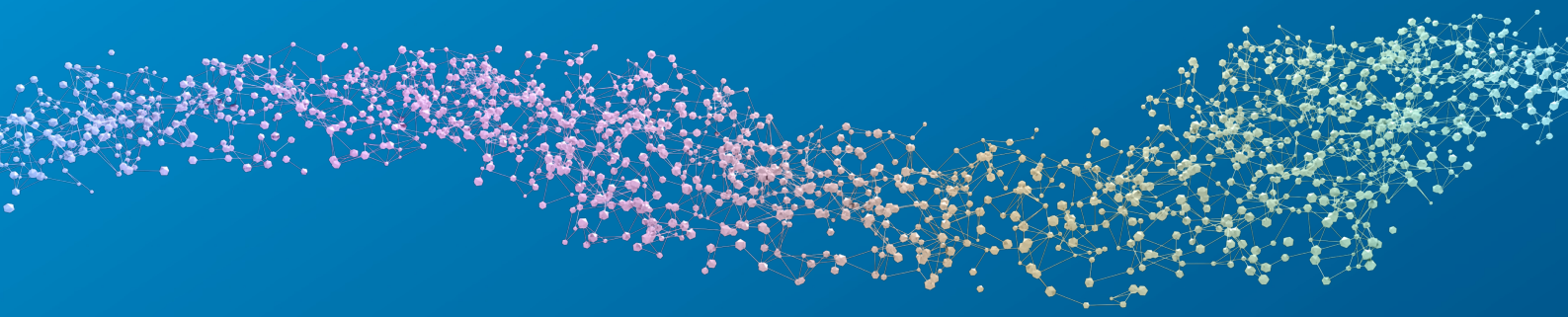
ALITHEA GENOMICS



MERCURIUS™

Multiplexed Plant RNA-seq Library Preparation Kit (mRNA) for 96 and 384 Samples

PN 10715, 11613

User Guide

October 2025

Related Products

Kit name	Kit PN	Modules	Module PN
Mercurius™ Multiplexed Plant RNA-seq Library Preparation (mRNA) 96 kit	10715	Barcoded Oligo-dT Adapters Module 96 samples	10513
		Multiplexed Plant RNA-seq Library Preparation and UDI Module 96 samples	10661
Mercurius™ Multiplexed Plant RNA-seq Library Preparation (mRNA) 4x 96 kit	11613	Barcoded Oligo-dT Adapters Module 4x 96 samples	10513
		Multiplexed Plant RNA-seq Library Preparation and UDI Module 4x 96 samples	10667

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 10715)	4x 96 samples (PN 11613)	
Plate with 96 barcoded oligo-dT primers, set V5C (PN 10400)	96 V5C OdT	1 plate	4 plates	-20°C
Aluminium Seal	-	2 pcs	8 pcs	-20°C/RT

Multiplexed Plant RNA-seq Library Preparation and UDI Module (mRNA)

Component Name	Label	Cap colour	Volume, µL		Storage
			96 samples (PN 10661)	4x 96 samples (PN 10667)	
First Strand Buffer FL	FSB FL	magenta	245	4x 245	-20°C
Repair/RT Enzyme	RAE	magenta	140	4x 140	-20°C
Repair/RT Mix	RAM	magenta	930	4x 930	-20°C
Exonuclease I Enzyme	F-EXO	purple	10	10	-20°C
Exonuclease Buffer	F-EXB	purple	20	20	-20°C
Second Strand Enzyme FL	SSE FL	orange	25	25	-20°C
Second Strand Buffer FL	SSB FL	orange	30	30	-20°C
Adapter Ligation Buffer	ALB	blue	100	100	-20°C
BRB-compatible Adapter	BRB.AD	blue	10	10	-20°C
Library Amplification Mix FL	LAM FL	green	200	200	-20°C
UDI Adapter Mix 1	MF.UDI.1	transparent	10	10	-20°C
UDI Adapter Mix 2	MF.UDI.2	transparent	10	10	-20°C
UDI Adapter Mix 3	MF.UDI.3	transparent	10	10	-20°C
UDI Adapter Mix 4	MF.UDI.4	transparent	10	10	-20°C

Additional required reagents and equipment (supplied by the user)

Plasticware	Manufacturer	Product number
0.2 mL 8-Strip non-flex PCR tubes	Starlab	I1402-3700
Disposable pipetting reservoir 25 mL polystyrene	Integra	4382
Disposable pipetting reservoir 150 mL polystyrene	Integra	6318

Reagents	Manufacturer	Product number
NEBNext Poly(A) mRNA Magnetic Isolation Module, 96 rxns	NEB	E7490L
SPRI AMPure Beads	Beckman Coulter	A63881
or		
CleanNA Beads	CleanNA	CNGS0050D
Qubit™ dsDNA HS Assay Kit	Invitrogen	Q32851
High Sensitivity NGS Fragment Analysis Kit	Agilent	DNF-474
Ethanol, 200 proof	-	-
Nuclease-free water	Thermo Fisher	A57775

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
Magnetic stand for 5 mL tubes	Permagen	
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
VIAFLO instrument (optional)	Integra	6001
VIAFLO 96 channel pipetting head, 0.5-12.5 µL (optional)	Integra	6101

Protocol Overview and Timing

The MERCURIUS™ Multiplexed Plant RNA-seq (mRNA) kits allow the preparation of Illumina-compatible RNA sequencing libraries for up to 384 plant RNA samples in a time- and cost-efficient manner. This protocol is based on the enrichment of the mRNA pool, its further fragmentation with subsequent modifications, and the library preparation using Multiplexed RNA-seq technology.

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	10715	96WP	96	96	4
4x 96-sample	11613	96WP	96	384	4

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

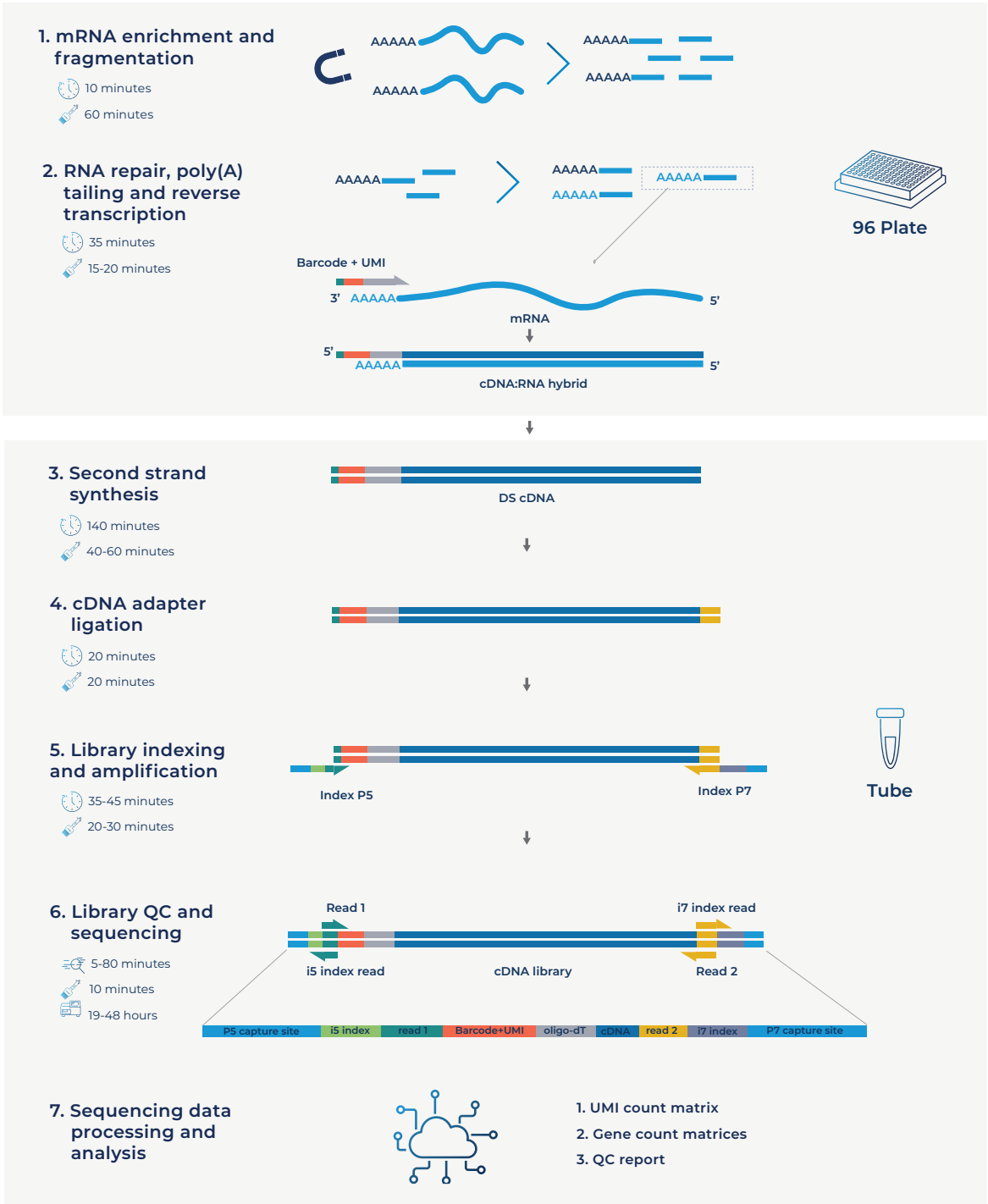
The Multiplexed RNA-seq technology can be used to generate high-quality sequencing data starting with 10 - 1000 ng of total purified RNA per sample. Notably, the kit can be used to pool any number of samples up to 96 with two considerations:

- The total RNA amount per pool should be at least 1000 ng.
- 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. 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, assuming that RNA samples have been previously quality-checked and diluted.

Protocol Workflow



Overall time

Incubation time: 4h-4h10.
Hands-on time: 2h45-3h20.
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).

Figure 1 Schematic illustration of the protocol workflow

Part 1. PREPARATION OF RNA SAMPLES

Guidelines for RNA samples

The Multiplexed RNA-seq protocol is based on early sample multiplexing; therefore, it is critical to ensure the quantity, purity, and quality of RNA are uniform for all samples before starting the library preparation. Individual sample quality checks and adjustments will not be possible after pooling.

Quantity

The tested range of total RNA amount per well is 10 – 1000 ng. The recommended starting amount is 100 ng per well for the library, prepared from at least eight wells pooled after RT. The more RNA used in the RT reaction, the higher the library complexity is observed after sequencing.

Purity

Plant RNA samples extracted using TRIzol™, phenol, or chloroform compounds are prone to residual contamination with organic solvents that may inhibit the reverse transcription reaction. This usually results in low cDNA library yields, loss of sequencing reads for a fraction of pooled samples, and uneven read distributions. To ensure the high purity of RNA, assess the 260/230 ratio of at least a few samples from the same RNA isolation batch using a spectrophotometer (i.e., Nanodrop). The 260/230 ratio values should be >1.5.

Uniformity

To ensure an even distribution of reads after sequencing, the RNA amount, integrity (RIN number), and 260/230 values of the starting RNA samples must be as uniform as possible, i.e., 50 ng (±10%) of starting material with RIN > 7 for every sample. To obtain such uniform amounts, we therefore recommend to:

- Measure the RNA concentration of all samples with a dye-based method (e.g., Qubit Quant-iT or RiboGreen for a large number of samples).
- Dilute samples to obtain the same RNA concentration in all wells (±10%).
- Re-measure the RNA concentration of all samples to confirm uniform concentrations across the samples.
- Ensure that the 260/230 ratios are >1.5 and similar across the samples.
- Ensure that the RIN values are similar across the samples, and > 7.

Contact info@alitheagenomics.com for technical support.

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. mRNA enrichment and fragmentation

At this step, mRNA molecules are enriched from every sample due to oligo-dT-based purification from total RNA. We highly recommend using the pipetting robot to minimize the variation between sample preparations due to different volumes, times of purification, drying, elution, etc.

Preparation

- Thaw the RNA samples on ice.
- Thaw the **FSB FL** reagent at room temperature and mix well before use.
- Prepare Program 1_Fragmentation on the thermocycler (set the lid at 100°C) and preheat it to 94°C

Step	Temperature, °C	Time
Incubation	94	3 min
Keep	4	pause

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

Procedure

We recommend using the NEBNext Poly(A) mRNA Magnetic Isolation kit (NEB #E7490) for the mRNA enrichment procedure.

Before starting, complete each sample volume with nuclease-free water to each sample to start with 50 µL for mRNA enrichment and proceed with the corresponding protocol until the elution step. For the latter, follow the procedure below:

2.1.1. Prepare the Elution Master mix (+10%) as follows:

Reagent	Volume, µL	
	Per well	96 wells +10%
FSB FL	2.2	242
Water	8.8	968
TOTAL	**11**	**1210**

2.1.2. Using a multichannel pipette or robot, transfer 11 µL of the Elution Master mix to each well and pipette a few times to ensure a proper resuspension of the beads.

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

2.1.4. Incubate in thermocycler Program 1_Fragmentation. Do not exceed the incubation time, which can lead to mRNA degradation.

2.1.5. Briefly spin the samples in the centrifuge, open the seal and pipette beads a few times.

2.1.6. Place the plate on the magnetic stand, wait 5 min, and carefully transfer 10 µL of supernatant with fragmented mRNA into a plate with barcoded oligo-dT primers directly (keep it on ice).

2.1.7. Pipette 3-5 times to ensure proper reconstitution of dried oligo-dT. The appearance of red in all wells indicates a proper and uniform reconstitution of oligos.

2.1.8. Carefully seal the plate and briefly spin it in the centrifuge.

2.1.9. Proceed immediately to step 2.2.

2.2. RNA repair, poly (A) tailing, and reverse transcription

At this step, fragmented mRNA molecules are repaired, poly-adenylated, and reverse-transcribed using the barcoded oligo-dT primers provided in a 96-well plate. Subsequently, all the barcoded samples can be pooled into one 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 to visualize reaction preparation and pooling better.

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 the reagents will enable the visualization of red color, confirming the presence of the oligos in all wells.

Preparation

- Thaw all tubes on ice and mix well before use.
- Prepare Program 2_Repair/RT on the thermocycler (set the lid at 90°C) and pre-heat it to 37°C:

Step	Temperature, °C	Time
Incubation	37	30 min
Inactivation	75	5 min
Keep	4	pause

Procedure

2.2.1. Keep the plate with RNA and oligo-dT on ice.

2.2.2. Depending on the number of samples, prepare the following Repair/RT Master mix (+10%) as follows:

Reagent	Volume, µL	
	Per well	96 wells +10%
RAM	8.75	927.5
RAE	1.25	132.5
TOTAL	**10**	**1060**

2.2.3. Using a multichannel pipette or robot, pipette 10 µL of Repairing/RT Master mix to each well and mix a few times.

2.2.4. Carefully seal the plate and briefly spin it in the centrifuge.

2.2.5. Incubate in a thermocycler Program 2_Repair/RT.

2.2.6. Proceed immediately to step 2.3.

2.3. Sample pooling and bead purification

After pooling, the barcoded RT samples can be purified using SPRI magnetic beads.

NOTE: Library normalization

The volume used for pooling from each well can be adjusted to re-equilibrate the proportion of samples in the pool, helping to improve the distribution of sequencing reads in the library, especially if some samples risk obtaining too many reads.

Shallow sequencing allows for assessing the coverage per sample unequivocally. For this approach, we recommend pooling only a fraction of the RT volume from each well (i.e., 10 µL out of 20 µL) for the library preparation. After the library QC by sequencing (see section 2.8), the volume used for pooling can be re-adjusted to reduce the variation at the sequencing stage.

Perform cDNA purification with SPRI magnetic beads using a 1:1.8 ratio of cDNA pool and beads slurry. Purifying large volumes (i.e., 2 mL of the pool if 20 µL of 96 samples are pooled) requires three to four 1.5 mL tubes and a corresponding magnetic stand (Permagen, MSR06).

If the pool's volume is higher than 500 µL, split it equally in the required number of 1.5- 2 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 900 µL of beads per tube).

Preparation

- Pre-warm the SPRI beads at room temperature for ~30 min.
- Prepare 5 mL of 80% ethanol.

Procedure

2.3.1. Using a multichannel pipette or robot, pool the RT samples in the reservoir (Integra, 4382 or 6318).

2.3.2. Transfer the collected pool into a 2 mL or 5 mL tube, depending on the pooled volume. Consider that the final volume will be almost three times higher due to the addition of the beads.

2.3.3. Add pre-warmed beads in a 1:1.8 ratio (i.e., for 960 µL of pooled samples, add 1728 µL of beads slurry), and mix by pipetting up and down ten times.

2.3.4. Incubate for 5 min at room temperature.

2.3.5. If needed, split the volume into a few tubes.

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

2.3.7. To wash the beads, pipette 1000 µL of freshly prepared 80% ethanol into the tube.

2.3.8. Incubate for 30 sec.

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

2.3.10. Repeat step 2.3.7 for a total of two washes.

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

2.3.12. Resuspend the beads in 21 µL of water and incubate for 1 min.

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

2.3.14. Immediately proceed to step 2.4

If the RT pool was split into several tubes at step 2.3.5, resuspend the beads in the **first tube** in 22 µL of water. Keep other tubes closed to avoid over-drying of the beads. Transfer the obtained elution to the next tube and resuspend the beads. Repeat this step for every tube.

2.4. Free primer digestion

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

Preparation

- Pre-warm the SPRI beads at room temperature for ~30 min.
- Prepare 5 mL of 80% ethanol.
- Label 0.2 mL PCR tubes corresponding to the number of pools prepared.
- Thaw the **F-EXB** reagent at room temperature.
- Keep the **F-EXO** reagent constantly on ice.
- Prepare Program 3_FPD on the thermocycler (set the lid at 90°C):

Step	Temperature, °C	Time
Incubation	37	6 min
Incubation	80	1 min
Keep	4	pause

Procedure

2.4.1. Transfer 17 µL of the eluate from step 2.3.13 into a new labeled 0.2 mL PCR tube.

2.4.2. Prepare the F-EXO reaction Master mix as follows (with 10% excess):

Reagent	Per reaction + 10%, µL
F-EXB	2.2
F-EXO	1.1
TOTAL	**3.3**

2.4.3. Transfer 3 µL of the F-EXO reaction mix into each PCR tube with purified cDNA.

2.4.4. Mix by pipetting up and down 5 times.

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

2.4.6. Incubate in a thermocycler Program 3_FPD.

2.4.7. Proceed to step 2.5.1 or keep the tube at 4°C overnight.

2.5. Second-strand synthesis and DNA repair

At this step, double-stranded full-length cDNA is generated and repaired.

Preparation

- Pre-warm the SPRI beads at room temperature for ~30 min.
- Prepare 5 mL of 80% ethanol.
- Thaw the **SSB FL** reagent at room temperature and mix well before use.
- Keep the **SSE FL** reagent constantly on ice.
- Prepare Program 4_SSS on the thermocycler (set the lid at 90°C):

Step	Temperature, °C	Time
Incubation	16	60 min
Incubation	70	20 min
Keep	4	pause

Procedure

2.5.1. Add 11 µL of water to the tube from step 2.4.7.

2.5.2. Prepare the SSS FL Master mix for the second-strand synthesis as follows (with 10% excess):

Reagent	Per reaction + 10%, µL
SSB FL	5.5
SSE FL	4.4
TOTAL	**9.9**

2.5.3. Transfer 9 µL of the SSS reaction mix to the tube from step 2.4.7 and mix well by pipetting up and down 5 times.

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

2.5.5. Incubate in thermocycler Program 4_SSS.

2.5.6. Proceed immediately to step 2.5.7

cDNA clean-up with SPRI beads

Perform the cDNA purification with SPRI magnetic beads using a 1.8x beads:cDNA ratio (i.e., 90 µL of bead slurry plus 50 µL of cDNA).

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

2.5.7. Complement the final reaction volume to 50 µL with water.

2.5.8. Add 90 µL of beads and mix by pipetting 10 times.

2.5.9. Incubate for 5 min at room temperature.

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

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

2.5.12. Incubate for 30 sec.

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

2.5.14. Repeat step 2.5.11 for a total of two washes.

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

2.5.16. Resuspend the beads in 21 µL of water.

2.5.17. Incubate for 1 min.

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

2.5.19. Use 2 µL to measure the concentration with Qubit (recommended).

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

2.6. cDNA adaptor ligation

At this step, the BRB-compatible adaptor is ligated to the cDNA fragments to facilitate the following amplification of the library with Unique Dual Indexing (UDI) primers.

Preparation

- Pre-warm the SPRI beads at room temperature for ~30 mins.
- Prepare 5 mL of 80% ethanol.
- Thaw the **ALB** and **BRB.AD** reagents on ice and mix well before use.
- Prepare Program 5_ADL on the thermocycler (set the lid at 90°C):

Step	Temperature, °C	Time
Incubation	20	15 min
Keep	4	pause

Procedure

2.6.1. Complement every sample from step 2.5.18 to 18.75 µL with nuclease-free water (if necessary). Then pipette **BRB.AD,** and then add **ALB** as indicated in the table below. It is **not recommended** to prepare a master mix for all samples.

Reagent	Per reaction, µL
cDNA	18.75
BRB.AD	1.25
ALB	20
TOTAL	**40**

2.6.2. **CRITICAL:** Mix well by pipetting up and down 10 times. This is essential to ensure a sufficient ligation. The presence of small bubbles will not interfere with performance.

2.6.3. Briefly spin down the tube in the bench-top centrifuge.

2.6.4. Incubate in a thermocycler Program 5_ADL.

2.6.5. **CRITICAL:** Proceed immediately to step 2.6.6.

cDNA clean-up with SPRI beads

Perform the cDNA purification with SPRI magnetic beads using 0.9x beads:cDNA ratio (i.e., 45 µL of bead slurry plus 50 µL of cDNA).

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

2.6.6. Complement the final reaction volume to 50 µL with water.

2.6.7. Add 45 µL of beads and mix by pipetting 10 times.

2.6.8. Incubate for 5 min at room temperature.

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

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

2.6.11. Incubate for 30 sec.

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

2.6.13. Repeat step 2.6.10 for a total of two washes.

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

2.6.15. Resuspend the beads in 21 µL of water.

2.6.16. Incubate for 1 min.

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

2.6.18. Proceed to step 2.7.1.

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

2.7. Library indexing and amplification

At this step, the library undergoes amplification using Unique Dual Indexing (UDI) primers. The kit contains four Illumina-compatible primer pairs to generate up to four UDI libraries. The index sequences are indicated in Table 2.

The number of amplification cycles required for library preparation is usually 10-14, depending on the number and quantity of RNA samples.

It is strongly recommended that the final library bead clean-up be performed twice to remove primer dimer fragments.

Preparation

- Pre-warm the SPRI beads at room temperature for ~30 min.
- Prepare 10 mL of 80% ethanol.
- Thaw the **LAM FL** reagents on ice and mix well before use.
- Thaw the required number of **MF.UDI Adapters** at room temperature; briefly spin them before use.
- Prepare Program 6_AMP (set the lid at 100°C) on the thermocycler:

Step	Temperature, °C	Time	Cycles
Initial denaturation	98	30 sec	1
Denaturation	98	10 sec	5-20*
Annealing and Extension	65	75 sec	
Final extension	65	5 min	1
Keep	4	pause	

*The required number of PCR cycles can be estimated based on the amount of cDNA used for adapter ligation (preferably) or the total RNA input used for the protocol. Follow the guidelines below.

Library amplification reaction setup

2.7.1. Prepare the PCR amplification reaction as follows:

Reagent	Per reaction, µL
LAM FL	25
MF.UDI Adapter	5
Ligated cDNA	20
TOTAL	**50**

2.7.2. Pipette up and down 5 times.

2.7.3. Put the tube in the PCR machine and start Program 6_AMP.

2.7.4. Set the required number of PCR cycles based on the amount of cDNA used for adaptor ligation (step 2.5.19).

cDNA used for library prep, ng	Number of cycles
80	6-7
40	8
20	9
10	10
5	11
2.5	12
1.25	13
0.6	14

2.7.5. If the cDNA amount is unavailable, please refer to the total RNA input (used for step 2.1).

Total RNA input, ng	Number of cycles
8’000	9-10
1’000	12-13
500	13-14
200	15-17
100	16-18
50	18-20

Indexed cDNA library clean-up with SPRI beads

Purify the final cDNA library with SPRI magnetic beads using a 0.9x ratio (45 µL of bead slurry for 50 µL cDNA library).

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

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

2.7.7. Add 45 µL of beads and mix by pipetting up and down 10 times.

2.7.8. Incubate for 5 min at room temperature.

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

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

2.7.11. Incubate for 30 sec.

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

2.7.13. Repeat step 2.7.10 for a total of two washes.

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

2.7.15. Resuspend the beads in 21 µL of water.

2.7.16. Incubate for 1 min.

2.7.17. 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.

2.7.18. Perform the bead clean-up once again by repeating the procedure from step 2.7.6.

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

2.8. Library quality control

Pooled library quality control

Before sequencing, the libraries should be subjected to fragment analysis (with a 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 300 – 700 bp range with a peak at 400-500 bp; see Figure 2 for an example of a standard Multiplexed RNA-seq library profile.

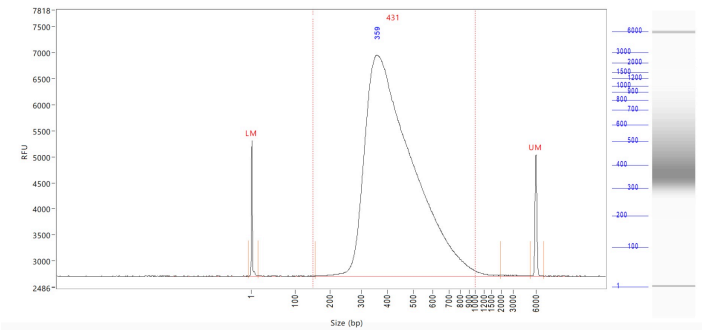
Importantly, the bead clean-up must be performed twice to remove primer dimer fragments, likely producing lower-quality sequencing data with reduced mappable reads (Figure 3). Therefore, it is strongly recommended that those peaks be removed by performing an additional round of SPRI bead purification with the 0.9x ratio (see step 2.7.6).

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 2.7.6 – 2.7.17 to remove the peaks <300 bp.

Figure 2 A successful library profile with fragments between 300-700 bp



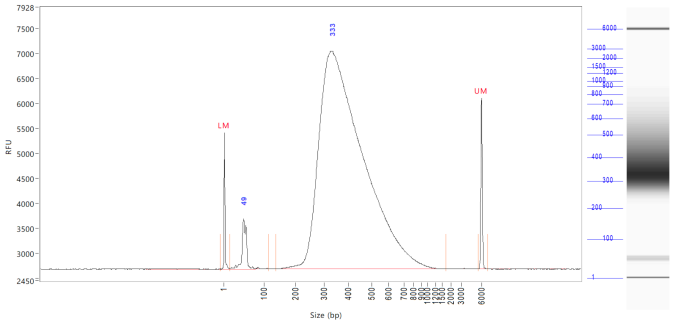


Figure 3 An example of a library profile after a single purification, demonstrating the small peak at 50 bp

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 2.3). This approach ensures that every sample will obtain enough reads required for the analysis. These 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™ Multiplexed RNA-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™ Multiplexed RNA-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 1 for the optimal sequencing structure and Table 2 for the list of i5 and i7 index sequences.

Given the Multiplexed RNA-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 1 Sequencing structure of Multiplexed RNA-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 2).

Name	Type	i7 index sequence	i5 index sequence Forward Workflow	i5 index sequence Reverse Workflow
MF.UDI.1	UDI (i7/i5)	GCTTGTCA	AGGCGAAG	CTTCGCCT
MF.UDI.2	UDI (i7/i5)	CAAGCTAG	TAATCTTA	TAAGATTA
MF.UDI.3	UDI (i7/i5)	AGTTCAGG	CAGGACGT	ACGTCCTG
MF.UDI.4	UDI (i7/i5)	GACCTGAA	GTACTGAC	GTCAGTAC

Table 2 UDI adapter sequences

NOTE: Sequencing depth

1. The recommended sequencing depth is 5-10 Mio reads per sample (see Figure 4). Overall, the higher the input, the more genes can be detected at the same sequencing depth (compare 10 ng and 100 ng RNA samples in Figure 4). We recommend getting 15-20 Mio reads per sample 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 sequencing instrument (see 3.1. and 3.2) and should be discussed with the sequencing facility or an experienced person.

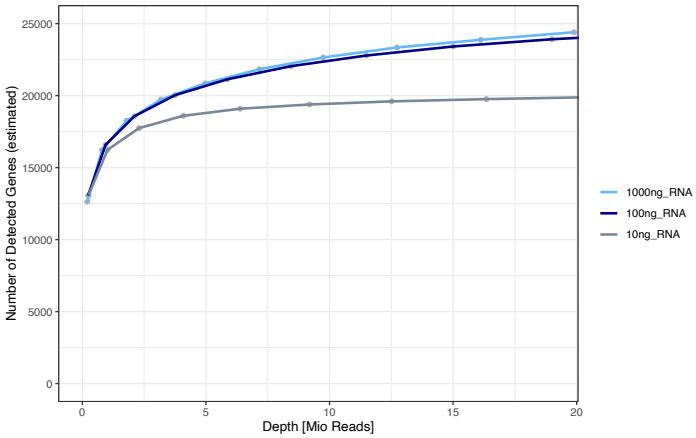


Figure 4 Number of detected genes as a function of the sequencing depth for a different quantity of starting RNA per well (Universal Human Reference RNA, Thermofisher, QS0639)

3.1. Sequencing on the Illumina instruments

The loading concentration for the Illumina instruments is indicated in Table 5. Please refer to Appendix 1 for the list of Illumina instruments with 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 Multiplexed RNA-seq libraries sequencing. We recommend a prior dilution of the libraries to 0.8 nM before denaturation.

Table 3 Reference loading concentrations for various Illumina instruments

3.2. Sequencing on the Element AVITI instrument

For the most optimal results, the MERCURIUS™ Multiplexed RNA-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 4).

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 4 Loading concentration for Cloudbreak AVITI 2x75 High Output sequencing kit

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 transcriptome-level count matrices from raw FASTQ files.

To obtain the data ready for analysis, the user needs to demultiplex the sequencing reads by sample barcodes and perform transcriptome pseudo-alignment and quantification, as well as perform an alignment of the sequencing reads to the genome, and perform the gene/UMI read count generation.

For manual data processing, the user requires a terminal and a server or a 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. Merging FASTQ files from individual lanes and/or libraries (Optional)

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
```

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
```

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 (i.e., L001, L002, L003, L004, not L002, L001 ...) to avoid issues when demultiplexing the samples.

4.3. Sequencing data quality check

Perform basic quality control checks on raw sequencing reads to assess read quality, GC content, duplication levels, adapter contamination, and other key metrics before downstream processing. Run FastqQC 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.

Input:

- Raw FASTQ files from sequencing (e.g., mylibrary_R1.fastq.gz, mylibrary_R2.fastq.gz)

Output:

- HTML and .zip QC reports in the specified output directory
  (e.g., fastqc_out_dir/mylibrary_R1_fastqc.html, fastqc_out_dir/mylibrary_R2_fastqc.html)

Command line:

```
> fastqc --outdir fastqc_out_dir/ mylibrary_R1.fastq.gz
> fastqc --outdir fastqc_out_dir/ mylibrary_R2.fastq.gz
```

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 there is some *adapter contamination* or *overrepresented sequence* 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.

4.4. Pseudo-alignment and transcriptome quantification

This section explains how to process Multiplexed RNA-seq (mRNA) sequencing data to obtain transcript-level expression estimates. This approach uses **pseudo-alignment,** which provides a much faster and resource-efficient method for determining transcript abundances from sequencing reads.

4.4.1. Demultiplex FASTQ files

Multiplexed RNA-seq libraries are generated by pooling barcoded samples prior to sequencing. The first step is to **split this pooled data into individual samples** using the inline barcodes embedded in Read 1 or Read 2. `fqtk` — a tool designed for barcode and UMI extraction and demultiplexing.

Input:

- Raw paired-end FASTQ files (e.g., mylibrary_R1.fastq.gz, mylibrary_R2.fastq.gz)
- Barcode reference file (tab-separated: sample_id <TAB> barcode sequence)
- CIGAR strings that specify barcode/UMI layout in reads

Output:

- Demultiplexed paired FASTQ files per sample (e.g., sample1_R1.fq.gz, sample1_R2.fq.gz)

Command line:

```
> fqtk demux -i mylibrary_R1.fastq.gz
  mylibrary_R2.fastq.gz \
>            -r 14B14M 90T \
>            -o dmx_fastq \
>            -s barcode_ref.txt
```

NOTE: 14B14M means a 14-nt cell barcode followed by a 14-nt UMI in Read 1, and 90T is a fixed 90-nt length of genomic Read 2.

4.4.2. Build transcriptome index

Transcript quantification tools like Kallisto require a **prebuilt index** of the transcriptome. This index maps k-mers to known transcripts and is used to **efficiently pseudoalign reads.**

Input:

- cDNA FASTA file (transcript sequences)
- (Optional) Corresponding GTF annotation (not needed for index, but useful for downstream analysis)

Output:

- Kallisto index file (e.g., Homo_sapiens.GRCh38.idx)

Command Example:

```
> kallisto index -i Homo_sapiens.GRCh38.idx
  Homo_sapiens.GRCh38.cdna.all.fa
```

How to Obtain the FASTA File:

Download the cDNA reference FASTA from Ensembl:

```
> wget ftp://ftp.ensembl.org/pub/release-
104/fasta/homo_sapiens/cdna/Homo_sapiens.GRCh38.cdna.all.fa.gz
```

```
> gunzip Homo_sapiens.GRCh38.cdna.all.fa.gz
```

4.4.3. Quantify transcript abundance

This step estimates **transcript-level expression** for each sample using Kallisto's fast and pseudoalignment algorithm. Outputs include abundance estimates, effective lengths, TPM expression values, and bootstrapping results.

Input:

- Sample-specific FASTQ files from demultiplexing step 4.4.1.
- Kallisto index file

Output:

- Abundance files for each sample (e.g., abundance.tsv, run_info.json)

Command Example:

```
> kallisto quant -i Homo_sapiens.GRCh38.idx  -o quant/sample1 -l 550 -
  s 150 -b 5 sample1_R1.fq.gz sample1_R1.fq.gz -t 30
```

NOTE:

`-l` and `-s` indicate estimated fragment length and estimated standard deviation of fragment length in a **single-end** run — only one input FASTQ should be listed.

`-b 5` adds bootstrapping for variance estimation.

`-t 30` uses 30 threads for performance.

4.4.4. Assemble transcriptome counts

To facilitate downstream analysis (e.g., differential expression, clustering), you can **combine quantification results across all samples** into a single matrix of estimated transcript counts.

Input:

- Quant/ folder containing all abundance.tsv files

Output:

- Count matrix (e.g., transcript_count_matrix.csv)

Example R script:

```
    # R script for collecting transcriptome data

     cbind_vec2matrix <- function(list_vectors, row_names,
col_names){
         df_ = data.frame(do.call(cbind, list_vectors))
         colnames(df_) = col_names
         rownames(df_) = row_names
         return(df_)
     }

     list_dirs = list.dirs("quant/", recursive = F)
     est_counts_l = list()
     tmp_l = list()
     sample_name_l = list()
     for(i in 1:length(list_dirs)){
         this_dir = list_dirs[[i]]
         abundance_file = paste0(this_dir, '/abundance.tsv')
         if(file.exists(abundance_file)){
             abundance_tab = read.table(abundance_file, header=T)
             est_counts_l[[i]] = abundance_tab[['est_counts']]
             tmp_l[[i]] = abundance_tab[['tpm']]
             sample_name_l[[i]] = gsub("^\\\\/", "", gsub("$in_dir",
'', this_dir))
         }
     }

     df_counts = cbind_vec2matrix(est_counts_l, row_names =
abundance_tab[["target_id"]], col_names = unlist(sample_name_l))
     df_tpm    = cbind_vec2matrix(tmp_l,        row_names =
abundance_tab[["target_id"]], col_names = unlist(sample_name_l))

     lib_name = gsub("_kallisto_out","","$in_dir")
     write.csv(df_counts, paste0(lib_name,".counts.txt"), quote=F)
     write.csv(df_tpm, paste0(lib_name,".tpm.counts.txt"), quote=F)
```

NOTE: Ensure that you are running the script from the same directory that contains the `quant/` folder.

4.5. Alignment and gene quantification

While transcriptome pseudo-alignment provides fast quantification at the transcript level, full-length data also benefits from traditional genome alignment to support additional analyses—such as quality control metrics, gene body coverage, and gene-level quantification. This section explains how to align sequencing reads to a reference genome using the **STAR** aligner, and how to use the aligned reads to generate gene-level counts.

4.5.1. Preparing the reference genome

Before aligning sequencing reads, a genome index must be generated from a reference genome and a corresponding gene annotation file. This is a one-time step per genome version. STAR uses this index to efficiently map RNA-seq reads, including for multiplexed libraries (e.g., Multiplexed RNA-seq), using its STARsolo mode, which also generates count matrices.

Input:

- Reference genome FASTA file (e.g., Homo_sapiens.GRCh38.dna.primary assembly.fa)
- Gene annotation GTF file (e.g., Homo_sapiens.GRCh38.108.gtf).

Output:

- A directory containing STAR genome index files, including:
  - SA
  - Genome
  - sjdbList.fromGTF.out.tab
  - and other supporting files

Download reference files (Ensembl, Human GRCh38):

```
> # Download and decompress reference genome (FASTA)
> wget https://ftp.ensembl.org/pub/release-
  108/fasta/homo_sapiens/dna/Homo_sapiens.GRCh38.dna.pri
  mary_assembly.fa.gz
> gzip -d Homo_sapiens.GRCh38.dna.primary_assembly.fa.gz
>
> # Download and decompress annotation file (GTF)
> 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
```

Recommendation:

- Use the `primary_assembly` FASTA file when available (avoid 'sm' or 'rm' tags).
- For GTF, select the version without `chr` prefixes or `abinitio` tags for compatibility.

Index Generation Command:

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

Parameter details:

- `--genomeDir`: Output directory where STAR will write the index files.
- `--genomeFastaFiles`: Full path to the decompressed FASTA file.
- `--sjdbGTFfile`: Full path to the decompressed GTF annotation file.
- `--runThreadN`: Number of CPU cores to use for parallel processing.

NOTE:

- Adjust `--runThreadN` based on available CPU cores (higher values = faster indexing).
- STAR requires ~32–40 GB RAM depending on genome size — ensure sufficient memory.
- The generated index can be reused for all analyses with the same genome/annotation.

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

After the reference genome index is prepared, sequencing reads (FASTQ files) can be aligned to the genome using STAR. In this context, the STARsolo mode is used, which not only performs the alignment

but also generates gene and UMI (unique molecular identifier) count matrices directly. This step is tailored for multiplexed libraries such as those used in the Multiplexed RNA-seq (mRNA) protocol.

Input:

- Paired-end FASTQ files (e.g., mylibrary_R1.fastq.gz and mylibrary_R2.fastq.gz)
- STAR genome index (from Step 1.5.1)
- Barcode whitelist file (text file with one barcode sequence per line). Example:

```
TACGTTATTCCGAA
AACAGGATAACTCC
ACTCAGGCACCTCC
ACGAGCAGATGCAG
```

Output:

- Aligned BAM files (sorted by coordinate)
- Gene and UMI count matrices in Matrix Market (MTX) format
- Demultiplexing statistics (Solo.out/Barcodes.stats)
- Alignment summary (Log.final.out)

Command example:

```
> STAR --runMode alignReads \
>      --outSAMmapqUnique 60 \
>      --runThreadN 8 \
>      --outSAMunmapped Within \
>      --soloStrand Forward \
>      --quantMode GeneCounts \
>      --outBAMsortingThreadN 8 \
>      --genomeDir /path/to/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 /path/to/bamdir/libraryname/
  \
>      --readFilesIn mylibrary_R2.fastq.gz
  mylibrary_R1.fastq.gz
```

Parameter details:

- `--genomeDir`: Path to the STAR genome index directory (e.g., `/path/to/genomeDir`).
- `--readFilesIn`: Order is important. R2 (genomic reads) first, R1 (barcodes and UMI) second.
- `--soloCBwhitelist`: Text file with one barcode per line; use the version appropriate for your kit (e.g., Mercurius V5).
- `--soloCBstart, --soloCBlen`: Barcode position and length in R1 (e.g., start = 1, length = 14).
- `--soloUMIstart, --soloUMIlen`: UMI starts after barcode (e.g., start = 15, length = 14).
- `--soloUMIdedup: NoDedup`: produces read count matrix; `1MM_Directional:` adds UMI count matrix

- `--outFileNamePrefix:` Output prefix directory, e.g., `/path/to/bamdir/libraryname/`

Additional NOTE:

- **Output folder**: BAM files and matrices will be saved in `/path/to/bamdir/libraryname/`
- **Demultiplexing statistics**: Found in `/path/to/bamdir/libraryname/Solo.out/Barcodes.stats`
- **Alignment summary**: Found in `/path/to/bamdir/libraryname/Log.final.out`

IMPORTANT: 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.

4.5.3. 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:

```
> library(data.table)
> library(Matrix)
> matrix_dir <- "/path/to/bamdir/libraryname/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.

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

Once you obtain a multiplexed BAM file from STARsolo, you can use **FastReadCounter** to extract gene-level read counts for each sample based on the associated barcodes, ready for downstream differential expression analysis.

Input:

- **BAM file**: A multiplexed alignment file (e.g., /path/to/bamdir/libraryname/Aligned.sortedByCoord.out.bam)
- **GTF file**: Genome annotation file (e.g., Homo_sapiens.GRCh38.108.gtf)
- **Barcode file**: A tab-separated file listing the expected barcodes. Example of expected format:

```
TACGTTATTCCGAA sample_1
AACAGGATAACTCC sample_2
ACTCAGGCACCTCC sample_3
ACGAGCAGATGCAG sample_4
```

Output:

- **Gene read count matrix**: Located in the specified output folder (counts/)
  - One file per barcode/sample
  - Summary statistics per sample
  - Global matrix across all barcoded samples

Command line example:

```
> #!/bin/bash
```

```
>
> gtf_file=homo_sapience.gtf    # Genome annotation file in GTF
  format
> output_folder=counts/         # Output directory for final results
> bam_path=/path/to/bamdir/libraryname/Aligned.sortedByCoor
  d.out.bam     # Path to directory and prefix of the BAM file
> barcode_file=barcode_frc.txt # File listing expected barcodes
>
> FastReadCounter-1.0.jar \
>    --bam ${bam_path} \
>    --gtf ${gtf_file} \
>    --umi-dedup none \
>    --barcodeFile ${barcode_file} \
>    -o ${output_folder}
```

NOTE:

- The BAM file should be sorted and indexed if required by your downstream tools.
- The barcode file must match the barcodes used during Multiplexed RNA-seq library preparation.
- If you do **not have the barcode sequences**, please contact info@alitheagenomics.com, including the barcode set name and the Product Number (PN) of your barcode module

4.5.5. 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 (e.g., GEO or ArrayExpress), or for running a bulk RNA-seq analysis pipeline. This can be done using the **Picard** tool.

Input:

- Aligned.sortedByCoord.out.bam — multiplexed BAM file generated by STARsolo
- barcode_brb.txt — tab-separated file with two columns: sample_id and barcode sequence. Example:

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

Output:

- A BAM file for each sample in the specified output directory.

Command line example:

```
> #!/bin/bash
>
> demultiplexed_bam_out_dir=/path/to/output_bams
> input_bam=/path/to/bamdir/libraryname/Aligned.sortedByCoord.
  out.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"
```

Appendix 1. Compatible Illumina instruments

Illumina instruments can use two workflows for sequencing 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