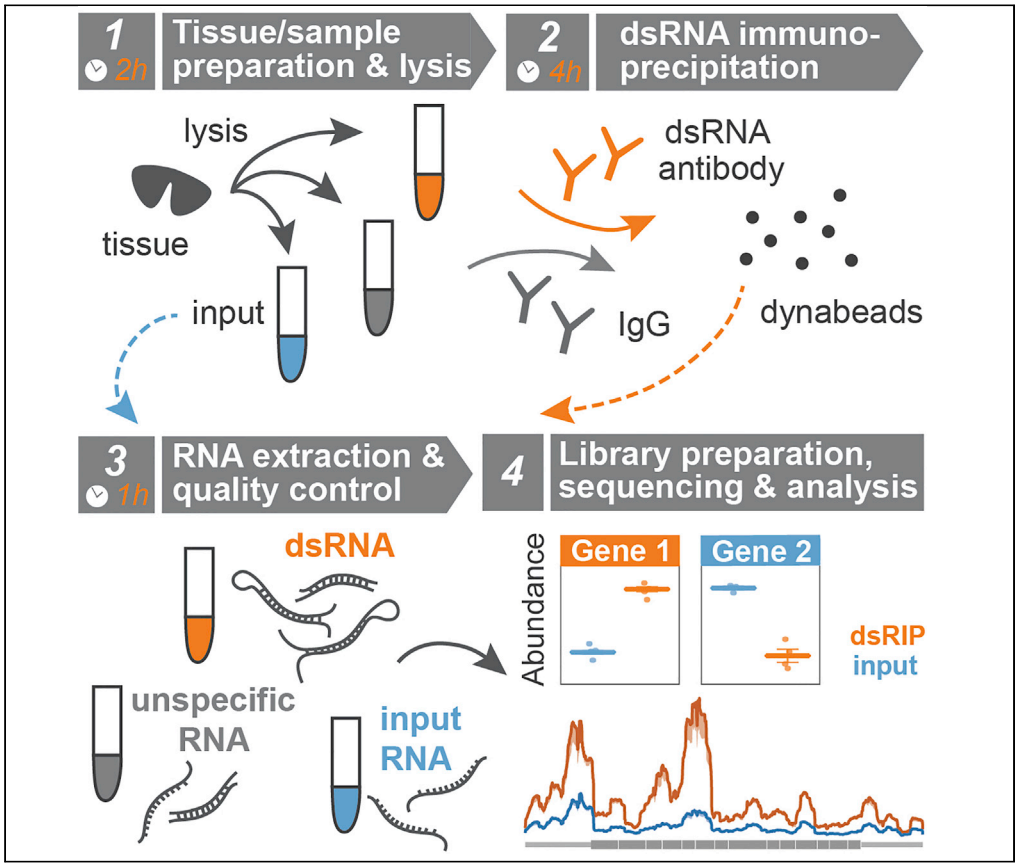
STAR Protocols

CellPress
OPEN ACCESS

Protocol

# Transcriptome-wide quantification of double-stranded RNAs in live mouse tissues by dsRIP-Seq




Double-stranded RNAs (dsRNAs) are abundantly present in cells, playing multiple regulatory functions. dsRNAs of viral origin activate innate immune responses. Since RNA editing and modifications affect the structure and recognition of RNAs, their alteration can result in the accumulation of aberrant endogenous dsRNAs inducing a deleterious innate immune response. Here, we present a complete protocol for the measurement of dsRNAs in a live mouse tissue using dsRIP-Seq. This protocol focuses on tissue isolation, dsRNA immunoprecipitation and downstream computational analysis.

Yimeng Gao, Shirui Chen, Stephanie Halene, Toma Tebaldi

stephanie.halene@yale.edu (S.H.)
toma.tebaldi@yale.edu (T.T.)

HIGHLIGHTS

Purification and sequencing of double-stranded RNAs (dsRNAs) in live tissues

Open-source computational framework for the quantification and comparison of dsRNAs

Gao et al., STAR Protocols 2, 100366
March 19, 2021 © 2021 The Author(s).
https://doi.org/10.1016/j.xpro.2021.100366

STAR Protocols

CellPress
OPEN ACCESS

Protocol

# Transcriptome-wide quantification of double-stranded RNAs in live mouse tissues by dsRIP-Seq

Yimeng Gao,[1,2,5] Shirui Chen,[3] Stephanie Halene,[1,2,6,*] and Toma Tebaldi[1,2,4,5,*]

[1]Section of Hematology, Yale Cancer Center and Department of Internal Medicine, Yale University School of Medicine, New Haven, CT 06520, USA

[2]Yale Stem Cell Center and Yale RNA Center, Yale University School of Medicine, New Haven, CT 06520, USA

[3]Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0032, Japan

[4]Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, Trento 38123, Italy

[5]Technical contact

[6]Lead contact

*Correspondence: stephanie.halene@yale.edu (S.H.), toma.tebaldi@yale.edu (T.T.)
https://doi.org/10.1016/j.xpro.2021.100366

## SUMMARY

**Double-stranded RNAs (dsRNAs) are abundantly present in cells, playing multiple regulatory functions. dsRNAs of viral origin activate innate immune responses. Since RNA editing and modifications affect the structure and recognition of RNAs, their alteration can result in the accumulation of aberrant endogenous dsRNAs inducing a deleterious innate immune response. Here, we present a complete protocol for the measurement of dsRNAs in a live mouse tissue using dsRNA immunoprecipitation and sequencing (dsRIP-Seq). This protocol focuses on tissue isolation, dsRNA immunoprecipitation and downstream computational analysis.**

**For complete details on the use and execution of this protocol, please refer to Gao et al. (2020).**

## BEFORE YOU BEGIN

### Timed breeding

Timing: 15–20 days

1. Timed breeding to obtain fetal liver.
   a. Set up mating between one male and two female mice (8–16 weeks old) per cage in the afternoon 15–20 days before experiment. Set up as many cages as needed to obtain desired number of fetal livers per genotype, taking into account breeding success for mouse strain and season.
   b. Check for vaginal plug each morning after mating has been set up. Move female mouse with plug to a separate cage. The morning a plug is detected is counted as day E0.5.
   c. On day E14.5 proceed with fetal liver isolation. Usually, one E14.5 fetal liver will yield approximately 30 μg RNA (~40–50 million nucleated cells), sufficient for a single experimental sample.

### Isolation of murine fetal livers

Timing: 2 h

2. Thoroughly clean the bench and dissection tools with 70% ethanol and a cleaning agent to remove RNases (e.g., RNase Away™).


This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

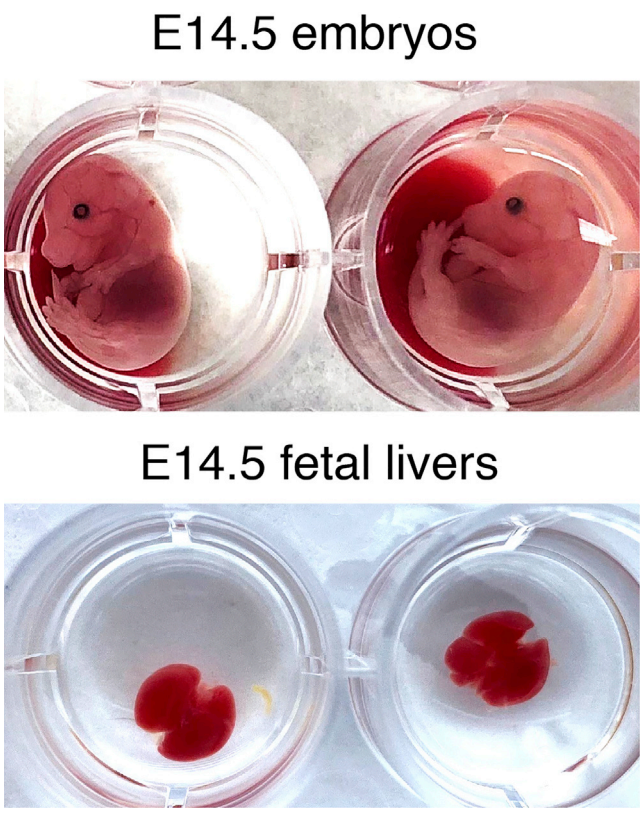



**Figure 1. Morphology of murine E14.5 embryos and fetal livers**

3. Prepare sterile petri dishes, dissection scissors, microdissection forceps (2) and tweezers (2, 4 1/4" extra fine point), 24-well plates, 1.5 mL Eppendorf tubes.
4. Isolate fetal livers from the pregnant mouse at E14.5.
   a. Prepare two to four matching 24-well plates per pregnant female and fill each well with 1 mL 1× PBS. Label each well with mouse identifier and fetal liver number. In general, expect 4–10 fetal livers per pregnant female.
   b. Sacrifice the pregnant mouse by isofluorane anesthesia and cervical dislocation per institutional animal care and use committee guidelines. Disinfect the abdominal wall with 70% ethanol.
   c. Using forceps, tent skin of the lower abdomen and cut up towards sternum to expose internal organs. Remove bilateral uterine horns and place in 10 cm diameter petri dish with 1× PBS to rinse maternal blood.
   d. Gently remove one fetal-placental unit from uteri at a time using extra-fine point tweezers and separate fetus from placenta and place fetus in a 24-well filled with 1× PBS. Place the 24-well plate on ice.
   e. Dissect the fetal livers from the fetuses. Gently hold the body of the fetus with one tweezer while gently pulling the fetal liver (recognizable by its dark red color) away from the body with the second tweezer. Place fetal liver in matching well of second 24-well plate, also on ice. Remove remnant connective tissue from fetal livers.
5. Store the fetal livers in the 24-well plate on ice for further experiments (Figure 1).

⚠ CRITICAL: If genotyping of each fetal liver is required, clean utensils with 10% bleach followed by 70% ethanol in-between each fetus and fetal liver prep. After isolation of fetal livers, remove a small piece of fetus from each corresponding well for genotyping following standard tissue lysis and DNA genotyping protocol.

## KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Antibodies | | |
| Normal mouse IgG | MilliporeSigma | Cat# 12-371; RRID: AB_145840 |
| Mouse Monoclonal Antibody SCICONS J2 | English & Scientific Consulting Kft. | Cat# 10010200; RRID: AB_2651015 |
| Chemicals, peptides, and recombinant proteins | | |
| Sodium chloride, 5M solution | AmericanBio | Cat# AB13198-01000 |
| Magnesium chloride, 1M solution | AmericanBio | Cat# AB09006-00100 |
| Tris-HCl, 1M solution, pH 7.4 | AmericanBio | Cat# AB14044-01000 |
| IGEPAL CA-630 | Sigma-Aldrich | Cat# I8896-50ML |
| cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail | Sigma-Aldrich | Cat# 11836170001 |
| SUPERase● In™ RNase Inhibitor | Thermo Fisher Scientific | Cat# AM2696 |
| Water, RNase free | AmericanBio | Cat# AB02128-00500 |
| Dynabeads™ Protein G for Immunoprecipitation | Thermo Fisher Scientific | Cat# 10004D |
| TRIzol™ Reagent | Thermo Fisher Scientific | Cat# 15596026 |
| Chloroform | J.T. Baker | Cat# 9180-01 |
| GlycoBlue™ Coprecipitant | Thermo Fisher Scientific | Cat# AM9515 |
| Critical commercial assays | | |
| RNeasy Plus Mini Kit | QIAGEN | Cat# 74134 |
| KAPA RNA HyperPrep Kit with RiboErase | Roche | Cat# KK8560 |
| Deposited data | | |
| Mouse genome and transcript annotation | Gencode M15 | https://www.gencodegenes.org/mouse/release_M15.html |
| Transcript sequences | Ensembl 90 | https://useast.ensembl.org/info/website/archives/index.html |
| Experimental models: organisms/strains | | |
| Mouse: Vav-Cre; Mettl3$^{fl/fl}$ | (Gao et al., 2020) | N/A |
| Mouse: C57BL/6J | In house colony | JAX: 000664 |
| Software and algorithms | | |
| STAR v2.3.3a | N/A | https://github.com/alexdobin/STAR |
| edgeR | Bioconductor | https://bioconductor.org/packages/release/bioc/html/edgeR.html |
| FastUniq v1.1 | N/A | https://sourceforge.net/projects/fastuniq/ |
| RNAfold v2.4.11 | ViennaRNA package | https://www.tbi.univie.ac.at/RNA/ |
| FastQC v0.11.5 | N/A | https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ |
| Other | | |
| TapeStation 2200 | Agilent Technologies | Cat# G2964AA |
| HiSeq 2500 Sequencing System | Illumina | Cat# SY–401–2501 |

## MATERIALS AND EQUIPMENT

**dsRIP lysis buffer**

| Reagent | Final concentration | Volume (100 mL) |
|---|---|---|
| NaCl | 100 mM | 2 mL |
| Tirs-HCl pH 7.4 | 50 mM | 5 mL |
| $MgCl_2$ | 3 mM | 0.3 mL |
| IGEPAL CA-630 | 0.5% | 0.5 mL |
| $H_2O$, RNase-free | n/a | 92.2 mL |
| **Total** | **n/a** | **100 mL** |

On the day of the experiment, to 10 mL dsRIP lysis buffer add 0.4 mL of 25× Complete Proteinase Inhibitor and 20 μL SUPERase• In™ RNase Inhibitor for immediate use; maintain at 4°C.

⚠ CRITICAL: The dsRIP lysis buffer can be stored at 4°C for one month. Make fresh lysis buffer if longer time period elapses between experiments.

## STEP-BY-STEP METHOD DETAILS

### Double-stranded RNA immunoprecipitation (dsRIP) from fetal livers

Timing: 4 h

This protocol allows for the isolation of double-stranded RNAs (dsRNAs) from tissues, such as murine fetal livers. It is predicted to work with other tissues, such as bone marrow, and cell pellets of cell lines or primary cells. Cell numbers equivalent to ~30 μg RNA should be used. Lysates are incubated with the mouse monoclonal anti-dsRNA antibody J2 followed by isolation with magnetic beads (Protein G Dynabeads™) (Blango and Bass, 2016; Lybecker et al., 2014; Son et al., 2015).

1. Set aside approximately one fifth of each fetal liver for preparation of total RNA as input control for dsRNA analysis. The total RNA can be extracted per standard protocol of the QIAGEN RNeasy Plus Mini Kit (Cat# 74134) or following the standard TRIzol isolation procedure.
2. Place the remainder of each fetal liver (~30 million nucleated fetal liver cells) into a 1.5 mL Eppendorf tube with 600 μL dsRIP lysis buffer. Vortex and incubate on ice for exactly 5 min. If the tissues are firm and hard to dissociate and lyse by lysis buffer, they should be minced prior to lysis into smaller fragments.
3. Spin down cell lysates in a tabletop centrifuge at maximum speed (13,000 × *g*) for 10 min at 4°C.
4. Carefully pipet equal volumes of the supernatant from each fetal liver into two new 1.5 mL tubes placed on ice. Add 5 μL anti-dsRNA J2 antibody (1 μg/μL) to one tube and 5 μg mouse IgG to the second tube and incubate at 4°C for 2 h with rotation. During this step, dsRNA will be recognized and bound by J2 antibody while non-specific binding is expected to occur with mouse IgG (Schonborn et al., 1991). The mouse monoclonal anti-dsRNA J2 antibody may be replaced by other antibodies of interest, but readers should ensure specificity and that appropriate Dynabeads are used, according to the species antibodies are raised in.
5. Towards the end of the incubation period of cell lysates with anti-dsRNA J2 antibody, prepare 25 μL Protein G Dynabeads for each sample.
   a. Aliquot 25 μL Dynabeads Protein G each into two 1.5 mL tubes for each fetal liver and add 500 μL dsRIP lysis buffer and vortex.
   b. Place the tubes on a magnetic rack and allow beads to collect on side wall of Eppendorf tube for 2 min.
   c. Carefully pipette off supernatant without dislodging beads while tube remains on the magnetic rack. Repeat the wash step × 1.
   d. Discard the supernatant and keep the beads for the next step.
6. At the end of the two-hour lysate/antibody incubation period, transfer lysates to the tubes with the Protein G Dynabeads and incubate at 4°C for 1 h with rotation.
7. At the end of the incubation period, place the tubes on the magnetic rack and allow beads to collect on side wall of tubes, remove supernatant and wash beads as under 5b–5d 4 times with 500 μL dsRIP lysis buffer. This step should be performed in a cold room.
8. Place the tubes with washed beads on ice for RNA extraction.

### Extraction of immunoprecipitated dsRNAs

Timing: 1 h

This protocol allows for isolation of J2-immunoprecipitated dsRNAs from fetal livers.

9. Add 1 mL TRIzol to each tube with Protein G Dynabeads™ and mix well by pipetting up and down.
10. Add 200 μL chloroform to the TRIzol solution, mix well and incubate on ice for 5 min.
11. Spin at 13,000 × *g* for 15 min at 4°C.
12. Carefully remove aqueous phase (530 μL) without disrupting protein layer or disturbing phenol phase. Add 200 μL chloroform and incubate on ice for 5 min.

*Note:* The second round of incubation with chloroform serves to reduce contamination with phenol, which is a common problem in the extraction of low amount of RNAs inhibiting downstream reactions (Toni et al., 2018).

13. Spin at 4°C and 13,000 × *g* for 15 min.
14. Transfer 480 μL of the upper aqueous phase as above into equal volume (480 μL) isopropanol. Add 1 μL GlycoBlue or RNAse-free glycogen of choice, invert by hand 10 times to mix and incubate at −20°C for 20 min.
15. Spin down at 4°C and 13,000 × *g* for 10 min to precipitate RNA.

*Note:* Addition of glycogen serves to make the RNA pellet visible following precipitation and to prevent loss of RNA during the next wash and precipitation steps.

16. Pipet off and discard the supernatant and add 800 μL 75% ethanol in nuclease-free water to the RNA pellet.
17. Centrifuge at 4°C and 8000 × *g* for 5 min.
18. Pipet off and discard the supernatant.
19. Pulse spin samples at ~23°C.
20. Carefully remove the remainder of the supernatant without disturbing the RNA pellet.
21. Air-dry the pellet with the lid open at ~23°C for 5 min.
22. Add 15–20 μL RNase-free water to the pellet and dissolve dsRNAs.
23. Quantify RNA using a Nanodrop or other instrument to determine concentration and purity. Usually, one sample can produce around 2 ng/μL dsRNA corresponding to ~30 ng in total. Clear absorbance peaks at 260 nm and 280 nm should be seen.
24. Perform quality control analysis on a TapeStation System using High Sensitivity RNA ScreenTape (Figure 2).

**Pause point:** The extracted RNA can be stored at −80°C for use in subsequent steps.

### Library preparation and next-generation sequencing

**Timing: 2–5 days**

25. At this step, submit both INPUT and experimental (J2-dsRIP ± IgG-dsRIP) samples for library generation. Note that IgG immunoprecipitation may not yield RNA sufficient for library preparation (in our hands, IgG controls yielded concentrations lower than 0.1 ng/μL and less than 2 ng in total, insufficient for sequencing). In this case only INPUT and J2-dsRNA samples should be processed further.
26. Ribosomal RNA (rRNA) depletion and library construction are performed following the instructions of the KAPA RNA HyperPrep Kit with RiboErase.
27. Samples are ready for high-throughput sequencing on Illumina or other platforms. Generally, 60 million reads/sample should be sequenced to yield good results.

rRNA depletion instead of mRNA selection is important for the dsRIP-seq to capture all possible dsRNAs that may or may not have poly-A tails, such as circular RNA (circRNA), for sequencing.

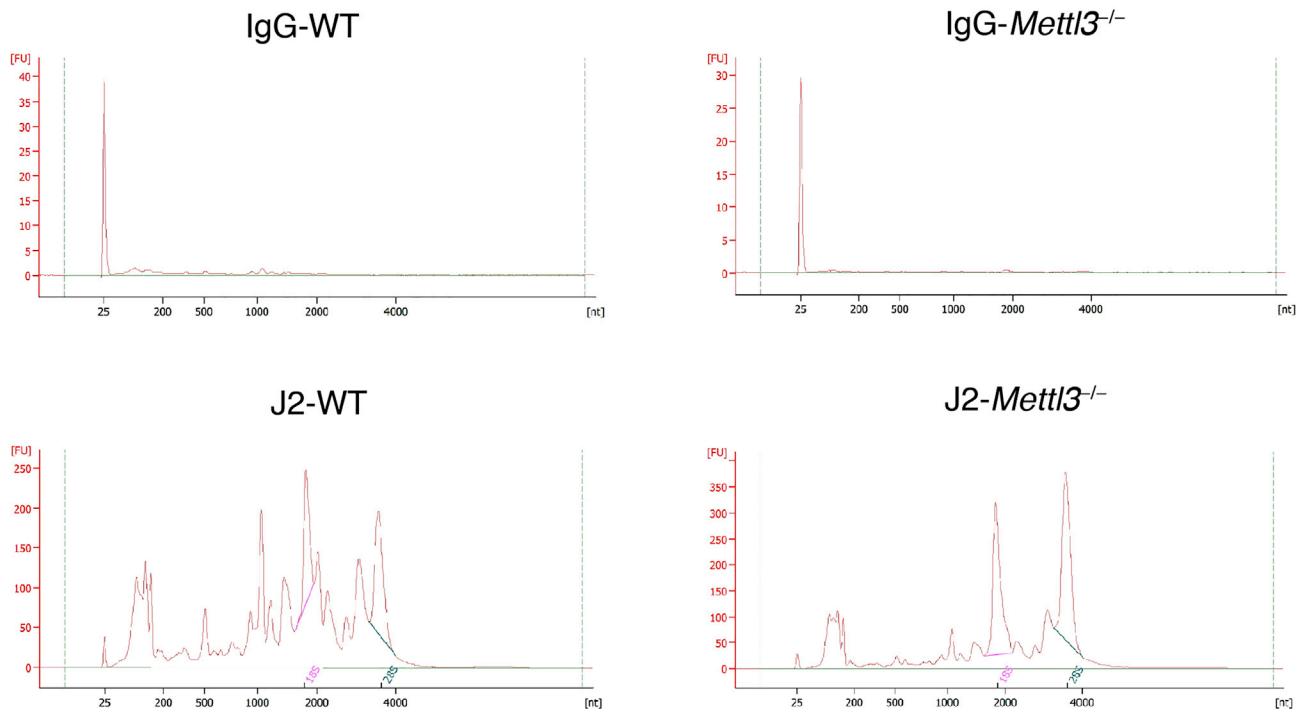



**Figure 2. Examples of RNA profiles obtained by dsRIP on the TapeStation System**

## EXPECTED OUTCOMES

The results of your TapeStation quality control analysis, performed prior to rRNA depletion and library preparation, will likely depend on your tissues and experimental conditions. We observed a random distribution of RNA ranging from 50 nucleotides (nt) to 4000 nt length on the TapeStation system with a peak around 200 nt in the wildtype samples. We observed sharp peaks corresponding to the sizes of the 18S and 28S rRNAs (Figure 2).

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Pre-processing of next-generation sequencing data

This section lists all the computational steps necessary to obtain gene or transcript specific read counts starting from sequencing raw data (FASTQ). In the simplest experimental design, dsRIP and INPUT libraries with ideally 3 or more biological replicates are expected. In our worked example, we sequenced libraries with Illumina, 2×100 paired end sequencing, obtaining approximately 60 million of read pairs for each sample.

28. Perform quality control of FASTQ files with FastQC or similar tools. QC metrics should resemble those obtained by RNA-seq of the same tissue, in our worked example we only observed a higher percentage of duplicated reads (Table 1).
29. Optionally, remove duplicated reads from FASTQ files with FastUniq (Xu et al., 2012) or similar tools. In our worked example we decided to perform this step, prompted by the relatively high amount of duplicates. To be sure to prevent the removal of natural duplicates, UMIs (Unique Molecular Identifiers) should be used (Kivioja et al., 2011) UMIs, or molecular barcodes, are standard in almost all single-cell RNA-seq and CLIP-seq protocols.
30. Align reads to a reference genome with a splice-aware aligner. In our worked example, reads were aligned to the mouse genome, assembly GRCm38.p5, with STAR version 2.5.3a (Dobin et al., 2013) and default parameters (see https://github.com/alexdobin/STAR for a full tutorial on this aligner).

**Table 1. Sequencing statistics in our worked example**

| Sample | Sequenced read pairs | After deduplication | Duplication % |
|---|---|---|---|
| dsRIP_ctrl_rep1 | 55162796 | 15529442 | 72% |
| dsRIP_ctrl_rep2 | 65336514 | 15878260 | 76% |
| dsRIP_trt_rep1 | 60185224 | 15682440 | 74% |
| dsRIP_trt_rep2 | 48709914 | 13363489 | 73% |
| INPUT_ctrl_rep1 | 65554565 | 12143560 | 81% |
| INPUT_ctrl_rep2 | 67722509 | 13159257 | 81% |
| INPUT_trt_rep1 | 66321387 | 13885129 | 79% |
| INPUT_trt_rep2 | 53744307 | 12558777 | 77% |

31. Quantify read counts per gene or per transcript from BAM files. In our worked example, gene-specific read counts were obtained running STAR with the –quantMode GeneCounts option. We used the Gencode M15 transcript annotation as transcriptome guide. Alternative tools can be used, such as FeatureCounts or HTSeq.
32. Load the table of read counts, with each column associated with a sample and each row associated with a gene or transcript, in the R software environment. Filter out genes or transcripts with low or absent signals. In our worked example, we selected genes with at least 10 raw reads in all the replicates of at least 1 condition (dsRIP or INPUT). This filter selected 13779 out of 52640 annotated genes. The specific R code used for the worked example is available and enables the reproduction of the analyses shown in this protocol: https://github.com/tomateba/dsRIP-Seq.
33. Normalize counts across samples. In our worked example, we used the TMM method implemented in the edgeR Bioconductor package (McCarthy et al., 2012) (see https://www.bioconductor.org/packages/release/bioc/html/edgeR.html for a full tutorial on edgeR). Alternatively, similar packages such as DESeq2 can be used for the normalization and for the identification of dsRNA (see https://bioconductor.org/packages/release/bioc/html/DESeq2.html for a tutorial on DESeq2).

### Identification of dsRNAs by comparing dsRIP and INPUT normalized signals

In the simplest experimental design, not involving differential conformation analysis, genes or transcripts that are more likely to be in dsRNA conformation can be identified by comparing dsRIP and INPUT signals (Figure 3).

34. Identify dsRIP enriched genes or transcripts. In our worked example, we used generalized linear models implemented in edgeR (*glmQLFTest* function). The contrast used for this analysis is: $(dsRIP - INPUT)$. After applying this contrast, select genes or transcripts in dsRNA conformation applying a significance threshold ($p-value<0.05$) and a fold enrichment threshold ($\log_2 dsRIP/INPUT)>0$.

### Differential analysis of dsRNAs comparing two conditions

In this experimental design, the aim of the analysis is to identify changes in the dsRNA conformation of genes or transcripts by comparing two conditions. In the general case, these conditions can be named control (ctrl) and treatment (trt). Genes or transcripts with enriched dsRNA conformation in the treatment condition can be identified applying 3 different criteria, explained below (Figure 3).

35. Identify dsRIP enriched genes or transcripts, considering both control and treatment conditions. This criterion selects genes or transcripts that are more likely to be in a dsRNA conformation. In our worked example, we used generalized linear models implemented in edgeR (glmQLFTest function). The contrast used for this criterion is:

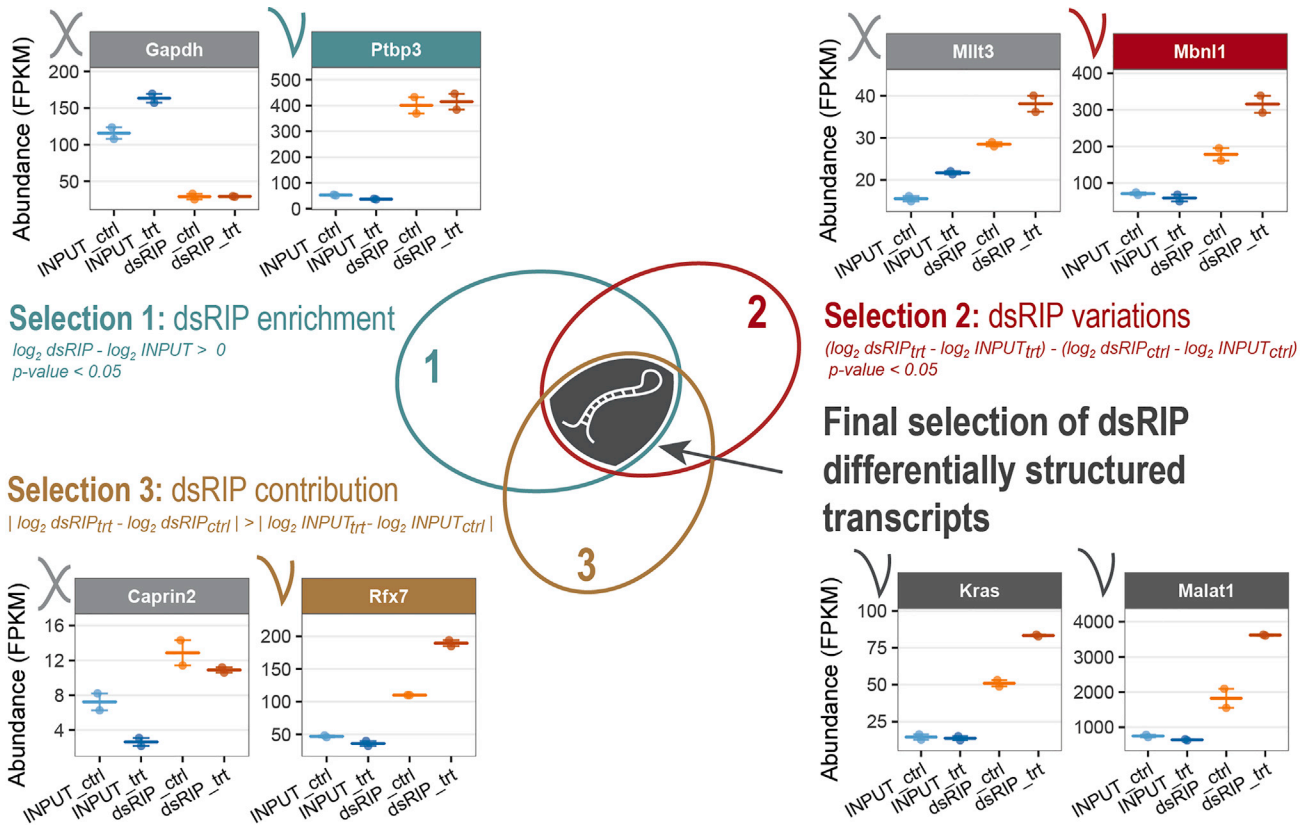



**Figure 3. Computational analysis to identify and compare dsRNAs, with examples**

$(dsRIP_{ctrl} + dsRIP_{trt}) - (INPUT_{ctrl} + INPUT_{trt})$ After applying this contrast, select genes or transcripts in dsRNA conformation applying a significance threshold ($p - value$<0.05) and a fold enrichment threshold ($\log_2 dsRIP/INPUT$)>0.

36. Identify genes or transcripts showing significant changes in dsRNA conformation comparing treatment and control condition, normalizing for INPUT. This criterion selects transcripts that are specifically altered in the dsRIP signal of the treatment versus control condition, normalizing for variations in the INPUT, that could be due to expression changes. Using generalized linear models implemented in edgeR (glmQLFTest function), the contrast used for this criterion is:

$(dsRIP_{trt} - dsRIP_{ctrl}) - (INPUT_{trt} - INPUT_{ctrl})$, that is also equivalent to the contrast:

$(dsRIP_{trt} - INPUT_{trt}) - (dsRIP_{ctrl} - INPUT_{ctrl})$

After applying this contrast, select genes or transcripts applying a significance threshold ($p - value$<0.05) and a fold change threshold, for example $(\log_2 dsRIP_{trt} - \log_2 INPUT_{trt}) > (\log_2 dsRIP_{ctrl} - \log_2 INPUT_{ctrl})$ to look for increased dsRNA conformation in the treatment condition, or $(\log_2 dsRIP_{trt} - \log_2 INPUT_{trt}) < (\log_2 dsRIP_{ctrl} - \log_2 INPUT_{ctrl})$ to look for decreased dsRNA conformation in the treatment condition.

37. Identify genes or transcripts with major contribution of dsRIP changes over INPUT changes. This criterion selects genes or transcripts for which the change in dsRIP levels (treatment vs control) is larger, in absolute values, than the change in INPUT levels. This criterion is important to exclude genes that, for example, decrease in both dsRIP and INPUT, but more in the latter. These genes

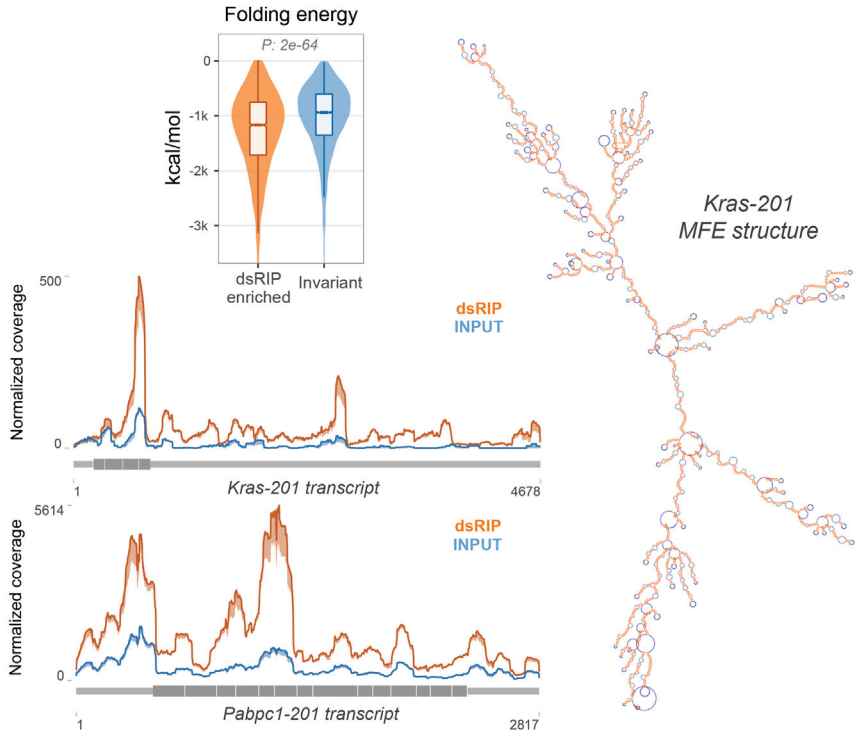



**Figure 4. Examples of downstream analysis of the identified dsRNAs**

could result enriched in the analysis described at point 34, just because their decrease in dsRIP is lower than their decrease in INPUT. Yet for these genes variations in expression levels are larger than variations in conformations, therefore they should be considered with caution or excluded from the analysis. Considering average expression values among replicates, select genes or transcripts with $|\log_2 dsRIP_{trt} - \log_2 dsRIP_{ctrl}| > |\log_2 INPUT_{trt} - \log_2 INPUT_{ctrl}|$. In our worked example, this criterion was not associated with an edgeR contrast, therefore it is not dependent on a p value.

## Downstream structural and functional analysis of dsRNAs

The populations of relevant genes or transcripts with dsRNA conformation identified by dsRIP-Seq can be further characterized with structural of functional computational analysis tools (Figure 4).

38. Predict the secondary structure of the identified genes or transcripts. In our worked example, we used the RNAfold algorithm contained in the ViennaRNA package (v2.4.11) (Lorenz et al., 2011) (see https://www.tbi.univie.ac.at/RNA/RNAfold.1.html for a tutorial on RNAfold). This algorithm requires as input a fasta file of transcript sequences and enables the prediction of the minimum free energy (MFE) structure, with the corresponding free folding energy, and also more advanced features such as thermodynamic ensemble prediction and the ensemble diversity, describing the flexibility of the expected RNA structure by measuring the diversity of the ensemble structures. In our worked example, we executed RNAfold starting with the fasta file of whole transcript sequences with the following code: *RNAfold –infile=input.fasta –MEA -d2 -p*. Numerical features such as folding free energy and ensemble diversity can be used to compare the population of dsRIP enriched vs non-enriched transcripts. Predicted secondary structures for each transcript, in the commonly used dot-bracked notation, can be displayed with multiple online tools such as forna (http://rna.tbi.univie.ac.at/forna/) or further dissected to identify long dsRNA regions and structural motifs.

## LIMITATIONS

The approach presented here is based on the J2 antibody to specifically recognize dsRNAs and perform the immunoprecipitation. Since it has been reported that J2 antibody mainly recognizes dsRNAs longer than 40 bp, dsRNAs which are shorter than 40 bp may be missed in this analysis. Besides fetal livers, other tissues and cells can also be used for the dsRIP-Seq, as long as they can provide enough RNA for the immunoprecipitation. As far as we have tested, a E14.5 fetal liver can yield more than 20 μg RNA. Therefore, any fresh tissue or cultured cell lines that can provide similar RNA quantities, can be considered as a feasible starting material. While the dsRIP-Seq approach enables the identification of RNAs enriched in double stranded conformation, the specific location of dsRNA regions within long transcripts cannot be directly sequenced, as the RNA fragmentation step is performed after RNA immunoprecipitation. RNA fragmentation or partial digestion before immunoprecipitation would theoretically enrich the library with specific dsRNA regions, but this approach would require more starting material, a more complex experimental protocol and computational workflows based on peak callers such as those used in the context of identifying RNA-protein interactions (CLIP-seq) or $m^6A$ enrichment sites ($m^6A$-RIP).

## TROUBLESHOOTING

### Problem

Uncertainty of fetal liver timing.

### Potential solution

It is best to check for vaginal plugs early in the morning, especially in multigravida mice as it can fall out. If a plug is not obvious, a blunt surgical probe can be used to detect a plug that is still situated deeper in the female mouse's vagina.

### Problem

Insufficient RNA is isolated after dsRNA immunoprecipitation.

### Potential solution

Sterile technique and RNAse free surfaces, tools, and reagents are key. Incubation steps should be exactly timed. Avoid unnecessary delays between steps. Long-term storage of RNA should be avoided prior to sequencing. A positive control to confirm the dsRNA immunoprecipitation process is1 μg polyinosinic:polycytidylic acid (pI:pC). 1 μg can be used to monitor and confirm the efficiency of each step, as pI:pC is a well-characterized dsRNA which can be recognized by J2 antibody.

### Problem

Contamination of phenol after RNA extraction identified by an unexpected peak at 270 nm during TapeStation system quality control analysis of isolated RNAs.

### Potential solution

RNA may be contaminated with Phenol, inhibiting downstream steps. RNA resuspended in 20 μL RNAse free water can be re-extracted by adding 1 mL TRIzol followed by careful execution of steps 10–20. Especially steps 16–20 serve to eliminate all remnant phenol contamination. With a tweezer manipulate a clean Kimwipes to wipe the side of the Eppendorf tube after each chloroform step to further eliminate phenol. Be careful not to touch the RNA pellet visualized by addition of GlycoBlue. If a second clean-up step fails, the experiment should be repeated from the beginning starting with new tissue to obtain high-quality RNA for downstream sequencing.

### Problem

Interpretation of TapeStation quality control results and choice of dsRNA antibody.

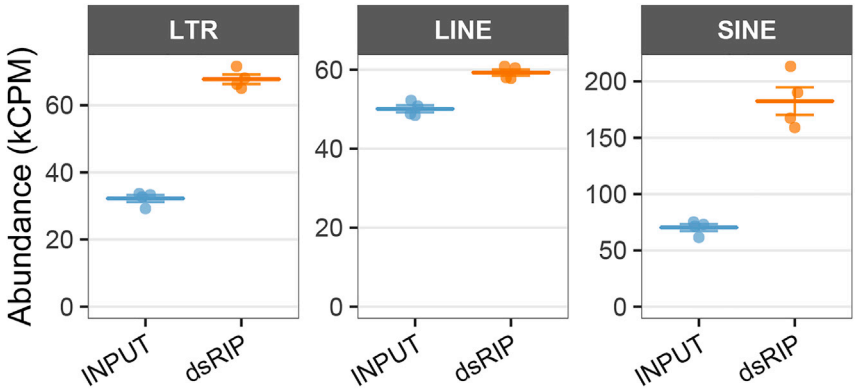



**Figure 5. Example of analysis and dsRNA quantification of transposable elements**

### Potential solution

If performed for the first time and under experimental conditions where expected results are unknown, it is essential to strictly adhere to RNAse free conditions, ensure high quality of reagents and buffers and exact documentation of experimental steps. It is advised that samples from control and experimental condition (such as wildtype versus knockout) are always performed in parallel. At least 3 biologic replicates should be performed for each condition. The quality of the anti-dsRNA J2 or other antibodies are essential for the specific immunoprecipitation of dsRNA. The provided information for the J2 antibody reflects the one with best performance in our hands. At the time of receipt, researchers should resuspend according to manufacturer instruction and generate multiple aliquots to avoid multiple freeze-thaw cycles. J2 antibodies produced by other companies or provided in different forms may yield differing results. Although other dsRNA antibodies may produce similar results to the ones with J2 antibody, we suggest researchers start with J2 antibody since it is the most widely used and well-characterized anti-dsRNA antibody.

### Problem

Analysis of transposable elements

### Potential solution

Transposable elements (TEs) comprise approximately half of the mammalian genomes. If the analysis of repetitive sequences transcribed from transposable elements such as retrotransposons is an important aspect of the experiment, the dsRIP protocol may require some modifications, both in the experimental and the computational parts. First, unique molecular identifiers (UMIs) should be ideally included in library preparation to discriminate between technical duplicates (due to PCR) and natural duplicates, such as repetitive elements (Kivioja et al., 2011). Libraries should be sequenced with a paired end approach, as in our worked example, trying to maximize read length. Sequenced reads coming from repeated elements also introduce ambiguities in the mapping step. Therefore, mapping parameters can be adjusted to allow higher number of multi-mapped reads. Finally, the quantification of transposable elements can be performed with dedicated tools, recently reviewed in (Teissandier et al. 2019). In our worked example, the analysis of Transposable Elements, including their classification in classes and families, was based on GTF files curated by the Hammell Lab: (http://labshare.cshl.edu/shares/mhammelllab/www-data/TEtranscripts/TE_GTF/mm10_rmsk_TE.gtf.gz). Reads mapped on transposable elements were identified and quantified with the Rsubread Bioconductor package. Our results suggest that some retrotransposon classes are enriched in dsRNA conformation (Figure 5)

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Stephanie Halene (stephanie.halene@yale.edu).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

This study did not generate any unique datasets or code. Raw and analyzed sequencing data for the worked example were previously deposited under GEO: GSE148882.

Software used for the analyses are described and referenced in the quantification and statistical analysis subsections and are listed in the key resources table. The specific R code used for the analysis of the worked example are available at https://github.com/tomateba/dsRIP-Seq.

## ACKNOWLEDGMENTS

This study was supported in part by the State of Connecticut under the Regenerative Medicine Research Fund (to S.H.; its contents are solely the responsibility of the authors and do not necessarily represent the official views of the State of Connecticut or Connecticut Innovations), the NIH/NIDDK R01DK102792 (to S.H.), and The Frederick A. Deluca Foundation (to S.H.); Y.G. was supported by the James Hudson Brown-Alexander Brown Coxe Postdoctoral Fellowships. T.T. was supported by a pilot grant from the Yale Cooperative Center of Excellence in Hematology (YCCEH) (NIDDK U54DK106857) and by AIRC under MFAG 2020 - ID. 24883 project.

## AUTHOR CONTRIBUTIONS

Conceptualization, Y.G., T.T., and S.H.; experimental methodology, Y.G. and S.C.; computational methodology, T.T.; writing, Y.G., T.T., and S.H.; funding acquisition and supervision, S.H.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## REFERENCES

Blango, M.G., and Bass, B.L. (2016). Identification of the long, edited dsRNAome of LPS-stimulated immune cells. Genome Res. *26*, 852–862.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Gao, Y., Vasic, R., Song, Y., Teng, R., Liu, C., Gbyli, R., Biancon, G., Nelakanti, R., Lobben, K., Kudo, E., et al. (2020). m(6)A modification prevents formation of endogenous double-stranded RNAs and deleterious innate immune responses during hematopoietic development. Immunity *52*, 1007–1021.e8.

Kivioja, T., Vähärautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S., and Taipale, J. (2011). Counting absolute numbers of molecules using unique molecular identifiers. Nat. Methods *9* (1), 72–74.

Lorenz, R., Bernhart, S.H., Honer Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms Mol. Biol. *6*, 26.

Lybecker, M., Zimmermann, B., Bilusic, I., Tukhtubaeva, N., and Schroeder, R. (2014). The double-stranded transcriptome of Escherichia coli. Proc. Natl. Acad. Sci. U S A *111*, 3134–3139.

McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. *40*, 4288–4297.

Schonborn, J., Oberstrass, J., Breyel, E., Tittgen, J., Schumacher, J., and Lukacs, N. (1991). Monoclonal-antibodies to double-stranded-Rna as probes of Rna structure in crude nucleic-acid extracts. Nucleic Acids Res. *19*, 2993–3000.

Son, K.N., Liang, Z., and Lipton, H.L. (2015). Double-stranded RNA is detected by immunofluorescence analysis in RNA and DNA virus infections, including those by negative-stranded RNA viruses. J. Virol. *89*, 9383–9392.

Teissandier, A., Servant, N., Barillot, E., and Bourc'his, D. (2019). Tools and best practices for retrotransposon analysis using high-throughput sequencing data. Mob. DNA *10*, 52.

Toni, L.S., Garcia, A.M., Jeffrey, D.A., Jiang, X., Stauffer, B.L., Miyamoto, S.D., and Sucharov, C.C. (2018). Optimization of phenol-chloroform RNA extraction. MethodsX *5*, 599–608.

Xu, H., Luo, X., Qian, J., Pang, X., Song, J., Qian, G., Chen, J., and Chen, S. (2012). FastUniq: a fast de novo duplicates removal tool for paired short reads. PLoS One *7*, e52249.