Feb 19, 2020 Version 3

Smart-seq3 Protocol V.3

In 1 collection

DOI
dx.doi.org/10.17504/protocols.io.bcq4ivyw

Michael Hagemann-Jensen[1], Christoph Ziegenhain[1], Ping Chen[1], Daniel Ramsköld[1], Gert-Jan Hendriks[1], Anton J.M Larsson[1], Omid R. Faridani[1], Rickard Sandberg[1]

[1]Karolinska Institute Stockholm

Human Cell Atlas Method … single cell transcriptome

Michael Hagemann-Jensen
Karolinska Institutet

OPEN ACCESS

DOI: dx.doi.org/10.17504/protocols.io.bcq4ivyw

Protocol Citation: Michael Hagemann-Jensen, Christoph Ziegenhain, Ping Chen, Daniel Ramsköld, Gert-Jan Hendriks, Anton J.M Larsson, Omid R. Faridani, Rickard Sandberg 2020. Smart-seq3 Protocol. **protocols.io**
https://dx.doi.org/10.17504/protocols.io.bcq4ivyw

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: February 19, 2020

Last Modified: February 19, 2020

Protocol Integer ID: 33276

Guidelines

• If you are pro, and want to change the volumes indicated in the protocol, make sure that the reaction concentrations stay similar. Also ensure that the spillover concentrations into the next reaction also stay similar. Failing to do so can lead to decreased performance of the protocol. In terms of the tagmentation reaction make sure to keep the ratio of cDNA input to Nextera TN5 amount, if you contemplate to miniaturize this reaction.

• The current protocols is based on the TN5 from Nextera Xt kit. However Illumina TDE1, works as well.

• For determining PCR cycles, a good general guideline is to add 1-2 cycles more than using Smart-seq2. However as always, this is good to empirically test first, before running important samples.

List of oligos:

	Oligo	Vendor	Purification	Working concentration	Sequence	
	Smartseq3_OligodT30VN	IDT	HPLC	100uM	/5Biosg/ACGAGCATCAGCAGCATACGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTVN	
	Smartseq3_N8_TSO	IDT	RNase-Free HPLC	100uM	/5Biosg/AGAGACAGATTGCGCAATGNNNNNNNNrGrGrG	
	Fwd_PCR_primer	IDT	HPLC	100uM	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATTGCGCAA*T*G	
	Rev_PCR_primer	IDT	HPLC	100uM	ACGAGCATCAGCAGCATAC*G*A	

* phosphorothioate bonds

• It is absolute fine to use standard desalting instead of HPLC, both works fine in our hands, same goes for hand-mixed vs machine mixed degenerate bases. Using the regular DNA oligos service at IDT should provide based on their QC full length oligos.

• We use custom Nextera Indexes primers (standard 25 nmol oligo preps from IDT, delivered at 200 nM concentration in IDTE buffer) and we typically get performance that is indistinguishable from Illumina's primers.

For making your own primers, we recommend using the "DNABarcodes" R package. using the following settings:
Barcode length: 10 bp (or 8bp like Illumina primers, depending on the amount of cells you need indexed and sequenced at the same time)
Minimal levenshtein distance: 3
Filter out homopolymers >= 3
Filter for uneven GC content

Additionally, there seems to be an artifact on the NovaSeq platform for i5 index primers starting with the bases "AC", so we recommend to avoid those too!
(see supplementary information in this paper: https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-018-4703-0#Sec13)

- For troubleshooting: feel free to leave comments or message directly.

Materials

MATERIALS

- Triton X-100 **Sigma Aldrich Catalog #**T8787-50ML
- Agilent High Sensitivity DNA Kit **Agilent Technologies Catalog #**5067-4626
- Nextera XT DNA Library Preparation Kit **illumina Catalog #**FC-131-1096
- QIAquick Gel Extraction Kit **Qiagen Catalog #**28704
- DNA LoBind Tubes, 1.5 mL **Eppendorf Catalog #**0030108051
- Recombinant RNAse Inhibitor **Takarabio Catalog #**2313A
- Dithiothreitol (DTT) **Thermo Fisher Scientific Catalog #**707265ML
- dNTP Set 100 mM Solutions **Thermo Fisher Scientific Catalog #**R0182
- UltraPure™ DNase/RNase-Free Distilled Water **Thermo Fisher Catalog #**10977035
- EDTA (0.5 M), pH 8.0, RNase-free **Thermo Fisher Catalog #**AM9260G
- SDS, 10% Solution, RNase-free **Thermo Fisher Catalog #**AM9822
- Maxima H Minus Reverse Transcriptase (200 U/µL) **Thermo Fisher Catalog #**EP0751
- Poly Ethylene Glycol (PEG) 8000 **Sigma Aldrich Catalog #**89510-250G-F
- Sodium Chloride (5M) **Invitrogen - Thermo Fisher Catalog #**AM9760G
- Magnesium Chloride (1M Solution) **Invitrogen - Thermo Fisher Catalog #**AM9530G
- GTP (Tris buffered solution 100mM) **Thermo Scientific Catalog #**R1461
- Trizma-base **Sigma Aldrich Catalog #**T6791-100G
- KAPA HiFi Hotstart PCR kit **Roche Catalog #**KK2502
- Phusion High-Fidelity DNA Polymerase (2 U/µL) **Thermo Scientific Catalog #** F530L
- Sera-Mag Speed Beads **Ge Healthcare Catalog #**65152105050250
- Sodium Azide **Sigma Aldrich Catalog #**S2002-100G
- IGEPAL® CA-630 **Sigma Aldrich Catalog #**I8896
- Armadillo PCR Plate 96-well clear semi-skirted white wells **Thermo Scientific Catalog #**AB3596
- Armadillo PCR Plate 384-well **Thermo Scientific Catalog #** AB2384B
- QuantiFluor® dsDNA System **Promega Catalog #**E2670
- E-Gel™ EX Agarose Gels 2% **Thermo Scientific Catalog #**G402002
- 2-propanol **Sigma Aldrich Catalog #**I9516
- NN-Dimethylformamide **Sigma Aldrich Catalog #**D4551

Before starting

1 This protocol should be carried out in a clean environment. Use ethanol, RNAseZAP, DNA-OFF, or similar to prepare work bench before start.

Work quickly and preferably on On ice .

Prepare master-mixes right before use.

Use multichannel pipettes, liquid dispensers etc. to dispense the master-mixes. Avoid pipetting up and down, to minimize the loss of material.

• Take a look at the Guidelines section for more info about Oligos etc. used in this protocol.

Prepare lysis plates

2 Prepare **lysis buffer mix**

	Reagent	**Reaction conc.**	**uL per. reaction**	**96 well plate**	**384 well plate**
	Guanidine Hydrochloride (8000mM ; Optional)	0mM - 50mM	0.00	-	-
	Poly-ethylene Glycol 8000 (50% solution)	5%	0.40	44	164
	Triton X-100 (10% solution)	0.1%	0.03	3.3	12.3
	ERCC spike-ins (Optional)	-	-	-	-
	RNAse Inhibitor (40u/uL)	0.5u/uL	0.04	4.1	15.4
	OligodT30VN (100uM)	0.5uM	0.02	2.2	8.2
	dNTPs (25mM/each)	0.5mM/each	0.08	8.8	32.8
	Nuclease Free Water		2.43	267.6	997.3
	Total		**3 uL**	**330 uL**	**1230 uL**

Reaction concentrations for PEG8000, OligodT30VN and dNTPs, are adjusted to and reflect their concentration in the reverse transcription reaction (4uL)

The lysis master-mix contains PEG! Ensure that PEG is fully mixed into solution, by either pipetting up and down until the liquid is clear, or start with vortexing the required master-mix volume of water and PEG together before adding the remaining reagents.

Add 3 µL lysis buffer to each well of a 96/384 well plate.

Quick centrifugation to collect the lysis buffer before storage until use.

Sample collection

3 Sort single cells into 3 µL lysis in either 96 or 384 wells.

Seal with appropriate seals (-80 C to >100 C) and centrifuge the finished sorted plate immediately after, and transfer it to a -80 °C freezer or dry-ice.

Cell lysis

4 Remove the plate of sorted cells from the -80 freezer and incubate in a thermocycler with heated lid at 72 °C for 00:10:00 , followed by a 4 °C hold (keeping the storage seal sheet on the plate, unless damaged or loose).

Reverse Transcription

5 While the plate is incubating as per **step 4**, prepare the following **Reverse transcription master-mix**.

	Reagent	Reaction conc.	uL per reaction	96 well plate	384 well plate
	Tris-HCl pH 8.3 (1M)	25mM	0.1	11	41
	NaCl (1M)	30mM	0.12	13.2	49.2
	MgCl2 (100 mM)	2.5mM	0.1	11	41
	GTP (100 mM)	1mM	0.04	4.4	16.4
	DTT (100 mM)	8mM	0.32	35.2	131.2
	RNase Inhibitor (40u/uL)	0.5u/uL	0.05	5.5	20.5

TSO (100uM)	2uM	0.08	8.8	32.8
Maxima H-minus RT enzyme (200 U/uL)	2u/uL	0.04	4.4	16.4
Nuclease Free Water		0.15	16.5	61.5
Total		**1uL**	**110uL**	**410uL**

Add 1 µL RT mix to each well of a 96/384 well plate.

Replace the storage seal with a PCR seal. Ensure that the plate is properly sealed, to avoid evaporation.

Do a quick centrifugation to collect reaction at the bottom, before incubating the plate in a thermocycler at;

42 °C	90 min	1x
50 °C	2 min	10x
42 °C	2 min	
85 °C	5 min	1x

Preamplification PCR

6

Start preparing the **PCR mix,** when the incubation of the reverse transcription reaction is near completion, by combining the following components.

! Note that the KAPA DNA polymerase has a 3-5' exonuclease activity that is not HotStart. Therefore add polymerase just before using the master-mix.

Reagent	**Reaction conc.**	**uL per. reaction**	**96 well plate**	**384 well plate**
Kapa HiFi HotStart buffer (5X)	1X	2.0	220	820

dNTPs (25mM/each)	0.3mM/each	0.12	13.2	49.2
MgCl2 (100mM)	0.5mM	0.05	5.5	20.5
Fwd Primer (100uM)	0.5uM	0.05	5.5	20.5
Rev Primer (100uM)	0.1uM	0.01	1.1	4.1
Polymerase (1U/uL)	0.02 U/uL	0.2	22	82
Nuclease Free Water		3.57	392.7	1463.7
Total		**6uL**	**660uL**	**2460uL**

Add 6 µL PCR mix to each well of a 96/384 well plate.

Quick centrifugation to collect reaction at the bottom, before running the following PCR program in a thermocycler.

Step	**Temperature**	**Time**	**Cycles**
Initial denaturation	98 °C	3 min	1x
Denaturation	98 °C	20 sec	18-25 x
Annealing	65°C	30 sec	
Elongation	72 °C	4 min	
Final Elongation	72 °C	5 min	1x
Hold	4 °C	Hold	

The PCR cycle number depends on the input, and is very cell-type specific. See the Guidelines & warnings for help determining PCR cycles needed.

cDNA purification (preferable but optional)

7 Before purification prepare **22% PEG Clean-up Beads** used for cleaning up the preamplified cDNA**.** These beads perform similar to Ampure XP beads. Beads are prepared as per mcSCRB-seq protocol

CITATION

Johannes Bagnoli, Christoph Ziegenhain, Aleksandar Janjic, Lucas Esteban Wange, Beate Vieth, Swati Parekh, Johanna Geuder, Ines Hellmann, Wolfgang Enard. mcSCRB-seq protocol. Nature Communications.

LINK

dx.doi.org/10.17504/protocols.io.p9kdr4w

	Reagent	**Amount**
	PEG 8000	11 g
	NaCl, 5M	10 mL
	Tris-HCL, 1 M, pH 8.0	500 µL
	EDTA, 0.5M	100 µL
	IGEPAL, 10% solution	50 µL
	Sodium Azide, 10% solution	250 µL
	UltraPure Water	up to 49 mL
	Total	**49 mL**

Add all ingredients into a 50 mL falcon tube, but do not add the total amount of water until after PEG is completely solubilized.

Incubate at 40 °C and **vortex regularly until PEG is completely dissolved.**

Resuspend bead stock carefully (Sera-Mag Speed Beads).

Pipette 1000 µL of bead suspension into a 1.5 mL tube.

Place on magnet stand Remove supernatant.

Add 1000 µL 10 mM Tris-HCl, pH 8.0, 1 mM EDTA (TE), and resuspend beads. Place on magnet stand.

Remove supernatant. Repeat wash one more time.

Add 900 µL 10 mM Tris-HCl, pH 8.0, 1 mM EDTA (TE), and resuspend beads.

Add to PEG solution above and mix well.

8 1. To purify cDNA add **0.6:1** ratio of 22% PEG beads to sample, and mix by gently pipetting up and down.

2. Incubate at Room temperature for 00:08:00 .

3. Place on magnet and allow beads to settle. Roughly 00:05:00 .

4. Discard supernatant, and wash once with 20 µL / 100 µL of **freshly prepared 80% Ethanol** for 384 / 96 well plates respectively.

5. Remove Ethanol and let the beads air dry for 00:02:00 - 00:05:00

6. Elute cDNA in 12 µL UltraPure Water, resuspend beads and incubate for 00:05:00 .

DNA concentration measurement and normalization (Optional, but recommended)

9 1. Prepare 1X TE buffer by either diluting the 20X TE buffer from the QuantiFluor® dsDNA kit or by preparing a solution of 10mM Tris-HCl, 0.1mM EDTA, pH 8.

2. Dilute the QuantiFluor® dsDNA Dye 1:400 in 1X TE buffer and mix.

3. Prepare dsDNA standards for plate read-out, according to manufacturers protocol.

4. Dispense 49 µL / 99 µL per well of the ready Quantiflour dye mix into black, flat-bottom 384/96 well plates, respectively.

5. Add Standards to a separate plate.

6. Add 1 µL of cDNA to each well. Incubate asssay for 00:05:00 at Room temperature

7. Use a plate reader, to measure fluorescence (504nM Excitation/ 531nM Emission)

8. Calculate cDNA concentration.

9. Calculate water needed to dilute 1 µL cDNA to [M] 100 pg/uL .

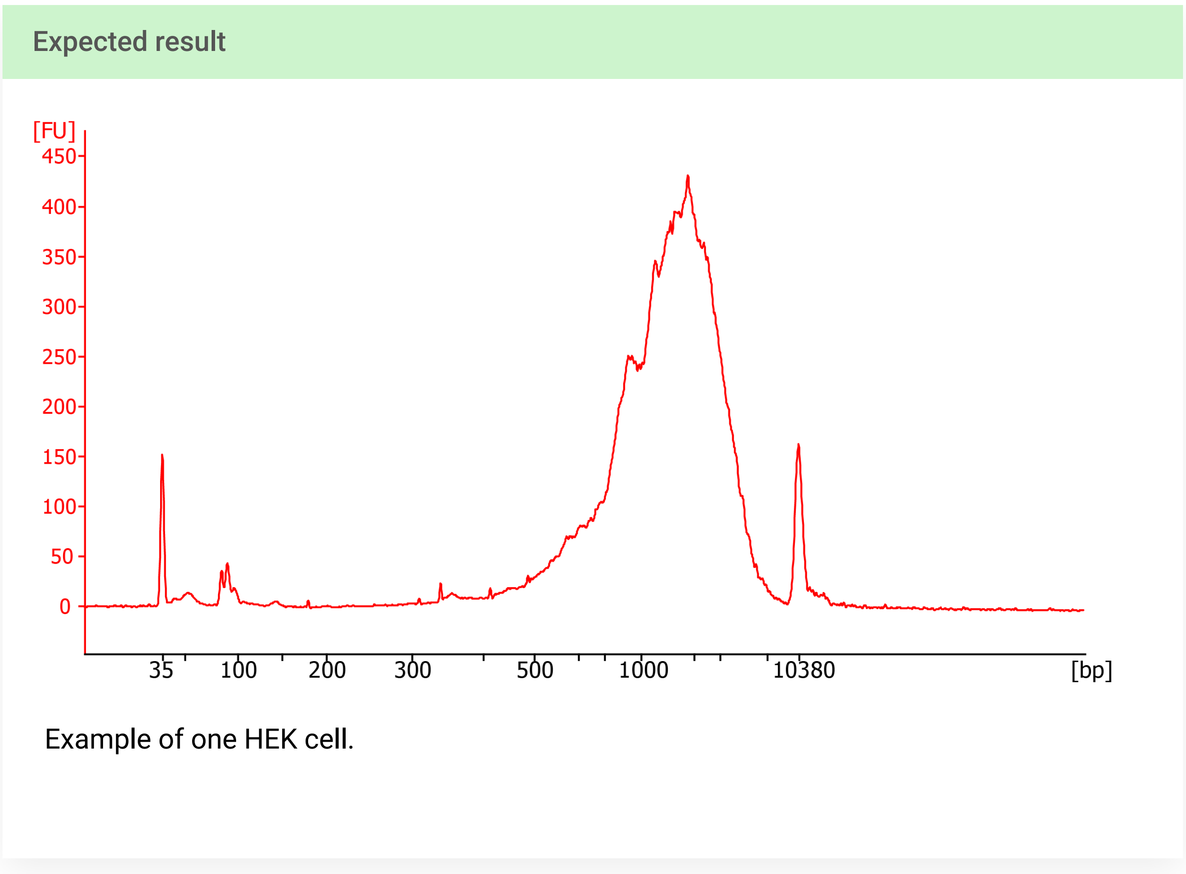
10 1. Prepare **normalization plate** by adding the calculated water volumes to each well.

2. Add 1 µL of preamplified cDNA to each well.

Quality Control check (Optional)

11 Check the cDNA preamplification library content and quality on a Agilent Bioanalzyer, using High Sensitivity DNA Analysis chips.

Expected result



Example of one HEK cell.

Tagmentation

12

1. Prepare 4x Tagmentation buffer as following. Aliquots of 4xTD buffer can be stored for later use. The TD buffer (2x) from Nextera Kits can also be used, however with the current small amount of ATM used, the Illumina TD buffer will at some point run out.

Safety information

Dimethylformamide (DMF) should be handled in a fume hood and according to local safety regulations.

	Reagent	**Amount (uL)**	**Concentration in 4X**	
	Tris-HCl pH 7.5 (1 M)	40	40mM	
	MgCl2 (100	200	20mM	

mM)			
Dimethylformamide (DMF)	200	20%	
UltraPure Water	560		
Total	**1000 uL**		

2. Prepare **Tagmentation mix.**

• Please note that the ATM amount is a suggested starting point for 100pg/uL input, and some optimization might be necessary to reach a desired UMI-read to Internal-read ratio, based on input and celltype..

Reagent	**Reaction conc.**	**uL per. reaction**	**96 well plate**	**384 well plate**
Tagmentation buffer (4x)	1X	0.5	55	205
Amplicon Tagmentation Mix (Tn5)		0.08	8.8	32.8
UltraPure water		0.42	46.2	172.2
Total		**1uL**	**110uL**	**410uL**

3. Dispense 1 µL of Tagmentation mix to a new 96 or 384 well plate.

4. Add 1 µL of **normalized 100pg/uL cDNA (step 10)** to the plate containing tagmentation mix.

5. Apply a quick spin-down of the plate before incubation in a thermocycler at 55 °C for 00:10:00 .

6. To strip off the Tn5 from the DNA, add 0.5 µL of 0.2% SDS to each well. Centrifuge quickly and incubate for 00:05:00 .

7. **Concerning Nextera Index primers: We highly suggest to design or order custom Nextera Index primers. This ensures higher flexibility while also being much cheaper in the long run! The following protocol is designed as such. If using Nextera index primers purchased from Illumina, dilute all primers 5x with UltraPure water, and proceed to use similar volume as follows.**

8. Add 1.5 µL **Nextera Index primers t**o each well as follows

	Reagent	Reaction conc.	uL per. reaction		
	Custom S50X index primer (0.5uM)	0.1uM	0.75		
	Custom N70X index primer (0.5uM)	0.1uM	0.75		

9. Prepare **Tagmentation PCR mix**.

	Reagent	Reaction conc.	uL per. reaction	96 well plate	384 well plate
	Phusion HF buffer (5X)	1X	1.4	154	574
	dNTPs (25 mM/each)	0.2mM/each	0.06	6.2	23
	Phusion HF (2U/uL)	0.01 U/uL	0.04	3.9	14.4
	H2O		1.51	166	618.7

	Total		3uL	330uL	1230uL

10. Add 3 µL of **Tagmentation PCR mix** to each well, centrifuge quickly and incubate in a thermocycler using the following PCR program.

	Step	Temperature	Time	Cycles
	Gap-filling	72 °C	3 min	1x
	Initial denaturation	98 °C	3 min	1x
	Denaturation	98 °C	10 sec	12x
	Annealing	55 °C	30 sec	
	Elongation	72 °C	30 sec	
	Final Elongation	72 °C	5 min	1x
	Hold	4 °C	Hold	

Library clean-up

13 For the final library clean-up, pool all the Tagmented cDNA (step 12) sample wells in a 1.5mL or 5mL eppendorf tube.

1. Add **0.6:1 22% PEG beads to final volume of the pooled tagmentation cDNA**. Mix gently by pipetting and incubate for 00:08:00 at Room temperature

2. Place on magnet and allow beads to settle. Roughly 00:05:00 .

4. Discard supernatant, and wash twice with >= 1000 µL **freshly prepared 80% Ethanol.**

5. Remove Ethanol and let the beads air dry for at least 00:05:00

6. Elute cDNA in 40 µL UltraPure Water, resuspend beads and incubate for 00:05:00

.

(Optional) Size selection via Gel-cutting and extraction

14 To further select for longer tagmented fragments, an optional step including a size selection step can be included.

1. Load 20 µL of the eluted tagmented library from step 13 into a **2% Agarose E-Gel EX** together with 50bp DNA ladder.

2. Run gel for 00:12:00

3. After finished run, open the gel casing and cut the gel between **550bp - 2kb** using a clean scalpel or blade.

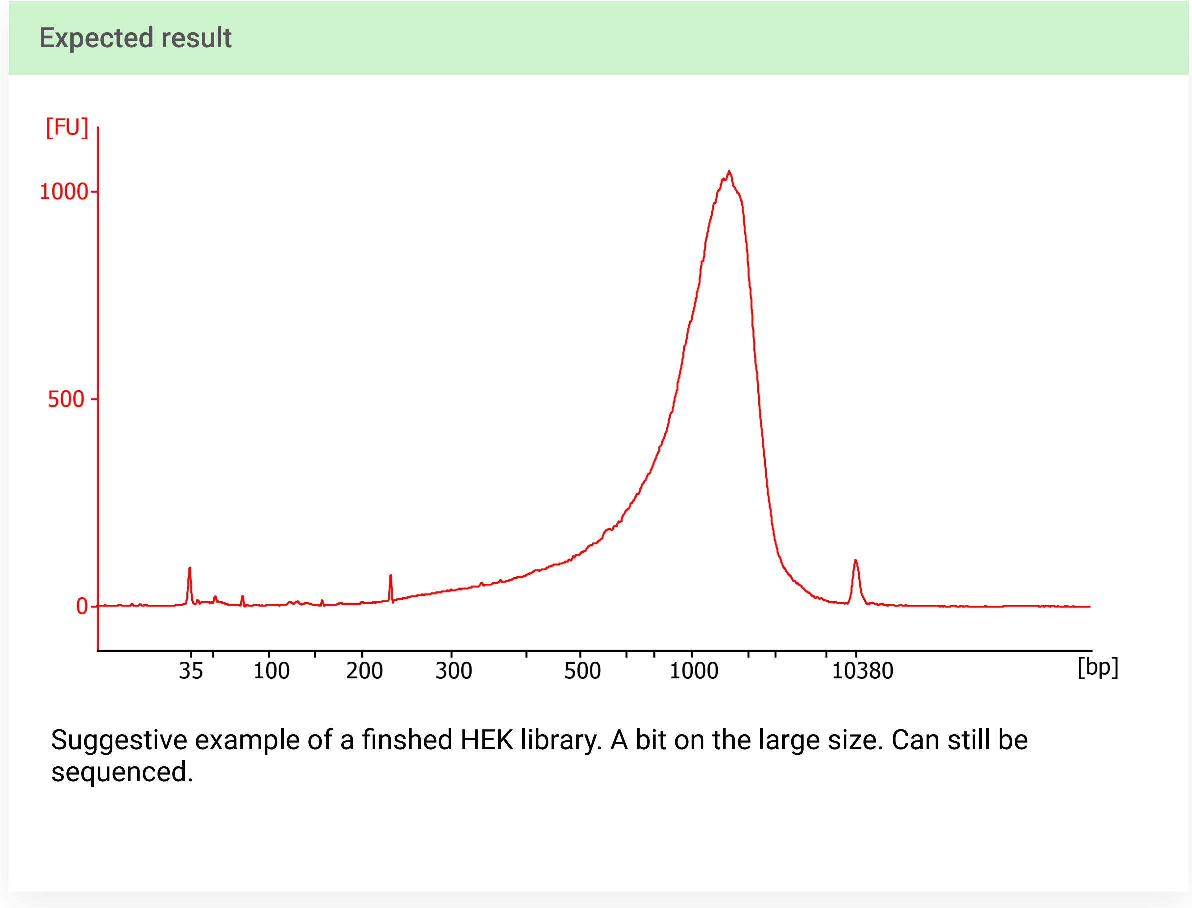
4. Purify the excised gel slice using Qiagen QIAquick Gel extraction kit according to manufacturers protocol.

Final Library Quantification

15

Run the final library on a Agilent Bioanalyzer (High Sensitivity DNA), to inspect the quality and median base-pair length of your library.

Expected result



Suggestive example of a finshed HEK library. A bit on the large size. Can still be sequenced.

Use Qubit fluorometer or similar to quantify the library.

Calcutate the final library concentration, using above metrics.

Sequencing

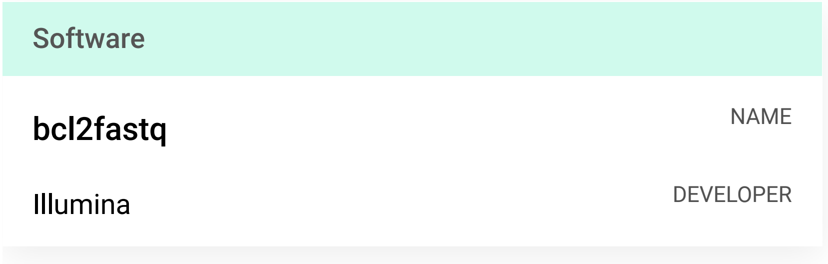
16 The sequencing ready library should be sequenced on any Illumina compatible sequencer, either Single-end or Paired-end, depending on the question and need.

For final library whether gel cut/size selected or not, the expected median base-pair should be around or above 1kb. In our experience NovaSeq/HiSeq sequncers are more tolerant towards wider or longer size fragment distributions, than the NextSeq. Because of this consider increasing the loading concentration a bit to ensure proper cluster density. However "your milage may vary". Empircal investigation or a pilot run is always adviced, if possible.

Primary Data processing

17 After sequencing has completed successfully, binary base-call files (BCL) need to be converted to fastq.

For this, bcl2fastq should be used in the latest version (bcl2fastq v2.20).



At this stage, demultiplexing into per-cell fastq files is not necessary - a sample sheet is thus not needed.
Be sure to adjust the base mask to represent your sequencing layout and the length of your barcode reads.
Remove the option --no-lane-splitting if the same cell barcodes have been reused for different libraries on different lanes of the flow cell.
You may restrict the number of cores used with the following options:
--loading-threads
--processing-threads
--writing-threads

Command

bcl2fastq: 2x150bp dual-index

```
bcl2fastq --use-bases-mask Y150N,I8,I8,Y150N --no-lane-splitting --
create-fastq-for-index-reads -R
/mnt/storage1/NextSeqNAS/191011_NB502120_0154_AHVG7JBGXB
```

-R denotes the runfolder and you may redirect the fastq output to a different folder with the -o option.

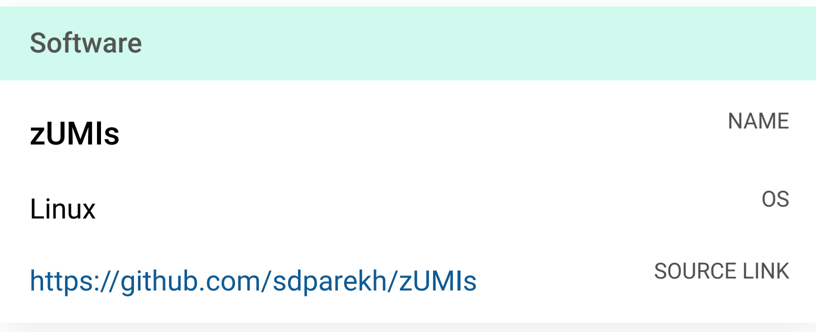
18 After generating fastq files, the zUMIs pipeline should be used to process Smart-seq3 data to ensure correct handling of UMI reads and internal reads.

CITATION

Parekh S, Ziegenhain C, Vieth B, Enard W, Hellmann I (2018). zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs.. GigaScience.

LINK

https://doi.org/10.1093/gigascience/giy059



We recommend the newest version v2.5.6 at the time of this protocol.

All options are set in a configuration file following the YAML format. Here is a best practice example:

Command

Smartseq3.yaml


```
project: Smartseq3
sequence_files:
  file1:
    name: /smartseq3/fastq/Undetermined_S0_R1_001.fastq.gz
    base_definition:
      - cDNA(23-150)
      - UMI(12-19)
    find_pattern: ATTGCGCAATG
  file2:
    name: /smartseq3/fastq/Undetermined_S0_R2_001.fastq.gz
    base_definition:
      - cDNA(1-150)
  file3:
    name: /smartseq3/fastq/Undetermined_S0_I1_001.fastq.gz
    base_definition:
      - BC(1-8)
  file4:
    name: /smartseq3/fastq/Undetermined_S0_I2_001.fastq.gz
    base_definition:
      - BC(1-8)
reference:
  STAR_index: /resources/genomes/Mouse/STAR5idx_noGTF/
  GTF_file: /resources/genomes/Mouse/Mus_musculus.GRCm38.91.gtf
  additional_STAR_params: '--clip3pAdapterSeq CTGTCTCTTATACACATCT'
  additional_files:
    - /resources/genomes/spikes/ERCC92.fa
out_dir: /smartseq3/zUMIs/
num_threads: 20
mem_limit: 50
filter_cutoffs:
  BC_filter:
    num_bases: 3
    phred: 20
  UMI_filter:
    num_bases: 2
    phred: 20
barcodes:
  barcode_num: ~
  barcode_file: /smartseq3/expected_barcodes.txt
  automatic: no
```

```
    BarcodeBinning: 1
    nReadsperCell: 100
    demultiplex: no
  counting_opts:
    introns: yes
    downsampling: '0'
    strand: 0
    Ham_Dist: 1
    write_ham: no
    velocyto: no
    primaryHit: yes
    twoPass: no
  make_stats: yes
  which_Stage: Filtering
  samtools_exec: samtools
  pigz_exec: pigz
  STAR_exec: STAR
  Rscript_exec: Rscript
```

Be sure to use full paths to all files and folders. For further descriptions of the individual options visit the **zUMIs GitHub repository wiki**

Now, simply start zUMIs with the following command:

Command

Invoke zUMIs

```
zUMIs-master.sh -y Smartseq3.yaml
```

Citations

Step 18

Parekh S, Ziegenhain C, Vieth B, Enard W, Hellmann I. zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs.
https://doi.org/10.1093/gigascience/giy059

Step 7

Johannes Bagnoli, Christoph Ziegenhain, Aleksandar Janjic, Lucas Esteban Wange, Beate Vieth, Swati Parekh, Johanna Geuder, Ines Hellmann, Wolfgang Enard. mcSCRB-seq protocol
dx.doi.org/10.17504/protocols.io.p9kdr4w