Methods for improved throughput and quality in long-read RNA sequencing workflows

Christine Sumner (sumner@neb.com), New England Biolabs, **Christine Sumner**, New England Biolabs, **Evan Janzen**, New England Biolabs, **Michael Sproviero**, New England Biolabs, **Gautam Naishadham**, New England Biolabs, **Jian Sun**, New England Biolabs, Inc., **Bradley Langhorst**, New England Biolabs, **Keerthana Krishnan**, New England Biolabs

The ability to accurately sequence and quantify the abundance of RNA transcripts has become increasingly valuable in research and medicine. Recent technological advances in long-read sequencing promise to further aid the characterization of RNA by enabling the analysis of full-length transcript isoforms. Full-length reads are increasingly in-demand, with applications to genome annotation, organism development, and disease etiology.

Technical challenges have limited the wide-spread adoption of these methods, as full-length RNA sequencing often requires large amounts of high-quality RNA and reverse transcription with high processivity and fidelity. To overcome these challenges, we have optimized methods for preparing high-quality RNA libraries which allow sequencing of long, full-length transcripts.

Using high-capacity oligo(dT)25 derivatized magnetic beads, we have developed a high-input poly(A) isolation method optimized for bulk-enrichment of poly(A) RNA from as much as 50 µg of total RNA in a single reaction. This method facilitates processing of multiple samples to enrich high-quality poly(A) RNA with high recovery and low background, yielding sufficient RNA for sequencing with little or no amplification. We have additionally incorporated InduroTM Reverse Transcriptase in library preparation workflows, producing full-length cDNA from transcripts >10 kb with high-fidelity, even for highly structured or modified transcripts. Using these methods, we have produced libraries from human, mouse, and yeast samples, with good library yields, consistent coverage across transcripts, and high mapping rates. Combining optimized poly(A) selection and a high-processivity reverse transcriptase, we enable high-quality RNA sequencing data from long-read platforms.