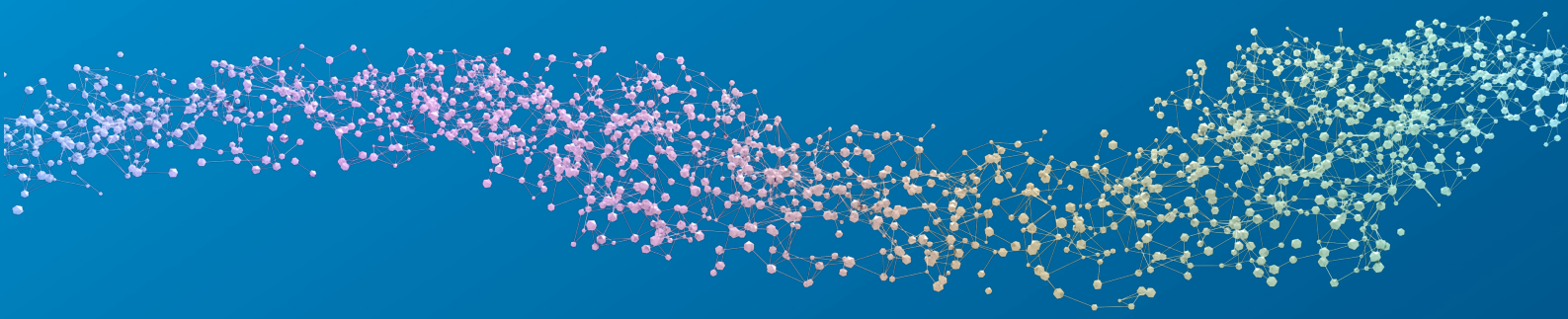
ALITHEA GENOMICS



MERCURIUS™

Spheroid DRUG-seq Library Preparation Kit for 96, 384, and 1’536 Samples

PN 10870, 10875, 11670, 11675

User Guide

November 2025
(Early-Access)

Related Products

Kit name	Kit PN	Modules	Module PN
Mercurius™ Spheroid DRUG-seq Library Preparation 96 kit	10870	Barcoded Oligo-dT Adapters Module 96 samples	10513
		Spheroid DRUG-seq Library Preparation and UDI Module 96 samples	10543
Mercurius™ Spheroid DRUG-seq Library Preparation 384 kit	10875	Barcoded Oligo-dT Adapters Module 384 samples	10515
		Spheroid DRUG-seq Library Preparation and UDI Module 384 samples	10544
Mercurius™ Spheroid DRUG-seq Library Preparation 384 (4x 96) kit	11670	Barcoded Oligo-dT Adapters Module 4x 96 samples	10513
		Spheroid DRUG-seq Library Preparation and UDI Module 4x 96 samples	10557
Mercurius™ Spheroid DRUG-seq Library Preparation 1'536 (4x 384) kit	11675	Barcoded Oligo-dT Adapters Module 4x 384 samples	10515
		Spheroid DRUG-seq Library Preparation and UDI Module 4x 384 samples	10559

Table of Contents

Kit Components

Reagents supplied

Barcoded Oligo-dT Adapters Set V5 Module

Component name	Label	Amount provided per kit				Storage
		96 samples (PN 10870)	384 samples (PN 10875)	4x 96 samples (PN 11670)	4x 384 samples (PN 11675)	
Plate with 96 barcoded oligo-dT primers, set V5D (PN 10419)	96 V5D OdT	1 plate	-	4 plates	-	-20°C
Plate with 384 barcoded oligo-dT primers, set V5D (PN 10420)	384 V5D OdT	-	1 plate	-	4 plates	-20°C
Aluminium Seal	-	3 pcs	3 pcs	12 pcs	12 pcs	-20°C/RT

Spheroid DRUG-seq Library Preparation and UDI Module

Component name	Label	Cap colour	Volume, µL				Storage
			96 samples (PN 10543)	384 samples (PN 10544)	4x 96 samples (PN 10557)	4x 384 samples (PN 10559)	
RT Enzyme	RTE	magenta	11	43	43	4x 43	-20°C
RT Buffer	RTB	magenta	550	1100	2x 1100	4x 1100	-20°C
Exonuclease I Enzyme	EXO	purple	10	10	10	10	-20°C
Exonuclease Buffer	EXB	purple	20	20	20	20	-20°C
Second Strand Enzyme	SSE	orange	20	20	20	20	-20°C
Second Strand Buffer	SSB	orange	45	45	45	45	-20°C
Tagmentation Enzyme	TE	red	12	12	12	12	-20°C
Tagmentation Buffer	TAB	red	40	40	40	40	-20°C
Library Amplification Mix	LAM	green	200	200	200	200	-20°C
UDI Adapter Mix 1	MQ.UDI.1	transparent	10	10	10	10	-20°C
UDI Adapter Mix 2	MQ.UDI.2	transparent	10	10	10	10	-20°C
UDI Adapter Mix 3	MQ.UDI.3	transparent	10	10	10	10	-20°C
UDI Adapter Mix 4	MQ.UDI.4	transparent	10	10	10	10	-20°C
Spheroid Dissociation Reagent	SDR	yellow	650	650	4x 650	4x 650	-20°C
Spheroid Lysis Buffer	SLB	yellow	1150	2x 1150	4x 1150	8x 1150	-20°C
Spheroid RNase Inhibitor	SRI	yellow	520	520	4x 520	4x 520	-20°C

Additional required reagents and equipment (supplied by the user)

Plasticware	Manufacturer	Product number
15 mL conical tubes	Greiner	188271
0.2 mL 8-strip Non-Flex PCR tubes	Starlab	I1402-3700
Zymo-Spin IIICG columns (optional)	Zymo	C1006-50-G
25 mL reservoir for spin columns	Zymo	C1039-25
Disposable pipetting reservoir 25 mL polystyrene	Integra	4382
Solution reservoir 150 mL polystyrene (optional)	Integra	6318

Reagents	Manufacturer	Product number
DNA Clean and Concentrator-5 kit	Zymo	D4014
Magnetic beads:		
• SPRI AMPure Beads	Beckman Coulter	A63881
• CleanNA Beads	CleanNA	CNGS0050D
• Mag – Bind TotalPure NGS	Omega Bio-tek	M1378-00
• MagFlo NGS	Integra	7000
• MagMax PureBind	Applied biosystems	A58521
Qubit™ dsDNA HS assay kit	Invitrogen	Q32851
High sensitivity NGS fragment analysis kit	Agilent	DNF-474
SYBR Green	Thermo Fisher	S7563
Ethanol, 200 proof	-	-
Nuclease-free water	Thermo Fisher	A57775
DPBS, cell culture grade	Gibco	10010023
ERCC RNA spike-in mix	Thermo Fisher	4456740

Equipment	Manufacturer	Product number
Benchtop centrifuge for plates	-	-
Benchtop centrifuge for 1.5 mL tubes	-	-
Single and multichannel pipettes	-	-
Fragment Analyser / Bioanalyzer / TapeStation	Agilent	M5310AA
Qubit™	Invitrogen	Q33238
Magnetic stand for 0.2 mL tubes	Permagen	MSR812
Magnetic stand for 1.5 mL tubes	Permagen	MSR06
12-channel pipette, 0.5-12.5 µL VIAFLO or similar	Integra	4631
12-channel pipette, 5-125 µL VIAFLO or similar	Integra	4632
8-channel adjustable tip spacing pipette, VOYAGER, 2 – 50 µL	Integra	4726
Pipetboy	Integra	155 000
Vacuum manifold Qiagen Qiavac 24 Plus or similar	Qiagen	19413
Vacuum pump	-	-
Real-Time PCR instrument (optional)	-	-
VIAFLO instrument (optional)	Integra	6001
VIAFLO 96/384 channel pipetting head, 0.5-12.5 µL (optional)	Integra	6101/6131

Protocol Overview and Timing

The MERCURIUS™ Spheroid DRUG-seq kits allow the preparation of Illumina-compatible 3' RNA sequencing libraries for up to 1536 spheroid samples in a time and cost-efficient manner. The kits include a three-component lysis buffer for efficient digestion of spheroids and subsequent cell lysis. Crude lysates obtained can be used directly in the RT reaction, skipping the tedious RNA purification step. **Important!** This protocol is designed to work exclusively with **non-matrix-grown spheroids** and small organoids, grown in U-shaped or specific culture formats (i.e., Akura™ plates).

The kits are provided in the following formats:

Kit format	PN	PCR plate format	Maximum number of samples in one pool	Maximum number of samples processable	Number of UDI libraries
96-sample	10870	96WP	96	96	4
384-sample	10875	384WP	384	384	4
4x 96-sample	11670	96WP	96	384	4
4x 384-sample	11675	384WP	384	1536	4

Every kit contains barcoded MERCURIUS™ Oligo-dT primers, designed to tag polyA+ RNA from lysate samples with individual barcodes during the first-strand synthesis reaction. This enables the pooling of the resulting cDNA samples from each experimental group into a single tube, facilitating streamlined sequencing library preparation.

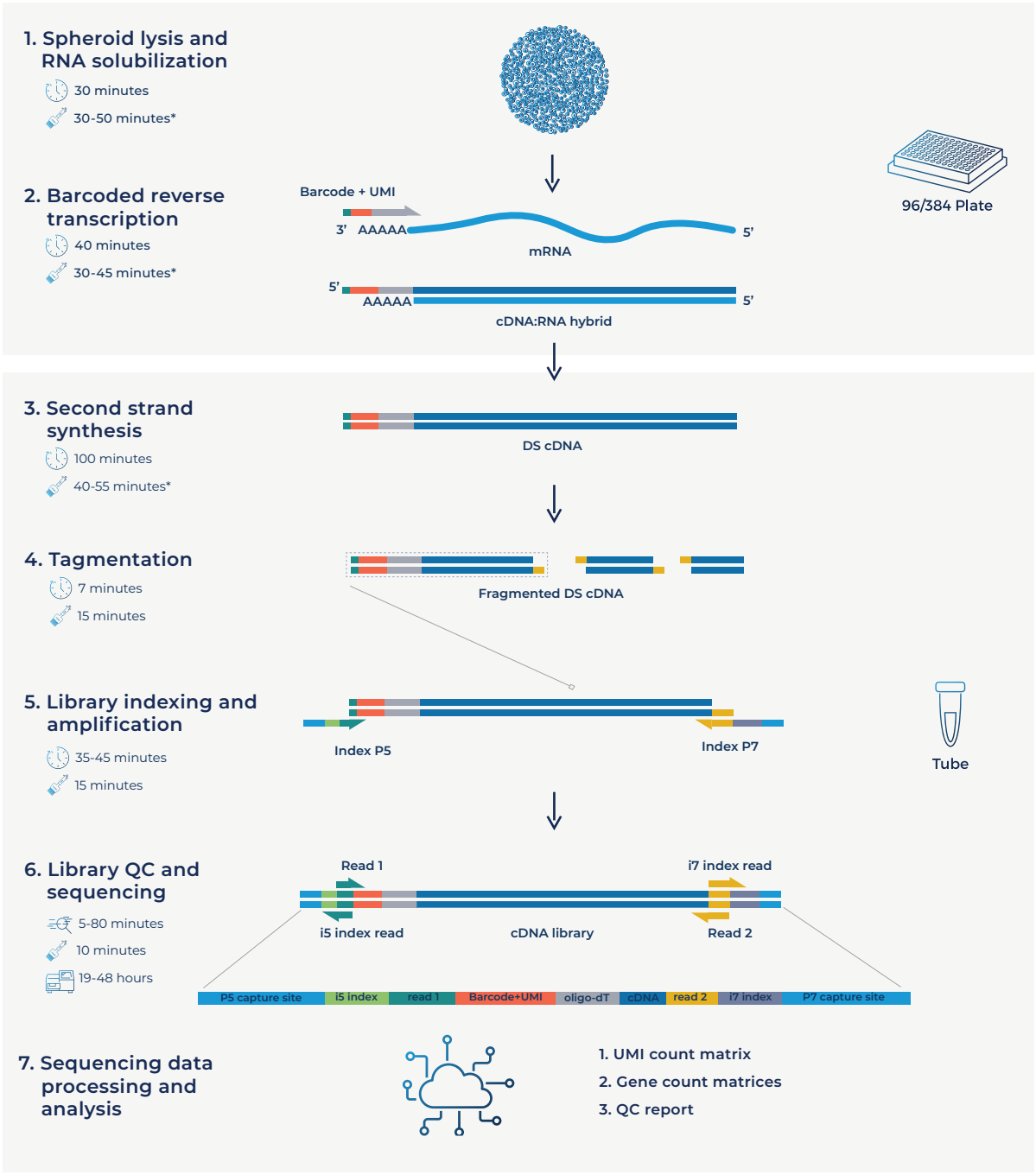
The DRUG-seq technology can be used to generate high-quality sequencing data from 2'000 – 50'000 mammalian cells per well. Notably, the kit can be used to pool any number of samples up to the capacity of the provided plate (96 or 384) with two considerations:

- The total cell number per pool should be at least 80000 cells.
- Pooling less than eight samples may result in low-complexity reads during sequencing, decreasing the overall sequencing quality. If necessary, the latter can be improved by increasing the proportion of PhiX spike-in control during sequencing (see Part 3).

Each library indexing is performed using a Unique Dual Indexing (UDI) strategy, which minimizes the risk of barcode misassignment after NGS. The kits contain four UDI adapters. Every adapter can be used to prepare an individual library. Libraries with different UDI adapters can be pooled and sequenced in a single flow cell.

Figure 1 provides an estimation of the time required to accomplish each step of the protocol.

Protocol Workflow



Overall time

Incubation time: 3h30-3h45.
Hands-on time: 2h20-3h10.
QC time: 5min-1h20 (depending on the instrument used: Qubit or Fragment Analyzer).
Sequencing time: 19-48 hours (depending on instrument and sequencing run settings).

* The first provided time is for a 96 well-plate, followed by the time for a 384 well-plate.

Figure 1 Schematic illustration of the protocol workflow

Part 1. PREPARATION OF SPHEROID LYSATE SAMPLES

1.1. Essential considerations for input material

- The spheroid lysate protocol was validated using spheroids cultured in U-shaped or microtissue-specific 96-well (**96WP**) or 384-well (**384WP**) plates (e.g., Akura^{TM}, BIOFLOAT^{TM}).
- The recommended input range of cells in spheroids is 5000-50000 cells/well of **96WP** and 2000-10000 cells/well of **384WP** on the day of harvesting.
- Spheroids should be plated a few days in advance for optimal results.
- Depending on the type of spheroids (human, mouse, metastatic, or primary cells) and experimental design (e.g., drug treatment, induction of apoptosis, cell cycle arrest, etc.), consider the cell doubling time and potential effect of the treatment on cell quality and quantity.
- To ensure an even distribution of reads after sequencing, the amount of starting material must be as uniform as possible. For this, we suggest using automated cell seeding instruments or double-verified cell counts.

1.2. ERCC Spike-in Controls (Optional)

To ensure the uniformity of sequencing reads across samples and to assess the impact of sample and library preparation steps on this, we recommend adding External RNA Controls Consortium (ERCC) Spike-Ins to the lysate buffer (Thermo Fisher, 4456740). Please refer to Appendix 1 for detailed information before proceeding with the lysis step.

1.3. Spheroid pellet preparation

At this step, plated spheroids are washed with DPBS and frozen at -80°C for at least 5 minutes. If possible, snap-freeze the plate with dry ice or liquid nitrogen beforehand.

Preferably, perform the lysis step procedure just before the reverse transcription reaction.

NOTE: The freezing step is required to achieve a higher exon mapping.

1.3.1. Gently aspirate culture media from the plate and wash Spheroids by adding DPBS (per well):

- **96WP** (U-bottom plates): 80-100 µL DPBS
- **96WP** (Microtissue-specific plates) and **384WP**: 20µL DPBS

1.3.2. Centrifuge at 300x g for 3 min, if necessary.

1.3.3. Gently tap the plate and aspirate as much DPBS as possible without disturbing the spheroids' structure.

1.3.4. Seal the plate well with an Aluseal and immediately transfer it to a -80°C freezer for storage. If possible, snap-freeze the plate with dry ice beforehand.

NOTE: If multiple plates need to be processed, perform the procedure with each plate individually, one at a time, to prevent prolonged storage at room temperature.

1.3.5. Proceed to step 1.4.1 for spheroid lysis.

1.4. Spheroid lysate preparation

At this step, frozen spheroids are lysed directly in a 96- or 384-well plate by adding 1x Spheroid Lysis Buffer to the wells. The lysates can be used directly for the downstream RT.

Depending on the type of plates, we recommend using:

- **96WP** (U-bottom plates): 30 µL
- **96WP** (Microtissue-specific plates) and **384WP:** 10µL

Preparation

- Thaw the **SDR, SLB,** and **SRI** tubes on ice.
- Mix well and briefly spin down before use.
- Prepare a working solution of 1x Spheroid Lysis Buffer according to the type and number of wells. Add the water to a 15-mL Falcon tube, followed by the SDR, SLB, and SRI (in this particular order).

Reagent	96WP (U-bottom), µL		96WP (Microtissue-specific), 384WP, µL	
	Per well	96 wells	Per well	384 wells
SDR	3	330	1	430
SLB	10.2	1122	3.4	1462
SRI	2.4	264	0.8	344
Water	14.4	1584	4.8	2064
TOTAL	30	3300	10	4300

Gently pipette a prepared mix a few times and briefly spin the tube. Keep the mix on ice until further use.

Procedure for spheroid lysis

1.4.1. Remove the plate with frozen spheroids from the -80°C and put it on ice.

1.4.2. Quickly spin it down at 300x g for 1 min to ensure that all material is collected on the bottom.

1.4.3. Using a multi-dispenser in every well, distribute the prepared 1x SLB:

- **96WP:** 30 µL per well
- **384WP**: 10 µL per well

1.4.4. Centrifuge the plate at 300x g for a few seconds to ensure that 1x SLB is uniformly distributed across the surface of each well.

1.4.5. Incubate the plate at +37°C for 5 min.

1.4.6. Centrifuge the plate at 300x g for a few seconds to collect all the liquid.

1.4.7. Gently pipette the SLB 10 times. To minimize bubbles, avoid excessive pipetting.

1.4.8. Incubate the plate at +37°C for 5 min.

1.4.9. Centrifuge the plate at 300x g for a few seconds to collect all the liquid.

1.4.10. Perform a second round of pipetting.

1.4.11. Centrifuge the plate at 300x g for 3 min.

1.4.12. Label a new 96- or 384-well PCR plate.

1.4.13. Transfer the required volume of spheroid lysates from every well to a corresponding well of the provided 96- or 384-well plate with oligos (follow step 2.1.1 for the volumes). Preferably, use a multichannel pipette and avoid transferring any cell pellet.

Part 2. LIBRARY PREPARATION PROTOCOL

NOTE: Before starting every step, briefly spin down the tubes and plates before opening them to ensure that all liquid or particles are collected at the bottom of the tube/plate.

2.1. Reverse transcription

Each spheroid lysate sample is reverse-transcribed at this step using the barcoded oligo-dT primers provided in a 96- or 384-well plate format, depending on the kit type. Subsequently, all barcoded samples can be pooled into a single tube.

NOTE: Barcoded oligo-dT primers are provided lyophilized with the addition of dye. The dye has no impact on the enzymatic reactions and is used solely for better visualization of reaction preparation and pooling.

Despite variations in appearance caused by the drying process, wells may exhibit traces of dried dye ranging from dispersed to solid dots on the bottom. The following addition of RT reagents will enable the visualization of red color, confirming the presence of the oligos in all wells.

Preparation

- Keep the spheroid lysate samples on ice
- Thaw the **RTB** reagent at room temperature and mix well before use.
- Briefly spin down the **96WP** or **384WP** plate containing dried oligo-dT primers. This plate will be referred to as the RT plate.
- Prepare program 1_RT on the thermocycler (set the lid at 90°C):

Step	Temperature, °C	Time
Incubation	50	30 min
Inactivation	85	10 min
Keep	4	pause

NOTE: All manipulations with spheroid lysates and RT enzyme should be performed in an RNase-free environment, using RNase-free consumables and filter tips, on ice, and with gloves.

Procedure for 96WP and 384WP

2.1.1. Keep the RT plate on ice. Using a multichannel pipette, transfer the following volume of Spheroid lysate directly to the corresponding wells and pipette 3-5 times to ensure proper reconstitution of dried oligo-dT:

- **96WP**: 14.9 µL
- **384WP**: 7.4 µL

2.1.2. The appearance of red color in all wells indicates a proper and uniform reconstitution of oligos.

2.1.3. Carefully re-seal the RT plate and briefly spin it in the centrifuge.

2.1.4. Leave the RT plate on ice for 5 min.

2.1.5. Prepare the Master Mix for the RT reaction (with 10% excess) as follows:

Reagent	96WP, µL		384WP, µL	
	Per well	96 wells	Per well	384 wells
RTB	**5**	**550**	**2.5**	**1075**
RTE	**0.1**	**11**	**0.1**	**43**
TOTAL	**5.1**	**561**	**2.6**	**1118**

2.1.6. Keep the RT plate on ice and, using a multichannel pipette, transfer the following volume of Master Mix to each well containing the spheroid lysate sample:

- **96WP**: 5.1 µL
- **384WP**: 2.6 µL

2.1.7. Carefully re-seal the RT plate and briefly spin it in the centrifuge.

2.1.8. Transfer the plate to the thermocycler and start Program 1_RT.

Safe stop: After this step, the RT plate can be kept at 4°C overnight.

2.2. Sample pooling and column purification

After pooling, the barcoded RT samples can be purified using either column-based Zymo Clean & Concentration Kit (Zymo, D4014) or SPRI magnetic beads (Beckman, A63881). Both approaches produce comparable outcomes and can be used interchangeably. Depending on the availability of 3rd party reagents and instruments, the corresponding method should be applied.

NOTE: The pool may contain some cell debris, which could block a column membrane during purification leading to a long waiting time. To avoid this, it is recommended to perform a pre-cleaning of the RT pool by passing it through the Zymo column (see below) before mixing it with 7x DNA Binding buffer.

The procedure of cDNA pre-cleaning and purification using the column-based method

After the cDNA from each well is pooled in a reservoir, mix it with a 7x volume of DNA binding buffer (Zymo, D4004-1-L). We strongly recommend using a vacuum manifold for cDNA purification to avoid damage to the column membrane from multiple centrifugation rounds. A high-capacity Zymo-Spin IIICG column (Zymo, C1006-50-G) is required to purify large volumes resulting from 384 sample pooling.

Plate format	Pipetting strategy	Zymo-Spin column type	First strand cDNA		DNA binding buffer		TOTAL	
			Per well, µL	Per plate, mL	Per well, µL	Per plate, mL	Per well, µL	Per plate, mL
96WP	multichannel pipette or pipetting robot	I (#D4014)	20	**1.92**	140	**13.44**	160	**15.36**
384WP	pipetting robot	IIICG (C1006-50-G)	10	**3.84**	70	**26.88**	80	**30.72**

Table 1 Overview of the recommended pipetting strategy, plasticware, and reagent volumes to be used depending on the number of pooled samples

Preparation

- Make sure Zymo DNA Wash buffer has Ethanol added.

Procedure

2.2.1. According to Table 1, use a multichannel pipette, or pipetting robot, to transfer the entire RT volume (20 µL for **96WP**, 10 µL for **384WP**) of each sample into a specific reservoir (25 mL or 100 mL).

2.2.2. Mix the pool well and transfer it to a Falcon tube using a pipette.

2.2.3. For RT pool pre-cleaning, place a Zymo column in a new 2 mL tube, add 800 µL of the collected pool, and briefly centrifuge.

2.2.4. Collect the cleaned flow-through in a new Falcon tube. Repeat step 2.2.3 until all the pool passes through the Zymo column. Discard the column.

2.2.5. Using a pipette, measure the volume of the pool after cleaning, transfer it to a 50 mL Falcon tube, and add 7x DNA Binding buffer accordingly (see Table 1). The color of the mix should turn yellow.

2.2.6. Connect the 25 mL funnel (Zymo, C1039-25) to a Zymo column suitable for purification volume (Table 1) and place it on a vacuum manifold.

2.2.7. Gently mix the cDNA in the binding buffer mixture and transfer it to 25 ml funnel using a pipetboy.

2.2.8. Turn on the vacuum pump and let the liquid pass through the column.

2.2.9. Transfer any remaining volume to the funnel. Do not let the membrane overdry.

2.2.10. After the entire pool mix has passed through the column, add 200 µL of DNA Wash buffer (with Ethanol added) directly to the membrane of the column.

2.2.11. Repeat step 2.2.10 once the wash buffer passed through the filter.

2.2.12. Remove the column from the vacuum manifold, place it in a 1.5 mL tube, and centrifuge for 1 min to remove any remaining wash buffer.

2.2.13. Depending on the Zymo-Spin column type used, perform the following:

- For the type **I** column used with ≤96 samples (**96WP**), add 20 µL of water to the column matrix and incubate for 1 min.
- For the type **IIICG** column used with 384 samples (**384WP**), add 38 µL of water to the column matrix and incubate for 1 min.

2.2.14. Transfer the column into a new labeled 1.5 mL tube and centrifuge for 30 sec.

2.2.15. Immediately proceed to step 0.

The procedure of cDNA purification using the SPRI bead-based method

Perform cDNA purification using SPRI magnetic beads with a 1:1 ratio of cDNA pool to beads slurry. The purification of large volumes (i.e., 2 mL from **96WP** and 4 mL from **384WP**) requires three to six 1.5 mL tubes and a corresponding magnetic stand (Permagen, MSR06).

If the volume of the pool is higher than 750 µL, split it equally in the required number of 1.5 mL tubes and add the identical volume of beads (i.e., a pool of 1 mL split in 2 tubes with 500 µl per tube and add 500µL of beads per tube).

NOTE: Please use SPRI-beads only if the solution is clear and has no visible debris.

Preparation

Pre-warm beads at room temperature and vortex them vigorously before pipetting.

2.2.16. Pool the RT samples as described in Table 1.

2.2.17. Transfer the collected pool to a 2 mL or 15 mL tube, depending on the pooled volume. Consider that the final volume will be twice higher due to the addition of the beads.

2.2.18. Add pre-warmed beads in a 1:1 ratio (i.e., for 960 µL of pooled samples, add 960 µL of beads slurry), and mix by pipetting up and down ten times.

2.2.19. Incubate for 5 min at room temperature.

2.2.20. Place the tube on the magnetic stand, wait 5 min, and carefully remove and discard the supernatant.

2.2.21. To wash the beads, pipette 1 mL of freshly prepared 80% ethanol into the tube.

2.2.22. Incubate for 30 sec.

2.2.23. Carefully remove the ethanol without touching the bead pellet.

2.2.24. Repeat step 2.2.21 for a total of two washes.

2.2.25. Remove the tube from the magnetic stand and let the beads dry for 1-2 min.

2.2.26. Resuspend the beads in 37 µL of water and incubate for 1 min.

2.2.27. Place tubes on the magnetic stand, wait 5 min, and carefully transfer 35 µL of supernatant to a new tube to avoid bead carry-over.

2.2.28. Immediately proceed to step 2.3.

If the RT pool was split into several tubes at step 2.2.16, use one of the following options:

- **[Two tubes only]**, resuspend the beads in **both tubes** in 20 µL/tube, and combine both elutions in one tube;
- **[Two or more tubes]** resuspend the beads in the **first tube** in 40 µL of water. Keep other tubes closed to avoid over-drying of the beads. Transfer obtained elution to the next tube and resuspend beads. Repeat this step for every tube;
- **[Two or more tubes]**, resuspend **every** tube in 37 µL. Combine all elutions in one tube and perform one additional purification of the pool adding beads slurry accordingly to the pool volume (steps 2.2.18 - 2.2.27). Elute in 37 µL of water and collect 35 µL in a new tube.

2.3. Free primer digestion

It is recommended to perform non-incorporated primer digestion immediately after pooling.

Preparation

- Label 0.2 mL PCR tubes corresponding to the number of pools prepared.
- Thaw the **EXB** reagent at room temperature.
- Keep the **EXO** reagent on ice.
- Prepare program 2_FPD on the thermocycler (set the lid at 90°C):

Step	Temperature, °C	Time
Incubation	37	30 min
Incubation	80	20 min
Keep	4	pause

Procedure

2.3.1. Depending on the cDNA volume obtained from steps 2.2.15 or 2.2.27, transfer 17 µL or 35 µL of the eluate from each tube into a new labeled 0.2 mL PCR tube.

2.3.2. Prepare the EXO reaction mix as follows (with 10% excess):

Reagent	Per reaction, µL	
	For 17 µL elution	For 35 µL elution
EXB	2.2	4.4
EXO	1.1	1.1
TOTAL	3.3	5.5

2.3.3. According to the table, transfer **3 µL** or **5 µL** of EXO reaction mix into each PCR tube with purified cDNA.

2.3.4. Mix by pipetting up and down 5 times.

2.3.5. Briefly spin down in the bench-top centrifuge.

2.3.6. Incubate in the thermocycler Program 2_FPD.

2.3.7. Proceed immediately to step 2.4 or keep the tube at 4°C overnight.

Safe stop: After this step, the tube(s) can be kept at 4°C overnight.

2.4. Second-strand synthesis

At this step, double-stranded full-length cDNA is generated and purified using magnetic beads.

Preparation

- Pre-warm the SPRI beads at room temperature for ~30 min.
- Prepare 5 mL of 80% ethanol.
- Thaw the **SSB** reagent at room temperature and mix well before use.
- Keep the **SSE** reagent constantly on ice.
- Prepare program 3_SSS on the thermocycler (set the lid at 70°C):

Step	Temperature, °C	Time
Incubation	37	20 min
Incubation	65	30 min
Keep	4	pause

Procedure

2.4.1. Prepare the SSS reaction mix for the second strand synthesis as follows (with 10% excess):

Reagent	Per reaction, µL	
	For 17 µL elution	For 35 µL elution
SSB	5.5	7.7
SSE	2.2	3.3
TOTAL	7.7	11

2.4.2. Transfer **7 µL** or **10 µL** of the SSS reaction mix to the tube from step 2.3.7 and mix well by pipetting up and down 5 times.

2.4.3. Incubate in the thermocycler Program 3_SSS.

2.4.4. Proceed immediately to step 2.4.5.

cDNA clean-up with SPRI beads

Perform the double-stranded cDNA purification with SPRI magnetic beads using a 0.6x ratio (i.e., 30 µL of bead slurry plus 50 µL of cDNA):

NOTE: Use pre-warmed beads and vortex them vigorously before pipetting.

2.4.5. Complement the final volume to 50 µL with water.

2.4.6. Add 30 µL of beads and mix by pipetting up and down 10 times.

2.4.7. Incubate for 5 min at room temperature.

2.4.8. Place the tube on the magnetic stand, wait 5 min, and carefully remove and discard the supernatant.

2.4.9. To wash the beads, pipette 200 µL of freshly prepared 80% ethanol into the tube.

2.4.10. Incubate for 30 sec.

2.4.11. Carefully remove the ethanol without touching the bead pellet.

2.4.12. Repeat step 2.4.9 for a total of two washes.

2.4.13. Remove the tube from the magnetic stand and let the beads dry for 1-2 min.

2.4.14. Resuspend the beads in 21 µL of water.

2.4.15. Incubate for 1 min.

2.4.16. Place the tube on the magnetic stand, wait 5 min, and carefully transfer 20 µL of supernatant into a new tube to avoid bead carry-over.

2.4.17. Use 2 µL to measure the concentration with Qubit.

Safe stop: At this step, the cDNA can be safely kept at -20°C for a few weeks.

2.5. Tagmentation

At this step, the full-length cDNA is tagmented using a Tn5 transposase pre-loaded with adapters for library amplification. It is recommended to use 20 ng of cDNA for tagmentation to obtain a higher library complexity with less PCR amplification.

Preparation

- Pre-warm the SPRI beads at room temperature for ~30 min.
- If needed, prepare fresh 5 mL of 80% ethanol.
- Thaw the **TAB** reagent at room temperature and mix well before use.
- Keep the **TE** reagent constantly on ice.
- Set the PCR machine at 55°C incubation.

Procedure

2.5.1. Prepare the Tagmentation mix on ice in a PCR tube. If several cDNA samples are tagmented, prepare a Master mix as follows:

Reagent	Per reaction, µL, for cDNA inputs		
	≤ 9 ng	10-14 ng	15-20 ng
TAB	4	4	4
TE	1	2	3
cDNA	X	X	X
Water	15.0 – X	14.0 – X	13.0 – X
TOTAL	20	20	20

2.5.2. Keep the mix on ice and pipette up and down 10 times with the pipette set at 5 µL. Pay attention to thoroughly mixing the reaction volume.

2.5.3. Incubate at 55°C for 7 min in the PCR machine.

2.5.4. Immediately place the tube on ice and add 30 µL of water to achieve a final volume of 50 µL.

Tagmented cDNA clean-up with SPRI beads

Purify the tagmented cDNA with SPRI magnetic beads using a 0.6x ratio.

NOTE: Use pre-warmed beads and vortex them vigorously before pipetting.

2.5.5. Add 30 µL of beads to 50 µL of tagmented cDNA and mix by pipetting up and down 10 times.

2.5.6. Incubate for 5 min at room temperature.

2.5.7. Place the tubes on the magnetic stand, wait 5 min, and carefully remove and discard the supernatant.

2.5.8. To wash the beads, pipette 200 µL of freshly prepared 80% ethanol into the tube.

2.5.9. Incubate for 30 sec.

2.5.10. Carefully remove the ethanol without touching the bead pellet.

2.5.11. Repeat steps 2.5.8 - 2.5.10 for a total of two washes.

2.5.12. Remove the tube from the magnetic stand and let the beads dry for 1-2 minutes.

2.5.13. Resuspend the beads in 21 µL of water.

2.5.14. Incubate for 1 min.

2.5.15. Place the tube on the magnetic stand, wait 5 min, and carefully transfer 20 µL of the supernatant into a new tube to avoid bead carry-over.

2.5.16. Proceed immediately to step 2.6.

2.6. Library indexing and amplification

At this step, 5' terminal fragments are amplified using the Unique Dual Indexing (UDI) adapter primers. The kit contains 4 Illumina-compatible primer pairs to generate the UDI libraries. The index sequences are indicated in Table 4.

The number of amplification cycles required for library preparation typically ranges from 12 to 17. The precise number may depend on the RNA samples and the input cDNA amount used for tagmentation. To determine the optimal number of cycles, follow the Library quantification protocol below.

It is strongly recommended to perform the final library clean-up step twice to effectively remove primer dimer fragments.

Preparation

- Pre-warm the SPRI beads at room temperature for ~30 mins.
- Prepare 10 mL of 80% ethanol.
- Thaw the **LAM** reagent on ice and mix well before use.
- Thaw the required number of **MQ.UDI Adapters** at room temperature and briefly spin before use.
- Prepare the Program 4_TN5AMP (set the lid at 100°C) on the thermocycler (*The exact number of PCR cycles should be determined following the Library quantification protocol below)

Step	Temperature, °C	Time	Cycles
Initial denaturation	98	1 min	
Denaturation	98	10 sec	
Annealing	63	30 sec	15 or TBD*
Extension	72	1 min	
Final extension	72	3 min	
Keep	4	pause	

Procedure

2.6.1. Prepare the PCR amplification reaction as follows:

Reagent	Per reaction, µL
LAM	25
MQ.UDI Adapter	5
Tagmented cDNA	20
TOTAL	50

2.6.2. Pipette up and down 5 times.

2.6.3. Put the tube in the PCR machine and do one of the following:

- Determine the optimal number of amplification cycles using real-time PCR (highly recommended; follow steps 2.6.5 - 2.6.12); or
- start Program 4_TN5AMP with the default 15 PCR cycles (less preferable).

2.6.4. After the PCR is done, proceed to the library bead clean-up (2.6.13).

2.6.5. In the case of the cycle number optimization, perform 5 cycles of library preamplification (Program 4_TN5AMP).

2.6.6. Place the tube on ice.

2.6.7. Use a 5 µL aliquot from the PCR reaction to prepare a qPCR reaction mix in the appropriate PCR tube or plate as follows:

Reagent	Per reaction, µL
5 cycles pre-amplified library	5
SYBR 100x*	0.1
LAM	2.5
Water	2.4
TOTAL	10

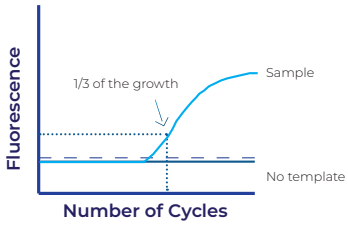
*Prepare 100x dilution with nuclease-free water from 10'000x stock

2.6.8. Place the tube in the qPCR machine and run with the following program:

Step	Temperature	Time	Cycles
Initial denaturation	98°C	1 min	
Denaturation	98°C	10 sec	
Annealing	63°C	30 sec	25
Extension + acquisition	72°C	1 min	

2.6.9. Determine the cycle number depending on the growth curve in the multicomponent plot, as in Figure 2.

Figure 2 Determination of the additional number of amplification cycles with qPCR



2.6.10. Place the tube containing the remaining 45 µL of the pre-amplified library in the PCR machine.

2.6.11. Set the number of amplification cycles determined in step 2.6.9.

2.6.12. Typically, 10 ng of cDNA is sufficient to obtain 20-40 ng of DNA library after 13-14 cycles of amplification. Table 2 shows the approximate amount of amplification cycles and the expected library yield.

Tagmented cDNA, ng	# PCR cycles	Expected library yield, ng
5	14-15	
10	13-14	20-40
20	11-12	

Table 2 Expected yield of DRUG-seq libraries from different amounts of tagmented cDNA

Indexed cDNA library clean-up with SPRI beads

Purify the final cDNA library with SPRI magnetic beads using a 0.7x ratio (35 µL of bead slurry for 50 µL cDNA library).

NOTE: Use pre-warmed beads and vortex them vigorously before pipetting.

2.6.13. Adjust the library volume to 50 µL with water.

2.6.14. Add 35 µL of beads and mix by pipetting up and down 10 times.

2.6.15. Incubate for 5 min at room temperature.

2.6.16. Place the tubes on the magnetic stand, wait 5 min, carefully remove, and discard the supernatant.

2.6.17. To wash the beads, pipette 200 µL of freshly prepared 80% ethanol into the tube.

2.6.18. Incubate for 30 sec.

2.6.19. Carefully remove the ethanol without touching the bead pellet.

2.6.20. Repeat step 2.6.17 for a total of two washes.

2.6.21. Remove tubes from the magnetic stand and let the beads dry for 1-2 min.

2.6.22. Resuspend the beads in 21 µL of water.

2.6.23. Incubate for 1 min.

2.6.24. Place tubes on the magnetic stand, wait 5 minutes, and carefully remove 20 µL of supernatant into a new tube to avoid bead carry-over.

2.6.25. Perform the bead clean-up once again by repeating the procedure from step 2.6.13.

Safe stop: At this step, the cDNA libraries can be safely kept at -20°C for a few weeks.

2.7. Library quality control

Pooled library quality control

Before sequencing, the libraries should be subjected to fragment analysis (with Fragment analyzer, Bioanalyzer, or TapeStation) and quantification (with Qubit). This information is required to assess the molarity of the libraries and prepare the appropriate library dilution for sequencing. A successful library contains fragments in the range of 300 – 1000 bp with a peak at 400-700 bp; see Figure 3 for an example of a standard DRUG-seq library profile.

Importantly, libraries with primer dimer peaks at 180 bp and ranging at 250 – 290 bp will likely produce lower-quality sequencing data with a reduced proportion of mappable reads (Figure 4). Therefore, it is strongly recommended to remove those peaks by performing an additional round of SPRI beads purification with the 0.7x ratio (see step 2.6.13).

Undertagmented libraries have a broader fragment range distribution with a peak at 1000-1200 bp (Figure 5). Only a fraction of such libraries contains fragments that can be efficiently sequenced; therefore, it is recommended to re-tagment the cDNA for the best results.

Library quantification can also be done unbiasedly by qPCR using standard Illumina library quantification kits (i.e., KAPA HiFi, Roche).

Pre-sequencing library QC:

- Use 2 µL of the library to measure the concentration with Qubit.
- Use 2 µL of the library to assess the profile with the Fragment Analyzer instrument or similar.
- If necessary, re-purify the libraries by following steps 2.6.13 – 2.6.24 to remove the peaks <300 bp.

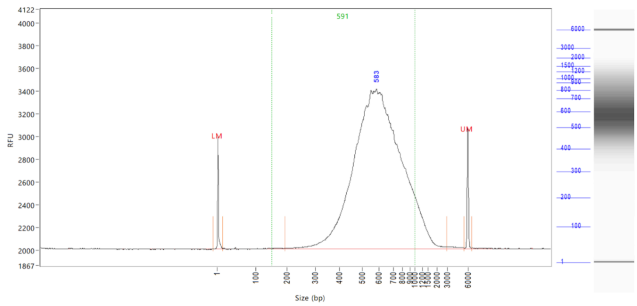


Figure 3 A successful library profile with fragments between 300-1000 bp

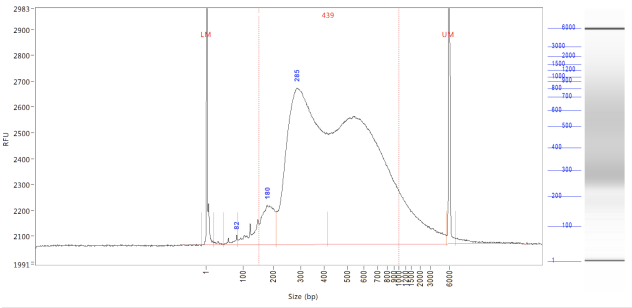


Figure 4 An example of an over-tagmented library profile with a peak at 290 bp and an adapter peak at 160 bp

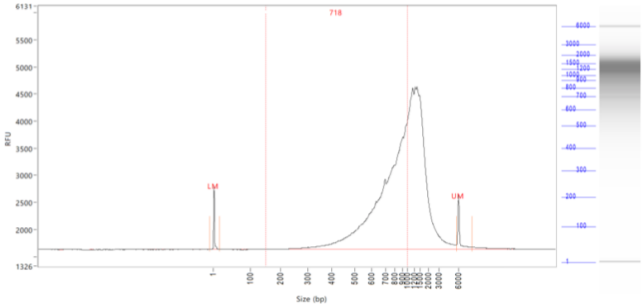


Figure 5 An example of an under-tagmented library profile with a major peak > 1000 bp

Assessing uniformity of read distribution across the samples

For projects involving highly heterogeneous RNA samples, it is recommended to validate the uniformity of read coverage across the samples by shallow library sequencing (see step 2.2). This approach ensures that every sample will obtain enough reads required for the analysis. DRUG-seq libraries can be added as spike-ins to the compatible sequencing run (see Part 3). For this validation 0.5-1M sequencing reads per library is sufficient to assess the fraction of reads attributed to every sample.

Part 3. LIBRARY SEQUENCING

The libraries prepared with the MERCURIUS™ DRUG-seq kit carry Illumina- and AVITI-compatible adapter sequences. They can be processed on any Illumina instrument (e.g., HiSeq, NextSeq, MiSeq, iSeq, and NovaSeq) or in the Element AVITI System with Adept Workflow.

The MERCURIUS™ DRUG-seq libraries are Unique Dual-Indexed and can potentially be pooled in a sequencing run with other libraries if the sequencing structure is compatible. Please refer to Table 3 for the optimal sequencing structure and Table 4 for the list of i5 and i7 index sequences.

Given the DRUG-seq library structure, the optimal number of cycles for Read 1 is 28 (and 29 for AVITI). The following cycles, 29-60, will cover the homopolymer sequence, which may result in a significant drop of Q30 values reflecting sequencing quality. Standard paired-end run setups on Illumina platforms (e.g., 100 PE or 150 PE) are not suitable due to the low performance of the sequencing machine for homopolymer sequences.

However, on the AVITI platform, a custom setup with Read1 at 200 bp would be sufficient to read through the oligo-dT sequence and into the cDNA, and Read2 at 100 bp is recommended and compatible.

Read	Length (cycles)		Comment
	for Illumina	for AVITI	
Read 1	28	29	Sample barcode (14 nt) and UMI (14 nt); +1 extra base for AVITI
Index 1 (i7) read	8	8	Library Index
Index 2 (i5) read	8*	8*	Library Index (*optional and valid for UDI libraries)
Read 2	60-90	101	Gene fragment

Table 3 Sequencing structure of DRUG-seq libraries

The Unique Dual Indexing (UDI) strategy ensures the highest library sequencing and demultiplexing accuracy and complies with the best practices for Illumina sequencing platforms. UD-indexed libraries have distinct index adapters for i7 and i5 index reads (Table 4).

Name	Type	i7 index sequence	i5 index sequence Forward Workflow	i5 index sequence Reverse Workflow
MQ.UDI.1	UDI (i7/i5)	TAAGGCGA	TATAGCCT	AGGCTATA
MQ.UDI.2	UDI (i7/i5)	CGTACTAG	ATAGAGGC	GCCTCTAT
MQ.UDI.3	UDI (i7/i5)	AGGCAGAA	CCTATCCT	AGGATAGG
MQ.UDI.4	UDI (i7/i5)	GCGTTGGA	TTGGACTT	AAGTCCAA

Table 4 UDI adapter sequences

NOTE: Sequencing depth

1. The recommended sequencing depth is 1-5 Mio reads per sample. Deeper sequencing can also be performed to detect very lowly expressed genes.

2. If only one library is sequenced in a flow cell, the Index reads can be skipped.

3. The loading molarity for the library depends on the type of sequencing instrument (see 3.1 and 3.2) and should be discussed with the sequencing facility or an experienced person.

3.1. Sequencing on the Illumina instruments

The loading concentration for the Illumina instruments is indicated in Table 5. Please refer to Appendix 2 for the list of Illumina instruments with a forward or reverse workflow.

Instrument	Final loading concentration	PhiX
MiSeq	20 pM	1 %
iSeq	100 pM	1 %
NextSeq 500/550/550Dx	2.2 pM	1 %
NextSeq 2000, manual denature	85 pM	1%
NextSeq 2000, onboard denature	850 pM	1%
NovaSeq Standard Workflow*	160 pM	1 %
NovaSeq XP Workflow	100 pM	1 %
HiSeq4000	270 pM	1 %

* - adjusted molarity for DRUG-libraries sequencing. We recommend a prior dilution of the libraries to 0.8 nM before denaturation

Table 5 Reference loading concentrations for various Illumina instruments

3.2. Sequencing on the Element AVITI instrument

For the most optimal results, the MERCURIUS™ DRUG-seq libraries can be sequenced with the Element Biosciences AVITI System using Cloudbreak AVITI 2x75 High Output sequencing kits (#860-00004).

Libraries must be converted with the Adept PCR-Plus module (#830-00018) for linear loading (Table 6).

Type	Loading molarity, pM	Library starting amount for denaturation, nM	PhiX control	PhiX, %
Cloudbreak	14	1*	PhiX Control Library, Adept	2 %

* - requires 4nM of library before conversion

Table 6 Loading concentration for Cloudbreak AVITI 2x75 High Output sequencing kit

NOTE: Sequencing depth

Please note that the **Cloudbreak AVITI 2x75 High Output sequencing** yields 1’000 Mio reads. Therefore, for the 384-sample library, the sequence depths would be limited to a maximum of 2.5 Mio reads/sample.

Part 4. SEQUENCING DATA PROCESSING

Following Illumina sequencing and standard library index demultiplexing, the user obtains raw read1 and read2 *fastq* sequencing files (e.g., *mylibrary_R1.fastq.gz* and *mylibrary_R2.fastq.gz*).

This section explains how to generate ready-for-analysis gene and UMI read count matrices from raw *fastq* files.

This section explains how to generate ready-for-analysis gene, and UMI read count matrices from raw *fastq* files.

To obtain the data ready for analysis, the user needs to align the sequencing reads to the genome and perform the gene/UMI read count generation, which can be done in parallel with the sample demultiplexing.

For manual data processing, the user requires a terminal and a server or powerful computer with an installed set of standard bioinformatic tools.

4.1. Required software

Tool	Description	Version
fastQC	Software for QC of *fastq* or *bam* files. This software is used to assess the quality of the sequencing reads, such as the number of duplicates, adapter contamination, repetitive sequence contamination, and GC content. The software is freely available from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ The website also contains informative examples of good and poor-quality data.	>0.11.9
STARsolo from STAR	Software for read alignment on reference genome (Dobin et al., 2013). It can be downloaded from Github (https://github.com/alexdobin/STAR)	2.7.9
FastReadCounter	Software for counting genome-aligned reads for genomic features	>1.1
Picard	Collections of command-line utilities to manipulate with BAM files. Used in this user guide for demultiplexing of BAM files.Java version 8 or higher	>2.17.8
Samtools	Collections of command-line utilities to manipulate with BAM files. Used in this user guide for sorting and indexing of BAM files.	>1.9
fqtk	Demultiplexes pooled FASTQs based on inline barcodes	0.3.1
kallisto	Pseudoaligns reads to a transcriptome for quantification	0.48.0
R	Data processing and visualization. Required packages installed: data.table and Matrix. Please refer to the official R documentation for instructions on installing R packages: https://cran.r-project.org/doc/manuals/r-patched/R-admin.html#Installing-packages	>3.0.0

IMPORTANT: Please refer to the official webpages of each software tool mentioned to review system requirements before installation.

4.2. Data processing

4.2.1. Merging *fastq* files from individual lanes and/or libraries (Optional)

4.2.1.1 Depending on the type of instrument used for sequencing, one or multiple R1/R2 *fastq* files per library may result from individual lanes of a flow cell. The *fastq* files from individual lanes should be merged into single *R1.fastq* and single *R2.fastq* files to simplify the following steps. This is an example of *fastq* files obtained from HiSeq 4 lane sequencing:

```
>  mylibrary_L001_R1.fastq.gz, mylibrary_L002_R1.fastq.gz,
   mylibrary_L003_R1.fastq.gz, mylibrary_L004_R1.fastq.gz
>  mylibrary_L001_R2.fastq.gz, mylibrary_L002_R2.fastq.gz,
   mylibrary_L003_R2.fastq.gz, mylibrary_L004_R2.fastq.gz
```

4.2.1.2 To merge the *fastq* files from different lanes use a `cat` command in a terminal. This will generate two files: *mylibrary_R1.fastq.gz* and *mylibrary_R2.fastq.gz*, containing the information of the entire library.

```
> cat mylibrary_L001_R1.fastq.gz mylibrary_L002_R1.fastq.gz
  mylibrary_L003_R1.fastq.gz mylibrary_L004_R1.fastq.gz >
  mylibrary_R1.fastq.gz
> cat mylibrary_L001_R2.fastq.gz mylibrary_L002_R2.fastq.gz
  mylibrary_L003_R2.fastq.gz mylibrary_L004_R2.fastq.gz >
  mylibrary_R2.fastq.gz
```

4.2.1.3 Move these 2 *fastq* files into a new folder, which will be referenced in this manual as $fastqfolder.

NOTE: This step can also be done if you sequenced your library in multiple sequencing runs.

Warning: The order of merging files should be kept the same (e.g., L001, L002, L003, L004, not L002, L001 ...) to avoid issues when demultiplexing the samples.

4.2.2. Sequencing data quality check

4.2.2.1 Run fastQC on both R1 and R2 fastq files. Use –outdir option to indicate the path to the output directory. This directory will contain HTML reports produced by the software.

```
> fastqc --outdir $QCdir/ mylibrary_R1.fastq.gz
> fastqc --outdir $QCdir/ mylibrary_R2.fastq.gz
```

4.2.2.2 Check fastQC reports to assess the quality of the samples (see Software and materials).

NOTE:

- The report for the R1 *fastq* file may contain some "red flags" because it contains barcodes/UMIs. Still, it can provide useful information on the sequencing quality of the barcodes/UMIs.
- The main point of this step is to check the R2 *fastq* report. Of note, *per base sequence content* and *kmer content* are rarely green. If some *adapter contamination* or *overrepresented sequence* is detected in the data, it may not be an issue (if the effect is limited to <10~20%). These are lost reads, but most of them will be filtered out during the next step.

4.2.3. Preparing the reference genome

The *fastq* files must be aligned (or “mapped”) on a reference genome. The STAR (Dobin et al., 2013[1]) aligner is one of the most efficient tools for RNA-seq reads mapping. It contains a “soft-clipping” tool that automatically cuts the beginning or the end of reads to improve the mapping efficiency, thus allowing the user to skip the step of trimming the reads for adapter contamination. Moreover, STAR has a mode called STARsolo, designed to align multiplexed data (such as BRB-seq) and directly generate count matrices.

The STAR aligner requires a genome assembly together with a genome index file. The index file generation is a time-consuming process that is only performed once on a given genome assembly so that it can be completed in advance and the index files can be stored on the server for subsequent analyses.

4.2.3.1 Download the correct genome assembly fasta file (e.g. Homo_sapiens.GRCh38.dna.primary_assembly.fa) and gene annotation file in gtf format (e.g. Homo_sapiens.GRCh38.108.gtf) from Ensembl or UCSC repository. Below is an example for a human assembly:

```
> wget https://ftp.ensembl.org/pub/release-
  108/fasta/homo_sapiens/dna/Homo_sapiens.GRCh38.dna.primary_assembly.fa.gz
> gzip -d Homo_sapiens.GRCh38.dna.primary_assembly.fa.gz # unzip
> wget https://ftp.ensembl.org/pub/release-
  108/gtf/homo_sapiens/Homo_sapiens.GRCh38.108.gtf.gz
> gzip -d Homo_sapiens.GRCh38.108.gtf.gz # unzip
```

[1] Alexander Dobin, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, Thomas R. Gingeras, STAR: ultrafast universal RNA-seq aligner, *Bioinformatics*, Volume 29, Issue 1, January 2013, Pages 15–21, https://doi.org/10.1093/bioinformatics/bts635

NOTE: It's recommended to download the primary_assembly fasta file when possible (without the 'sm' or 'rm' tags). If not available, download the top_level assembly. For the *gtf*, download the one that does not have the 'chr' or 'abinitio' tags.

4.2.3.2 Use STAR to create an index for the genome assembly. Indicate the output folder name containing the index files using *--genomeDir* option:

```
> STAR --runMode genomeGenerate --genomeDir /path/to/genomeDir --
  genomeFastaFiles Homo_sapiens.GRCh38.dna.primary_assembly.fa --sjdbGTFfile
  Homo_sapiens.GRCh38.108.gtf --runThreadN 8
```

NOTE:

- The --runThreadN parameter can be modified depending on the number of cores available on your machine. The larger this number is, the more parallelized/fast the indexing will be.
- Depending on the genome assembly, STAR can use up to 32-40Gb of RAM. So, you should use a machine that has this RAM capacity.

4.2.4. Aligning to the reference genome and generation of count matrices

After the genome index is created, both R1 and R2 *fastq* files can be aligned to this reference genome. For this step, use the "solo" mode of STAR, which not only aligns the reads to the reference genome but also creates the gene read count and UMI (unique molecular identifier) count matrices.

The following parameters should be adjusted according to the sequencing information:

- *--soloCBwhitelist*: a text file with the list of barcodes (one barcode sequence per lane) which is used by STAR for demultiplexing. This file is provided according to version of the MERCURIUS kit used. Example of "barcodes_96_V5D_star.txt":

```
> TACGTTATTCCGAA
> AACAGGATAACTCC
> ACTCAGGCACCTCC
> ACGAGCAGATGCAG
```

- *--soloCBstart*: Start position of the barcode in the R1 fastq file, equal to 1.
- *--soloCBlen*: Length of the barcode. This value should match the length of the barcode sequence in the file specified by –soloCBwhitelist. The length of the barcode depends on the version of the oligo-dT barcodes provided in the kit. For the barcode plate set V5, the default value is 14.
- *--soloUMIstart*: Start position of the UMI, it's *soloCBlen* + 1 since the UMI starts right after the barcode sequence.
- *--soloUMIlen*: The length of UMI. This parameter depends on the version of the oligo-dT barcodes in the kit and the number of sequencing cycles performed for Read1. For the barcode plate set V5 the default value is 14.
- --readFilesIn: name and path to the input *fastq* files.

The order of the fastq files provided in the script is important. The first *fastq* must contain genomic information, while the second the barcode and UMI content. Thus, files should be provided for STARsolo in the following order: --readFilesIn mylibray_R2 mylibrary_R1.

- --genomeDir: a path to the genome indices directory generated before ($genomeDir).

Output count matrix parameters:

By default, STARsolo produces a UMI count matrix, i.e., containing unique non-duplicated reads per sample for each gene. This type of count data is a standard for single-cell RNA-seq analysis. For bulk RNA-seq analysis, a gene read count matrix is usually used. The following parameters will enable the generation of the output of interest.

--soloUMIdedup **NoDedup**, will generate a read count matrix output

--soloUMIdedup **NoDedup 1MM_Directional,** will generate both UMI and read count matrices in mtx format.

This step will output *bam* files and count matrices in the folder $bamdir.

```
> STAR --runMode alignReads --outSAMmapqUnique 60 --runThreadN 8 --
  outSAMunmapped Within --soloStrand Forward --quantMode GeneCounts --
  outBAMsortingThreadN 8 --genomeDir $genomeDir --soloType CB_UMI_Simple --
  soloCBstart 1 --soloCBlen 14 --soloUMIstart 15 --soloUMIlen 14 --
  soloUMIdedup NoDedup 1MM_Directional --soloCellFilter None --soloCBwhitelist
  barcodes.txt --soloBarcodeReadLength 0 --soloFeatures Gene --
```

```
outSAMattributes NH HI nM AS CR UR CB UB GX GN sS sQ sM --
outFilterMultimapNmax 1 --readFilesCommand zcat --outSAMtype BAM
SortedByCoordinate --outFileNamePrefix $bamdir --readFilesIn
mylibrary_R2.fastq.gz mylibrary_R1.fastq.gz
```

The demultiplexing statistics can be found in the "*bamdir/Solo.out/Barcodes.stats*" file.

The alignment quality and performance metrics can be found in the "*bamdir/Log.final.out*" file.

NOTE: The most important statistic at this step is the proportion of "Uniquely mapped reads" which is expected to be greater than 70% (for human, mouse or drosophila).

4.2.5. Generating the count matrix from .mtx file

STARsolo will generate a count matrix (*matrix.mtx* file) located in the *bamdir/Solo.out/Gene/raw* folder. This file is a sparse matrix format that can be transformed into a standard count matrix using an R script provided below:

```
> #Myscript.R
> library(data.table)
> library(Matrix)
> matrix_dir <- "$bamdir/Solo.out/Gene/raw"
> f <- file(paste0(matrix_dir, "matrix.mtx"), "r")
> mat <- as.data.frame(as.matrix(readMM(f)))
> close(f)
> feature.names = fread(paste0(matrix_dir, "features.tsv"), header = FALSE,
  stringsAsFactors = FALSE, data.table = F)
> barcode.names = fread(paste0(matrix_dir, "barcodes.tsv"), header = FALSE,
  stringsAsFactors = FALSE, data.table = F)
> colnames(mat) <- barcode.names$V1
> rownames(mat) <- feature.names$V1
> fwrite(mat, file = umi.counts.txt, sep = "\t", quote = F, row.names = T,
  col.names = T)
```

The resulting UMI/gene count matrix can be used for a standard expression analysis following conventional bioinformatic tools.

4.2.6. Generating the read count matrix with per-sample stats (Optional)

Given a multiplex BAM file obtained with STARsolo and a set of barcodes, the software FastReadCounter produces a read count matrix with per-sample statistics with the following code:

```
> #!/bin/bash
>
> gtf_file=homo_sapience.gtf                ### GTF genome annotation file
> output_folder=counts/              ### Name of the final count output file
> bam_dir=mypath/bam_demult          ### Directory with demultiplexed bam
  files
> barcode_file=V5D_96_frc.txt        ### Barcode reference file
>
> FastReadCounter-1.0.jar" --bam ${bam_dir}/${bam_dir}.bam \
>                          --gtf "${gtf_file}" \
>                          --umi-dedup none \
>                          --barcodeFile \${barcode_file} \
>                          -o ${output_folder}
```

The resulting read count matrices can be used for subsequent gene expression analysis using established pipelines and tools.

NOTE: Please contact us at info@alitheagenomics.com in case you don't have the barcode sequences (in your email, please indicate the name of the barcode set and the PN of the barcode module)

4.2.7. Demultiplexing bam files (Optional)

Generation of demultiplexed bam files, i.e., individual bam files for each sample, might be needed in some cases, for example, for submitting the raw data to an online repository that does not accept multiplexed

data (for example, GEO or ArrayExpress), or for running an established bulk RNA-seq data analysis pipeline.

For this purpose, the Picard tool can be used with the following parameters:

- $out_dir, The output directory for demultiplexed bam files
- $path_to_bam, the path to multiplexed single bam file
- $barcode_brb.txt, the tab-delimited file containing 2 columns: sample_id and barcode seq. Example of barcode_96_V5D_brb.txt:

```
>  Sample1      TACGTTATTCCGAA
>  Sample2      AACAGGATAACTCC
>  Sample3      ACTCAGGCACCTCC
>  Sample4      ACGAGCAGATGCAG
```

NOTE: This file differs from the list of barcode files provided to STAR.

Run the following Picard script

```
>  #!/bin/bash
>  demultiplexed_bam_out_dir=$out_dir
>  input_bam=$path_to_bam
>  barcode_info=$barcode_brb.txt
>
>  while IFS=$'\t' read -r -a line
>  do
>    sample_id="${line[0]}"
>    tag_value="${line[1]}"
>
>  java -jar /path/to/picard.jar FilterSamReads  \
>          I=${input_bam} \
>          O=${demultiplexed_bam_out_dir}/${sample_id}.bam \
>          TAG=CR TAG_VALUE=${tag_value}  \
>          FILTER=includeTagValues
>  done < "$barcode_info"
```

NOTE: Please contact us at info@alitheagenomics.com in case you don't have the barcode sequences (in your email, please indicate the name of the barcode set and the PN of the barcode module).

Appendix 1. ERCC Spike-In Control

The current protocol includes the addition of External RNA Controls Consortium (ERCC) Spike-Ins to the lysis buffer.

Prepare a 1:100 dilution of the ERCC RNA Spike-In mix in nuclease-free water. Mix 990 µL of pre-chilled water with 10 µL of ERCC. Pipette well and aliquot the dilution into 50 µL aliquots, keeping them at -20°C.

The working solution of 1x Spheroid Lysis Buffer with ERCC Spike-In controls consists of the following (with 105 excess):

Reagent	96WP, µL		384WP, µL	
	Per well	96 wells	Per well	384 wells
SDR	3	330	1	430
SLB	10.2	1122	3.4	1462
SRI	2.4	264	0.8	344
Water	14.1	1551	4.7	2021
ERCC* (1:100)	0.3	33	0.1	43
TOTAL	30	3300	10	4300

**The final ERCC is 1:1000 in a 384-type well (equal to 50 ng of RNA/well) and 1:250 in a 96-type well (150-200 ng of RNA/well)*

1x Spheroid Lysis Buffer (SLB) preparation with ERCC

1. Thaw the SDR, SLB, and ERCC tubes on ice and avoid storing them for long-term use.
2. Keep the nuclease-free water on ice to maintain a cold temperature.
3. Spin down all the tubes before pipetting.
4. First, add the water to a 15-mL Falcon tube, then the SDR, SLB, SRI, and the ERCC (in this particular order).
5. Pipette the prepared mix a few times and briefly spin the tube. Keep it on ice until further use.
6. Follow the main protocol for the Spheroid lysis procedure (step 1.4.1)

Appendix 2. Compatible Illumina instruments

Illumina instruments can use two workflows for sequencing the i5 index (see the details in Indexed Sequencing Overview Guide on Illumina's website).

Forward strand workflow instruments:

- NovaSeq 6000 with v1.0 reagents
- MiSeq with Rapid reagents
- HiSeq 2500, HiSeq 2000

Reverse strand workflow instruments:

- NovaSeq 6000 with v1.5 reagents
- iSeq 100
- MiniSeq with Standard reagents
- NextSeq
- HiSeq X, HiSeq 4000, HiSeq 3000

ALITHEA GENOMICS

info@alitheagenomics.com

www.alitheagenomics.com

Alithea Genomics SA
Batiment Serine
Rue de la Corniche 8, level -2
1066 Epalinges
Switzerland
Tel: +41 78 830 3139

Alithea Genomics Inc.
321 Ballenger Center Drive
Suite 125 C/O Alithea Genomics Inc.
Frederick, MD 21703
USA
Tel: +1 240 656 3142