nature protocols

PROTOCOL

https://doi.org/10.1038/s41596-022-00752-0

Check for updates

Optimized single-nucleus transcriptional profiling by combinatorial indexing

Beth K. Martin[1]✉, Chengxiang Qiu[1], Eva Nichols[1], Melissa Phung[1,2], Rula Green-Gladden[1,3], Sanjay Srivatsan[1,4], Ronnie Blecher-Gonen[5], Brian J. Beliveau[1], Cole Trapnell[1,6,7], Junyue Cao[8] and Jay Shendure[1,6,7,9]✉

Single-cell combinatorial indexing RNA sequencing (sci-RNA-seq) is a powerful method for recovering gene expression data from an exponentially scalable number of individual cells or nuclei. However, sci-RNA-seq is a complex protocol that has historically exhibited variable performance on different tissues, as well as lower sensitivity than alternative methods. Here, we report a simplified, optimized version of the sci-RNA-seq protocol with three rounds of split-pool indexing that is faster, more robust and more sensitive and has a higher yield than the original protocol, with reagent costs on the order of 1 cent per cell or less. The total hands-on time from nuclei isolation to final library preparation takes 2–3 d, depending on the number of samples sharing the experiment. The improvements also allow RNA profiling from tissues rich in RNases like older mouse embryos or adult tissues that were problematic for the original method. We showcase the optimized protocol via whole-organism analysis of an E16.5 mouse embryo, profiling ~380,000 nuclei in a single experiment. Finally, we introduce a 'Tiny-Sci' protocol for experiments in which input material is very limited.

Introduction

Single-cell combinatorial indexing (sci-*) combines in situ molecular indexing and a 'split-pool' framework to uniquely label an exponentially scalable number of cells or nuclei with a unique combination of nucleic acid barcodes. After its demonstration in the context of chromatin accessibility[1,2] in 2015, we and others additionally developed sci-* methods for profiling gene expression[3–7], genome sequence[8,9], genome architecture[10], genome-wide methylation[11], co-assays of mRNA and chromatin accessibility[12], transcriptional dynamics[13], transcription factor occupancy[14,15], surface proteins[14], small molecule exposures[16] and spatial locations[17], all at single-cell resolution.

A split-pool method for profiling gene expression, sci-RNA-seq, was first reported in 2017[3], and an improved three-level version, sci-RNA-seq3[3,4], in 2019. Among other applications, sci-RNA-seq3 has been applied to generate the largest atlases of single-cell gene expression for both mice[3] (~2 million cells) and humans[18] (~4 million cells) to date. Both of these datasets were largely generated within a single laboratory, each within a few weeks and by one or two individuals. Nonetheless, the underlying protocol remains cumbersome. Here, we describe the culmination of extensive efforts to simplify and optimize the protocol, enabling the transcriptional profiling of previously difficult tissues. In addition, we describe a modified version of the protocol (Box 1), called 'Tiny-Sci', for experiments in which input material is very limited. For example, we recently applied Tiny-Sci to a series of individual embryonic day (E) 8.5 (E8.5) mouse embryos to generate a sci-RNA-seq3 dataset with somite-level resolution[19]. These embryos were only a few millimeters in length, such that Tiny-Sci extends the range of sci-RNA-seq3 to samples for which only small amounts of tissue are available.

Briefly, the sci-RNA-seq3 protocol (Fig. 1) starts by allocating fixed cells or nuclei to the wells of one or more 96-well plates. The first index is introduced during reverse transcription with barcoded oligo-dT primers. Cells or nuclei are then pooled and split to a new set of one or more 96-well plates. The second index is ligated onto the end of the first index, and then the cells or nuclei are pooled and split again. In the third set of plates, second-strand synthesis occurs, and the double-stranded product is then tagmented

[1]Department of Genome Sciences, University of Washington, Seattle, WA, USA. [2]Department of Biology, Case Western Reserve University, Cleveland, OH, USA. [3]Division of Hematology/Oncology, Seattle Children's Hospital, Seattle, WA, USA. [4]Medical Scientist Training Program, University of Washington, Seattle, WA, USA. [5]The Crown Genomics Institute of the Nancy and Stephen Grand Israel National Center for Personalized Medicine, Weizmann Institute of Science, Rehovot, Israel. [6]Brotman Baty Institute for Precision Medicine, Seattle, WA, USA. [7]Allen Discovery Center for Cell Lineage Tracing, Seattle, WA, USA. [8]Laboratory of Single-Cell Genomics and Population Dynamics, The Rockefeller University, New York, NY, USA. [9]Howard Hughes Medical Institute, Seattle, WA, USA. ✉e-mail: martin91@uw.edu; shendure@uw.edu

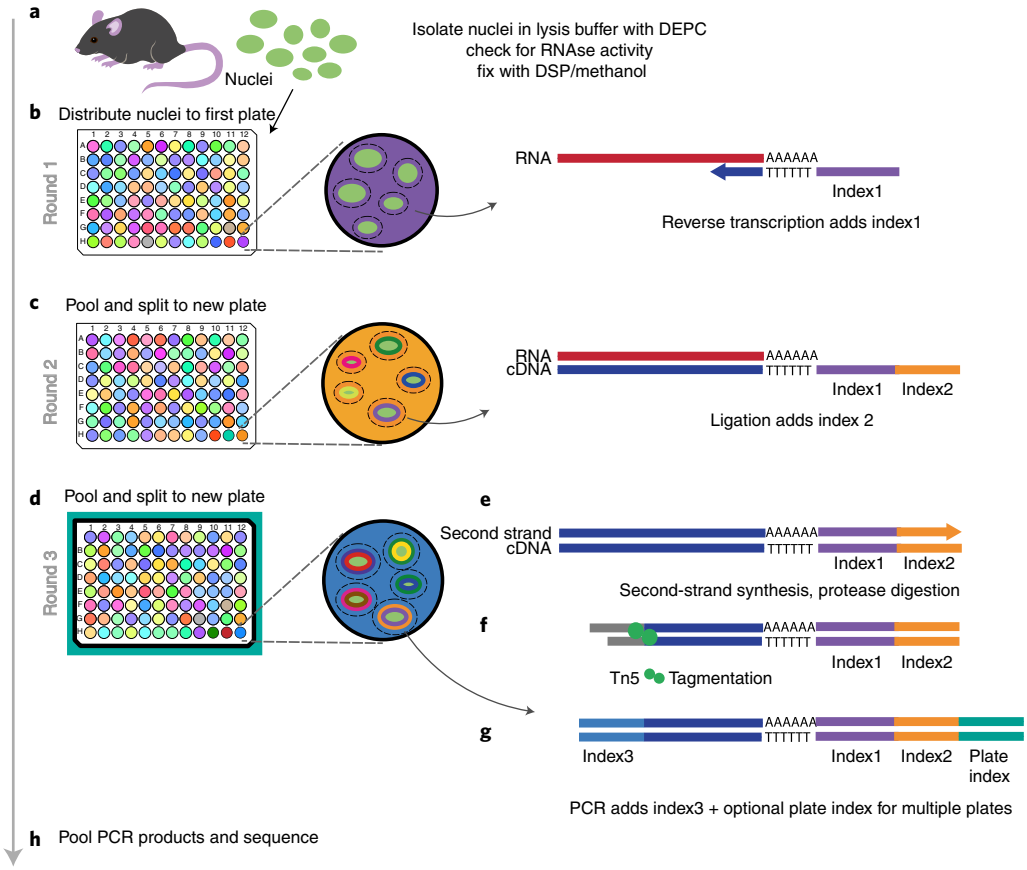

Fig. 1 | Summary of the optimized sci-RNA-seq3 protocol. Colors in plate wells represent 96 unique index sequences for each of the three rounds of indexing. RNA transcript in red, DNA in dark blue. **a**, Nuclei (in green) are isolated in lysis buffer with diethyl pyrocarbonate (DEPC) and then fixed with dithiobis(succinimidyl propionate) (DSP) and methanol. **b**, Nuclei are then distributed to a 96-well plate for reverse transcription, where the first index (in the purple well) is introduced. If desired, nuclei from different samples can be deposited to different wells during this first round of indexing, facilitating multi-sample processing while minimizing batch effects. **c**, After indexed reverse transcription, the nuclei are pooled and split into a new plate to add a second index (in the orange well) via ligation. **d**, The nuclei are pooled and split again. After second-strand synthesis (**e**), protease digestion and tagmentation (**f**), the third index is added by PCR (in the blue well), with an optional plate index (aqua) (**g**). Finally, the library is purified and sequenced (**h**).

with Tn5 transposase. PCR amplification adds the third index, and, finally, the library is purified and sequenced. A detailed schematic of primers and adaptor sequences is shown in Fig. 2.

Development of the protocol

Here, we describe the optimized sci-RNA-seq3 protocol, primarily focusing on nuclei (although to apply the protocol to cells, one need only omit the lysis step; see Limitations below). To expand the range of tissues that can be processed with sci-RNA-seq3, a first set of changes was directed at better neutralizing endogenous RNases found in older embryonic and adult tissues. A secondary consequence of these changes has been an increase in the number of unique molecular identifiers (UMIs) obtained per nucleus. Specifically, diethyl pyrocarbonate (DEPC) is now used to inactivate RNases during the lysis step. DEPC has commonly been used to deactivate RNases in water and solutions used for RNA work by modifying the RNase protein, but it has been unclear whether it negatively affects RNA itself[20]. We compared the effectiveness of DEPC versus commercial RNase inhibitors on RNase activity in the lysate of an E13.5 mouse embryo. The differences were stark (Supplementary Figs. 1 and 2). Using an RNase Alert kit (IDT), we quickly realized why older embryos often failed the sci-RNA-seq3 protocol. The expensive SUPERase•In RNase Inhibitor that was used in the original sci-RNA-seq3 protocol was ineffective at inactivating the RNases in older tissues, and the change to DEPC alone has had the biggest impact on sci-RNA-seq3 success. Moreover, the resulting data indicate that DEPC is not detrimental to the transcriptional readout, at least not for sci-RNA-seq3. Of

Fig. 2 | Detailed schematic of the sci-RNA-seq3 combinatorial indexing strategy. a, The indexed primer sets used. '#' indicates the index sequence, and 'N' are random bases incorporated into the primer for the unique molecular identifier (UMI). There are 96–384 indexed reverse transcription (RT) primers, 96–384 indexed ligation primers, 96–384 indexed PCR P7 primers and indexed PCR P5 primers. The P5 primer can be used as a plate index for multiple plates or, optionally, for a full set of 96–384 for fully dual-indexed PCRs. **b**, In the first round of indexing, the oligodT of the RT primer binds to the polyA tail of mRNA, extending to make the complementary cDNA strand. **c**, After pooling and redistribution of the nuclei, the ligation primer, which can form a hairpin, anneals to the 6-bp linker on the RT primer, allowing ligation to the phosphorylated 5′ end of the RT primer. **d**, Nuclei are pooled and redistributed into the third and final plate, and during second-strand synthesis, the RNA is nicked and used as primers to create the second strand with DNA Polymerase I. **e**, Tn5 tagmentation fragments the now double-stranded DNA and adds adaptors for the PCR primers. **f**, PCR adds the final indexes and sequencing adapters. **g**, The final product can be sequenced on Illumina platforms with their standard primers. Lig, ligation; rc, reverse complement.

note, DEPC can phase-separate at the concentration used here, but the detergent in the lysis buffer seems to help keep it emulsified and effective.

The RNase Alert checkpoint step is now included to ensure that no RNase activity still exists before proceeding with the bulk of the protocol, preventing wasted time and reagents. The lysis buffer had to be changed from a Tris-based buffer to a hypotonic phosphate buffer to accommodate DEPC. Another new buffer, 0.3 M SPBSTM, replaces the original nuclei suspension buffer and enables better nuclei recovery during washes and spins. It contains sucrose for osmolarity and cushioning of the nuclei during spins, PBS to keep the buffer isotonic and Triton-X 100 to make the nuclei pellet more cleanly, and magnesium chloride is necessary for nuclei integrity.

Dithiobis (succinimidyl propionate) (DSP)/methanol fixation replaces the need for a separate permeabilization step and results in more UMIs per cell compared to paraformaldehyde fixation in the original protocol. DSP is an amine-crosslinker that does not alter RNA but unfortunately precipitates in aqueous solutions. Methanol both fixes and permeablizes nuclei well for access to the transcripts but often leaves nuclei too fragile to tolerate the complete protocol. When methanol is used on its own, this results in large, unusable clumps of hundreds of nuclei at the end of the protocol. However, the two fixatives work well together—the DSP is easily dissolved in the methanol and confers integrity to the fixed cells. The resulting nuclei are stable, accessible and less prone to clumping compared with the paraformaldehyde fixation in the original protocol.

We have eliminated the USER (uracil-specific excision reagent) enzyme step from the original protocol, and deoxyU is no longer needed in the ligation primer. We realized that the second-strand synthesis step can open up the hairpin, and thus the USER step is not necessary.

Table 1 | Cost comparison between original[4] and optimized sci-RNA-seq3 protocols

Enzyme	Cost/µl ($)	Original amount/plate (µl)	Original cost/plate ($)	Optimized amount/plate (µl)	Optimized cost/plate ($)
Superscript IV	7.11	205	1,458	55	391
RNaseOUT	1.56	205	320	Not used	0
SuperaseIN	1.00	250	250	Not used	0
Quick Ligase	2.59	215	557	Not used	0
T4 DNA Ligase	1.04	Not used	0	65	68
USER	1.21	110	133	Not used	0
Second-strand synthesis	2.95	73.3	216	35	103
AmpureXP	0.014	3,840	53	230	3
Tn5	24.50	0.92	22	4.6	112
NEBNext 2× PCR mix	0.058	1,920	111	2,200	127
Total enzyme cost per plate (96 × 96 × 96)	-	-	3,120	-	804
Total enzyme cost per four plates (384 × 384 × 384)	-	-	12,480	-	3,207

We focus on the most expensive reagents here (all enzymes) because other reagent costs are comparatively negligible.

After the nuclei are in their third and final plate and have undergone second-strand synthesis, we have found that extracting the nuclei before tagmentation allows better access for the Tn5. However, we have simplified this extraction with a simple digestion with a heat-inactivatable protease, thereby removing any need for a whole-plate AMPure/SPRI (solid-phase reversible immobilization) bead cleanup, significantly reducing the hands-on time and cost for the final library preparation steps.

These changes result in a streamlined protocol that allows transcripts to be recovered from RNase-rich tissues that were previously problematic for sci-RNA-seq3, while also mitigating nuclei losses for precious samples, because more nuclei are able to tolerate the entire process. In addition, we have drastically reduced the cost of conducting these experiments as compared with the original protocol (Table 1), which was already considerably less expensive than commercial alternatives.

Application of the method

Other single-cell techniques may benefit from the nuclei preparation and fixation method presented here, particularly with older embryos and tissues, which are rendered challenging by the abundant presence of RNases. Using DEPC as an RNase inhibitor has been very successful in our hands, but it has a short half-life in water and is broken down by Tris-containing buffers. It is needed only until the nuclei are fixed, so it might be useful in other methods to replace more expensive inhibitors if you can substitute buffers until the RNase inactivation is complete. Additional modifications to streamline or even remove the tagmentation step may further improve the protocol, and the method can also potentially be combined with oligo-based hashing techniques[16,17]. Using DSP as a fixative might prohibit some single-cell techniques because of the ability of dithiothreitol (commonly found in reaction buffers) to reverse the cross-links. For such protocols, one might try using disuccinimidyl suberate instead, which acts similarly to DSP but with cross-links that are irreversible.

Comparison to other single-cell methods

Key contrasts between sci-RNA-seq3 and widely used commercial solutions such as the Single Cell Gene Expression kit from 10x Genomics include that sci-RNA-seq3 incurs a considerably lower cost per cell or nucleus when used at scale, does not require a kit, is open source and is exponentially scalable. The optimized sci-RNA-seq3 protocol presented here is written here as the basic 'one-plate version', with 96 reverse transcription indexes, 96 ligation indexes and 96 PCR indexes. In our hands, a 96 × 96 × 96 experiment typically nets ~100,000 nuclei. However, there are usually enough nuclei to fill an additional three plates or more for the final round of indexing, such that it is straightforward to boost the number of nuclei profiled to ~400,000 (i.e., 96 × 96 × 384). Further scaling to a four-plate experiment (384 indexes at each level) enables increasing the number of nuclei to one or several million or more per experiment. As with all sci-* protocols, cells or nuclei from different samples can

be deposited to different wells during the first round of indexing, facilitating multi-sample processing while minimizing batch effects[3].

Limitations

In our hands, 'organism-scale' transcriptional profiling is much more straightforward and less biased when performed on nuclei, rather than cells, because whereas each tissue might require a different protocol to dissociate cells, cell lysis is a 'universal' means of freeing nuclei. Of course, a drawback is that one loses cytoplasmic transcripts. However, we note that it is straightforward to apply the protocol presented here to whole cells, simply by omitting the nuclear lysis step. We have done this successfully with mammalian cell lines, achieving similar performance (B.K.M., Sam Regalado and Jean-Benoit Lalanne, unpublished observations), although we note that we have not fully characterized how the subsequent buffers affect cell integrity.

Overview of the protocol

The three main sections of the protocol can be split over multiple days. If there is only one sample, you can combine days 1 and 2 into a single day.

1 *Day 1*. Nuclei preparation and fixing. Fixed samples can be stored at −80 °C.
2 *Day 2*. Reverse transcription, ligation and second-strand synthesis (overnight).
3 *Day 3*. Protease digestion, tagmentation, PCR, final library cleanup and loading on the sequencer.

Experimental design

Scaling up

This protocol is written for a small experiment with one plate of primers at each indexing round (which typically yields ~100,000 single-cell profiles). It can be scaled up easily to use as many as four plates of primers at each round (which would be expected to yield over ~1 million single-cell profiles)—just multiply each step by 4, combine all four plates when pooling and put more nuclei (4,000 per well) into the final plates.

Samples

Samples can be cell lines, tissues or whole embryos. Ideally, samples (e.g., embryos or other samples) are isolated and frozen individually in 1.5-ml microcentrifuge tubes with very little extra fluid with them. For a 96 × 96 × 96 sci-RNA-seq3 experiment, you will need 2 million fixed cells to distribute into the first plate. If you are using a cell line, count out ≥4–6 million cells to start with and wash cells in PBS before lysis. For tissues, a good size is 0.5–0.75 cm^3 or ~200 mg. Bigger chunks of tissue will necessitate more lysis buffer. For instance, an E13.5 mouse embryo, which is ≤1 cm long, can be lysed in 5 ml of buffer with one filter, but an E18.5 mouse at ~3 cm long will need to be split into four tubes of 20 ml of lysis buffer each and four filters, because there is so much more debris involved. Alternatively, you can powder a large sample, keeping it frozen, and save the extra powder for lysing later if needed. If you do not have enough cells for 2 million in the first plate, share with another sample to fill out the plate for 2 million nuclei total. The reverse transcription (RT) indexes in each well will allow you to sort the different samples later in the analysis.

If you are working with extremely small samples a few millimeters in size (such as E8.5 mouse embryos), see the 'Tiny-Sci' section (Box 1) for alternative nuclei isolation and fixing steps. This is the same sci- protocol, but with the lysis and fixation scaled significantly down and with fewer transfers that may contribute to nuclei loss. We have used this method for E8.5 single embryos, which are ~2–3 mm long, but it can more generally be used for instances in which starting material is very limited.

Counting nuclei

Nuclei prepared from frozen tissue are not simple to count, because debris can interfere with interpretation. Staining with Yoyo-1 dye helps to discern nuclei from debris and is visualized on the GFP channel. Mix 10 µl of diluted Yoyo-1 (see Reagent setup) with 10 µl of nuclei (twofold dilution) or do a 10–20-fold dilution if the nuclei are concentrated.

For the example experiment described in the Procedure below, nuclei were counted manually on a Countess Cell counter, because the automatic cell counting was unreliable for nuclei (Fig. 3). On the GFP channel, with the view zoomed all the way out, Yoyo-1–stained nuclei were counted in a 6 cm × 6 cm square. This count was multiplied by the dilution factor ×10,000 to get an approximate number of nuclei per milliliter.

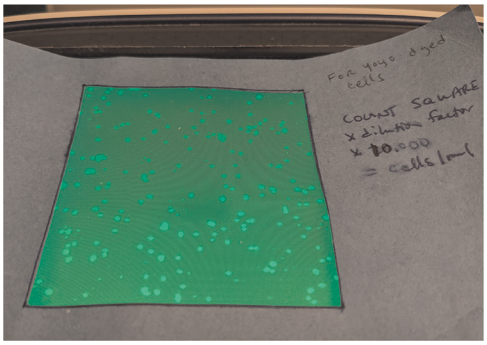

Fig. 3 | Yoyo-1 stained nuclei from an E16.5 mouse embryo visualized on a Countess Cell Counter. Nuclei are counted by hand in a 6-cm × 6-cm square. The method is perhaps inelegant but in our hands, fast and remarkably consistent.

Cell lysis

Tissue is dissociated by simply smashing it on dry ice (Fig. 4 and Supplementary Video 1). At this point, RNA is going to be especially vulnerable to RNases that are released by the cells, so a sufficient volume of DEPC-containing lysis buffer is necessary to inactivate them. This is the step that will make or break your experiment, so be sure to do an RNaseAlert check before proceeding to fixation.

Lysis buffer selection

There are two lysis buffers to choose from, depending on the tissues involved. Older mouse embryos perform better with buffer A, which is sucrose based and lacks BSA. Most other tissues, like younger embryos, isolated adult tissues (including mouse and human tissues) and cell lines, should be lysed in buffer B with BSA. We are not exactly sure why the older mouse embryos benefit from a lack of BSA; it could be that because they have so much protein of their own already, adding more somehow limits DEPC's effectiveness.

Reverse transcription

This step will take you longer than you think because you will need to get the right amount of nuclei loaded into the first plate, especially if you are including multiple samples. Allow plenty of time for counting and resuspending samples at the correct concentrations. For a single sample, you will need 2 million nuclei to fill out a plate.

Sequencing

For a one-plate experiment with 100,000 nuclei at the end, it will be sufficient to run the final library on a NextSeq550 (400 million reads) or NextSeq 2000 P2 (800 million reads). For large experiments, like the four-plate version with 1.6 million cells at the end, we need the read capacity (16–20 billion reads) from a NovaSeq 6000 S4 to capture most of the unique transcripts in each cell.

Data analysis

The pipeline for generating a single-cell digital expression matrix is posted on Github (https://github.com/JunyueC/sci-RNA-seq3_pipeline). The processing is slightly different from the pipeline used for analyzing data that are generated by other technologies (e.g., 10x Genomics): base calls are first converted to fastq format followed by demultiplexing based on PCR i5 and i7 barcodes. Reads are filtered for legitimate RT and ligation indexes, and these indexes are added to the read information. After mapping to the reference genome, reads are split into constituent cellular indices by further demultiplexing reads by using the RT index and ligation index, and deduplicated by the UMI. The pipeline is based on implementing the *qsub* command on a Linux system (e.g., Centos 7), and we recommend running it on a high-performance computing cluster. Of note, it will create some intermediate profiles that are space consuming (around three to four times bigger than the original data).

We tested the pipeline on sample data (created from a mouse embryo at E16.5) with ~7 billion reads. To speed up the processing, we split the whole data set into seven batches after demultiplexing by PCR barcodes and then run the pipeline on each batch in parallel. Each batch used 10 cores (Intel Xeon Gold 6238) with 20 GB of RAM for each core.

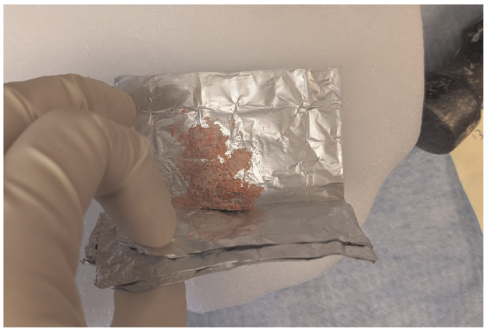

Fig. 4 | Smashing tissue in a foil packet on a slab of dry ice with a hammer. The foil packet with sample must stay on the dry ice until the powdered tissue is added to the lysis buffer.

Materials

Biological materials

- *E16.5 or E18.5 C57BL/6J embryos*. C57BL/6J (wild-type (WT)) mice (*Mus musculus*) were maintained in a specific pathogen-free environment in the Animal Research and Care Facility at the University of Washington. Mice were kept on a 12-h light/dark cycle with ad libitum access to food and water. Mice were originally bought from the Jackson Laboratory and bred in our facility. Male WT mice ≥6 weeks of age were bred to female WT mice ≥7 weeks of age; the first day of plug positive was marked as E0.5. On specific gestational days (E16.5 or E18.5), dams were euthanized and embryos were collected. Embryos were rinsed in 1× PBS buffer, dabbed dry and immediately flash-frozen in liquid nitrogen. Male or female embryos went into the sci-RNA-seq workflows **! CAUTION** Experiments using animals must conform to national and institutional regulations. All procedures were approved by the University of Washington Institutional Animal Care and Use Committee.
- *E8.5 embryos*. E8.5 mouse embryos were harvested as recently described[19] and dissected free of extraembryonic membranes. Embryos were transferred from the dissecting dish within a minimal volume of PBS (~5 µl), placed on the wall of individual cryovials and snap-frozen in liquid nitrogen. Samples were stored at −80 °C until further processing.

Reagents

- DSP (Lomant's reagent; Thermo Fisher, cat. no. 22586 or PG82081) **▲ CRITICAL** DSP is sensitive to water and should be used immediately after dissolving in dimethyl sulfoxide (DMSO).
- Methanol (Millipore Sigma, cat. no. 494437-2L)
- DMSO (Millipore Sigma, cat. no. D2438-5X10ML) **▲ CRITICAL** DMSO used for dissolving DSP should be new and unopened so that water is not introduced. These smaller bottles are useful for this reason.
- Sodium phosphate dibasic (Millipore Sigma, cat. no. S3264-250G)
- Sodium phosphate monobasic monohydrate (Millipore Sigma, cat. no. 71507-250G)
- Potassium phosphate monobasic (Millipore Sigma, cat. no. P9791-100G)
- Sodium chloride (Millipore Sigma, cat. no. S3014-500G)
- Potassium chloride (Millipore Sigma, cat. no. P9541-500G)
- Magnesium chloride solution, 2 M (Millipore Sigma, cat. no. 68475-100ML-F)
- Igepal CA-630 (Millipore Sigma, cat. no. I8896-50ML)
- BSA, 20 mg/ml (New England Biolabs, cat. no. B9000S)
- DEPC (Millipore Sigma, cat. no. D5758-25ML) **! CAUTION** Handle DEPC, and samples containing it, in a fume hood.
- Sucrose (VWR, cat. no. 97061-428)
- Triton X-100 (Millipore Sigma, cat. no. T8787-100ML)
- Tween 20 (Thermo Fisher, cat. no. BP-337-100)
- 10× Dulbecco's PBS (10× DPBS; Thermo Fisher, cat. no. 14200075)
- Superscript IV reverse transcriptase (Thermo Fisher, cat. no. 18090200) **▲ CRITICAL** This protocol has not been tested with other reverse transcriptases.
- T4 DNA Ligase (New England Biolabs, cat. no. M0202L) **▲ CRITICAL** The previous version of this protocol used Quick Ligase, but the buffer that is included with that enzyme interferes with the pelleting of the nuclei during centrifugation.

- Tagmentase (Tn5 transposase), unloaded (Diagenode, cat. no. C01070010-20) ▲ **CRITICAL** The amount of tagmentase added to Step 57 has been determined for this brand only.
- Tn5-N7 oligo (5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3′, high-purity salt-free; Eurofins)
- Mosaic End (ME) oligo (5′-/5Phos/CTGTCTCTTATACACATCT-3′, high-purity salt-free; Eurofins)
- NEBNext mRNA second-strand synthesis module (New England Biolabs, cat. no. E6111L)
- NEBNext high fidelity 2× PCR master mix (New England Biolabs, cat. no. M0541L)
- dNTP mix (New England Biolabs, cat. no. N0447L)
- Agencourt AMPure XP (Beckman Coulter, cat. no. A63882)
- Yoyo dye (Thermo Fisher, cat. no. Y3601)
- RNaseAlert kit (IDT, cat. no. 11-02-01-02)
- RNaseZap (Thermo Fisher, cat. no. AM9780)
- Elution buffer (EB, 10 mM Tris pH 8.5; Qiagen, cat. no. 19086)
- Protease (Qiagen, cat. no. 19157) ▲ **CRITICAL** Do not use any other protease/proteinase. This one can be heat-inactivated at the temperature and time listed in the protocol.
- 6% TBE Novex PAGE gels (10 wells; Thermo Fisher, cat. no. EC6262BOX)
- UltraPure agarose (Thermo Fisher, cat. no. 16500-500)
- Qubit double-stranded DNA high-sensitivity quantitation kit (Thermo, cat. no. Q32851)
- sci-RNA-seq3 indexed primer plates at 10 μM dilution (standard desalting for purification; random bases do not need hand-mixing). The complete list of primers is found in Supplementary Table 1. The plates include the following: plate(s) of indexed oligo-dT RT primers (5′- /5Phos/CAGAGCNNNNNNNN[10bpRTindex]TTTTTTTTTTTTTTTTTTTTTTTTTTTTTT-3′, where 'N' is any base; IDT), plate(s) of indexed ligation primers (5′- GCTCTG[9- or 10-bp ligation index]TACGACGCTCTTCCGATCT[reverse complement of ligation index]-3′), plate of indexed PCR P7 primers (5′-CAAGCAGAAGACGGCATACGAGAT[PCR P7 index]GTCTCGTGGGCTCGG-3′; IDT); and PCR P5 primers (this primer does not need to be indexed if you do only one plate of PCR) (5′-AATGATACGGCGACCACCGAGATCTACAC[PCR P5 index]ACACTCTTTCCCTACACGACGCTCTTCCGATCT-3′; IDT).

Equipment

- Hammer
- DNA/RNA LoBind tubes (Eppendorf, cat. no. 022431021)
- Refrigerated centrifuges that hold 1.5-ml microcentrifuge tubes, microwell plates and 15- and 50-ml conical tubes
- Chemical fume hood
- Multichannel pipettes and tips
- FloMi filter, 40 μm (VWR, cat. no. 10032-802)
- Falcon cell strainer, 40 μm (VWR, cat. no. 21008-949)
- Pestle for cell strainer (Midsci, cat. no. SG-PEST)
- 96-well plates (Eppendorf, cat. no. 951020401)
- 96-well LoBind plates (Eppendorf, cat. no. 30129512) or FrameStar 96-well skirted PCR plates (Thomas Scientific, cat. no. 1149V59) if LoBind plates are not available
- Thermomixer
- Sonicator (Diagenode Bioruptor Plus)
- Cell counter with GFP channel, or a hemocytometer and an inverted microscope that allows visualization with GFP
- Electrophoresis chambers for PAGE and agarose gels

Reagent setup

10× PBS-hypotonic stock solution

Mix 5.45 g of Na_2HPO_4 (dibasic), 3.1 g of $NaH_2PO_4{\cdot}H_20$, 1.2 g of KH_2PO_4, 1 g of KCl and 3 g of NaCl in nuclease-free water and bring to a final volume of 500 ml. This stock solution will have a pH of ~6.8, but when diluted to 1×, should end up at pH 7.0–7.4. The buffer can be stored at room temperature (20–23 °C) for 6 months.

Hypotonic lysis buffer solution A

This buffer is used for whole mouse embryos E16.5 and older. Mix 5 ml of the 10× PBS-hypotonic stock solution, 5.7 g of sucrose, 75 μl of 2 M $MgCl_2$ and nuclease-free water to a final volume of 50 ml

to make the lysis base solution. Right before lysis, for every 1 ml of lysis buffer needed, add 2.5 µl of 10% (vol/vol) igepal (vol/vol) and 10 µl of DEPC and then vortex the solution to disperse the DEPC throughout. For example, if a sample needs 5 ml of lysis buffer, take a 5-ml aliquot of lysis buffer stock solution and add 12.5 µl of 10% (vol/vol) igepal and 50 µl of DEPC. Keep the buffer on ice. Make fresh for each experiment. **! CAUTION** DEPC needs to be used in a fume hood. **▲ CRITICAL** DEPC has a short half-life in aqueous solutions, so adding it to the buffer just before the cells are added is important.

Hypotonic lysis buffer solution B

This buffer is used for Tiny-Sci, cell lines, mouse embryos under E16.5 and isolated tissues. Mix 5 ml of the 10× PBS-hypotonic stock solution, 75 µl of 2 M $MgCl_2$ and nuclease-free water to a final volume of 50 ml to make the lysis base solution. Right before lysis, for every 1 ml of lysis buffer needed, add 40 µl of BSA (20 mg/ml), 2.5 µl of 10% (vol/vol) igepal and 10 µl of DEPC and then vortex the solution to disperse the DEPC throughout. For example, if a sample needs 5 ml of lysis buffer, take a 5-ml aliquot of lysis buffer base solution and add 200 µl of BSA, 12.5 µl of 10% (vol/vol) igepal and 50 µl of DEPC. Keep the buffer on ice. Make fresh for each experiment. **! CAUTION** DEPC needs to be used in a fume hood. **▲ CRITICAL** DEPC has a short half-life in aqueous solutions, so adding it to the buffer just before the cells are added is important.

0.3 M SPBSTM (sucrose PBS TritonX $MgCl_2$)

This is the main buffer used throughout the protocol for washing and diluting nuclei. Dissolve 28.5 g of sucrose in 25 ml of 10× DPBS (regular DPBS, not the hypotonic version) and 125 ml of nuclease-free water (about half the volume of water that you will need). Once the sucrose has dissolved, add 2.5 ml of 10% (vol/vol) Triton X-100, 375 µl of 2 M $MgCl_2$ and more water to the final volume of 250 ml. Store this buffer at 4 °C for ≤3 months.

DSP 50 mg/ml stock

Dissolve a 50-mg vial of DSP in 1 ml of anhydrous DMSO (use a new vial of DMSO), because DSP will precipitate in aqueous solutions. Dissolved DSP should be used immediately.

Yoyo-1 dye for counting

Dilute 1 µl of Yoyo-1 dye in 1 ml of 0.3 M SPBSTM in a dark or amber microcentrifuge tube and store the reagent at 4 °C for ≤3 months. This will be used to dilute nuclei for counting.

Annealed N7 oligos

The annealed N7 oligos are Tn5-N7 (5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3′) and ME (5′-[phos]CTGTCTCTTATACACATCT-3′). Resuspend both oligos to 100 µM in annealing buffer (50 mM NaCl, 40 mM Tris-HCl pH 8.0). Mix one volume of Tn5-N7 with one volume of ME. This creates a working stock at 50 µM. Anneal them with the following PCR program: 95 °C for 5 min, cool to 65 °C (0.1 °C/s), 65 °C for 5 min, cool to 4 °C (0.1 °C/s). Store annealed oligos at 4 °C for 6 months or divide them into aliquots and freeze them at −20 °C for ≤1 year.

N7-loaded Tn5

To 20 µl of Tn5, add 20 µl of annealed N7 oligos. Place in a thermomixer and shake at 350 rpm and 23 °C for 30 min. Add 20 µl of glycerol. Store at −20 °C for ≤6 months.

Tagment DNA buffer (2×)

To 38.75 ml of nuclease-free water, add 1 ml of 1 M Tris pH 7.6, 250 µl of 2 M $MgCl_2$ and 10 ml of dimethylformamide. The final volume is 50 ml. Make 550-µl aliquots and store them at −20 °C for ≤6 months.

Indexed primer plates

Primers for reverse transcription, ligation and PCR indexing steps are ordered at 100 µM. Working dilutions are made to 10 µM in EB (10 mM Tris-Cl pH 8.5) and kept at 4 °C for ≤6 months. **▲ CRITICAL** Spin the primer plates down before opening, but do not spin the plates with nuclei until the second-strand synthesis stage (Step 50).

10% (vol/vol) igepal

Dilute 5 ml of igepal in 45 ml of nuclease-free water. Store at room temperature for ≤6 months.

10% (vol/vol) Triton X-100
Dilute 5 ml of Triton X-100 in 45 ml of nuclease-free water. Store at room temperature for ≤6 months.

10% (vol/vol) Tween 20
Dilute 5 ml of Tween 20 in 45 ml of nuclease-free water. Store at room temperature for ≤6 months.

Protease
Add 7 ml of water to a bottle of lyophilized Qiagen protease (Qiagen, cat. no. 19157). Make 200-µl aliquots and store them at −20 °C for ≤6 months. Do not freeze–thaw. ▲ **CRITICAL** This must be the specified protease, not proteinase K.

Procedure

▲ **CRITICAL** Everything is to be kept cold at all times. Have lots of ice ready and pre-cool centrifuges to 4 °C. Pre-cool all tubes on ice before you put nuclei in them.
▲ **CRITICAL** Clean the workspace and fume hood area ahead of time; wipe pipettes, racks and centrifuges down with RNAseZap and change gloves often.

Nuclei isolation ● Timing 2 h

▲ **CRITICAL** LoBind (DNA/RNA) microcentrifuge tubes are preferred for all steps in which they are used for collecting nuclei.

1 Prepare two ice buckets with wet ice, a bucket with crushed dry ice to hold your frozen tissues and a thick, flat slab of dry ice for smashing tissues. Pre-cool both a centrifuge that will hold 50-ml tubes and a microcentrifuge to 4 °C.

2 Determine how much lysis buffer you will need for the tissue you will be processing. (For extremely small samples, skip to Box 1: Tiny-Sci.) An E13.5 mouse embryo (~200 mg) requires 5 ml of lysis buffer. An E16.5 embryo (~500 mg) will need 20 ml. An adult mouse heart needs 5 ml. Adult mouse kidneys need 5 ml per kidney. Adult mouse liver needs 20 ml. Adult mouse pancreas needs 15 ml. Adult tissues and tissues high in RNases will necessitate a bigger lysis volume. The buffer is inexpensive to make, so do not worry about using too much.
! CAUTION DEPC is flammable and toxic. Avoid breathing vapors. The following steps should be performed in the chemical hood from this point until the DEPC is washed from the sample (Step 17).

Box 1 | Tiny-Sci ● Timing 1 h

Procedure

1 Samples should be no more than 2–3 mm in size, should be frozen individually in a microcentrifuge tube and kept on dry ice until lysis. The key to these small samples is to limit pipetting and tube transfers that will leave behind precious nuclei.
2 You will need 100 µl of lysis buffer B per embryo, but make up enough for several samples.
3 Right before lysis, for every 1 ml of lysis buffer needed, add 2.5 µl of 10% (vol/vol) igepal, 40 µl of BSA (20 mg/ml) and 10 µl of DEPC and then vortex the solution to disperse the DEPC throughout.
! CAUTION You must work in the hood because the DEPC is toxic.
4 Add 100 µl of complete lysis buffer (with BSA/DEPC/igepal) to the tube with the frozen embryo and make sure that the embryo is actually in the buffer. Let it sit a couple of minutes on ice and then triturate the embryo slowly and gently with a pipette set to 50 µl with a yellow (200 µl) tip. You should not see any chunks left. Lysis time is only ~5 min.
5 (Optional) You can take 1 µl of this and mix it with 9 µl of diluted Yoyo dye to quickly check under a microscope with a GFP filter to make sure that you are seeing intact, separate nuclei.
6 Mix fixative: 400 µl of ice-cold methanol + 10 µl of DSP stock solution. Add 400 µl to the embryo on ice dropwise over 1–2 min. Do not pipette up and down; instead, flick gently to mix and repeat occasionally over 5–10 min. You may see some clumping now.
7 Add 1 ml of 0.3 M SPBSTM dropwise, slowly and mixing gently by flicking. Do not pipette up and down.
8 Spin at 500*g* for 3 min at 4 °C. You should see a very tiny pellet. Remove all but ~50 µl of the supernatant and discard, without disturbing the pellet.
9 Add 500 µl of 0.3 M SPBSTM. Set a 1-ml pipette to 100 µl and resuspend by pipetting gently.
10 If you got clumps from fixing, sonicate the tube for 12 s on low at 4 °C.
11 Spin again at 500*g* for 3 min at 4 °C. Carefully remove the supernatant and resuspend the nuclei pellet gently in SPBSTM so that the volume is 42.5 µl. Add 2.5 µl of 10 mM dNTPs and put 5 µl into each well of one column on a plate, noting into which wells the nuclei went.
12 Fill up the plate with more embryos or some other nuclei of which you have a lot and continue with the sci procedure at Step 24 of the reverse transcription.

3 For every 1 ml of lysis buffer needed, add 2.5 µl of 10% (vol/vol) igepal and 10 µl of DEPC to the hypotonic lysis buffer solution (and 40 µl of BSA if using lysis buffer B) and then vortex the solution to disperse the DEPC throughout. Have complete lysis buffer in a 50-ml tube for each sample on ice ready to go.

4 Fold a piece of aluminum foil four times so that you have a small pouch with eight layers of foil on each side. Place this on a slab of dry ice to chill.

5 Place your frozen tissue inside this foil and hold it firmly closed on the dry ice and smash it with a hammer (Fig. 4 and Supplementary Video 1). You want to be gentle enough not to tear the foil, but thorough enough to make a powder of the tissue. Do not let the tissue thaw.

6 Use the foil to guide your powdered tissue into the tube of the lysis buffer. It will stick a bit; pipette some of the lysis buffer from the tube to rinse the sample from the foil into the tube. Try to make sure that the sample is thawing only if it is in lysis buffer.

7 Cap the 50-ml tube and shake it to disperse the chunks in the buffer. Let it sit on ice for 10 min. Triturate the chunks with a 1-ml pipette tip to help tease them apart a bit.

8 Set up another 50-ml tube on ice with a 40-µm cell strainer on top. Pour your lysate through that; there will still be a lot of chunks. Use a disposable pestle to coax the tissue through the filter. Do not worry about getting all of it through.

9 Take a 45-µl sample of the filtered lysate and check for RNase activity with the IDT RnaseAlert kit. The RNaseAlert test will guide you on whether to proceed or not. There should not be any RNase detected, and if there is, you will have to restart with a new sample and adjust either the sample size or the volume of lysis buffer so that there is enough DEPC to inactivate the RNases. You cannot continue with a sample that has RNases detected at this point; the damage is already done.

? TROUBLESHOOTING

10 While the RNaseAlert sample is incubating, spin down the remainder of the lysate (500*g*, 3 min, 4 °C). Keep the nuclei in the 50-ml tube. Resuspend the nuclei in 1 ml of 0.3 M SPBSTM with 10 µl of DEPC added (or more buffer if there are a lot of nuclei—roughly 1 ml of buffer per 200–500 mg of starting material at minimum).

? TROUBLESHOOTING

Nuclei fixation ● Timing 1 hour

11 For each sample, prepare fixative. In a 5-ml tube, add 100 µl of 50 mg/ml DSP stock solution to 4 ml of ice-cold methanol for every 1 ml of nuclei with which you are starting.

12 Add the prepared fixative gradually to the 50-ml tube containing the nuclei, swirling as you are adding it.

13 Fix on ice for 15 min, swirling occasionally.

14 Add 2 volumes of 0.3 M SPBSTM gradually, 2–3 ml at a time, swirling between additions to rehydrate the nuclei. For instance, with 1 ml of nuclei and 4 ml of fixative, you would need 10 ml of buffer to rehydrate.

15 Spin down the nuclei at 500*g* for 3 min at 4 °C.

16 Carefully remove the supernatant and dispose properly. The nuclei pellet is at the bottom and should look a little white-ish from the DSP.

17 Resuspend the nuclei in ≥1 ml of 0.3 M SPBSTM. Triturate gently with a pipette tip to separate nuclei.

18 Divide fixed nuclei into aliquots of ~2 million nuclei (Fig. 5) in microcentrifuge tubes. Spin at 500*g* for 3 min at 4 °C and remove and discard the supernatant. Continue with the protocol or snap-freeze tubes in liquid nitrogen and store at −80 °C.

■ PAUSE POINT Fixed samples may be stored at −80 °C for ≤6 months.

19 (Optional) If continuing with the protocol the same day and there are obvious clumps at this point that will not tease apart, you will need to sonicate them to break them up. Sonicate (Diagenode Bioruptor Plus) on low intensity for 12 s only at 4 °C. Spin and resuspend the nuclei in 500 µl of 0.3 M SPBSTM as in Step 21.

? TROUBLESHOOTING

Reverse transcription ● Timing 2–3 h

20 Follow the chart below to determine how many starting nuclei you need and their volume. If you are doing sci on only one sample (filling all 12 columns of the 96-well plate with that

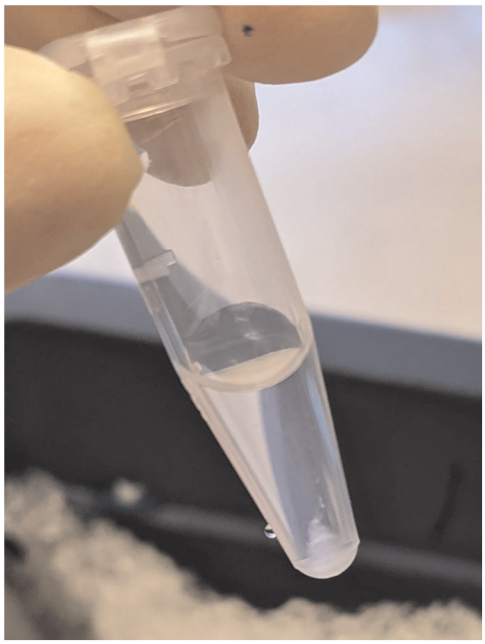

Fig. 5 | Nuclear pellet size. This is approximately the size of the nuclei pellet (~2 million nuclei) needed for one plate of RT. Extra fixed nuclei can be divided into aliquots and snap-frozen.

sample), then follow the volumes for 2 million cells. Figure 5 shows the approximate size of a pellet with 2 million nuclei needed for a whole plate. If you are dividing the plate among multiple samples, then the chart will outline the cell number you need depending on how many columns in the plate each sample will occupy. For example, if you will have two samples on one plate, each one will take six columns, you will need 1 million cells of each sample in a tube to start with in a volume of 250 µl each, and you will add 28 µl of dNTPs to each before distributing to their RT plate wells.

No. of cells	**2,000,000**	**1,000,000**	**800,000**	**500,000**	**400,000**	**200,000**
No. of columns	12	6	4	3	2	1
Nuclei volume (µl)	500	250	170	125	85	42.5
10 mM dNTP (µl)	56	28	19	14	9.5	4.75

21 Resuspend an aliquot of frozen nuclei in 500 µl of 0.3 M SPBSTM to start. Count. Dilute nuclei if necessary to get an accurate count. If nuclei are clumpy, sonicate (Diagenode Bioruptor Plus) on low intensity for 12 s only at 4 °C. If there is still an excess of clumps, put the nuclei over a 40-µm Flow-mi pipette tip filter before counting. Use of a Flow-mi filter is a last resort because it results in nuclei loss but is helpful if you have an excess of nuclei.

22 Pull out the desired amount of nuclei into a new tube and spin. Remove the supernatant and resuspend nuclei in the necessary volume of 0.3 M SPBSTM determined by the chart above and add the appropriate amount of dNTPs.

23 Distribute 5-µl aliquots of nuclei+dNTP mix to each well of the plate on ice.

24 Quickly spin the plate of three-level RT primers (10 µM) (5′-/5Phos/CAGAGCNNNNNNNN[10-bp RT index]TTTTTTTTTTTTTTTTTTTTTTTTTTTTTT-3′, where 'N' is any base)

25 Add 2 µl of primer to each well. Do not pipette up and down to mix; just stir gently with the pipette tips.

26 Incubate the plate at 55 °C for 5 min in a PCR machine (heated lid set to 65 °C) and then immediately place on ice.

27 While this is incubating, make the reaction mix. Note: we do not include dithiothreitol in this mix, because it will undo the DSP cross-links (it is not necessary for the RT to work).

RT mix per plate	1 well	×120
5× Superscript IV buffer (µl)	2	240
Superscript IV (200 U/µl) (µl)	0.5	60
Water (µl)	0.5	60

28 Put 3 µl of reaction mix into each well (45 µl of mix × 8 in a strip tube for multichannel pipetting), stirring gently with tips. The total volume per well is now 10 µl.
29 Incubate at 55 °C for 10 min (heated lid at 65 °C) and then immediately place on ice.
30 Ice plates until they are cold (10–15 min). Add 5 µl of cold 0.3 M SPBSTM per well. To maximize recovery, pool wells by using a 12-multichannel pipette with 200-µl tips to pipette gently up and down (the pipetting up and down is important to dislodge the nuclei, but try to avoid creating excessive bubbles) and combine each row of the plate into the bottom row. You can use the same tips for the whole plate. Then, collect these wells into two cold microcentrifuge tubes. (The solution will be bubbly, so it is difficult to squeeze into one tube.)
31 Spin at 500*g* for 3 min at 4 °C. The pellet will be small, but you should be able to see it. Remove and discard the supernatant.
32 Combine tubes and wash once more in 1 ml of cold 0.3 M SPBSTM. Spin at 500*g* for 3 min at 4 °C. Remove and discard the supernatant.

Ligation ● Timing 1.5 h

33 Resuspend nuclei in 1,200 µl of 0.3 M SPBSTM.
34 Distribute 11 µl to each well of a new plate on ice.
35 Quick-spin the plate of three-level ligation primers (10 µM) (5′-GCTCTG[9- or 10-bp ligation index]TACGACGCTCTTCCGATCT[reverse complement of ligation index]-3′).
36 Add 2 µl of primer to each well. Do not pipette up and down.
37 Make a 3:1 mix of 10× T4 ligation buffer and T4 DNA ligase (195 µl of 10× buffer + 65 µl of T4 DNA ligase).
38 Add 2 µl of ligase mix to each well (32 µl × 8 in a strip tube for multichannel pipetting). The total volume per well is now 15 µl.
39 Incubate for 20 min at room temperature.
40 Ice plates until cold.
41 Add 10 µl of cold SPBSTM to each well. This helps keep the nuclei from clumping and allows for more nuclei to be recovered.
42 Pool wells by using a 12-multichannel to pipette gently up and down (the pipetting up and down is important to dislodge the nuclei) and combine each row of the plate into the bottom row. Then, collect these wells into two cold microcentrifuge tubes.
43 Spin at 500*g* for 3 min at 4 °C. Remove and discard the supernatant.
44 Combine the two tubes and wash twice more with 1 ml of 0.3 M SPBSTM per wash.
45 Resuspend in 1 ml of 0.3 M SPBSTM to count. If nuclei are clumpy and cannot be teased apart with gentle pipetting, sonicate on low intensity for 6 s only at 4 °C and recount. Note that this is a shorter sonication step than before.
? TROUBLESHOOTING

Final distribution ● Timing 1 h

▲ **CRITICAL** In the final plate, you will want 1,000 nuclei/well (or 4,000/well if you have scaled up the experiment to 384 × 384 × 384). You should have enough nuclei to freeze multiple plates if you like.
46 Make 400 µl of 1× second-strand synthesis buffer for each plate in the final distribution. Dilute 40 µl of 10× second-strand buffer in 360 µl of water to get 1× concentration.
47 Spin down 100,000 nuclei for each plate desired for the final distribution (400,000 per plate if this is a 384 × 384 × 384 experiment). For each plate/100,000, resuspend in 400 µl of 1× second-strand synthesis buffer.
48 Put 4 µl of nuclei into each well of a regular, not LoBind, plate on ice.
49 Cover with foil seals and freeze plates at −80 °C or proceed with second-strand synthesis.
■ **PAUSE POINT** Plates may be kept frozen at −80 °C for ≤6 months.

Second-strand synthesis ● Timing 3 h (or overnight)

50 Thaw plate on ice.

51 Make second-strand synthesis mix on ice as follows:

Reaction mix per plate	1 well	×140
Water (µl)	0.675	94.5
Second-strand buffer (10×) (µl)	0.075	10.5
Second-strand enzyme (20×) (µl)	0.25	35

52 Put 1 µl of second-strand synthesis mix into each well (17 µl of mix ×8 in a strip tube for multichannel pipetting). The total volume per well is now 5 µl.

53 Incubate for 2.5 h at 16 °C in a PCR machine (no heated lid).

■ **PAUSE POINT** Keep the plate at 4 °C for ≤24 h.

Protease digestion ● Timing 2 h

54 Add 1 µl of protease to each well.

▲ **CRITICAL STEP** This is not proteinase K. Using Qiagen protease (cat. no. 19157) is important because it can be heat-inactivated. See Reagent setup.

55 Incubate at 37 °C for 30 min in a PCR machine (47 °C heated lid). Check 1 µl on a microscope: mix a 1-µl sample with 2 µl of diluted Yoyo-1 dye, put this on a slide and check on the GFP channel. You should see whisps of DNA instead of intact nuclei (Fig. 6). If not, incubate for another 10 min before proceeding to heat inactivation.

56 Heat-inactivate the protease at 75 °C for 20 min (85 °C heated lid).

▲ **CRITICAL STEP** Do not lower this temperature and do not shorten this time. Qiagen lists different conditions for heat-inactivating their protease, but it is not sufficient. Put the plate on ice after inactivating.

■ **PAUSE POINT** Store the plate at 4 °C for ≤1 week.

Tagmentation ● Timing 1 h

57 On ice, make tagmentation mix as follows:

Reaction mix per plate	1 well	×110
Tagment DNA buffer (µl)	5	550
N7-loaded Tn5 (µl)	0.125	13.75

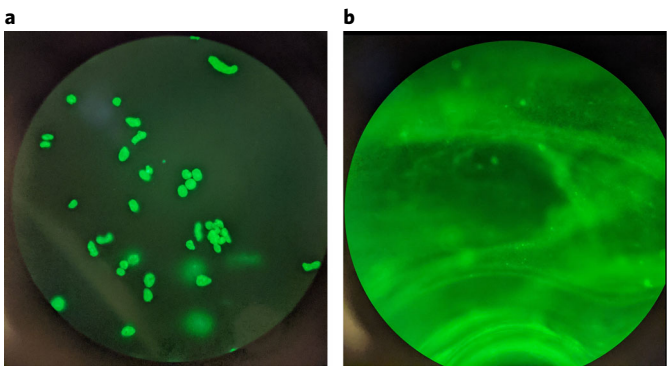

Fig. 6 | Visualizing protease digestion of the nuclei by Yoyo-1 staining. Stained nuclei as seen through the eyepiece of an inverted microscope at 200× magnification on the GFP channel. **a**, Nuclei after ~10 min of protease digestion, swelling and starting to lose integrity. **b**, Nuclei after 30 min of protease digestion; DNA has been released, and now the protease can be heat-inactivated.

58 On ice, add 5 µl of tagmentation mix to each well. The total volume per well is now ~10–11 µl.
59 Incubate for 5 min at 55 °C in a PCR machine (65 °C heated lid). Do not put on ice afterwards; just keep on the bench at room temperature as you add the next reaction mix, or else the SDS will come out of solution at the next step.
60 Remove the transposases with this buffer (keep at room temperature):

Reaction mix per plate	1 well	×120
1% (wt/vol) SDS (µl)	0.4	48
BSA (µl)	0.4	48
Water (µl)	1.8	216
Total (µl)	2.6	312

61 Add 2.6 µl to each well and mix (39 µl ×8 into a strip tube for multichannel pipetting).
62 Incubate for 15 min at 55 °C in a PCR machine (65 °C heated lid).
63 Quench SDS by adding 2 µl of 10% (vol/vol) Tween 20 to each well.

PCR amplification ● Timing 1 h for PCR, 2–3 h for gel purification, 1 h for loading the sequencer

64 Assemble the PCR master mix. PCR is done with 96 indexed P7 primers and 1 P5 primer. You can also add an optional index sequence on the P5 primer for multiplexing multiple plates. The primers are as follows: plate of 96 indexed PCR P7 primers (5′-CAAGCAGAAGACGGCATACGAGAT[PCR P7 index]GTCTCGTGGGCTCGG-3′) and TruSeqP5 primer(s) (5′-AATGATACGGCGACCACCGAGATCTACAC[PCR P5 index]ACACTCTTTCCCTACACGACGCTCTTCCGATCT-3′).

Reaction mix per plate	1 well	×110
2× NEBNext (µl)	20	2,200
TruSeqP5 primer (100 µM) (µl)	0.2	22
Water (µl)	3.2	352
Total (µl)	23.4	2,574

65 Add 2 µl of indexed P7 primers (10 µM) to each well.
66 Add 23.4 µl of PCR master mix to each well.
67 Amplify in a PCR machine by using 16 cycles with a pre-extension step in the following program:

1	70 °C	3 min
2	98 °C	30 s
3	98 °C	10 s
4	63 °C	30 s
5	72 °C	1 min
6	go to Step 3, 15 more times	
7	72 °C	5 min

68 Run 1.5 µl of a few wells on a 6% PAGE gel in TBE buffer at 200 V for 30 min to check amplification. You should see a smear of products with primer-dimers underneath (Fig. 7a). We isolate a section of the smear centered on 400 bp.
? TROUBLESHOOTING
69 *Concentration of library and agarose gel purification.* Pool 3 µl of each well and do a 0.8× AMPure XP cleanup (230 µl of beads) by following the manufacturer's protocol. (Save the remaining plate by covering with foil and storing at room temperature or at 4 °C for ≤6 months,in case you need to redo cleanup or if you anticipate needing more library for a large NovaSeq run.) Wash the AMPure

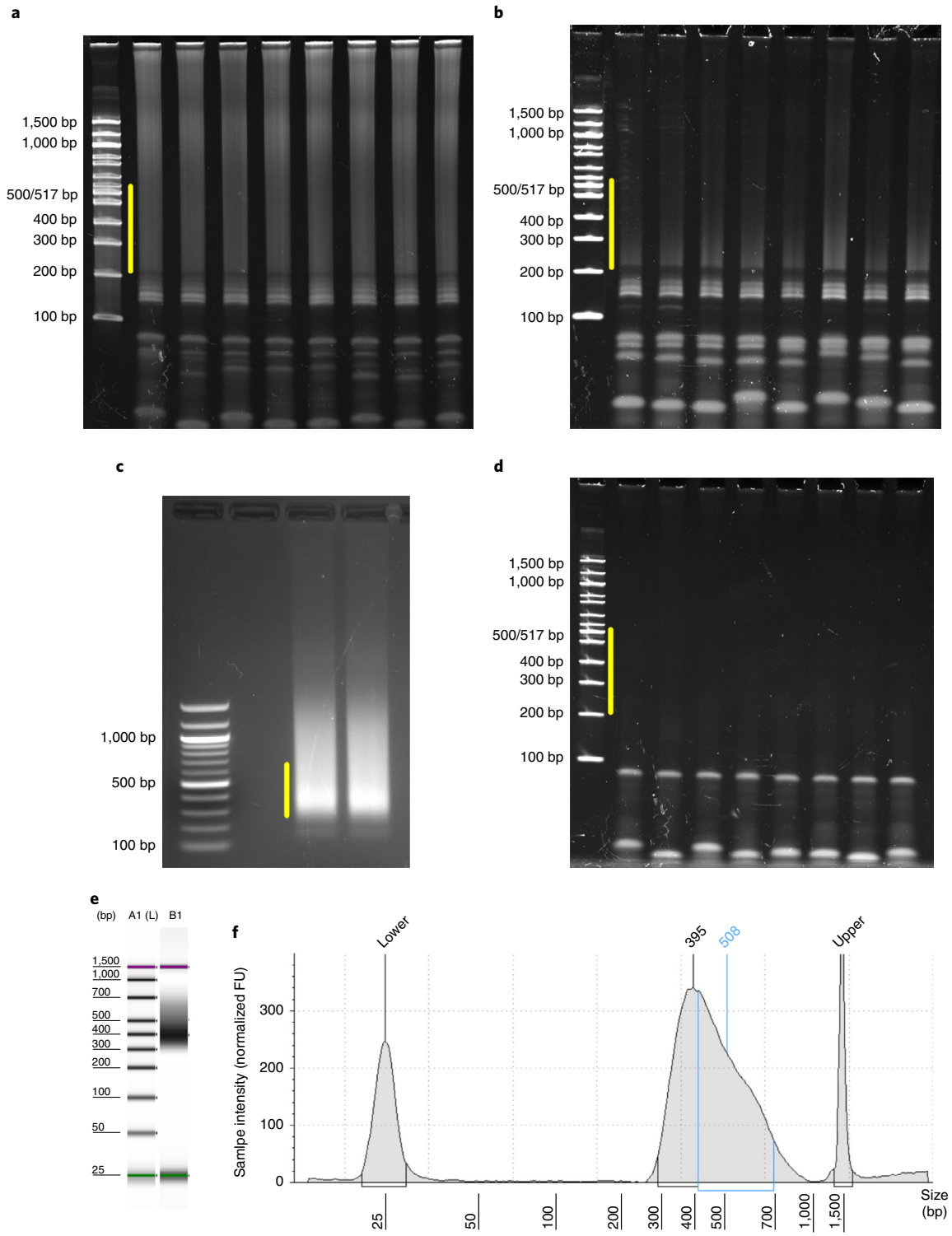

Fig. 7 | Evaluating the libraries after PCR. Yellow bars indicate the size range to look for. **a**, A good-quality library-sampling of eight wells from the PCR plate. 1.5 µl of each well is run on a 6% PAGE gel at 200 V for 30 min. A bright, long smear of products above the 200-bp marker indicates a robust library. The bands under 200 bp are probably empty adapter products, and those under 100 bp are primer-dimers. This one is from a 384 × 384 experiment with 4,000 nuclei plated in each well of the last plate. **b**, A medium-quality library with fainter smears. This was from a 96 × 96 experiment that underperformed; the quality of the data was good, and the number of UMIs per cell was as expected, but the experiment had fewer total cells overall, most likely due to counting estimates during the last round of plating. **c**, Gel size selection. Libraries have been concentrated via AMPure purification and run on a 1% agarose gel at 100 V for ~1 h. Most of the primer dimers are gone, and the libraries are cut out of the gel at the region indicated by the yellow line. This is a 384 × 384 experiment, so these smears are very intense. **d**, A failed experiment. The smaller number of primer-dimer bands suggests that the failure resulted from incomplete protease digestion. **e** and **f**, The final, gel-purified library fragment distribution as measured on an Agilent tapestation. Primer-dimers have been removed, and you should be left with a library ranging from 300 to 700 bp in size.

bead pellet twice gently with 70% (vol/vol) ethanol and elute the pool in 50 µl. Load this into a single 1-cm well on a 1% (wt/vol) agarose gel and run at 100 V for ~1 h (Fig. 7c). Cut out the smear between ~250 and 600 bp and use the NEB gel extraction kit. Use extra dissolving buffer, because the gel piece will be bigger than a normal slice, and run it all through the same purification column. Wash twice with 200 µl of NEB wash buffer and elute in 20 µl of EB. Quantitate the library with a Qubit double-stranded DNA high-sensitivity quantitation kit.

70 Run the library on NextSeq (or NovaSeq depending on the final cell numbers or sequencing depth desired) by using standard primers: Read1, 34 cycles; Index, 10 cycles; Read2, 48 cycles. If you have also used a P5 index for PCR, then add a second Index read of 10 bp.

Troubleshooting

Troubleshooting advice can be found in Table 2.

Table 2 | Troubleshooting table

Step	Problem	Possible reason	Solution
9	The RNase check is positive for RNase	The sample/lysis buffer volume ratio is too large, or tissue is especially rich in RNases	Redo cell lysis with fresh sample and add more lysis buffer, or decrease sample size, until you find a ratio that shows that all RNase has been deactivated. More lysis buffer is always better
10	Nuclei are clumping when they are lysed	Some cells are more delicate, and clumping indicates that DNA is leaking out of the nuclei	Try a shorter lysis, less igepal in the lysis buffer or more BSA in the lysis buffer. Make sure that you included $MgCl_2$ in the lysis buffer
19	Nuclei looked fine after lysis but clumped after fixation	Some cells are more delicate, and clumping indicates that DNA is leaking out of the nuclei	If sonicating the nuclei does not break up the clumps, try a shorter lysis, less igepal in the lysis buffer and/or more BSA in the lysis buffer. Try fixing and rehydrating more gradually
45	Losing too many nuclei when washing	Losing half the nuclei from the 2 million that started in the RT plate is not unexpected	Recovery can be maximized by being mindful that every transfer of the nuclei will result in some loss of nuclei to the walls of the tubes and pipettes. The supernatant does not always have to be completely removed for the washes if there is a chance to disturb the pellet. The biggest losses seem to happen when pooling wells, so at those steps make sure to gently pipette up and down a few times to dislodge settled nuclei before pulling them out of the well
45	Cells are sticking after ligation	DSP not fixing	Make a fresh batch of DSP; be sure to use anhydrous DMSO for a solvent. Some lots of DSP have been problematic. Some little clumps are expected. If you have a lot of cells, put them over a Flow-mi filter before continuing, to remove the clumps
68	No smear of the library on the gel	Sample quality is the biggest factor; bad reagents	If you have saved extra fixed nuclei, or extra final plates, you can retry from RT at Step 20 or second-strand synthesis at Step 48 with fresh reagents. Optionally, take an aliquot of frozen nuclei and bulk RNA extract to make sure that you are seeing RNA at all. Retry the experiment with less tissue in the lysis
	Spotty wells, or missing smears	Spotty wells can indicate that the protease was incompletely inactivated	If it is only a few bad wells, it is probably okay, but if it looks like ≥50% of wells do not have good smears, you may want to try again. If you have extra final plates, start again at Step 50 with second-strand synthesis, and when you finish the protease digestion, try a longer incubation
	The smear is mostly too high or mostly too low	Under- or over-tagmenting	If you have extra plates, you might try increasing or decreasing your Tn5 twofold. But do not worry if the library smear is not perfectly sized—with all the different lengths of transcripts, this is difficult to achieve. Any smear centered at ~400–600 is great

Timing

Steps 1–10, nuclei isolation: 2 h
Steps 11–19, nuclei fixation: 1 h
Steps 20–32, RT: 2–3 h
Steps 33–45, ligation: 1.5 h
Steps 46–49, final distribution: 1 h
Steps 50–53, second-strand synthesis: 3 h (or overnight)
Steps 54–56, protease digestion: 2 h
Steps 57–63, tagmentation: 1 h
Steps 64–69, PCR and gel purification: 3–4 h
Step 70, loading the sequencer: 1 h
Sequencing: 13–20 h
Bioinformatic analysis: 8–10 h

Anticipated results

The success of the experiment can be evaluated at the end of library generation after the third-round PCR step. You should see a smear of products on 6% TBE polyacrylamide gel above 200 bp. (Fig. 7a). The intensity of the smear is a good indicator of the cell/UMI number that you will see in the final

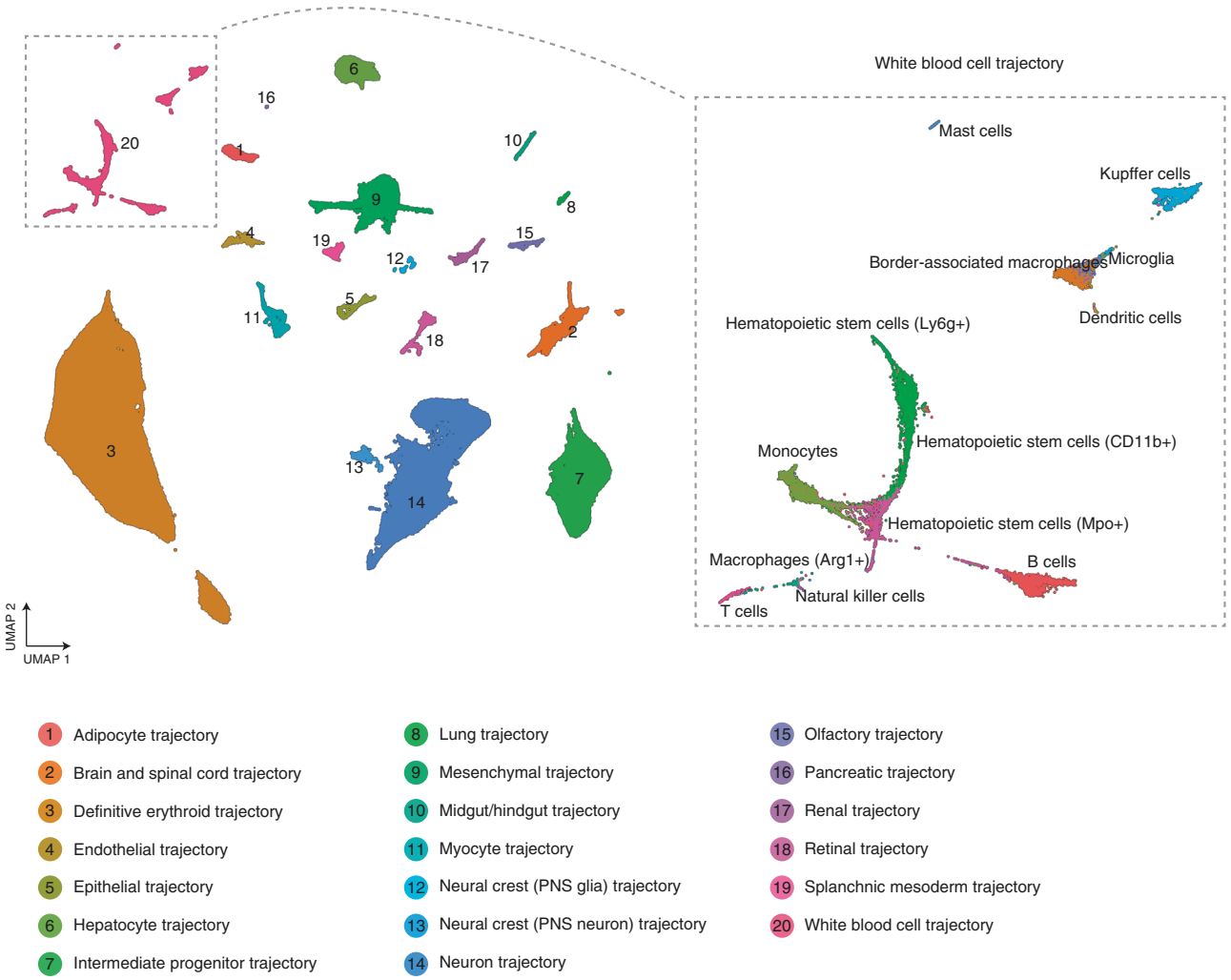

Fig. 8 | High-quality data of E16.5 mouse embryo generated by application of the optimized sci-RNA-seq3 protocol. 2D UMAP visualization of the new E16.5 dataset (GSE186824). All nuclei colored by each of the 20 cell trajectories are shown on the left. A subview of global 2D UMAP visualization highlighting subpopulations of the white blood cells' trajectory is shown on the right. See Supplementary Information for details of data analysis and cell type annotation.

data (compare Fig. 7a,b). An absence of a smear (Fig. 7d) can result from poor-quality tissue in which the RNA has been degraded or a failure somewhere in the protocol. Primer dimers are expected and, after gel purification, should be completely removed, as visualized on the Tapestation (Fig. 7e,f).

In bioinformatic analysis, during the UMI Attach step in which the reads are filtered for the expected indexes, you should see a filter rate of ~75–80% or higher. Aligning reads with STAR[21] should see a mapping rate over 70% (okay) to 80% (better). For a single NextSeq 550 High sequencing run of one plate of PCR at 1,000 cells/well, you should expect 300,000–400,000 reads, with <30% PCR duplicates. At 4,000 cells/well, you should see a duplication rate <20%.

To evaluate the performance of the improved protocol, we used the sci-RNA-seq3 method on an E16.5 mouse in a 192 × 192 experiment (two plates of RT indexes and two plates of ligation indexes) with 3.5 plates of PCR indexes. 2,000 nuclei/well were placed in the final PCR plates, for potentially 700,000 nuclei in total. The pooled library was sequenced on a NovaSeq and after filtering, resulted in 381,888 nuclei with high quality (for these, the median UMI count per cell was 2,648; the median gene count detected per cell was 1,493). The resulting 2D UMAP (uniform manifold approximation and projection) visualization of cells with their cell-type annotation labels is shown in Fig. 8. The details of data analysis and more results can be found in Supplementary Information.

Data availability

Raw data from the E16.5 mouse embryo is available for download from the NCBI Gene Expression Omnibus repository with accession number GSE186824. Original photos and gels have been deposited at Figshare (https://doi.org/10.6084/m9.figshare.c.5915834).

References

1. Domcke, S. et al. A human cell atlas of fetal chromatin accessibility. *Science* **370**, eaba7612 (2020).
2. Cusanovich, D. A. et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* **348**, 910–914 (2015).
3. Cao, J. et al. Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667 (2017).
4. Cao, J. et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
5. Rosenberg, A. B. et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360**, 176–182 (2018).
6. Datlinger, P. et al. Ultra-high throughput single-cell RNA sequencing by combinatorial fluidic indexing. *Nat. Methods* **18**, 635–642 (2021).
7. Lareau, C. A. et al. Droplet-based combinatorial indexing for massive-scale single-cell chromatin accessibility. *Nat. Biotechnol.* **37**, 916–924 (2019).
8. Vitak, S. A. et al. Sequencing thousands of single-cell genomes with combinatorial indexing. *Nat. Methods* **14**, 302–308 (2017).
9. Yin, Y. et al. High-throughput single-cell sequencing with linear amplification. *Mol. Cell* **76**, 676–690.e10 (2019).
10. Ramani, V. et al. Massively multiplex single-cell Hi-C. *Nat. Methods* **14**, 263–266 (2017).
11. Mulqueen, R. M. et al. Highly scalable generation of DNA methylation profiles in single cells. *Nat. Biotechnol.* **36**, 428–431 (2018).
12. Cao, J. et al. Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science* **361**, 1380–1385 (2018).
13. Cao, J., Zhou, W., Steemers, F., Trapnell, C. & Shendure, J. Sci-fate characterizes the dynamics of gene expression in single cells. *Nat. Biotechnol.* **38**, 980–988 (2020).
14. Hwang, B. et al. SCITO-seq: single-cell combinatorial indexed cytometry sequencing. *Nat. Methods* **18**, 903–911 (2021).
15. Wang, Q. et al. CoBATCH for high-throughput single-cell epigenomic profiling. *Mol. Cell* **76**, 206–216.e7 (2019).
16. Srivatsan, S. R. et al. Massively multiplex chemical transcriptomics at single-cell resolution. *Science* **367**, 45–51 (2020).
17. Srivatsan, S. R. et al. Embryo-scale, single-cell spatial transcriptomics. *Science* **373**, 111–117 (2021).
18. Cao, J. et al. A human cell atlas of fetal gene expression. *Science* **370**, eaba7721 (2020).
19. Qiu, C. et al. Systematic reconstruction of cellular trajectories across mouse embryogenesis. *Nat. Genet.* **54**, 328–341 (2022).
20. Ehrenberg, L., Fedorcsak, I. & Solymosy, F. Diethyl pyrocarbonate in nucleic acid research. *Prog. Nucleic Acid Res. Mol. Biol.* **16**, 189–262 (1976).
21. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

Acknowledgements

We thank Bridget Kulesakara for the sucrose buffer advice, Jase Gehring for fixative expertise and Riza Daza for helpful feedback. We thank Diana O'Day, Mai Le, Roshella Gomes, Saskia Ilcisin and Dana Jackson for protocol testing, and the entire Shendure Lab for support and encouragement. This work was funded in part by NIH R01HG010632 to J.S. and C.T., UM1HG011586 to J.S. and B.J.B., R35GM137916 to B.J.B. and T32HG000035 to E.N. J.S. is an Investigator of the Howard Hughes Medical Institute.

Author contributions

B.K.M. developed the improved protocol with input and testing from E.N., M.P., R.G.-G., S.S., R.B.-G. and J.C. C.Q. performed all data analysis. E.N. collected and staged mouse embryos. J.C. developed the original protocol. B.J.B., C.T. and J.S. supervised aspects of the work, with J.S. providing overall oversight. B.K.M., C.Q. and J.S. wrote the paper, with input from all authors.

Competing interests

J.S. is a scientific advisory board member, consultant and/or cofounder of Cajal Neuroscience, Guardant Health, Maze Therapeutics, Camp4 Therapeutics, Phase Genomics, Adaptive Biotechnologies and Scale Biosciences. C.T. is a founder of Scale Biosciences. All other authors have no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41596-022-00752-0.

Correspondence and requests for materials should be addressed to Beth K. Martin or Jay Shendure.

Peer review information *Nature Protocols* thanks Linas Mazutis, Samantha Morris and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.2

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 2 February 2022; Accepted: 30 June 2022;
Published online: 19 October 2022

Related links

Key references using this protocol

Qiu, C. et al. *Nat. Genet.* **54**, 328–341 (2022): https://doi.org/10.1038/s41588-022-01018-x
Cao, J. et al. *Science* **357**, 661–667 (2017): https://doi.org/10.1126/science.aam8940
Cao, J. et al. *Nature* **566**, 496–502 (2019): https://doi.org/10.1038/s41586-019-0969-x