Methods 67 (2014) 64–73

Contents lists available at SciVerse ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Improved genome-wide mapping of uncapped and cleaved transcripts in eukaryotes—GMUCT 2.0

Matthew R. Willmann [a,b], Nathan D. Berkowitz [a,b,c], Brian D. Gregory [a,b,c,d,*]

[a] Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA
[b] PENN Genome Frontiers Institute, University of Pennsylvania, Philadelphia, PA 19104, USA
[c] Genomics and Computational Biology Graduate Program, University of Pennsylvania, Philadelphia, PA 19104, USA
[d] Cell and Molecular Biology Graduate Program, University of Pennsylvania, Philadelphia, PA 19104, USA

ARTICLE INFO

Article history:
Available online 15 July 2013

Keywords:
Post-transcriptional regulation
RNA degradation
RNA silencing
microRNAs
Small interfering RNAs

ABSTRACT

The advent of high-throughput sequencing has led to an explosion of studies into the diversity, expression, processing, and lifespan of RNAs. Recently, three different high-throughput sequencing-based methods have been developed to specifically study RNAs that are in the process of being degraded. All three methods—genome-wide mapping of uncapped and cleaved transcripts (GMUCT), parallel analysis of RNA ends (PARE), and degradome sequencing—take advantage of the fact that Illumina sequencing libraries use T4 RNA ligase 1 to ligate an adapter to the 5′ end of RNAs that have a free 5′-monophosphate. This condition for T4 RNA ligase 1 substrates means that mature mRNAs are not substrates of the enzyme because they have a 5′-cap moiety. As a result, these sequencing libraries are specifically made up of clones of decapped or degrading mRNAs resulting from 5′-to-3′ or nonsense-mediated decay (NMD) and the 3′ fragment of cleaved microRNA (miRNA) and small interfering RNA (siRNA) target RNAs. Here, we present a massively streamlined protocol for GMUCT that takes 2–3 days, can be initiated with as little as 5 μg of starting total RNA, and involves only one gel size-selection step. We show that the resulting datasets are similar to those produced using the previous GMUCT and PARE protocols. In total, our results suggest that this method will be the preferable approach for future studies of RNA degradation intermediates and small RNA-mediated cleavage in eukaryotic transcriptomes.

1. Introduction

Genomic studies of gene expression have primarily focused on analyses of total RNA abundance and regulation of RNA expression at the level of transcription. The advent of high-throughput sequencing has led to an explosion of genomic techniques for studying very specific populations of RNAs, allowing for a greater understanding of how gene expression can be affected globally by the rate of transcription, alternative splicing, RNA silencing, RNA stability, translation, and chromatin modifications.

Recently, three different high-throughput sequencing based methods (genome-wide mapping of uncapped and cleaved transcripts (GMUCT), parallel analysis of RNA ends (PARE), and degradome sequencing) have been developed to specifically study the degradation of mRNAs [1–3]. The pathways that degrade mRNAs are highly conserved in eukaryotes. Most turnover of normal, functional mRNAs occurs by 5′-to-3′ degradation, 3′-to-5′ decay, or RNA silencing [4–8]. Aberrant mRNAs are degraded by one of three pathways—nonsense-mediated mRNA decay (NMD), non-stop decay (NSD), and no-go decay (NGD) [9–12]. Before discussing the high-throughput sequencing methods, we will briefly consider each of the decay pathways in turn.

5′-to-3′ and 3′-to-5′ decay of normal mRNAs both involve exoribonucleases, RNases that chew the ends of RNAs processively. Normal mRNAs are stabilized and protected from degradation by a 5′-m7GpppN cap and a 3′-poly-A tail. Most methods of mRNA degradation begin with the removal of one or the other of these protective elements. For instance, 5′-to-3′ decay begins with deadenylation of the poly-A tail in processing bodies (p-bodies) [4,9,13]. Then, the 5′-cap is also removed in p-bodies by DCP2 with the help of DCP1 and VCS, yielding an mRNA with a 5′-monophosphate [9,14,15]. The 5′-to-3′ RNases, XRN family members, then degrade the deadenylated and decapped mRNA by processively removing one nucleotide at a time from the 5′ end. Each intermediate product in the degradation of the mRNA by XRN RNases is an RNA with a 5′-monophosphate [4,9,16]. By contrast, 3′-to-5′ decay requires deadenylation but does not require decapping. After deadenylation, most 3′-to-5′ mRNA decay is done by the exosome, a multisubunit protein complex that has 3′-to-5′ ribonuclease activity that

* Corresponding author. Address: Department of Biology, University of Pennsylvania, 433 S. University Ave., Philadelphia, PA 19104, USA. Fax: +1 215 898 8780.
E-mail addresses: willmann@sas.upenn.edu (M.R. Willmann), nberk@mail.med.upenn.edu (N.D. Berkowitz), bdgregor@sas.upenn.edu (B.D. Gregory).

1046-2023/$ - see front matter © 2014 Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.ymeth.2013.07.003

removes one nucleotide at a time from the 3′ end [4,5,9,17]. The final step in this process is the hydrolysis of the 5′-cap moiety by DcpS [4,5].

The turnover of functional mRNAs can be initiated by endoribonucleases as well as exoribonucleases [5–9]. The most well-studied pathway of endoribonuclease-mediated degradation of normal mRNAs is RNA silencing. The endoribonucleases involved in RNA silencing are Argonaute (AGO) family proteins. In RNA silencing, small RNAs (smRNAs) including microRNAs (miRNAs) and small interfering RNAs (siRNAs) can initiate AGO-mediated cleavage of target RNAs with high complementarity through specific base pairing interactions [5–8]. The 3′ product of this cleavage reaction has a resulting monophosphate on its 5′ end, while the 5′ intermediate lacks a poly-A tail [18]. Thus, both cleavage products are susceptible to further degradation by either 5′-to-3′ or 3′-to-5′ decay, respectively [5–8].

All three mechanisms that act on aberrant mRNAs recognize specific mRNAs as abnormal during the process of translation and then elicit their degradation. NMD-targeted mRNAs include those that have early termination codons resulting in a truncated open reading frame, upstream open reading frames (uORFs), a mutated poly-A site preventing proper polyadenylation, mRNAs subject to alternative polyadenylation due to the presence of multiple poly-A sites, introns in the 3′-UTR, and very long 3′-UTRs, though not all mRNAs with lengthy 3′-UTRs are equally susceptible to NMD [9,10,12,19]. NMD occurs by first decapping the mRNA, which is done by the mRNA decapping complex as described above for 5′-to-3′ degradation of normal mRNAs. The decapped mRNA is then degraded by members of the XRN family of 5′-to-3′ exoribonucleases, as for 5′-to-3′ decay of normal mRNAs, or becomes a target of RNA-dependent RNA polymerases, which convert them into double-stranded RNA [9,10,12,19]. The double-stranded RNA is then subject to Dicer-like endoribonucleases that cleave it into 21–24 nucleotide (nt) small-interfering RNAs that can go on to silence *cis* targets with perfect complementarity [9].

Non-stop decay (NSD) and no-go decay (NGD) also recognize aberrant mRNAs but act on very different classes of mRNAs than NMD [5,7,11,20,21]. For instance, NSD acts on mRNAs for whom translation continues through the poly-A tail due to the absence of a stop codon [5,11,21]. Translation through the poly-A tail results in the addition of a string of lysines to the polypeptide because the codon AAA encodes a lysine, and the ribosome is not released from the 3′ end of the mRNA due to the lack of a stop codon. The string of lysines initiates NSD in eukaryotes by stimulating the degradation of the nascent polypeptide by the proteasome, endonucleolytic cleavage of the mRNA near the ribosome, and the degradation of the mRNA by exosomes without the removal of the 5′-cap. NGD recognizes ribosomes that stall in the process of translation elongation, and, like NSD, initiates the degradation of the polypeptide, endonucleolytic cleavage of the mRNA near the stalled ribosome, and 3′-to-5′ decay of the mRNA by the exosome. Like NSD, the 5′-cap is not removed during NGD [11,20].

In total, the GMUCT, PARE, and degradome approaches are useful for studying 5′-to-3′ decay, small RNA-mediated target cleavage, and NMD, but not 3′-to-5′ decay, NSD, or NGD because a key enzyme (T4 RNA ligase 1) used to construct the library of molecules to be sequenced requires an available 5′-monophosphate on every RNA to be cloned. This requirement also means that mature, functional mRNAs are not substrates of these high-throughput methods because they have a 5′-cap. As a result, these sequencing libraries are specifically made up of clones of decapped or degrading mRNAs from 5′-to-3′ exoribonucleases, the 3′ fragment of cleaved miRNA and siRNA targets, and NMD decaying mRNAs.

Here, we present a streamlined method for GMUCT that we call GMUCT 2.0. The new method reduces the time necessary to make a GMUCT library from 5–6 to 2–3 days and decreases the amount of starting total RNA from 50 μg to 5 μg [1,22]. We use this protocol to make libraries of degrading and cleaved RNAs isolated from plant tissue and human cell lines. We also compare the data generated from this improved, more time-efficient method to data obtained using the original GMUCT [1], PARE [3], and degradome sequencing [2] approaches to show that the data produced is similar. Thus, we have significantly improved a methodology that will be widely useful for future studies of mRNA turnover in all eukaryotic organisms.

2. GMUCT 2.0 protocol

The new method for making GMUCT sequencing libraries is summarized in Fig. 1, where it is also compared to the original GMUCT method [1,22].

2.1. Materials

This protocol is for making GMUCT 2.0 libraries for Illumina TruSeq sequencing. The following materials are required.

Item	Manufacturer	Catalog Number
miRNeasy Mini Kit	Qiagen	217004 for 50. For smaller samples, use the miRNeasy Micro Kit (217084)
Dynabeads mRNA DIRECT Kit	Invitrogen	61011 for 20 preps or 61012 for 40 preps. A magnetic stand that holds 1.5 ml tubes is required to use this kit.
TruSeq Small RNA Sample Prep Kit	Illumina	Each kit is for 24 samples and includes primers for 12 of the available 48 indices. RS-200-0012 (comes with indices RPI1-12), RS-200-0024 (comes with indices RPI13-24), RS-200-0036 (comes with indices RPI25-36), RS-200-0048 (comes with indices RPI37-48). Each kit also comes with the RA5 adapter, T4 DNA ligase 1, ligation buffer HML, 10 mM ATP, PCR mix PML, RP1 primer, RNase inhibitor.
HPLC-purified random hexamer primer that also includes the TruSeq RA3 3′-adapter on the 5′ end of the primer 5′-ctggagttccttggcacccgagaattccannnnnn-3′		Special primer order from Integrated DNA Technologies (IDT)

(*continued on next page*)