Accurate detection of small non-coding RNAs using NEBNext® Low-bias Small RNA Library Prep Kit

Heather M Raimer Young, Deyra N Rodriguez, Brittany S Sexton, Gautam Naishadham, Bradley W Langhorst, Louise Williams | New England Biolabs, Inc.

Introduction

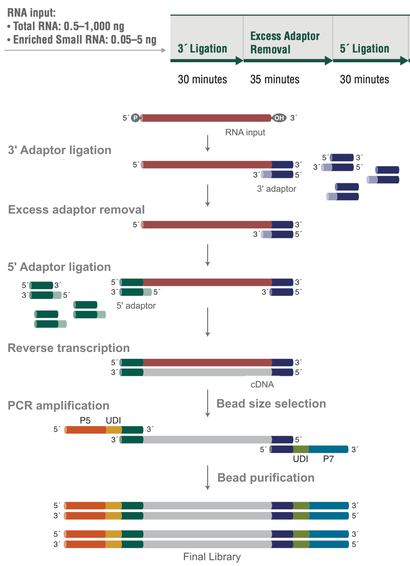
Changes in small non-coding RNA (sncRNA) expression have been implicated in the development and progression of cancers, neurological, and cardiovascular diseases. Use of sncRNAs as disease biomarkers requires their precise and sensitive detection. High-throughput sequencing is a powerful tool for the sequence characterization of sncRNAs, However, library preparation methods often limit the accuracy and sensitivity of detection. Bias is often introduced in ligation-based methods, which obscures the true sncRNA composition. Improvements in library preparation methods are essential for using sncRNAs as clinical biomarkers.

We have developed a novel, ligation-based small RNA library preparation method that is characterized by reduced bias in addition to increased detection of sncRNAs. Libraries can be made in a single day using a streamlined protocol with bead-based size-selections and cleanups. The robustness of this method is demonstrated with its compatibility across a broad input range as well as in challenging sample types such as formalin-fixed paraffin-embedded (FFPE) RNA.

Even representation of sncRNAs using this low-bias sncRNA library preparation method was confirmed using a pool of synthetic miRNAs. Approximately 90% of miRNAs were within 2-fold of the expected number, compared to less than 30% with other methods. Additionally, miRNA detection was consistent using 0.5 ng to 1,000 ng of total RNA from human brain. Furthermore, this low-bias method robustly detected 2'-O-methylated sncRNAs, such as piRNAs and plant miRNAs, without any protocol modifications. To address more challenging samples, libraries were made from low quality FFPE total RNA, 1 ng to 100 ng, resulting in consistent yields and miRNA detection. Regardless of input or sample, this method shows a robust capability to generate high quality libraries with increased confidence in the detection of sncRNAs and therefore the potential identification of disease biomarkers.

Methods

Overview of NEBNext Low-bias Small RNA Library Preparation



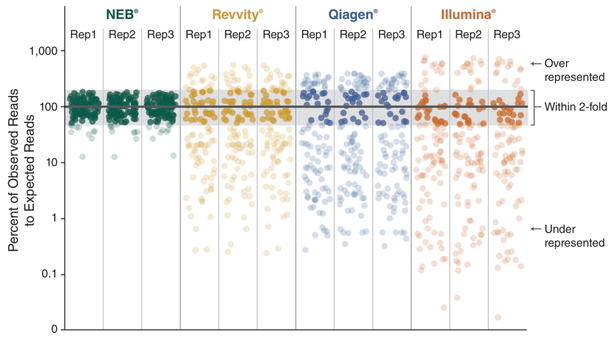
RNA Samples and Inputs:
- 1,000–0.5 ng of human brain total RNA
- 100 ng human testis total RNA
- 100–1 ng FFPE RNA (breast and lung)
- 100 ng plant total RNA (Arabidopsis, rice, soy)
- 0.3 ng of a synthetic miRNA mix (100 miRNA-like sequences)

Vendor Kits:
- 3 ligation-based competitor kits were tested
  - Revvity NEXTFLEX Small RNA Sequencing Kit V4
  - Qiagen QIAseq miRNA Library Kit
  - Illumina TruSeq Small RNA Library Prep
- Manufacturers' RNA input recommendations and protocols were followed.

Sequencing and Analysis
- Sequenced on NextSeq 500 or NovaSeq 6000 using 1x56 base reads or 1x72 base reads (for Qiagen + UMI).
- 30 million reads to 5 million reads were used in a down-sampling analysis.
- 15 million reads were typically used in analysis.
- Reads were adaptor trimmed (Flexbar) and aligned to respective genomes (human – hg38; Arabidopsis – Arabidopsis thaliana TAIR 10; rice - IRGSP-1.0; soy - Glycine max v2.1) the hg38 genome with STAR (v2.7.8a). The STAR reference was built using gencode v35 main annotations (contains 1,881 miRNAs) supplemented with gencode tRNA annotations, rRNA annotations for subunits not included in gencode, and piRNA annotations from piRNAdb v1.7.6 that did not overlap other annotations.

Results

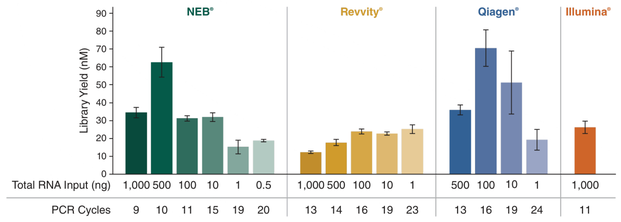
Even Representation of Synthetic miRNAs



NEBNext Low-bias Small RNA Library Prep Kit produces libraries with the lowest bias. NEBNext Low-bias Small RNA libraries were made using a mix of 100 synthetic control miRNAs, including five that had 3' 2'-O-methyl ends. Expected reads were calculated from total reads mapped to the synthetic controls, divided by the total number of control sequences, represented here by the black line at 100%. Percent of observed reads to expected reads was calculated for each control sequence and plotted across replicates. For the NEBNext libraries, 90% of the synthetic miRNAs were located within 2-fold of expected, whereas the competitor kits had between 19–30%. Competitor kits had more underrepresented sequences (50–65%) than overrepresented (15–24%). Libraries produced by the NEBNext Low-bias Small RNA Kit reliably generated libraries with the lowest bias.

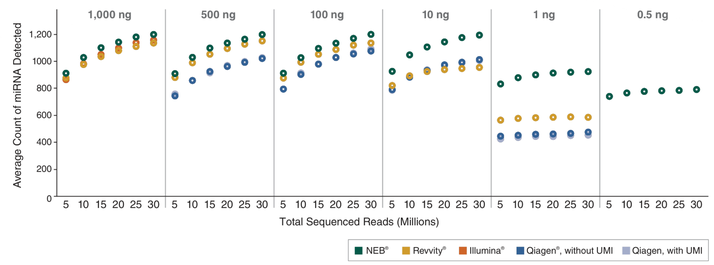
Results

Robust Small RNA Library Prep



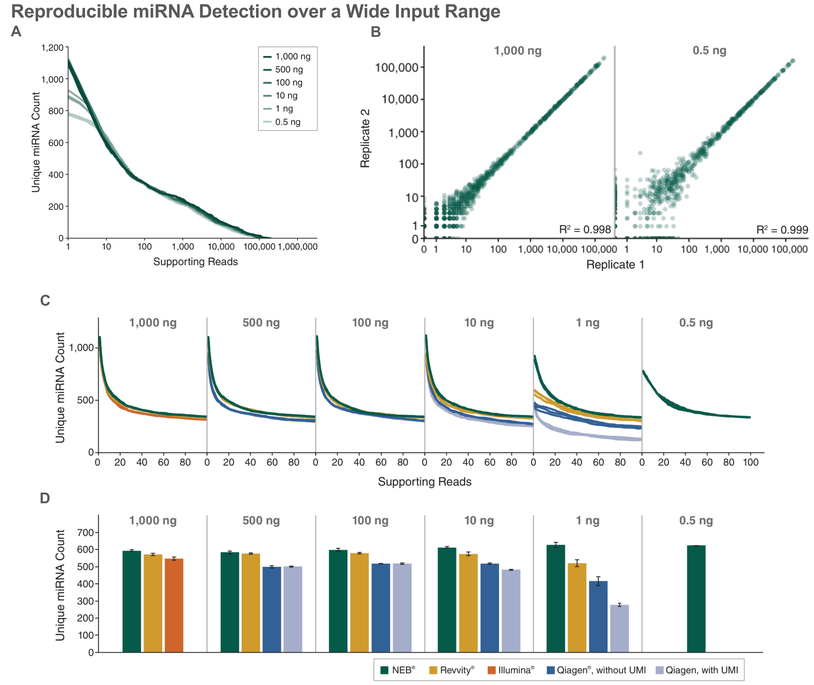
NEBNext Low-bias Small RNA Library Prep Kit produces robust library yields across a broad input range, with fewer PCR cycles. NEBNext Low-bias Small RNA libraries were made using 1,000–0.5 ng of human brain total RNA, with yields similar to or higher than competitors kits. NEBNext libraries typically were prepared using 2 – 5 fewer PCR cycles than recommended by competitor kits for the same input. NEBNext libraries were prepared following the bead size-selection protocol after library amplification. Library yields were determined on the Agilent® TapeStation® using the High Sensitivity D1000 tape. Yields represent three technical replicates, and error bars are standard deviation.

miRNA Detection with Increased Sequencing Depth



NEBNext Low-bias Small RNA Library Prep Kit detects the most miRNAs across a broad input range and sequencing depth. The number of miRNAs detected over increasing sequencing depths are shown for small RNA libraries made using 1,000–0.5 ng of human brain total RNA. NEBNext Low-bias Small RNA libraries consistently detected more miRNAs at all sequencing depths and inputs when compared to competitor kits. The number of miRNAs observed with at least one read was averaged among three technical replicates for each down-sampling.

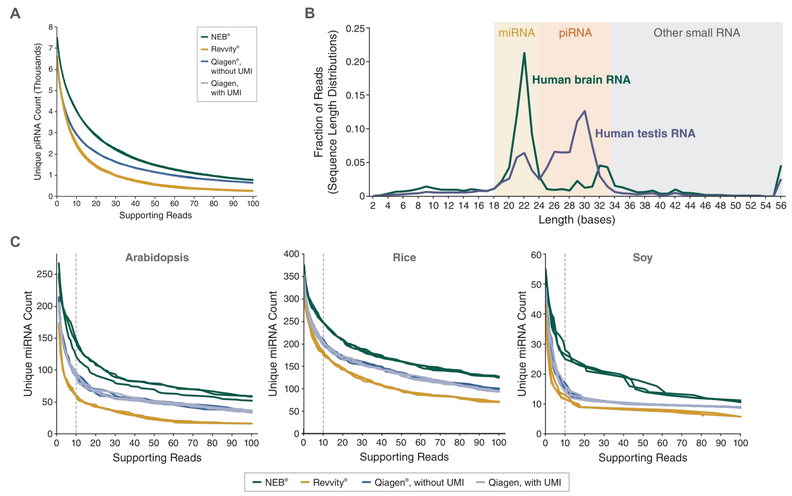
Reproducible miRNA Detection over a Wide Input Range



NEBNext Low-bias Small RNA Library Prep Kit consistently and reproducibly detects more miRNAs. (A) The number of miRNAs detected in NEBNext Low-bias Small RNA libraries are shown on a cumulative-detection histogram for mapped miRNAs at different supporting-read thresholds. Libraries made using 1,000–10 ng of human brain total RNA had similar miRNA profiles and detected ~1,110 miRNAs with at least 1 read. The 1 ng and 0.5 ng human brain total RNA inputs detected slightly fewer miRNAs, with ~900 miRNAs and ~780 miRNAs, respectively. The gencode annotation contains 1,881 miRNAs. (B) Linear correlations of mapped miRNAs between two technical replicates for libraries made using 1,000 ng and 0.5 ng human brain total RNA demonstrate high reproducibility of NEBNext Low-bias Small RNA libraries. Each point represents an individual miRNA, axes are shown in log scale, and R^2 are shown for each input. (C) A cumulative-detection histogram for mapped miRNAs at different supporting read thresholds across all inputs and library prep kits indicates the number of different miRNAs detected. (D) A histogram of miRNAs detected using a threshold of 10 supporting reads across inputs and kits is shown. The NEBNext Low-bias Small RNA libraries identify the most miRNAs across all inputs. Values represent three technical replicates, and error bars represent standard deviation.

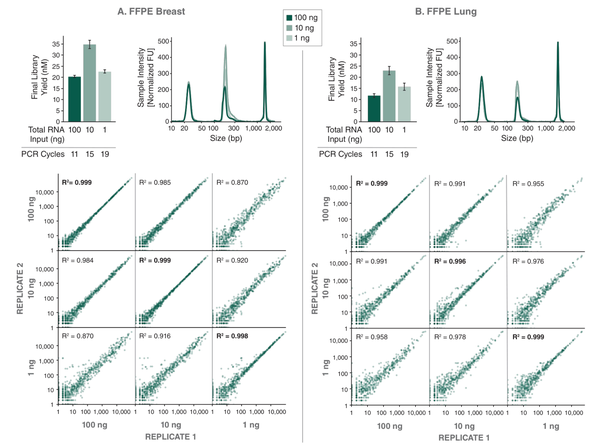
Results

Robust Detection of 2'-O-methylated Small RNAs



NEBNext Low-bias Small RNA Library Prep Kit robustly detects 2'-O-methylated small RNAs.
(A) NEBNext Low-bias Small RNA libraries better represent the 2'-O-methylated piRNAs in human testis. The cumulative-detection histogram for mapped piRNAs demonstrates that NEB libraries consistently detected more 2'-O-methylated piRNAs than competitor kits. (B) The NEBNext Low-bias Small RNA Library Prep Kit accurately represents the difference in small RNA fractions between biological samples as shown in the sequenced insert length of small RNAs prepared using either 100 ng human brain or testis total RNA. The low-bias nature of the NEBNext libraries is demonstrated by the expected shift in the proportion of miRNAs compared to 2'-O-methylated piRNAs between human brain and testis. (C) The cumulative-detection histograms indicate the number of mapped 2'-O-methylated miRNAs detected from libraries prepared from 100 ng of Arabidopsis, rice and soy total RNA. The NEB libraries consistently detected more 2'-O-methylated plant miRNAs than competitor kits.

Prepares Libraries from Low Input FFPE RNA Samples



NEBNext Low-bias Small RNA Library Prep Kit prepares robust libraries from FFPE RNA, even at low inputs. NEBNext Low-bias Small RNA libraries were successfully produced using total RNA isolated from formaldehyde-fixed, paraffin-embedded (FFPE) samples of (A) human breast (RIN 2.4 DV200 30-50%) and (B) human lung (RIN 2.6 DV200 < 30%). RNA integrity of FFPE RNA was determined using the Agilent® Bioanalyzer® RNA 6000 Pico Kit. Robust libraries were generated using 100–1 ng of the input RNA following the bead size-selection protocol after library amplification. Yields represent three technical replicates, and error bars are standard deviation. Library profiles remain consistent between inputs (Agilent TapeStation High Sensitivity D1000). miRNAs detected using NEBNext Low-bias Small RNA libraries show high levels of reproducibility between replicates and input amounts. Correlations for two replicates of each input for breast and lung FFPE libraries were generated. Each point represents an individual miRNA, axes are shown in log scale, and R^2 are shown for each comparison.

Conclusions

- The NEB Low-bias Small RNA Kit has significantly lower bias than competitor kits enabling accurate representation of small RNA abundance
- The kit is compatible with a wide input range (1000–0.5 ng total RNA; 50 ng–5 pg enriched small RNA) and utilizes bead-based size selection for all input amounts
- Consistent numbers of miRNAs detected across inputs
- Efficient capture of 2'-O-methylated small RNAs (piRNAs and plant miRNAs) without modifying the workflow
- Detects miRNAs from RNA isolated from FFPE samples
- Single day workflow done in PCR strip tubes with minimal tube transfers
- Compatible with NEBNext's 480 LV unique dual index primer pairs

Acknowlegements

We thank Dora Posfai, Kristen Augulewicz, Harry Bell, and Rebecca Gawron of the sequencing core for their technical support and members of NEBNext for productive discussions.