PRO-seq: Precise Mapping of Engaged RNA Pol II at Single-Nucleotide Resolution

Claudia A. Mimoso[1,2,3] and Seth R. Goldman[1,2,3]

[1]Department of Biological Chemistry and Molecular Pharmacology, Blavatnik Institute, Harvard Medical School, Boston, Massachusetts, USA
[2]These authors contributed equally to this work.
[3]Corresponding authors: *cmimoso@g.harvard.edu*; *Seth_Goldman@hms.harvard.edu*

Published in the Molecular Biology section

Gene regulation is dependent on the production of mRNAs and a repertoire of non-coding RNAs by RNA polymerase II (RNAPII). Precision run-on sequencing (PRO-seq) maps the position of engaged RNAPII complexes at single-nucleotide resolution and can reveal direct targets of regulation, locations of enhancers, and transcription mechanisms that are difficult or impossible to measure by analysis of total cellular RNA. Briefly, this method first involves permeabilizing cells with mild detergents to remove intracellular NTPs and halt transcription. Transcription is then resumed in the presence of biotin-NTPs and sarkosyl to allow transcriptional incorporation of a single biotinylated NTP by RNAPII. The biotin moiety is then bound to streptavidin beads to stringently enrich for nascent RNAs. Sequencing libraries are then generated such that the first base read corresponds to the 3′ end of the nascent transcript. Here, we describe our current protocol for generating PRO-seq libraries from metazoan cells, including adaptations of previously published protocols to incorporate unique molecular identifiers, reduce ligation bias, and improve library yields. Additional commentary describes quality control and processing of PRO-seq data and references for more advanced downstream analysis such as gene and enhancer identification.

Basic Protocol 1: Cell permeabilization for PRO-seq
Basic Protocol 2: Construction of PRO-seq libraries
Support Protocol: Adenylation of 3′ adapter

Keywords: gene expression • nascent RNA • PRO-seq • RNA polymerase II • transcription

How to cite this article:
Mimoso, C. A., & Goldman, S. R. (2023). PRO-seq: Precise mapping of engaged RNA Pol II at single-nucleotide resolution. *Current Protocols*, *3,* e961. doi: 10.1002/cpz1.961

INTRODUCTION

RNA mediates diverse cellular functions ranging from directing protein synthesis (mRNA) and synthesizing proteins (rRNA) to a vast array of regulatory functions (non-coding RNAs, e.g., small nuclear RNAs, long non-coding RNAs, micro RNAs, etc.) (Dunham et al., 2012; Kaikkonen & Adelman, 2018; Morris & Mattick, 2014). RNA is transcribed from DNA templates by RNA polymerases (RNAPs), a family of highly conserved enzymes found in all domains of life. Eukaryotes possess three RNA polymerases: RNAPI, which synthesizes rRNA; RNAPII, which primarily synthesizes mRNA; and

Current Protocols e961, Volume 3
Published in Wiley Online Library (wileyonlinelibrary.com).
doi: 10.1002/cpz1.961

RNAPIII, which primarily synthesizes tRNA (Roeder, 2019). Dysregulation of transcription can be fatal or contribute to diseases such as cancer (Anastasiadou et al., 2017; Bradner et al., 2017; Dhanasekaran et al., 2022; Futreal et al., 2004; Lee & Young, 2013; Sur & Taipale, 2016; van der Lee et al., 2020; Vervoort et al., 2021; Vogt, 2001). Thus, each step of transcription and gene expression must be tightly controlled. RNAP activity is modulated by numerous transcription factors, regulatory proteins, and modular subunits (Core & Adelman, 2019; Jonkers et al., 2014; Malik & Roeder, 2023). Dissecting the molecular mechanisms of transcription can reveal a wealth of information about gene expression, gene regulation, and cellular physiology.

Analysis of total cellular RNA is a powerful tool to understand gene regulation, but steady-state measurements are insufficient to reveal the direct effects of cellular perturbations on transcription and are poorly suited to analyze short, unstable RNAs such as enhancer RNAs (eRNAs) and long non-coding RNAs (lncRNAs). In response to these limitations, a number of methods have been developed to analyze nascent RNAs, including methods that metabolically label newly synthesized RNAs (e.g., SLAM-seq, TT-seq) and methods that enrich for RNAPII-associated transcripts (e.g., GRO-seq, PRO-seq, mNET-seq) (Core et al., 2008; Herzog et al., 2017; Kwak et al., 2013; Mahat et al., 2016; Nojima et al., 2015; Schwalb et al., 2016). Metabolic labeling methods, in which cells are incubated with a ribonucleotide analog that is incorporated into RNA by RNAPII, have relatively low spatial resolution and cannot differentiate true nascent transcripts from those recently synthesized. In contrast, methods that enrich for RNAPII-associated transcripts can identify nascent transcripts and track the position of RNAPII at single-nucleotide resolution. In particular, the single-nucleotide resolution afforded by methods that enrich for actively engaged RNAPII (e.g., PRO-seq) offers mechanistic insights into transcription processes such as polymerase pausing and elongation.

The laboratories of John Lis and Leighton Core have been at the forefront of developing nascent RNA analysis methods, and precision run-on sequencing (PRO-seq) is their refinement of the earlier GRO-seq method (Core & Adelman, 2019; Core et al., 2008; Kwak et al., 2013; Mahat et al., 2016). PRO-seq is based on an in vitro run-on transcription assay in which engaged RNAPII complexes incorporate a single biotinylated nucleotide analog at the 3′ end of nascent RNA chains. The incorporation of biotinylated NTPs is enabled by the presence of Sarkosyl in the run-on reaction. Sarkosyl removes negative elongation factors from RNAPII and chromatin and also prevents new initiation events of RNAPII during the run-on reaction (Chang & Luse, 1997; Core et al., 2012; Gariglio et al., 1974; Hawleys & Roederfj, 1985; Li et al., 2013; Rougvie & Lis, 1988; Shmookler et al., 1974; Wissink et al., 2019). The biotinylated nucleotides functionally block additional elongation due to steric hindrance (Kwak et al., 2013). This property is the basis of PRO-seq's single-nucleotide resolution. The use of a biotinylated nucleotide also permits stringent enrichment of nascent RNA using the high-affinity biotin-streptavidin interaction. PRO-seq libraries effectively capture nascent RNAs from protein coding genes and a repertoire of non-coding regions, e.g., eRNAs (Core & Adelman, 2019; Core et al., 2008; Jonkers et al., 2014; Kwak et al., 2013; Mahat et al., 2016). As a result, PRO-seq can reveal alterations in gene expression and enhancer activity in a single dataset.

Despite the power of PRO-seq, its wider adoption has been hampered by its famously laborious and troublesome protocol. A number of groups, in particular the Core and Lis labs, have made strides in streamlining the protocol (Judd et al., 2020; Mahat et al., 2016). To these, we add our version of the PRO-seq protocol that builds on previous versions and offers extensive tips and commentary. Here, we provide a stepwise protocol for construction of PRO-seq libraries that is suitable for a wide variety of cell types (Fig. 1). Briefly, this protocol starts by preparing permeabilized metazoan cells (see Basic Protocol 1) and continues through preparation to quality control of final libraries (see Basic Protocol 2).

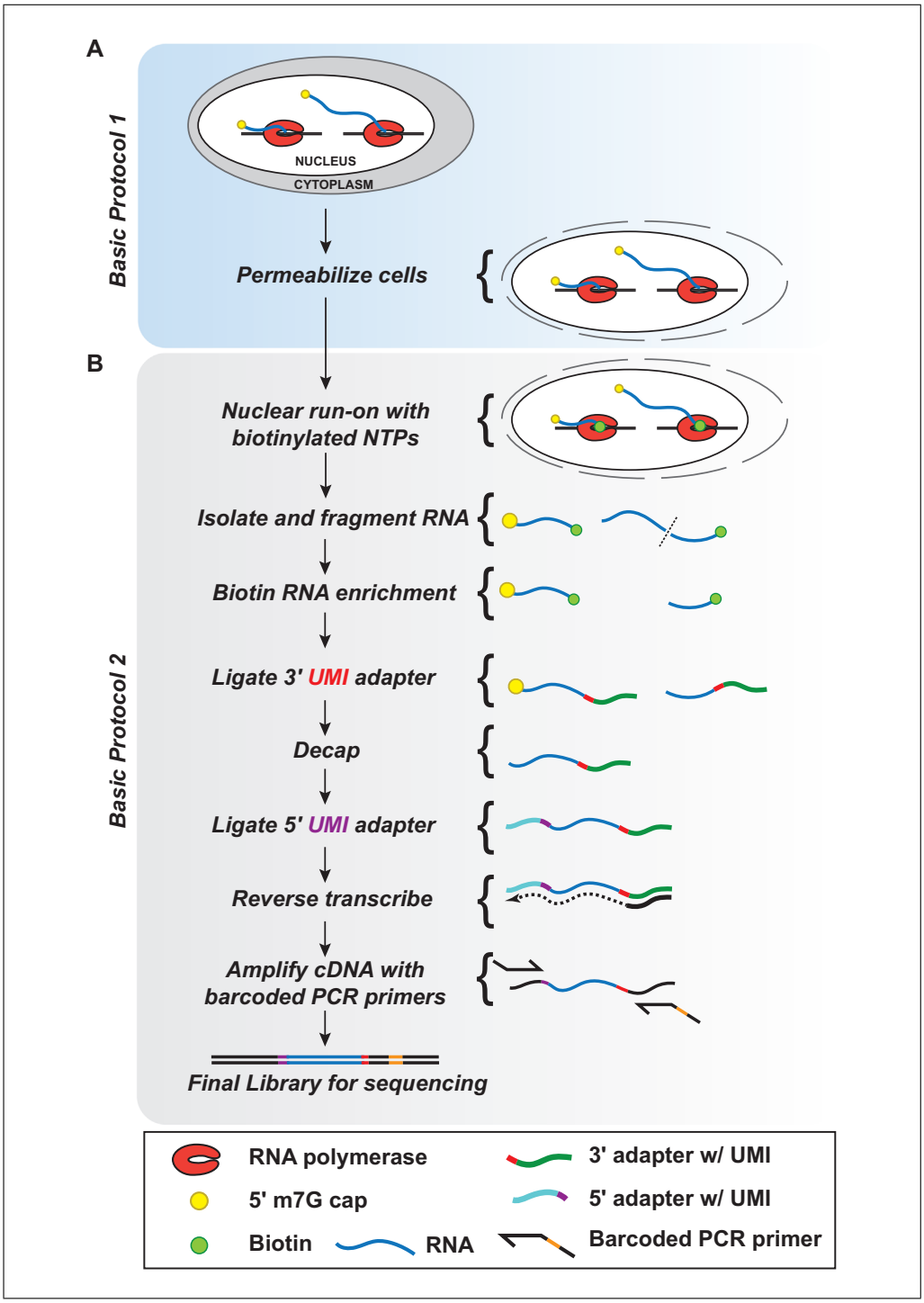

Figure 1 Overview of PRO-seq library construction. (**a**) Cell of interest are permeabilized in Basic Protocol 1 to halt transcription and deplete endogenous NTPs. (**b**) Sequence of steps for PRO-seq library construction in Basic Protocol 2. UMI, unique molecular identifier.

Additionally, a Support Protocol is provided to generate pre-adenylated 3′ adapters used in Basic Protocol 2 for library construction.

CAUTION: The protocol is not amenable to convenient in-line diagnostics to pinpoint the precise step(s) responsible for failed libraries. It is thus critical to follow the protocol strictly and to document that all steps are performed as intended. Great care must be taken to protect all reagents and stocks used in PRO-seq from contamination or mishandling. If applicable, it is recommended to include a positive control sample when executing Basic Protocol 2.

NOTE: It is best to use the time during enzyme incubations to prepare reagents for the next step. For enzyme master mixes, all components except the enzymes should be combined and kept on ice. Enzymes should be added immediately before use to maximize activity.

NOTE: All reagents, chemicals, and consumables (except cell culture medium) must be nuclease free. Chemicals should be of American Chemical Society (ACS) grade or higher. The work area and pipets should be decontaminated by wiping with an RNase inactivation reagent (e.g., RNaseZap, Invitrogen, cat. no. AM9784) before starting Basic Protocol 2 and then periodically as needed.

NOTE: All protocols involving animals must be reviewed and approved by the appropriate Animal Care and Use Committee and must follow regulations for the care and use of laboratory animals. Appropriate informed consent is necessary for obtaining and use of human study material.

BASIC PROTOCOL 1

CELL PERMEABILIZATION FOR PRO-seq

Permeabilized cells are used in the biotin run-on assay step of PRO-seq (Fig. 1a). To generate suitable samples for PRO-seq, cells are harvested immediately into ice-cold buffer to halt transcription and subsequently treated with mild detergents to create holes in the cell membranes, releasing intracellular NTP pools into solution and further stalling transcription elongation (Kwak et al., 2013; Mahat et al., 2016). The protocol described below is widely applicable to a wide variety of metazoan cell types and culture conditions. However, optimization of spin duration and/or *g* force may be necessary to obtain optimal quality for a given cell type (see Critical Parameters for more details).

Construction of PRO-seq libraries (see Basic Protocol 2) requires two types of permeabilized cells: experimental samples and a spike-in control. *Drosophila* S2 cells are commonly used to spike mammalian cells (and vice versa) because their genomes are sufficiently different to allow the majority of reads to be unambiguously mapped to the source genome. Permeabilized spike-in cells can be generated in larger batches using this protocol and frozen as single-use aliquots.

Materials

Cells of interest, adherent or suspension
Dulbecco's phosphate-buffered saline (DPBS) without calcium and magnesium (Corning, cat. no. 45000-434), ice cold
Trypsin (Corning, cat. no. 45000-660 or equivalent) *or* ESGRO Complete Accutase (Millipore, cat. no. SF006 or equivalent)
DMEM/FBS: Dulbecco's modified Eagle medium (Corning, cat. no. 45000-312 or equivalent) with 10% fetal bovine serum (Fisher Scientific, cat. no. 16-000-044 or equivalent), ice cold
Buffer W (wash; see recipe)
Buffer P (permeabilize; see recipe)
Buffer F (freeze; see recipe)
0.4% trypan blue (Logos Biosystems, cat. no. T13011 or equivalent)

Serological pipets
50-ml conical centrifuge tubes (VWR, cat. no. 10026-078 or equivalent)
Refrigerated tabletop centrifuge with swinging bucket rotor (Allegra X 15R or equivalent)
1.5-ml low-binding tubes (DNA LoBind, Eppendorf, cat. no. 022431021)
Filtered, low-retention pipet tips (Thomas Scientific, cat. no. 1149J70/1149J71/1138W95/1145N16 or equivalent)
Automated cell counter or hemocytometer
Liquid nitrogen or crushed dry ice

Harvest and permeabilize cells

1. Harvest cells as described below:

   For adherent cells: Aspirate medium from plate and rinse once with room-temperature 1× DPBS. Apply trypsin or Accutase at room temperature to release adherent cells, using conditions typical for your cell type. After 1-2 min (or when cells detach), quench with ice-cold DMEM/FBS (for trypsin) or ice-cold medium (for Accutase).

   For suspension cells: Tap flasks against your hand to detach cells and use a serological pipet as necessary to resuspend cells.

2. Collect cells in a 50-ml conical tube and place immediately on ice. Pool cells from multiple plates, if appropriate.

   To recover any remaining cells, rinse plate/flask with additional DMEM/FBS or medium, as used in step 1, and add to the 50-ml tube.

   We achieve the highest efficiency when a 50-ml tube contains 10-20 million mammalian cells or up to 50 million Drosophila S2 cells. Optimal and maximal cell inputs for other cell lines may vary.

   Since the pellet after permeabilization can be broad and translucent, it can be difficult to see and resuspend properly if cell numbers are very low. Conversely, if there are too many cells in a tube, they can aggregate after permeabilization and shear during resuspension.

   There is no need to determine the cell number precisely at this stage. If you know the expected cell number based on plating, plate size, or just looking at the dish, that should be sufficient.

   For lower cell inputs, we recommend using 5-ml or 1.5-ml low-binding tubes and scaling the reagent volumes down proportionally.

   From this point forward, cells must always be kept on ice.

3. Spin cells at 300 × *g* for 4 min at 4°C.

4. Remove supernatant and gently resuspend pellet in 10 ml ice-cold 1× DPBS.

   If the same cells will be used for multiple purposes (e.g., PRO-seq and RNA-seq), they should be counted at this step and aliquoted for each purpose before proceeding. When practicing the protocol, the count at this step is the "input" cell count used to calculate percent recovery (see Critical Parameters for more details).

5. Spin cells at 300 × *g* for 4 min at 4°C.

6. Remove supernatant and resuspend cells in 250 µl buffer W to obtain a single-cell suspension.

   It is critical that cells are in single-cell suspension at this stage. If cells are still clumped when buffer P is added, they can become a gunky mass that is sheared to pieces during resuspension and results in very low final cell counts.

7. Add 10 ml buffer P and incubate sample(s) on ice for 5 min.

   Add buffer P at a medium speed on an automatic pipettor, dispensing along the wall of the tube. It may help to set the tube diagonally in an ice bucket and let buffer P flow along the lower wall. This will enable mixing but avoid sloshing or foaming that would disrupt the cells.

   Keep the amount of time in buffer P consistent across samples, especially when processing multiple samples at the same time.

8. Spin at 400 × *g* for 8 min at 4°C. Check that the supernatant is clear and the pellet is compact to avoid cell loss.

If the supernatant is cloudy and the pellet has a hazy/wispy appearance at the top, spin again before removing the supernatant.

9. Remove supernatant from the first tube and immediately add 1 ml buffer W without resuspending. Repeat one tube at a time for all samples.

   The idea is to dilute any residual buffer P quickly.

10. Completely resuspend all pellets by pipetting.

11. Add an additional 9 ml buffer W to each tube.

    Add buffer W at medium speed on your automatic pipettor, as described in step 7.

12. Spin at 400 × *g* for 4 min at 4°C.

13. Remove supernatant and carefully pipet off any remaining buffer without disturbing the cell pellet.

14. Add 200 µl buffer F to each sample (without resuspending or mixing) and incubate on ice for 1 min.

    Allowing the pellet to equilibrate gently in buffer F before resuspending reduces clumping.

15. Resuspend samples in buffer F and transfer to a 1.5-ml low-binding tube.

16. Rinse the bottom of the conical tube with an additional 200 µl buffer F and add to the 1.5-ml tube for a final volume of 400 µl per sample.

17. Pipet the pooled cells to resuspend completely.

    Avoid introducing bubbles. The final suspension should look uniformly cloudy and free of clumps. If you draw the cells into a pipet tip and hold it up to the light, you should not see any "snowflakes" or particles. Permeabilized cells must be well resuspended for accurate counting.

    Some cells tend to clump at high cell densities in buffer F. If cells do not resuspend well by pipetting alone, try adding another 100-200 µl buffer F and mixing. Additional buffer F may be added, if necessary.

Count cells and prepare for PRO-seq

NOTE: It is critical to obtain an accurate cell count as this will determine the PRO-seq input.

18. Working one tube at a time, prepare 1:10 dilutions in DPBS. Pipet the sample to resuspend, transfer 10 µl to a tube containing 90 µl DPBS, and pipet thoroughly with a 200-µl pipet to mix.

19. Obtain a total cell count of the diluted sample (without trypan blue) using a cell counter or hemocytometer.

    If counts are outside the range of your counter, dilute as needed to obtain accurate counts.

20. Pipet an equal volume of trypan blue dye into the tube, tap gently two or three times to mix, and count again. Determine the percentage of trypan-positive cells.

    Trypan-positive cells are successfully permeabilized. A sample that stains >95% trypan-positive is successful.

    Some automatic cell counters give errors or report zero cells if almost all cells are trypan-positive. If the unstained counts are reasonable, such errors indicate that the cells are >98% permeabilized. It is good practice to always inspect each counting chamber on a microscope to help catch any quality issues.

21. Aliquot 1 million permeabilized cells per 45 µl buffer F.

    If cells are too dilute: Spin at 400 × g for 4 min at 4°C. Without disturbing the pellet, carefully pipet off the volume of supernatant necessary to increase the concentration to 1 million cells per 45 µl. Resuspend cells gently but thoroughly in the remaining volume before aliquoting.

    If cells are too concentrated: Add buffer F to decrease the cell concentration to 1 million cells per 45 µl. Resuspend gently but thoroughly in the remaining volume before aliquoting.

    We recommend using 1 million permeabilized cells for the biotin run-on reaction in Basic Protocol 2 (see Critical Parameters for more details). Aliquoting 1 million cells in 45 µl leaves room for the addition of spike-in cells and streamlines the run-on setup. The total volume of permeabilized cells for the biotin run-on reaction, including cells of interest and spike, must be 50 µl.

    For consistency of spike-in cells over time, stocks are best prepared in large batches and frozen as single-use aliquots. Each aliquot should contain sufficient volume for a typical batch of PRO-seq (six to twelve samples).

22. Snap freeze permeabilized cells in liquid nitrogen and store at −80°C (stable at least 6 months).

CONSTRUCTION OF PRO-seq LIBRARIES

BASIC PROTOCOL 2

The protocol below describes how to use the permeabilized cells generated in Basic Protocol 1 to construct PRO-seq libraries (Fig. 1b). After thawing, permeabilized sample cells are spiked by mixing in permeabilized cells from another species whose transcripts have minimal homology to the sample (spike in). For example, permeabilized mammalian cells can be spiked with permeabilized *Drosophila* S2 cells and vice versa. A 5% ratio of spike to sample cells (e.g., 5×10^4 spike cells for 10^6 mammalian cells) generally yields an ∼5% ratio of spike to sample reads. Sample cells with lower- or higher-than-average transcriptional activity will yield more or fewer spike reads, respectively. The mixed permeabilized cells are reconstituted in run-on reaction buffer containing Sarkosyl and biotin-NTPs to resume transcription. The "bulky" biotin moiety allows for incorporation of a single biotinylated NTP by engaged RNAPII, allowing for single-nucleotide resolution mapping of RNAPII. The presence of Sarkosyl promotes elongation by dissociating inhibitory factors from RNAPII and chromatin and by blocking initiation of new transcription (Chang & Luse, 1997; Core et al., 2012; Gariglio et al., 1974; Hawleys & Roederfj, 1985; Li et al., 2013; Rougvie & Lis, 1988; Shmookler et al., 1974; Wissink et al., 2019).

After the run-on reaction is complete, the RNA is purified and fragmented, and then the nascent RNAs containing a single biotin-NTP are efficiently captured onto streptavidin magnetic beads and washed stringently. A series of enzymatic steps are then performed to add adapters to the RNA 3′ and 5′ ends. After samples are reverse-transcribed, the final library is amplified by PCR. The majority of these steps are performed while nascent RNAs are bound to streptavidin beads for faster processing.

Library quantification

The 5′ end of the 3′ adapter and 3′ end of the 5′ adapter consist of six randomized nucleotides, referred to as a unique molecular identifier (UMI; Fig. 1). UMIs are transcript-level barcodes that enable discrimination between biological and amplification duplicates. PRO-seq libraries are enriched in transcripts indistinguishable from amplification duplicates (e.g., promoter proximally paused RNAs will often have identical 5′ and 3′ ends). As a result, UMIs are required to accurately quantify a substantial fraction of the

transcripts in the library. For optimal UMI performance, the number of possible UMI sequences, equal to $4^{(\text{number of randomized bases})}$, must be much larger than the number of biological duplicate RNAs for any locus. This ensures that biological duplicates are highly unlikely to receive the same UMI. In addition to allowing correction of amplification bias, UMIs also reduce ligation biases (Fuchs et al., 2015; Zhuang et al., 2012) that might otherwise distort apparent RNA abundance.

UMI-based quantification works by counting unique UMIs associated with transcripts rather than transcripts themselves. For example, assume a sample has ten copies of transcript A and two of B, all with unique UMIs. After amplification, the library contains 100 copies each of transcripts A and B. Without UMIs, transcripts A and B appear to have equal expression. However, with UMIs we observe that the 100 copies of A and B are comprised of only ten and two UMIs, respectively, which reveals the original relative abundances. (For additional references on the application of UMIs, see Fu et al., 2018; Hong & Gresham, 2017; Kivioja et al., 2012; Saunders et al., 2020; Smith et al., 2017.)

Timeline

The complete protocol is performed in three working days, but these do not need to be consecutive. Day 1 includes preparation of streptavidin beads, the run-on reaction, RNA purification and fragmentation, enrichment of biotin-labeled fragments using streptavidin beads, and 3′ adapter ligation. Day 2 includes a second round of streptavidin bead binding, 5′ decapping and 5′ hydroxyl repair, 5′ adapter ligation, reverse transcription, and partial amplification of the resulting cDNA library. Day 3 begins with a series of test amplifications to determine the optimal conditions for final library amplification. A full amplification is then performed, followed by library cleanup and quantification. After days 1 and 2, samples can be stored for several days at –80°C before proceeding to the next step.

Bead washes

The protocol contains multiple bead washing steps. The following technique should be used any time beads are washed:

Resuspend beads in buffer by pipetting
Rotate beads for 1 min at room temperature
Quick spin beads
Place beads on magnet for 1 min
Carefully remove supernatant and discard

Be sure to check that no beads are in the tip before discarding the supernatant. If any beads are aspirated, dispense them back to the sample tube. Also, never allow beads to dry out, as dry beads will clump irreversibly and reduce yields. Add the next solution promptly after removing the supernatant.

Materials

DynaBeads MyOne Streptavidin C1 (Thermo Fisher Scientific, cat. no. 65-001)
Decon solution (see recipe)
5 M NaCl BioUltra (Sigma-Aldrich, cat. no. 71386)
Binding buffer (see recipe)
SUPERase·In RNase inhibitor (Thermo Fisher Scientific, cat. no. AM2696)
Nuclease-free water (VWR, cat. no. E476-1l or equivalent)
1 M Tris·Cl, pH 8.0 (Fisher Scientific, cat. no. 15568025 or equivalent)
1 M $MgCl_2$ (Sigma-Aldrich, cat. no. M1028 or equivalent)
1 M dithiothreitol (DTT, Sigma-Aldrich, cat. no. 43816 or equivalent)
KCl (Millipore Sigma, cat. no. 60135-250ML or equivalent)
Biotin-11-NTPs (A/G/C/U; Perkin Elmer, cat. no. NEL54[2,3,4,5]001)
Permeabilized cells, experimental and spike-in (see Basic Protocol 1)

20% sarkosyl (Sigma-Aldrich, cat. no. L7414)
Total RNA Purification Kit (Norgen Biotek, cat. no. 37500) including Buffer RL, Wash Solution A, columns, and collection tubes
Ethanol, 200 proof (VWR, cat. no. TX89125170HU or equivalent)
5× fragmentation buffer (see recipe), ice cold
0.5 M EDTA, pH 8.0 (Invitrogen, cat. no. 15-575-020 or equivalent)
High-salt wash buffer (see recipe)
Low-salt wash buffer (see recipe)
TRIzol (Invitrogen, cat. no. 15596018 or equivalent)
Chloroform (≥99.8%, unstabilized, VWR, cat. no. MK444004)
GlycoBlue (Thermo Fisher Scientific, cat. no. AM9516)
Adenylated 3′ adapter (App-VRA3-6N, see Support Protocol)
T4 RNA ligase 2, truncated KQ (New England Biolabs, cat. no. M0373) with 10× buffer and PEG 8000
Betaine binding buffer (see recipe)
mRNA decapping enzyme (MDE, New England Biolabs, cat. no. M0608S) with 10× buffer
T4 polynucleotide kinase (PNK, New England Biolabs, cat. no. M0201) with 10× buffer
10 mM adenosine triphosphate (ATP, New England Biolabs, cat. no. P0756)
T4 RNA ligase 1, high concentration (New England Biolabs, cat. no. M0437) with 10× buffer and PEG 8000
5′ adapter (VRA5-6N, Integrated DNA Technologies, Table S1)
SuperScript IV reverse transcriptase (SSIV, Invitrogen, cat. no. 18090200) with 5× buffer and 0.1 M DTT
RP1 primer (Integrated DNA Technologies, Table S1)
25 mM dNTP mix (Thermo Fisher Scientific, cat. no. R1121)
RPI-X primers (Integrated DNA Technologies, Table S1)
NEBNext Ultra II Q5 Master Mix (New England Biolabs, cat. no. M0544)
P5 and P7 primers (Integrated DNA Technologies, Table S1)
SeaKem Agarose LE (Lonza, cat. no. 50004)
10× TBE buffer (Bio-Rad, cat. no. 1610770 or equivalent)
6× loading dye
MassRuler Low Range DNA Ladder (Thermo Fisher Scientific, cat. no. SM0383)
SYBR Gold Nucleic Acid Gel Stain (Thermo Fisher Scientific, cat. no. S11494)
ProNex Size-Selective Purification System (Promega, cat. no. NG2001), including ProNex Size-Selective Chemistry, wash buffer, and elution buffer
NEBNext Library Quant Kit for Illumina (New England Biolabs, cat. no. E7630) with master mix, primer mix, dilution buffer, six standards, and ROX passive reference dye

1.5-ml low-binding tubes (DNA LoBind, Eppendorf, cat. no. 022431021)
Low-retention, filtered pipet tips
Magnetic tube racks for 1.5-ml tubes and PCR strips
Mini Tube Rotator (Thomas Scientific, cat. no. 1217H25 or equivalent)
Heating block or water bath
200-µl wide-bore, low-retention, filtered pipet tips (Thermo Fisher, cat. no. 2069G)
PCR tubes and strip tubes
Thermocycler (Bio-Rad C1000, T100 or similar)
Aluminum foil
384-well qPCR plate, white wells (Bio-Rad HSP-3805 or similar)
Plate seals for qPCR (Bio-Rad MSB-1001 or similar)
qPCR machine (Bio-Rad CFX384 or similar)
NEBioCalculator webtool (*nebiocalculator.neb.com*)

Additional reagents and equipment for agarose gel electrophoresis, staining, and imaging

Day 1

Prepare streptavidin magnetic beads

Streptavidin beads are used to isolate nascent RNAs. They should be prepared fresh at the beginning of each experiment and stored at 4°C.

1. Vortex DynaBeads MyOne Streptavidin C1 until fully resuspended. Aliquot 20 µl beads to a 1.5-ml low-binding tube for each sample.
2. Place tubes on a bead magnet for 1 min and discard supernatant.
3. Resuspend beads in 500 µl Decon solution and rotate tubes for 2 min at room temperature.
4. Place tubes on magnet for 1 min and discard supernatant.
5. Repeat steps 3-4 once for a total of two washes.
6. Wash beads twice with 500 µl of 100 mM NaCl.
7. Wash beads twice with 500 µl binding buffer.
8. Resuspend beads in 99 µl binding buffer.
9. Add 1 µl SUPERase·In.

   There are two streptavidin bead binding steps, and each requires 50 µl prepared beads as input.

10. Store beads at 4°C.

Perform biotin run-on

11. Prepare 25 µl of 4× run-on master mix in a 1.5-ml low-binding tube:

    11.0 µl nuclease-free water
    0.5 µl 1 M Tris·Cl, pH 8.0
    0.5 µl 1 M $MgCl_2$
    0.5 µl 0.1 M DTT
    7.5 µl 2 M KCl
    1.0 µl 1 mM biotin-11-ATP
    1.0 µl 1 mM biotin-11-GTP
    1.0 µl 1 mM biotin-11-CTP
    1.0 µl 1 mM biotin-11-UTP
    1.0 µl SUPERase·In

    For multiple reactions, prepare a single master mix and aliquot 25 µl into the appropriate number of reaction tubes.

    The biotin-11-CTP and -UTP stocks are 10 mM and should be diluted to 1 mM in nuclease-free water before use. Final concentrations in the 1× run-on reaction are 5 mM Tris·Cl, 5 mM $MgCl_2$, 0.5 mM DTT, 150 mM KCl, 10 µM of each biotin-NTP, and 0.2 U/µl SUPERase·In.

12. Preheat 4× master mix to 37°C in a heating block or water bath.
13. Thaw permeabilized sample and spike-in cells on ice.
14. Mix spike cells well by pipetting and inspect them visually to ensure that there are no clumps.
15. Add 5 µl spike-in cells to the 45-µl experimental sample (total 50 µl). Keep on ice.

Accurate pipetting is critical. Use a fine-bore 10- or 20-µl pipet tip.

Discard the spike aliquot after use. Do not refreeze.

It is important to use the same batch of spike cells for all samples that will be compared with each other. Subtle differences in cell growth and permeabilization could make it difficult to compare spike returns between spike batches.

If you are not using spike-in cells, bring cell volume to 50 µl with cold buffer F.

16. Add 25 µl of 2% sarkosyl to the preheated 4× master mix (total 50 µl at 2×). Return to 37°C.

    The final concentration of sarkosyl in the 1× run-on reaction is 0.5%. Sarkosyl can precipitate out of solution at 4°C and thus should be added to prewarmed master mix that is then kept at room temperature or higher.

17. Remove permeabilized cells from ice and equilibrate at room temperature for exactly 5 min.

    If processing two or more samples at the same time, stagger them by 30-60 s so that incubation times are kept identical. Start the timer counting up when the first reaction begins. Example:

    Sample A:
    0:00 Move cells to RT (step 17)
    5:00 Start run-on (step 18)
    7:30 Flick to mix (step 19)
    10:00 Stop run-on (step 20)

    Sample B:
    0:30 Move cells to RT (step 17)
    5:30 Start run-on (step 18)
    8:00 Flick to mix (step 19)
    10:30 Stop run-on (step 20)

    For most users, up to nine samples can be reasonably processed in one batch. We do not recommend staggering samples by less than 30 s.

18. Using a wide-bore 200-µl pipet tip, briefly mix cells, transfer the entire volume (50 µl) to the tube of prewarmed master mix, and quickly pipet 15 times to mix (avoiding bubbles). Return promptly to 37°C to start the run-on reaction.

    Permeabilized cells can settle at the bottom of the tube during the 5-min incubation at room temperature and should be mixed before being added to the master mix. The reaction will become very viscous during pipetting due to the presence of sarkosyl, so it is also critical to mix the reaction thoroughly after adding cells to distribute the biotinylated NTPs uniformly.

19. Incubate 5 min at 37°C. Halfway through the incubation, flick the tube gently to mix.

Stop run-on and purify RNA

This section was adapted from the Norgen Biotek Total RNA Purification Kit. All steps are performed at room temperature. All spins are at 14,000 × *g* unless otherwise noted.

20. Stop the run-on reaction by adding 350 µl Buffer RL and vortexing for 10 s.

    If processing multiple samples, proceed with the next step after all reactions have been stopped.

21. Quick spin all reactions.

22. Add 240 µl of 100% ethanol (room-temperature), vortex 10 s to mix, and quick spin.

23. Assemble one Norgen column in a 2-ml collection tube for each sample.
24. Pipet the entire volume of each run-on lysate (~700 µl) onto its respective column.
25. Spin 1 min at 3500 × *g* to bind RNA and discard the flowthrough.
26. Wash column three times with 400 µl Wash Solution A. For each wash, apply buffer, spin 1 min, and discard the flowthrough.

    Ethanol must be added to Wash Solution A or samples will be lost during purification. See manufacturer's instructions.

27. Dry the column by spinning for 2 min.
28. Transfer column to a 1.5-ml low-binding tube.
29. Apply 50 µl water directly to the center of the matrix and spin as follows to collect the eluate:

    2 min at 200 × *g*
    1 min at 14,000 × *g*

30. Repeat step 29, collecting the second eluate in the same tube (total 100 µl).
31. Measure sample volumes with a pipet to verify that all volumes are similar. If needed, add water to any low volumes to match the rest of the set to obtain uniform hydrolysis.

Fragment RNA

32. Transfer sample to a fresh PCR tube and incubate on ice for 2 min.
33. Add 25 µl ice-cold 5× fragmentation buffer and pipet to mix.
34. Incubate for *exactly* 5 min at 94°C in a thermocycler with lid closed. After 5 min, place samples immediately on ice.

    The chemical fragmentation is very time sensitive. To maintain consistency between samples, it is important to move as quickly as possible between steps. To do so, place the mini centrifuge, ice bucket, etc. next to the thermocycler. If multiple samples are being processed at once, the steps below can be performed at the same time in a PCR strip.

    Fragmentation time may require additional optimization (see Troubleshooting).

35. Add 125 µl of 0.1 M EDTA to each sample and pipet to mix.
36. Quick spin and incubate immediately on ice for 2 min.
37. Quick spin.

Perform first bead binding and precipitate nascent RNA

REMINDER: Refer to the discussion about proper bead washing in the protocol introduction.

38. Resuspend prepared beads (step 10) evenly by pipetting and add 50 µl to each sample.
39. Rotate for 20 min at room temperature.
40. Place on the magnet and remove supernatant.
41. Perform the following washes in order, removing the supernatant on the magnet after each wash:

    2× with 400 µl high-salt wash
    2× with 400 µl binding buffer
    2× with 400 µl low-salt wash

42. Resuspend beads in 500 µl TRIzol.

    CAUTION: *TRIzol and similar reagents contain phenol and are toxic by inhalation and skin contact. Contact causes severe tissue damage. Wear gloves, a lab coat, and eye protection and handle in a chemical fume hood.*

43. Vortex samples for 20 s and then incubate 2.5 min at 65°C.

44. Vortex samples for 20 s, spin quickly, and incubate another 2.5 min at 65°C.

45. Place on magnet for 1 min, then transfer TRIzol/supernatant to a fresh tube.

46. Repeat steps 42-45, pooling the eluate in the same tube (total 1 ml).

    It is not a problem if a few beads are carried over into the TRIzol. They will pellet during phase separation.

47. Add 200 µl chloroform.

    CAUTION: *Chloroform is toxic by inhalation and skin absorption. Acute exposure can result in loss of consciousness, and prolonged exposure may be carcinogenic. Wear gloves, a lab coat, and eye protection and handle in a chemical fume hood.*

48. Vortex at least 20 s, then incubate 1 min at room temperature.

49. Spin 5 min at max speed at 4°C.

50. Transfer aqueous layer (400-450 µl) to a fresh low-binding tube.

51. Add 1.5 µl GlycoBlue and vortex to mix.

52. Add 1100 µl ice-cold 100% ethanol.

53. Incubate >1 hr at −20°C.

    As alternative end of day 1, precipitate samples overnight at −20°C, then begin day 1.5 at step 54.

54. Spin 30 min at max speed at 4°C.

55. Remove supernatant.

56. Wash pellet with 1 ml ice-cold 75% ethanol.

    The pellet must be dislodged from the tube wall during the wash, but should not be broken up. This can be done by inverting the tube or gently pipetting the ethanol at the pellet. Do not pipet the pellet up and down.

    IMPORTANT: *Check the pellet composition before moving forward. If the pellet is flat (like a feather), stuck to the side of the tube after step 54, and falls slowly to the bottom of the tube in 75% ethanol during step 56, you may proceed to the next step. If the pellet looks like a blue jellybean (i.e., is a ball of semi-solid material), does not adhere to the tube after step 54, and sinks rapidly to the bottom of the tube in step 56, the samples should be reprecipitated before moving forward. To reprecipitate, perform step 57, removing as much ethanol as possible. Then, resuspend the pellet with 200 µl of 0.3 M sodium acetate, pH 5.5 (prepared by 1:10 dilution of 3 M sodium acetate stock), add 600 µl ice-cold 100% ethanol, repeat steps 53-56, and re-evaluate pellet quality.*

57. Spin 10 min at 14,000 × *g* at 4°C. Carefully remove supernatant.

    Remove as much ethanol as possible with a 1000-µl pipet tip, then with a 200-µl tip. Spin quickly, then remove any remaining ethanol with a 20- or 10-µl tip.

58. Air-dry samples for ∼10 min.

Ligate 3′ adapter

59. For each sample, add 1 µl of 10 µM 3′ adapter (App-VRA3-6N) to 4 µl water.
60. Resuspend air-dried pellet in 5 µl diluted 3′ adapter.
61. Incubate 30 s at 65°C, then place on ice.
62. Prepare 15 µl of 3′ ligation master mix for each sample:

    6 µl 50% PEG 8000
    2 µl 10× T4 RNA ligase buffer
    0.5 µl SUPERase·In
    1 µl T4 RNA ligase 2, truncated KQ
    5.5 µl water

63. Add 15 µl 3′ ligation master mix to each sample and incubate overnight at 16°C.

    After the 3′ ligation is complete, samples may be stored at −80°C for several days before proceeding with day 2.

Day 2

Perform second bead binding and 5′ decapping

64. Remove samples from 16°C. If samples were frozen after the ligation reaction, thaw them on ice.
65. Add 180 µl betaine binding buffer and pipet to mix.
66. Heat 5 min at 65°C, then place on ice for 1 min.
67. Fully resuspend prepared streptavidin beads (step 10), add 50 µl to each sample, and pipet to mix.
68. Rotate for 20 min at room temperature.
69. During incubation, prepare the following for the decapping reaction:

    100 µl (per sample) of 1× MDE buffer
    17 µl (per sample) of incomplete MDE master mix containing 15 µl water and 2 µl of 10× MDE Buffer

    To save time, we recommend preparing all enzyme master mixes in this protocol without enzyme (i.e., incomplete) during the incubation. The incomplete master mix can be stored on ice. To maximize activity, enzyme and SUPERase·In are added just before use.

70. Place samples on magnet for 1 min. Discard the supernatant.
71. Resuspend beads in 400 µl high-salt wash and transfer the entire bead suspension to a new 1.5-ml low-binding tube.

    This step greatly reduces the amount of adapter dimer in the final library by eliminating unligated 3′ adapter adhered to the tube walls. This is necessary despite the stringent washes and low-binding plastics.

72. Perform the following washes:

    2× with 400 µl binding buffer
    1× with 400 µl high-salt wash
    2× with 400 µl low-salt wash

73. Wash 1× with 100 µl of 1× MDE Buffer but do *not* remove the supernatant.

    The beads should not be left to dry out.

74. Complete the MDE master mix by adding 1 µl MDE and 1 µl SUPERase·In per sample.

75. Remove 1× MDE buffer from the beads and discard.

76. Resuspend beads in 19 µl complete MDE master mix and rotate samples for 30 min at 37°C.

    Rotate tubes horizontally rather than end-over-end. The reaction should stay in the bottom of the tube rather than sloshing up and down the sides.

Perform 5′ hydroxyl repair

77. Prepare 100 µl of 1× PNK buffer per sample.

78. Prepare 17 µl incomplete PNK master mix per sample:

    13 µl water
    2 µl 10× PNK buffer
    2 µl 10 mM ATP

79. Quick spin samples, then place on magnet for 1 min and discard the supernatant.

80. Perform the following washes:

    1× with 400 µl high-salt wash
    1× with 400 µl low-salt wash
    1× with 100 µl 1× PNK buffer

81. Add 1 µl of T4 PNK and 1 µl SUPERase·In (per sample) to the incomplete PNK master mix.

82. Resuspend beads in 19 µl complete PNK master mix and rotate 1 hr at 37°C.

    Follow the same rotation guidelines in step 76.

Ligate 5′ adapter

83. For each sample, dilute 2.5 µl of 10× T4 RNA ligase buffer in 97.5 µl water to make a 0.25× ligase buffer wash.

84. For each sample, add 1 µl of 10 µM 5′ adapter (VRA5-6N) to 4 µl water.

85. Place samples on the magnet for 1 min. Remove supernatant.

86. Perform the following washes:

    1× with 400 µl high-salt wash
    1× with 400 µl low-salt wash
    1× with 100 µl 0.25× ligase buffer wash

87. Resuspend beads in 5 µl diluted 5′ adapter and incubate for 30 s at 65°C.

88. Incubate 10 min at room temperature.

89. During the 10-min incubation, make 15 µl ligation master mix per sample:

    3 µl water
    2 µl 10× T4 RNA ligase 1 buffer
    2 µl 10 mM ATP
    6 µl 50% PEG 8000
    1 µl SUPERase·In
    1 µl T4 RNA ligase 1 (high concentration)

90. Add 15 µl ligation master mix to each sample, pipet well to mix, and rotate for 2 hr at room temperature.

    Follow the same rotation guidelines in step 76.

RPI-X primers will be used in the next set of reactions to assign a specific index to each library. Use this 2-hr incubation to determine the appropriate index primers for samples.

Perform reverse transcription

91. Prepare 100 µl of 0.25× SSIV buffer per sample.
92. Add 1 µl of 10 µM RP1 primer to 6.5 µl water per sample.
93. Quick spin samples, place on magnet for 1 min, and discard supernatant.
94. Perform the following washes:

    2× with 400 µl high-salt wash
    2× with 400 µl binding buffer
    2× with 400 µl low-salt wash
    1× with 100 µl 0.25× SSIV buffer

95. Resuspend beads in 7.5 µl diluted RP1 primer.
96. Heat 3 min at 65°C, then place samples immediately on ice.
97. Prepare 7.5 µl RT master mix per sample:

    3 µl 5× SSIV buffer
    0.75 µl 0.1 M DTT
    0.75 µl 12.5 mM dNTP mix
    0.75 µl SUPERase·In
    1 µl SSIV RTase
    1.25 µl water

    The 12.5 mM dNTP mix is prepared by diluting the 25 mM stock 1:1 in nuclease-free water. Aliquots can be stored for weeks at −20°C.

98. Transfer samples to a PCR tube strip.
99. Add 7.5 µl RT master mix to each sample and run the reaction in a thermocycler under the following conditions (lid temperature = 59°C):

    10 min at 48°C
    15 min at 54°C
    hold at 4°C

Elute cDNA library and perform initial amplification

100. Pipet samples to resuspend beads.
101. Preheat a thermocycler to 95°C with lid at 100°C.
102. Incubate samples in the thermocycler for 30 s.
103. Remove samples, spin down quickly, and place on a magnetic stand for 30 s.
104. Transfer entire supernatant (~15 µl) to a new PCR strip tube, taking care to minimize the transfer of beads, and place on ice.
105. Resuspend beads in 11.5 µl water.
106. Repeat steps 102-104, pooling the second supernatant with the first.

    If residual beads are accidentally carried over, place the PCR strip on a magnetic stand for 1 min and transfer the entire supernatant (~26 µl) to a new PCR strip.

107. Add 1 µl of 10 µM RPI-X to each sample (using a different index primer for each library).

108. Prepare 28.5 µl PCR master mix by adding 1 µl of 10 µM RP1 primer to 27.5 µl Ultra II Q5 Master Mix per sample.

109. Add 28.5 µl PCR master mix to each sample and pipet to mix.

110. Amplify under the following conditions (lid temperature = 105°C):

1 cycle:	3 min at 98°C
5 cycles:	15 s at 98°C
	75 s at 65°C
1 cycle:	5 min at 65°C
Hold:	4°C

111. Store partially amplified libraries at −20°C.

Performing the initial and final amplifications separately (see steps 123-125) allows day 2 to be completed in ∼8 hr instead of 12 hr or more. Converting the cDNA to dsDNA in the initial amplification also increases library stability for storage. Samples may be kept at −20°C for several days before proceeding to day 3.

Day 3

Perform test amplifications

112. Thaw partially amplified libraries on ice. Mix gently by flicking tubes, then quick spin and return to ice.

113. Combine 3.3 µl partially amplified library (6% of total reaction) with 4.7 µl water.

114. Make four 4-fold dilutions by adding 2 µl of each dilution to 6 µl water. Discard 2 µl from the final dilution so that all five dilutions contain 6 µl.

115. Prepare 84 µl test amplification master mix for each library:

60 µl Ultra II Q5 Master Mix
12 µl primer mix containing 10 µM each P5 + P7
12 µl water

Each test amplification requires 14 µl master mix. Preparing 84 µl per library will provide enough master mix for all five dilutions plus one extra.

116. Add 14 µl test amplification master mix to each library dilution (total 20 µl) and pipet to mix.

117. Amplify under the following conditions (lid temperature = 105°C):

1 cycle:	3 min at 98°C
14 cycles:	15 s at 98°C
	75 s at 65°C
1 cycle:	5 min at 65°C
Hold:	4°C

118. Prepare a 2.2% agarose gel in 1× TBE.

The gel will be post-stained SYBR Gold. Do not cast gel with ethidium bromide or SYBR Gold.

Not all agarose brands perform equally well for post-staining with SYBR Gold. In our hands, SeaKem Agarose LE works well.

Gel thickness, staining conditions, and imaging exposure should be kept consistent to facilitate comparison between PRO-seq batches performed on different days.

119. Add 4 µl of 6× loading dye to each sample and load on the gel along with a low-range DNA ladder. Run gel for 30 min at 120 V.

120. Place gel in a plastic tray, cover the gel with 1× SYBR Gold in 1× TBE, and then cover the tray with foil.

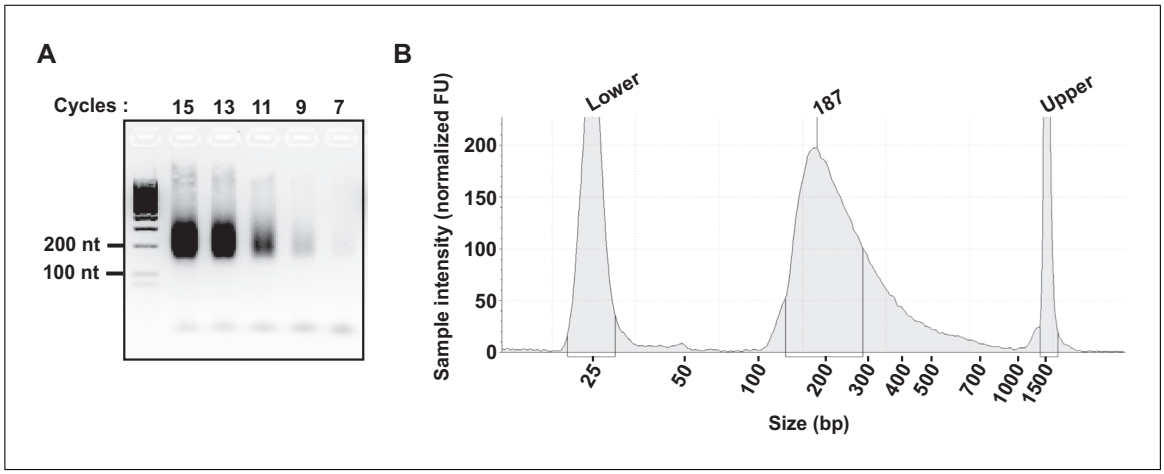

Figure 2 Results of PRO-seq test and final amplifications. (**a**) Test amplification dilution series resolved on a 2.2% agarose gel and stained with SYBR Gold. Lane 1 corresponds to 15 total cycles of amplification (5 cycles preliminary amplification + 10 additional cycles). Lanes 2-5 correspond to the total number of cycles shown above. In this example, 7 to 9 total cycles (i.e., 2-4 additional cycles) would be appropriate. (**b**) Example size distribution of a PRO-seq library after final amplification and library cleanup. Samples were run on the 4200 TapeStation using a High Sensitivity DNA Screen Tape (Agilent, cat. no. 5067-5584) per manufacturer's instructions and signal was scaled per sample.

For uniform staining, the container should be large enough to allow the gel to move a little. Prepare 100 ml staining solution or enough to fully submerge the gel.

SYBR gold is light sensitive. The solution and stained gel should be protected from light.

121. Incubate 10-40 min with gentle agitation in the dark.

122. Image the gel and identify the dilution with the optimal library product (visible library smear, PCR primers <50% depleted, no excess adapter dimers, etc.).

The most dilute sample corresponds to seven cycles of amplification total (five cycles of preliminary amplification + two additional cycles). The least dilute sample corresponds to fifteen total cycles (five cycles of preliminary amplification + ten additional cycles). Intermediate dilutions correspond to nine, eleven, and thirteen total cycles. An example test amplification gel is shown in Figure 2a.

Select the lowest number of total cycles that still produces a visible library with good characteristics. In general, if a band is visible on the gel, it is usually safe to amplify using the number of cycles indicated by that dilution or one fewer. Each library should be amplified to the extent indicated by its test amplification.

It is safe to proceed with libraries that require twelve or fewer cycles. Libraries that require thirteen or more cycles will likely result in low-quality libraries containing amplification biases, and a considerable loss of sequencing reads due to PCR duplication. As a result, we recommend rebuilding libraries to improve library yields (for tips, see Critical Parameters).

Perform full amplification

123. Add 3.3 µl Ultra II Q5 Master Mix to each sample to replace the volume removed for test amplification. Pipet to mix thoroughly.

124. Using the cycle number determined in the test amplifications (*n*), amplify samples using the following conditions (lid temperature = 105°C):

1 cycle:	3 min at 98°C
n cycles:	15 s at 98°C
	75 s at 65°C
1 cycle:	5 min at 65°C
Hold:	4°C

125. When the reaction is complete, proceed to library cleanup or store samples up to 1 week at −20°C.

Perform library cleanup

126. Bring ProNex Chemistry (bead suspension) to room temperature for ∼30 min before use.
127. Transfer samples to a 1.5-ml low-binding tube and check volumes with a 200-µl pipet tip.
128. Resuspend warmed ProNex beads completely by vortexing.
129. Add 2.6× vol. ProNex beads to each sample and pipet up and down 10× to mix.
130. Incubate 20 min at room temperature.
131. Place on the magnet for 2 min, and carefully remove and discard supernatant.
132. Apply 200 µl wash buffer, incubate for 30 s, and remove and discard supernatant. Repeat for a total of two washes.
133. Remove any residual wash buffer using a 20-µl pipet tip.
134. Air-dry beads for ∼5-10 min on the magnet.
135. Remove samples from magnet, add 12.5 µl elution buffer or water, and pipet until beads are fully resuspended.
136. Incubate 10 min.
137. Place on magnet for 2 min and transfer supernatant (the final size-selected library) to a new low-binding tube.

    This marks a stopping point. Samples can be stored for several months at −20°C.

138. If available, run 1-2 µl of the final library on a TapeStation to check the library size distribution.

    An example library TapeStation profile is shown in Figure 2b.

Quantify qPCR library

This section summarizes the use of the NEBNext Library Quant Kit for Illumina. Add the provided ROX passive reference dye to the master mix if necessary for your qPCR instrument. Comparable kits are available from other vendors and should be used according to the manufacturer's instructions.

139. Prepare 1.2 ml of 1× dilution buffer in nuclease-free water for each sample.
140. Prepare the following sample dilutions:

    Add 1 µl library to 999 µl of 1× dilution buffer (1:1000 dilution).
    Add 10 µl of 1:1000 dilution to 90 µl of 1× dilution buffer (1:10,000 dilution).
    Add 10 µl of 1:10,000 dilution to 90 µl of 1× dilution buffer (1:100,000 dilution).

141. Load a 384-well qPCR plate with 8 µl Library Quant Master Mix (with primers, per manufacturer instructions) per well.
142. Load 2 µl of each standard and sample to corresponding wells containing master mix:

    DNA standards (100, 10, 1, 0.1, 0.01, and 0.001 pM)
    No-template control (nuclease-free water)

1:10,000 library dilution × number of samples
1:100,000 library dilution × number of samples

For best results, run all standards and samples in technical triplicate.

143. Seal plate and spin down for 1 min at 200 × *g*.

144. Cover plate with foil and incubate 10-40 min at 4°C.

145. Remove foil and spin down for 1 min at 200 × *g*.

146. Run the following program in a real-time thermocycler using the SYBR Green channel.

1 cycle:	2 min at 95°C
35 cycles:	15 s at 98°C
	45 s at 63°C

147. Export data as a spreadsheet and enter the values in the NEBioCalculator webtool (*nebiocalculator.neb.com*) to determine sample concentrations.

Library yields typically range from 5 to 40 nM. Yields greater than ∼50 nM often indicate some degree of over-amplification and rates of UMI-deduplication may be elevated during data processing.

148. Follow the submission instructions for your sequencing facility.

SUPPORT PROTOCOL

ADENYLATION OF 3′ ADAPTER

The 3′ adapter ligation uses truncated T4 RNA ligase 2, which lacks the ability to form adenylated reaction intermediates and thus requires pre-adenylated adapters. Using pre-adenylated adapters and truncated T4 RNA ligase 2 offers advantages: (1) reduced ligation bias and (2) reduced production of undesired ligation byproducts such as circularized or concatamerized inserts. As a result, we and others (Hafner et al., 2011; Munafó & Robb, 2010; Viollet et al., 2011) have found that T4 RNA ligase 2 generates higher ligation efficiencies than T4 RNA ligase 1. The protocol below describes the preparation and quality control of pre-adenylated RNA adapters. Adenylation and QC must be complete prior to starting Basic Protocol 2.

Materials

3′ RNA adapter: VRA3-6N (Integrated DNA Technologies, Table S1)
5′ DNA adenylation kit (New England Biolabs, cat. no. E2610S), including Mth RNA ligase, 10× buffer, 1 mM ATP
3 M sodium acetate (NaOAc), pH 5.5, RNase-free (Thermo Fisher Scientific, cat. no. AM9740)
GlycoBlue (Thermo Fisher Scientific, cat. no. AM9516)
Ethanol, 200 proof, USP grade (Koptec, cat. no. V1016TP or equivalent)
RNA loading dye (Thermo Fisher, cat. no. AM8547)
15% precast TBE-urea gels (Thermo Fisher, cat. no. EC68855BOX)
10× TBE buffer (Bio-Rad, cat. no. 1610770 or equivalent)
SYBR Gold Nucleic Acid Gel Stain (Thermo Fisher Scientific, cat. no. S11494)

PCR tubes
Low-retention, filtered pipet tips
Thermocycler (Bio-Rad C1000, T100 or similar)
1.5-ml low-binding tubes (DNA LoBind, Eppendorf, cat. no. 022431021)
Nanodrop spectrophotometer

Additional reagents and equipment for denaturing PAGE and staining

Prepare adenylated adapter stock

1. Prepare 50 µl adenylation master mix in a PCR tube on ice:

   2.5 µl 100 µM 5′ phosphorylated adapter
   5 µl 10× 5′ DNA adenylation reaction buffer
   5 µl 1 mM ATP
   5 µl Mth RNA ligase
   32.5 µl water

   The reaction may be scaled up proportionally to increase yield.

2. Run the following program in a thermocycler with lid heated to 95°C:

   1 hr at 65°C
   5 min at 85°C
   Hold at 4°C

3. Combine the following in a 1.5-ml low-binding tube (total 200 µl):

   50 µl completed adenylation reaction
   130 µl water
   20 µl 3 M NaOAc, pH 5
   1 µl GlycoBlue

4. Add 500 µl ice-cold 100% ethanol and vortex to mix.

5. Incubate overnight at −80°C.

   This marks a stopping point.

6. Centrifuge samples for 30 min at max speed at 4°C.

7. Remove supernatant and add 1 ml freshly prepared ice-cold 80% ethanol.

8. Spin samples for 5 min at max speed at 4 °C.

9. Carefully remove supernatant.

10. Repeat steps 7-9 for a total of two washes.

    Remove as much ethanol as possible after the second wash, using a 1000-µl pipet tip and then a 200-µl tip. Perform a quick spin then remove any remaining ethanol with a 20- or 10-µl tip.

11. Air dry pellet for ∼5 min at room temperature.

12. Resuspend pellet with 25 µl water.

    For scaled-up reactions, pool suspensions into a single tube.

    If insoluble material is present after resuspending the adapter, centrifuge sample(s) for 5 min at max speed at 4°C and transfer the supernatant to a new tube.

13. Measure concentration of the adapter stock using a microvolume spectrophotometer (e.g., Nanodrop).

14. Calculate the adapter stock molarity.

    The molecular weight of adenylated adapter (App-VRA3-6N) = 10,640 Da. The micromolarity of the oligo = (ng/µl RNA × 1000)/10,640.

15. If necessary, normalize the adapter stock concentration to 10 µM.

    This is the prepared adenylated 3′ adapter stock.

16. Store in 20-µl aliquots at −20°C (short term, ∼2 months) or −80°C (longer term).

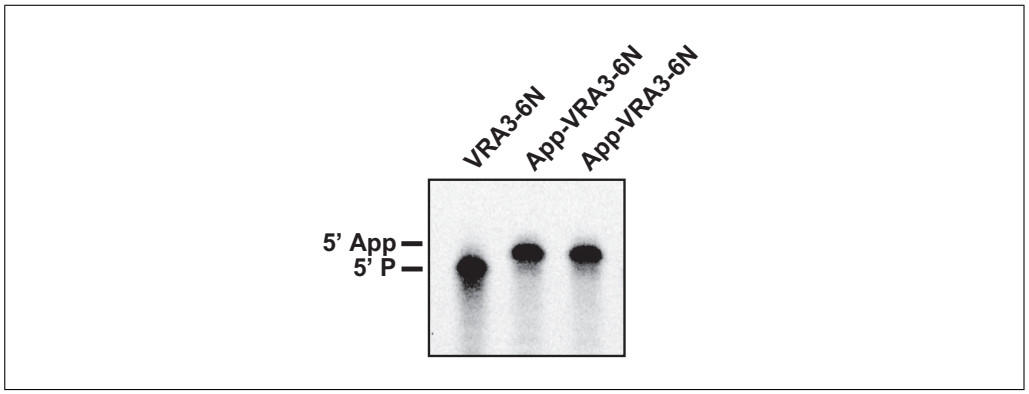

Figure 3 Result of 3′ adapter adenylation. Denaturing PAGE gel stained with SYBR Gold depicting the unadenylated 3′ adapter (VRA3-6N) and two independent batches of adenylated 3′ adapter (App-VRA3-6N). The migration of 5′ phosphorylated (5′ P) and 5′ adenylated (5′ App) species are labeled.

Aliquots should be discarded after two to three freeze-thaw cycles. The aliquot size can be adjusted accordingly to avoid waste.

Confirm adenylation by denaturing PAGE (strongly recommended)

17. Prepare 10 µl each of 2 µM adenylated adapter and 2 µM 5′ phosphorylated adapter.

    If available, we recommend including adenylated oligo from a previously validated adenylation batch as a positive control.

18. Combine 1 µl of each 2 µM oligo with 4 µl RNA loading dye. Place samples on ice.
19. Set up a precast Novex 15% TBE-urea gel (10 or 15 wells) in a gel tank with 1× TBE running buffer, per manufacturer instructions.
20. Use a 1000-µl pipet to flush any unpolymerized acrylamide out of the wells.
21. Pre-run the gel at 200 V for 10-20 min.

    This will warm the gel to maintain denaturing conditions during the run and further remove unpolymerized material from the gel matrix.

22. A few minutes before the pre-run is finished, heat samples at 80°C for 2 min, then place at room temperature.
23. Stop the pre-run and flush the wells again with a 1000-µl pipet to remove the urea that diffused into the wells.
24. Carefully load each sample (5 µl) using a 10-µl pipet tip.

    Minimize the vertical height of the sample in the well by dispensing slowly at the well bottom.

25. Run the gel at 80 V for 10 min, then at 200 V until the bromophenol blue tracking dye has exited the gel (~60-90 min).
26. Prepare 10 ml SYBR Gold staining solution by adding 1 µl SYBR Gold stock to 10 ml of 1× TBE.

    If the solution is made during the gel run, it should be stored in the dark until needed.

27. Carefully disassemble the gel, keeping track of the sample orientation, and place the gel into the staining solution.
28. Cover the gel with foil and agitate or rock gently for 10 min.
29. Image the gel.

    The adenylated oligo is ~1 nt longer than the unadenylated oligo and will migrate slightly higher up the gel, as shown in Figure 3.

REAGENTS AND SOLUTIONS

Betaine binding buffer

1.46 g betaine (Sigma-Aldrich, cat. no. 61962, final ~1.25 M)
10 ml binding buffer (see recipe)

Transfer 1.46 g betaine to a 15-ml conical tube and add 10 ml binding buffer. Vortex until betaine is fully dissolved. Filter sterilize with a 0.2-µm filter. Store up to several months at room temperature.

Binding buffer

Prepare using nuclease-free water
10 mM 1 M Tris·Cl, pH 7.4 (VWR, cat. no. 1185-53-1)
300 mM NaCl
0.1% Triton X-100 (Sigma-Aldrich, cat. no. T8787)
1 mM EDTA
Store at room temperature or 4°C (stable for many months)
Before use, add 0.2 µl SUPERase·In per ml

Aliquots containing SUPERase·In should be kept on ice and discarded at the end of the day.

Buffer F (freeze)

Prepare using nuclease-free water
50 mM Tris·Cl, pH 8.0 (Fisher Scientific, cat. no. 15568025)
40% (v/v) glycerol
5 mM $MgCl_2$
1.1 mM EDTA
Filter sterilize using a 0.2-µm filter into RNase-free plastic bottles
Store up to several months at 4°C
Before use, add:
0.5 mM DTT
SUPERase·In to 1 µl/ml final

Aliquots containing DTT and inhibitors should be discarded after use.

Buffer P (permeabilize)

Prepare buffer W (see recipe) with:
0.1% (v/v) Igepal CA-630
0.05% (v/v) Tween-20

Buffers P and W are identical except for the addition of detergents for permeabilization. It may be convenient to make a large stock of buffer W, split into two aliquots, and add the detergents to one.

Buffer W (wash)

Prepare using nuclease-free water
10 mM Tris·Cl, pH 8.0 (Fisher Scientific, cat. no. 15568025)
10 mM KCl
250 mM sucrose
5 mM $MgCl_2$
1 mM EGTA
10% (v/v) glycerol
Filter sterilize using a 0.2-µm filter into RNase-free plastic bottles
Store up to several months at 4°C
Before use, add:

0.5 mM DTT
SUPERase·In to 0.2 µl/ml final

100% Tween-20 and 100% Igepal CA-630 stocks are viscous and slow to dissolve in aqueous solutions. It may be convenient to first prepare 10% stock solutions in water. A positive displacement pipet such as an Eppendorf repeater with Combi-tips can help dispense accurately and completely. Roll or invert the stock solution until all the detergent is fully dissolved, then filter sterilize through a 0.2-µm membrane.

Aliquots containing DTT and inhibitors should be discarded after use.

Decon solution

Prepare using nuclease-free water
0.1 M NaOH
100 mM NaCl
Store single-use aliquots up to several months at −20°C

Fragmentation buffer, 5×

Prepare using nuclease-free water
375 mM Tris·Cl, pH 8.3 (Teknova, T1083)
562.5 mM KCl
22.5 mM $MgCl_2$
Store indefinitely at room temperature.

High-salt wash buffer

Prepare using nuclease-free water
50 mM 1 M Tris·Cl, pH 7.4 (VWR, cat. no. 1185-53-1)
2 M NaCl
0.5% Triton X-100 (Sigma-Aldrich, cat. no. T8787)
1 mM EDTA
Store at room temperature or 4°C (stable for many months)
Before use, add 0.2 µl SUPERase·In per ml

Aliquots containing SUPERase·In should be kept on ice and discarded at the end of the day.

Low-salt wash buffer

Prepare using nuclease-free water
5 mM Tris·Cl, pH 7.4 (VWR, cat. no. 1185-53-1)
0.1% Triton X-100
1 mM EDTA
Store at room temperature or 4°C (stable for many months)
Before use, add 0.2 µl SUPERase·In per ml

Aliquots containing SUPERase·In should be kept on ice and discarded at the end of the day.

COMMENTARY

Background Information

PRO-seq is performed to map the position of engaged RNAPII at single-nucleotide resolution. This technique is often applied to evaluate transcriptional dynamics under baseline conditions, across development, after cell perturbations, etc. Some of the major advantages to this technique are (1) stringent enrichment of nascent RNAs with minimal non-nascent or other contaminants such as those commonly observed with chromatin-associated RNA sequencing methods, and (2) the ability to derive the strand-specific, single-nucleotide position of RNAPII from relatively small numbers of cells (see below). Since PRO-seq is dependent on the incorporation of a biotinylated NTP by RNAPII, the position of arrested and backtracked polymerases, where the RNA

3′ end is not in the polymerase active site for NTP incorporation, will not be captured. However, sites of prominent transcriptional backtracking and arrest will result in a loss of PRO-seq signal, so it may be possible to infer that these processes have occurred.

Critical Parameters

Cell permeabilization

High-quality permeabilized cells are the single most important factor for obtaining good PRO-seq data. Permeabilization can be influenced by cell type, cell handling, temperature, and time of exposure to detergents. These must be controlled as precisely and reproducibly as possible. Although Basic Protocol 1 is effective for many cell types, the conditions were optimized using selected lines of mammalian or insect cells in culture, so additional optimization may be required for other cell types or tissues. Notably, centrifugation speeds may need to be altered based on cell delicacy to avoid bursting or clumping of cells. If other optimization steps fail to yield high-quality cells, bovine serum albumen (BSA) may be added to buffers W and P at a final concentration of 0.1-0.5%. BSA itself does not inhibit nuclear run-on reactions, but some BSA preparations contain substantial amounts of nucleic acids or nucleases that could lead to failed libraries. The BSA must therefore be of highly purified, molecular biology grade (fraction V or similar) and certified nuclease-free. Lastly, low-quality spike stocks are especially pernicious, as problems with spiking can only be detected by sequencing. Great care must be taken to ensure that spike stocks are of the highest quality.

Clumping

Cell clumps and debris are common problems that must be resolved before starting Basic Protocol 2. Critically, cell clumps and debris often result in inaccurate cell counts. This will lead to inaccurate addition of spike-in cells prior to the run-on reaction and confound downstream data normalization. If cell clumps cannot be resolved by resuspending the sample, we recommend the following changes to Basic Protocol 1. (1) When resuspending cells in buffer F, pipet the buffer onto the pellet and incubate for 1 min without mixing. Then use a 200-μl tip (not 1000-μl) to resuspend the disc-like pellet, being careful to not let it fold over onto itself, as this will generate large clumps that cannot be resuspended. (2) It may be that the spin speeds are too harsh for the cells of interest. In this case, we recommend reducing the *g* force and increasing the spin time during Basic Protocol 1, steps 8 and 12.

Percent permeabilization

The permeabilization protocol typically yields >99% permeabilization regardless of cell type. Samples with low permeabilization should be treated with suspicion, as this may reflect technical problems in permeabilization, e.g., rough handing, sample aeration, or errors in counting. If cells are not thoroughly permeabilized, endogenous NTPs from intact cells may compete with biotinylated NTPs during the nuclear run-on and lower the resolution of RNAPII mapping. Furthermore, since downstream data analysis relies on each sample containing the same number of permeabilized cells in the run-on, variable permeabilization among samples will complicate the downstream data processing and interpretation. Generally, samples with ≤90% permeabilization are not optimal for PRO-seq and should be prepared fresh.

Yield of permeabilized cells

Yield, calculated as the total number of permeabilized cells compared to the starting number of cells, is often a priority when cells are precious or if yields are so low that you cannot practically generate enough permeabilized cells to perform the experiment. If yields are low, the first step is to identify the step(s) where cells are lost in Basic Protocol 1. At each step, we recommend taking aliquots of the supernatants and cell resuspensions for cell counting.

Basic Protocol 1 uses 50-ml conical tubes for the high-volume steps (PBS, buffers W and P) and 1.5-ml tubes for the low-volume steps (buffer F). However, 15-ml conical tubes or 5-ml low-binding tubes (Eppendorf, cat. no 0030108310 or equivalent) may improve yields, especially when starting with a limited number of cells. For 5-ml tubes, reduce the volume of buffers W and P by half. Depending on your cells, this may also require scaling down the number of input cells per tube.

Lastly, keep in mind that yield and the tendency for cells to clump are affected by some of the same parameters. If it is not feasible to obtain both high yield and high quality, it is better to have a lower yield of high-quality cells than a higher yield of low-quality cells.

Batch effects

The steps between buffer P addition and buffer F addition must be performed as

consistently as possible across all samples during Basic Protocol 1, steps 7-11. Ideally, all samples should be prepared by the same person using the same buffer stock solutions.

Run-on cell inputs

For most cell types, Basic Protocol 2 is robust with inputs of 0.5–1 $\times$ 10^6 permeabilized cells. Libraries can be constructed with 1×10^5 or fewer permeabilized cells (Judd et al., 2020; Mahat et al., 2016), but we recommend determining the minimum input for experiments empirically. For example, we might build libraries in parallel starting from 5×10^5, 2.5×10^5, and 1×10^5 of the cells being tested. Importantly, when using fewer than 1×10^6 cells, the amount of 5′ and 3′ adapters should be reduced proportionally (i.e., for 5×10^5 cells, use 0.25 μM of each adapter instead of 0.5 μM). If your cells of interest have characteristically low transcriptional activity, more than 1×10^6 cells per sample may be necessary for each library construction. In this case, we recommend decreasing the percentage of spike cells to avoid over-representation of spike reads in the final libraries. For example, 2×10^6 low-activity cells would be spiked with 5×10^4 spike cells (2.5%) to maintain an ∼5% spike return in the final data.

Run-on timing

Precise timing of the run-on assay is critical for reproducible results. Before starting the run-on, make sure all materials are within easy reach of the work area and you have a printed timing chart corresponding to the number of samples you are processing.

RNA fragmentation

Nascent RNAs will vary in size depending on the position of RNAPII, i.e., promoter proximally paused RNAPII yields short nascent RNAs (∼20-50 nt), whereas RNAPII in the downstream gene body yields longer nascent RNAs (Fig. 1). The goal of the fragmentation is to truncate the longer nascent RNAs to the same size range as the shorter transcripts. If the longer transcripts are not fragmented sufficiently, they will be under-represented in the final library due to amplification bias favoring small amplicons, and will cluster poorly on the sequencer flow cell due to inefficient bridge amplification. In contrast, over-fragmented libraries will bias against shorter transcripts as fragmentation of short nascent RNAs will yield products that are too short to map uniquely to the reference genome.

In our experience, the 5-min fragmentation time suggested in Basic Protocol 2 reproducibility fragments RNAs for a wide variety of cell types. If, however, the final library size distribution is bimodal or skewed toward larger or smaller inserts, we recommend optimizing the fragmentation time by building test libraries using different fragmentation times.

When optimizing fragmentation, bear in mind that over-hydrolyzed samples will not yield high-quality PRO-seq libraries, as a large proportion of sequencing reads will either fail quality filters or fail to map uniquely to the reference genome. It may be more practical to reconstruct libraries with a reduced fragmentation time than sequence low-quality libraries. Notably, RNA degradation at any point prior to the 5′ adapter ligation will also display an over-fragmented size distribution. It is critical to practice proper RNase-free techniques throughout the protocol.

RNA precipitation and 3′ adapter ligation

The efficiency of the 3′ ligation is dramatically reduced by organic solvents, chaotropic salts, or ethanol remaining in the RNA pellet. These contaminants can be detected by an altered appearance of the RNA-glycogen pellet. Carefully observe the morphology and behavior of each pellet before proceeding to the 3′ adapter ligation (see Basic Protocol 2, step 56). If there is any uncertainty regarding pellet quality, reprecipitate.

Troubleshooting

Table 1 describes common problems a user can encounter when performing cell permeabilization or constructing PRO-seq libraries, along with potential causes and solutions.

Statistical Analysis

All custom scripts described herein for PRO-seq data processing are available on the Adelman Lab GitHub (*https://github.com/AdelmanLab/NIH_scripts*).
Dual, 6nt UMIs were extracted from read pairs using UMI-tools [10.1101/gr.209601.116]. Read pairs were trimmed using cutadapt 1.14 to remove adapter sequences (-O 1 –match-read-wildcards -m 26). The UMI length was trimmed off the end of both reads to prevent read-through into the mate's UMI, which will happen for shorter fragments. An additional nucleotide was removed from the end of read 1 (R1), using seqtk trimfq (*https://github.com/lh3/seqtk*) to preserve a single mate orientation during alignment. The paired end reads were then mapped to a

Table 1 Troubleshooting Guide for Cell Permeabilization and Library Construction

Problem	Possible cause	Solution
Cell permeabilization (Basic Protocol 1)		
Low yield	Cell handling	Take aliquots after each suspension step and count to calculate step yields.
Cell Clumping	Excessive *g* force	Reduce g-force and increase spin time.
	Inherently sticky or clumpy cells	Incubate samples in buffer F for 1 min prior to resuspension; Resuspend samples with a P200; Use a cell strainer prior to permeabilization
Variable Permeabilization Between Sample Batches	Delays in sample processing	Work as quickly and as gently as possible during steps 7-11; Process fewer samples at a time; Keep buffer stocks consistent
PRO-seq library construction (Basic Protocol 2)		
No library	Low cell quality/quantity, low transcriptional activity	Increase the number of input cells and/or concentration of biotin-11-NTPs.
	Inefficient 3′ adapter ligation step due to trace organics or chaotropes in RNA pellet.	Reprecipitate samples prior to the 3′ adapter ligation
	RNA degradation	Test reagents for RNase contamination; Replace reagents and decontaminate workstation.
	Low enzymatic activity	Replace expired enzymes; Test each enzyme's activity and replace as necessary.
	RNA was over-fragmented	Optimize fragmentation time.
Low library yield	Under-amplification	Amplify the final library with an additional 2 cycles and repurify.
	ProNex bead mishandling; overdried beads or residual ethanol	Reduce drying time; Remove residual ethanol before elution.
Adapter dimers in final library	Poor run-on efficiency	Confirm quality of permeabilized cells and that run-on reaction was performed exactly as described.
	Low cell input/transcriptional activity	
	Inefficient 3′ adapter ligation step due contaminants in RNA pellet.	Increase the number of input cells and/or the concentration of biotin-11-NTPs.
	Cumulative inefficiencies and errors	Reprecipitate samples prior to 3′ adapter ligation
		Avoid bead loss during washes, errors in enzyme reaction setup, and carrying over wash buffer into downstream enzymatic reactions.
Variable final library sizing among samples	Inconsistent fragmentation setup; Uneven final concentrations of $MgCl_2$	Manually check all sample volumes after the Norgen RNA purification. If necessary, add water to bring the final volume to 100 µl for each sample.
Small inserts	Over-fragmentation	Optimize fragmentation time.
	RNA degradation	Practice proper RNase-free technique and decontaminate workstation. If necessary, increase concentration of SUPERase·In in buffers.

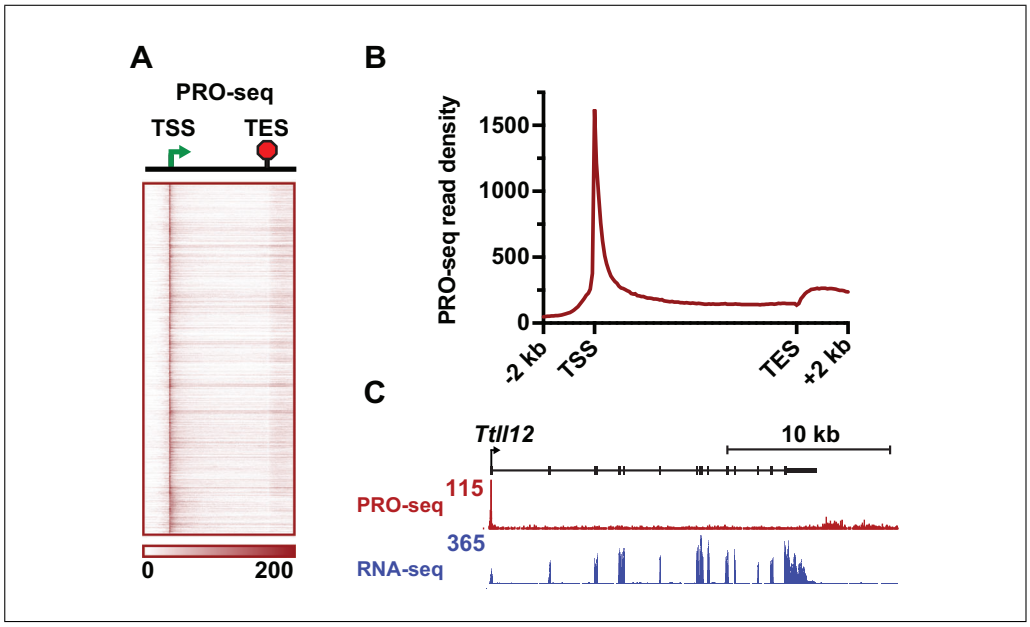

Figure 4 Examples of PRO-seq data representation. (**a**) Representative heatmap of PRO-seq read density at active genes ($N = 12{,}598$; >1 kb in length) from 2 kb upstream of the TSS to 2 kb downstream of the TES. The region between the TSS and TES was scaled by gene length into 100 bins and the signal in each bin was converted to reads per kilobase for visualization. (**b**) Data in A shown as a gene average metagene plot. (**c**) PRO-seq and RNA-seq for an example gene, *Ttll12*.

combined genome index, including both the spike genome build and primary genome build genomes using bowtie2 [10.1038/nmeth.1923]. Properly paired reads were retained. These read pairs were then separated based on the genome to which they mapped (i.e., spike-in vs. primary), and both these spike and primary reads were independently deduplicated, again using UMI-tools. Reads mapping to the reference genome were separated according to whether they were R1 or R2, sorted via samtools 1.3.1 (-n), and subsequently converted to bedGraph format using a custom script (bowtie2stdBedGraph.pl). We note that this script counts each read once at the exact 3′ end of the nascent RNA. Because R1 in PRO-seq reveals the position of the RNA 3′ end, the + and − strands were swapped to generate bedGraphs representing 3′ end positions at single-nucleotide resolution.

For downstream statistical analysis, PRO-seq is compatible with most of the common tools used for other NGS techniques to evaluate changes in gene expression (e.g., DEseq2, EdgeR). PRO-seq can also be used to generate a refined gene annotation (custom script, get_gene_annotations.sh available on the Adelman Lab GitHub at *https://github.com/AdelmanLab/GetGeneAnnotation_GGA*) and report enhancer location and activity (Danko et al., 2015; Wang et al., 2019).

Understanding Results

As mentioned above, PRO-seq data reports the 3′ end position of the read, which represents the position of the active site of engaged RNAPII at single-nucleotide resolution. We first recommend evaluating the distribution of PRO-seq signals across genes, for example, as a heatmap (Fig. 4a), metagene plot (Fig. 4b), and/or at individual example genes (Fig. 4c). We expect to observe an enrichment of signal ∼25-50 nt downstream of the transcription start site (TSS); this corresponds to the promoter proximal pause and release into productive elongation (as reviewed in Core & Adelman, 2019). Additionally, PRO-seq signal increases downstream of the transcript end site (TES; Fig. 4a-c) due to an increase in the residence time of RNAPII at each nucleotide position, in agreement with reports of slower elongation after cleavage of the nascent RNA by the cleavage and polyadenylation machinery at TESs (Core et al., 2008; Muniz et al., 2021; Schwalb et al., 2016).

Time Considerations

All potential stopping points and storage conditions are listed in the corresponding protocols. Below, is a general timeline for each Basic Protocol, but actual times will vary with the number of samples being processed.

The time required for cell permeabilization (Basic Protocol 1) is 30-60 min. PRO-seq (Basic Protocol 2) is typically performed in 3-4 days. The first day (run-on to 3′ adapter ligation) takes ∼7 hr, second day (second streptavidin binding to preliminary amplification) takes ∼8 hr, and the final library processing (test amplification, final amplification, qPCR, and library pooling) can be done flexibly in 1-2 days with a total processing time of ∼5 hr.

Acknowledgments

We thank K. Adelman and all members of the Adelman Lab and the Nascent Transcriptomics Core for helpful discussions on the protocol. We thank T. Henriques, E.F. Ablondi, M. Jones, and Z. Best for discussions and assistance on method optimization and G. Nelson for advice on data analysis and pipeline development. We additionally thank the HMS Biopolymers Facility and Bauer Core Facility at Harvard University for sequencing services. This work was supported by the National Institutes of Health (NIH R01GM139960 to K. Adelman). The Nascent Transcriptomics Core is grateful for support from the Harvard Medical School Foundry and the Dean's Epigenetics and Gene Dynamics Initiative. C.A.M was supported by the National Science Foundation Graduate Research Fellowship (DGE1745303) and the Sophia H. Y. Chang Fellowship.

Author Contributions

Claudia A. Mimoso: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing—original draft; writing—review and editing. **Seth R. Goldman:** Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing—original draft; writing—review and editing.

Conflict of Interest

C.A.M and S.R.G report no conflicts of interest.

Data Availability Statement

The data, tools, and materials (or their sources) that support the protocol are available from the authors upon reasonable request.

Supporting Information

cpz1961-sup-0001-TableS1.pdf
Table S1. Oligos used in this protocol

Literature Cited

Anastasiadou, E., Jacob, L. S., & Slack, F. J. (2017). Non-coding RNA networks in cancer. *Nature Reviews Cancer*, *18*, 5–18. https://doi.org/10.1038/nrc.2017.99

Bradner, J. E., Hnisz, D., & Young, R. A. (2017). Transcriptional addiction in cancer. *Cell*, *168*, 629–643. https://doi.org/10.1016/j.cell.2016.12.013

Chang, C.-H., & Luse, D. S. (1997). The H3/H4 tetramer blocks transcript elongation by RNA polymerase II in vitro. *Journal of Biological Chemistry*, *272*, 23427–23434. https://doi.org/10.1074/jbc.272.37.23427

Core, L., & Adelman, K. (2019). Promoter-proximal pausing of RNA polymerase II: A nexus of gene regulation. *Genes & Development*, *33*, 960–982. 10.1101/gad.325142.119

Core, L. J., Waterfall, J. J., Gilchrist, D. A., Fargo, D. C., Kwak, H., Adelman, K., & Lis, J. T. (2012). Defining the status of RNA polymerase at promoters. *Cell Reports*, *2*, 1025–1035. https://doi.org/10.1016/j.celrep.2012.08.034

Core, L. J., Waterfall, J. J., & Lis, J. T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science*, *322*, 1845–1848. https://doi.org/10.1126/science.1162228

Danko, C. G., Hyland, S. L., Core, L. J., Martins, A. L., Waters, C. T., Lee, H. W., Cheung, V. G., Kraus, W. L., Lis, J. T., & Siepel, A. (2015). Identification of active transcriptional regulatory elements from GRO-seq data. *Nature Methods*, *12*, 433–438. https://doi.org/10.1038/nmeth.3329

Dhanasekaran, R., Deutzmann, A., Mahauad-Fernandez, W. D., Hansen, A. S., Gouw, A. M., & Felsher, D. W. (2022). The MYC oncogene—The grand orchestrator of cancer growth and immune evasion. *Nature Reviews Clinical Oncology*, *19*, 23–36. https://doi.org/10.1038/s41571-021-00549-2

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., Epstein, C. B., Frietze, S., Harrow, J., Kaul, R., Khatun, J., Lajoie, B. R., Landt, S. G., Lee, B. K., Pauli, F., Rosenbloom, K. R., Sabo, P., Safi, A., Sanyal, A., … The ENCODE Project Consortium. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*, 57–74. https://doi.org/10.1038/nature11247

Fu, Y., Wu, P. H., Beane, T., Zamore, P. D., & Weng, Z. (2018). Elimination of PCR duplicates in RNA-seq and small RNA-seq using unique molecular identifiers. *BMC Genomics*, *19*, 531. https://doi.org/10.1186/s12864-018-4933-1

Fuchs, R. T., Sun, Z., Zhuang, F., & Robb, G. B. (2015). Bias in ligation-based small RNA sequencing library construction is determined by adaptor and RNA structure. *PLoS ONE*, *10*, e0126049. https://doi.org/10.1371/journal.pone.0126049

Futreal, P. A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Rahman, N., &

Stratton, M. R. (2004). A census of human cancer genes. *Nature Reviews Cancer*, *4*, 177–183. https://doi.org/10.1038/nrc1299

Gariglio, P., Buss, J., & Green, M. H. (1974). Sarkosyl activation of RNA polymerase activity in mitotic mouse cells. *FEBS Letters*, *44*, 330–333. https://doi.org/10.1016/0014-5793(74)81170-1

Hafner, M., Renwick, N., Brown, M., Mihailović, A., Holoch, D., Lin, C., Pena, J. T. G., Nusbaum, J. D., Morozov, P., Ludwig, J., Ojo, T., Luo, S., Schroth, G., & Tuschl, T. (2011). RNA-ligase-dependent biases in miRNA representation in deep-sequenced small RNA cDNA libraries. *RNA*, *17*, 1697–1712. https://doi.org/10.1261/rna.2799511

Hawleys, D. K., & Roederfj, R. G. (1985). Separation and partial characterization of three functional steps in transcription initiation by human RNA polymerase II. *Journal of Biological Chemistry*, *260*, 8163–8172. https://doi.org/10.1016/S0021-9258(17)39577-7

Herzog, V. A., Reichholf, B., Neumann, T., Rescheneder, P., Bhat, P., Burkard, T. R., Wlotzka, W., von Haeseler, A., Zuber, J., & Ameres, S. L. (2017). Thiol-linked alkylation of RNA to assess expression dynamics. *Nature Methods*, *14*, 1198–1204. https://doi.org/10.1038/nmeth.4435

Hong, J., & Gresham, D. (2017). Incorporation of unique molecular identifiers in TruSeq adapters improves the accuracy of quantitative sequencing. *BioTechniques*, *63*, 221–226. https://doi.org/10.2144/000114608

Jonkers, I., Kwak, H., & Lis, J. T. (2014). Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife*, *3*, e02407. https://doi.org/10.7554/eLife.02407

Judd, J., Wojenski, L. A., Wainman, L. M., Tippens, N. D., Rice, E. J., Dziubek, A., Villafano, G. J., Wissink, E. M., Versluis, P., Bagepalli, L., Shah, S. R., Mahat, D. B., Tome, J. M., Danko, C. G., Lis, J. T., & Core, L. J. (2020). A rapid, sensitive, scalable method for precision run-on sequencing (PRO-seq). *bioRxiv*, 2020.05.18.102277. https://doi.org/10.1101/2020.05.18.102277

Kaikkonen, M. U., & Adelman, K. (2018). Emerging roles of non-coding RNA transcription. *Trends in Biochemical Sciences*, *43*, 654–667. https://doi.org/10.1016/j.tibs.2018.06.002

Kivioja, T., Vähärautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S., & Taipale, J. (2012). Counting absolute numbers of molecules using unique molecular identifiers. *Nature Methods*, *9*, 72–74. https://doi.org/10.1038/nmeth.1778

Kwak, H., Fuda, N. J., Core, L. J., & Lis, J. T. (2013). Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science*, *339*, 950–953. https://doi.org/10.1126/science.1229386

Lee, T. I., & Young, R. A. (2013). Transcriptional regulation and its misregulation in disease. *Cell*, *152*, 1237–1251. https://doi.org/10.1016/j.cell.2013.02.014

Li, J., Liu, Y., Rhee, H. S., Ghosh, S. K. B., Bai, L., Pugh, B. F., & Gilmour, D. S. (2013). Kinetic competition between elongation rate and binding of NELF controls promoter-proximal pausing. *Molecular Cell*, *50*, 711–722. https://doi.org/10.1016/j.molcel.2013.05.016

Mahat, D., Kwak, H., Booth, G. T., Jonkers, I. H., Danko, C. G., Patel, R. K., Waters, C. T., Munson, K., Core, L. J., & Lis, J. T. (2016). Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). *Nature Protocols*, *11*(8), 1455–1476. https://doi.org/10.1038/nprot.2016.086

Malik, S., & Roeder, R. G. (2023). Regulation of the RNA polymerase II pre-initiation complex by its associated coactivators. *Nature Reviews Genetics*, *24*(11), 767–782. https://doi.org/10.1038/s41576-023-00630-9

Morris, K. V., & Mattick, J. S. (2014). The rise of regulatory RNA. *Nature Reviews Genetics*, *15*, 423–437. https://doi.org/10.1038/nrg3722

Munafó, D. B., & Robb, G. B. (2010). Optimization of enzymatic reaction conditions for generating representative pools of cDNA from small RNA. *RNA*, *16*, 2537–2552. https://doi.org/10.1261/rna.2242610

Muniz, L., Nicolas, E., & Trouche, D. (2021). RNA polymerase II speed: A key player in controlling and adapting transcriptome composition. *The EMBO Journal*, *40*, e105740. https://doi.org/10.15252/embj.2020105740

Nojima, T., Gomes, T., Grosso, A. R. F., Kimura, H., Dye, M. J., Dhir, S., Carmo-Fonseca, M., & Proudfoot, N. J. (2015). Mammalian NET-Seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell*, *161*, 526–540. https://doi.org/10.1016/j.cell.2015.03.027

Roeder, R. G. (2019). 50+ years of eukaryotic transcription: An expanding universe of factors and mechanisms. *Nature Structural & Molecular Biology*, *26*, 783–791. https://doi.org/10.1038/s41594-019-0287-x

Rougvie, A. E., & Lis, J. T. (1988). The RNA polymerase II molecule at the 5′ end of the uninduced *hsp70* gene of *D. melanogaster* is transcriptionally engaged. *Cell*, *54*, 795–804. https://doi.org/10.1016/S0092-8674(88)91087-2

Saunders, K., Bert, A. G., Dredge, B. K., Toubia, J., Gregory, P. A., Pillman, K. A., Goodall, G. J., & Bracken, C. P. (2020). Insufficiently complex unique-molecular identifiers (UMIs) distort small RNA sequencing. *Scientific Reports*, *10*, 14593. https://doi.org/10.1038/s41598-020-71323-0

Schwalb, B., Michel, M., Zacher, B., Hauf, K. F., Demel, C., Tresch, A., Gagneur, J., & Cramer, P. (2016). TT-seq maps the human transient transcriptome. *Science*, *352*, 1225–1228. https://doi.org/10.1126/science.aad9841

Shmookler, R. J., Buss, J., & Green, M. H. (1974). Properties of the polyoma virus transcription complex obtained from mouse nuclei. *Virology*, *67*, 122–127. https://doi.org/10.1016/0042-6822(74)90113-5

Smith, T., Heger, A., & Sudbery, I. (2017). UMI-tools: Modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. *Genome Research*, *27*, 491–499. https://doi.org/10.1101/gr.209601.116

Sur, I., & Taipale, J. (2016). The role of enhancers in cancer. *Nature Reviews Cancer*, *16*, 483–493. https://doi.org/10.1038/nrc.2016.62

van der Lee, R., Correard, S., & Wasserman, W. W. (2020). Deregulated regulators: Disease-causing cis variants in transcription factor genes. *Trends in Genetics*, *36*, 523–539. https://doi.org/10.1016/j.tig.2020.04.006

Vervoort, S. J., Devlin, J. R., Kwiatkowski, N., Teng, M., Gray, N. S., & Johnstone, R. W. (2021). Targeting transcription cycles in cancer. *Nature Reviews Cancer*, *22*, 5–24. https://doi.org/10.1038/s41568-021-00411-8

Viollet, S., Fuchs, R. T., Munafo, D. B., Zhuang, F., & Robb, G. B. (2011). T4 RNA ligase 2 truncated active site mutants: Improved tools for RNA analysis. *BMC Biotechnology*, *11*, 72. https://doi.org/10.1186/1472-6750-11-72

Vogt, P. K. (2001). Jun, the oncoprotein. *Oncogene*, *20*(19), 2365–2377. https://doi.org/10.1038/sj.onc.1204443

Wang, Z., Chu, T., Choate, L. A., & Danko, C. G. (2019). Identification of regulatory elements from nascent transcription using dREG. *Genome Research*, *29*, 293–303. https://doi.org/10.1101/gr.238279.118

Wissink, E. M., Vihervaara, A., Tippens, N. D., & Lis, J. T. (2019). Nascent RNA analyses: Tracking transcription and its regulation. *Nature Reviews. Genetics*, *20*, 705. https://doi.org/10.1038/s41576-019-0159-6

Zhuang, F., Fuchs, R. T., Sun, Z., Zheng, Y., & Robb, G. B. (2012). Structural bias in T4 RNA ligase-mediated 3′-adapter ligation. *Nucleic Acids Research*, *40*, e54. https://doi.org/10.1093/nar/gkr1263