Combined mRNA & microRNA NGS Library Prep Enables a more Complete Characterization of Cell-free RNA

K.D. Allen; A.R. Morris; S.N. Piehl; K. Kushiro; A. Kothandaraman
Kevin.Allen@perkinelmer.com

Bioo Scientific, a PerkinElmer company
7050 Burleson Rd.
Austin, TX 78744

ABSTRACT

Historically, most RNA-seq experiments have been performed with the goal of studying mRNAs. Recently, deep sequencing of small RNAs, in particular miRNAs, has gained in popularity as researchers have begun to appreciate that miRNA expression can also be used to create gene signatures and that miRNAs have significant impact on regulation of gene expression. Researchers are increasingly interested in profiling both mRNA and miRNA content of a sample, but no simple method currently exists for a combined library preparation. The methods currently in existence typically require splitting a sample into two fractions, isolating mRNA in one fraction, fragmenting the mRNA, treating the fragmented mRNA with T4 PNK, then recombining the fractions and using a small RNA library preparation protocol with direct ligation of adapters to RNA molecules.

Here, a combined mRNA/miRNA library preparation protocol is described that is streamlined, does not require rRNA depletion or poly(A) selection, and does not require splitting the sample into separate fractions. It relies upon a novel technique in which poly(A)-tailed RNA species are selectively reverse transcribed, then sheared by RNase H into fragments of useful length. Sequencing results show high concordance of RNA abundance between this method and traditional mRNA-Seq protocols, as well as high replicability.

This method will be especially valuable to researchers studying circulating, or cell-free RNA. Both mRNA and miRNA biomarkers are known to be present in biofluids, however due to their low concentration, no practical method yet exists for sequencing both RNA species simultaneously. As a proof of principle, we show that this previously presented method can also be employed to generate combined mRNA/miRNA libraries from human plasma. We expect that this method will become a powerful new tool for researchers interested in biomarker discovery in this first report on data generated using this method.

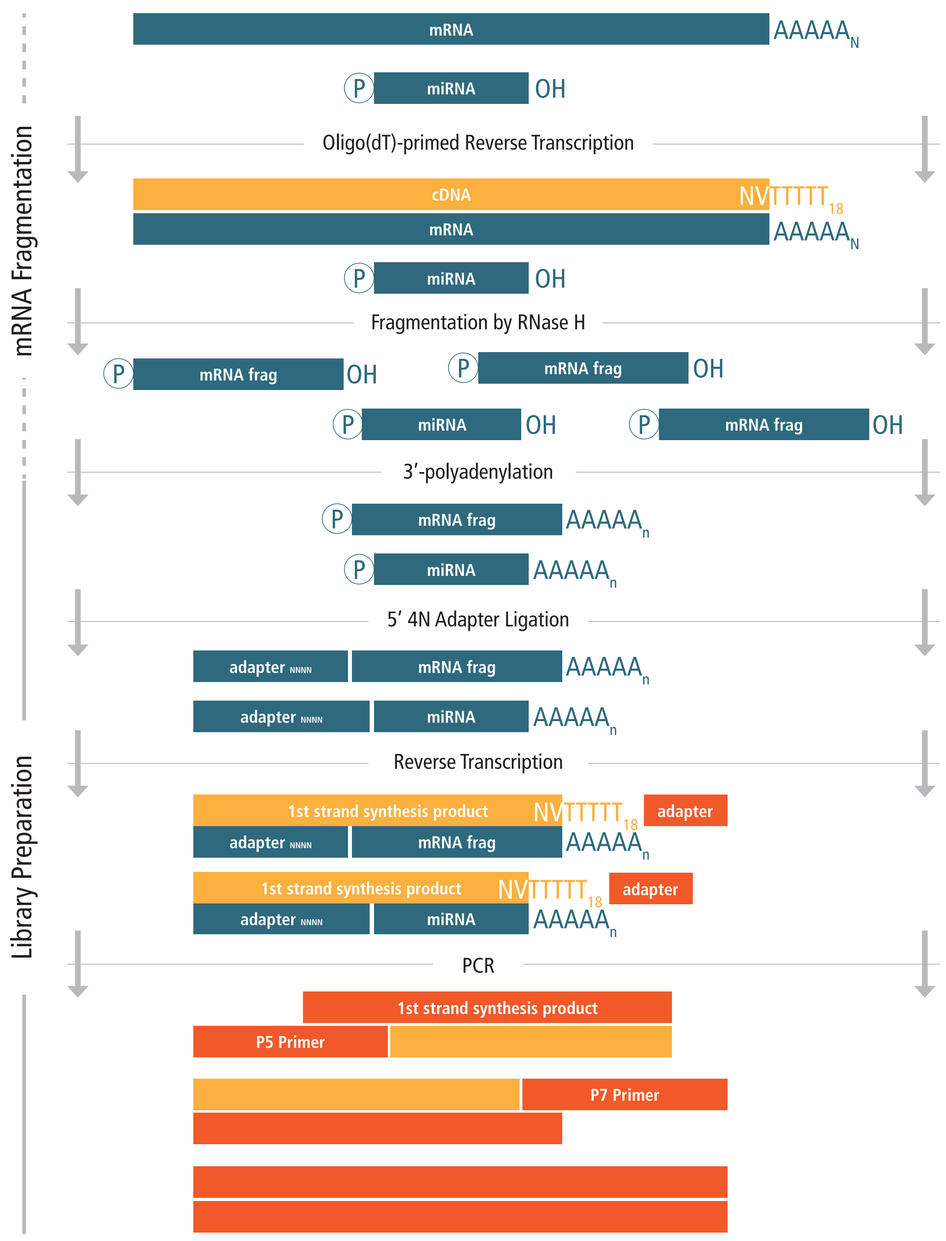
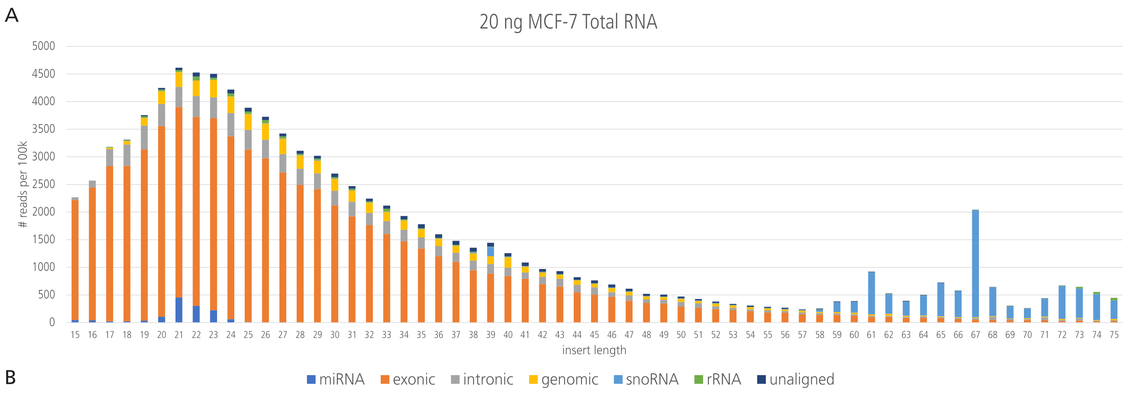


Figure 1. The NEXTFLEX® Combo-Seq™ Kit is a simple, streamlined, and novel workflow for combined mRNA/miRNA NGS library preparation that preserves strand orientation. *The workflow enables the user to generate combined mRNA/miRNA NGS libraries from 1 ng – 100 ng of total RNA in approximately 7 hours. The workflow consists of two modules: mRNA fragmentation and library preparation. During the mRNA fragmentation module, oligo(dT)-primed reverse transcription first generates DNA:RNA duplexes from poly(A)-tailed RNA species in the sample. Next, an attenuated RNase H reaction specifically shears RNA molecules that are complexed with cDNA. The mRNA fragments generated by RNase H feature a 5′-monophosphate and a 3′-hydroxyl, which allows them to be ligated and polyadenylated in the following steps along with miRNAs and other small RNAs.*

During the library preparation module, mRNA fragments and small RNAs are first 3′-polyadenylated. NEXTFLEX® Combo-Seq™ 5′ 4N adapters featuring bias-reducing randomized ends are then ligated to their 5′ ends. First strand synthesis is performed using an anchored oligo(dT) primer with an overhanging adapter sequence. Finally, PCR amplifies the first-strand synthesis product and adds Unique Dual Indices (UDIs) along with other sequences necessary for Illumina sequencing.

Following library sequencing, small RNAs and mRNA fragments can be distinguished from one another computationally via sequence similarity, as outlined in Figure 2. Because distinct 3′ and 5′ adapters are added to either end of an mRNA fragment or miRNA, this method is inherently directional and is capable of identifying the specific DNA strand a given RNA transcript was derived from with >99.9% strand specificity. The workflow described in this figure has been presented previously [1, 2].

A



B

miRNA	exonic	intronic	genomic	snoRNA	rRNA	unaligned
1.4%	67.3%	9.3%	6.4%	9.7%	3.7%	2.2%

Figure 2a,b. The NEXTFLEX® Combo-Seq™ library prep method yields a high mapping rate to mRNAs, miRNAs, and other small RNAs while minimizing rRNA alignment and unaligned reads. *Libraries were prepared from 20 ng of MCF-7 total RNA and sequenced on an Illumina MiSeq. Reads were adapter-trimmed with cutadapt, and reads with inserts <15 nt in length were excluded, resulting in a loss of 6.3% of reads. Trimmed reads were aligned consecutively to the following references, with aligned reads removed at each step: human miRNAs from miRBase release 21, concatenated exons from the longest isoform of each transcript in GENCODE release 27, concatenated exons and introns from the longest isoform of each transcript in GENCODE release 27, human snoRNAs from snoRNABase release 3, human rRNAs from the UCSC Table Browser, and the human genome annotation hg38. All alignments except the alignment to hg38 were performed using bowtie2 in "no-rc" mode, which allows only properly directional reads to align. **(a)** Mapped identity of reads is shown as a function of insert length. **(b)** Overall alignment rates to each reference are shown.*

The data described in this figure are representative of results that have been presented previously [1, 2].

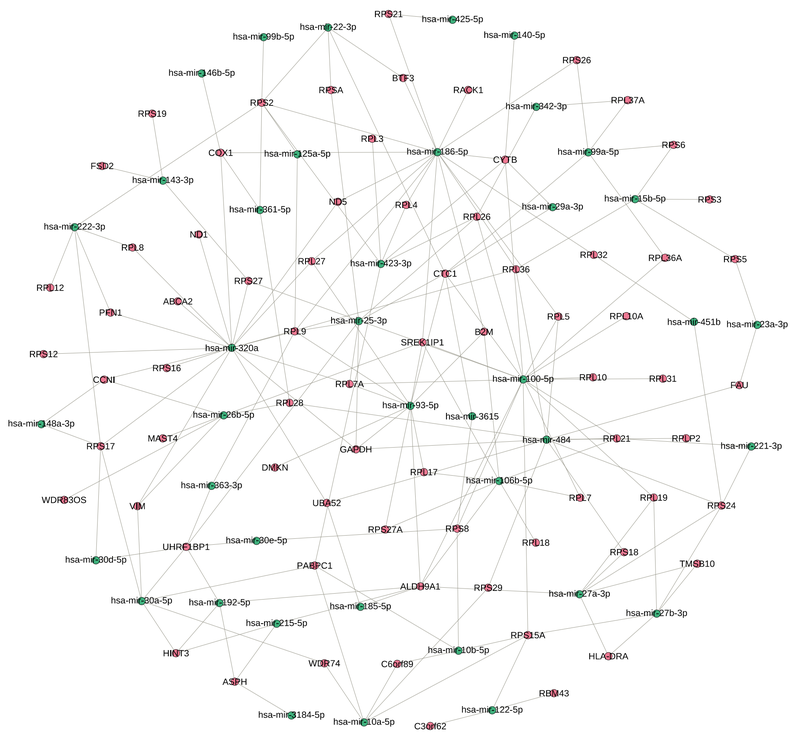


Figure 3. Sequencing data generated by the NEXTFLEX® Combo-Seq™ mRNA/miRNA Kit can be used to simultaneously probe the mRNA and miRNA content of cell-free RNA. *RNA was isolated from 600 uL of human plasma using the NextPrep™ Magnazol™ cfRNA Isolation Kit. Libraries were prepared according to the NEXTFLEX® Combo-Seq™ protocol, following the procedure for depletion of abundant tRNA fragments and Y RNA fragments using the included NEXTFLEX® tRNA/YRNA Blocker. Reads were sequenced on an Illumina® MiSeq® instrument, then processed and mapped as described in the previous figure.*

mRNA and miRNA transcripts found to be abundant at >= 5 reads per 100k were then referenced against known and predicted miRNA-target interactions using miRNet [3]. The largest contiguous interaction network that was generated is shown here.

References

1. *Allen K.D., Morris A.R., Combined mRNA & miRNA NGS Library Preparation in a Single-Tube Format without rRNA Depletion or Poly(A) Selection. Presented at the Advances in Genome Biology and Technology General Meeting, February 13, 2018, Orlando, Florida.*
2. *Allen K.D., Morris A.R., Combined mRNA & miRNA NGS Library Preparation in a Single-Tube Format without rRNA Depletion or Poly(A) Selection. Presented at the Association of Biomolecular Research Facilities Annual Meeting, April 23, 2018, Myrtle Beach, South Carolina.*
3. *Fan Y, Siklenka, K., Arora, SK., Ribeiro, P., Kimmins, S. and Xia, J. (2016) miRNet - dissecting miRNA-target interactions and functional associations through network-based visual analysis. Nucl. Acids Res. 44 W135–141.*

CONCLUSIONS

- **Libraries generated from low-input starting material: 1 ng – 100 ng total RNA or 600 uL plasma**
- **Entire workflow can be completed in approximately 7 hours**
- **Low rRNA mapping rate without the need for poly(A) selection or rRNA depletion**
- **Preserves strand orientation information with >99.9% accuracy**
- **Sequence mRNAs as well as small RNAs such as miRNAs, piRNAs, and snoRNAs**
- **Included NEXTFLEX® tRNA/YRNA Blocker reduces mapping to abundant species in cfRNA**