# RNA SEQUENCING METHODS COLLECTION

An overview of recent RNA-Seq publications featuring Illumina® technology
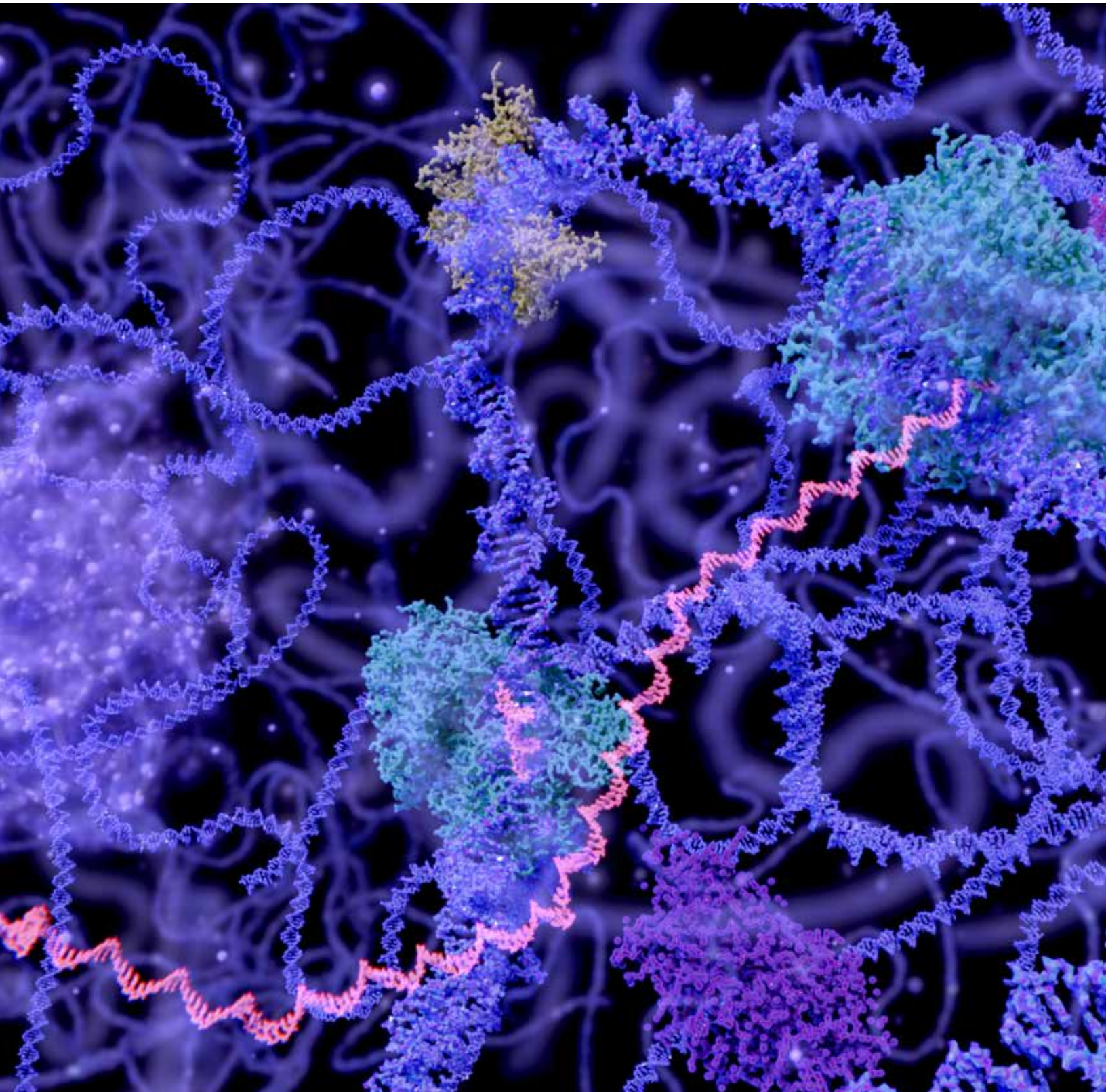
# TABLE OF CONTENTS

# INTRODUCTION

This publication is a collection of next-generation sequencing (NGS) methods for RNA sequencing, compiled from the scientific literature. It is both a tribute to the creativity of the users and the versatility of the technology. We hope it will inspire researchers to use these methods or to develop new ones to address new scientific challenges.

A method refers to the processing steps between extracting the nucleic acids (sample preparation) and the addition of oligonucleotide adapters for sequencing (library preparation). With a few extra processing steps, a wide range of scientific questions can be addressed by this technology.

When using this publication, consider the following points:

- New methods are continually being developed. For the most up-to-date list of methods, visit: www.illumina.com/science/sequencing-method-explorer.html
- Only the most recent references are provided, typically for the past 2 years.
- With few exceptions, the capitalization, punctuation, and special characters in the method name are exactly the same as in the original, referenced publication. This is important, because methods such as CapSeq[1] and CAP-seq[2] are quite different and refer to RNA and DNA method, respectively.
- The methods are arranged according to their similarity, so the most similar methods should be adjacent. To visually compare the methods, refer to http://www.illumina.com/methods-explorer
- When methods are essentially identical, and can be represented by a single diagram, they are grouped together, such as GRO-seq[3] and BRIC-seq.[4]
- The diagrams are stylized depictions of the processes and are intended to depict only the key elements of the method.
- These methods were developed by users, so readers should refer to the original publications for detailed descriptions and protocols.

Have we missed anything? Contact us if you are aware of a protocol that should be listed at scientificaffairs@illumina.com.

1. Gu W., Lee H. C., Chaves D., et al. CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as *C. elegans* piRNA precursors. *Cell*. 2012;151: 1488-1500
2. Illingworth R. S., Botting C. H., Grimes G. R., Bickmore W. A. and Eskeland R. PRC1 and PRC2 are not required for targeting of H2A.Z to developmental genes in embryonic stem cells. *PLoS One*. 2012;7: e34848
3. Core L. J., Waterfall J. J. and Lis J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science*. 2008;322: 1845-1848
4. Tani H., Mizutani R., Salam K. A., et al. Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. *Genome Res*. 2012;22: 947-956

# RNA TRANSCRIPTION

The regulation of RNA transcription and processing directly affects protein synthesis. Proteins, in turn, mediate cellular functions to establish the phenotype of the cell. Dysregulated RNAs are the cause of some diseases and cancers.[5,6] Sequencing RNA provides information about both the abundance and sequence of the RNA molecules. Careful analysis of the results, along with adaptation of the sample preparation protocols, can provide remarkable insight into all the various aspects of RNA processing and control of transcription. Examples of these measures include: post-translational modifications, RNA splicing, RNA bound to RNA-binding proteins (RBPs), RNA expressed at various stages, unique RNA isoforms, RNA degradation, and regulation of other RNA species.[7,8] Studies of RNA transcription and translation are leading to a better understanding of the implications of RNA production, processing, and regulation for cellular phenotype.

Scientists have discovered a link between long-term memory and protein synthesis in the brain.[9,10]

## Reviews

Conesa A., Madrigal P., Tarazona S., et al. A survey of best practices for RNA-Seq data analysis. *Genome Biol.* 2016;17: 13

Campbell Z. T. and Wickens M. Probing RNA-protein networks: biochemistry meets genomics. *Trends Biochem Sci.* 2015;40: 157-164

Flynn R. A., Martin L., Spitale R. C., et al. Dissecting noncoding and pathogen RNA-protein interactomes. *RNA*. 2015;21: 135-143

Kolodziejczyk A. A., Kim J. K., Svensson V., Marioni J. C. and Teichmann S. A. The Technology and Biology of Single-Cell RNA Sequencing. *Mol Cell.* 2015;58: 610-620

Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38: 226-236

Reuter J. A., Spacek D. V. and Snyder M. P. High-Throughput Sequencing Technologies. *Mol Cell.* 2015;58: 586-597

Hausser J. and Zavolan M. Identification and consequences of miRNA-target interactions--beyond repression of gene expression. *Nat Rev Genet.* 2014;15: 599-612

5.  Kloosterman W. P. and Plasterk R. H. The diverse functions of microRNAs in animal development and disease. *Dev Cell.* 2006;11: 441-450
6.  Castello A., Fischer B., Hentze M. W. and Preiss T. RNA-binding proteins in Mendelian disease. *Trends Genet.* 2013;29: 318-327
7.  McGettigan P. A. Transcriptomics in the RNA-Seq era. *Curr Opin Chem Biol.* 2013;17: 4-11
8.  Feng H., Qin Z. and Zhang X. Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. *Cancer Lett.* 2013;340: 179-191
9.  Davis H. P. and Squire L. R. Protein synthesis and memory: a review. *Psychol Bull.* 1984;96: 518-559
10. Holt C. E. and Schuman E. M. The central dogma decentralized: new perspectives on RNA function and local translation in neurons. *Neuron*. 2013;80: 648-657

# RNA-Seq: RNA Sequencing

RNA-Seq describes the abundance and sequence of RNA transcripts. The method, first published by several groups in 2008[11-15] has effectively displaced older methods such as serial analysis of gene expression (SAGE)[16] and massively parallel signature sequencing (MPSS).[17] With over 8000 references in PubMed, RNA-Seq is by far the most abundantly cited NGS method.[18] All RNA-Seq methods are based on the use of a reverse transcriptase to convert the RNA to cDNA. Two basic variations use either random primers or oligo(dT) primers for this reaction. Oligo(dT) primers are highly 3' biased and mostly suitable for mRNA abundance (expression) analysis. Random primers also result in some bias, which can be reduced by fragmentation of the input RNA.[19]
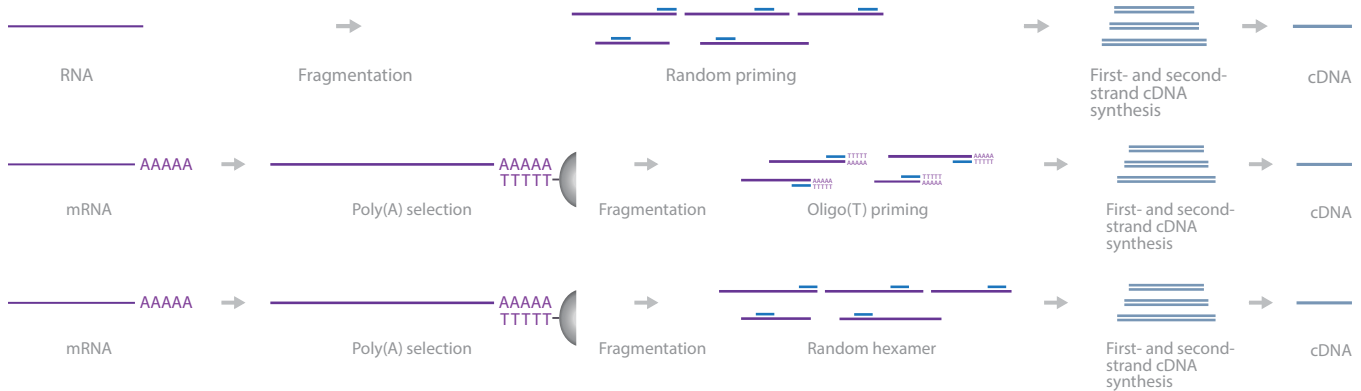


A schematic overview of RNA-Seq.

Additional refinements, such as the use of Moloney murine leukemia virus reverse transcriptase (MMLV-RT) and template-switching oligonucleotides produce a higher yield of full-length transcripts for gene annotation and splice-variant detection. These methods include those based on switch mechanism at the 5' end of RNA templates (Smart), such as Smart-Seq[20] and Smart-seq2;[21] cap-analysis gene expression (CAGE), such as NanoCAGE[22] and CAGEscan;[23] RNA-Seq from single nuclei (snRNA-Seq);[24] and sequencing of fixed and recovered intact single-cell RNA (FRISCR).[25]

| Advantages | Disadvantages |
|---|---|
| • Highly suitable for discovering novel exons, genes, and splice isoforms | • Primer bias[26] |
| • High dynamic range in expression analysis, compared to microarrays | • Reverse transcriptase may introduce sequencing errors |

11. Wilhelm B. T., Marguerat S., Watt S., et al. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature*. 2008;453: 1239-1243
12. Sultan M., Schulz M. H., Richard H., et al. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science*. 2008;321: 956-960
13. Marioni J. C., Mason C. E., Mane S. M., Stephens M. and Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res*. 2008;18: 1509-1517
14. Mortazavi A., Williams B. A., McCue K., Schaeffer L. and Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008;5: 621-628
15. Nagalakshmi U., Wang Z., Waern K., et al. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science*. 2008;320: 1344-1349
16. Velculescu V. E., Zhang L., Vogelstein B. and Kinzler K. W. Serial analysis of gene expression. *Science*. 1995;270: 484-487
17. Harbers M. and Carninci P. Tag-based approaches for transcriptome research and genome annotation. *Nat Methods*. 2005;2: 495-502
18. Based on a PubMed search for RNA-Seq on 11/11/2016
19. Mortazavi A., Williams B. A., McCue K., Schaeffer L. and Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008;5: 621-628
20. Ramskold D., Luo S., Wang Y. C., et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol*. 2012;30: 777-782
21. Picelli S., Bjorklund A. K., Faridani O. R., Sagasser S., Winberg G. and Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods*. 2013;10: 1096-1098
22. Plessy C., Bertin N., Takahashi H., et al. Linking promoters to functional transcripts in small samples with nanoCAGE and CAGEscan. *Nat Methods*. 2010;7: 528-534
23. Plessy C., Bertin N., Takahashi H., et al. Linking promoters to functional transcripts in small samples with nanoCAGE and CAGEscan. *Nat Methods*. 2010;7: 528-534
24. Grindberg R. V., Yee-Greenbaum J. L., McConnell M. J., et al. RNA-sequencing from single nuclei. *Proc Natl Acad Sci U S A*. 2013;110: 19802-19807
25. Thomsen E. R., Mich J. K., Yao Z., et al. Fixed single-cell transcriptomic characterization of human radial glial diversity. *Nat Methods*. 2016;13: 87-93
26. Hansen K. D., Brenner S. E. and Dudoit S. Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Res*. 2010;38: e131

## Reviews

Conesa A., Madrigal P., Tarazona S., Gomez-Cabrero D., Cervera A., et al. A survey of best practices for RNA-Seq data analysis. *Genome Biol.* 2016;17: 13

## References

Pareek C. S., Smoczynski R., Kadarmideen H. N., et al. Single Nucleotide Polymorphism Discovery in Bovine Pituitary Gland Using RNA-Seq Technology. *PLoS One*. 2016;11: e0161370

The authors used strand-specific RNA-Seq to detect single-nucleotide polymorphisms (SNPs) in the bovine pituitary gland transcriptome. The experiment was conducted on a variety of young, growing bull breeds, such as Polish Holstein-Friesian, Polish Red, and Hereford breeds, at various developmental ages. Using the vast data they gathered, the researchers were able to identify *KCNIP4, CCSER1, DPP6, MAP3K5,* and *GHR* as candidate genes with the highest SNPs hit loci across all 3 breeds and developmental ages.

Illumina Technology: NextSeq 500 System, TruSeq DNA Sample Prep Kit

Bennett C. G., Riemondy K., Chapnick D. A., et al. Genome-wide analysis of Musashi-2 targets reveals novel functions in governing epithelial cell migration. *Nucleic Acids Res.* 2016;44: 3788-3800

Musashi-2 (Msi2) is an RBP that maintains self-renewal of stem cells and promotes oncogenesis through modulation of cell-proliferation mechanisms in hematopoietic and gastrointestinal tissues. This study aimed at gaining a deeper understanding of the mechanisms underlying Msi2 and its oncogenic capabilities. Using RNA-Seq with high-throughput sequencing of RNA isolated by crosslinking immunoprecipition (HITS-CLIP), the researchers were able to identify Msi2 targets and also discovered that Msi2 primarily recognizes 3'-untranslated regions (UTRs) of mRNAs at sites with multiple UAG motifs. Msi2 also promotes specific mRNA decay with no effect on translational efficiency. The authors identified Msi2 targets as key regulators of cell motility, extracellular matrix-receptor interaction, cell growth, and cell survival.

Illumina Technology: HiSeq 2000 System, ARTseq/TruSeq Ribo Profile Kit, Ribo-Zero Gold rRNA Removal Kit

Suarez-Vega A., Gutierrez-Gil B., Klopp C., Tosser-Klopp G. and Arranz J. J. Comprehensive RNA-Seq profiling to evaluate lactating sheep mammary gland transcriptome. *Sci Data.* 2016;3: 160051

Davila J. I., Fadra N. M., Wang X., et al. Impact of RNA degradation on fusion detection by RNA-seq. *BMC Genomics.* 2016;17: 814

Chen M. J., Chen L. K., Lai Y. S., et al. Integrating RNA-Seq and ChIP-Seq data to characterize long non-coding RNAs in Drosophila melanogaster. *BMC Genomics*. 2016;17: 220

Arnold W. K., Savage C. R., Brissette C. A., Seshu J., Livny J. and Stevenson B. RNA-Seq of Borrelia burgdorferi in Multiple Phases of Growth Reveals Insights into the Dynamics of Gene Expression, Transcriptome Architecture, and Noncoding RNAs. *PLoS One.* 2016;11: e0164165

Joshi R. K., Megha S., Rahman M. H., Basu U. and Kav N. N. A global study of transcriptome dynamics in canola (Brassica napus L.) responsive to Sclerotinia sclerotiorum infection using RNA-Seq. *Gene.* 2016;590: 57-67

Seo M., Caetano-Anolles K., Rodriguez-Zas S., et al. Comprehensive identification of sexually dimorphic genes in diverse cattle tissues using RNA-seq. *BMC Genomics.* 2016;17: 81

Lukoszek R., Feist P. and Ignatova Z. Insights into the adaptive response of *Arabidopsis thaliana* to prolonged thermal stress by ribosomal profiling and RNA-Seq. *BMC Plant Biol.* 2016;16: 221

Zhang Q., Lai M. M., Lou Y. Y., Guo B. H., Wang H. Y. and Zheng X. Q. Transcriptome altered by latent human cytomegalovirus infection on THP-1 cells using RNA-seq. *Gene*. 2016;594: 144-150

Choi S. Y., Park B., Choi I. G., et al. Transcriptome landscape of Synechococcus elongatus PCC 7942 for nitrogen starvation responses using RNA-seq. *Sci Rep.* 2016;6: 30584

Zhang Z. X., Zhao S. N., Liu G. F., et al. Discovery of putative capsaicin biosynthetic genes by RNA-Seq and digital gene expression analysis of pepper. *Sci Rep.* 2016;6: 34121

Weissbein U., Schachter M., Egli D. and Benvenisty N. Analysis of chromosomal aberrations and recombination by allelic bias in RNA-Seq. *Nat Commun.* 2016;7: 12144

Chakraborty S., Britton M., Martinez-Garcia P. J. and Dandekar A. M. Deep RNA-Seq profile reveals biodiversity, plant-microbe interactions and a large family of NBS-LRR resistance genes in walnut (Juglans regia) tissues. *AMB Express.* 2016;6: 12

# CaptureSeq: RNA Capture Sequencing

CaptureSeq is a targeted RNA sequencing method that is able to provide higher sequencing coverage for selected regions of the genome.[27]

This method follows the TruSeq RNA sample preparation protocol, in which mRNA is first isolated from total RNA by poly(A) selection and then fragmented. Double-stranded cDNA copies of the fragments are generated using reverse transcriptase and then ligated to p5 and p7 adapters. Next, these cDNA library fragments are amplified by the polymerase chain reaction (PCR). To increase specificity, custom capture probes are hybridized to the cDNA and bound to an array while other transcripts are washed away prior to PCR amplification. This process leaves the targeted fragments that are ready for sequencing.



A schematic overview of CaptureSeq.

| Advantages | Disadvantages |
| --- | --- |
| • Highly suitable for discovering novel exons, genes, and splice isoforms | • Requires large amount of total RNA (5 µg) to yield 250 ng of amplified cDNA for capture<br>• Coverage accuracy drops for transcripts with repetitive sequences |

## Reviews

Marbaniang C. N. and Vogel J. Emerging roles of RNA modifications in bacteria. *Curr Opin Microbiol.* 2016;30: 50-57

Gloss B. S. and Dinger M. E. The specificity of long noncoding RNA expression. *Biochim Biophys Acta.* 2015;

Luciano D. J. and Belasco J. G. NAD in RNA: unconventional headgear. *Trends Biochem Sci.* 2015;40: 245-247

## References

Bussotti G., Leonardi T., Clark M. B., et al. Improved definition of the mouse transcriptome via targeted RNA sequencing. *Genome Res.* 2016;26: 705-716

The authors used CaptureSeq on mouse tissue samples to refine transcript annotations of the GRCm38 mouse reference genome. Targeted sequencing revealed novel coding and long noncoding RNA (lncRNA) transcripts, and improved annotations of splice junctions and alternative splice sites.

Illumina Technology: HiSeq System, TruSeq Stranded mRNA Library Prep Kit

Clark M. B., Mercer T. R., Bussotti G., et al. Quantitative gene profiling of long noncoding RNAs with targeted RNA sequencing. *Nat Methods.* 2015;

Mercer T. R., Gerhardt D. J., Dinger M. E., et al. Targeted RNA sequencing reveals the deep complexity of the human transcriptome. *Nat Biotechnol.* 2012;30: 99-104

## Associated Kits

TruSeq Exome Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

27.    Mercer T. R., Clark M. B., Crawford J., et al. Targeted sequencing for gene discovery and quantification using RNA CaptureSeq. *Nat Protoc.* 2014;9: 989-1009

# RASL-Seq: RNA-Mediated Oligonucleotide Annealing, Selection, and Ligation with Next-Generation Sequencing

RNA-mediated oligonucleotide annealing, selection, and ligation with next-generation sequencing (RASL-Seq) is a 2-dimensional RNA sequencing method to quantify expression profiles of several hundred genes, under thousands of different conditions.[28]

Custom probe pairs are designed for each gene of interest. A pair of probes needs to contain: 1) One probe containing a common index primer on its 3' end and a 20 nt oligonucleotide corresponding to the targeted exon sequence with a phosphate on the 5' end; and 2) another probe with a P5 adapter on its 5' end with a 20 nt sequence complementary to the exon that is adjacent to the other probe. The probe pairs are hybridized to the mRNA and separated from total RNA using oligo(dT)-biotin beads. A ligation step joins the probe pairs into a single PCR amplicon probe. The biotinylated mRNA strands are subsequently attached to streptavidin magnetic beads to elute the probe fragments. Next, P7 adapters are attached to the 3' index primer, and the library undergoes PCR amplification before sequencing. The library is sequenced from the 40 nt ligated P5 primer, followed by sequencing from the P7 primer oligonucleotide.
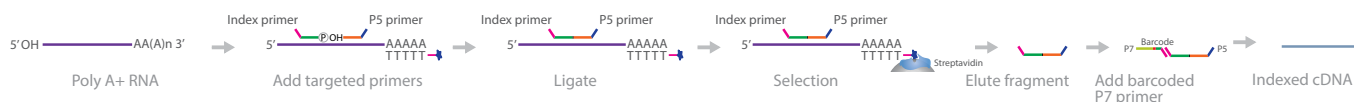


A schematic overview of RASL-seq.

| Advantages | Disadvantages |
|---|---|
| • Quantify genetic expression in large gene panels under thousands of different experimental conditions<br>• Effective on low total RNA amounts (10 ng for about 1000 cells)<br>• Can be performed on isolated RNA samples or cell lysates<br>• Can be performed manually or automatically using a custom robot | • Quantify genetic expression in large gene panels under thousands of different experimental conditions<br>• Effective on low total RNA amounts (10 ng for about 1000 cells)<br>• Can be performed on isolated RNA samples or cell lysates<br>• Can be performed manually or automatically using a custom robot |

## Reviews
Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38: 226-236

## References
Qiu J., Zhou B., Thol F., et al. Distinct splicing signatures affect converged pathways in myelodysplastic syndrome patients carrying mutations in different splicing regulators. *RNA.* 2016;

The authors used RASL-Seq to identify pre-mRNA alternative splicing events from human bone marrow and blood samples with myelodysplastic syndrome (MDS). They found that certain splicing signatures in genes involved in cell cycle regulation and DNA damage remediation were correlated to MDS samples. Alternative splicing patterns can be used as biomarkers for rapid MDS identification.

Illumina Technology: HiSeq 2500 System

Kralovicova J., Knut M., Cross N. C. and Vorechovsky I. Identification of U2AF(35)-dependent exons by RNA-Seq reveals a link between 3' splice-site organization and activity of U2AF-related proteins. *Nucleic Acids Res.* 2015;

Shao C., Yang B., Wu T., et al. Mechanisms for U2AF to define 3' splice sites and regulate alternative splicing in the human genome. *Nat Struct Mol Biol.* 2014;21: 997-1005

Zhou Z., Qiu J., Liu W., et al. The Akt-SRPK-SR axis constitutes a major pathway in transducing EGF signaling to regulate alternative splicing in the nucleus. *Mol Cell.* 2012;47: 422-433

## Associated Kits
TruSeq RNA Library Prep Kit v2

TruSeq Stranded mRNA Library Prep Kit

---

28.  Li H., Qiu J. and Fu X. D. RASL-Seq for massively parallel and quantitative analysis of gene expression. *Curr Protoc Mol Biol.* 2012;Chapter 4: Unit 4.13.11-19

# ClickSeq: RNA-Seq Libraries from Stochastically Terminated 3'-azido-blocked cDNA Fragments

ClickSeq is an RNA sequencing technique that uses bioconjugation as an alternative to fragmentation in the library preparation step, to produce lower error rates than standard sequencing methods.[29]

First, RNA is reverse-transcribed into cDNA, in a process similar to Sanger sequencing, with 3'-azido-2',3'dideoxynucleotides (AzNTPs). This process induces chain termination and semirandom primers (6 random nucleotides followed by an Illumina 3' P7 adapter) to sequence random positions on the RNA template. Single-stranded cDNA is purified, and the 3'-ends are click-ligated with alkyne-modified (5'hexynyl) P5 adapters. After PCR amplification, the cDNA library is ready to be sequenced.



A schematic overview of ClickSeq.

| Advantages | Disadvantages |
|---|---|
| • Significantly reduced artifactual recombination rate due to elimination of the fragmentation step<br>• Highly suitable for detecting rare recombination events<br>• No fragmentation step | • 3'-azido blocked cDNA fragments are converted into double-stranded DNA with low efficiency<br>• Read-through of AzNTP is still highly inefficient<br>• Further optimization required for single-cell or single-molecule studies |

## Reviews
None available yet

## References
Routh A., Head S. R., Ordoukhanian P. and Johnson J. E. ClickSeq: Fragmentation-Free Next-Generation Sequencing via Click Ligation of Adaptors to Stochastically Terminated 3'-Azido cDNAs. *J Mol Biol.* 2015;427: 2610-2616

This study compares sequencing data obtained using ClickSeq to that obtained from a standard RNA-Seq library preparation kit. The authors used Flock house virus RNA was used to generate cDNA libraries for the comparison. Alignment frequencies and read coverage of ClickSeq sequence reads were similar to standard RNA-Seq reads. However, the absence of RNA fragmentation in ClickSeq resulted in a dramatic reduction of inter-RNA recombination events by over 99%.

Illumina Technology: HiSeq System

## Associated Kits
TruSeq RNA Library Prep Kit v2

TruSeq Stranded mRNA Library Prep Kit

29.    Routh A., Head S. R., Ordoukhanian P. and Johnson J. E. ClickSeq: Fragmentation-Free Next-Generation Sequencing via Click Ligation of Adaptors to Stochastically Terminated 3'-Azido cDNAs. *J Mol Biol.* 2015;427: 2610-2616

# 3Seq: 3'-End Sequencing for Expression Quantification

3'-end Sequencing for Expression Quantification (3Seq) is able to isolate highly degraded mRNA from formalin-fixed paraffin-embedded (FFPE) tissue samples for quantitative genome-wide expression analysis.[30]

First, mRNA is isolated from total RNA by poly(A) selection. Next, an oligo(dT)-P7 RT primer is annealed to the 3' end of the mRNA fragment to synthesize the first cDNA strand via reverse transcription (RT). A second cDNA strand is generated from the first strand, and P5 adapters are ligated on the opposite end from the P7 adapters. The fragments containing the adapters are amplified by PCR and are ready for sequencing. 3Seq can also be used for fresh-frozen tissue samples.
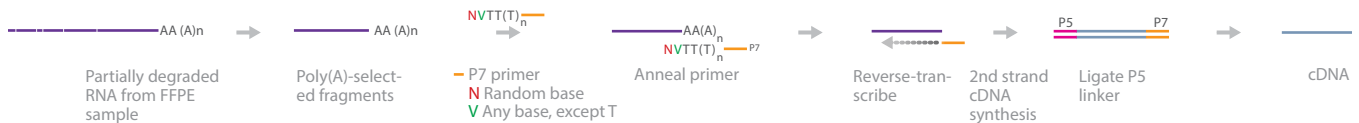


A schematic overview of 3Seq.

| Advantages | Disadvantages |
|---|---|
| • Sequences highly degraded mRNA from FFPE or fresh-frozen tissue samples | • High internal priming rate<br>• Homopolymeric nature of poly(A) tail often causes polymerase slippage |

## Reviews
Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci*. 2015;38: 226-236

## References
Riemondy K., Wang X. J., Torchia E. C., Roop D. R. and Yi R. MicroRNA-203 represses selection and expansion of oncogenic Hras transformed tumor initiating cells. *Elife*. 2015;4:

The authors studied the ability of microRNA (miRNA)-203 to repress tumorigenesis induced by the HRASG12V mutation. They used 3Seq because it is able to quantify mRNA expression levels and identify alternative 3'UTR isoforms. They performed 3Seq on murine primary keratinocytes to investigate changes in the transcriptome upon HRASG12V transduction and miR-203 induction or knockout. The results revealed the importance of miR-203 in restricting clonal selection and Hras-mutated tumorigenesis. The authors also used high-throughput sequencing of crosslinking and immunoprecipitation (CLIP) cDNA library (HITS-CLIP) and microRNA sequencing (miRNA-Seq) to enhance their understanding of miR-203 targets.

Illumina Technology: HiSeq 2000 System, Ribo-Zero Gold rRNA Removal Kit

Wei G., Luo H., Sun Y., et al. Transcriptome profiling of esophageal squamous cell carcinoma reveals a long noncoding RNA acting as a tumor suppressor. *Oncotarget*. 2015;6: 17065-17080

Guo X., Forgo E. and van de Rijn M. Molecular subtyping of leiomyosarcoma with 3' end RNA sequencing. *Genom Data*. 2015;5: 366-367

## Associated Kits
TruSeq RNA Access Library Prep Kit

---

30.    Beck A. H., Weng Z., Witten D. M., et al. 3'-end sequencing for expression quantification (3SEQ) from archival tumor samples. *PLoS One*. 2010;5: e87

# cP-RNA-Seq: 2',3'-cyclic phosphate (cP) RNA Sequencing

2',3'-cyclic phosphate (cP) RNA sequencing (cP-RNA-Seq) is a method to isolate and sequence RNA species protected by cP at their 3' terminus, which usually prevents adapter ligation.[31] The researchers originally developed this technique to identify transfer RNA (tRNA) species that generate sex hormone–dependent tRNA-derived RNAs (SHOT-RNAs) in human breast cancer cells, but it can be used for other RNA species protected by cP at the 3' end.[32]

Total RNA from samples is isolated and gel-purified to the desired length. This mixture contains RNA species with 3' hydroxy, 3'phosphate, and 3'-cP. Phosphatase treatment removes phosphates from the 5' and 3' ends. The RNA is treated subsequently with periodate to cleave the 3'-hydroxyl ends into 2'3'-dialdehydes, while leaving the 3'-cP ends intact. RNA strands with 2'3'-dialdehydes at their 3' terminus are inert to adapter ligation. Strands with 3'-cP are cleaved with T4 polynucleotide kinase (PNK) and ligated to sequencing adapters. After flanking both ends with adapters, the RNA strands are reverse-transcribed, amplified, purified, and sequenced.
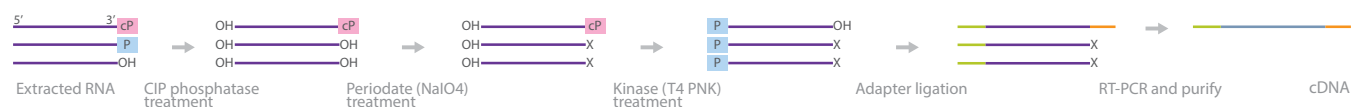


A schematic overview of cP-RNA-Seq.

| Advantages | Disadvantages |
| --- | --- |
| • Selective isolation, amplification, and sequencing of RNA species protected by cP<br>• Considerable improvement in efficiency and specificity compared to an earlier method using recombinant tRNA ligase and multiple gel purifications[33] | • False positives will arise when sequencing RNA species with 2'-O-methylribose modifications—such as plant miRNAs, plant and animal short interfering RNAs (siRNAs), and animal PIWI-interacting RNAs (piRNAs)—due to periodate cleavage specificity[34] |

## Reviews
None available yet

## References
Honda S., Loher P., Shigematsu M., Palazzo J. P., Suzuki R., et al. Sex hormone-dependent tRNA halves enhance cell proliferation in breast and prostate cancers. *Proc Natl Acad Sci U S A*. 2015;112: E3816-3825

This study uncovered a novel type of tRNA, termed SHOT-RNA, that is highly expressed in estrogen receptor–positive breast cancer and androgen receptor–positive prostate cancer cell lines. The authors used cP-RNA-Seq to isolate, amplify, and sequence SHOT-RNAs. They discovered numerous 5'-SHOT-RNAs and found that 5'-SHOT-RNAs, but not their 3' counterparts, have significant involvement in cell proliferation. Their results indicate that SHOT-RNAs can serve as potential biomarkers for breast and prostate cancers.

Illumina Technology: HiSeq 2000 System, TruSeq Small RNA Library Preparation Kit

## Associated Kits
TruSeq Small RNA Library Prep Kit

Ribo-Zero rRNA Removal Kits

31.  Honda S., Morichika K. and Kirino Y. Selective amplification and sequencing of cyclic phosphate-containing RNAs by the cP-RNA-Seq method. *Nat Protoc.* 2016;11: 476-489
32.  Honda S., Loher P., Shigematsu M., et al. Sex hormone-dependent tRNA halves enhance cell proliferation in breast and prostate cancers. *Proc Natl Acad Sci U S A.* 2015;112: E3816-3825
33.  Honda S., Morichika K. and Kirino Y. Selective amplification and sequencing of cyclic phosphate-containing RNAs by the cP-RNA-Seq method. *Nat Protoc.* 2016;11: 476-489
34.  Honda S., Morichika K. and Kirino Y. Selective amplification and sequencing of cyclic phosphate-containing RNAs by the cP-RNA-Seq method. *Nat Protoc.* 2016;11: 476-489

# 3P-Seq: Poly(A)-Position Profiling by Sequencing

Poly(A)-position profiling by sequencing (3P-Seq) is used to identify 3'-UTRs in mRNA.[35]

Poly(A) selection is used to isolate mRNA from total RNA, and biotinylated-splint primers are annealed and splint-ligated to the end of the mRNA poly(A) tail. The RNA-primer complex is partially digested by RNase T1, bound to streptavidin, and washed to purify the 3' fragments. Primers corresponding to the 3'end of the poly(A) tail are annealed and reverse-transcribed with deoxythymidine triphosphate (dTTP) as the only dNTP to generate a cDNA strand complementary to the poly(A) tail. The biotin-bound poly(A) RNA fragments are released by RNase H digestion and purified. Next, P7 and P5 adapters are attached to the RNA fragments, and they are reverse-transcribed to generate cDNA fragments.
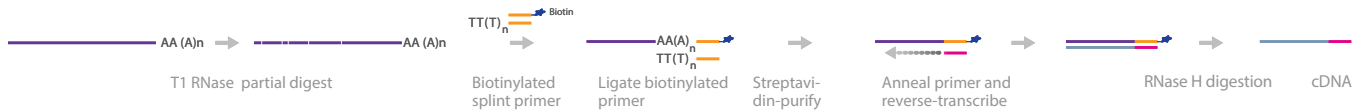


A schematic overview of 3P-Seq.

| Advantages | Disadvantages |
|---|---|
| • Reliable for UTR isoform discoveries<br>• Prevents polymerase slippage by using splint-ligation primer in poly(A) capture<br>• Prevents internal priming and is specific to 3' ends of poly(A) RNAs | • Requires large amounts of RNA<br>• Technically challenging to perform |

## Reviews

Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38: 226-236

## References

Lakshmanan V., Bansal D., Kulkarni J., et al. Genome-Wide Analysis of Polyadenylation Events in Schmidtea mediterranea. *G3 (Bethesda).* 2016;

The authors profiled post-transcriptional regulation mediated by 3'UTRs in Schmidtea mediterranea, a freshwater planarian species. They used 3P-Seq to identify 31,377 polyadenylation sites and annotate 3'UTRs in the protein-coding genes. The study also highlighted the prevalence of alternate polyadenylation in planarian gene expression.

Illumina Technology: HiSeq 1000 System, Genome Analyzer System, TruSeq Small RNA Library Prep Kit

Hezroni H., Koppstein D., Schwartz M. G., Avrutin A., Bartel D. P. and Ulitsky I. Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. *Cell Rep.* 2015;11: 1110-1122

Jambor H., Surendranath V., Kalinka A. T., Mejstrik P., Saalfeld S. and Tomancak P. Systematic imaging reveals features and changing localization of mRNAs in Drosophila development. *Elife.* 2015;4:

Spies N., Burge C. B. and Bartel D. P. 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. *Genome Res.* 2013;23: 2078-2090

## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Stranded mRNA Library Prep Kit

35.    Jan C. H., Friedman R. C., Ruby J. G. and Bartel D. P. Formation, regulation and evolution of Caenorhabditis elegans 3'UTRs. *Nature*. 2011;469: 97-101

# 2P-Seq: Poly(A)-Tail-Primed Sequencing

Poly(A)-tail-primed sequencing (2P-Seq) is designed to quantify mRNA stability and translational efficiency by characterizing 3'UTR isoforms.[36]

First, poly(A)+ mRNAs are isolated from total RNA and partially digested by RNase T1 to cleave the non-UTR regions. Poly (A)+ mRNA fragments are reverse-transcribed with a 20T-VN primer, which contains a stretch of 20 Ts and a VN anchor (N represents a fully degenerate base, and V is any base except for T). The resultant cDNAs are circularized and PCR-amplified using barcoded primers. The cDNA is sequenced with a custom primer ending with 20 Ts. To avoid 2P tags enriched for adenosine, analyses should be restricted to 2P tags mapping to within 20 bases of 3P-Seq–annotated poly(A) sites.
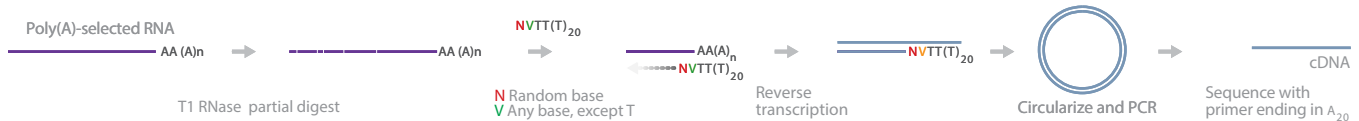


A schematic overview of 2P-Seq.

| Advantages | Disadvantages |
| --- | --- |
| • Quantifies mRNA stability and translational efficiency<br>• Accurately measures half-lives and length of each UTR isoform | • 3P-Seq datasets are needed to limit results specifically to poly(A) sites<br>• Poly(A) primer can misprime with A-rich regions elsewhere in the mRNA |

## Reviews
Kim M., You B.-H. and Nam J.-W. Global estimation of the 3' untranslated region landscape using RNA sequencing. *Methods*. 2015;83: 111-117

## References
Spies N., Burge C. B. and Bartel D. P. 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. *Genome Res.* 2013;23: 2078-2090

The authors quantified the effects of alternate 3'UTRs on mRNA stability and translation performance in murine 3T3 cells. They used 3P-Seq to isolate poly(A)+ RNA strands from total RNA to create accurate poly(A) site annotations before determining mRNA half-life and translational efficiency via 2P-Seq. Their 2P-Seq data showed that 3'UTR isoforms have minute effects on mRNA stability and translation rate.

Illumina Technology: Unspecified Illumina sequencing system

## Associated Kits
TruSeq RNA Library Prep Kit v2

TruSeq Stranded mRNA Library Prep Kit

---

36. Spies N., Burge C. B. and Bartel D. P. 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. *Genome Res*. 2013;23: 2078-2090

# 3'-Seq: Quantitatively Measure Abundance of 3'UTR Isoforms

3'-Seq was designed to measure the abundance of 3'-UTR isoforms quantitatively in a wide array of human tissue types.[37]

The first cDNA strand is generated by reverse-transcribing total RNA using an oligo(dT) primer containing a VN-anchor (where V is dA, dC, or dG), P5 adapter sequence, a uridine, and biotin that is bound to a streptavidin-coated magnetic bead. The second cDNA strand is synthesized using DNA polymerase I. To initiate nick translation with DNA polymerase I, RNase HII is used to introduce a nick at the uridine. This creates another nick that is 50–75 bases away from the 3' end. A blunt end is created at the position of the new nick, and double-stranded P7 adapters are ligated at this position. The double-stranded cDNA is PCR-amplified, purified, and ready for sequencing.



A schematic overview of 3'-seq.

| Advantages | Disadvantages |
|---|---|
| • Measures 3'UTR isoform abundance to detect protein expression differences in multiple tissue types | • Technically challenging; nick translation requires precise conditions |

## Reviews

Bangru S. and Kalsotra A. Advances in analyzing RNA diversity in eukaryotic transcriptomes: peering through the Omics lens. *F1000Research*. 2016;5: 2668

Miura P., Sanfilippo P., Shenker S. and Lai E. C. Alternative polyadenylation in the nervous system: To what lengths will 3' UTR extensions take us? *BioEssays*. 2014;36: 766-777

## References

**Berkovits B. D. and Mayr C. Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. *Nature*. 2015;522: 363-367**

This study explores the role of alternative polyadenylation in regulating localization of membrane proteins in human cell lines. The authors used 3'-Seq to quantify the abundance of 3'UTR mRNA isoforms of CD47, CD44, and TNFRSF13C in naïve B cells, B-lymphoblastoid cells, and HEK293 cells. They showed that 3'UTR isoform length changes protein function in surface-membrane proteins.

**Illumina Technology: HiSeq System, TruSeq RNA Library Prep Kit**

## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Stranded mRNA Library Prep Kit

---

37.  Lianoglou S., Garg V., Yang J. L., Leslie C. S. and Mayr C. Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes Dev.* 2013;27: 2380-2396

# TIF-Seq: Transcript Isoform Sequencing

Transcript isoform sequencing (TIF-Seq) identifies transcript isoforms by selective sequencing of full-length mRNA molecules with 5'caps and poly(A) tails.[38]

Capped mRNA molecules are selected by substituting the 5'caps with oligonucleotides. To achieve this result, 5'-phosphate groups are removed from non-capped RNAs to differentiate them from their capped counterparts. The caps are removed by tobacco acid pyrophosphatase (TAP) treatment to expose the 5'-phosphate groups for ligation with oligonucleotides. mRNAs are separated into 2 different tubes and reverse-transcribed to generate full-length cDNA (flcDNA). flcDNA in each tube is annealed to barcoded 5'-biotinylated primers and 3' primers. The primers contain barcode sequences unique to each tube, as a control mechanism against chimeric fragments and intermolecular ligation. The 2 tubes are combined, digested with NotI enzyme to produce sticky ends, and ligated to form circular double-stranded cDNA, which is fragmented subsequently. Fragments containing the biotinylated 3' and 5' ends are isolated with streptavidin. Multiplexing barcodes are added to the 3' and 5' ends of the purified cDNA fragments to create a cDNA library for sequencing.



A schematic overview of TIF-Seq.

| Advantages | Disadvantages |
| --- | --- |
| • Identifies transcript isoforms by sequencing the 5' and 3' end of the same RNA strand<br>• Chimera-control barcodes filter out intermolecular ligation of cDNA fragments | • Requires a large amount of full-length RNA in the sample<br>• Use of NotI enzyme to introduce sticky ends is only viable for AT-rich genomes like yeast<br>• Strong bias toward short RNA strands[39] |

## Reviews

Anamika K., Verma S., Jere A. and Desai A. Transcriptomic Profiling Using Next Generation Sequencing - Advances, Advantages, and Challenges. 2016;

Bagchi D. N. and Iyer V. R. The Determinants of Directionality in Transcriptional Initiation. *Trends Genet*. 2016;32:322-333

Solé C., Nadal-Ribelles M., de Nadal E. and Posas F. A novel role for lncRNAs in cell cycle control during stress adaptation. *Current Genetics.* 2015;61:299-308

## References

Pelechano V., Wei W. and Steinmetz L. M. Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature*. 2013;497:127-131

The authors highlight the depth of transcript isoform heterogeneity in yeast (Saccharomyces cerevisiae) using TIF-Seq. Data comparison showed that TIF-Seq holds an advantage over 3'end or 5'start site capture techniques in identifying short, overlapping 3'end transcripts, as well as distinguishing monocistronic from bicistronic transcripts.

Illumina Technology: HiSeq 2000 System

Pelechano V., Wei W., Jakob P. and Steinmetz L. M. Genome-wide identification of transcript start and end sites by transcript isoform sequencing. *Nat Protoc*. 2014;9:1740-1759

## Associated Kits

TruSeq RNA Library Prep Kit v2

38.    Pelechano V., Wei W. and Steinmetz L. M. Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature*. 2013;497:127-131
39.    Pelechano V., Wei W., Jakob P. and Steinmetz L. M. Genome-wide identification of transcript start and end sites by transcript isoform sequencing. *Nat Protoc*. 2014;9:1740-1759

# PEAT: Paired-end Analysis of Transcription Start Sites

PEAT characterizes transcription start sites (TSS) in mRNA using a technique similar to TIF-Seq.[40]

First, poly(A)+ RNAs are enriched from total RNA and the caps are removed with TAP. The 5' ends of uncapped mRNAs are ligated to chimeric linkers containing MmeI restriction endonuclease sites prior to RT. The RT primers also contain an MmeI site, resulting in single-stranded cDNA flanked by MmeI sites. The fragments are PCR-amplified and circularized into circular single-stranded cDNA, which is amplified further by rolling-circle amplification. MmeI is used to cut circular cDNA at the 2 MmeI sites to create linear, double-stranded cDNA fragments that are 93–95 bp long. The fragments are ligated to paired-end adapters, amplified, and sequenced.
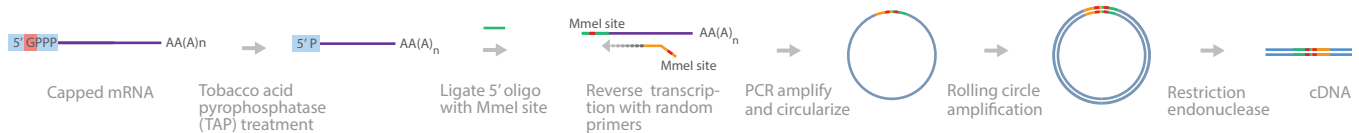


A schematic overview of PEAT.

| Advantages | Disadvantages |
|---|---|
| • Maps transcription initiation patterns using paired-end sequencing<br>• Improved accuracy and alignment yield compared to older, single-end TSS mapping strategies | • Does not distinguish between capped and noncapped RNA as in TIF-Seq<br>• Not designed to sequence 3'-UTRs as in TIF-Seq<br>• May produce chimeric fragments |

## Reviews
Megraw M., Cumbie J. S., Ivanchenko M. G. and Filichkin S. A. Small Genetic Circuits and MicroRNAs: Big Players in Polymerase II Transcriptional Control in Plants. *The Plant Cell*. 2016;28:286-303

Murakawa Y., Yoshihara M., Kawaji H., et al. Enhanced Identification of Transcriptional Enhancers Provides Mechanistic Insights into Diseases. *Trends Genet*. 2016;32:76-88

## References
Morton T., Petricka J., Corcoran D. L., et al. Paired-end analysis of transcription start sites in Arabidopsis reveals plant-specific promoter signatures. *Plant Cell*. 2014;26:2746-2760

This study focused on building a robust data set for TSS in plants. The authors used PEAT to map TSS locations in root samples of Arabidopsis thaliana because of the technique's high resolution, accuracy, and paired-end mapping ability. The PEAT TSS data set provided a significant increase in resolution over annotation data from The Arabidopsis Information Resource release 10 (TAIR10). The authors also developed a custom TSS peak-prediction model, called 3PEAT, for TSS prediction directly from the sequence.

Illumina Technology: HiSeq 2000 System

## Associated Kits
TruSeq RNA Library Prep Kit v2

TruSeq Stranded mRNA Library Prep Kit

40. Ni T., Corcoran D. L., Rach E. A., et al. A paired-end sequencing strategy to map the complex landscape of transcription initiation. *Nat Methods*. 2010;7:521-527

# SMORE-Seq: Simultaneous Mapping of RNA Ends with Sequencing

SMORE-Seq is an RNA-Seq method to identify TSS and polyadenylation sites (PAS) by sequencing the 5' and 3' ends simultaneously.[41]

Poly(A)+ RNA is first isolated from total RNA, and the 5' caps are removed with TAP. The de-capped 5' ends of the mRNA are ligated to P5 adapters. Next, the mRNA is fragmented before ligating P7 adapters to the 3' ends of the fragments. The mRNA fragments are reverse-transcribed, PCR amplified, and size-selected to ~250 bp fragments. The selected fragments are PCR-amplified and ready for sequencing.
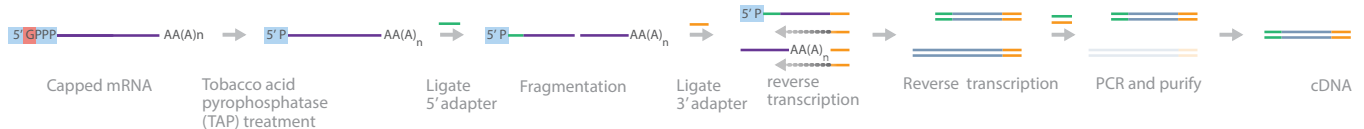


A schematic overview of SMORE-seq.

| Advantages | Disadvantages |
| --- | --- |
| • Identifies both TSS and PAS from the same RNA library dataset<br>• More accurate in identifying TSS regions than RNA-Seq due to 5' de-capping with TAP<br>• Mapping PAS using degradation intermediates is possible because a phosphatase treatment before TAP is omitted | • Full-length mRNA samples are preferred, which can be rare in highly degraded samples |

### Reviews
Bagchi D. N. and Iyer V. R. The Determinants of Directionality in Transcriptional Initiation. Trends in Genetics. 2016;32:322-333

### References
Park D., Morris A. R., Battenhouse A. and Iyer V. R. Simultaneous mapping of transcript ends at single-nucleotide resolution and identification of widespread promoter-associated non-coding RNA governed by TATA elements. *Nucleic Acids Res.* 2014;42:3736-3749

This study focused on understanding the effect of TSS and PAS on gene regulation factors. The authors used SMORE-Seq in yeast RNA to identify TSS and PAS simultaneously, with single-nucleotide resolution, from the same RNA library. The resulting data revealed mis-annotation of TSS in at least 150 genes from the Saccharomyces Genome Database. Unlike TIF-Seq that maps TSS and PAS from the same mRNA strand, SMORE-Seq identifies them separately but from the same RNA population. This method enables annotation of the most prominent TSS and PAS from each gene..

Illumina Technology: HiSeq 2000 System

### Associated Kits
TruSeq Small RNA Library Prep Kit

---

41.  Park D., Morris A. R., Battenhouse A. and Iyer V. R. Simultaneous mapping of transcript ends at single-nucleotide resolution and identification of widespread promoter-associated non-coding RNA governed by TATA elements. *Nucleic Acids Res.* 2014;42:3736-3749

# TL-Seq: Transcript Leader Sequencing

TL-Seq targets and enriches for the sequence around the 5'-UTRs of 5'-capped mRNA molecules before sequencing.[42]

Poly(A)+ RNA is fragmented and selected for 50–80 nt fragments. RNA fragments with phosphorylated 5' ends are dephosphorylated, using calf intestinal phosphatase (CIP), to distinguish them from capped mRNA fragments. Next, TAP strips the capped mRNAs and exposes the phosphate on the 5' end for P5 adapter ligation. The adapter-ligated fragments are gel-purified by molecular weight before ligation of P7 adapters. The 3'-end adapters can be attached by poly(A) tailing and ligation, or directly if using preadenylated adapters. After RT and PCR amplification, the RNA library is sequenced.
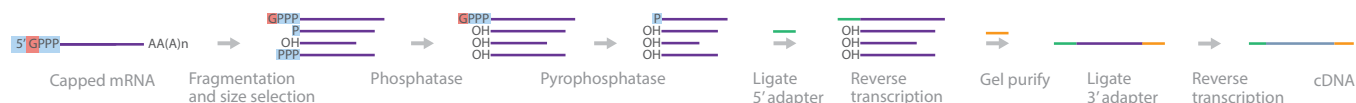


A schematic overview of TL-seq.

| Advantages | Disadvantages |
| --- | --- |
| • Sequences 5' UTRs and identifies variants<br>• Able to associate transcript leader function in translation when combined with translation-associated TL-Seq (TATL-Seq) | • Selects short nucleotide fragments (50–80 nt)<br>• Labor-intensive; requires large amounts of starting material |

## Reviews

Bangru S. and Kalsotra A. Advances in analyzing RNA diversity in eukaryotic transcriptomes: peering through the Omics lens. *F1000Research*. 2016;5:2668

Smith J. E. and Baker K. E. Nonsense-mediated RNA decay – a switch and dial for regulating gene expression. *BioEssays*. 2015;37:612-623

## References

Arribere J. A. and Gilbert W. V. Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. *Genome Res.* 2013;23:977-987

This study explored the role of 5'-UTRs in regulating post-transcriptional gene expression. Using TL-Seq on yeast cells, the authors mapped TSS on the majority of yeast genes. Their sequencing data also highlighted the role of short 5'-UTRs (less than 12 nt) in activating the nonsense-mediated mRNA decay pathway. The authors also developed TATL-Seq to investigate the effect of 5'-UTRs on translational efficiency. They discovered that short 5'-UTR isoforms in the CRZ1 gene have significantly higher translational efficiency than longer isoforms.

Illumina Technology:  Genome Analyzer System, HiSeq System

## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Stranded mRNA Library Prep Kit

---

42.    Arribere J. A. and Gilbert W. V. Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. *Genome Res*. 2013;23:977-987

# TATL-Seq: Translation-Associated Transcript Leader Sequencing

TATL-Seq works in conjunction with TL-Seq to target the sequence around the 5'-UTRs of mRNA attached to polysomes.[43] RNA from polysome gradient fractions is extracted, poly(A)-selected, and fragmented. Next, the TL-Seq protocol is followed to isolate and sequence the 5'-UTRs.
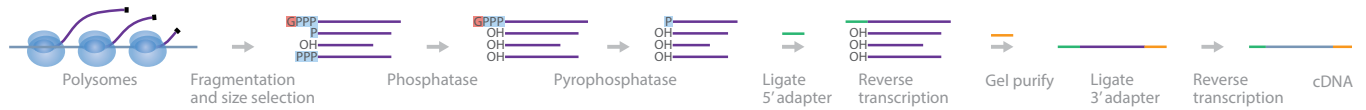


A schematic overview of TATL-seq.

| Advantages | Disadvantages |
| --- | --- |
| • Sequences 5'-UTRs and identifies variants<br>• TATL-Seq enables *de novo* transcript leader annotation while simultaneously testing their translational activity in a single experiment | • Selects short nucleotide fragments (50–80 nt)<br>• Labor-intensive; requires large amounts of starting material |

## Reviews

Bangru S. and Kalsotra A. Advances in analyzing RNA diversity in eukaryotic transcriptomes: peering through the Omics lens. *F1000Research*. 2016;5:2668

Smith J. E. and Baker K. E. Nonsense-mediated RNA decay – a switch and dial for regulating gene expression. *BioEssays*. 2015;37:612-623

## References

Arribere J. A. and Gilbert W. V. Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. *Genome Res.* 2013;23:977-987

This study explored the role of 5'-UTRs in regulating post-transcriptional gene expression. Using TL-Seq on yeast cells, the authors mapped TSS on the majority of yeast genes. Their sequencing data also highlighted the role of short 5'-UTRs (less than 12 nt) in activating the nonsense-mediated mRNA decay pathway. The authors also developed TATL-Seq to investigate the effect of 5'-UTRs on translational efficiency. They discovered that short 5'-UTR isoforms in the CRZ1 gene have significantly higher translational efficiency than longer isoforms.

Illumina Technology:  Genome Analyzer System, HiSeq System

## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Stranded mRNA Library Prep Kit

43.    Arribere J. A. and Gilbert W. V. Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. *Genome Res.* 2013;23:977-987

# RARseq: Restriction Site Associated RNA Sequencing

RARseq is a cDNA-based, genotype-by-sequencing method for identifying RNA SNPs and polymorphic loci for population genomics.[44] This method performs read alignment using sequencing data from single and dual restriction enzyme digestion libraries and aligns them using both *de novo* and reference-based genome assembly.

In RARseq, total RNA is isolated from samples. Next, first- and second-strand cDNA are generated by RT. The double-stranded cDNA is digested with selected restriction enzymes (MseI and MseI-StyI-HF). The double-stranded cDNA fragments are ligated to sequencing adapters before purification and size-selection to 200 bp. The samples are PCR-amplified, purified, and sequenced. Reads from both single- and double-digestion libraries are used for *de novo* and reference-based genome assembly.



A schematic overview of RARseq.

| Advantages | Disadvantages |
|---|---|
| • Identifies RNA SNPs and polymorphic loci for population genomics studies<br>• Reduces reads from nongenic regions arising from highly methylated genomes<br>• Uses both *de novo* and reference-based genome assembly for read alignment<br>• More accurate SNP and allele discovery than genotyping by sequencing methods<br>• Restriction enzyme selection is flexible and can be customized, depending on the genome | • Not yet adopted widely by the scientific community<br>• Has only been performed on plant genomes |

## Reviews
None available yet

## References
Alabady M. S., Rogers W. L. and Malmberg R. L. Development of Transcriptomic Markers for Population Analysis Using Restriction Site Associated RNA Sequencing (RARseq). *PLoS One.* 2015;10:e0134855

The authors developed RARseq as an improvement in identifying SNPs and loci polymorphisms, compared to genotype by sequencing (GBS) and standard RNA-Seq methods. RARseq offers an advantage over standard GBS methods by providing higher SNP and allelic discovery, due to its cDNA-based genotyping process. The authors performed RARseq to assemble F2 pitcher plant genomes using MseI and MseI-StyI-HF restriction enzymes. Utilizing both *de novo* and reference genome assembly, they found over 500 and 1800 SNPs per individual, respectively.

Illumina Technology: MiSeq System

## Associated Kits
TruSeq RNA Library Prep Kit v2

---

44.	Alabady M. S., Rogers W. L. and Malmberg R. L. Development of Transcriptomic Markers for Population Analysis Using Restriction Site Associated RNA Sequencing (RARseq). *PLoS One*. 2015;10:e0134855

# TAIL-Seq: Poly(A) Tail Sequencing

TAIL-Seq focuses on sequencing the very ends of mRNA molecules (3'-UTRs and poly(A) tail regions) to explore their role in mRNA half-life, stability, their impact on translational efficiency, and to discover other aspects surrounding 3'-terminome function.[45]

Ribosomal RNA (rRNA) is first removed from total RNA by affinity-based depletion. After purification, mRNAs are ligated to biotinylated 3' adapters prior to RNase T1 fragmentation and purified further by streptavidin pull-down. The 5' ends are phosphorylated and size-selected for 500–1000 nt fragments to prevent short noncoding RNA (ncRNA) fragments from contaminating the sequence data. The phosphorylated 5' ends are ligated to 5' adapters, reverse-transcribed, PCR-amplified, and sequenced. TAIL-Seq uses a unique paired-end run system to correlate the genes corresponding to the 3'-end sequence reads. By separating the paired-end reads, read 1 sequences 52 nt from the 5' end of the fragment to map the genome and identify transcripts, while read 2 sequences 251 nt from the 3' end specifically for sequence determination.
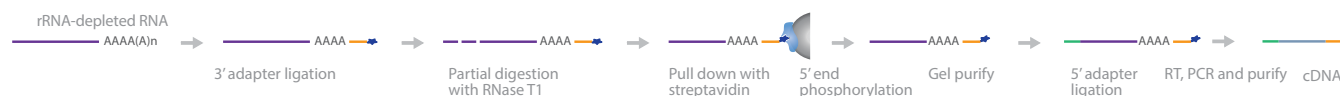


A schematic overview of TAIL-seq.

| Advantages | Disadvantages |
| --- | --- |
| • Quantifies poly(A) tail length on mRNA samples with high accuracy using a special fluorescence analysis method<br>• Eliminates bias toward long poly(A) because it does not use oligo(dT) enrichment<br>• Able to detect modified 3' ends, unlike poly(A)-tail length profiling by sequencing (PAL-Seq)[46] | • PCR amplification is unfavorable for homopolymeric sequences |

## Reviews

Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38:226-236

Hrdlickova R., Toloue M. and Tian B. RNA-Seq methods for transcriptome analysis. *Wiley Interdisciplinary Reviews: RNA.* 2016;n/a-n/a

Bangru S. and Kalsotra A. Advances in analyzing RNA diversity in eukaryotic transcriptomes: peering through the Omics lens. *F1000Research.* 2016;5:2668

Viegas S. C., Silva I. J., Apura P., Matos R. G. and Arraiano C. M. Surprises in the 3' end: 'U' can decide too! *FEBS Journal.* 2015;282:3489-3499

## References

Zuber H., Scheer H., Ferrier E., et al. Uridylation and PABP Cooperate to Repair mRNA Deadenylated Ends in Arabidopsis. *Cell Rep.* 2016;14:2707-2717

This study sought to enhance understanding of the complex process of uridylation in eukaryotes. The researchers used TAIL-Seq on wild-type and mutant Arabidopsis seedlings to obtain high-quality sequencing reads on uridylation  mediated by uridyl transferase 1 (URT1). TAIL-Seq was effective in isolating and sequencing oligoadenine strands with up to 30 adenines and its ability to detect modified 3' ends. The authors discovered that URT1-mediated uridylation is not only able to promote mRNA degradation but also poly(A) tail repair via recruitment of poly(A)-binding proteins.

Illumina Technology: HiSeq 2500 System, TruSeq Small RNA Library Prep Kit

Park J. E., Yi H., Kim Y., Chang H. and Kim V. N. Regulation of Poly(A) Tail and Translation during the Somatic Cell Cycle. *Mol Cell.* 2016;62:462-471

Lapointe C. P., Wilinski D., Saunders H. A. J. and Wickens M. Protein-RNA networks revealed through covalent RNA marks. Nature Methods. 2015;

## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Stranded mRNA Library Prep Kit

Ribo-Zero Gold rRNA Removal Kits

PhiX Sequencing Control v3

---

45.    Chang H., Lim J., Ha M. and Kim V. N. TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. *Mol Cell.* 2014;53:1044-1052

46.    Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38:226-236

# PAL-Seq: Poly(A)-Tail Length Profiling by Sequencing

PAL-Seq measures poly(A)-tail length by incorporating fluorescent tags on biotinylated deoxyuridine triphosphate (dUTPs) and using signal intensity to quantify poly(A)-tail length.[47] Similar to 3P-Seq RNA library preparation, a splint oligonucleotide containing a 3'-adapter sequence is ligated to the 3' end of polyadenylated RNA and partially digested with RNase T1.

To separate mRNA from total RNA, the sample is size-selected by gel purification and bound to streptavidin beads before phosphorylating the 5' ends for adapter ligation. Before cluster generation, the mRNA fragments are reverse-transcribed into cDNA, released from the beads, and purified by size selection through a gel. Sequencing primers are attached to the 3' end of the poly(A) sequence and extended using deoxythymidine triphosphate (dTTP) and biotinylated dUTP. To map the fragments, regions near the 5' end of the poly(A) tails are sequenced. Fluorescent-labeled streptavidin molecules are attached to the biotin-dUTPs, and their signal intensity is measured to quantify the length of the adenine homopolymers in each cluster.
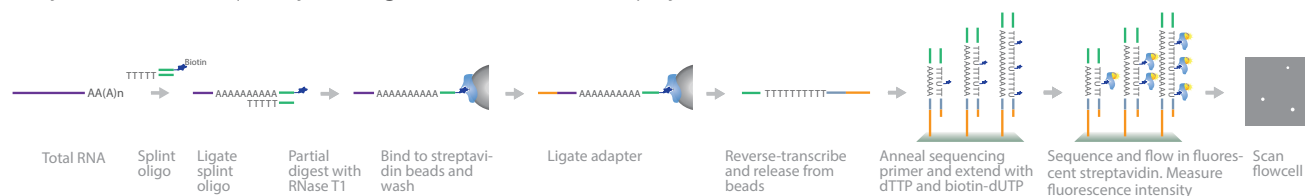


A schematic overview of PAL-seq.

| Advantages | Disadvantages |
|---|---|
| • Accurate measurement of poly(A)-tail length, regardless of length<br>• Avoids direct sequencing of the poly(A) tail | • Technically complex to perform<br>• Efficiency-related issues may arise during the biotin-dUTP extension step<br>• Only captures 3' ends that consist purely of adenines |

## Reviews

Bangru S. and Kalsotra A. Advances in analyzing RNA diversity in eukaryotic transcriptomes: peering through the Omics lens. *F1000Research*. 2016;5:2668

Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38:226-236

## References

Eichhorn S. W., Subtelny A. O., Kronja I., Kwasnieski J. C., Orr-Weaver T. L. and Bartel D. P. mRNA poly(A)-tail changes specified by deadenylation broadly reshape translation in Drosophila oocytes and early embryos. *Elife*. 2016;5:

The authors investigated the relationship between poly(A)-tail length and translation efficiency (TE) in *Drosophila melanogaster* throughout the course of oocyte maturation into embryos. They used PAL-Seq to determine poly(A)-tail length and RNA-Seq, in conjunction with ribosome-footprint profiling, to determine TE. PAL-Seq offers an advantage due to its ability to identify short and long 3'-UTR isoforms, as well as cleavage and poly(A) site mapping. The results indicate that poly(A)-tail length plays a major part in determining TE up to the gastrulation phase.

Illumina Technology: HiSeq System, Genome Analyzer System

Harrison P. F., Powell D. R., Clancy J. L., et al. PAT-seq: a method to study the integration of 3'-UTR dynamics with gene expression in the eukaryotic transcriptome. *RNA*. 2015;21:1502-1510

Kappel C., Trost G., Czesnick H., et al. Genome-Wide Analysis of PAPS1-Dependent Polyadenylation Identifies Novel Roles for Functionally Specialized Poly(A) Polymerases in Arabidopsis thaliana. *PLoS Genet.* 2015;11:e1005474

## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Stranded mRNA Library Prep Kit

---

47.    Subtelny A. O., Eichhorn S. W., Chen G. R., Sive H. and Bartel D. P. Poly(A)-tail profiling reveals an embryonic switch in translational control. Nature. 2014;508:66-71

# FRT-Seq: Flowcell Reverse Transcription Sequencing

Flow-cell surface reverse transcription sequencing (FRT-Seq) is an transcriptome sequencing technique developed in 2010 by Mamanova et al.[48] The method is strand-specific, free of amplification, and compatible with paired-end sequencing.

To begin with, poly(A) RNA samples are fragmented by metal-ion hydrolysis and dephosphorylated. Next, P7 primer/adapters are ligated to the 3' end of the fragments. The 3' adapter sequence starts at the 5' terminus with 20 nt of RNA followed by DNA nucleotides. The adapters are also 5' phosphorylated and blocked with dideoxycytosine (ddC) at the 3' end. After 3' adapter ligation, the fragments are size-selected for nucleotide fragments longer than the adapter. The 5' ends of the fragments are phosphorylated and ligated to P5 adapters. These adapters are blocked with an amino-C6 linker at the 5' end. Now that the fragments are flanked by adapters, they are hybridized to the flow cell and reverse-transcribed before cluster generation and sequencing.



A schematic overview of FRT-seq.

| Advantages | Disadvantages |
| --- | --- |
| • Strand-specific poly(A) mRNA sequencing for transcriptome analysis<br>• No amplification step—gives more accurate representation of the total mRNA population, preventing amplification biasl | • Requires a large amount of poly(A) RNA material (250 ng)<br>• Selects only poly(A) mRNA samples |

## Reviews

van Dijk E. L., Jaszczyszyn Y. and Thermes C. Library preparation methods for next-generation sequencing: Tone down the bias. *Experimental Cell Research.* 2014;322:12-20

McGettigan P. A. Transcriptomics in the RNA-seq era. Current Opinion in Chemical Biology. 2013;17:4-11

## References

Vergara-Irigaray M., Fookes M. C., Thomson N. R. and Tang C. M. RNA-seq analysis of the influence of anaerobiosis and FNR on Shigella flexneri. *BMC Genomics.* 2014;15:438:

This study investigated the effects of anaerobic environments on the transcriptome of Shigella flexneri, a Gram-negative bacterium notorious for causing acute human rectocolitis. The authors used RNA-Seq, along with FRT-seq, to quantify gene expression changes triggered upon anaerobiosis and to identify genes moderated by the anaerobic fumarate and nitrate reduction regulator (FNR). The results showed that, upon anaerobiosis, genes involved in carbon metabolism, DNA regulation, host interaction, and survival were upregulated. Most notably, many T3SS-related genes were downregulated in an FNR-dependent manner.

Illumina Technology: HiSeq 2000 System

Mills J. D., Kawahara Y. and Janitz M. Strand-Specific RNA-Seq Provides Greater Resolution of Transcriptome Profiling. *Curr Genomics.* 2013;14:173-181

Kroger C., Dillon S. C., Cameron A. D., et al. The transcriptional landscape and small RNAs of Salmonella enterica serovar Typhimurium. *Proc Natl Acad Sci U S A.* 2012;109:E1277-1286

## Associated Kits

TruSeq Stranded mRNA Library Prep Kit

48.    Mamanova L., Andrews R. M., James K. D., et al. FRT-seq: amplification-free, strand-specific transcriptome sequencing. *Nat Methods*. 2010;7:130-132

# ChIRP: Chromatin Isolation by RNA Purification

ChIRP, also commonly referred to as ChIRP-seq, is a protocol to detect the locations on the genome where ncRNAs, such as lncRNAs, and their proteins are bound.[49]

In this method, samples are first crosslinked and sonicated. Biotinylated tiling oligos are hybridized to the RNAs of interest, and the complexes are captured with streptavidin magnetic beads. After treatment with RNase H, the DNA is extracted and sequenced. Deep sequencing can determine the lncRNA/protein interaction site at single-base resolution.



A schematic overview of ChIRP.

| Advantages | Disadvantages |
|---|---|
| • Identifies binding sites anywhere on the genome<br>• Enables discovery of new binding sites<br>• Allows selection of specific RNAs of interest | • Nonspecific oligonucleotide interactions can lead to misinterpretation of binding sites<br>• Chromatin can be disrupted during the preparation stage<br>• The sequence of the RNA of interest must be known |

## Reviews

Tomita S., Abdalla M. O. A., Fujiwara S., et al. Roles of long noncoding RNAs in chromosome domains. Wiley Interdisciplinary Reviews: *RNA*. 2016;n/a-n/a

Schmitt Adam M. and Chang Howard Y. Long Noncoding RNAs in Cancer Pathways. *Cancer Cell*. 2016;29:452-463

Simon M. D. Insight into lncRNA biology using hybridization capture analyses. Biochimica et Biophysica Acta (BBA) - *Gene Regulatory Mechanisms*. 2016;1859:121-127

Yang Y., Wen L. and Zhu H. Unveiling the hidden function of long non-coding RNA by identifying its major partner-protein. *Cell & Bioscience*. 2015;5:59

Chu C., Spitale R. C. and Chang H. Y. Technologies to probe functions and mechanisms of long noncoding RNAs. *Nat Struct Mol Biol*. 2015;22:29-35

## References

**Flynn R. A., Do B. T., Rubin A. J., et al. 7SK-BAF axis controls pervasive transcription at enhancers. *Nat Struct Mol Biol*. 2016;23:231-238**

The researchers use a combination of ChIRP-seq, global run-on sequencing (GRO-Seq), and assay for transposase accessible chromatin (ATAC-Seq) to show that the 7SK small nuclear RNA (snRNA) modulates nucleosome positions to inhibit enhancer transcription. They were able to selectively recover 7SK-bound regions in an RNA-dependent manner and showed that 7SK uses various mechanisms to counteract super-enhancers, enhancers, and promoters.

Illumina Technology: HiSeq 2500 System

Huang W., Thomas B., Flynn R. A., et al. DDX5 and its associated lncRNA Rmrp modulate TH17 cell effector functions. *Nature*. 2015;528:517-522

Luo M., Jeong M., Sun D., et al. Long Non-Coding RNAs Control Hematopoietic Stem Cell Function. *Cell Stem Cell*. 2015;16:426-438

Wongtrakoongate P., Riddick G., Fucharoen S. and Felsenfeld G. Association of the Long Non-coding RNA Steroid Receptor RNA Activator (SRA) with TrxG and PRC2 Complexes. *PLoS Genet*. 2015;11:e1005615

Li Z., Chao T. C., Chang K. Y., et al. The long noncoding RNA THRIL regulates TNFalpha expression through its interaction with hnRNPL. *Proc Natl Acad Sci U S A*. 2014;111:1002-1007

## Associated Kits

ScriptSeq Complete Kit

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

---

49.   Chu C., Qu K., Zhong F. L., Artandi S. E. and Chang H. Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol Cell*. 2011;44:667-678

# CHART: Capture Hybridization Analysis of RNA Targets

CHART maps genomic binding sites of ncRNAs by isolating and sequencing the DNA regions where the crosslinked RNA-DNA-protein complexes are bound.[50] CHART differs from other crosslinked-complex purification techniques, such as ChIRP, due to the use of biotinylated 24 nt oligonucleotides (C-oligos) that are highly sensitive and unique to the ncRNA of interest.[51]

An RNase H mapping assay is used to design the 24 nt sequence of the C-oligos. First, nuclei samples are crosslinked and fragmented. Next, C-oligos are hybridized to the complex and bound to streptavidin beads. The mixture is washed and the complex eluted. The DNA is isolated and sequenced, and the proteins involved in the complex are isolated and analyzed by Western blots.



| | ncRNA | Crosslink and lyse cells | Hybridize biotinylated probes | Capture, wash and elute | Extract DNA or Western blot analysis of proteins | DNA |
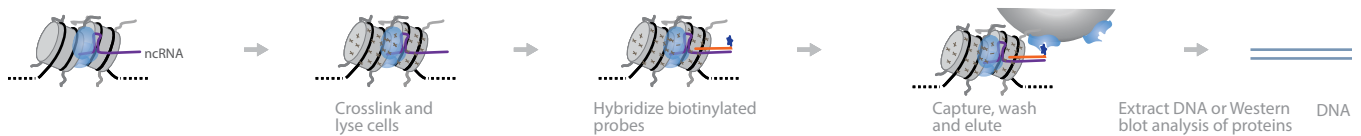
A schematic overview of ChIRP.

| Advantages | Disadvantages |
|---|---|
| • Maps genomic binding sites of lncRNAs<br>• Simultaneously identifies proteins associated with the lncRNA complex | • Needs large amount of nuclei (1 x 10$^9$ cells) |

## Reviews
Hassan M. Q., Tye C. E., Stein G. S. and Lian J. B. Non-coding RNAs: Epigenetic regulators of bone development and homeostasis. *Bone*. 2015;81:746-756

## References
Lee N., Moss W. N., Yario T. A. and Steitz J. A. EBV noncoding RNA binds nascent RNA to drive host PAX5 to viral DNA. *Cell*. 2015;160:607-618

This study focused on expanding the understanding of the complex interaction of EBER2—a highly expressed nuclear ncRNA in the Epstein-Barr virus (EBV) with high oncogenic potential—with the host's system. The authors used antisense oligos in CHART to isolate EBER2 and sequence the target DNA. They found that EBER2 required the recruitment of the host's PAX5 protein to colocalize with the nascent EBV genome at its terminal tandem-repeat regions. Knockouts of EBER2 decreased the lytic replication of EBV.

Illumina Technology: HiSeq 2000 System

Torres M., Becquet D., Blanchard M. P., et al. Circadian RNA expression elicited by 3'-UTR IRAlu-paraspeckle associated elements. *Elife*. 2016;5:

Vance K. W., Sansom S. N., Lee S., et al. The long non-coding RNA Paupar regulates the expression of both local and distal genes. *EMBO* J. 2014;33:296-311

## Associated Kits
TruSeq DNA PCR-Free Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

50.   Simon M. D., Wang C. I., Kharchenko P. V., et al. The genomic binding sites of a noncoding RNA. *Proc Natl Acad Sci U S A.* 2011;108:20497-20502
51.   Kashi K., Henderson L., Bonetti A. and Carninci P. Discovery and functional analysis of lncRNAs: Methodologies to investigate an uncharacterized transcriptome. *Biochim Biophys Acta*. 2015;

# RAP: RNA Antisense Purification

RAP isolates lncRNAs and maps the sequence of their target DNA through a probe-capture mechanism.[52]

First, the cells are crosslinked and lysed before DNase I chromatin digestion to 100–300 bp DNA fragments. Biotinylated RNA probes, antisense to the lncRNA, are hybridized and captured with streptavidin. The biotin-RNA probes are 120 nt and are tiled every 15 nt over the span of the lncRNA. The captured complexes are eluted and prepared for sequencing. RNA library preparation is done through RAP-RNA, and DNA library preparation by standard chromatin immuniprecipitation (ChIP).



| lncRNA | Crosslink and lyse cells | Hybridize biotinylated 120 bp antisense probes | Capture, wash and elute | PCR and reverse-tran-scribe | cDNA |
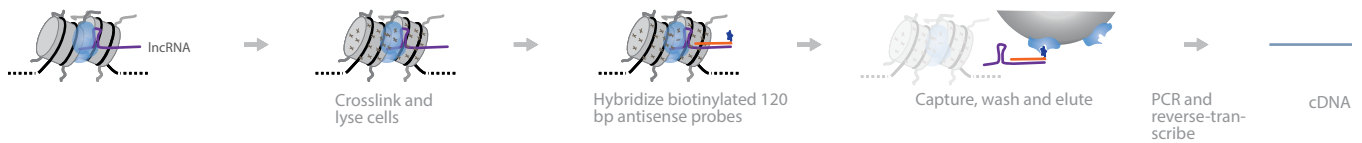
A schematic overview of RAP.

| Advantages | Disadvantages |
|---|---|
| • Genomic mapping of lncRNA targets<br>• Possible to sequence RNA and DNA from the purification products<br>• Long RNA probe length provides high binding affinity to the target lncRNA[53]<br>• Minimal amplification steps during RNA sequencing after purification of the lncRNA complex | • Requires the RNA sequence to be known |

## Reviews
Kashi K., Henderson L., Bonetti A. and Carninci P. Discovery and functional analysis of lncRNAs: Methodologies to investigate an uncharacterized transcriptome. *Biochim Biophys Acta*. 2015;

Simon M. D. Insight into lncRNA biology using hybridization capture analyses. *Biochim Biophys Acta.* 2016;1859:121-127

## References
Chen C. K., Blanco M., Jackson C., et al. Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science*. 2016;

This study showed the direct interaction of Xist lncRNA with the lamina B receptor in mediating X-chromosome inactivation. The authors used RAP and DNA sequencing to isolate Xist lncRNA in ΔLBS-Xist and wild-type mouse embryonic stem cells (ESCs). They found that Xist localization was significantly reduced in ΔLBS-Xist ESCs.

Illumina Technology: HiSeq System, MiSeq System

Boque-Sastre R., Soler M., Oliveira-Mateos C., et al. Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. *Proc Natl Acad Sci U S A.* 2015;

McHugh C. A., Chen C.-K., Chow A., et al. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*. 2015;521:232-236

Engreitz J. M., Sirokman K., McDonel P., et al. RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent Pre-mRNAs and chromatin sites. *Cell*. 2014;159:188-199

## Associated Kits
TruSeq Stranded Total RNA Library Prep Kit

TruSeq ChIP Library Prep Kit

52.   Engreitz J. M., Pandya-Jones A., McDonel P., et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science*. 2013;341:1237973
53.   Kashi K., Henderson L., Bonetti A. and Carninci P. Discovery and functional analysis of lncRNAs: Methodologies to investigate an uncharacterized transcriptome. *Biochim Biophys Acta*. 2015;

# GRO-seq: Global Run-on Sequencing

GRO-Seq maps the binding sites of transcriptionally active RNA polymerase II (RNAPII).[54]

In this method, active RNAPII is allowed to run on in the presence of 5-bromouridine 5'-triphosphate (Br-UTP). RNAs are hydrolyzed and purified using beads coated with antibodies to 5-bromo-2-deoxyuridine (BrdU). After cap removal and end repair, the eluted RNA is reverse-transcribed to cDNA. Deep sequencing of the cDNA identifies RNAs that are actively transcribed by RNAPII.



Run-on with analog    Isolate and hydrolyze    Bead coated with α-BrdUTP antibody    Elute Cap removal End Repair    Reverse-transcription, amplification    cDNA
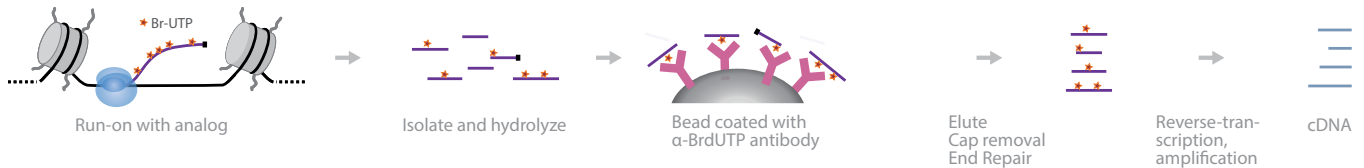
A schematic overview of GRO-seq.



Br-UTP
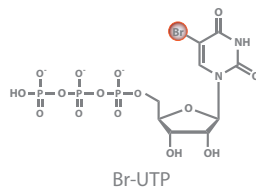
The structure of 5-bromouridine 5'-triphosphate.

| Advantages | Disadvantages |
|---|---|
| • Maps position of transcriptionally engaged RNA polymerases<br>• Determines relative activity of transcription sites<br>• Detects sense and antisense transcription<br>• Detects transcription anywhere on the genome<br>• No prior knowledge of transcription sites is needed<br>• Provides robust coverage of enhancer- and promoter-associated RNAs[55] | • Limited to cell cultures and other artificial systems, due to the requirement for incubation in the presence of labeled nucleotides<br>• Artifacts may be introduced during the preparation of the nuclei[56].<br>• New initiation events may occur during the run-on step<br>• Physical impediments may block the polymerases<br>• Resolution is only 30–100 nt[57]<br>• Requires nascent RNAs of at least 18 nt[58] |

## Reviews

Murakawa Y., Yoshihara M., Kawaji H., Nishikawa M., Zayed H., et al. Enhanced Identification of Transcriptional Enhancers Provides Mechanistic Insights into Diseases. *Trends Genet*. 2016;32:76-88

Li Y., Chen C. Y., Kaye A. M. and Wasserman W. W. The identification of cis-regulatory elements: A review from a machine learning perspective. *Biosystems*. 2015;138:6-17

Jonkers I. and Lis J. T. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol*. 2015;16:167-177

54.    Core L. J., Waterfall J. J. and Lis J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science*. 2008;322:1845-1848

55.    Melnik S., Caudron-Herger M., Brant L., et al. Isolation of the protein and RNA content of active sites of transcription from mammalian cells. *Nat Protoc*. 2016;11:553-565

56.    Adelman K. and Lis J. T. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet*. 2012;13:720-731

57.    Nojima T., Gomes T., Carmo-Fonseca M. and Proudfoot N. J. Mammalian NET-seq analysis defines nascent RNA profiles and associated RNA processing genome-wide. *Nat Protoc*. 2016;11:413-428

58.    Mayer A. and Churchman L. S. Genome-wide profiling of RNA polymerase transcription at nucleotide resolution in human cells with native elongating transcript sequencing. *Nat Protoc*. 2016;11:813-833

# References

Schwer B., Wei P. C., Chang A. N., et al. Transcription-associated processes cause DNA double-strand breaks and translocations in neural stem/progenitor cells. *Proc Natl Acad Sci U S A.* 2016;113:2258-2263

The authors used GRO-Seq to identify actively transcribed regions in primary neural stem/progenitor cells. They then used high-throughput genome-wide translocation sequencing (HTGTS) to show that activated B cells in culture generate genomic DNA double-strand breaks (DSBs) at active TSS.

Illumina Technology: MiSeq System

Chen Y. C., Stuwe E., Luo Y., et al. Cutoff Suppresses RNA Polymerase II Termination to Ensure Expression of piRNA Precursors. *Mol Cell.* 2016;63:97-109

Czimmerer Z., Varga T., Kiss M., et al. The IL-4/STAT6 signaling axis establishes a conserved microRNA signature in human and mouse macrophages regulating cell survival via miR-342-3p. *Genome Med*. 2016;8:63

Day D. S., Zhang B., Stevens S. M., et al. Comprehensive analysis of promoter-proximal RNA polymerase II pausing across mammalian cell types. *Genome Biol.* 2016;17:120

de Dieuleveult M., Yen K., Hmitou I., et al. Genome-wide nucleosome specificity and function of chromatin remodellers in ES cells. *Nature*. 2016;530:113-116

Flynn R. A., Do B. T., Rubin A. J., Calo E., Lee B., et al. 7SK-BAF axis controls pervasive transcription at enhancers. *Nat Struct Mol Biol.* 2016;23:231-238

Korkmaz G., Lopes R., Ugalde A. P., et al. Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nat Biotechnol*. 2016;34:192-198

Melnik S., Caudron-Herger M., Brant L., Carr I. M., Rippe K., et al. Isolation of the protein and RNA content of active sites of transcription from mammalian cells. *Nat Protoc.* 2016;11:553-565

Nojima T., Gomes T., Carmo-Fonseca M. and Proudfoot N. J. Mammalian NET-seq analysis defines nascent RNA profiles and associated RNA processing genome-wide. *Nat Protoc.* 2016;11:413-428

Petryk N., Kahli M., d'Aubenton-Carafa Y., et al. Replication landscape of the human genome. *Nat Commun*. 2016;7:10208

Woolnough J. L., Atwood B. L., Liu Z., Zhao R. and Giles K. E. The Regulation of rRNA Gene Transcription during Directed Differentiation of Human Embryonic Stem Cells. *PLoS One*. 2016;11:e0157276

## Associated Kits

ScriptSeq Complete Kit

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

# Bru-Seq: Bromouridine Sequencing

Bru-Seq maps nascent RNA transcripts using bromouridine tagging.[59] Active RNAPII synthesizes RNA in the presence of Br-UTP. Tagged RNA transcripts are immunoseparated from total RNA using magnetic beads coated with anti-BrdU antibodies. The captured RNA transcripts are eluted and fragmented before synthesis of cDNA strands via RT and PCR amplification. The resultant cDNA strands are prepared for sequencing using an Illumina TruSeq RNA Library Prep Kit.
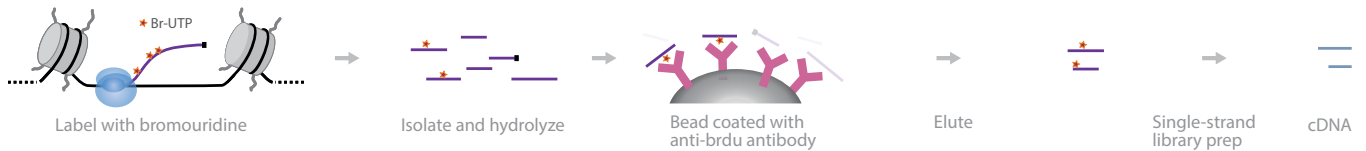


A schematic overview of Bru-Seq.

| Advantages | Disadvantages |
| --- | --- |
| • Maps sequences of nascent RNA transcripts and determines relative transcription rate<br>• Detects lncRNAs[60]<br>• Detects transcription anywhere on the genome | • Limited to cell cultures and other artificial systems, due to the requirement for incubation in the presence of labeled nucleotides |

## Reviews

Andrade-Lima L., Veloso A. and Ljungman M. Transcription Blockage Leads to New Beginnings. *Biomolecules*. 2015;5:1600

## References

Lefkofsky H. B., Veloso A. and Ljungman M. Transcriptional and post-transcriptional regulation of nucleotide excision repair genes in human cells. *Mutat Res*. 2015;776:9-15

This study investigated transcriptional and post-transcriptional regulation of 29 genes involved in nucleotide excision repair (NER), an important mechanism in repairing ultraviolet (UV)-induced DNA damages. The authors used Bru-Seq and bromouridine pulse-chase and sequencing (BruChase-Seq) in 13 human cell lines to quantify gene expression levels and RNA stability along with turnover rate, respectively. Among the 29 genes, ERCC6 stood out as a strong potential marker for DNA damage due to its large size and high RNA turnover rate.

Illumina Technology: HiSeq 2000 System, TruSeq RNA Library Prep Kit

Kocab A. J., Veloso A., Paulsen M. T., Ljungman M. and Duckett C. S. Effects of physiological and synthetic IAP antagonism on c-IAP-dependent signaling. *Oncogene*. 2015;

## Associated Kits

TruSeq RNA Library Prep Kit v2

---

59.  Paulsen M. T., Veloso A., Prasad J., et al. Coordinated regulation of synthesis and stability of RNA during the acute TNF-induced proinflammatory response. *Proc Natl Acad Sci U S A*. 2013;110:2240-2245

60.  Paulsen M. T., Veloso A., Prasad J., et al. Use of Bru-Seq and BruChase-Seq for genome-wide assessment of the synthesis and stability of RNA. *Methods*. 2014;67:45-54

# BruChase-Seq: Bromouridine Pulse-chase Sequencing

BruChase-Seq uses bromouridine tagging to map and quantify the relative stability of nascent RNA transcripts.[61] RNAPII synthesizes RNA in the presence of Br-UTP. Untagged uridine is added to chase out bromouridine from RNAs with high turnover. Long-lived RNA transcripts will remain tagged.

The tagged RNA transcripts are immunoseparated from total RNA using magnetic beads coated with anti-BrdU antibodies. The captured RNA transcripts are eluted and fragmented before synthesis of cDNA strands via RT and PCR amplification. The resultant cDNA strands are prepared for sequencing using an Illumina TruSeq RNA Library Prep Kit.



A schematic overview of BruChase-Seq.

| Advantages | Disadvantages |
| --- | --- |
| • Determines RNA stability through pulse-chase with bromouridine<br>• Predicts nonsense and frameshift mutations[62]<br>• Detects transcription anywhere on the genome<br>• No prior knowledge of transcription sites is needed | • Limited to cell cultures and other artificial systems, due to the requirement for incubation in the presence of labeled nucleotides |

## Reviews

Andrade-Lima L., Veloso A. and Ljungman M. Transcription Blockage Leads to New Beginnings. *Biomolecules*. 2015;5:1600

## References

Lefkofsky H. B., Veloso A. and Ljungman M. Transcriptional and post-transcriptional regulation of nucleotide excision repair genes in human cells. *Mutat Res.* 2015;776:9-15

This study investigated transcriptional and post-transcriptional regulation of 29 genes involved in NER, an important mechanism in repairing UV-induced DNA damages. The authors used Bru-Seq and BruChase-Seq in 13 human cell lines to quantify gene expression levels and RNA stability along with turnover rate, respectively. Among the 29 genes, ERCC6 stood out as a strong potential marker for DNA damage due to its large size and high RNA turnover rate.

Illumina Technology: HiSeq 2000 System, TruSeq RNA Library Prep Kit

Kocab A. J., Veloso A., Paulsen M. T., Ljungman M. and Duckett C. S. Effects of physiological and synthetic IAP antagonism on c-IAP-dependent signaling. *Oncogene*. 2015;

## Associated Kits

TruSeq RNA Library Prep Kit v2

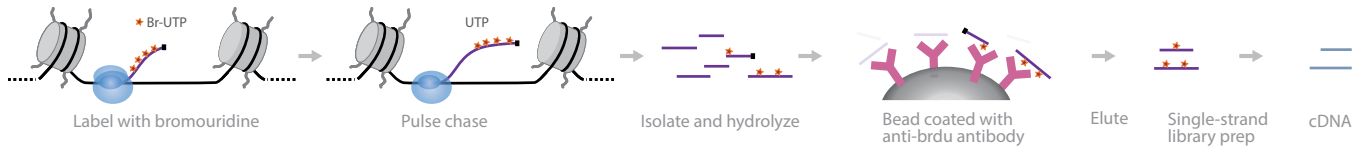61.   Paulsen M. T., Veloso A., Prasad J., et al. Coordinated regulation of synthesis and stability of RNA during the acute TNF-induced proinflammatory response. *Proc Natl Acad Sci U S A*. 2013;110:2240-2245
62.   Paulsen M. T., Veloso A., Prasad J., et al. Use of Bru-Seq and BruChase-Seq for genome-wide assessment of the synthesis and stability of RNA. *Methods*. 2014;67:45-54

# 5'-GRO-Seq: 5' Global Run-on Sequencing

5'-GRO-Seq maps the sequences of nascent RNA with a 7-methylguanylate (m7G) cap at any given time using labeled nucleotides.[63] This method was originally developed to map and detect instabilities in TSS due to the presence of enhancer RNA.

Much like GRO-seq, 5'-GRO-Seq starts with the addition of Br-UTP and sarkosyl to the lysed nuclear material. The Br-UTP acts as a marker for isolating the RNA, while sarkosyl inhibits binding of additional RNAPII to the DNA. After the reaction is stopped, DNase I is added and the RNA products fragmented. The 3' ends of RNA are dephosphorylated with T4 PNK and fragments containing bromouridine are captured with anti-BrdU antibodies. The isolated RNAs are dephosphorylated with CIP, and the 5' caps are removed using TAP. Next, 3' and 5' adapters are ligated using RNA ligase 2 and RNA ligase, respectively. The fragments are reverse-transcribed, the resultant cDNA is isolated and amplified, and the cDNA is size-selected for 60–110 bp fragments. The fragments are isolated from the gel and sequenced.
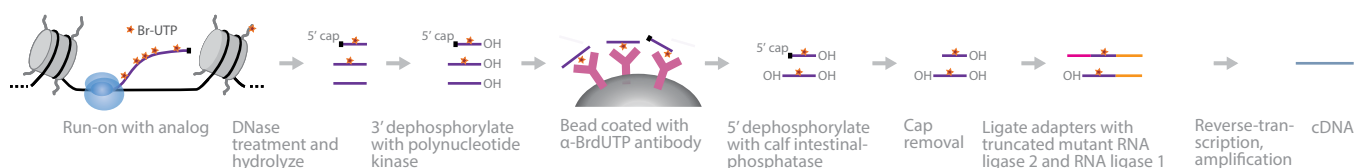


A schematic overview of 5'-GRO-Seq.

| Advantages | Disadvantages |
|---|---|
| • Maps nascent capped 5' RNA sequence at any given time<br>• Determines activity of transcription sites<br>• No prior knowledge of transcription sites needed | • Limited to cell cultures and other artificial systems, due to the requirement for incubation in the presence of labeled nucleotides |

## Reviews

Lam M. T. Y., Li W., Rosenfeld M. G. and Glass C. K. Enhancer RNAs and regulated transcriptional programs. *Trends in Biochemical Sciences.* 2014;39:170-182

## References

Duttke S. H., Lacadie S. A., Ibrahim M. M., et al. Human promoters are intrinsically directional. *Mol Cell.* 2015;57:674-684

This study used 5'-GRO-Seq to investigate directionality in human promoter sequences by mapping nascent RNA 5' ends in HeLa cells. The authors discovered that the basal transcription mechanism is inherently unidirectional, and divergent transcription sites have their own promoters that are adjacent to the edges of nucleosome-free regions. Surprisingly, the sequences of these regions are similar to their forward counterparts.

Illumina Technology: TruSeq RNA Library Prep Kit

## Associated Kits

TruSeq Stranded Total RNA Library Prep Kit2

63.    Lam M. T., Cho H., Lesch H. P., et al. Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature.* 2013;498:511-515

# BruDRB-Seq: Bromouridine 5,6-dichlorobenzimidazole1-β-D-ribofuranoside Sequencing

BruDRB-Seq reports the elongation rate of RNAPII.[64]

5,6-dichlorobenzimidazole 1-β-D-ribofuranoside (DRB) is added to cells before elongation to inhibit RNAPII transiently, allowing synchronized transcriptional initiation throughout the genome. Upon removal of DRB, Br-UTP is added instead of UTP, along with other nucleotides. After cell lysis, RNA is isolated and fragmented. Next, bromouridine-tagged RNA is immunoseparated from total RNA using magnetic beads coated with anti-BrdU antibodies. cDNA libraries are generated, using the TruSeq RNA library preparation protocol, and then sequenced.
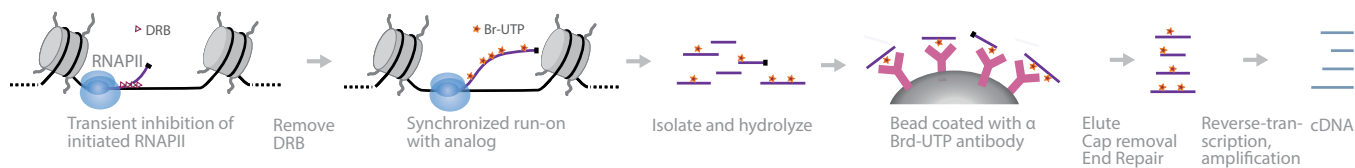


A schematic overview of BruDRB-seq.

| Advantages | Disadvantages |
|---|---|
| • Quantifies RNA elongation rate throughout the whole genome<br>• Newly transcribed RNA molecules are tagged through their entire length | • Limited to cell cultures and other artificial systems, due to the requirement for incubation in the presence of labeled nucleotides<br>• Bru-Seq must be performed in conjunction with Bru-DRB-Seq for accurate data analysis |

## Reviews
Jonkers I. and Lis J. T. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol.* 2015;16:167-177

## References
Veloso A., Kirkconnell K. S., Magnuson B., Biewen B., Paulsen M. T., et al. Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. *Genome Res.* 2014;24:896-905

This is the original publication describing the development of BruDRB-seq. The authors studied the intrinsic mechanism of transcription control and used BruDRB-Seq to calculate the elongation rate of over 2000 different genes in human fibroblast cell lines. The results showed that genes with histone methylation markers H3K79me2 and H4K20me1 have higher transcription elongation rates.

Illumina Technology: HiSeq 2000 System, TruSeq RNA Library Prep Kit

## Associated Kits
TruSeq RNA Library Prep Kit v2

---

64.  Veloso A., Kirkconnell K. S., Magnuson B., et al. Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. *Genome Res*. 2014;24:896-905

# 4sUDRB-Seq: 4-Thiouridine and 5,6-Dichlorobenzimidazole 1-β-D-Ribofuranoside Sequencing

4sUDRB-Seq investigates initiation frequencies and RNA elongation rates throughout the genome using 4-thiouridine (4-SU) and DRB.[65]

Cells are first treated with DRB to inhibit RNA elongation and arrest RNAPII at TSS. The cells are lysed, cleansed of DRB, and immediately incubated with 4-SU to label newly transcribed RNA molecules. Biotin is added to bind with 4-SU, and the tagged RNA is subsequently captured using streptavidin beads. The biotinylated RNA fragments are eluted and prepared for sequencing according to the protocol in the TruSeq RNA Library Preparation Kit.
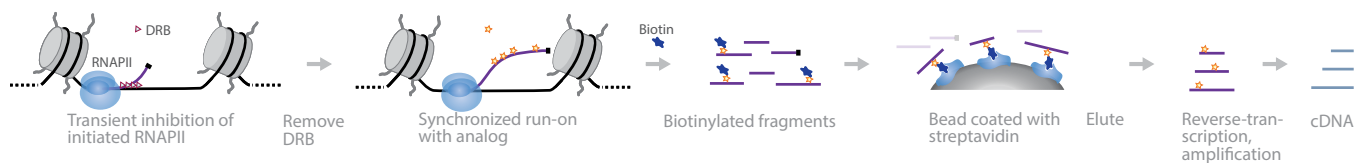


A schematic overview of 4sUDRB-seq.

| Advantages | Disadvantages |
|---|---|
| • Measures RNA elongation rate and transcription initiation rate simultaneously<br>• Sequencing reads behave dynamically throughout the transcription wave—useful in determining transcription initiation rate accurately[66]<br>• No previous knowledge of the sequence is needed | • High toxicity of 4-SU can debilitate experiments with slow elongation rates[67]<br>• Measuring genome-wide transition of RNAPII into active elongation can be cost-inefficient, due to the high sequencing depth needed[68]<br>• Dynamic sequence reads makes identification of transcript ends challenging[69]<br>• Limited to cell cultures and other artificial systems, due to the requirement for incubation in the presence of labeled nucleotides |

## Reviews

Rabani M., Raychowdhury R., Jovanovic M., et al. High-Resolution Sequencing and Modeling Identifies Distinct Dynamic RNA Regulatory Strategies. *Cell*. 159:1698-1710

Duffy E. E., Rutenberg-Schoenberg M., Stark C. D., Kitchen R. R., Gerstein M. B. and Simon M. D. Tracking Distinct RNA Populations Using Efficient and Reversible Covalent Chemistry. *Mol Cell.* 2015;59:858-866

Jonkers I. and Lis J. T. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol.* 2015;16:167-177

## References

Zhang Y., Xue W., Li X., et al. The Biogenesis of Nascent Circular RNAs. Cell Rep. 2016;15:611-624

Pre-mRNA back-splicing produces circular RNAs (circRNAs), which have been mapped to thousands of genomic loci in mammals and are important to gene regulation. However, the detailed production mechanisms of circRNAs are poorly understood. In this study, the authors used 4sUDRB-Seq in human cell lines. They discovered that circRNAs are processed post-transcriptionally in most cases and are stable.

Illumina Technology: HiSeq 2000 System

65.  Fuchs G., Voichek Y., Benjamin S., Gilad S., Amit I. and Oren M. 4sUDRB-seq: measuring genomewide transcriptional elongation rates and initiation frequencies within cells. *Genome Biol.* 2014;15:R69
66.  Fuchs G., Voichek Y., Rabani M., et al. Simultaneous measurement of genome-wide transcription elongation speeds and rates of RNA polymerase II transition into active elongation with 4sUDRB-seq. *Nat Protoc.* 2015;10:605-618
67.  Fuchs G., Voichek Y., Rabani M., et al. Simultaneous measurement of genome-wide transcription elongation speeds and rates of RNA polymerase II transition into active elongation with 4sUDRB-seq. *Nat Protoc.* 2015;10:605-618
68.  Fuchs G., Voichek Y., Rabani M., et al. Simultaneous measurement of genome-wide transcription elongation speeds and rates of RNA polymerase II transition into active elongation with 4sUDRB-seq. *Nat Protoc.* 2015;10:605-618
69.  Fuchs G., Voichek Y., Rabani M., et al. Simultaneous measurement of genome-wide transcription elongation speeds and rates of RNA polymerase II transition into active elongation with 4sUDRB-seq. *Nat Protoc.* 2015;10:605-618

Jaenicke L. A., von Eyss B., Carstensen A., et al. Ubiquitin-Dependent Turnover of MYC Antagonizes MYC/PAF1C Complex Accumulation to Drive Transcriptional Elongation. Mol Cell. 2016;61:54-67

Fuchs G., Rosenthal E., Bublik D.-R., Kaplan T. and Oren M. Gene body H2B monoubiquitylation regulates gene-selective RNA Polymerase II pause release and is not rate limiting for transcription elongation. bioRxiv. 2015;

## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Stranded Total RNA Library Prep Kit

# PRO-Seq: Precision Nuclear Run-on Sequencing

Precision nuclear run-on sequencing (PRO-Seq) maps RNAPII pause sites with base-pair resolution during RNA transcription.[70, 71] This approach is similar to GRO-Seq, but it provides the added benefit of single-base resolution. RNAPII initiation sites can be mapped using a modified protocol named PRO-cap.

In PRO-seq, 4 separate run-on reactions, each with only 1 type of biotin-NTP and sarkosyl, are carried out on nuclear lysates. Incorporation of the single biotin-NTP halts further elongation of nascent RNA strands by RNAPII. The RNA strands are extracted, fragmented, and purified through streptavidin pull-down. Next, 3' adapters are ligated directly to the purified sample before another streptavidin purification step. The 5' ends are repaired using TAP and PNK before ligating 5' adapters. The adapter-flanked RNA fragments are enriched through another streptavidin pull-down process before RT and PCR amplification. The resultant cDNA strands are sequenced from the 3' end.



Single biotin-NTP run-on. Four reactions with 1 nucleotide each → Extract and Hydrolyze → Bead coated with streptavidin → Purify → Add 3' adapters → Streptavidin pulldown → Purify → Repair ends and add 5' adapter → Streptavidin pulldown → Reverse-transcription, amplification → cDNA for each nucleotide
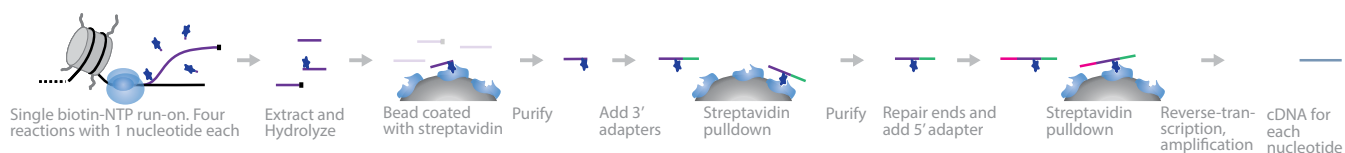
A schematic overview of PRO-seq.

| Advantages | Disadvantages |
|---|---|
| • Maps RNAPII pausing sites with base-pair resolution<br>• Separate run-on reactions limit the addition of nucleotides other than the provided biotin-NTP<br>• Multiple biotin enrichment steps before PCR<br>• Initiation sites can be mapped using PRO-cap | • Unable to detect arrested or backtracked RNAPII complexes[72]<br>• Limited to *in vitro* reactions |

## Reviews

Brent M. R. Past Roadblocks and New Opportunities in Transcription Factor Network Mapping. *Trends Genet.* 2016;32:736-750

Engreitz J. M., Haines J. E., Perez E. M., et al. Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature*. 2016;539:452-455

## References

Wang I. X., Core L. J., Kwak H., et al. RNA-DNA differences are generated in human cells within seconds after RNA exits polymerase II. *Cell Rep.* 2014;6:906-915

The authors used a combination of GRO-Seq and PRO-Seq to study RNA-DNA sequence differences (RDD) in nascent RNA strands. They discovered that RDDs start occurring approximately 55 nt from the RNAPII active site. This proximal occurrence of RDDs was incompatible with known deaminase-mediated RNA editing mechanisms. Further analysis revealed that the 55 nt delay does not arise during RNAPII-directed synthesis, and it is not due to modified base incorporation.

Illumina Technology: HiSeq 2500 System, TruSeq RNA Library Prep Kit

Danko C. G., Hyland S. L., Core L. J., et al. Identification of active transcriptional regulatory elements from GRO-seq data. *Nat Methods.* 2015;

Core L. J., Martins A. L., Danko C. G., Waters C. T., Siepel A. and Lis J. T. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat Genet.* 2014;

Pagano J. M., Kwak H., Waters C. T., et al. Defining NELF-E RNA binding in HIV-1 and promoter-proximal pause regions. *PLoS Genet.* 2014;10:e1004090

## Associated Kits

TruSeq RNA Library Prep Kit v2

70.    Kwak H., Fuda N. J., Core L. J. and Lis J. T. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science*. 2013;339:950-953
71.    Mahat D. B., Kwak H., Booth G. T., et al. Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). *Nat Protoc.* 2016;11:1455-1476
72.    Weber C. M., Ramachandran S. and Henikoff S. Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol Cell.* 2014;53:819-830

# PRO-Cap: Precision Nuclear Run-on Sequencing for RNA Polymerase II Start Sites

PRO-cap maps RNAPII initiation sites during RNA transcription with base-pair resolution. This approach is a variation of the PRO-Seq method, which maps RNAPII pause sites.[73]

A nuclear run-on reaction with biotin-NTP and sarkosyl is carried out on nuclear lysates. Incorporation of the first biotin-NTP halts further elongation of nascent RNA strands by RNAPII. The RNA strands are extracted and purified through streptavidin pull-down. Next, 3' adapters are ligated directly to the purified sample before another streptavidin purification step. The 5' ends are repaired using Antarctic phosphatase and TAP before ligating 5' adapters. The adapter-flanked RNA fragments are enriched through another streptavidin pull-down process before RT and PCR amplification. The resultant cDNA strands are sequenced from the 5' end, and RNAPII pause sites are mapped.
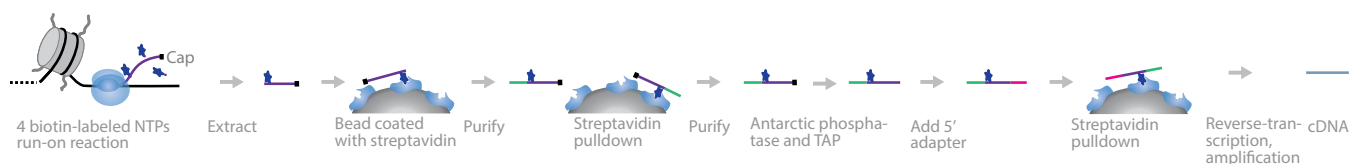


A schematic overview of PRO-cap.

| Advantages | Disadvantages |
|---|---|
| • Maps RNAPII initiation sites with base-pair resolution<br>• Multiple biotin enrichment steps before PCR<br>• Pause sites mapped using PRO-seq | • Limited to *in vitro* reactions |

## Reviews

Brent M. R. Past Roadblocks and New Opportunities in Transcription Factor Network Mapping. *Trends Genet.* 2016;32:736-750

Engreitz J. M., Haines J. E., Perez E. M., Munson G., Chen J., et al. Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature*. 2016;539:452-455

## References

Wang I. X., Core L. J., Kwak H., Brady L., Bruzel A., et al. RNA-DNA differences are generated in human cells within seconds after RNA exits polymerase II. *Cell Rep*. 2014;6:906-915

The authors used a combination of GRO-Seq and PRO-Seq to study RDDs in nascent RNA strands. They discovered that RDDs start occurring approximately 55 nt from the RNAPII active site. This proximal occurrence of RDDs was incompatible with known deaminase-mediated RNA editing mechanisms. Further analysis revealed that the 55 nt delay does not arise during RNAPII-directed synthesis, and it is not due to modified base incorporation.

Illumina Technology: HiSeq 2500 System, TruSeq RNA Library Prep Kit

Danko C. G., Hyland S. L., Core L. J., Martins A. L., Waters C. T., et al. Identification of active transcriptional regulatory elements from GRO-seq data. *Nat Methods.* 2015;

Core L. J., Martins A. L., Danko C. G., Waters C. T., Siepel A., et al. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat Genet.* 2014;

## Associated Kits

TruSeq RNA Library Prep Kit v2

---

73.    Kwak H., Fuda N. J., Core L. J. and Lis J. T. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science*. 2013;339:950-953

# CAGE: Cap Analysis Gene Expression Sequencing

CAGE measures RNA expression and maps TSS in promoters.[74]

In this method, RNA is first reverse-transcribed using random primers. The RNA cap and 3' ends are biotinylated. Nonhybridized, single-stranded RNAs are digested with RNase, leaving 5' complete cDNAs that are captured using streptavidin beads. The cDNA is processed for sequencing.



A schematic overview of CAGE.

| Advantages | Disadvantages |
|---|---|
| • Measures RNA expression levels and maps TSS in promoter regions<br>• Provides precise mapping of TSS with single-nucleotide resolution | • Only works on total mature RNA[75]<br>• Detection is biased toward TSS of long-lived transcripts[76] |

## Reviews
Haberle V. and Lenhard B. Promoter architectures and developmental gene regulation. Semin Cell Dev Biol. 2016;57:11-23

## References
**Horie M., Yamaguchi Y., Saito A., et al. Transcriptome analysis of periodontitis-associated fibroblasts by CAGE sequencing identified DLX5 and RUNX2 long variant as novel regulators involved in periodontitis. *Sci Rep.* 2016;6:33666**

This study examined periodontitis using periodontitis-associated fibroblasts (PAFs) from affected human cells. The authors used CAGE to analyze transcription initiation profiles. PAFs from affected cells displayed distinctive patterns upon comparison with CAGE profiles from other organs. The authors also identified important transcription factors, such as BARX1, PAX9, LHX8, and DLX5, that are markers for the development of periodontitis.

**Illumina Technology: HiSeq 2500 System, TruSeq Stranded Total RNA Library Prep Kit, Ribo-Zero Gold rRNA Removal Kit**

Poletti V., Delli Carri A., Malagoli Tagliazucchi G., et al. Genome-Wide Definition of Promoter and Enhancer Usage during Neural Induction of Human Embryonic Stem Cells. *PLoS One.* 2015;10:e0126590

Andersson R., Gebhard C., Miguel-Escalada I., et al. An atlas of active enhancers across human cell types and tissues. *Nature*. 2014;507:455-461

Brown J. B., Boley N., Eisman R., et al. Diversity and dynamics of the Drosophila transcriptome. *Nature*. 2014;512:393-399

Chen Z. X., Sturgill D., Qu J., et al. Comparative validation of the *D. melanogaster* modENCODE transcriptome annotation. *Genome Res.* 2014;24:1209-1223

Fort A., Hashimoto K., Yamada D., et al. Deep transcriptome profiling of mammalian stem cells supports a regulatory role for retrotransposons in pluripotency maintenance. *Nat Genet.* 2014;46:558-566

## Associated Kits
TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Nano DNA Library Prep Kit

TruSeq DNA PCR-Free Library Prep Kit

---

74. Takahashi H., Lassmann T., Murata M. and Carninci P. 5' end-centered expression profiling using cap-analysis gene expression and next-generation sequencing. Nat Protoc. 2012;7:542-561
75. Haberle V. and Lenhard B. Promoter architectures and developmental gene regulation. Semin Cell Dev Biol. 2016;57:11-23
76. Haberle V. and Lenhard B. Promoter architectures and developmental gene regulation. Semin Cell Dev Biol. 2016;57:11-23

# 3'NT Method: 3' End of Nascent Transcripts

The 3'NT method sequences nascent RNA transcripts to map the positions of elongating and arrested RNAPII complexes at nucleotide resolution.[77] This method effectively isolates the RNAPII-chromatin complex by capitalizing on its ability to remain bound to the DNA strand in the presence of high salt, urea, detergents, and polyanions.

Transcription is arrested in the sample while both the cells and nuclei are lysed. The nascent RNA strands are isolated from the purified RNAPII-chromatin complex and enriched for mRNA by selecting for strands with 5' 7-methylguanosine caps. cDNA libraries are prepared from nascent mRNA strands by following the protocol for native elongating transcript sequencing (NET-Seq), with slight modifications. The resulting cDNA libraries are sequenced to determine the precise locations of RNAPII during each experimental condition.



Transcriptome complex isolated in hypotonic buffers and nonionic detergents → Insoluble chromatin complex → RNA extraction → Reverse-transcription → cDNA

A schematic overview of 3'NT.

| Advantages | Disadvantages |
| --- | --- |
| • Tracks position of elongating and arrested RNAPII throughout the whole genome<br>• Does not require transgenes, solubilization, or immunopurification | • Unable to interrogate the relationship between RNAPII C-terminal domain modifications and nascent RNA strands[78] |

## Reviews

Zentner G. E. and Henikoff S. High-resolution digital profiling of the epigenome. *Nat Rev Genet.* 2014;

Teves S. S., Weber C. M. and Henikoff S. Transcribing through the nucleosome. *Trends Biochem Sci.* 2014;39:577-586

## References

Weber C. M., Ramachandran S. and Henikoff S. Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol Cell.* 2014;53:819-830

To study the effects of nucleosomes as transcriptome barriers, the authors developed 3'NT sequencing to isolate 5'-cap-protected nascent RNA strands from RNAPII transcription. They used the 3'NT method, along with micrococcal nuclease sequencing (MNase-Seq) and chromatin immunoprecipitation sequencing (ChIP-Seq) to study the locations of elongating and arrested RNAPII in the genome of Drosophila melanogaster. The sequence reads revealed the +1 nucleosome to be a barrier to RNAPII, while gene-body nucleosomes were low barriers that caused RNAPII stalling. The authors also discovered that depleting histone variant H2A.Z from nucleosome positions resulted in a higher barrier of entry to RNAPII.

Illumina Technology: HiSeq 2000 System, TruSeq ChIP Library Prep Kit

## Associated Kits

TruSeq ChIP Library Prep Kit

77.    Weber C. M., Ramachandran S. and Henikoff S. Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol Cell.* 2014;53:819-830
78.    Nojima T., Gomes T., Grosso A. R., et al. Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. *Cell.* 2015;161:526-540

# NET-Seq: Native Elongating Transcript Sequencing

NET-Seq detects nascent, actively transcribed RNAPII RNAs, through the capture of 3' RNA.[79]

In this method, the RNAPII elongation complex is immunoprecipitated, and RNA is extracted and reverse-transcribed to cDNA. Deep sequencing of the cDNA allows for 3'-end sequencing of nascent RNA, providing nucleotide-resolution mapping of transcripts.



A schematic overview of NET-seq.

| Advantages | Disadvantages |
|---|---|
| • Mapping of nascent RNA-bound protein<br>• Transcription is mapped at nucleotide resolution | • Antibodies not specific to the target will precipitate nonspecific complexes<br>• Requires nascent RNAs of at least 18 nt[80] |

## Reviews

Hrdlickova R., Toloue M. and Tian B. RNA-Seq methods for transcriptome analysis. Wiley Interdisciplinary Reviews: RNA. 2016;n/a-n/a

Murakawa Y., Yoshihara M., Kawaji H., Nishikawa M., Zayed H., et al. Enhanced Identification of Transcriptional Enhancers Provides Mechanistic Insights into Diseases. Trends Genet. 2016;32:76-88

Zhang J. and Landick R. A Two-Way Street: Regulatory Interplay between RNA Polymerase and Nascent RNA Structure. Trends Biochem Sci. 2016;41:293-310

Liu X., Kraus W. L. and Bai X. Ready, pause, go: regulation of RNA polymerase II pausing and release by cellular signaling pathways. Trends Biochem Sci. 2015;40:516-525

Lemay J. F. and Bachand F. Fail-safe transcription termination: Because one is never enough. RNA Biol. 2015;12:927-932

## References

Harlen K. M., Trotta K. L., Smith E. E., Mosaheb M. M., Fuchs S. M. and Churchman L. S. Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue. *Cell Rep.* 2016;15:2147-2158

This study aimed to enhance the understanding of the dynamic mechanism of the RNAPII C-terminal domain and its regulatory factors. The authors used NET-Seq to identify the transcription factors that are associated with an active, transcribing RNAPII. They also used it to analyze changes in RNAPII density in wild-type and mutant samples. The results revealed hundreds of different protein factors that were variably enriched, indicating a more complex interactome than expected.

Illumina Technology: NextSeq 500 System, MiSeq System, Ribo-Zero Gold rRNA Removal Kit

Mayer A. and Churchman L. S. Genome-wide profiling of RNA polymerase transcription at nucleotide resolution in human cells with native elongating transcript sequencing. *Nat Protoc.* 2016;11:813-833

Imashimizu M., Takahashi H., Oshima T., et al. Visualizing translocation dynamics and nascent transcript errors in paused RNA polymerases *in vivo*. *Genome Biol.* 2015;16:98

## Associated Kits

TruSeq ChIP Library Prep Kit TruSeq Ribo Profile Kit

Ribo-Zero rRNA Removal Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

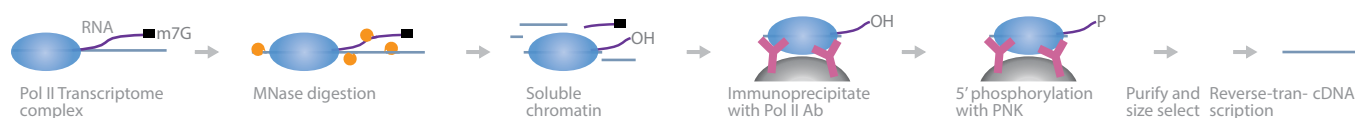TruSeq Targeted RNA Expression Library Prep Kit

79.    Churchman L. S. and Weissman J. S. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature*. 2011;469:368-373
80.    Mayer A. and Churchman L. S. Genome-wide profiling of RNA polymerase transcription at nucleotide resolution in human cells with native elongating transcript sequencing. *Nat Protoc.* 2016;11:813-833

# mNET-Seq: Native Elongating Transcript Sequencing Technology for Mammalian Chromatin

mNET-Seq generates profiles of nascent RNA and cotranscriptional RNA processing associated with different C-terminal domain (CTD) phosphorylation states throughout the whole genome.[81] mNET-Seq is able to provide precise sequence reads of RNAPII active sites during transcript elongation and also RNA processing intermediates.

First, elongating RNAPII complexes are isolated through chromatin fractionation. They are digested with MNase, breaking down all exposed DNA while leaving RNA strands protected by RNAPII or spliceosomes intact. The RNAPII complexes are immunoprecipitated using RNAPII antibodies and 5' phosphorylated by T4 PNK. Next, 3' linkers are ligated to the 3' hydroxyl end of the RNA strand still embedded within RNAPII. They are also incubated with radioactive ATP to facilitate size selection. Nascent RNA strands are size-selected for 35–100 nt, processed into cDNA sequencing libraries, and sequenced. The use of various RNAPII antibodies during purification raises the versatility and specificity of the technique in targeting CTDs of RNAPII.



| Pol II Transcriptome complex | MNase digestion | Soluble chromatin | Immunoprecipitate with Pol II Ab | 5' phosphorylation with PNK | Purify and size select | Reverse-tran- scription cDNA |

A schematic overview of mNET-seq.

| Advantages | Disadvantages |
|---|---|
| • Maps nascent RNA strands and cotranscriptional RNA processing during RNAPII elongation with phosphorylated CTDs<br>• Able to detect sense and antisense transcripts at TSS[82]<br>• No crosslinking—eliminates introduction of artifactual interactions<br>• MNase digestion is specific and efficient<br>• Various RNAPII-specific antibodies can be used to increase targeting accuracy | • Nascent RNAs shorter than 35 nt cannot be detected reliably<br>• RNA can degrade during RNAPII immunoprecipitation<br>• mNET-Seq peaks might be obscured by peaks from cotranscriptional RNA cleavage<br>• PCR amplification may give rise to peaks from amplification bias. |

## Reviews

Mellor J., Woloszczuk R. and Howe F. S. The Interleaved Genome. *Trends Genet.* 2016;32:57-71

Zaborowska J., Egloff S. and Murphy S. The pol II CTD: new twists in the tail. *Nat Struct Mol Biol.* 2016;23:771-777

## References

Nojima T., Gomes T., Grosso A. R., Kimura H., Dye M. J., et al. Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. *Cell*. 2015;161:526-540

The authors developed mNET-Seq to generate transcription profiles of nascent RNA from RNAPII and its processing intermediates. By using mNET-Seq on mammalian chromatin, they discovered that 5'-splice site cleavage by spliceosomes was associated with RNAPII CTD phosphorylation on the serine 5 position. Further, this phosphorylation affected transcription termination by restricting nonproductive RNA synthesis.

Illumina Technology: HiSeq 2000/2500 System, TruSeq Small RNA Library Prep Kit

## Associated Kits

TruSeq Small RNA Library Prep Kit

---

81.  Nojima T., Gomes T., Grosso A. R., et al. Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. *Cell*. 2015;161:526-540
82.  Nojima T., Gomes T., Carmo-Fonseca M. and Proudfoot N. J. Mammalian NET-seq analysis defines nascent RNA profiles and associated RNA processing genome-wide. *Nat Protoc*. 2016;11:413-428

# PARE-Seq: Parallel Analysis of RNA Ends Sequencing

Various RNA degradation processes impart characteristic sequence ends. By analyzing the cleavage sites, the degradation processes can be inferred.[83]

In PARE-Seq, the degraded uncapped mRNA is ligated to 5' adapters containing an MmeI restriction site and reverse-transcribed. The cDNA fragments are digested with MmeI, purified, ligated to 3' adapters, and PCR-amplified. Deep sequencing of the cDNA provides information about uncapped transcripts that undergo degradation.



A schematic overview of PARE-Seq.

| Advantages | Disadvantages |
|---|---|
| • Maps RNA degradation<br>• miRNA cleavage sites are identified<br>• No prior knowledge of the target RNA sequence is required | • Nonlinear PCR amplification can lead to biases, affecting reproducibility<br>• Amplification errors caused by polymerases will be represented and sequenced incorrectly |

## Reviews
Reuter J. A., Spacek D. V. and Snyder M. P. High-Throughput Sequencing Technologies. *Mol Cell.* 2015;58:586-597

## References
Yi F., Chen J. and Yu J. Global analysis of uncapped mRNA changes under drought stress and microRNA-dependent endonucleolytic cleavages in foxtail millet. *BMC Plant Biol.* 2015;15:241

Little is known about the role of uncapped mRNAs in plants during drought conditions. This study used PARE-Seq in Setaria italica to profile uncapped transcripts and identify miRNA-programmed cleavage sites for RNA-induced silencing complex (RISC). The authors discovered 4 types of mRNA decay patterns and identified 385 different miRNA-target interactions. In addition, they found 11 C4 enzymes related to photosynthesis enzymes that are encoded by drought-response genes.

Illumina Technology: HiSeq 2000 System

## Associated Kits
TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

---

83.    German M. A., Pillay M., Jeong D. H., et al. Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat Biotechnol.* 2008;26:941-946

# GMUCT: Genome-wide Mapping of Uncapped and Cleaved Transcripts

GMUCT is a method for constructing sequencing libraries made up of decapped or cleaved mRNAs.[84] There are 2 versions of GMUCT: version 1.0 was developed in 2008, while a more streamlined and efficient version 2.0 was designed in 2013. GMUCT 2.0 significantly decreases library preparation time by 3 days (it takes 2–3 days instead of 5–6 days) and requires 10 times less starting total RNA (5 µg instead of 50 µg.[85]

*GMUCT 1.0:* Starting with total RNA, mRNA is isolated by poly(A) selection. Next, 5' RNA adapters are ligated and the RNA is reverse-transcribed. The resultant first strand of cDNA is amplified using oligo(dT) and 5'-adapter primers. The double-stranded cDNA is fragmented, ligated to 3'- and 5'-sequencing adapters, PCR-amplified, and sequenced.
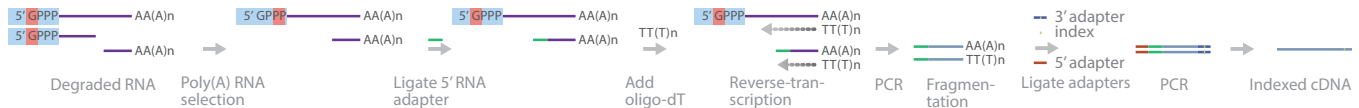


A schematic overview of GMUCT 1.0.

GMUCT 2.0: This version starts out similar to GMUCT 1.0, but deviates after the ligation of 5' RNA adapters to poly(A) RNAs. Another round of poly(A) selection is performed to further purify the desired mRNA away from any unligated RNA. RT is performed using primers with 3' adapters on their 5' terminus and a random hexamer on the 3' end. This modification adds 3' adapters during RT, resulting in cDNA strands flanked with adapters at both ends. The cDNA is PCR-amplified, to add sequencing indexes, and sequenced.



A schematic overview of GMUCT 2.0.

| Advantages | Disadvantages |
|---|---|
| • Sequences RNA degradation intermediates and uncapped RNAs<br>• GMUCT 2.0 only takes 2-3 days and requires just 5 µg of total RNA<br>• Can be modified to study cleavage of miRNA or siRNA targets | • Minimum size of 135 bp |

## Reviews
Ma X., Tang Z., Qin J. and Meng Y. The use of high-throughput sequencing methods for plant microRNA research. *RNA Biol.* 2015;12:709-719

## References
Vandivier L. E., Campos R., Kuksa P. P., Silverman I. M., Wang L. S. and Gregory B. D. Chemical Modifications Mark Alternatively Spliced and Uncapped Messenger RNAs in Arabidopsis. *Plant Cell.* 2015;27:3024-3037

The authors aimed to study the effect of post-transcriptional RNA modification on RNA maturation, stability, and function in plants. They used GMUCT, small-RNA sequencing (smRNA-Seq), and RNA-Seq to obtain sequence reads and fed them into the high-throughput annotation of modified ribonucleotides (HAMR) analysis pipeline. GMUCT was used to efficiently capture mRNA transcripts lacking 5' caps. The results revealed that modifications affecting Watson-Crick base pairing occur mostly in polyadenylated, small, and degrading RNAs. The results also hinted at additional functions of base-pair modifications in regulating alternative splicing and stress response.

Illumina Technology: HiSeq 2000 System, TruSeq Small RNA Library Prep Kit
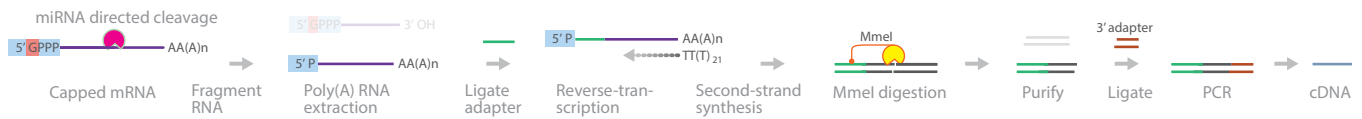
## Associated Kits
TruSeq Small RNA Library Prep Kit

84.    Gregory B. D., O'Malley R. C., Lister R., et al. A link between RNA metabolism and silencing affecting Arabidopsis development. *Dev Cell.* 2008;14:854-866
85.    Willmann M. R., Berkowitz N. D. and Gregory B. D. Improved genome-wide mapping of uncapped and cleaved transcripts in eukaryotes--GMUCT 2.0. *Methods.* 2014;67:64-73

# RNA-PROTEIN INTERACTIONS

RNA-protein interactions involve reading the RNA sequences occupied by RBPs during various post-transcriptional processes, including gene expression, alternative splicing, RNA export and localization, RNA stability, and translation. RBPs are implicated in various diseases, such neurological disorders or cancer.[86, 87] The dysregulation of RNA-binding capabilities may lead to disastrous effects at the cellular or even organismal level. This section highlights methods developed to read RNA sequences targeted by RBPs using NGS technology.

The method workflows for isolating RNA-RBP complexes described in this section follow a common scheme: 1) isolation of intact RNA-protein complexes; 2) crosslinking the complex; 3) reversing the RNA-protein crosslinks; 4) cDNA library preparation from the isolated RNA; and 5) reading the sequence of the RNA bound to the RBPs.
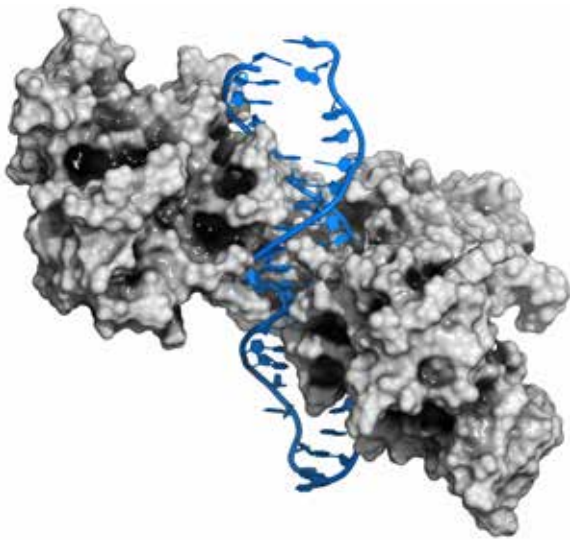


Three-dimensional structure of NF-kB protein interacting with DNA during transcription.

## Reviews

Haque N. and Hogg J. R. Easier, Better, Faster, Stronger: Improved Methods for RNA-Protein Interaction Studies. *Mol Cell.* 2016;62:650-651

Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38:226-236

Lu Z. and Chang H. Y. Decoding the RNA structurome. *Curr Opin Struct Biol.* 2016;36:142-148

Marchese D., de Groot N. S., Lorenzo Gotor N., Livi C. M. and Tartaglia G. G. Advances in the characterization of RNA-binding proteins. *Wiley Interdiscip Rev RNA.* 2016;

Cook K. B., Hughes T. R. and Morris Q. D. High-throughput characterization of protein-RNA interactions. *Brief Funct Genomics.* 2015;14:74-89

Popova V. V., Kurshakova M. M. and Kopytova D. V. Methods to study the RNA-protein interactions. *Molecular Biology.* 2015;49:418-426

Campbell Z. T. and Wickens M. Probing RNA-protein networks: biochemistry meets genomics. *Trends Biochem Sci.* 2015;40:157-164

Qu S., Yang X., Li X., et al. Circular RNA: A new star of noncoding RNAs. *Cancer Lett.* 2015;365:141-148

Burgess D. J. RNA. Detailed probing of RNA structure in vivo. *Nat Rev Genet.* 2015;16:255

86.  Haque N. and Hogg J. R. Easier, Better, Faster, Stronger: Improved Methods for RNA-Protein Interaction Studies. *Mol Cell.* 2016;62:650-651
87.  Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38:226-236

# Ribo-Seq or ARTSeq: Ribosome Profiling Sequencing

Ribosome profiling (Ribo-Seq), also called active mRNA translation sequencing (ARTseq), isolates RNA that is being processed by the ribosome in order to monitor the translation process.[88]

In this method, ribosome-bound RNA first undergoes digestion. The RNA is extracted, and the rRNA is depleted. The extracted RNA is reverse-transcribed to cDNA. Deep sequencing of the cDNA provides the sequences of RNAs bound by ribosomes during translation. This method has been refined to improve both the qualitative and quantitative nature of the results. Careful attention should be paid to: 1) generation of cell extracts in which ribosomes have been faithfully halted along the mRNA they are translating *in vivo*; 2) nuclease digestion of RNAs that are not protected by the ribosome, followed by recovery of the ribosome-protected mRNA fragments; and 3) quantitative conversion of the protected RNA fragments into a DNA library that can be analyzed by deep sequencing.[89] The addition of harringtonine (an alkaloid that inhibits protein biosynthesis) causes ribosomes to accumulate precisely at initiation codons and assists in their detection.



A schematic overview of Ribo-Seq/ARTseq.

| Advantages | Disadvantages |
| --- | --- |
| • Reveals a snapshot with the precise location of ribosomes on the RNA<br>• More closely reflects the rate of protein synthesis than mRNA levels<br>• No prior knowledge of the RNA or open-reading frames (ORFs) is required<br>• Surveys the whole genome<br>• Can identify protein-coding regions | • Initiation from multiple sites within a single transcript makes it challenging to define all ORFs<br>• Does not provide the kinetics of translational elongation |

## References

Zur H., Aviner R. and Tuller T. Complementary Post Transcriptional Regulatory Information is Detected by PUNCH-P and Ribosome Profiling. *Sci Rep.* 2016;6:21635

The authors used a combination of Ribo-Seq and PUNCH-P, which is based on mass-spectrometric analysis of only newly synthesized proteins. They found a significant overlap in the results obtained with the 2 approaches; however, combining them provided a better prediction of the steady-state changes in protein abundance.

Illumina Technology: Unspecified Illumina sequencing system

Arribere J. A., Cenik E. S., Jain N., et al. Translation readthrough mitigation. *Nature*. 2016;534:719-723

The researchers used a combination of RNA-Seq, Ribo-Seq, and gene editing with a clustered regularly interspaced short palindromic repeats (CRISPR)–Cas9 system in Caenorhabditis elegans to show how cells are able to block the accumulation of C-terminal–extended proteins that result from failure to terminate at stop codons. They found that 3'-UTRs may encode peptide sequences that destabilize the attached protein, to limit the impact of translation errors.

Illumina Technology: MiSeq System

Jeong Y., Kim J. N., Kim M. W., et al. The dynamic transcriptional and translational landscape of the model antibiotic producer Streptomyces coelicolor A3(2). *Nat Commun.* 2016;7:11605

Streptomyces species are used extensively to produce antibiotics and other bioactive compounds. However, these products are often not detectable under laboratory culture conditions, and the gene networks that produce the compounds are relatively unknown. The authors used a combination of single-stranded RNA-Seq (ssRNA-Seq) and Ribo-Seq to explore the transcriptional and translational regulatory networks at various growth phases. They identified 230 small RNAs and a considerable proportion (~21%) of leaderless mRNAs. Based on ribosome profiling, the translation efficiency of secondary metabolic genes appears to be correlated negatively with transcription. In addition, several important antibiotic regulatory genes are induced during translation.

Illumina Technology: MiSeq System

---

88. Ingolia N. T., Ghaemmaghami S., Newman J. R. and Weissman J. S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*. 2009;324:218-223

89. Ingolia N. T., Lareau L. F. and Weissman J. S. (2011) Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. *Cell* 147: 789-802

Nissley D. A., Sharma A. K., Ahmed N., et al. Accurate prediction of cellular co-translational folding indicates proteins can switch from post- to co-translational folding. *Nat Commun*. 2016;7:10341

The authors developed a model to predict how varying the translation rate at different codon positions along a transcript's coding sequence affects this self-assembly process. During the testing of the model, they performed Ribo-Seq experiments in yeast to confirm that stationary ribosome profile distributions occur between biological replicates, consistent with translation occurring at a steady state.

Illumina Technology: HiSeq 2000 System

Bennett C. G., Riemondy K., Chapnick D. A., Bunker E., Liu X., et al. Genome-wide analysis of Musashi-2 targets reveals novel functions in governing epithelial cell migration. *Nucleic Acids Res.* 2016;44:3788-3800

Dar D., Shamir M., Mellin J. R., et al. Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. *Science*. 2016;352:aad9822

Gawron D., Ndah E., Gevaert K. and Van Damme P. Positional proteomics reveals differences in N-terminal proteoform stability. *Mol Syst Biol.* 2016;12:858

Lin L., Jiang P., Park J. W., et al. The contribution of Alu exons to the human proteome. *Genome Biol.* 2016;17:15

Olexiouk V., Crappe J., Verbruggen S., Verhegen K., Martens L. and Menschaert G. sORFs.org: a repository of small ORFs identified by ribosome profiling. *Nucleic Acids Res.* 2016;44:D324-329

Sin C., Chiarugi D. and Valleriani A. Quantitative assessment of ribosome drop-off in E. coli. *Nucleic Acids Res.* 2016;44:2528-2537

Singh A. R., Sivadas A., Sabharwal A., et al. Chamber Specific Gene Expression Landscape of the Zebrafish Heart. *PLoS One.* 2016;11:e0147823

Tichon A., Gil N., Lubelsky Y., et al. A conserved abundant cytoplasmic long noncoding RNA modulates repression by Pumilio proteins in human cells. *Nat Commun*. 2016;7:12209

## Associated Kits

TruSeq Ribo Profile Kit

Ribo-Zero rRNA Removal Kits

# RIP-Seq: RNA Immunoprecipitation Sequencing

RIP-Seq maps the sites at which proteins are bound to the RNA within RNA-protein complexes.[90]

In this method, RNA-protein complexes are immunoprecipitated with antibodies targeted to the protein of interest. After RNase digestion, RNA protected by protein binding is extracted and reverse-transcribed to cDNA. The locations can then be mapped back to the genome. Deep sequencing of cDNA provides single-base resolution of protein-bound RNA.



RNA-protein complex → Immunoprecipitate RNA-protein complex → RNase digestion → RNA extraction → Reverse-transcription → cDNA

A schematic overview of RIP-seq.

| Advantages | Disadvantages |
| --- | --- |
| • Maps specific protein-RNA complexes, such as polycomb-associated RNAs<br>• Low background and higher resolution of binding site due to RNase digestion<br>• No prior knowledge of the RNA is required<br>• Genome-wide RNA screen | • Requires antibodies to the targeted proteins<br>• Nonspecific antibodies will precipitate nonspecific complexes<br>• Lack of crosslinking or stabilization of the complexes may lead to false negatives<br>• RNase digestion must be controlled carefully |

## Reviews

McFadden E. J. and Hargrove A. E. Biochemical Methods To Investigate lncRNA and the Influence of lncRNA:Protein Complexes on Chromatin. *Biochemistry*. 2016;55:1615-1630

Fang Y. and Fullwood M. J. Roles, Functions, and Mechanisms of Long Non-coding RNAs in Cancer. *Genomics Proteomics Bioinformatics.* 2016;14:42-54

Kim T. K. and Shiekhattar R. Diverse regulatory interactions of long noncoding RNAs. *Curr Opin Genet Dev.* 2016;36:73-82

Goeman F. and Fanciulli M. Application of RNA-Seq Technology in Cancer Chemoprevention. Cancer Chemoprevention: *Methods and Protocols*. 2016;31-43

Carrier M.-C., Lalaouna D. and Massé E. A game of tag: MAPS catches up on RNA interactomes. *RNA Biology.* 2016;13:473-476

## References

Cloonan S. M., Glass K., Laucho-Contreras M. E., et al. Mitochondrial iron chelation ameliorates cigarette smoke-induced bronchitis and emphysema in mice. *Nat Med.* 2016;22:163-174

The researchers combined RIP-seq, RNA-Seq, and functional enrichment clustering analysis to identify Irp2 as a regulator of mitochondrial function in the lungs of mice.

Illumina Technology: HiSeq 2000 System

Yabe-Wada T., Matsuba S., Takeda K., et al. TLR signals posttranscriptionally regulate the cytokine trafficking mediator sortilin. *Sci Rep.* 2016;6:26566

The researchers used RIP-Seq analysis to identify mRNAs that were recognized by poly(rC)-binding protein 1 (PCBP1) and to define elements bound by PCBP1. Motif analysis revealed significant enrichment of a CCCCNCCCCC motif in the RIP sample, which indicated that PCBP1 was bound to mRNAs with this sequence motif.

Illumina Technology: HiSeq 1500 System

Dugar G., Svensson S. L., Bischler T., et al. The CsrA-FliW network controls polar localization of the dual-function flagellin mRNA in Campylobacter jejuni. *Nat Commun.* 2016;7:11667

Ennajdaoui H., Howard J. M., Sterne-Weiler T., et al. IGF2BP3 Modulates the Interaction of Invasion-Associated Transcripts with RISC. *Cell Rep.* 2016;15:1876-1883

Faraji F., Hu Y., Yang H. H., et al. Post-transcriptional Control of Tumor Cell Autonomous Metastatic Potential by CCR4-NOT Deadenylase CNOT7. *PLoS Genet.* 2016;12:e1005820

## Associated Kits

| | | |
| --- | --- | --- |
| TruSeq Ribo Profile Kit | TruSeq Small RNA Library Prep Kit | TruSeq Targeted RNA Expression Library Prep Kit |
| Ribo-Zero rRNA Removal Kits | TruSeq Stranded mRNA Library Prep Kit | |
| TruSeq RNA Library Prep Kit v2 | TruSeq Stranded Total RNA Library Prep Kit | |

---

90.  Zhao J., Ohsumi T. K., Kung J. T., et al. Genome-wide identification of polycomb-associated RNAs by RIP-seq. Mol Cell. 2010;40:939-953
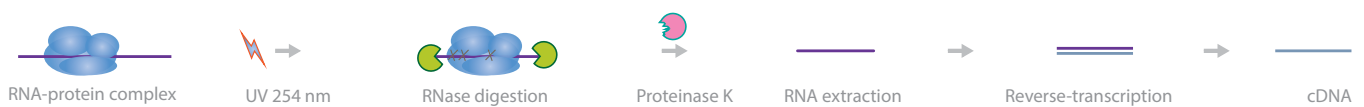
# CLIP-Seq or HITS-CLIP: High-Throughput Sequencing of CLIP cDNA Library

Crosslinking and immunoprecipitation (CLIP), with the use of RNase T1 trimming, was first described by Ule *et al*.[91] and later applied to high-throughput sequencing to map protein-RNA binding sites *in vivo*.[92-93] This approach is similar to RIP-Seq but uses crosslinking to stabilize the protein-RNA complexes.

In HITS-CLIP, RNA-protein complexes are UV-crosslinked and immunoprecipitated. The protein-RNA complexes are treated with RNase T1, followed by proteinase K. RNA is extracted and reverse-transcribed to cDNA. Deep sequencing of the cDNA provides single-base resolution mapping of protein binding sites on RNA.

An improvement on the HITS-CLIP protocol was published by Gillen et al., which reduces artifacts from mispriming occurences.[94]

Other versions: iCLIP, irCLIP, eCLIP, miCLIP



| RNA-protein complex | UV 254 nm | RNase digestion | Proteinase K | RNA extraction | Reverse-transcription | cDNA |

A schematic overview of HITS-CLIP.

| Advantages | Disadvantages |
| --- | --- |
| • Crosslinking stabilizes the protein-target binding<br>• UV crosslinking can be carried out in vivo<br>• Provides low background and higher resolution of binding site, due to RNase digestion<br>• No prior knowledge of the RNA is required<br>• Genome-wide RNA screen | • Over-representation of the RT complement due to mispriming[95]<br>• Antibodies not specific to the target may precipitate nonspecific complexes.<br>• UV crosslinking is not efficient and requires close protein-RNA interactions<br>• Artifacts may be introduced during the crosslinking process |

## Reviews
Cook K. B., Hughes T. R. and Morris Q. D. High-throughput characterization of protein-RNA interactions. Brief Funct *Genomics*. 2015;14:74-89

## References
Van Haute L., Dietmann S., Kremer L., et al. Deficient methylation and formylation of mt-tRNA(Met) wobble cytosine in a patient carrying mutations in NSUN3. *Nat Commun.* 2016;7:12039

Epigenetic modifications on the transcriptome are associated with human diseases. This study identified RNA targets for 5-methylcytosine (m5C) methyltransferase, NSun3, and the effects of m5C modifications on energy metabolism in samples with mitochondrial respiratory chain complex deficiency. The authors used methylation individual-nucleotide-resolution CLIP (miCLIP) to screen the transcriptome for NSun3-methylated nucleosides. In addition, they used HITS-CLIP in mitochondrial lysates of HEK293T cells to provide greater insights into NSun3 mitochondrial RNA targets.

Illumina Technology: HiSeq 2500 System, HiSeq 2000 System, MiSeq System, Ribo-Zero Gold rRNA Removal Kit

Vourekas A., Alexiou P., Vrettos N., Maragkakis M. and Mourelatos Z. Sequence-dependent but not sequence-specific piRNA adhesion traps mRNAs to the germ plasm. *Nature*. 2016;531:390-394

The researchers used Argonaute (AGO) HITS-CLIP to determine the processes involved in the transport of mRNAs to the oocyte by mid-oogenesis in Drosophila. They found that that Aubergine (Aub)-loaded piRNAs use partial base-pairing characteristics of AGO RNPs to bind mRNAs randomly in Drosophila, acting as an adhesive trap that captures mRNAs in the germplasm.

Illumina Technology: HiSeq System

91.    Ule J., Jensen K., Mele A. and Darnell R. B. CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods*. 2005;37:376-386
92.    Licatalosi D. D., Mele A., Fak J. J., et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature*. 2008;456:464-469
93.    Chi SW, Zang JB, Mele A, Darnell RB; (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 460: 479-86
94.    Gillen A. E., Yamamoto T. M., Kline E., Hesselberth J. R. and Kabos P. Improvements to the HITS-CLIP protocol eliminate widespread mispriming artifacts. *BMC Genomics*. 2016;17:338
95.    Gillen A. E., Yamamoto T. M., Kline E., Hesselberth J. R. and Kabos P. Improvements to the HITS-CLIP protocol eliminate widespread mispriming artifacts. *BMC Genomic*s. 2016;17:338

**Ji Z., Song R., Huang H., Regev A. and Struhl K. Transcriptome-scale RNase-footprinting of RNA-protein complexes.** *Nat Biotechnol.* **2016;34:410-413**

This publication describes the use of HITS-CLIP and a computational pipeline, Rfoot, to precisely map RNase-protected regions within small nucleolar RNAs, spliceosomal RNAs, miRNAs, tRNAs, lncRNAs, and 3'-UTRs of mRNAs in human cells. The authors determined that lncRNAs have multiple RNase-protected footprints, which are conserved evolutionarily.

Illumina Technology: HiSeq 2500 System

**Gosline S. J., Gurtan A. M., JnBaptiste C. K., et al. Elucidating MicroRNA Regulatory Networks Using Transcriptional, Post-transcriptional, and Histone Modification Measurements.** *Cell Rep.* **2016;14:310-319**

The researchers used RNA-Seq and CLIP-Seq to characterize post-transcriptional and epigenetic activity of miRNA-regulatory networks systematically in a pair of murine fibroblast cell lines, with and without Dicer expression. They found that most of the changes induced by global miRNA loss occur at transcription. This study provides an example of analyzing complex networks by combining diverse datasets.

Illumina Technology: HiSeq 2000 System

**Nutter C. A., Jaworski E. A., Verma S. K., et al. Dysregulation of RBFOX2 Is an Early Event in Cardiac Pathogenesis of Diabetes.** *Cell Rep.* **2016;15:2200-2213**

The genome-wide CLIP-Seq of human ESCs provides a genome-wide map of binding sites for the RBP RBFOX2, a key regulator of alternative exon splicing in the nervous system and other cell types. The authors showed that RBFOX2 controls the alternative splicing and transcriptome changes of genes associated with diabetic cardiomyopathy.

Illumina Technology: HiSeq 1000 System

Spengler R. M., Zhang X., Cheng C., et al. Elucidation of transcriptome-wide microRNA binding sites in human cardiac tissues by Ago2 HITS-CLIP. *Nucleic Acids Res.* 2016;

Gillen A. E., Yamamoto T. M., Kline E., Hesselberth J. R. and Kabos P. Improvements to the HITS-CLIP protocol eliminate widespread mispriming artifacts. *BMC Genomics.* 2016;17:338

Moore M. J., Zhang C., Gantman E. C., Mele A., Darnell J. C. and Darnell R. B. Erratum: Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis. *Nat Protoc.* 2016;11:616

Bennett C. G., Riemondy K., Chapnick D. A., Bunker E., Liu X., et al. Genome-wide analysis of Musashi-2 targets reveals novel functions in governing epithelial cell migration. *Nucleic Acids Res.* 2016;44:3788-3800

Vourekas A., Alexiou P., Vrettos N., Maragkakis M. and Mourelatos Z. Sequence-dependent but not sequence-specific piRNA adhesion traps mRNAs to the germ plasm. *Nature*. 2016;531:390-394

Zheng Q., Bao C., Guo W., et al. Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs. *Nat Commun.* 2016;7:11215

Preusse M., Theis F. J. and Mueller N. S. miTALOS v2: Analyzing Tissue Specific microRNA Function. *PLoS One*. 2016;11:e0151771

Kapeli K., Pratt G. A., Vu A. Q., et al. Distinct and shared functions of ALS-associated proteins TDP-43, FUS and TAF15 revealed by multisystem analyses. *Nat Commun.* 2016;7:12143

Valentin-Vega Y. A., Wang Y. D., Parker M., et al. Cancer-associated DDX3X mutations drive stress granule assembly and impair global translation. *Sci Rep*. 2016;6:25996

Chen S., Blank M. F., Iyer A., et al. SIRT7-dependent deacetylation of the U3-55k protein controls pre-rRNA processing. *Nat Commun.* 2016;7:10734

Zhao H., Chen M., Lind S. B. and Pettersson U. Distinct temporal changes in host cell lncRNA expression during the course of an adenovirus infection. *Virology.* 2016;492:242-250

Taliaferro J. M., Vidaki M., Oliveira R., et al. Distal Alternative Last Exons Localize mRNAs to Neural Projections. *Mol Cell.* 2016;61:821-833

Gaudreau M. C., Grapton D., Helness A., et al. Heterogeneous Nuclear Ribonucleoprotein L is required for the survival and functional integrity of murine hematopoietic stem cells. *Sci Rep.* 2016;6:27379

Dugar G., Svensson S. L., Bischler T., Waldchen S., Reinhardt R., et al. The CsrA-FliW network controls polar localization of the dual-function flagellin mRNA in Campylobacter jejuni. *Nat Commun.* 2016;7:11667

## Associated Kits

TruSeq Ribo Profile Kit

Ribo-Zero rRNA Removal Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

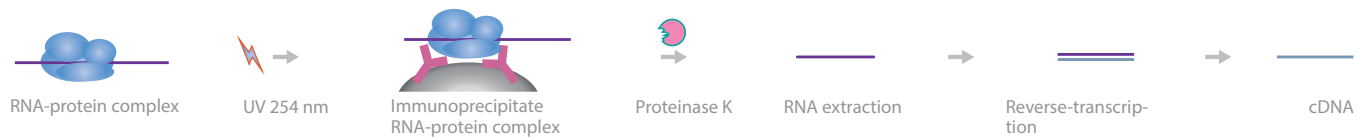TruSeq Stranded mRNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

# Pol II CLIP: Crosslinking and Immunoprecipitation of RNA Polymerase II

Pol II CLIP was developed to isolate and sequence noncoding RNAs involved in transcriptional regulation, particularly RNAPII-mediated transcription.

Briefly, cells are UV-crosslinked at 254 nm and lysed. Anti-RNAPII antibodies are used to immunoprecipitate the RNAPII complexes. Next, the RNAPII complexes are separated via proteinase K reverse-crosslinking. The released RNA released is isolated, reverse-transcribed, and prepared into sequencing libraries. The CLIP RNA samples are ready for sequencing.



| RNA-protein complex | UV 254 nm | Immunoprecipitate RNA-protein complex | Proteinase K | RNA extraction | Reverse-transcription | cDNA |

A schematic overview of Pol II CLIP.

| Advantages | Disadvantages |
| --- | --- |
| • Identifies ncRNA sequences involved in transcriptional regulation<br>• Simple procedure | • None reported yet |

## Reviews
Qu S., Yang X., Li X., Wang J., Gao Y., et al. Circular RNA: A new star of noncoding RNAs. *Cancer Lett.* 2015;365:141-148

## References
Li Z., Huang C., Bao C., Chen L., Lin M., et al. Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol.* 2015;

The authors developed Pol II CLIP to identify the role of a novel type of circRNA, named exon-intron circRNAs (ElciRNAs). They found that ElciRNAs were localized predominantly in the nucleus, interacted with U1 snRNP, and promoted transcription of their parental genes. ElciRNAs also increased expression of their parental genes in *cis*.

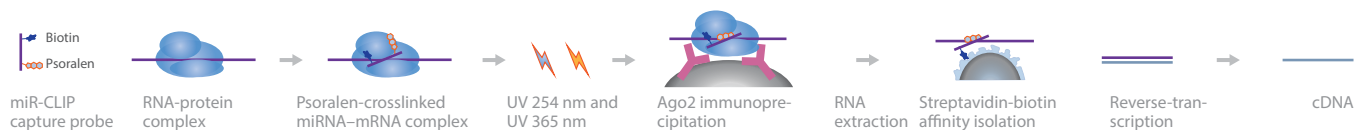Illumina Technology: Genome Analyzer System

## Associated Kits
TruSeq RNA Library Prep Kit v2

---

96.    Li Z., Huang C., Bao C., et al. Exon-intron circular RNAs regulate transcription in the nucleus. Nat Struct Mol Biol. 2015;

# miR-CLIP: MicroRNA Crosslinking and Immunoprecipitation

miR-CLIP was developed to identify mRNA targets of miRNA, using synthetic miRNAs.[97] The capture probes in miR-CLIP are pre-miRNAs conjugated to psoralen and biotin that can be crosslinked with mRNA to isolate their targets from total RNA.

First, a miR-CLIP capture probe needs to be designed and tested for functionality and strong binding affinity. Next, they are transfected into cells and UV-crosslinked at 254 nm and 365 nm, followed by immunoprecipitation of the crosslinked complex. The RNA is extracted and further purified by streptavidin pull-down. The isolated RNA targets can now be reverse-transcribed and processed into cDNA libraries for sequencing.



A schematic overview of miR-CLIP.

| Advantages | Disadvantages |
|---|---|
| • Identifies targets of miRNAs genome-wide<br>• Capture probe miRNA sequences can be easily modified to investigate different miRNA species | • Not yet adopted widely by the scientific community<br>• Requires careful design of miRNA capture probes |

## Reviews

Sandhu G. K., Milevskiy M. J. G., Wilson W., Shewan A. M. and Brown M. A. Non-coding RNAs in Mammary Gland Development and Disease. *Non-coding RNA and the Reproductive System.* 2016;121-153

Yu B. and Shan G. Functions of long noncoding RNAs in the nucleus. *Nucleus.* 2016;7:155-166

Jeker L. T. and Marone R. Targeting microRNAs for immunomodulation. *Curr Opin Pharmacol.* 2015;23:25-31

Tycowski K. T., Guo Y. E., Lee N., et al. Viral noncoding RNAs: more surprises. *Genes Dev.* 2015;29:567-584

## References

Imig J., Brunschweiger A., Brummer A., Guennewig B., Mittal N., et al. miR-CLIP capture of a miRNA targetome uncovers a lincRNA H19-miR-106a interaction. *Nat Chem Biol.* 2014;advance online publication:

The authors investigated the effects of miRNA on post-transcriptional gene regulation in human cell lines. They developed miR-CLIP to identify prospective targets of miR-106a and assessed the effects of overexpression on its functional targets. The results revealed that miR-106a and other members of its family of miRNAs bind to the H19 lncRNA. Further, the H19 lncRNA behaves as a buffer in regulating the gene expression of H19, the miR-17-5p family, and their mRNA targets.

Illumina Technology: Unspecified Illumina sequencing system

## Associated Kits

TruSeq RNA Library Prep Kit

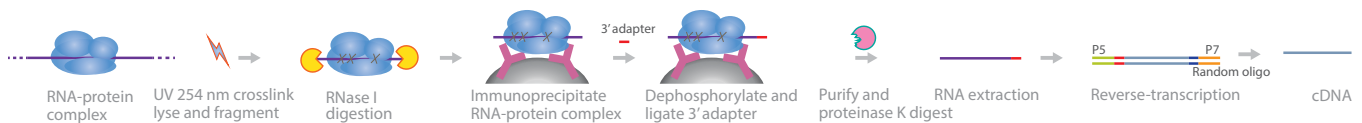TruSeq Small RNA Library Prep Kit

Ribo-Zero Gold rRNA Removal Kits

97.  Imig J., Brunschweiger A., Brummer A., et al. miR-CLIP capture of a miRNA targetome uncovers a lincRNA H19-miR-106a interaction. *Nat Chem Biol.* 2014;advance online publication:

# eCLIP: Enhanced Cross-linking Immunoprecipitation

eCLIP maps the binding sites of RBPs on their target RNAs using a modified individual nucleotide resolution CLIP (iCLIP) protocol, improving efficiency and decreasing execution complexity.[98] The hallmark of this method is the ligation of barcoded single-stranded DNA adapters, which reduce amplification bias significantly.

First, RNA and the protein of interest are UV-crosslinked, followed by cell lysis and RNase I digestion. Next, the protein-RNA complexes are immunoprecipitated and ligated to an RNA adapter on the 3' end of the target RNA. The bound protein is removed by proteinase K digestion, and the RNA is reverse-transcribed. The resulting cDNA is ligated to single-stranded DNA adapters on the 3' end that contain either an N5 or N10 sequence to serve as unique identifiers against PCR duplicates. Finally, the paired-end cDNA fragments are amplified and sequenced.

Similar methods: iCLIP, irCLIP, HITS-CLIP



A schematic overview of eCLIP.

| Advantages | Disadvantages |
|---|---|
| • High-throughput mapping of protein-RNA binding sites<br>• Barcoded adapters significantly reduce PCR duplicate reads and improve throughput<br>• Improved ligation efficiency by ~1000-fold<br>• Avoids usage of radioactive markers | • Antibodies not specific to the target will precipitate nonspecific complexes<br>• Crosslinking does not cover all RNA-binding domains[99] |

## Reviews

Marchese D., de Groot N. S., Lorenzo Gotor N., Livi C. M. and Tartaglia G. G. Advances in the characterization of RNA-binding proteins. *Wiley Interdisciplinary Reviews: RNA.* 2016;7:793-810

Bangru S. and Kalsotra A. Advances in analyzing RNA diversity in eukaryotic transcriptomes: peering through the Omics lens. *F1000Research*. 2016;5:2668

Manning K. S. and Cooper T. A. The roles of RNA processing in translating genotype to phenotype. *Nat Rev Mol Cell Biol.* 2016;advance online publication:

Sundararaman B., Zhan L., Blue S. M., et al. Resources for the Comprehensive Discovery of Functional RNA Elements. *Mol Cell.* 2016;61:903-913

## References

Conway Anne E., Van Nostrand Eric L., Pratt Gabriel A., et al. Enhanced CLIP Uncovers IMP Protein-RNA Targets in Human Pluripotent Stem Cells Important for Cell Adhesion and Survival. *Cell Reports.* 2016;

The authors explored the intricate mechanisms of post-transcriptional RNA networks for maintaining proliferation and survival in human pluripotent stem cells (hPSCs). Using eCLIP to identify the RNA targets of the IMP/IGF2BP family of RBPs, they found that IMP1 and IMP2 bound to a large number of 3'-UTR-enriched targets. They also used RNA Bind-N-Seq with recombinant full-length IMP1 and IMP2 to identify CA-rich motifs that were enriched in binding sites revealed by eCLIP.

Illumina Technology: Unspecified Illumina sequencing system

## Associated Kits

TruSeq Ribo Profile Kit

Ribo-Zero Gold rRNA Removal Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

98. Van Nostrand E. L., Pratt G. A., Shishkin A. A., et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods.* 2016;
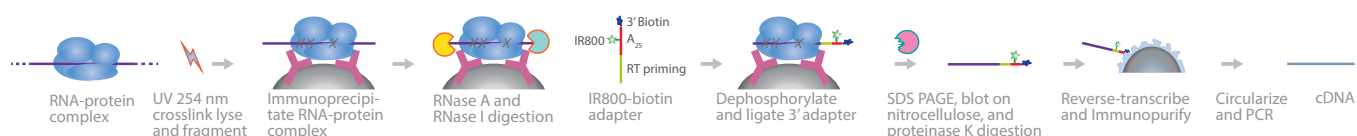
99. Martin G. and Zavolan M. Redesigning CLIP for efficiency, accuracy and speed. *Nat Methods.* 2016;13:482-483

# irCLIP: UV-C Crosslinking and Immunoprecipitation

irCLIP maps protein-RNA interaction sites using less sample material, time, and increased cDNA library quality compared to previous CLIP methods.[100] irCLIP was designed to tackle the issues in both iCLIP and HITS-CLIP, such as reverse-transcriptase halting and short cDNA library fragments, by using on-bead nuclease digestion.

First, RNA-protein complexes are UV-crosslinked and immunoprecipitated. Next, they are digested using RNase A and RNase I to excise both ends of the protein-bound RNA strand. An IR800-biotin adapter is ligated to the 3' end of the RNA fragment. After size selection, nitrocellulose blotting, and proteinase K digestion, the RNA strands of interest are purified. They are reverse-transcribed and immunopurified once again. The cDNA strands are circularized and PCR-amplified to produce cDNA libraries for sequencing.

Other versions: iCLIP, HITS-CLIP, eCLIP



A schematic overview of irCLIP..

| Advantages | Disadvantages |
|---|---|
| • Nonisotopic detection of protein-RNA interactions with minimal starting material<br>• On-bead nuclease digestion improves the length of RNA fragments for cDNA library prep<br>• Able to accommodate small cell samples and reveal novel RBP binding sites<br>• Uses thermostable reverse transcriptase at 60°C, reducing biases by melting secondary RNA structures[101] | • Not yet adopted widely by the scientific community<br>• Circularization step may introduce artifacts |

## Reviews

Martin G. and Zavolan M. Redesigning CLIP for efficiency, accuracy and speed. *Nat Meth.* 2016;13:482-483

Haque N. and Hogg J. R. Easier, Better, Faster, Stronger: Improved Methods for RNA-Protein Interaction Studies. *Mol Cell.* 2016;62:650-651

## References

Zarnegar B. J., Flynn R. A., Shen Y., Do B. T., Chang H. Y., et al. irCLIP platform for efficient characterization of protein-RNA interactions. *Nat Methods*. 2016;13:489-492

The authors developed irCLIP as an improvement over existing CLIP methods. They addressed issues in reverse-transcriptase halting and biases from secondary RNA structures by using on-bead nuclease digestion and a reverse transcriptase that is thermostable at 60°C. They used irCLIP in HeLa cells and were able to produce similar or even higher unique reads using far fewer cells (20,000 and 100,000 cells per immunoprecipitation) than iCLIP results (5,000,000 cells per immunoprecipitation).

Illumina Technology: NextSeq 500 System, TruSeq RNA Library Prep Kit

## Associated Kits
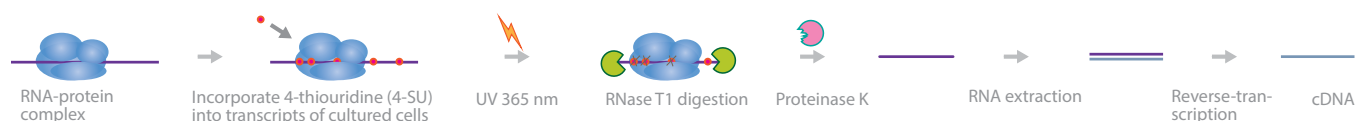
TruSeq RNA Library Prep Kit

---

100. Zarnegar B. J., Flynn R. A., Shen Y., Do B. T., Chang H. Y. and Khavari P. A. irCLIP platform for efficient characterization of protein-RNA interactions. *Nat Methods*. 2016;13:489-492
101. Martin G. and Zavolan M. Redesigning CLIP for efficiency, accuracy and speed. *Nat Meth.* 2016;13:482-483
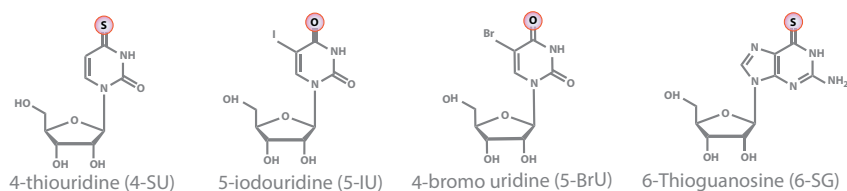
# PAR-CLIP: Photoactivatable Ribonucleoside–Enhanced Crosslinking and Immunoprecipitation

PAR-CLIP maps RBP sites on the target RNAs.[102] This approach is similar to HITS-CLIP and CLIP-Seq, but it uses much more efficient crosslinking to stabilize the protein-RNA complexes. The requirement to introduce a photoactivatable ribonucleoside limits this approach to cell culture and *in vitro* systems.

In this method, 4-SU and 6-thioguanosine (6-SG) are incorporated into transcripts of cultured cells. UV irradiation crosslinks 4-SU/6-SG–labeled transcripts to interacting RBPs. The targeted complexes are immunoprecipitated and digested with RNase T1, followed by proteinase K, before RNA extraction. The RNA is reverse-transcribed to cDNA and sequenced. Deep sequencing of cDNA accurately maps RBPs interacting with labeled transcripts.



| RNA-protein complex | Incorporate 4-thiouridine (4-SU) into transcripts of cultured cells | UV 365 nm | RNase T1 digestion | Proteinase K | RNA extraction | Reverse-tran-scription | cDNA |

A schematic overview of PAR-CLIP.



4-thiouridine (4-SU)    5-iodouridine (5-IU)    4-bromo uridine (5-BrU)    6-Thioguanosine (6-SG)

Photoactivatable ribonucleosides

| Advantages | Disadvantages |
|---|---|
| • Highly accurate mapping of RNA-protein interactions<br>• Labeling with 4-SU/6-SG improves crosslinking efficiency<br>• Legitimately crosslinked sequences can be identified based on the presence of mutated bases, and mispriming can be filtered bioinformatically[103] | • Antibodies not specific to the target may precipitate nonspecific complexes<br>• Limited to cell culture and *in vitro* systems<br>• Photoreactive nucleosides can be cytotoxic[104]<br>• High concentrations of 4-SU can inhibit rRNA synthesis and induce a nucleolar stress response[105] |

## Reviews

Cook K. B., Hughes T. R. and Morris Q. D. High-throughput characterization of protein-RNA interactions. Brief Funct *Genomics*. 2015;14:74-89

Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38:226-236

102.  Hafner M., Landgraf P., Ludwig J., et al. Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods*. 2008;44:3-12
103.  Gillen A. E., Yamamoto T. M., Kline E., Hesselberth J. R. and Kabos P. Improvements to the HITS-CLIP protocol eliminate widespread mispriming artifacts. *BMC Genomics*. 2016;17:338
104.  Kloetgen A., Munch P. C., Borkhardt A., Hoell J. I. and McHardy A. C. Biochemical and bioinformatic methods for elucidating the role of RNA-protein interactions in posttranscriptional regulation. Brief Funct *Genomics*. 2015;14:102-114
105.  Burger K., Muhl B., Kellner M., et al. 4-thiouridine inhibits rRNA synthesis and causes a nucleolar stress response. *RNA Biol.* 2013;10:1623-1630

# References

Liu N., Dai Q., Zheng G., He C., Parisien M. and Pan T. N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature*. 2015;518:560-564

Heterogeneous nuclear ribonucleoprotein C (HNRNPC) is an abundant nuclear RBP responsible for pre-mRNA processing. To map the $N^6$-methyladenosine (m$^6$A) sites around HNRNPC-binding sites, the researchers performed PAR-CLIP to isolate all HNRNPC-bound RNA regions. They then used anti-m$^6$A immunoprecipitation (MeRIP) to identify 39,060 m$^6$A switches among the HNRNPC-binding sites.

Illumina Technology: HiSeq 2000 System

Murakawa Y., Hinz M., Mothes J., et al. RC3H1 post-transcriptionally regulates A20 mRNA and modulates the activity of the IKK/NF-kappaB pathway. N*at Commun.* 2015;6:7367

The researchers used PAR-CLIP to identify ~3,800 mRNA targets with 416,000 binding sites for the human RBP, RC3H1. RC3H1 (also known as ROQUIN) promotes TNF-α mRNA decay via a 3'-UTR constitutive decay element (CDE). A large number of the sites recovered were distinct from the consensus CDE and revealed a structure-sequence motif with U-rich sequences embedded in hairpins.

Illumina Technology: HiSeq 2000 System

Matveeva E., Maiorano J., Zhang Q., et al. Involvement of PARP1 in the regulation of alternative splicing. *Cell Discovery*. 2016;2:15046

The authors showed that the PAR-CLIP product was precipitated with PARP1 antibodies. This provided additional proof that PARP1 may act as an adapter to locate RNA in specific regions of the chromatin.

Illumina Technology: HiSeq 2500 System


Tichon A., Gil N., Lubelsky Y., Havkin Solomon T., Lemze D., et al. A conserved abundant cytoplasmic long noncoding RNA modulates repression by Pumilio proteins in human cells. *Nat Commun*. 2016;7:12209

Vongrad V., Imig J., Mohammadi P., et al. HIV-1 RNAs are Not Part of the Argonaute 2 Associated RNA Interference Pathway in Macrophages. *PLoS One*. 2015;10:e0132127

Kang D., Skalsky R. L. and Cullen B. R. EBV BART MicroRNAs Target Multiple Pro-apoptotic Cellular Genes to Promote Epithelial Cell Survival. *PLoS Pathog*. 2015;11:e1004979

Porter D. F., Koh Y. Y., VanVeller B., Raines R. T. and Wickens M. Target selection by natural and redesigned PUF proteins. *Proc Natl Acad Sci U S A*. 2015;112:15868-15873

Wang Q., Taliaferro J. M., Klibaite U., Hilgers V., Shaevitz J. W. and Rio D. C. The PSI-U1 snRNP interaction regulates male mating behavior in Drosophila. *Proc Natl Acad Sci U S A.* 2016;113:5269-5274

Lee A. S., Kranzusch P. J. and Cate J. H. eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. Nature. 2015;522:111-114

Chu Y., Wang T., Dodd D., Xie Y., Janowski B. A. and Corey D. R. Intramolecular circularization increases efficiency of RNA sequencing and enables CLIP-Seq of nuclear RNA from human cells. *Nucleic Acids Res*. 2015;43:e75

Xie H., Lee L., Scicluna P., et al. Novel functions and targets of miR-944 in human cervical cancer cells. *Int J Cancer.* 2015;136:E230-241

Xiong X. P., Vogler G., Kurthkoti K., Samsonova A. and Zhou R. SmD1 Modulates the miRNA Pathway Independently of Its Pre-mRNA Splicing Function. *PLoS Genet.* 2015;11:e1005475

Chen T., Hao Y. J., Zhang Y., et al. m(6)A RNA Methylation Is Regulated by MicroRNAs and Promotes Reprogramming to Pluripotency. *Cell Stem Cell.* 2015;16:289-301

## Associated Kits

TruSeq Ribo Profile Kit

Ribo-Zero rRNA Removal Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit
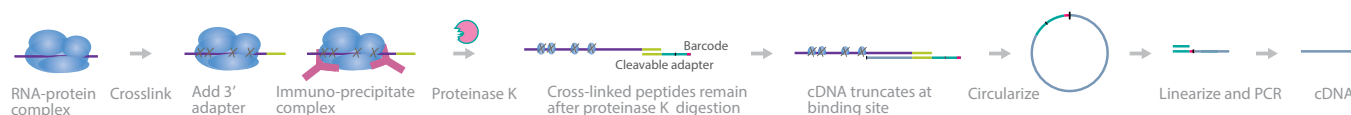
TruSeq Stranded Total RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

# iCLIP: Individual Nucleotide Resolution CLIP

iCLIP maps protein-RNA interactions, in a process similar to HITS-CLIP and PAR-CLIP.[106] This approach includes additional steps to digest the proteins after crosslinking and to map the crosslink sites with reverse transcriptase.

In iCLIP, specific crosslinked RNA-protein complexes are immunoprecipitated. The complexes are treated with proteinase K, as the protein crosslinked at the binding site remains undigested. Upon RT, cDNA truncates at the binding site and is circularized. These circularized fragments are linearized and PCR-amplified. Deep sequencing of these amplified fragments provides nucleotide resolution of the protein-binding site  eCLIP is an improvement over iCLIP that avoids circularizing the cDNA to reduce artifacts.[107]

Other versions: HITS-CLIP, PAR-CLIP, eCLIP, irCLIP



A schematic overview of iCLIP.

| Advantages | Disadvantages |
|---|---|
| • Nucleotide resolution of protein-binding sites<br>• Avoids the use of nucleases<br>• Amplification allows the detection of rare events | • Antibodies not specific to target will precipitate nonspecific complexes<br>• Nonlinear PCR amplification can lead to biases affecting reproducibility<br>• Artifacts may be introduced in the circularization step<br>• Radioisotopes required for visualization of UV-crosslinked protein-RNA complexes only label 5' ends[108]<br>• Circular ligation can be inefficient[109] |

## Reviews
Cook K. B., Hughes T. R. and Morris Q. D. High-throughput characterization of protein-RNA interactions. *Brief Funct Genomics*. 2015;14:74-89

## References
Ennajdaoui H., Howard J. M., Sterne-Weiler T., Jahanbani F., Coyne D. J., et al. IGF2BP3 Modulates the Interaction of Invasion-Associated Transcripts with RISC. *Cell Rep*. 2016;15:1876-1883

The authors studied insulin-like growth factor 2 mRNA-binding protein 3 (IGF2BP3), an RBP that promotes association of the RISC with specific transcripts. They used iCLIP to show that there are significant overlaps in binding sites for IGF2BP3 and miRNAs. Their results indicate that IGF2BP3 influences a malignancy-associated RNA regulon by modulating miRNA-mRNA interactions.

Illumina Technology: HiSeq 2500 System

Sutandy F. X., Hildebrandt A. and Konig J. Profiling the Binding Sites of RNA-Binding Proteins with Nucleotide Resolution Using iCLIP. *Methods Mol Biol*. 2016;1358:175-195

Misra A., Ou J., Zhu L. J. and Green M. R. Global analysis of CPSF2-mediated alternative splicing: Integration of global iCLIP and transcriptome profiling data. Genom Data. 2015;6:217-221

Wang Q., Taliaferro J. M., Klibaite U., Hilgers V., Shaevitz J. W., et al. The PSI-U1 snRNP interaction regulates male mating behavior in Drosophila. *Proc Natl Acad Sci U S A*. 2016;113:5269-5274

Gosline S. J., Gurtan A. M., JnBaptiste C. K., Bosson A., Milani P., et al. Elucidating MicroRNA Regulatory Networks Using Transcriptional, Post-transcriptional, and Histone Modification Measurements. *Cell Rep*. 2016;14:310-319

106.  Konig J., Zarnack K., Rot G., et al. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol. 2010;17:909-915
107.  Van Nostrand E. L., Pratt G. A., Shishkin A. A., et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods*. 2016;13:508-514
108.  Zarnegar B. J., Flynn R. A., Shen Y., Do B. T., Chang H. Y. and Khavari P. A. irCLIP platform for efficient characterization of protein-RNA interactions. *Nat Methods*. 2016;13:489-492
109.  Van Nostrand E. L., Pratt G. A., Shishkin A. A., et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods*. 2016;13:508-514

## Associated Kits

TruSeq Ribo Profile Kit

Ribo-Zero rRNA Removal Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

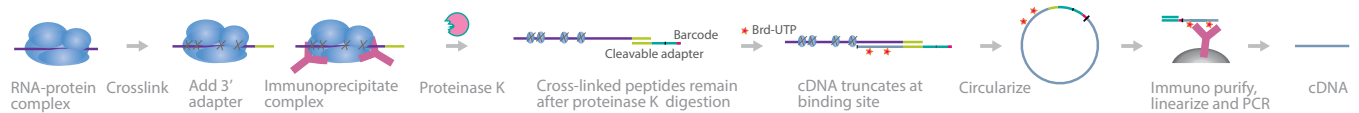TruSeq Stranded mRNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

# BrdU-CLIP: CLIP with BrdU Affinity Purification

BrdU-CLIP sequences the binding sites of RBPs with single-nucleotide resolution. BrdU-CLIP fixes a major problem in HITS-CLIP where 5' adapters are not attached to the cDNA due to premature termination of reverse transcription. BrdU-CLIP uses BrdUTPs and primers with 3' and 5' adapters during reverse transcription.[110] This method enables greater cDNA yield and attaches both 5' and 3' adapters in a single step.

First, RNA-protein complexes are UV-crosslinked, digested with RNase, and immunoprecipitated. 3' adapters are ligated to the RNA and reverse-crosslinked by proteinase K digestion. Some peptides from the crosslink remain bonded with the RNA strands, even after reverse-crosslinkage. RT is carried out using BrdUTPs and primers containing 3' and 5' adapters separated by an apurinic/apyrimidinic endonuclease (APE) cleavage site. The resultant cDNA is purified, circularized, and further purified by BrdU pull-down. Next, the cDNA is linearized by cleaving at the APE site, PCR-amplified, and sequenced.



A schematic overview of BrdU-CLIP.

| Advantages | Disadvantages |
|---|---|
| • Sequences RNA-binding sites of RBPs with single-nucleotide resolution<br>• Captures both truncated and nontruncated cDNAs, unlike HITS-CLIP<br>• Provides higher cDNA yield due to BrdUTP<br>• Attaches 3' and 5' adapters in one RT step | • Tube-column transfers can result in major loss of material[111]<br>• Not yet adopted widely by the scientific community |

## Reviews

Scotti M. M. and Swanson M. S. RNA mis-splicing in disease. *Nat Rev Genet*. 2016;17:19-32

Vuong C. K., Black D. L. and Zheng S. The neurogenetics of alternative splicing. *Nat Rev Neurosci*. 2016;17:265-281

Hsiao Y.-H. E., Cass A. A., Bahn J. H., Lin X. and Xiao X. Global Approaches to Alternative Splicing and Its Regulation—Recent Advances and Open Questions. *Transcriptomics and Gene Regulation*. 2016;37-71

Bangru S. and Kalsotra A. Advances in analyzing RNA diversity in eukaryotic transcriptomes: peering through the Omics lens. *F1000Research*. 2016;5:2668

Flynn R. A., Martin L., Spitale R. C., Do B. T., Sagan S. M., et al. Dissecting noncoding and pathogen RNA-protein interactomes. *RNA*. 2015;21:135-143

## References

Weyn-Vanhentenryck S. M., Mele A., Yan Q., Sun S., Farny N., et al. HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism. *Cell Rep*. 2014;6:1139-1152

The researchers developed BrdU-CLIP and used it, along with HITS-CLIP, to understand how the RBPs RBFOX1/2/3 regulate alternative splicing in the nervous system and the implications in autism. They found 2 guanines in the UGCAUG binding motif for RBFOX are crucial for the protein's interaction with RNA and during crosslinking. The study also revealed a correlation between altered RBFOX targets in autistic brains and the downregulation of all 3 RBFOX proteins.

Illumina Technology: HiSeq 2000 System, Genome Analyzer System, TruSeq Small RNA Library Prep Kit

## Associated Kits

TruSeq Small RNA Library Prep Kit

---
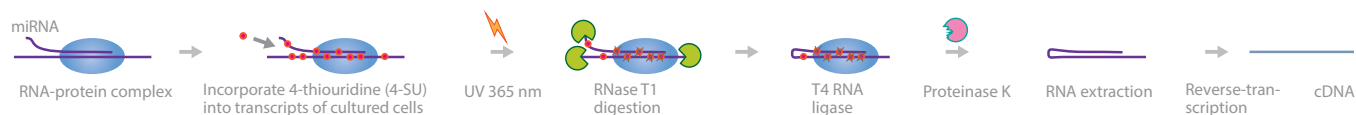
110. Weyn-Vanhentenryck S. M., Mele A., Yan Q., et al. HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism. *Cell Rep*. 2014;6:1139-1152
111. Zarnegar B. J., Flynn R. A., Shen Y., Do B. T., Chang H. Y. and Khavari P. A. irCLIP platform for efficient characterization of protein-RNA interactions. *Nat Methods*. 2016;13:489-492

# AGO-CLIP: Argonaute-Crosslinking and Immunprecipitation

Argonaute-crosslinking and immunoprecipitation (AGO-CLIP) is a protocol that maps miRNA binding to AGO sites in *C. elegans*.[112] In this method, the photoreactive nucleoside 4-SU is first incorporated into *C. elegans* RNA *in vivo*. The 4-SU–labeled RNA is crosslinked to bound protein, and the extracted lysate is treated with RNase T1, which shortens some miRNAs. After immunoprecipitation and washing, crosslinked RNA is treated with PNK, T4 RNA ligase, and finally proteinase K. The RNA is extracted, reverse-transcribed to cDNA, and sequenced. Deep sequencing provides single base-pair resolution of miRNA and AGO binding sites.



A schematic overview of AGO-CLIP.

| Advantages | Disadvantages |
|---|---|
| • Maps miRNA binding to AGO sites using CLIP<br>• Provides single base-pair resolution of the binding sites | • Low efficiency in hybrid capture<br>• Unable to provide quantitative analysis of the strength of miRNA-target interactions[113] |

## Reviews
Schirle N. T., Sheu-Gruttadauria J. and MacRae I. J. Structural basis for microRNA targeting. *Science*. 2014;346:608-613

## References
Hamilton M. P., Rajapakshe K. I., Bader D. A., et al. The Landscape of microRNA Targeting in Prostate Cancer Defined by AGO-PAR-CLIP. *Neoplasia*. 2016;18:356-370

The deregulation of miRNAs directly correlates with prostate cancer initiation and progression. The authors used AGO-CLIP in LNCaP cells that have been treated with the androgen receptor antagonist, enzalutamide, to identify the interactions of miRNAs with their targets in human prostate cancer disease progression. They also used AGO-PAR-CLIP in prostate cancer cell lines at various stages of disease progression. The results identified pathways of the dominant and novel drivers in the progression of prostate cancer.

Illumina Technology: HiSeq 2000 System, TruSeq RNA Library Prep Kit

Imig J., Brunschweiger A., Brummer A., et al. miR-CLIP capture of a miRNA targetome uncovers a lincRNA H19-miR-106a interaction. *Nat Chem Biol.* 2015;11:107-114

Luna J. M., Scheel T. K., Danino T., et al. Hepatitis C Virus RNA Functionally Sequesters miR-122. *Cell*. 2015;160:1099-1110

Wang P., Ning S., Zhang Y., et al. Identification of lncRNA-associated competing triplets reveals global patterns and prognostic markers for cancer. *Nucleic Acids Res*. 2015;

Grosswendt S., Filipchyk A., Manzano M., Klironomos F., Schilling M., et al. Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. *Mol Cell.* 2014;54:1042-1054

Newie I., Sokilde R., Persson H., et al. The HER2-encoded miR-4728-3p regulates ESR1 through a non-canonical internal seed interaction. *PLoS One.* 2014;9:e97200

## Associated Kits
TruSeq Small RNA Library Prep Kit

TruSeq RNA Library Prep Kit v2

TruSeq Nano DNA Library Prep Kit

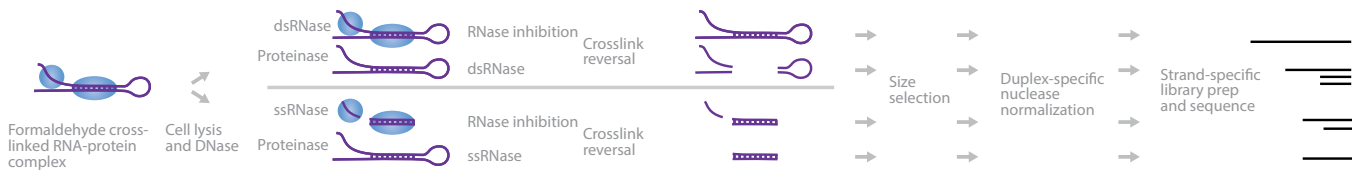TruSeq DNA PCR-Free Library Prep Kit

112. Grosswendt S., Filipchyk A., Manzano M., et al. Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. *Mol Cell*. 2014;54:1042-1054
113. Imig J., Brunschweiger A., Brummer A., et al. miR-CLIP capture of a miRNA targetome uncovers a lincRNA H19-miR-106a interaction. *Nat Chem Biol.* 2014; advance online publication:

# PIP-Seq: Protein Interaction Profile Sequencing

PIP-Seq sequences RBP interaction sites on both unprocessed and mature RNA species in crosslinked or noncrosslinked cells.[114,115] In PIP-seq, both RNase-sensitive and RNase-insensitive fragments are isolated and processed separately to differentiate which sequences are actually bound to RBPs and which are just insensitive to RNases.

First, crosslinked (through UV or formaldehyde) or noncrosslinked cells are lysed. The lysates are split into 2 batches: experimental and RNase-insensitivity control. The experimental/footprint samples are digested with double-stranded RNase (dsRNase) or single-stranded RNase (ssRNase) and subsequently treated with proteinase K to remove the RBPs. However, to screen for regions insensitive to RNases, the controls are first treated with proteinase K and then with ssRNase or dsRNase. Both sets of fragments are reverse-crosslinked and used to prepare strand-specific libraries. Each library is normalized using rehybridization and a thermostable duplex-specific nuclease, and then sequenced.



A schematic overview of PIP-seq.

| Advantages | Disadvantages |
| --- | --- |
| • Sequences RNA regions bound to RBPs in unprocessed or mature RNAs<br>• Identifies regions that are insensitive to RNases<br>• Can be used on tissues and whole organisms<br>• Strand-specific | • Limited resolution of small nucleotide bulges and loops[116]<br>• Formaldehyde crosslinking presents a risk of protein-protein linking, in addition to protein-RNA linking[117]<br>• Nucleases have limited diffusability into plant cells; an advantage of chemical probes, such as dimethyl sulfate (DMS)[118] |

## Reviews

Anderson S. J., Willmann M. R. and Gregory B. D. Protein Interaction Profile Sequencing (PIP-seq) in Plants. *Current Protocols in Plant Biology.* 2016;

Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38:226-236

Foley S. W., Vandivier L. E., Kuksa P. P. and Gregory B. D. Transcriptome-wide measurement of plant RNA secondary structure. *Curr Opin Plant Biol.* 2015;27:36-43

Popova V. V., Kurshakova M. M. and Kopytova D. V. Methods to study the RNA-protein interactions. *Molecular Biology.* 2015;49:418-426

## References

Gosai S. J., Foley S. W., Wang D., et al. Global analysis of the RNA-protein interaction and RNA secondary structure landscapes of the Arabidopsis nucleus. *Mol Cell.* 2015;57:376-388

The authors studied the complex interactions of nuclear RNAs in Arabidopsis seedlings. They used PIP-Seq to map RBP binding sites in the seedling nuclei. This study revealed that alternative splice sites have distinct RBP binding sites and also their own secondary structures. Further, using their PIP-Seq data, the authors discovered that chloroplast RBP CP29A interacts with nuclear mRNAs.

Illumina Technology: HiSeq 2000 System, TruSeq Small RNA Sample Prep Kit

## Associated Kits

TruSeq Stranded Total RNA Library Prep Kit

Ribo-Zero Gold rRNA Removal Kits

114. Silverman I. M., Li F., Alexander A., et al. RNase-mediated protein footprint sequencing reveals protein-binding sites throughout the human transcriptome. *Genome Biology.* 2014;15:R3
115. Silverman I. M. and Gregory B. D. Transcriptome-wide ribonuclease-mediated protein footprinting to identify RNA-protein interaction sites. *Methods.* 2014;
116. Foley S. W., Vandivier L. E., Kuksa P. P. and Gregory B. D. Transcriptome-wide measurement of plant RNA secondary structure. *Curr Opin Plant Biol.* 2015;27:36-43
117. Foley S. W., Vandivier L. E., Kuksa P. P. and Gregory B. D. Transcriptome-wide measurement of plant RNA secondary structure. *Curr Opin Plant Biol.* 2015;27:36-43
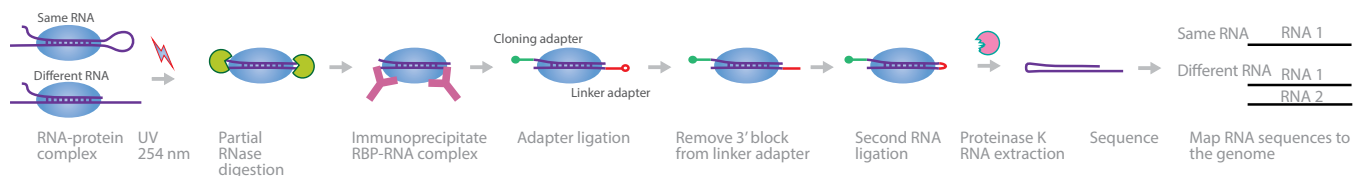118. Foley S. W., Vandivier L. E., Kuksa P. P. and Gregory B. D. Transcriptome-wide measurement of plant RNA secondary structure. *Curr Opin Plant Biol.* 2015;27:36-43

# hiCLIP: RNA Hybrid and Individual-Nucleotide Resolution Ultraviolet Crosslinking and Immunoprecipitation

hiCLIP sequences RNA duplexes bound to RBPs *in vivo*.[119] The unique cloning and linker-adapter system in hiCLIP identifies whether the RBP-bound duplex originates from the same RNA or different RNAs.

Similar to CLIP library preparation techniques, RNA and RBPs are UV-crosslinked, partially digested, and immunoprecipitated. The signature hiCLIP cloning and linker-adapters are ligated to both strands of the duplex. The 3' end of the linker-adapter is dephosphorylated and ligated to the 5' end of the other strand. The bound proteins are removed with proteinase K. The cDNA library is prepared in a similar dashion to iCLIP for high-throughput sequencing. The cloning and linker-adapters enable the mapping of the hybrid reads to the transcriptome and distinguish whether the duplex is formed by the same RNAs or different RNAs.



A schematic overview of hiCLIP.

| Advantages | Disadvantages |
|---|---|
| • Maps RBP-bound RNA duplexes *in vivo* <br> • Linker-adapter remediates physical and computational challenges seen in crosslinking, ligation, and sequencing of hybrids (CLASH) <br> • Detects long-range RNA duplex interactions[120] | • Does not identify non-RBP-bound RNA duplexes <br> • Technically complex procedure <br> • Uses radioactive labeling <br> • Artifacts may be introduced in the circularization step |

## Reviews

Lu Z. and Chang H. Y. Decoding the RNA structurome. *Curr Opin Struct Biol.* 2016;36:142-148

Weidmann C. A., Mustoe A. M. and Weeks K. M. Direct Duplex Detection: An Emerging Tool in the RNA Structure Analysis Toolbox. *Trends Biochem Sci.* 2016;41:734-736

Nawy T. Structural biology: RNA structure served *in vivo*. *Nature Methods.* 2015;12:383-383

Burgess D. J. RNA. Detailed probing of RNA structure *in vivo*. *Nat Rev Genet.* 2015;16:255

## References

Sugimoto Y., Vigilante A., Darbo E., Zirra A., Militti C., et al. hiCLIP reveals the *in vivo* atlas of mRNA secondary structures recognized by Staufen 1. *Nature*. 2015;519:491-494

The authors developed hiCLIP to gain a deeper understanding of the interaction between RNA structures bound by the RBP Staufen 1 in human cells. They discovered an abundance of intramolecular RNA duplexes, depletion of the duplexes in coding regions of highly translated mRNAs, a high prevalence of long-range duplexes in 3'-UTRs, and a low incidence of SNPs in regions where duplexes are formed.

Illumina Technology: HiSeq System, MiSeq System

## Associated Kits

TruSeq Ribo Profile Kit

Ribo-Zero Gold rRNA Removal Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

119. Sugimoto Y., Vigilante A., Darbo E., et al. hiCLIP reveals the *in vivo* atlas of mRNA secondary structures recognized by Staufen 1. *Nature*. 2015;519:491-494
120. Lu Z. and Chang H. Y. Decoding the RNA structurome. *Curr Opin Struct Biol.* 2016;36:142-148

# RBNS: RNA Bind-n-Seq

RBNS characterizes RBPs by high-throughput quantification of their binding affinity, dissociation constants, and the effects of secondary RNA structures on binding.[121] RBNS performs deep sequencing on random short RNA oligonucleotides bound to pools of fluorescently labeled RBPs of different concentrations.

First, RBPs with streptavidin tags are separated into different concentrations pools. Next, they are added to a pool of random RNA fragments. Each RNA oligonucleotide is made up of high- and low-affinity binding sites that are flanked by sequencing adapters at both ends. After the RBPs bind to their target RNAs, the complexes are purified by streptavidin pull-down and the bound RNAs are eluted out. Standard cDNA library preparation procedures are carried out to produce the sequencing library.

Similar methods: HiTS-RAP



A schematic overview of RBNS.

| Advantages | Disadvantages |
| --- | --- |
| • Characterizes sequence and specificity of RBPs<br>• Complements crosslinking-based methods | • Only uses a single round of selection, yielding shorter core RBP sites[122] |

## Reviews
Campbell Z. T. and Wickens M. Probing RNA-protein networks: biochemistry meets genomics. *Trends Biochem Sci.* 2015;40:157-164

## References
Taliaferro J. M., Lambert N. J., Sudmant P. H., et al. RNA Sequence Context Effects Measured *In Vitro* Predict *In Vivo* Protein Binding and Regulation. *Mol Cell.* 2016;64:294-306

The authors investigated the contextual features that determines differences between bound and unbound motifs in the interaction of RBPs and their RNA sequence motifs. They used RBNS with 12,000 mouse RNA sequences and RBPs MBNL1 and RBFOX2 to characterize the binding profiles of identical motifs in different transcripts, using naturally occurring intronic RNA sequences. The results showed that *in vitro* motif occurrences are correlated significantly with *in vivo* binding. RNA secondary structure also played a primary role in these interactions. The authors also performed selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq) to identify the structural characteristics of 588 oligonucleotides used in the RBNS experiments.

Illumina technology: HiSeq System

Kapeli K., Pratt G. A., Vu A. Q., Hutt K. R., Martinez F. J., et al. Distinct and shared functions of ALS-associated proteins TDP-43, FUS and TAF15 revealed by multisystem analyses. *Nat Commun.* 2016;7:12143

Conway Anne E., Van Nostrand Eric L., Pratt Gabriel A., Aigner S., Wilbert Melissa L., et al. Enhanced CLIP Uncovers IMP Protein-RNA Targets in Human Pluripotent Stem Cells Important for Cell Adhesion and Survival. *Cell Reports.* 2016;

## Associated Kits
TruSeq RNA Library Prep Kit v2

---

121. Lambert N., Robertson A., Jangi M., McGeary S., Sharp P. A. and Burge C. B. RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Mol Cell.* 2014;54:887-900
122. Campbell Z. T. and Wickens M. Probing RNA-protein networks: biochemistry meets genomics. *Trends Biochem Sci.* 2015;40:157-164

# TRIBE: Targets of RNA-Binding Proteins Identified by Editing

TRIBE identifies the target RNA sequences of RBPs *in vivo* by modifying the RNA sequence using fusion proteins.[123] The fusion protein consists of the RBP of interest, which binds to the target RNA, and the catalytic domain of adenosine deaminase acting on RNA (ADAR), which irreversibly modifies the proximal adenosine to inosine to serve as a marker during sequence analysis.

First, fusion proteins with the RBP of interest are cloned into the animal model and expressed along with fluorescent marker proteins. After activating the expression of the fusion proteins, mRNA from the cells of interest are isolated using oligo(dT) beads. mRNAs are reverse-transcribed using poly(dT)-T7 primers and random-T7 primers to reduce 3' bias. The nascent cDNA strands are input into the TruSeq RNA Library Prep Kit to generate the cDNA library for sequencing.



A schematic overview of TRIBE.

| Advantages | Disadvantages |
|---|---|
| • *In vivo* identification of RNA targets of RBPs in specific cell types<br>• Not restricted to the specificity of antibodies<br>• Can be performed in a small number of specific cells | • Requires assembly of fusion proteins<br>• Requires an adenosine proximal to the RBP binding site<br>• Catalytic domain of ADAR has a strong preference for double-stranded RNA |

## Reviews
None available yet

## References
McMahon A. C., Rahman R., Jin H., Shen J. L., Fieldsend A., et al. TRIBE: Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding Proteins. *Cell*. 2016;165:742-753

The authors developed TRIBE to identify *in vivo* targets of Hrp48, dFMR1, and NonA RBPs in *Drosophila melanogaster*. TRIBE can be performed in small number of specific cells, as little as 150 different fly neurons.

Illumina Technology: HiSeq System, MiSeq System, TruSeq RNA Library Prep Kit

## Associated Kits
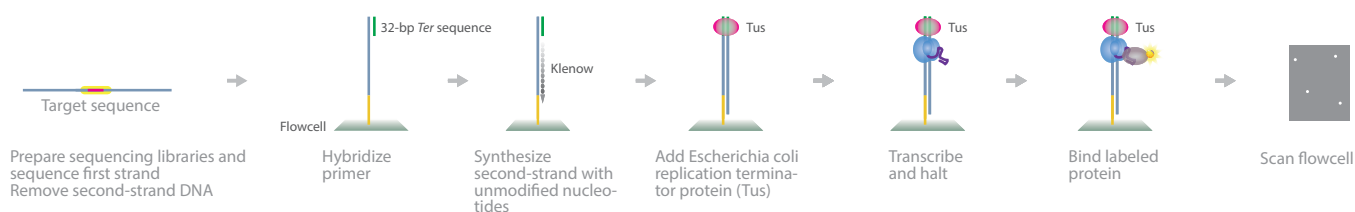TruSeq RNA Library Prep Kit v2

Ribo-Zero rRNA Removal Kits

---

123. McMahon A. C., Rahman R., Jin H., et al. TRIBE: Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding Proteins. Cell. 2016;165:742-753

# HiTS-RAP: High-Throughput Sequencing—RNA Affinity Profiling

HiTS-RAP is a quantitative method to evaluate binding interactions of RNA aptamers with their proteins at a massive scale.[124,125] HiTS-RAP transcribes DNA directly on the flow cell and measures the binding affinity of RNA aptamers with fluorescently labeled proteins.

First, sequencing libraries are prepared and sequenced on the Illumina flow cell. Next, the second strand of the DNA is removed and the single-stranded DNA attached to the flow cell is annealed to primers containing a 32 bp *Ter* sequence. An *Escherichia coli* replication terminator protein (Tus) is introduced into the system and binds to the *Ter* sequence. RNA transcription by T7 RNA polymerase begins and halts upstream of the Tus-bound *Ter* site. Because the RNA polymerase activity was halted abruptly due to the Tus bound to the DNA, the RNA strand is still linked to the RNA polymerase complex. Fluorescently labeled proteins are introduced and bind to the exposed RNA aptamers, producing a fluorescent signal that can be read by the sequencer.

Similar methods: RBNS, RNA-MaP



A schematic overview of HiTS-RAP.

| Advantages | Disadvantages |
|---|---|
| • Quantitatively evaluates binding affinity of RNA aptamers on a massively parallel scale.<br>• Can be adapted easily to nonprotein molecules<br>• Able to determine *de novo* binding specificity of RBPs | • Limited to analysis of RNA fragments 150 nt or shorter, unless paired-end sequencing protocols are used, which increases the limit to 500 nt<br>• Unable to measure binding kinetics, such as on and off rates<br>• Susceptible to steric hindrance effects usually present in large proteins |

## Reviews

Marchese D., de Groot N. S., Lorenzo Gotor N., Livi C. M. and Tartaglia G. G. Advances in the characterization of RNA-binding proteins. *Wiley Interdiscip Rev RNA*. 2016;

Campbell Z. T. and Wickens M. Probing RNA-protein networks: biochemistry meets genomics. *Trends Biochem Sci.* 2015;40:157-164

## References

Tome J. M., Ozer A., Pagano J. M., Gheba D., Schroth G. P., et al. Comprehensive analysis of RNA-protein interactions by high-throughput sequencing-RNA affinity profiling. *Nat Methods.* 2014;11:683-688

The authors developed HiTS-RAP to measure the binding affinities of NELF-E and GFP proteins with their RNA aptamers. Their results indicate that NELF-E–aptamer binding is heavily affected by mutations, because it relies on single-stranded RNA motifs for binding. However, GFP binding relies primarily on RNA secondary structure and is unaffected by mutations.

Illumina Technology: Genome Analyzer System

## Associated Kits

TruSeq DNA PCR-Free Library Prep Kit

---

124.  Tome J. M., Ozer A., Pagano J. M., Gheba D., Schroth G. P. and Lis J. T. Comprehensive analysis of RNA-protein interactions by high-throughput sequencing-RNA affinity profiling. *Nat Methods*. 2014;11:683-688

125.  Ozer A., Tome J. M., Friedman R. C., Gheba D., Schroth G. P. and Lis J. T. Quantitative assessment of RNA-protein interactions with high-throughput sequencing-RNA affinity profiling. *Nat Protoc.* 2015;10:1212-1233

# TRAP-Seq: Targeted Purification of Polysomal mRNA

Targeted purification of polysomal mRNA (TRAP-Seq) maps translating mRNAs under various conditions.[126] In this method, tagged ribosomal proteins are expressed in cells. The tagged ribosomal proteins are purified and the RNA isolated. The RNA is reverse-transcribed to cDNA. Deep sequencing of the cDNA provides single-base resolution of translating RNAs.



Cap    Polysomes                    Bead                           RNA              Reverse-transcription              cDNA

A schematic overview of TRAP-seq.

| Advantages | Disadvantages |
|---|---|
| • Allows detection of translating RNAs<br>• RNAs translated by specific targeted ribosomes can be assessed<br>• No prior knowledge of the RNA is required<br>• Provides a genome-wide RNA screen | • Not as specific as more recently developed methods, such as Ribo-Seq |

## Reviews

Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38:226-236

Maze I., Shen L., Zhang B., et al. Analytical tools and current challenges in the modern era of neuroepigenomics. *Nat Neurosci.* 2014;17:1476-1490

## References

Reynoso M. A., Juntawong P., Lancia M., Blanco F. A., Bailey-Serres J. and Zanetti M. E. Translating Ribosome Affinity Purification (TRAP) followed by RNA sequencing technology (TRAP-SEQ) for quantitative assessment of plant translatomes. *Methods Mol Biol.* 2015;1284:185-207

This publication provides an in-depth protocol for using TRAP-Seq to interrogate the plant translatome (mRNAs associated with at least one 80S ribosome).

Illumina Technology: HiSeq 2000/2500 System, NextSeq 500 System, ScriptSeq RNA-Seq Library Prep Kit, TruSeq RNA Library Prep Kit, TruSeq Small RNA Library Prep Kit

## Associated Kits

TruSeq Ribo Profile Kit

Ribo-Zero rRNA Removal Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

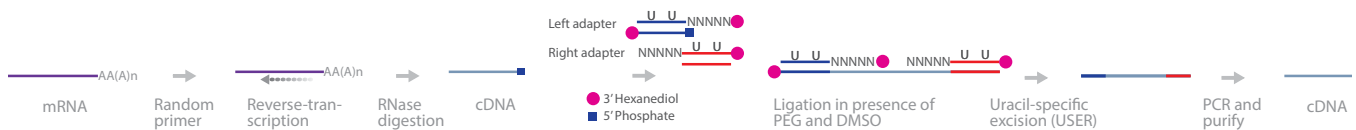TruSeq Stranded Total RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

---

126. Jiao Y. and Meyerowitz E. M. Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. *Mol Syst Biol.* 2010;6:419

# DLAF: Directly Ligate Sequencing Adapters to the First-Strand cDNA

DLAF is a strand-specific RNA sequencing method that avoids second-strand cDNA synthesis by directly attaching unique double-stranded adapters to the first-strand cDNA.[127]

First, mRNA is isolated and depleted of rRNA. Next, RNA fragments are partially hydrolyzed and annealed to random primers before RT. The resulting first-strand cDNA is purified and flanked with double-stranded adapters. One strand of each adapter, the annealing strand, contains uracils and an overhang with 5 or 6 random nucleotides that anneal to the cDNA. The other strand, the ligating strand, directly ligates to the last nucleotide of the cDNA strand. 3'-hexanediol is attached to the 3' ends of each adapter to reduce concatenation with other strands. The adapter-ligated cDNAs are purified and exposed to uracil-specific excision enzyme, to remove the annealing strands from the cDNA. This process results in single-stranded cDNA strands flanked by the ligated adapters. The cDNA is PCR-amplified, purified, and sequenced.



A schematic overview of DLAF.

| Advantages | Disadvantages |
| --- | --- |
| • Reads mRNA sequences solely from the first strand of cDNA<br>• Sequences the 5' and 3' ends of mRNAs<br>• Fewer steps ensure a high yield of cDNA library product | • Unable to differentiate between 5' capped and uncapped mRNA |

## Reviews

Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38:226-236

## References

Iwase S., Brookes E., Agarwal S., et al. A Mouse Model of X-linked Intellectual Disability Associated with Impaired Removal of Histone Methylation. *Cell Rep.* 2016;

The authors investigated the association of mutated chromatin modifiers with X-linked intellectual disability (XLID), a human neurological disorder. They focused on characterizing the importance of KDM5C, a demethylase specific for histone H3 di- and tri-methylated lysine 4 (H3K4me2/3), which is frequently mutated in XLID. They performed DLAF to attach sequencing adapters to first-strand cDNA from RNA-Seq libraries of the frontal cortex and amygdala of adult and cultured murine neurons. The results indicate that KDM5C predominantly repressed genes involved in key pathways that regulate the development and function of neuronal circuits. This study suggests that mutations in KDM5C have strong causal effects to XLID.

Illumina Technology: HiSeq 2000 System, TruSeq ChIP Library Prep Kit

## Associated Kits

ScriptSeq v2 RNA-Seq Library Prep Kit

Ribo-Zero rRNA Removal Kits

---

127.  Agarwal S., Macfarlan T. S., Sartor M. A. and Iwase S. Sequencing of first-strand cDNA library reveals full-length transcriptomes. *Nat Commun.* 2015;6:6002

# miTRAP: miRNA Trapping by RNA *in Vitro* Affinity Purification

miTRAP identifies the target RNA sequences of miRNA species *in vitro* utilizing MS2 stem loops attached to the target RNA sequence as bait for the miRNA.[128] MS2 is a 19 nt bacteriophage RNA sequence present at the ribosomal binding site of the MS2 replicase mRNA that folds into a hairpin loop structure.

First, MS2 bait transcripts are transcribed *in vitro*. The miRNA-protein-RNA complexes are captured by immobilization on amylose resin. The RNA-protein complexes are incubated with cellular extracts, and the extracted RNA-protein complexes are eluted using maltose solution. miRNAs are purified from the maltose solution by a phenol-chloroform extraction, while protein analysis is carried out after incubating the resin with SDS-sample buffer containing 10% β-mercaptoethanol. miRNA eluates are prepared into cDNA libraries, using standard small-RNA Library Prep Kit, and sequenced.

Similar methods: CLASH



A schematic overview of miTRAP.

| Advantages | Disadvantages |
|---|---|
| • Enables *in vitro* identification of miRNA targets<br>• Also identifies novel and noncanonical miRNA target sequences involved in gene regulation | • Not yet validated for analyzing trans-acting RBPs or lncRNAs that affect miRNA targeting<br>• Unable to incorporate UV or chemical-based protein-RNA crosslinking prior to cell lysis |

## Reviews

Seliger B. Role of microRNAs on HLA-G expression in human tumors. *Hum Immunol.* 2016;77:760-763

Elton T. S. and Yalowich J. C. Experimental procedures to identify and validate specific mRNA targets of miRNAs. *EXCLI J.* 2015;14:758-790

## References

Jasinski-Bergner S., Reches A., Stoehr C., et al. Identification of novel microRNAs regulating HLA-G expression and investigating their clinical relevance in renal cell carcinoma. *Oncotarget*. 2016;7:26866-26878

This study focused on identifying new therapeutic targets for renal cell carcinoma (RCC). The researchers used miTRAP combined with in silico profiling of miRNAs that regulate human leukocyte antigen G (HLA-G) expression in HEK293T and JEG-3 cell lines. They found 2 novel HLA-G–regulating miRNAs, namely miR-548q and miR-628-5p. In addition, overexpression of these miRNAs downregulated HLA-G mRNA and protein.

Illumina Technology: TruSeq Small RNA Library Prep Kit

Jasinski-Bergner S., Stehle F., Gonschorek E., et al. Identification of 14-3-3beta gene as a novel miR-152 target using a proteome-based approach. *J Biol Chem*. 2014;289:31121-31135

## Associated Kits

TruSeq Small RNA Library Prep Kit

---

128. Braun J., Misiak D., Busch B., Krohn K. and Hüttelmaier S. Rapid identification of regulatory microRNAs by miTRAP (miRNA trapping by RNA *in vitro* affinity purification). *Nucleic Acids Research.* 2014;

# CLASH: Crosslinking, Ligation, and Sequencing of Hybrids

CLASH maps RNA-RNA interactions.[129, 130]

In this method, RNA-protein complexes are UV-crosslinked and affinity-purified. RNA-RNA hybrids are ligated, isolated, and reverse-transcribed into cDNA. Deep sequencing of the cDNA provides high-resolution chimeric reads of RNA-RNA interactions.

Similar methods: miTRAP, SPLASH, hiCLIP, RAP, RPL



A schematic overview of CLASH.

| Advantages | Disadvantages |
|---|---|
| • Maps RNA-RNA interactions *in vivo*<br>• Provides binding site–level resolution[131] | • Hybrid ligation may be difficult between short RNA fragments<br>• Relatively low efficiency[132]<br>• Requires known bait protein[133] |

## Reviews

Chou C. H., Chang N. W., Shrestha S., et al. miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Res.* 2016;44:D239-247

Hausser J. and Zavolan M. Identification and consequences of miRNA-target interactions–beyond repression of gene expression. *Nat Rev Genet.* 2014;15:599-612

## References

Imig J., Brunschweiger A., Brummer A., Guennewig B., Mittal N., et al. miR-CLIP capture of a miRNA targetome uncovers a lincRNA H19-miR-106a interaction. *Nat Chem Biol.* 2014;advance online publication:

The authors investigated the effects of miRNA on post-transcriptional gene regulation in human cell lines. They developed miR-CLIP to identify prospective targets of the miR-106a miRNA and assessed the effects of overexpression on its functional targets. By comparing miR-CLIP and CLASH results, they found that miR-106a and other members of its family bind to the H19 lncRNA. Further, the H19 lncRNA behaves as a buffer in regulating the gene expression of H19, the miR-17-5p family, and their mRNA targets.

Illumina Technology: Unspecified Illumina sequencing system

## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

129.  Kudla G., Granneman S., Hahn D., Beggs J. D. and Tollervey D. Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. *Proc Natl Acad Sci U S A.* 2011;108:10010-10015

130.  Helwak A., Kudla G., Dudnakova T. and Tollervey D. Mapping the Human miRNA Interactome by CLASH Reveals Frequent Noncanonical Binding. *Cell.* 2013;153:654-665

131.  Imig J., Brunschweiger A., Brummer A., et al. miR-CLIP capture of a miRNA targetome uncovers a lincRNA H19-miR-106a interaction. *Nat Chem Biol.* 2015;11:107-114

132.  Hausser J. and Zavolan M. Identification and consequences of miRNA-target interactions--beyond repression of gene expression. *Nat Rev Genet.* 2014;15:599-612

133.  Lu Z., Zhang Q. C., Lee B., et al. RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure. *Cell.* 2016;165:1267-1279

# RNA MODIFICATIONS

RNA modifications provide another level of control to the transcriptome. These modifications occur post-transcriptionally on the RNA strands. The alterations can occur on mRNAs, lncRNAs, or ncRNAs. Modifications include $m^5A$ and $m^6A$ methylation, and pseudouridine changes.[134-136] A variety of human diseases involve $m^6A$ modifications, such as Alzheimer's disease, Parkinson's disease, and obesity. However, unlike the reversible $m^6A$ modifications, pseudouridine modifications are irreversible and may potentially function as regulators of nonsense-to-sense codon conversions during translation.

The methods in this section have varying workflows that are dependent on the modification of interest.
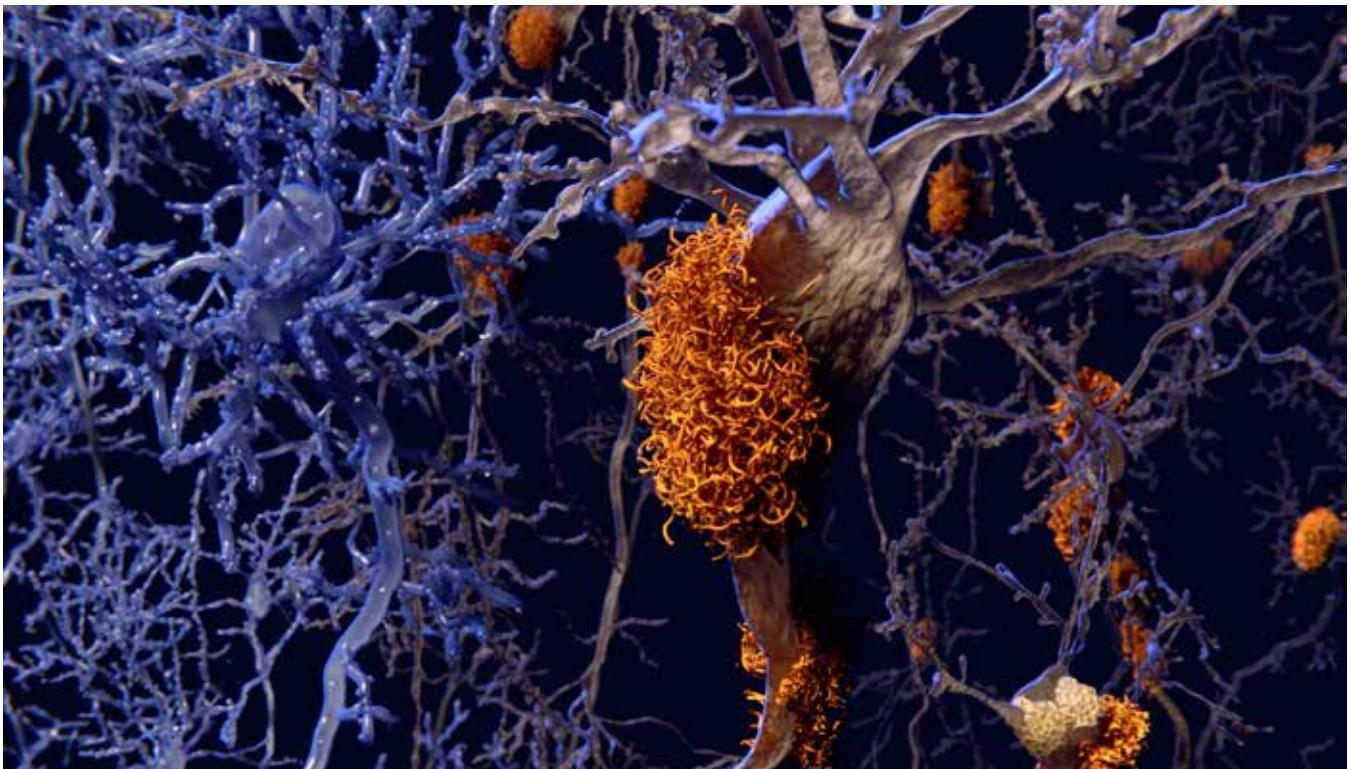


Post-transcriptional alterations play a role in the development of Alzheimer's Disease.

## Reviews

Frye M., Jaffrey S. R., Pan T., Rechavi G. and Suzuki T. RNA modifications: what have we learned and where are we headed? *Nat Rev Genet.* 2016;17:365-372

Maity A. and Das B. N6-methyladenosine modification in mRNA: machinery, function and implications for health and diseases. *FEBS J.* 2016;283:1607-1630

Zaringhalam M. and Papavasiliou F. N. Pseudouridylation meets next-generation sequencing. *Methods.* 2016;107:63-72

Ramaswami G. and Li J. B. Identification of human RNA editing sites: A historical perspective. *Methods.* 2016;107:42-47

134. Frye M., Jaffrey S. R., Pan T., Rechavi G. and Suzuki T. RNA modifications: what have we learned and where are we headed? *Nat Rev Genet.* 2016;17:365-372
135. Maity A. and Das B. N6-methyladenosine modification in mRNA: machinery, function and implications for health and diseases. *FEBS J.* 2016;283:1607-1630
136. Zaringhalam M. and Papavasiliou F. N. Pseudouridylation meets next-generation sequencing. *Methods.* 2016;107:63-72

# MeRIP-Seq: m6A-Specific Methylated RNA Immunoprecipitation Sequencing

MeRIP-Seq[137] maps m6A-methylated RNA. In this method, m6A-specific antibodies are used to immunoprecipitate RNA. RNA is reverse-transcribed to cDNA and sequenced. Deep sequencing provides high-resolution reads of m6A-methylated RNA.

Similar methods: miCLIP, m6A-LAIC-Seq



A schematic overview of MeRIP-Seq.

| Advantages | Disadvantages |
|---|---|
| • Maps m$^6$A-methylated RNA | • Antibodies not specific to the target will precipitate nonspecific RNA modifications. |

## Reviews
Bian Q. and Cahan P. Computational Tools for Stem Cell Biology. *Trends Biotechnol.* 2016;

## References
Gokhale N. S., McIntyre A. B., McFadden M. J., et al. N6-Methyladenosine in Flaviviridae Viral RNA Genomes Regulates Infection. *Cell Host Microbe.* 2016;

This study explored the effects of m6A modification in hepatitis C virus (HCV), a member of the Flaviviridae family. Using MeRIP-Seq, the authors identified m6A peaks in the HCV RNA genome. Depletion of m6A methyltransferases increased the production of infectious HCV particles, while reduction of m6A demethylase decreased production. They also discovered that inactivation of m6A in a specific viral genomic region increased the viral titer while having no effect on RNA replication. The results from this study can be used to identify conserved regulatory regions in Flaviviridae viruses such as dengue, Zika, yellow fever, and West Nile virus.

Illumina Technology: HiSeq 2500 System, TruSeq RNA Sample Prep Kit

Lin S., Choe J., Du P., Triboulet R. and Gregory R. I. The m(6)A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. *Mol Cell.* 2016;62:335-345

Meng J., Lu Z., Liu H., et al. A protocol for RNA methylation differential analysis with MeRIP-Seq data and exomePeak R/Bioconductor package. *Methods.* 2014;69:274-281

## Associated Kits
TruSeq DNA Methylation Kit

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

---

137. Meyer K. D., Saletore Y., Zumbo P., Elemento O., Mason C. E. and Jaffrey S. R. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell.* 2012;149:1635-1646

# miCLIP-m6A: N6-methyladenosine (m6A Individual-Nucleotide-Resolution Crosslinking and Immunoprecipitation

miCLIP-m6A maps m6A locations in the transcriptome with single-nucleotide resolution.[138] In this method, anti-m6A antibodies are crosslinked to mRNA sequences, and a cDNA library is prepared and sequenced. The cDNA library preparation in miCLIP follows the iCLIP protocol closely.

Starting with total RNA, mRNA strands are isolated and fragmented. Anti-m6A antibodies are introduced and UV-crosslinked. The RNA-antibody complexes are immunoprecipitated and purified. Following the iCLIP cDNA library preparation protocol, the 3' ends of the RNA are dephosphorylated and ligated to 3' adapters. The RNA complexes are purified again before digesting the bound anti-m6A antibodies with proteinase K. The freed RNA strands are reverse-transcribed, and the resulting cDNA strands are circularized, relinearized, and PCR-amplified before sequencing. The m6A residues are identified accurately by analyzing the cDNA truncation and cytosine-to-thymine substitution patterns from the peptide residues on the RNA.



| Sheared cellular RNA | Add anti-m6A antibody | UV 365 nm | Immunoprecipitate | Reverse-transcription truncated at binding sites | cDNA |

A schematic overview of miCLIP-m6A.

| Advantages | Disadvantages |
|---|---|
| • Maps m⁶A locations transcriptome-wide with single-nucleotide resolution<br>• Identifies m6A in small-RNA species<br>• Able to identify m6Am in addition to m⁶A<br>• Does not involve pretreating cells with 4-SU, as in PAR-CLIP | • Dependent on the consistency of antibodies used in producing C-to-T substitution patterns<br>• cDNA library preparation uses radioactive labeling |

## Reviews

Maity A. and Das B. N6-methyladenosine modification in mRNA: machinery, function and implications for health and diseases. *FEBS J.* 2016;283:1607-1630

## References

Meyer K. D., Patil D. P., Zhou J., et al. 5' UTR m(6)A Promotes Cap-Independent Translation. *Cell*. 2015;163:999-1010

The mechanism of protein translation from transcripts in a cap-independent manner is poorly understood. This study discovered that mRNAs with m6A in their 5'-UTRs can  be translated in a cap-independent manner. The authors used miCLIP-m6A  to observe the effects of heat shock on the m6A landscape with single-nucleotide resolution. They found that translation of mRNAs with 5'-UTR m6A can be reduced by inhibiting adenosine methylation in select regions. Additionally, Hsp70 mRNA showed increased m6A levels after exposure to heat shock. The results from this study showed that 5'-UTR m6A modification is able to bypass the 5'-capping process to speed up translation during stress conditions.

Illumina Technology: HiSeq 2500 System

## Associated Kits

TruSeq Ribo Profile Kit

Ribo-Zero Gold rRNA Removal Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

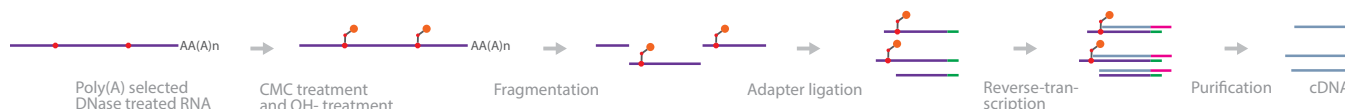TruSeq Targeted RNA Expression Library Prep Kit

---

138.   Linder B., Grozhik A. V., Olarerin-George A. O., Meydan C., Mason C. E. and Jaffrey S. R. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods.* 2015;

# PSI-Seq: Pseudouridine Site Identification Sequencing

PSI-Seq identifies RNA sequences containing pseudouridine sites using high-throughput sequencing.[139] PSI-Seq uses *N-Cyclohexyl-N'*-(2-morpholinoethyl)carbodiimide (CMC) to modify pseudouridines selectively, effectively halting reverse transcription. The cDNA libraries are prepared by the ARTseq method.

Briefly, samples are poly(A)-selected, treated with DNase, and fragmented. CMC is added to modify existing pseudouridines, and the 3' ends of the RNA are ligated to linker-adapters. Next, the RNA fragments are reverse-transcribed to cDNA; however, upon encountering CMC-modified pseudouridines, reverse transcription is halted. cDNA strands of 20–80 nt are isolated and processed into cDNA libraries using the Ribo-Seq/ARTseq method before high-throughput sequencing.

Similar methods: Pseudo-seq, Ψ-Seq, CeU-Seq



| Poly(A) selected DNase treated RNA | CMC treatment and OH- treatment | Fragmentation | Adapter ligation | Reverse-tran-scription | Purification | cDNA |

A schematic overview of PSI-seq.

| Advantages | Disadvantages |
| --- | --- |
| • Identifies pseudouridylation sites in ncRNAs<br>• Single-base resolution<br>• Uses a regression analysis to compare reads in a specific location between treated and mock-treated libraries[140] | • None reported yet |

## Reviews
Zaringhalam M. and Papavasiliou F. N. Pseudouridylation meets next-generation sequencing. *Methods*. 2016;107:63-72

## References
Lovejoy A. F., Riordan D. P. and Brown P. O. Transcriptome-Wide Mapping of Pseudouridines: Pseudouridine Synthases Modify Specific mRNAs in S. cerevisiae. *PLoS One.* 2014;9:e110799

The authors developed PSI-Seq to map pseudouridylation sites in cellular RNAs across the entire transcriptome. They validated the efficacy of the technique by detecting all 43 known pseudouridines in the yeast 18S and 25S ribosomal RNA. By genetically deleting select pseudouridine synthase enzymes (Pus1), they discovered that Pus1 was necessary and sufficient for pseudouridylation of RPL11a mRNA. Additionally, Pus4 interacts with TEF1 mRNA, and Pus6 is involved in the pseudouridylation of KAR2 mRNA. The authors also found indications that certain pseudouridylation sites are conserved evolutionarily throughout the yeast lineage.

Illumina Technology: Genome Analyzer II System

## Associated Kits
TruSeq Ribo Profile Kit

139.  Lovejoy A. F., Riordan D. P. and Brown P. O. Transcriptome-Wide Mapping of Pseudouridines: Pseudouridine Synthases Modify Specific mRNAs in S. cerevisiae. *PLoS One.* 2014;9:e110799
140.  Zaringhalam M. and Papavasiliou F. N. Pseudouridylation meets next-generation sequencing. *Methods*. 2016;107:63-72

# Pseudo-Seq: Method for Genome-Wide Identification of Pseudouridylation Sites

Pseudo-Seq detects pseudouridylation sites in ncRNAs with single-nucleotide resolution using high-throughput sequencing.[141] Pseudo-Seq is very similar to PSI-seq, in that both methods use CMC to modify pseudouridines selectively and halt reverse transcription. However, Pseudo-Seq circularizes cDNA strands before PCR amplification and purification, instead of using ARTseq.

Briefly, poly(A)-selected RNA is fragmented and treated with CMC. The RNA is dephosphorylated with CIP and PNK, and size-selected. Next, 3' adapters are ligated to RNA strands and reverse transcription is initiated. The truncated cDNAs resulting from CMC-modified pseudouridines are purified, circularized, and PCR-amplified. The purified cDNA libraries are sequenced by an NGS method.

Similar methods: PSI-seq, Ψ-Seq, CeU-Seq



A schematic overview of Pseudo-seq.

| Advantages | Disadvantages |
|---|---|
| • Identifies pseudouridylation sites in ncRNAs<br>• Provides single-nucleotide resolution<br>• Identifies peaks by computationally calculating the ratio of reads at the initial mapped position to the total number of reads covering that position[142] | • Circularization step may introduce additional bias<br>• Not yet adopted widely by the scientific community |

## Reviews
Zaringhalam M. and Papavasiliou F. N. Pseudouridylation meets next-generation sequencing. *Methods*. 2016;107:63-72

## References
Carlile T. M., Rojas-Duran M. F., Zinshteyn B., Shin H., Bartoli K. M., et al. Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature*. 2014;

The authors developed Pseudo-Seq to study pseudouridylation in the yeast and human transcriptomes. They identified existing sites of modification and discovered novel sites in ncRNAs and mRNAs. They were able to assign the novel sites to 7 conserved pseudouridine synthases, Pus 1–4, 6, 7, and 9. Most pseudouridines detected in mRNAs were involved in gene regulation as a response to environmental signals, such as nutrient deprivation.

Illumina Technology: HiSeq 2000 System

## Associated Kits
TruSeq Small RNA Library Prep Kit

---

141. Carlile T. M., Rojas-Duran M. F., Zinshteyn B., Shin H., Bartoli K. M. and Gilbert W. V. Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature*. 2014;

142. Zaringhalam M. and Papavasiliou F. N. (2016) Pseudouridylation meets next-generation sequencing. *Methods* 107: 63-72
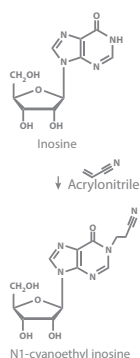
# ICE: Inosine Chemical Erasing

ICE[143-145] followed by NGS identifies adenosine-to-inosine editing.

In this method, RNA is treated with acrylonitrile, while control RNA is untreated. Control and treated RNAs are reverse-transcribed and PCR-amplified. Inosines in RNA fragments treated with acrylonitrile cannot be reverse-transcribed. Deep sequencing of the cDNA prepared from control and treated RNA provides high-resolution reads of inosines in RNA fragments.
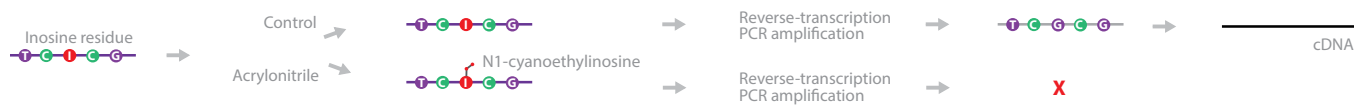


A schematic overview of ICE.



RNA modification during the ICE process.

| Advantages | Disadvantages |
| --- | --- |
| • Provides mapping of adenosine-to-inosine editing<br>• Can be performed with limited material | • Nonlinear PCR amplification can lead to biases, affecting reproducibility<br>• Amplification errors caused by polymerases will be represented and sequenced incorrectly |

## Reviews

Ramaswami G. and Li J. B. Identification of human RNA editing sites: A historical perspective. *Methods*. 2016;107:42-47

Frye M., Jaffrey S. R., Pan T., Rechavi G. and Suzuki T. RNA modifications: what have we learned and where are we headed? *Nat Rev Genet.* 2016;17:365-372

## References

Ishida K., Miyauchi K., Kimura Y., et al. Regulation of gene expression via retrotransposon insertions and the noncoding RNA 4.5S RNA. *Genes Cells.* 2015;

The authors studied the role of short interspersed elements (SINEs) in gene expression regulation using HeLa cells. They discovered that Myodonta clade–specific 4.5S RNAH (4.5SH), an abundant nuclear ncRNA homologous to retrotransposon SINE B1, is crucial in controlling the expression of a reporter gene that contains the antisense insertion of SINE B1. They used cyanoethylation, as described in the ICE-Seq protocol, to observe nuclear retention upon depletion of endogenous 4.5SH. In addition, they found that cellular growth was decreased upon knockdown of 4.5SH in the cytoplasm.

Illumina Technology: MiSeq System, Nextera XT DNA Library Prep Kit

## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

---

143. Sakurai M., Yano T., Kawabata H., Ueda H. and Suzuki T. Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome. *Nat Chem Biol.* 2010;6:733-740
144. Sakurai M., Ueda H., Yano T., et al. A biochemical landscape of A-to-I RNA editing in the human brain transcriptome. *Genome Res.* 2014;24:522-534
145. Suzuki T., Ueda H., Okada S. and Sakurai M. Transcriptome-wide identification of adenosine-to-inosine editing using the ICE-seq method. *Nat Protoc.* 2015;10:715-732

---

# RNA STRUCTURE

lRNA has the ability to form secondary structures that can either promote or inhibit RNA-protein or protein-protein interactions.[146,147] The most diverse secondary and tertiary structures are found in tRNAs and are thought to play a major role in modulating protein translation. RNA structures were first studied in *Tetrahymena thermophilia* using X-ray crystallography, but those studies are inherently cumbersome and limited.[148] Sequencing not only provides information on secondary structures, but it can also determine point-mutation effects on RNA structures in a large number of samples. Recent studies have shown that sequencing is a powerful tool to identify RNA structures and determine their significance.
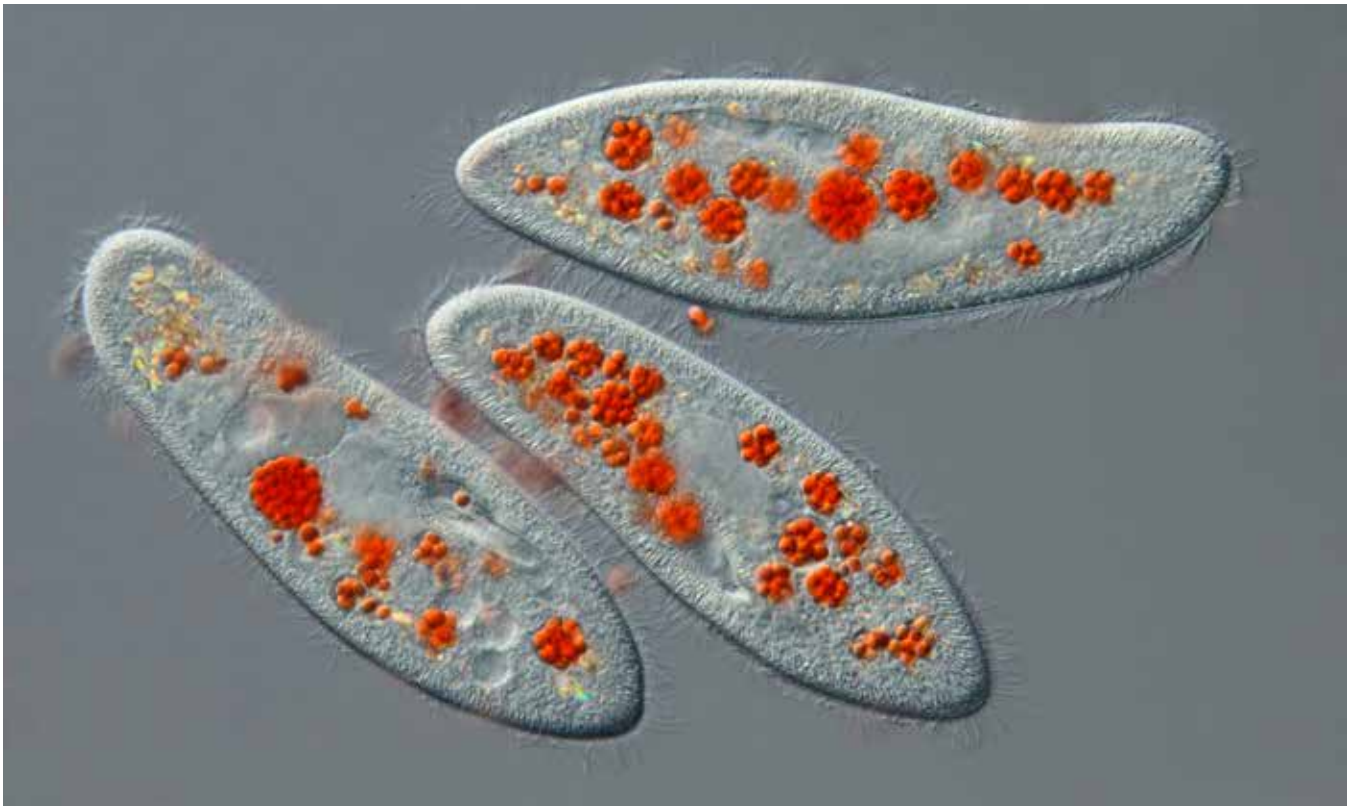


Paramecia species were one of the first model organisms used to study tRNA structure.

## Reviews

Fang Y. and Fullwood M. J. Roles, Functions, and Mechanisms of Long Non-coding RNAs in Cancer. *Genomics Proteomics Bioinformatics*. 2016;14:42-54

Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38:226-236

Ma X., Tang Z., Qin J. and Meng Y. The use of high-throughput sequencing methods for plant microRNA research. *RNA Biology.* 2015;12:709-719

Reuter J. A., Spacek D. V. and Snyder M. P. High-Throughput Sequencing Technologies. *Mol Cell.* 2015;58:586-597

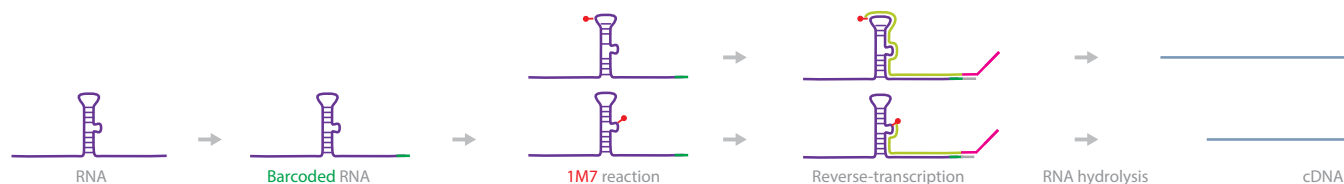146. Osborne R. J. and Thornton C. A. RNA-dominant diseases. *Hum Mol Genet.* 2006;15 Spec No 2:R162-169
147. Thapar R., Denmon A. P. and Nikonowicz E. P. Recognition modes of RNA tetraloops and tetraloop-like motifs by RNA-binding proteins. *Wiley Interdiscip Rev RNA.* 2014;5:49-67
148. Rich A. and RajBhandary U. L. Transfer RNA: molecular structure, sequence, and properties. *Annu Rev Biochem.* 1976;45:805-860
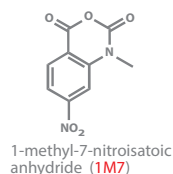
# SHAPE-Seq: Selective 2'-Hydroxyl Acylation Analyzed by Primer Extension Sequencing

SHAPE-Seq[149,150] provides structural information about RNA.

In this method, a unique barcode is first added to the 3' end of RNA, and the RNA is allowed to fold under pre-established *in vitro* conditions. The barcoded and folded RNA is treated with a SHAPE reagent, 1-methyl-7-nitroisatoic anhydride (1M7), which blocks RT. The RNA is reverse-transcribed to cDNA. Deep sequencing of the cDNA provides single-nucleotide sequence information for the positions occupied by 1M7. The structural information of the RNA can then be deduced.



A schematic overview of SHAPE-Seq.



1-methyl-7-nitroisatoic anhydride (1M7)

The structure of 1M7.

| Advantages | Disadvantages |
|---|---|
| • Provides RNA structural information <br> • Multiplexed analysis of barcoded RNAs provides information for multiple RNAs <br> • Effect of point mutations on RNA structure can be assessed <br> • Alternative to mass spectrometry, nuclear magnetic resonance (NMR), and crystallography | • Requires positive and negative controls to account for transcriptase drop-off <br> • Requires pre-established conditions for RNA folding <br> • Folding *in vitro* may not reflect actual folding *in vivo* |

## Reviews

Fang Y. and Fullwood M. J. Roles, Functions, and Mechanisms of Long Non-coding RNAs in Cancer. *Genomics Proteomics Bioinformatics.* 2016;14:42-54

## References

Taliaferro J. M., Lambert N. J., Sudmant P. H., Dominguez D., Merkin J. J., et al. RNA Sequence Context Effects Measured *In Vitro* Predict *In Vivo* Protein Binding and Regulation. *Mol Cell.* 2016;64:294-306

The authors investigated the contextual features that determines differences between bound and unbound motifs in the interaction of RBPs and their RNA sequence motifs. They used 12,000 mouse RNA sequences with RBP MBNL1 and RBFOX2. Using SHAPE-Seq, they identified the structural characteristics of 588 oligonucleotides. The results showed that *in vitro* motif occurrences are significantly correlated with *in vivo* binding. RNA secondary structure was also a primary player in these interactions. The authors used the RBNS pipeline to characterize binding profiles of identical motifs in different transcripts, using naturally occurring intronic RNA sequences.

Illumina Technology: HiSeq System

Watters K. E., Abbott T. R. and Lucks J. B. Simultaneous characterization of cellular RNA structure and function with in-cell SHAPE-Seq. *Nucleic Acids Res.* 2015;

## Associated Kits

TruSeq Small RNA Library Prep Kit

149. Lucks J. B., Mortimer S. A., Trapnell C., et al. Multiplexed RNA structure characterization with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Proc Natl Acad Sci U S A.* 2011;108:11063-11068
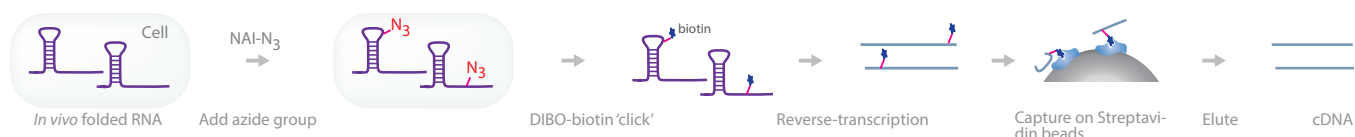150. Watters K. E., Yu A. M., Strobel E. J., Settle A. H. and Lucks J. B. Characterizing RNA structures *in vitro* and *in vivo* with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Methods.* 2016;

# icSHAPE: *In Vivo* Click Selective 2'-Hydroxyl Acylation and Profiling Experiment

icSHAPE provides accurate predictions of RNA-protein interactions and m6A modification *in vivo* by combining SHAPE-Seq with click chemistry for enhanced isolation.[151,152] Secondary RNA structures are modified by the addition of a custom 2-methylnicotinic acid imidazolide (NAI) probe, termed NAI-N$_3$. The modifed RNA is marked selectively with dibenzocyclooxtyne (DIBO)-biotin through copper-free click chemistry, enabling purification by streptavidin pull-down.

Briefly, NAI-N3 is added to RNA *in vivo* to mark it selectively for DIBO-biotin tagging. The cells are lysed, the RNA is poly(A)-selected, tagged with DIBO-biotin, and fragmented. The RNA strands are 3'-end-repaired with T4 PNK and ligated to 3' adapters. After size-selection, the RNA strands are reverse-transcribed, and both the RNA and first-strand cDNA are captured on streptavidin beads. Another cDNA size-selection step is carried out before circularization and PCR amplification. The samples are ready for NGS.

Similar methods: SHAPE-Seq



A schematic overview of icSHAPE.

| Advantages | Disadvantages |
|---|---|
| • Provides accurate predictions of RNA-protein interactions and m6A modifications *in vivo* <br> • Can be applied to *ex vivo* applications with slight modifications <br> • Chemical modification is applicable to all nucleotides, unlike DMS which only modifies adenines and cytosines | • Circularization may introduce additional bias |

## Reviews
No reviews yet

## References
Flynn R. A., Do B. T., Rubin A. J., et al. 7SK-BAF axis controls pervasive transcription at enhancers. *Nat Struct Mol Biol*. 2016;

This study examined RNA functions at enhancer regions, specifically the 7SK snRNA. The authors found that 7SK snRNA inhibits enhancer transcription by modulation of the nucleosome position. Using icSHAPE on Hexim1 or BAF-associated 7SK snRNA, they discovered that 7SK physically interacts with the BAF chromatin-remodeling complex. In turn, it recruits BAF to enhancers and inhibits transcription of enhancers through modulation of the chromatin structure. In conjunction with icSHAPE, the authors used sequencing of start site-associated RNAs (Start-Seq), GRO-seq, ChIRP-seq, ATAC-Seq, and ChIP-Seq.

Illumina Technology: HiSeq 2500 System, NextSeq 500 System, NEBNext DNA Library Prep Master Mix Set for Illumina

Lu Z., Zhang Q. C., Lee B., Flynn R. A., Smith M. A., et al. RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure. *Cell*. 2016;165:1267-1279

## Associated Kits
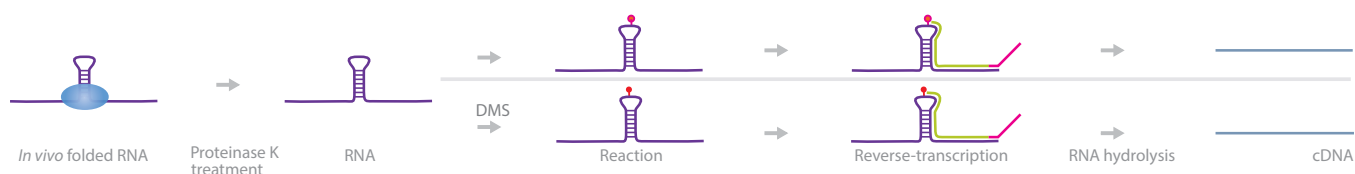TruSeq Small RNA Library Prep Kit

---

151.  Spitale R. C., Flynn R. A., Zhang Q. C., et al. Structural imprints *in vivo* decode RNA regulatory mechanisms. *Nature*. 2015;519:486-490
152.  Flynn R. A., Zhang Q. C., Spitale R. C., Lee B., Mumbach M. R. and Chang H. Y. Transcriptome-wide interrogation of RNA secondary structure in living cells with icSHAPE. *Nat Protoc.* 2016;11:273-290

# CIRS-Seq: Chemical Inference of RNA Structures

CIRS-Seq was developed to investigate the complexity of secondary RNA structures in the mammalian transcriptome.[153] CIRS-Seq uses DMS to methylate the N1 of adenosine and N3 of cytosine residues, and CMC to modify pseudouridines selectively, but only when they are in single-stranded conformation. Modifications on these nucleotide residues halt the RT process, effectively producing a truncated cDNA as a marker for the location of secondary RNA structures.

Briefly, cells are lysed and treated with proteinase K to dissociate protein-bound RNAs while leaving RNA secondary structures intact. The lysates are separated into 3 different treatment lines—DMS, CMC, and no treatment. In all 3 treatment lines, total RNA is extracted and reverse-transcribed using random primers. The resulting cDNA is isolated, ligated to sequencing adapters, and subjected to high-throughput sequencing. Reads from the DMS and CMC lines are used to identify the locations of secondary RNA structures, while the control is used to reduce background noise.



A schematic overview of CIRS-seq.

| Advantages | Disadvantages |
|---|---|
| • Accurately predicts secondary RNA structures, and reveals features of mRNAs and ncRNAs<br>• Provides single-base resolution<br>• Can identify structural requirements for RBPs | • CMC and DMS may react with non–secondary RNA structures |

## Reviews

Zucchelli S., Patrucco L., Persichetti F., Gustincich S. and Cotella D. Engineering Translation in Mammalian Cell Factories to Increase Protein Yield: The Unexpected Use of Long Non-Coding SINEUP RNAs. *Computational and Structural Biotechnology Journal.* 2016;

Kwok C. K., Tang Y., Assmann S. M. and Bevilacqua P. C. The RNA structurome: transcriptome-wide structure probing with next-generation sequencing. *Trends Biochem Sci.* 2015;40:221-232

Strobel E. J., Watters K. E., Loughrey D. and Lucks J. B. RNA systems biology: uniting functional discoveries and structural tools to understand global roles of RNAs. Curr Opin Biotechnol. 2016;39:182-191

## References

Incarnato D., Anselmi F., Morandi E., et al. High-throughput single-base resolution mapping of RNA 2'-O-methylated residues. *Nucleic Acids Res.* 2016;

The authors developed a new method for identifying 2'-O-methylation (2'-OMe) of the ribose moiety, one of the most abundant post-transcriptional RNA modifications. 2'-OMe sequencing uses a modified version of CIRS-Seq in its library preparation step. They were able to discover 12 new 2'-OMe sites across 18S and 28S rRNAs, 11 of which are conserved in human and mouse cells.

Illumina Technology: NextSeq 500 System, TruSeq Small RNA Library Prep Kit

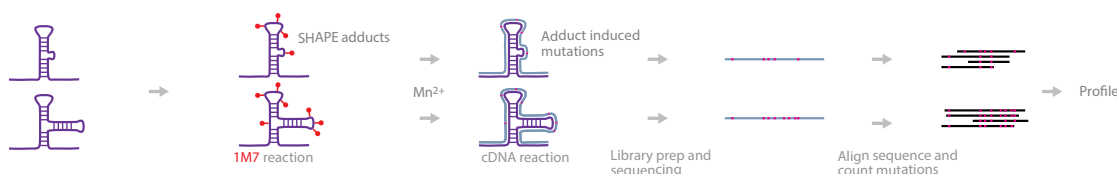## Associated Kits

TruSeq Small RNA Library Prep Kit

---

153.  Incarnato D., Neri F., Anselmi F. and Oliviero S. Genome-wide profiling of mouse RNA secondary structures reveals key features of the mammalian transcriptome. *Genome Biology*. 2014;15:491

# SHAPE-MaP: Selective 2'-Hydroxyl Acylation Analyzed by Primer Extension and Mutational Profiling

SHAPE-MaP sequences secondary RNA structures at various levels on a massively parallel scale. The method can be customized to interrogate small RNAs, amplicons, or rare RNA species accurately in a mixture of RNAs.[154,155] As implied in the name, SHAPE-Map uses the SHAPE-Seq 1M7 reaction to mark RNA ribose 2'-OH groups to identify secondary RNA structures. This reaction is followed by mutational profiling (MaP) to induce noncomplementary nucleotide mutations on SHAPE adducts during RT. These MaP mutations are analyzed after sequencing using powerful informatics tools developed specifically for this method.

Briefly, SHAPE electrophiles are added to the sample containing folded RNAs. The samples are divided into 3 different reaction lines: +reagent, –reagent, and a denaturing control, to correct against intrinsic background mutation rates from reverse transcription. After SHAPE mutations are introduced, RT primers are selected depending on the RNA type of interest (amplicon, small RNA, or specific RNA species profiling). The primers are added to the reaction and the RNA is reverse-transcribed. Subsequent library preparation steps differ according to the RNA of interest. The small RNA workflow involves standard PCR amplification with appropriate primers. The workflows for amplicons and specific RNA species follow the Nextera XT DNA Library Preparation Kit protocol. The barcoded samples are ready for sequencing.

Similar methods: SHAPE-Seq, icSHAPE



A schematic overview of SHAPE-MaP.

| Advantages | Disadvantages |
|---|---|
| • Identifies RNA structures at single-nucleotide resolution using a combination of SHAPE-Seq and MaP techniques<br>• Library preparation step is highly customizable for different applications: amplicon, small RNA, or rare RNA species profiling<br>• Avoids common issues encountered with adapter-ligation methods | • Length of the RNA strand must be at least ~150 nt for the randomer and native workflow, and at least ~40 nt for the small-RNA workflow<br>• Difficult to distinguish MaP mutations from background noise in samples with low levels of RNA |

## Reviews

Schmitz S. U., Grote P. and Herrmann B. G. Mechanisms of long noncoding RNA function in development and disease. *Cellular and Molecular Life Sciences*. 2016;73:2491-2509

Kwok C. K. Dawn of the *in vivo* RNA structurome and interactome. *Biochemical Society Transactions*. 2016;44:1395-1410

Smola M. J., Calabrese J. M. and Weeks K. M. Detection of RNA–Protein Interactions in Living Cells with SHAPE. *Biochemistry*. 2015;54:6867-6875

154. Siegfried N. A., Busan S., Rice G. M., Nelson J. A. and Weeks K. M. RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). Nat Methods. 2014;11:959-965
155. Smola M. J., Rice G. M., Busan S., Siegfried N. A. and Weeks K. M. Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. Nature Protocols. 2015;10:1643-1669

## References

Mauger D. M., Golden M., Yamane D., et al. Functionally conserved architecture of hepatitis C virus RNA genomes. *Proc Natl Acad Sci U S A.* 2015;112:3692-3697;

The authors studied various features in the RNA genome of HCV. They used SHAPE-MaP to model the infectious HCV RNA genome structures from genotypes 1a, 1b, and 2a. They identified multiple RNA structural regions that were conserved and also novel structures that were required for optimal viral fitness.

Illumina Technology: MiSeq System

Lu Y. F., Mauger D. M., Goldstein D. B., Urban T. J., Weeks K. M. and Bradrick S. S. IFNL3 mRNA structure is remodeled by a functional non-coding polymorphism associated with hepatitis C virus clearance. *Sci Rep.* 2015;5:16037

Taylor A. I., Pinheiro V. B., Smola M. J., et al. Catalysts from synthetic genetic polymers. *Nature*. 2014;
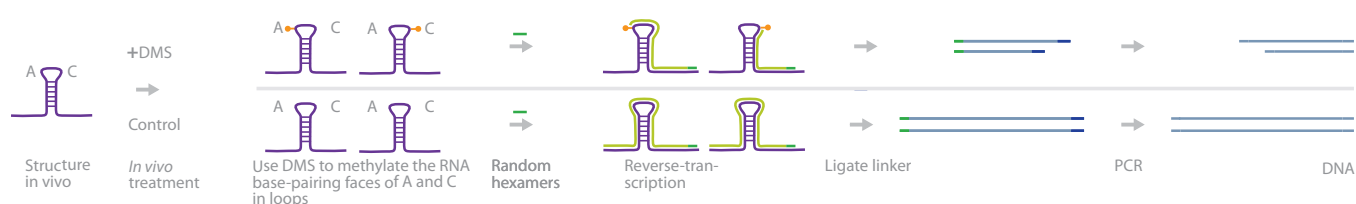
## Associated Kits
Nextera XT DNA Library Prep Kit

# Structure-Seq/DMS-Seq: Use Dimethyl Sulphate Methylation of Unprotected Adenines and Cytosines

Structure-Seq profiles RNA structures for *in vivo* or *in vitro* applications with single-nucleotide resolution.[156,157] This method identifies secondary RNA structures by using the chemical modification induced by DMS on unpaired adenines and cytosines.

Briefly, samples are treated with DMS to mark secondary RNA structures *in vivo*. The RNA is extracted, poly(A)-selected, and treated with DNase. Using random hexamers as primers, reverse transcription is initiated. The resulting single-stranded DNA is ligated to single-stranded DNA linkers and self-circularized using CircLigase enzyme. Next, the DNA is PCR-amplified, size-selected, and sequenced.

Similar methods: icSHAPE, Mod-Seq, DMS-seq, PARS, Frag-seq, dsRNA-Seq



A schematic overview of Structure-seq.

| Advantages | Disadvantages |
|---|---|
| • Provides genome-wide profiling of RNA structures at single-nucleotide resolution<br>• Can be used for *in vivo* and *in vitro* applications<br>• DMS is cell-permeable and can be used for *in vivo* applications<br>• One RT primer synthesis retrieves information for tens of thousands of RNA structures<br>• Random-hexamer primers minimize 3' end bias | • RBPs can block DMS modification *in vivo*<br>• Circularization may introduce additional bias |

## Reviews

Fang Y. and Fullwood M. J. Roles, Functions, and Mechanisms of Long Non-coding RNAs in Cancer. *Genomics Proteomics Bioinformatics.* 2016;14:42-54

Reuter J. A., Spacek D. V. and Snyder M. P. High-Throughput Sequencing Technologies. *Mol Cell.* 2015;58:586-597

## References

Fang R., Moss W. N., Rutenberg-Schoenberg M. and Simon M. D. Probing Xist RNA Structure in Cells Using Targeted Structure-Seq. *PLoS Genet.* 2015;11:e1005668

The authors developed a modified version of Structure-seq, called Targeted Structure-Seq, to study the role of the lncRNA Xist. They identified a new element near the C repeats region that is crucial for the functionality of Xist. Targeted Structure-Seq is able to achieve high coverage of the target RNA with few reads.

Illumina Technology: HiSeq 2500 System, TruSeq RNA Library Prep Kit

## Associated Kits

TruSeq RNA Library Prep Kitt

---

156. Ding Y., Tang Y., Kwok C. K., Zhang Y., Bevilacqua P. C. and Assmann S. M. *In vivo* genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature.* 2014;505:696-700
157. Ding Y., Kwok C. K., Tang Y., Bevilacqua P. C. and Assmann S. M. Genome-wide profiling of *in vivo* RNA structure at single-nucleotide resolution using structure-seq. *Nat Protoc.* 2015;10:1050-1066

# SPARE: Specific Parallel Amplification of 5' RNA Ends

SPARE identifies genome-wide miRNA processing intermediates in plants.[158] SPARE infers the directionality of miRNA processing (base-to-loop or loop-to-base) by analyzing the resultant cDNA sequences. If only the first cleavage position is detected, the miRNA was processed in a base-to-loop fashion; however, if all cleavage intermediates are detected, it was processed in a loop-to-base fashion.

Briefly, total RNA is depleted of rRNA, and RNA adapters are ligated to the 5' ends of uncapped RNAs. These ligated RNAs are used as the template in reverse transcription using miRNA precursor–specific primers with generic adapter tails as RT primers. The resultant cDNAs are PCR-amplified, size-selected, and sequenced.

Similar methods: PARE, 5' RACE



A schematic overview of SPARE.

| Advantages | Disadvantages |
| --- | --- |
| • Enables genome-wide identification of miRNA intermediates<br>• Able to infer processing directionality through sequence analysis<br>• Less time-consuming than existing methods<br>• Optimized for plant genomes | • Not yet adopted widely by the scientific community |

## Reviews
Ma X., Tang Z., Qin J. and Meng Y. The use of high-throughput sequencing methods for plant microRNA research. *RNA Biology.* 2015;12:709-719

## References
Schmidt S. A., Foley P. L., Jeong D. H., et al. Identification of SMG6 cleavage sites and a preferred RNA cleavage motif by global analysis of endogenous NMD targets in human cells. *Nucleic Acids Res.* 2015;43:309-323

This study focused on the identification of endonuclease SMG6 cleavage sites, which are often the first target during nonsense-mediated mRNA decay (NMD) in metazoans. The authors used SPARE to identify 5'-terminus decay intermediates that are dependent on SMG6 and the universal NMD factor UPF1. Depletion of SMG6 for numerous SMG6 substrates resulted in an increased accumulation of decapped transcripts.

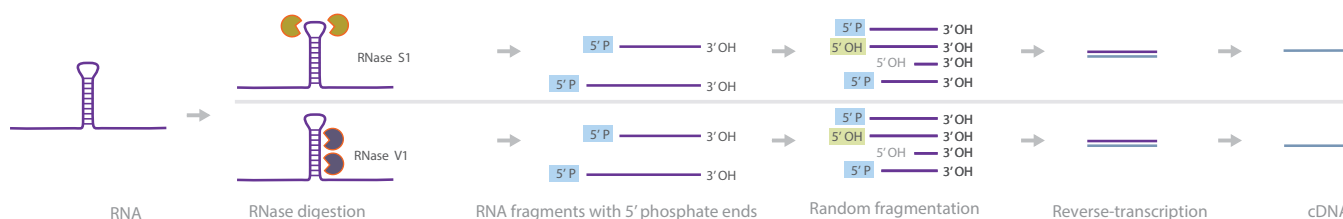Illumina Technology: HiSeq System, TruSeq RNA Library Prep Kit

## Associated Kits
TruSeq Small RNA Library Prep Kit

---

158. Schapire A. L., Bologna N. G., Moro B., Zhai J., Meyers B. C. and Palatnik J. F. Construction of Specific Parallel Amplification of RNA Ends (SPARE) libraries for the systematic identification of plant microRNA processing intermediates. *Methods*. 2013;64:283-291

# PARS-Seq: Parallel Analysis of RNA Structure

PARS-Seq[159] mapping provides information about the secondary and tertiary structure of RNA. In this method, RNA is digested with RNases that are specific for double-stranded and single-stranded RNA, respectively. The resulting fragments are reverse-transcribed to cDNA. Deep sequencing of the cDNA provides high-resolution sequences of the RNA. The RNA structure can be deduced by comparing the digestion patterns of the various RNases.



A schematic overview of PARS-Seq.

| Advantages | Disadvantages |
| --- | --- |
| • Provides RNA structural information<br>• Distinguishes between paired and unpaired bases<br>• Alternative to mass spectrometry, NMR, and crystallography | • Enzyme digestion can be nonspecific<br>• Digestion conditions must be carefully controlled<br>• RNA can be overdigested<br>• Limited to *in vitro* applications |

## Reviews
Nussbacher J. K., Batra R., Lagier-Tourenne C. and Yeo G. W. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci.* 2015;38:226-236

## References
Righetti F., Nuss A. M., Twittenhoff C., et al. Temperature-responsive *in vitro* RNA structurome of Yersinia pseudotuberculosis. *Proc Natl Acad Sci U S A.* 2016;113:7237-7242

The authors used PARS-Seq to determine single- and double-stranded regions in *Yersinia pseudotuberculosis* across a temperature gradient. They found structures for more than 1750 RNAs across 3 temperatures and discovered that, on average, mRNAs around the ribosome-binding sites tend to be unstructured. PARS-Seq results also showed that 5'-UTRs fold at low temperatures and identified novel thermoresponsive RNA structures across various gene categories.

Illumina Technology: HiSeq 2500 System, TruSeq RNA Library Prep Kit

## Associated Kits
TruSeq Small RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

---

159.   Wan Y., Qu K., Ouyang Z. and Chang H. Y. Genome-wide mapping of RNA structure using nuclease digestion and high-throughput sequencing. *Nat Protoc.* 2013;8:849-869

# Cap-Seq: CXXC Affinity Purification Plus Deep Sequencing

CAP-seq[160] maps the 5' end of RNAs anchored to RNAPII.

In this method, RNA transcripts are treated sequentially with Terminator exonuclease, CIP, and TAP, followed by linker ligation and RT to cDNA. Deep sequencing of the cDNA provides high-resolution sequences of RNAPII transcripts.
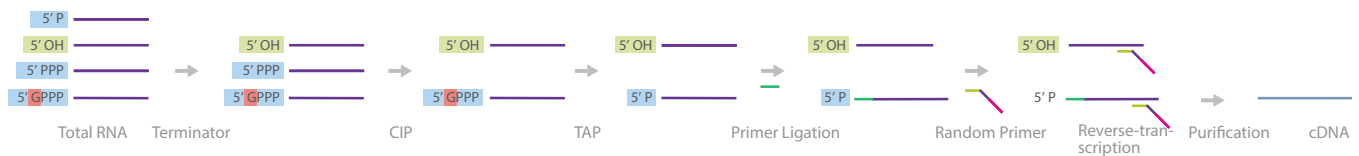


A schematic overview of CAP-seq.

| Advantages | Disadvantages |
| --- | --- |
| • Maps RNAs anchored to RNAPII | • Multiple steps and treatments can lead to loss of material |

## Reviews
No reviews yet

## References
Illingworth R. S., Gruenewald-Schneider U., Webb S., Kerr A. R., James K. D., et al. (2010) Orphan CpG islands identify numerous conserved promoters in the mammalian genome. PLoS Genet 6: e1001134

The authors developed CAP-Seq to study the roles of CpG islands (CGI) in human and mouse cells. The results showed that CGI abundance in both species is very similar. They also found that CpG densities were correlated positively with H3K4 trimethylation.

Illumina Technology:  Genome Analyzer System

## Associated Kits
TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit
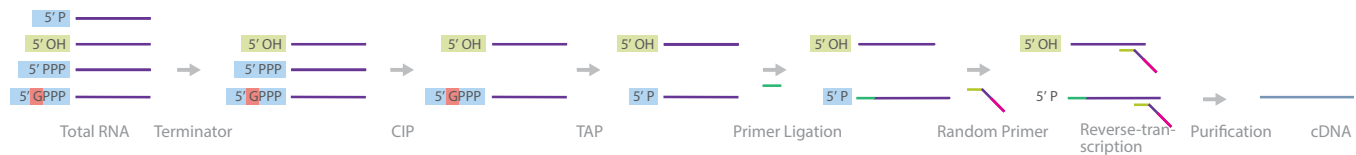
TruSeq Stranded mRNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

160.  Illingworth R. S., Gruenewald-Schneider U., Webb S., et al. Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS Genet*. 2010;6:e1001134

# CIP-TAP: Alkaline Phosphatase, Calf Intestine-Tobacco Acid Pyrophosphatase Sequencing

CIP-TAP maps capped small RNAs.[161]

In this method, RNA is treated with CIP followed by 3'-end linker ligation. Next, the RNA is treated with TAP, followed by 5'-end linker ligation. The fragments are reverse-transcribed to cDNA, PCR-amplified, and sequenced. Deep sequencing provides single-nucleotide resolution reads of the capped small RNAs.



A schematic overview of CIP-TAP.

| Advantages | Disadvantages |
|---|---|
| • Identifies capped small RNAs missed by CapSeq<br>• High throughput | • Nonlinear PCR amplification can lead to biases affecting reproducibility<br>• Amplification errors caused by polymerases |

## References

Gu W., Lee H. C., Chaves D., Youngman E. M., Pazour G. J., et al. CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as *C. elegans* piRNA precursors. *Cell*. 2012;151:1488-1500

The authors used CIP-TAP in conjunction with CapSeq to identify TSS for piRNAs. CapSeq was unable to identify the majority of known piRNA loci. However, the results showed that piRNA precursors are likely to be ~26 nt capped small RNAs that are initiated 2 nt upstream of mature RNAs.

Illumina Technology: Genome Analyzer System, HiSeq 2000 System

## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

161. Gu W., Lee H. C., Chaves D., et al. CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as *C. elegans* piRNA precursors. *Cell*. 2012;151:1488-1500

# LOW-LEVEL RNA DETECTION

Low-level RNA detection refers to both the detection of rare RNA molecules in a cell-free environment—such as circulating tumor RNA—or the expression patterns of single cells. Tissues consist of a multitude of different cell types, each with a distinctly different set of functions. Even within a single cell type, the transcriptomes are highly dynamic and reflect temporal, spatial, and cell cycle–dependent changes. Cell harvesting, handling, and technical issues with sensitivity and bias during amplification add an additional level of complexity. To resolve this multi-tiered complexity would require the analysis of many thousands of cells. The use of unique barcodes has greatly increased the number of samples that can be multiplexed and pooled, with little to no decrease in reads associated with each sample. Recent improvements in cell capture and sample preparation will provide more information, faster, and at lower cost.[162] These improvements promise to expand our understanding of cell function fundamentally, with significant implications for research and human health.[163]
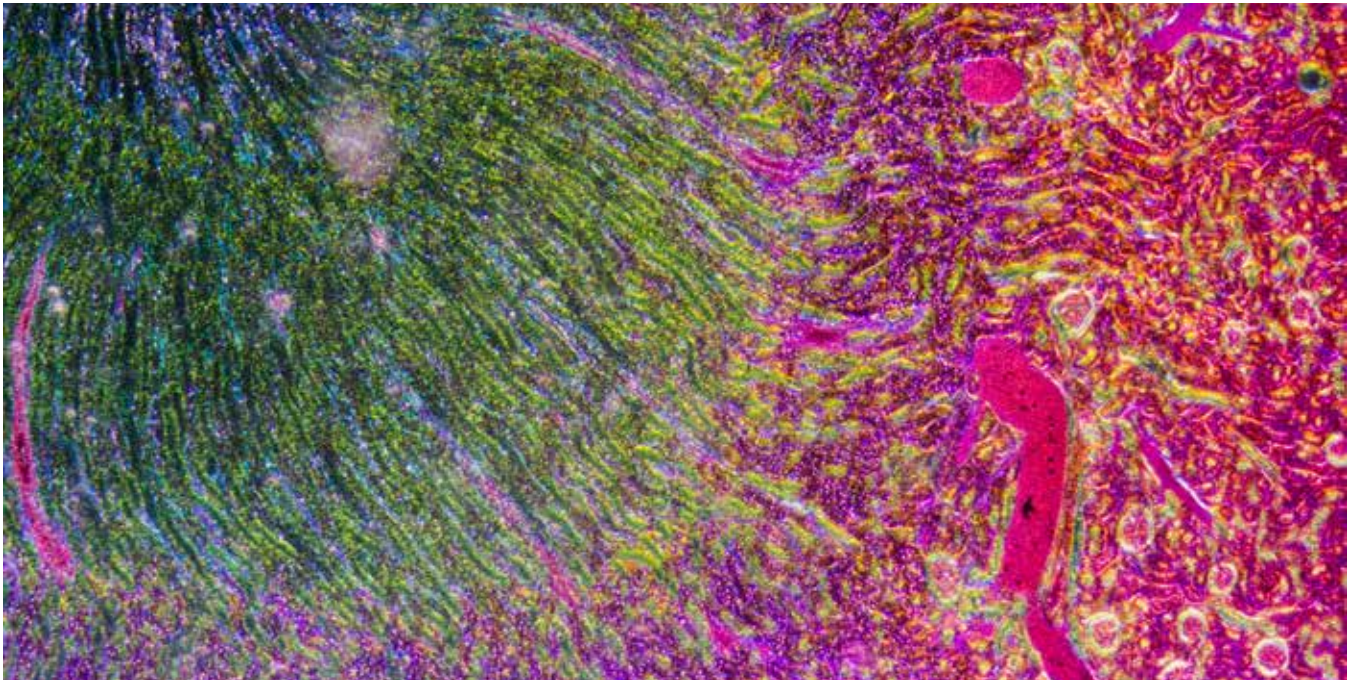


**Organs, such as the kidney depicted in this cross-section, consist of a myriad of phenotypically distinct cells.** Single-cell transcriptomics can characterize the function of each of these cell types.

## Reviews

Liu S. and Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges. *F1000Res*. 2016;5:

Faridani O. R. and Sandberg R. Putting cells in their place. *Nat Biotechnol.* 2015;33:490-491

Kanter I. and Kalisky T. Single cell transcriptomics: methods and applications. *Front Oncol.* 2015;5:53

Kolodziejczyk A. A., Kim J. K., Svensson V., Marioni J. C. and Teichmann S. A. The Technology and Biology of Single-Cell RNA Sequencing. *Mol Cell.* 2015;58:610-620

Trapnell C. Defining cell types and states with single-cell genomics. *Genome Res.* 2015;25:1491-1498

Wang Y. and Navin N. E. Advances and Applications of Single-Cell Sequencing Technologies. *Mol Cell.* 2015;58:598-609

---

162. Shapiro E., Biezuner T. and Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet.* 2013;14:618-630
163. Spaethling J. M. and Eberwine J. H. Single-cell transcriptomics for drug target discovery. *Curr Opin Pharmacol.* 2013;13:786-790

# References

Hou Y., Guo H., Cao C., et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. Cell Res. 2016;26:304-319

To fully understand the mechanisms by which the genome, transcriptome, and DNA methylome interact at the single-cell level, these 3 separate methods ideally should be applied to the same individual cell. In this study, the authors report single-cell triple-omics sequencing (scTrio-Seq), a method that can analyze genomic copy-number variants (CNVs), the DNA methylome, and the transcriptome of an individual mammalian cell simultaneously. They used scTrio-Seq in 25 individual hepatocellular carcinoma primary cells to identify 2 subpopulations of cells. They also found that large-scale CNVs can cause proportional changes in RNA expression in subsets of genes, but the CNVs did not affect DNA methylation in the relevant genomic regions.

Illumina Technology: HiSeq 2000/2500 System

Habib N., Li Y., Heidenreich M., Swiech L., Trombetta J. J., et al. (2016) Div-Seq: A single nucleus RNA-Seq method reveals dynamics of rare adult newborn neurons in the CNS. bioRxiv

Krishnaswami S. R., Grindberg R. V., Novotny M., et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat Protoc.* 2016;11:499-524

Mora-Castilla S., To C., Vaezeslami S., et al. Miniaturization Technologies for Efficient Single-Cell Library Preparation for Next-Generation Sequencing. *J Lab Autom.* 2016;21:557-567

Achim K., Pettit J. B., Saraiva L. R., et al. High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nat Biotechnol.* 2015;33:503-509

Dueck H., Khaladkar M., Kim T. K., et al. Deep sequencing reveals cell-type-specific patterns of single-cell transcriptome variation. *Genome Biol.* 2015;16:122
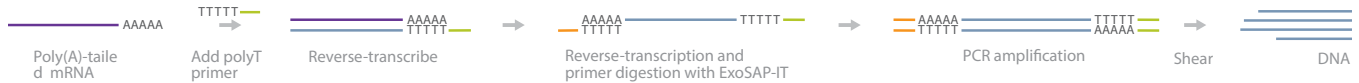
Kim J. K., Kolodziejczyk A. A., Illicic T., Teichmann S. A. and Marioni J. C. Characterizing noise structure in single-cell RNA-seq distinguishes genuine from technical stochastic allelic expression. *Nat Commun.* 2015;6:8687Fan

Scialdone A., Natarajan K. N., Saraiva L. R., et al. Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods*. 2015;85:54-61

# scRNA-Seq: Single-Cell mRNA Sequencing

scRNA-Seq provides deeper insight to the multi-tiered complexity of different cells within the same tissue type. scRNA-Seq has now been adapted widely into other methods in the single-cell RNA sequencing field.

In this method, single cells are isolated manually under a microscope and lysed. Next, mRNAs are purified and primed with a poly(T) primer for reverse transcription. Unreactive primers are removed by exonuclease I digestion. Poly(A) tails are added to the first strand cDNA at the 3' end and annealed to poly(T) primers for second-strand cDNA generation. Finally, the cDNAs are PCR-amplified, sheared, and prepared into sequencing libraries.



Poly(A)-tailed mRNA | Add polyT primer | Reverse-transcribe | Reverse-transcription and primer digestion with ExoSAP-IT | PCR amplification | Shear | DNA

A schematic overview of scRNA-seq.

| Advantages | Disadvantages |
| --- | --- |
| • Single-cell–resolution transcriptomic analysis <br> • Able to detect unknown splice junctions | • Very low throughput <br> • No molecular barcodes used <br> • Can be expensive to scale up |

## Reviews

Moignard V. and Gottgens B. Dissecting stem cell differentiation using single cell expression profiling. *Curr Opin Cell Biol.* 2016;43:78-86

Woodhouse S., Moignard V., Gottgens B. and Fisher J. Processing, visualising and reconstructing network models from single-cell data. *Immunol Cell Biol.* 2016;94:256-265

Wagner A., Regev A. and Yosef N. Revealing the vectors of cellular identity with single-cell genomics. *Nat Biotechnol.* 2016;34:1145-1160

Stegle O., Teichmann S. A. and Marioni J. C. Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet.* 2015;16:133-145

Kolodziejczyk A. A., Kim J. K., Svensson V., Marioni J. C. and Teichmann S. A. The Technology and Biology of Single-Cell RNA Sequencing. *Mol Cell.* 2015;58:610-620

## References

Lee J. H., Daugharthy E. R., Scheiman J., et al. Fluorescent *in situ* sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat Protoc.* 2015;10:442-458

Single-cell mRNA sequencing (scRNA-Seq) can profile gene expression over the entire cell transcriptome, but cell isolation typically results in loss of spatial context. In situ hybridization is an excellent technique for identifying the location of gene expression, but it is restricted to a fixed number of genes. In this study, the authors present a protocol for in situ profiling of gene expression in cells and tissues. In this approach, RNA is converted into crosslinked cDNA amplicons and sequenced manually on a confocal microscope. The approach has the benefit of enriching for context-specific transcripts over housekeeping/structural genes, while preserving the tissue architecture for transcript localization.

Illumina Technology: MiSeq System, Nextera XT DNA Sample Prep Kit

Padovan-Merhar O., Nair G. P., Biaesch A. G., et al. Single Mammalian Cells Compensate for Differences in Cellular Volume and DNA Copy Number through Independent Global Transcriptional Mechanisms. *Mol Cell.* 2015;58:339-352

Individual mammalian cells display a wide diversity in cellular size and volume, yet they maintain the same absolute DNA content. Therefore, it is likely that compensatory mechanisms exist to maintain a constant concentration of gene expression products, despite differences in DNA concentration. In this study, the authors used single-molecule counting and single-cell image analysis to demonstrate that individual human primary foreskin fibroblasts globally control transcription to compensate for variability in the ratio of DNA to cellular content. They performed scRNA-Seq using the NextSeq 500 system and found that ubiquitously expressed housekeeping genes exhibited lower levels of expression noise than other genes.

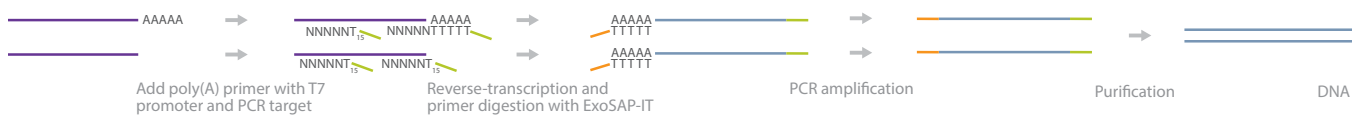Illumina Technology: NextSeq 500 System, Nextera XT DNA Sample Prep Kit

## Associated Kits

TruSeq RNA Library Prep Kit v2

---

164.  Tang F., Barbacioru C., Wang Y., et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods.* 2009;6:377-382

# SUPeR-Seq: Single-Cell Universal Poly(A)-Independent RNA Sequencing

SUPeR-Seq sequences non-poly(A) and poly(A) RNAs from single cells. It is designed particularly for mapping circular RNA (circRNA) species.[165]

RNA samples from lysed single cells are annealed to random primers with universal anchor sequences (AnchorX-T15N6) and reverse-transcribed to generate the first strand of cDNA. Unreacted primers are digested, to avoid primer-dimers. Next, a poly(A) tract is added to the 3' end of the cDNA by introducing dATP and ddATP in a 100:1 ratio, respectively. A second set of random primers, also with a universal anchor sequence (AnchorY-T24), anneals to the newly synthesized poly(A) tract. A second cDNA strand is generated by RT, and the cDNA is purified by gel electrophoresis. The purified cDNA molecules are PCR-amplified using 5'-amine-terminated primers, prepared by the TruSeq DNA library preparation protocol, and sequenced. circRNAs are identified from the dataset by finding 2 exonic reads that are distal in the reference genome, but adjacent to each other in the dataset with 1 inverted, signifying the circularization of the RNA.



A schematic overview of SUPeR-seq.

| Advantages | Disadvantages |
|---|---|
| • Identifies circRNAs from single cells<br>• Avoids 3' bias by using random primers with anchor sequences<br>• Able to identify novel circRNAs due to random primers | • Relies on dataset analysis to identify circRNAs |

## References

Dang Y., Yan L., Hu B., et al. Tracing the expression of circular RNAs in human pre-implantation embryos. *Genome Biol.* 2016;17:130

This study focused on enhancing the understanding of circRNAs, a novel type of poly(A)– RNA, in human and mouse embryos. Using SUPeR-seq, the authors identified 10,032 circRNAs from 2974 host genes through *de novo* sequencing. Additionally, they compared human and mouse embryos and discovered that human cells generate more types of circRNAs than mouse.

Illumina Technology:  HiSeq 2500 System, TruSeq DNA Library Prep Kit
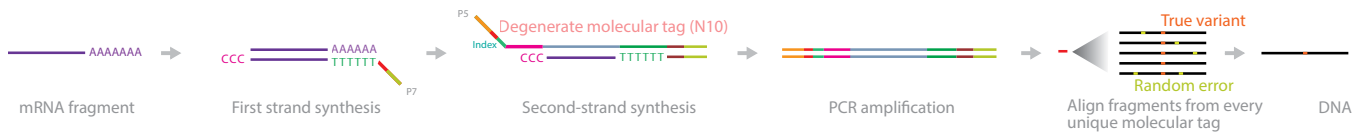
## Associated Kits

TruSeq Nano DNA Library Prep Kit

165.   Fan X., Zhang X., Wu X., et al. Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. *Genome Biol.* 2015;16:148

# UMI: Unique Molecular Identifiers

Unique molecular identifiers (UMI) are molecular tags that are used to detect and quantify unique mRNA transcripts.[166]

In this method, mRNA libraries are generated by fragmentation and reverse-transcribed to cDNA. Oligo(dT) primers with specific sequencing linkers are added to the cDNA. Another sequencing linker with a 10 bp random label and an index sequence is added to the 5' end of the template, which is amplified and sequenced. Sequencing allows for high-resolution reads, enabling accurate detection of true variants.



| mRNA fragment | First strand synthesis | Second-strand synthesis | PCR amplification | Align fragments from every unique molecular tag | DNA |

A schematic overview of UMI.

| Advantages | Disadvantages |
|---|---|
| • Can sequence unique mRNA transcripts<br>• Can detect transcripts occurring at low frequencies<br>• Transcripts can be quantified based on sequencing reads specific to each barcode<br>• Can be applied to multiple platforms to karyotype chromosomes | • Targets smaller than 500 bp are preferentially amplified by polymerases during PCR |

## References
This method has been widely integrated into various sequencing techniques due to its high versatility.

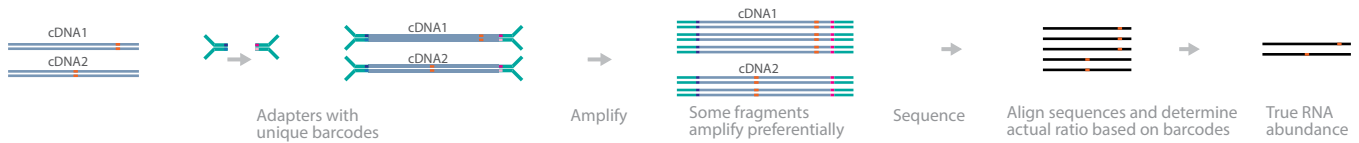## Associated Kits
TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

166. Kivioja T., Vaharautio A., Karlsson K., et al. Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods.* 2012;9:72-74

# Digital RNA Sequencing

Digital RNA sequencing is an approach to RNA-Seq that removes sequence-dependent PCR amplification biases by barcoding the RNA molecules before amplification.[167] RNA is reverse-transcribed to cDNA, and an excess of adapters, each with a unique barcode, is added to the preparation. This barcoded cDNA is amplified and sequenced. Deep sequencing reads are compared, and the barcodes are used to determine the actual distribution of RNA abundance.



A schematic overview of digital RNA sequencing.

| Advantages | Disadvantages |
|---|---|
| • Low amplification bias during PCR<br>• Provides information about abundance of RNA<br>• Detection of low-copy–number RNA<br>• Single-copy resolution | • Some amplification bias still persists<br>• Barcodes may miss targets during ligation |

## References
This method has been widely integrated into various sequencing techniques due to its high versatility.

## Associated Kits
TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Stranded mRNA Library Prep Kit

TruSeq Stranded Total RNA Library Prep Kit

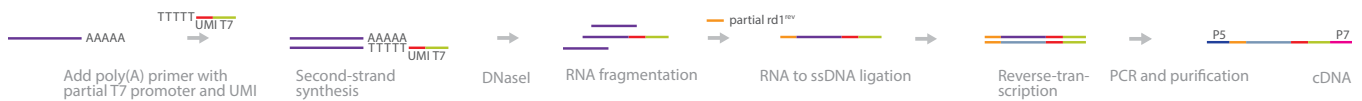TruSeq Targeted RNA Expression Library Prep Kit

167. Shiroguchi K., Jia T. Z., Sims P. A. and Xie X. S. Digital RNA sequencing minimizes sequence-dependent bias and amplification noise with optimized single-molecule barcodes. *Proc Natl Acad Sci U S A.* 2012;109:1347-1352

# MARS-Seq: Massively Parallel RNA Single-Cell Sequencing Framework

MARS-Seq profiles the transcriptional dynamics of single cells in an automated and massively parallel workflow with high resolution.[168] MARS-Seq can be used with *in vivo* samples containing a wide variety of different cell subpopulations.

Single cells are first isolated into individual wells using FACS. Each cell is lysed, and the 3' ends of mRNAs are annealed to unique molecular identifiers containing a T7 promoter. The mRNA is reverse-transcribed to generate the first cDNA strand and treated with exonuclease I to remove leftover RT primers. Next, the cellular lysates are pooled together and converted to double-stranded cDNA. The DNA strands are transcribed to RNA and treated with DNase to remove leftover DNA templates in the mixture. The RNA strands are fragmented and annealed to sequencing adapters, followed by RT to generate barcoded cDNA libraries that are ready for sequencing.

Similar methods: CEL-Seq, Quartz-Seq, Drop-seq, CytoSeq, inDrop



A schematic overview of MARS-Seq.

| Advantages | Disadvantages |
|---|---|
| • High-throughput transcriptional profiling of single cells<br>• *In vivo* sampling of thousands of cells<br>• Three barcode levels (molecular, cellular, and plate-level tags) facilitate robust multiplexing capabilities<br>• Processes 100 to 1000 single cells<br>• Pooling all single cells into 1 flow cell reduces the cost to less than 50 cents per cell[169] | • 3' bias can occur during the purification step<br>• Fragmentation step eliminates strand-specific information[170] |

## Reviews

Hrdlickova R., Toloue M. and Tian B. RNA-Seq methods for transcriptome analysis. *Wiley Interdisc Rev RNA*. 2016;

Wen L. and Tang F. Single-cell sequencing in stem cell biology. *Genome Biol.* 2016;17:71

Jaitin D. A., Keren-Shaul H., Elefant N. and Amit I. Each cell counts: Hematopoiesis and immunity research in the era of single cell genomics. *Semin Immunol.* 2015; 27:67-71

## References

Paul F., Arkin Y., Giladi A., et al. Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell*. 2015;163:1663-1677

The authors used MARS-Seq to study single-cell transcriptomic profiles of bone marrow myeloid progenitor cells. The results revealed unexpected mechanisms of transcriptional priming toward 7 differentiation fates. They also observed that mixed lineage states do not appear when regulation is disturbed. This study provides a new reference model for single-cell hematopoietic cell research.

Illumina Technology: HiSeq 1500 System, NextSeq 500 System

Baruch K., Deczkowska A., Rosenzweig N., et al. PD-1 immune checkpoint blockade reduces pathology and improves memory in mouse models of Alzheimer's disease. Nat Med. 2016;22:135-137

## Associated Kits

TruSeq RNA Library Prep Kit v2

168. Jaitin D. A., Kenigsberg E., Keren-Shaul H., et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science*. 2014;343:776-779
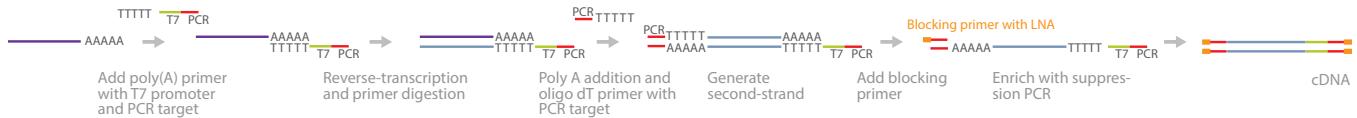169. Jaitin D. A., Keren-Shaul H., Elefant N. and Amit I. Each cell counts: Hematopoiesis and immunity research in the era of single cell genomics. *Semin Immunol*. 2015;27:67-71
170. Hrdlickova R., Toloue M. and Tian B. RNA-Seq methods for transcriptome analysis. *Wiley Interdiscip Rev RNA*. 2016;

# Quartz-Seq: Whole-Transcript Amplification for Single Cells

The Quartz-Seq method optimizes whole-transcript amplification (WTA) of single cells.[171] In this method, an RT primer with a T7 promoter and PCR target is first added to the extracted mRNA. RT synthesizes first-strand cDNA, after which the RT primer is digested by exonuclease I. Next, a poly(A) tail is added to the 3' ends of first-strand cDNA, along with a poly(dT) primer containing a PCR target. After second-strand generation, a blocking primer is added to ensure PCR enrichment in sufficient quantity for sequencing. Deep sequencing allows for accurate, high-resolution representation of the whole transcriptome of a single cell.

Similar methods: CEL-Seq, Drop-seq, MARS-Seq, CytoSeq, inDrop, Hi-SCL



A schematic overview of Quartz-Seq.

| Advantages | Disadvantages |
|---|---|
| • Single-tube reaction suitable for automation | • PCR biases can underrepresent GC-rich templates |
| • Digestion of RT primers by exonuclease I eliminates amplification of byproducts | • Amplification errors caused by polymerases will be represented and sequenced incorrectly |
| • Short fragments and byproducts are suppressed during enrichment | • Targets smaller than 500 bp are preferentially amplified by polymerases during PCR |

## Reviews

Zhang X., Marjani S. L., Hu Z., Weissman S. M., Pan X. and Wu S. Single-Cell Sequencing for Precise Cancer Research: Progress and Prospects. *Cancer Research.* 2016;

Poulin J. F., Tasic B., Hjerling-Leffler J., Trimarchi J. M. and Awatramani R. Disentangling neural cell diversity using single-cell transcriptomics. *Nat Neurosci.* 2016;19:1131-1141

Sun H. J., Chen J., Ni B., Yang X. and Wu Y. Z. Recent advances and current issues in single-cell sequencing of tumors. *Cancer Lett.* 2015;365:1-10

Grun D. and van Oudenaarden A. Design and Analysis of Single-Cell Sequencing Experiments. *Cell.* 2015;163:799-810

Navin N. E. Cancer genomics: one cell at a time. *Genome Biol.* 2014;15:452

Liang J., Cai W. and Sun Z. Single-Cell Sequencing Technologies: Current and Future. *J Genet Genomics.* 2014;41:513-528

## References

Takeuchi M., Yamaguchi S., Sakakibara Y., Hayashi T., Matsuda K., et al. (2016) Gene expression profiling of granule cells and Purkinje cells in the zebrafish cerebellum. J Comp Neurol

The authors studied the neural circuitry in the cerebellum of transgenic zebrafish larvae. They used a modified Quartz-Seq method to perform WTA based on cell mass instead of single cells. The results revealed that more developmental genes are expressed in granule cells, while neuronal-function genes are increasingly expressed in Purkinje cells.

Illumina Technology: HiSeq 1500 System

Archer N., Walsh M. D., Shahrezaei V. and Hebenstreit D. Modeling Enzyme Processivity Reveals that RNA-Seq Libraries Are Biased in Characteristic and Correctable Ways. *Cell Syst.* 2016;

Scialdone A., Natarajan K. N., Saraiva L. R., et al. Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods.* 2015;
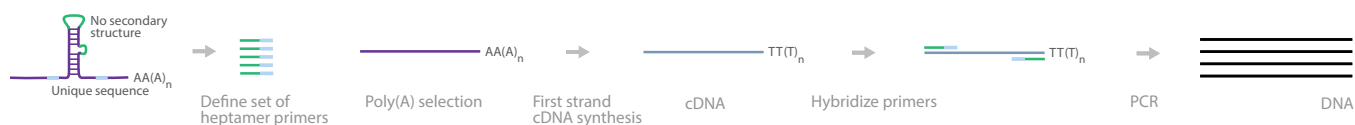
## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

---

171. Sasagawa Y., Nikaido I., Hayashi T., Danno H., Uno K. D., et al. (2013) Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome Biol* 14: R31

# DP-Seq: Designed Primer–Based RNA Sequencing

DP-Seq amplifies mRNA from limited starting material, as low as 50 pg.[172] In this method, a specific set of heptamer primers is designed. The enriched poly(A)-selected mRNA undergoes first-strand cDNA synthesis. Next, the designed primers are hybridized to the first-strand cDNA, followed by second-strand synthesis and PCR. Deep sequencing of the amplified DNA allows for accurate detection of specific mRNA expression at the single-cell level.



A schematic overview of DP-seq.

| Advantages | Disadvantages |
| --- | --- |
| • As little as 50 pg of starting material can be used<br>• Little transcript-length bias | • Sequences of the target areas must be known to design the heptamers<br>• Exponential amplification during PCR can lead to primer-dimers and spurious PCR products[173]<br>• Some read-length bias |

## Reviews

Friedmann-Morvinski D., Bhargava V., Gupta S., Verma I. M. and Subramaniam S. Identification of therapeutic targets for glioblastoma by network analysis. *Oncogene*. 2015;

Kolodziejczyk A. A., Kim J. K., Svensson V., Marioni J. C. and Teichmann S. A. The Technology and Biology of Single-Cell RNA Sequencing. *Mol Cell.* 2015;58:610-620

Head S. R., Komori H. K., LaMere S. A., et al. Library construction for next-generation sequencing: overviews and challenges. *Biotechniques*. 2014;56:61-64, 66, 68, passim

## References

Bhargava V., Head S. R., Ordoukhanian P., Mercola M. and Subramaniam S. Technical variations in low-input RNA-seq methodologies. *Sci Rep*. 2014;4:3678

This study compared the quality of sequencing libraries and reads generated from 3 different amplification-based, low-level RNA-Seq methods: Smart-Seq, DP-seq, and cell expression by linear amplification and sequencing (CEL-Seq). When compared to Smart-Seq, DP-Seq produced less PCR biases, more consistent relative transcript abundance, and it was the most cost-effective of the 3 methods. Despite having higher representation of short transcripts, Smart-Seq gave high transcriptome coverage for different amounts of mRNA and uniform coverage along the transcript length. CEL-Seq, however, had the highest technical variations and distortions among the 3 methods.

Illumina Technology: HiSeq 2000 System, TruSeq RNA Library Prep Kit

## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

---

172. Bhargava V., Ko P., Willems E., Mercola M. and Subramaniam S. (2013) Quantitative transcriptomics using designed primer-based amplification. *Sci Rep* 3: 1740
173. Bhargava V., Head S. R., Ordoukhanian P., Mercola M. and Subramaniam S. Technical variations in low-input RNA-seq methodologies. *Sci Rep.* 2014;4:3678

# Smart-Seq and Smart-Seq2: Switch Mechanism at the 5' End of RNA Templates

Smart-Seq was developed as a single-cell sequencing protocol with improved read coverage across transcripts.[174] Complete coverage across the genome allows the detection of alternative transcript isoforms and SNPs.

There are 2 versions of Smart-Seq: Smart-Seq and Smart-seq2. Smart-seq2 includes several improvements over the original Smart-Seq protocol.[175,176] The new protocol includes a locked nucleic acid (LNA), an increased MgCl2 concentration, betaine, and elimination of the purification step to improve the yield significantly.
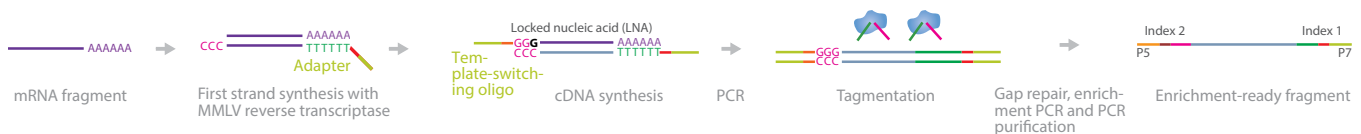
Smart-Seq: Cells are lysed, and the RNA is hybridized to an oligo(dT)-containing primer. The first strand of the cDNA is synthesized with the addition of a few untemplated C nucleotides. This poly(C) overhang is added exclusively to full-length transcripts. An oligonucleotide primer is hybridized to the poly(C) overhang and used to synthesize the second strand. Full-length cDNAs are PCR-amplified to obtain nanogram amounts of DNA. The PCR products are purified for sequencing.



A schematic overview of Smart-Seq.

| Advantages | Disadvantages |
|---|---|
| • mRNA sequence does not have to be known<br>• As little as 50 pg of starting material can be used<br>• Improved coverage across transcripts<br>• High level of mappable reads | • Not strand-specific<br>• No early multiplexing<br>• Applicable only to poly(A)+ RNA |

Smart-seq2: Single cells are lysed in a buffer that contains free dNTPs and oligo(dT)-tailed oligonucleotides with a universal 5'-anchor sequence. RT is performed, which adds 2–5 untemplated nucleotides to the cDNA 3' end. A template-switching oligo (TSO) is added, carrying 2 riboguanosines and a modified guanosine to produce a LNA as the last base at the 3' end. After the first-strand reaction, the cDNA is amplified using a limited number of cycles. Next, tagmentation is used to construct sequencing libraries quickly and efficiently from the amplified cDNA.



A schematic overview of Smart-seq2.

| Advantages | Disadvantages |
|---|---|
| • As little as 50 pg of starting material can be used<br>• mRNA sequence does not have to be known<br>• Improved coverage across transcripts<br>• High level of mappable reads | • Not strand-specific<br>• No early multiplexing[177]<br>• Transcript length bias, with inefficient transcription of reads over 4 Kb[178]<br>• Preferential amplification of high-abundance transcripts<br>• Purification step may lead to loss of material<br>• Could be subject to strand-invasion bias[179] |

174. Ramskold D., Luo S., Wang Y. C., et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol.* 2012;30:777-782

175. Picelli S., Bjorklund A. K., Faridani O. R., Sagasser S., Winberg G., et al. (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods* 10: 1096-1098

176. Picelli S., Faridani O. R., Björklund Å. K., Winberg G., Sagasser S., et al. (2014) Full-length RNA-Seq from single cells using Smart-seq2. *Nat. Protocols* 9: 171-181

177. Shapiro E., Biezuner T. and Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet.* 2013;14:618-630

178. Bhargava V., Head S. R., Ordoukhanian P., Mercola M. and Subramaniam S. Technical variations in low-input RNA-seq methodologies. *Sci Rep.* 2014;4:3678

179. Tang D. T., Plessy C., Salimullah M., et al. Suppression of artifacts and barcode bias in high-throughput transcriptome analyses utilizing template switching. *Nucleic Acids Res.* 2013;41:e44

## References

This method has been widely integrated into various sequencing techniques due to its high versatility.

## Associated Kits

Nextera DNA Library Prep Kit
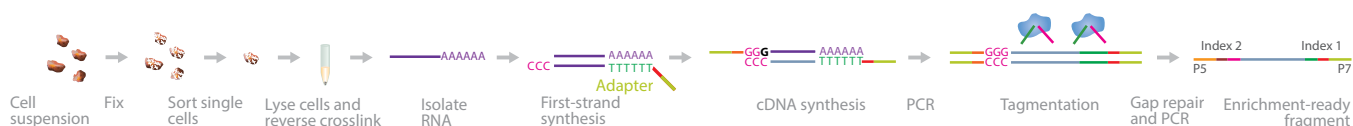
TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

# FRISCR: Fixed and Recovered Intact Single-Cell RNA

FRISCR characterizes transcriptome profiles from fixed and stained single cells.[180] The method uses a combination of molecular barcodes and Tn5 tagmentation to identify each cDNA fragment uniquely from every cell.

The cell suspension is fixed with paraformaldehyde, permeabilized, and immunostained. Individual cells are sorted into tubes by FACS. These cells are lysed and reverse-crosslinked by incubation at 56°C for 1 hour. mRNA from the cells is isolated by $dT_{25}$ magnetic bead pull-down. The mRNA sequencing library is prepared according to the Smart-seq2 procedure: 1) template-switching RT with Moloney murine leukemia virus reverse transcriptase; 2) the resulting cDNAs are PCR amplified; and 3) the cDNA library is generated using the Nextera XT Library Preparation Kit. The cDNA fragments are flanked with adapters and are ready for sequencing.



A schematic overview of FRISCR.

| Advantages | Disadvantages |
|---|---|
| • Full-length mRNA transcriptome profiling from fixed and stained single cells<br>• Immunostaining enables targeting of rare cell populations<br>• Generates full-length mRNA reads<br>• Significantly more mRNA recovered compared to fixed cells from Triton-X100 lysis | • Possible 3' to 5' bias |

## Reviews

Wen L. and Tang F. Single-cell sequencing in stem cell biology. Genome Biol. 2016;17:71

## References

Thomsen E. R., Mich J. K., Yao Z., Hodge R. D., Doyle A. M., et al. (2016) Fixed single-cell transcriptomic characterization of human radial glial diversity. *Nat Methods* 13: 87-93

The authors developed FRISCR to identify radial glia (RG) accurately. These cells are rare and can be defined only through a combination of intracellular markers, position, and morphology. They performed FRISCR on primary human RG, which are represented in only 1% of the midgestation cortex. Using FRISCR, the researchers identified markers and molecular profiles for ventricular zone–enriched RG (vRG) and outer subventricular zone–localized RG (oRG).

Illumina Technology: MiSeq System, Nextera XT DNA Library Prep Kit
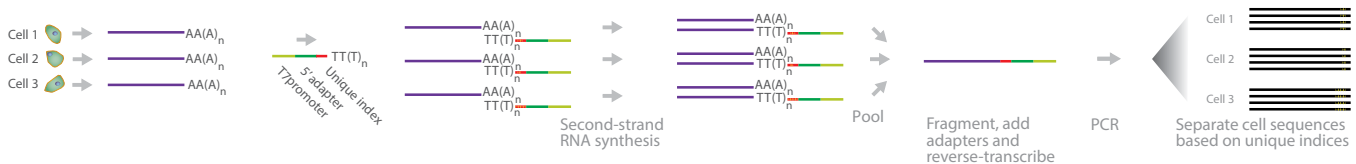
## Associated Kits

Nextera XT DNA Library Prep Kit

---

180. Thomsen E. R., Mich J. K., Yao Z., et al. Fixed single-cell transcriptomic characterization of human radial glial diversity. *Nat Methods*. 2016;13:87-93

# CEL-Seq: Cell Expression by Linear Amplification Sequencing

CEL-Seq uses barcoding and pooling of RNA to overcome challenges from low input.[181] In this method, each cell undergoes RT with a unique barcoded primer in its individual tube. After second-strand synthesis, cDNAs from all reaction tubes are pooled and PCR-amplified. Paired-end deep sequencing of the PCR products allows for accurate detection of sequence information derived from both strands.

Similar methods: CEL-Seq2, Quartz-Seq, Drop-seq, MARS-Seq, CytoSeq, inDrop, Hi-SCL



A schematic overview of CEL-Seq.

| Advantages | Disadvantages |
|---|---|
| • Barcoding and pooling allow for multiplexing and studying many different single cells at a time<br>• Cross-contamination is greatly reduced due to using 1 tube per cell<br>• Fewer steps than single-cell tagged reverse-transcription sequencing (STRT-Seq)<br>• Very little read-length bias[182]<br>• Strand-specific | • Strongly 3' biased[183]<br>• Abundant transcripts are amplified preferentially<br>• Requires at least 400 pg of total RNA |

## Reviews

Zhang X., Marjani S. L., Hu Z., Weissman S. M., Pan X., et al. Single-Cell Sequencing for Precise Cancer Research: Progress and Prospects. *Cancer Research*. 2016;

Grun D. and van Oudenaarden A. Design and Analysis of Single-Cell Sequencing Experiments. *Cell*. 2015;163:799-810

Kolodziejczyk A. A., Kim J. K., Svensson V., Marioni J. C. and Teichmann S. A. The Technology and Biology of Single-Cell RNA Sequencing. *Mol Cell*. 2015;58:610-620

Liang J., Cai W. and Sun Z. Single-Cell Sequencing Technologies: Current and Future. *J Genet Genomics*. 2014;41:513-528

## References

Levin M., Anavy L., Cole A. G., et al. The mid-developmental transition and the evolution of animal body plans. *Nature*. 2016;

The authors aimed to enhance the understanding of the phylotypic period—a stage during the development of species widely used in morphological and molecular analyses for species categorization— that is conserved in many species across different phyla. They used CEL-Seq to analyze the transcriptome profile of individual embryos from 10 distantly related species in the animal kingdom (from poriferans to tardigrades). Transcriptome comparison revealed a variety of signaling pathways and transcription factors that are specific to each species during the phylotypic period.

Illumina Technology: HiSeq 2000 System, TruSeq RNA Library Prep Kit

Bose S., Wan Z., Carr A., Rizvi A. H., Vieira G., et al. (2015) Scalable microfluidics for single-cell RNA printing and sequencing. *Genome Biol* 16: 120

In this study, the authors present a new scalable, high-density microfluidic platform for solid-phase capture of RNA on glass coverslips or polymer beads. They trapped single-cell lysates in sealed picoliter microwells capable of printing RNA on glass or capturing RNA on beads. They combined this sample-preparation approach with a scalable technology for scRNA-Seq based on CEL-Seq. The technology is relatively inexpensive, with consumable costs of $0.10–0.20 per cell and is capable of processing hundreds of individual cells in parallel.

Illumina Technology: NextSeq 500 System, HiSeq 2500 System, TruSeq RNA Library Prep Kit

181. Hashimshony T., Wagner F., Sher N. and Yanai I. (2012) CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep* 2: 666-673
182. Bhargava V., Head S. R., Ordoukhanian P., Mercola M. and Subramaniam S. Technical variations in low-input RNA-seq methodologies. *Sci Rep*. 2014;4:3678
183. Shapiro E., Biezuner T. and Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet*. 2013;14:618-630

Seillet C., Mielke L. A., Amann-Zalcenstein D. B., et al. Deciphering the Innate Lymphoid Cell Transcriptional Program. Cell Rep. 2016;17:436-447

Mooijman D., Dey S. S., Boisset J. C., Crosetto N. and van Oudenaarden A. Single-cell 5hmC sequencing reveals chromosome-wide cell-to-cell variability and enables lineage reconstruction. Nat Biotechnol. 2016;

Thomsen E. R., Mich J. K., Yao Z., Hodge R. D., Doyle A. M., et al. Fixed single-cell transcriptomic characterization of human radial glial diversity. Nat Methods. 2016;13:87-93

Grun D., Lyubimova A., Kester L., et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature. 2015;

Klein A. M., Mazutis L., Akartuna I., et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell. 2015;161:1187-1201

Grun D., Kester L. and van Oudenaarden A. Validation of noise models for single-cell transcriptomics. Nat Methods. 2014;11:637-640

Bhargava V., Head S. R., Ordoukhanian P., Mercola M. and Subramaniam S. Technical variations in low-input RNA-seq methodologies. Sci Rep. 2014;4:3678

Hashimshony T., Feder M., Levin M., Hall B. K. and Yanai I. Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. Nature. 2014;
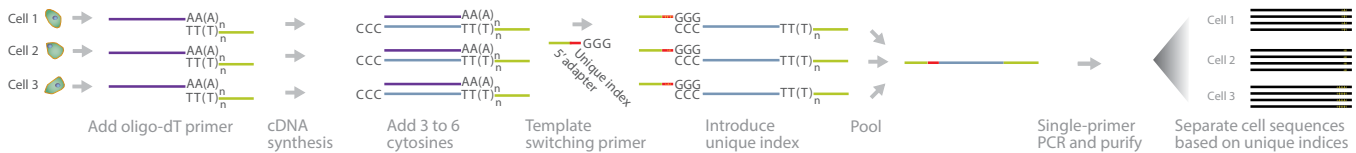
## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Small RNA Library Prep Kit

TruSeq Targeted RNA Expression Library Prep Kit

# STRT-Seq: Single-Cell Tagged Reverse Transcription Sequencing

STRT-Seq is a method similar to CEL-Seq that involves unique barcoding and sample pooling to overcome the challenges of samples with limited material.[184,185] In this method, single cells are first picked in individual tubes, where first-strand cDNA synthesis occurs using an oligo(dT) primer with the addition of 3–6 cytosines. A helper oligonucleotide promotes template switching, which introduces the barcode into the cDNA. The barcoded cDNA is amplified by single-primer PCR. Deep sequencing allows for accurate transcriptome determination of individual cells.



A schematic overview of STRT-Seq.

| Advantages | Disadvantages |
|---|---|
| • Barcoding and pooling allows for multiplexing and studying many different single cells at a time<br>• Sample handling and the potential for cross-contamination are greatly reduced due to using a single tube per cell | • PCR biases can underrepresent GC-rich templates<br>• Nonlinear PCR amplification can lead to biases affecting reproducibility<br>• Amplification errors caused by polymerases will be represented and sequenced incorrectly<br>• Loss of accuracy due to PCR bias<br>• Targets smaller than 500 bp are amplified preferentially by polymerases during PCR |

## Reviews

Macaulay I. C. and Voet T. Single cell genomics: advances and future perspectives. *PLoS Genet*. 2014;10:e1004126

Kolodziejczyk A. A., Kim J. K., Svensson V., Marioni J. C. and Teichmann S. A. The Technology and Biology of Single-Cell RNA Sequencing. *Mol Cell.* 2015;58:610-620

Sun H. J., Chen J., Ni B., Yang X. and Wu Y. Z. Recent advances and current issues in single-cell sequencing of tumors. *Cancer Lett.* 2015;365:1-10

Grun D. and van Oudenaarden A. Design and Analysis of Single-Cell Sequencing Experiments. *Cell*. 2015;163:799-810

Zhang X., Marjani S. L., Hu Z., Weissman S. M., Pan X., et al. Single-Cell Sequencing for Precise Cancer Research: Progress and Prospects. *Cancer Research*. 2016;

Navin N. E. Cancer genomics: one cell at a time. *Genome Biol.* 2014;15:452

## References

Korber I., Katayama S., Einarsdottir E., et al. Gene-Expression Profiling Suggests Impaired Signaling via the Interferon Pathway in Cstb-/- Microglia. *PLoS One.* 2016;11:e0158195

This study explored the effects of loss-of-function mutations in the gene for cysteine protease inhibitor cystatin B (CSTB) to progressive myoclonus epilepsy of Unverricht-Lundborg type (EPM1). The authors performed modified STRT-Seq and microarray hybridization to gain a transcriptomic analysis of microglial RNA samples. They observed 184 differentially expressed genes in CSTB–/– microglial cells. Deeper investigation of CSTB deficiency uncovered a potential function of CSTB in regulating chemotaxis, antigen presentation, and immune- and defense-response through the JAK-STAT pathway.

Illumina Technology: HiSeq 2000 System, TruSeq Targeted RNA Expression Library Prep Kit

Töhönen V., Katayama S., Vesterlund L., et al. Novel PRD-like homeodomain transcription factors and retrotransposon elements in early human development. *Nature Communications*. 2015;6:8207

Katayama S., Skoog T., Jouhilahti E. M., et al. Gene expression analysis of skin grafts and cultured keratinocytes using synthetic RNA normalization reveals insights into differentiation and growth control. *BMC Genomics.* 2015;16:476

## Associated Kits

TruSeq Targeted RNA Expression Library Prep Kit

184. Islam S., Kjallquist U., Moliner A., Zajac P., Fan J. B., et al. (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res* 21: 1160-1167
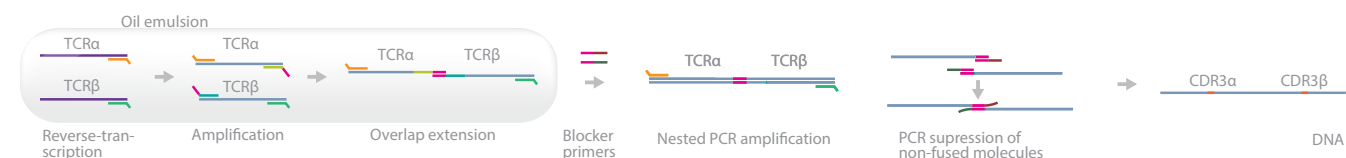185. Islam S., Kjallquist U., Moliner A., et al. Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. *Nat Protoc.* 2012;7:813-828

# TCR Chain Pairing: Identification of T-Cell Receptor (TCR) $\alpha$-$\beta$ Chain Pairing in Single Cells

This method identifies TCR-$\alpha$ and -$\beta$ chain pairing in single cells using cell-based emulsion technology for isolation, followed by NGS.[186] TCR chain pairing resolves one of the biggest challenges of identifying coexpressed gene pairs—random, nonspecific overlap extension of nonfused molecules—by introducing a unique PCR suppression technique during post–emulsion amplification reactions.

First, single T cells are isolated into oil emulsion droplets containing RT primers for $\alpha$ and $\beta$ chain mRNA strands. The resultant cDNA is amplified, and the RT primer extensions overlapped to connect TCR-$\alpha$ and TCR-$\beta$ strands, which are now called fused molecules. The cDNA products are extracted from the emulsion and introduced to blocking primers. These primers anneal to the 3' end of nonfused cDNA strands, preventing them from being amplified; the authors name this technique "PCR suppression." Finally, the fused molecules are PCR-amplified and sequenced.

Similar methods: TCR-LA-MC PCR



A schematic overview of TCR chain pairing.

| Advantages | Disadvantages |
| --- | --- |
| • Identifies TCR-$\alpha$-$\beta$ chain pairing using NGS<br>• PCR suppression of nonfused molecules significantly reduces random, nonspecific overlap extension during post–emulsion amplification reactions<br>• Simple and straightforward protocol | • Heat-shock during cell lysis may reduce enzymatic activity during RT[187]<br>• Amplification product is not suitable for cloning[188] |

## Reviews

Friedensohn S., Khan T. A. and Reddy S. T. Advanced Methodologies in High-Throughput Sequencing of Immune Repertoires. *Trends in Biotechnology*. 2016;

## References

Munson D. J., Egelston C. A., Chiotti K. E., et al. Identification of shared TCR sequences from T cells in human breast cancer using emulsion RT-PCR. *Proceedings of the National Academy of Sciences*. 2016;113:8272-8277

The authors wanted to determine if breast cancer immunotherapeutic strategies can be improved by increasing the targeting specificity of T cells. They used TCR chain pairing to link and amplify TCR pairs, and they observed ~85% accurate pairing fidelity. They further discovered that, for any TCR pair, a dominant $\alpha$- or $\beta$-binding partner comprised ~90% of the total binding partners.

Illumina Technology: MiSeq System

## Associated Kits

TruSeq RNA Library Prep Kit v2

186. Turchaninova M. A., Britanova O. V., Bolotin D. A., et al. Pairing of T-cell receptor chains via emulsion PCR. *Eur J Immunol.* 2013;43:2507-2515
187. Friedensohn S., Khan T. A. and Reddy S. T. Advanced Methodologies in High-Throughput Sequencing of Immune Repertoires. *Trends in Biotechnology*. 2016;
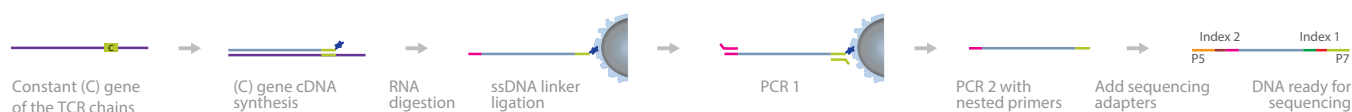188. Sprouse M. L., Blahnik G., Lee T., et al. Rapid identification and expression of human TCRs in retrogenic mice. *J Immunol Methods.* 2016;

# TCR-LA-MC PCR: TCR Ligation-Anchored Magnetically Captured PCR

TCR-LA-MC PCR identifies TCR-α and -β chains from T cells and uses sequencing to analyze the catalog of clonal TCR *in vivo* or *in vitro* from blood or tissue samples.[189] This technique can be performed with as little as 10 ng of cDNA, while providing great sensitivity and accuracy for analyzing the diversity and mechanisms that affect TCR clonality.

First-strand cDNA is generated using biotinylated primers that anneal to the constant gene of the TCR chains. RNA strands are removed by RNA digestion, and the single-stranded cDNA is captured magnetically using streptavidin beads. Single-stranded linker cassettes (ssLCs) containing primer sequences are ligated to the cDNA, and the samples are PCR-amplified. The double-stranded cDNAs are flanked with sequencing adapters and are ready for sequencing.

Similar methods: TCR chain pairing



A schematic overview of TCR-LA-MC PCR.

| Advantages | Disadvantages |
|---|---|
| • Identifies TCR diversity without sequence-associated or quantitative restrictions<br>• Can be used to study the diversity and mechanism of TCRs that influences clonality<br>• As little as 10 ng of cDNA can be used as input<br>• Identifies T-cells with 1:10,000 resolution capacity, or even single cells | • None reported yet |

## Reviews

Hou D., Chen C., Seely E. J., Chen S. and Song Y. High-Throughput Sequencing-Based Immune Repertoire Study during Infectious Disease. *Front Immunol.* 2016;7:336

## References

Oliveira G., Ruggiero E., Stanghellini M. T. L., et al. Tracking genetically engineered lymphocytes long-term reveals the dynamics of T cell immunological memory. *Science Translational Medicine*. 2015;7:317ra198-317ra198

This study used TCR-LA-MC PCR to analyze TCR-α and TCR-β chain sequences in genetically modified human T lymphocytes. The results showed that dominant long-term clonotypes originated from infused TSCM and TCM clones. The authors suggest that the original phenotype of infused cells and exposure to antigens influence long-term persistence in gene-modified memory T cells.

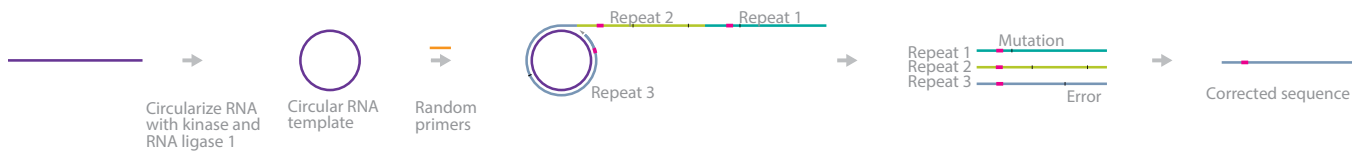Illumina Technology: MiSeq System

## Associated Kits

TruSeq RNA Library Prep Kit v2

189.   Ruggiero E., Nicolay J. P., Fronza R., et al. High-resolution analysis of the human T-cell receptor repertoire. *Nat Commun.* 2015;6:8081

# CirSeq: Identification of Low-Abundance RNA Viruses with Circular Sequencin

CirSeq accurately identifies ultra-rare and low-frequency genetic variants in RNA viruses. The method uses a unique step for circularization of fragmented viral RNAs, followed by rolling-circle RT.[190,191] CirSeq corrects for mutations introduced during its amplification steps by aligning the tandem-repeat sequences with each other and excluding those reads using informatics tools.

First, single-stranded RNAs are fragmented using $Zn^{2+}$ and size-selected to no more than one-third of the sequencing read length. Next, they are circularized and reverse-transcribed using random primers. Rolling-circle RT is used to generate tandem-repeat cDNA strands. The first-strand cDNAs are amplified, generating double-stranded cDNAs, followed by end repair, poly(A) tailing, and adapter ligation. The cDNA libraries are ready for sequencing.



A schematic overview of CirSeq.

| Advantages | Disadvantages |
|---|---|
| • Detects ultra-rare and low-frequency genetic variants in RNA viruses <br> • Rolling-circle RT creates tandem-repeat cDNAs that can be used to correct artificial mutations <br> • Error rates reported are far below those observed using standard RNA virus sequencing methods | • Whole process takes ~5 days <br> • Unsuitable for sequencing clinical isolates, because it needs large quantities of purified viral RNAs <br> • Not applicable for *de novo* sequencing of viral RNAs |

## Reviews

Posada-Cespedes S., Seifert D. and Beerenwinkel N. Recent advances in inferring viral diversity from high-throughput sequencing data. *Virus Res.* 2016;

Andino R. and Domingo E. Viral quasispecies. *Virology*. 2015;479-480C:46-51

Gordon A. J., Satory D., Halliday J. A. and Herman C. Lost in transcription: transient errors in information transfer. *Curr Opin Microbiol.* 2015;24C:80-87

## References

Acevedo A., Brodsky L. and Andino R. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature*. 2014;505:686-690

The authors developed CirSeq to interrogate low-frequency variants in viral genetic structures. They used CirSeq to define mutation rates and unravel the mutational landscape of poliovirus. Analysis of fluctuations in variant frequencies enabled them to assign fitness values for thousands of mutations across the genome. Further, they were able to map the fitness values to 3-dimensional structures of viral proteins, giving unprecedented power to correlate protein structure with function.

Illumina Technology: HiSeq System, MiSeq System

Wang K., Ma Q., Jiang L., et al. Ultra-precise detection of mutations by droplet-based amplification of circularized DNA. *BMC Genomics.* 2016;17:214
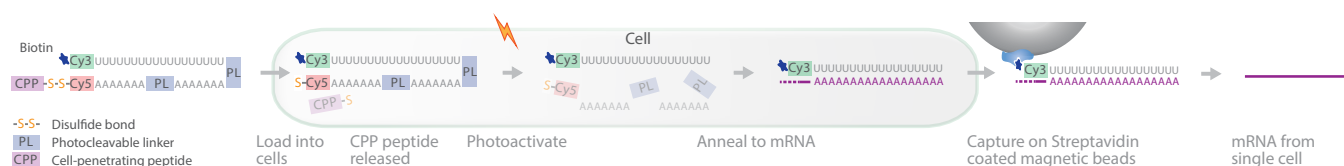
## Associated Kits

TruSeq RNA Library Prep Kit v2

---

190. Acevedo A., Brodsky L. and Andino R. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. Nature. 2014;505:686-690
191. Acevedo A. and Andino R. Library preparation for highly accurate population sequencing of RNA viruses. Nat Protoc. 2014;9:1760-1769

# TIVA: Transcriptome *In Vivo* Analysis

TIVA is a protocol that captures mRNA from live cells.[192]

In this method, a photoactivatable TIVA tag is loaded into cells. Selective photoactivation exposes the mRNA-capturing portion of the tag, allowing it to hybridize to the poly(A) tails of mRNA. The biotin-bound mRNA is captured using streptavidin-coated magnetic beads and transcribed into cDNA. Sequencing the cDNA provides transcriptome analysis of RNA from single cells in complex tissues.

A schematic overview of TIVA.

| Advantages | Disadvantages |
|---|---|
| • *In vivo* transcriptome analysis using photoactivation in neuronal cells<br>• Noninvasive method for capturing mRNAs in their natural microenvironment | • Limited to a small number of cells |

## Reviews

Zhang X., Marjani S. L., Hu Z., Weissman S. M., Pan X., et al. Single-Cell Sequencing for Precise Cancer Research: Progress and Prospects. *Cancer Research*. 2016;

Binan L., Mazzaferri J., Choquet K., et al. Live single-cell laser tag. *Nat Commun*. 2016;7:11636

Bian Q. and Cahan P. Computational Tools for Stem Cell Biology. *Trends Biotechnol*. 2016;

Liu S. and Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges. *F1000Res*. 2016;5:

Achim K., Pettit J. B., Saraiva L. R., et al. High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nat Biotechnol*. 2015;

Kolodziejczyk A. A., Kim J. K., Svensson V., Marioni J. C. and Teichmann S. A. The Technology and Biology of Single-Cell RNA Sequencing. *Mol Cell*. 2015;58:610-620

Baslan T. and Hicks J. Single cell sequencing approaches for complex biological systems. *Curr Opin Genet Dev*. 2014;26C:59-65

## References

Lovatt D., Ruble B. K., Lee J., Dueck H., Kim T. K., et al. Transcriptome *in vivo* analysis (TIVA) of spatially defined single cells in live tissue. *Nat Methods*. 2014;11:190-196

The authors developed TIVA to analyze transcriptome variance in single neurons in culture, as well as mouse and human neurons *in vivo*. The results showed that the microenvironment in tissues controls the transcriptomic landscape down to the level of individual cells.

Illumina Technology: HiSeq System, TruSeq RNA Library Prep Kit

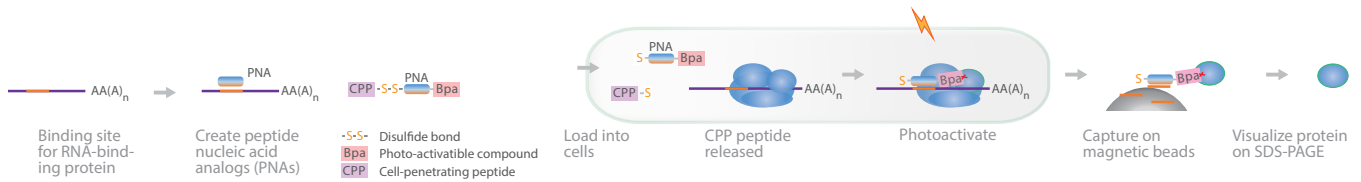## Associated Kits

TruSeq RNA Library Prep Kit v2

TruSeq Nano DNA Library Prep Kit

TruSeq DNA PCR-Free Library Prep Kit

---

192.  Lovatt D., Ruble B. K., Lee J., et al. Transcriptome in vivo analysis (TIVA) of spatially defined single cells in live tissue. *Nat Methods*. 2014;11:190-196

---

# PAIR: Peptide Nucleic Acid (PNA)-Assisted Identification of RNA Binding Proteins

PAIR uses PNAs to capture RBPs in vivo.[193,194] The PNAs are coupled to a cell membrane-penetrating peptide (CPP) to deliver PNAs efficiently into living cells, as well as a photoactivatable compound, p-benzoylphenylalanine (Bpa). The cells are illuminated with UV light, activating the Bpa on the PNA to form covalent bonds with the RBP. Next, the cells are lysed, and the RNA complexes are captured on magnetic beads. The proteins can be reverse-crosslinked and visualized by denaturing gel electrophoresis, while the RNA strands are sequenced.



A schematic overview of PAIR.

| Advantages | Disadvantages |
|---|---|
| • Identifies RNA-protein interactions in vivo<br>• Commercial PNAs can be purchased, eliminating the need for PNA design | • Custom PNA design can cause problems with low protein yield<br>• Not yet adopted widely by the scientific community |

## Reviews

Jazurek M., Ciesiolka A., Starega-Roslan J., Bilinska K. and Krzyzosiak W. J. Identifying proteins that bind to specific RNAs - focus on simple repeat expansion diseases. *Nucleic Acids Res.* 2016;44:9050-9070

McHugh C. A., Russell P. and Guttman M. Methods for comprehensive experimental identification of RNA-protein interactions. *Genome Biology*. 2014;15:203

## References

Zielinski J., Kilk K., Peritz T., Kannanayakal T., Miyashiro K. Y., et al. In vivo identification of ribonucleoprotein-RNA interactions. *Proc Natl Acad Sci U S A.* 2006;103:1557-1562

The authors developed PAIR to gain a deeper understanding of RBPs in gene expression regulation. They performed PAIR on ankylosis mRNAs and observed specific regions in the 3'- and 5'-UTRs that were targeted by the PNAs. Additionally, they discovered a variety of proteins that were associated with the PNAs through the crosslinking step in PAIR.

Illumina Technology: Unspecified Illumina sequencing system

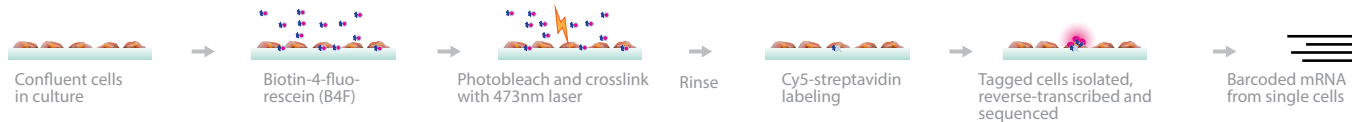## Associated Kits

TruSeq RNA Library Prep Kit v2

---

193.  Zielinski J., Kilk K., Peritz T., et al. In vivo identification of ribonucleoprotein-RNA interactions. *Proc Natl Acad Sci U S A.* 2006;103:1557-1562
194.  Zeng F., Peritz T., Kannanayakal T. J., et al. A protocol for PAIR: PNA-assisted identification of RNA binding proteins in living cells. *Nat Protoc.* 2006;1:920-927

# CLaP: Cell Labeling via Photobleaching

CLaP is a noninvasive, laser-based labeling technique for single cells.[195] The method uses lasers to crosslink specific cells with fluorescent tags before isolating individual cells for sequencing.

In CLaP, cells of interest are tagged by crosslinking biotin-4-fluorescein (B4F) to the cell membrane with laser irradiation. Next, streptavidin-conjugated fluorescent labels are bound to biotinylated cells. These steps can be repeated to tag multiple cell types with a variety of fluorescent tags. The tagged cells are subsequently isolated and processed to generate cDNA libraries for sequencing.



| Confluent cells in culture | Biotin-4-fluo-rescein (B4F) | Photobleach and crosslink with 473nm laser | Rinse | Cy5-streptavidin labeling | Tagged cells isolated, reverse-transcribed and sequenced | Barcoded mRNA from single cells |

A schematic overview of PAIR.

| Advantages | Disadvantages |
|---|---|
| • Noninvasive, targeted, laser-based single-cell labeling<br>• Enables automated, image-based cell selection<br>• Fluorescence-based tags can be substituted with other labels, such as electron-dense molecules<br>• Multicolored fluorescent stains can be used | • Diffusion of reagents through the extracellular matrix and continuous laser illumination limit the procedure for 3-dimensional environments/tissues<br>• Cellular specificity may be decreased slightly in primary cell cultures.<br>• Image-based selection limits the potential for high-throughput applications |

## Reviews

Gurwitz D. Human iPSC-derived neurons and lymphoblastoid cells for personalized medicine research in neuropsychiatric disorders. *Dialogues Clin Neurosci*. 2016;18:267-276

## References

Binan L., Mazzaferri J., Choquet K., Lorenzo L. E., Wang Y. C., et al. Live single-cell laser tag. *Nat Commun.* 2016;7:11636

The authors developed CLaP to tag individual cells instantly, based on shape, behavior, or positional information. They showed that the tag is stable, nontoxic, retained in cells for days, and is transferred during cell division but not to neighboring cells in the same culture. They also combined CLaP with microfluidics-based single-cell capture techniques to increase throughput. Further, they showed that CLaP can be used to induce transient cell adhesion to different substrates in microengineered culture environments with spatially patterned cell types.

Illumina Technology: HiSeq 2500 System, Nextera XT DNA Library Prep Kit

## Associated Kits

Nextera XT DNA Library Prep Kit

---

195. Binan L., Mazzaferri J., Choquet K., et al. Live single-cell laser tag. *Nat Commun.* 2016;7:11636

# CytoSeq: Gene Expression Cytometry

CytoSeq enables gene expression profiling of thousands of single cells.[196] In this method, single cells are randomly deposited into wells. A combinatorial library of beads with specific capture probes is added to each well. After cell lysis, mRNAs hybridize the to beads, which are pooled subsequently for RT, amplification, and sequencing. Deep sequencing provides accurate, high-coverage gene expression profiles of several single cells.

Similar methods: CEL-Seq, Quartz-Seq, MARS-Seq, inDrop, Hi-SCL

A schematic overview of CytoSeq.

| Advantages | Disadvantages |
|---|---|
| • Can readily scale to 10,000s or 100,000s of cells<br>• Complements and expands the capabilities of fluorescence or mass spectrometry–based cytometry<br>• Detects any transcribed mRNA without the limitations of antibody availability<br>• Enables rare cell characterization on small samples with insufficient cells for traditional flow cytometry<br>• Allows direct analysis of complex samples of heterogeneous cell size and shape | • Sequencing depth requires large number of reads (eg, 200,000 transcripts per cell requires 2 million reads for 10X coverage: 2 billion reads for 1000 cells)<br>• Single run can be relatively expensive and time-consuming<br>• Trade-off between depth of sequencing and differential gene expression |

## Reviews

Friedensohn S., Khan T. A. and Reddy S. T. Advanced Methodologies in High-Throughput Sequencing of Immune Repertoires. *Trends in Biotechnology.* 2016;

Kolodziejczyk A. A., Kim J. K., Svensson V., Marioni J. C. and Teichmann S. A. The Technology and Biology of Single-Cell RNA Sequencing. *Mol Cell.* 2015;58:610-620

Saadatpour A., Lai S., Guo G. and Yuan G. C. Single-Cell Analysis in Cancer Genomics. *Trends Genet.* 2015;31:576-586

## References

Fan H. C., Fu G. K. and Fodor S. P. Expression profiling. Combinatorial labeling of single cells for gene expression cytometry. *Science.* 2015;347:1258367

The authors developed CytoSeq to combine molecular barcoding with high-throughput single-cell sequencing. They performed CytoSeq on human hematopoietic cells to characterize their response to *in vitro* stimulation. The method proved to be robust in detecting low-frequency transcripts and rare cell populations. CytoSeq was able to analyze thousands of cells simultaneously and is easily scalable for 10,000s or 100,000s of cells.

Illumina Technology: MiSeq System, TruSeq Nano DNA Library Prep Kit

## Associated Kits

TruSeq Nano DNA Library Prep Kit

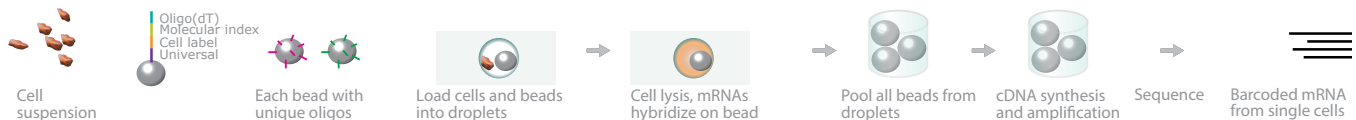TruSeq DNA PCR-Free Library Prep Kit

196. Fan H. C., Fu G. K. and Fodor S. P. Expression profiling. Combinatorial labeling of single cells for gene expression cytometry. *Science.* 2015;347:1258367

# Drop-Seq: Analysis of mRNA Transcripts from Individual Cells in Droplets

Drop-Seq analyzes mRNA transcripts from droplets of individual cells in a highly parallel fashion.[197]

This single-cell sequencing method uses a microfluidic device to compartmentalize droplets containing a single cell, lysis buffer, and a microbead covered with barcoded primers. Each primer contains: 1) a 30 bp oligo(dT) sequence to bind mRNAs; 2) an 8 bp molecular index to identify each mRNA strand uniquely; 3) a 12 bp barcode unique to each cell and 4) a universal sequence identical across all beads. Following compartmentalization, cells in the droplets are lysed and the released mRNA hybridizes to the oligo(dT) tract of the primer beads. Next, all droplets are pooled and broken to release the beads within. After the beads are isolated, they are reverse-transcribed with template switching. This generates the first cDNA strand with a PCR primer sequence in place of the universal sequence. cDNAs are PCR-amplified, and sequencing adapters are added using the Nextera XT Library Preparation Kit. The barcoded mRNA samples are ready for sequencing.

Similar methods: CEL-Seq, Quartz-Seq, MARS-Seq, CytoSeq, inDrop, Hi-SCL



A schematic overview of Drop-seq.

| Advantages | Disadvantages |
|---|---|
| • Analyzes sequences of single cells in a highly parallel manner<br>• Unique molecular and cell barcodes enable cell- and gene-specific identification of mRNA strands<br>• RT with template-switching PCR produces high-yield reads from single cells<br>• Low cost: $0.07 per cell ($653 per 10,000 cells) and fast library preparation (10,000 cells per day) | • Requires custom microfluidics device to perform droplet separation<br>• Low gene-per-cell sensitivity compared to other scRNA-Seq methods[198]<br>• Limited to mRNA transcripts |

## Reviews

Bowen J. R., Ferris M. T. and Suthar M. S. Systems biology: A tool for charting the antiviral landscape. *Virus Res.* 2016;

Chaitankar V., Karakulah G., Ratnapriya R., Giuste F. O., Brooks M. J. and Swaroop A. Next generation sequencing technology and genomewide data analysis: Perspectives for retinal research. *Prog Retin Eye Res.* 2016;

Zhao Q.-Y., Gratten J., Restuadi R. and Li X. Mapping and differential expression analysis from short-read RNA-Seq data in model organisms. *Quantitative Biology.* 2016;4:22-35

Conesa A., Madrigal P., Tarazona S., Gomez-Cabrero D., Cervera A., et al. A survey of best practices for RNA-seq data analysis. *Genome Biol*. 2016;17:13

Friedensohn S., Khan T. A. and Reddy S. T. Advanced Methodologies in High-Throughput Sequencing of Immune Repertoires. *Trends in Biotechnology.* 2016;

Poulin J. F., Tasic B., Hjerling-Leffler J., Trimarchi J. M. and Awatramani R. Disentangling neural cell diversity using single-cell transcriptomics. *Nat Neurosci.* 2016;19:1131-1141

Grun D. and van Oudenaarden A. Design and Analysis of Single-Cell Sequencing Experiments. *Cell*. 2015;163:799-810

Saadatpour A., Lai S., Guo G. and Yuan G. C. Single-Cell Analysis in Cancer Genomics. *Trends Genet.* 2015;31:576-586

Alizadeh A. A., Aranda V., Bardelli A., et al. Toward understanding and exploiting tumor heterogeneity. *Nat Med.* 2015;21:846-853

197. Macosko E. Z., Basu A., Satija R., et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*. 2015;161:1202-1214
198. Ziegenhain C., Parekh S., Vieth B., et al. Comparative analysis of single-cell RNA-sequencing methods. *bioRxiv*. 2016;

## References

Ziegenhain C., Parekh S., Vieth B., Smets M., Leonhardt H., et al. Comparative analysis of single-cell RNA-sequencing methods. *bioRxiv*. 2016;

The authors performed comparative analysis of Drop-seq, single-cell RNA barcoding and sequencing (SCRB-Seq), Smart-Seq, and Smart-seq2 using 447 mouse ESCs. They determined that Drop-Seq was the most preferable method, in terms of cost efficiency, to quantify transcriptomes for a large number of cells at varying depth at 80% power. Smart-seq2 gave unparalleled sensitivity in detecting the most genes per cells, but also the most amplification noise. SCRB-Seq was preferable for transcriptomic quantification in low numbers of cells.

Illumina Technology: HiSeq 1500 System, HiSeq 2500 System, Nextera XT DNA Library Prep Kit


## Associated Kits
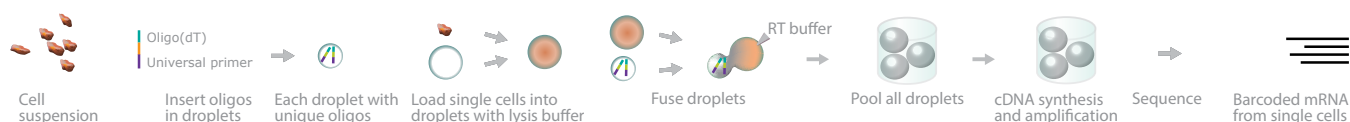
TruSeq Nano DNA Library Prep Kit

TruSeq DNA PCR-Free Library Prep Kit

# Hi-SCL: High-Throughput Single-Cell Labeling

Hi-SCL generates transcriptome profiles for thousands of single cells using a custom microfluidics system, similar to Drop-Seq and inDrop.[199]

Single cells from cell suspension are isolated into droplets containing lysis buffer. After cell lysis, cell droplets are fused with a droplet containing cell-specific barcodes and another droplet with enzymes for RT. The droplets from all the wells are pooled and subjected to isothermal reactions for RT. The barcodes anneal to poly(A)$^+$ mRNAs and act as primers for reverse transcriptase. Now that each mRNA strand has cell-specific barcodes, the droplets are broken, and the cDNA is purified. The 3' ends of the cDNA strands are ligated to adapters, amplified, annealed to indexed primers, and amplified further before sequencing.

Similar methods: CEL-Seq, Drop-seq, MARS-Seq, CytoSeq, inDrop, Quartz-Seq



A schematic overview of Hi-SCL.

| Advantages | Disadvantages |
|---|---|
| • High-throughput, single-cell transcriptome profiling using a microfluidics system<br>• Low cost: $0.1 per cell (for experiment with 100 cells)<br>• Highly scalable to larger cell quantities<br>• No fragmentation step | • Lack of UMI in oligonucleotides may create amplification noise<br>• Droplets may contain 2 cells or 2 different types of barcodes |

## Reviews

Mato Prado M., Frampton A. E., Stebbing J. and Krell J. Single-cell sequencing in cancer research. *Expert Rev Mol Diagn*. 2016;16:1-5

## References

Rotem A., Ram O., Shoresh N., Sperling R. A., Schnall-Levin M., et al. (2015) High-Throughput Single-Cell Labeling (Hi-SCL) for RNA-Seq Using Drop-Based Microfluidics. *PLoS One* 10: e0116328

The authors developed Hi-SCL to tackle the problem of high cost associated with single-cell sequencing. Using droplet-based microfluidics, barcoded oligonucleotide libraries, and in-drop cDNA synthesis, they were able to bring the cost down to $0.1 per cell in an experiment with 100 cells. This study was published to validate the data generated from the technique and prove its potential for experiments that require a large number of cells.

Illumina Technology: MiSeq System, HiSeq System

## Associated Kits

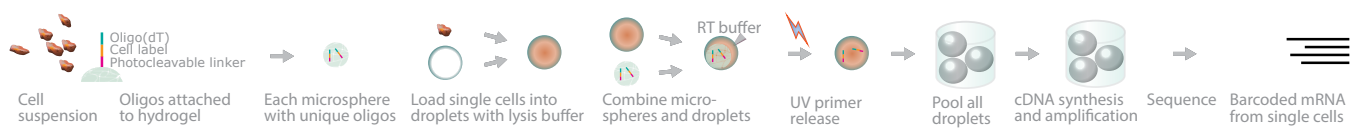TruSeq RNA Library Prep Kit v2

---

199. Rotem A., Ram O., Shoresh N., et al. High-Throughput Single-Cell Labeling (Hi-SCL) for RNA-Seq Using Drop-Based Microfluidics. *PLoS One*. 2015;10:e0116328

# InDrop: High-Throughput Single-Cell Labeling with Indexing Droplets

InDrop is used for high-throughput single-cell labeling.[200] This approach is similar to Drop-seq, but it uses hydrogel microspheres to introduce the oligonucleotides.

Single cells from a cell suspension are isolated into droplets containing lysis buffer. After cell lysis, cell droplets are fused with a hydrogel microsphere containing cell-specific barcodes and another droplet with enzymes for RT. Droplets from all the wells are pooled and subjected to isothermal reactions for RT. The barcodes anneal to poly(A)+ mRNAs and act as primers for reverse transcriptase. Now that each mRNA strand has cell-specific barcodes, the droplets are pooled and broken, and the cDNA is purified. The 3' ends of the cDNA strands are ligated to adapters, amplified, annealed to indexed primers, and amplified further before sequencing.

Similar methods: CEL-Seq, Drop-seq, MARS-Seq, CytoSeq, Quartz-Seq, Hi-SCL



A schematic overview of inDrop.

| Advantages | Disadvantages |
|---|---|
| • High-throughput, single-cell transcriptome profiling using a microfluidics system<br>• Highly scalable to larger cell quantities<br>• No fragmentation step | • Low mRNA capture efficiency of ~7%<br>• Droplets may contain 2 cells or 2 different types of barcodes |

## Reviews

Bian Q. and Cahan P. Computational Tools for Stem Cell Biology. *Trends Biotechnol.* 2016;

Zhao Q.-Y., Gratten J., Restuadi R. and Li X. Mapping and differential expression analysis from short-read RNA-Seq data in model organisms. *Quantitative Biology.* 2016;4:22-35

Grun D. and van Oudenaarden A. Design and Analysis of Single-Cell Sequencing Experiments. *Cell*. 2015;163:799-810

Saadatpour A., Lai S., Guo G. and Yuan G. C. Single-Cell Analysis in Cancer Genomics. *Trends Genet.* 2015;31:576-586

## References

Derr A., Yang C., Zilionis R., et al. End Sequence Analysis Toolkit (ESAT) expands the extractable information from single-cell RNA-seq data. *Genome Res.* 2016;26:1397-1410

The authors developed End Sequence Analysis Toolkit (ESAT) to remedy computational artifacts produced during global RNA-Seq. They performed inDrop on 1000 individual pancreatic islet cells and were able to identify 9 distinct cell types, 3 distinct  -cell types, and complex interactions between hormone secretion and vacularization.

Illumina Technology: HiSeq System

Klein A. M., Mazutis L., Akartuna I., Tallapragada N., Veres A., et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*. 2015;161:1187-1201

## Associated Kits

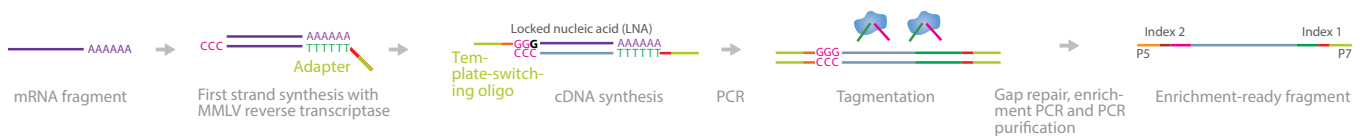Nextera XT DNA Library Prep Kit

---

200.  Klein A. M., Mazutis L., Akartuna I., et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*. 2015;161:1187-1201

---

# snRNA-Seq: Single-Nuclei RNA Sequencing

snRNA-Seq uses a mild and quick nuclear dissociation protocol to isolate and sequence RNA within the nucleus. The method minimizes technical issues that can arise from common dissociation protocols, especially in studying immediate early gene (IEG) behavior.[201]

In this method, the cell suspension is lysed gently and the nuclei are separated from cytoplasmic lysates by centrifugation. Single cells/nuclei are sorted into individual wells using FACS. Individual nuclei are amplified using a microfluidics-assisted machinery that performs cell capture and reaction chemistry. The nuclear RNA contents are processed into cDNA libraries using a Nextera XT DNA Library Prep Kit. Pools of 40 samples are collected and purified using magnetic beads. The cDNA library is ready for sequencing.

Similar methods: Div-Seq, Nuc-seq



A schematic overview of snRNA-seq.

| Advantages | Disadvantages |
|---|---|
| • Rapid dissociation protocol prevents technical issues arising from protease digestion, heating, and spurious gene expression by cytoplasmic ribosomes<br>• Prevents dendritic loss that commonly occurs during protease dissociation step | • None reported yet |

## Reviews

Poulin J.-F., Tasic B., Hjerling-Leffler J., Trimarchi J. M. and Awatramani R. Disentangling neural cell diversity using single-cell transcriptomics. *Nat Neurosci.* 2016;19:1131-1141

Gagliano S. A. It's All in the Brain: A Review of Available Functional Genomic Annotations. *Biol Psychiatry.* 2016;

Goncalves J. T., Schafer S. T. and Gage F. H. Adult Neurogenesis in the Hippocampus: From Stem Cells to Behavior. *Cell.* 2016;167:897-914

## References

Lacar B., Linker S. B., Jaeger B. N., et al. Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat Commun.* 2016;7:11022

The authors studied the transcriptomics of activated neurons upon exposure to different experiences and environments. To reduce experimental errors and noise from traditional nuclear dissociation protocols, they developed a milder and quicker protocol. snRNA-Seq was able to produce expression profiles not only for the activation of IEGs but also activation states beyond IEGs.

Illumina Technology: HiSeq 2500 System, Nextera XT DNA Library Prep Kit

## Associated Kits

Nextera XT DNA Library Prep Kit

---

201.  Lacar B., Linker S. B., Jaeger B. N., Krishnaswami S., Barron J., et al. (2016) Nuclear RNA-Seq of single neurons reveals molecular signatures of activation. *Nat Commun* 7: 11022

# Nuc-Seq: A Single-Nucleus RNA-Seq from Frozen Tissues

Nuc-Seq is an RNA sequencing technique optimized for isolating and sequencing nuclear RNA from frozen tissue samples.[202] Nuc-Seq avoids proteolytic treatment during nuclei dissociation from cells, minimizing gene expression changes resulting from common protease dissociation procedures; this step is similar to the one used in snRNA-seq. Next, Smart-seq2 is used for cDNA synthesis, increasing the full-length cDNA yield due to its template-switching mechanism. Finally, the sequencing library is prepared using Tn5 transposase tagmentation.

Similar methods: snRNA-seq, Div-Seq



A schematic overview of Nuc-seq.

| Advantages | Disadvantages |
|---|---|
| • Nonproteolytic nuclear dissociation minimizes gene expression changes during isolation of nuclei<br>• Template-switching mechanism of Smart-seq2 during cDNA synthesis increases the yield of full-length cDNA strands<br>• FACS isolation increases throughput<br>• RNA from frozen nuclei gives better cDNA quality than cytoplasmic RNA<br>• Can be used on frozen tissues | • Excludes any information from cytoplasmic RNA<br>• Low-copy transcripts are hard to detect due to low amounts of isolated nucleolar RNA<br>• ncRNAs and other short non-poly(A) RNA species are challenging to detect |

## Reviews

Jennings C. G., Landman R., Zhou Y., et al. Opportunities and challenges in modeling human brain disorders in transgenic primates. *Nat Neurosci*. 2016;19:1123-1130

## References

**Krishnaswami S. R., Grindberg R. V., Novotny M., Venepally P., Lacar B., et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat Protoc.* 2016;11:499-524**

This study provides an in-depth protocol for Nuc-seq, a sensitive approach for isolating and deep sequencing of nuclei from frozen tissue samples.

Illumina Technology: MiSeq System, Nextera XT DNA Library Prep Kit

## Associated Kits

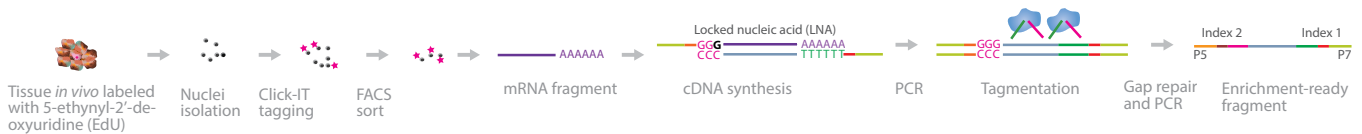Nextera XT DNA Library Prep Kit

---

202. Krishnaswami S. R., Grindberg R. V., Novotny M., et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat Protoc.* 2016;11:499-524

# Div-Seq: Nuc-Seq with EdU-Mediated Labeling of Proliferating Cells

Div-Seq is a single-nucleus RNA sequencing technique that improves upon Nuc-Seq by incorporating 5-ethynyl-2'-deoxyuridine (EdU) labeling to identify dividing cells during their different cell stages.[203] EdU labeling also enables identification of different cell types in complex tissue samples and rare cell populationsduring FACS.

Briefly, samples are labeled *in vivo* with EdU, dissected, and fixed before isolation into single nuclei. Individual nuclei are tagged fluorescently and sorted by FACS. From this step, the procedure follows the Nuc-Seq method: cDNA synthesis is performed using the Smart-seq2 protocol, while the cDNA library is prepared with Tn5 transposase–mediated tagmentation.

Similar methods: Nuc-seq, snRNA-seq



A schematic overview of Div-Seq.

| Advantages | Disadvantages |
|---|---|
| • Tracks transcriptome dynamics in single nuclei<br>• Detects rare cell populations<br>• Compatible with fresh, frozen, or fixed sample types<br>• EdU labeling gives unbiased identification of different types of dividing cells and their current stage in the cell cycle<br>• Mild nuclear dissociation technique minimizes gene expression changes commonly seen in protease-mediated dissociation | • Excludes any information from cytoplasmic RNA |

## Reviews

Wagner A., Regev A. and Yosef N. Revealing the vectors of cellular identity with single-cell genomics. Nat Biotech. 2016;34:1145-1160

Goncalves J. T., Schafer S. T. and Gage F. H. Adult Neurogenesis in the Hippocampus: From Stem Cells to Behavior. Cell. 2016;167:897-914

## References

Habib N., Li Y., Heidenreich M., Swiech L., Avraham-Davidi I., et al. Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science*. 2016;

The authors developed Div-Seq as an improvement to Nuc-seq, incorporating pulse labeling of proliferating cells with EdU. They were able to track transcriptome profiles of nascent neurons in the adult hippocampal neurogenic environment. Using Div-Seq, they also identified rare newborn GABAergic neurons in the adult spinal cord.

Illumina technology: NextSeq 500 System, Nextera XT DNA Library Prep Kit
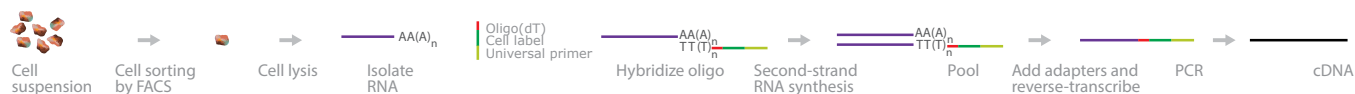
## Associated Kits

Nextera XT DNA Library Prep Kit

---

203.  Habib N., Li Y., Heidenreich M., et al. Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science*. 2016;

# SCRB-Seq: Single-Cell RNA Barcoding and Sequencing

SCRB-Seq is a cost-efficient, multiplexed, single-cell mRNA sequencing technique.[204]

SCRB-Seq isolates single cells into wells using FACS. After cell lysis, poly(A)+ mRNAs are annealed to a custom primer containing a poly(T) tract, UMI, well barcode, and biotin. Template-switching RT and PCR amplification reactions are carried out on the mRNA, generating barcoded, full-length cDNA. cDNA strands from all wells are pooled together to be purified. They are PCR-amplified and purified further. The cDNA libraries are prepared using the Nextera XT library preparation protocol, with modified i5 primers. The resultant cDNA fragments are size-selected for 300–800 bp and sequenced.



A schematic overview of SCRB-Seq.

| Advantages | Disadvantages |
| --- | --- |
| • Cost-efficient, high-throughput, single-cell transcriptome profiling<br>• Highly sensitive gene-detection results compared to popular scRNA-Seq techniques[205] | • Template-switching RT is heavily biased to full-length mRNA[206] |

## Reviews

Ziegenhain C., Parekh S., Vieth B., Smets M., Leonhardt H., et al. Comparative analysis of single-cell RNA-sequencing methods. *bioRxiv*. 2016;

Shapiro E., Biezuner T. and Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet.* 2013;14:618-630

## References

Cacchiarelli D., Trapnell C., Ziller M. J., et al. Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of Induced Pluripotency. *Cell.* 2015;162:412-424

To study induced pluripotency in human cells, the authors performed modified SCRB-Seq to obtain genomic and epigenomic profiles of single cells. The results revealed unique patterns of gene network activation and ordered reactivation of developmental regulators. Further, analyses of complementary functions in the reprogramming process uncovered new regulators of the system.

Illumina Technology: HiSeq 2000 System, HiSeq 2500 System, MiSeq System, TruSeq RNA Library Prep Kit, TruSeq Small RNA Library Prep Kit

Anahtar M. N., Byrne E. H., Doherty K. E., et al. Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. Immunity. 2015;42:965-976

## Associated Kits
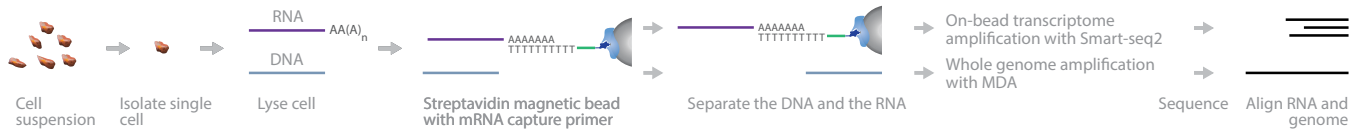
Nextera XT DNA Library Prep Kit

204. Soumillon M., Cacchiarelli D., Semrau S., van Oudenaarden A. and Mikkelsen T. S. Characterization of directed differentiation by high-throughput single-cell RNA-Seq. 2014;
205. Ziegenhain C., Parekh S., Vieth B., et al. Comparative analysis of single-cell RNA-sequencing methods. *bioRxiv*. 2016;
206. Shapiro E., Biezuner T. and Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet.* 2013;14:618-630

# G&T-Seq: Genome and Transcriptome Sequencing

G&T-Seq can separate and sequence genomic DNA and full-length mRNA from single cells.[207] In this method, single cells are isolated and lysed. RNA is captured using biotinylated oligo(dT) capture primers and separated from DNA using streptavidin-coated magnetic beads. Smart-seq2 is used to amplify captured RNA on the bead, while multiple displacement amplification (MDA) is used to amplify DNA. After sequencing, integrating the DNA and RNA sequences provides insights into the gene expression profiles of single cells.



A schematic overview of G&T-seq.

| Advantages | Disadvantages |
|---|---|
| • Compatible with any whole-genome amplification method<br>• No 3'-end bias in sequence reads because full-length transcripts are captured<br>• Because DNA and RNA are physically separated and amplified independently, there is no need to mask coding sequences during analysis | • Physical separation of DNA and RNA can increase the risk of sample loss or contamination<br>• Physical separation of DNA and RNA increases handling time |

## Reviews

Bock C., Farlik M. and Sheffield N. C. Multi-Omics of Single Cells: Strategies and Applications. *Trends in Biotechnology.* 2016;34:605-608

Clark S. J., Lee H. J., Smallwood S. A., Kelsey G. and Reik W. Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity. *Genome Biology*. 2016;17:72

Zhang X., Marjani S. L., Hu Z., Weissman S. M., Pan X. and Wu S. Single-Cell Sequencing for Precise Cancer Research: Progress and Prospects. *Cancer Research*. 2016;76:1305-1312

Vieira Braga F. A., Teichmann S. A. and Chen X. Genetics and immunity in the era of single-cell genomics. *Human Molecular Genetics.* 2016;25:R141-R148

Lu L., Lv B., Huang K., Xue Z., Zhu X. and Fan G. Recent advances in preimplantation genetic diagnosis and screening. *Journal of Assisted Reproduction and Genetics.* 2016;33:1129-1134

## References

Angermueller C., Clark S. J., Lee H. J., et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods.* 2016;13:229-232

Multiparameter single-cell sequencing is a powerful tool that has uncovered relationships among genomic, transcriptional, and epigenetic heterogeneity. In this study, the authors developed single-cell methylome and transcriptome sequencing (scM&T-Seq), a multiparameter sequencing method that allows methylome and transcriptome profiling in the same cell. They used the G&T-Seq protocol to purify single-cell DNA that was subjected to single-cell bisulfite conversion (scBS-Seq). The authors performed scM&T-Seq on 61 mouse ESCs. They found that gene expression levels of many pluripotency factors were negatively associated with DNA methylation. These data demonstrate that epigenetic heterogeneity is an important mechanism of fluctuating pluripotency in ESCs. They also demonstrate that scM&T-Seq can illuminate the poorly understood relationship between transcriptional and DNA-methylation heterogeneity in single cells.

Illumina Technology: HiSeq 2000 System, Nextera XT DNA Library Prep Kit

## Associated Kits

Nextera XT DNA Library Prep Kit

TruSeq RNA Library Prep Kit v2

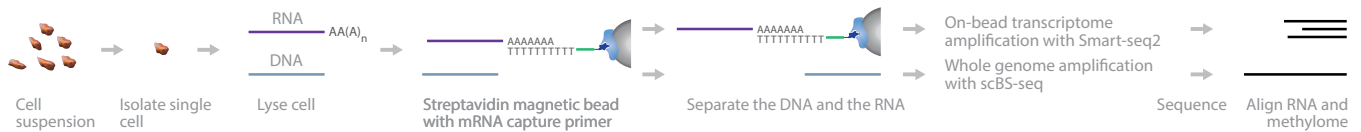TruSeq Small RNA Library Prep Kit

TruSeq Nano DNA Library Prep Kit

TruSeq DNA PCR-Free Library Prep Kit

---

207. Macaulay I. C., Haerty W., Kumar P., et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods.* 2015;

# scM&T-Seq: Single-Cell Methylome and Transcriptome Sequencing

scM&T-Seq allows parallel analysis of both epigenetic and gene expression patterns from single cells using Smart-seq2 and scBS-seq.[208] scM&T-Seq is built upon G&T-seq, but instead of using MDA for DNA sequencing, it uses scBS-Seq to determine DNA methylation patterns.

Single cells are isolated and individually lysed. The mRNAs are captured with streptavidin-coupled mRNA capture primers to separate them physically from the DNA strands. Smart-seq2 uses RT with template switching and tagmentation to generate cDNA libraries from the mRNA. DNA libraries are prepared via scBS-seq, which involves bisulfite conversion of DNA strands to identify methylated cytosines. Both libraries are ready for sequencing.



A schematic overview of scM&T-seq.

| Advantages | Disadvantages |
|---|---|
| • Investigates links between epigenetic and transcriptional heterogeneity in single cells<br>• Because DNA and RNA are physically separated and amplified independently, there is no need to mask coding sequences during analysis | • Smart-seq2 is not strand-specific and applicable to only poly(A)$^+$ RNA<br>• Does not distinguish between 5mC and 5hmC |

## Reviews

Clark S. J., Lee H. J., Smallwood S. A., Kelsey G. and Reik W. Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity. *Genome Biol.* 2016;17:72

Wen L. and Tang F. Single-cell sequencing in stem cell biology. *Genome Biol.* 2016;17:71

## References

Hu Y., Huang K., An Q., et al. Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol.* 2016;17:88

In this study, the authors developed a method that simultaneously profiles the methylome and the transcriptome of the same individual cell (scM&T-Seq). The method is very similar to scM&T-seq, except that scM&T-Seq uses single-cell reduced-representation bisulfite sequencing (scRRBS) for methylome analysis. The authors validated this method by simultaneously profiling the transcriptome and DNA methylome in individual sensory neurons of the dorsal root ganglion (DRG). Their data identified transcriptome and DNA methylome heterogeneity in DRG neurons. They also found that gene methylation and expression are positively correlated, but only for those genes that contain CpG island promoters.

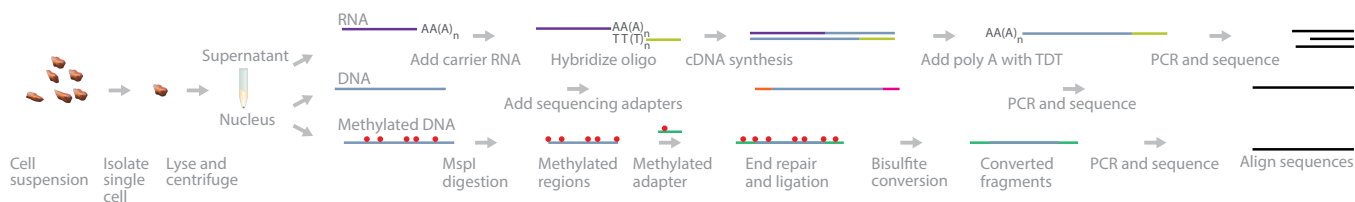Illumina Technology: HiSeq 2500 System, MiSeq System

## Associated Kits

Nextera XT DNA Library Prep Kit

---

208.  Angermueller C., Clark S. J., Lee H. J., et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods.* 2016;13:229-232

# scTrio-seq: Single-Cell Triple Omics Sequencing

scTrio-Seq can analyze genomic CNVs, the DNA methylome, and the transcriptome of an individual mammalian cell simultaneousl.[209] This approach is an extension of previous methods, such as scMT-seq.[210]



A schematic overview of scTrio-seq.

| Advantages | Disadvantages |
| --- | --- |
| • Accurately analyzes the mechanism by which the transcriptome, genome, and DNA methylome regulate each other<br>• CNVs can be identified reliably using scRRBS data | • Lower transcriptome coverage than scMT-seq[213]<br>• Results in 3'-biased transcriptome |

## Reviews

Bock C., Farlik M. and Sheffield N. C. Multi-Omics of Single Cells: Strategies and Applications. *Trends in Biotechnology.* 2016;34:605-608

Poirion O. B., Zhu X., Ching T. and Garmire L. Single-Cell Transcriptomics Bioinformatics and Computational Challenges. *Frontiers in Genetics.* 2016;7:163

Cheow L. F., Courtois E. T., Tan Y., et al. Single-cell multimodal profiling reveals cellular epigenetic heterogeneity. *Nat Meth.* 2016;13:833-836

Picelli S. Single-cell RNA-sequencing: The future of genome biology is now. *RNA Biology.* 2016;1-14

Qian M., Wang D. C., Chen H. and Cheng Y. Detection of single cell heterogeneity in cancer. *Seminars in Cell & Developmental Biology.*

## References

Hou Y., Guo H., Cao C., Li X., Hu B., et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res.* 2016;26:304-319

The authors used scTrio-Seq to analyze 25 single cells derived from a human hepatocellular carcinoma (HCC) tissue sample. They found 2 subpopulations distinct in DNA copy numbers, DNA methylation, and RNA expression levels. By comparing the differences between the 2 HCC subpopulations, the researchers discovered that the minor subpopulation I, which harbored more copy-gain CNVs, expressed more invasive cell markers and was more likely to evade immune surveillance.

Illumina Technology: HiSeq 2000 System, HiSeq 2500 System

## Associated Kits

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

Nextera Rapid Capture Exome/Custom Enrichment Kit

209. Hou Y., Guo H., Cao C., et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res.* 2016;26:304-319
210. Hu Y., Huang K., An Q., et al. Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol.* 2016;17:88

Scientific Publication Reviews can be accessed at
www.illumina.com/pubreviews

**DNA Sequencing Methods Collection**
This publication is a collection of next-generation sequencing (NGS) methods for DNA sequencing, compiled from the scientific literature. It is both a tribute to the creativity of the users and the versatility of the technology. We hope it will inspire researchers to use these methods or to develop new ones to address new scientific challenges.

A method refers to the processing steps between extracting the nucleic acids (sample preparation) and the addition of oligonucleotide adapters for sequencing (library preparation). With a few extra processing steps, a wide range of scientific questions can be addressed by this technology.
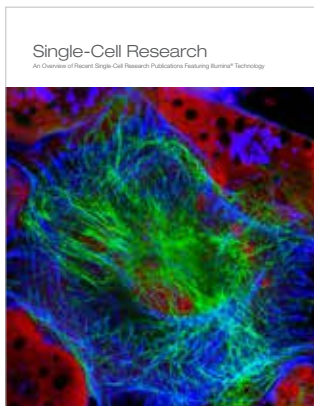
**Single-Cell Research Review**
Most of the impetus for single-cell tissue sequencing has come from cancer research, where cell lineage and the detection of residual disease are of paramount concern. The same approaches are being used to improve our understanding of massively complex biological systems, such as neural development and immunology.

This document highlights recent publications that demonstrate the use of Illumina technology for single-cell sequencing and very low input applications and techniques.

**Gene Editing Research Review**
CRISPR-Cas9 is a recently developed genome editing technique that allows scientists to perform precise genomic manipulation quickly and conveniently. This technology has a vast spectrum of applications. As any molecular biology technique, it is crucial that the obtained results have high levels of specificity. This review highlights recent publications that demonstrate the use of genomic technologies and high-throughput sequencing in CRISPR-Cas9 experiments for checking specificity and genome-wide off target effects.

This Scientific Publication Review is brought to you by Illumina, Inc.

illumına®