2O and adjusting pH to 7.8 by adding 1M KOH dropwise. Store at room temperature. This is a time-consuming step.

Buffer recipes provided are calculated assuming preparation of 8 samples for capture using a full Chromium chip.

Optiprep™ diluent (altered as per Kriaucionis et al., 2009)[32]: Combine the following in a 250 mL glass bottle. Store at room temperature.

Component	Volume (in µL)	Final concentration
1M KCl	15000	150 mM
1M $MgCl_2.6H_2O$	500	5 mM
0.5M Tricine-KOH (pH 7.8)	4000	20 mM
Deionized water	80500	–
Total volume	100000	–

Optiprep™ solutions: Using Optiprep™ diluent solution dilute the Optiprep™ reagent to make 50% volume/volume and 29% volume/volume solutions of Optiprep™ solution in separate 50 mL centrifuge tubes. Protect from light and store at room temperature.

Lysis buffer (LB): Combine the listed components in a 50 mL centrifuge tube. This buffer should be made fresh and kept at 4^{o}C or on ice.

Component	Volume (in µL)	Final concentration
1M Tris pH 7.4	200	10 mM
1M NaCl	200	10 mM
100 mM $MgCl_2$	600	3 mM
NP-40 (10%)	100	0.05% (v/v)
Deionized water	18900	–
Total volume	20000	–

Nuclei wash buffer (NWB): Combine the listed components in a 250 mL glass bottle. This buffer should be made fresh and kept at 4^{o}C or on ice.

Component	Volume (in µL)	Final concentration
10% BSA	100000	5% (w/v)
Glycerol	500	0.25% (v/v)
Protector RNAse inhibitor	200	40 units/mL
1X PBS	~100000	0.5X
Total volume	~200000	–

EQUIPMENT SETUP

- Precool both centrifuges to 4^{o}C.
- Set up EVOS FL Auto microscope with 10X magnification, bright field and DAPI channels.

Computational requirements

Linux OS with minimum requirements for running CellRanger as described on the 10X Genomics' webpage must be met.

Procedure

Tissue preparation (Timing 1-4 hours)

1. Cut using a scalpel and weigh 30-50mg of frozen tissue per sample. Keep tissue on dry ice while cutting to minimize degradation. Transfer to a 1.5mL microcentrifuge tube using a spatula and place back on dry ice. Clean scalpel and spatula with 70% ethanol between

samples. Use a fresh weigh boat for each sample. Alternatively, this step could be replaced by cryosectioning a fresh frozen histology grade dissection of tissue and collecting several sections such that the total weight is between 30-50 mg.
CAUTION: Post-mortem human tissue can contain pathogens. Take precautions including wearing PPE and seek medical attention if the scalpel breaks your skin.

Nuclei extraction (Timing 2-3 hours)

2. Transfer tissue using spatula to douncing tube on ice. Add 3mL of ice-cold lysis buffer and dounce with loose pestle 10 times and 5 more times with the tight pestle.
   CRITICAL: Use proper douncing technique to ensure proper mechanical breakdown of tissue. Proceed slowly and avoid bubbles. Grind tissue against the bottom of the tube using the douncer with each stroke.
3. Transfer homogenized tissue to a 15mL centrifuge tube by pouring and add 2 mL of chilled lysis buffer. Incubate on ice for 5 minutes gently swirling to mix 2 times during incubation.
4. Add 5 ml of chilled wash buffer to lysed tissue to quench lysis. Swirl to mix.
5. Place 30 μm MACs SmartStrainer on a 15mL centrifuge tube. Pipette lysed tissue suspension on top of filter to remove cell debris and large clumps. In case of blocked flow through the filter, tap filter gently to encourage the suspension to flow through.
6. In the precooled Allegra-14X centrifuge, spin down the lysed tissue suspensions at 500g for 5 minutes at 4°C.
7. Decant supernatant into a waste beaker without disrupting the nuclei pellet.
   CAUTION: The supernatant should be treated as biohazardous waste and treated with bleach before disposal.
   CRITICAL: Pour out supernatant in a single motion as repeated pouring motions can dislodge the pellet. If the pellet dislodges during decanting slowly remove the supernatant using a pipette.
8. Using a 10 mL serological pipette add 10mL of nuclei wash buffer to the pelleted nuclei and gently pipet 8-10 times to mix.
9. Repeat step 5-7 using the resuspended nuclei.
10. Using a 10 mL serological pipette add 5mL of nuclei wash buffer to the pelleted nuclei and gently pipet 8-10 times to mix.
11. Repeat steps 6-7 using the resuspended nuclei.
12. Using a 1000 μL pipette tip, add 1mL of nuclei wash buffer to pelleted nuclei and gently pipet 8-10 times to mix.
13. Add 1 mL of 50% working solution of Optiprep™ and mix well to make 2 mL of 25% iodixanol solution containing nuclei.
14. Prepare Optiprep™ cushion of 2ml of 29% Optiprep™ solution in a 5ml Eppendorf centrifuge tube.
15. Gently add the 2ml nuclei suspension on top of the Optiprep™ cushion by pipetting slowly against the wall of the tube to avoid mixing.
16. In the precooled Eppendorf centrifuge, spin the tubes containing nuclei layered over Optiprep™ cushion at 10,000g for 30 minutes at 4°C.
17. Carefully pour out the supernatant leaving the least amount of volume in the tube without disrupting the pellet. **?TROUBLESHOOTING**

18. Using a 1000 µL pipette tip, resuspend the nuclei pellets in 500 µL or less of nuclei wash buffer. Gently pipette 8-10 times or until nuclei are resuspended.
19. For a quick estimate of nuclei concentration, mix 10 µL of the nuclei suspension with 10 µL of Trypan blue in a separate tube. Load 10 µL of the mixture onto a Countess hemocytometer slide. Count nuclei on the Countess hemocytometer and measure range of sizes. For human neurons the average diameter of the nuclei should be around 10 µm. Trypan blue is a live dead stain and properly isolated nuclei should have be marked as dead cells.
20. Using the estimated count from the Countess dilute cells to around 500,000 cells/ mL or 500 cells/ µL by adding an appropriate volume of nuclei wash buffer. It may be possible to increase these concentrations for capturing more nuclei.
    CRITICAL: If the concentration of nuclei is too high it can result in aggregation which will prevent efficient capture of single nuclei in subsequent steps.
21. Add Hoechst stain to the resuspended nuclei at a 1:2000 dilution.

Nuclei concentration assessment and capture (Timing 45 mins)

Immediately prior to loading the Hoechst stained nuclei on the Chromium system, check to make sure the nuclei are well segregated (nuclei may clump) and recheck sample concentration using a fluorescent microscope like the Evos FL Auto (Thermofisher)

22. OPTIONAL: Use a 1mL pipette to take a minimum of 200 µL of the sample and filter it through a 40µm Flowmi pipette tip filter before counting and loading. This will get rid of the clumped nuclei and large debris that can clog the microfluidics of Chromium chips.
23. Load 10µL of sample onto a disposable hemocytometer slide and into the Evos.

    CRITICAL: Make sure to pipette the full volume of the nuclei suspension several times before loading for accurate counts.

24. Set the Evos to 10X magnification and image the entire hemocytometer grid field of view in DAPI and bright field. This image makes counting nuclei easier and serves as a record as well (**Fig. 3g**).
25. In parallel to steps 22-24, make the RT Master Mix and aliquot appropriate volumes of master mix and water into PCR tubes according to the number of nuclei to be targeted for capture, referring to the Chromium protocol CG00052 Rev. D or later. **?TROUBLESHOOTING**

    CRITICAL: The Chromium capture rate for nuclei from archived tissue is lower (~20% less) than the capture rate for cells. To account for this, it is necessary to adjust the count used to determine loading volume. We empirically determined that choosing the loading volume by using a count that is 30% less than observed count worked best for our samples. This adjustment may vary from tissue type to tissue type. For example, if the sample has a concentration of 500 nuclei/µL, the sample volume should be loaded as if it has 350 cells/µL (70% of 500 cells) in order to get the targeted nuclei to recover.

    CRITICAL: Resuspend nuclei by pipette mixing the full volume several times immediately before loading to prevent aggregation of nuclei.

26. Load the Chromium Chip and harvest the nuclei captured in droplet (i.e. GEMs) according to the Chromium protocol CG00052 Rev. D

Library preparation and sequencing (Timing)

27. Reverse transcription, cDNA amplification, and library preparation are performed according to the Chromium protocol CG00052 Rev. D. Libraries can be sequenced on an Illumina sequencer. Sequencing two samples per lane of a HiSeq 4000 machine can yield ~100,000 reads per nucleus based on default CellRanger parameters and provide sufficient information for cell-type identification and differential expression analysis. Downstream analysis options for the sequencing results are described in Box 1.

Box 1: Downstream data analysis

Alignment, Demultiplexing, and Generation of Counts Matrix (Time: variable)
Since our experiments utilized human nuclei, we built a pre-mRNA reference using the cellranger mkref (Cellranger version 2.0.1) command. Default parameters were used, starting with the refdata-cellranger-GRCh38-1.2.0 transcriptome and as per the instructions provided on the 10X Genomics website. For mouse tissue the corresponding pre-mRNA reference would need to be created for the mouse genome. We demultiplexed reads by sample index using the cellranger mkfastq command (Cellranger v2.1.0), aligned FASTQ files to the custom transcriptome, demultiplexed cell barcodes, counted the UMIs corresponding to genes using the cellranger count command and default parameters. These steps may be performed with custom code if desired.

Custom Filtering to Recover Low Transcript Number Cell Types (Time: variable)
While there are many options for software to be used for downstream analysis of snRNA-seq data such as scater[33], SC3[34], Monocle3[35], etc., we used the Seurat R package (version 2.2.0, 2.3.0)[36]. Unfiltered gene barcode matrices for each sample were loaded into R using the Read10X function. At this step, cell names can be modified such that the subject name, batch, and biological condition are appended to them. Seurat objects were created corresponding to each sample using the CreateSeuratObject function with the imported unfiltered gene-barcode matrices provided as the raw data. Individual Seurat objects for each sample were combined sequentially using the MergeSeurat function. No filtering or normalization was performed up to this step. Since we were working with a single nucleus dataset, all mitochondrial genes that are transcribed from the mitochondrial genome were removed, along with genes not detected in any cell. More recently several methods have been developed to align multiple datasets of snRNA-seq and other single-cell level data[28,37] which can be used for combining the data from individual subjects if inter-individual variability or batch effects are deemed to have a large influence on the results.

For preliminary filtering, some nuclei with very low number of genes detected (>110) and nuclei with very high numbers of UMIs detected (in the top 0.5%) were removed as low-quality nuclei and potential multiplets respectively. These cut-offs are arbitrary but can be based upon the distribution of the data. For example, in our dataset there was a sharp increase in the number of UMIs from 16,393 at the 99.5th percentile to 102,583 at the maximum which probably represents the multiplets in the dataset.

If the dataset contains multiple cell-types which are expected to be heterogenous in terms of the

number of molecules of RNA present per nucleus (such as when the nuclei of different cell-types are known to be of very different size), the following approach can be used for removing low quality cells without unduly biasing the filtering against nuclei which biologically contain fewer molecules . For our dataset, given the known trend for higher number of RNA molecules in neuronal nuclei compared to glial nulcei[10,11,38], the distribution of number of UMIs was fit with three normal distributions using the normalmixEM function from the mixtools[39] package. The rationale is that the filtered barcodes contain a population of low quality "noise" barcodes that have a very low numbers of UMIs on average, a population of non-neuronal cells that have an intermediate numbers of UMIs and a population of neuronal cells that have a high numbers of UMIs. After fitting the normal distributions, only the barcodes with a high probability (> 0.95) of belonging to either the putative "non-neuronal" or putative "neuronal" distributions, and a low probability (<0.05) of belonging to the "noise" distribution were retained for further analysis. As an example, in our publication, for a subset of 20 subjects, applying our custom filtering approximately doubled the total number of cells but increased the number of non-neuronal cells by almost 6 times[20].

?TROUBLESHOOTING

Troubleshooting advice can be found in Table 1.

Table 1: Troubleshooting table.

Step	Problem	Possible reason	Solution
17	Absence of visible pellet after Optiprep™ cushion centrifugation	A large pellet may indicate presence of excessive debris rather than high nuclei yield and absence of a pellet is not necessarily cause for concern. However, in certain cases it may indicate very low yield of nuclei.	Continue with downstream steps assuming the location of the nuclei based on the direction in which the tube is placed within the centrifuge and assess nuclei yield under the microscope. If very low nuclei yield is observed, consider increasing the amount of input material.
25	Number of nuclei captured does not meet the expected number based on the table provided by the 10x loading guidelines	The capture rate for nuclei may not be the same as that for cells.	Empirically determine the difference between the capture rate expected and observed and adjust loading volume accordingly.

TIMING

Step 1, 1 hour for 8 samples if cutting pieces using a scalpel, up to 4 hours if collecting cryosections

Step 2-21, 2-3 hours

Step 22-26, 45 minutes

Step 27, can be split into two 4 hours blocks on 2 days

Anticipated Results

We expect our single-nuclei extraction protocol to produce high quality single-nucleus suspensions **(Fig. 3)** from frozen archived post-mortem tissues. The nuclear suspensions are relatively free from debris and do not show substantial aggregation of nuclei even after 16 hours of refrigeration, upon visual inspection **(Fig. 3-f)**. Capture of single-nuclei using these nuclear suspensions on a microfluidic device for droplet-based snRNA-seq or snATAC-seq reproducibly produces high-quality libraries with sufficient cDNA yield for sequencing **(Fig. 4)**. Moreover, the variability in sample parameters such as PMI, archival time, pH and RIN did not affect most of the quality metrics of snRNA-seq results with this nuclei extraction protocol (**Fig. 2**).

Acknowledgements

GT holds a Canada Research Chair (Tier 1) and a NARSAD Distinguished Investigator Award. He is supported by grants from the Canadian Institute of Health Research (CIHR) (FDN148374 and EGM141899),

We acknowledge the expert help of the Douglas-Bell Canada Brain Bank staff (Josée Prud'homme, Maâmar Bouchouka and Annie Baccichet), and the technology development team at the McGill University and Genome Quebec Innovation Centre. The Douglas-Bell Canada Brain Bank is supported in part by funding from the Canada First Research Excellence Fund, awarded to McGill University for the Healthy Brains for Healthy Lives project, and from the

Fonds de recherche du Québec - Santé (FRQS) through the Quebec Network on Suicide, Mood Disorders and Related Disorders.

Author Contributions

MM and C. Nagy developed nuclei extraction protocol, prepared nuclei, and wrote the manuscript. YCW preformed 10X snRNA-seq protocol. C. Nascimento performed 10X snATAC-seq protocol. MS and JFT guided bioinformatic analysis. NM contributed to tissue processing and data interpretation. JR provided technical single-cell expertise and experimental support. GT provided general oversight, including in experimental design. All authors contributed to manuscript preparation.

Figures

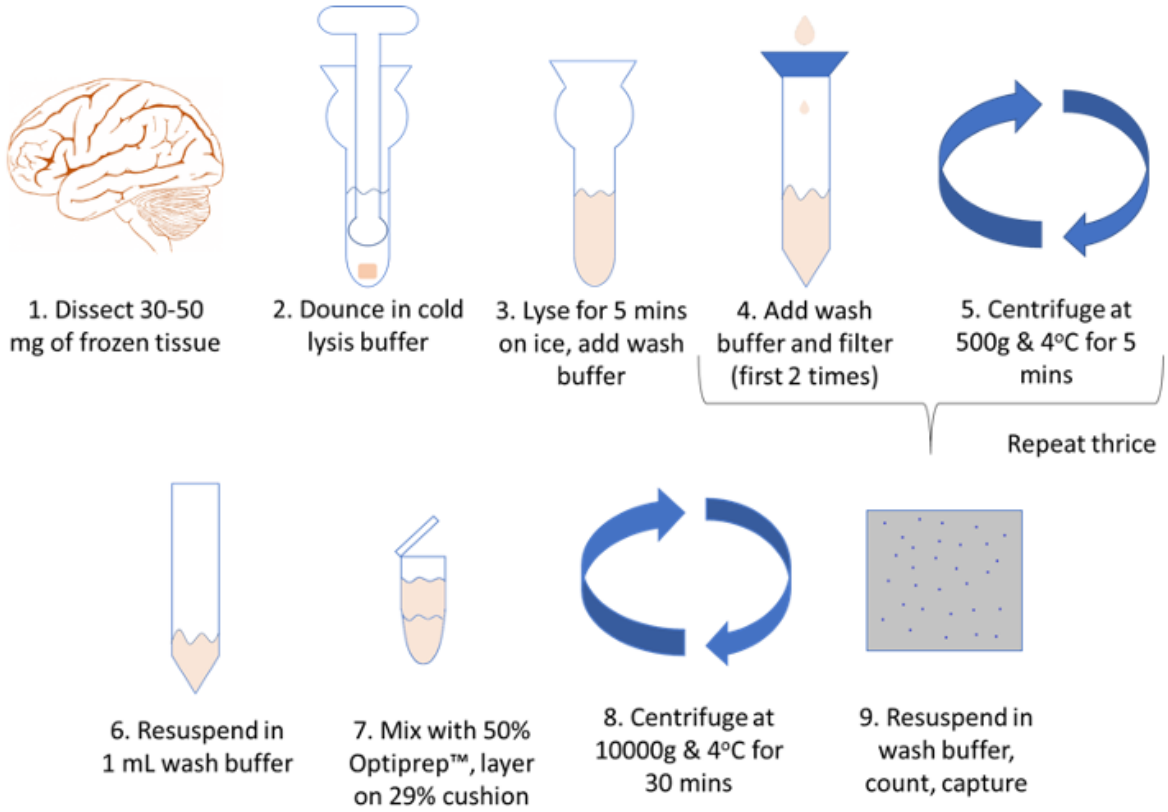


Figure 1: Schematic representation of the steps of the protocol. Frozen tissue is dissected, homogenized by douncing, lysed, and then washed, filtered, and centrifuged several times until a single-nucleus suspension is obtained.

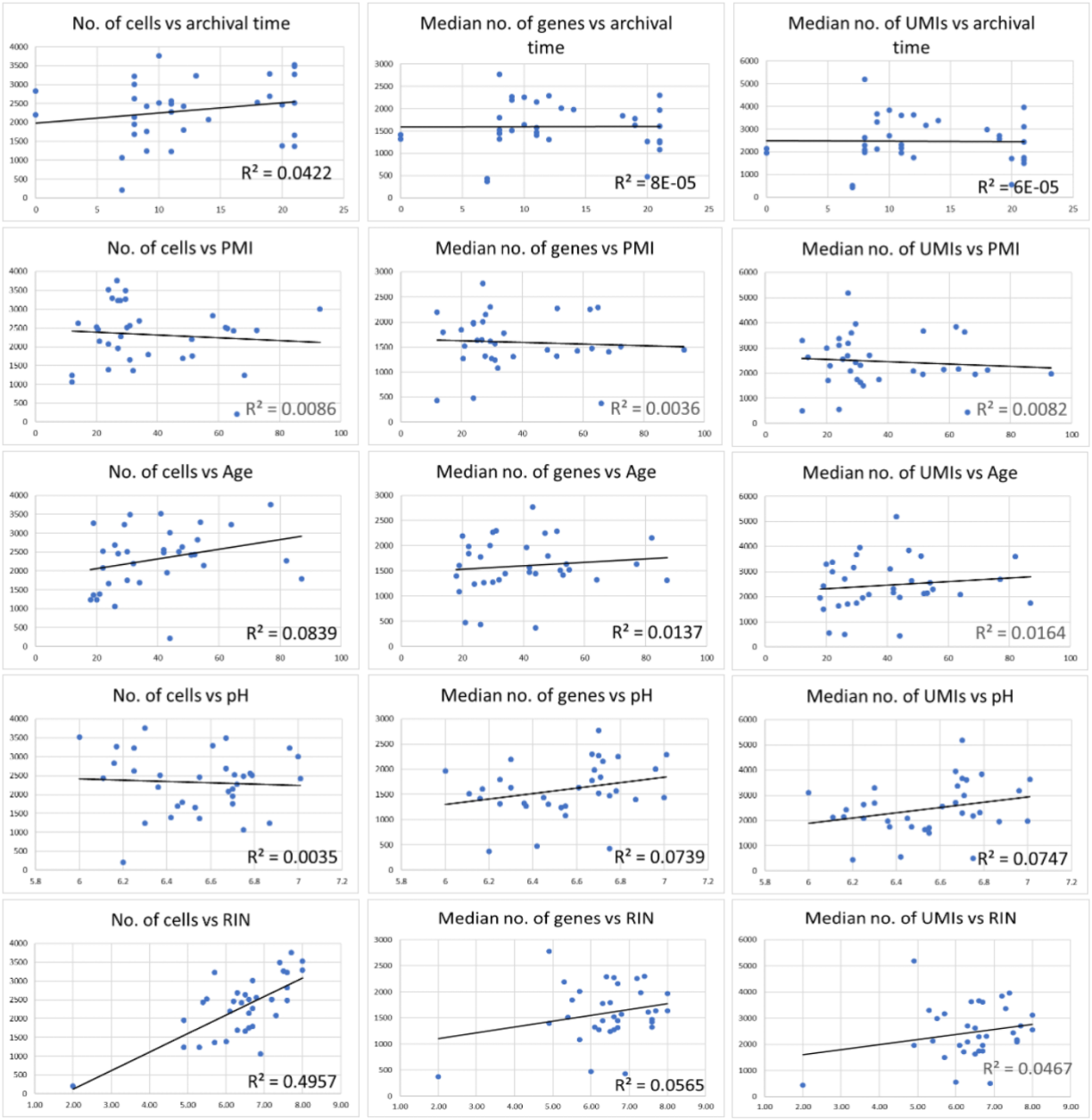


Figure 2: Effect of sample quality parameters on single-nucleus capture and sequencing metrics. The archival time, post mortem interval (PMI), age, and pH of the brains accounted for less than 10% of the variation in number of cells (nuclei) retained after filtering, median number of genes per nucleus, and median no of UMIs (unique molecular identifiers) per nucleus. The RIN of the samples had had a significant negative effect on the number of nuclei captured, especially for RIN < 4, but did not have a large effect on the median numbers of genes or UMIs.

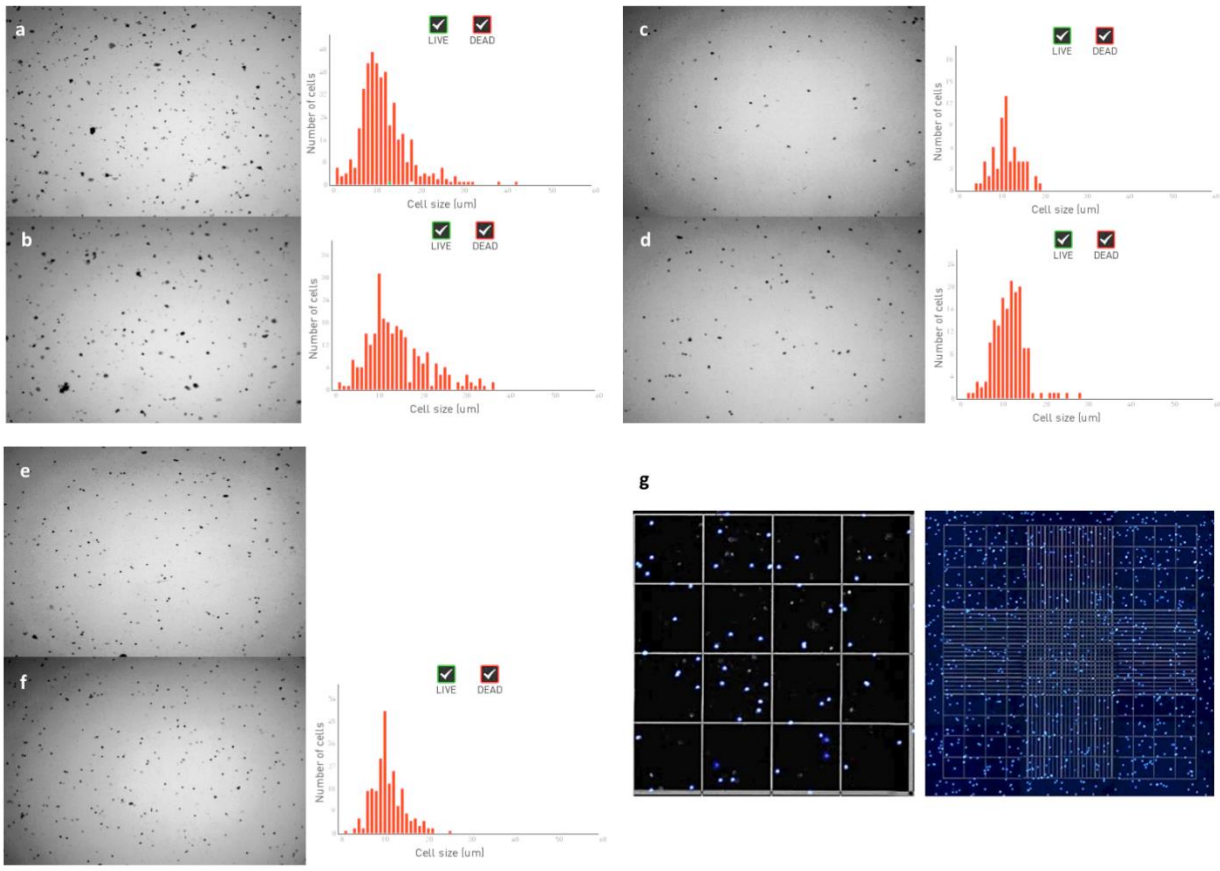
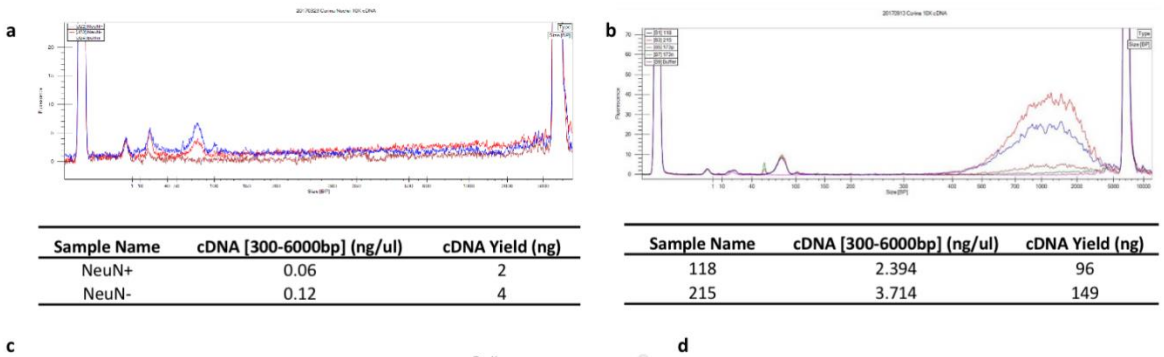


Figure 3: Images of extracted nuclei. (a-d) Before optimization, using the 10X Genomics demonstrated protocol, the extracted nuclei from two different samples (a and b) show large amounts of debris and the size distribution is skewed towards larger sizes (> 10 µm). After optimization, representative images of nuclei extracted from two samples (c and d) show much

less debris and a size-distributions are centered around 10 µm, as expected for human brain nuclei. Images were acquired with the Countess Cell Counter using Trypan blue for staining. Note that extracted nuclei should be marked as dead cells, as seen. (e-f) Extracted nuclei do not tend to aggregate even after (e) 2.5 hours or (f) 16 hours of storage at 4^{o}C. Note that the size distribution after 16 hours is still centered around 10 µm, indicate an absence of aggregation. (g) Representative images of extracted nuclei stained with Hoechst (1:2000) acquired at 10X magnification on the Evos microscope.



Sample Name	cDNA [300-6000bp] (ng/ul)	cDNA Yield (ng)
NeuN+	0.06	2
NeuN-	0.12	4

Sample Name	cDNA [300-6000bp] (ng/ul)	cDNA Yield (ng)
118	2.394	96
215	3.714	149

Sample Name	Number of reads	Estimated no. of cells	Mean reads per cell	Median genes per cell	Median UMI count per cell
NeuN+	165,749,748	1,507	109,986	80	124
215	144,125,115	1,486	96,988	3,501	7,287

Sample Name	cDNA [163-700bp] (pg/ul)	cDNA Yield (ng)
84	6140	

Figure 4: cDNA traces and quality metrics for snRNA-seq libraries before and after optimization of nuclei extraction. (a) A FANS based nuclei isolation of single-nucleus suspensions resulted in very low yield cDNA libraries whereas (b) the optimized nuclei extraction protocol resulted in good yield of cDNA in the expected size range. Perkin Elmer

Caliper traces are shown for snRNA-seq cDNA libraries. (c) With similar numbers of sequencing reads and median reads per cell, the libraries produced using the optimized nuclei extraction protocol has much higher median numbers of genes and UMIs per cell as can be seen from the elbow blot produced by Cellranger as well as the tabulated summary metrics. (d) Our nuclei extraction protocol is compatible with 10X Genomics' commercial snATAC-seq protocol, as can be seen from the Tapestation trace of a successfully constructed snRNA-seq library from archived post-mortem human brain tissue.

References:

1. Macosko, Evan Z*., et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202-1214 (2015).
2. Zilionis, R*., et al.* Single-cell barcoding and sequencing using droplet microfluidics. *Nat Protoc* **12**, 44-73 (2017).
3. Zheng, G.X.Y*., et al.* Massively parallel digital transcriptional profiling of single cells. **8**, 14049 (2017).
4. Haber, A.L*., et al.* A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333 (2017).
5. Wang, Y*., et al.* Pulmonary alveolar type I cell population consists of two distinct subtypes that differ in cell fate. *Proceedings of the National Academy of Sciences* **115**, 2407-2412 (2018).
6. Sathyamurthy, A*., et al.* Massively Parallel Single Nucleus Transcriptional Profiling Defines Spinal Cord Neurons and Their Activity during Behavior. *Cell reports* **22**, 2216-2225 (2018).
7. Hrvatin, S*., et al.* Single-cell analysis of experience-dependent transcriptomic states in the mouse visual cortex. *Nature neuroscience* **21**, 120-129 (2018).
8. Hochgerner, H., Zeisel, A., Lönnerberg, P. & Linnarsson, S. Conserved properties of dentate gyrus neurogenesis across postnatal development revealed by single-cell RNA sequencing. *Nature neuroscience* **21**, 290-299 (2018).
9. Lake, B.B*., et al.* Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586-1590 (2016).
10. Habib, N*., et al.* Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nature Methods* **14**, 955 (2017).
11. Lake, B.B*., et al.* Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat Biotechnol* **36**, 70-80 (2018).
12. Sullivan, P.F., Neale, M.C. & Kendler, K.S. Genetic epidemiology of major depression: review and meta-analysis. *The American journal of psychiatry* **157**, 1552-1562 (2000).
13. Lacar, B*., et al.* Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nature Communications* **7**, 11022 (2016).
14. van den Brink, S.C*., et al.* Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nature Methods* **14**, 935 (2017).
15. Grindberg, R.V*., et al.* RNA-sequencing from single nuclei. *Proceedings of the National Academy of Sciences* **110**, 19802-19807 (2013).
16. Lake, B.B*., et al.* A comparative strategy for single-nucleus and single-cell transcriptomes confirms accuracy in predicted cell-type expression from nuclear RNA. *Scientific reports* **7**, 6031 (2017).
17. Cutler, A.A., Jackson, J.B., Corbett, A.H. & Pavlath, G.K. Non-equivalence of nuclear import among nuclei in multinucleated skeletal muscle cells. *Journal of Cell Science* **131**(2018).
18. Berridge, B.R., Bolon, B. & Herman, E. Chapter 10 - Skeletal Muscle System. in *Fundamentals of Toxicologic Pathology (Third Edition)* (eds. Wallig, M.A., Haschek, W.M., Rousseaux, C.G. & Bolon, B.) 195-212 (Academic Press, 2018).
19. Krishnaswami, S.R*., et al.* Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat Protoc* **11**, 499-524 (2016).
20. Nagy, C*., et al.* Single-nucleus RNA sequencing shows convergent evidence from different cell types for altered synaptic plasticity in major depressive disorder. *bioRxiv* (2018).
21. Jessa, S*., et al.* Stalled developmental programs at the root of pediatric brain tumors. *Nature genetics* **51**, 1702-1713 (2019).

22. Mathys, H*., et al.* Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332-337 (2019).
23. Velmeshev, D*., et al.* Single-cell genomics identifies cell type–specific molecular changes in autism. *Science* **364**, 685 (2019).
24. Jäkel, S*., et al.* Altered human oligodendrocyte heterogeneity in multiple sclerosis. *Nature* **566**, 543-547 (2019).
25. Sorrells, S.F*., et al.* Immature excitatory neurons develop during adolescence in the human amygdala. *Nature Communications* **10**, 2748 (2019).
26. Proudfoot, N.J., Furger, A. & Dye, M.J. Integrating mRNA Processing with Transcription. *Cell* **108**, 501-512 (2002).
27. Amamoto, R*., et al.* Probe-Seq enables transcriptional profiling of specific cell types from heterogeneous tissue by RNA-based isolation. *eLife* **8**, e51452 (2019).
28. Welch, J.D*., et al.* Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity. *Cell* **177**, 1873-1887.e1817 (2019).
29. Skelly, D.A*., et al.* Single-Cell Transcriptional Profiling Reveals Cellular Diversity and Intercommunication in the Mouse Heart. *Cell reports* **22**, 600-610 (2018).
30. Qu, K*., et al.* Individuality and variation of personal regulomes in primary human T cells. *Cell systems* **1**, 51-61 (2015).
31. Cusanovich, D.A*., et al.* The cis-regulatory dynamics of embryonic development at single-cell resolution. *Nature* **555**, 538-542 (2018).
32. Kriaucionis, S. & Heintz, N. The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science* **324**, 929-930 (2009).
33. McCarthy, D.J., Campbell, K.R., Lun, A.T.L. & Wills, Q.F. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics (Oxford, England)* **33**, 1179-1186 (2017).
34. Kiselev, V.Y*., et al.* SC3: consensus clustering of single-cell RNA-seq data. *Nature Methods* **14**, 483-486 (2017).
35. Cao, J*., et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496-502 (2019).
36. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology* **36**, 411 (2018).
37. Stuart, T*., et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e1821 (2019).
38. Darmanis, S*., et al.* A survey of human brain transcriptome diversity at the single cell level. *Proceedings of the National Academy of Sciences* **112**, 7285-7290 (2015).
39. Benaglia, T., Chauveau, D., Hunter, D.R. & Young, D.S. mixtools: An R Package for Analyzing Mixture Models. *2009* **32**, 29 (2009).