Published in final edited form as:

Methods. 2014 May 1; 67(1): 45–54. doi:10.1016/j.ymeth.2013.08.015.

Use of Bru-Seq and BruChase-Seq for genome-wide assessment of the synthesis and stability of RNA

Michelle T. Paulsen^a, Artur Veloso^{a,b}, Jayendra Prasad^a, Karan Bedi^{a,c}, Emily A. Ljungman^a, Brian Magnuson^a, Thomas E. Wilson^{d,e}, and Mats Ljungman^{a,c,*}

^a Department of Radiation Oncology, University of Michigan Comprehensive Cancer Center and Translational Oncology Program, University of Michigan, Ann Arbor, MI, USA

^b Bioinformatics Program, University of Michigan, Ann Arbor, MI, USA

^c Department of Environmental Health Sciences, School of Public Health, University of Michigan, Ann Arbor, MI, USA

^d Department of Pathology, University of Michigan, Ann Arbor, MI, USA

^e Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA

Abstract

Gene expression studies commonly examine total cellular RNA, which only provides information about its steady-state pool of RNA. It remains unclear whether differences in the steady-state reflects variable rates of transcription or RNA degradation. To specifically monitor RNA synthesis and degradation genome-wide, we developed Bru-Seq and BruChase-Seq. These assays are based on metabolic pulse-chase labeling of RNA using bromouridine (Bru). In Bru-Seq, recently labeled RNAs are sequenced to reveal spans of nascent transcription in the genome. In BruChase-Seq, cells are chased in uridine for different periods of time following Bru-labeling, allowing for the isolation of RNA populations of specific ages. Here we describe these methodologies in detail and highlight their usefulness in assessing RNA synthesis and stability as well as splicing kinetics with examples of specific genes from different human cell lines.

Keywords

Transcription; RNA stability; RNA splicing

1. Introduction

The steady-state level of a particular RNA in a cell is a balance between its rates of production and degradation. Production of RNA is regulated by epigenetic marks, transcription factors binding to enhancer and promoter elements and by release of RNA

© 2013 Elsevier Inc. All rights reserved.

* Corresponding author at: Department of Radiation Oncology, University of Michigan Comprehensive Cancer Center and Translational Oncology Program, University of Michigan, Ann Arbor, MI, USA. ljungman@umich.edu (M. Ljungman)..

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ymeth.2013.08.015>.

polymerases from transcription pause sites. Regulation of transcriptional elongation may also influence the output of RNA [1]. Post-transcriptional regulation of RNA is mediated by the binding of miRNA or specific RNA-binding proteins to 3'-UTR sequences of selected transcripts to direct the recruitment of factors involved in RNA degradation [2]. Knowing the relative contribution of RNA synthesis and degradation to the steady-state level of particular transcripts is critical in order to better understand the mechanisms of regulation of these transcripts. Furthermore, when cell homeostasis is changed by environmental stimuli or stress and the steady-state levels of certain RNAs are altered, it would be of great interest to explore whether the ensuing gene expression changes were the result of altered RNA synthesis, stability or both.

A number of techniques have recently been developed to assess nascent RNA synthesis genome-wide. In global run-on and sequencing (GRO-Seq), nuclei from a cell sample are isolated and initiated RNA polymerases are allowed to "run-on" *in vitro* in the presence of bromouridine [3]. RNA polymerases that were arrested *in vivo* are released *in vitro* revealing which promoters harbored arrested RNA polymerases. GRO-Seq also allows for the detection of unstable RNAs such as promoter divergent transcripts since very little RNA degradation takes place in the *in vitro* run-on assay [4]. In native elongating transcript sequencing (NET-Seq), nascent RNA is isolated by immunoprecipitation of the RNA polymerase II elongation complex followed by deep sequencing of the 3' ends of nascent transcripts associated with the RNA polymerases [5]. This technique allows for nucleotide-level resolution of nascent transcription and has revealed that RNA polymerase II frequently pauses and backtracks when encountering nucleosomes in the bodies of genes [5]. Nascent-Seq is based on the isolation of chromatin-bound nascent RNA obtained from the lysis of cells and washing of cell nuclei with NUN buffer consisting of high concentrations of NaCl, urea and NP-40 [6]. This technique has been used to monitor the efficiency of intron splicing and has provided evidence that not all splicing events occur co-transcriptionally. A different approach to assess nascent transcription is through metabolic labeling of RNA with tagged ribonucleotides followed by isolation and analysis using microarrays or deep sequencing [7–11]. This approach has been extended to also estimate the half-lives of transcripts by computationally comparing nascent and steady-state levels of RNA.

Bromouridine sequencing (Bru-Seq) and bromouridine-chase sequencing (BruChase-Seq) are based on the metabolic pulse-chase labeling of nascent RNA with bromouridine. Bromouridine has been used to label steady state RNA [12] and nascent RNA [7,13] both *in vitro* and in cells [3]. While other ribonucleotide analogs, such as 4-thiouridine (4sU) and ethynyluridine (EU), can be used to specifically label and isolate nascent RNA, bromouridine is less toxic to cells than these other analogs [12]. Furthermore, the low cost of bromouridine and the availability of excellent anti-BrdU antibodies make bromouridine labeling of nascent RNA an attractive approach to study transcriptional and post-transcriptional regulation.

Following labeling, Bru-containing RNA is specifically captured using anti-BrdU antibodies conjugated to magnetic beads. cDNA libraries are then produced from the isolated Bru-RNA and subjected to deep sequencing [14]. By chasing Bru-labeled cells with uridine for different periods of time, RNA populations of defined ages can be isolated and analyzed.

This allows for the estimation of the relative stability of all transcripts and splicing kinetics of all introns. We recently used these techniques to obtain signatures of the TNF-induced acute inflammatory response in human fibroblasts and found a complex pattern of altered synthesis and/or stability of specific RNAs [14]. We also found interesting patterns of synthesis, stability and splicing in untreated cells suggesting that steady-state RNA levels are controlled by intricate transcriptional and post-transcriptional regulation.

Here we describe Bru-Seq and BruChase-Seq in detail and show examples of how the stability of transcripts vary in a cell type-specific manner. Furthermore, we show that BruChase-Seq can be used to predict nonsense and frameshift mutations in genes by revealing increased mRNA turnover rates. Finally, using segmentation analysis of nascent transcription spans we show how Bru-Seq can detect unannotated, long non-coding RNAs (lncRNA) with a highly cell type-specific expression pattern.

2. Description of methods

The Bru-Seq and BruChase-Seq techniques were recently described [14]. We will here provide a more detailed description of the materials and procedures involved in the different steps of these techniques.

2.1. Materials

Buffer	Composition
RPMI growth medium	RPMI 1640, 10% FBS, 100 U/ml penicillin, 100 U/ml streptomycin
DMEM growth medium	DMEM, 10% FBS, 100 U/ml penicillin, 100 U/ml streptomycin
MEM growth medium	Minimal Essential Medium, 10% FBS, 1× MEM Amino Acids, 1× Non-Essential Amino Acids, 2 mM l-glutamine, 1X antibiotic-antimycotic, 1× MEM vitamin mixture, 0.15% (w/v) sodium bicarbonate
6× gel loading buffer	10 mM Tris, pH 7.6, 60% glycerol, 60 mM EDTA, 0.03% bromophenol blue

Material	Supplier, catalog number
(-)-5-Bromouridine	Sigma-Aldrich, 850187
Uridine	Sigma, U3750
Trizol reagent	Invitrogen, 15596-018
Chloroform	Fisher, BP1145
Isopropanol	Sigma-Aldrich, 190764
Diethyl pyrocarbonate (DEPC)	Sigma, D5758
Bovine Serum Albumin (BSA)	Roche, 03116999001
Dynabeads Goat anti-Mouse IgG	Invitrogen, 110.33
Mouse anti-BrdU	BD Pharmingen, 555627
RNaseOUT, Ribonuclease Inhibitor	Invitrogen, 10777-019
Superscript II	Invitrogen, 18064-014
Random primers (3 µg/µl)	Invitrogen, 48190-011
100 mM dNTP set	Invitrogen, 10297-018
ActinomycinD	Sigma, A9415
AmPure RNAClean beads	Fisher, APN000494

Material	Supplier, catalog number
10× NEBuffer 2	New England Biolabs, B7002S
dUTP	Roche, 11934554001
RNase H	Invitrogen, 18021-014
DNA polymerase I	Invitrogen, 18010-017
AmPure XP beads	Fisher, NC9933872
TruSeq RNA preparation kit	Illumina, RS-122-2001
NuSieve 3:1 agarose	Lonza, 50090
10× TAE Buffer	Lonza, 50844
50 bp ladder	Invitrogen, 10416-014
Gel excision tips	The Gel Company, PKB6.5-R
QIAEX II gel extraction kit	Qiagen, 20021
USER enzyme	New England Biolabs, M5505L

2.2. Procedures

2.2.1. Cell culturing

1. Grow cells in appropriate growth medium. For this study, RPMI (BxPC3), DMEM (Panc1, MiaPaCa2, HeLa) and MEM (NF) were used.
2. Follow normal cell culture protocols to expand cells. For most cell lines, we recommend using 2–3 10-cm plates, or a minimum of 4×10^6 cells per sample.
3. Cells are grown to approximately 80% confluency before the addition of bromouridine.

2.2.2. Bromouridine labeling

- Make a stock solution of 50 mM Bromouridine in PBS.
- Make a stock solution of 1 M uridine in PBS for chase (stability analysis).
- Use conditioned media for all treatments.
 1. Remove 3–4 ml of media from each plate of cells to a clean tube and add BrU to a final concentration of 2 mM. Discard remaining media from plate, or save to use for a uridine chase.
 2. Add back BrU-containing media to plate and incubate at 37 °C for 30 min.
 3. If doing a chase, after the 30 min incubation, rinse plate twice with PBS, then add back saved media containing 20 mM uridine and incubate for desired time period (6 h may be an appropriate time to start with).
 4. To collect cells (pooling plates as necessary), either add Trizol directly to the plate to lyse cells, or trypsinize cells, spin to pellet and resuspend in 3–5 ml Trizol. Vortex until no cell pellet is visible. Store samples at –80 °C if not isolating RNA immediately. We recommend collecting the cells in 14 ml round-bottom centrifuge tubes (e.g. BD 352059).

2.2.3. Isolation of RNA

1. To each Trizol-lysed sample, add 0.2 ml chloroform per 1 ml of Trizol used initially. Cap tube and shake vigorously for 5–10 s. Remove cap, cover tube with parafilm and centrifuge at 4 °C for 15 min at 12,000g in a Sorvall RC5C floor centrifuge with SS-34 rotor (or equivalent).
2. Transfer the upper aqueous layer to a new round-bottom 14 ml tube and add 0.5 ml isopropanol per 1 ml of Trizol used initially. Cover tube with parafilm and mix gently. Incubate at room temperature for 10 min before centrifuging at 4 °C for 10 min at 12,000g to pellet RNA.
3. Remove supernatant and wash pellet by adding 1 ml of 75% ethanol per 1 ml of Trizol used initially. Cover tube with parafilm and centrifuge at 4 °C for 5 min at 7500g.
4. Remove supernatant and invert tube to allow pellet to dry slightly. Resuspend the pellet in 200 µl DEPC-water and incubate at 55 °C for 10 min to ensure RNA is fully dissolved. Store RNA at –80 °C unless immediately proceeding to isolation of Bru-RNA.

2.2.4. Preparation of magnetic beads conjugated with anti-BrdU antibodies

1. Transfer 50 µl of anti-mouse IgG magnetic Dynabeads (Invitrogen) per sample to a 1.5 ml microfuge tube. Capture beads with a magnetic stand (Novagen) and aspirate storage buffer.
2. Add 200 µl 0.1% BSA in DEPC-PBS, flick the tube to resuspend beads, capture beads on the magnetic stand and aspirate supernatant. Repeat 2 more times for a total of 3 washes. After the final wash, resuspend each bead pellet in 200 µl 0.1% BSA in DEPC-PBS and add 0.5 µl RNaseOUT.
3. To each tube, add 4 µl (2 µg) anti-BrdU antibody and 0.5 µl (20 U) RNaseOUT. Incubate with gentle rotation for 1 h at room temperature.
4. Wash beads 3 times with 200 µl 0.1% BSA in DEPC-PBS as detailed above. After the final wash, resuspend conjugated beads in 200 µl 0.1% BSA in DEPC-PBS and add 0.5 µl RNaseOUT.

2.2.5. Isolation of Bru-labeled RNA

1. Heat isolated RNA in an 80 °C heat block for 10 min, then immediately put samples on ice.
2. Remove 90% of the sample (180 µl) and add to prepared beads. Flick tube to ensure sample is mixed well and place on rotator for 1 h at room temperature.
3. Wash beads with 0.1% BSA in DEPC-PBS for 5 min on the rotator. Do 2 additional brief washes with 0.1% BSA in DEPC-PBS, making sure to completely remove the final wash.

4. Resuspend the bead pellet in 40 μ l DEPC-water and incubate for 10 min in a 95 $^{\circ}$ C heat block to elute Bru-RNA from the beads.
5. Centrifuge tubes briefly, then capture beads in the magnetic stand.
6. Remove the supernatant to a clean 1.5 ml microfuge tube, quantitate Bru-RNA concentrations (Nanodrop, Thermo Scientific), and store at -80° C if not using immediately for library preparation.

2.2.6. cDNA library preparation

- Start with at least 250 ng Bru-RNA
- Unless otherwise stated, use a Thermomixer R (Eppendorf) for all incubations.

2.2.6.1. Fragment mRNA

1. Pre-mix (per sample):
 - a. 8 μ l 5 $\times$ First –strand buffer (comes with Superscript II).
 - b. 1 μ l Random primer (3 μ g/ μ l).
2. Add 9 μ l of pre-mix to each PCR tube.
3. Add 16 μ l RNA to each tube.
4. Incubate in PCR machine at 85 $^{\circ}$ C for 10 min. Cool down to 4 $^{\circ}$ C.

2.2.6.2. Synthesize first strand cDNA (for strand specificity)

1. Mix the following reagents (per sample):
 - i. 4.0 μ l 100 mM DTT (comes with Superscript II).
 - ii. 0.8 μ l 25 mM dNTP.
 - iii. 0.5 μ l RNaseOUT.
 - iv. 0.8 μ l ActinomycinD (2.5 μ g/ μ l stock).
 - v. 6.9 μ l ddH₂O.
 - vi. 2.0 μ l Superscript II.
2. Add 15 μ l mixture to each 25 μ l RNA sample.
3. Incubate samples on thermal cycler using the following program:
 - i. 25 $^{\circ}$ C for 10 min.
 - ii. 42 $^{\circ}$ C for 50 min.
 - iii. 70 $^{\circ}$ C for 15 min.
 - iv. Hold at 4 $^{\circ}$ C.

2.2.6.3. Purify first strand DNA with AMPure RNAClean beads

1. Mix 40 μ l cDNA mixture with 72 μ l RNAClean beads.

2. Bind at room temp for at least 10 min (with 500 rpm shaking).
3. Wash beads twice with 80% ethanol. Spin briefly and remove any additional ethanol.
4. Dry beads at 37 °C for ~3 min.
5. Add 42 µl of 5 mM Tris, pH 8.0, mix well.
6. Incubate at ~28 °C for 10–15 min (with shaking).
7. Capture beads and transfer 40 µl supernatant to a new tube.

2.2.6.4. Synthesize second strand cDNA

1. Mix the following reagents:
 - i. 20 µl 10× NEBuffer 2.
 - ii. 1.2 µl 25 mM dG + dA + dU + dC mix.
 - iii. 35.3 µl ddH₂O.
 - iv. 1 µl RNase H.
 - v. 5 µl DNA polymerase I.
2. Add 60 µl mixture to each 40 µl First strand sample.
3. Incubate samples on thermal cycler at 16 °C for 2.5 h.

2.2.6.5. Clean up with AMPure beads

1. Vortex AMPure beads and add 150 µl to each 100 µl sample.
2. Pipette up and down to mix thoroughly.
3. Incubate at room temp with shaking for 10–15 min.
4. Capture beads for 5 min.
5. Remove and discard supernatant from each sample.
6. With the tubes still in the stand, add 200 µl freshly prepared 80% ethanol to each sample without disturbing the beads.
7. Incubate at room temp for 30 s, then remove and discard supernatant from each sample.
8. Repeat steps 6 and 7 for a total of two ethanol washes.
9. Spin samples briefly and remove any remaining ethanol
10. Incubate samples at 37 °C until dry.
11. Add 62 µl Resuspension Buffer (from TruSeq kit) to each sample.
12. Pipette up and down to mix thoroughly.
13. Incubate samples at room temp (with shaking) for 10–15 min.

14. Place samples on magnetic stand for 5 min.

15. Transfer 60 µl of the supernatant (ds cDNA) to a new tube.

Can stop here and store samples at –20 °C for up to seven days

From this step (End Repair) forward, the reagents will be from Illumina's TruSeq Kit.

2.2.6.6. Perform end repair

- Thaw End Repair Mix, resuspension buffer at room temperature.
- Make sure AMPure beads are at room temperature.
- Pre-heat thermomixer to 30 °C.
 1. Add 40 µl End Repair Mix to each 60 µl sample.
 2. Adjust pipette to 100 µl and pipette up and down to mix thoroughly.
 3. Incubate samples in thermomixer (no shaking) at 30 °C for 30 min.

2.2.6.7. Clean up with AMPure beads

1. Vortex AMPure beads and add 160 µl to each sample.
2. Pipette up and down to mix thoroughly.
3. Incubate at room temp (with shaking) for 10–15 min.
4. Capture beads for 5 min or until liquid appears clear.
5. Remove and discard supernatant from each sample.
6. With the tubes still in the stand, add 200 µl freshly prepared 80% ethanol to each sample without disturbing the beads.
7. Incubate at room temp for 30 s, then remove and discard supernatant from sample.
8. Repeat steps 7 and 8 for a total of two ethanol washes.
9. Spin samples briefly and remove any remaining ethanol.
10. Incubate samples at 37 °C until dry.
11. Resuspend the dried pellet with 20 µl Resuspension Buffer.
12. Pipette up and down to mix thoroughly.
13. Incubate samples at room temp (with shaking) for 10–15 min.
14. Place samples on magnetic stand for 5 min, or until liquid appears clear
15. Transfer 17.5 µl of the supernatant to a new tube.

Can stop here and store samples at –20 °C for up to seven days

2.2.6.8. Adenylate 3' ends

- Thaw Resuspension Buffer and A-Tailing Mix at room temp.

- Pre-heat Thermomixer to 37 °C.
 1. Add 12.5 µl A-Tailing Mix to each sample.
 2. Pipette up and down to mix thoroughly.
 3. Incubate samples at 37 °C for 30 min (no shaking).

2.2.6.9. Ligate Adaptors

- Thaw RNA adaptor index tubes, stop ligation buffer, and resuspension buffer at room temp.
- Make sure AMPure beads are at room temp.
- Pre-heat Thermomixer to 30 °C.
 1. Add 2.5 µl resuspension buffer to each sample.
 2. Add 2.5 µl Ligation Mix to each sample. (Remove Ligation Mix from –20 °C just before using and return to –20 °C immediately after using).
 3. Add 2.5 µl desired RNA adaptor index to appropriate sample.
 4. Adjust pipette to 37.5 µl and pipette up and down to mix thoroughly.
 5. Incubate samples in Thermomixer at 30 °C for 10 min (no shaking).
 6. Add 5 µl stop ligation buffer to each sample.
 7. Adjust pipette to 42.5 µl and pipette up and down to mix thoroughly.

2.2.6.10. Clean Up with AMPure beads

1. Vortex AMPure beads and add 65 µl to each sample.
2. Pipette up and down to mix thoroughly.
3. Incubate at room temp for 15 min.
4. Capture beads for 5 min or until liquid appears clear.
5. Remove and discard supernatant from each sample.
6. With the tubes still in the stand, add 200 µl freshly prepared 80% ethanol to each sample without disturbing the beads.
7. Incubate at room temp for 30 s, then remove and discard supernatant from sample.
8. Repeat steps 6 and 7 for a total of two ethanol washes.
9. Spin samples briefly and remove any remaining ethanol.
10. Incubate samples at 37 °C until dry.
11. Resuspend the dried pellet with 32 µl 5 mM Tris.
12. Pipette up and down to mix thoroughly.
13. Incubate samples at room temp (with shaking) for 10–15 min.

14. Place samples on magnetic stand for 5 min, or until liquid appears clear.

15. Transfer 30 µl of the supernatant to a new tube.

Can stop here and store samples at –20 °C for up to seven days

2.2.6.11. Size Selection by Agarose gel electrophoresis

- Cast 3% gel using NuSieve 3:1 agarose.
- Remove buffer from wells before loading.
- Load order: ladder, sample, ladder, sample, ladder, sample, etc.
- Run gel in 1XTAE and do not cover gel with buffer.
 1. Add 5 µl 6× gel loading buffer to each sample.
 2. Use 50 bp ladder.
 3. Load the gel and run at 65 V for 1 h 40 min.
 4. Rinse gel with distilled water
 5. Excise gel slices in the 300 bp region using a gel excision tip with a 1000 µl pipettor. Cut out backup gel slice as well of a 350 bp size.
 6. Purify the gel slices with QIAEXII kit as follows:
 - a. Add 900 µl QX buffer and 10 µl QIAEX II suspension beads and mix well.
 - b. Incubate at 40 °C with shaking for 15 min.
 - c. Spin 13,000 rpm for 30 s. Remove supernatant.
 - d. Add 500 µl QX, vortex, spin at 13,000 rpm for 30 s.
 - e. Remove supernatant, add 500 µl PE buffer, spin at 13,000 rpm for 30 s.
 - f. Repeat PE wash.
 - g. Remove supernatant, spin again at 13,000 rpm.
 - h. Remove supernatant, dry beads at 37 °C until they turn white.
 7. Add 22 µl Resuspension Buffer to elute DNA. Mix well and incubate at room temperature (with shaking) for 10–15 min.
 8. Transfer 20 µl of the supernatant to a PCR tube.

Can stop here and store samples at –20 °C for up to seven days

2.2.6.12. Uridine Digestion/Enrich DNA Fragments

- Thaw PCR Master Mix and PCR Primer Cocktail at room temp and spin briefly.
- Make sure AMPure beads are at room temperature.

- 1 Mix the following reagents (per sample):
 - i. 25.0 µl PCR Master Mix.
 - ii. 5.0 µl PCR Primer Cocktail.
 - iii. µl USER enzyme.
- 2 Add 31 µl to each sample.
- 3 Pipette up and down to mix thoroughly
- 4 Incubate samples on thermal cycler using the following program:
 - i. 37 °C for 15 min (uridine digestion).
 - ii. 98 °C for 30 s.
 - iii. 15 cycles of:
 - 1 98 °C for 10 s.
 - 2 60 °C for 30 s.
 - 3 72 °C for 30 s.
 - iv 72 °C for 5 min
 - v Hold at 10 °C

2.2.6.13. Clean Up with AMPure beads

1. Vortex AMPure beads and add 50 µl to each sample.
2. Pipette up and down to mix thoroughly.
3. Incubate at room temp with shaking for 10–15 min.
4. Capture beads for 5 min or until liquid appears clear.
5. Remove and discard supernatant from each sample.
6. With the tubes still in the stand, add 200 µl freshly prepared 80% ethanol to each sample without disturbing the beads.
7. Incubate at room temp for 30 s, then remove and discard supernatant from sample.
8. Repeat steps 6 and 7 for a total of two ethanol washes.
9. Spin samples briefly and remove any remaining ethanol.
10. Incubate samples at 37 °C until dry.
11. Resuspend the dried pellet with 27 µl 5 mM Tris Buffer.
12. Pipette up and down to mix thoroughly.
13. Incubate at room temp with shaking for 10–15 min.
14. Place samples on magnetic stand for 5 min, or until liquid appears clear
15. Transfer 25 µl of the supernatant to a new PCR tube.

2.2.6.14. Validation

1. Use 3 μ l of each library to run on a thin 1.5% agarose gel to ensure there is a single band running around 300 bp.
2. Quantitate libraries using Nanodrop. Set 20 μ l of sample aside and save. The rest of the samples are now ready for sequencing.

2.3. Deep sequencing

Sequencing can be performed using any preferred platform. We use Illumina HiSeq 2000 via the University of Michigan Sequencing Core. We also take advantage of cost-saving associated with sample indexing. Acceptable results can often be obtained for most expressed genes when the reads (~40 million at a time, in our case) are distributed across multiple samples.

2.4. Data analysis pipeline

The conceptual bioinformatics approaches used in Bru-Seq and BruChase-Seq were recently described [14]. Our data analysis pipeline, which uses common bioinformatics tools for sequence read analysis (section 2.4.1) as well as custom scripts, is implemented using the q pipeline manager (<http://sourceforge.net/projects/qpln-mngr/>). Supplementary File 1 (see also <http://tewlab.org/q/BruSeq.html>) provides a q-generated HTML report that includes master scripts and log files for example jobs for each of the major steps outlined below, as well as all worker scripts called by those masters. A description of our custom extensions of BED and BAM file formats is provided in Supplementary File 2.

2.4.1. Major programs used—

Program	Version
TopHat [15]	v1.4.1
Bowtie [16]	v0.12.8
BEDTools [17]	v2.16.2
Samtools [18]	v0.1.18
R [19]	v2.15.1
DESeq [20]	v1.4.1

2.4.2. Read mapping (q master ‘map’)

1. Map reads to the human ribosomal DNA complete repeating unit (U13369.1) using Bowtie; keep rRNA read counts and non-rRNA read sequences.
2. Map non-rRNA reads to the human reference genome assembly hg19/GRCh37, or other appropriate genome, using TopHat.
 - i. Keep only reads that map uniquely, with up to two read segment mismatches.

- ii. Reads are allowed to split between exons in RefSeq or another preferred transcript annotation, but *de novo* splice junction calling is not performed since nascent RNA reads are mainly intronic.
- iii. Duplicate reads are maintained and expected in mature RNA samples where reads cluster in exons.

2.4.3. Genome annotation

1. In preparation for counting, condense the RefSeq transcript isoforms of genes into one BED file of non-redundant intron and exon spans, using `create_transcriptome_map.pl` (<http://tewlab.path.med.umich.edu/software/utilities/utilities.html>) or another utility, so that genome bases will have only one assigned identity.
 - i. When isoforms conflict, give priority to annotation as an exon to prevent a stable exon from being annotated as an intron.
 - ii. Overlapping regions of different genes are termed ambiguous and ignored when determining the expression level of the involved genes.

2.4.4. Expression scoring (q master 'map')

1. Determine the strand-specific coverage over each genome base so that a read might be fractionally attributed to different exons, bins or other features.
 - i. Count the number of reads in a given orientation overlapping each base, using BEDTools.
 - ii. Divide by the length of the sequenced reads.
2. Sum the base coverages across a given feature to determine its read coverage.
 - i. For gene expression in Bru-Seq samples, calculate the RPKM using all introns and exons.
 - ii. For gene expression in BruChase-Seq samples, calculate the RPKM using all, and only, exons.
3. Similarly sum the base coverages and calculate RPKM for each 1 Kb genome bin, or other desired bin size, in preparation for segmentation.

2.4.5. Combining replicates (q master 'merge')

1. Sum the fractional base coverages and bin coverages over all replicate samples and recalculate feature and bin RPKM as in 2.4.4.

2.4.6. Genome segmentation (q master 'segment')

1. Normalize the 1 Kb genome bins by discarding unmappable bins and dividing remaining bin RPKM values by the fractional mappability, determined using `extractKmers.pl` (<http://tewlab.path.med.umich.edu/software/utilities/utilities.html>)

or another utility, to prevent unmappable regions from breaking contiguous transcription units.

2. Apply wavelet smoothing to the normalized bin RPKM using `smooth.pl` (<http://sourceforge.net/projects/smooth-stream/>) or another utility.
3. Establish the emission probabilities of a hidden Markov model:
 - i. Score the bins by rounding each into one of 17 logarithmically distributed RPKM input states.
 - ii. Score annotated genes by rounding each into one of 10 logarithmically distributed RPKM output states.
 - iii. Assign emission probabilities as the frequency of input bin states observed over all gene output states. See Supplementary Fig. 1.
4. Set the transition probability to 0.005 for all bins.
5. Solve the model using the Viterbi algorithm using `segment.pl` (<http://sourceforge.net/projects/segment-stream/>) or another utility to establish the most likely bin expression states.
6. Fuse adjacent bins of the same state into genome segments of sustained contiguous expression.

2.4.7. Rank synthesis and stability (q master 'assemble')

1. Score relative synthesis by ranking the RPKM of genes, segments, or other features obtained from nascent RNA collected immediately after Bru labeling.
2. Score relative stability by ranking the ratio of the RPKM of features obtained from aged RNA samples over paired nascent RNA samples.
3. Score splicing extent as intron retention, i.e. the RPKM of a specific intron divided by the RPKM of all exons of the same gene.

2.4.8. Inter-sample comparisons (q master 'compare')

1. Apply the DESeq R package to compare the replicates that gave rise to one merged sample to the replicates of a different merged sample of the same type (nascent or aged).
 - i. Use genes and other annotated feature as is.
 - ii. Split hidden Markov segments at all inter-segment boundaries encountered in either sample to provide the complete set of potentially divergent transcription segments.
2. Compare Bru-Seq samples to explore differences in nascent RNA synthesis for genes, individual exons, or other features.
3. Compare BruChase-Seq samples to explore differences in RNA stability, but only when both samples were split from the same Bru-labeled cell stock prior to a

manipulation that might affect transcript stability, so that the input nascent RNA is identical.

3. Results

3.1. BruChase-Seq reveals cell type-specific regulation of RNA stability

RNA degradation is regulated by specific miRNA and RNA-binding proteins that bind to the 3'-UTR or internal sequences of mature transcripts. By comparing the amount of exonic RNA reads present 6 h after Bru-labeling with the total amount of reads for the entire gene directly after Bru-labeling, an estimation of the relative stability of each transcript can be made. To test whether RNA stability is transcript-specific or whether the stabilities of specific RNAs differ in a cell type-specific manner, we performed Bru-Chase-Seq on a set of human cell lines. As can be seen in Fig. 1a, the *NFKB1* transcript showed robust stability in normal fibroblasts as determined by the relatively high exonic signal at 6 h (red) compared to the nascent transcript level (blue). However, the exonic signal was found to be much less prominent at 6 h in either HeLa or H146 cells (Fig. 1b and c). In contrast, the transcript of *DPC2* exhibited a low relative level of exonic signal in human fibroblasts (Fig. 1d) while the relative exonic signal was much higher in K562 and H146 cell lines suggesting that this transcript is more stable in these cancer cell lines (Fig. 1e and f). These results demonstrate the usefulness of BruChase-Seq in monitoring the relative stability of transcripts and indicate that the stability of some transcripts varies in a cell type-specific manner.

3.2. Stability of the MYC transcript is elevated in some cancer cell lines

The expression of the oncoprotein MYC is frequently upregulated in human cancers. This upregulation is sometimes caused by amplification of the *MYC* gene and in the case of HeLa cells is due to integration and amplification of viral regulatory regions proximal to the *MYC* gene [21,22]. To test whether *MYC* transcripts may be regulated at the level of stability in cancer cells we used BruChase-Seq. The *MYC* transcript was found to be quite unstable in both human fibroblasts and HeLa cells as previously shown for many cell lines [23] (Fig. 2a and b). However, in the pancreatic cancer cell lines BxPC3 and MiaPaCa2, the exonic reads at 6 h were quite robust suggesting that the *MYC* transcript is more stable in these cell lines as compared to human fibroblasts and HeLa cells. The BruChase-Seq technique thus revealed that the *MYC* transcript shows differential stability in different cell lines, suggesting that regulation of transcript stability may contribute to MYC overexpression in human tumors.

3.3. Nonsense and frame-shift mutated transcripts show low stabilities

Transcripts containing premature translation termination codons are targeted by the nonsense-mediated decay (NMD) pathway, removing these defective transcripts through RNA degradation during attempted translation [24]. To test whether BruChase-Seq can reveal a higher rate of degradation of transcripts from genes bearing nonsense mutations, we compared the relative stabilities of *RBI* mRNA in human fibroblasts, where the *RBI* gene is wild-type, and in H146 cells where it carries a nonsense mutation (CCLE, Broad Institute). The wild-type *RBI* transcript had high exonic reads at 6 h in human fibroblasts (Fig. 3a).

However, the mutated *RB1* transcript in H146 cells had a low level of exonic reads at 6 h (Fig. 3b). Frameshift mutations can also lead to the formation of premature translation termination codons activating NMD [25,26]. The *TP53* gene in H146 cells contains such a frame-shift mutation [27]. To explore whether it causes the *TP53* transcripts to become unstable, we used BruChase-Seq comparing the exonic reads of *TP53* in human fibroblasts and in H146 cells. The wild-type *TP53* transcripts had a high level of exonic reads after a six-hour chase in the fibroblasts while the mutant *TP53* transcript in H146 cells had much lower levels of exonic reads, demonstrating that it is much less stable (Fig. 3c and d). Using the Bru-Chase-Seq approach, we have confirmed that the stabilities of mutant *RB1* and *TP53* transcripts in H146 cells are reduced, presumably through the activation of NMD.

3.4. Using BruChase-Seq to explore splicing kinetics

A variation of BruChase-Seq is to use different durations of uridine chase to allow RNAs of defined ages to be isolated and analyzed. This approach is very useful when exploring post-transcriptional processing of primary transcripts. Four different ages of RNA expressed from the *CD44* gene are shown in Fig. 4. At 0 h (nascent RNA), the complete and predominantly unspliced primary transcript can be observed. After a 2 h chase, most of the intronic signal has disappeared as a function of splicing and degradation, while the reads covering exons and the 3'-UTR are enhanced. One region of the *CD44* transcript, however, is not processed with the same kinetics as other regions. This represents the so called “variable region” of the *CD44* transcript, which is commonly omitted in the mature transcript of most cell types [28]. The variable region, notably, was still present after 4 h, but by 6 h it had nearly disappeared. This “retention” of the variable region over time may be due to slow splicing of the introns in this region, which would suggest that this process occurs post-transcriptionally. Alternatively, most of the transcripts may have their variable regions spliced out co-transcriptionally, but the transcripts retaining the variable region are subjected to accelerated degradation and are lost more quickly than the fully-spliced population. A third possibility is that the variable region is co-transcriptionally spliced but that this spliced intron possesses a non-coding function allowing it to escape degradation. Regardless of the mechanism, BruChase-Seq allows determination of intron retention genome-wide and should provide new insights into the regulation of splicing in human cells.

3.5. Bru-Seq reveals cell type-specific expression of long, non-coding RNAs

The use of RNA-Seq technology has revealed a myriad of lncRNAs generated throughout the genome [29]. The functions of these RNA species are mostly unknown. By employing a hidden Markov model-based segmentation analysis, we were able to identify transcription units independently of prior gene annotation. We applied this method across five different cell types (human whole blood, normal human fibroblasts and the cancer cell lines K562, BxPC3 and Panc1) and identified numerous unannotated lncRNAs, some of which were very large, spanning over 100 Kb in length (Fig. 5a). When compared across the five cell types, these lncRNAs exhibited distinctive cell type-specific expression patterns. The lncRNA shown in Fig. 5a was only found to be expressed in BxPC3 cells while the lncRNA shown in Fig. 5b was expressed only in whole blood cells. Human fibroblasts contained the least (21) and K562 cells contained the most (153) previously unannotated lncRNAs of the five cell types analyzed (Fig. 5c). Strikingly, most of the lncRNAs were uniquely expressed

in each cell line and it appeared that the lncRNA transcripts were either highly expressed or not expressed at all (Fig. 5 a and b). We conclude that Bru-Seq can be used to identify non-annotated lncRNA genome-wide and our results suggest that these RNAs have a very strong cell type-specific expression pattern.

4. Conclusions

To better understand the underlying mechanisms regulating the steady state levels of RNA in cells, the contributions of both RNA synthesis and RNA degradation must be taken into account. To study both RNA synthesis and degradation in living cells, we developed the Bru-Seq and BruChase-Seq approaches based on metabolic pulse-chase labeling of nascent RNA with bromouridine [14]. We have here outlined these techniques in detail and have provided examples of their usefulness in assessing RNA synthesis and RNA stability for selected genes across multiple cell lines. The results obtained with these techniques suggest that transcript stability is differentially regulated in different cell types. For example, the *NFKB1* transcript was turned over faster in the cancer cell lines HeLa and H146 than in normal human fibroblasts while the *DCP2* transcript was more stable in the cancer lines K562 and H146 than in the human fibroblasts. The *MYC* transcript was very unstable in human fibroblasts and HeLa cells but much more stable in the cancer cell lines BxPC3 and MiaPaCa2. Bru-Chase-Seq also confirmed that mutations causing premature translation termination codons generate highly unstable transcripts, probably due to the activation of NMD to degrade these defective transcripts.

Comparing RNA populations of different ages, obtained from cells chased for different periods of time following bromouridine labeling, allows for the detailed exploration of post-transcriptional processing events such as splicing. As an example, BruChase-Seq revealed a much slower removal of the variable region of the CD44 compared to adjacent introns. The reason for this is not clear but it could be due to (i) slow splicing, (ii) a faster decay of transcripts containing a retained variable region or (iii) enhanced stability of the spliced intron, perhaps due to the presence of a putative functional sequence. We are currently exploring the regulation of splicing and intron retention genome-wide using Bru-Chase-Seq.

Finally we showed examples of the power of Bru-Seq to identify nascent RNA transcripts outside of any prior annotation via segmentation analysis. We found that most unannotated lncRNA species were expressed in a very cell type-specific manner when comparing across five different cell lines. The functions of these lncRNAs and how they are regulated are poorly understood but of great interest. and the Bru-Seq technique is ideally suited to identify and characterize expression of these RNA species in diverse cell types and during cellular processes such as differentiation and transformation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful for the assistance by Manhong Dai and Fan Meng for administration and maintenance of the University of Michigan Molecular and Behavioral Neuroscience Institute (MBNI) computing cluster and by the personnel at the University of Michigan Sequencing Core. This work has been supported by funds from University of Michigan Bioinformatics Program, University of Michigan Biomedical Research Council, the Will and Jeanne Caldwell Endowed Research Fund of the University of Michigan Comprehensive Cancer Center, University of Michigan School of Public Health (NIEHS P30), Department of Defense, Uniting Against Lung Cancer, University of Michigan Nathan Shock Center, University of Michigan Office of the Vice President of Research, National Cancer Institute (5R21CA150100), National Institute of Environmental Sciences (1R21ES020946) and National Human Genome Research Institute (1R01HG006786).

References

1. Danko CG, Hah N, Luo X, Martins AL, Core L, Lis JT, Siepel A, Kraus WL. Mol. Cell. 2013; 50:212–222. [PubMed: 23523369]
2. Houseley J, Tollervey D. Cell. 2009; 136:763–776. [PubMed: 19239894]
3. Core LJ, Waterfall JJ, Lis JT. Science. 2008; 322:1845–1848. [PubMed: 19056941]
4. Core LJ, Waterfall JJ, Gilchrist DA, Fargo DC, Kwak H, Adelman K, Lis JT. Cell Rep. 2012; 2:1025–1035. [PubMed: 23062713]
5. Churchman LS, Weissman JS. Nature. 2011; 469:368–373. [PubMed: 21248844]
6. Khodor YL, Rodriguez J, Abruzzi KC, Tang CH, Marr MT 2nd, Rosbash M. Genes Dev. 2011; 25:2502–2512. [PubMed: 22156210]
7. Ohtsu M, Kawate M, Fukuoka M, Gunji W, Hanaoka F, Utsugi T, Onoda F, Murakami Y. DNA Res. 2008; 15:241–251. [PubMed: 18611946]
8. Rabani M, Levin JZ, Fan L, Adiconis X, Raychowdhury R, Garber M, Gnirke A, Nusbaum C, Hacohen N, Friedman N, Amit I, Regev A. Nat. Biotechnol. 2011; 29:436–442. [PubMed: 21516085]
9. Schwanhaussner B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. Nature. 2011; 473:337–342. [PubMed: 21593866]
10. Schwalb B, Schulz D, Sun M, Zacher B, Dumcke S, Martin D, Cramer P, Tresch A. Bioinformatics. 2012
11. Sun M, Schwalb B, Schulz D, Pirkel N, Etzold S, Lariviere L, Maier KC, Seizl M, Tresch A, Cramer P. Genome Res. 2012; 22:1350–1359. [PubMed: 22466169]
12. Tani H, Mizutani R, Salam KA, Tano K, Ijiri K, Wakamatsu A, Isogai T, Suzuki Y, Akimitsu N. Genome Res. 2012; 22:947–956. [PubMed: 22369889]
13. Haider SR, Juan G, Traganos F, Darzynkiewicz Z. Exp. Cell Res. 1997; 234:498–506. [PubMed: 9260920]
14. Paulsen MT, Veloso A, Prasad J, Bedi K, Ljungman EA, Tsan YC, Chang CW, Tarrier B, Washburn JG, Lyons R, Robinson DR, Kumar-Sinha C, Wilson TE, Ljungman M. Proc. Natl. Acad. Sci. USA. 2013; 110:2240–2245. [PubMed: 23345452]
15. Trapnell C, Pachter L, Salzberg SL. Bioinformatics. 2009; 25:1105–1111. [PubMed: 19289445]
16. Langmead B, Trapnell C, Pop M, Salzberg SL. Genome Biol. 2009; 10:R25. [PubMed: 19261174]
17. Quinlan AR, Hall IM. Bioinformatics. 2010; 26:841–842. [PubMed: 20110278]
18. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. Bioinformatics. 2009; 25:2078–2079. [PubMed: 19505943]
19. R.C. Team. 2012. <http://www.R-project.org/><http://www.R-project.org/>
20. Anders S, Huber W. Genome Biol. 2010; 11:R106. [PubMed: 20979621]
21. Lazo PA, DiPaolo JA, Popescu NC. Cancer Res. 1989; 49:4305–4310. [PubMed: 2545339]
22. Macville M, Schrock E, Padilla-Nash H, Keck C, Ghadimi BM, Zimonjic D, Popescu N, Ried T. Cancer Res. 1999; 59:141–150. [PubMed: 9892199]
23. Dani C, Blanchard JM, Piechaczyk M, El Sabouty S, Marty L, Jeanteur P. Proc. Natl. Acad. Sci. USA. 1984; 81:7046–7050. [PubMed: 6594679]
24. Brogna S, Wen J. Nat. Struct. Mol. Biol. 2009; 16:107–113. [PubMed: 19190664]

25. Micale L, Muscarella LA, Marzulli M, Augello B, Tritto P, D'Agruma L, Zelante L, Palumbo G, Merla G. J. Biomed. Biotechnol. 2009; 2009:860761. [PubMed: 20145706]
26. Pereira FJ. Br. J. Haematol. 2006; 133:98–102. [PubMed: 16512835]
27. Forbes S, Clements J, Dawson E, Bamford S, Webb T, Dogan A, Flanagan A, Teague J, Wooster R, Futreal PA, Stratton MR. Br. J. Cancer. 2006; 94:318–322. [PubMed: 16421597]
28. Tolg C, Hofmann M, Herrlich P, Ponta H. Nucleic Acids Res. 1993; 21:1225–1229. [PubMed: 8464707]
29. Wang KC, Chang HY. Mol. Cell. 2011; 43:904–914. [PubMed: 21925379]

Fig. 1.

BruChase-Seq reveals differential RNA stabilities across cell lines. (a) High exonic reads at 6 h for the *NFKB1* transcript in human fibroblasts indicating high relative stability. Low exonic reads at 6 h of *NFKB1* transcripts in (b) HeLa and (c) H146 cells indicating low relative stability. (d) Low exonic reads at 6 h for the *DCP2* transcript in human fibroblasts indicating low relative stability. High level of exonic reads at 6 h of the *DCP2* transcript in (e) K562 and (f) H146 cells indicating relative high stability. Nascent RNA corresponds to the blue trace and the 6-h old RNA corresponds to the red trace. The gene maps are from RefSeq Genes hg19 (UCSC genome browser <http://genome.ucsc.edu/>).

Fig. 2.

The *MYC* transcript show enhanced stability in the pancreatic cancer lines BxPC3 and MiaPaCa2 as assessed by BruChase-Seq. Low relative stability of *MYC* transcripts found in (a) human fibroblasts and (b) HeLa cells assessed by the low level of exonic reads at 6 h in these two cell lines. High relative stability of *MYC* transcripts in (c) BxPC3 and (d) MiaPaCa2 assessed by BruChase-Seq. Nascent RNA corresponds to the blue trace and the 6-h old RNA corresponds to the red trace. The gene maps are from RefSeq Genes hg19 (UCSC genome browser <http://genome.ucsc.edu/>).

Fig. 3.

Results obtained with BruChase-Seq show reduced stability of mutant transcripts. (a) The wild-type *RB1* transcript is stable in human fibroblasts as assessed by the high exonic reads from the 6-h RNA sample. (b) The *RB1* transcript containing a nonsense mutation in H146 cells shows low stability as assessed by the low exonic reads from the 6-h RNA. (c) The wild-type *TP53* transcript in human fibroblasts shows high relative stability while (d) the *TP53* transcript with a frameshift mutation in H146 cells shows low stability. Nascent RNA corresponds to the blue trace and the 6-h old RNA corresponds to the red trace. The gene maps are from RefSeq Genes hg19 (UCSC genome browser <http://genome.ucsc.edu/>).

Fig. 4.

Use of BruChase-Seq to assess splicing kinetics of the CD44 transcript. HeLa cells were incubated with 2 mM bromouridine for 30 min to label nascent RNA followed by chases in uridine for 0, 2, 4 and 6 h to generate RNA populations of different ages. It can be noted that intronic sequences that are in the region termed “variable region” are removed/degraded much slower than adjacent intronic sequences. The gene maps are from RefSeq Genes hg19 and only one of the many isoforms of the CD44 gene are shown for simplicity (UCSC genome browser <http://genome.ucsc.edu/>).

Fig. 5.

Cell type-specific expression of non-annotated lncRNAs identified using Bru-Seq (a) Example of lncRNA exclusively expressed in BxPC3 cells. (b) Example of lncRNA exclusively expressed in whole blood. (c) Table listing cell types examined, the number of non-annotated lncRNA expressed in the different cell types and the number of lncRNAs shared among the different cell types. For simplicity, the reads in (a) and (b) are shown as positive values with transcription going from right to left.