P362

Rapid, robust and sensitive DNA, RNA and small RNA library construction methods for evaluation of various samples and species

Scott Adams, Anetta Nowosielska-Vecchietti, Karen McKay, Chelsea Pinegar, Kaylinnette Pinet, Chen Song, Pingfang Liu, V K Chaithanya Ponnaluri, Daniel J Evanich, Heather M Raimer Young, Deyra N Rodriguez, Brittany S Sexton, Matthew Campbell, Gautam Naishadham, Bradley W Langhorst, Louise Williams, Christine Sumner, Keerthana Krishnan

New England Biolabs, Inc. 240 County Road, Ipswich, MA 01938, USA

INTRODUCTION

Ongoing research in agricultural genetics aims to transform crop cultivation and livestock production through advanced breeding methods. Next-generation sequencing offers rapid, reliable, and highly sensitive tools to unravel the complex genomes of plants and animals, enabling precise data analysis.

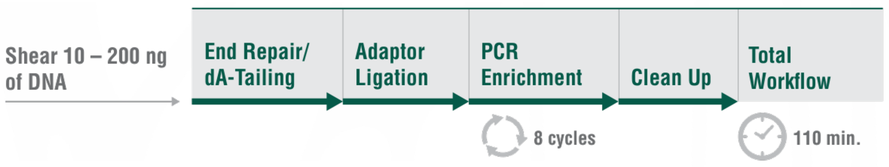
We developed multiple protocols that reliably and rapidly construct high-quality sequencing libraries from diverse agricultural animal and plant species, accommodating a broad range of input quantities. NEBNext UltraExpress® DNA, NEBNext UltraExpress® FS DNA and NEBNext UltraExpress® RNA library prep kits provide rapid (2-3 hours) and reliable library construction, utilizing a single protocol for all input amounts (10 ng - 200 ng DNA and 25 ng - 250 ng RNA). The NEBNext® Low-bias Small RNA library prep kit delivers a fast (~3.5 hours) workflow, generating libraries from a broad range of inputs (0.5 ng - 1 µg total RNA or 50 pg – 5 ng enriched small RNA).

Sequencing data generated from our NEBNext UltraExpress DNA library construction methods for bovine, pig, maize, and Arabidopsis show uniform GC coverage as well as robust library yield and complexity. In addition, the NEBNext UltraExpress RNA library prep for Arabidopsis, rice and maize demonstrates consistent coverage and sensitivity for transcript detection. Fewer protocol steps, consumables, and cleanups provide a simple, streamlined process without sacrificing quality. Further sequencing data from our Low-bias Small RNA library prep method in Arabidopsis, rice, and soy highlight its strong reproducibility across technical replicates and input amounts. The workflow consistently detects microRNAs with high sensitivity and accuracy, generating libraries that capture diverse small RNA species while reducing sequence-specific biases common in traditional methods.

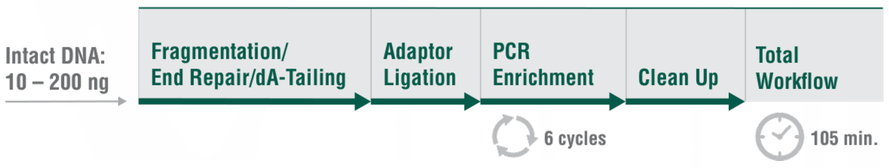
These streamlined NEBNext workflows empower researchers to efficiently generate high-quality sequencing libraries across diverse plant and animal sample types. By combining speed, sensitivity, and consistency, as well as compatibility with automation, they offer a powerful solution for accelerating genomic studies in agricultural research.

METHODS

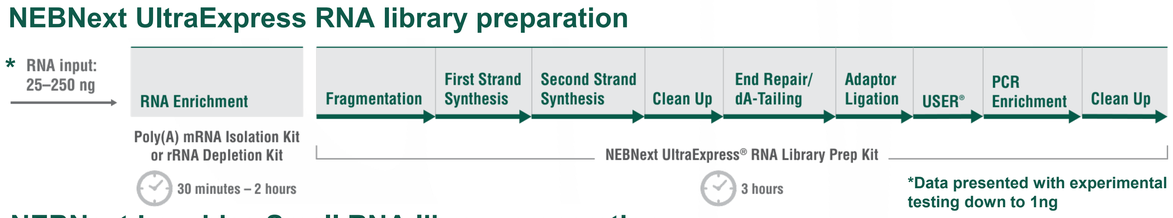
NEBNext UltraExpress DNA library preparation



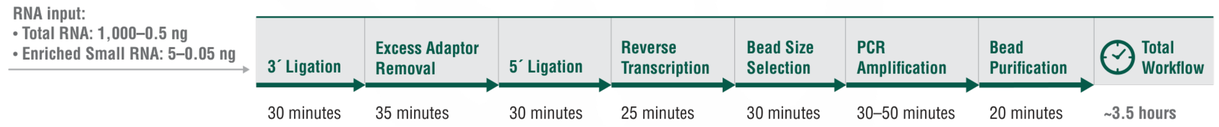
NEBNext UltraExpress FS DNA library preparation



NEBNext UltraExpress RNA library preparation



NEBNext Low-bias Small RNA library preparation



RESULTS

NEBNext UltraExpress DNA and NEBNext UltraExpress FS DNA: One protocol fits a wide input range and diverse sample types

(A) NEBNext UltraExpress DNA and FS DNA with Bovine and Pig

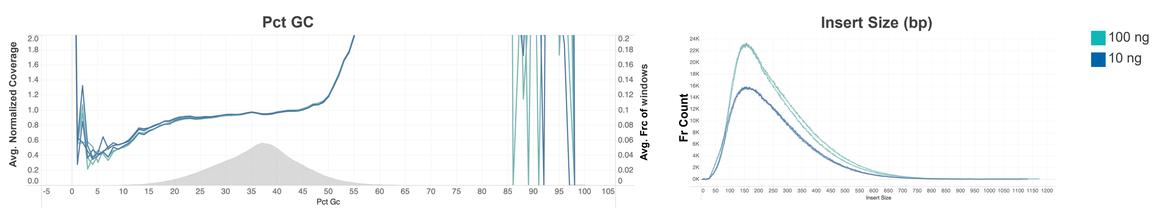
Libraries prepared from 10 and 200 ng Bovine and Pig DNA (BioChain Genomic DNA) using Covaris (NEBNext UltraExpress DNA) or enzymatic shearing (NEBNext UltraExpress FS DNA).

(B) NEBNext UltraExpress DNA with Maize (*Z. mays*)

Libraries were prepared from 10 and 100 ng of maize DNA *(BioChain Genomic DNA)* using Covaris shearing (NEBNext UltraExpress DNA).

(A&B) Bovine and pig libraries were sequenced on Illumina NextSeq® 500 (2 x 76 bases) with 1.9 million paired end reads. Maize libraries were sequenced on MiSeq® with 110,000 paired end reads. All sample types showed an even representative GC coverage and the expected insert sizes. The apparent over-representation at >55% GC in Bovine is likely due to a collapsed repeat. All libraries were sampled (seqtk v1.3), adapter-trimmed (seqprep v0.1) and mapped (bowtie2 v2.4.5) to the Bovine (bosTau8), Pig (susScr11) or Maize (*Zea Mays)* reference genome. GC Coverage and Insert size metrics were assessed using Picard Collect InsertSizeMetrics and CollectGCBiasMetrics (Galaxy version 1.56.0).

(C) NEBNext UltraExpress FS DNA with Arabidopsis (*A. thaliana*)

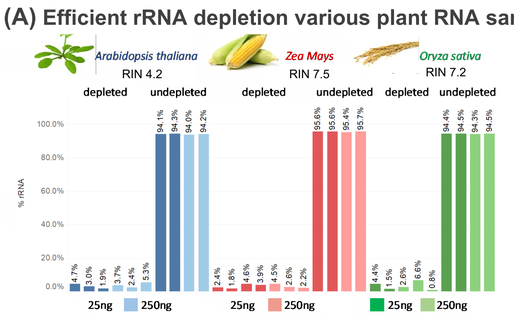


(C) The same fragmentation and library preparation protocol was used to make libraries from 10 ng and 100 ng of *A. thaliana* and sequenced on NextSeq500 (2 x 76 bases). Coverage of most frequent GC windows of the genome (gray histogram) is even in the relevant GC content ranges. The apparent over-representation at 35-40% GC in *A. thaliana* can be eliminated by excluding alignments with many mismatches (AS < -5), which indicates a mismatch between the reference genome and the specific *A. thaliana* sample used in this experiment, likely due to different repeat copy number. 500,000 paired end reads were sampled (seqtk v1.3), adapter-trimmed (seqprep v0.1) and aligned to respective reference genomes (bowtie2 v2.4.5). GC Bias and Insert size metrics were assessed using Picard CollectInsertSizeMetrics and CollectGCBiasMetrics (Galaxy version 1.56.0).

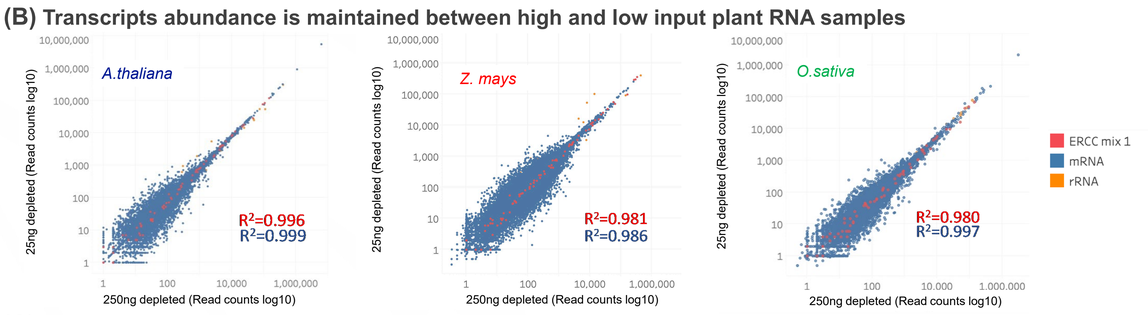
RESULTS Continued

NEBNext UltraExpress RNA: Superior data quality with one protocol for various sample types

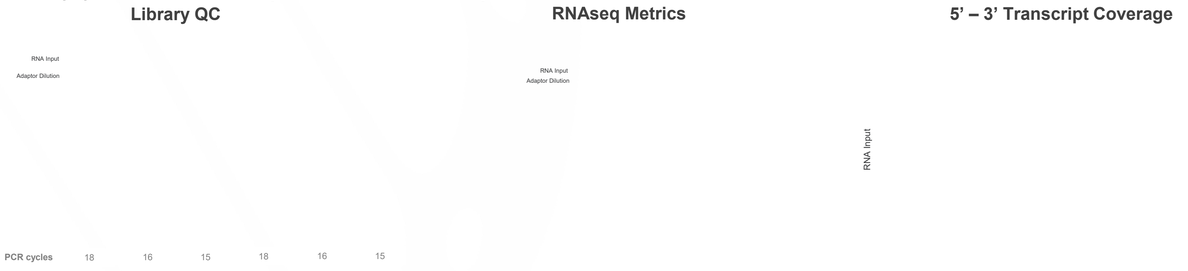
(A) Efficient rRNA depletion various plant RNA samples



(B) Transcripts abundance is maintained between high and low input plant RNA samples



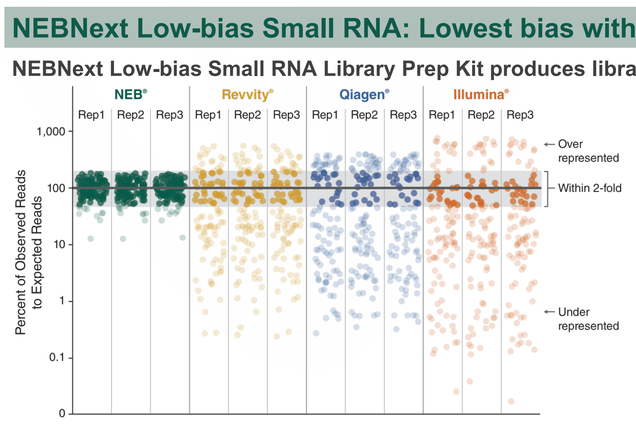
(C) NEBNext UltraExpress RNA for low input Human RNA



(C) Poly(A)-containing mRNA was isolated from Universal Human Reference RNA (UHRR) (Agilent®) using the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB #E7490). Ribosomal RNA (rRNA) was depleted from UHRR by the NEBNext rRNA Depletion Kit (Human/Mouse/Rat (NEB #E7400). Final library sizes and yields were assessed by Agilent Tapestation High Sensitivity DNA ScreenTape Analysis. Libraries were sequenced on Illumina NovaSeq® 6000 (2 x 100 bases) to 5 million paired end reads and mapped to the GRCh38 reference genome using RNA STAR v2.7.8a. Read pairs were identified as ribosomal using Mirabait (6 or more, 25-mers). 5´ to 3´ Transcript coverage was calculated from the top 1,000 transcripts using the CollectRnaSeqMetrics (Picard) tool v2.18.2.2.

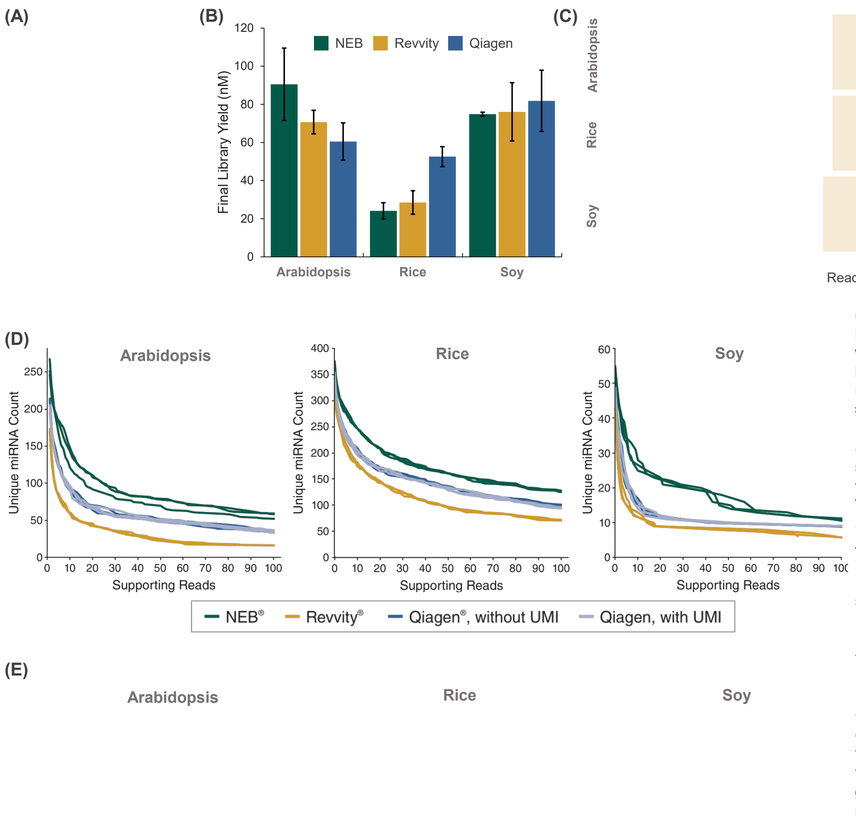
NEBNext Low-bias Small RNA: Lowest bias with one protocol for various qualities of samples

NEBNext Low-bias Small RNA Library Prep Kit produces libraries with the lowest bias



NEBNext Low-bias Small RNA libraries were made using an equimolar mix of 100 synthetic control miRNAs, including five that had 3´ 2´-O-methyl ends. Libraries were prepared from 0.3 ng of the synthetic miRNA mix to compare library-generated bias between the NEBNext Low-bias Small RNA Library Prep Kit and small RNA kits from Revvity (NEXTFLEX Small RNA Sequencing Kit V4), Qiagen (QIAseq miRNA Library Kit) and Illumina (TruSeq Small RNA Library Prep). Libraries were sequenced on an Illumina NextSeq 500 (1 x 56 bases), and 5 million reads from each library were adaptor trimmed (Flexbar) and aligned to the synthetic sequences using STAR (v2.7.8a). 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 control mix miRNAs were located within 2-fold of expected, whereas the competitor kits had between 19–30% fall within that range. Competitors' 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.

NEBNext Low-bias Small RNA Library Prep Kit identifies more unique 2'-O-methylated miRNAs in plants



(A) Agilent Bioanalyzer RNA 6000 Pico gel images of input RNA for Arabidopsis, rice, and soy total RNAs with RIN values indicated for each RNA. **(B)** Final library yields shown for small RNA libraries prepared using 100 ng of total RNA (Arabidopsis: BioChain #R1634310; rice: BioChain #R1634360; soy: BioChain #R1634370) with the NEBNext Low-bias Small RNA Library Prep Kit (following the size selection protocol) or competitors' small RNA kits from Revvity (NEXTFLEX Small RNA Sequencing Kit V4) and Qiagen (QIAseq miRNA Library Kit). **(C)** Sequencing profiles of small RNA libraries from Arabidopsis, rice, or soy total RNA. Profiles represent read lengths after adaptor trimming. Highlighted boxes for each plant species indicates the expected size range for miRNAs. Libraries were sequenced on Illumina NovaSeq 6000 (1 x 56 bases; or 1 x 72 bases for Qiagen UMI), and reads were adaptor trimmed using Flexbar. **(D)** NEBNext Low-bias Small RNA libraries identify 2'-O-methylated miRNAs in plants. The graphs indicate the number of mapped 2'-O-methylated miRNAs. The NEB libraries consistently detected more 2'-O-methylated miRNAs than competitors' kits. Libraries were downsampled to 15 million reads and aligned to their respective genomes (Arabidopsis: Arabidopsis thaliana TAIR 10; rice: IRGSP-1.0; soy: Glycine max v2.1). The STAR references for Arabidopsis, rice and soy were built with annotations from EnsemblPlants release 52, supplemented with miRNA annotations from miRbase for Arabidopsis and rice. **(E)** A histogram of miRNAs detected at 10 or more supporting reads for all kits is shown. Values represent three technical replicates, and error bars represent standard deviation.

CONCLUSIONS

- Shorter hands-on time
- Fewer steps and consumables
- Flexibility with sample types and protocols
- Automation friendly volumes
- High sensitivity and accuracy

NEBNext UltraExpress DNA and FS:
- Sample to library in under 2 hours
- Single tube library prep to reduce consumables
- One protocol fits wide input range and diverse sample types
- Pipetting/reaction volume automation friendly

NEBNext UltraExpress RNA:
- High quality directional libraries in 3.5-5 hours
- Ability to use a single condition for all inputs
- Compatibility with multiple RNA samples types
- Recommendations available for lower inputs

NEBNext Low-bias Small RNA RNA:
- Total or enriched small RNA to library in ~3.5 hours
- Straightforward gel-free protocol with a wide input range
- Analyze all RNA species present, without library prep bias
- Multiplexable (up to 480 UDI primer pairs, available separately)

Authors would like to acknowledge the technical assistance provided by Dora Posfai, Kristen Augulewicz, Harry Bell, and Rebecca Gawron at the New England Biolabs' Sequencing Core Facility.