Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

APEX-seq: RNA subcellular localization by proximity labeling

Furqan M. Fazal (ffazal08@gmail.com)
Stanford University https://orcid.org/0000-0003-0758-4582
Shuo Han
Stanford University
Kevin R. Parker
Stanford University
Pornchai Kaewsapsak
Stanford University
Jin Xu
Stanford University
Alistair Boettiger
Stanford University
Howard Y. Chang (howchang@stanford.edu)
Stanford University
Alice Y. Ting (ayting@stanford.edu)
Stanford University

Keywords: APEX-seq, promixity labeling, vicinal, RNA-seq, sequencing, subcellular, RNA, localization, spatial, transcriptomics

Posted Date: May 15th, 2019

DOI: https://doi.org/10.21203/rs.2.1857/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

We introduce APEX-seq, a method for RNA sequencing based on direct proximity labeling of RNA using the peroxidase enzyme APEX2. APEX-seq in nine distinct subcellular locales produced a nanometer-resolution spatial map of the human transcriptome as a resource, revealing extensive and exquisite patterns of localization for diverse RNA classes and transcript isoforms. APEX-seq should be widely applicable to many systems, enabling comprehensive investigations of the spatial transcriptome.

Introduction

General notes:

1. The APEX-seq protocol takes roughly 4 days to complete:

    Day 1: Labeling of cells and extraction of RNAs

    Day 2: Enrichment for biotinylated RNAs

    Days 3-4: Preparation of sequencing libraries

2. This protocol assumes cells containing APEX-fusion proteins have been prepared, and the localization of the fusion construct has been confirmed by microscopy. See Hung *et al.*[1] for details on designing APEX-fusion constructs, transfecting constructs into cells, and imaging cells.

3. Use proper techniques for handing RNA, and avoiding contamination with RNases. These include using a dedicated bench area for RNA, using filtered pipette tips, wiping all surfaces with RNase Zap (Invitrogen), using certified RNase-free buffers and reagents, and testing buffers for RNase contamination using RNase Alert (Ambion). When appropriate, add ~1-2 µL of Ribolock RNase inhibitor (Thermo Fischer) per 100-200 µL of buffer/solution.

4. For all water, we use ultrapure distilled water (Invitrogen). However, any source of certified/verified RNase-free water may be used.

5. The current APEX-seq protocol recovers total RNA or polyA-selected RNAs. We select for RNAs longer than 100 nt. We have not tested modifications to these selection criteria.

6. All experiments were carried out at room temperature, unless otherwise mentioned.

Reagents

Equipment

Procedure

Day 0: Growing Cells

Details described below are for adherent human embryonic kidney (HEK293) cells. Modify as appropriate for cells from other tissues and/or organisms.

1. Prepare 10-cm plates for growing HEK293 cells by coating with fibronectin. To do so:

(a) Dilute 1 mg/mL fibronectin (EMD Millipore, Catalog No. FC010) by 100 fold in PBS.

(b) Filter solution using 0.10 µm filter (0.22 µm okay too; Steriflip EMD Millipore).

(c) Fully cover each 10-cm plate with ~3 mL fibronectin mix.

(d) Incubate in 37°C incubator for ~1 hour.

(e) Aspirate off remaining liquid, and add cells to plate.

We typically grow/expand cells without fibronectin, and only use fibronectin for plates that will be used for labeling.

2. Grow cells overnight at 37°C (5% CO_2) incubator to ~70-80% confluency. For HEK293 we use Gibco DMEM (Thermo Fischer) with 10% FBS (filtered) and 1% Gibco PenStrep (100X, Thermo Fischer). If cells grow too quickly, a 1:1 mix of DMEM: MEM may be used instead. However, the growth conditions and media should be kept constant for the project.

Note: For APEX-seq, we typically prepare 6 plates per condition – 3 plates for labeled “targets”, and 3 plates for unlabeled “controls”. Depending on the biological question, 4 plates (2 targets, 2 controls)

might be sufficient, though it's better to have multiple biological replicates (ideally 3+ each) for both targets and controls.

Day 1: Labeling Cells

1. Wash cells once with Gibco PBS (Thermo Fischer).

2. Incubate cells at 37°C for 30 minutes with 3 mL media (or PBS) containing 0.5 mM biotin-phenol (Iris Biotech GMBH). To do so:

(a) Prepare a 1000X biotin-phenol (BP) stock (500 mM) by adding 0.37 g BP to 2 mL DMSO. Any surplus stock can be aliquoted and stored in the -80°C for >6 months.

(b) Add 3 µL 1000X BP stock to 3 mL media (or PBS).

3. Label the target plates (but not the control plates) with 1 mM H_2O_2 for exactly 1 minute. To do so:

(a) Prepare a ~83X H_2O_2 stock solution by adding 10 µL 30% H_2O_2 solution (Sigma Aldrich, 8.8 M) to 1 mL water. Add 36 uL of this solution to the center of each target plate containing the BP solution. Shake the plate gently to mix and incubate for exactly 1 minute, followed by addition of 3 mL quenching solution.

(b) To prepare quenching solution, add 10 mM sodium ascorbate, 10 mM sodium azide and 5 mM Trolox to PBS.

(c) To prepare 1 M (100X) sodium ascorbate, dissolve 198 mg (Sigma Aldrich) in 1 mL water. For 1 M (100X) sodium azide, dissolve 65 mg (Sigma Aldrich) in 1 mL water. For 500 mM Trolox (100X), add 125 mg to 1 mL DMSO. Prepare these reagents fresh each time.

(d) To stop labeling after 1 minute treatment with H_2O_2, pour out BP solution containing H_2O_2 from 10-cm plates and immediately add ~3 mL quenching solution. Shake well and pour out quenching solution. Quickly add another ~3 mL quenching solution, shake well and pour out again. Finally add ~1 mL quenching solution and shake to cover cells. Be careful not to lose or detach cells during washing.

(e) Add ~4 uL RNase inhibitor to each plate.

(f) Proceed to RNA extraction as soon as possible. If additional time is required to label/process multiple plates, we keep the plates on ice in the meantime (no longer than 1 hour).

(g) To prepare control samples, exclude 1 minute H_2O_2 addition step and carry out all subsequent steps (including quenching-solution treatment) identically to target samples.

4. Scrape the cells carefully using cell lifters (Corning), and transfer contents of 10-cm plate (cells and residual quenching buffer) to 1.5 mL Eppendorf tubes. Transfer to benchtop centrifuge, and spin at ~300G for 3-5 minutes. Remove most of the solution using a pipette, keeping the last ~50-100 µL quenching solution in the tube.

Note: We want to inhibit any further reaction of biotin-phenol catalyzed by APEX2, so we keep cells in residual quenching solution until we have lysed cells.

5. Follow instructions to extract total RNA from animal cells using the RNeasy plus mini kit (Qiagen). A few notes:

(a) Add ß-mercaptoethanol to Buffer RLT plus, as recommended by the manufacturer (10 uL ß-ME per 1 mL Buffer).

(b) Use 600 uL Buffer RLT Plus per 10-cm plate.

(c) Homogenize the lysate carefully to prevent clogging the column, and transfer lysate to a gDNA eliminator spin column.

(d) Instead of using Buffer RW1, we use Buffer RWT (Qiagen).

(e) Carry out the optional step of centrifuging at full speed for 1 minute before eluting.

(f) Elute in 60-100 µL water. We usually elute by running 30-50 µL water twice through the column to maximize yield.

(g) After eluting into supplied 1.5 mL tubes, transfer RNA to a fresh 1.5 mL Eppendorf tube.

(h) Check concentration of RNA. We use the Nanodrop (Thermo Fischer) for quantification, and typically recover 60-100 ug RNA per plate.

6. Bioanalyze samples (or run a gel) to confirm RNA is intact. We aim for a RIN (RNA integrity number) > 8.5. Proceed with caution for RNA with RIN = 7.5 - 8.5. Check for RNase contamination and repeat

protocol if RNA is degraded (RIN <= 7.5). We use the RNA pico assay on the Agilent Bioanalyzer 2100 to determine RNA integrity.

7. Recovered RNA may be stored overnight at –20°C. However, beyond 1-2 days at –20°C, RNA should be stored at –80°C.

Note: For some cell lines and/or purposes, it may be necessary or desirable to have an additional step of DNA elimination. To do so:

(a) Incubate RNA with TURBO DNase (Thermo Fischer) for 30 minutes at 37°C following the manufacture's protocol.

(b) Purify RNA by sending through RNeasy plus mini kit.

Day 2: Enrichment of Biotinylated RNAs

Below are details to enrich for biotinylated RNAs. It is very important to maintain RNase-free conditions, as we typically do not carry out any further quality-control measure until we have sequencing libraries.

1. We usually use ~½ (~30-50 µg) of our recovered RNA from the previous day for enrichment, but no less than 25 µg. The remaining ½ is kept in case it is necessary to repeat downstream measurements, to make a technical replicate, or to generate libraries using different approaches (for example, ½ of original used to prepare total-RNA libraries, and the other half libraries from polyA-selected RNAs). Start by adding water to make the RNA volume ~125 µL.

Note: If carrying out APEX-seq experiment for first time, it is advisable to use >= 50 µg RNA.

2. Prepare streptavidin beads to enrich for biotinylated RNAs. To do so:

(a) We use 10 µL pierce streptavidin magnetic beads (Thermo Fischer) per 25 µg RNA. Adjust volume of beads appropriate to RNA amount. Mix beads well by vortexing the bottle prior to aliquoting.

(b) Add ~500 µL binding and washing (B&W) buffer (5 mM Tris-HCl, pH = 7.5; 0.5 mM EDTA; 1 M NaCl; 0.1% TWEEN 20 (Sigma Aldrich)) to beads in 1.5 mL Eppendorf tube.

(c) Put beads on magnet for 5 minutes. Discard supernatant and remove tubes from magnet.

(d) Resuspend beads in 500 µL B&W buffer, place on magnet for 2 minutes, and then discard supernatant.

(e) Repeat step (d) twice more, for a total of 3 washes.

(f) Repeat washing step twice more, but this time using solution A (100 mM NaOH, 50 mM NaCl) each time.

(g) Repeat washing step again, but this time with solution B (100 mM NaCl) once.

(h) Resuspend beads in 125 uL solution B containing ~3 uL Ribolock RNase inhibitor.

(i) Add 125 uL RNA from step 1 to ~125 uL beads and incubate tube on a rotator for 2 hours at 4°C.

Note: Do not rotate overnight – we have noticed a decreased yield upon overnight incubation.

3. After incubation, to get rid of unbound RNAs we carry out high-salt washes using B&W buffer. To do so:

(a) Place tubes from step 3 (j) on magnet for 5 minutes and discard supernatant.

(b) Remove tubes from magnet and resuspend beads in 500 µL B&W buffer.

(e) Repeat steps (a) and (b) above twice for a total of 3 washes.

(f) After the third wash, keep beads on magnet and remove any residual liquid remaining using a P20 (20 uL micropipette) tip.

(g) Resuspend beads in 54 uL water, transfer to a new 1.5 mL Eppendorf tube, and proceed to next step immediately.

4. To release biotinylated RNAs bound to streptavidin beads, incubate beads with proteinase K. To do so:

(a) Prepare 3X proteinase digestion buffer (1.1 mL) fresh as follows: 330 µL 10X PBS, pH = 7.4 (Ambion); 330 µL 20% N-laurylsarcosine sodium solution (Sigma Aldrich); 66 µL of 0.5 M ETDA; 16.5 uL of 1 M DTT; 357.5 uL water).

(b) For each reaction, mix 33 uL 3X proteinase digestion buffer, 10 µL proteinase K (20 mg/ml, Thermo Fischer), 3 µL Ribolock RNase inhibitor. Add this mix to the 54 uL beads (total volume = 100 µL).

(c) Incubate beads with vigorous mixing at 42°C for 1 hour, followed by 55°C for 1 hour. We incubate on a thermomixer while shaking at 900 rpm.

(d) After incubation, place tubes containing beads on magnet for 5 minutes. Carefully extract supernatant containing released RNAs, without disturbing beads. Place liquid into new 1.5 mL Eppendorf tube.

5. Purify recovered RNAs using RNA clean and concentrator -5 kit (Zymo Research). A few notes:

(a) Follow the manufacture protocol, but carry out an extra centrifuge step for 1 minute with no liquid added, before eluting RNA with water. This step ensures complete removal of the wash buffer.

(b) Elute in 6 µL water for downstream total RNA libraries, and 12 µL water for polyA-selected RNA libraries. For 12 µL volume, send 6 µL twice through column to increase yield.

6. Transfer RNA to 0.2 mL PCR tubes (Thermo Fischer). Store RNA overnight at –20°C, or at –80°C for longer duration.

Days 3-4: Preparation of TruSeq Libraries

1. Prepare RNA-seq libraries using the Illumina TruSeq stranded mRNA library prep kit (previously RS-122-2101/2102; recently Illumina has begun offering the library preparation kit, catalog number 20020594, and index adapters separately.). The kit allows preparation of libraries from total RNA, or alternatively polyA-selected RNA using the Oligo-dT beads (supplied). Follow the LS protocol to generate stranded RNA-seq libraries.

2. Note that the expected yield (~10 ng) from day 2 is less than the recommended amount for TruSeq, but we get high-quality libraries using the kit. Nonetheless we recommend following the guidelines carefully to minimize material loss.

3. For the first-strand synthesis mix, we use Superscript III (instead of Superscript II). So far we have not tested Superscript IV in the APEX-seq protocol.

4. When ligating barcodes to cDNA libraries, we recommend using 0.55 µL adapter index mix (and not the 2.5 µL recommended by protocol). Unreacted adaptor can lead to downstream issues.

5. Typical amplification is ~16-24 cycles, and the default 15 cycles recommended is often insufficient amplification. We aim for final concentrations of ~10-20 nM, as determined by bioanalyzer. However, we usually do not recover high-quality libraries from samples that require >25 cycles amplification. In our experience, the control samples need ~2-3 extra cycles of amplification relative to the labeled target samples.

6. To PCR all APEX-seq libraries to the same amount for convenient pooling, we recommend replacing the suggested Illumina 15-cycle PCR step with the following protocol:

(a) PCR 1: Add 20 µL DNA, 5 µL PCR Primer Cocktail, and 25 µL PCR Mix. Carry out 10 cycles of PCR following TruSeq recommendations: step 1: 98°C for 30 seconds; step 2: 10 cycles of 98°C 10 seconds, 60°C 30 seconds, 72°C 30 seconds; step 3: 72°C for 5 seconds; step 4: Hold at 4°C. Clean up following TruSeq protocol using 50 µL Ampure XP beads (Thermo Fischer), and finally elute in 22 µL buffer.

(b) To quantify libraries, carry out a qPCR step using 0.4 µL 10 µM Illumina general primers, 10 µL 2X SYBR green master mix (Thermo Fischer), 8.6 µL water, and 1 µL amplified DNA. qPCR this 20 µL mix for 30 cycles using the following settings: step 1 (preincubation): 95°C for 5 minutes; step 2 (amplification, x30 cycles): 95°C for 10 seconds, 60°C for 20 seconds, 72°C for 30 seconds; step 3 (melting curve): follow qPCR-machine-manufacture protocol; step 4 (cooling): hold at 37°C for 30 seconds. Carry out all qPCR measurements in duplicate or triplicate. The sequence of Illumina primers are as follows:

AATGATACGGCGACCACCGAGAT

CAAGCAGAAGACGGCATACGA

(c) From step (b), choose the cycle number that gives 1/6 maximum fluorescence (*N*), and subtract by two to determine the number of cycles (= $N - 2$) for amplification. We use the Lightcycler 480 (Roche) for qPCR experiments.

(d) PCR 2: Add remaining DNA (19-20 µL), 5 µL PCR Primer Cocktail, and 25 µL 2X NEBNext PCR master mix (high fidelity, New England Biolabs). PCR each sample for the number of cycles determined in step (c). Clean up following the TruSeq protocol using 50 µL XP Ampure beads, and eluting in 30 µL buffer.

(e) For accurate pooling do a second qPCR step using 1/1000 of the DNA from step (d). Calculate the delta Ct values and pool as desired. We typically aim to sequence all APEX-seq libraries to the same depth, which usually means the libraries are pooled at the same amount. Quantify absolution concentration using the bioanalyzer “high-sensitivity” setting (Agilent Bioanalyzer 2100).

7. Prepare a 4 nM 10 µL pooled-library mix for subsequent sequencing. For sequencing, we typically check library quality on the Illumina MiSeq. For deep sequencing, we recommend paired-end reads (2 x 75) using the Illumina Hiseq 4000 or Illumina NextSeq. We aim for ~25-40 million reads per sample.

Note: There are many library-preparation kits on the market to prepare sequencing libraries. While other kits could work as well, we have not tested these kits.

Note: We do not carry out any gel purification of DNA libraries. However, it is very important that contamination with unreacted adapters is minimal, as determined by the bioanalyzer traces in step 6 (e). If there is substantial contamination, additional purification of the libraries may be necessary.

Days 5+: Analysis of APEX-seq Libraries

Please refer to the methods section of Fazal *et al.*[2] for suggestions on analyzing libraries.

1. See summarized methods for details on trimming data (optional if contamination with unreacted adapters is absent/minimal), mapping data (using STAR[3]), counting reads (using HTSEQ[4]) and analyzing data (using DESEQ2[5]).

2. Transcripts that are differentially expressed in targets relative to controls are identified using DESEQ2. Suggestions to analyze data may be gathered from APEX proteomics experiments (Hung et al.[1]) using true-positive and false-positive lists, and by referring to the “ERM APEX-seq extended analysis section” of the methods.

3. If no prior information is known or assumed about expected transcripts, a reasonable starting point is to consider transcripts with a $\log_2$foldchange > 0.75, and an FDR < 0.05, as enriched by APEX-seq.

Troubleshooting

Time Taken

Anticipated Results

References

1 Hung, V. *et al.* Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat. Protoc.* **11**, 456-475, doi:10.1038/nprot.2016.018 (2016).

2 Fazal, F. M. *et al.* Atlas of Subcellular RNA Localization Revealed by APEX-seq. *bioRxiv*, 454470, doi:10.1101/454470 (2018).

3 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).

4 Anders, S., Pyl, P. T. & Huber, W. (Cold Spring Harbor Laboratory, 2014).

5 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, doi:10.1186/s13059-014-0550-8 (2014).

Acknowledgements

Figures

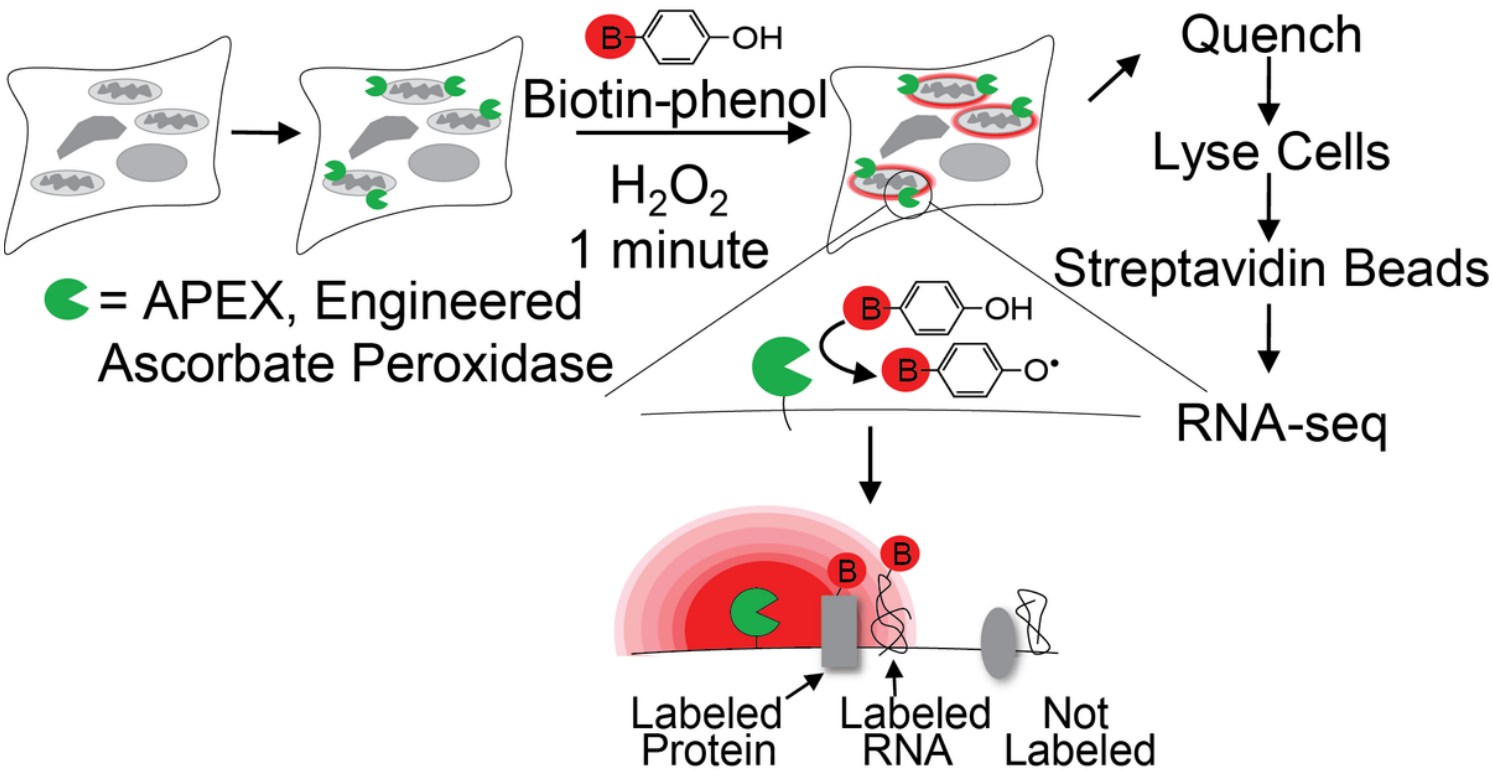

Figure 1

Schematic diagram of the APEX-seq protocol