Methods in Molecular Biology (2022) 2520: 233–259
DOI 10.1007/7651_2021_435

Published online: 19 October 2021

A Modified SMART-Seq Method for Single-Cell Transcriptomic Analysis of Embryoid Body Differentiation

Jianqun Zheng, Ying Ye, Qiushi Xu, Wei Xu, Wensheng Zhang, and Xi Chen

Abstract

Embryoid bodies (EBs) are aggregate of cells that contain three embryonic germ layers. They can be formed by direct differentiation from pluripotent embryonic stem cells (ESCs), which serves as a useful model for understanding early embryo development. Due to the mixture of different cell types, it is necessary to investigate EBs at the single-cell level. Here, we describe a robust and straightforward method for single-cell gene expression profiling during mouse EB differentiation from mouse ESCs (mESCs). The protocol is modified from a widely used method in the SMART-seq family, which only requires standard molecular biology techniques and lab equipment. It allows for accurate 3′ counting of transcript at the single-cell level, which helps reveal cellular identities during EB formation. Combined with perturbation experiments, the method provides an opportunity for mechanistic studies of embryo development at the single-cell level.

Key words Mouse embryonic stem cells, Embryoid body, Genomics, Single-cell RNA-seq

1 Introduction

Mouse embryonic stem cells (mESCs) are pluripotent cells that have the ability of self-renewal and the potential to differentiate into almost all cell types of a mouse [1]. Under appropriate culture conditions, they can be maintained in vitro indefinitely while keeping the state of pluripotency [2–4]. Importantly, cultured mESCs can be directed to differentiate in petri dishes into embryoid bodies (EBs) that contain a mixture of cells from all three germ layers (mesoderm, endoderm, and ectoderm). This is a process that recapitulates several key events, such as gastrulation, during early mouse embryo development [5]. Therefore, the in vitro EB differentiation system serves as an invaluable platform to study embryo development. It has many advantages including easy genetic manipulation and unlimited cell numbers, which allows mechanistic studies that are often difficult and time-consuming in vivo. Due to the heterogeneous mixture of cell types in EBs, it is necessary to

characterize EBs at the single-cell level to pinpoint different responses of different cell types during the differentiation process.

With the development of the genomic technologies, it is now possible to sequence the genome, the transcriptome and the epigenome of single cells [6–8]. Early single-cell studies focused on analyzing transcriptomes of rare cells, such as oocytes and early embryonic cells [9, 10]. Current studies mainly use single-cell technologies to characterize the cell composition of complex tissues, annotate cell types and put cells back to the spatial context [6]. There are two main types of technologies for transcriptome profiling at the single-cell level. The first type is droplet- or nanowell-based high-throughput technologies, where thousands of cells can be captured in droplets or nanowells by barcoded oligos. Cell barcodes are added at the reverse transcription stage, and only 3′ or 5′ of the mRNA is sequenced. The commonly used methods of this type include Drop-seq [11], inDrop [12], Seq-Well [13], Microwell-seq [14], and 10× Genomics [15]. Gene expressions of thousands of single cells can be profiled per experiment using these methods, but special equipment or expensive commercial platform is needed. The second type is plate-based methods, where single cells are isolated into each well of 96-well or 384-well plates. Alternatively, a combinatorial indexing strategy is used to avoid single-cell isolation, and single cells can be identified later by combination of different levels of barcodes. MARS-seq [16], STRT-seq [17], SMART-seq [18–20], Quartz-seq [21], CEL-seq [22], SCRB-seq [23], sci-RNA-seq [24], and SPLiT-seq [25] belong to this type. These methods can be implemented in any lab with basic molecular equipment, but they either have low throughput (limited number of cells) or low sensitivity (limited number of detected genes).

The single-cell technologies have proved to be very useful in the field of stem cell research. For example, they have been widely used to characterize the trajectories of early embryo development [10, 26–29] and investigate the transcriptome heterogeneity of stem cells under different in vitro culture conditions [30, 31]. In a study where the 10× Genomics platform [15] was used to profile the single-cell gene expressions during EB differentiation, the authors found several distinct cell populations by day 4 and day 6 and described a detailed differentiation trajectories in the process [32]. Interestingly, three different populations of cells on day 4 are revealed which possess distinct pluripotent signatures [32]. A more recent study combined CEL-seq2 [22] with genetic barcodes to profile gene expression and the lineage information of single cell at the same time during the process of EB formation [33]. The study found several branching points and provided a detailed lineage

relationship among different cell types during the differentiation process [33].

Among all the single-cell transcriptome profiling technologies, SMART-seq3 is one of most sensitive methods and has a unique advantage of being able to interrogate full-length mRNA [20]. In SMART-seq3, all procedures, from cell lysis to the final library preparation, are done separately for each individual cell in the wells of 96-well or 384-well plates. Libraries are pooled just before sequencing [20]. Though sensitive and powerful, the entire workflow is laborious and expensive.

Here, we described a modified version of SMART-seq3 done in 384-well format, where we introduced barcoded reverse transcription (RT) primers (Fig. 1a) at the very beginning of the protocol. All cells are pooled after cDNA amplification, and the final library of all cells is done in one tube. Finally, the 3′ end of the transcript is sequenced, and gene expressions are quantified by the unique molecular identifiers (UMIs) in the RT primers (Fig. 1b). The modified method can serves as a useful alternative when full-length transcript is not needed. It reduces a lot of hands-on time and saves a significant amount of reagent cost. Only basic lab equipment is needed to execute the protocol.

2 Materials

Prepare all solutions using ultrapure water (ddH_2O) and molecular biology grade reagents. Room temperature means 23–25 °C.

2.1 mESCs Maintenance and Embryoid Body Differentiation

1. Plate Coating Solution: dilute 1% gelatin in DPBS to a final concentration of 0.1%.
2. Stem Cell Culture Media: DMEM + 10% fetal bovine serum + 100 U/mL penicillin + 100 μg/mL streptomycin + 0.1 mM 2-mercaptoethanol + 1000 U/mL LIF.
3. EB Differentiation Media: Stem Cell Culture Media without LIF.

2.2 Single-Cell RNA-Seq

1. Oligonucleotides: dissolve all oligonucleotides in ddH_2O (*see* Table 1) to a concentration of 100 μM. Store in −80 °C.
2. 50% (wt/vol) polyethylene glycol 8000 (PEG-8000) solution: dissolve 25 g PEG 8000 powder in 25 mL ddH_2O. Once the powder is fully dissolved, top up the volume to 50 mL with ddH_2O.

a

Barcoded RT Primer:
5'- CTACACGACGCTCTTCCGATCT[8-bp well barcode][10-bp UMI]TTTTTTTTTTTTTTTTTTTTTTTTTTVN -3'

Template Switching Oligo (TSO):
5'- AAGCAGTGGTATCAACGCAGAGTACATrGrGrG -3'

cDNA PCR Fwd Primer:
5'- CTACACGACGCTCTTCCGATCT -3'

cDNA PCR Rev Primer:
5'- AAGCAGTGGTATCAACGCAGAG -3'

Library PCR Universal Primer:
5'- AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTC -3'

Library PCR Index Primer:
5'- CAAGCAGAAGACGGCATACGAGAT[8-bp plate barcode]GTCTCGTGGGCTCGG -3'

Colour code:
Illunima P5
Trueseq Read 1
Well barcode
UMI
TSO
Nextera Read 2
Plate barcode
Illumina P7

b

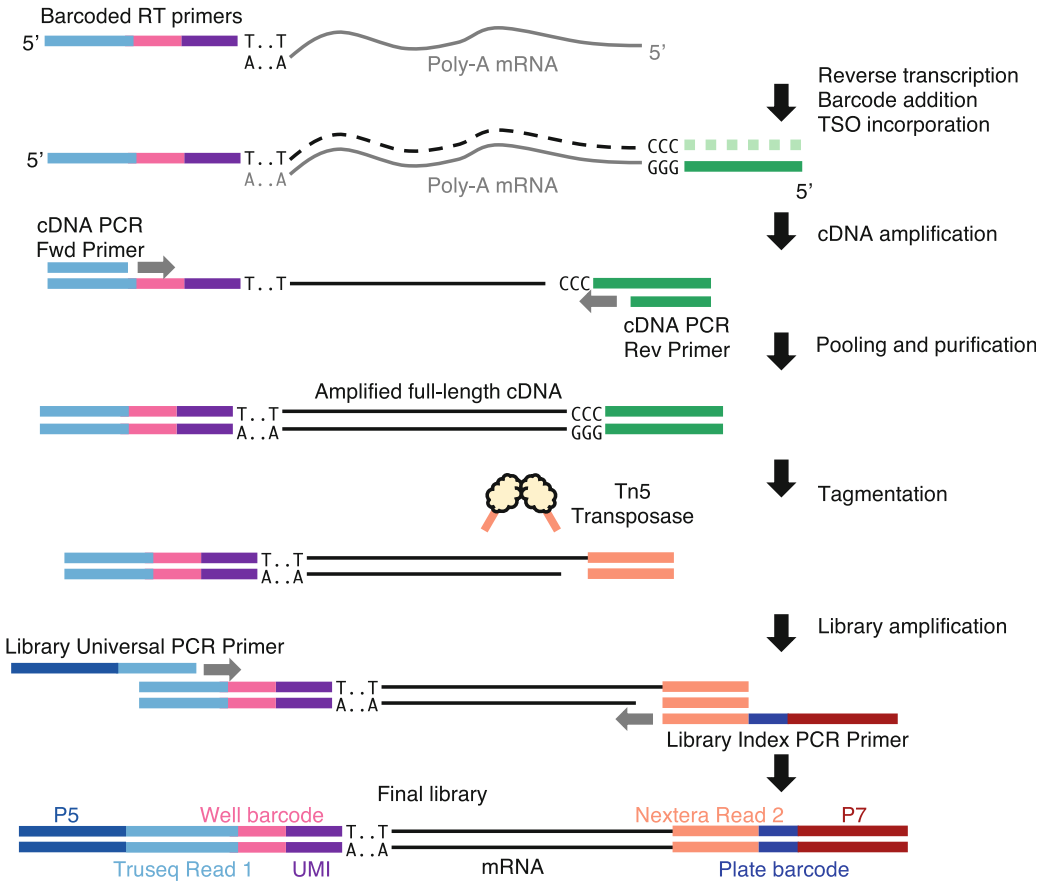

Fig. 1 Schematic view of the sequencing library preparation workflow. (**a**) Oligonucleotide information of the protocol. The identities of different parts of the oligo are color-coded. (**b**) The library preparation steps starting from reverse transcription to final library

Table 1
Oligonucleotides

Name	Sequence (5′– >3′)	Comments
Barcoded RT Primer	CTACACGACGCTCTTCCGATCT[8-bp well barcode]NNNNNNNNNNTTTTTTTTTTTTTTTTTTTTTTTTTVN	*See* Table 2 for the well barcode
Template Switching Oligo (TSO)	AAGCAGTGGTATCAACGCAGAGTACATrGrGrG	rG indicates riboguanosine
cDNA PCR Fwd Primer	CTACACGACGCTCTTCCGATCT	
cDNA PCR Rev. Primer	AAGCAGTGGTATCAACGCAGAG	
Library PCR Universal Primer	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTC	
Library PCR Index Primer	CAAGCAGAAGACGGCATACGAGAT[8-bp plate barcode]GTCTCGTGGGCTCGG	*See* Table 3 for the plate index

3. FACS buffer: dissolve 1 g bovine serum albumin (BSA) powder in 100 mL DPBS. Pass through 0.22 μm filter and store in 4 °C.
4. Lysis Buffer (freshly prepared).

Stock	For one well	Final Concentration
50% PEG-8000	0.4 μL	6.67%
10% Triton X-100	0.03 μL	0.1%
RNase Inhibitor (40 U/μL)	0.04 μL	0.53 U/μL
dNTPs (25 mM/each)	0.08 μL	0.67 mM/each
ddH_2O	1.45 μL	N/A
Total	**2 μL**	N/A

5. Reverse Transcription Mix (4×) (freshly prepared).

Stock	For one well	Final concentration
1 M Tris–HCl, pH 8.0	0.1 μL	100 mM
1 M NaCl	0.12 μL	120 mM

(continued)

Stock	For one well	Final concentration
100 mM $MgCl_2$	0.1 μL	10 mM
100 mM GTP	0.04 μL	4 mM
100 mM DTT	0.32 μL	32 mM
RNase Inhibitor (40 U/μL)	0.05 μL	2 U/μL
Template Switching Oligo (TSO) (100 μM)	0.08 μL	8 μM
Maxima H-minus RT Enzyme (200 U/μL) (Thermo Fisher, cat. mo. EP0751)	0.04 μL	8 U/μL
ddH_2O	0.15 μL	N/A
Total	**1 μL**	N/A

6. cDNA Amplification Mix (freshly prepared).

Stock	For one well	Final concentration
KAPA HiFi HotStart ReadyMix (2×) (KAPA, cat. no. KK2601)	5 μL	1.67×
cDNA PCR Fwd Primer (100 μM)	0.05 μL	0.83 μM
cDNA PCR Rev. Primer (100 μM)	0.05 μL	0.83 μM
ddH_2O	0.9 μL	N/A
Total	**6 μL**	N/A

7. Column Wash Buffer.

Stock	For 50 mL	Final concentration
1 M Tris–HCl, pH 7.5	0.5 mL	10 mM
100% Ethanol	40 mL	80%
ddH_2O	9.5 mL	N/A
Total	**50 mL**	N/A

8. TD Buffer (2×), store in −20 °C and discard after 3 months.

Stock	For 1 mL	Final concentration
1 M Tris–HCl, pH 7.5	20 μL	20 mM
1 M $MgCl_2$	10 μL	10 mM
Dimethylformamide (DMF)	200 μL	20%
ddH_2O	770 μL	N/A
Total	**1 mL**	N/A

Table 2
Well barcodes

A01	**B01**	**C01**	**D01**	**E01**	**F01**	**G01**	**H01**
AGTAAACC	CCGTTTAG	GACGCCGA	TTACGGTT	TCATGAAA	AGGGACGT	CACATGCG	GTTCCTTC
I01	**J01**	**K01**	**L01**	**M01**	**N01**	**O01**	**P01**
CAGTACTG	AGACTACT	TCTACTGC	GTCGGGAA	GAATCATA	ACTGTGGG	TTCAGCAT	CGGCATCC
A02	**B02**	**C02**	**D02**	**E02**	**F02**	**G02**	**H02**
TCACCTAG	CTGAACGA	AATTGGCT	GGCGTATC	ACATAGCG	TGGACAGC	GATCTCAA	CTCGGTTT
I02	**J02**	**K02**	**L02**	**M02**	**N02**	**O02**	**P02**
ACCTCTGT	CAACTATA	TGGGACCG	GTTAGGAC	GGAGATGC	ACCTTACA	CATCGCAG	TTGACGTT
A03	**B03**	**C03**	**D03**	**E03**	**F03**	**G03**	**H03**
CTTTAAGA	GAGCCTCG	AGAAGGAC	TCCGTCTT	AGGGTTTC	GACAGAAA	TCACACGG	CTTTCGCT
I03	**J03**	**K03**	**L03**	**M03**	**N03**	**O03**	**P03**
GACCGGAC	ATGATCTT	CCTTCAGG	TGAGATCA	GTTCCACT	AAGGAGAC	TGATGTGA	CCCATCTG
A04	**B04**	**C04**	**D04**	**E04**	**F04**	**G04**	**H04**
AAGATTAC	CTTCCGGA	GCCGAACT	TGATGCTG	GAAGAGTA	CGCACAGG	ACGTGTCC	TTTCTCAT
I04	**J04**	**K04**	**L04**	**M04**	**N04**	**O04**	**P04**
TAATACAC	CCCGCACA	GTTATGGT	AGGCGTTG	TATGAAGT	GTCATCTC	CCACGGCA	AGGTCTAG
A05	**B05**	**C05**	**D05**	**E05**	**F05**	**G05**	**H05**
CCATTATT	TGCCCTGG	ATGAGGCA	GATGACAC	GATTAACG	AGCGTGGC	CTGAGTAA	TCACCCTT
I05	**J05**	**K05**	**L05**	**M05**	**N05**	**O05**	**P05**
TGAGGTTT	ACCAAGGC	GAGTCCAG	CTTCTACA	AGCACTTT	CAGGTAGC	TTACAGCA	GCTTGCAG
A06	**B06**	**C06**	**D06**	**E06**	**F06**	**G06**	**H06**
AGTTACAG	TCGCTAGA	CACACTCT	GTAGGGTC	CATACGGT	ATCTAATC	TCAGTCAG	GGGCGTCA
I06	**J06**	**K06**	**L06**	**M06**	**N06**	**O06**	**P06**

(continued)

Table 2
(continued)

TGAGGAAC	GCGTACCT	CATACTTA	ATCCTGGG	TGGTGGTA	AAACTTAG	CTTGACCC	GCCACAGT
A07	**B07**	**C07**	**D07**	**E07**	**F07**	**G07**	**H07**
ATAAGTGG	CGCCAGTT	TATGCCAA	GCGTTACC	GGTCTAGG	ACAGAGAT	TAGAGCTA	CTCTCTCC
I07	**J07**	**K07**	**L07**	**M07**	**N07**	**O07**	**P07**
ACGGCTGA	GGCCTCTG	TATTGACC	CTAAAGAT	TGAATTGT	GCTGCGCA	AAGTGATG	CTCCACAC
A08	**B08**	**C08**	**D08**	**E08**	**F08**	**G08**	**H08**
TCAAGTCG	AGTCTGTA	CACGCAAT	GTGTACGC	GAAGTAAT	AGTTCGCA	CCGAATGC	TTCCGCTG
I08	**J08**	**K08**	**L08**	**M08**	**N08**	**O08**	**P08**
CGAGAGAC	GAGAGATT	TCCCTCCG	ATTTCTGA	TTCTCAAC	CCAGATCT	AGTATCGA	GAGCGGTG
A09	**B09**	**C09**	**D09**	**E09**	**F09**	**G09**	**H09**
TTCTGCGC	GGTAAGAT	ACACCATA	CAGGTTCG	CACTGAGA	ATAGTCTC	GCTAAGCG	TGGCCTAT
I09	**J09**	**K09**	**L09**	**M09**	**N09**	**O09**	**P09**
AGATCCTC	TATGGTCT	GCGCAGGA	CTCATAAG	AAACGAGA	CGCTAGCC	GCGGTCAT	TTTACTTG
A10	**B10**	**C10**	**D10**	**E10**	**F10**	**G10**	**H10**
TCCGAGTG	AATTCAGC	GTACTTCA	CGGAGCAT	CCTTGTTA	GAGGAACC	TGCATGAT	ATACCCGG
I10	**J10**	**K10**	**L10**	**M10**	**N10**	**O10**	**P10**
AGTAATGT	TACCCAAA	GTGGGCTG	CCATTGCC	TGTTAGGG	ATCGGAAT	GCAATCCA	CAGCCTTC
A11	**B11**	**C11**	**D11**	**E11**	**F11**	**G11**	**H11**
GTGACGAG	TGCTGATC	CCTGTTGT	AAACACCA	ATCGCATG	GAATATCA	TCGATCAC	CGTCGGGT
I11	**J11**	**K11**	**L11**	**M11**	**N11**	**O11**	**P11**
TAGCAAAT	GGTGTCTA	CCCTGTGG	ATAACGCC	TTTGTTGC	GACAACTA	CGATGGCG	ACGCCAAT
A12	**B12**	**C12**	**D12**	**E12**	**F12**	**G12**	**H12**
CATCTCCT	ACCACATC	GGATGTAG	TTGGAGGA	GGGTATTG	CATAGACA	TTCGTGGT	ACACCCAC
I12	**J12**	**K12**	**L12**	**M12**	**N12**	**O12**	**P12**

CCGCAAAG	TTTGTGCA	GACTGCTT	AGAACTGC	CATTGTGC	GTACCAAG	ACGGTGCA	TGCAACTT
A13	**B13**	**C13**	**D13**	**E13**	**F13**	**G13**	**H13**
GTTTACCA	ACCCTTTC	CGAGCAGT	TAGAGGAG	GGTACCAC	CTATAGTA	TACCTTGT	ACGGGACG
I13	**J13**	**K13**	**L13**	**M13**	**N13**	**O13**	**P13**
CAATACCT	ACTAGGAG	GGCCTTGA	TTGGCATC	AGGGCGAA	GCCCATCC	CATTGATT	TTAATCGG
A14	**B14**	**C14**	**D14**	**E14**	**F14**	**G14**	**H14**
CGCTAATG	TCAGCGAA	ATTCTTGT	GAGAGCCC	AGCCGCAG	TTTGAGTC	GAGTTTCT	CCAACAGA
I14	**J14**	**K14**	**L14**	**M14**	**N14**	**O14**	**P14**
AAGGCGTG	CTATATAC	GCCCGAGA	TGTATCCT	GTAACTAT	ACTCAGCA	TACGTAGG	CGGTGCTC
A15	**B15**	**C15**	**D15**	**E15**	**F15**	**G15**	**H15**
GGAAACAA	CCTCCTCC	TTGTTAGG	AACGGGTT	GCACATTT	ATTTGCCC	CGGATAGA	TACGCGAG
I15	**J15**	**K15**	**L15**	**M15**	**N15**	**O15**	**P15**
CAGCGCTT	ATCAAACG	TGTGCTAC	GCATTGGA	GAGCCGGT	ACTAACTC	CTAGGACG	TGCTTTAA
A16	**B16**	**C16**	**D16**	**E16**	**F16**	**G16**	**H16**
GCTGCAAC	TAATTCCA	CTCCATTG	AGGAGGGT	CATGTAAA	GTATCGCG	ACCCGCGT	TGGAATTC
I16	**J16**	**K16**	**L16**	**M16**	**N16**	**O16**	**P16**
CACCTGAA	ATGTCCGT	TCAAGATC	GGTGATCG	CTATTGGG	AGCGACAC	GCGACTCT	TATCGATA
A17	**B17**	**C17**	**D17**	**E17**	**F17**	**G17**	**H17**
ACGTAGTC	CTCGCTAG	GATATAGA	TGACGCCT	GTGCTCCG	AATAGGTC	TCCTAAGT	CGAGCTAA
I17	**J17**	**K17**	**L17**	**M17**	**N17**	**O17**	**P17**
TCTGCACG	ATCTTGTT	CGAAGCGA	GAGCATAC	TGTTCATG	GCGAGTGA	ATCCAGCT	CAAGTCAC
A18	**B18**	**C18**	**D18**	**E18**	**F18**	**G18**	**H18**
GGAGCTTG	CAGTGAAC	TTTCACGA	ACCATGCT	AGCCAAGC	GTTTGTTT	TAAGCCCG	CCGATGAA
I18	**J18**	**K18**	**L18**	**M18**	**N18**	**O18**	**P18**
ACTCTCGC	TGAACGTA	GTGGGACT	CACTATAG	ATGCATCA	TCAGTAGC	GGCAGGTG	CATTCCAT

(continued)

Table 2
(continued)

A19	**B19**	**C19**	**D19**	**E19**	**F19**	**G19**	**H19**
AGATTCAT	CTGAGATC	GCTCCTGG	TACGAGCA	TAGAATCA	CTCCTAGT	AGTGGCTG	GCATCGAC
I19	**J19**	**K19**	**L19**	**M19**	**N19**	**O19**	**P19**
GAATGAGG	CGGATGCT	ATTGCCAC	TCCCATTA	CACATCGC	TTGTATCT	AGTGGAAA	GCACCGTG
A20	**B20**	**C20**	**D20**	**E20**	**F20**	**G20**	**H20**
TCTTGCTC	ATCACAGA	GAACTGCG	CGGGATAT	CGCGTCAG	GTACGACC	AAGAAGGA	TCTTCTTT
I20	**J20**	**K20**	**L20**	**M20**	**N20**	**O20**	**P20**
TGCATACC	ATTTCGAG	GAAGGTGT	CCGCACTA	AGCTCATA	GCTATCGG	TTGGGTAT	CAACAGCC
A21	**B21**	**C21**	**D21**	**E21**	**F21**	**G21**	**H21**
ACGTCCTA	TGTGGGAT	GACATTCC	CTACAAGG	GCTGGCGA	CTAACATT	AGCTATCG	TAGCTGAC
I21	**J21**	**K21**	**L21**	**M21**	**N21**	**O21**	**P21**
AACGATAA	GCTCTGCT	TGGAGATG	CTATCCGC	CTTAGAAT	TAATCGGG	AGCCTCCA	GCGGATTC
A22	**B22**	**C22**	**D22**	**E22**	**F22**	**G22**	**H22**
TGACATAC	GTCTGACA	CCGACGTG	AATGTCGT	GGTCATTA	TAAGGCAT	ATCTCGGC	CCGATACG
I22	**J22**	**K22**	**L22**	**M22**	**N22**	**O22**	**P22**
CATCTTGG	TCGGGCCT	GTCACGTA	AGATAAAC	GGAATGGC	AATTCTTG	TCCGGCAA	CTGCAACT
A23	**B23**	**C23**	**D23**	**E23**	**F23**	**G23**	**H23**
TGTAGTGG	CCAGAATC	AAGCCGCA	GTCTTCAT	TTATCCTA	GACAAAGG	CCGCGTAC	AGTGTGCT
I23	**J23**	**K23**	**L23**	**M23**	**N23**	**O23**	**P23**
TGATAGCT	GTTATATG	CACCCTGA	ACGGGCAC	ATCAACAA	GGTTGAGC	CCACTTTG	TAGGCGCT
A24	**B24**	**C24**	**D24**	**E24**	**F24**	**G24**	**H24**
AACAGTGT	CCATCCTG	GTTCAGCC	TGGGTAAA	GCAATTAC	AAGCGACT	TTCTCGTG	CGTGACGA
I24	**J24**	**K24**	**L24**	**M24**	**N24**	**O24**	**P24**
TACTCGCC	ATAGAAGT	CGTATTTG	GCGCGCAA	ATGCTGTC	TGTACAAA	CACGGCCT	GCATATGG

Table 3
Plate index

N701	TCGCCTTA	**N702**	CTAGTACG	**N703**	TTCTGCCT	**N704**	GCTCAGGA
N705	AGGAGTCC	**N706**	CATGCCTA	**N707**	GTAGAGAG	**N710**	CAGCCTCG
N711	TGCCTCTT	**N712**	TCCTCTAC	**N714**	TCATGAGC	**N715**	CCTGAGAT
N716	TAGCGAGT	**N718**	GTAGCTCC	**N719**	TACTACGC	**N720**	AGGCTCCG
N721	GCAGCGTA	**N722**	CTGCGCAT	**N723**	GAGCGCTA	**N724**	CGCTCAGT
N726	GTCTTAGG	**N727**	ACTGATCG	**N728**	TAGCTGCA	**N729**	GACGTCGA

3 Methods

3.1 Thawing and Culturing mESCs in Feeder-Free Condition

1. Coat the flask with **Plate Coating Solution**: add just enough solution to cover the bottom of the plastic. Leave at room temperature for at least 5 min. Remove just before plating the cells.
2. Thaw frozen vials of mESCs in 37 °C water batch.
3. Centrifuge the cells at 250 × *g* for 3 min and remove supernatant.
4. Wash the cell pellet with prewarmed **Stem Cell Culture Media**, centrifuge 250 × *g* for 3 min and remove supernatant.
5. Resuspend the cell pellet in prewarmed **Stem Cell Culture Media**, and plate cells into the coated flask.
6. Shake the flask back and forth gently to distribute cells evenly.
7. Incubate the cells in a 37 °C, 5% CO_2 incubator. Change the medium once every day until they reach 80% confluency.

3.2 Passaging Feeder-Free mESCs

1. Aspirate the medium, wash the cells twice with prewarmed DPBS.
2. Add just enough prewarmed 0.05% trypsin to cover all cells.
3. Incubate the cells at room temperature for 3 min, and check under a microscope whether cells are detached from the bottom of the plate or dish.
4. Add equal volume of **Stem Cell Culture Media** to inactivate trypsin and pipette ups and downs to gently dissociate cells.
5. Transfer the cell suspension into a centrifuge tube, spin down at 250 × *g* for 3 min, and remove the supernatant.
6. Resuspend the cell pellet with prewarmed **Stem Cell Culture Media** by gently pipetting ups and downs.
7. Passage the cells at 1:5 or 1:8 ratios in coated flasks.

3.3 EB Differentiation

1. Prewarm the **EB Differentiation Media** in a 37 °C water batch.
2. Follow **steps 1–5** from **Subheading 3.2** to obtain mESC pellet with **EB Differentiation Media**.
3. Resuspend the cell pellet with prewarmed **EB Differentiation Media** by gently pipetting ups and downs.
4. Count cells using a cell counter. Then seed 2×10^6 of them onto a 9-cm low-attachment dish containing **EB Differentiation Media**.
5. After 1 day (day 1), transfer EB suspension into a centrifuge tube, spin down at $250 \times g$ for 3 min, and remove the supernatant.
6. Resuspend the EBs gently with prewarmed **EB Differentiation Media** by gently pipetting ups and downs.
7. Transfer 1/10 EB suspension to a 9-cm low-attachment dish containing **EB Differentiation Media**.
8. Culture the EBs to the desired time points and renew the media every other day.

3.4 Prepare Single-Cell Suspension

1. Transfer EB suspension at the desired time points to a centrifugation tube. Let EBs settle down to the bottom of the tube by gravity, and carefully remove the supernatant without disturbing EBs.
2. Add DPBS to wash EBs, let them settle for 1 min, and carefully remove the supernatant.
3. Add 0.05% trypsin to EBs and incubate at room temperature for 5 min. Carefully remove the supernatant.
4. Add **EB Differentiation Media** to neutralize the leftover trypsin. Let EBs settle for 1 min, and carefully remove the supernatant.
5. Add **FACS buffer** to EBs and dissociate the cells by gently pipetting ups and downs.
6. Pass the cell suspension through a 30 μm filter to get single-cell suspension, and count cells using a cell counter.
7. Adjust cell concentration to 1×10^6/mL using **FACS buffer** and transfer the cell suspension to an FACS tube.
8. Add DAPI at a final concentration of 1 μg/mL to the cell suspension, leave the tube on ice and proceed to FACS.

3.5 Single-Cell Sorting

1. Dilute the 100 μM **Barcoded RT Primer** stock into 2 μM concentration in a 384-well plate with ddH_2O. Seal the plate and it can be stored in −80 °C for up to 6 months (*see* **Note 1**).
2. Prepare the lysis plate: this can be done by preparing enough **Lysis Buffer**, and aliquot 2 μL into each well of a 384-well

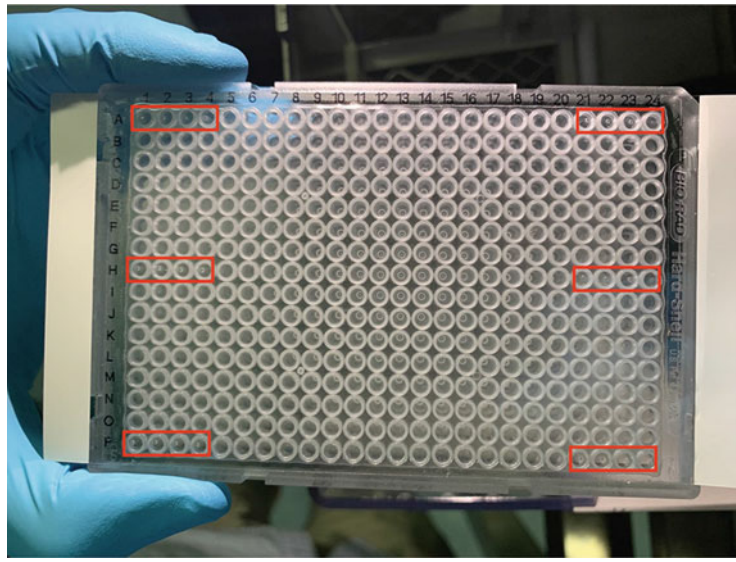

Fig. 2 A picture showing an example of adjusting the FACS machine. Note all the liquid drops, containing 30 single cells, are around the center of those wells (red rectangles)

plate. Then add 1 μL **Barcoded RT Primer** (2 μM) to each well to reach a final volume of 3 μL per well. This lysis plate can be stored in −80 °C for up to 6 months (*see* **Note 2**).

3. On the day of the experiment, take the lysis plate out of the −80 °C freezer, and thaw on ice.
4. Adjust the FACS machine setting and nozzle position: put an empty plate with seal in the FACS machine, and sort 20–30 single cells per well to the wells A1–4, A21–24, H1–4, H21–24, P1–4, P21–24. Check the drop on the seal, and adjust the nozzle position such that all drops are in the center of those wells (Fig. 2) (*see* **Note 3**).
5. After the nozzle adjustment, sort DAPI negative live single cells (Fig. 3) (*see* **Note 3** for detailed comments) into all the wells (one cell per well), except the well P24 which is used as a negative control.
6. Seal the plate and immediately spin down the plate using a plate centrifuge at 1000 × *g* for 1 min. The plate can be stored in −80 °C for up to 3 months (*see* **Note 4**).

3.6 Single-Cell cDNA Amplification

1. Take the lysis plate with sorted cells out of −80 °C freezer and thaw on ice. Briefly spin down to collect any evaporation. Incubate on a thermocycler at 72 °C for 10 min followed by 4 °C hold, with the lid temperature set to 105 °C.
2. Prepare enough **Reverse Transcription Mix (4×)**, remove the seal and add 1 μL to each well. Seal the plate again, and briefly spin down the plate.

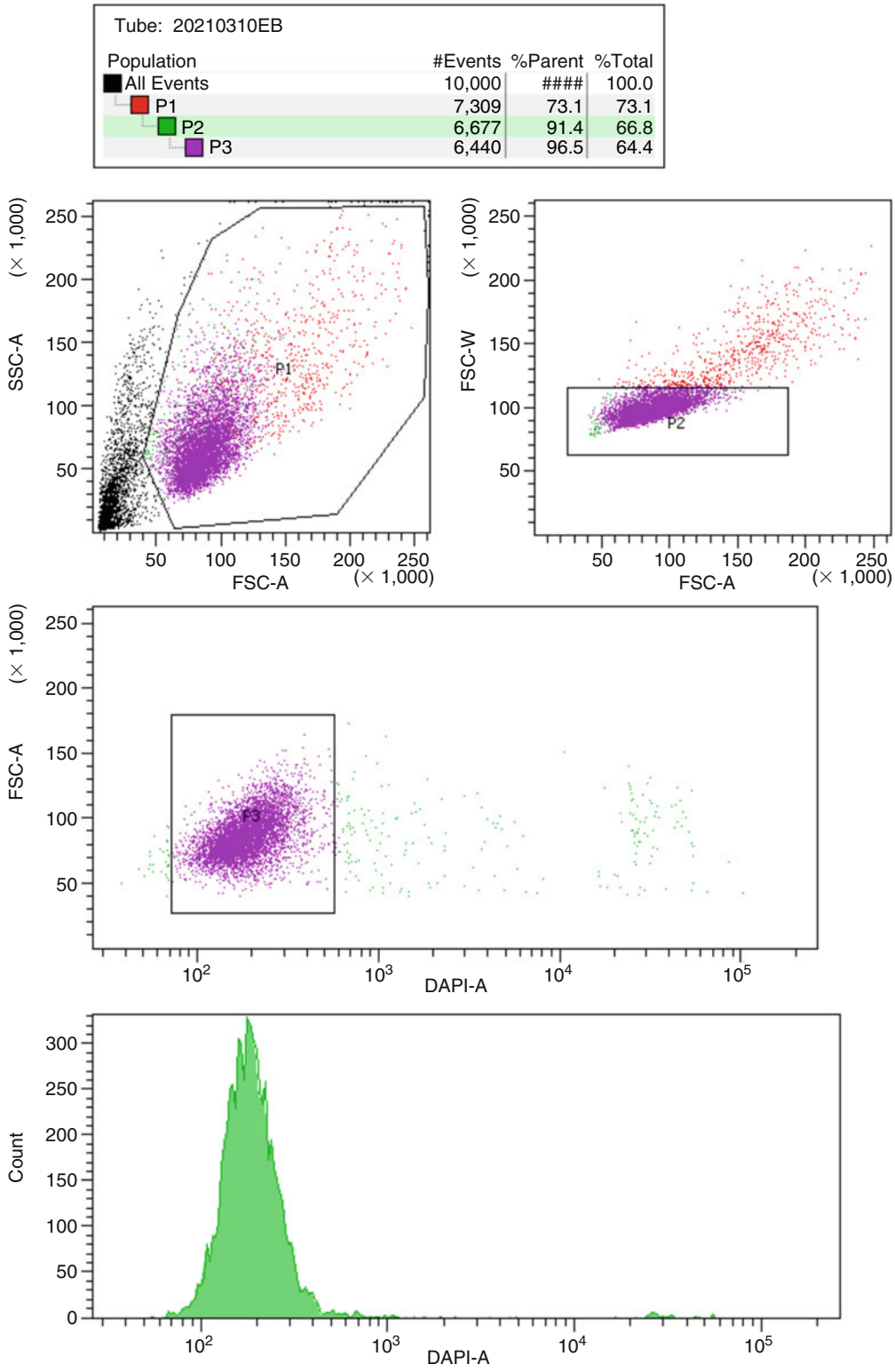

Fig. 3 FACS results of single cells from day 4 mouse EBs. First, FSC-A vs SSC-A was used to gate out the cell debris. Only intact cells (P1) were retained. Within P1, FSC-A vs FSC-W was used to gate out the cell multiplets. Only singlets (P2) were retained. Within P2, DAPI stain was used to gate out dead cells. Only DAPI negative live cells (P3) were sorted into the wells

3. Reverse transcription and template switching: put the plate on a thermocycler with the following program, with the lid temperature set to 105 °C.

Temperature (°C)	Duration	Cycle number
42	90 min	1
50 42	2 min 2 min	10
85	5 min	1

4. Prepare enough **cDNA Amplification Mix**, remove the seal and add 6 μL to each well. Seal the plate again, and briefly spin down the plate.
5. cDNA amplification: put the plate on a thermocycler with the following program, with the lid temperature set to 105 °C.

Temperature (°C)	Duration	Cycle number
98	5 min	1
98 65 72	20 s 30 s 4 min	16
72	5 min	1
10	Hold	N/A

6. The plate can be stored in −20 °C for up to a week.

3.7 Single-Cell Pooling, cDNA Purification, and Quality Check

1. Take the plates out of the −20 °C freezer, thaw on ice, and pool all reactions from a plate into one 50-ml tube. This can be done by using either a multichannel pipette (Fig. 4a) or a plate reservoir (e.g., Clickbio, cat. no. VBLOK200) (Fig. 4b). Each plate should be done separately, and do not mix wells from different plates at this stage (*see* **Note 5**).
2. Add 5 volumes of Buffer PB (Qiagen) to the pooled reaction, and mix by inverting the tube until the solution becomes homogeneous. A full 384-well plate normally yield a total of around 3.8 mL reaction. Therefore, 19 mL Buffer PB should be added.
3. Pass each plate pool through a single column from the QIAquick PCR Purification kit (Qiagen, cat. no. 28104) by putting a 20 mL Extender Tube (e.g., Angen Biotech, cat. no. D50071) or a 50 mL tube (with a puncture at the bottom) (Fig. 5a) on top of the column and connecting the column to a vacuum (Fig. 5b).
4. Pass 25 mL column wash buffer to wash the column.

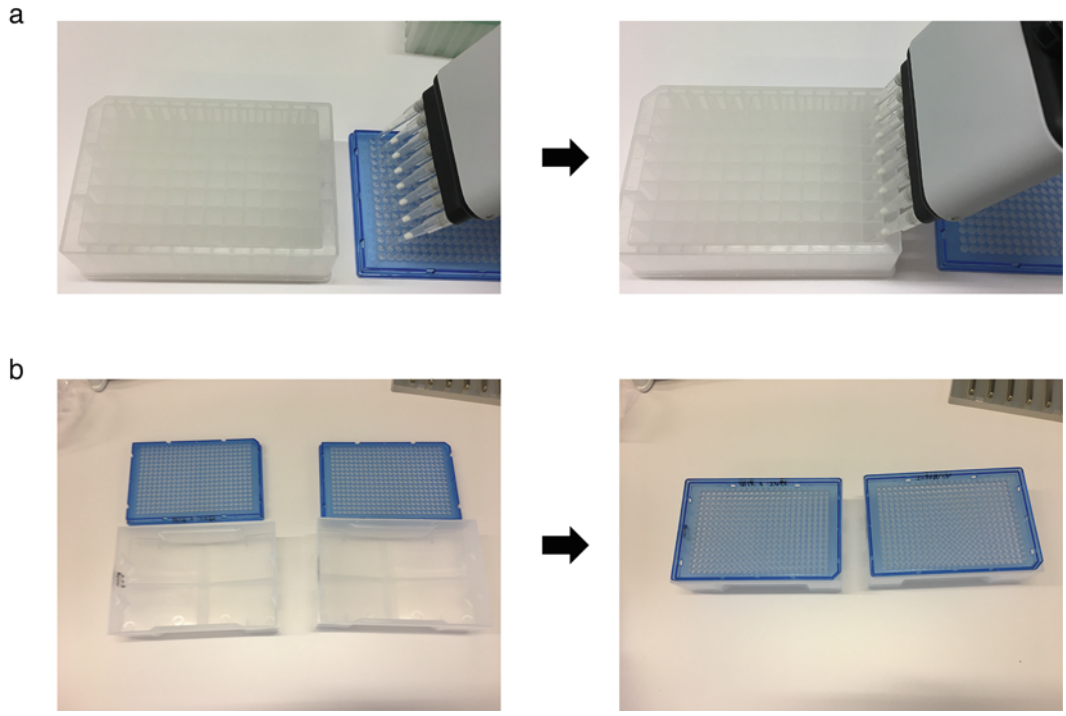

Fig. 4 Demonstration of the plate pooling procedure. (**a**) Plate pooling using a multichannel pipette. (**b**) Plate pooling using a plate reservoir

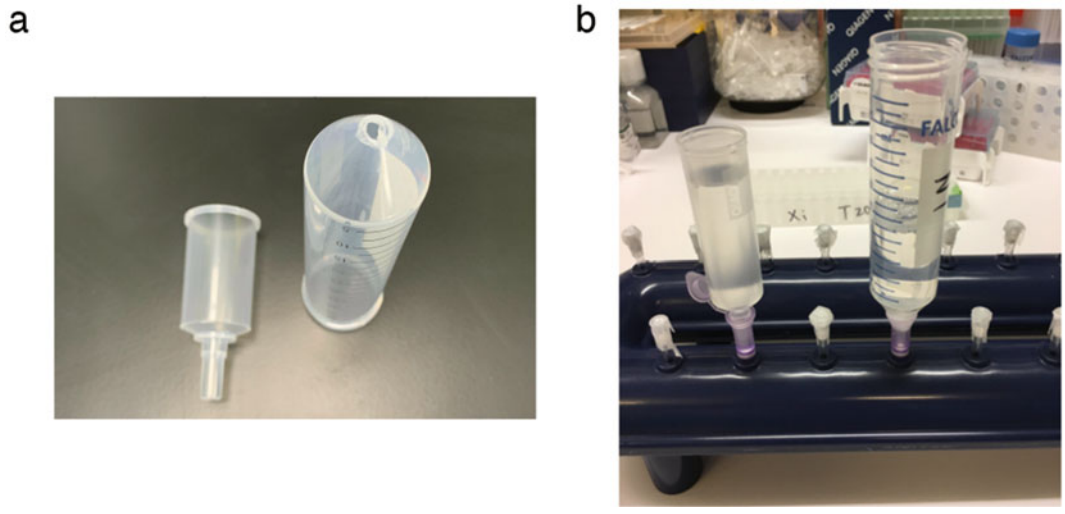

Fig. 5 Demonstration of the purification of large volume of cDNA using a single column. (**a**) Examples of the extender tube and home-made 50-mL tube with a puncture at the bottom. (**b**) An example showing how to connect the extender tube or 50-mL tube with the column and vacuum

5. Take the column off the vacuum, put on to a 2 mL collection tube and centrifuge on a table top centrifuge at top speed for 1 min to remove trace of ethanol on the column.
6. Add 50 μL 10 mM Tris–HCl, pH 8.5 to the center of the column, leave at room temperature for 1 min.
7. Put the column on a 1.5 mL Eppendorf tube and centrifuge on a table top centrifuge at top speed for 1 min to elute the cDNA. This will yield about 45 μL purified cDNA and can be stored in −80 °C indefinitely.
8. Use Exo I to remove the excess of primers by assembling the following reaction.

Stock	Amount	Final concentration
10× NEB 3.1 Buffer	5 μL	1×
Exonuclease I (NEB, cat. no. M0568)	2.5 μL	1 U/μL
Purified cDNA from above	42.5 μL	N/A
Total	**50 μL**	N/A

9. Incubate at 37 °C for 30 min to digest the primer left-over, and then at 80 °C for 2 min to inactivate the Exonuclease I (*see* **Note 6**).
10. Purify the full-length cDNA using SPRI beads (*see* **Note 7**) by adding 35 μL (0.7 volumes) to the Exo I digested reaction and mix well by pipetting ups and downs.
11. Leave at room temperature for 5 min, and put the tube on a magnetic stand.
12. Wait until the supernatant becomes clear, which usually takes a few minutes, and remove the supernatant.
13. While the tube is on the magnet, add 200 μL 80% ethanol to the beads and wait for 20 s.
14. Carefully remove the ethanol without disturbing the beads.
15. Repeat the ethanol wash for a total of three washes.
16. Let the beads air-dry until there is no shiny reflections of liquid on the surface of the beads.
17. Remove the tube from the magnet, and resuspend the beads in 20.5 μL 10 mM Tris–HCl, pH 8.5.
18. Leave at room temperature for 2 min, and then put the tube on a magnetic stand.
19. Wait until the supernatant becomes clear, and transfer 20 μL supernatant to a new 1.5 mL Eppendorf tube.
20. Measure the concentration of the cDNA using a Qubit dsDNA HS Assay Kits according to the manufacturer's instructions. A

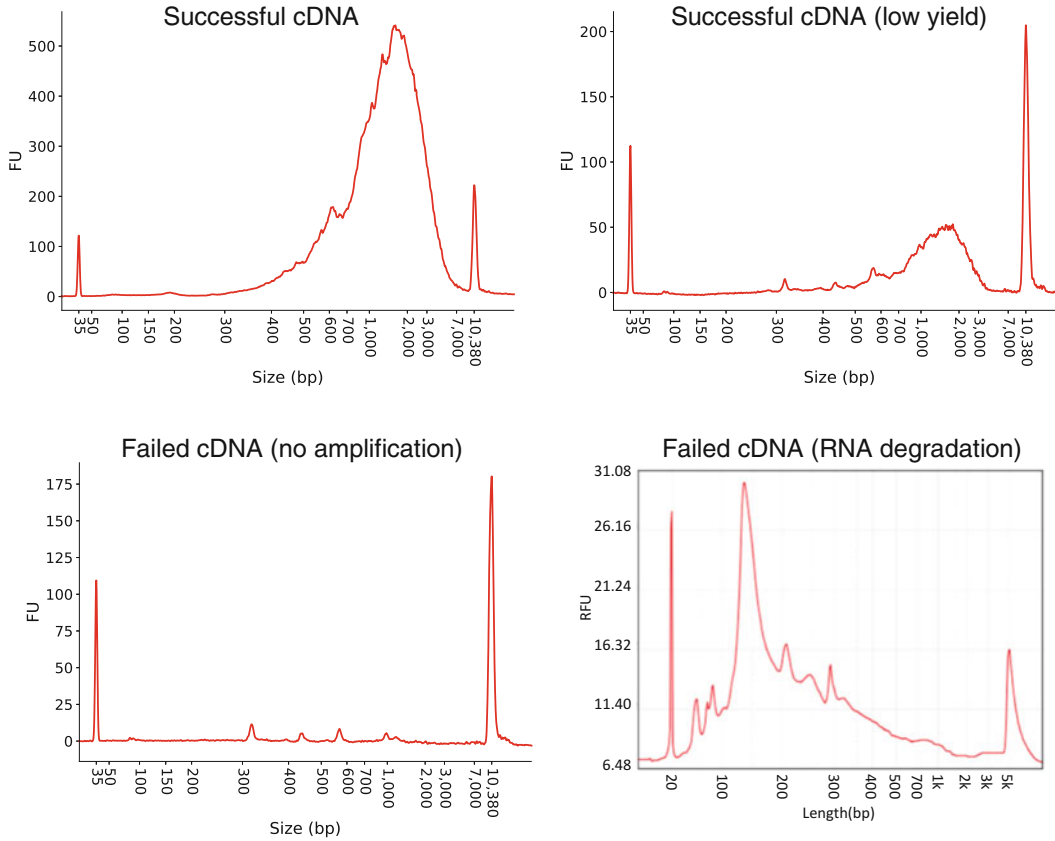

Fig. 6 Examples of successful and failed cDNA profiles

typical concentration is 20–40 ng/μL from a full 384-well plate of cells of EBs (*see* **Note 8**).

21. If needed, dilute the cDNA to a concentration of 1–10 ng/μL. Run 1 μL of cDNA on a Agilent Bioanalyzer to check the integrity of the cDNA (Fig. 6). In Fig. 6, examples of both successful and failed cDNA profiles are shown. *See* **Note 9** for detailed comments.

3.8 Sequencing Library Preparation and Quality Check

1. Perform tagmentation of full-length cDNA by assembling the following reaction in a 1.5 mL Eppendorf tube (*see* **Note 10**).

Component	Amount	Final concentration
Purified full-length cDNA	X μL (50 ng total cDNA)	10 ng/μL
Tn5 transposase	1 μL	N/A
2× TD Buffer	25 μL	1×
ddH_2O	24–X μL	N/A
Total	**50 μL**	N/A

2. Incubate the reaction at 55 °C for 5 min.
3. Stop the tagmentation reaction by adding 12.5 μL 0.2% SDS, mix by pipetting ups and downs and incubate at 55 °C for 5 min.
4. Purify the cDNA after tagmentation by adding 50 μL (0.8 volumes) SPRI beads to the tagmentation reaction and mix well by pipetting ups and downs.
5. Follow **steps 11–19** from Subheading 3.7 to purify the tagmented cDNA. This yields 20 μL purified tagmented cDNA.
6. Prepare the sequencing library by PCR amplification by assembling the reaction as follows (*see* **Note 11**).

Component	Amount	Final concentration
Purified tagmented cDNA	20 μL	N/A
Library Universal PCR Primer (10 μM)	2.5 μL	0.5 μM
Library Index PCR Primer (10 μM)	2.5 μL	0.5 μM
KAPA HiFi HotStart ReadyMix (2×) (KAPA, cat. no. KK2601)	25 μL	1×
Total	**50 μL**	N/A

7. Amplify using the following PCR program.

Temperature (°C)	Duration	Cycle number
72	5 min	1
98	2 min	1
98 65 72	10 s 30 s 20 s	8
10	Hold	N/A

8. Purify the library after PCR by adding 40 μL (0.8 volumes) SPRI beads the reaction and mix well by pipetting ups and downs.
9. Follow **steps 11–19** from Subheading 3.7 to purify the library. This yields 20 μL purified library per plate.
10. Run 1 μL on an Agilent Bioanalzyer to see the size distribution of the library (Fig. 7) (*see* **Note 12**).
11. Send the library for pair-end sequencing with the following setting (*see* **Note 13**):

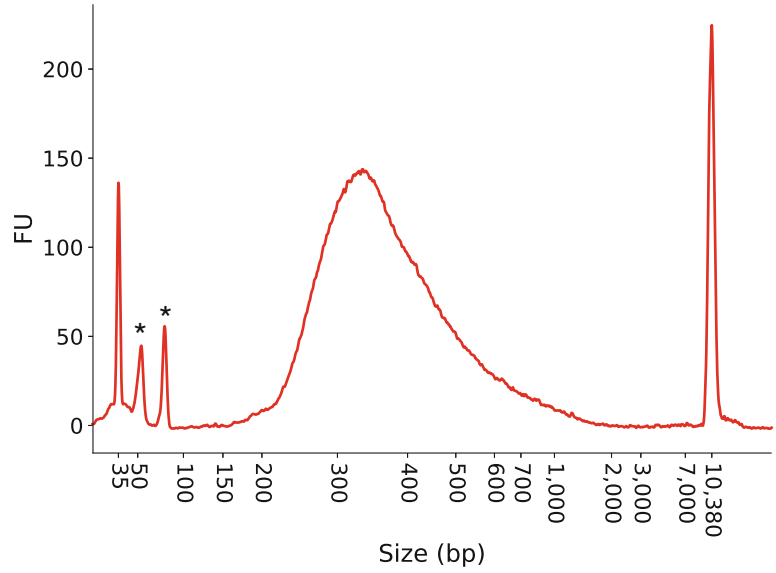

Fig. 7 An example of a successful sequencing library. Asterisks indicate primer leftover, which needs to be removed by another round of SPRI beads purification before sequencing

Read	Cycle number (read length)	Identity of the read
Read 1	At least 18 cycles (18 bp)	Well barcode and UMIs
Index 1	8 cycles (8 bp)	Plate barcode
Index 2	8 cycles (8 bp)	N/A
Read 2	At least 50 cycles (50 bp)	cDNA reads

3.9 Data Processing and Quality Control

1. Once the sequencing is done, two fastq files are returned per plate. In this example, single cells from EBs at day 4 of differentiation were profiled. For example, in this protocol, the file names associated with the experiments are "mEB_day4_r1.fq.gz" (Read 1 file) and "mEB_day4_r2.fq.gz" (Read 2 file).
2. STARsolo [34] can be used to process the data to get the gene expression matrix, containing the UMI count for each gene in each cell (*see* **Note 14**).
3. Build the genome index with STAR by running the following command (*see* **Note 15**):

```
STAR --runThreadN 20 --runMode genomeGenerate --genomeDir
<star_index> --genomeFastaFiles <genome.fa> --sjdbGTFfile
<genes.gtf>
```

4. Get the count matrix by running the following command (*see* **Note 16**):

```
STAR --genomeDir <star_index> --readFilesCommand zcat --read-
FilesIn mEB_day4_r2.fq.gz mEB_day4_r1.fq.gz --soloCBstart 1
--soloCBlen 8 --soloUMIstart 9 --soloUMIlen 10 --soloType
CB_UMI_Simple --soloCBwhitelist whitelist.csv --runThreadN
20 --outSAMattributes CB UB --outSAMtype BAM SortedBy
Coordinate
```

5. Once the program finishes, a BAM file called "Aligned.sortedByCoord.out.bam" and a directory named "Solo.out" should appear (*see* **Note 17**). The gene expression matrix is inside the "Solo.out" directory and can be easily analysed by other scRNA-seq analysis package, such as Seurat [35] and Scanpy [36].
6. Plot the some quality control related metrics (Fig. 8a, b), perform dimensionality reduction using UMAP [37] (Fig. 8c), carry out cell clustering using the Leiden algorithm [38] (Fig. 8c) and identify marker genes of each cluster. Markers of three germ layers, such as *Cdx2*, *Gata2* and *Gata4*, should start expressing (Fig. 8d). *See* **Note 18** for a detailed comments.

4 Notes

1. We find slow evaporation still happens in −80 °C, possibly due to the opening and closing of the freezer door. Make sure to centrifuge the plate to collect evaporation every time before use. When removing the seal, be very careful to avoid splash and cross-well contaminations.
2. We routinely use 384-well plates as units for experiments, but 96-well plates also work. The lysis plates can be prepared in bulk in advance and stored in −80 °C. This saves time on the experimental day. For a pilot experiment to test the protocol, you do not need to perform a full plate. 12 wells should be more than enough to test if the experiments work or not.
3. It is very important to make sure the drops are all in the centre of those 24 wells. During the actual sorting, we use FSC-A vs SSC-A to gate out the cell debris, and use FSC-A vs FSC-W to remove doublets. Finally, we only sort DAPI negative live single cells (gate P3 in Fig. 3). In most places, FACS machines are operated by specialists. Check with your local FACS experts for the gating strategy and sorting accuracy adjustment.
4. At this stage, the cells are in lysis buffer. Although it is safe to store the plate in −80 °C, we still recommend proceed to the

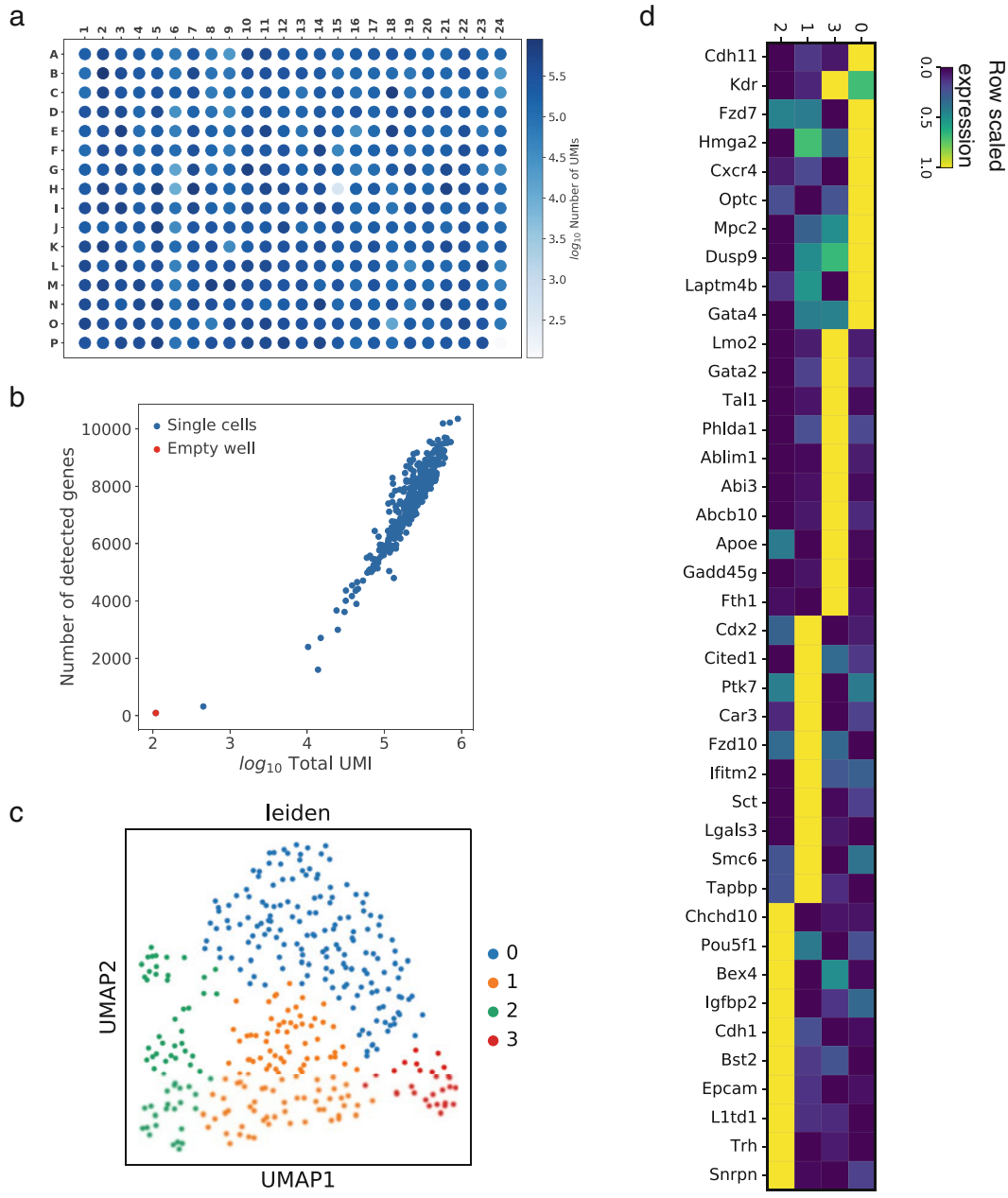

Fig. 8 Results of single cells from day 4 EBs from the sequencing data. (**a**) The number of UMIs (in log scale) in each well on a plate. Note there are very few UMIs in the empty well (P24). (**b**) Scatter plot showing the number of UMIs (log) and the number of detected genes per well. Note the empty wells and a few failed single cells have orders of magnitude difference in terms of UMIs comparing to the majority of single cells. (**c**) UMAP representation of each single cell. The cells are colored by the Leiden cluster. (**d**) Row scaled expression of top 10 marker genes in each cluster

next stage as soon as possible. When you have multiple plates for sorting, put the plates with cells on dry ice while waiting for other plates to finish sorting. Put all plates together into −80 °C when all of them are done sorting.

5. At this stage, the 3′ end of the cDNA is tagged by the well barcodes, but the plate barcode has not been incorporated yet. Therefore, you can only pool wells from the same plate. Different plate pools should be kept separately.
6. This step is critical, as primer leftover results in the mispriming of oligos from different wells in the subsequent PCR reaction. Digestion of the excess of primers eliminate index swapping in the final library amplification.
7. There are quite a few choices of the beads, and we have successful experience with AmpureXP beads and VAHTS DNA clean beads.
8. With mESCs, we normally get hundreds of nanograms of cDNA, which is more than enough to make the final library. Cells with low RNA content might yield less cDNA. When the total cDNA is less than 10 ng, a few more PCR cycles (**step 5** from Subheading 3.6) can be done using the **cDNA PCR Fwd/Rev Primers**. If the amount of cDNA is between 10 ng and 50 ng, the tagmentation reaction can be proportionally scaled down in the **step 1** from Subheading 3.8.
9. In the profile from the Bioanalyzer, the main peak of cDNA should be above 1000 bp. The two successful examples shown in Fig. 6 represent typical cDNA size profiles. The exact amount of cDNA varies across cell types and experiments, but intact cDNA should always have a major peak larger than 1000 bp. There should be nearly flat or very few peaks with low intensity below 1000 bp. No visible peaks means failure of cDNA amplification. Many small peaks below 1000 bp indicates RNA degradation. Two examples of failed ones are shown in Fig. 6.
10. This is an important step where cDNA is "tagmented," where the Tn5 transposases cut the cDNA and paste the partial sequencing adaptors to cDNA. The key component here is the Tn5 transposase. There are a few choices for the Tn5 Transposase. It can be purchased from Illumina (Illumina Tagment DNA TDE1 Enzyme and Buffer Kit, cat. no. 20034197). You can also get it from other vendors, such as Fapon (cat. no. NK001) or Vazyme (TD-501). One can also makes the Tn5 transposase in house following the procedure described by Picelli et al. [39].
11. The Library Index PCR Primer is basically Illumina's Nextera XT Index PCR Primer. This introduces plate barcode at the right-hand side of the library. Use different primers for

different plates if you intend to pool multiple plates to sequence together.

12. A typical library profile should be a smooth bell shaped curve with the peak around 300–400 bp. Small sharp peak at 200 bp indicates over-tagmentation, and large peak above 700 bp suggest under-tagmentation. The most efficient way of solving the problem is to adjust the amount of Tn5 accordingly. In general, we found under-tagmentation is not a big problem and it still produces very successful sequencing results.
13. In most places, sequencing is done in a genomic core facility. Talk to your local sequencing specialist for the library requirement. The exact sequencing mode will depend on the machine and whether you are sequencing alone or together with libraries from others. In general, the first read (Read 1) contain 8 bp well barcode and 10 bp UMI at the beginning. Therefore, you need at least 18 cycles for Read 1. Index 1 is the plate index that has 8 bp in length. You need 8 cycles here. Index 2 is optional in the current protocol due to the absence of an index at the left-hand side. Read 2 is the actual mRNA read, at least 50 bp are needed, and we normally perform 75 cycles (75 bp) for Read 2.
14. To convert the fastq files to gene expression matrix, a flexible program that allows you to specify the position of barcodes and UMIs is very important. Both STARsolo [34] and kallisto-bustools [40] can do this. In this protocol, STARsolo is demonstrated as it provides the alignment files which might be useful for other purposes.
15. The command described here should be in one single line. The files <genome.fa> and <genes.gtf> can be obtained from the UCSC genome browser [41] and GENCODE [42], respectively. The number of CPUs to run the job can be specified by the "`--runThreadN`" option.
16. The command described here should be in one single line. After the "`--readFilesIn`" flag, two fastq files should be provided. The first fastq file should contain the mRNA reads, and the second fastq file the barcode and UMIs. In our method, Read 1 contains the barcode and UMIs, and Read 2 the mRNA reads. Therefore, in the example, Read 2 file "`mEB_day4_r2.fq.gz`" appears before the Read 1 file "`mEB_day4_r1.fq.gz`". The "whitelist.csv" file is a simple text file which basically contains the 8-bp well barcode in the **Barcoded RT Primers**, one barcode per line. *See* Table 1 for the sequence information.
17. In the "Solo.out" directory, there are a few text files that contain some basic quality control information. The files inside "Gene/raw" and "Gene/filtered" contain expression values of

each gene in each cell. The files in the "Gene/raw" directory contain the gene expression of every barcode in the whitelist. The files in the "Gene/filtered" contain similar information with barcodes that have too few reads removed.

18. Simple plots of some basic metrics, such as number total UMIs and number detected genes, like shown in Fig. 8a, b are useful to check if the technique works or not. When plotting the total number of UMIs in a 384-well plate layout (Fig. 8a), the empty well at the bottom right (P24) should have extremely faint colour comparing to other wells. If pipetting or sorting error happens, clear patterns will be visible; for example, some rows or columns have very few UMIs. When draw a scatter plot like shown in Fig. 8b, the empty well should have orders of magnitude fewer reads comparing to the cells. Failed cells also have very few UMIs.

Acknowledgments

We thank all members from the Chen lab and the Zhang lab for the helpful comments on the manuscript. The work was supported by Shenzhen Science And Technology Innovation Committee (ZDSYS20200811144002008) and National Natural Science Foundation of China (31970812). The computational work was supported by Center for Computational Science and Engineering at Southern University of Science and Technology.

References

1. Smith AG (2001) Embryo-derived stem cells: of mice and men. Cell Dev Biol 17:435–462. https://doi.org/10.1146/annurev.cellbio.17.1.435
2. Evans MJ, Kaufman MH (1981) Establishment in culture of pluripotential cells from mouse embryos. Nature 292:154–156. https://doi.org/10.1038/292154a0
3. Martin GR (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A 78:7634–7638. https://doi.org/10.1073/pnas.78.12.7634
4. Ying Q-L, Wray J, Nichols J et al (2008) The ground state of embryonic stem cell self-renewal. Nature 453:519–523. https://doi.org/10.1038/nature06968
5. Doetschman TC, Eistetter H, Katz M et al (1985) The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. J Embryol Exp Morph 87:27–45. https://doi.org/10.1242/dev.87.1.27
6. Chen X, Teichmann SA, Meyer KB (2018) From tissues to cell types and back: single-cell gene expression analysis of tissue architecture. Annu Rev Biomed Data Sci 1:1–23. https://doi.org/10.1146/annurev-biodatasci-080917-013452
7. Clark SJ, Lee HJ, Smallwood SA et al (2016) Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity. Genome Biol 17:72. https://doi.org/10.1186/s13059-016-0944-x
8. Wang Y, Navin NE (2015) Advances and applications of single-cell sequencing technologies. Mol Cell 58:598–609. https://doi.org/10.1016/j.molcel.2015.05.005
9. Tang F, Barbacioru C, Wang Y et al (2009) mRNA-seq whole-transcriptome analysis of a single cell. Nat Methods 6:377–382. https://doi.org/10.1038/nmeth.1315

10. Tang F, Barbacioru C, Bao S et al (2010) Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-seq analysis. Cell Stem Cell 6:468–478. https://doi.org/10.1016/j.stem.2010.03.015
11. Macosko EZ, Basu A, Satija R et al (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 161:1202–1214. https://doi.org/10.1016/j.cell.2015.05.002
12. Klein AM, Mazutis L, Akartuna I et al (2015) Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 161:1187–1201. https://doi.org/10.1016/j.cell.2015.04.044
13. Gierahn TM, Wadsworth MH, Hughes TK et al (2017) Seq-well: portable, low-cost RNA sequencing of single cells at high throughput. Nat Methods 14:395–398. https://doi.org/10.1038/nmeth.4179
14. Han X, Wang R, Zhou Y et al (2018) Mapping the mouse cell atlas by Microwell-seq. Cell 172:1091–1107.e17. https://doi.org/10.1016/j.cell.2018.02.001
15. Zheng GXY, Terry JM, Belgrader P et al (2017) Massively parallel digital transcriptional profiling of single cells. Nat Commun 8:14049. https://doi.org/10.1038/ncomms14049
16. Jaitin DA, Kenigsberg E, Keren-Shaul H et al (2014) Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science 343:776–779. https://doi.org/10.1126/science.1247651
17. Islam S, Kjällquist U, Moliner A et al (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res 21:1160–1167. https://doi.org/10.1101/gr.110882.110
18. Ramsköld D, Luo S, Wang Y-C et al (2012) Full-length mRNA-seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol 30:777–782. https://doi.org/10.1038/nbt.2282
19. Picelli S, Björklund ÅK, Faridani OR et al (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods 10:1096–1098. https://doi.org/10.1038/nmeth.2639
20. Hagemann-Jensen M, Ziegenhain C, Chen P et al (2020) Single-cell RNA counting at allele and isoform resolution using Smart-seq3. Nat Biotechnol 38:708–714. https://doi.org/10.1038/s41587-020-0497-0
21. Sasagawa Y, Danno H, Takada H et al (2018) Quartz-Seq2: a high-throughput single-cell RNA-sequencing method that effectively uses limited sequence reads. Genome Biol 19:29. https://doi.org/10.1186/s13059-018-1407-3
22. Hashimshony T, Senderovich N, Avital G et al (2016) CEL-Seq2: sensitive highly-multiplexed single-cell RNA-seq. Genome Biol 17:77. https://doi.org/10.1186/s13059-016-0938-8
23. Bagnoli JW, Ziegenhain C, Janjic A et al (2018) Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. Nat Commun 9:2937. https://doi.org/10.1038/s41467-018-05347-6
24. Cao J, Packer JS, Ramani V et al (2017) Comprehensive single-cell transcriptional profiling of a multicellular organism. Science 357:661–667. https://doi.org/10.1126/science.aam8940
25. Rosenberg AB, Roco CM, Muscat RA et al (2018) Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science 360:eaam8999. https://doi.org/10.1126/science.aam8999
26. Scialdone A, Tanaka Y, Jawaid W et al (2016) Resolving early mesoderm diversification through single-cell expression profiling. Nature 535:289–293. https://doi.org/10.1038/nature18633
27. Mohammed H, Hernando-Herraez I, Savino A et al (2017) Single-cell landscape of transcriptional heterogeneity and cell fate decisions during mouse early gastrulation. Cell Rep 20:1215–1228. https://doi.org/10.1016/j.celrep.2017.07.009
28. Cao J, Spielmann M, Qiu X et al (2019) The single-cell transcriptional landscape of mammalian organogenesis. Nature 566:496–502. https://doi.org/10.1038/s41586-019-0969-x
29. Pijuan-Sala B, Griffiths JA, Guibentif C et al (2019) A single-cell molecular map of mouse gastrulation and early organogenesis. Nature 566:490–495. https://doi.org/10.1038/s41586-019-0933-9
30. Kolodziejczyk AA, Kim JK, Tsang JCH et al (2015) Single cell RNA-sequencing of pluripotent states unlocks modular transcriptional variation. Cell Stem Cell 17:471–485. https://doi.org/10.1016/j.stem.2015.09.011
31. Gao X, Nowak-Imialek M, Chen X et al (2019) Establishment of porcine and human expanded potential stem cells. Nat Cell Biol 21:687–699. https://doi.org/10.1038/s41556-019-0333-2
32. Spangler A, Su EY, Craft AM, Cahan P (2018) A single cell transcriptional portrait of embryoid body differentiation and comparison to

progenitors of the developing embryo. Stem Cell Res 31:201–215. https://doi.org/10.1016/j.scr.2018.07.022

33. Kim IS, Wu J, Rahme GJ et al (2020) Parallel single-cell RNA-seq and genetic recording reveals lineage decisions in developing embryoid bodies. Cell Rep 33:108222. https://doi.org/10.1016/j.celrep.2020.108222
34. Kaminow B, Yunusov D, Dobin A (2021) STARsolo: accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data. Biorxiv 2021.05.05.442755. https://doi.org/10.1101/2021.05.05.442755
35. Satija R, Farrell JA, Gennert D et al (2015) Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 33:495–502. https://doi.org/10.1038/nbt.3192
36. Wolf FA, Angerer P, Theis FJ (2018) SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 19:15. https://doi.org/10.1186/s13059-017-1382-0
37. Becht E, McInnes L, Healy J et al (2019) Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol 37:38–44. https://doi.org/10.1038/nbt.4314
38. Traag VA, Waltman L, van Eck NJ (2019) From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep 9:5233. https://doi.org/10.1038/s41598-019-41695-z
39. Picelli S, Björklund ÅK, Reinius B et al (2014) Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. Genome Res 24:2033–2040. https://doi.org/10.1101/gr.177881.114
40. Melsted P, Booeshaghi AS, Liu L et al (2021) Modular, efficient and constant-memory single-cell RNA-seq preprocessing. Nat Biotechnol 39(7):1–6. https://doi.org/10.1038/s41587-021-00870-2
41. Kent WJ, Sugnet CW, Furey TS et al (2002) The human genome browser at UCSC. Genome Res 12:996–1006. https://doi.org/10.1101/gr.229102
42. Frankish A, Diekhans M, Ferreira A-M et al (2018) GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res 47(D1):D766–D773. https://doi.org/10.1093/nar/gky955