Bioo Scientific®

Advances in Small RNA Library Preparation Allow Combination of Bias Reduction with Gel-free or Low Input Protocols

Adam Morris, Ph.D., Kevin Allen, Suk Ho Eun, Ph.D.,
Masoud Toloue, Ph.D.*
Bioo Scientific Corp., 7050 Burleson Rd.
Austin, TX 78744 USA
mtoloue@biooscientific.com

INTRODUCTION

Small RNA discovery and profiling by next generation sequencing (NGS) is an ideal application of NGS technology, particularly for poorly-annotated species. Unlike methods like qPCR and microarray, NGS can be used to study small RNA species with no a priori information on the sequences of the small RNAs in that species, making it ideal to study small RNAs in non-model organisms.

However, small RNA sequencing has typically suffered from three major drawbacks:

1. Severe bias, such that sequencing data does not reflect original miRNA abundances
2. The need to gel purify final libraries
3. Lack of low-input protocols

Bias is introduced into small RNA libraries during the ligation steps. Small RNA library preparation involves ligating adapters directly on to the 3' and 5' ends of the RNA molecule in two separate steps, and each of these ligation steps has been shown to introduce severe bias into library preparation. In an effort to reduce this bias, it was discovered that inserting random bases onto the ends of the adapters greatly reduces bias in comparison to using non-randomized adapters [1-4].

Another major drawback of small RNA sequencing is the need to purify final libraries by gel, typically by PAGE gel. This is due to the small difference in the size of adapter-dimer molecules versus insert-containing molecules following the PCR step of library preparation. In typical DNA or RNA library prep, insert-containing molecules are at least 100 bp larger than adapter-dimer molecules, and thus can be removed using SPRI magnetic beads. However, since insert-containing molecules are only ~20 bp larger than adapter-dimer molecules in small RNA libraries, SPRI size selection is not feasible, and gel-based selection must be performed. The need for gel-based size selection greatly limits both throughput and automation potential of small RNA library preparation, as only a limited number of libraries can be run on a single gel and it is a labor-intensive process that is not amenable to automation.

The lack of low-input protocols for small RNA-Seq is also related to adapter-dimer formation. Small RNA sequencing is somewhat unique in that additional PCR cycles result in negligible bias; thus it should theoretically be possible to create low-input small RNA libraries by using a high number of PCR cycles. However, adapter-dimer present in the libraries will also be greatly amplified, which eventually leads to a library where adapter-dimer products are extremely abundant, making it difficult to isolate insert-containing products and leading to sequencing data where very few of the reads are useful. A number of methods have been developed to reduce adapter-dimer formation in small RNA library preparation, but unfortunately none are effective at reducing adapter-dimer formation to such an extent that gel-free or low-input small RNA library preparation is possible.

In response to the drawbacks associated with small RNA sequencing, Bioo Scientific has developed a small RNA library preparation method that utilizes randomized adapters for greatly reduced bias and features gel-free or low-input protocols, which are possible due to significant reduction in adapter-dimer formation. This method is currently available as the NEXTflex Small RNA-Seq Kit v3.

1. Jayaprakash, A.D., et al., Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. Nucleic Acids Res, 2011. 39(21): p. e141.
2. Sun, G., et al., A bias-reducing strategy in profiling small RNAs using Solexa. RNA, 2011. 17(12): p. 2256-62.
3. Sorefan, K., et al., Reducing ligation bias of small RNAs in libraries for next generation sequencing. Silence, 2012. 3(1): p. 4.
4. Zhang, Z., et al., High-efficiency RNA cloning enables accurate quantification of miRNA expression by deep sequencing. Genome Biol, 2013. 14(10): p. R109.

RESULTS

A dual adapter-dimer reduction strategy allows gel-free or low input small RNA-Seq library preparation

Figure 1 shows a flowchart of or improved method using a dual method for adapter-dimer reduction and incorporating randomized adapters to significantly reduce bias.

Randomized adapters show more equal coverage of synthetic miRNA standards

Figures 2 and 3 demonstrate more equal coverage of miRNAs mixed in equimolar amounts when using randomized adapters for library preparation.

Randomized adapters allow greater small RNA discovery at lower sequencing depth in total RNA

Figure 4 demonstrates that fewer reads are required to detect/discover more small RNAs when using randomized adapters.

Enhanced adapter-dimer reduction allows low-input library preparation

Figure 4 also demonstrates that libraries prepared from low input samples using randomized adapters and enhanced adapter-dimer reduction also show greater detection/discovery rates than libraries prepared with 10x more RNA and standard adapters and protocols.

Additional PCR cycles do not contribute to bias in small RNA-Seq library preparation

Table 1 demonstrates that additional PCR cycles add negligible bias to small RNA library preparation.

Figure 1. Flowchart of the protocol used for reduced bias and gel-free or low-input Small RNA-Seq library preparation using the NEXTflex Small RNA-Seq Kit v3. A combination of bead-based excess 3' adapter depletion and enzymatic excess 3' adapter inactivation allow either gel-free or low-input protocols.

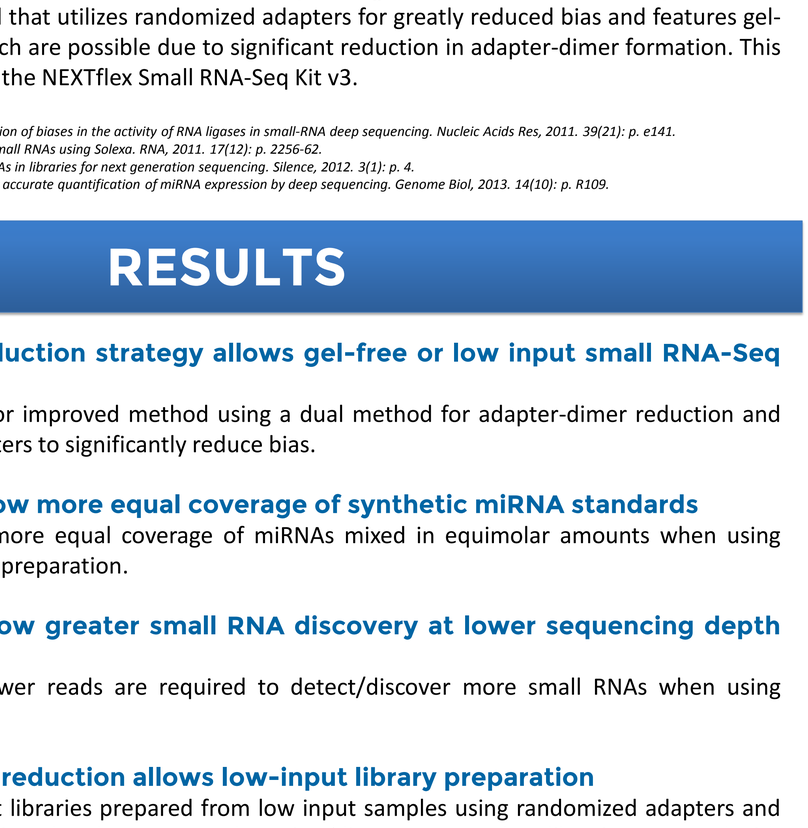


Figure 2. Sequencing results from small RNA libraries created in triplicate from 1 ng of miRNA Calibrator, an equimolar mixture of 24 miRNAs. Values farther from 1 indicate more bias. The inset shows the Coefficient of Variation of the 24 miRNAs in each sample.

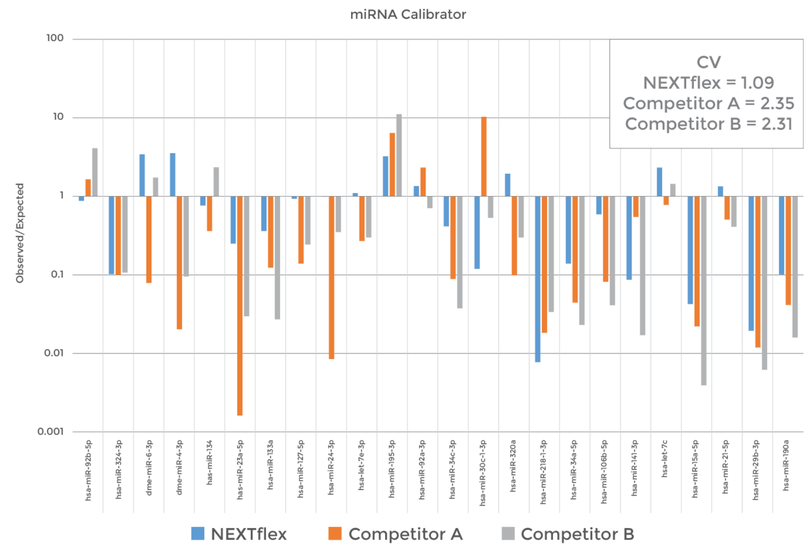


Figure 3. Sequencing results from small RNA libraries created in triplicate from 1 ng of Miltenyi miRXplore Universal Reference, an equimolar mixture of 963 miRNAs. The number of miRNAs detected at various thresholds is shown. The inset shows the Coefficient of Variation of the 963 miRNAs in each sample.

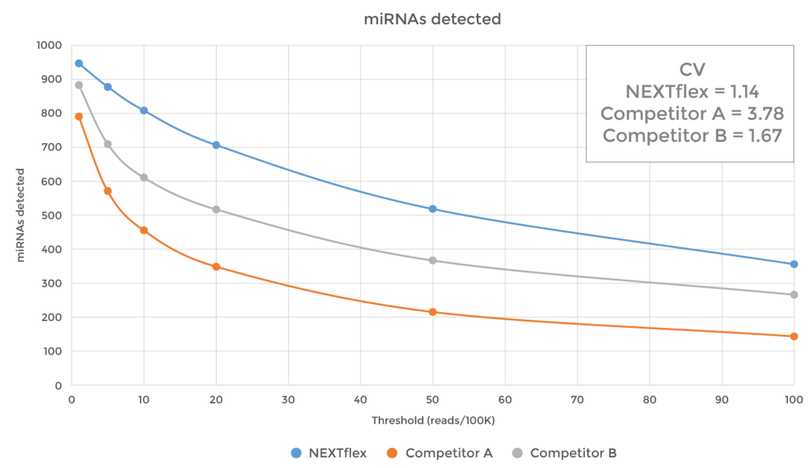
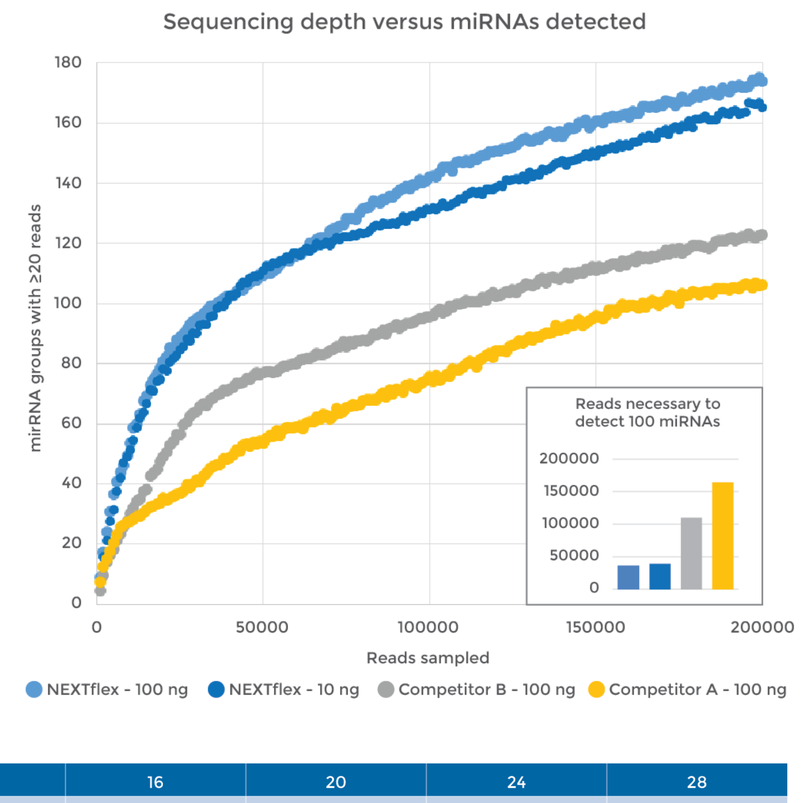


Figure 4. Small RNA libraries were created in triplicate from the indicated amount of human brain total RNA. The indicated number of reads were sampled from each library and the average number of miRNA groups with ≥20 reads determined. The inset shows the number of reads required to detect 100 miRNA groups at a threshold of ≥20 reads.



PCR Cycles	12	16	20	24	28
12	1	0.9999	0.9994	0.9966	0.9807
16		1	0.9996	0.9970	0.9813
20			1	0.9975	0.9824
24				1	0.9816
28					1

Table 1. Correlation of miRNA abundance in libraries created using serial 10x dilutions of cDNA and the indicated number of PCR cycles; 12 PCR cycles were used for undiluted cDNA, 16 cycles for 1/10 diluted cDNA, 20 cycles for 1/100 diluted cDNA, etc. The Pearson correlation coefficients calculated from the Log10(reads) values of miRNAs with ≥ 10 reads in all samples are shown.

MATERIALS & METHODS

Libraries were created in duplicate or triplicate from the indicated starting material using the manufacturer's instructions. Libraries were sequenced on an Illumina MiSeq. Data pre-processing included first clipping adapter sequences, then trimming the 4 random bases from the 5' and 3' ends of the reads for libraries created using randomized adapters. For libraries created from synthetic miRNA mixtures, reads were aligned to the appropriate reference using Bowtie2 and counts were determined. For libraries created from total RNA, reads were aligned to miRBase human miRNAs using Bowtie2, and the resulting alignments were processed with mirUtils (mirutils.sourceforge.net) in order to determine counts for miRNA groups, which contain miRNAs with identical mature sequences.

CONCLUSIONS

Small RNA-Seq library preparation has historically suffered from three major drawbacks; severe bias, the need for gel-based purification, and the lack of low-input protocols. Here we describe a method that addresses all of these drawbacks with a library preparation kit that uses patent-pending technology to offer reduced bias and allow gel-free or low-input library preparation. Both gel-free and low-input libraries created with our new method show greatly reduced bias in comparison to libraries created with other commonly used adapters and protocols, resulting in greater depth of discovery and sequencing data that more accurately represents the starting material.