Methods for sequencing cell-free DNA and cell-free RNA from human plasma

Bradley W. Langhorst, Maggie Heider, Chen Song, Anagha Kadam, Jian Sun, Gautam Naishadham, Keerthana Krishnan

New England Biolabs, Inc.

Poster# 440

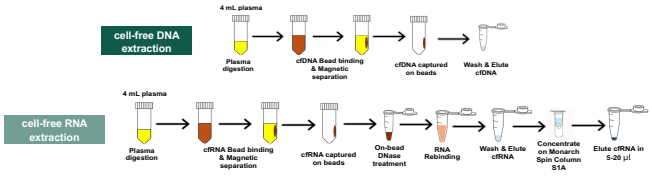
NEW ENGLAND Biolabs®

Introduction

Cell-free DNA (cfDNA) and cell-free RNA (cfRNA) are nucleic acid fragments that freely circulate in blood serum. Cell-free nucleic acids (cfNAs) can originate from various types of cells, including normal cells, cancer cells, and fetal cells. A great deal of biological information can be gathered from the sequencing of cfNAs. Thus, cfNA sequencing can serve as a non-invasive approach for the monitoring of atypical and typical biological processes ranging from cancer to ageing. The ability to analyze cfDNA and cfRNA from blood samples facilitates longitudinal studies and enables researchers to track molecular changes over time via liquid biopsies. Here we demonstrate methods for extracting and sequencing cfDNA and cfRNA isolated from healthy donor plasma samples. A modified protocol of the Monarch® Mag Viral DNA/RNA Extraction Kit (NEB#T4010) was used with supplemental Monarch reagents for optimal cell-free nucleic acid recovery. Protocols using NEBNext® DNA and RNA library prep kits were optimized to accommodate the size and quality of the cfDNA and cfRNA samples. Using these in-development methods, we obtained consistent, high-quality sequencing metrics, including high library yield, even coverage, and expected library complexity from varied individual samples. These libraries are compatible with whole genome sequencing (WGS), whole transcriptome, and target enrichment workflows. These protocols for cfDNA and cfRNA sequencing are relevant to a broad range of research and diagnostic scenarios, enabling scientists interested in the biological insights available from cell-free nucleic acids.

Methods

Cell-free DNA and cell-free RNA extraction



Consult with NEB for the latest version as we continue to optimize this protocol

Cell-free DNA and cell-free RNA Library Prep

DNA Library Prep

End Repair/ dA-Tailing → Adaptor Ligation → Clean Up → PCR Enrichment → Clean Up

- The NEBNext Ultra II DNA Library Prep Kit is compatible with up to 50 µl input volume while the NEBNext UltraExpress DNA kit is compatible with up to 15 µl input volume
- Input quantity (ng) constant for experiments shown
- Modified ligation conditions improve library conversion efficiency

NEBNext® Ultra II™ DNA Library Prep Kit
NEBNext UltraExpress™ DNA Library Prep Kit

RNA Library Prep

Fragmentation → First Strand Synthesis → Second Strand Synthesis → Clean Up → End Repair/ dA-Tailing → Adaptor Ligation → USER® → Clean Up → PCR Enrichment → Clean Up

NEBNext® Ultra II™ Directional RNA Library Prep Kit
NEBNext UltraExpress™ RNA Library Prep Kit

- Input cfRNA volume is kept constant for each kit to evaluate variations between sample source and donors.
- Protocol modifications include different adaptor concentration, addition of post-ligation cleanup, and sample-specific PCR cycle count.

Results

Effective cfDNA and cfRNA extraction from plasma

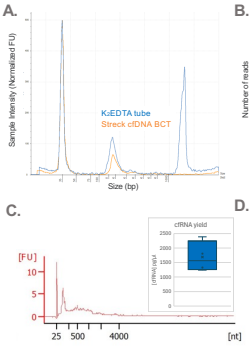


Figure 1. Effective extraction of cfDNA and cfRNA from plasma. (A) cfDNA was extracted from 4 ml of normal donor plasma stored in either K2EDTA tubes (blue) (Innovative Research) or Streck Cell-free DNA BCT (orange) (Quadrant Health Strategies, Inc.) to either 33% or 86% cfDNA purity using a custom protocol for Monarch Mag Viral DNA/RNA Extraction Kit. (B) Libraries were prepared using either 10 ng (K2EDTA) or 2 ng (Streck cfDNA) extracted cfDNA using the NEBNext Ultra II DNA Library Prep Kit and the NEBNext Unique Dual Index Primers. Libraries were sequenced on the NovaSeq6000 2x100 bp, downsampled to 10M pairs, Reads were trimmed using seqprep (v0.1), aligned using bowtie2 (v2.5.0) to the GRCh38 reference, duplicates were marked using MarkDuplicates (v1.56.0), and insert size calculated using Picard Insert size metrics (1.56.0). Extracted cfDNA showed the expected fragment sizes for normal donors. (C) cfRNA was extracted from 4 ml of normal donor plasma stored in K2EDTA tubes (Innovative Research) using a custom protocol for the Monarch Mag Viral DNA/RNA Extraction Kit and visualized on an Agilent Bioanalyzer RNA 6000 Pico chip. cfRNA yield varies between four donors. (D) 2.5 µL of cfRNA was used to prepare libraries using a protocol modification of the NEBNext UltraExpress RNA Library Prep Kit. Libraries were sequenced on the NextSeq2000 2x75 bp, downsampled to 3M read pairs. Reads were trimmed using fastp (v.0.20.0), mapped to the GRCh38 reference genome using RNA STAR v2.7.8a. Duplicates were marked using MarkDuplicates (v1.56.0), and insert size calculated using Picard Insert size metrics (1.56.0). Extracted cfRNA showed an insert size peak at 155bp.

Results

Sensitive variant detection in cfDNA

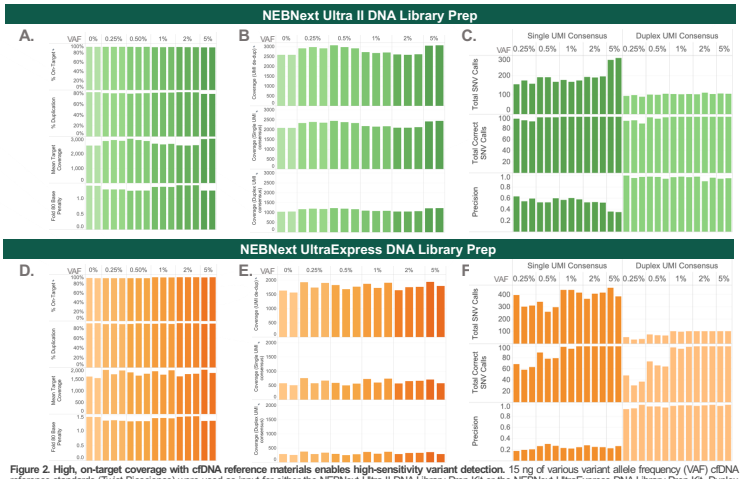


Figure 2. High, on-target coverage with cfDNA reference materials enables high-sensitivity variant detection. 15 ng of various variant allele frequency (VAF) cfDNA reference standards (Twist Bioscience) were used as input for either the NEBNext Ultra II DNA Library Prep Kit or the NEBNext UltraExpress DNA Library Prep Kit. Duplex UMI adaptors (Integrated DNA Technologies) and 7 PCR cycles were used to generate libraries for 6-plex capture using a custom cancer panel (Twist Bioscience). Captured libraries were sequenced on the NovaSeq6000 2x100 bp, downsampled to 30M read pairs, trimmed using fastp (v.0.20.0), and aligned using bwa-mem (v.0.7.17) to the GRCh38 reference. Duplicates were marked using Picard MarkDuplicates with UMI (v2.18.29), and capture metrics assessed using Picard HS metrics (v2.18.29). Duplex UMIs were processed using fgbio (v2.3.0), depth calculated using mosdepth (v.0.2.6), and somatic variant calling performed using VarDict (v.1.8.3). Both NEBNext Ultra II DNA (A) and NEBNext UltraExpress DNA (D) produce high on-target, uniform captured libraries. Effective coverage with different UMI approaches is shown, with effective coverage reduced upon duplex consensus building for NEBNext Ultra II DNA (B) and NEBNext UltraExpress DNA (E). NEBNext Ultra II DNA Library Prep enables ≥ 1000X duplex coverage (B). Somatic variant calling for NEBNext Ultra II DNA (C) and NEBNext UltraExpress DNA (F) are shown, with precision improving for both kits when using duplex UMI consensus as opposed to single UMI consensus.

High quality RNA-seq library and transcript detection in cfRNA

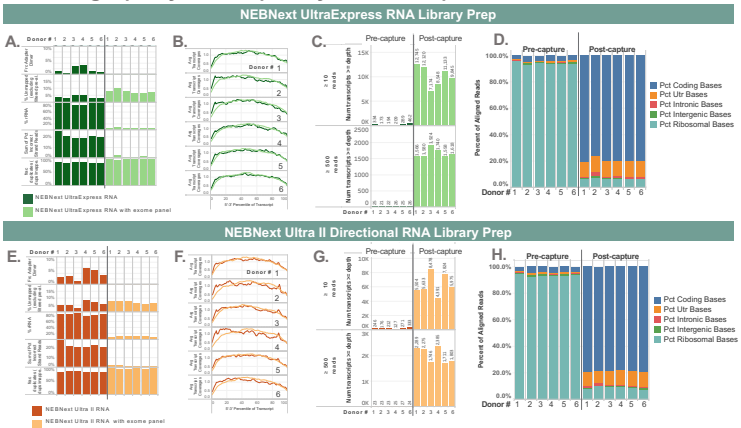


Figure 3. High quality RNA-seq metrics and transcripts detection from cfRNA. Cell-free RNA was extracted from six healthy donor plasma samples using Monarch Mag Viral DNA/RNA Extraction Kit. 2.5 µl or 5 µl of extracted cfRNA were used as input for either the NEBNext UltraExpress RNA Library Prep Kit (A-D) or the NEBNext Ultra II Directional RNA Library Prep Kit (E-H), respectively. NEBNext Multiplex Oligos for Illumina (Unique Dual Index UMI Adaptors RNA Set 1) and 18 PCR cycles were used to generate libraries for an 8-plex capture with the RNA exome panel (Twist Bioscience). Both pre-capture and post-capture libraries were sequenced on the NovaSeq6000 2x100 bp, downsampled to 5M read pair, trimmed using fastp (v.0.20.0), mapped to the hg38 reference genome using RNA STAR v2.7.8a and 5' to 3' Transcript coverage was calculated from the top 1,000 transcripts using the CollectRnaSeqMetrics (Picard) tool v2.18.2.2. Salmon v1.5.1 was used for mapping and quantification of all Gencode v38 transcripts. The percent of ribosomal RNA (rRNA) reads was calculated using bbduk v39.01 by identifying reads containing at least six kmers (k=25) from rRNA sequences. Both NEBNext Ultra II Directional RNA (A) and NEBNext UltraExpress RNA (E) produce high quality sequencing metrics with exome panel captured libraries containing significantly lower rRNA portion and increased directionality. 5'-3' transcript coverage, (B, F), are representative of the cfRNA profile with variations from donor to donor. Captured libraries showed more even transcript coverage. Transcripts detected with ≥ 10 reads and ≥ 500 reads, (C, G), were assessed. Aligned reads (D, H) represent the alignment proportion of each base category. Pre-captured libraries contain mostly ribosomal reads, and post-captured libraries greatly increased the coding and UTR fraction.

Conclusions

- The NEB Monarch Mag Viral DNA/RNA Extraction kit can be used with a modified protocol to produce high quality cell-free nucleic acids from plasma stabilized in different blood collection tubes.
- The NEBNext Ultra II DNA Library Prep Kit and the NEBNext UltraExpress DNA Library Prep Kit produce high quality libraries for target enrichment from cfDNA and enable high sensitivity and precision in variant calling.
- The NEBNext Ultra II Directional RNA Library Prep Kit and the NEBNext UltraExpress RNA Library Prep Kit produce high quality libraries from limited cfRNA input and enable target enrichment for transcripts profiling in plasma.
- Please reach out to NEB technical support if you are interested in applying these in-development protocols.