Published in final edited form as:

Methods. 2019 February 15; 155: 41–48. doi:10.1016/j.ymeth.2018.10.022.

RPAD (RNase R Treatment, Polyadenylation, and Poly(A)+ RNA Depletion) Method to Isolate Highly Pure Circular RNA

Poonam R. Pandey¹, Pranita K. Rout², Aniruddha Das^{2,3}, Myriam Gorospe^{1,*}, and Amaresh C. Panda^{1,2,*}

¹Laboratory of Genetics and Genomics, National Institute on Aging, National Institutes of Health, Biomedical Research Center, Baltimore, Maryland, USA

²Institute of Life Sciences, Bhubaneswar, India

³School of Biotechnology, KIIT University, Bhubaneswar, India.

Abstract

Recent developments in high-throughput RNA sequencing methods coupled with innovative bioinformatic tools have uncovered thousands of circular (circ)RNAs. CircRNAs have emerged as a vast and novel class of regulatory RNAs with potential to modulate gene expression by acting as sponges for microRNAs (miRNAs) and RNA-binding proteins (RBPs). The biochemical enrichment of circRNAs by exoribonuclease treatment or by depletion of polyadenylated RNAs coupled with deep-sequencing is widely used for the systematic identification of circRNAs. Although these methods enrich circRNAs substantially, they do not eliminate efficiently non-polyadenylated and highly-structured RNAs. Here, we describe a method we termed RPAD, based on initial RNase R treatment followed by Polyadenylation and poly(A)⁺ RNA Depletion. These joint interventions drastically depleted linear RNAs leading to isolation of highly pure circRNAs from total RNA pools. By facilitating the isolation of highly pure circRNAs, RPAD enables the elucidation of circRNA biogenesis, sequence, and function.

Keywords

circRNAs; backsplice sequence; RNase R; Polyadenylation; RT-qPCR; linear RNA depletion

*Correspondence: Amaresh C Panda, Institute of Life Sciences, Nalco Square, Bhubaneswar-751023, Odisha, India, Tel: +91 674 2304 314; Fax: +91 674 2300 728, amaresh.panda@ils.res.in or Myriam Gorospe, Laboratory of Genetics and Genomics, National Institute on Aging, National Institutes of Health, 251 Bayview Blvd., Baltimore, MD 21224, USA, Tel: +1 410 558 8443; Fax: +1 410 558 8331, GorospeM@grc.nia.nih.gov.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. INTRODUCTION

For decades, the vast majority of RNAs were considered to convey genetic information encoded in the DNA for the synthesis of specific proteins [1]. However, with widespread use of RNA sequencing technologies, we now recognize that >80% of the human genome is transcribed but only 2% of the RNA in a human cell encodes protein, the rest is noncoding RNA (ncRNA) [2, 3]. The major classes of ncRNAs include ribosomal (r)RNAs, transfer (t)RNAs, micro (mi)RNAs, pseudogene-encoded RNAs, long (l)ncRNA and circular (circ)RNA [3-5]. Although circRNAs were first identified three decades ago by electron microscopy [6, 7], they were considered to be splicing anomalies and were not studied in depth due to a lack of technologies to isolate, identify, and functionally characterize them. With the advent of high-throughput RNA sequencing methods coupled with novel bioinformatic analysis tools we now recognize that eukaryotic cells contain tens of thousands of distinct circRNAs [8, 9]. They are believed to arise from pre-RNAs by covalent ligation of downstream 3' splice donor site of an exon with the upstream 5' splice acceptor site by a process called backsplicing, mediated by the spliceosomal machinery, although direct 3'-5' ligation mediated by ribozymes can also generate circRNAs [9-11] [12, 13]. Besides being abundant, circRNAs are found in all cells and some show evolutionary conservation; however, many display tissue-specific expression patterns [8, 9]. Transcription elongation velocity and the presence of inverted repeat sequences including ALU elements in introns flanking the circularizing exons can promote circRNA biogenesis by backsplicing [9, 14, 15]. RNA-binding proteins (RBPs) and splicing factors like NF90/NF110, muscleblind (MBL), and quaking (QKI) are known to regulate circRNA formation [13, 16, 17].

Over 100,000 different circRNAs have been identified in humans. They have been implicated in several physiologic states, such as myogenesis, development, and the immune response, as well as in pathologies including cancer, neurodegeneration, and cardiovascular disease [18-20]. As studies on circRNAs have begun to uncover the value of circRNAs in diagnosis, prognosis, and therapy [18, 21], there is growing need to investigate circRNAs systematically. Based on their origin, circRNAs can be exonic (E)circRNA if they contain only exons, exon-intron (EI)circRNAs, if they comprise exonic and intronic sequences, circular intronic (ci)RNAs if they are generated from intronic lariats, and intronic (I)circRNA if they contain intronic sequences [8, 22-24]. Given that circRNAs lack 3' or 5' ends and hence are exceptionally stable [9], there is escalating interest in identifying their potential functions. One line of inquiry has focused on the factors that interact with circRNAs. Some circRNAs showing robust association with miRNAs and RBPs have been termed competing endogenous (ce)RNAs and they are capable of regulating gene expression by acting as a sponges for miRNAs or RBPs [5]. Accordingly, several ceRNAs have been found to regulate key cellular events including cell growth, oncogenesis, and neuronal function by acting as a sponge for miRNAs and RBPs. For example, the ceRNA CDR1as, identified as a sponge for miR-7 (with >60 miR-7 binding sites) influences insulin production, neurotransmitter production, and cancer [10, 11] [5, 11, 25, 26], while nuclear ceRNAs circEIF3J and circPAIP2 promoted host gene transcription by interacting with *U1* snRNP [22]. Although circRNAs are believed to be noncoding in nature, a few circRNAs

has been reported be translated into proteins mediated by N6-methyladenosine (m6A) modifications and internal ribosome entry sites (IRESs) [27, 28]; examples of translated circRNAs include circZNF609 and CircMbl, which encode functional proteins in muscle cells and in the fruit fly, respectively [28, 29].

CircRNAs can be distinguished from their host linear RNA counterparts by the unique circRNA backsplice junction sequence, which is critical for their identification and quantification [9-11]. Since the discovery of circRNAs, many biochemical methods have been developed for identification and validation of circRNAs including reverse transcription (RT) followed by quantitative (q)PCR amplification, *in situ* hybridisation, and analyses by Northern blotting, microarray, and RNA sequencing (RNA-seq) [8-11, 25]. At present, the most reliable and widely used method for large-scale identification of circRNAs is high-throughput RNA-seq analysis followed by bioinformatic survey of backsplice reads. Although several bioinformatic and statistical tools have been developed to identify and quantify circRNAs, none of them is able to identify true circRNAs with absolute confidence [30]. For example, backsplice reads present in the RNA-seq data originating from trans-splicing of pre-mRNAs, tandem duplication in the genome, or template switching during PCR amplification can complicate the identification of true circRNAs [31]. CircRNAs can also be detected using microarrays that employ probes spanning back-spliced junctions [25]. Depletion of rRNA coupled with depletion of poly(A)+ RNA or exonuclease (RNase R) treatment to deplete the majority of linear RNAs are the most favored methods to enrich for circRNAs among a pool of total RNA; these enriched circRNAs are then used for RNA-seq, microarray or RT-PCR analyses to detect and quantify true circRNAs. However, some linear non-polyadenylated and RNase R-resistant RNAs with short 3' overhangs, such as snRNAs, escape this enrichment step, and can interfere with downstream analyses [32]. To resolve this major obstacle, we have developed the method 'RNase R treatment followed Poly(A) tailing and poly(A) Depletion' (RPAD) to isolate highly pure circRNA populations from total RNA [24]. This method allows for significant depletion of linear RNAs of all types, including RNase R-resistant snRNAs. The highly pure circRNA population isolated through RPAD has been successfully used in high-throughput RNA-seq after rRNA depletion to identify circRNAs of very low abundance and novel circRNAs.

2. DESCRIPTION OF THE METHOD

Various approaches are available to enrich circRNAs from total RNA pool including rRNA removal, poly(A)+ RNA depletion, and exonuclease treatment. Here we describe in detail the method RPAD, designed to isolate high-pure circRNAs from total RNA. The basic outline of RPAD is illustrated (Fig. 1). The schematic was adapted from a recent report in which we used RPAD followed by high-throughput RNA-seq to identify novel circRNAs and circRNA isoforms.

2.1. Step 1: Cell culture and total RNA isolation

HeLa cells were cultured for 24-48 hours before harvest for RNA isolation. After lysis using the cell lysis reagent provided in the RNA isolation kit, total RNA was isolated for further analysis as explained in Step 2.

2.1.1. Materials & Equipments

- Human cervical carcinoma (HeLa) (ATCC, USA)
- Dulbecco's Modified Eagle Medium (DMEM) (Thermo Fisher Scientific)
- Fetal Bovine Serum (FBS) (Thermo Fisher Scientific)
- Dulbecco's phosphate-buffered saline (DPBS) (Thermo Fisher Scientific)
- miRNeasy Kit (QIAGEN)
- Cell scrapers
- RNase-free 1.5-mL microcentrifuge tubes
- Vortexer
- Microcentrifuge (Eppendorf)
- NanoDrop spectrophotometer (Thermo Fisher Scientific)

2.1.2. Detailed protocol

1. Plate ~2 million HeLa cells per 100-mm culture dish in DMEM supplemented with 10% FBS and antibiotics.
2. Allow the cells to grow for 24-48 hour till they reach a 70-80% confluency.
3. Wash cells with cold DPBS 3× and use a cell scraper to collect them into a 1.7-mL tube.
4. Centrifuge tubes at $500 \times g$ for 5 min at 4 °C to collect the cells.
5. Dissolve the cell pellet by pipetting 700 µL of QIAzol Lysis Reagent provided in the miRNeasy Kit.
6. Use the manufacturer's protocol to prepare the total RNA using the miRNeasy Kit.
7. Quantify the isolated total RNA with a NanoDrop spectrophotometer for use Step 2.

2.2. Step 2: Linear RNA depletion by RNase R treatment

Total RNA isolated from Step 1 contains all types of RNA expressed in HeLa cells, including mRNA, rRNA, tRNA, snRNA, miRNA, and circRNA. Selective degradation of linear RNAs led to specific enrichment of circRNA populations from total RNA. Seven Ribonucleases (RNases) have been reported to degrade linear RNAs, including RNase II, RNase R, and polynucleotide phosphorylase (PNPase). Recently, a strong 3' to 5' exonuclease, RNase R, was widely used to enrich the circRNAs and intronic lariats, both of which are resistant to RNase R because they lack of 3' overhangs [32, 33]. In this step, we describe in detail the procedure to deplete linear RNAs by RNase R treatment.

2.2.1. Materials and equipment

- Total RNA isolated from HeLa cells

- Nuclease-free microcentrifuge tubes
- RNase R (Lucigen, Epicentre)
- TURBO DNA-free™ Kit (Thermo Fisher Scientific)
- RiboLock RNase inhibitor (Thermo Fisher Scientific)
- Thermomixer (Eppendorf)
- miRNeasy Kit (QIAGEN)
- Vortexer
- Microcentrifuge (Eppendorf)
- NanoDrop spectrophotometer (Thermo Fisher Scientific)

2.2.2. Detailed protocol

1. Measure the concentration of RNA using the NanoDrop spectrophotometer.
2. Prepare a 20- μ L reaction containing 2-4 μ g of total RNA, 2 μ L 10 $\times$ RNase R reaction buffer (provided with RNase R), 1 μ L RiboLock, 1 μ L of RNase R, and 1 μ L of TURBO DNase.
3. Prepare another reaction without RNase R to serve as a control for RNase R treatment.
4. Mix the reactions well by flicking with the finger followed by a short spin to bring everything to the bottom of the tube.
5. Incubate the control as well as RNase R reactions for 30 min at 37 °C.
6. Add 500 μ L of QIAzol lysis reagent to the reaction and vortex for few seconds to dissolve the samples.
7. Elute control RNA and RNase R-treated RNA in 50 μ L of nuclease-free water.
8. Use a fraction of the prepared RNA for validation of linear RNA degradation as discussed in Step 4 (below) or store at –20°C or –80°C for further use in Step 3.

2.3. Step 3: Polyadenylation followed by depletion of polyadenylated RNAs

Currently, RNase R digestion together with rRNA and/or mRNA depletion is used to enrich circRNA pools before RNA-seq analysis. RNase R digests most linear RNAs, but does not digest intronic lariat RNAs, circular RNAs, or double-stranded RNAs with 3' overhangs shorter than 7 nucleotides [32, 33]. Accordingly, since snRNAs are highly structured RNAs and have 3' overhangs spanning only a few nucleotides, they are generally resistant to RNase R digestion. In this step, we add a poly(A) tail to any RNA which has a 3' end, including the RNase R-resistant snRNAs and partially digested linear RNAs. Thus, depletion of poly(A)-tailed RNAs following polyadenylation removes all linear RNAs including partially degraded rRNAs and RNase R-resistant snRNAs, leaving a highly enriched population of circRNAs.

2.3.1. Materials and equipment

- Poly(A) tailing kit (Thermo Fisher Scientific)
- RiboLock RNase inhibitor (Thermo Fisher Scientific)
- Poly(A)Purist™ MAG Kit (Thermo Fisher Scientific)
- Nuclease-free water (Thermo Fisher Scientific)
- Magnetic separation stand (New England Biolabs)
- Thermomixer (Eppendorf)
- miRNeasy Kit (QIAGEN)
- Vortexer
- Microcentrifuge (Eppendorf)

2.3.2. Detailed protocol

1. Prepare a 40-μL polyadenylation reaction with 20 μL of RNase R-treated RNA, 4 μL of 10× poly(A) polymerase buffer, 1 μL of poly(A) polymerase (E-PAP), 4 μL of 10 mM ATP solution (provided with Poly(A) tailing kit), 4 μL of 25 mM MnCl₂ (provided in the Poly(A) tailing kit) and 1 μL of RiboLock RNase inhibitor.
2. Prepare a control reaction exactly the same as the polyadenylation reaction but without E-PAP.
3. Mix well by flicking the tube with a finger and incubate reactions at 37°C for 30 min.
4. While the polyadenylation reaction is ongoing, dissolve 10 μL of well-suspended oligo-dT magnetic beads from Poly(A)Purist™ MAG Kit with 100 μL of 1 × binding buffer provided in the kit and incubate on the magnetic stand for 1 min. Discard the supernatant and repeat this step 2 additional times.
5. Dissolve the washed beads in 40 μL of 2× binding buffer provided with Poly(A)Purist™ MAG Kit and add to the 40-μL polyadenylation reaction.
6. Incubate the polyadenylation reaction with magnetic beads and control reaction for 5 min at 75 °C followed by incubation for 20 min at 25°C with shaking to allow the oligo-dT beads to bind the poly(A)-tailed RNAs.
7. Incubate the RNA sample with Oligo-dT Dynabeads on the magnetic stand for 1 min and collect the supernatant for RNA isolation.
8. Dissolve the supernatant in 500 μL of QIAzol lysis reagent and use the miRNeasy Mini Kit to prepare RNA from the control and RPAD samples and dissolve it in 40 μL of nuclease-free water.

2.4. Step 4: RT-qPCR analysis to test if linear RNAs are depleted

RNase R is known to degrade linear RNAs leaving circRNAs intact. To validate the depletion of linear RNAs and the resistance of circRNAs and snRNAs to RNase R treatment in Step 2, we performed RT-qPCR analysis. In Step 3, we expect to deplete all linear RNAs including the snRNAs and partially digested RNAs by carrying out the addition of a poly(A) tail and the removal of poly(A)⁺ RNAs. To validate the depletion of linear RNAs by RPAD, RT-qPCR was performed for a subset of linear RNAs including the snRNAs on the RPAD-processed RNAs isolated in Step 3. This is the most crucial step in validating the depletion of linear RNAs before proceeding to any downstream analysis.

2.4.1. Materials and equipment

1. Control RNA, RNase R-treated RNA, and RPAD-prepared RNA
2. Random primers (Thermo Fisher Scientific)
3. dNTP mix (Thermo Fisher Scientific)
4. Ribolock RNase inhibitor (Thermo Fisher Scientific)
5. Maxima Reverse Transcriptase (Thermo Fisher Scientific)
6. Mir-X™ miRNA First Strand Synthesis Kit (Clontech, Takara Bio)
7. MicroAmp® Optical 96-Well Reaction Plate (Thermo Fisher Scientific)
8. MicroAmp® Optical Adhesive Film (Thermo Fisher Scientific)
9. MPS 1000 Mini Plate Spinner
10. KAPA SYBR® FAST qPCR mix (Thermo Fisher Scientific)
11. Thermomixture (Eppendorf)
12. Veriti® 96-Well Thermal Cycler (Thermo Fisher Scientific)
13. QuantStudio 5 Real-Time PCR System (Thermo Fisher Scientific)

2.4.2. Detailed protocol

1. Design forward and reverse primers for linear RNAs using the Primer 3 online tool (<http://bioinfo.ut.ee/primer3/>) [34]. Design the junction-spanning divergent primers for the circRNA of interest using the CircInteractome web tool (http://circinteractome.irp.nih.gov/Divergent_Primers/divergent_primers.html) [35]. Prepare a 1 µM working primer mix of forward and reverse primers in nuclease-free water for using in the qPCR reaction.
2. Take equal volumes (2 µL) of control RNA and RNase R-treated RNA samples, and 3.75 µL of RNase R-treated RNA, control RNA, and RPAD-prepared RNA for small RNA cDNA synthesis using the Mir-X™ miRNA first strand synthesis kit following the manufacturer's protocol (Fig. 2). Briefly, prepare a 10-µL reaction mix containing 5 µL of 2× mRQ buffer and 1.25 µL of mRQ enzyme mix provided with the kit along with the RNA samples and nuclease-free water to adjust the volume to 10 µL. Mix the reaction well and incubate the reaction at

37°C for 1 h for cDNA synthesis. Terminate the cDNA synthesis by incubating at 85°C for 5 min to inactivate the enzymes. Dilute the cDNA with 500 µL of nuclease-free water.

3. For long cDNA synthesis from long RNAs, take equal volume (5 to 10 µL) of control RNA, RNase R-treated RNA, and RPAD-treated RNA to perform cDNA synthesis using Maxima reverse transcriptase. Prepare a 20-µL RT reaction mix with the above amounts of prepared RNA, 4 µL 5× RT buffer, 1 µL dNTP mix, 1 µL RiboLock, 1 µL random hexamer primers, and 1 µL Maxima reverse transcriptase. Mix the reaction gently and incubate it for 10 min at 25 °C and for 30 min at 50 °C. Inactivate the reverse transcriptase by incubating the reaction at 85 °C for 5 min. Dilute the cDNA with 500 µL of nuclease-free water [36].
4. Prepare a 20-µL qPCR reaction containing 5 µL cDNA, 5 µL primer mix, and 10 µL of 2× SYBR Green mix.
5. Seal the plate with the optical adhesive film and mix the reaction plate well for a few seconds followed by a short spin.
6. Set up the qPCR reaction cycle as follows: 40 cycles of 2 min at 95 °C, 2 sec at 95 °C, and 10 sec at 60 °C using QuantStudio 5 Real-Time PCR System.
7. Use the delta Ct method to calculate the percentage (%) of linear RNA left after RNase R digestion and RPAD treatment [36].

3. EXPERIMENTAL RESULTS AND DISCUSSION

Depletion of linear RNA

The exonuclease RNase R has been used widely to deplete the linear RNAs in the total RNA pool leaving circRNAs and intronic lariats intact [32]. RNase R treatment of total RNA from HeLa cells degraded most of the linear RNAs. RT-qPCR analysis method was used to validate the extent of elimination of linear RNAs following RNase R digestion. A limited screening was done for linear and circRNAs by RT-qPCR analysis using convergent and divergent primers, respectively. Although RNase R was expected to degrade the linear RNAs completely, we found that 5-35% of linear RNAs remained undigested (Fig. 3A). Since circRNAs are not expected to be affected by RNase R treatment, we calculated the mRNA levels normalized to their corresponding circRNAs in RNase R-treated samples to account for the variability introduced during RNase R treatment or RNA isolation or cDNA synthesis (Fig. 3A). As highly structured RNAs like snRNAs without long 3' overhangs are expected to be resistant to RNase R digestion, we measured the depletion of small RNAs by RT-qPCR analysis. Even after using higher concentration of RNase R, we were unable to degrade *U1* snRNA, while other small RNAs like *U6* snRNA and *5S* rRNA were depleted down to 20% compared with no RNase R treatment (Fig. 3B). As small RNAs such as *5S* rRNA, snRNAs, and tRNAs constitute the majority of the small-RNA fraction in cells and circRNAs are generally low-abundance, the presence of residual undigested linear RNA could hamper downstream analysis like RNA-seq for detecting circRNAs.

Isolation of highly pure circRNAs

As RNase R treatment alone is not sufficient to deplete the linear RNAs completely, an additional step of polyadenylation and depletion of the poly(A)⁺ RNAs was introduced; we termed it 'RNase R, polyadenylation and poly(A)⁺ RNA depletion' (RPAD) (Fig. 1). The RNA remaining after RNase R treatment is expected to contain circRNAs and a fraction of partially digested and RNase R-resistant linear RNAs. The linear and circRNAs left after RPAD treatment were assessed by RT-qPCR analysis. Poly(A)-tailing followed by poly(A)⁺ RNA depletion was able to deplete mRNAs and rRNAs further, leaving only 10-60% of the RNAs present after digestion with RNase R (Fig. 4A). Furthermore, RPAD treatment depleted small RNAs down to 10-50% of the levels remaining after RNase R treatment (Fig. 4B). Thus, the RNAs left after RPAD contained a far more highly pure circRNA population for downstream analysis.

4. TROUBLESHOOTING

4.1. Low signals or no amplification.

The target RNA may not be expressed or may be present in very low amounts, or the primers may not be suitable for amplification. Use different target RNAs for validation, change primer pairs, or use low-stringency settings in the PCR amplification cycle. Use well-established and previously validated mRNAs and circRNAs that are known to be expressed in the cells used for validation.

4.2. No depletion of mRNA after RNase R treatment.

The target linear mRNA may be highly structured and the 3' end may not be accessible for the exonucleolytic activity of RNase R. Heat-denaturation of total RNA prior to RNase R digestion can help. It is particularly important to recognize that convergent primers targeting mRNAs may be amplifying the body of one or more circRNAs originating from the same pre-mRNA. Using a primer set directed at mRNA sequences outside of the circRNA sequence may result in specific detection of mRNAs.

4.3. mRNA and snRNA remaining after RPAD treatment.

Residual linear RNAs may remain in the RNA sample due to inadequate polyadenylation or inefficient depletion of poly(A)⁺ RNAs. Increasing the polyadenylation time or changing the concentration of poly(A) polymerase would help to increase in polyadenylation of linear RNAs. Inefficient depletion of linear RNA after polyadenylation may be due to reduced binding of oligo-dT beads to the polyadenylated RNAs. Shaking the beads during the incubation with polyadenylated RNA and increasing the time of incubation at 25 °C will increase the capture of poly(A)⁺ RNAs by the beads.

4.4. Low RNA amounts left after RPAD treatment.

As the RPAD method involves several enzymatic treatments and RNA isolation steps, there is a high chance of getting RNase contamination or losing RNAs during the RNA isolation steps. Cleaning the work area with RNaseZap (Thermo Fisher Scientific) or 0.1 M NaOH

will remove some of the RNase contamination. Increasing the initial amount of total RNA used for RNase R digestion will likely increase the yield of RNA after RPAD treatment.

5. ADAPTATIONS OF THIS METHODOLOGY

5.1. True circRNA identification.

As RPAD depletes the linear RNAs to minimal levels, the chance of identifying low-abundance and novel circRNAs increases. RPAD-selected RNAs may be used directly to prepare cDNA libraries and perform high-throughput RNA-seq analysis to identify novel circRNAs, as the RNA-seq reads coming from linear RNA has been reduced drastically. In our recent report, we identified thousands of novel circRNAs in HeLa cells and in mouse C2C12 myoblasts which were not previously identified in total RNA or RNase R-treated RNAs [24].

5.2. Identification of mature circRNA sequences.

CircRNA sequences are bioinformatically predicted considering the intervening exons present in between the backsplice junction coordinate of the circRNA. However, the complete sequence of the mature circRNA after backsplicing may be different in depending on the cell or tissue leading to different circRNA isoform biogenesis. The assembly of RNA-seq reads between the backsplice junction coordinates can accurately identify the mature full-length body sequence of circRNA. Using this method in HeLa and C2C12 cells, we identified several circRNAs that had different mature sequences than those previously reported sequences [24].

5.3. Identification of the *in vitro* circRNA-RBP interactome.

To-date, circRNA-RBP interactions were only predicted from CLIP-type datasets (PAR-CLIP, iCLIP, etc) performed on whole-cell transcriptome. Such RBP binding sites in these datasets may belong to linear RNAs or circRNAs. RPAD-purified RNAs can be used to perform *in vitro* CLIP analyses to identify the binding sites of RBPs on circRNAs.

5.4. *In vitro* circRNA translation.

Although circRNAs are believed not to be translated due to a lack of 5'-cap and poly(A) tail, there is emerging evidence that some circRNAs can be translated into functional proteins. To identify the whole pool of translated circRNAs, RPAD-enriched circRNAs can be translated in a cell-free system. Mass spectroscopy to identify novel polypeptides in *in vitro* translation reactions followed by bioinformatic identification of potential circRNAs can be used to identify translated circRNAs.

5.5. Identification of circRNAs in non-model organisms.

Conventional methods of circRNA identification by RNA-seq include enrichment of circRNA population by RNase R treatment and by depletion of rRNA. Currently, the rRNA depletion kits use probes specific for rRNAs of model organisms like human, mouse, and rat, but for many non-model organisms, rRNA depletion kits are not available at present. For these reasons, RNase R treatment alone is not sufficient to deplete all the linear RNAs which

may interfere with the downstream RNA-seq analysis of circRNAs. Given that RPAD treatment efficiently removes linear RNAs regardless of the RNA sequence, RPAD-enriched circRNA pool can be used for high-throughput RNA-seq to identify true circRNAs in non-model organisms.

6. LIMITATIONS AND CONCLUDING REMARKS

The list of circRNAs is on the rise with advances in high-throughput RNA-seq technologies and bioinformatic tools. There is escalating interest to identify and characterize the circRNAs, with emerging evidence that they can regulate various cellular events important for disease diagnosis and treatment. As RNase R is unable to remove linear RNAs completely and circRNAs are generally less abundant than their counterpart linear RNAs, small percentages of residual undigested linear RNAs in RNase R-treated RNA pools could hamper the identification of low-abundance circRNAs as well as the identification of the full-length sequence of circRNAs. The RPAD method described here could potentially overcome these issues and help identify novel circRNAs as well as full-length sequence of mature circRNAs. A noteworthy drawback is that since RPAD involves many enzymatic and RNA isolation steps, it may not be the strictly accurate for quantification of circRNA abundance. In closing, while RNase R digestion may eliminate mRNAs efficiently, RPAD represents an improvement for eliminating other linear RNAs, particularly those bearing extensive secondary structures that render them resistant to RNase R digestion. In sum, the RPAD method enables the systematic identification and characterization of highly enriched, highly pure populations of circRNAs.

ACKNOWLEDGEMENTS

PP, ACP, and MG were supported by the National Institute on Aging Intramural Research Program, National Institutes of Health, in Baltimore, Maryland, USA. PR and AD were supported by the Institute of Life Sciences Intramural Research Program, Department of Biotechnology, Government of India. ACP was supported by the Science & Engineering Research Board, a statutory body of the Department of Science & Technology (DST), Government of India (SERB/F/6890/2017-18).

REFERENCES

- [1]. Crick F, Central dogma of molecular biology, *Nature* 227(5258) (1970) 561–3. [PubMed: 4913914]
- [2]. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhata R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J, Guigo R, The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression, *Genome Res* 22(9) (2012) 1775–89. [PubMed: 22955988]
- [3]. Palazzo AF, Lee ES, Non-coding RNA: what is functional and what is junk?, *Front Genet* 6 (2015) 2. [PubMed: 25674102]
- [4]. Wery M, Kwapisz M, Morillon A, Noncoding RNAs in gene regulation, *Wiley Interdiscip Rev Syst Biol Med* 3(6) (2011) 728–38.
- [5]. Panda AC, Grammatikakis I, Munk R, Gorospe M, Abdelmohsen K, Emerging roles and context of circular RNAs, *Wiley Interdiscip Rev RNA* 8(2) (2017).
- [6]. Sanger HL, Klotz G, Riesner D, Gross HJ, Kleinschmidt AK, Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures, *Proc Natl Acad Sci U S A* 73(11) (1976) 3852–6. [PubMed: 1069269]

- [7]. Hsu MT, Coca-Prados M, Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells, *Nature* 280(5720) (1979) 339–40. [PubMed: 460409]
- [8]. Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO, Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types, *PLoS One* 7(2) (2012) e30733. [PubMed: 22319583]
- [9]. Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF, Sharpless NE, Circular RNAs are abundant, conserved, and associated with ALU repeats, *RNA* 19(2) (2013) 141–57. [PubMed: 23249747]
- [10]. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J, Natural RNA circles function as efficient microRNA sponges, *Nature* 495(7441) (2013) 384–8. [PubMed: 23446346]
- [11]. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, Loewer A, Ziebold U, Landthaler M, Kocks C, le Noble F, Rajewsky N, Circular RNAs are a large class of animal RNAs with regulatory potency, *Nature* 495(7441) (2013) 333–8. [PubMed: 23446348]
- [12]. Starke S, Jost I, Rossbach O, Schneider T, Schreiner S, Hung LH, Bindereif A, Exon circularization requires canonical splice signals, *Cell Rep* 10(1) (2015) 103–11. [PubMed: 25543144]
- [13]. Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S, circRNA biogenesis competes with pre-mRNA splicing, *Mol Cell* 56(1) (2014) 55–66. [PubMed: 25242144]
- [14]. Zhang Y, Xue W, Li X, Zhang J, Chen S, Zhang JL, Yang L, Chen LL, The Biogenesis of Nascent Circular RNAs, *Cell Rep* 15(3) (2016) 611–624. [PubMed: 27068474]
- [15]. Ivanov A, Memczak S, Wyler E, Torti F, Porath HT, Orejuela MR, Piechotta M, Levanon EY, Landthaler M, Dieterich C, Rajewsky N, Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals, *Cell Rep* 10(2) (2015) 170–7. [PubMed: 25558066]
- [16]. Conn SJ, Pillman KA, Toubia J, Conn VM, Salmanidis M, Phillips CA, Roslan S, Schreiber AW, Gregory PA, Goodall GJ, The RNA binding protein quaking regulates formation of circRNAs, *Cell* 160(6) (2015) 1125–34. [PubMed: 25768908]
- [17]. Li X, Liu CX, Xue W, Zhang Y, Jiang S, Yin QF, Wei J, Yao RW, Yang L, Chen LL, Coordinated circRNA Biogenesis and Function with NF90/NF110 in Viral Infection, *Mol Cell* 67(2) (2017) 214–227 e7. [PubMed: 28625552]
- [18]. Haque S, Harries LW, Circular RNAs (circRNAs) in Health and Disease, *Genes (Basel)* 8(12) (2017).
- [19]. Abdelmohsen K, Panda AC, De S, Grammatikakis I, Kim J, Ding J, Noh JH, Kim KM, Mattison JA, de Cabo R, Gorospe M, Circular RNAs in monkey muscle: age-dependent changes, *Aging (Albany NY)* 7(11) (2015) 903–10. [PubMed: 26546448]
- [20]. Ghosal S, Das S, Sen R, Basak P, Chakrabarti J, Circ2Traits: a comprehensive database for circular RNA potentially associated with disease and traits, *Front Genet* 4 (2013) 283. [PubMed: 24339831]
- [21]. Han YN, Xia SQ, Zhang YY, Zheng JH, Li W, Circular RNAs: A novel type of biomarker and genetic tools in cancer, *Oncotarget* 8(38) (2017) 64551–64563. [PubMed: 28969093]
- [22]. Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, Zhong G, Yu B, Hu W, Dai L, Zhu P, Chang Z, Wu Q, Zhao Y, Jia Y, Xu P, Liu H, Shan G, Exon-intron circular RNAs regulate transcription in the nucleus, *Nat Struct Mol Biol* 22(3) (2015) 256–64. [PubMed: 25664725]
- [23]. Zhang Y, Zhang XO, Chen T, Xiang JF, Yin QF, Xing YH, Zhu S, Yang L, Chen LL, Circular intronic long noncoding RNAs, *Mol Cell* 51(6) (2013) 792–806. [PubMed: 24035497]
- [24]. Panda AC, De S, Grammatikakis I, Munk R, Yang X, Piao Y, Dudekula DB, Abdelmohsen K, Gorospe M, High-purity circular RNA isolation method (RPAD) reveals vast collection of intronic circRNAs, *Nucleic Acids Res* 45(12) (2017) e116. [PubMed: 28444238]
- [25]. Abdelmohsen K, Panda AC, Munk R, Grammatikakis I, Dudekula DB, De S, Kim J, Noh JH, Kim KM, Martindale JL, Gorospe M, Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1, *RNA Biol* 14(3) (2017) 361–369. [PubMed: 28080204]

- [26]. Panda AC, Grammatikakis I, Kim KM, De S, Martindale JL, Munk R, Yang X, Abdelmohsen K, Gorospe M, Identification of senescence-associated circular RNAs (SAC-RNAs) reveals senescence suppressor CircPVT1, *Nucleic Acids Res* 45(7) (2017) 4021–4035. [PubMed: 27928058]
- [27]. Yang Y, Fan X, Mao M, Song X, Wu P, Zhang Y, Jin Y, Yang Y, Chen LL, Wang Y, Wong CC, Xiao X, Wang Z, Extensive translation of circular RNAs driven by N(6)-methyladenosine, *Cell Res* 27(5) (2017) 626–641. [PubMed: 28281539]
- [28]. Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, Hanan M, Wyler E, Perez-Hernandez D, Ramberger E, Shenzis S, Samson M, Dittmar G, Landthaler M, Chekulaeva M, Rajewsky N, Kadener S, Translation of CircRNAs, *Mol Cell* 66(1) (2017) 9–21 e7. [PubMed: 28344080]
- [29]. Legnini I, Di Timoteo G, Rossi F, Morlando M, Briganti F, Sthandier O, Fatica A, Santini T, Andronache A, Wade M, Laneve P, Rajewsky N, Bozzoni I, Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis, *Mol Cell* 66(1) (2017) 22–37 e9. [PubMed: 28344082]
- [30]. Szabo L, Salzman J, Detecting circular RNAs: bioinformatic and experimental challenges, *Nat Rev Genet* 17(11) (2016) 679–692. [PubMed: 27739534]
- [31]. Jeck WR, Sharpless NE, Detecting and characterizing circular RNAs, *Nat Biotechnol* 32(5) (2014) 453–61. [PubMed: 24811520]
- [32]. Vincent HA, Deutscher MP, Substrate recognition and catalysis by the exoribonuclease RNase R, *J Biol Chem* 281(40) (2006) 29769–75. [PubMed: 16893880]
- [33]. Suzuki H, Zuo Y, Wang J, Zhang MQ, Malhotra A, Mayeda A, Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing, *Nucleic Acids Res* 34(8) (2006) e63. [PubMed: 16682442]
- [34]. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG, Primer3--new capabilities and interfaces, *Nucleic Acids Res* 40(15) (2012) e115. [PubMed: 22730293]
- [35]. Dudekula DB, Panda AC, Grammatikakis I, De S, Abdelmohsen K, Gorospe M, CircInteractome: A web tool for exploring circular RNAs and their interacting proteins and microRNAs, *RNA Biol* 13(1) (2016) 34–42. [PubMed: 26669964]
- [36]. Panda AC, Gorospe M, Detection and Analysis of Circular RNAs by RT-PCR, *Bio Protoc* 8(6) (2018).

- Key RPAD steps are RNase R digestion and polyadenylation and poly(A)+ RNA depletion
- Through these strategies, RPAD enables the isolation of highly pure circRNA
- RPAD eliminates linear RNA more effectively than RNase R or poly(A) RNA removal
- RNA-seq analysis of RPAD-isolated RNA can be used to uncover new circRNAs
- RNA-seq analysis of RPAD-isolated RNA can inform on mature circRNA sequences

Figure 1. Schematic of the RPAD method.

Methodology discussed in this report. Step 1, total RNA is isolated from cultured cells. Step 2, linear RNAs are depleted by RNase R digestion. Step 3, RNA left after RNase R digestion is polyadenylated and the poly(A)+ RNA is depleted using oligo(dT) beads, yielding a pool of RNA highly enriched in circRNAs. Step 4, RNAs isolated from above steps are used in RT-qPCR for validation of linear RNA depletion.

Figure 2. Schematic of cDNA synthesis for small RNAs.

In RPAD, the cDNA from small RNAs are prepared using the Mir-X™ miRNA First Strand Synthesis Kit. cDNA is synthesized in a single-step reaction where the RNAs are polyadenylated using poly(A) polymerase followed by the hybridization of the modified oligo-dT primer and cDNA synthesis by using reverse transcriptase. Quantitative (q)PCR analysis is performed using the specific primer pairs (Table 1).

Figure 3. Depletion of linear RNA by RNase R digestion.

RT-qPCR analysis of subsets of linear rRNA, mRNAs and circRNAs (**A**), as well as snRNAs and 5S rRNA (**B**) in HeLa cell RNA samples left untreated or treated with RNase R. In (**A**) since the circRNA levels were not affected by RNase R treatment, mRNAs left after RNase R treatment were normalized to the levels of the corresponding circRNAs. Data represent the means $\pm$ SEM from 4 or more experiments.

Figure 4. Validation of RPAD.

Subsets of rRNAs, circRNAs and mRNAs (**A**), as well as snRNAs and 5S rRNA (**B**) from HeLa cell RNAs were quantified by RT-qPCR analysis following RPAD. In (**A**) since circRNAs lack poly(A) tails, mRNAs left after RPAD treatment were normalized to the levels of the corresponding circRNAs. The data represent the means $\pm$ SEM from 3 or more experiments.

Table 1.

Sequences of the DNA primers used to detect HeLa cell RNAs by qPCR analysis.

Oligos for target RNAs (circBase ID)	(5'–3') sequences
<i>18S rRNA-F</i>	CGAACGTCTGCCCTATCAACTT
<i>18S rRNA-R</i>	ACCCGTGGTCACCATGGTA
<i>GAPDH-F</i>	TGCACCACCAACTGCTTAGC
<i>GAPDH-R</i>	GGCATGGACTGTGGTCATGAG
<i>circHIPK3 (hsa_circ_0000284)-F</i>	GGCAGCCTTACAGGGTTAAA
<i>circHIPK3 (hsa_circ_0000284)-R</i>	GGGTAGACCAAGACTTGTGAGG
<i>HIPK3-F</i>	GACCTGAGGAGATCAAGCCG
<i>HIPK3-R</i>	CTAGGGCGAGGGGTAGTGG
<i>circSPECC1 (hsa_circ_0000745)-F</i>	TTGCAATGAGCTCAGACAAGA
<i>circSPECC1 (hsa_circ_0000745)-R</i>	ACGTGGCACAGACCTCTCTC
<i>SPECC1-F</i>	TTCTTTGACTGGAGCGGACC
<i>SPECC1-R</i>	TTGTAAAGGCCCTGGTTCC
<i>U6-F</i>	CTCGCTTCGGCAGCACATATACT
<i>U6-R</i>	ACGCTTCACGAATTTGCGTGTC
<i>U1-F</i>	TACTTACCTGGCAGGGGAGATAC
<i>U1-R</i>	GAACGCAGTCCCCACTAC
<i>5S rRNA-F</i>	GGCCATACCACCCTGAACGC
<i>5S rRNA-R</i>	CAGCACCCGGTATTCCCAGG