Jan 09, 2026

RNA sequencing (RNA-seq) and analysis

DOI
https://dx.doi.org/10.17504/protocols.io.8epv55jkjv1b/v1

Fanxuan Tian[1]

[1]Department of Gastroenterology, Chongqing Key Laboratory of Digestive Malignancies, Daping Hospital, Chongqing 400042, China

Fanxuan Tian

OPEN ACCESS

DOI: https://dx.doi.org/10.17504/protocols.io.8epv55jkjv1b/v1

Protocol Citation: Fanxuan Tian 2026. RNA sequencing (RNA-seq) and analysis. **protocols.io** https://dx.doi.org/10.17504/protocols.io.8epv55jkjv1b/v1

Protocol status: Working
We use this protocol and it's working

Created: January 09, 2026

Last Modified: January 09, 2026

Protocol Integer ID: 238277

Keywords: data analysis for cultured murine cell, cultured murine cell, analysis rna, rna, sequencing, cell, seq

Abstract

RNA sequencing and data analysis for cultured murine cells, such as B16F10.

1 Total RNA is extracted from cultured cells using TRIzol reagent (15596018CN, Invitrogen).

2 Total RNA is quantified using a NanoDrop spectrophotometer and an Agilent 2100 bioanalyzer (Thermo Fisher).

3 RNA library construction and subsequent RNA sequencing are performed using BGI-Shenzhen (Shenzhen, China) software.

4 First-strand cDNA is generated by random hexamer-primed reverse transcription followed by second-strand cDNA synthesis.

5 A-Tailing Mix and RNA Index Adapters are added by incubation for end repair.

6 The cDNA fragments obtained from the previous step are amplified by PCR and the products are purified using Ampure XP Beads and dissolved in an EB solution.

7 The product is validated using an Agilent Technologies 2100 Bioanalyzer for quality control.

8 Double-stranded PCR products from the previous step are denatured and circularized using splint oligo sequences to obtain the final library.

9 Single-stranded circular DNA is used as the final library. The final library is amplified with phi29 to form a DNA nanoball, which has over 300 copies of one molecule and is loaded into the patterned nanoarray, and single end 100 bases reads are generated on the BGIseq500 platform (BGI-Shenzhen).

10 SOAPnuke (v1.5.6)

Citation

Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, Li Y, Ye J, Yu C, Li Z, Zhang X, Wang J, Yang H, Fang L, Chen Q
(2018)
. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data.
GigaScience.

https://doi.org/10.1093/gigascience/gix120 LINK

is used to filter out adapter sequences and eliminate poor-quality bases from raw sequencing data. A read is excluded if its sequence closely matched the adapter sequences.

11 Reads are excluded if they contained > 20% low-quality bases (base quality < 15) or > 5% unknown bases.

12 Clean reads are mapped to Ensemble GRCm38 using HISAT.

Citation

Kim D, Langmead B, Salzberg SL (2015)
. HISAT: a fast spliced aligner with low memory requirements. Nature methods.

https://doi.org/10.1038/nmeth.3317 LINK

13 To obtain gene expression data, RNA-Seq by Expectation Maximization (RSEM)

Citation

Li B, Dewey CN (2011)
. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome.
BMC bioinformatics.

https://doi.org/10.1186/1471-2105-12-323 LINK

is used to count the number of genes and quantify gene expression by calculating the Fragments Per Kilobase of exon per Million mapped reads (FPKM) for downstream analysis.

Citations

Step 10

Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, Li Y, Ye J, Yu C, Li Z, Zhang X, Wang J, Yang H, Fang L, Chen Q. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data.
https://doi.org/10.1093/gigascience/gix120

Step 12

Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements.
https://doi.org/10.1038/nmeth.3317

Step 13

Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome.
https://doi.org/10.1186/1471-2105-12-323