



## Combined Analysis of Transcriptome and T-Cell Receptor Alpha and Beta (TRA/TRB) Repertoire in Paucicellular Samples at the Single-Cell Level

Nicolle H. R. Litjens, Anton W. Langerak, Zakia Azmani, Xander den Dekker, Michiel G. H. Betjes, Rutger W. W. Brouwer, and Wilfred F. J. van IJcken

### Abstract

With the advent of next-generation sequencing (NGS) methodologies, the total repertoires of B and T cells can be disclosed in much more detail than ever before. Even though many of these strategies do provide in-depth and high-resolution information of the immunoglobulin (IG) and/or T-cell receptor (TR) repertoire, one clear disadvantage is that the IG/TR profiles cannot be connected to individual cells. Single-cell technologies do allow to study the IG/TR repertoire at the individual cell level. This is especially relevant in cell samples in which much heterogeneity of the cell population is expected. By combining the IG/TR repertoire with transcriptome data, the reactivity of the B or T cell can be associated with activation or maturation stages. An additional advantage of such single-cell technologies is that the combination of both IG and both TR chains can be studied on a per cell basis, which better reflects the antigen receptor reactivity of cells. Here we present the ICELL8 single-cell method for the parallel analysis of the TR repertoire and transcriptome, which is especially useful in samples that contain relatively few cells.

**Key words** T-cell receptor alpha, T-cell receptor beta, Repertoire, Transcriptome, Single cell, Next-generation sequencing

---

## 1 Introduction

T cells recognize antigens via unique T-cell receptor (TCR) molecules. Approximately 95% of T cells express a TCR $\alpha\beta$  receptor, consisting of a TCR $\alpha$  and a TCR $\beta$  chain, whereas the remaining 5% possess a TCR $\gamma\delta$  receptor, consisting of a TCR $\gamma$  and a TCR $\delta$  chain. All four TCR chains are highly diverse in their variable domains. Diversity in these variable domains arises from complex recombination processes involving V, D, and J genes in the TCR chain-encoding loci [1]. In this way the V(D)J recombination

process generates a huge TR repertoire diversity, which is especially apparent in the V(D)J junction. The V(D)J junction is one of the complementarity-determining regions (i.e., CDR3) of the variable domain, which collectively mediate the specific recognition of antigens. Estimates of the number of possible different TCR $\alpha\beta$  receptors amount to  $10^{12}$  molecules [2, 3]. Importantly, whereas antigen-inexperienced or naïve T cells have a broad, unselected TCR repertoire [4], antigen-experienced or memory T cells generally contain more narrow TCR repertoires, mostly consisting of particular antigen-selected specificities.

Historically, TCR repertoire diversity assays have mostly focused on TCR $\beta$  (TRB) chain profiling. Varying from DNA- [5] or RNA-based [6] TRB bulk sequencing assays to flow cytometry-based single-cell TCRV $\beta$  approaches [7], all suffer from drawbacks. A major disadvantage of bulk sequencing approaches is the large number of cells required, whereas flow cytometry-based TCRV $\beta$  assays suffer from the limitation that the 24 different TCRV $\beta$  antibodies collectively cover only 70% of the normal human TCRV $\beta$  repertoire. Moreover, neither of these approaches allows to evaluate the actual composition of the total TCR $\alpha\beta$  receptor, as no information on TCR $\alpha$  (TRA) profiles is obtained. Most importantly perhaps, with any of these approaches, it remains difficult to examine changes in TCR $\alpha\beta$  repertoire diversity within a heterogeneous pool of T cells or low-abundant population like antigen-specific T cells without purifying them first and/or acquiring large enough numbers of cells.

Over the last 5 years, single-cell transcriptomics has become a popular approach, as it allows to detect the heterogeneity in gene expression among individual cells and the discovery of small subpopulations [8]. The combination of single-cell transcriptomics with TR transcript sequencing provides gene expression and TCR repertoire information at the single-cell level. Several platforms exist for single-cell-combined TCR repertoire and transcriptomics analysis, including 10 $\times$  Genomics and more recently the ICELL8 single-cell system [9, 10]. Typically, single-cell transcriptomics requires 5–10 K cells [11–13], but little is known about the possibilities of single-cell-based molecular tools for questioning clinically relevant paucicellular samples [9, 10].

Here we describe a method for the combined evaluation of the transcriptome and TRA/TRB repertoire at the single-cell level in clinical samples with low cell numbers. The method covers all the steps from cell dispensation using the ICELL8 single-cell system, double cDNA preparation at the single-cell level, parallel sequencing of transcript and TRA/TRB sequencing libraries, to data evaluation.

---

## 2 Materials

### 2.1 Sample Preparation

1. 15 mL polypropylene tubes.
2. 5 mL polystyrene tube with cell strainer cap.
3. RPMI-1640 medium without L-glutamine.
4. Penicillin/streptomycin (pen/strep) ( $10^4$  U and  $10^4$   $\mu\text{g}/\text{mL}$  stock).
5. DNase I (Sigma-Aldrich; 10 mg/mL stock).
6. Phosphate buffered Saline ( $1\times$ PBS) without  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  pH 7.4 (Invitrogen), sterilized and degassed for at least 1 h using a vacuum system.
7. Fetal bovine serum (FBS) heat-inactivated (HI, 30 min at  $56^\circ\text{C}$ ).
8. PAN T cell isolation kit (Miltenyi Biotec).
9. AutoMacs rinsing solution (Miltenyi Biotec).
10. AutoMacs washing solution (Miltenyi Biotec).
11. Bovine serum albumin (BSA, sterile filtered, Sigma-Aldrich; 15% stock solution in AutoMacs rinsing solution).
12. Miltenyi Biotec Automacs Pro Cell Sorter.
13. BD FACSCANTO II flow cytometer (Becton Dickinson).
14. Brilliant Violet (BV)510-labelled antihuman CD3 (Becton Dickinson).
15. Trypan blue (Sigma-Aldrich; 0.4% 0.2  $\mu\text{M}$  filtered before use).
16. Burkert counting chamber.

### 2.2 Cell Dispensation

1. ICELL8 single-cell system (Takara Bio).
2. ICELL8 Collection Kit (Takara Bio).
3. ICELL8 Loading Kit (Takara Bio).
4. Biometra Advanced Thermocycler (Westburg).
5. HulaMixer Sample Mixer (Thermo Fisher Scientific).
6. 2100 Bioanalyzer instrument (Agilent Technologies).
7. Varioskan 3001 microplate reader (Thermo Fisher Scientific).
8. MSND 384-well plates and seals, 20 packs (Takara Bio).
9. Nuclease-free 0.2 mL PCR tubes.
10. Biotix nuclease-free LoBind 1.5 mL microcentrifuge tubes (VWR).
11. 15 mL polypropylene tubes.
12. Magnetic separator/PCR strip (Takara Bio).
13. ICELL8 Human TCR a/b Profiling Reagent Kit (Takara Bio).

14. ICELL8 Human TCR a/b Profiling/Indexing Primer Set (Takara Bio).
15. ICELL8 TCR chip (Takara Bio).
16. Nextera XT DNA Library Preparation Kit (24 samples; Illumina).
17. Nextera XT Index Kit (24 indexes, 96 samples; Illumina).
18. Ethanol absolute ( $\geq 99.8\%$ ).
19. AccuGENE molecular biology water (Westburg).
20. Helium ( $\geq 99.9\%$  purity).
21. Phosphate buffered saline ( $1 \times$  PBS) without  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  pH 7.4 (Invitrogen).
22. ReadyProbes Cell Viability Imaging Kit: blue/red contains Hoechst 33342 and propidium iodide (Thermo Fisher Scientific).
23. NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel).
24. Agencourt AMPure XP (Beckman Coulter).
25. Quant-iT™ dsDNA Assay Kit, high sensitivity (Thermo Fisher Scientific).
26. Agilent High-Sensitivity DNA Kit (Agilent Technologies).

### 2.3 Sequencing

1. Illumina (MiSeq system: HiSeq 2500, 3000, or 4000 or Next-Seq 550, 1000, or 2000 system).
2. Paired end 600-Cycle Sequencing Kit (Illumina).
3. HiSeq Rapid SBS Kit v2 (50 cycles) or equivalent (for sequencers other than HiSeq1500, 2500) (Illumina).
4. PhiX (Illumina).

### 2.4 Analysis

Linux (virtual) machine with at least 16 GB RAM memory and the following software installed:

1. Python3 (version  $\geq 3.6$ ).
2. pyngs (<https://github.com/erasmus-center-for-biomics/pyngs>).
3. pysc (<https://github.com/erasmus-center-for-biomics/pysc>).
4. Snakemake (version  $\geq 5.28.0$ ).
5. BCL2Fastq2 (<https://emea.support.illumina.com/downloads/bcl2fastq-conversion-software-v2-20.html>).
6. AdapterTrimmer (<https://github.com/erasmus-center-for-biomics/AdapterTrimmer>).
7. IgBLAST (Ye et al., 2013; Reference [14]).
8. Pear (Zhang et al., 2014; Reference [15]).

9. Biomics TCR workflows (<https://github.com/erasmus-center-for-biomics/tcr-workflows>).
10. R (version  $\geq 4.0$ ).
11. tidyverse.
12. Seurat v3 (Stuart et al., 2019; Reference [16]).
13. wesanderson.
14. scales.
15. R studio (version  $>1$ ).
16. R analysis scripts (<https://github.com/erasmus-center-for-biomics/MiMB-R>).

---

### 3 Methods

#### 3.1 Sample Preparation

1. Prepare DNase medium by adding 5 mL of p/s and 5 mL DNase stock solution to 500 mL of RPMI-1640.
2. Bring DNase medium and FBS-HI to 37 °C.
3. Prepare AutoMacs running buffer by adding 50 mL of 15% BSA to 1450 mL of AutoMacs rinsing solution, and bring buffer to room temperature (*see Note 1*).
4. Start up the AutoMacs Pro Cell Sorter according to manufacturer's instruction.
5. Add 5 mL of DNase medium and 1 mL of FBS-HI per vial (1.8 mL) of 10–20 million PBMC to a 15 mL polypropylene tube (use maximal two vials of PBMC per 15 mL tube).
6. Take the required number of vials of peripheral blood mononuclear cells (PBMC) from liquid nitrogen storage or  $-150$  °C freezer (*see Note 2*).
7. Thaw PBMC at 37 °C until only a small clump of ice remains.
8. Add the PBMC suspension dropwise to the 15 mL tube containing 5 mL of DNase medium and 1 mL of FBS-HI.
9. Centrifuge at  $900 \times g$  for 10 min.
10. Discard supernatant, resuspend pellet in DNase medium, add 5 mL of DNase medium.
11. Centrifuge at  $900 \times g$  for 10 min.
12. Discard supernatant and resuspend pellet in 2 mL AutoMacs running buffer.
13. Take 20  $\mu$ L of this cell suspension, mix it with 20  $\mu$ L trypan blue solution, and add 20  $\mu$ L of this mixture to a Burker counting chamber.
14. Count the number of cells and evaluate their viability (*see Note 3*).

15. Fill up the tube using AutoMacs running buffer, and centrifuge at  $900 \times g$  for 10 min.
16. Discard supernatant, resuspend cell pellet, and label PBMC to enrich for T cells in an untouched manner using the PAN T-cell isolation kit according to manufacturer's instruction (*see Note 4*).
17. Following labeling, add AutoMacs running buffer to fill up the tube to 15 mL.
18. Centrifuge cells for 10 min at  $900 \times g$ .
19. Discard supernatant, and resuspend cells in AutoMacs running buffer to a concentration of 20–25 million of cells/mL (*see Note 5*).
20. Filter cells in order to have a single-cell suspension using 5 mL round bottom polystyrene tube with cell strainer snap cap (*see Note 6*).
21. Enrich using the DEPLETES protocol according to manufacturer's instruction.
22. Collect enriched sample from Automacs Pro Cell Sorter, and fill up the tube to 15 mL by adding sterile PBS.
23. Take 20  $\mu\text{L}$  to determine cell number within the enriched fraction similar to **step 13**.
24. Take another 20  $\mu\text{L}$  to evaluate purity using flow cytometry (*see Note 7*).
25. Discard supernatant, and resuspend in degassed PBS to a concentration of  $2\text{--}5 \times 10^4/\text{mL}$  with a minimum of 270–350  $\mu\text{L}/$  sample.

### 3.2 Cell Dispensation

#### 3.2.1 Dispense Instrument Pre-checks

1. Before operating the ICELL8 instrument, check the system according to manufacturer's instructions. Check the water level in the pressure reservoir, humidifier, and wash bottle.
2. Check the helium tank pressure. The regulator should be set to a supply input of  $>435$  psi (30 bar). Replace the tank if the pressure is  $<435$  psi (30 bar) (*see Note 8*).
3. The regulator should have a supply output of 20–30 psi (1.3–2.0 bar) (*see Note 9*).
4. Before starting the cell dispensation, initialize the MSND, preform a daily warm-up, and start the ICELL8 Imaging System according to manufacturer's instructions (*see Note 10*).
5. Pre-freeze the empty chip holder at  $-80$  °C.

#### 3.2.2 Staining of Cell Suspension

1. Mix the cell suspension gently by inverting the tube five times, and transfer the required volume from the center of the tube.
2. Stain the transferred cells with Hoechst 33342 and propidium iodide. Add 40  $\mu\text{L}$  of each dye per mL of washed cells.

3. Incubate and mix the cell stain suspension with the HulaMixer at room temperature for 20 min in the dark.
4. Settings for the HulaMixer: orbital rotation at 15 rpm for 15 s, reciprocal rotation at 45° for 10 s, and the vortexing off (*see Note 11*).

### 3.2.3 Dilution and Dispensation of Cells

1. Thaw the second diluent (100×), ICELL8 fiducial mix (1×), and nuclease-free water on ice.
2. Prepare 100 μL of 5 pg/50 nL positive control Jurkat total RNA by mixing 1 μL second diluent (100×), 1 μL RNase inhibitor (40 U/μL), 97 μL PBS (1× Ca<sup>2+</sup> and Mg<sup>2+</sup> free), and 1 μL control Jurkat RNA (10 ng/μL). Keep the dilution on ice.
3. Prepare 100 μL of the negative control mix by mixing 1 μL second diluent (100×), 1 μL RNase inhibitor (40 U/μL), and 98 μL PBS (1× Ca<sup>2+</sup> and Mg<sup>2+</sup> free). Keep the dilution on ice.
4. Prepare 200 μL of a 2 cell/50 nl (4 × 10<sup>3</sup> cells/mL) cell suspension by mixing 1 μL second diluent (100×), 1 μL RNase inhibitor (40 U/μL), 97 μL PBS (1× Ca<sup>2+</sup> and Mg<sup>2+</sup> free), and 1 μL control Jurkat RNA (10 ng/μL) (*see Note 12*). Keep the dilution on ice.
5. Prepare the 384-well source plate, and start the run to dispense 50 nL of the prepared suspensions into the nanowells according to manufacturer's instructions.

### 3.2.4 Imaging of Cells

1. In this section, images of all 5184 nanowells of the ICELL8 TCR chip are acquired. *See* manufacturer's instructions (Chapter C) on [https://www.takarabio.com/documents/User Manual/ICELL8 Human TCR ab Profiling User Manual/ICELL8 Human TCR ab Profiling User Manual\\_072219.pdf](https://www.takarabio.com/documents/User Manual/ICELL8 Human TCR ab Profiling User Manual/ICELL8 Human TCR ab Profiling User Manual_072219.pdf) (*see Note 13*).

### 3.2.5 Analyzing Nanowells (Blank Chip)

1. In this section, CellSelect Software is used to analyze the images of the ICELL8 TCR chip in order to identify the Poisson value of each sample position. *See* manufacturer's instructions (Chapter D) on [https://www.takarabio.com/documents/User Manual/ICELL8 Human TCR ab Profiling User Manual/ICELL8 Human TCR ab Profiling User Manual\\_072219.pdf](https://www.takarabio.com/documents/User Manual/ICELL8 Human TCR ab Profiling User Manual/ICELL8 Human TCR ab Profiling User Manual_072219.pdf).
2. Proceed up to "Specify sample names," then return to the Summary tab, and determine the Poisson value for each sample position (*see Note 14*).

### 3.2.6 Analyzing Nanowells (Printed Chip)

1. In this section, CellSelect Software is used to analyze the images of the ICELL8 TCR chip in order to identify nanowells containing viable single cells that are suitable for further

processing and analysis via RT-PCR. See manufacturer's instructions (Chapter D) on [https://www.takarabio.com/documents/User Manual/ICELL8 Human TCR ab Profiling User Manual/ICELL8 Human TCR ab Profiling User Manual\\_072219.pdf](https://www.takarabio.com/documents/User Manual/ICELL8 Human TCR ab Profiling User Manual/ICELL8 Human TCR ab Profiling User Manual_072219.pdf) (*see Note 15*).

2. When instructed to use the Manual triage function, use this function for the following:
  - Exclude some wells that were falsely marked as candidate wells.
  - Include a lot of wells that were not included by the software:
    - (a) Go to Wells tab.
    - (b) Sort the wells as follows:
      - “State” (*see Note 16*)
      - “HasDeadCells”
      - “Cells1”.
    - (c) Click on “Manual triage.”
    - (d) If the well is selected for dispense by the software but needs manual exclusion, click on “Reject – Next Well.”
    - (e) If the well is not selected by the by the software but needs manual inclusion, click on “Use – Next Well.”
3. Save the file using a different filename (<Chip ID>\_analysis1.wcd). *See Note 17*.
4. The TCR printed chip contains barcodes in triplicates. Exclude wells that have been selected for dispense, so only one unique barcoded well remains.
  - (a) Go to the “Wells” tab and select all wells.
  - (b) Copy (Ctrl-c) and paste (Ctrl-v) the date into a spreadsheet.
  - (c) In the spreadsheet, sort by “For dispense,” and delete rows that contain “For dispense- FALSE.”
  - (d) Highlight wells that have duplicate barcodes by using the conditional format function.
  - (e) Select the wells that need to be excluded.
  - (f) Switch to the CellSelect software and go to “Wells” tab.
  - (g) Manually highlight the wells that needs to be excluded, and exclude these wells.
5. Save the file using a different filename (<Chip ID>\_final.wcd).
6. Copy the filter file (<Chip ID>\_final.wcd) to the MSND. It is required for dispensing the RT reaction mix.



### 3.3 First and Second Strand cDNA Synthesis

1. Thaw the required components for the first and second strand cDNA synthesis (except the enzyme) on ice. Vortex and spin down briefly.
2. Thaw the chip (without holder) at room temperature for 10 min. Centrifuge the chip at  $3220 \times g$  for 3 min at 4 °C.
3. Take the SMARTScribe reverse transcriptase and cDNA amplification polymerase from the freezer just before use. Gently mix, do not vortex, and spin down briefly.
4. Prepare the RT-PCR mix by mixing 56  $\mu\text{L}$  GC Melt (5 M), 24  $\mu\text{L}$  cDNA amplification dNTP mixture (25 mM), 3.2  $\mu\text{L}$   $\text{MgCl}_2$ , 8.8  $\mu\text{L}$  DTT (100 mM), 30.9  $\mu\text{L}$  SMARTScribe buffer (10 $\times$ ), 33.3  $\mu\text{L}$  cDNA amplification buffer (2 $\times$ ), 5.3  $\mu\text{L}$  Triton X-100 (3%), 3.8  $\mu\text{L}$  Oligo dT Amp primer (100  $\mu\text{M}$ ), and 8.8  $\mu\text{L}$  Amp primer (10  $\mu\text{M}$ ). Mix well by vortexing until the Triton is dissolved.
5. Add 48  $\mu\text{L}$  SMARTScribe reverse transcriptase (100 U/ $\mu\text{L}$ ) and 9.6  $\mu\text{L}$  cDNA amplification polymerase (2 $\times$ ) to the RT-PCR mix. Mix by pipetting up and down six times.
6. Prepare the RT-PCR 384-well source plate, and start the run to dispense 50 nL of RT-PCR reaction mix into selected nanowells according to manufacturer's instructions.
7. Run the RT-PCR reaction in a preheated SmartChip cycler with a heated lid. Use the following PCR program: 3 min at 50 °C; 5 min at 4 °C; 90 min at 42 °C; 2 min at 50 °C and 2 min at 42 °C (two cycles); 15 min at 70 °C; 1 min at 95 °C; 10 s at 98 °C, 30 s at 65 °C, and 3 min at 68 °C (24 cycles); and 10 min at 72 °C and 4 °C on hold. The amplified reactions can be stored at 4 °C overnight.
8. Collect the full-length cDNA extraction from the chip to a clean PCR tube with the collection module according to manufacturer's instructions.
9. Measure the volume of the collected cDNA extraction. The measured volume may contain up to 15% less than the expected volume. If more than 15% loss is observed, it gives an indication that cDNA synthesis could have been performed suboptimally. Proceed with steps in Subheadings 3.4 and 3.5 (cDNA cleanup and cDNA QC validation) (*see Note 18*).
10. The collected full-length cDNA can be stored at -20 °C.

### 3.4 Cleanup and Concentration After Full-Length cDNA Extraction

1. Purify and concentrate the amplified full-length cDNA extraction with the Gel and PCR Cleanup Kit according to manufacturer's instructions.
2. Purify the concentrated full-length cDNA eluted from the Gel and PCR Cleanup Kit with the AMPure XP beads.

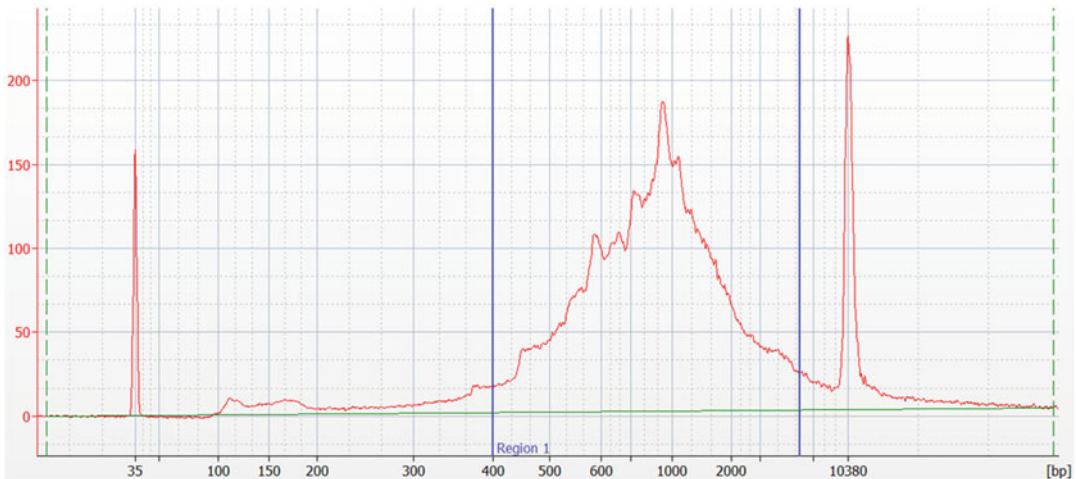
3. Add 30  $\mu\text{L}$  ( $0.6\times$ ) of AMPure XP beads to the cleaned and concentrated cDNA extraction. Mix by pipetting and spin down briefly.
4. Follow the wash steps of the purification according to manufacturer's instructions.
5. Resuspend the dried beads with 15  $\mu\text{L}$  nuclease-free water. Mix by pipetting and spin down briefly.
6. Incubate the sample at room temperature for 5 min.
7. Place the sample on the magnet and incubate for 2 min or until the solution is clear.
8. Transfer 14  $\mu\text{L}$  supernatant containing purified full-length cDNA to a clean PCR tube. The sample can be stored at  $-20\text{ }^{\circ}\text{C}$ .

### **3.5 Validation and Quantification of cDNA**

1. Use 1  $\mu\text{L}$  of the purified cDNA product and dilute to 1:3. Then use 1  $\mu\text{L}$  of the diluted cDNA for quantitation using the Quant-iT™ High-Sensitivity dsDNA Assay Kit. Use the manufacturer's instructions on [https://assets.thermofisher.com/TFS-Assets/LSG/manuals/Quant\\_iT\\_dsDNA\\_HS\\_Assay\\_UG.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/manuals/Quant_iT_dsDNA_HS_Assay_UG.pdf). See **Note 19**.
2. Using the results of the Quant-iT™ High-Sensitivity dsDNA Assay Kit, normalize 1  $\mu\text{L}$  of the purified cDNA product to 2 ng/ $\mu\text{L}$ .
3. Use 1  $\mu\text{L}$  of the normalized cDNA (2 ng/ $\mu\text{L}$ ) to load the Agilent 2100 BioAnalyzer high-sensitivity DNA chip. Use manufacturer's instructions on [https://www.agilent.com/cs/library/usermanuals/public/2100\\_Bioanalyzer\\_HSDNA\\_QSG.pdf](https://www.agilent.com/cs/library/usermanuals/public/2100_Bioanalyzer_HSDNA_QSG.pdf).
4. Determine cDNA QC: compare the results for your sample with Fig. 1 to verify if the sample is suitable for further processing. Proper cDNA synthesis and purification should yield a broad peak spanning ~400 bp to ~6000 bp (Fig. 1; see **Note 20**).

### **3.6 Preparation of TCR a/b Library by Semi-Nested PCR**

1. Thaw the required PCR components for the first PCR reaction (except the enzyme) on ice. Vortex and spin down briefly.
2. Dilute the cDNA to 500 pg/ $\mu\text{L}$ . Transfer 1  $\mu\text{L}$  in a 0.2 mL tube. Keep on ice.
3. Prepare the TCRA +TCRb human primer 1 mix by mixing 4  $\mu\text{L}$  of the TCRA human primer 1 and 2  $\mu\text{L}$  of the TCRb human primer 1. Mix well by vortexing and spin down briefly.
4. Take the TCR amplification polymerase from the freezer just before use. Gently mix, do not vortex, and spin down briefly.



**Fig. 1** Typical BioAnalyzer output of full-length cDNA, showing a broad peak spanning ~400 bp to ~6000 bp

5. Prepare 49  $\mu\text{L}$  PCR 1 mastermix in a 0.2 mL tube by combining 10  $\mu\text{L}$  TCR amplification buffer ( $5\times$ ), 4  $\mu\text{L}$  TCR amplification dNTP mixture (2.5 mM each), 1.25  $\mu\text{L}$  primer P5 (5  $\mu\text{M}$ ), 3  $\mu\text{L}$  TCRA + TCRb human primer 1 premix, 1  $\mu\text{L}$  TCR amplification polymerase, and 29.75  $\mu\text{L}$  nuclease-free water (*see Note 21*).
6. Mix by gently vortexing, and spin down briefly.
7. Add 1  $\mu\text{L}$  of cDNA (500  $\text{pg}/\mu\text{L}$ ) to the PCR 1 mastermix. Mix by pipetting and spin down briefly.
8. Incubate the reaction in a pre-heated thermal cycler with a heated lid. Use the following PCR program: 1 min at 95  $^{\circ}\text{C}$  and 10 s at 98  $^{\circ}\text{C}$ , 15 s at 60  $^{\circ}\text{C}$ , and 45 s at 68  $^{\circ}\text{C}$  (16 cycles) and 4  $^{\circ}\text{C}$  on hold. The tubes may be stored at 4  $^{\circ}\text{C}$  overnight.
9. Thaw the required PCR components for the second PCR reaction (except the enzyme) on ice. Vortex and spin down briefly.
10. Prepare the ICELL8 TCRA +TCRb human primer 2 mix by mixing 4  $\mu\text{L}$  of the TCRA Human Primer 2 Forward HT Index and 2  $\mu\text{L}$  of the TCRb Human Primer 2 Forward HT Index. Mix well by vortexing and spin down briefly.
11. Take the TCR amplification polymerase from the freezer just before use. Gently mix, do not vortex, and spin down briefly.
12. Prepare 46  $\mu\text{L}$  PCR 2 mix in a 0.2 mL tube by combining 10  $\mu\text{L}$  TCR amplification buffer ( $5\times$ ), 4  $\mu\text{L}$  TCR amplification dNTP mixture (2.5 mM each), 1.25  $\mu\text{L}$  primer P5 (5  $\mu\text{M}$ ), 1  $\mu\text{L}$  TCR amplification polymerase, and 29.75  $\mu\text{L}$  nuclease-free water. Mix by gently vortexing and spin down briefly.

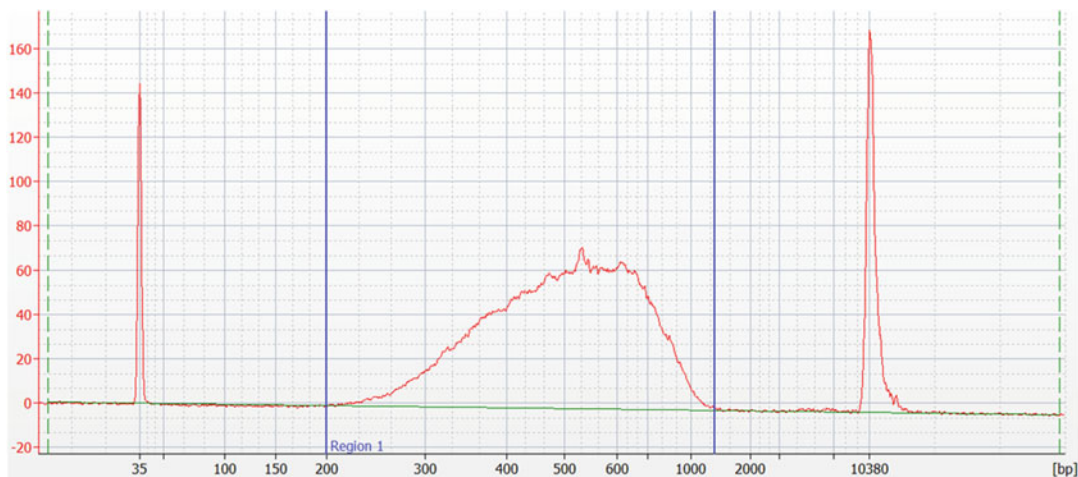
13. Add 1  $\mu\text{L}$  of the amplified product from the first PCR reaction and 3  $\mu\text{L}$  of the TCRA + TCRb Human Primer 2 Reverse HT Index Mix (*see Note 22*).
14. Mix by pipetting and spin down briefly.
15. Incubate the reaction in a pre-heated thermal cycler with a heated lid. Use the following PCR program: 1 min at 95 °C and 10 s at 98 °C, 15 s at 60 °C, and 45 s at 68 °C (14 cycles) and 4 °C on hold. The amplified reactions may be stored at 4 °C overnight.

### 3.7 Purification of TCR a/b Library

1. Purify and size-select the TCR library with the AMPure XP beads.
2. Add 22.5  $\mu\text{L}$  (0.45 $\times$ ) of AMPure XP beads to the TCR library to remove fragments larger than ~900 bp. Mix by pipetting and spin down briefly.
3. Incubate the sample at room temperature for 5 min to let the fragments ~900 bp bind to the beads.
4. Place the sample on the magnet, and incubate for 2 min or until the solution is clear.
5. Transfer the supernatant to a clean PCR tube, and add 10  $\mu\text{L}$  of AMPure XP beads. Mix by pipetting and spin down briefly.
6. Incubate the sample at room temperature for 5 min to let the fragments between ~400 and 900 bp bind to the beads.
7. Follow the wash steps of the purification according to manufacturer's instructions.
8. Resuspend the dried beads in 17.5  $\mu\text{L}$  nuclease-free water. Mix by pipetting and spin down briefly.
9. Incubate the sample at room temperature for 5 min.
10. Place the sample on the magnet, and incubate for 2 min or till the solution is clear.
11. Transfer the supernatant containing purified and size-selected TCR library to a clean PCR tube.
12. The purified TCR library can be stored at  $-20$  °C or keep at 4 °C, and proceed directly to Subheading 3.8 (validation and quantification of TCR a/b library).

### 3.8 Validation and Quantification of TCR a/b Library

1. Use 1  $\mu\text{L}$  of the purified TCR library and dilute to 1:20. Then use 1  $\mu\text{L}$  for quantitation using the Agilent 2100 BioAnalyzer high-sensitivity DNA chip. Use manufacturer's instructions on [https://www.agilent.com/cs/library/usermanuals/public/2100\\_Bioanalyzer\\_HSDNA\\_QSG.pdf](https://www.agilent.com/cs/library/usermanuals/public/2100_Bioanalyzer_HSDNA_QSG.pdf) (*see Note 23*). Determine TCR library QC: compare the results for your samples with Fig. 2 to verify if the sample is suitable for further processing. A proper purified TCR library should yield a broad



**Fig. 2** Typical BioAnalyzer output of purified TCR library, showing a broad peak (550–1200 bp) with a maximum between ~700 bp and ~900 bp

peak spanning 550–1200 bp, with a maximum between ~700 bp and ~900 bp (Fig. 2).

2. Determine TCR library molarity: set the region table to measure between 550 and 1200 bp, and obtain the molarity in pmol/L.
3. Store the TCR library at  $-20^{\circ}\text{C}$  until sequencing (*see Note 24*).

### 3.9 Preparation of 5' differential expression (5' DE) Library

1. Prepare the tagmentation reaction in a 0.2 mL tube (*see Table 1*); keep on ice.
2. Incubate the reaction in a thermal cycler, using the program in Table 2.
3. Add 5  $\mu\text{L}$  of neutralize tagment (NT) buffer to each well.
4. Pipette up and down five times to mix.
5. Incubate for 5 min at room temperature.
6. Prepare the Nextera XT PCR reaction in a 0.2 mL tube (Table 3), and vortex to mix (*see Note 25*).
7. Briefly centrifuge the 0.2 mL tube, and incubate in a thermal cycler, using the program in Table 4.

### 3.10 Purification of 5' DE Library

1. Purify the amplified 5' DE library with the AMPure XP beads (use triple purification).
2. First purification: add 1.0 $\times$  volume (50  $\mu\text{L}$ ) of AMPure XP beads to the previous PCR product (~50  $\mu\text{L}$ ). Mix by pipetting and spin down briefly.
3. Incubate the sample at room temperature for 5 min.

**Table 1**  
**Tagmentation reaction**

Component	Volume ( $\mu\text{L}$ )
Nextera XT, tagment DNA (TD) buffer	10.0
Sample, purified cDNA (0.2 ng/ $\mu\text{L}$ )	5.0
Nextera XT, amplicon tagmentation Mix (ATM)	5.0
<i>Total</i>	<i>20.0</i>

**Table 2**  
**Thermal cycler program of tagmentation reaction**

Step	Number of cycles	Temperature	Time (mm:ss)
1	1	55 °C	5:00
2	1	10 °C	On hold

**Table 3**  
**Nextera XT PCR reaction**

Component	Volume ( $\mu\text{L}$ )
Nextera XT, PCR mastermix (NPM)	15.0
Nextera XT, I7 index primer (orange cap)	5.0
Primer P5 (5 $\mu\text{M}$ ) (TCR Kit)	5.0
Neutralized sample	25.0
<i>Total</i>	<i>50.0</i>

**Table 4**  
**Thermal cycler program for Nextera XT PCR reaction**

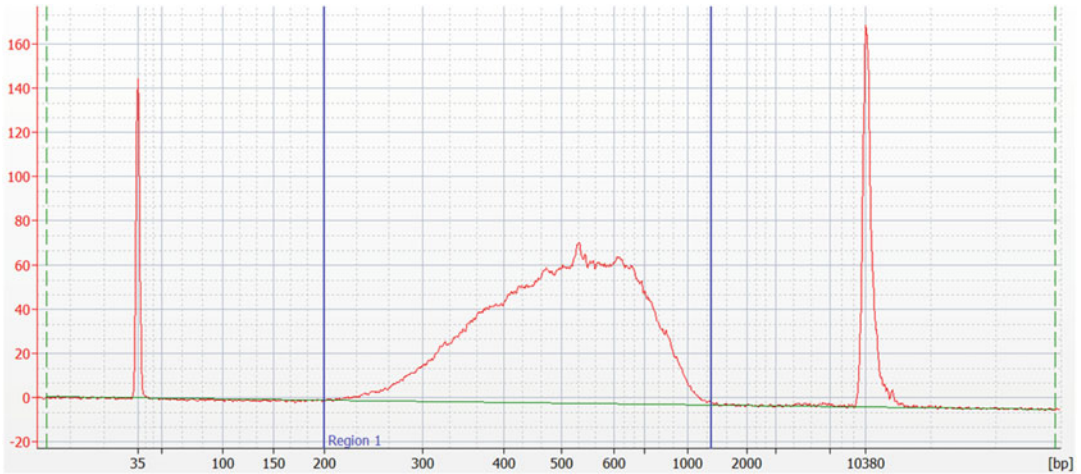
Step	Number of cycles	Temperature	Time (mm:ss)
1	1	72 °C	03:00
2	1	95 °C	00:30
3	12	95 °C	00:10
4		55 °C	00:30
5		72 °C	00:30
6	1	72 °C	05:00
7	1	10 °C	On hold

4. Place the sample on the magnet, and incubate for 2 min or until the solution is clear.
5. Follow the wash steps of the purification according to manufacturer's instructions.
6. Resuspend the beads with 51  $\mu\text{L}$  nuclease-free water, and transfer eluate ( $\sim 50 \mu\text{L}$ ) to a clean tube.
7. Second purification: add  $0.5\times$  volume (25  $\mu\text{L}$ ) of AMPure XP beads to the previous PCR eluate ( $\sim 50 \mu\text{L}$ ) (*see Note 26*). Mix by pipetting and spin down briefly.
8. Incubate the sample at room temperature for 5 min.
9. Place the sample on the magnet, and incubate for 2 min or until the solution is clear.
10. Transfer eluate ( $\sim 75 \mu\text{L}$ ) to a clean tube.
11. Third purification: add  $0.2\times$  volume (15  $\mu\text{L}$ ) of AMPure XP beads to the previous PCR product ( $\sim 75 \mu\text{L}$ ). Mix by pipetting and spin down briefly.
12. Incubate the sample at room temperature for 5 min.
13. Place the sample on the magnet, and incubate for 2 min or until the solution is clear.
14. Follow the wash steps of the purification according to manufacturer's instructions.
15. Resuspend the beads with 13  $\mu\text{L}$  nuclease-free water.
16. The purified 5' DE library can be stored at  $-20^\circ\text{C}$  or keep at  $4^\circ\text{C}$ , and proceed directly to Subheading 3.11 (validation and quantification of 5' DE library).

### **3.11 Validation and Quantification of 5' DE Library**

1. Use 1  $\mu\text{L}$  of the purified 5' DE library and dilute to 1:10. Then use 1  $\mu\text{L}$  for quantitation using the Agilent 2100 BioAnalyzer high-sensitivity DNA chip, following the manufacturer's instructions on [https://www.agilent.com/cs/library/usermanuals/public/2100\\_Bioanalyzer\\_HSDNA\\_QSG.pdf](https://www.agilent.com/cs/library/usermanuals/public/2100_Bioanalyzer_HSDNA_QSG.pdf).
2. *See Note 27.*
3. Determine 5' DE library QC: Compare the results for your sample with Fig. 3 to verify if the sample is suitable for further processing. A proper purified 5' DE library should yield a broad peak spanning 200–1400 bp.
4. Determine 5' DE library molarity: set the region table to measure between 200 and 1400 bp, and obtain the molarity in pmol/L.
5. Store the 5' DE Library at  $-20^\circ\text{C}$  until sequencing (*see Note 28*).





**Fig. 3** Typical BioAnalyzer output of purified 5' DE library, showing a broad peak spanning 200–1400 bp

### 3.12 Next-Generation Sequencing

1. Thaw the TCR and the 5' DE sequencing libraries, and prepare them according to the instructions of Illumina ([www.illumina.com](http://www.illumina.com)).
2. In the case you want to sequence multiple samples on one flow cell, check that the libraries have different P7 indices (*see Note 29*).
3. Pool the selected samples in equal molarities in a single tube.
4. The TCR library has to be sequenced Paired End 300 (600 cycle kit) on a MiSeq, whereas the 5' DE library has to be sequenced either single-read 50 or 75 base pairs on a HiSeq or NextSeq sequencer. Sequencing instructions can be found at [www.illumina.com](http://www.illumina.com).
5. Both single index or dual index sequencing can be applied (*see Note 30*).
6. Add the PhIX control sample to both the TCR, and the 5' DE sequencing runs at 1% of the total output for each flow cell. This enables quantity and quality checks.
7. Proceed with the sequencing procedure as described by the manufacturer. Alternatively, the sequencing procedure can be outsourced to a sequence service provider.
8. Check after sequencing the quality parameters of the PE300 and SR50 sequencing runs. Q30 values should be above 75% for the PE300 reads and above 80% for the single 50 or 75 bp reads.
9. Check after sequencing the yield of the sequencing run. Expected yields vary per sequencer. We advise to obtain at least 5 M clusters for the PE300 run to ensure sufficient data



for subsequent TCR analysis (*see* **Note 31**). For the transcriptome library, a minimal yield of 50 M clusters is advised to ensure sufficient data for transcriptome analysis.

### 3.13 Data Analysis

In the following sections, code will be typeset in a monospaced font.

#### 3.13.1 Primary Analysis of Single-Cell TCR Data

1. Demultiplex the TCR data using the `bcl2fastq2` program from Illumina per the instructions in the BCL2FASTQ2 manual. The I5 indices for these libraries are TCTTTCCC on an Illumina MiSeq sequencer. For other sequencers the reverse complement sequence may be needed. This procedure will result in two files of which the names are constructed as follows: `{x}_S{y}_L001_R{z}_001.fastq.gz`. In this filename, `x` denotes the sample ID, `y` is an arbitrary sample number in the sample sheet, and `z` is the number of the read (1 for forward, 2 for reverse) (*see* **Note 32**).
2. Make a new working directory in which to perform the TCR assignment, and copy the newly generated FastQ files and the well list obtained during well selection over.
3. Go into the new working directory.
4. Create a directory labeled demultiplexed.
5. Create individual FastQ files per well using `pysc` (<https://github.com/erasmusmc-center-for-biomics/pysc>). Both the data start (base 14) and barcode position (read 1 bases 0 to 10) will need to be specified while running this tool as well as a format where to put the output reads. The following command will place the FastQ files per well in the demultiplexed directory:

```
python3/data/Software/python/pysc/bin/pysc demultiplex \
--read_1 TCR_S1_L001_R1_001.fastq.gz \
--read_2 TCR_S1_L001_R2_001.fastq.gz \
--well-list welllist.txt \
--output-read-1 "demultiplexed/{sample}_{row}_{column}_R1.
fastq" \
--output-read-2 "demultiplexed/{sample}_{row}_{column}_R2.
fastq" \
--well-barcode-read 1 \
--well-barcode-start 0 \
--well-barcode-end 10 \
--data-start 14
```

6. Run the TCR snakemake workflow (<https://github.com/erasmusmc-center-for-biomics/tcr-workflows>) which removes sequence adapters introduced during the sample preparation,

merges the forward and reverse reads using PEAR [15], and runs IG-BLAST [14] to align the merged sequences to human V, D, and J genes for each well individually. This workflow yields clonotype reports containing both TRA and TRB structures and their CDR3 sequences per well.

### 3.13.2 Primary Analysis of RNA-Seq Data

1. Demultiplex the single-cell 5' DE RNA-seq data using `bcl2fastq2` in a similar manner to the TCR data.
2. Make a new directory in which to perform the scRNA-seq quantification, and copy the well list and scRNA-seq FastQ files over and move into.
3. In the new directory, make a demultiplexed sub-directory.
4. Assign reads to wells using `pysc`, and place them in a single file per sample. To this end run the following command:

```
python3/data/Software/python/pysc/bin/pysc demultiplex \
--read_1 $file \
--well-list welllist.txt \
--output-read-1 "demultiplexed/{sample}_R1.fastq" \
--well-barcode-start 0 \
--well-barcode-end 10 \
--data-start 14
```

5. Run the expression analysis snakemake workflow (<https://github.com/erasmusmc-center-for-biomics/tcr-workflows>) which removes the adapters introduced during the sample preparation, aligns the reads with HISAT2 [17], converts mapped regions to BED format, and intersects these with transcripts with `bedtools` [18]. Finally, the gene expression per gene per well is quantified.

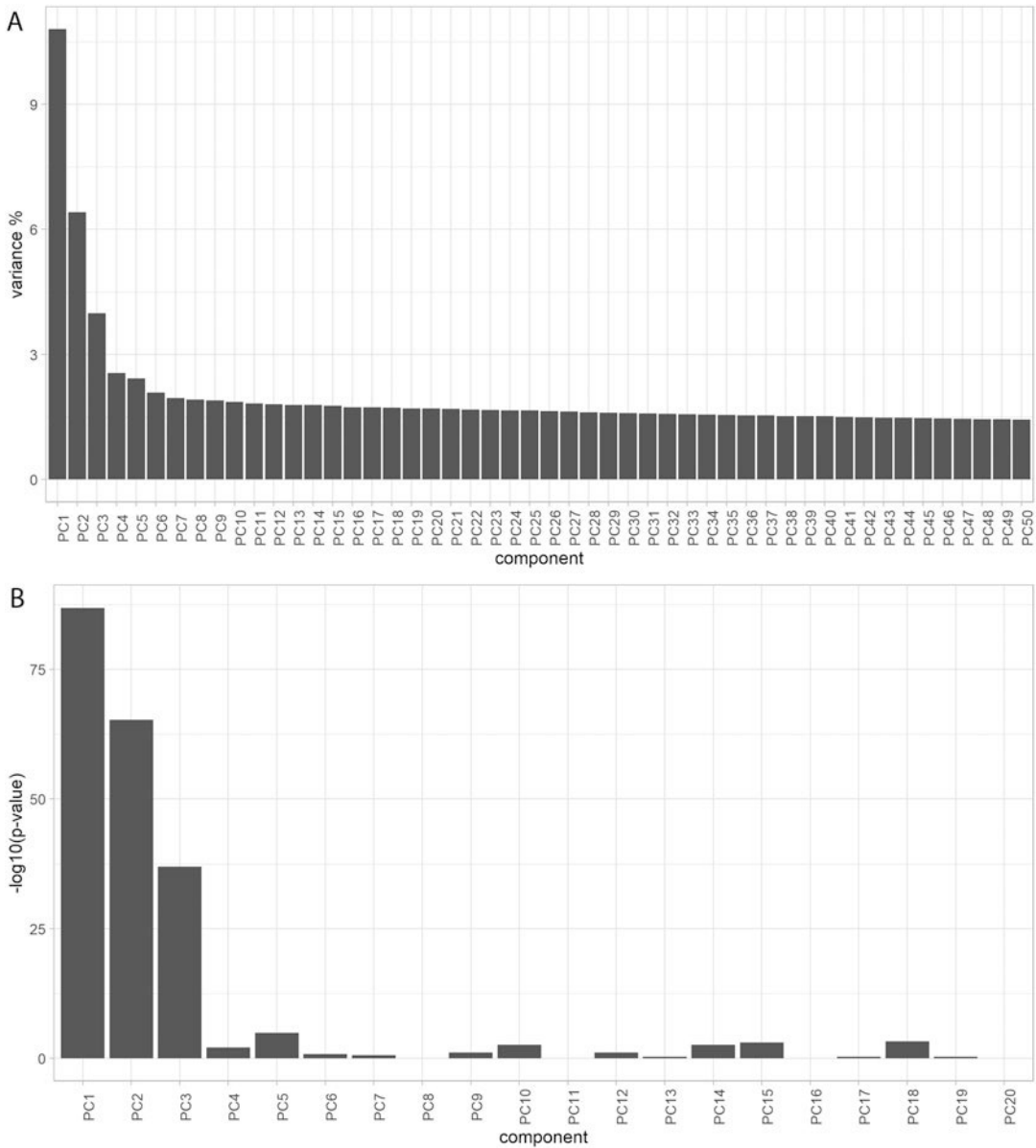
### 3.13.3 Downstream Analysis in R

1. Make a new directory for the downstream analysis and go into this. It is strongly advised to give this folder a meaningful name to distinguish it from other experiments.
2. Make an R sub-directory and data sub-directory, and in this create a directory named `Expression` and a directory named `TCR`.
3. Copy the `sample.exon.tsv.gz` to the `Expression` directory and the `sample_row_column.report.csv` files to the `TCR` directory.
4. Copy the well list to the data directory as `welllist.txt`.
5. Start the RStudio (<https://www.rstudio.com/>), and make a new project from an existing directory. For the directory choose the downstream analysis folder.
6. Copy the files `environment.R`, `seurat_analysis_part1.R`, `seurat_analysis_part2.R` and `tcr_analysis.R` from <https://github.com/erasmusmc-center-for-biomics/MiMB-R> to the R directory.

7. Run the environment.R script by typing `source("R/environment.R")` in the R console. This script will make an R environment with the well list, expression, and TCR data loaded.
8. Run the `seurat_analysis_part1.R` script by typing `source("R/seurat_analysis_part1.R")` in the R console. This script will load the data assembled with environment.R and perform normalization, scaling, principal component analysis, and a Jackstraw analysis. It will create a directory named output with figures and tables.
9. Open the `pca_variance_explained.png` and `jackstraw_score.png` figures in the output directory (Fig. 4), and determine at which principal component the scores decrease sharply. Up to this component, biologically relevant variation is contained (*see Note 33*).
10. Set the `ndim` variable to the component number previously obtained by issuing the following command `ndim <- n`, where *n* should be larger than 1.
11. Run the `seurat_analysis_part2.R` script by typing `source("R/seurat_analysis_part2.R")` in the R console. This script will run both UMAP and *t*-SNE which will place the cells based on their expression profiles on a two-dimensional plane (*see Note 34*). Figures depicting the cell layout are made in the output directory as well as tables with the coordinates per cell obtained for both methods (Fig. 5).
12. Run the TCR analysis script by executing `source("R/tcr_analysis.R")` in the R console. This script will summarize and filter the alignments to the TCR and yield a receptor to number of cells table (*see Note 35*). Two more tables are created in the output directory: one table with the filtered T-cell receptor data (`filtered_tcr.txt`) and one table with the TCR versus the number of cells (`tcr_to_cells.txt`) which gives an indication of sample complexity. Furthermore, the TCR complexity is plotted in a figure in the file `tcr_to_cells.png` (Fig. 6).

### 3.13.4 Projecting Gene Expression Data on Cell Coordinates

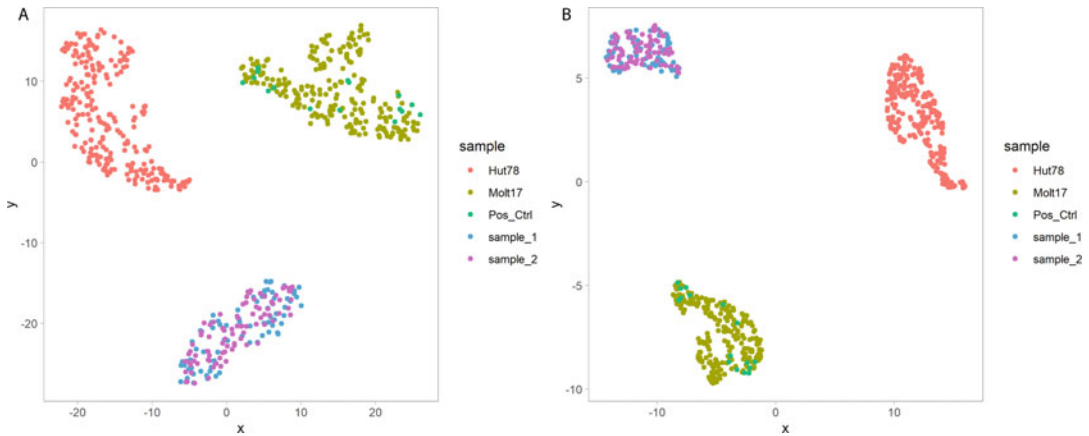
1. Create a variable named `gene` with identifier of the gene of which the expression will be displayed, for example, `gene <- "ENSG00000167286"` for CD3d.
2. Make a variable `coordinate_type` to use either the *t*-SNE (`tsne`) or UMAP (`umap`) coordinates.
3. Run the `project_expression.R` script with `source("R/project_expression.R")`. A file will be created in the output directory with the gene identifier and the coordinate type. Opening this file will display the *t*-SNE or UMAP projection with the expression depicted over it (Fig. 7). The expression is scaled from 0 to 1 with 0 being the cell(s) with the lowest gene expression and 1 the cell(s) the highest expression.



**Fig. 4** Results from the PCA (a) and Jackstraw (b) analyses

3.13.5 *Projecting VDJ Usage Data on Cell Coordinates*

1. Create a variable named locus with either TRA to display the alpha chains or TRB to display the beta chains.
2. Make a variable coordinate\_type to use either the *t*-SNE (tsne) of UMAP (umap) coordinates.
3. Run the project\_vdj.R script with source(“R/project\_vdj.R”). A file will be created in the output directory with the locus and the coordinate type. Opening this file will display the *t*-SNE or UMAP projection with the V(D)J composition projected as the colors (Fig. 8).

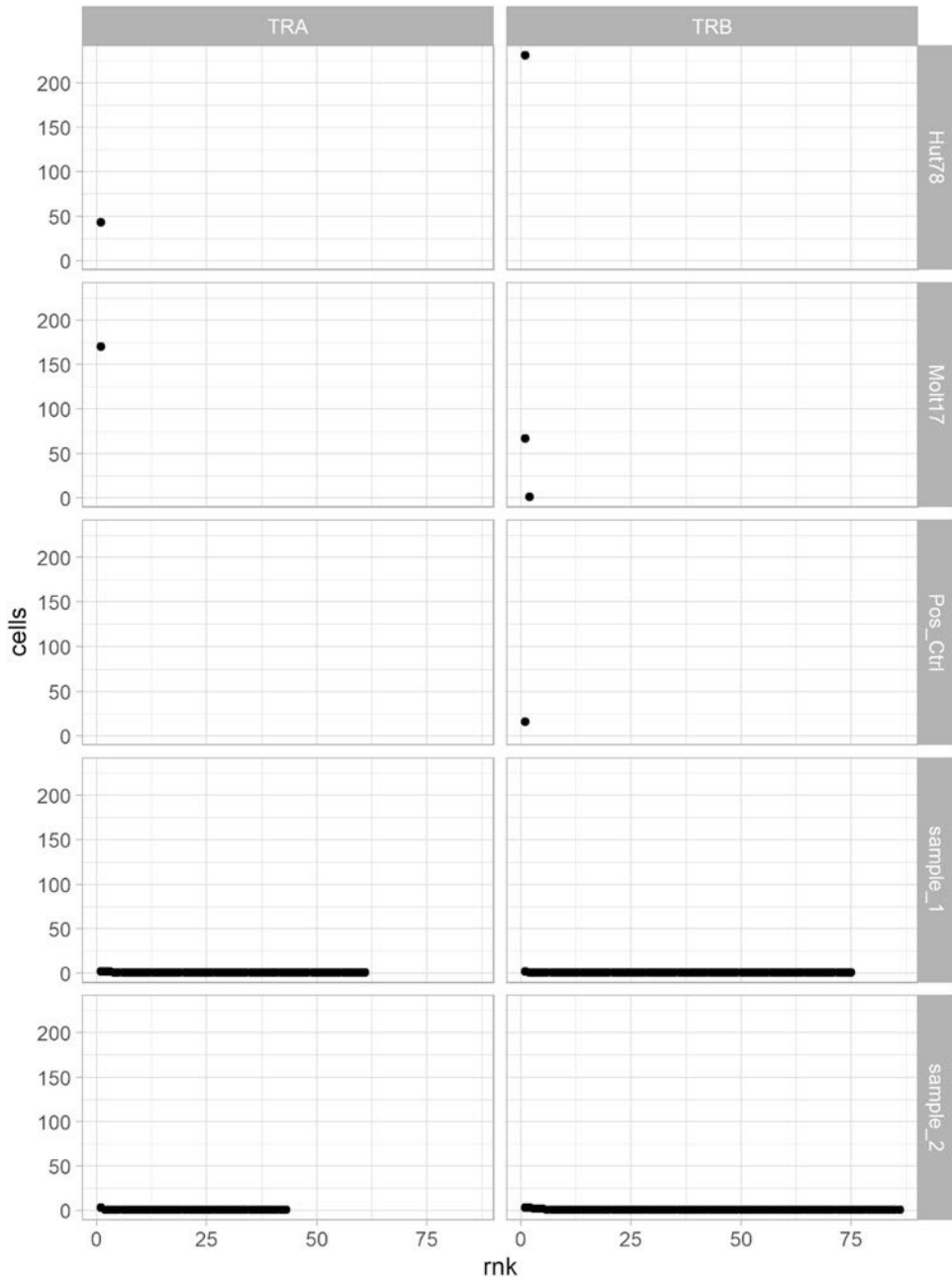


**Fig. 5** Projections of single cells on a two-dimensional plane based on their expression profiles using *t*-SNE (a) and UMAP (b)

## 4 Notes

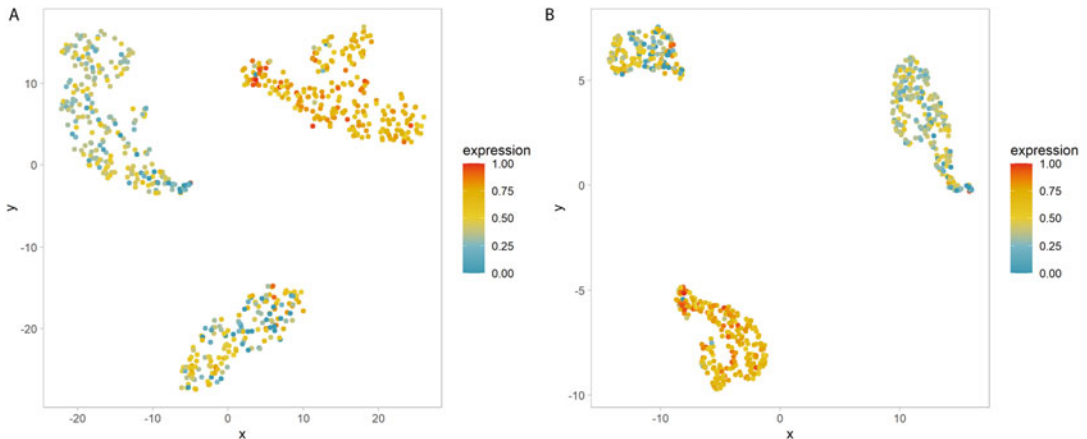
1. This buffer needs to be stored at 4 °C and expires 1 month after adding the BSA stock solution (15%).
2. Use freshly isolated PBMC (obtained following Ficoll gradient centrifugation as described before [19] instead of frozen PBMC for the enrichment of T cells. Moreover one can also enrich for activated (antigen-specific) T cells following stimulation with antigens using, for example, CD137, a costimulatory molecule upregulated upon the interaction of a TCR with antigen presented by antigen-presenting cells using flowcytometry-based cell sorting [20].
3. Determine the number of unstained (living) and blue (dead) cells in duplicate by evaluating number of cells in 25 squares. The number of cells multiplied by 2 (dilution factor) and  $10^4$  return the number of cells/mL. Multiplication by the volume results in total cell number. The number of living cells divided by the total cell count results the fraction of living cells. The number of dead cells divided by the total cell count returns the fraction of dead cells.

Samples containing a lot of dead cells, for example, 50%, do not perform well in the single-cell RNA sequencing. Typically viability should be more than 80%.
4. Both untouched and positive selection of T cells/cells of interest work for downstream processing/combined analysis of transcript and TCRA/TRB sequencing libraries.



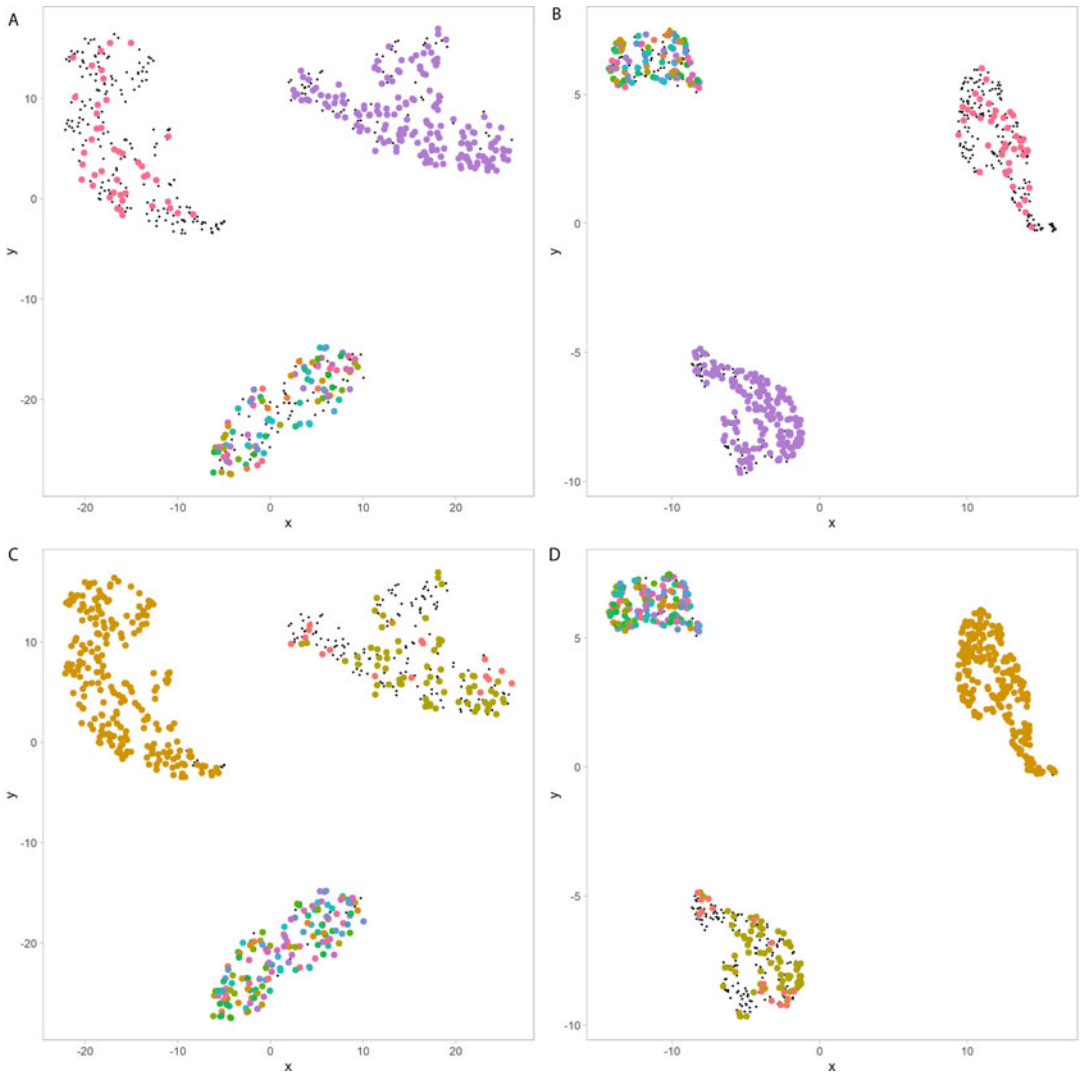
**Fig. 6** The number of cells with the same TCR ordered from the most to the least abundant receptor per sample

5. When having more than 25 million of cells, consider to run a second separation tube instead of increasing the volume of the tube. This will result in better purities of enriched fractions. For more than one separation, use a quick rinse in between samples when they are of the same origin, and rinse if you want to separate a sample of a different origin.



**Fig. 7** Expression of CD3d (ENSG00000167286) projected on cells placed with *t*-SNE (a) and UMAP (b)

6. Only filter if clumps of cells are visible by eye to minimize cell loss due to filtration of cell suspensions. Prewet the filter using AutoMacs running buffer before adding the cell suspension, and rinse the tube afterward. The nylon mesh (pore) size of the cell strainer is 35  $\mu\text{m}$ .
7. Purity evaluation is done by adding 30  $\mu\text{L}$  1  $\times$  PBS to 20  $\mu\text{L}$  of cell suspension and staining using 2  $\mu\text{L}$  of BV510-labelled antihuman CD3 for 15 min at room temperature. Following a wash ( $900 \times g$  for 5 min), the supernatant is discarded, pellet resuspended, and the sample measured on the BD FACSCanto II. The proportion of T cells within the live gate represents the purity of the cell sample. Purities over 95% give the best results as most of the cells represent the cells of interest, i.e., T cells.
8. One dispense uses  $\sim 73$  psi (5 bar). Replace earlier if more experiments are planned.
9. Be aware that 35 psi (2.4 bar) should be the very max. Pressure above 35 psi (2.4 bar) will blow the overpressure valve and cause a Helium leakage.
10. The fluorescence light source requires a warming up period of  $\sim 5$  min. After switching it on, wait at least 5 min before switching it off.
11. Do not vortex or centrifuge the cells. Only use 200  $\mu\text{L}$  and 1000  $\mu\text{L}$  pipets.
12. The manual advises to make a 1 cell/50 nL solution; however, in practice cell counts are usually overestimated; therefore, we make a 2 cell/50 nL solution. After the blank chip, the solution will be diluted based on the Poisson information. Prepare enough volume for the blank and printed chip; 100  $\mu\text{L}$  is required per source plate well. If there are  $< 8$  biological samples for the chip, for example four, prepare at least 300  $\mu\text{L}$  cell



**Fig. 8** TCR V(D)J composition projected on cells placed with *t*-SNE (**a, c**) and UMAP (**b, d**) for TRA (**a, b**) and TRB (**c, d**)

suspension per sample (100  $\mu$ L for the blank chip and 200  $\mu$ L for the barcoded chip). The blank chip can be re-used. Keep the unused wells in the source plate empty. After dispense the wells in the chip that, corresponding to the empty wells from the source plate, will stay clean for the next blank dispense.

13. When prompted for “Run CellSelect with images from: C:\Wafergen\WafergenData For Chip: <Chip ID>?”, click NO. After imaging keep the blank chip. It might contain some empty sample positions, or use as balance chip during centrifugation. Store printed chip at  $-80\text{ }^{\circ}\text{C}$ ; this step can also be performed with non-pre-chilled holders.



14. A Poisson value of 0.8–1.0 means that most wells are filled with a single cell. If the Poisson value on the blank chip is within the 0.1–3.0 range, adjust the cell concentration to 0.8 before dispensing on a printed chip. If the Poisson value is outside the 0.1–3.0 range, adjust cell concentration, and perform new blank chip dispense.
15. Before starting the analysis, make a copy of the data folder containing the images:

From: C:\Wafergen\WafergenData\ < Chip ID>

To: C:\Users\ICELL8\Desktop\Analyzed Images\ < Chip ID>

A maximum of 1728 wells can be selected for further dispense. The downselect function is generally not required.

16. Overview of “States” that might contain wells to include or exclude wells for dispense:
  - Cluster (+LowConfidence): might contain some cells to include
  - Good: might contain some cells to exclude (empty or duplicate)
  - HasDeadCells (+LowConfidence): might contain many cells to include (Cells are sometimes incorrectly marked as “dead” due to false detection of well center; dead cells could be included for analysis as well.)
  - Inconclusive: not likely to contain cells to include
  - LowConfidence: might contain some cells to include
  - MultipleCells (+LowConfidence): not likely to contain cells to include
  - NoCells (+LowConfidence): not likely to contain cells to include
  - TooManyCells: not likely to contain cells to include
17. Additional selection (**steps 19–23** in the manufacturer’s instructions) is generally not required.
18. An additional volume of 3–10% was observed in our experiments; there are no indications that this affects cDNA quality. Proceed with steps in Subheadings 3.4 and 3.5 (cDNA cleanup and cDNA QC validation).
19. Instead of the Quant-iT™ High-Sensitivity dsDNA Assay Kit, the Denovix instrument (<https://www.denovix.com/>) is also able to quantify the undiluted purified cDNA product. Use

1  $\mu\text{L}$  to load onto the Denovix instrument. This concentration is overestimated because of optical impurities. To compensate for this effect, the concentration outcome must be divided by 2.

20. If the Denovix was used to normalize the purified cDNA product to 2  $\text{ng}/\mu\text{L}$ , then the final concentration must be determined by using the results of the Agilent 2100 BioAnalyzer high-sensitivity DNA chip: set the region table to measure between 400 and 6000 bp, and obtain the yield in  $\text{pg}/\mu\text{L}$ .
21. If the concentration of the cDNA is  $<500$   $\text{pg}/\mu\text{L}$ , add more volume of cDNA and less volume of the nuclease-free water.
22. The index is present in the TCRA and the TCRB human primer 2 reverse primers. When preparing more than one chip, use TCRA and TCRB primers with different indices to enable pooled sequencing.
23. Alternatively, the libraries can be quantified by qPCR using the NGS Library Quantification Kit (Takara Bio, Cat No. 638324). Please refer to the corresponding user manuals for detailed instructions.
24. Following validation, the TCR library is ready for sequencing on the Illumina platforms. It is advised to determine the final molarity of the library by sequencing a small amount of the library first to adjust the sample molarity and perform additional sequencing to the required amount.
25. Do not use the I5 index primer supplied with the Nextera XT Index Kit.
26. In the second purification of the manufacturer's instructions, "50  $\mu\text{L}$  of eluate" is listed. This volume is incomplete as it contains 75  $\mu\text{L}$  (25  $\mu\text{L}$  AMPure XP beads and 50  $\mu\text{L}$  eluate from the first purification). We have used the full eluate volume from the second purification and adjusted the volumes in the third purification: from 10  $\mu\text{L}$  of AMPure XP beads added to ~50  $\mu\text{L}$  eluate to 15  $\mu\text{L}$  of AMPure XP beads added to ~75  $\mu\text{L}$  eluate.
27. Alternatively, the libraries can be quantified by qPCR using the NGS Library Quantification Kit (Takara Bio, Cat No. 638324). Please refer to the corresponding user manuals for detailed instructions.
28. Following validation, the 5' DE library is ready for sequencing on Illumina platforms. Determine final molarity by sequencing a small amount of the library first. Then normalize the sample molarity and perform further sequencing.
29. Preferred index combinations can be found in the sequencing manuals of your sequencing provider.

30. In the case of dual index sequencing, the second index read will be TCTTTCCC, which is part of the NexteraXT adaptor sequence. This sequence is not an actual index as the TCR and 5' DE sequencing libraries are single-indexed libraries.
31. The advised amount of the TCR reads is based on our previously published experiments [21], where 5 M clusters yielded 71% of the single a TCR signature. Increasing the yield will likely increase the percentage of TCR alpha and beta clonotypes but will also increase cost per sample. The advised amount of 50 M cluster for the 5' DE sequencing libraries is based on a minimal amount of 50 k clusters per cell multiplied with 1000 single cells, which is a usual number of cells that can be obtained from a Takara TCR and 5' DE chip.
32. In the unfortunate event that a flow cell yielded an insufficient number of clusters, less than 5 M for a typical experiment, the data files of multiple flow cells can be merged. First, make a new directory to hold the output file, and use the following commands to merge the data over multiple flow cells:

```
zcat \
  {fc_1}/{x}_S{y}_L001_R{z}_001.fastq.gz \ {fc_n}/{x}_S{y}_
_L001_R{z}_001.fastq.gz | \ gzip -c > {dir}/{x}_R{z}.fastq.gz
```

Please note that neither samples nor reads 1 and 2 should be merged together, and the order of the flow cells remains the same while merging reads 1 and 2. The sample number is variable between flow cells and should not be taken into account while merging.

33. At least two dimensions need to be selected, otherwise the subsequent UMAP will fail. For the visualization, a slightly larger number of dimensions will introduce some noise in the figure, but only in extreme cases will it change the overall topology. Choosing too few dimensions with which to continue will result in a loss of variation which is likely biologically significant. For these reasons, often times a few more dimensions are chosen than the absolute minimum based on the PCA and Jackstraw analyses. If the PCA and Jackstraw analyses are not in accordance, choose the largest number of dimensions.
34. The exact placement of the cells on the plane is arbitrary and uses random number generation. To make the results more reproducible, the seed of the random number generator can be set using the command `set.seed`. For example, executing `set.seed(42)` sets the seed of the random number generator to 42.

35. TCR sequences are considered identical, when they are composed of the same V, D, and J genes and have the same CDR3 sequence. As cells should not be counted twice, only the most highly abundant sequence for either TRA or TRB is taken into account.

---

## Acknowledgments

We would like to thank Maaïke de Bie, Amy van der List, and Mariska Klepper for technical assistance.

## References

- Garcia KC, Teyton L, Wilson IA (1999) Structural basis of T cell recognition. *Annu Rev Immunol* 17:369–397
- Nikolich-Zugich J, Slifka MK, Messaoudi I (2004) The many important facets of T-cell repertoire diversity. *Nat Rev Immunol* 4: 123–132
- Miles JJ, Douek DC, Price DA (2011) Bias in the alpha beta T-cell repertoire: implications for disease pathogenesis and vaccination. *Immunol Cell Biol* 89:375–387
- Kohler S, Wagner U, Pierer M, Kimmig S, Oppmann B, Mowes B et al (2005) Post-thymic in vivo proliferation of naive CD4+ T cells constrains the TCR repertoire in healthy human adults. *Eur J Immunol* 35:1987–1994
- van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, Lavender FL et al (2003) Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 concerted action BMH4-CT98–3936. *Leukemia* 17: 2257–2317
- Li B, Li T, Pignon JC, Wang B, Wang J, Shukla SA et al (2016) Landscape of tumor-infiltrating T cell repertoire of human cancers. *Nat Genet* 48:725–732
- Langerak AW, van Den Beemd R, Wolvers-Tettero IL, Boor PP, van Lochem EG, Hooijkaas H et al (2001) Molecular and flow cytometric analysis of the V beta repertoire for clonality assessment in mature TCR alpha beta T-cell proliferations. *Blood* 98:165–173
- Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E et al (2017) Human cell atlas meeting, the human cell atlas. *eLife* 6: 1–30
- Valihrach L, Androvic P, Kubista M (2018) Platforms for single-cell collection and analysis. *Int J Mol Sci* 19:807
- Papalexi E, Satija R (2018) Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol* 18:35–45
- Aarts M, Georgilis A, Beniazza M, Beolchi P, Banito A, Carroll T et al (2017) Coupling shRNA screens with single-cell RNA-seq identifies a dual role for mTOR in reprogramming-induced senescence. *Genes Dev* 31: 2085–2098
- Bergiers I, Andrews T, Vargel Bolukbasi O, Buness A, Janosz E, Lopez-Anguita N et al (2018) Single-cell transcriptomics reveals a new dynamical function of transcription factors during embryonic hematopoiesis. *eLife* 7: e29312
- Goldstein LD, Chen YJ, Dunne J, Mir A, Hubschle H, Guillory J et al (2017) Massively parallel nanowell-based single-cell gene expression profiling. *BMC Genomics* 18:519
- Ye J, Ma N, Madden TL, Ostell JM (2013) IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res* 41(W1):W34–W40
- Zhang JK, Kobert K, Flouri T, Stamatakis A (2014) PEAR: a fast and accurate illumina paired-end ReAd MergeR. *Bioinformatics* 30(5):614–620
- Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Satija R (2019) Comprehensive integration of single-cell data. *Cell* 177(7):1888–1902.e21
- Daehwan K, Paggi JM, Park C, Bennett C, Salzberg SL (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 37(8): 907–915

18. Quinlan AR (2014) BEDTools: the Swiss-Army tool for genome feature analysis. *Curr Protoc Bioinformatics* 47:11.12.1–11.1234
19. Litjens NH, Huisman M, Hijdra D, Lambrecht BM, Stittelaar KJ, Betjes MG (2008) IL-2 producing memory CD4<sup>+</sup> T lymphocytes are closely associated with the generation of IgG-secreting plasma cells. *J Immunol* 181(5):3665–3673
20. Litjens NH, de Wit EA, Baan CC, Betjes MG (2013) Activation-induced CD137 is a fast assay for identification and multi-parameter flow cytometric analysis of alloreactive T cells. *Clin Exp Immunol* 174(1):179–191
21. Litjens NHR, Langerak AW, van der List ACJ, Klepper M, de Bie M, Azmani Z, den Dekker AT, Brouwer RWW, Betjes MGH, Van IJcken WFJ (2020) Validation of a combined transcriptome and T cell receptor alpha/beta (TRA/TRB) repertoire assay at the single cell level for paucicellular samples. *Front Immunol* 11:1999

**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.

