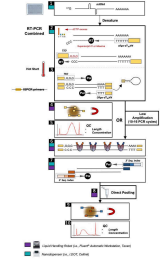


Mar 18, 2022 Version 2

# FLASH-seq protocol V.2

DOI

[dx.doi.org/10.17504/protocols.io.kxygxzkrwv8j/v2](https://dx.doi.org/10.17504/protocols.io.kxygxzkrwv8j/v2)



Simone Picelli<sup>1</sup>, Vincent Hahaut<sup>1</sup>

<sup>1</sup>Institute of Molecular and Clinical Ophthalmology Basel (IOB)

Human Cell Atlas Metho...

The Single Cell Ninjas



**Simone Picelli**

Institute of Molecular and Clinical Ophthalmology Basel (IOB...)

OPEN ACCESS



DOI: [dx.doi.org/10.17504/protocols.io.kxygxzkrwv8j/v2](https://dx.doi.org/10.17504/protocols.io.kxygxzkrwv8j/v2)

External link: <https://www.biorxiv.org/content/10.1101/2021.07.14.452217v1>

**Protocol Citation:** Simone Picelli, Vincent Hahaut 2022. FLASH-seq protocol . **protocols.io**  
<https://dx.doi.org/10.17504/protocols.io.kxygxzkrwv8j/v2> Version created by **Simone Picelli**

**License:** This is an open access protocol distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

**Protocol status:** Working

**We use this protocol and it's working**

**Created:** March 18, 2022

**Last Modified:** March 18, 2022

**Protocol Integer ID:** 59611

**Keywords:** FLASH-Seq, single-cell, RNA-sequencing, full-length, tagmentation

## Abstract

The single-cell RNA-sequencing (scRNA-seq) field has evolved tremendously since the first paper was published back in 2009. While the first methods analysed just a handful of cells, the throughput and performance rapidly increased over a very short timespan. However, it was not until the introduction of emulsion droplets methods, that the robust and reproducible analysis of thousands of cells became feasible. Despite generating data at a speed and a cost per cell that remains unmatched by full-length protocols like Smart-seq, scRNA-seq in droplets still comes with the drawback of addressing only the terminal portion of the transcripts, thus lacking the required sensitivity for comprehensively analyzing the transcriptome of individual cells. Building upon the existing Smart-seq2/3 workflows, we developed FLASH-seq (FS), a new full-length scRNA-seq method capable of detecting a significantly higher number of genes than both previous versions, requiring limited hands-on time and with a great potential for customization.

## Materials

### REAGENTS - CELL LYSIS MIX

☒ dNTP-Set 1 Carl Roth Catalog #K039.2

☒ Triton X-100 Sigma Aldrich Catalog #X100-100ML

☒ Recombinant RNase Inhibitor (40 U/uL) Takara Bio USA, Inc. Catalog #2313B

☒ dCTP (100 mM solution) Thermo Fisher Scientific Catalog #10217016

### REAGENTS - RT-PCR MIX

☒ KAPA HiFi HotStart ReadyMix (2x) Roche Catalog #KK2602

☒ SuperScript™ IV Reverse Transcriptase Thermo Fisher Scientific Catalog #18090050

☒ Magnesium Chloride (1M Solution) Invitrogen - Thermo Fisher Catalog #AM9530G

### REAGENTS - MAGNETIC BEADS SOLUTION PREPARATION

☒ Polyethylenglycol (MW=8000) Sigma Aldrich Catalog #89510-1KG-F

☒ Sodium chloride Sigma Aldrich Catalog #59222C-1000ML

☒ Sera-Mag SpeedBead Carboxylate-Modified Magnetic Particles (Hydrophobic), 15 mL Ge Healthcare Catalog #65152105050250

☒ Sodium azide Sigma Aldrich Catalog #S2002-25G

☒ EDTA (0.5 M), pH 8.0 Life Technologies Catalog #AM9260G

☒ Tris-HCl pH 8.0 (1M solution) Thermo Fisher Scientific Catalog #15568025

☒ Tween-20 Sigma-aldrich Catalog #P-7949

If a commercial solution for sample cleanup is preferred, choose the following product:

☒ Agencourt AMPure XP Beckman Coulter Catalog #A63880

### REAGENTS - SAMPLE & LIBRARY QC

☒ Quant-iT™ PicoGreen® dsDNA Assay Kit Life Technologies Catalog #P11496

☒ Nunc™ F96 MicroWell™ Polystyrene Plate black Thermo Fisher Scientific Catalog #237105

☒ Qubit™ 1X dsDNA HS Assay Kit Invitrogen - Thermo Fisher Catalog #Q33231

☒ Qubit™ Assay Tubes Invitrogen - Thermo Fisher Catalog #Q32856

☒ Agilent High Sensitivity DNA Kit Agilent Technologies Catalog #5067-4626

### REAGENTS - TAGMENTATION WITH NEXTERA XT KIT

☒ Nextera XT DNA Library Preparation Kit illumina Catalog #FC-131-1096

☒ Nextera XT Index Kit v2 (set A B C D) illumina Catalog #FC-131-2001; FC-131-2002; FC-131

## REAGENTS - TAGMENTATION WITH IN-HOUSE Tn5 TRANSPOSASE

- ☒ KAPA HiFi plus dNTPs Roche Catalog #KK2102
- ☒ NN-Dimethylformamide (DMF) solution Millipore Sigma Catalog #D4551
- ☒ SDS, 10% Solution Life Technologies Catalog #AM9822
- ☒ TAPS Sigma Aldrich Catalog #T9659-100G
- ☒ Sodium Hydroxide (pellet purity 98%) Sigma Aldrich Catalog #71690-1KG

## GENERAL CONSUMABLES

- ☒ RNase AWAY<sup>®</sup>; Spray Bottle, RNase in spray bottle; 475mL Thermo Fisher Catalog #7002
- ☒ Adhesive PCR Plate Seals Thermo Fisher Scientific Catalog #AB0558
- ☒ Aluminium foil seals for -80°C storage VWR International Catalog #391-1281
- ☒ Twin.Tec<sup>®</sup> PCR plates 384 (LoBind colourless) Eppendorf Catalog #EP0030129547
- ☒ UltraPure<sup>™</sup> DNase/RNase-Free Distilled Water Thermo Fisher Scientific Catalog #10977023
- ☒ DNA LoBind<sup>®</sup> 1.5 mL (PCR clean colourless) Eppendorf Catalog #30108051
- ☒ Ethanol for molecular biology Sigma Aldrich Catalog #51976-500ML-F

## OLIGONUCLEOTIDES - RT-PCR

A	B	C
Oligo ID	Sequence (5' → 3')	Purification / synthesis scale
Smart dT30VN	/5Biosg/AAGCAGTGGTATCAACGCAGAGTACTTT TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTVN	desalted or HPLC
FS TSO	/5Biosg/AAGCAGTGGTATCAACGCAGAGTACrGr GrG	desalted or HPLC

/5Biosg/ = C6-linker biotin

## OLIGONUCLEOTIDES - IN-HOUSE Tn5 PROTOCOL ONLY

A	B	C
Oligo ID	Sequence (5' → 3')	Comments
TN5MErev	/5Phos/ CTGTCTCTTATACACATCT	2 μM scale - desalted*
TN5ME-A	TCGTCGGCAGCGTCAGATGTGTATAAGAGAC AG	1 μM scale - desalted*

A	B	C
TN5ME-B	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG	1 $\mu$ M scale - desalted*

\* It is important to follow these recommendations. Ordering oligos at this scale but choosing “HPLC purification” will result in insufficient material for Tn5 loading. The scale indicated here is sufficient for producing 20-25 ml of loaded Tn5.

### OLIGONUCLEOTIDES - ALL TAGMENTATION PROTOCOLS (when not ordering the Nextera Index Kit)

One can order the 4 Nextera XT Index Kit v2 (set A, B, C, D) sets, as described above or, alternatively, get them manufactured by any oligonucleotide provider. Below is the list of 24 N7xx and 16 S5xx adaptors required to multiplex 384 samples.

Prepare working dilution plates containing unique combinations of N7xx + S5xx adaptors, each with a final concentration of 5  $\mu$ M.

A	B
Oligo ID	Sequence (5' → 3')
Nextera_v2_N714	/5Biosg/CAAGCAGAAGACGGCATAACGAGATTCATGAGCGTCTCGTGGGCTCG*G
Nextera_v2_N715	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCCTGAGATGTCTCGTGGGCTCG*G
Nextera_v2_N716	/5Biosg/CAAGCAGAAGACGGCATAACGAGATTAGCGAGTGTCTCGTGGGCTCG*G
Nextera_v2_N718	/5Biosg/CAAGCAGAAGACGGCATAACGAGATGTAGCTCCGTCTCGTGGGCTCG*G
Nextera_v2_N719	/5Biosg/CAAGCAGAAGACGGCATAACGAGATTACTACGCGTCTCGTGGGCTCG*G
Nextera_v2_N720	/5Biosg/CAAGCAGAAGACGGCATAACGAGATAGGCTCCGGTCTCGTGGGCTCG*G
Nextera_v2_N721	/5Biosg/CAAGCAGAAGACGGCATAACGAGATGCAGCGTAGTCTCGTGGGCTCG*G
Nextera_v2_N722	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCTGCGCATGTCTCGTGGGCTCG*G
Nextera_v2_N723	/5Biosg/CAAGCAGAAGACGGCATAACGAGATGAGCGCTAGTCTCGTGGGCTCG*G
Nextera_v2_N724	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCGCTCAGTGTCTCGTGGGCTCG*G
Nextera_v2_N726	/5Biosg/CAAGCAGAAGACGGCATAACGAGATGTCTTAGGGTCTCGTGGGCTCG*G
Nextera_v2_N727	/5Biosg/CAAGCAGAAGACGGCATAACGAGATACTGATCGGTCTCGTGGGCTCG*G
Nextera_v2_N728	/5Biosg/CAAGCAGAAGACGGCATAACGAGATTAGCTGCAGTCTCGTGGGCTCG*G
Nextera_v2_N729	/5Biosg/CAAGCAGAAGACGGCATAACGAGATGACGTCGAGTCTCGTGGGCTCG*G

A	B
Nextera_v2_S502	/5Biosg/AATGATACGGCGACCACCGAGATCTACACCTCTCTATTTCGTCCGGCAGCGT*C
Nextera_v2_S513	/5Biosg/AATGATACGGCGACCACCGAGATCTACACTCGACTAGTCGTCCGGCAGCGT*C
Nextera_v2_S503	/5Biosg/AATGATACGGCGACCACCGAGATCTACACTATCCTCTTCGTCCGGCAGCGT*C
Nextera_v2_S515	/5Biosg/AATGATACGGCGACCACCGAGATCTACACTTCTAGCTTCGTCCGGCAGCGT*C
Nextera_v2_S505	/5Biosg/AATGATACGGCGACCACCGAGATCTACACGTAAGGAGTCGTCCGGCAGCGT*C
Nextera_v2_S516	/5Biosg/AATGATACGGCGACCACCGAGATCTACACCCTAGAGTTCGTCCGGCAGCGT*C
Nextera_v2_S506	/5Biosg/AATGATACGGCGACCACCGAGATCTACACACTGCATATCGTCCGGCAGCGT*C
Nextera_v2_S517	/5Biosg/AATGATACGGCGACCACCGAGATCTACACGCGTAAGATCGTCCGGCAGCGT*C
Nextera_v2_S507	/5Biosg/AATGATACGGCGACCACCGAGATCTACACAAGGAGTATCGTCCGGCAGCGT*C
Nextera_v2_S518	/5Biosg/AATGATACGGCGACCACCGAGATCTACACCTATTAAGTCGTCCGGCAGCGT*C
Nextera_v2_S508	/5Biosg/AATGATACGGCGACCACCGAGATCTACACCTAAGCCTTCGTCCGGCAGCGT*C
Nextera_v2_S520	/5Biosg/AATGATACGGCGACCACCGAGATCTACACAAGGCTATTCGTCCGGCAGCGT*C
Nextera_v2_S510	/5Biosg/AATGATACGGCGACCACCGAGATCTACACCGTCTAATTCGTCCGGCAGCGT*C
Nextera_v2_S521	/5Biosg/AATGATACGGCGACCACCGAGATCTACACGAGCCTTATCGTCCGGCAGCGT*C
Nextera_v2_S511	/5Biosg/AATGATACGGCGACCACCGAGATCTACACTCTCTCCGTCCGTCCGGCAGCGT*C
Nextera_v2_S522	/5Biosg/AATGATACGGCGACCACCGAGATCTACACTTATGCGATCGTCCGGCAGCGT*C

All oligonucleotides carry a 5'-biotin (/5Biosg/) and a phosphorothioate bond (\*) between the last and the second last nucleotide. For cost reasons, we ordered desalted oligos and not HPLC purified.

### OLIGONUCLEOTIDES - ALL TAGMENTATION PROTOCOLS (when not ordering the Nextera Index Kit)

To increase the multiplex capabilities, we designed an additional set of 32 S5xx and 48 N7xx adaptors (non-UDI).

A	B
Oligo ID	Sequence
Nextera_extra_i7_1	/5Biosg/CAAGCAGAAGACGGCATAACGAGATGCCTATCAGTCTCGTGGGCTCG*G

A	B
Nextera_extra_i7_2	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCTTGGATGGTCTCGTGGGC TCG*G
Nextera_extra_i7_3	/5Biosg/CAAGCAGAAGACGGCATAACGAGATAGTCTCACGTCTCGTGGGCT CG*G
Nextera_extra_i7_4	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCTCATCAGGTCTCGTGGGC TCG*G
Nextera_extra_i7_5	/5Biosg/CAAGCAGAAGACGGCATAACGAGATTGTACCGTGTCTCGTGGGCT CG*G
Nextera_extra_i7_6	/5Biosg/CAAGCAGAAGACGGCATAACGAGATAAGTCGAGGTCTCGTGGGC TCG*G
Nextera_extra_i7_7	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCACGTTGTGTCTCGTGGGC TCG*G
Nextera_extra_i7_8	/5Biosg/CAAGCAGAAGACGGCATAACGAGATTCACAGCAGTCTCGTGGGC TCG*G
Nextera_extra_i7_9	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCTACTTGGGTCTCGTGGGC TCG*G
Nextera_extra_i7_10	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCCTCAGTTGTCTCGTGGGC TCG*G
Nextera_extra_i7_11	/5Biosg/CAAGCAGAAGACGGCATAACGAGATTCCTACCTGTCTCGTGGGC TCG*G
Nextera_extra_i7_12	/5Biosg/CAAGCAGAAGACGGCATAACGAGATATGGCGAAGTCTCGTGGGC TCG*G
Nextera_extra_i7_13	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCTTACCTGGTCTCGTGGGC TCG*G
Nextera_extra_i7_14	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCTCGATACGTCTCGTGGGCT CG*G
Nextera_extra_i7_15	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCCCGTGAAGTCTCGTGGGC TCG*G
Nextera_extra_i7_16	/5Biosg/CAAGCAGAAGACGGCATAACGAGATTAGAGCTCGTCTCGTGGGC TCG*G
Nextera_extra_i7_17	/5Biosg/CAAGCAGAAGACGGCATAACGAGATTGACTGACGTCTCGTGGGC TCG*G
Nextera_extra_i7_18	/5Biosg/CAAGCAGAAGACGGCATAACGAGATTAGACGTGGTCTCGTGGGC TCG*G
Nextera_extra_i7_19	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCCGGAATTGTCTCGTGGGC TCG*G
Nextera_extra_i7_20	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCTCCTAGAGTCTCGTGGGC TCG*G
Nextera_extra_i7_21	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCAACGGATGTCTCGTGGGC TCG*G
Nextera_extra_i7_22	/5Biosg/CAAGCAGAAGACGGCATAACGAGATTGGCTATCGTCTCGTGGGCT CG*G
Nextera_extra_i7_23	/5Biosg/CAAGCAGAAGACGGCATAACGAGATCGGTCATAGTCTCGTGGGCT CG*G
Nextera_extra_i7_24	/5Biosg/CAAGCAGAAGACGGCATAACGAGATTCCAATCGGTCTCGTGGGC TCG*G

A	B
Nextera_extra_i7_2 5	/5Biosg/CAAGCAGAAGACGGCATACGAGATGAGCTTGTGTCTCGTGGGC TCG*G
Nextera_extra_i7_2 6	/5Biosg/CAAGCAGAAGACGGCATACGAGATGAAGTTTCGTCTCGTGGGC TCG*G
Nextera_extra_i7_2 7	/5Biosg/CAAGCAGAAGACGGCATACGAGATATCTCGCTGTCTCGTGGGCT CG*G
Nextera_extra_i7_2 8	/5Biosg/CAAGCAGAAGACGGCATACGAGATAGTTACGGGTCTCGTGGGC TCG*G
Nextera_extra_i7_2 9	/5Biosg/CAAGCAGAAGACGGCATACGAGATGTGTCTGAGTCTCGTGGGC TCG*G
Nextera_extra_i7_3 0	/5Biosg/CAAGCAGAAGACGGCATACGAGATTGACTTCGGTCTCGTGGGC TCG*G
Nextera_extra_i7_3 1	/5Biosg/CAAGCAGAAGACGGCATACGAGATTGGATCACGTCTCGTGGGC TCG*G
Nextera_extra_i7_3 2	/5Biosg/CAAGCAGAAGACGGCATACGAGATACACCAGTGTCTCGTGGGC TCG*G
Nextera_extra_i7_3 3	/5Biosg/CAAGCAGAAGACGGCATACGAGATCAGGTTAGGTCTCGTGGGC TCG*G
Nextera_extra_i7_3 4	/5Biosg/CAAGCAGAAGACGGCATACGAGATAGTTGGCTGTCTCGTGGGC TCG*G
Nextera_extra_i7_3 5	/5Biosg/CAAGCAGAAGACGGCATACGAGATTCAACTGGGTCTCGTGGGC TCG*G
Nextera_extra_i7_3 6	/5Biosg/CAAGCAGAAGACGGCATACGAGATCTGCACTTGTCTCGTGGGC TCG*G
Nextera_extra_i7_3 7	/5Biosg/CAAGCAGAAGACGGCATACGAGATACACGGTTGTCTCGTGGGC TCG*G
Nextera_extra_i7_3 8	/5Biosg/CAAGCAGAAGACGGCATACGAGATAATACGCGGTCTCGTGGGC TCG*G
Nextera_extra_i7_3 9	/5Biosg/CAAGCAGAAGACGGCATACGAGATTGCGAACTGTCTCGTGGGC TCG*G
Nextera_extra_i7_4 0	/5Biosg/CAAGCAGAAGACGGCATACGAGATGCTGACTAGTCTCGTGGGC TCG*G
Nextera_extra_i7_4 1	/5Biosg/CAAGCAGAAGACGGCATACGAGATGTGGTGTGTCTCGTGGGC TCG*G
Nextera_extra_i7_4 2	/5Biosg/CAAGCAGAAGACGGCATACGAGATGTGCTTACGTCTCGTGGGC TCG*G
Nextera_extra_i7_4 3	/5Biosg/CAAGCAGAAGACGGCATACGAGATTCAAGGACGTCTCGTGGGC TCG*G
Nextera_extra_i7_4 4	/5Biosg/CAAGCAGAAGACGGCATACGAGATTGAACCTGGTCTCGTGGGC TCG*G
Nextera_extra_i7_4 5	/5Biosg/CAAGCAGAAGACGGCATACGAGATAGTGTGGGTCTCGTGGGC TCG*G
Nextera_extra_i7_4 6	/5Biosg/CAAGCAGAAGACGGCATACGAGATGACTCTCGTCTCGTGGGCT CG*G
Nextera_extra_i7_4 7	/5Biosg/CAAGCAGAAGACGGCATACGAGATCCGTATCTGTCTCGTGGGCT CG*G



A	B
Nextera_extra_i7_4 8	/5Biosg/CAAGCAGAAGACGGCATACGAGATCGAAGAACGTCTCGTGGGC TCG*G

All oligonucleotides carry a 5'-biotin (/5Biosg/) and a phosphorothioate bond (\*) between the last and the second last nucleotide. For cost reasons, we ordered desalted oligos and not HPLC purified.

Prepare working dilution plates containing unique combinations of N7xx + S5xx adaptors, each with a final concentration of 5 µM.

A	B
Oligo ID	Sequence
Nextera_extra_i5_1	/5Biosg/AATGATACGGCGACCACCGAGATCTACACCGACCATTTTCGTCCG CAGCGT*C
Nextera_extra_i5_2	/5Biosg/AATGATACGGCGACCACCGAGATCTACACGATAGCGATCGTCCGC AGCGT*C
Nextera_extra_i5_3	/5Biosg/AATGATACGGCGACCACCGAGATCTACACAATGGACGTCTCGTCCG CAGCGT*C
Nextera_extra_i5_4	/5Biosg/AATGATACGGCGACCACCGAGATCTACACCGCTAGTATCGTCCGC AGCGT*C
Nextera_extra_i5_5	/5Biosg/AATGATACGGCGACCACCGAGATCTACACTCTCTAGGTCTCGTCCG CAGCGT*C
Nextera_extra_i5_6	/5Biosg/AATGATACGGCGACCACCGAGATCTACACACATTGCGTCTCGTCCG CAGCGT*C
Nextera_extra_i5_7	/5Biosg/AATGATACGGCGACCACCGAGATCTACACTGAGGTGTTTCGTCCG CAGCGT*C
Nextera_extra_i5_8	/5Biosg/AATGATACGGCGACCACCGAGATCTACACAATGCCTCTCGTCCG CAGCGT*C
Nextera_extra_i5_9	/5Biosg/AATGATACGGCGACCACCGAGATCTACACCTGGAGTATCGTCCGC AGCGT*C
Nextera_extra_i5_10	/5Biosg/AATGATACGGCGACCACCGAGATCTACACGTATGCTGTCTCGTCCGC AGCGT*C
Nextera_extra_i5_11	/5Biosg/AATGATACGGCGACCACCGAGATCTACACTGGAGAGTTTCGTCCG CAGCGT*C
Nextera_extra_i5_12	/5Biosg/AATGATACGGCGACCACCGAGATCTACACCGATAGAGTCTCGTCCG CAGCGT*C
Nextera_extra_i5_13	/5Biosg/AATGATACGGCGACCACCGAGATCTACACCTCATTGCTCTCGTCCG CAGCGT*C
Nextera_extra_i5_14	/5Biosg/AATGATACGGCGACCACCGAGATCTACACACCAGCTTTTCGTCCG CAGCGT*C
Nextera_extra_i5_15	/5Biosg/AATGATACGGCGACCACCGAGATCTACACGAATCGTGTCTCGTCCG CAGCGT*C
Nextera_extra_i5_16	/5Biosg/AATGATACGGCGACCACCGAGATCTACACAGGCTTCTTCGTCCG CAGCGT*C
Nextera_extra_i5_17	/5Biosg/AATGATACGGCGACCACCGAGATCTACACCAGTTCTGTCTCGTCCG CAGCGT*C

A	B
Nextera_extra_i5_1_8	/5Biosg/AATGATACGGCGACCACCGAGATCTACACTTGGTGAGTCGTCGG CAGCGT*C
Nextera_extra_i5_1_9	/5Biosg/AATGATACGGCGACCACCGAGATCTACACCATTCCGTTTCGTCGG CAGCGT*C
Nextera_extra_i5_2_0	/5Biosg/AATGATACGGCGACCACCGAGATCTACACTGTGAAGCTCGTCGG CAGCGT*C
Nextera_extra_i5_2_1	/5Biosg/AATGATACGGCGACCACCGAGATCTACACTAAGTGGCTCGTCGG CAGCGT*C
Nextera_extra_i5_2_2	/5Biosg/AATGATACGGCGACCACCGAGATCTACACACGTGATGTCGTCGG CAGCGT*C
Nextera_extra_i5_2_3	/5Biosg/AATGATACGGCGACCACCGAGATCTACACGTAGAGCATCGTCGG CAGCGT*C
Nextera_extra_i5_2_4	/5Biosg/AATGATACGGCGACCACCGAGATCTACACGTCAGTTGTCGTCGG CAGCGT*C
Nextera_extra_i5_2_5	/5Biosg/AATGATACGGCGACCACCGAGATCTACACATTCGAGGTCGTCGG CAGCGT*C
Nextera_extra_i5_2_6	/5Biosg/AATGATACGGCGACCACCGAGATCTACACGATACTGGTCGTCGGC AGCGT*C
Nextera_extra_i5_2_7	/5Biosg/AATGATACGGCGACCACCGAGATCTACACGCCTTGTTTCGTCGG CAGCGT*C
Nextera_extra_i5_2_8	/5Biosg/AATGATACGGCGACCACCGAGATCTACACTTGGTCTCTCGTCGG CAGCGT*C
Nextera_extra_i5_2_9	/5Biosg/AATGATACGGCGACCACCGAGATCTACACCCGACTATTCGTCGG CAGCGT*C
Nextera_extra_i5_3_0	/5Biosg/AATGATACGGCGACCACCGAGATCTACACGTCCTAAGTCGTCGG CAGCGT*C
Nextera_extra_i5_3_1	/5Biosg/AATGATACGGCGACCACCGAGATCTACACACCAATGCTCGTCGG CAGCGT*C
Nextera_extra_i5_3_2	/5Biosg/AATGATACGGCGACCACCGAGATCTACACGATGCACTTCGTCGG CAGCGT*C

All oligonucleotides carry a 5'-biotin (/5Biosg/) and a phosphorothioate bond (\*) between the last and the second last nucleotide. For cost reasons, we ordered desalted oligos and not HPLC purified.

Prepare working dilution plates containing unique combinations of N7xx + S5xx adaptors, each with a final concentration of 5 µM.

## Before start

The protocol should be carried out in a clean environment, ideally on a dedicated PCR workstation or on a separate bench used only for this purpose. Before starting, clean the bench and wipe any piece of equipment with RNaseZAP or 0.5% sodium hypochlorite. Rinse with nuclease-free water to avoid corrosion of delicate equipment.

Work quickly and preferably on ice.

Reagent mixes should be prepared shortly before use.

Mix thoroughly each mix before dispensing. For higher accuracy use liquid handling robots and/or nanodispensers whenever possible. In FLASH-Seq we use the I.DOT (Dispendix) for all the dispensing steps and the Fluent 780 liquid handling robot (Tecan) for sample cleanup, reagent transfers and pooling.

The protocol described below is meant to be carried out in 384-well plates. When using 96-well plates, we recommend using 5 times larger volume to guarantee successful cell sorting and prevent evaporation issues.


Always use LoBind plates and tubes (especially for long-term storage) to prevent the cDNA/DNA from sticking to plastic.

## Prepare lysis mix


1

Prepare the following lysis mix:

A	B	C	D
Reagent	Reaction concentration	Volume ( $\mu$ l)	384-well plate
Triton-X100 (10% v/v)	0.2%	0.020	8.448
dNTP mix (25 mM each)	6 mM	0.240	101.376
SMART dT30VN (100 $\mu$ M)	1.8 $\mu$ M	0.018	7.603
RNAse inhibitor (40 U/ $\mu$ L)	1.2 U/ $\mu$ l	0.030	12.672
DTT (100 mM)	1.2 mM	0.012	5.069
FS TSO (100 $\mu$ M)	9.2 $\mu$ M	0.092	38.861
dCTP (100 mM)	9 mM	0.090	38.016
Betaine (5 M)	1 M	0.200	84.480
Nuclease-free water	-	0.298	125.875
<b>Total volume (<math>\mu</math>l)</b>		<b>1.000</b>	<b>422.400</b>

Add  1  $\mu$ L lysis mix to each well of a 384-well plate

Seal the plate with a PCR seal and quickly spin it down to collect the lysis mix to the bottom.


Proceed immediately to the next step or store the plate at  -20  $^{\circ}$ C long-term. Plates that are going to be used on the same day can be stored in the fridge or kept on ice.

### Note

**SAFE STOPPING POINT** - Plates containing lysis buffer can be stored for >6 months at

 -20  $^{\circ}$ C

## Sample collection

2 Sort single cells into 384-well plates containing  1  $\mu$ L lysis mix.

Seal the plate with an aluminium seal. If processing multiple plates at once, keep each plate on dry ice until ready to transfer them all at  $-80\text{ }^{\circ}\text{C}$  for long-term storage. Plates containing single cells should ideally be processed within 6 months.

## Cell lysis

3m

- Remove the plates from the  $-80\text{ }^{\circ}\text{C}$  freezer and check that the aluminium seal is still intact. If damaged or not sticking to the plate anymore, wait a few minutes for the plate to partially thaw, remove the damaged foil and replace it with a new one.

3m

Place the plate in a thermocycler with a heated lid and incubate for 00:03:00 at  $72\text{ }^{\circ}\text{C}$ , followed by a  $4\text{ }^{\circ}\text{C}$  hold step.

Spin down any condensation droplets that may have formed during the incubation and return the plate to a cool rack. Proceed quickly to the next step. If not ready with the RT-PCR mix, keep the plate on the cool rack at all times.

## RT-PCR reaction

15m

- While the plate is in the thermocycler, prepare the following RT-PCR mix:

A	B	C	D
Reagent	Reaction concentration	Volume ( $\mu\text{l}$ )	384-well plate
DTT (0.1 M)	4.8 mM	0.238	100.531
MgCl <sub>2</sub> (1 M)	9.2 mM	0.046	19.430
Betaine (5 M)	800 mM	0.800	337.920
RNAse inhibitor (40 U/ $\mu\text{l}$ )	0.8 U/ $\mu\text{l}$	0.096	40.550
SuperScript IV (200 U/ $\mu\text{l}$ )	2.00 U/ $\mu\text{l}$	0.050	21.120
KAPA HiFi HotStart ReadyMix (2 x)	1 x	2.500	1056.000
Nuclease-free water	-	0.270	114.048
<b>Total volume (<math>\mu\text{l}</math>)</b>		<b>4.000</b>	<b>1689.600</b>

Add  $4\text{ }\mu\text{L}$  RT-PCR mix into each well of the 384-well plate.


Seal the plate with a PCR seal, gently vortex and spin down to collect the liquid at the bottom.

Place it in a thermocycler with heated lid and start the following RT-PCR program:

A	B	C	D	E
Step		Temperature	Time	Cycles
RT		50°C	60 min	1 x
PCR	initial denaturation	98°C	3 min	1 x
	denaturation	98°C	20 sec	18-21 x*
	annealing	67°C		
	elongation	72°C		
		15°C	Hold	

\*Adjust the number of cycles according to the cell type used. We recommend 18-19 cycles for HEK 293T cells and 21 cycles for hPBMC. As a rule of thumb, we typically start with 1-2 less PCR cycles compared to the Smart-seq2 protocol.

**Note**

**SAFE STOPPING POINT** - Amplified cDNA before purification can be stored for several months at  -20 °C

## Magnetic beads working solution preparation

1h

- 5 You can either use AMPure XP beads, SPRI beads or prepare your own solution of SeraMag beads containing 18% w/v PEG to reduce costs. A detailed protocol for making your own magnetic bead solution is described in:

**CITATION**

Picelli S (2019). Full-Length Single-Cell RNA Sequencing with Smart-seq2.. Methods in molecular biology (Clifton, N.J.).

LINK

[https://doi.org/10.1007/978-1-4939-9240-9\\_3](https://doi.org/10.1007/978-1-4939-9240-9_3)

Below is a short description of how to prepare 50 ml of working solution:

A	B	C
Reagent	Amount to add	Final concentration
Sodium chloride	2.92 gr	1 M
Tris-HCl, pH = 8.0 (1 M)	500 $\mu$ l	10 mM
EDTA (500 mM)	100 $\mu$ l	1 mM
PEG (MW=8000)	9.5 gr	18% w/v
Nuclease-free solution	to a final volume of 50 ml	-

Add all components to a 50-ml Falcon tube but do not add the total amount of nuclease-free water yet.

Solubilise the PEG by stirring and heating the solution at  $37\text{ }^{\circ}\text{C}$ .

While the PEG is dissolving, prepare the Sera-Mag SpeedBeads™. Vortex thoroughly to ensure complete resuspension and then withdraw  $1\text{ mL}$  Sera-Mag SpeedBeads™ stock solution. Transfer it into a new 1.5-ml tube.

Pellet the beads by placing the tube on a magnetic stand. Wait until the solution clears and then discard the supernatant.

Add  $1\text{ mL}$  10 mM Tris-HCl pH 8.0 + 1 mM EDTA (TE buffer) and resuspend the beads off the magnet.

Pellet the beads again, wait until the solution is clear, discard the supernatant and resuspend off the magnet with  $0.9\text{ mL}$  TE buffer.

Once the PEG solution is clear, add the resuspended beads prepared in the previous step.

Add  $50\text{ }\mu\text{L}$  Tween-20 (10% v/v) and  $250\text{ }\mu\text{L}$  sodium azide ( $\text{NaN}_3$ , 10% w/v) and adjust the volume to  $50\text{ mL}$  with nuclease-free water.

Store at  $4\text{ }^{\circ}\text{C}$ . Do not freeze.

### Note

Confirm that the beads have been properly prepared by cleaning-up a control sample (i.e., amplified cDNA from total RNA) and running a High Sensitivity DNA chip on the Agilent Bioanalyzer. Batch-to-batch variations in PEG concentration will influence size-cutoffs.

### Note



Add Tween-20 at the end, to prevent foaming during PEG resuspension.

### Safety information

Sodium azide is extremely toxic and should be handled under a fume hood.



## cDNA purification

25m



- 6 Remove the Sera-Mag SpeedBeads™ working solution (or AMPure XP beads or SPRI beads when using a commercial solution) from the  4 °C storage and equilibrate it at room temperature for  00:15:00 .


29m

### Note

We recommend adding extra nuclease-free water to each sample, to increase the volume, simplify the handling and improve recovery rate. We generally add  10 µL nuclease-free water to  5 µL amplified cDNA.


Add a 0.8 x volume ratio Sera-Mag SpeedBeads™ working solution to each well. Mix thoroughly by pipetting or vortexing.

Incubate the plate off the magnetic stand for  00:05:00 at  Room temperature .


Place the plate on the magnetic stand and leave it for  00:05:00 or until the solution appears clear.



Remove the supernatant without disturbing the beads.


Remove the plate from the magnetic stand, add  15  $\mu\text{L}$  nuclease-free water and mix well by pipetting or vortexing to resuspend the beads. Do not let the bead pellet to dry completely, as it can decrease the final cDNA yield!

Incubate  00:02:00 off the magnetic stand.

Place the plate back on the magnetic stand and incubate for  00:02:00 or until the solution appears clear.

Remove  14  $\mu\text{L}$  of the supernatant and transfer it to a new plate.

#### Note

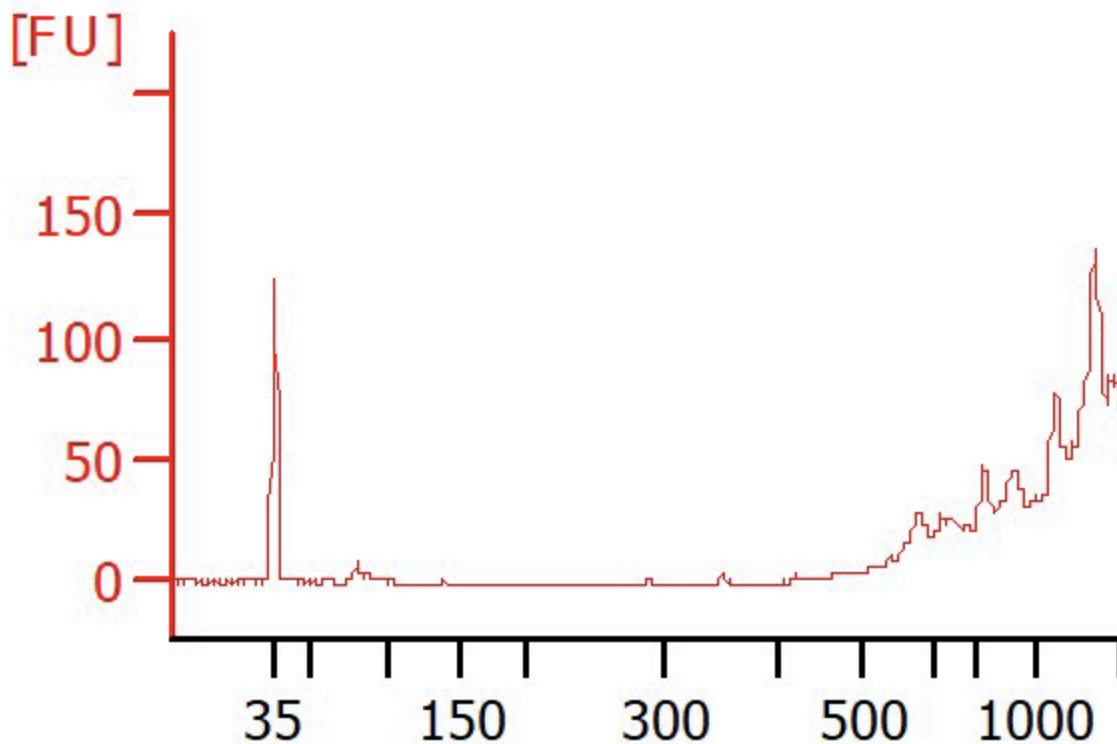
**SAFE STOPPING POINT** - Amplified and purified cDNA can be stored for several months at   $-20\text{ }^{\circ}\text{C}$ . We recommend using LoBind plates to avoid material losses upon long-term storage.

## Quality control check (highly recommended!)

45m

- 7 Check the cDNA quality on Agilent Bioanalyzer High Sensitivity DNA chip. A good sample is characterised by a low proportion of fragments  $<400\text{ bp}$ , absence of residual primers (ca.  $100\text{ bp}$ ) and an average cDNA size of  $1.8\text{--}2.2\text{ Kb}$ .

Expected result



Example of amplified cDNA from a single hPBMC (21 cycles)

### cDNA quantification (optional but recommended)

15m

- 8 Allow the Quant-iT PicoGreen reagent to warm to  before opening the vial. PicoGreen is light sensitive; while thawing, wrap in aluminium foil.

2m

Prepare a 1 x working solution TE using 20 x TE (supplied) and nuclease-free water.

Prepare a 1:400 dilution PicoGreen solution and always use a plastic vessel.

Prepare the standard curve using Lambda DNA standard (supplied at a concentration of ) , with the PicoGreen kit) and 1 x TE in 8 tubes, as below. The stock tubes

can be used multiple times, keep any leftover in the fridge at  $4\text{ }^{\circ}\text{C}$  between experiments.

Vortex well and spin down the DNA standards before every use. Not vortexing thoroughly the standards is going to negatively affect the standard curve and your readings! Serial dilutions should be prepared as shown in the table below.

A	B	C	D
Tube no.	Contents	Concentration	Final volume
1	90 $\mu\text{l}$ TE + 10 $\mu\text{l}$ Lambda DNA stock	10 ng/ $\mu\text{l}$	100 $\mu\text{l}$
2	50 $\mu\text{l}$ from Tube 1 + 50 $\mu\text{l}$ TE	5 ng/ $\mu\text{l}$	100 $\mu\text{l}$
3	50 $\mu\text{l}$ from Tube 2 + 50 $\mu\text{l}$ TE	2.5 ng/ $\mu\text{l}$	100 $\mu\text{l}$
4	50 $\mu\text{l}$ from Tube 3 + 50 $\mu\text{l}$ TE	1.25 ng/ $\mu\text{l}$	100 $\mu\text{l}$
5	50 $\mu\text{l}$ from Tube 4 + 50 $\mu\text{l}$ TE	0.625 ng/ $\mu\text{l}$	100 $\mu\text{l}$
6	50 $\mu\text{l}$ from Tube 5 + 50 $\mu\text{l}$ TE	0.3125 ng/ $\mu\text{l}$	100 $\mu\text{l}$
7	50 $\mu\text{l}$ from Tube 6 + 50 $\mu\text{l}$ TE	0.15625 ng/ $\mu\text{l}$	100 $\mu\text{l}$
8	TE only	blank	-

Prepare the PicoGreen solution by pipetting  $0.5\text{ }\mu\text{L}$  PicoGreen dye +  $99.5\text{ }\mu\text{L}$  1X TE for each sample. Vortex to mix.

Pipette  $1\text{ }\mu\text{L}$  each of the 7 standards + 1 Blank into a black, flat-bottom Nunc™ F96 MicroWell™ plate. Place the standards on one column.

Pipet  $1\text{ }\mu\text{L}$  of your samples into the center of each well of the Nunc™ F96 MicroWell™ polystyrene plate.

Add  $99\text{ }\mu\text{L}$  PicoGreen + TE mix into every well. There is no need to mix.


Cover the plate with the provided plastic (transparent) lid to prevent possible contaminations.

Allow  $00:02:00$  for the dye to bind the DNA. Protect from light but keep at room temperature. For optimal results, the plate should be read within the next hour.

Use a plate reader to measure fluorescence (excitation: 485 nm; emission: 530 nm; read from top; endpoint reading).

## Plate normalisation

10m

- 9 Prepare a normalisation plate by adding  1  $\mu\text{L}$  purified cDNA and nuclease-free water to a final concentration of  $[M]$  150  $\text{pg}/\mu\text{l}$  .

## Tagmentation and enrichment PCR

1h

- 10 This step can be carried out either by using the commercially available **Nextera XT kit** or a **in-house Tn5 transposase**, as described below. Indexing primers can be purchased from Illumina (Nextera XT index kit v2) or ordered from your local oligo manufacturer. In the "Materials" section we have added additional sequences for higher multiplexing.

### 10.1 Tagmentation with the Nextera XT kit

13m


#### Note




Please note that the volumes described here are a mere suggestion. For example, decreasing the final volume by a factor 2 (while still using the same amount of cDNA and ATM) would give comparable results but significantly reduce costs.



Prepare the tagmentation mix as described below:

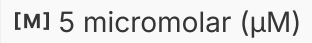
A	B
Reagent	Volume ( $\mu\text{l}$ )
ATM (Amplification Tagment Mix)	0.250
TD (Tagmentation DNA buffer)	2.000
Nuclease-free water	0.500
<b>Total volume (<math>\mu\text{l}</math>)</b>	<b>3.000</b>


Dispense  3  $\mu\text{L}$  tagmentation mix in a new 96-well or 384-well plate.

Add  1  $\mu\text{L}$  normalized cDNA (  $[M]$  150  $\text{pg}/\mu\text{l}$  ) to each well containing the tagmentation mix.

Seal the plate, vortex, spin down, and carry out the tagmentation reaction:  55  $^{\circ}\text{C}$  for  00:08:00 ,  4  $^{\circ}\text{C}$  hold. Upon completion proceed immediately to the next step.

Add  1  $\mu\text{L}$  NT buffer to each well. Seal the plate, vortex, spin down and incubate  00:05:00 at room temperature. Do not put the plate back on ice.


Add  2  $\mu\text{L}$  N7xx + S5xx index adaptors (  5 micromolar ( $\mu\text{M}$ ) each).

Add  3  $\mu\text{L}$  NPM solution to each well.

Seal the plate, vortex, spin down, and place it in a thermocycler and carry out the enrichment PCR reaction. Adjust the number of PCR cycles according to the number of processed cells.

A	B	C	D	E
<b>Step</b>		<b>Temperature</b>	<b>Time</b>	<b>Cycles</b>
gap filling		72°C	3 min	1 x
enrichment PCR	initial denaturation	95°C	30 sec	1 x
	denaturation	95°C	10 sec	12-14 x
	annealing	55°C	30 sec	
	elongation	72°C	30 sec	
		15°C	hold	


Note

**SAFE STOPPING POINT** - The final unpurified sequencing library can be stored for several months at  -20 °C

## 10.2 Tagmentation with in-house Tn5 transposase

8m

Note


Please note that the Tn5 transposase amount indicated below is a suggested starting point for tagmenting  150  $\text{pg}/\mu\text{l}$  cDNA. Optimisation might be necessary, depending on the specific activity of each batch of Tn5.


Prepare the tagmentation mix as described below:




A	B	C
<b>Reagent</b>	<b>Volume (µl)</b>	<b>Final concentration</b>
TAPS-Mg buffer, pH=7.3 (5 x)	0.800	10 mM TAPS, 5 mM MgCl <sub>2</sub>
Dimethylformamide (DMF) (100%)	0.800	20%
Tn5 transposase (2 µM working dil.)	0.125	62.5 nmol
Nuclease-free water	2.275	
<b>Total volume (µl)</b>	<b>3.000</b>	


### Safety information

Dimethylformamide (DMF) is toxic and should be handled under the hood according to local safety regulations.

Dispense  3 µL tagmentation mix in a new 384-well plate.

Add  1 µL normalized cDNA ( [IM] 150 pg/µl ) to each well containing the tagmentation mix.

Seal the plate, vortex, spin down, and carry out the tagmentation reaction:  55 °C for  00:08:00 ,  4 °C hold. Upon completion proceed immediately to the next step.

Add  1 µL 0.2% SDS to each well. Seal the plate, vortex, spin down and incubate 5 min at room temperature. Do not put the plate back on ice.

Add  2 µL N7xx + S5xx index adaptors ( [IM] 5 micromolar (µM) each).

Add  3 µL enrichment PCR mix to each well:

A	B	C
<b>Reagent</b>	<b>Volume (µl)</b>	<b>Final concentration</b>
KAPA HiFi enzyme (1 U/µl)	0.200	0.02 U/µl
KAPA HiFi buffer (5 x)	2.000	1 x
dNTPs (10 mM)	0.300	300 nM
Nuclease-free water	0.500	


A	B	C
<b>Total volume (μl)</b>	<b>3.000</b>	

Seal the plate, vortex, spin down, and place it in a thermocycler and carry out the enrichment PCR reaction. Adjust the number of PCR cycles according to the number of processed cells.



A	B	C	D	E
	<b>Step</b>	<b>Temperature</b>	<b>Time</b>	<b>Cycles</b>
	gap filling	72°C	3 min	1 x
enrichment PCR	initial denaturation	98°C	30 sec	1 x
	denaturation	98°C	10 sec	12-14 x
	annealing	55°C	30 sec	
	elongation	72°C	30 sec	
		15°C	hold	

## Library cleanup and quantification



30m


- 11 Take an aliquot from each sample for the final library cleanup (i.e. 5 μl). and transfer it to a 1.5-ml Eppendorf tube. The rest of the library can be stored long-term at  -20 °C .

31m 30s

Remove the Sera-Mag SpeedBeads™ working solution from the  4 °C storage and equilibrate it at room temperature for  00:15:00 .


Add Sera-Mag SpeedBeads™ working solution to a final ratio of 0.8 x and mix well to homogenisation.


Incubate the tube off the magnetic stand for  00:05:00 at  Room temperature .


Place the tube on the magnetic stand and leave it for  00:05:00 or until the solution appears clear.

Remove the supernatant without disturbing the beads.


Recommended: wash the pellet with  1 mL 80% v/v ethanol. Incubate



 00:00:30 without removing the tube from the magnetic stand.

Remove any trace of ethanol and let the bead pellet dry for  00:02:00 or until small cracks appear. Do not cap the tube or remove it from the magnetic stand during this time. Do not completely air-dry the beads.

Remove the tube from the magnetic stand, add  50  $\mu$ L nuclease-free water and mix well by pipetting or vortexing to resuspend the beads.

Incubate  00:02:00 off the magnetic stand.

Place the tube back on the magnetic stand and incubate for  00:02:00 or until the solution appears clear.

Remove  49  $\mu$ L of the supernatant and transfer it to a new 1.5-ml LoBind tube. Store the cDNA at  -20 °C long-term or until ready for sequencing.

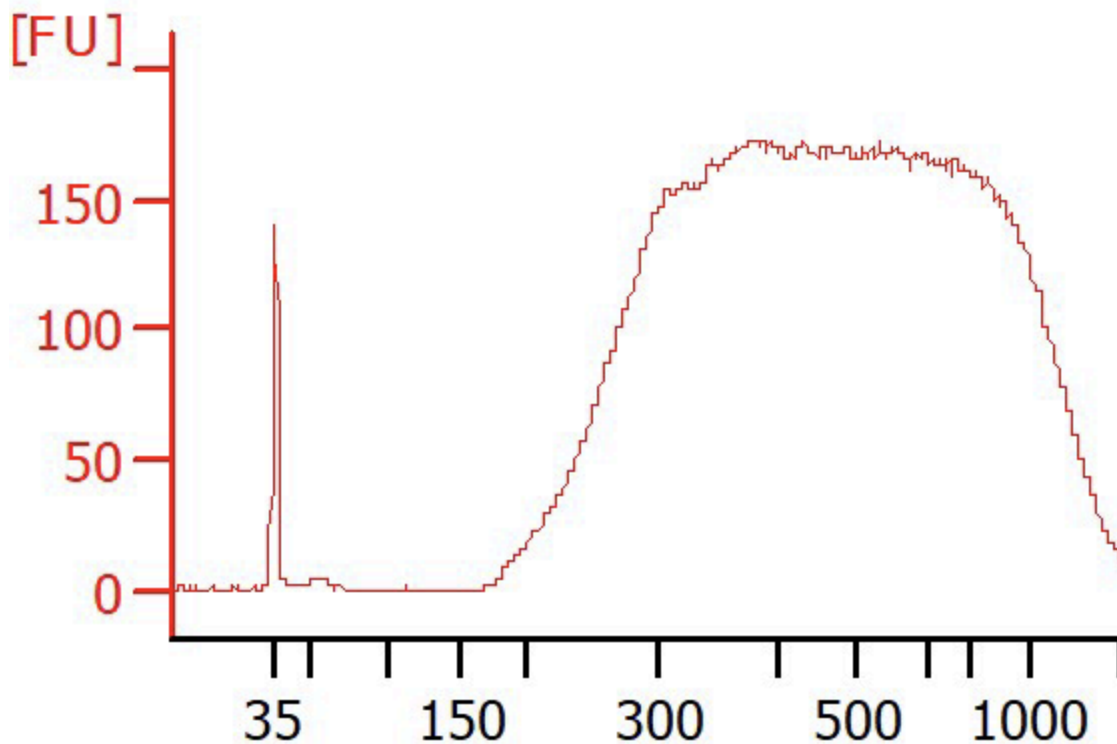
Use Qubit fluorometer to quantify the library. Library yield can vary depending on the number of cells being pooled.

Check the final library size on the Agilent Bioanalyzer.

Use the average size indicated on the Bioanalyzer and the concentration reported after Qubit measurement to determine the exact molarity required for sequencing.




### Expected result



Example of sequencing-ready library (pool of 384 HEK 293T cells).

### Note

**SAFE STOPPING POINT** - The final purified sequencing library can be stored for several months at  -20 °C .

## Pooling and sequencing

- 12 The purified library can be sequenced on any Illumina sequencer. Follow the specifications reported for each instrument. Single-End 75 bp is generally sufficient but longer read modes or Paired-End sequencing can be an option, depending on the question at hand.

## Data processing

- 13 These instructions briefly describe the data processing of the sequencing results. The final pipeline will likely have to be adapted to the question at hand. The following lines assume that all the programs and their dependencies are installed on your machine and that the data are single-end reads (75 bp). Some values, such as the number of threads and RAM usage may have to be adapted to your machine settings.

It should be noted that there are many other ways to analyse full-length single-cell RNA-sequencing data. Pseudo-alignment tools (e.g., Salmon or Kallisto) or automatic pipelines (zUMIs) could be used as well.

Requirements (tested version):

- bcl2fastq (v2.20)
- STAR (v2.7.3)
- FeatureCounts (v1.6.5)
- BBMAP (v38.86)
- samtools (v1.9)
- IGV

### 13.1 Sample demultiplexing

Sequencing results will be delivered as demultiplexed FASTQ or raw bcl2 files. To convert bcl2 files to FASTQ, bcl2fastq program (Illumina) can be used.

#### Command

```
# 0. Variables
BASECALL_DIR="/path/to/flowcell/Data/Intensities/BaseCalls/"
OUTPUT_DIR="/path/to/output_folder/"
SAMPLESHEET="/path/to/Demultiplexing_SampleSheet.csv"
```

## Command

```
# 1. Bcl2fastq
ulimit -n 10000
cd /path/to/flowcell/
bcl2fastq --input-dir $BASECALL_DIR --output-dir $OUTPUT_DIR --sample-sheet $SAMPLESHEET --create-fastq-for-index-reads --no-lane-splitting
```

When sequencing on a NextSeq550 instrument, the sample sheet should contain the following information in a csv file:

```
[Header]
IEMFileVersion      5
Investigator Name   Your Name
Experiment Name      Experiment Title
Date                11/12/2021
Workflow            GenerateFASTQ
Application          NextSeq FASTQ Only
Instrument Type      NextSeq/MiniSeq
Assay               Nextera XT
Index Adapters       Nextera XT Index Kit (24 Indexes 96 Samples)
Chemistry            Amplicon

[Reads]
75

[Settings]
Adapter             CTGTCTCTTATACACATCT

[Data]
Sample_ID  Sample_Name      Sample_Plate      Sample_WellI7_Index_ID  index      I5_Index_ID  index2      Sample_Project
Sample1_A1  Sample1_A1 001      A1      Nextera_extra_i7_25      ACAAGCTC  Nextera_extra_i5_17      CAGAACTG
```

Illumina Experiment Manager can be used to assist you in creating the sample sheet.

We recommend exploring the barcode combinations left in the undetermined reads looking to confirm that all the cells have been properly demultiplexed.

#### Command

```
zcat Undetermined_S0_I1_001.fastq.gz | awk -F' 1:N:0:' 'NR%4==1{print $2}' | sort | uniq -c > left_index.txt  
sort -k1,1 left_index.txt
```

as well as the read distribution between samples:

#### Command

```
for file in ./out/*R1*  
do  
zcat $file | wc -l  
done
```

## 13.2 Index the genome

The reference genome needs to be indexed prior to any mapping. The FASTA and GTF references can be obtained from ENSEMBL, Gencode, UCSC, ...

#### Command

```
# 0. Variables  
OUTPUTREF="/path/to/STAR_indexed_genome/"  
FASTA="GRCh38.primary_assembly.genome.fa"  
GTF="gencode.v34.primary_assembly.annotation.gtf"
```



#### Command

```
# 1. Genome indexing
# sjdbOverhang should be adapted based on the read length
(read_length - 1)
mkdir $OUTPUTREF
```

#### Command

```
STAR --runThreadN 15 --runMode genomeGenerate --genomeDir $OUTPUTREF -
-genomeFastaFiles $FASTA --sjdbGTFfile $GTF --sjdbOverhang 74
```

### 13.3 **FASTQ trimming (optional)**

If you observe sequencing primer left-overs the FASTQ files can be trimmed using BBDUK or Trimmomatic.

#### Command

```
bbduk.sh -Xmx48g in=sample.fastq.gz out=cleaned.left.fastq t=32
ktrim=1 ref=adapters.fa k=23 mink=7 hdist=1 hdist2=0 tbo
bbduk.sh -Xmx48g in=cleaned.left.fastq out=cleaned.fastq t=32 ktrim=r
ref=adapters.fa k=23 mink=7 hdist=1 hdist2=0 tbo
```

**Command**

```
mv FASTQ/cleaned.fastq FASTQ/sample.R1.fastq.gz
```

## 13.4 Mapping

The FASTQ file can then be mapped onto the reference genome. Example for one sample, use a loop or parallelise this task to process all the cells:

**Command**

```
# 0. Variables  
GENOME="/path/to/STAR_indexed_genome/"  
FASTQ="/path/to/sample.R1.fastq.gz"  
ID="sample_id"
```

**Command**

```
# 1. Mapping  
STAR --runThreadN 30 --limitBAMsortRAM 20000000000 --genomeLoad  
LoadAndKeep --genomeDir "$GENOME" --readFilesIn "$FASTQ" --  
readFilesCommand zcat --limitSjdbInsertNsj 2000000 --  
outFilterIntronMotifs RemoveNoncanonicalUnannotated --outSAMtype BAM  
SortedByCoordinate --outFileNamePrefix "$ID_"
```

#### Command

```
# 2. SAM to sorted BAM
# -F 260 filters out unmapped and secondary alignments
samtools view -@ 30 -Sb -F 260 "$ID"_Aligned.sortedByCoord.out.bam >
"$ID"_Aligned.sortedByCoord.filtered.bam
samtools index "$ID"_Aligned.sortedByCoord.filtered.bam
```

### 13.5 **Data visualization (optional)**

Once the reads have been mapped we highly recommend using the Integrated Genome Viewer (IGV) to visualise the mapping results and ensure that the results make sense. As a quick check-up visualise a few housekeeping genes (i.e., ACTB, GAPDH, ...) and cell specific markers to look for reads mapping to exon, intron, exon-intron junctions. Look for abnormalities such as read piles falling in intergenic or centromeric regions.

No single-cell RNA sequencing protocol is perfect and non-specific priming, genomic DNA contaminations, ... can happen but should represent rare events.

Recurrent soft-clipping could also indicate the presence of sequencing adaptor left-overs that could affect the mapping rate.

### 13.6 **Count matrix**

Finally, the number of reads associated with each gene can be obtained as follows:

#### Command

```
featureCounts -T 1 -t exon -g gene_name --fracOverlap 0.25 -a "$GTF" -
o "$ID"_ReadCount.featureCounts.gencode.txt
"$ID"_Aligned.sortedByCoord.filtered.bam
```

### 13.7 **Post-processing**



The post-processing steps will vary depending on the question at hand. The online book “Orchestrating Single-Cell Analysis with Bioconductor” (<https://bioconductor.org/books/release/OSCA/>) is a gold mine of information that can be used to help you design your own pipeline. Alternatively, Seurat (R, <https://satijalab.org/seurat/>) or scanpy (python, <https://scanpy.readthedocs.io/en/stable/>) provide tools compatible with FLASH-seq data. Given their similarities, we currently recommend using Smart-seq2 guidelines when processing FLASH-seq data.

## Citations

### Step 5

Picelli S. Full-Length Single-Cell RNA Sequencing with Smart-seq2.  
[https://doi.org/10.1007/978-1-4939-9240-9\\_3](https://doi.org/10.1007/978-1-4939-9240-9_3)