HAL
open science

# Single-Cell Analysis and Tracking of Antigen-Specific T Cells: Integrating Paired Chain AIRR-Seq and Transcriptome Sequencing: A Method by the AIRR Community

Nidhi Gupta, Ida Lindeman, Susanne Reinhardt, Encarnita Mariotti-Ferrandiz, Kevin Mujangi-Ebeka, Kristen Martins-Taylor, Anne Eugster

## To cite this version:

Nidhi Gupta, Ida Lindeman, Susanne Reinhardt, Encarnita Mariotti-Ferrandiz, Kevin Mujangi-Ebeka, et al.. Single-Cell Analysis and Tracking of Antigen-Specific T Cells: Integrating Paired Chain AIRR-Seq and Transcriptome Sequencing: A Method by the AIRR Community. Immunogenetics, 2453, Springer US, pp.379-421, 2022, Methods in Molecular Biology, 10.1007/978-1-0716-2115-8_20 . hal-03838013

**HAL Id: hal-03838013**

**https://hal.sorbonne-universite.fr/hal-03838013**

Submitted on 6 Apr 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Chapter 20

# Single-Cell Analysis and Tracking of Antigen-Specific T Cells: Integrating Paired Chain AIRR-Seq and Transcriptome Sequencing: A Method by the AIRR Community

**Nidhi Gupta, Ida Lindeman, Susanne Reinhardt, Encarnita Mariotti-Ferrandiz, Kevin Mujangi-Ebeka, Kristen Martins-Taylor, and Anne Eugster**

## Abstract

Single-cell adaptive immune receptor repertoire sequencing (scAIRR-seq) offers the possibility to access the nucleotide sequences of paired receptor chains from T-cell receptors (TCR) or B-cell receptors (BCR). Here we describe two protocols and the downstream bioinformatic approaches that facilitate the integrated analysis of paired T-cell receptor (TR) alpha/beta (TRA/TRB) AIRR-seq, RNA sequencing (RNAseq), immunophenotyping, and antigen-binding information. To illustrate the methodologies with a use case, we describe how to identify, characterize, and track SARS-CoV-2-specific T cells over multiple time points following infection with the virus. The first method allows the analysis of pools of memory $CD8^+$ cells, identifying expansions and contractions of clones of interest. The second method allows the study of rare or antigen-specific cells and allows studying their changes over time.

**Key words** Single-cell sequencing, TR gene, IG gene, Rearrangement, Transcriptome, 10x Genomics, SMART-seq, Multi-omic analysis

## 1 Introduction

Single-cell adaptive immune receptor repertoire sequencing (scAIRR-seq) aims at describing the sequences of T-cell receptor (TR) or immunoglobulin (IG) rearrangements at the single-cell level. scAIRR-seq has been used since the mid-1990s [1, 2] and has seen rapidly increasing adoption by the scientific community over the last few years [3]. This has been facilitated by a plethora of protocols, commercial kits, and platforms as well as by the associated software tools developed in the last decade as discussed in detail in the two AIRR Community commentary chapters in this

Nidhi Gupta, Ida Lindeman, and Susanne Reinhard are shared first authors.

Anton W. Langerak (ed.), *Immunogenetics: Methods and Protocols*, Methods in Molecular Biology, vol. 2453, https://doi.org/10.1007/978-1-0716-2115-8_20, © The Author(s) 2022

volume. Workflows for scAIRR-seq distinguish themselves from bulk methods by several features: First, they preserve the chain-pairing information of the complete IG/TR, which is critical for the experimental reconstruction and measurement of receptor reactivities. Second, they allow an unbiased view of clonal expansion, as observed sequences can be attributed to and normalized by the individual cell from which they originate. Third, using the individual cell as a common reference point, they allow the integration with other single-cell resolution data, such as transcriptome, cell-surface phenotype as well as antigen-specificities. All of this, however, comes at a reduced throughput and lower sensitivity for the detection of rare clones when compared to bulk sequencing.

The vast majority of currently used scAIRR-seq methods, including the two presented here (Fig. 1), maintains the compartmentalization of the cells throughout the process, either physically or via barcoding [4]. The 10x Genomics Chromium is a microfluidic-based platform, which allows the encapsulation and barcoding of up to 3000 to 10,000 cells at a time. The Chromium Next GEM Single Cell V(D)J Reagent Kits described here allow the generation of three libraries (Fig. 2): (1) full-length, paired AIRR sequences, (2) the cell transcriptome (derived from all polyadenylated transcripts), and (3) feature barcodes linked to the surface protein expression and antigen specificity (e.g., CITE-seq) as well as barcoding of libraries for multiplexing (e.g., hash-tagging). Single-cell SMART-seq is a method to collect single-cell AIRR-seq and gene expression data from cells which are sorted into 96-well plates (Fig. 3 and Fig. 4), allowing the analysis of rare cells. Single-cell SMART-seq is based on the SMART® (switching mechanism at 5′ end of RNA template) technology. The SMART-Seq Single Cell Kit used to generate mRNA-seq libraries is particularly useful for the analysis of cells with very low RNA content, such as PBMCs.

In both methods AIRR-seq as well as transcriptome sequences are obtained from RNA, making use of the template-switching activity of reverse transcriptase to enrich for full-length cDNAs and to add defined PCR adapters directly to both ends of the first-strand cDNA. This ensures that the final cDNA libraries contain the 5′ end of the mRNA and maintain a true representation of the original mRNA transcripts. These factors are critical for AIRR-seq and transcriptome sequencing.

A use case for these two methods is to identify, characterize, and track SARS-CoV-2- or other virus-specific T cells over multiple time points following infection, remittance, or vaccination. The Chromium Next GEM Single Cell V(D)J method can be used to broadly analyze thousands of memory $CD8^+$ cells from several time points of an individual, before and after the immunizing event, in a multiplexed manner (through hash-tagging). This allows identifying expansions and contractions of clones of interest (through AIRR-seq), phenotyping them (through transcriptome analysis

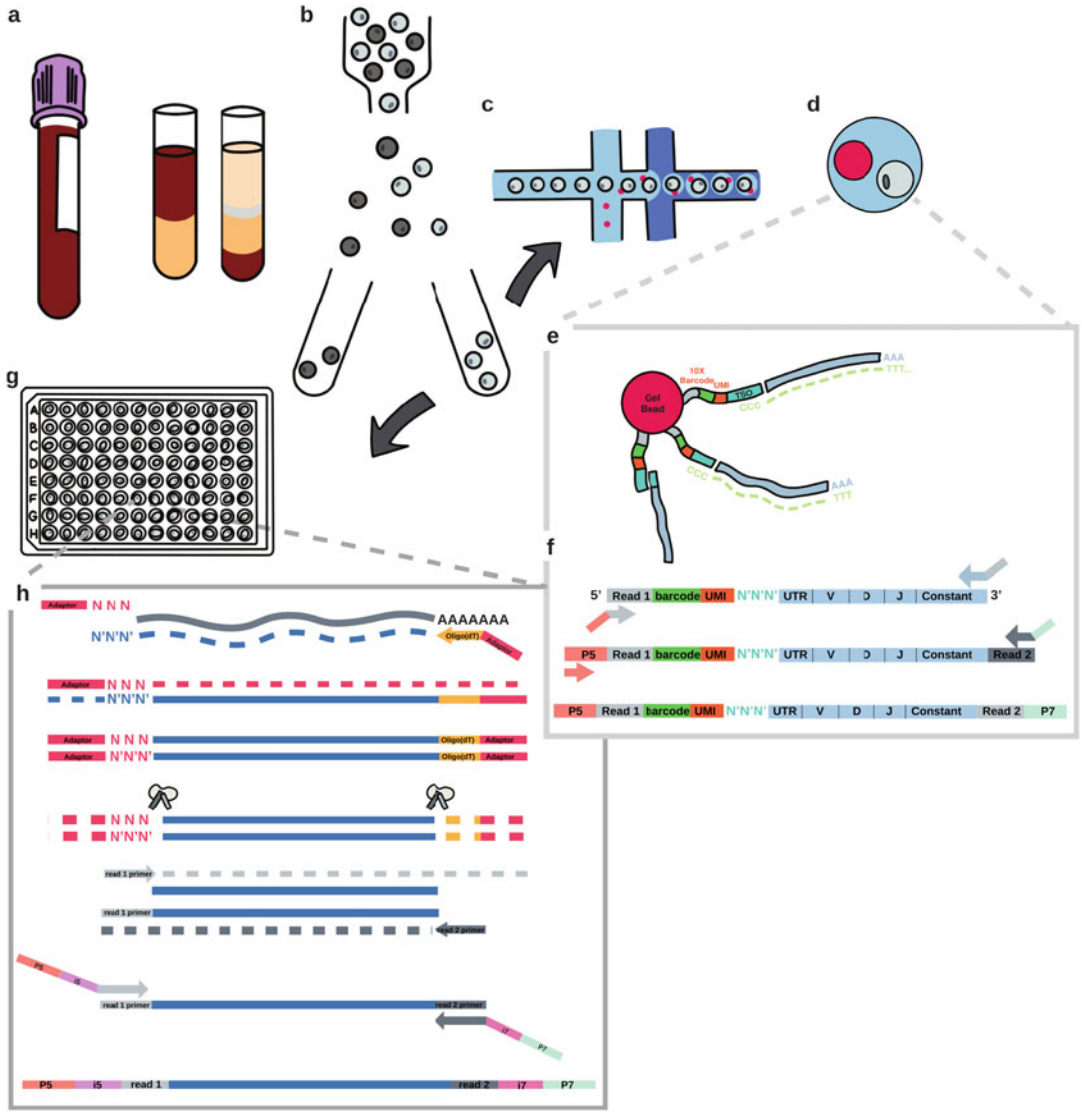


**Fig. 1** Schematic illustrating the main characteristics of single cell paired chain AIRR-seq. (**a**) A blood sample is processed by Ficoll gradient centrifugation to obtain PBMC. (**b**) PBMCs are stained and cells of interest sorted by FACS. As described in Subheading 3.1, using the 10x Genomics fluidics system (panel **c**), cells are processed for transcript barcoding. After encapsulation in a droplet, a "GEM" is created (panel **d**). Gel bead primers containing a 10× barcode, a UMI, and a template switch oligo (TSO) bind to the transcripts after cell lysis (panel **e**). Gel bead primers also capture the cell surface feature barcodes. Barcoded transcripts and feature barcodes are then reverse transcribed, and through size selection and enrichment, a library containing amplified AIRR (panel **f**), a library containing whole cell transcriptome, and one containing feature barcodes are prepared through the sequential additions of primers containing the P5 and P7 sites required for sequencing. As described in Subheading 3.1, cells are deposited into a plate (panel **g**). Here reverse transcription, cDNA amplification, and preparation of a library containing amplified AIRR and of a library containing whole cell transcriptome take place through the sequential additions of primers that include the P5 and P7 sites required for sequencing by the SMARTseq method (panel **h**)

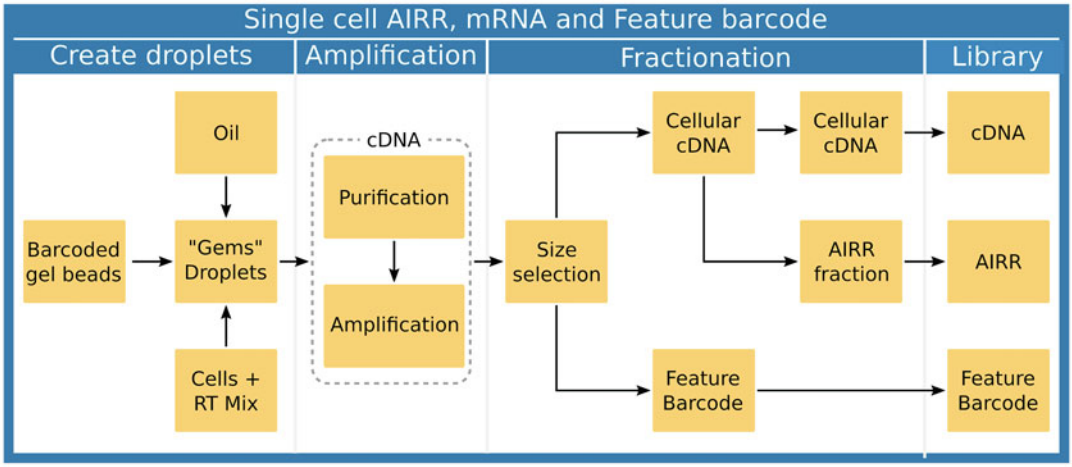


**Fig. 2** Overview of the main steps of the Chromium Next GEM Single-Cell procedure: The creation of droplets (left), in which the RNA is captured and barcoded is followed by breaking the GEMs, the amplification of cDNA, the fractionation that allows separation of cDNA from feature barcodes from cellular cDNA, and finally by the preparation of the three libraries

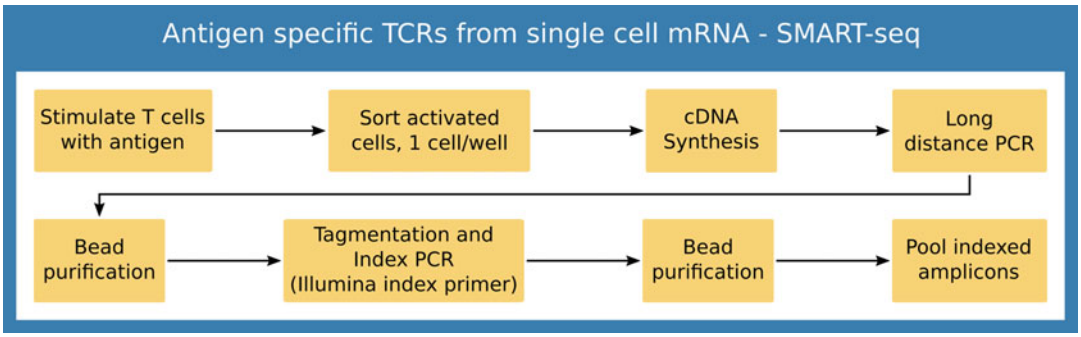


**Fig. 3** Overview of main steps for SMART-seq scTCR procedure: Stimulated cells are sorted into PCR plates, followed by cDNA synthesis, two rounds of PCR and purification, and are finally pooled to prepare the library for sequencing

and feature barcode analysis), and correlating them to disease status. Clones or cells of interest can be defined through their activation state or their antigen specificity and are isolated by flow cytometry after surface marker or multimer staining, respectively. Clones or cells of interest can be activated $CD8^+CD25^+CD137^+$ T cells [5] from several time points during and after a viral infection, or cells stimulated with an antigen of interest. Single-cell SMART-seq of these often rare clones after isolation gives access to their AIRR data that can then be matched to the data obtained from Chromium Next GEM Single Cell V(D)J. Here we provide protocols and detailed information for the generation, processing, and analysis of scAIRR-seq- and associated data produced with the two platforms described (Fig. 5).

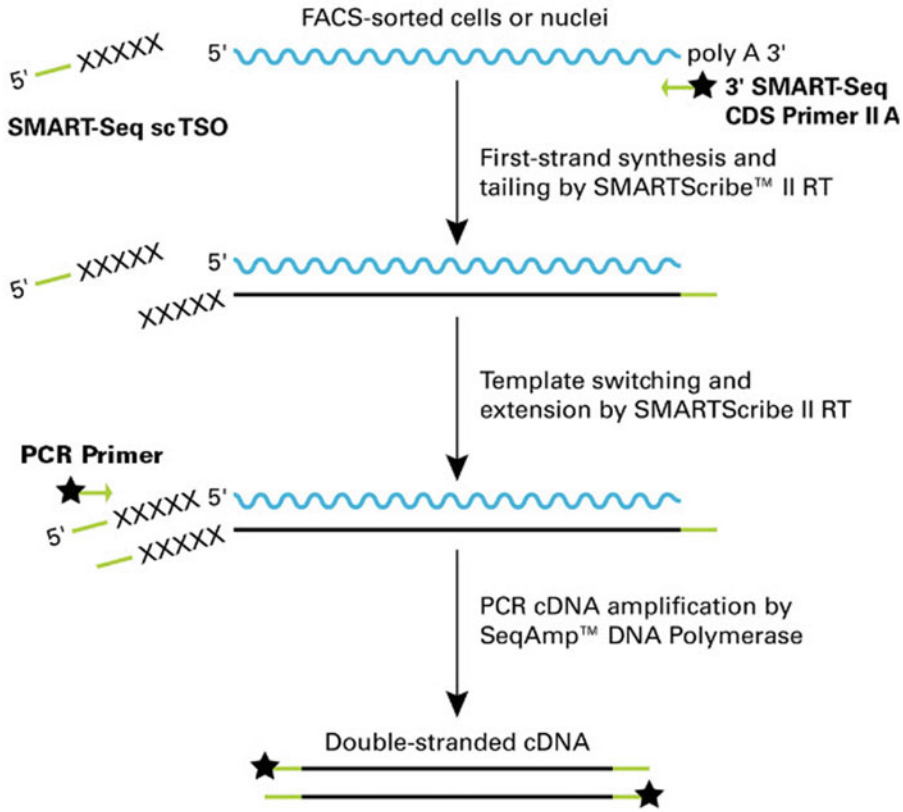


**Fig. 4** Schematic of the technology in the SMART-Seq Single-Cell Kit. Non-templated nucleotides (indicated by Xs) added by the SMARTScribe II reverse transcriptase (RT) hybridize to the SMART-Seq single-cell template-switching oligonucleotide (SMART-Seq sc TSO), which provides a new template for the RT. The SMART adapters used for amplification during PCR added by the oligo(dT) primer (3′ SMART-Seq CDS Primer II A) and TSO are indicated in green. Chemical modifications to block ligation (if using a ligation-based library preparation method) are present on some primers (indicated by black stars)

## 2 Materials

### *2.1 10x Genomics Chromium Next GEM Single-Cell V(D)J Kit*

1. Single-channel pipette: 10 μl, 20 μl, 200 μl, and 1000 μl.
2. 8-channel or 12-channel pipette (recommended): 20 μl and 200 μl.
3. Filter pipette tips: 2 μl, 20 μl, 200 μl, and 1000 μl.
4. Minicentrifuge for 1.5-ml tubes.
5. Minicentrifuge for 0.2-ml tubes or strip.
6. Bioanalyzer or TapeStation for library validation.
7. All the related cell-sorting equipment.

cDNA Synthesis and Amplification

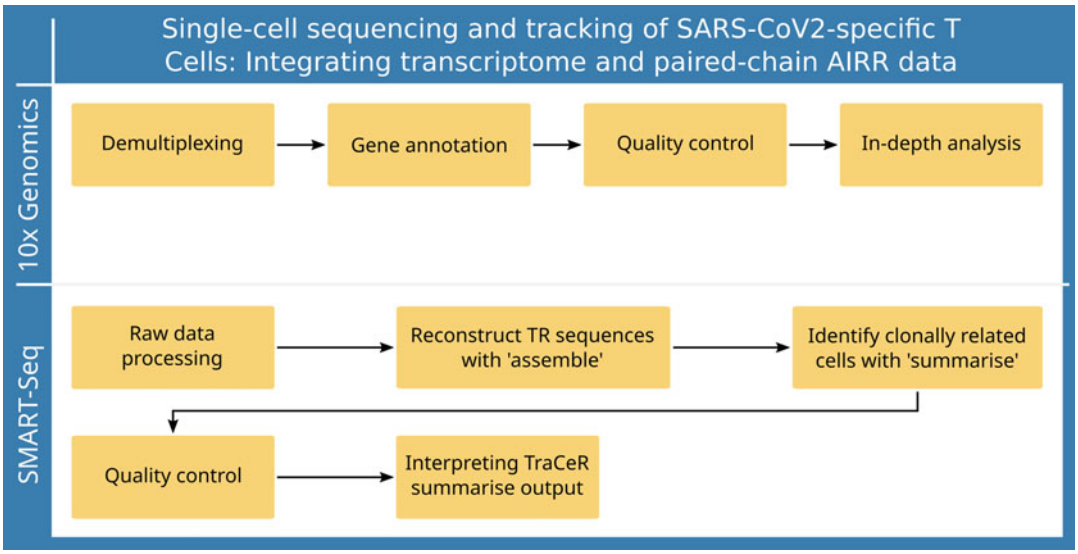


**Fig. 5** Overview of the main steps of the analysis of single-cell AIRR-seq data. Libraries created with the 10x Genomics technology (upper panel) are processed using the CellRanger software. In brief, sequencing libraries are demultiplexed before TR sequences are extracted and annotated. The quality of each library is assessed, and TR sequences may be combined with transcriptional and feature libraries for an in-depth integrated analysis. For libraries created by plate-based sequencing technologies such as SMART-seq (lower panel), TR sequences are computationally reconstructed with TraCeR. Low-quality cells or potential duplets may be filtered out, before clonally related cells are identified and visualized in clonal networks

8. Two thermal cyclers with heated lids (*see* **Note 1**): one thermal cycler used only for first-strand cDNA synthesis; one thermal cycler used only for double-stranded cDNA amplification by PCR.
9. 96-well semi-skirted plates (Thermo Fisher Scientific) or 8-tube strips (Thermo Fisher Scientific).
10. Thermo Scientific Adhesive PCR Plate Seals (Thermo Fisher Scientific) for 96-well plates or flat cap strips (Thermo Fisher Scientific) for 96-well plates or 8-tube strips.

10x Genomics Kits and Reagents (10X Genomics, Unless Mentioned)

10. Chromium Next GEM Single Cell 5′ Library and Gel Bead Kit v1.1, 16 rxns.
11. DynaBeads® MyOne™ Silane Beads (Thermo Fisher Scientific).
12. Chromium Single Cell 5’ Library Construction Kit, 16 rxns.
13. Chromium Single Cell 5’ Feature Barcode Library Kit, 16 rxn.

14. Chromium™ Single Cell V(D)J Enrichment Kit, human T cell/mouse T cell.
15. Chromium™ Single Cell V(D)J Enrichment Kit, human B cell/mouse B cell.
16. Chromium Next GEM Chip G Single Cell Kit, six chips.
17. Single Index Kit N Set A, Single Index Kit T Set A.

Other Supplies

18. Nuclease-free water.
19. Low TE buffer: 10 mM Tris–HCl pH 8.0, 0.1 mM EDTA.
20. Ethanol, pure (200 Proof, anhydrous).
21. Tween 20.
22. Glycerin (glycerol), 50% (v/v) aqueous solution.
23. Elution buffer (EB).
24. NGS HS Fragment Analysis Kit (Agilent) or comparable chemistry to run QC.
25. Bovine serum albumin (BSA).
26. SPRIselect Reagent Kit (Beckman Coulter).

Bead Purifications

27. NucleoMag NGS cleanup and size select (*see* **Note 2**) Takara Bio) or the AMPure XP PCR Purification Kit (Beckman Coulter).

For cDNA and Illumina Library Quantification and Preparation

28. High-Sensitivity DNA Kit (Agilent) for bioanalyzer or equivalent high-sensitivity electrophoresis method (may be used in Sections V.D and VI.D).
29. Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific) or Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific).
30. Library Quantification Kit (Takara Bio).
31. Nuclease-free, PCR-grade, thin-wall PCR strips or 96-well plates.
32. Nuclease-free, low-adhesion 1.5-ml tubes.
33. Benchtop cooler, such as VWR CryoCoolers.

Cell Preparation

34. Benzonase (10 U/ml).

### 2.2 Single-Cell SMART-Seq

The same equipment and supplies are used as for 10× Chromium, except for the following:

General Lab Equipment

1. 96-well PCR chiller rack, such as IsoFreeze PCR Rack, or 96-well aluminum block.

Sample Preparation

2. Nuclease-free, PCR-grade 8-tube strips secured in a PCR rack, or 96-well plates that have been validated to work with your FACS instrument.
3. Microplate film for sealing tube strips/plates before sorting.
4. Aluminum single-tab foil seal or cap strips for sealing tubes/plates after sorting.
5. Low-speed benchtop centrifuge for 96-well plates or tube strips.
6. Dry ice in a suitable container for flash freezing cell.

Bead Purifications

7. 80% ethanol: freshly made for each experiment from molecular-biology-grade 100% ethanol.
8. Strong magnetic separation device for plates that accommodates 96 samples in 96-well V-bottom plates (500 μl; VWR).
9. Low-speed benchtop centrifuge for a 96-well plate.
10. Strong magnetic separation device for tubes.
11. Nuclease-free, PCR-grade 8-tube strips secured in a PCR rack, or 96-well plates.

Sequencing Library Generation

12. Nextera XT DNA Library Preparation Kit (Illumina).
13. Nextera XT Index Kit (Illumina) or other Nextera-compatible indexes.
14. 96-well PCR plate.
15. Microseal "B" adhesive seals.
16. Freshly prepared 80% ethanol (EtOH), as above.
17. 96-well 0.8-ml polypropylene deep-well storage plate (midi plate).
18. Microseal "F" foil seals.
19. Nuclease-free water.

cDNA Synthesis (Takara Bio Unless Otherwise Specified)

20. 1 μg/μl control total RNA.
21. 10× lysis buffer.
22. 40 U/μl RNase inhibitor.
23. Nuclease-free water.
24. SMART-Seq sc TSO.
25. 3′ SMART-Seq CDS Primer II A.

26. SMARTScribe II reverse transcriptase.
27. SMART-Seq sc First-Strand Buffer (5×).
28. SeqAmp DNA polymerase.
29. 2× SeqAmp CB PCR buffer.
30. PCR primer.
31. 10 mM Tris-Cl elution buffer (pH 8.5).

Nextera Library Preparation (Illumina Unless Otherwise Mentioned)

32. Amplicon Tagment Mix (ATM).
33. Tagment DNA Buffer (TD).
34. Neutralize Tagment Buffer (NT).
35. Nextera PCR Master Mix (NPM).
36. Resuspension Buffer (RSB).

Nextera Indices

37. Index 1 (i7) Adapters (N7xx—Nextera XT Index Kit v2, Nextera Index Kit): N701; TCGCCTTA, N702; CTAGTACG, N703; TTCTGCCT, N704; GCTCAGGA, N705; AGGAGTCC, N706; CATGCCTA, N707; GTAGAGAG, N710; CAGCCTCG, N711; TGCCTCTT, N712; TCCTCTAC, N714; TCATGAGC, N715; and CCTGAGAT (i7 index name; bases in adapter).
38. Index 2 (i5) adapters (S5xx—Nextera XT Index Kit v2): CTCTCTAT, S503; TATCCTCT, S505; GTAAGGAG, S506; ACTGCATA, S507; AAGGAGTA, S508; CTAAGCCT, S510; CGTCTAAT, S511; and TCTCTCCG (i5 index name; bases in adapter S502).

### 2.3 10x Genomics Data Processing and Analysis

1. Cell Ranger software, provided by 10x Genomics free of charge, required for raw data handling and various secondary analysis.
2. Linux workstation (minimum 8 cores, 64 GB RAM, 1 TB storage) running a recent version of CentOS/RHEL or Ubuntu, to run Cell Ranger.
3. Loupe Brower and Loupe VDJ Browser, also available via the 10x Genomics website, which provide a complementary set of analysis tools.
4. Windows or Macintosh operating systems to run Loupe Browser.

### 2.4 Single-Cell SMART-Seq Data Processing and Analysis

1. TraCeR, installed as a Docker container, with all of its dependencies installed and properly configured. Alternatively, TraCeR can be installed from GitHub.
2. Dependencies, installed and configured by the user (see https://github.com/Teichlab/tracer), (if installing TraCer from GitHub).

## 3 Methods

### 3.1 10x Genomics Chromium Next GEM Single-Cell V(D)J Kit

Before starting, please refer to considerations regarding the kits used (*see* **Note 3**), sample multiplexing (*see* **Note 4**), and surface protein detection (*see* **Note 5**).

#### 3.1.1 Coat Tubes for Cell Sort and Count Cells

1. Fill PCR tubes with 1% BSA in PBS and incubate them overnight. Remove the BSA completely by shortly centrifuging the tubes after the first removal and removing the collected liquid. Prepare one tube for each sample and pre-lye 1 μl PBS.
2. After staining and washing (*see* **Note 6**), sort cells into the tube (*see* **Note 7**), and use 1–2 μl from the sample to verify the cell quality and number (*see* **Note 8**) under a light microscope. Proceed to loading the chip taking into account the time required for the sort (*see* **Note 9**).

#### 3.1.2 Load Next GEM Chip G

To avoid contamination, this section should be carried out on a separate bench dedicated to RNA/cell work.

1. Equilibrate Single-Cell VDJ 5' Gel Beads v1.1 (−80 °C), RT reagent B, additive A, poly-dT-RT primer (−20 °C), 50% glycerol solution (max. 1.4 ml per chip) to room temperature for 30 min; place RT Enzyme Mix B and cell suspension on ice.
2. Prepare RT mix, by mixing 18.8 μl of RT Reagent B (blue lid), 6.4 μl poly-dT RT primer (blue lid), 2 μl additive A (blue-green lid), and 10 μl RT Enzyme Mix B (white lid), resulting in a total volume of 37.2 μl. Mix 15 times, centrifuge, and keep on ice.
3. Place a Next GEM G Chip in a 10× chip holder, and add 50% glycerol in the following order: (1) 70 μl in row 1, (2) 50 μl in row 2, and (3) 45 μl in row 3.
4. Dispense 34.8 μl RT Mix into an eight-stripe tube per sample, and add the amount of water needed to dilute to the final cell concentration and to obtain 35.5 μl cell suspension; resuspend the cells very slowly and add them to the RT mix. Each tube should contain a total volume of 70.3 μl.
5. With a pipette set to 70 μl slowly mix the cell-RT mix five times, and finally transfer them to the wells in row 1 in the chip. Do not introduce bubbles.
6. Vigorously vortex gel beads for 30 s, and flick sharply to collect beads at the bottom (meanwhile the cell RT mix is priming on the chip). Slowly load 50 μl of the viscous gel beads with a multichannel pipette into row 2 in the chip (avoid bubbles).
7. Pipette 45 μl partitioning oil into row 3 of the chip (*see* **Note 10**).

8. Attach the 10× gasket. The notched cut is at the left top corner. Check that the holes are aligned with the wells. Avoid touching the smooth gasket side. Do not press on top of the gasket! Avoid wetting. Keep assembly horizontal while loading the chip. Tap the touchscreen to eject and insert tray and start the run. Proceed immediately after the completion of the run (~18 min).
9. Remove the chip from the Chromium controller, discard the gasket, and open the chip holder in a 45-degree angle position (*see* **Note 11**). Check for volume uniformity in gel bead and sample wells.
10. With a multichannel pipette transfer very slowly 100 μl uniformly opaque GEMs from row 1 (backwards triangle) into a precooled eight-tube PCR strip. Avoid air bubbles as they may compromise the GEMs. The presence of excess *clear* partitioning oil indicates a potential clog. Discard the used Next GEM chip G.
11. Perform the reverse transcription as follows (lid temperature: 53 °C, reaction volume 125 μl): 20°C, hold; 53°C, 45 min; 85°C, 5 min, 4°C, hold.

    Stopping point: At this point the samples can be stored at 4 °C for up to 72 h or at −20 °C up to a week. If samples were frozen, keep them at room temperature for 10 min before continuing. The aqueous phase will look translucent (rather than clear).

#### 3.1.3 Post GEM Cleanup and cDNA Amplification

To avoid contamination, this section should be carried out on a separate bench dedicated to RNA work.

1. Equilibrate DynaBeads MyOne silane beads (4 °C, white lid), additive A (−20 °C), amplification master mix (−20 °C), and SC5′ feature cDNA primers (yellow lid) from 5' FBC Kit to room temperature for 30 min.
2. Thaw buffer sample cleanup 1 (−20 °C) for 10 min at 65 °C, and vigorously mix until no precipitate is visible anymore, and cool down to room temperature.
3. Apply 130 μl recovery agent dropwise on top of the post incubation GEMs, and wait 2 min until you see a clear aqueous upper phase. If biphasic separation is incomplete, mix by inverting the capped tube strip 5×, and centrifuge briefly.
4. Slowly remove 130 μl of the pink lower phase and discard. Be careful not to aspirate the clear aqueous phase containing your cDNA. A small volume of recovery reagent will remain.
5. Vortex DynaBeads MyOne silane beads, and prepare cleanup mix by mixing 5 μl nuclease-free water, 182 μl buffer sample cleanup 1 (green), 8 μl DynaBeads MyOne silane (white), and 5 μl additive A (blue-green) to a total of 200 μl.

6. Add 200 μl cleanup mix to each sample, and pipette five times. Incubate 10 min at room temperature.
7. Prepare 2.5 ml fresh 80% ethanol per sample.
8. Prepare elution solution I by mixing 98 μl EB buffer (Qiagen), 1 μl 10% Tween 20, and 1 μl additive A (blue-green lid) to a total of 100 μl for each reaction, and mix thoroughly and centrifuge.
9. After a 10 min incubation, place the tube strip into the 10× magnet (HIGH position) until the supernatant is clear (a white interface may appear between the phases). Carefully remove and discard the supernatant.
10. Add 300 μl 80% ethanol to the pellet, wait 30 s, and remove 300 μl volume.
11. Add 200 μl 80% ethanol to the pellet, wait 30 s, and remove 200 μl volume.
12. Pulse-spin tube, and place it on the 10× magnet (low position). Remove residual ethanol. Air-dry beads for 2 min.
13. Elute sample in 35 μl elution solution I. DynaBeads might be difficult to resuspend. Let beads rehydrate for 1 min.
14. Place the strip in a 10× magnet (low position). Transfer 35 μl of GEM-RT product into a new tube.

#### 3.1.4 cDNA and Feature Barcode Amplification

1. Prepare cDNA (*see* **Note 12**) and feature barcode amplification reaction mix on ice by mixing 50 μl amplification master mix (blue lid) and 15 μl SC5′ feature cDNA primers to a total of 65 μl.
2. Add 65 μl of amplification mix to each tube containing 35 μl GEM-RT product. Mix and centrifuge.
3. Perform amplification as follows (lid temperature: 105 °C, reaction volume 100 μl).

   98°C, 45 s; 14 cycles of (98°C, 20 s; 68°C, 30 s; 72°C,1 min); 72°C, 1 min; 4°C, hold. The number of cycles depends on cell size and the number of cells recovered.

   Stopping point: At this point the samples can be stored at 4 °C for up to 72 h.

#### 3.1.5 Feature Barcode and cDNA Fractionation by Size Selection

In this section the amplified feature barcode fraction is separated from the amplified cDNA by size selection, so that both fractions can be further processed separately. These steps should be carried out on a bench dedicated to cDNA work.

1. Add 60 μl (0.6×) of resuspended SPRIselect beads to the amplification tube. Mix well, pulse-spin the tube, and incubate for 5 min at room temperature. Place the tube on a 10× magnet (high position) to separate beads from supernatant.

2. Transfer 80 μl of the supernatant containing the feature barcode fraction into a new clean tube, and keep the remaining supernatant on the beads with the cDNA (incubate for further 5 min).
3. Add 70 μl (2×) SPRIselect beads to the feature barcode supernatant. Mix well. Pulse-spin the tube. Incubate for 5 min at room temperature.
4. Place the cDNA and feature barcode tubes onto a 10× magnet (high position). Remove supernatant (without disturbing the beads).
5. Wash the samples *twice* with 200 μl of freshly prepared 80% EtOH while the tube remains on the 10× magnet. Incubate for at least 30 s, and then remove and discard all of the supernatant.
6. After the final ethanol wash, pulse-spin the beads, and place them on the magnet (low position). Remove any residual ethanol. Air-dry beads. Do not exceed 2 min as this will lead to decreased elution efficiency.
7. Resuspend beads with 45 μl EB. Quickly spin the tube, and incubate for 2 min at room temperature.
8. Place the sample in the 10× magnet (low position), and transfer 45 μl supernatant to a fresh PCR tube.
9. Quantify the cDNA using the NGS Standard Sensitivity Kit for the Fragment Analyzer (Agilent) in a region of 400–5500 bp. Alternatively, an Agilent Bioanalyzer High-Sensitivity chip can be used.

   Stopping point: At this point the samples can be stored at 4 °C for up to 72 h or at −20 °C up to 4 weeks.

#### *3.1.6 Library Construction*

The following steps describe the preparation of three types of libraries: a feature barcode library (that will yield information on cell surface proteins (features) or hash-tags, made from the purified feature barcode fraction) (a), AIRR libraries (TR and/or IG) that require target enrichment (made from the purified cDNA fraction), and 5′ gene expression libraries (made from the purified cDNA fraction) (b, c).

##### Feature Barcode Library Construction by Index-PCR and Purification

1. Prepare the index-PCR by mixing 10 μl amplification master mix (blue lid), 2.5 μl SI Primer, and 2.5 μl Single-Index Kit N Set A to a total of 15 μl, for each sample. Record the assignment of the used indices.
2. Add 5 μl feature barcode sample fraction, mix, pulse-spin, and start the following PCR program (lid temperature: 105 °C): 20°C, Hold; 98°C, 45 s; 98°C, 20 s; 54°C, 30 s; 72°C, 20 s (eight cycles); 72°C, 1 min; 4°C, hold.

The following steps should be carried out on a separate, post-PCR-dedicated bench.

3. Vortex SPRIselect beads and add 24 μl (1.2×) beads to each sample. Incubate for 5 min at room temperature. Place the tube on a 10× magnet (high position). Remove supernatant.
4. Wash samples *twice* with 200 μl freshly prepared 80% EtOH, while the tube remains on the 10× magnet. Incubate for at least 30 s, and then remove and discard all of the supernatant. After the final ethanol wash, pulse-spin the beads, and place them on the 10× magnet (low position). Remove any residual EtOH. Air-dry the beads. Do not exceed 1 min to ensure maximum elution efficiency.
5. Resuspend beads with 15 μl nuclease-free water. Quickly spin the tube and incubate for 2 min at room temperature.
6. Place the sample in the 10× magnet (low position). Transfer 15 μl supernatant to a new PCR tube.
7. Measure concentration with Qubit and dilute library appropriately with TE. Check the size and concentration with the FA NGS Standard Sensitivity (region of 190–210 bp).

   Stopping point: At this point the samples can be stored at 4 °C for up to 72 h or at −20 °C for long-term storage.

Target Enrichment for AIRR Libraries

1. Prepare the enrichment reaction mix 1 on ice by mixing 5 μl amplified cDNA, 35 μl nuclease-free water, 50 μl amplification master mix (blue lid), 5 μl cDNA additive (pink lid), and 5 μl T-cell mix 1 (red lid) **or** B-cell mix 1 (blue lid). Pipet 95 μl of enrichment mix to each tube, and add 5 μl amplified cDNAμ to a total of 100 μl. Mix by pipetting and centrifuge.
2. Perform amplification as follows (lid temperature: 105 °C, reaction volume 100 μl): 98°C, 45 s; x* cycles of (98°C, 20 s; 67°C, 30 s; 72°C, 1 min); 72°C, 1 min; 4°C, hold. *Six cycles for IG and ten cycles for TR.

   Stopping point: At this point the samples can be stored at 4 °C for up to 72 h.

   The following steps should be carried out on a bench dedicated to amplified cDNA.
3. Purification of amplified cDNA: Add 80 μl (0.8×) resuspended SPRIselect beads to the cDNA. Mix well. Pulse-spin the tube. Incubate for 5 min at room temperature. Place the tube on a 10× magnet (high). Discard supernatant without disturbing the beads.
4. Wash the samples *twice* with 200 μl of freshly prepared 80% EtOH, while the tube remains in the magnetic rack. Incubate for at least 30 s, and then remove and discard all of the supernatant.

5. After the final ethanol wash, pulse-spin the beads, place them back on the rack (low position), and remove any residual ethanol. Air-dry the beads. Do not exceed 2 min as this will lead to decreased elution efficiency.
6. Resuspend beads with 35 μl EB. Quickly spin the tube, and incubate for 2 min at room temperature.
7. Place the sample in the 10× magnet (low). Transfer supernatant to a clean nuclease-free PCR tube.
8. Prepare enrichment reaction mix 2 on ice by mixing 5 μl nuclease-free water, 50 μl amplification master mix (blue lid), 5 μl cDNA additive (pink lid), and 5 μl T-cell mix 2 (red lid) **or** B-cell mix 2 (blue lid) to a total of 65 μl.
9. Pipet 65 μl of enrichment mix to each tube and add 35 μl amplified cDNA. Mix by pipetting and centrifuge.
10. Perform amplification as follows (lid temperature: 105 °C, reaction volume 100 μl):

    20°C, hold; 98°C, 45 s; 98°C, 20 s and 67°C, 30 s and 72°C, 1 min (six cycles); 72°C, 1 min; 4 °C, hold.

    Stopping point: At this point the samples can be stored at 4 °C for up to 72 h.
11. Perform size selection on AIRR-enriched cDNA: Vortex SPRIselect beads, and add 50 μl (0.5×) beads to each sample and incubate for 5 min at room temperature. Place the tube on a 10× magnet (high position). Keep supernatant.
12. Transfer 145 μl supernatant to a new tube strip, and add 30 μl (0.8×) SPRIselect reagent to each sample. Mix and incubate for 5 min at room temperature.
13. Place the tube on a 10× magnet (high). Discard supernatant without disturbing the beads.
14. Wash the samples *twice* with 200 μl of freshly prepared 80% EtOH, while the tube remains in the magnetic rack. Incubate for at least 30 s, and then remove and discard all of the supernatant.
15. After the final EtOH wash, pulse-spin the beads, and place them on the 10× magnet (low). Remove any residual ethanol. Air-dry beads. Do not exceed 1 min to ensure maximum elution efficiency.
16. Resuspend beads with 22 μl nuclease-free water (NFW). Quickly spin the tube and incubate for 2 min at room temperature.
17. Place the sample in the 10× magnet (low). Transfer 22 μl supernatant to a new PCR tube.

18. Quantify samples using the FA NGS Standard Sensitivity in a region of 200–5500 bp. Alternatively, an Agilent Bioanalyzer High-Sensitivity chip can be used.

    Stopping point: At this point the samples can be stored at −20 °C for up to 1 week (or 4 °C for up to 72 h).

5′ Gene Expression and AIRR Library Construction: Fragmentation, Adaptor Ligation, and Library Amplification (*See* **Note 13**)

1. Equilibrate the fragmentation buffer to room temperature for 30 min, and verify that there is no precipitate; otherwise vortex (the same should be done later for the ligation buffer, adaptor mix, and SI-PCR primer as well). Keep fragmentation enzyme blend (and later DNA ligase, amplification master mix (−20 °C)) on ice.
2. Prepare your cDNA samples: You will need 50 ng (maximally) in 20 μl in PCR tubes. Keep on ice. Prepare PCR tubes with 20 μl AIRR fraction.
3. Vortex the fragmentation buffer (check if precipitate is still visible), and prepare fragmentation mix (the volume required for both cDNA and AIRR fraction fragmentation) by mixing 20 μl nuclease-free water, 5 μl fragmentation buffer (white lid), and 10 μl fragmentation enzyme blend (purple lid) to a total of 30 μl.
4. Start cDNA fragmentation: Dispense 30 μl of the fragmentation mix into the PCR tubes containing the 50 ng cDNA.
5. Incubate the samples under the following conditions (lid temperature: 65 °C, reaction volume 50 μl): pre-cooling block (4 °C, Hold); fragmentation (32 °C, 5 min); end Repair and A-tailing (65 °C, 30 min); 4 °C, hold.
6. Once cDNA fragmentation has started, add enriched AIRR fraction to the corresponding fragmentation mix tubes. Immediately after cDNA fragmentation, proceed to size selection!
7. cDNA size selection: Vortex SPRIselect beads, and add 30 μl beads to each sample (0.6×). Incubate for 5 min at room temperature. Place the tube on a magnetic rack (high position). Keep supernatant!
8. Meanwhile, add enriched AIRR fraction to the corresponding fragmentation mix tubes, and start AIRR fragmentation: Incubate the samples under the following conditions (lid temperature: 65 °C, reaction volume 50 μl): pre-cooling block (4 °C, hold); fragmentation (32 °C, 2 min); end repair and A-tailing (65 °C, 30 min); 4 °C, hold.
9. Continue cDNA purification: Place the tube on a magnetic rack (high position). Keep supernatant.
10. Transfer 75 μl cDNA supernatant to a new tube. Add 10 μl SPRIselect beads (0.8×) to each sample and mix. Incubate for 5 min at room temperature. Place the tube on a 10× magnet (high). Discard 80 μl supernatant.

11. Wash the beads *twice* with 150 μl of freshly prepared 80% EtOH, while the tube remains in the magnetic rack. Incubate for at least 30 s, and then remove and discard all of the supernatant.
12. After the final EtOH wash, pulse-spin the beads, place them on the 10× magnet (low). Remove any residual ethanol. Air-dry beads. Do not exceed 1 min to ensure maximum elution efficiency.
13. Resuspend beads with 50.5 μl EB. Quickly spin the tube, and incubate for 2 min at room temperature.
14. Place the sample on the 10× magnet (low). Transfer 50 μl supernatant to a new PCR tube.
15. Prepare adapter ligation mix for cDNA and AIRR fraction by mixing 20 μl ligation buffer (green lid), 17.5 μl nuclease-free water, 10 μl DNA ligase (yellow lid), and 2.5 μl adaptor mix (blue-green lid) to a total of 50 μl (*see* **Note 14**).
16. Add 50 μl ligation mix to each sample, mix, and briefly spin down the sample. Incubate as follows (lid temperature: 30 °C, reaction volume 100 μl): 20°C, 15 min; 4°C, hold.
17. Proceed immediately to the cleanup: Add 80 μl (0.8×) of resuspended SPRIselect beads to the sample. Mix well. Pulse-spin the tube. Incubate for 5 min at room temperature. Place the tube on a 10× magnet (high). Discard supernatant.
18. Wash the samples *twice* with 200 μl of freshly prepared 80% EtOH, while the tube remains in the magnetic rack. Incubate for at least 30 s, and then remove and discard all of the supernatant.
19. After the final ethanol wash, pulse-spin the beads, and place them on the 10× magnet (low). Remove any residual ethanol. Air-dry beads. Do not exceed 2 min as this will lead to decreased elution efficiency.
20. Resuspend beads with 30.5 μl EB. Quickly spin the tube and incubate for 2 min at room temperature.
21. Place the sample on the 10× magnet (low). Transfer 30 μl supernatant to a new PCR tube.
22. Prepare the index-PCR by mixing 50 μl amplification master mix (blue lid), 8 μl nuclease-free water, and 2 μl SI-PCR primer (purple lid) to a total of 60 μl.
23. Add 60 μl PCR mix and 10 μl of Single-Index Kit T Set A (PN 1000213) primer (12.5 μM) to each 30 μl sample, and record the assignment of the used indices.
24. Mix, briefly spin down the sample, and start the following PCR program (lid temperature: 105 °C): 20°C, hold; 98°C, 45 s; 98°C, 20 s and 54°C, 30 s and 72°C, 20 s (x* cycles); 72°C,

1 min; 4°C, hold. *Cycles: cDNA: adjust to your input: 1–25 ng cDNA: 14–16 cycles, 26–50 ng cDNA: 10–14 cycles; AIRR: 8 cycles.

Stopping point: At this point the samples can be stored at 4 °C for up to 72 h.

25. Perform size selection on final cDNA libraries and AIRR libraries, followed by a 0.8 x purification using SPRIselect beads: Vortex SPRIselect beads, add 60 μl (0.6×) beads to each cDNA sample, and incubate for 5 min at room temperature. Place the tube on a 10× magnet (high position). Keep supernatant.
26. Vortex SPRIselect beads, add 55 μl (0.55×) beads to each BCR/TCR sample, and incubate for 5 min at room temperature. Place the tube on a 10× magnet (high position). Keep supernatant.
27. Transfer 150 μl of the cDNA supernatant to a new tube, and add 20 μl (0.8×) SPRIselect reagent to each sample. Mix and incubate for 5 min at room temperature.
28. Transfer 150 μl of the IG/TR supernatant to a new tube, and add 20 μl (0.75×) SPRIselect reagent to each sample. Mix and incubate for 5 min at room temperature.
29. Place all tubes (cDNA and IG/TR) on a 10×magnet (high). Discard 165 μl supernatant without disturbing the beads.
30. Wash the samples *twice* with 200 μl of freshly prepared 80% EtOH, while the tube remains in the magnetic rack. Incubate for at least 30 s, and then remove and discard all of the supernatant.
31. After the final ethanol wash, pulse-spin the beads, and place them back on the rack (low). Remove any residual ethanol. Air-dry beads not longer than 1 min to ensure maximum elution efficiency.
32. Resuspend beads with 50 μl NFW. Quickly spin the tube, and incubate for 2 min at room temperature.
33. Vortex SPRIselect beads. Add 40 μl (0.8×) beads to each sample. Incubate for 5 min at room temperature. Place the tube on a 10× magnet (high). Discard 85 μl supernatant (without beads).
34. Wash the samples *twice* with 200 μl of freshly prepared 80% EtOH, while the tube remains in the magnetic rack. Incubate for at least 30 s, and then remove and discard all of the supernatant.
35. After the final ethanol wash, pulse-spin the beads, and place them back on the rack (low). Remove any residual ethanol. Air-dry beads not longer than 1 min to ensure maximum elution efficiency.

36. Resuspend beads with 15 μl NFW. Quickly spin the tube and incubate for 2 min at room temperature.
37. Place the sample in the 10× magnet (low). Transfer 15 μl of supernatant to a clean PCR tube.
38. Quantify libraries with Qubit and with the correct dilution load the Fragment Analyzer using the NGS High-Sensitivity Kit. Calculate concentration in a region of 200–700 bp (Fig. 6).

#### 3.1.7 Sequencing

Prepare libraries for Illumina sequencing (*see* **Note 15**).

### 3.2 Single-Cell SMART-Seq

#### 3.2.1 Cell Sorting and cDNA Synthesis

Due to the sensitivity of these protocols, cells should be collected under clean-room conditions to avoid contamination. The whole process of cDNA synthesis should be carried out in a PCR clean workstation under clean-room conditions.

Buffer Preparations

1. At room temperature, thaw the SMART-Seq sc First-Strand Buffer. On ice, thaw all the remaining reagents (except the enzyme) needed for first-strand cDNA synthesis: 10× lysis buffer, RNase inhibitor, nuclease-free water, SMART-Seq sc TSO (*see* **Note 16**), and 3′ SMART-Seq CDS Primer II A. Gently vortex each reagent to mix and spin down briefly. Store all reagents on ice except the SMART-Seq sc First-Strand Buffer (*see* **Note 17**).
2. Assemble the plain sorting solution (PSS; without 3′ SMART-Seq CDS Primer II A) by mixing 104.5 μl 10× lysis buffer (*see* **Note 18**), 5.5 μl RNase inhibitor, and 1155 μl nuclease-free water to a total of 1265 μl for 96 wells, volume includes ~10% extra for overage (*see* **Note 19**).
3. Mix briefly, and then spin down.
4. Aliquot 11.5 μl of PSS into the appropriate number of wells of PCR tube strips or a 96-well plate (*see* **Note 20**).
5. Seal the plate/tube strips, and briefly spin to ensure the PSS collects at the bottom of the wells (*see* **Note 21**).
6. Store the plate/tube strips at −20 °C for 10 min at a minimum and for up to 24 h. As the volume of PSS is small, the tubes/plate should be kept at −20 °C until just before sorting.

Cell Sorting

1. When ready to sort, unseal the prepared plate/tube strips, and sort cells into the sorting solution according to the FACS system manual and desired parameters.
2. Seal the plate/tube strips with an aluminum foil seal or PCR strip caps. Ensure the plate/tube strips are sealed firmly to minimize any evaporation.
3. Immediately after sorting the cells and sealing the plate, spin briefly to collect the cells at the bottom of each well in the PSS.
4. Place the plate on dry ice to flash-freeze the sorted cells (*see* **Note 22**).

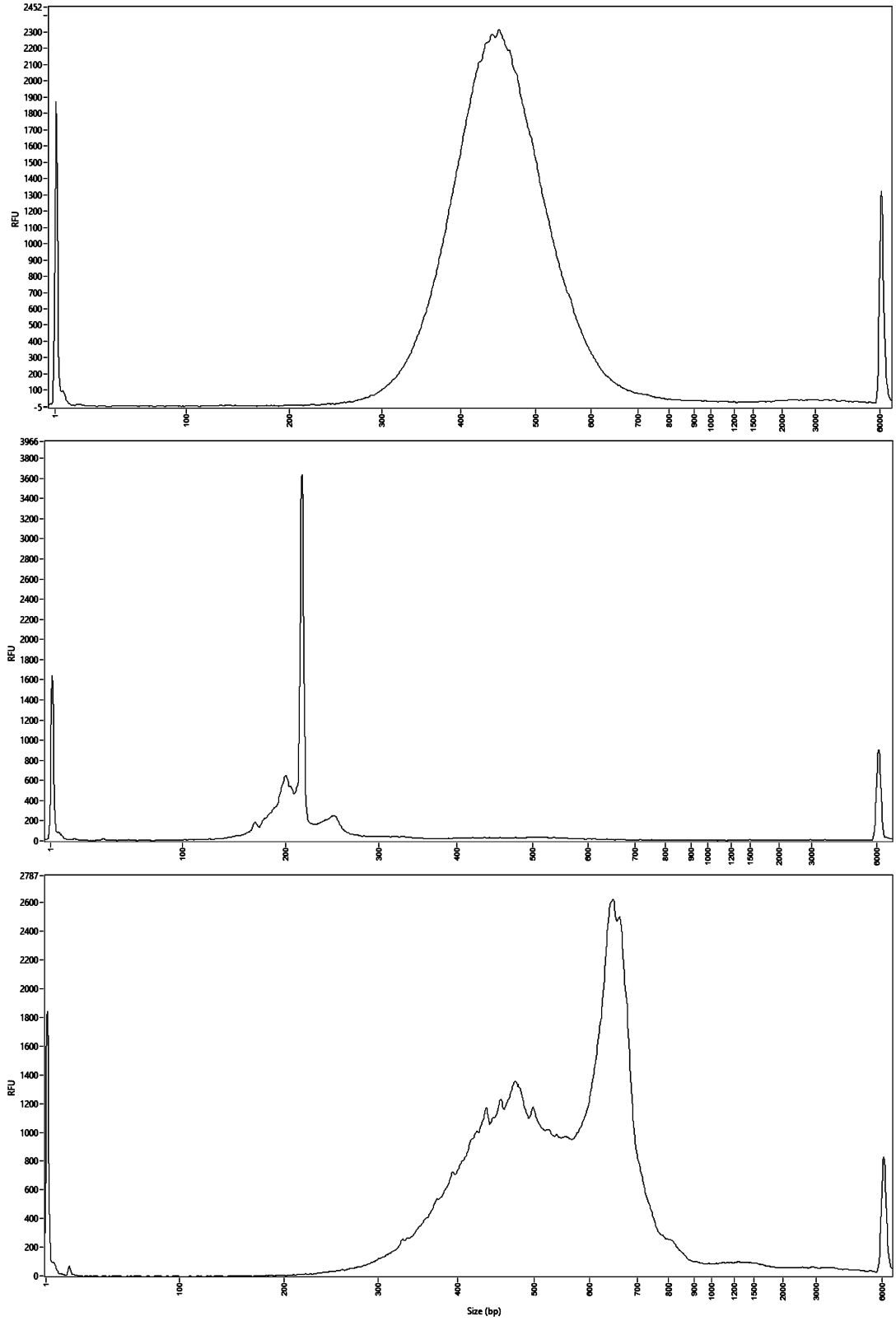


**Fig. 6** Fragment analysis of the final libraries. Top: example of the distribution of a cDNA library. Middle: example of the distribution of a feature barcode library. Bottom: example of the distribution of an AIRR library

Preparing Controls

See below for guidelines on setting up your positive and negative controls alongside your cell samples.

1. Prepare each control in 8-well strips or a 96-well plate by mixing 9.5 μl nuclease-free water, 2 μl control sample (cells or RNA [*see* **Note 23**]), and 1 μl 3′ SMART-Seq CDS Primer II A to a total volume of 12.5 μl.

cDNA Synthesis

1. When ready to start, remove the samples (plate or PCR strips containing the sorted cells) from the freezer and briefly spin to collect the contents at the bottom of the tubes.
2. Place the samples on ice, and add any necessary remaining reagents, including 1 μl of 3′ SMART-Seq CDS Primer II A (*see* **Note 24**). Mix well by gently vortexing, and then spin the tube(s) briefly to collect the contents at the bottom of the tube (*see* **Note 25**).
3. Immediately incubate the tubes at 72 °C in a preheated, hot-lid thermal cycler for 3 min.
4. Prepare RT master mix while the samples are incubating. Prepare enough for all the reactions, plus 10% of the total reaction mix volume, by mixing at room temperature 422.4 μl SMART-Seq sc First-Strand Buffer, 105.6 μl SMART-Seq sc TSO, 52.8 μl RNase inhibitor (40 U/μl), and 211.2 μl SMART-Scribe II reverse transcriptase to a total volume of 792 μl for 96 wells (Includes 10% overage). Add the SMARTScribe II reverse transcriptase just prior to use (in **step 7** part (b) of this section).
5. Immediately after the 3-min incubation at 72 °C, place the samples on ice for at least 2 min (but no more than 10 min).
6. Preheat the thermal cycler to 42 °C.
7. Add the SMARTScribe II reverse transcriptase to the RT master mix. Mix well by gently vortexing, and then spin the tube briefly in a mini-centrifuge to collect the contents at the bottom of the tube.
8. Add 7.5 μl of the RT master mix to each sample. Mix the contents of the tubes by gently vortexing, and spin briefly to collect the contents at the bottom of the tubes.
9. Place the tubes in a thermal cycler with a heated lid, preheated to 42 °C. Run the following program: 42°C, 180 min; 70°C, 10 s; 4°C, hold.

   Stopping point: The tubes can be stored at 4°C overnight.

### 3.2.2 cDNA Amplification by LD PCR

The PCR primers amplify the cDNA by priming to the sequences introduced by the 3′ SMART-Seq CDS Primer II A and the SMART-Seq sc TSO.

1. Thaw SeqAmp CB PCR buffer and PCR primer on ice. Do not thaw SeqAmp DNA polymerase. Gently vortex each reagent tube to mix and spin down briefly. Store on ice.
2. Prepare enough PCR master mix for all the reactions by mixing 2640 μl SeqAmp CB PCR buffer (2×), 105.6 μl PCR primer, 105.6 μl SeqAmp DNA polymerase, and 316.8 μl nuclease-free water to a total volume of 3168 μl for 96 wells (including 10% overage).

   Remove the SeqAmp DNA polymerase from the freezer, gently mix the tube without vortexing, and add to the master mix just before use. Mix the master mix well by vortexing gently, and spin the tube briefly to collect the contents at the bottom of the tube.
3. Add 30 μl of PCR master mix to each tube containing 20 μl of the first-strand cDNA product. Mix well by gently vortexing, and briefly spin to collect the contents at the bottom of the tube. Transfer the samples from the PCR clean work station to the general lab. All downstream processes should be performed in the general lab.
4. Place the tubes in a preheated thermal cycler with a heated lid, and run with the following conditions: 95 °C, 1 min; 98 °C,10 s; 65 °C, 30 s and 68 °C, 3 min (X* cycles); 72 °C, 10 min; 4 °C, hold. PCR cycle number guidelines depend on the cell type (******see* **Note 26**).

   Stopping point: The tubes may be stored at 4 °C overnight.

*3.2.3 Purification of Amplified cDNA*

1. If purification is performed directly in the PCR tubes or strips using the Takara Bio SMARTer-Seq Magnetic Separator-PCR Strip, add 40 μl of beads (*see* **Note 27**) to each sample (*see* **Note 28**). Mix thoroughly by vortexing for 3–5 s or pipetting the entire volume up and down at least ten times. Proceed to **step 3** of this section.
2. If you are performing purification with the Thermo Fisher Magnetic Stand-96 (recommended if processing 48–96 samples), cDNA samples need to be transferred to a 96-well V-bottom plate. Distribute 40 μl of beads (*see* **Note 28**) to each well of the 96-well V-bottom plate, and then use a multichannel pipette to transfer the cDNA. Pipette the entire volume up and down at least ten times to mix thoroughly. Proceed to **step 3** of this section.
3. Incubate the beads-cDNA mixture at room temperature for 8 min to let the cDNA bind to the beads.
4. Briefly spin the samples to collect the liquid from the side of the tubes or plate wells (centrifugation is generally not necessary if using a 96-well V-bottom plate).

5. Place the samples on the magnetic separation device for ~5 min or longer, until the liquid appears completely clear and there are no beads left in the supernatant.
6. While the samples are on the magnetic separation device, remove and discard the supernatant. Take care not to disturb the beads.
7. Keep the samples on the magnetic separation device. Add 200 μl of freshly made 80% ethanol to each sample without disturbing the beads (*see* **Note 27**). Incubate for 30 s. Then, carefully remove and discard the supernatant, taking care not to disturb the beads. The cDNA remains bound to the beads during the washing process.
8. Repeat the ethanol wash (**step 7** of this section) once more.
9. Briefly centrifuge the samples to collect the liquid from the side of the tubes or plate wells. Place the samples on the magnetic separation device for 30 s, and then remove any residual ethanol with a pipette.
10. Incubate the samples at room temperature for ~2–2.5 min, until the pellet is no longer shiny, but before cracks appear (*see* **Note 29**).
11. Once the beads are dry, remove the samples from the magnetic separation device, and add 17 μl of elution buffer to cover the bead pellet. Mix thoroughly by pipetting or gently vortexing to resuspend the beads.
12. Incubate at room temperature for at least 2 min to rehydrate.
13. Briefly spin the samples to collect the liquid from the side of the tubes or plate wells. Place the samples back on the magnetic separation device for 1 min or longer, until the solution is completely clear.
14. Transfer clear supernatant (~15 μl) containing purified cDNA from each tube/well to a new tube/plate. Do not pool samples at this point. Take care not to carry over any beads with your sample.
15. Proceed to validation immediately or store at −20 °C.

#### *3.2.4 Validation Using the Agilent 2100 Bioanalyzer*

1. Aliquot 1 μl of the amplified cDNA for validation using the Agilent 2100 Bioanalyzer and Agilent's High-Sensitivity DNA Kit. See the Agilent High-Sensitivity DNA Kit User Manual for instructions.
2. Compare the results for your samples and controls to verify whether the sample is suitable for further processing. Successful cDNA synthesis and amplification should yield no product in the negative control, and a distinct peak spanning 400 bp to 10,000 bp, peaked at ~2500 bp for the positive control RNA sample, yielding approximately 3.4–17 ng of cDNA (depending on the input type and amount).

### 3.2.5 Library Preparation for Next-Generation Sequencing

The following sections describe a modified Illumina Nextera XT DNA library preparation protocol that has been fully validated to work with the SMART-Seq Single-Cell Kit. The reaction size has been reduced to a quarter volume of what is recommended by Illumina.

#### Dilute and Prepare cDNA for Tagmentation

1. Dilute each cDNA to 100 pg/μl with nuclease-free water in a plate or PCR strips (*see* **Note 30**). Do not pool at this step. Vortex at medium speed for 20 s and centrifuge at 350 × *g* for 1 min.
2. Warm Tagment DNA Buffer and NT Buffer to room temperature. Visually inspect NT Buffer to ensure that there is no precipitate. If there is a precipitate, vortex the buffer until all particles are resuspended.
3. After thawing, gently invert the tubes 3–5 times, followed by centrifuging the tubes briefly, to ensure all reagents are adequately mixed.
4. Label a new 96-well PCR plate "Library Prep."
5. In a 1.5-ml PCR tube, prepare tagmentation premix by mixing 300 μl Tagment DNA Buffer and 150 μl amplification tagment mix to a total volume of 450 μl (calculated based on a 25% excess). Vortex gently for 20 s and centrifuge the tube briefly.
6. Distribute 3.75 μl of the tagmentation premix into each well of the "Library Prep" plate.
7. Transfer 1.25 μl of each diluted cDNA sample to the "Library Prep" plate.
8. Seal the plate and vortex at medium speed for 20 s. Centrifuge at 2000 × *g* for 5 min to remove bubbles.
9. Place the "Library Prep" plate in a thermal cycler with a heated lid, and run the following program: 55 °C, 10 min; 10 °C, hold.
10. Once the thermal cycler reaches 10 °C, pipette 1.25 μl of NT Buffer into each of the tagmented samples to neutralize the samples (*see* **Note 31**).
11. Seal the plate and vortex at medium speed, and then centrifuge at 2000 × *g* for 1 min.
12. Incubate at room temperature for 5 min.

#### Amplify the Tagmented cDNA

1. Pipette 3.75 μl of Nextera PCR Master Mix (NPM) into each well of the "Library Prep" plate using an eight-channel pipette (*see* **Note 32**).

   Select appropriate Index 1 (N7xx) and Index 2 (S5xx) primers for the number of samples in your experiment (*see* **Note 33**).

2. Pipette 1.25 μl of Index 1 Primers (N7xx) into the corresponding wells of each row of the "Library Prep" plate. As a result, each of the 12 wells in row "A" will contain different Index 1 Primers.
3. Pipette 1.25 μl of Index 2 Primers (S5xx) into the corresponding wells of each column of the "Library Prep" plate. As a result, each of the 8 wells in column "1" will contain different Index 2 Primers.
4. Seal the plate with adhesive film and vortex at medium speed for 20 s. Centrifuge at 2000 × *g* for 2 min. Place the "Library Prep" plate into a thermal cycler, and perform PCR amplification using the following program: 72°C, 3 min; 95°C, 30 s; 95°C, 10 s; 55°C, 30 s; and 72°C, 60 s (12 cycles); 72°C, 5 min; 10°C, hold.

   Samples can be left overnight in the thermal cycler at 4 °C. If not processed within the next day, freeze the PCR products at −20 °C.

### Pooling and Purification of Amplified Libraries

1. Determine the number of libraries to be pooled based on the desired sequencing depth and sequencer throughput. If preferred, clean up libraries individually before pooling (*see* **Note 34**).
2. Pool the libraries by pipetting a fixed volume from each sample into a 1.5-ml tube or PCR tube. Volumes between 2 μl and 8 μl are appropriate. Do not use less than 2 μl per sample to ensure greater accuracy (e.g., to pool 96 libraries, add 2 μl of each library (total 192 μl) and 154 μl of bead volume to a 1.5 ml tube. The bead volume is approximately 80% of the total pool volume).
3. Add a volume of beads representing 80% of the volume of the pooled libraries. If cleaning up libraries individually, add 40 μl of beads to each 50-μl sample.
4. Mix well by vortexing or pipetting the entire mixture up and down ten times (*see* **Note 35**).
5. Incubate at room temperature for 5 min to let the cDNA libraries bind to the beads.
6. Briefly spin the sample to collect the liquid from the side of the tube. Place the tube on a magnetic stand for ~2 min or until the liquid appears completely clear, and there are no beads left in the supernatant.
7. While the samples are on the magnetic separation device, remove and discard the supernatant. Take care not to disturb the beads.
8. Keep the samples on the magnetic separation device. Add 200 μl of fresh 80% ethanol to each sample without disturbing

the beads. Incubate for 30 s, and then remove and discard the supernatant, taking care not to disturb the beads. The cDNA remains bound to the beads during washing.

9. Repeat the ethanol wash (**step 8** of part c) of this section) once more.
10. Briefly centrifuge the samples to collect the liquid from the side of the tube or plate well. Place the samples on the magnetic separation device for 30 s, and then remove any residual ethanol with a pipette.
11. Incubate the samples at room temperature for ~5–15 min, until the pellet is no longer shiny, but before cracks appear (*see* **Note 29**). The pooled samples requiring higher bead volumes take longer to dry.
12. Once the beads are dry, elute the pooled, purified libraries by adding the required volume of nuclease-free water (provided), based on the number of samples pooled (*see* **Note 36**).
13. Remove from the magnetic separation device, and vortex the tube for 3 s to mix thoroughly. Incubate at room temperature for ~5 min to rehydrate the beads.
14. Briefly spin to collect the liquid from the side of the tube. Place the tube back on the magnetic separation device for ~2 min or longer until the solution is completely clear.
15. Transfer the clear supernatant containing purified libraries to a nuclease-free, low-adhesion tube. Label each tube with sample information. The purified libraries can be stored at −20 °C.

#### 3.2.6 Sequencing

Sequence the SMART-Seq single-cell library with Illumina sequencing (*see* **Note 37**).

### 3.3 10x Genomics Chromium Next GEM Single-Cell V(D)J Kit Data Processing and Analysis

10X Genomics data is analyzed using the Cell Ranger software, provided by 10x Genomics free of charge. Cell Ranger allows (1) sequencing raw data demultiplexing, (2) quality control of the raw data obtained, (3) raw data alignment to the reference genome, and (4) data matrix preparation for further in-depth analyses using dimensionality reduction methods. Three Cell Ranger pipelines are now available: `cellranger count` (for transcriptome and feature data), `cellranger vdj` (for AIRR data) and `cellranger multi` (which does the integrative analysis of transcriptome, AIRR, and feature data). Loupe Brower and Loupe VDJ Browser, which are also available via the 10x Genomics website, provide a complementary set of analysis tools. Note that the Loupe Browsers are only available for Windows or macOS environments.

#### 3.3.1 Setup

Cell Ranger can be installed in a folder named "cellranger" in the home directory. Before running Cell Ranger, ensure that this folder is included in the PATH environment variable: `export PATH = $PATH:$HOME/cellranger.`

#### 3.3.2 Demultiplexing

Libraries can be demultiplexed using the following command (*see* **Note 38**): `cellranger mkfastq --id = LBA-01 --run = data/BCL --samplesheet = cellranger-tiny-bcl-samplesheet-1.2.0.csv --output-dir = data/VDJ.`

Arguments:

--id: The folder name that will be created by cellranger mkfastq (here: LBA-01).

--run: Path of Illumina BCL Sample run folder.

--output: Path to where your folder will be located (here: data/VDJ) (optional).

--samplesheet: Path to an Illumina Experiment Manager (IEM) sample sheet format which contains all the information needed for describing samples.

--csv: Path to a csv file which contains some information for describing samples. It is an alternative if your Illumina Experiment Manager is not provided.

#### 3.3.3 Alignment

V(D)J genes, gene transcripts, and feature barcodes can be analyzed by running either "cellranger vdj" (a) or "cellranger count" (b) using the following commands (*see* **Note 39**):

1. `cellranger vdj --id LBA-01-VDJ --sample LBA-01-Sample --fastqs data/VDJ --reference data/refdata/refdata-cellranger-vdj-GRCh38-alts-ensembl-5.0.0.`

   Arguments (*see* **Note 40**):

   (a) --id: The folder that will contain the output of the pipeline (here: LBA-01-VDJ)

   (b) --sample: Sample name as specified in the FASTQ file (here: LBA-01-Sample)

   (c) --fastqs: Path to where your FASTQ file is located (here: data/VDJ)

   (d) --reference: Path to a Cell Ranger compatible VDJ reference (*see* **Notes 41** and **42**).

2. `cellranger count --id LBA-01-GEX --transcriptome data/refdata/refdata-gex-GRCh38-2020-A --fastqs data/GEX --sample LBA-01-GEX-Sample`

   Arguments:

   (a) --id: The folder that will contain the output of the pipeline (here: LBA-01-GEX)

   (b) --sample: Sample name as specified in the FASTQ file (here: LBA-01-GEX-Sample)

(c) --fastqs: Path to where your FASTQ file is located (here: data/GEX)

(d) --transcriptome: Path to a Cell Ranger compatible GEX reference (*see* **Notes 41** and **42**).

#### 3.3.4 Quality Control

Visualize the HTML report file generated by Cell Ranger on your computer. Ensure that there are no warnings regarding the RNA content, the number of gene/VDJ count, the number of cells detected, or the average number of genes per cell. When feature barcoding is used, ensure the absence of aggregates, of unbound antibodies as this may dramatically reduce the cells to be analyzed. In case you encounter major warnings, please reconsider your protocol (antibody dilution, washing steps, etc.). This step is critical to ensure the quality of the data that will be further analyzed.

#### 3.3.5 Exploratory Analysis

Loupe Browser (version 5.0.1) software can be employed for a first analysis using the vloupe.vloupe file to explore the aligned results using classical single-cell approaches and determine whether expected variables have been detected (and therefore sequenced) (*see* **Note 43**).

#### 3.3.6 In-Depth Analysis

Further analysis can be performed using additional tools (*see* **Note 44**).

### 3.4 Single-Cell SMART-Seq Data Processing and Analysis

The steps below outline how to use TraCeR [6] as a Docker container using the test data in the TraCeR GitHub repository. The TraCeR pipeline consists of two main steps, which are run separately: assemble and summarise (*see* **Note 45**).

#### 3.4.1 Raw Data Processing

1. Consider whether your data corresponds to the demands for TraCeR (*see* **Note 46**).
2. Check raw data files (*see* **Note 47**).
3. Perform standard quality control, and trim sequencing reads for bad quality reads/nucleotides and adapter sequences by using tools such as Trim Galore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore) and Cutadapt [7].

#### 3.4.2 Obtain the Software

1. Download or clone the TraCeR GitHub repository by running: `git clone` https://github.com/Teichlab/tracer.
2. Pull the Docker container from DockerHub by running: `docker pull teichlab/tracer`.
3. Increase the Docker memory limit to 6–8 GB to avoid TraCeR running out of memory during the *assemble* step. Details on how to increase the Docker memory limit are found at https://docs.docker.com/docker-for-windows/#advanced for Windows and https://docs.docker.com/docker-for-mac/#advanced for Mac.

4. Run a TraCeR test to verify that the installation of TraCeR and all its dependencies runs smoothly and gives the expected outputs. To do this, enter the main directory of the downloaded GitHub TraCeR repository, and run `docker run -it --rm -v $PWD:/scratch -w /scratch teichlab/tracer test -o test_data`.
5. Inspect output of the TraCeR test by comparing the output in *test_data/results/filtered_TCR_summary* with the expected results in *test_data/expected_summary*. The test output should consist of three cells, of which Cell 1 and Cell 2 are clonally related. Each cell should have a productive TRA, a nonproductive TRA, a productive TRB, and a nonproductive TRB rearrangement.

#### 3.4.3 Reconstruct TR Sequences with the Assemble Mode

1. Run TraCeR following commands followed by any appropriate arguments, from the directory containing the input data, to ensure that TraCeR runs on Docker (*see* **Note 48**):

```
docker run -it --rm -v $PWD:/scratch -w /scratch teichlab/
tracer assemble [options] <file_1> [<file_2>] <cell_name>
<output_directory>
```

The main arguments expected by TraCeR are as follows:

<*file_1*>: FASTQ file containing #1 mates from paired-end sequencing or all reads from single-end sequencing. If paired-end sequencing is used, provide #2 mates after the #1 mates FASTQ file.

<*cell_name*>: Name of the cell chosen by the user. This name will be used in file names and labels.

<*output_directory*>: Directory for output. The cell-specific output from the *assemble* mode will be found in <output_directory>/<cell_name>.

TraCeR also accepts several options, which are detailed at https://github.com/Teichlab/tracer and in [8].

2. Reconstruct TRA, TRB, TRD, and TRG rearrangements from paired-end data using one processor core by running the following command (here for a hypothetical example dataset consisting of T cells from humans):

```
docker run -it --rm -v $PWD:/scratch -w /scratch teichlab/
tracer assemble cell_1_R1.fq.gz cell_1_R2.fq.gz cell_1 Exp_1
-c my_config_file -s Hsap --loci A B G D -m assembly
```

#### 3.4.4 Identify Clonally Related Cells with the Summarise Mode

1. Remove low-quality cells, defined by standard quality control pipelines for single-cell RNA-seq data, to obtain the most accurate results, before running the TraCeR *summarise* mode.
2. Define cell populations to analyze. The output of the *assemble* mode of TraCeR is a directory for each cell. Before running the *summarise* mode, create a new directory for cells you want to analyze together, and move the relevant TraCeR result directories for these cells into the new directory. If cells from multiple donors are present in the dataset, run the summarise mode separately for each donor in order to define true clonally related cells.
3. Run TraCeR in *summarise* mode on Docker with the following command, where <*input_dir*> is the path to the directory containing the output of TraCeR *assemble* mode for all the cells to be summarized together:

```
docker run -it --rm -v $PWD:/scratch -w /scratch teichlab/
tracer summarise [options] <input_dir>
```

For a hypothetical example dataset consisting of T cells from humans, the following command could be run using one processor core:

```
docker run -it --rm -v $PWD:/scratch -w /scratch teichlab/
tracer summarise Exp_1 -c my_config_file -s Hsap --loci A B G D
-g svg -u
```

#### 3.4.5 Quality Control

1. Create a new directory for filtering out the *assemble* result directories for suspicious cells.
2. Remove likely cell doublets/multiplets affecting clonal assignments with more or less strict criteria depending on the dataset and biological questions (*see* **Note 49**) by visually inspecting the clonal graph outputs created by TraCeR *summarise* run with the `-u` flag. Likely doublets/multiplets can be seen as cells with two or more sets of rearranged TRA and TRB chains (or TRD and TRG), connecting smaller clone groups that otherwise do not share rearranged sequences with each other. If such likely doublet/multiplets exist in the data, we recommend to remove the result directories from *assemble* for these cells and rerun TraCeR *summarise* mode without the `-u` flag.
3. Remove likely cell doublets/multiplets/contaminations by opening `TCR_summary.txt` in the unfiltered summary folder and looking at the section named `#Cells with more than two recombinants for a locus`. Move the *assemble* result folder for any cell with more than three reconstructed TR

rearrangements for any locus to the directory containing cells to be filtered out.

4. Remove potential cell doublets/contaminations by opening `<cell_name>/unfiltered_TCR_seqs/unfiltered_TCRs.txt` for each cell with three reconstructed TR rearrangements for a locus. Discard a cell if all reconstructed sequences for the locus in question are substantially different from each other and have nonzero expression values. If two or more rearrangements use the same gene segments, they have probably not been collapsed to one sequence due to PCR errors or misassemblies, and the sequence with the highest expression value is likely the true rearranged sequence. In such cases, the cells need not be filtered out.
5. Consider removing other possible cell doublets: While the expression of two different TRA rearrangements is not uncommon in T cells, it is less common for TRB rearrangements [9]. Thus, cells expressing two TRA and two TRB rearrangements may be cell doublets and can be filtered out depending on the desired balance of discarding false cell doublets versus keeping potential cell doublets.

#### 3.4.6 Interpreting TraCeR Summarise Output

The output of the *summarise* step is written to `<input_dir>/filtered_TCR_summary` or `<input_dir>/unfiltered_TCR_summary` depending on whether the `-u` flag was used. The most useful output files of `summarise` are as follows:

`TCR_summary.txt`: Summary statistics for TR reconstruction and a list of clonally related cells.

`recombinants.txt`: List of TR identifiers, lengths, productivities, and CDR3 sequences reconstructed for each cell.

`clonotype_sizes.pdf` and `clonotype_sizes.txt`: Distribution of clone sizes as bar graph and text file.

`clonotype_network_[with|without]_identifiers.<graph_format>`: Graphical representation of TR rearrangements with full identifiers or just lines indicating presence/absence of rearrangements of the different loci (Fig. 7). The graph without identifiers gives an overview of the degree of clonality in the cell population (A). The graph with identifiers only shows clonally expanded cells, with details on the shared TR rearrangements within each clone (B). Edges between nodes represent cells that share one or more reconstructed TR rearrangements, colored by the TR locus (*see* **Note 50**).

#### 3.4.7 In-Depth Analysis

Further analysis can be performed using additional tools (*see* **Note 44**).

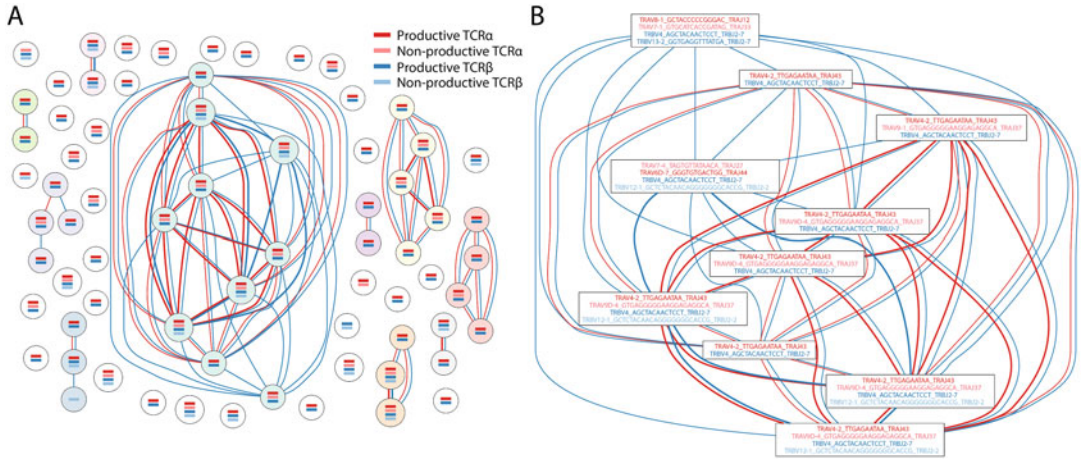


**Fig. 7** Overview of clonality of a T-cell population based on TR sequences reconstructed by TraCeR. Each cell is represented by a node, while each reconstructed TR sequence is represented by a horizontal line (**a**) or a sequence identifier (**b**; only showing the largest clone group), colored according to chain type. Cells sharing identical TR sequences are connected with edges colored by chain type

## 4 Notes

1. The thermal cycler should always be used with the heated lid option turned on. If prompted to input a specific temperature, use 105 °C. Most thermal cyclers with heated lids will automatically adjust the lid temperature just above the highest block temperature within a cycling program. However, if the thermal cycler does make this automatic adjustment, one may want to follow the manufacturer's instructions to choose a lower lid temperature for the reverse transcription step.
2. The kit has been specifically validated with the beads listed above. Any substitutions may lead to unexpected results. Beads need to come to room temperature before the container is opened. We strongly recommend aliquoting the beads into 1.5-ml tubes upon receipt and then refrigerating the aliquots. Individual tubes can be removed for each experiment, allowing them to come to room temperature more quickly (~30 min). Aliquoting is also instrumental in decreasing the chances of bead contamination. Immediately before use, vortex the beads until they are well dispersed. The color of the liquid should appear homogeneous. Confirm that there is no remaining pellet of beads at the bottom of the tube. Mix well to disperse before adding the beads to your reactions. The beads are viscous, so pipette them slowly.
3. Please consult the Chromium Next GEM Single-Cell V(D)J Kit manual (v1.1 or 2) from the 10x Genomics website (https://support.10xgenomics.com/single-cell-vdj/overview/doc/user-guide-chromium-single-cell-vdj-reagent-kits-

v11-chemistry) carefully, and follow all instructions regarding general reagent handling, Chromium Next GEM Chip handling, assembly, loading, and all other technical instructions. The evolving kit versions (v1., v1.1, v2) differ with respect to certain volumes and concentrations. The following protocol is based on v1.1.

4. Multiplexing samples (hash-tagging): multiple small cell samples can be multiplexed into one large sample. This is accomplished by labelling an ubiquitously expressed cell surface protein with an antibody conjugated to a feature barcode oligonucleotide, followed by the direct capture of the feature barcode by the gel bead primer and amplification of the feature barcode (such "hashtag" antibodies are provided by BioLegend (https://www.biolegend.com/totalseq)). A lipid anchored hash-tagging strategy will be available from 10x Genomics in the future (Cellplex).
5. Measuring the expression of cell surface proteins or bound antigenic peptides. This is accomplished by labeling cell surface proteins with antibodies conjugated to a feature barcode oligonucleotide, followed by the direct capture of the feature barcode by the gel bead primer. Cells can also be labeled using a feature barcode oligonucleotide conjugated to a MHC dextramer, such as a dCODE dextramer (detailed information can be found here from 10x Genomics:

   https://support.10xgenomics.com/single-cell-gene-expression/sample-prep/doc/demonstrated-protocol-cell-surface-protein-labeling-for-single-cell-rna-sequencing-protocols

   https://support.10xgenomics.com/single-cell-vdj/sample-prep/doc/demonstrated-protocol-cell-labeling-with-dextramers-for-single-cell-rna-sequencing-protocols).
6. This protocol starts from a cell suspension, usually isolated by flow cytometry. If thawing cells before staining and sorting, 10 μl benzonase per 10 ml cell suspension must be used. Incubation with FACS antibodies and/or TotalSeq-C antibodies is done on ice for 30 min. FACS sorting buffer washing steps should be repeated three times after staining.
7. Sorting for example 20,000 cells using the 70 μm nozzle and four-way purity mode yields 26.5 μl of sorted cell suspension. If the total sample is loaded, the volume should be topped to a final volume of 35.5 μl required to be loaded onto the chip. The final number of cells that are recovered depends on the cell concentration (*see* Table 1).
8. Counting the obtained cell suspension (after FACS sorting) under a microscope (or with a cell counter such as a Vi-Cell instrument) is crucial. A small excess of cells should be sorted, to allow for the removal of a small aliquot for cell counting (2 μl will be sufficient).

**Table 1**
**Cell recovery as a function of cell concentration**

| | Cell recovery (in thousands) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Cells/μl | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 800 | **4.1**<br>33.7 | **6.2**<br>31.6 | **8.3**<br>29.5 | **10.3**<br>27.5 | **12.4**<br>25.4 | **14.4**<br>23.4 | **16.5**<br>21.3 | **18.6**<br>19.2 | **20.6**<br>17.2 |
| 900 | **3.7**<br>34.1 | **5.5**<br>32.3 | **7.3**<br>30.5 | **9.2**<br>28.6 | **11.0**<br>26.8 | **12.8**<br>25.0 | **14.7**<br>23.1 | **16.5**<br>21.3 | **18.3**<br>19.5 |
| 1000 | **3.3**<br>34.5 | **5.0**<br>32.8 | **6.6**<br>31.2 | **8.3**<br>29.5 | **9.9**<br>27.9 | **11.6**<br>26.3 | **13.2**<br>24.6 | **14.9**<br>23.0 | **16.5**<br>21.3 |
| 1100 | **3.0**<br>34.8 | **4.5**<br>33.3 | **6.0**<br>31.8 | **7.5**<br>30.3 | **9.0**<br>28.8 | **10.5**<br>27.3 | **12.0**<br>25.8 | **13.5**<br>24.3 | **15.0**<br>22.8 |
| 1200 | **2.8**<br>35.1 | **4.1**<br>33.7 | **5.5**<br>32.3 | **6.9**<br>30.9 | **8.3**<br>29.5 | **9.6**<br>28.2 | **11.0**<br>26.8 | **12.4**<br>25.4 | **13.8**<br>24.0 |

Top rows (bold): μl cell suspension; bottom rows: μl PBS or nuclease-free water (NFW)

9. If the sorting process of three samples takes <20 min for each, the samples can be loaded onto the same chip, whereas if sorting of each sample takes >30 min, the samples should be run on separate chips.
10. Failure to add partitioning oil can damage the Chromium controller!
11. Avoid storing GEMs longer than 1 h on ice! Do not centrifuge!
12. The cDNA contains AIRR transcripts and all the remaining cellular transcripts.
13. The time required for fragmentation of the AIRR library is shorter, and no size selection is performed between fragmentation and adapter ligation. Thus, you can start with the cDNA fragmentation, and while purifying the fragments, the IG/TR fraction can be fragmented. The master mix can be prepared for both libraries together.
14. The viscosity of the ligation mix is higher than for other mixes. Please ensure thorough mixing of all components, otherwise the ligation efficiency could be reduced.
15. AIRR, gene expression, and feature barcode libraries (BCL) are standard Illumina paired-end constructs which begin with P5 and end with P7. Libraries are prepared for Illumina sequencing with the following components: using version v1.1 with Single-Index Kit N Set A (for feature barcode libraries); Single-Index Kit T Set A (for gene expression and BCR/TCR libraries) and using v2 with dual indices (Dual Index Kit TN Set A (for feature barcode libraries); and Dual Index Kit TT Set A (for gene expression and IG/TR libraries).

These libraries include a P5 part that binds to the flow cell, the primer binding site for read 1 which contains a 16 bp 10× barcode to identify the cell assignment, followed by a 10 mer UMI for counting the transcripts, the TSO, and the poly-A-stretch. The transcript insert follows and is sequenced in read 2, followed by a region for the sequencing primer, the i7 index, and the P7 part that binds to the flow cell. The minimum sequencing lengths are 26 bp for read 1 and 91 bp for read 2. Sequencing these libraries produces standard Illumina BCL data. The optimal sequencing depth is 25–30 K reads/cell for cDNA libraries and 10 K reads/cell for AIRR and feature barcode libraries. For library loading we recommend the following: MiSeq (2× 150 bp reads): 15 pM; NovaSeq in XP mode: cDNA lib 250 pM; feature barcode library: 190–200 pM; AIRR library: 300 pM; NovaSeq in standard mode: cDNA lib 450 pM; and feature barcode library 300 pM: AIRR library 500 pM.

16. First-strand cDNA synthesis is primed by the 3′ SMART-Seq CDS Primer II A and uses the SMART-Seq sc TSO for template switching at the 5′ end of the transcript.
17. The SMART-Seq sc First-Strand Buffer forms precipitates. Keep this buffer at room temperature until you use it. Vortex before using it to ensure all components to be dissolved.
18. The lysis buffer contains a detergent, avoiding bubbles when mixing.
19. Due to small pipetting volumes, prepare no less than 250 μl of plain sorting solution (PSS) (enough for 18 wells). Scale up as needed. Be sure to count any negative control reactions you wish to include. If you need to sort large numbers of cells compared to the number of cDNA reactions you plan to prepare, you have the option to purchase the 10× lysis buffer (Takara Bio) separately.
20. To minimize bubble formation, set single- or multichannel pipettes to 11.6 μl, and pipette only to the first stop when aliquoting. Changing tips often also minimizes bubble formation.
21. If using PCR strips, strip caps can be used instead of aluminum foil, but are not practical when sorting a large number of samples.
22. If using PCR strips, leave them secured on the PCR rack for freezing. Store sorted samples at −80 °C until ready to proceed with cDNA synthesis. To use PCR strips sealed with an aluminum foil seal, use a clean razor blade to separate the individual strips, and then push up slightly on the tubes from under the PCR rack to loosen them before taking out the desired number of strips. Long-term storage at −80 °C may impact the

efficiency of cDNA synthesis; however, it is safe to store the cells for several weeks prior to cDNA synthesis.

23. Control cells should be in PSS described above. PSS does not contain the 3′ SMART-Seq CDS Primer II A, so it must be added here.

    The Control Total RNA is supplied at a concentration of 1 μg/μl. It should be diluted to match the concentration of your test sample using serial dilutions. For positive and negative controls, replace the cell sample with 2 μl of the diluted control RNA and water, respectively.

24. Since the PSS does not include the 3′ SMART-Seq CDS Primer II A, one needs to add it when thawing the samples.

25. If more than one reagent needs to be added (i.e., the 10× reaction buffer, the 3′ SMART-Seq CDS Primer II A and extra nuclease-free water), they can be pooled in a master mix and incorporated into each sample as a single addition.

26. Cycle guideline: When using good-quality control RNA, such as the Control Total RNA from the kit (mouse brain total RNA), 2 pg will generate 500–1500 pg/μl if using 19 cycles. It is best to choose a number of cycles that will yield <1500 pg/μl.

    Cycling guidelines based on cell type and pg RNA per cell: PBMCs (1–5 pg), 20 cycles; Jurkat cells (5 pg), 17 cycles; and lymphoblastoid cells (2–15 pg), 17–19 cycles.

27. Aliquot beads into 1.5-ml tubes upon receipt in the laboratory. Before each use, bring bead aliquots to room temperature for at least 30 min, and mix well by vortexing. Here the bead: sample ratio is 0.8:1. Prepare fresh 80% ethanol for each experiment. One will need 400 μl per sample. Use a magnetic separation device for 0.2-ml tubes, strip tubes, or a 96-well plate.

28. Do not pool the samples at the cDNA purification step. If pooling is desired, it can only be performed after library preparation.

29. Be sure to dry the pellet only until it is just dry. The pellet will look matte with no shine. If the pellet is under-dried, ethanol will remain in the sample wells. The ethanol might reduce the amplified cDNA recovery rate and ultimately the cDNA yield. Allow the plate to sit at room temperature until the pellet is dry. If the pellet is over-dried, there will be cracks in the pellet. It will take longer than 2 min to rehydrate (in the next step) and may reduce amplified cDNA recovery and yield.

30. The optimal cDNA input for Nextera XT library preparation is 100–300 pg. A larger amount of cDNA will generate libraries that are too large for sequencing on an Illumina instrument. The protocol below uses 125 pg of cDNA (in a volume of

1.25 μl), but any input between 100 and 300 pg will work. If all samples are correctly quantified and normalized to a uniform input amount before Nextera XT library preparation, sequencing libraries can be pooled before cleanup, and a relatively uniform amount of sequencing reads will be obtained. However, sample-to-sample read coverage varies, and one may observe some underrepresented or overrepresented samples with the pooling option. Always use a minimum of 2 μl of cDNA to make dilutions. Samples containing less than 100 pg/μl can still be used without dilution, but one may get fewer reads than for other samples if pooled for cleanup. If negative controls are going to be sequenced, they should be used without dilution.

31. If processing a large volume of samples, aliquot equal amounts of Tagmentation Premix into each tube of an eight-tube strip and then use an eight-channel pipette to distribute the Tagmentation Premix.
32. If processing a large number of samples, aliquot equal amounts of NPM into each tube of an eight-tube strip, and then use an eight-channel pipette to distribute the NPM.
33. Consult Illumina literature (Index Adapters Pooling Guide 1000000041074, available at https://support.illumina.com/downloads/index-adapters-pooling-guide-1000000041074.html) for proper index primer selection before proceeding to PCR amplification of the tagmented cDNA.
34. PCR-amplified libraries can be purified individually, or, optionally, the libraries can be pooled if the input cDNA was quantified and normalized to a uniform input amount before library preparation. The libraries are then purified by immobilization on NucleoMag NGS cleanup and size select (available from Takara Bio) beads. The beads are then washed with 80% ethanol, and then the cDNA is eluted with nuclease-free water.
35. The beads are viscous; pipette the entire volume and push it out slowly.
36. If 96 libraries were pooled, elute using 96 μl of nuclease-free water. Nuclease-free water volume is 50% of the original pool volume. If libraries were cleaned up individually, elute in 25 μl of nuclease-free water.
37. The SMART-Seq Single-Cell library is generated using Nextera XT and can be treated as a regular Nextera XT library. The libraries contain p5-i5-read 1 and p7-i7-read 2 sequences on 5′ and 3′ ends, respectively. The recommended read length is 75 bp. The sequencing depth depends on the number of pooled libraries and sample complexity. General considerations regarding sequencing depth and QC are discussed in AIRR Community method chapters "Bulk gDNA Sequencing of

Antibody Heavy-Chain Gene Rearrangements for the Detection and Analysis of B-Cell Clone Distribution" and "Bulk Sequencing from mRNA with UMI for Evaluation of B-Cell Isotype and Clonal Evolution," in **Note 15** and on the Illumina website.

38. Several libraries are usually sequenced together, and each library is individually barcoded. The raw sequencing data obtained from Illumina sequencers are in BCL file format, which contains the raw base calls for all the libraries of the sequencing run. The first step of the demultiplexing consists of splitting the whole BCL file into individual FASTQ files. Each of the resulting FASTQ files will contain the sequence reads linked to a quality score for each base. As sequencing is performed in a paired-end manner and on two lanes of the sequencer, the command will generate four FASTQ files for each library (two lanes x two read directions). To increase reproducibility, it is recommended to start the processing from the original BCL files, even if demultiplexed FASTQ files are provided by the sequencing provider. It is recommended to initially perform the analysis strategy separately for each sequencing lane, until quality control steps. If no issues have been found on the different lanes, the analysis can be repeated on the complete dataset. To extract the reads from one or more specific lanes, argument `--lanes=1` (for lane 1) and `--lanes=2` (for lane 2) can be used.
39. The commands cellranger vdj and cellranger count will generate several outputs, including an HTML report for each type of library, a vloupe.vloupe file for an exploratory analysis, a detailed description of each identified VDJ rearrangements in the files airr_rearrangement.tsv (annotated contigs and consensus sequences of VDJ rearrangements in the AIRR TSV format) and clonotype.csv (high-level descriptions of each clonotype), and a features.tsv.gz file containing the gene expression counts. The last two files will be used for in-depth analysis.
40. For a complete reference, see the 10x Genomics support website: https://support.10xgenomics.com/single-cell-vdj/software/pipelines/latest/using/vdj.
41. The example mentioned here is available on the 10x Genomics website. The command will provide the VDJ and UMI counts for each cell within the library. The main pipeline output files will be in a subfolder named "outs" with the folder indicated by "`--id`". Likewise, the quantification of gene expression can be performed using cellranger count, which will provide the gene and UMI counts per gene for each cell within the library.

42. The choice of reference libraries and cell numbers should be carefully considered when processing 10× sequencing data. Reference libraries: The read alignment procedures require appropriate reference sequences for VDJ and gene expression profiling. For human transcriptome data, "refdata-gex-GRCh38–2020-A" should be used, and for human AIRR-seq data, "refdata-cellranger-vdj-GRCh38-alts-ensembl-5.0.0" is recommended; both libraries are available from 10x Genomics. Cell numbers: When performing a single-cell experiment, a given number of cells have been loaded for the library preparation. Cell Ranger automatically detects unique cells, assuming a high RNA content per cell. However, in heterogeneous populations, cells with low RNA content could accidentally be discarded. To circumvent this, Cell Ranger allows to indicate the expected cell number using the --expect-cells parameter (default: 3000 cells).
43. The Loupe Browser (version 5.0.1) software is a complementary tool provided by 10×. This tool uses the vloupe.vloupe file and can run on MacOS or Windows.
44. In-depth analysis can be done using various tools, mainly R packages. The Seurat package [10] can be used to analyze the gene expression data obtained. The TR sequences can be analyzed using the scRepertoire package [11]. The input for the scRepertoire R package is the *filtered_contig_annotations.csv* file from Cell Ranger. The individual files can be combined into the list element, and the list is then used as an input for the *combineTCR/BCR* function, which combines individual chains into clones for each cell. One can also exclude single-chain clones, or filter the chains based on their UMI count in the case of multiple chains of the same type in a single cell by setting different parameters of the function. There are several definitions of a clone based on the user's needs, which one can choose with the *cloneCall* parameter of the various functions. The package provides analysis for a number of modalities, including calling clones, clonal space/homeostasis, clonal diversity, and repertoire overlap between samples. Moreover, the package allows the investigation of the clones with regard to the cellular gene expression by combining the clones with Seurat objects by using *combineExpression* function. One needs to pay attention to consistent naming of the barcodes between the Seurat and scRepertoire objects. Scirpy is another recently developed Python toolkit that allows simple and straightforward analysis of and visualization of scAIRR data and also allows integration with transcriptome data [12]. VDJView is also a recently published tool providing similarly multi-omic integration [13].

45. The two main steps in the TraCeR pipeline are assemble and summarise. *Assemble*. Sequencing reads derived from TRs are extracted from single-cell RNA-seq data by aligning the reads from each cell against synthetic TR rearrangements made up of all possible combinations of V and J genes with a masked junctional region. TR-derived reads are then assembled into paired full-length TR sequences and characterized by the V and J gene usage and junctional sequence. TraCeR reports both the full nucleotide sequence of reconstructed TR rearrangements, as well as a shorter sequence identifier (such as TRBV31_AG TCTTGACACAAGA _TRBJ2–5) for each rearrangement. *Summarize*. Sequence identifiers for each cell are used to identify groups of clonally related T cells. Both productive and nonproductive TR rearrangements are considered clonally related if they have the same sequence identifier (thus the same V gene, J gene, and junctional nucleotide sequence). Clonal relationships are visualized as clonal networks, where each cell is represented by a node in the graph. Reconstructed TR rearrangements are represented within nodes by horizontal lines or sequence identifiers colored according to locus and productivity. Edges between the nodes represent shared TR rearrangements and are also colored according to locus.
46. Data usable in TraCeR needs to fulfil the following demands. TraCeR requires data from a sequencing library protocol that generates sequencing reads from full-length mRNA transcripts, for example, SMART-Seq. Sequencing approaches that only provide partial coverage of transcript ends are not suitable for TR reconstruction by TraCeR. TraCeR works with both paired-end and single-end reads, although paired-end reads give the highest accuracy and reconstruction rate. The library must be sequenced with a minimum read length of 50 bases. The recommended read depth is 250,000–500,000 paired-end reads per cell (in total 0.5–1 million reads per cell), depending on the cell activation state and how much information about the rest of the transcriptome is needed.
47. TraCeR expects raw reads to be demultiplexed according to cell of origin. This means that two files (one for R1 and one for R2) for each cell are expected for paired-end output and one file per cell for single-end output. TraCeR accepts FASTQ (fastq or fastq.gz) files as input.
48. TraCeR reconstructs TR sequences from raw reads when run in *assemble* mode. The reconstruction is performed separately for

each cell but may be performed for multiple cells in parallel if running TraCeR on a computational cluster.

49. It is important to remove cell duplets or multiplets from single-cell data. TraCeR should only be applied to single-cell data, as the tool assumes all input reads provided to the *assemble* step to derive from one cell. A single T cell should have no more than two reconstructed TR rearrangements per TR locus, and the number of reconstructed TR rearrangements for a given locus may thus be used to identify cell duplets/multiplets or potential contaminations.
50. Whether sharing of rearranged TR sequences between cells is defined as evidence of clonality depends on the type of TR chain and how strict the definition of clonality should be. All TR rearrangements present in a given cell may not always be reconstructed as the detection sensitivity is never 100% and depends on experimental as well as biological factors. Thus, clone groups can be inferred based on various levels of evidence. We therefore recommend that clone groups should be interpreted by the user depending on the specific biological questions of interest. Some typical patterns are as follows: (1) Clone groups consisting of cells sharing the same TR rearrangements, including a productive TRA and TRB (or TRD and TRG), are clonally related with a high confidence. Sharing of additionally rearranged sequences when these are detected in cells (such as a nonproductive rearrangement) further increases the confidence. (2) Sometimes clone groups consisting of subclones may exist due to the possibility of a TRB rearrangement to be found in combination with different TRA rearrangements due to the TRB rearrangement preceding cell proliferation and TRA rearrangement. Such patterns can be seen as cells all sharing the same TRB rearrangement, with clusters of cells within the clone sharing TRA rearrangements. Such patterns may indicate that the TRB is particularly important for antigen specificity of the clone. It is up to the user whether such composite clones should be divided into subclones sharing the same TRA or not. (3) Because of unsuccessful reconstruction of all TR rearrangements in some cells, it may occur that the only shared TR rearrangement between a cell and a clone group is nonproductive or a TRA. In such cases it cannot be certain that the cell belongs to the clone group, but there is also no evidence to the contrary. Such cells may be excluded manually by the user if strict definitions of clones are desired.

## Acknowledgments

We thank Magnolia Bostick, Christian Busse, Eline T. Luning Prak, Gloria Kraus, Chaim Schramm, Nicolas Tchitchek, Ulrik Stervbo, and Johannes Trück for helpful discussions and inspiration around the manuscript and for editing and proofreading and Elaine Chen for contributing figures. EMF and KME were funded by iMAP (ANR-16-RHUS-0001), Transimmunom LabEX (ANR-11-IDEX-0004-02), TriPoD ERC Research Advanced Grant (Fp7-IdEAS-ErC-322856), AIR-MI (ANR-18-ECVD-0001), iReceptor-Plus (H2020 Research and Innovation Programme 825821), and SirocCo (ANR-21-CO12-0005-01) grants. AE is supported by grants from the Deutsche Forschungsgemeinschaft (BO 3429/3-1 and BO 3429/4-1) and the BMBF (RESET-AID). IL is funded by KG Jebsen (project SKGJ-MED-017). Conflict of interest: AE, EMF, KME, IL, NG, and SR declare no conflict of interest. KM is an employee at 10x Genomics, Pleasanton, CA, USA, and NG is an employee at Takara Bio, Mountain View, CA, USA. Both companies produce a kit described in this protocol.

### References

1. Kantor AB, Merrill CE, Herzenberg LA, Hillson JL (1997) An unbiased analysis of V(H)-D-J(H) sequences from B-1a, B-1b, and conventional B cells. J Immunol 158:1175–1186
2. Brezinschek HP, Brezinschek RI, Lipsky PE (1995) Analysis of the heavy chain repertoire of human peripheral B cells using single-cell polymerase chain reaction. J Immunol 155: 190–202
3. de Simone M, Rossetti G, Pagani M (2018) Single cell T cell receptor sequencing: techniques and future challenges. Front Immunol 9: 1638
4. Wardemann H, Busse CE (2017) Novel approaches to analyze immunoglobulin repertoires. Trends Immunol 38:471–482
5. Fuchs YF, Sharma V, Eugster A, Kraus G, Morgenstern R, Dahl A, Reinhardt S, Petzold A, Lindner A, Löbel D, Bonifacio E (2019) Gene expression-based identification of antigen-responsive CD8+ T Cells on a single-cell level. Frontiers in Immunology. https://doi.org/10.3389/fimmu.2019.02568
6. Stubbington MJT, Lönnberg T, Proserpio V, Clare S, Speak AO, Dougan G et al (2016) T cell fate and clonality inference from single-cell transcriptomes. Nat Methods 13:329–332
7. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10
8. Lindeman I, Stubbington MJT (2019) Antigen receptor sequence reconstruction and clonality inference from scRNA-seq data. Methods Mol Biol 1935:223–249
9. Schuldt NJ, Binstadt BA (2019) Dual TCR T Cells: Identity Crisis or Multitaskers? J Immunol 202:637–644
10. Satija R, Farrell JA, Gennert D, Schier AF, Regev A (2015) Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 33:495–502

11. Borcherding N, Bormann NL, Kraus G (2020) scRepertoire: an R-based toolkit for single-cell immune receptor analysis. F1000Res 9:47
12. Sturm G, Szabo T, Fotakis G, Haider M, Rieder D, Trajanoski Z (2020) Scirpy: a Scanpy extension for analyzing single-cell T-cell receptor-sequencing data. Bioinformatics 36: 4817–4818
13. Samir J, Rizzetto S, Gupta M, Luciani F (2020) Exploring and analysing single cell multi-omics data with VDJView. BMC Med Genomics 13:29

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.