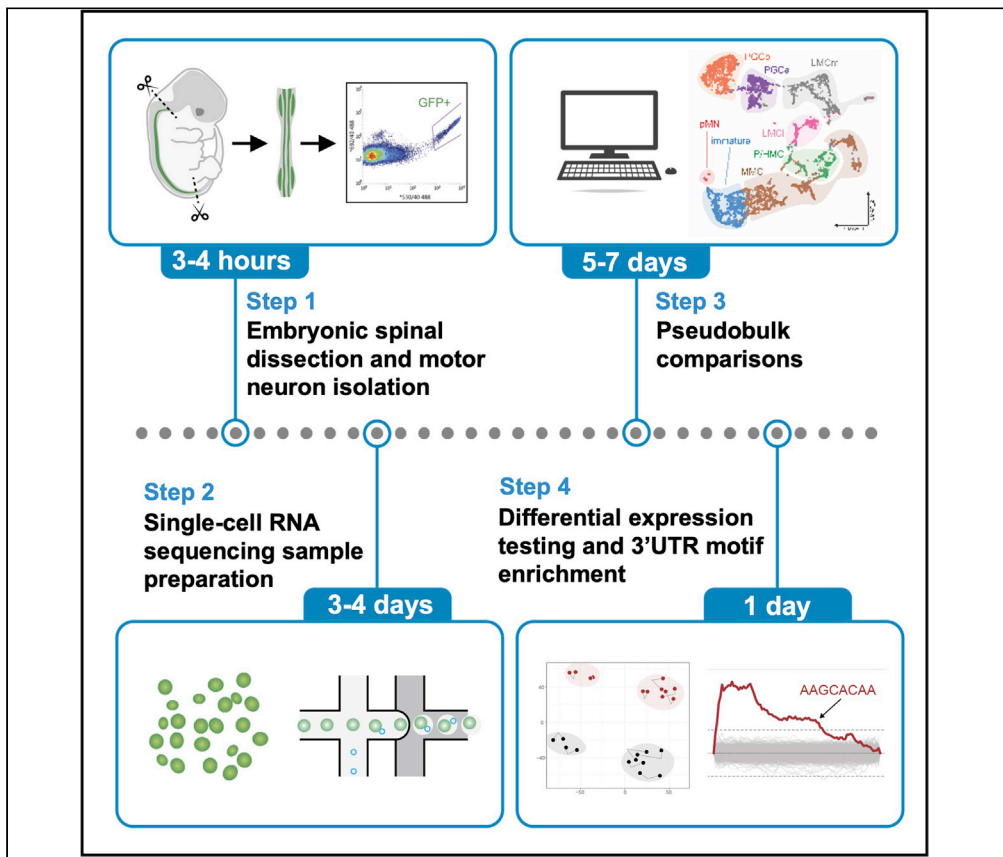


Protocol

Detecting microRNA-mediated gene regulatory effects in murine neuronal subpopulations



microRNAs (miRNAs) have unique gene regulatory effects in different neuronal subpopulations. Here, we describe a protocol to identify neuronal subtype-specific effects of a miRNA in murine motor neuron subpopulations. We detail the preparation of primary mouse spinal tissue for single cell RNA sequencing and bioinformatics analyses of pseudobulk expression data. This protocol applies differential gene expression testing approaches to identify miRNA target networks in heterogeneous neuronal subpopulations that cannot otherwise be captured by bulk RNA sequencing approaches.

Neal D. Amin,
Gokhan Senturk,
Marito Hayashi,
Shawn P. Driscoll,
Samuel L. Pfaff

nealamin@stanford.edu (N.D.A.)
pfaff@salk.edu (S.L.P.)

Highlights

Careful removal of the meninges during embryonic mouse spinal cord dissection

Fluorescence-activated cell sorting of motor neurons

Perform differential expression testing between cell populations identified *in silico*

Hypergeometric enrichment analysis detects dysregulated miRNA targets

Amin et al., STAR Protocols 3, 101130
March 18, 2022 © 2022 The Authors.
<https://doi.org/10.1016/j.xpro.2022.101130>



Protocol

Detecting microRNA-mediated gene regulatory effects in murine neuronal subpopulations

Neal D. Amin,^{1,2,4,*} Gokhan Senturk,¹ Marito Hayashi,^{1,3} Shawn P. Driscoll,¹ and Samuel L. Pfaff^{1,5,*}¹Gene Expression Laboratory, Salk Institute for Biological Studies, 10010 North Torrey Pines, La Jolla, CA 92037, USA²Present address: Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA 94305, USA³Present address: Howard Hughes Medical Institute, Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA⁴Technical contact⁵Lead contact*Correspondence: nealamin@stanford.edu (N.D.A.), pfaff@salk.edu (S.L.P.)
<https://doi.org/10.1016/j.xpro.2022.101130>

SUMMARY

microRNAs (miRNAs) have unique gene regulatory effects in different neuronal subpopulations. Here, we describe a protocol to identify neuronal subtype-specific effects of a miRNA in murine motor neuron subpopulations. We detail the preparation of primary mouse spinal tissue for single cell RNA sequencing and bioinformatics analyses of pseudobulk expression data. This protocol applies differential gene expression testing approaches to identify miRNA target networks in heterogeneous neuronal subpopulations that cannot otherwise be captured by bulk RNA sequencing approaches.

For complete details on the use and execution of this protocol, please refer to Amin et al. (2021).

BEFORE YOU BEGIN

microRNAs (miRNAs) are post-transcriptional repressors of gene expression and have important roles in regulating nervous system development, mature neuronal function, and disease states. miRNAs target the 3' untranslated regions (3'UTRs) of mRNA to reduce mRNA levels and prevent protein translation (Gebert and MacRae, 2019). Bioinformatics models have been successful at predicting mRNA targets based upon a miRNA's unique seed sequence and reverse-complementary binding sites on mRNA's 3'UTRs. When predictive models are developed with experimental bulk RNA sequencing data, a miRNA's target genes can be identified with high confidence (Agarwal et al., 2015). However, miRNAs can have different effects in neuronal subtypes due to intrinsic differences in expressed mRNAs, levels of miRNA expression, and context-specific effects (Carroll et al., 2013; Amin et al., 2021). Neuronal subtype-specific miRNA effects are lost by bulk RNA sequencing methods in which heterogeneous cell populations are combined in a single sample. The extent and character of neuronal subtype-specific miRNA gene regulatory effects are not well understood. A deeper understanding of miRNA effects on a single cell and subpopulation level may reveal novel paradigms in miRNA targeting with widespread implications for neurodevelopment and disease states under miRNA control.

Motor neuron cell bodies are located in the spinal cord and brainstem and they project axons into the periphery to control muscle contractions. In developmental and degenerative conditions affecting motor neurons such as spinal muscular atrophy and amyotrophic lateral sclerosis, motor neurons are lost resulting in paralysis. Motor neurons are known to selectively and abundantly express microRNA-218 (miR-218) which regulates neuromuscular synaptogenesis and motor neuron survival via the repression of hundreds of identified target genes (Amin et al., 2015). Furthermore,



mice lacking miR-218 die at birth due to paralysis (Amin et al., 2015). Nevertheless, motor neurons are highly heterogeneous with regards to the expression of transcription factors, neurotransmitter profiles, electrophysiological characteristics, and axonal projection patterns (Stifani, 2014).

Here, we describe a protocol to use single cell RNA sequencing data to characterize miR-218's motor neuron subtype-specific effects. We provide technical details on the preparation of embryonic spinal tissues for FACS-isolation of motor neurons for single cell RNA sequencing. We then describe the identification of gene expression changes resulting from the loss of miR-218 using complementary bioinformatics methodologies and provide example results. This protocol can be applied to other microRNAs expressed in other cell populations relevant to a wide range of biological questions.

All mouse work was conducted in accordance with IACUC guidelines of the Salk Institute for Biological Studies.

Prepare *Hb9::gfp* mice

⌚ Timing: 2–3 months

1. Set up crosses of mice carrying desired genetic mutations (e.g., miR-218 knockout alleles) with a motor neuron fluorescent reporter (e.g., *Hb9::gfp*) to generate breeding pairs for timed mating.

Note: It is advisable to plan the breeding strategy in advance to optimize the chance of getting multiple desired alleles in the same animal. For animals in which multiple alleles are combined, such as the five alleles of *Hb9::gfp*; miR-218-1^{-/-}; miR-218-2^{-/-} (double knockout, DKO) embryos, develop a breeding strategy to increase the yield of the desired genotype. For example, using breeding pairs of *Hb9::gfp*; miR-218-1^{-/-}; miR-218-2^{+/-} and miR-218-1^{-/-}; miR-218-2^{+/-} mice will be expected to yield embryos in which one eighth are *Hb9::gfp*; miR-218-1^{-/-}; miR-218-2^{-/-} (DKO mice with a *Hb9::gfp* motor neuron reporter). For these complex breeding strategies, it is important to plan many months in advance.

⚠ **CRITICAL:** For fluorescence activated cell sorting (FACS) experiments, it is important to consider the signal differences between embryos carrying one versus two copies of a genomically encoded fluorescent reporter. Embryos with two fluorescent *Hb9::gfp* alleles exhibit brighter fluorescence signals, thereby introducing variance in captured cells from the gating set on the fluorescence activated cell sorter (FACS) machine. We advise that only one animal from each breeding pair carries the *Hb9::gfp* fluorescent reporter. Considering the effect of fluorescence upon FACS gating and capture of cells for bulk or single cell RNA sequencing experiments is critical to avoid introducing irrevocable biases that will affect later bioinformatics analyses (see step 17).

2. Genotype mice with highly reliable methods, such as qPCR-based strategies employed by Transnetyx, Inc.
3. Set up timed mouse mating 12 days prior to the experiment.

Note: It is advisable to set up multiple breeding pairs on the same day to increase the chance that at least one female becomes pregnant.

Buffer preparation and reagent collection

⌚ Timing: 1 h

4. Prepare tools and tubes.

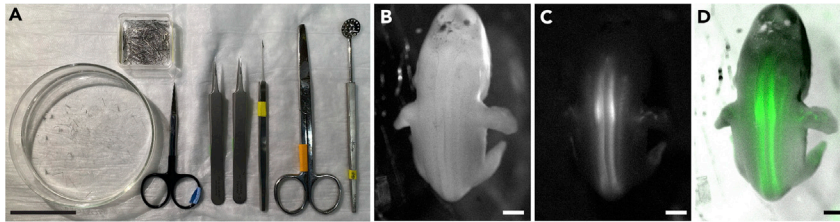


Figure 1. Preparation for spinal cord dissection

(A) Display of dissection tools.
 (B) Brightfield dorsal view of an E12 mouse embryo.
 (C) GFP signal expressed from the *Hb9::gfp* allele identifies the motor neuron cell bodies in spinal motor columns and axons projecting into peripheral tissues.
 (D) Inverted brightfield and GFP merged view. Scale bar (A) 5 mm; (B-D) 1 mm.

- a. Add PBS to three 10 cm plastic dishes.

Note: Keeping PBS chilled at 4°C may preserve cellular gene expression better during the time between mouse euthanasia and removal of spinal cord tissues.

- b. Label 1.5 mL tubes and accompanying PCR strip tubes for each potential embryonic spinal cord and tail snip, respectively.
- c. Collect scissors, dissection knife, spatula, Sylgard dissection dish with dissection pins, microdissection scissors, and forceps and spray with 70% ethanol and wipe down (Figure 1A).

5. Reconstitute papain solutions according to the manufacturer's instructions and equilibrate in 5% CO₂ in an incubator at 37°C.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
PBS, pH 7.4	Thermo Fisher Scientific	Cat#10010023
Neurobasal without phenol red	Thermo Fisher Scientific	Cat#12348017
DMEM/F12 without phenol red	Thermo Fisher Scientific	Cat#21041025
Heat inactivated horse serum	Thermo Fisher Scientific	Cat#26050088
Trypan Blue	Bio-Rad	Cat#1450021
Critical commercial assays		
QuickExtract DNA Extraction Solution	Lucigen	Cat #QE09050
Papain Dissociation Kit	Worthington Biochemical	Cat# LK003153
DNase (from papain dissociation kit) 2000 U/mL	Worthington Biochemical	Cat# LK003153
Gotaq 2x master mix	Promega	Cat#M7122
Chromium Single Cell 3' Library and Gel Beads Kit v2	10x Genomics	Cat#PN-120267
Deposited data		
Bulk RNA seq - E12 Hb9::gfp	Amin et al. (2021)	E12_WT (A-C)
Bulk RNA seq - E12 Hb9::gfp;miR218-1-/-	Amin et al. (2021)	E12_1KO (A-C)
Bulk RNA seq - E12 Hb9::gfp;miR218-2-/-	Amin et al. (2021)	E12_2KO (A-C)
Bulk RNA seq - E12 Hb9::gfp;miR218-1-/-;miR218-2-ΔILO	Amin et al. (2021)	E12_1KO;2-ΔILO (A-C)
Bulk RNA seq - E12 Hb9::gfp;miR218-1-/-;miR218-2-/-	Amin et al. (2021)	E12_DKO (A-C)
scRNAseq - E12 Hb9::gfp	Amin et al. (2021)	MNscWT(1,2)
scRNAseq - E12 Hb9::gfp;miR218-1-/-;miR218-2-/-	Amin et al. (2021)	MNscDKO(1,2)
The bulk and single cell RNA sequencing data sets	Amin et al. (2021)	ArrayExpress accession: E-MTAB-10571.

(Continued on next page)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oligonucleotides		
Genotyping forward primer for miR-218-1	Integrated DNA Technologies	accctcttttcaatccctga
Genotyping reverse primer for miR-218-1	Integrated DNA Technologies	ctttctgctgatccgattctg
Genotyping forward primer for miR-218-2	Integrated DNA Technologies	ggcagtcctgtgtcacaag
Genotyping reverse primer for miR-218-2	Integrated DNA Technologies	tggatcatcggggttcatcgg
Experimental models: Organisms/strains		
Mus musculus Hb9::gfp allele	Lee et al. (2004)	N/A
Mus musculus miR-218-1 KO (1KO) allele	Amin et al. (2015)	N/A
Mus musculus miR-218-2 KO (2KO) allele	Amin et al. (2015)	N/A
Mus musculus miR-218-2-ΔILO KO allele	Amin et al. (2021)	N/A
Software and algorithms		
RStudio	www.rstudio.com	Version 1.2.1335
Galaxy	Afgan et al. (2018)	Version 0.43.1.3
Sylamer	van Dongen et al. (2008)	N/A
bcl2fastq	Illumina	v2.18.0.12
Cellranger	10X Genomics	V3
DESeq	Bioconductor	1.26.0
ClustVis	Metsalu and Vilo (2015)	N/A
Other		
HiSeq 4000	Illumina	SY-401-4001
Minutien Pins	Fine Science Tools	Cat# FST 26002-20
Fine Scissor	Fine Science Tools	Cat# FST 14058-09
Dumont #5 Forceps	Fine Science Tools	Cat# FST 11295-00
Dissection Knife	Fine Science Tools	Cat# FST 10055-12
Surgical Scissor	Fine Science Tools	Cat# FST 14001-12
Moria Perforated Spoon	Fine Science Tools	Cat# FST 10370-18
Sylgard Dissection Dish	Living Systems Instrumentation	Cat# DD-90-S
Falcon Cell Strainer	Fisher Scientific	Cat# 08-771-1
Stereomicroscope	Zeiss	Model: Lumar V12
Dissection stereomicroscope	Zeiss	Model: Stemi SV6
FACSVantage SE DiVa using Coherent Sapphire 488 nm	Becton Dickinson	Model: SE DiVa
10× Chromium Controller	10× genomic	Cat#: 120263
Petri dishes, polystyrene (100 mm)	Sigma-Aldrich	Cat#: P5606-400EA

STEP-BY-STEP METHOD DETAILS

Spinal cord dissection

⌚ Timing: 2–3 h for a litter of embryos

During these steps, spinal cords are microdissected from mouse embryos collected from pregnant female mice. Specific care is taken to remove the meninges from spinal tissues. A tail snip from each embryo is also taken for genotyping.

1. Euthanize the pregnant dam on embryonic day (E) 12 following your institutionally approved method.
2. Remove embryos.
 - a. Spray the mouse abdomen with 70% ethanol until hair is matted down.
 - b. Using scissors, make an incision along the abdominal midline, and remove the uterine horns.
 - c. Transfer uterine horns to a 10 cm dish and rinse twice with ice cold PBS. Then transfer to a new 10 cm dish with PBS on ice.
 - d. Using dissection scissors and forceps, remove the E12 embryos from the uterus and transfer them into a new 10 cm dish with PBS on ice.

Note: PBS may become cloudy with blood or other tissue debris. Prepare three 10 cm dishes with PBS per mouse. Transfer the uterine horns to a new dish as needed to clearly visualize and cleanly dissect the embryos.

3. Select the embryos desired for spinal dissection (15 min).
 - a. Visualize the embryos under a fluorescent stereoscope and identify embryos that have fluorescent signals (Figures 1B–1D).
 - b. Save one *Hb9::gfp* negative embryo as a negative control sample to set the FACS gating.

Note: Embryos may need to be taken off ice for brief periods of time (several minutes) for this step. Try to limit the time embryos are at room temperature.

Optional: GFP goggles may also be used to identify embryos carrying the *Hb9::gfp* reporter.

- c. Transfer *Hb9::gfp* embryos to individually numbered wells of a 12 well plate with 2 mL PBS on ice.
- d. Clean forceps thoroughly with 70% ethanol.
- e. Remove a piece of the tail from each embryo with forceps and transfer to a pre-labeled PCR strip tube for genotyping.

△ CRITICAL: Clean the forceps by spraying them with 70% ethanol once and wiping down. This should be repeated for each embryo to avoid cross-contaminating samples for genomic DNA extraction.

Note: You may also submit tail samples for genotyping by RT-PCR based methods (e.g. as provided by Transnetyx, Inc.) as a more accurate *post hoc* verification of genotypes.

4. Perform rapid genotyping (2 h).
 - a. Extract genomic DNA with QuickExtract according to the manufacturer's protocol.
 - i. Add 100 uL of QuickExtract buffer to each PCR tube with tail snips.
 - ii. Heat the samples on a thermocycler to 65C for 8 min.
 - iii. Remove from the thermocycler and vortex for 30 s.
 - iv. Heat the samples on a thermocycler to 98C for 2 min.
 - v. Remove the samples from the thermocycler and vortex for 30 s.

Note: Only a very small amount of starting material is needed for genotyping, such as the last 0.5 mm of the tail of an E12 embryo. Using excess tissue (such as 4–5 mm of tail length) may interfere with genomic DNA extraction or subsequent PCR.

Note: DNA can also be isolated with commercially available DNA extraction kits. Many commonly used kits will produce DNA with high purity, though these kits typically take more time to complete.

- b. Perform genotyping PCR for miR-218-1 and miR-218-2 alleles
 - i. If there is any large cellular debris remaining in the genomic DNA, spin the samples down and use the supernatant
 - ii. Follow the manufacturer's protocol for 2× GoTaq Master Mix using the following primers:

Gene	Forward primer	Reverse primer	Band sizes	Deletion size
miR-218-2	ggcagtcctgtgtcacaag	Tggatctcggggttcacgg	WT: 998 bp/KO: 712 bp	286 bp deletion
miR-218-1	accctctttcaatccctga	Ctttctgctgatccgattctg	WT: 841 bp/KO: 482 bp	359 bp deletion

iii. Start the thermocycler using the following protocol

PCR cycling conditions			
Steps	Temperature	Time	Cycles
Initial Denaturation	95°C	2 min	1 ×
Denaturation	95°C	30 s	30 ×
Annealing	60°C	30 s	
Extension	72°C	1 min	
Final extension	72°C	1 min	1 ×
Hold	4°C	Indefinitely	

iv. Run the PCR product on an agarose gel to resolve WT and KO bands

Note: The expected frequency of embryos carrying the desired alleles may be 1/8 or lower. Thus, having genotyping results prior to performing FACS will reduce the time required on the FACS machine and the time that samples spend on ice prior to processing. It is helpful to have another lab member set up the genotyping PCR while the dissection and dissociation are ongoing. Performing genotyping prior to loading cells onto the 10× controller is important to reduce the cost of preparing samples of the incorrect genotype.

5. Prepare the embryo for removal of spinal tissue (1–3 min per embryo).
 - a. Remove the head and discard by cutting at the indicated position (Figures 2A and 2B).
 - b. Pin embryo ventral side up onto a Sylgard dish filled with ice cold PBS (Figure 2C).
 - c. Eviscerate the contents of the thoracic and abdominal cavities using forceps to first remove the overlying ectoderm, then pinching and pulling organs until they are released from their attachment points (Figure 2D).

Note: These steps are difficult to perform at 4°C since they require working under a stereomicroscope which often does not accommodate ice buckets. Dissection dishes can be pre-chilled on ice and filled with ice-cold PBS to reduce the temperature of the tissues during dissection.

6. Remove the spinal cord (2–5 min per embryo).
 - a. Remove dissection pins and place the eviscerated embryo dorsal side up.
 - b. Securely pin each limb to the dish (Figure 3A).
 - c. Using a microdissection knife, cut along the dorsal midline along the full length of the embryo to open the dorsal neural tube (Figure 3B).
 - d. Cut along the left and right sides of the spinal cord along the full length of the embryo to free the spinal cord from the rest of the embryo (Figures 3C and 3D). Note that the left and the right hemi-cords are still connected ventrally.
 - e. Remove the meninges.

Note: Removal of the meninges is technically challenging and may require some practice. The meninges will adhere more tightly to the spinal cord as embryos sit out over time, making meninges removal very challenging. Therefore, it is ideal to be time efficient with dissections.

- i. Place the spinal cord on its side and identify the meninges on the ventral surface (Figure 4A).
- ii. Using forceps, gently loosen the meninges from the rostral region.
- iii. Using one set of forceps, grip the spinal cord gently but securely, so as not to crush the tissue.
- iv. With another set of forceps, secure the meninges tightly (Figure 4B).
- v. Quickly and smoothly detach the meninges by pulling in the caudal direction (Figure 4C). [Problem 1.](#)

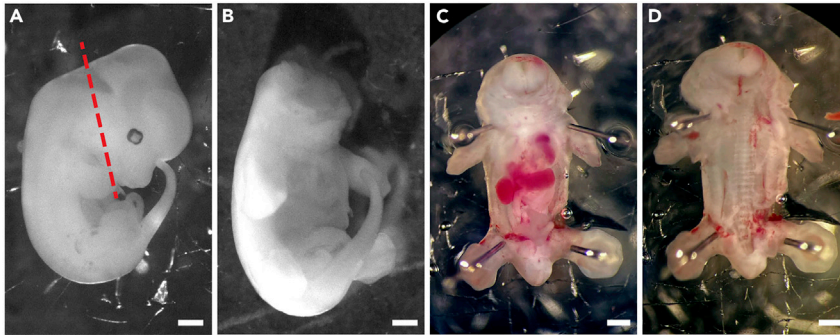


Figure 2. Decapitation and evisceration of thoraco-abdominal contents

(A) Brightfield image of a lateral view of an E12 embryo. The dotted line designates the optimal location for decapitation to separate the meninges in later steps.

(B) Embryo after decapitation.

(C) The embryo was pinned to a Sylgard dish ventral side up using dissection pins securely placed in the proximal area of the limbs.

(D) Same embryo after evisceration of the thoracic and abdominal contents. Scale bar 1 mm.

vi. Discard the meninges and any non-spinal tissues.

△ **CRITICAL:** Inspect the spinal cord to ensure all the spinal cord regions are intact by placing pins in the ends of the spinal cord. Loss of particular rostro-caudal spinal regions or regions of the spinal cord will affect the capture of the same motor neuron populations across replicates (Figures 4D–4G).

f. Transfer the spinal cord to a labeled Eppendorf tube with 500 uL PBS and keep on ice.

7. Repeat steps 5 and 6 for the remainder of the embryos.

Spinal tissue dissociation and FACS purification

⌚ **Timing:** 2–3 h (1 h for dissociation and 1–2 h for FACS purification of all samples)

The goal of these steps is to dissociate the spinal tissue into single cells for FACS purification of GFP+ motor neurons.

8. Prepare Papain dissociation reagents according to the manufacturer's protocol.
9. Aspirate PBS from spinal cord samples.
10. Add 1mL reconstituted papain solution (20 units per mL) to each spinal sample, place at 37°C, and set timer for 30 min.
11. Triturate the spinal tissue with a P1000 pipette tip about 10 times. If large tissue chunks remain, triturate with a P200 or incubate at 37°C for another 10 min and triturate again until no large pieces remain.
12. Centrifuge dissociated cells at 200 rcf for 5 min at room temperature (20°C–22°C) in a spinning bucket centrifuge and discard the supernatant.
13. Add 1 mL reconstituted and diluted albumin-ovomucoid inhibitor solution (10 mg per mL) with DNase (10 units DNase) per the manufacturer's protocol.
14. Centrifuge cells at 200 rcf for 5 min at room temperature in a spinning bucket centrifuge and discard the supernatant.
15. Resuspend cell pellet in sorting buffer (1:1 Neurobasal:DMEM/F12 without phenol red with 3% Horse Serum and 10 units per mL DNase solution).

△ **CRITICAL:** Sorting buffer must contain DNase or cells may clump, causing yield to significantly decrease.

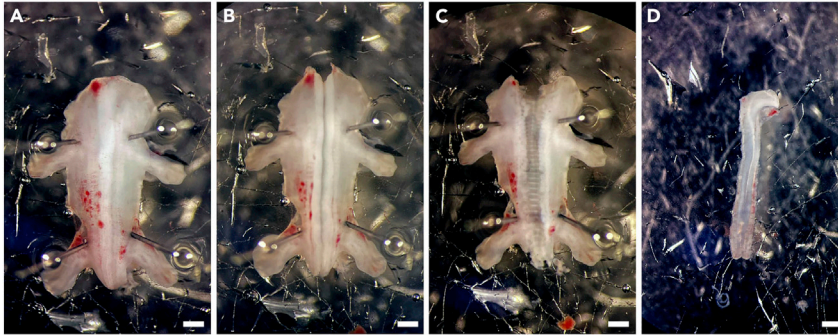


Figure 3. Removal of spinal cord from the embryo

(A) The embryo was reoriented dorsal side up and again securely attached to the Sylgard dish with dissection pins. (B) The dorsal axis of the spinal cord is cut vertically using a dissection knife. (C and D) After cutting both sides of the spinal cord with microdissection scissors, the spinal cord (D) is released leaving the rest of the embryo attached to the Sylgard (C). Scale bar 1mm.

Note: The addition of 3 μ M DAPI to the sorting buffer can aid in the exclusion of dead or dying cells by FACS gating. Ensure the cells have been incubating in DAPI for at least 15 min. DAPI can be left in the sorting buffer during FACS.

16. Pass cells through a 40 μ m cell strainer.
17. Sort dissociated cells from each embryo separately on a FACS machine.

△ CRITICAL: Setting the gating on FACS machine and maintaining the same gates for all samples is important to reduce variability in motor neuron capture. Motor neuron subpopulations express different levels of endogenous Hb9 protein and, resultingly, of the *Hb9::gfp* reporter. Therefore, changing the gating settings may enrich or deplete some motor neuron pools.

18. Collect GFP+ motor neurons into 1 mL chilled sorting buffer (1:1 Neurobasal:DMEM/F12 without phenol red with 3% Horse Serum and 10 units per mL DNase solution) in an Eppendorf tube for single cell RNA sequencing preparation.

Note: We typically capture 5,000–10,000 motor neurons per E12 spinal cord. The number of motor neurons captured is affected by the age of the embryo, precision of the spinal dissection, and completeness of the papain dissociation.

10 \times genomics cell preparation

⌚ Timing: 2–3 days

19. Quantify motor neuron cells per volume.
 - a. Calculate the volume of media required to obtain the desired concentration of cells.
 - b. Spin collected cells at 200 rcf for 5 min at 4°C.
 - c. Gently remove media to desired volume with 10 μ L excess for hemocytometer cell counting.
 - i. Mix 5 μ L of the cell suspension with 5 μ L trypan blue in a separate tube, load onto a hemocytometer, and count cells.
 - ii. Repeat once for confirmation of cell count. [Problem 2:](#)
20. Load the resuspended cells into the 10 \times Chromium Controller following the manufacturer's instructions.

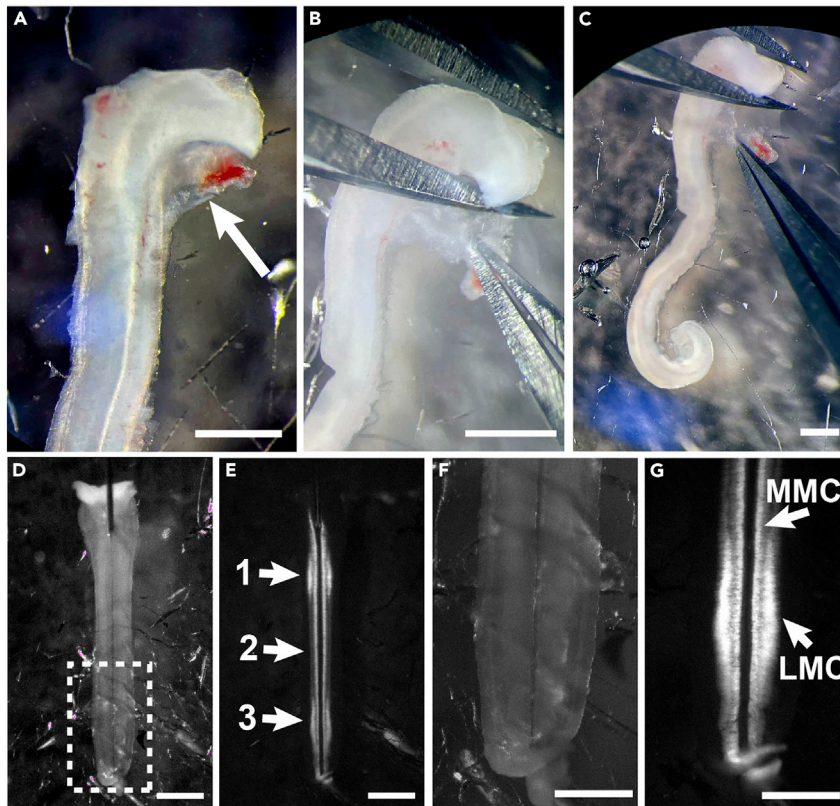


Figure 4. Removal of the meninges from the spinal cord

(A) The rostral meninges (arrow) is loosely attached to the spinal cord.

(B) One forceps was used to gently secure the rostral end of the spinal cord and the other forceps was used to firmly grasp the meninges on the ventral side.

(C) With the spinal cord secured, the meninges were pulled down in a firm and steady motion until completely removed.

(D) Brightfield image of the isolated spinal cord that is secured to the Sylgard dish with insect pins.

(E) GFP expression from the *Hb9::gfp* allele enables visualization of cervical (1), thoracic (2) and lumbar (3) regions of the spinal cord.

(F and G) Brightfield (F) and GFP (G) views of the lumbar spinal cord at higher magnification enable visualization of the lateral (LMC) and medial motor columns (MMC). Scale bar (A–E) 1 mm; (F–G) 0.5 mm.

21. Prepare sequencing libraries following the manufacturer's instructions (10× Genomics). In [Amin et al. \(2021\)](#), 14 cycles were used for cDNA amplification, and 10 cycles were used for library amplification using V2 reagents.

▣▣ **Pause point:** Libraries can be stored at -20°C prior to sequencing.

22. Sequence the resulting sequencing library following the manufacturer's instructions. In [Amin et al. \(2021\)](#), the libraries were sequenced with paired end reads, with a Read 1 of 26 basepairs and a Read 2 of 98 basepairs, on an Illumina HiSeq 4000 at the UCSD IGM Genomics Center.

Note: When performing differential expression with single cell RNA sequencing data, it is important to have biological replicates to limit the detection of spurious associations ([Squair et al., 2021](#)). We found a WT versus DKO ($n = 2,2$) comparison with single cell RNA sequencing resulted in pseudobulk differential expression results that detected many of the same differentially expressed target genes as our WT versus DKO ($n = 3,3$) bulk RNA sequencing results ([Amin et al., 2021](#)). The ability to accurately identify differentially expressed genes will be significantly higher when having even more replicates but will substantially increase the cost.

Preparing single cell RNA sequencing data for downstream analysis

⌚ Timing: 1 week

The goal is to filter out low quality cells and prepare the data for downstream analysis.

23. Demultiplex raw sequencing data. In [Amin et al. \(2021\)](#), bcl2fastq v2.18.0.12 was used.
24. Align and quantify reads from the samples separately for gene expression using Cellranger version 3 (10X Genomics) using the mm10 mouse transcriptome annotation.
25. Combine samples into a single expression matrix using Cellranger's 'aggr' program by equalizing aligned read counts. Use the estimated cell count from Cellranger for downstream analysis.
26. Import the resulting UMI matrix and sample identities into R.
27. Filter out low quality cells and doublets.
 - a. We generated a histogram of the number of detected features per-cell. We then selected lower and upper bounds that retained the dominant population of cells in order to eliminate low quality cells and doublets. [Problem 3](#).
 - b. We also generated a histogram of the percentage of mitochondrial gene expression per cell and selected an upper limit of 8%.
 - c. These two filters removed 796 cells, leaving 8,247 cells for downstream analysis.

Note: Lower sequencing depths are often sufficient for identifying cellular populations using single cell RNA sequencing, however, higher sequence depths are likely needed for effective differential gene expression testing. We sequenced a total of 1,128,753,214 reads across four samples. After filtering steps, we detected 4762 (median) genes per cell, 21,268 (median) unique molecular identifiers (UMIs) per cell, and 19,410 total genes detected (minimum of three cells).

28. Perform dimensionality reduction.
 - a. Identify highly variable genes (HVG) using a method previously described ([Brennecke et al., 2013](#)). We used FDR = 0.1.

Note: If using software such as Seurat, p values may not be provided. You can test a range of HVGs in these steps to see if and how it changes the interpretation of your analyses. Adjusting the HVG cutoff can help reveal detail in the data that you would otherwise miss if all genes were included, or if you use all genes over certain expression levels (as is often done with bulk RNA-seq analyses).

- b. Perform normalization by dividing each cell's UMI counts by its respective total UMI count across all genes and scaling by 10,000, resulting in an expression value comparable to TPM in bulk RNA sequencing.
- c. Perform PCA on normalized and log₂+1 scaled UMI using the identified HVGs.
- d. Identify significant components by comparing the eigenvalues to the 95th percentile of the appropriate Wishart distribution.

Note: We used the Wishart method but it's advisable to observe a scree plot of the eigenvalues and find the "elbow" in the curve as a second opinion. Sometimes the Wishart method will over-estimate the number of significant components.

- e. Generate UMAP. We used 'n_neighbors=20'.

Note: Changing the number of neighbors used in the UMAP algorithm will affect the outcome. Lower values will result in smaller and more numerous groupings of cells (better for fine detail analysis such as finding subtypes of a single cell type) while higher values will result in larger

and fewer groups of cells (better for large scale differences such as different cell types). These setting alterations will need to be adjusted to address the specific biological question being asked and the inherent diversity of the cell population that was sequenced.

- f. Perform dimensionality reduction separately for all cells and the interneuron-excluded subset (if contaminating interneurons are captured).

Note: *Hb9::gfp* may have background levels of expression in some spinal interneuron populations and may also be expressed in Hb9 spinal interneurons. Thus, you may capture some interneuron populations despite optimizing FACS gating settings. The transcriptomic identity of interneurons is very distinct from motor neurons and can be clearly distinguished by UMAP clusters or expression of neurotransmitters and transcription factor marker genes.

Pseudo-bulk gene expression comparison

⌚ Timing: 1 day

The goal of these steps is to compare and perform differential gene expression of a subset of neurons isolated from mice of different genotypes (e.g., differential expression of medial motor column neurons from control and miR-218 DKO backgrounds).

29. We summed the expression of individual cells within an identified cluster (e.g., medial motor column or phrenic/hypaxial motor neurons) on the gene level into pseudo-bulk expression tables. Each gene's un-normalized UMI counts were summed into bins based on sample identity and cluster assignment (e.g., medial motor column or phrenic/hypaxial motor neurons) resulting in a pseudo-bulk count matrix (4 samples per cluster assignment) [Problem 4](#).
30. The pseudo-bulk expression data were treated like bulk RNA-seq for expression normalization and differential expression testing performed using DESeq.
31. Perform Sylamer ([Bartonicsek and Enright, 2010](#); [van Dongen et al., 2008](#))
 - a. Use DESeq normalized gene expression for each population by genotype as data input.
 - b. Limit analysis to the top 10,000 most highly expressed genes. [Problem 5](#):

Note: If the quality and depth of sequencing reads is low, the number of genes included in Sylamer analysis might need to be reduced so that only the genes with the most accurately determined gene expression are included. We used this cutoff (i.e. top 10,000 most highly expressed genes) for both bulk RNA sequencing analysis and pseudo-bulk RNA sequencing analysis from single cell data for consistency.

- c. Rank the genes from highest enriched to lowest enriched in a given genotype (i.e., most highly enriched in WT medial motor column compared with DKO medial motor column neurons).
- d. Obtain a FASTA file of 3'UTR sequences. This can be obtained from UCSC Table Browser for the GENCODE_VM18 database with masked repeats with "N".

Note: 3'UTR sequences often contain low complexity repeated nucleotides that can reduce the quality of detection of enriched motifs. Masking repeats is an effective way to reduce this source of spurious correlations.

- e. For genes with multiple isoforms, the longest UTR sequence among annotated transcripts per gene was retained.

Note: For a given mRNA, there can be many annotated isoforms that vary with their 3'UTR sequence length. In any cell, multiple RNA isoforms may be present at varying abundance for a

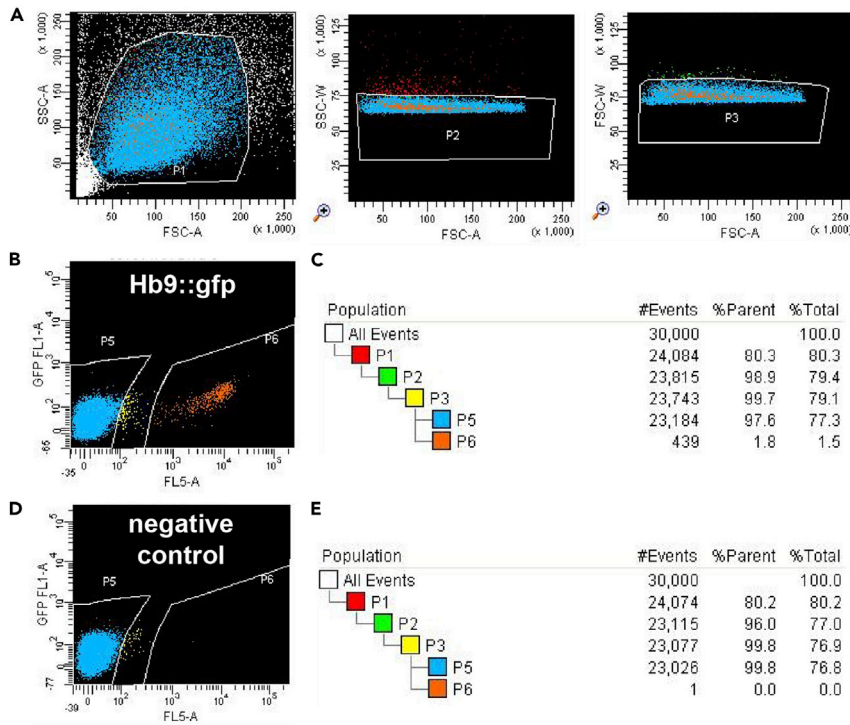


Figure 5. Example of FACS results

(A) Side and forward scatter plots of dissociated spinal cells and representative gating.
 (B) Collection of GFP+ motor neurons expressing *Hb9::gfp* (P6).
 (C) 1.5% of total cells are *Hb9::gfp*+.
 (D and E) A spinal cord without *Hb9::gfp* was run to validate gates.

single gene. Despite the complexity and diversity of RNA isoforms, we have found it effective to use the longest annotated UTR sequence for Sylamer analysis. It is also possible to generate isoform abundances (either de novo or annotated) from RNA sequencing data using available software, though we have not found this necessary to obtain robust data from Sylamer.

- f. FASTA sequence files and ranked gene lists were used with Sylamer software (Enright Lab) with the following settings: 8mer search (Markov correction: 5), 7mer search (Markov correction: 4), and 6mer search (Markov correction: 4) including known miRNA seed sequences.

EXPECTED OUTCOMES

An ideal outcome of the spinal dissection protocol is shown in Figure 4D. The spinal cord should be free of the meninges and surrounding tissues. It should also be intact without missing pieces. The rostral edge of the spinal cord should capture the *Hb9::gfp* neurons from the cervical spinal cord. Most cranial motor neurons do not express endogenous Hb9 protein and thus will not express the reporter. The caudal end of the spinal cord should be kept consistent between embryos so there is limited variability in the capture of motor neurons from the sacral regions.

Results of FACS purification of motor neurons is shown in Figures 5A–5E. Typically, 0.5%–2% of spinal neurons at E11.5–12.5 are GFP+. At later embryonic time points, this percentage decreases substantially due to relative expansion of other spinal neuron populations and the decreased intensity of transgene expression over time.

In our dataset, we identified 8,247 total cells from 4 samples, a minority of which were interneurons (Figure 6A). Among motor neurons, we identified major known subtypes of motor neurons across the

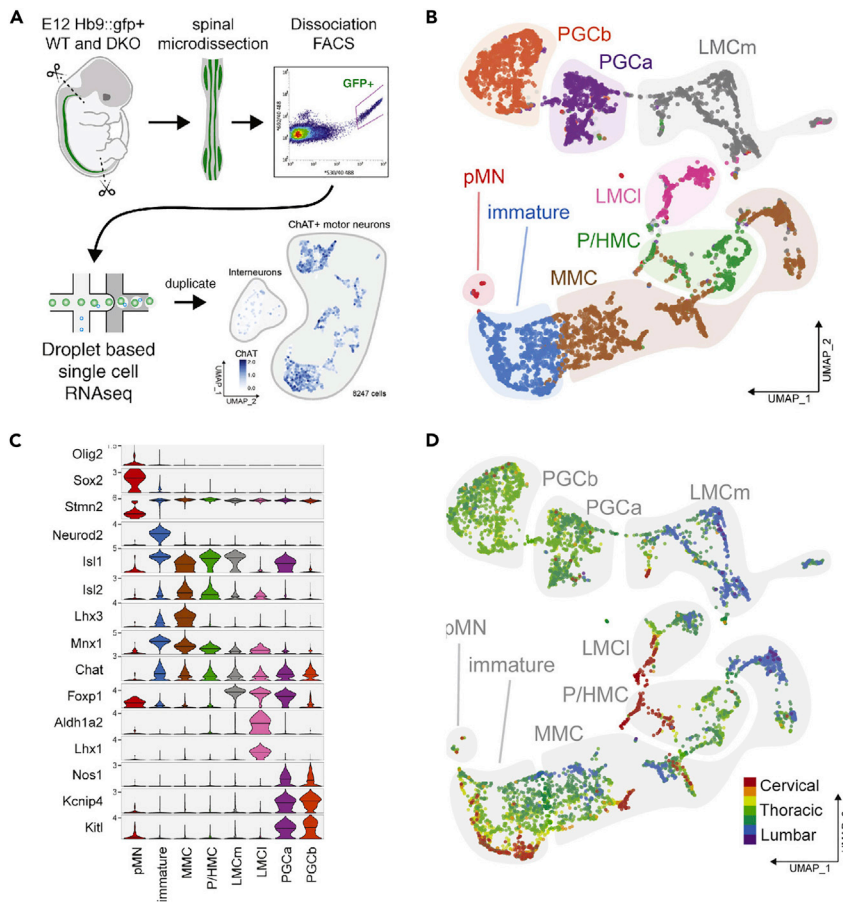


Figure 6. Representative results of single cell analysis (Amin et al., 2021)

(A) Experimental diagram showing the capture of *Hb9::gfp+* motor neurons from embryos using droplet based single cell RNA sequencing. A minority of cells are interneurons while the remainder are ChAT+ motor neurons. (B–D) Identification of spinal motor neurons columnar and rostro-caudal identities by canonical transcription factor expression and Hox gene expression.

rostral caudal axis and a novel segregation of preganglionic motor neurons that we named PGCa and PGCb (Figures 6B–6D).

Because we performed both bulk and single cell RNA sequencing on FACS-isolated *Hb9::gfp+* WT and DKO motor neurons (Amin et al., 2021), we could validate our pseudo-bulk differential expression approach. We compared Sylamer analysis using bulk ($n = 3,3$) and pseudobulk data ($n = 2,2$) and find that pseudobulk data is also able to detect the highly specific enrichment of miR-218's 8-mer seed sequence in 3'UTR sequences enriched in DKO compared with WT motor neurons (Figures 7A and 7B). The statistical significance of this enrichment was not as high as bulk RNA sequencing data, potentially reflecting the fewer number of replicates or inherently higher technical noise of single cell data. We also used pseudobulk expression data to perform analyses that are common to bulk RNA sequencing data, such as principal components analysis (PCA) and differential expression testing (Figures 7C and 7D).

LIMITATIONS

Mouse spinal microdissections are technically challenging and require practice to ensure the steps are completed reliably and within reasonable time frames to preserve tissue integrity and cell survival. This protocol requires the availability of a FACS machine and single cell RNA sequencing materials

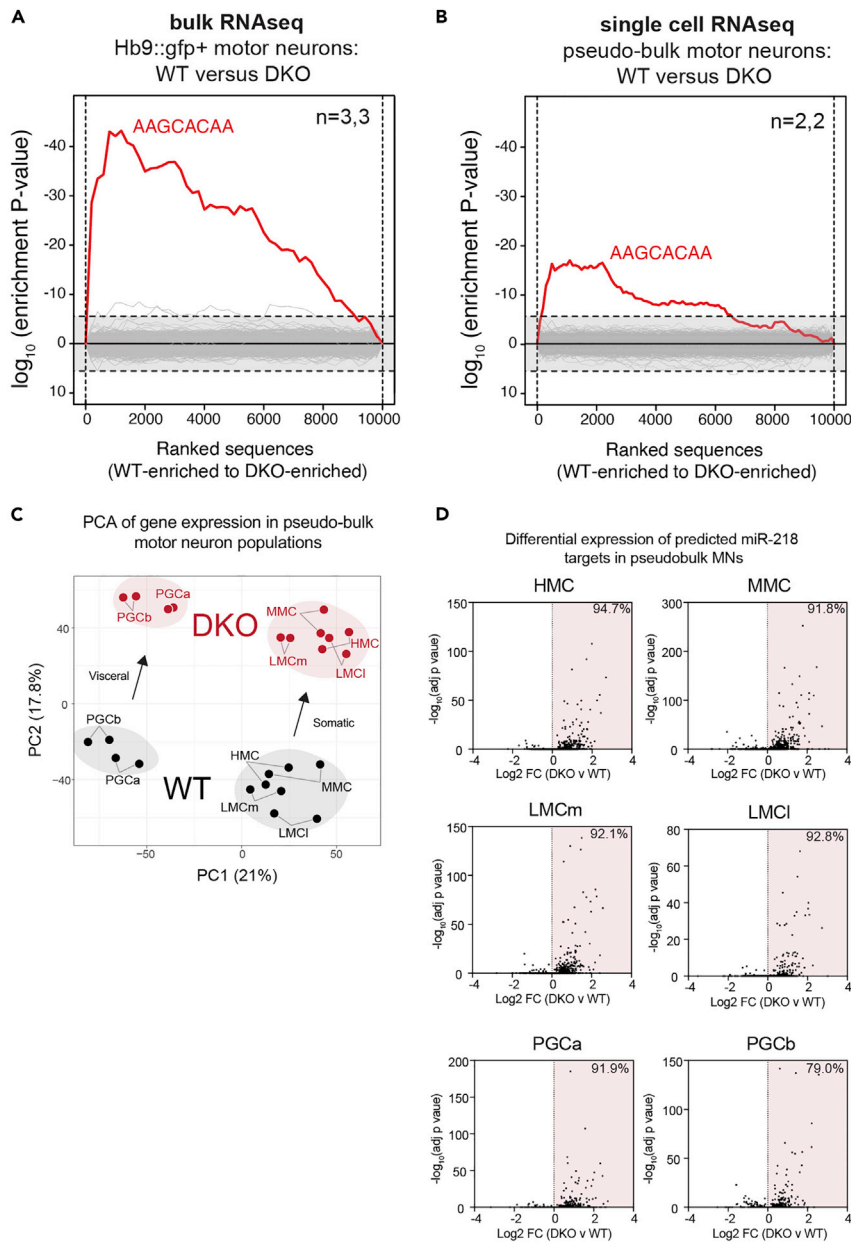


Figure 7. Example of gene expression analyses using single cell RNA sequencing pseudobulk data (Amin et al., 2021)

(A and B) We used Sylamer to detect the enrichment of 8 bp motifs in 3'UTRs enriched in WT versus DKO motor neurons using both (A) bulk and (B) single cell pseudo-bulk RNA sequencing data. In both analyses, only miR-218's seed match (AAGCACAA) is statistically enriched, though the magnitude of enrichment is greater with bulk RNA sequencing data.

(C) Principal components analysis of pseudo-bulk data from each sample, genotype, and motor column demonstrates that replicates cluster by genotype.

(D) Differential expression testing of miR-218's predicted mRNA targets by motor neuron subtype. The percentage of differentially expressed genes that are de-repressed is displayed.

that are expensive. Single cell RNA sequencing data is limited by intrinsic signal noise, dropouts of genes, and inability to detect specific mRNA isoforms that may be otherwise detected in bulk RNA sequencing data or qPCR. The quality of pseudo-bulk data sets for differential expression testing is unlikely to be as high as bulk RNA sequencing data, even with high sequencing depths.

TROUBLESHOOTING

Problem 1

The meninges is difficult to remove from the spinal cord without destroying the tissue (step 6).

Potential solution

The meninges become harder to remove as *ex vivo* tissues sit out over time. It is important to work quickly and practice the method of removing the meninges on control embryos prior to performing the experiment. If the meninges still cannot be removed from the spinal cord, you may resort to micro-dissecting away tissues adjacent to the spinal cord prior to incubating in papain for dissociation. Without an efficient dissection of the spinal cord, dissociation of the spinal cord into single cells will be prolonged and the percentage of GFP+ motor neurons will be very low (less than 0.5% of all cells) resulting in prolonged time on the FACS machine.

Problem 2

The number of cells captured for single cell RNA sequencing is much lower than expected (step 19).

Potential solution

After FACS collection of motor neurons, some neurons may adhere to the sides of Eppendorf tubes causing consequential cell loss. Some cells may also be lost when spinning down the samples during cell concentration. For these reasons, it is important to perform accurate cell counting with a hemocytometer after these steps have been performed. Cells may settle when sitting on the bench for any amount of time, so adequate resuspension of cells prior to loading them onto the Chromium Controller is important. You may use low protein binding tubes minimizing cell loss due to adherence to the walls of the tube. Furthermore, not all captured cells will pass bioinformatics quality control steps, so it is important to anticipate loading slightly more cells than needed. Be aware that capturing many more cells than desired will result in much higher sequencing costs to achieve the desired sequencing depth needed for differential gene expression testing.

Problem 3

Low quality cells and doublets are difficult to identify from the histogram of the number of features per cell (step 27).

Potential solution

There will always be some population of low quality of cells in single cell sequencing data that should be excluded from downstream analyses. Low quality cells may result from shearing of cell contents, cell death, or other causes. Close examination of the histogram of features per cell can be informative and should be considered in context of the cell population you are examining. For example, bimodal histograms can indicate two populations (one with higher average RNA content and one with lower) but they can also be two populations of the same cells (one high quality and one lower quality). This is exemplified by spinal cord samples in which motor neurons intrinsically contain significantly more RNA than interneurons, reflected by the overall larger soma size of motor neurons compared with other spinal neurons. Differences in RNA content are also observed among proliferating cells and cell types of different lineages. Lower quality cells characteristically tend to form clusters without any strongly enriched genes, unlike high quality cells with lower RNA content. If you have knowledge of the expected gene expression of the cells you sequenced, then low-quality cells can be filtered out via clustering and cell-type calling. By examining these characteristics of cells with low numbers of features, you can more effectively identify low quality cells for exclusion.

Problem 4

Identification of motor neuron subpopulations is ambiguous for some cells (step 29).

Potential solution

Motor neurons are highly diverse along the rostro-caudal axes, over developmental time, and within columns and pools. It is important to examine gene expression using multiple known marker genes of motor subtypes. For example, the phrenic motor column and hypaxial motor column have similar motor neuron transcription factor marker genes but can be distinguished by Hox code expression patterns. Referencing known databases of gene expression profiles for neuronal subtypes is important (Alaynick et al., 2011) for motor neuron identification from single cell RNA sequencing data.

Problem 5

Sylamer is not detecting the de-repression of miRNA targets (step 31).

Potential solution

Intrinsic noise from cellular preparation, sequencing, and transcript identification can limit the detection of miRNA target derepression. Furthermore, Sylamer may not be able to detect the repression of miRNAs that are expressed at low levels due to poor signal to noise ratios. Be sure to search for 6, 7 and 8-mer sequences and vary the Markov correction to eliminate spurious sequence associations that can occur with hypergeometric enrichment analyses. Depending on the sequencing quality, you can modify the number of transcripts analyzed. For example, if your data quality is lower, low expressed genes will likely exhibit very high variance causing differential expression testing methods to be unsuccessful. We consistently found that 10,000 genes were sufficient to detect miR-218 related changes across sequencing modalities and samples. However, this cutoff may have to be lowered for other miRNAs that are not as highly expressed or for lower quality data sets. Also, be sure to mask repeats in your FASTA files so that you do not end up picking up nucleotide repeats of low complexity instead of signals from miRNA binding motifs.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Sam Pfaff (spfaff@salk.edu).

Materials availability

Mouse lines generated in this study are available upon request.

Data and code availability

The bulk and single cell RNA sequencing datasets generated during this study are publicly available at ArrayExpress accession: E-MTAB-10571. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We thank: Karen Lettieri and Miriam Gullo (Salk Institute) for mouse husbandry; Flow Cytometry Core Facility (Salk Institute; NIH-NCI CCSG: P30 014195 and Shared Instrumentation Grant S10-OD023689 (Aria Fusion cell sorter); The Salk Institute Next Generation Sequencing Core (Nasun Hah and other members; NIH-NCI CCSG: P30 014195, the Chapman Foundation and the Helmsley Charitable Trust), the UCSD IGM Genomic Center (Kristen Jepsen and other members); S.L.P. is a Benjamin H. Lewis chair in neuroscience. N.D.A. is supported by the NIMH T32 Training Grant T32MH019938, NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation, and the Stanford Department of Psychiatry and Behavioral Sciences. Additional funding sources: National Institutes of Health (1 R03 NS121480-01 and 1 R21 NS121846-01), TargetALS, and The Department of Defense (ALSRP (TIA) W81XWH1810120).

AUTHOR CONTRIBUTIONS

N.D.A. optimized spinal cord isolation protocol for FACS, performed the biological experiments with M.H. G.S. acquired images. N.D.A. analyzed bulk RNA sequencing data and S.P.D. and N.D.A. analyzed single cell sequencing data. N.D.A. and S.L.P. wrote and edited the manuscript with G.S. and M.H.

DECLARATION OF INTERESTS

The authors note a related US patent related to motor neuron specific expression vectors (US patent US Patent 10,729,790).

REFERENCES

- Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Gruning, B.A., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46, W537–W544. <https://doi.org/10.1093/nar/gky379>.
- Agarwal, V., Bell, G.W., Nam, J.W., and Bartel, D.P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *Elife* 4, e05005. <https://doi.org/10.7554/eLife.05005>.
- Alaynick, W.A., Jessell, T.M., and Pfaff, S.L. (2011). SnapShot: spinal cord development. *Cell* 146, 178–178.e171. <https://doi.org/10.1016/j.cell.2011.06.038>.
- Amin, N.D., Bai, G., Klug, J.R., Bonanomi, D., Pankratz, M.T., Gifford, W.D., Hinckley, C.A., Sternfeld, M.J., Driscoll, S.P., Dominguez, B., et al. (2015). Loss of motoneuron-specific microRNA-218 causes systemic neuromuscular failure. *Science* 350, 1525–1529. <https://doi.org/10.1126/science.aad2509>.
- Amin, N.D., Senturk, G., Costaguta, G., Driscoll, S., O’Leary, B., Bonanomi, D., and Pfaff, S.L. (2021). A hidden threshold in motor neuron gene networks revealed by modulation of miR-218 dose. *Neuron* 109, 3252–3267.e6. <https://doi.org/10.1016/j.neuron.2021.07.028>.
- Bartonicek, N., and Enright, A.J. (2010). SylArray: a web server for automated detection of miRNA effects from expression data. *Bioinformatics* 26, 2900–2901. <https://doi.org/10.1093/bioinformatics/btq545>.
- Brennecke, P., Anders, S., Kim, J.K., Kolodziejczyk, A.A., Zhang, X., Proserpio, V., Baying, B., Benes, V., Teichmann, S.A., Marioni, J.C., and Heisler, M.G. (2013). Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* 10, 1093–1095. <https://doi.org/10.1038/nmeth.2645>.
- Carroll, A.P., Tooney, P.A., and Cairns, M.J. (2013). Context-specific microRNA function in developmental complexity. *J. Mol. Cell Biol.* 5, 73–84. <https://doi.org/10.1093/jmcb/mjt004>.
- Gebert, L.F.R., and MacRae, I.J. (2019). Regulation of microRNA function in animals. *Nat. Rev. Mol. Cell Biol.* 20, 21–37. <https://doi.org/10.1038/s41580-018-0045-7>.
- Lee, S.K., Jurata, L.W., Funahashi, J., Ruiz, E.C., and Pfaff, S.L. (2004). Analysis of embryonic motoneuron gene regulation: derepression of general activators function in concert with enhancer factors. *Development* 131, 3295–3306. <https://doi.org/10.1242/dev.01179>.
- Metsalu, T., and Vilo, J. (2015). ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* 43, W566–W570. <https://doi.org/10.1093/nar/gkv468>.
- Squair, J.W., Gautier, M., Kathe, C., Anderson, M.A., James, N.D., Hutson, T.H., Hudelle, R., Qaiser, T., Matson, K.J.E., Barraud, Q., et al. (2021). Confronting false discoveries in single-cell differential expression. *Nat. Commun.* 12, 5692. <https://doi.org/10.1038/s41467-021-25960-2>.
- Stifani, N. (2014). Motor neurons and the generation of spinal motor neuron diversity. *Front. Cell. Neurosci.* 8, 293. <https://doi.org/10.3389/fncel.2014.00293>.
- van Dongen, S., Abreu-Goodger, C., and Enright, A.J. (2008). Detecting microRNA binding and siRNA off-target effects from expression data. *Nat. Methods* 5, 1023–1025. <https://doi.org/10.1038/nmeth.1267>.