Protocol
Transcriptional profiling of mouse projection neurons with VECTORseq

Existing techniques for transcriptional profiling of projection neurons could be applied to only one neuronal population per experiment. To increase throughput, we developed VECTORseq, which repurposes retrogradely infecting viruses to deliver multiplexable RNA barcodes, enabling projection anatomy to be read out in single-cell datasets. In this protocol, we describe the delivery of viral barcodes to mouse brain to label different projection neurons. We then detail single-cell or nuclei isolation for sequencing, followed by the analysis of single-cell sequencing data.

Publisher’s note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.
Protocol

Transcriptional profiling of mouse projection neurons with VECTORseq

Victoria Cheung,1,2,5 Philip Chung,3 and Evan H. Feinberg1,4,6,*

1Department of Anatomy, University of California, San Francisco, San Francisco, CA 94158, USA
2Tetrad Graduate Program, University of California, San Francisco, San Francisco, CA 94158, USA
3Department of Anesthesiology & Pain Medicine, University of Washington, Seattle, WA 98195, USA
4Kavli Institute for Fundamental Neuroscience, University of California, San Francisco, San Francisco, CA 94158, USA
5Technical contact: victoria.cheung@gmail.com
6Lead contact
*Correspondence: evan.feinberg@ucsf.edu
https://doi.org/10.1016/j.xpro.2022.101625

SUMMARY

Existing techniques for transcriptional profiling of projection neurons could be applied to only one neuronal population per experiment. To increase throughput, we developed VECTORseq, which repurposes retrogradely infecting viruses to deliver multiplexable RNA barcodes, enabling projection anatomy to be read out in single-cell datasets. In this protocol, we describe the delivery of viral barcodes to mouse brain to label different projection neurons. We then detail single-cell or nuclei isolation for sequencing, followed by the analysis of single-cell sequencing data.

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

BEFORE YOU BEGIN

This protocol is used to prepare mice for stereotaxic surgeries and subsequently single-cell or single-nuclei isolation. However, it is likely that the VECTORseq approach could be applied to examine projection neurons in other species. This study was approved by the Institutional Animal Care and Use Committee at the University of California, San Francisco. All animal studies must be approved by an Institutional Animal Care and Use Committee (IACUC) and performed in accordance with IACUC guidelines.

KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Antibodies          |        |            |
| Rabbit anti-NeuN conjugated to Alexa Fluor 488 | Abcam | Cat# ab190195 |
| Bacterial and virus strains | Addgene | |
| AAVrg-EF1a-mCherry-IRES-Cre | Cat# 55632-AAVrg |
| AAVrg-EF1a-FLPo | Cat# 55637-AAVrg |
| AAVrg-hSyn-Dre | Cat# 50363-AAVrg |
| AAV1-CAG-FLEX-tdTomato | Cat# 28306-AAV1 |
| Critical commercial assays | | |
| Papain Dissociation System Protocol (includes Ovomucoid inhibitor, Earle’s Balanced Salt solution, DNase) | Worthington Biochemical Corporation | Cat# LK003150 |

(Continued on next page)
**Continued**

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| 5′ v2 Library prep kit | 10x Genomics | Cat# 1000265 |
| Chromium Next GEM Chip K Single Cell Kit, 16 rxns | 10x Genomics | Cat# 1000287 |
| Chromium Next GEM Single Cell 5′ Library and Gel Bead Kit v1.1, 4 rxns | 10x Genomics | Cat# 1000167 |
| Chromium Next GEM Chip G Single Cell Kit, 16 rxns | 10x Genomics | Cat# 1000127 |
| Neuron isolation kit | Miltenyi | Cat# 130-126-603 |

**Deposited data**

| Deposited data | Source | Identifier |
|----------------|--------|------------|
| Raw data | Cheung et al. (2021) | NCBI GEO (GSE189907) |
| Analysis | Cheung et al. (2021) | [https://github.com/vic-cheung/vectorseq](https://github.com/vic-cheung/vectorseq) [https://doi.org/10.5281/zenodo.5703724](https://doi.org/10.5281/zenodo.5703724) |

**Experimental models: Organisms/strains**

| Adult male or female (8–12 weeks) | C57BL/6J | The Jackson Laboratory | Cat# 000664 |

**Software and algorithms**

| Software and algorithms | Source | Identifier |
|-------------------------|--------|------------|
| Cellranger 6.0.0 | Zheng et al. (2017) | [http://10xgenomics.com/](http://10xgenomics.com/) |
| Scanpy 1.7.2 | Wolf et al. (2018) | [https://github.com/theislab/scanpy](https://github.com/theislab/scanpy) |
| Scrublet 0.2.1 | Wolock et al. (2019) | [https://github.com/swolock/scrublet](https://github.com/swolock/scrublet) |
| Python 3.9 | Python Software Foundation | [https://www.python.org/downloads/](https://www.python.org/downloads/) |

**Chemicals, peptides, and recombinant proteins**

| Chemicals, peptides, and recombinant proteins | Source | Identifier |
|-----------------------------------------------|--------|------------|
| Triton X-100 | Sigma-Aldrich | Cat# X100 CAS: 9002-93-1 |
| Kollidon VA64 | BASF Pharma | CAS: 25086-89-9 |
| Nuclease-Free Water | HyClone Products | Cat# 82007-334 CAS: 7732-18-5 |
| MgCl₂ (1 M stock) | Crescent Chemical Co Inc | Cat# 50248458 |
| Glucose | Supelco Inc | Cat# NC0651806 |
| K₂SO₄ | Alfa Aesar | Cat# AAA139750B CAS: 7778-80-5 |
| Na₂SO₄ | Chem-Impex Intl Inc | Cat# 50491082 CAS: 7757-82-6 |
| RNAs Inhibitor Lucigen | Lucigen Corporation | Cat# NC1819281 |
| HEPES | MilliporeSigma | Cat# 80502-686 CAS: 7365-45-9 |
| MgSO₄ | Sigma-Aldrich | Cat# 746452 CAS: 7487-88-9 |
| CaCl₂ (1 M stock) | Sigma-Aldrich | Cat# 21115-100ML CAS: 10043-52-4 |
| CaCl₂ (powder) | Fisher Scientific | Cat# 575069 CAS: 10035-04-8 |
| Pyruvic Acid | Sigma-Aldrich | Cat# 107360 CAS: 127-17-3 |
| Ascorbic Acid | Sigma-Aldrich | Cat# 795437 CAS: 50-81-7 |
| NMDG | Sigma-Aldrich | Cat# M2004 CAS: 6284-40-8 |
| HCI | Sigma-Aldrich | Cat# 320331 CAS: 7647-01-0 |
| KCI | Tocris | Cat# 31471KG CAS: 7447-40-7 |
| NaH₂PO₄ | Research Products International Corp | Cat# 50489015 CAS: 10049-21-5 |
| NaHCO₃ | Aldon Corp SE | Cat# 470302-438 CAS: 144-55-8 |

(Continued on next page)
| REAGENTS or RESOURCES | SOURCE | IDENTIFIER |
|------------------------|--------|------------|
| Kynurenic Acid (KYNA)  | Tocris  | Cat# 3694100R  
CAS: 1969901 |
| DAPI                   | Thermo Fisher Scientific | Cat# 62248 |
| autoMACS Rinsing Solution | Miltenyi | Cat# 130-091-222 |
| Magnetic-Activated Cell Sorting Bovine Serum Albumin (MACS BSA) Stock Solution | Miltenyi | Cat# 130-091-376 |
| Red Blood Cell Lysis Solution (10x) | Miltenyi | Cat# 130-094-183 |
| NaCl                   | Neta Scientific Inc. | Cat# 7647-14-5 |
| Carbogen               | Airgas  | Z020X9512000033 |

**Other**

| Model 940 Small Animal Stereotaxic Instrument | Kopf | N/A |
| UMP3 UltraMicroPump | World Precision Instruments | UMP3-1 |
| Hand Warmer, 2-1/4 In. X 3-1/2 In. | HotHands | 26KF14 |
| Streile drapes | Henry Schein | 9004686 |
| Nair Hair Remover | Church & Dwight | 22339 |
| BD Aria II FACS Sorter | BD Biosciences | N/A |
| Vibratome | Leica | Cat# VT1200s |
| NovaSeq 6000 | Illumina | N/A |
| -20°C freezer | SCI | Cat# SCGP21OW1AF |
| Spring Scissors – 8 mm Cutting Edge, Straight, Sharp | World Precision Instruments | Cat# 15024-10 |
| Large Scissors for Decapitation; Blunt 6.5 cm scissors | VWR | Cat# 82027-594 |
| Curved Dissecting Forceps | Fisher Scientific | Cat# 508097 |
| Stainless Steel 19 cm Spoon | VWR | Cat# 89259-968 |
| Razor Blade | Uline | Cat# 130013 |
| Disposable Vacuum Filter/Storage Systems, Pore Size: 0.2um, Membrane Material: Cellulose Acetate, Volume: 1000mL | Corning | Cat# 976140 |
| Filter Paper – 90 mm diameter | Whatman | Cat# 1001-090 |
| Falcon Bacteriological Petri Dishes with Lid; 100 x 15mm | Falcon | Cat# 08757100D |
| Super Glue Gel Control | Loctite | Cat# 234790 |
| Safety syringe, fixed needle, 1 mL 26G, 5/8" length | Thomas Scientific | Cat# 1145H04 (BX/1) |
| Gas Dispersion Tubes | Pyrex | Cat# 39533-12C |
| 96-Well PCR Plate Block for use with Compact Dry Blocks | Thermo Scientific | Cat# EW-36403-44 |
| Falcon 50 mL Conical Centrifuge Tubes | Falcon | Cat# 1495949A |
| Falcon 15 mL Conical Centrifuge Tubes | Falcon | Cat# 1495953A |
| 0.5 mL Microcentrifuge Tube | Corning Axygen | Cat# 14222292 |
| 2 mL Low-Retention Microcentrifuge Tube | Fisher | Cat# 02681321 |
| 12-Well Cell Culture Plate | Corning Costar | Cat# 0720082 |
| 40 Micron Cell Strainer | Fisher Scientific | Cat# 22363547 |
| Hemocytometer | INCYTO | Cat# DHCN015 |
| Mini Centrifuge | Eppendorf AG 5424 | Cat# 22620401 |
| Swing Bucket Centrifuge | Sorvall ST16 | Cat# 75004240 |
| VWR Rocking platform Model 200 | VWR | Cat# 40000-304 |
| VWR Digital Water Bath Model SL | VWR | Cat# 97025-130 |
| 2 mL Tubes Low-Retention | Fisher Scientific | Cat# 2681321 |
| LS Columns and 13 mL tubes set | Miltenyi | Cat# 130-122-729 |
| MidiMACS™ Separator | Miltenyi | Cat# 130-042-302 |
| MACS MultiStand | Miltenyi | Cat# 130-042-303 |
| Pre-Separation Filters | Miltenyi | Cat# 130-095-823 |
| MACS Acrylic Tube Rack | Miltenyi | Cat# 130-041-406 |
### MATERIALS AND EQUIPMENT

#### 10x ACSF stock (1 L)

| Item                                | Concentration | Amount   |
|-------------------------------------|---------------|----------|
| MilliQ H₂O (0.22 μm filtered deionized H₂O) |               | 800 mL   |
| NaCl                                | 1.2 M         | 70.1 g   |
| KCl                                 | 30 mM         | 3.73 g   |
| NaHCO₃                              | 260 mM        | 21.84 g  |
| NaH₂PO₄ (monobasic monohydrate)     | 12.5 mM       | 1.72 g   |

**Note:** Vacuum filter solution with 0.22 μm filter into a sterile 1 L container prior to use/for storage. (Corning Vacuum Filter in the KRT). Store at 4°C. Discard solution if it becomes cloudy.

#### 1 M MgSO₄

| Item | Amount |
|------|--------|
| MgSO₄ | 6.08 g |
| MilliQ H₂O | 50 mL |

**Note:** Store at room temperature (~20°C). Discard solution if it becomes cloudy.

#### 1 M CaCl₂

| Item | Amount |
|------|--------|
| CaCl₂ | 5.55 g |
| MilliQ H₂O | 50 mL |

**Note:** Or use 1 M stock solution listed in KRT—stored at room temperature. Discard solution if it becomes cloudy.

#### 1x ACSF (1 L)

| Item                                | Concentration | Amount   |
|-------------------------------------|---------------|----------|
| 10x ACSF Stock (above)              |               | 100 mL   |
| Glucose                             | 10 mM         | 1.8 g    |
| 1 M MgSO₄                           | 2 mM          | 2 mL     |
| 1 M CaCl₂                           | 2 mM          | 2 mL     |
| Pyruvic Acid                        | 0.1 mM        | 10 mg    |
| Ascorbic Acid                       | 0.4 mM        | 70 mg    |
| MilliQ H₂O                          | Dilute up to 1 L |

**Note:** Vacuum filter solution with 0.22 μm filter into a sterile 1 L container prior to use/for storage. (Corning Vacuum Filter in the KRT). Store at 4°C. Discard solution if it becomes cloudy. Do not store for more than 7 days.

#### NMDG buffer solution (1 L)

| Item                                | Concentration | Amount   |
|-------------------------------------|---------------|----------|
| NMDG                                | 110 mM        | 21.47 g  |
| 13 N HCl (stock concentration)     | 110 mM        | 8.46 mL  |
| KCl                                 | 10 mM         | 2.03 g   |
| NaH₂PO₄ (monobasic monohydrate)     | 1 mM          | 14.2 mg  |
| 1 M stock CaCl₂                     | 5 mM          | 0.5 mL   |

(Continued on next page)
**Vacuum filter solution** with 0.22 μm filter into a sterile 1 L container prior to use/for storage. (Corning Vacuum Filter in the KRT). Store at 4°C. Discard solution if it becomes cloudy. Do not store for more than 7 days.

**Density Gradient Solution** (All reagents are included in the Worthington Kit along with instructions on how to prepare required solutions)

| Item                                    | Concentration | Amount |
|-----------------------------------------|---------------|--------|
| EBSS                                    | 2.7 mL        |        |
| Ovomucoid Inhibitor                     | 1 mg/mL       | 300 μL |
| Dnase                                   | 100 units/mL  | 150 μL |
| 2 mM Kynurenic Acid (KYNA)              | 2 μM          | 3 μL   |
| Store at 4°C                            |               |        |

**Dissociation Buffer (DB)**

| Reagent                  | Concentration | Amount |
|--------------------------|---------------|--------|
| Na2SO4                   | 82 mM         | 5.83 g |
| K2SO4                    | 30 mM         | 2.615 g|
| Glucose                  | 10 mM         | 0.905 g|
| HEPES                    | 10 mM         | 1.2 g  |
| 1 M MgCl2                | 2.5 mM        | 2.5 mL |
| Nuclease-Free Water      | 500 mL        |        |
| Store at 4°C             |               |        |

**Extraction Buffer (ExB)**

| Reagent                  | Concentration | Amount |
|--------------------------|---------------|--------|
| DB                       | 15 mL         |        |
| Kollidon VA64            | 10 mg/mL      | 150 mg |
| Triton X-100 (1.7M)      | 1%            | 150 μL |
| 10% BSA                  | 0.01%         | 15 μL  |
| RNase inhibitor, Lucigen, 40,000 units/mL | 667 units/mL | 250 μL |
| Store at 4°C             |               |        |

**Note:** The other components can be mixed in advance but add the RNase inhibitor at the last moment.

**Wash Buffer (WB)**

| Reagent                  | Concentration | Amount |
|--------------------------|---------------|--------|
| DB                       | 30 mL         |        |
| 10% BSA                  | 0.01%         | 30 μL  |
| RNase inhibitor, Lucigen, 40,000 units/mL | 27 units/mL | 20 μL |
| Store at 4°C             |               |        |
**Note:** The other components can be mixed in advance but add the RNase inhibitor at the last moment.

| FACS Capture Buffer (CDB)                                      | Concentration | Amount |
|----------------------------------------------------------------|---------------|--------|
| DB                                                             |               | 20 mL  |
| 10% BSA                                                        | 0.01%         | 20 μL  |
| Store at 4°C.                                                  |               |        |

| FACS Buffer                                                   | Concentration | Amount |
|---------------------------------------------------------------|---------------|--------|
| 10% BSA stock solution                                        | 0.0025%       | 10 μL  |
| autoMACS Rinsing Solution                                     |               | 40 mL  |
| HEPES                                                         | 25 mM         | 0.3g   |
| RNase inhibitor, Lucigen, 40,000 units/mL                     | 35 units/mL   | 35 μL  |
| Store at 4°C.                                                 |               |        |

**Note:** The other components can be mixed in advance but add the RNase inhibitor at the last moment.

**STEP-BY-STEP METHOD DETAILS**

**Stereotaxic injections**

© Timing: 2–3 weeks

This step delivers the viral barcodes to label different projection neurons. It is important to ensure that the retrograde viruses used infect the intended cell types.

Surgery: 1–5 h/mouse for surgeries

Passive wait time: 2–3 weeks for virus to traffic and express in source location

1. Administer preoperative analgesics according to approved IACUC protocol.
2. Prepare:
   a. Viral aliquot(s) in styrofoam ice box.
   b. Microinjection system.
   c. Scalpel blade.
   d. Biohazard bag for biohazard disposal.
   e. Heating pad.
   f. Sterile drapes to place on stereotaxic frame and over mouse.
   g. Sterilize all surgical tools in the bead sterilizer for approximately 10 s.
      i. Place on sterile drape.
   h. Sterile gloves.
   i. Hair removal agent(s).
3. Place the mouse in the stereotaxic frame, remove the hair from the scalp, make an incision through the scalp, and level the head in the frame.
4. Inject each projection target with a retrograde virus encoding a different transgene (Figure 1).

**Note:** To reduce the likelihood of spillover and increase the efficiency of labeling, try to choose stereotaxic coordinates near the center of each target structure.

5. Close incision, recover mice from anesthesia, and return to home cage.
Tissue slice preparation

- **Timing:** 30 min–1 h

- **CRITICAL:** Make sure all surfaces and tools are free of fixative.

This step slices the brain to enable subsequent microdissection of the regions of interest.

**Pre-steps:**

6. Prepare artificial cerebrospinal fluid (ACSF) and N-Methyl-D-glucamine (NMDG) buffer solution before starting.
   a. The NMDG solution used with the vibratome should be put in a -20°C freezer and shaken every 10–15 min to prevent clumping. Freeze and shake until you achieve a slurry/slushy consistency. This step will take approximately 1 h.
   b. Warm approximately 200 mL of NMDG in a glass dish set inside a warm 32°C–34°C water bath. Place a cell strainer into the dish of NMDG.

**Note:** Before use, take care to bubble the warming NMDG solution with carbogen for at least 30 min by a gas diffusion tube connected to a carbogen source directly into the solution (Figure 2B).
7. Prepare the station in which you will sacrifice the mouse and retrieve the brain.
   a. Biohazard bag.
      i. Dissection tools: Large decapitation scissors, small spring scissors, forceps, spatula, and razor blade.
8. Deeply anesthetize the mouse to euthanize according to IACUC guidelines.
9. Transcardially perfuse mice with cold (4°C) 1x ACSF using gravity perfusion (approximately 5 min per mouse).
   a. It takes approximately 15 mL of 1x ACSF to perfuse one mouse.
10. Decapitate with large scissors.
    a. Quickly but carefully explant entire the brain. We used small spring scissors, forceps, and spatula for this step.
    b. Immediately place into NMDG slurry and wait 30 s.
11. Place a filter paper (e.g., a 90 mm Whatman filter disc) on a clean surface (e.g., a Petri dish).
    a. Remove the brain from the slurry and place directly onto the filter.
    b. Create a flat surface on the coronal plane by cutting the end of the brain that is not needed.
       We used a razor blade.
   
   **Note:** We cut off brainstem/cerebellum and adhered the brain directly onto the vibratome platform using Loctite glue with the olfactory bulb pointing up and our ROI oriented closer to the blade. Depending on the region of interest, the brain may need to be oriented differently.

   c. Orient the brain on the platform such that your region of interest is closest to the blade.
   i. We used the Paxinos mouse brain atlas as a guide to identify landmarks close to our ROI.
12. Cut 300 μm vibratome sections (Figure 2A).
    a. Use high speeds initially (e.g., 0.3 mm/s) to quickly trim away sections that do not contain areas that will be sequenced and slower speeds (e.g., 0.12 mm/s) when cutting sections containing the regions to be sequenced.
13. Recover slices in warm 32°C–34°C NMDG solution in the water bath. Continue bubbling with carbogen for 30–45 min.
   
   **Note:** Use cell strainers floating in the NMDG dish to suspend slices (Figure 2B).
a. While slices are recovering, prepare solutions for single-cell isolation or single-nuclei isolation.

14. Move slices into RT ACSF bubbled with carbogen. Slices can be kept alive up to 16 h.

15. Microdissect the region of interest under a dissection scope.

Single-cell isolation

© Timing: 1.5–2 h

This step microdissects brain regions of interest and isolates single cells for sequencing.

Note: Prepare Worthington Papain Dissociation Kit + EBSS without Phenol Red, Miltenyi Adult Neuron Isolation Kit, 1x ACSF.

TISSUE DISSOCIATION AND FILTRATION OF CELLS –modified Worthington protocol

16. Transfer one slice at a time to a Petri dish and dissect region of interest from vibratome sections and return to ACSF. Once all slices have been microdissected, place regions of interest in a 2 mL low-retention centrifuge tube.

17. Follow the Papain Dissociation System Protocol (https://worthington-biochem.com/PDS/) from Worthington Biochemical Corporation to make the different solutions. Some steps are summarized below.

a. Add 32 mL of Earle’s Balanced Salt Solution (EBSS) to the albumin-ovomucoid inhibitor mixture. Can be stored at 2°C–8°C after reconstitution. Discard if there is discoloration or cloudiness that could indicate microbial growth.

b. Add 5 mL of EBSS to the papain vial. Place papain in a 37°C water bath for 10 min until dissolved.

c. Add 500 µL EBSS to DNAse vial, mix gently to prevent shear denaturation.

   i. Add 250 µL of this solution to the vial containing papain.

   d. Place tissue in the papain solution.

18. Incubate at 37°C with agitation for 1–1.5 h triturating with a transfer pipette or 200 µL pipette after 20 min and every 10 min thereafter. Be very gentle.

19. Once the dissociation is complete, centrifuge at:

   a. 300–350 g for 5–10 min, depending on the tissue.

      i. Timing should be tested on different ROIs. We recommend spinning at 7 min as a start and adjusting the timing from there.

   b. Meanwhile, prepare the density gradient solution (step 9 of protocol summary for Papain Dissociation System).

20. Discard the supernatant and prepare the density gradient.
a. DENSITY GRADIENT
   i. Add 500 µL of the EBSS-reconstituted albumin-ovomucoid inhibitor mixture into a micro-
   cenrifuge tube.  

   **Note:** If the cell pellet after centrifugation is large, (example shown in Figure 3), fill more
   tubes since the goal is to maximize the number of cells passing through the density gradient
   solution. In most cases (e.g., isolating cells from a ~1 mm³ structure isolated from 3-5 mice)
   1–2 tubes should suffice.

   ii. Add 300 µL of density gradient solution to the cell pellet to resuspend and layer carefully on
   top of the 500 µL of EBSS-reconstituted albumin-ovomucoid to form a gradient. Do not mix.

   **Note:** if the cell resuspension appears viscous (noticeably more viscous than water, similar to
   maple syrup) or opaque after adding the initial 300 µL of the density gradient solution, or if
   the cell pellet was large (Figure 3), add another 300 µL of the density gradient solution and fill
   another microcentrifuge tube with 500 µL of the EBSS-reconstituted albumin-ovomucoid in-
   hibitor mixture. The result will be 2 density gradient tubes. Perform step iii.

   iii. Centrifuge at 100 g for 5–10 min to pellet the cells while leaving debris in the supernatant.
   For example, we used 6 min for ventral midbrain and 7 min for primary visual cortex (V1)
   and superior colliculus (SC).

   **Note:** Users should troubleshoot which spin time works best for their tissue type.

   iv. Aspirate and discard the supernatant and reconstitute the cell pellet in 1 mL of 1x ACSF.

21. Count the cells using a hemocytometer; using trypan blue to assess viability.

   **Note:** We only did this for cortex in the published paper. At this stage, we had ~10⁶ cells and
   had 1.4*10⁵ cells after filtering.

22. Magnetic Labeling and Separation with LS Column: Perform all steps on ice. Directions are
    adapted from the instructions in the kit.

   **Neuron Isolation Kit Data Sheet**
   **Midi MACS Data Sheet**
   **LS Columns Data Sheet**

   **Note** the cell count from Step 21.

   a. Centrifuge cell suspension at 300 g for 10 min. Discard the supernatant.

   i. During centrifugation, prepare d-MACS Buffer as follows: for each sample, make 10 mL of
   a 1:20 dilution of the MACS BSA Stock solution in Dulbecco’s Phosphate Buffered Saline
   (DPBS).

   ii. Prepare MACS Buffer as follows: for each sample, make 5 mL of a 1:20 dilution of the
   MACS BSA Stock solution in autoMACS Rinsing Solution.

   **Note:** Keep all solutions on ice. Always use freshly diluted buffer.

   b. Add 80 µL of d-MACS buffer per 10⁷ total cells to the cell pellet.

   c. Add 20 µL of the Adult Non-Neuronal Cell Biotin-Antibody cocktail per 10⁷ total cells.

   d. Mix well but **DO NOT VORTEX**. Incubate for 5 min on ice.

   e. Wash cells by adding 2 mL of MACS buffer per 10⁷ cells and centrifuge at 300 g for 10 min.
   Discard the supernatant.

   f. Add 80 µL of d-MACS buffer per 10⁷ cells to the cell pellet.

   g. Add 20 µL of Anti-Biotin Microbeads per 10⁷ total cells.

   h. Mix well by inverting the tube and **DO NOT VORTEX**. Incubate for 10 min on ice.
i. Adjust volume up to 500 μL for up to 10^7 cells using the buffer. For more cells, scale up buffer volume accordingly.

j. Proceed to magnetic separation.
   i. Place LS column in the magnetic field referred to in the LS Columns Data Sheet.
   ii. Wash column twice with 1 mL of MACS buffer. Discard the eluate.
   iii. Apply cell suspension to the column by passing through the pre-separation filters. Collect filtrate into the collection tube provided with the LS columns.
   iv. Wash column 2x with 1 mL of d-MACS. Collect filtrate into the collection tube provided with the LS columns.

**Note:** Do not use the plunger as this would remove the magnetically bound cells. Unbound cells are the neurons.

k. Transfer filtrate into a 2 mL microcentrifuge tube.

l. Pellet the cells at 300 g for 10 min using a microcentrifuge at 4°C and remove supernatant for single-cell isolation.

m. Resuspend cells in 1 mL of 1x ACSF and note the count using a hemocytometer.
   i. Concentrate or dilute cells as necessary for 10x library preparation.

---

**Single-nuclei isolation**

© Timing: 3–4 days

© Timing: 20–30 min to overnight for pre-steps

© Timing: 45 min–1 h active time for nuclei isolation

© Timing: 2.5–4 h for FACS

© Timing: 5–20 min for nuclei counting

© Timing: 1–2 days for library prep

This step isolates single nuclei from brain regions of interest for single-cell sequencing.

23. **PRE-CHILL FOR AT LEAST 20 min at –20°C or OVERNIGHT at 4°C:**

**Note:** “Precoated with 1% BSA” means that 1% BSA solution was used to coat the item and the excess fluid was removed. Refer to Martin et al. (2020) protocol for extra detail. 1% BSA was made by using 10% BSA stock diluted in the autoMACS Rinsing Solution.

| Item                                              | Quantity |
|---------------------------------------------------|----------|
| 26G syringe                                       | 1        |
| 96-well metal PCR plate                           | 1        |
| 50 mL Falcon tube precoated with 1% BSA           | 1        |
| 50 mL Falcon tube                                 | 1        |
| 15 mL Falcon tube precoated with 1% BSA           | 2        |
| 0.5 mL Eppendorf microcentrifuge tubes precoated with 1% BSA | 5        |
| 12-well cell culture plate coated with 1% BSA     | 1        |
| 40 μm cell strainer                               | 1        |
| 2 mL Eppendorf microcentrifuge tube               | 1        |
| Collection tubes with 20 μL of FACS Capture Buffer (CDB) each (tubes provided by FACS core at Gladstone) | 2        |
24. Solutions to Prepare (Night before or morning of dissociation)

1x ACSF

Chill in –20°C freezer to cool quickly (but be sure not to freeze).

△ CRITICAL: All solutions need to be chilled to 4°C and kept on ice prior to use. All steps are performed on ice or in a 4°C walk-in except for cell counting on a hemocytometer.

25. NUCLEI ISOLATION (~45 min- 1 h) based on Martin et al. (2020) protocol and Krishnaswami et al. (2016)

a. Add 1 mL ExB (Extraction Buffer) to tissue region of interest and deposit into a well of a 12-well plate.

b. Perform four trituration sets.
   i. A single trituration set is defined as 20 times of trituration followed by a 2-min wait period on ice. Triturate slowly with a 1 mL pipette tip.

△ CRITICAL: Do not create bubbles.

   ii. Repeat trituration cycle a 5th time only if the dissociation does not appear complete (e.g., large visible chunks).

c. Pass the entire volume twice through a 26 G needle into the same well.

d. Transfer into the pre-coated and pre-chilled 50 mL Falcon tube.
   i. Add 30 mL of WB (Wash Buffer).
   ii. Split volume evenly into the 2 1% BSA pre-coated, pre-chilled 15 mL Falcon tubes.
   iii. Spin down at 600 g for 10 min at 4°C using a swinging bucket rotor.
   iv. Remove supernatant until only ~500 µL remains in each tube. Combine the samples (pooled volume = 1 mL).

e. Pass the suspension through the pre-chilled 40 µm cell strainer and filter using only gravity. Do not plunge or apply pressure.

f. Transfer to a chilled 2 mL Eppendorf tube.

g. Count nuclei on a hemocytometer. Make sure the concentration is roughly 10⁶ nuclei/mL.
   i. To attain this concentration range, dilute or concentrate nuclei with 1x ACSF. Nuclei can be concentrated by following the gentle centrifugation protocol outlined in step 25j, removing the supernatant, and then adding desired volume of 1x ACSF.

h. Reserve some nuclei for FACS negative control before staining (minimum volume required is usually around 200 µL). Keep on ice.
   i. Stain remaining nuclei with rabbit anti-NeuN antibody conjugated to Alexa Fluor 488 (ab190195).

Note: Use 0.1–10 µg/mL anti-NeuN Ab. A dilution series should be tested to optimize concentration. We found that for our ROIs, 0.1 µg/mL was sufficient.

   i. Incubate in the dark at 4°C for 30–45 min on a gentle rocker or tube rotator.
   ii. Wash three times with 200 µL FACS buffer (total 600 µL) and gentle centrifugation (15–20 min total).

j. GENTLE CENTRIFUGATION PROTOCOL:
   i. Place a BSA-coated, pre-chilled 0.5 mL tube in a 2.0 mL tube inside a centrifuge.
   ii. Add nuclei to the 0.5 mL tube. Use a maximum volume of 200 µL.
   iii. Centrifuge at 200 g for 1 min at 4°C.
   iv. After reaching a complete stop, centrifuge again at 200 g for 1 min at 4°C.
v. Aspirate supernatant. Resuspend nuclei in FACS buffer at \(10^7\)/mL, add DAPI at 1 mg/mL or 1:1000 dilution.

k. Proceed to FACS.

i. Bring the stained sample and the negative control (unstained) sample from step h.

ii. Fill collection tubes with 20 \(\mu\)L of FACS capture buffer (CDB) in each tube.

**Note:** Number of tubes are dependent on number of cells to be counted.

26. **FACS** (~1–4 h depending on sample size)

a. Use DAPI and anti-NeuN labeling to isolate neuronal nuclei (Figure 4). The NeuN labeling distribution is bimodal and we gate to include the peak with higher labeling.

b. To concentrate nuclei, perform GENTLE CENTRIFUGATION PROTOCOL described in step 25j. Centrifuge at 200 g for 1 min at 4°C.

27. **COUNT NUCLEI AFTER FACS** (~5 min–20 min).

a. Count nuclei on hemocytometer.

i. Prepare a 1:10 dilution of part of the sample (18 \(\mu\)L of chilled DB (Dissociation Buffer) with 2 \(\mu\)L of nuclei from the PCR tube).

ii. Load into hemocytometer.

iii. Examine fluorescent channels and the DAPI channel.

iv. Dilute nuclei with 1x ACSF or concentrate nuclei via gentle centrifugation protocol in a 4°C walk-in to 17,000 nuclei per 40 \(\mu\)L for the 10x Genomics library prep protocol. Proceed as soon as possible to 10x library preparation. Note: We have not waited longer than 45 min post-FACS (including counting).
28. **LIBRARY PREP (1–2 days)**
   a. Ensure that the reagents are shipped ahead of time.
   b. Prepare libraries using the 10x Genomics Chromium single cell 5' mRNA workflow following standard protocol (https://support.10xgenomics.com/single-cell-gene-expression/sample-prep/doc/demonstrated-protocol-isolation-of-nuclei-for-single-cell-rna-sequencing-and-tissues-for-single-cell-rna-sequencing) and sequence libraries.
      i. Depending on the desired sequencing depth, we used either NovaSeq (2 \( \times 10^{10} \) reads) or NextSeq (1.3–2.6 \( \times 10^8 \) reads). We aimed for 20–50,000 reads/cell and used standard paired-end 150 bp read lengths.

---

29. **Genome alignment**

   © **Timing:** Varies from 15 min–1 day depending on dataset size and hardware

   This section describes methods used to align single-cell sequencing data with the reference genome and the addition of viral transgenes to the reference genome.

   Raw gene sequence data outputs are converted to cell count tables using Cell Ranger provided by 10x Genomics.

30. **Hardware:** We recommend a computer with at least 8–16 CPU cores, 128 GB RAM, and 1–2 TB of disk storage depending on data size.

31. **Operating System:**
   a. Install a Debian-based Linux distribution such as Ubuntu or Centos. For our publication, we used Ubuntu on an Amazon Web Service (AWS) Elastic Compute Cloud (EC2) instance in our examples.
   b. Using the package manager on the Linux distribution, install tools needed to configure the python environment, including wget, sh, bash.

---

31. **Cell Ranger:**
   a. Visit the 10x Genomics download links here to download software and reference genomes.
   b. Download Cell Ranger (for our publication, we used version 6.0.0) software. We used the 10x Genomics mm10 mouse reference genome.
   c. Follow instructions for building a custom reference genome here. Append custom reference information (such as viral transgenes) to the provided .gtf file within the reference genome. Then filter the .gtf file.
      i. create custom .gtf file
      ii. vim into the .gtf file and add in the transgene information. Example below.

```
touch transgene.gtf

transgene unknown exon 1 1443 . + . gene_id "transgene"; transcript_id "transgene"; gene_name "Transgene"
```

iii. Append into the custom transgene.gtf file to the end of the provided .gtf file.

```
cat transgene.gtf >> genes.gtf
```

iv. filter the .gtf file.
d. Append transgene.fa to the provided .fa file
   i. create custom .fa file

```sh
touch transgene.fa
```

   ii. vim into the .fa file and add in the transgene information. Example (for tdTomato sequence) below.

```sh
>transgene 1443
CGTGTTACCCCGACATGTTAAGGGCGTAGAAATGCTATGGGATTGAAGGATGGACAGTGCTACAGGGTCTACATGGGTGACGCTCATGAT
CTGAGTCGACTGAGATAGTGGAACTGGATACAGCAGTGGGTTGAGGGATGAGTGAGTTAGGATGCTATTGCAGAACAGTGCTTAGCT
AGTGATGAGGAGACACAGACAGAGAACTGGATAGGGTATTGAGGTGGATAGGATGCTTAGCTAGCTAGCTAGCTAGCTAGCTAGCTA
GCCGTGCTGGGTTGAGGGTATTGAGGTGGATAGGATGCTTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT
CTGAGTCGACTGAGATAGTGGAACTGGATACAGCAGTGGGTTGAGGGATGAGTGAGTTAGGATGCTATTGCAGAACAGTGCTTAGCT
AGTGATGAGGAGACACAGACAGAGAACTGGATAGGGTATTGAGGTGGATAGGATGCTTAGCTAGCTAGCTAGCTAGCTAGCTAGCTA
```

   iii. Append the custom transgene.fa file to the end of the provided .fa file.

```sh
cat transgene.fa >> genes.fa
```

e. Make the custom reference genome using the newly made genes.fa and genes.gtf files. Example code below:

```sh
cellranger mkref \
   -genome=mml0_custom_ref \
   -fasta=path_to_fa_file/genes.fa \
   -genes=path_to_gtf_file/genes.gtf
```

f. Run cellranger count to generate counts for each gene in a cell count table that will be saved in a .h5 file. Read more about this function here. We ran our parameters as follows:

```sh
cellranger count \
   -id=sample_id \
   -fastqs=path_to_fastq
```
g. If combining samples, refer to instructions here. Example code below:

```bash
touch samples_to_aggregate.csv
vim samples_to_aggregate.csv
```

vim into the .csv file and add the sample information.

```bash
sample_id,molecule_h5
sample_01,sample_01_molecule_info_directory/molecule_info.h5
sample_02,sample_02_molecule_info_directory/molecule_info.h5
sample_03,sample_03_molecule_info_directory/molecule_info.h5
cellranger aggr --id=custom_name --csv=samples_to_aggregate.csv
```

h. Resulting count matrix will end in .h5 file extension.
   i. example: filtered_feature_matrix.h5

**Data analysis**

**Timing: 1–3 d**

This section describes the procedures for clustering neurons, identifying markers, and overlaying viral barcode expression on clusters to identify projection neurons.

Single-cell sequencing analysis is commonly performed using ScanPy (Python package) or Seurat (R package). The below protocol describes steps using Python and ScanPy that were used in Cheung et al., (2021) (code). For clarity, our protocol references function names in ScanPy version 1.7.2, but please note that function names may change per authors and maintainers of the ScanPy library in future versions. Analogous steps can be performed using R and Seurat.

32. Hardware: We recommend a computer with at least 16 CPU cores, 64 GB RAM, and 1 TB of disk storage. The speed of some algorithms may benefit from additional CPU cores and it may be beneficial to use an on-demand cloud computing service to dynamically scale computing requirements since some (but not all) algorithms benefit from multi-core CPU acceleration.

33. Operating System:
   a. Install a Debian-based Linux distribution such as Ubuntu or Centos. We use Ubuntu in our examples.
   b. Using the package manager on the Linux distribution, install tools needed to configure the python environment, including wget, sh, bash.

34. Software:
   a. Download a python package manager such as pip or conda. We primarily used conda for package management and installations. Examples using conda below:

```bash
wget https://repo.continuum.io/miniconda/Miniconda3-latest-Linux-x86_64.sh -O ~/miniconda.sh
bash ~/miniconda.sh -b -p $HOME/miniconda
```
b. Create a conda environment `<env_name>` and then add bioconda and conda-forge channels. Then install packages. The specific packages shown below are used for the VECTORseq project. Specific versions of each package used can be found in the VECTORseq project code under the requirements.txt file.

```
conda config --add channels bioconda
conda config --add channels conda-forge
conda create --name <env_name> jupyter ipykernel nb_conda anndata==0.7.5 scanpy==1.7.2 leidenalg pysam pynndescent pandas==1.2.3 numpy scipy matplotlibLib tqdm black flake8 scikit-learn pyarrow fastparquet snappy seaborn
```

35. Analysis: The following data pipeline is used for data cleaning and generating optimal clusters with biological plausibility for both inhibitory and excitatory neuronal datasets described in Cheung et al., (2021).

a. Reformat: Convert data to AnnData format.
   i. Load the cell count table in the .h5 and convert it to the AnnData data format used by ScanPy using the function `scanpy.read_10x_h5`. This generates an in-memory `adata` object upon which downstream ScanPy functions can be applied.
   ii. In the VECTORseq project code, this step corresponds to the “reformat” pipeline stage.

b. Distribution Plots: Assess sequencing data quality.
   i. Compute quality control metrics using the function `scanpy.pp.calculate_qc_metrics`.
   ii. Generate scatter plots and violin plots for number of genes, number of counts, and percent mitochondrial genes. These plots can reveal any contamination of single cell data with dead cells, doublets, multiplets, etc.
   iii. The percent mitochondrial genes cutoff we used was 5% for brain tissue.
   iv. General rule of thumb to plot the distribution across counts/cell and genes/cell.

Note: For counts/cell, if there is a bimodal distribution, the initial filtration criteria is usually made on the value after the end of the first peak and at the beginning of the second one.

v. In the VECTORseq project code, this step corresponds to the “distribution_plots” pipeline stage.

c. Filter: Apply filtering cutoffs to remove dead cells, doublets, multiplets, and any unwanted artifacts created as a consequence of the single cell sequencing process that do not reflect biological phenomena.
   i. Apply count cutoffs using the function `scanpy.pp.filter_cells`.
   ii. You may also choose to selectively filter mRNA expression from chromosomal, ribosomal, and mitochondrial genes by manually subsetting `adata.X` and excluding a blacklist of known genes.
   iii. In the VECTORseq project code, this step corresponds to the “filter” pipeline stage.

d. Normalize: Normalize counts and convert viral transgenes from expression data to a meta-data label so that further analysis with Scanpy functions only operates upon endogenous genes.
   i. Log-normalize total counts (number of times gene detected) per cell using the functions `scanpy.pp.normalize_total` and `scanpy.pp.log1p`. This baseline count normalization is necessary for comparisons between cells. Apply any additional transforms as desired such as log transform, term frequency-inverse document frequency (TFIDF), depending on your application. TFIDF was used in Cheung et al., (2021) to highlight unique genes within each cell, including the unique viral transgenes that had been previously introduced.
   ii. Importantly, viral transgene expression should be removed from the cell count table and converted into a label since downstream steps are intended to be applied only to
endogenous gene expression. In ScanPy and AnnData, this can be accomplished by moving the viral transgenes and their corresponding counts from adata.X to adata.var which stores metadata corresponding to each cell. All further ScanPy methods will be applied to adata.X.

iii. Once adata.X contains only endogenous genes, several techniques are employed to reduce the number of genes to include in further analysis. TFIDF and log-normalization are repeated on only endogenous genes. Highly variable genes are computed using the function scanpy.pp.highly_variable_genes, both before and after TFIDF. Some genes may not be highly variable within a dataset but are useful for determination of cluster identity. To keep them in the dataset, we a priori whitelisted specific genes of interest. The final genes to be included in analysis are the set of genes in the whitelist combined with the union of highly variable genes before and after the TFIDF transform.

iv. Scale each gene to unit variance and clip outlier values using the function scanpy.pp.scale.

v. In the VECTORseq project code, these steps correspond to the “normalize” pipeline stage.

e. Cluster: Compute neighborhood graph and Leiden clusters.

i. Compute top \(n\) principal components of adata.X using scanpy.pp.pca, and then use scanpy.pp.neighbors to compute a neighborhood graph of observations using the top principal components. We selected an \(n\) of 50.

ii. Use the Leiden algorithm (Traag et al., 2019) with scanpy.tl.leiden to generate clusters based on the neighborhood graph. Various results can be obtained depending on how many \(n\) principal components are selected, the \(n\_neighbors\) used to generate the neighborhood graph, and the leiden\_resolution used for Leiden clustering. Using \(n = 50\), we ran a grid search for \(n\_neighbors\) and leiden\_resolution. The value of \(n\_neighbors\) is swept from 5 to 105 with increments of 5. The value of leiden\_resolution is swept from 0.05 to 1.25 with increments of 0.05.

iii. In the VECTORseq project code, this step corresponds to the “cluster” pipeline stage.

f. Cluster Metrics:

i. Compute various internal cluster validation metrics for each combination of \(n\_neighbors\) and leiden\_resolution used in the grid search. Internal cluster validation metrics include within-cluster sum of squared error, within-cluster variance, Davies-Bouldin Index, Average Silhouette Score, Calinski-Harabasz Index (Pseudo-F statistic), and Xie-Beni Index (Calinski and Harabasz, 1974; Davies and Bouldin, 1979; Rousseeuw, 1987; Singh et al., 2017).

ii. Visualize internal cluster validation metrics on heatmaps and 3D surface plots. Local minima and maxima on these plots correspond to more optimal hyperparameters.

Note: In general better separation of clusters is achieved by minimizing within-cluster sum of squared error, within-cluster variance, Davies-Bouldin Index, Xie-Beni Index, and by maximizing Calinski-Harabasz Index and Average Silhouette Score.

iii. Multiple local minima/maxima may exist for each of the internal cluster validation metrics. These local optima suggest good candidate options for hyperparameters \(n\_neighbors\) and leiden\_resolution.

Note: Final selection of these hyperparameters ultimately requires inspection of cluster gene expressions and correlation with biological plausibility.

iv. In the VECTORseq project code, this step corresponds to the “cluster\_metrics” pipeline stage.

g. Create UMAP: Generate Uniform Manifold Approximation and Projection (UMAP) visualizations.

i. Use UMAP to remap the high-dimensional gene expression manifold to a 2D representation with scanpy.tl.umap, then use scanpy.pl.umap to create a scatter plot. Color code
cluster identities generated from Leiden clustering onto the UMAP plot. A unique UMAP visualization is generated for every combination of n_neighbors and leiden_resolution used in the grid search.

ii. Using a combination of internal cluster validation metrics and inspection of UMAP visualization along with the biological plausibility of known genes from the whitelist of genes included during the Normalize step, optimal clustering schemes with specific n_neighbors and leiden_resolution can be selected for further downstream analysis.

**Note:** Selection of specific n_neighbors and leiden_resolution will vary between datasets. Cross-examining several different combinations via exploratory data analysis will be the most useful in determining which parameters to choose.

iii. In the VECTORseq project code, this step corresponds to the “create_umap” pipeline stage.

**h. Expression Plots: Generate gene expression plots.**

i. After an optimal combination of n_neighbors and leiden_resolution is chosen, use the cluster identities to generate a dendrogram plot using scanpy.tl.dendrogram and scanpy.pl.dendrogram.

ii. Relative gene expressions can be visualized with heatmaps using scanpy.pl.heatmap, matrix plots using scanpy.pl.matrixplot, and dot plots using scanpy.pl.dotplot. When calling these functions, group by the leiden cluster identities and include the computed dendrogram to help visualize cluster relationships.

iii. Most highly expressed genes within each cluster group can be selectively visualized. Genes within each cluster can be ranked using scanpy.pl.rank_genes_groups, then visualized as a heatmap using scanpy.pl.rank_genes_groups_heatmap, matrix plots using scanpy.pl.rank_genes_groups_matrixplot, and dot plots using scanpy.pl.rank_genes_groups_dotplot. When calling these functions, group by the Leiden cluster identities and include the computed dendrogram to help visualize relationships. You will also need to specify how many top ranked genes in each cluster you desire to visualize on the plots.

iv. At this stage, expression of viral transgenes can be overlaid to identify the clusters corresponding to different projection populations.

**Note:** These annotations were stored in adata.var during our normalization step. Viral overlays can be performed ScanPy’s built-in plotting functions by calling the column name used when storing the annotation in adata.var.

v. In the VECTORseq project code, this step corresponds to the “expression_plots” pipeline stage.

**i. Subset:** Subset data based on selected clusters to further explore expression data.

i. This is an optional step which can be recursively performed to target specific subset of cells based on cluster membership. For example, in Cheung et al., (2021), we extracted only excitatory neuron clusters based on gene expression profile, then performed further clustering and expression analysis on that subset by repeating the above steps Cluster, Cluster Metrics, Create UMAP, and Expression Plots. The same was done for inhibitory neuron clusters.

ii. In the VECTORseq project code, this step corresponds to the “subset” pipeline stage.

**EXPECTED OUTCOMES**

The expected outcome of this protocol is transcriptional delineation of the neuronal cell types within a structure with different projection types transcriptionally marked via transgene expression (Figure 5). In a successful experiment, viral transgenes will be selectively enriched in one or a few cell
types that project to the target structure; ubiquitous expression, if observed, is likely due to ambient mRNA contamination (see limitations).

LIMITATIONS

The largest concern in any VECTORseq experiment is viral tropism. It is essential to confirm that any virus and batch to be used in sequencing experiments infects the projection population of interest. Likewise, mixing a locally infecting virus encoding a fluorophore with the retrograde viruses allows validation of the site of infection in individual animals during tissue slicing, which is also useful for identifying mice within a cohort with injection failures or mistargeting that should be excluded from the pooled sequencing reaction. In addition, we found that spillover from viral expression within a projection target near the source structure yielded abundant ambient mRNA in all the sequenced populations. Therefore, when investigating projections to structures near the source structure, it may be necessary to ensure that dissections exclude the injection site or to use a retrograde virus such as HSV, which is reported to express stably in retrogradely infected neurons but only transiently in neurons at the injection site (Fenno et al., 2014). Similarly, when different cells within a source structure project to neighboring targets, it is important to ensure that viral injections into the target structure are focal, e.g., by limiting injection volume, in order to avoid spillover. Another concern is that stereotaxic targeting may be impaired in certain models of pathological conditions. Therefore, it is critical to validate the anatomy of the specific strain, age, and model under study. In addition, adjustments to the code and troubleshooting may be necessary for other experiments and brain areas.

A potential concern is that viruses may alter endogenous physiology and gene expression of infected cells. To this end, it is helpful to use viruses known to have minimal effects on cellular physiology such as AAV and HSV. In addition, one can take advantage of the fact that not every cell in a given population will be infected and compare gene expression between infected and uninfected cells within a given cluster; any changes in gene expression in infected cells, particularly induction of inflammatory or other immune genes, may indicate cellular perturbations by viral infection. Note that in our initial study, these analyses found that HSV and AAV did not alter endogenous gene expression in any of the cell types we analyzed, but similar analyses are worthwhile for each dataset. As an additional control for the effects of viral expression, one could sequence an uninfected cohort to compare the abundance of different cell clusters and gene expression within clusters.

Finally, although NeuN labeling can helpfully enrich for most neuronal nuclei, it is important to realize that not every neuronal cell type expresses NeuN and thus this enrichment approach could lose certain cell types.

TROUBLESHOOTING

Problem 1
Lack of viral infection or spillover from injection site (stereotaxic injections, step 4).

Potential solution
A combination of test retrograde and anterograde infections followed by histology to ensure that the viruses infect the cell types of interest and the lack of spillover from the injection site.
Problem 2
Differences in neuroanatomy across strains, ages, and/or disease models.

Potential solution
Confirm stereotaxic targeting of the structures of interest in the relevant model (stereotaxic injections, step 4).

Problem 3
Possibility that viral infections cause cytotoxicity and/or altered gene expression.

Potential solution
Perform within-cluster comparisons of gene expression in virally labeled and unlabeled cells or even, when available, to compare clusters with those obtained from sequencing of an uninfected cohort (data analysis, step 35hiv).

Problem 4
The most costly and time-intensive steps are preparing, sequencing, and analyzing the 10x library.

Potential solution
We recommend practicing the cell/nuclei isolation steps on the tissue of interest in uninfected animals before attempting to generate libraries and sequencing a VECTORseq cohort (single-cell isolation, single-nuclei isolation).

Problem 5
Potential problems with library preparation.

Potential solution
It is good practice to sequence each library shallowly and inexpensively at first to identify any issues with the library preparation before doing a deeper and more expensive sequencing run for final analyses (Library preparation).

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Evan Feinberg (evan.feinberg@ucsf.edu).

Materials availability
This study did not generate new unique reagents. Mice and all reagents used in this study are commercially available as indicated in the key resources table.

Data and code availability
● The raw and processed single-cell sequencing data are deposited at https://ucsf.app.box.com/v/vectorseq-data and can be found on NCBI GEO (GSE189907).
● All code for analysis is publicly available at https://github.com/vic-cheung/vectorseq and https://doi.org/10.5281/zenodo.5703724.
● Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS
This work was supported by departmental funds and grants from the E. M. Ziegler Foundation for the Blind, Sandler Foundation, Klingenstein-Simons Fellowship Award in Neuroscience, Brain and Behavior Research Foundation (NARSAD Young Investigator Awards 25337 and 27320), Whitehall Foundation, Simons Foundation (SFARI 574347), and US National Institutes of Health (DP2.
MH119426 and R01 NS109060) to E.H.F. We thank C. Cheung for help with software development and code refactoring. We thank J. McGuire and M. Bernardi at the Gladstone genomics core for 10X library preparation and sequencing and staff at the UCSF Center for Advanced Technology (CAT) for support with RNA sequencing. Flow cytometry was performed by the Gladstone flow cytometry core, which is supported by NIH S10 RR028962, James B. Pendleton Charitable Trust, and NIH P30 AI027763.

AUTHOR CONTRIBUTIONS
V.C.: Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing—original draft, Writing—Review & Editing, Visualization, Project Administration; P.C.: Software, Validation, Formal analysis, Writing—Review & Editing, Visualization; E.H.F.: Conceptualization, Methodology, Resources, Writing—original draft, Writing—Review & Editing, Visualization, Project Administration, Funding Acquisition, Supervision

DECLARATION OF INTERESTS
The authors declare no competing interests.

REFERENCES
Caliński, T., and Harabasz, J. (1974). A dendrite method for cluster analysis. Comm. Stats. - Theor. Methods 3, 1–27.

Cheung, V., Chung, P., Bjorni, M., Shvareva, V.A., Lopez, Y.C., and Feinberg, E.H. (2021). Virally Encoded Connectivity Transgenic Overlay RNA sequencing (VECTORseq) defines projection neurons involved in sensorimotor integration. Cell Rep. 37, 110131.

Davies, D.L., and Bouldin, D.W. (1979). A cluster separation measure. IEEE Trans. Pattern Anal. Mach. Intell. 1, 224–227.

Krishnaswami, S.R., Grindberg, R.V., Novotny, M., Venepally, P., Lacar, B., Bhutani, K., Linker, S.B., Pham, S., Erwin, J.A., Miller, J.A., et al. (2016). Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. Nat. Protoc. 11, 499–524.

Martin, C., Abdul, A., Vanderburg, C., Nadaf, N., Feirera, A., and Macosko, E. Frozen Tissue Nuclei Extraction (for 10xV3 snSEQ). protocols.io. https://dx.doi.org/10.17504/protocols.io.bck6iuze.

Rousseeuw, P.J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Maths. 20, 53–65. ISSN 0377-0427.

Singh, M., Bhattacharjee, R., Sharma, N., and Verma, A. (2017). An improved xie-beni index for cluster validity measure. In 2017 Fourth International Conference on Image Information Processing (ICIIP), pp. 1–5.

Traag, V.A., Waltman, L., and van Eck, N.J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. 91, 5233–5312.

Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 19, 15.

Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: computational identification of cell doublets in single-cell transcriptomic data. Cell Syst. 8, 281–291.e9.

Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun. 8, 14049–14112.