Abstract Reference Number: 04-002

Optimization of cfRNA detection via sequencing: Addressing workflow challenges

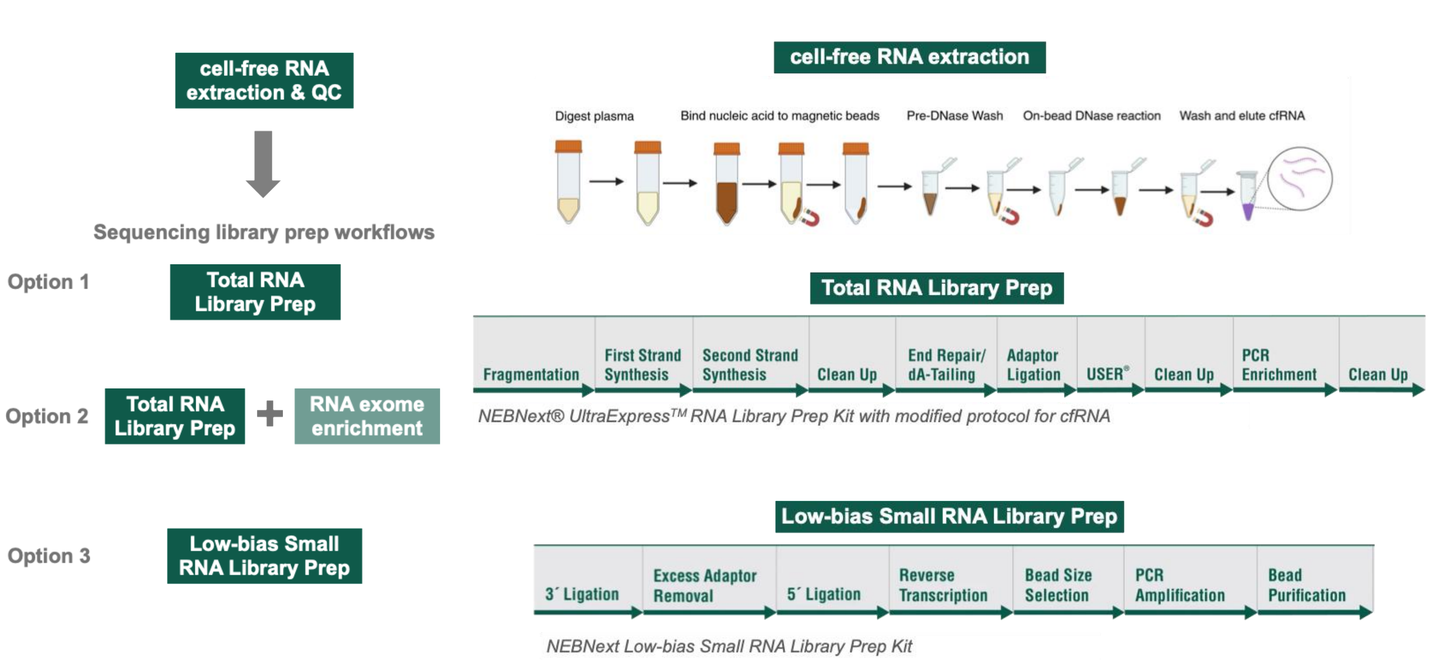
Chen Song, Chelsea Pinegar, Anagha Kadam, Heather Raimer Young, Gautam Naishadham, Keerthana Krishnan
New England Biolabs, Ipswich MA, USA

Background

Circulating cell-free RNA (cfRNA) has emerged as a promising biomarker for non-invasive diagnostics, offering dynamic insights into gene expression and disease states. Dynamic changes in cfRNA signatures can reflect both pathological and physiological processes, making it valuable for early disease detection, treatment monitoring, and prognostic assessments. However, clinical adoption of cfRNA-based diagnostics remains limited, largely due to technical and biological challenges throughout the workflow, including plasma processing, cfRNA extraction, quality control, and library preparation for sequencing. Here we show results from an optimized workflow for cfRNA detection, from extraction to sequencing.

Methods

Following standard plasma collection and storage methods, we used a magnetic bead-based cfRNA extraction protocol developed to maximize RNA yield and minimize DNA contamination. cfRNA quality and quantity were measured by automated electrophoresis and qPCR assays. Sequencing library prep assays were performed on these samples, to generate basic quality metrics, including total RNA-seq, target enrichment with RNA exome panel and small RNA library prep. The assays optimized for mRNA or small RNA were used for assessing the respective expected performances.



Please reach out to NEB technical support if you are interested in applying these in-development protocols.

Cell free RNA extraction and QC:
Blood from consented healthy donors was collected in K2EDTA tubes (Innovative Research) or Streck NABCT tubes and processed for plasma. 4 mL of plasma was used for cell-free RNA extraction using a custom NEB Monarch® solution that maximizes cfRNA concentration and minimizes DNA contamination. Agilent Bioanalyzer RNA 6000 Pico chip and RT-qPCR were used for cfRNA QC. A two step RT-qPCR assay was performed by using Luna® Script RT SuperMix Kit (NEB) and Luna Universal Probe qPCR Master Mix (NEB) with a rRNA 45S TaqMan assay (Thermo Fisher).

Total RNA library prep and exome capture:
2 µl of extracted cfRNA were used as input for the NEBNext UltraExpress RNA Library Prep Kit. 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 total RNA-seq and exome captured libraries were sequenced on the NovaSeq 6000 2x100 bp, downsampled to 30M read pairs. Unique reads were obtained by aligning to genome with STAR, and then deduplicated using picard MarkDuplicates (v1.56.0) by considering both the UMI sequence as well as alignment position. Reads were adapter trimmed using Flexbar (v.3.5.0), mapped to the hg38 reference genome using RNA STAR v2.7.8a. Salmon v1.10.1 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.

Low-bias small RNA library prep:
0.5 ng (Donors 1– 5) or 0.25 ng (Donors 6 & 7) of extracted cfRNA were used as input for the NEBNext Low-bias Small RNA Library Prep Kit. A 100-fold dilution of the NEBNext LB 3’ using the bead size selection protocol. Libraries were sequenced on the NextSeq 550 1x56 bp, downsampled to 10M reads. Reads were adapter trimmed using Flexbar (v.3.5.0), mapped to the hg38 reference genome using RNA STAR (v2.7.8a). STAR was used for mapping and quantification of Gencode v35 transcripts supplemented with rRNA, tRNA, and piRNA.

Results

cfRNA yields

cfRNA yields with electrophoresis assay

Donor	Collection tube	cfRNA concentration (pg/µl)	RIN	Yield ng/ml plasma
1	EDTA	692	2.4	3.5
2	EDTA	417	2.5	2.1
3	EDTA	297	2.5	1.5
4	EDTA	472	2.5	2.4
5	EDTA	311	2.5	1.6
6	EDTA	929	2.4	4.6
	Streck	35	1.8	0.2
7	EDTA	275	2.5	1.4
	Streck	111	1.5	0.6

cfRNA yields with RT-qPCR assay

Donor	Collection tube	cfRNA concentration (pg/µl)
1	EDTA	224.2
2	EDTA	103.3
3	EDTA	67.4
4	EDTA	243.6
5	EDTA	92.5
6	EDTA	699.2
	Streck	22.8
7	EDTA	113.5
	Streck	16.1

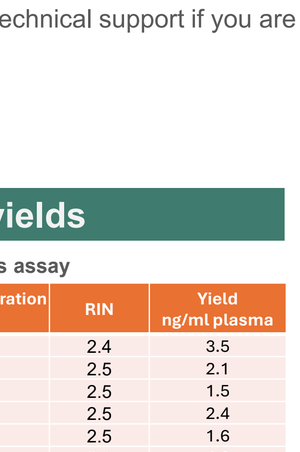


Figure 1. Efficient extraction of cfRNA from plasma assessed by two QC methods.
(A) Representative electrochromatographs from donor-7 using Agilent Bioanalyzer RNA 6000 Pico chip. RT-qPCR amplification curve **(B)** and standard curve **(C)** are shown targeting rRNA 45S region for cfRNA yield assessment. cfRNA yield varies between seven donors and collection tubes. Different quantification assays can show different results in cfRNA yield but have consistent trend in revealing sample variations.

Total RNA library prep and exome enrichment

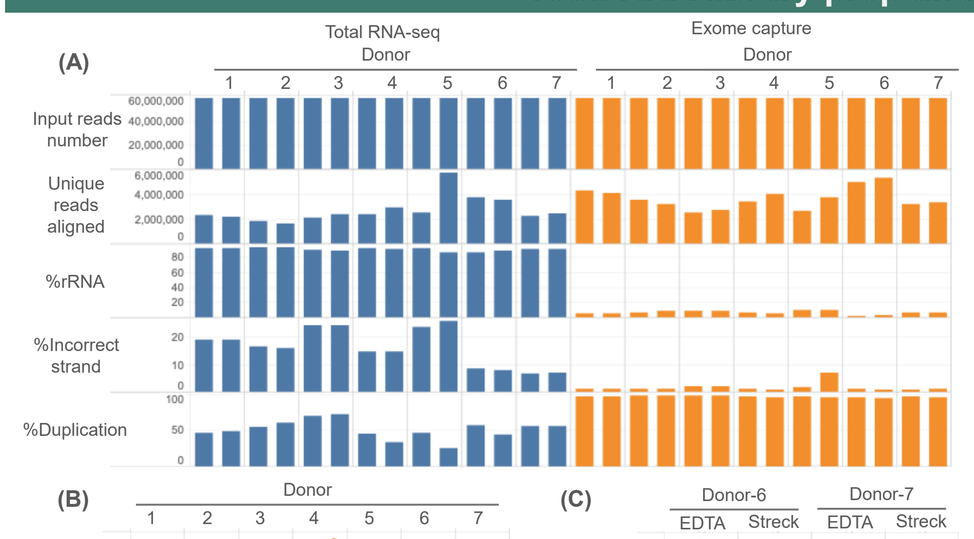


Figure 2. RNA-seq metrics from cfRNA.
(A) NEBNext UltraExpress RNA produces high quality sequencing metrics with exome panel captured libraries containing significantly lower rRNA portion and increased directionality. **(B)** Transcripts detected with ≥10 reads and ≥500 reads. Exome captured libraries detected higher number of transcripts. Samples collected in EDTA tubes are shown. **(C)** Transcript detection comparison between EDTA and Streck tube from the same donor. EDTA tube produced higher number of transcripts. Data from total RNA-seq is shown.

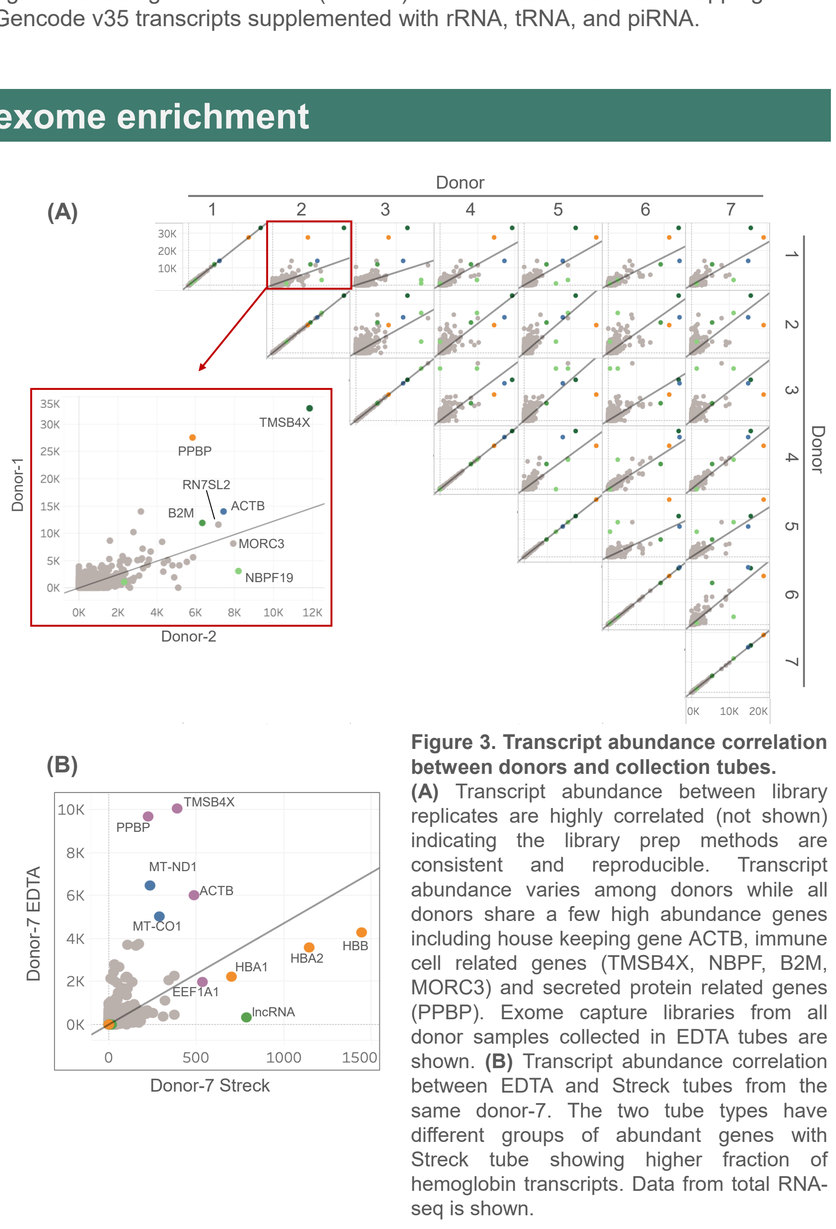
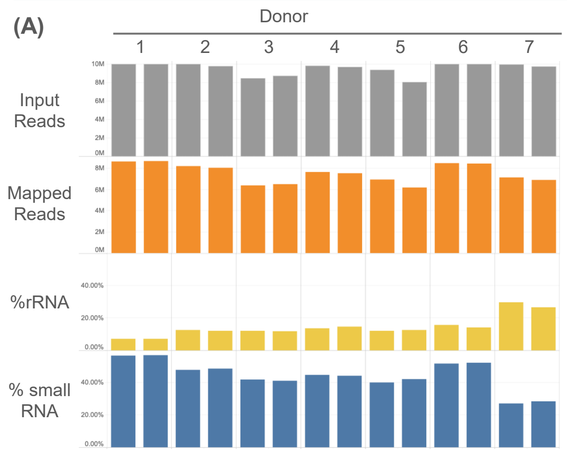
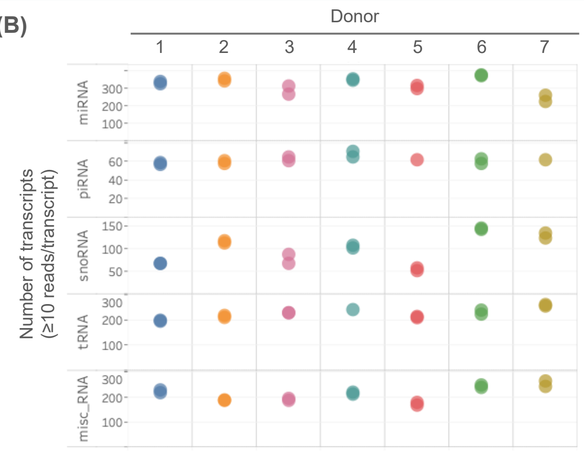


Figure 3. Transcript abundance correlation between donors and collection tubes.
(A) Transcript abundance between library replicates are highly correlated (not shown) indicating the library prep methods are consistent and reproducible. Transcript abundance varies among donors while all donors share a few high abundance genes including house keeping gene ACTB, immune cell related genes (TMSB4X, NBPF, B2M, MORC3) and secreted protein related genes (PPBP). Exome capture libraries from all donor samples collected in EDTA tubes are shown. **(B)** Transcript abundance correlation between EDTA and Streck tubes from the same donor-7. The two tube types have different groups of abundant genes with Streck tube showing higher fraction of hemoglobin transcripts. Data from total RNA-seq is shown.

Low-bias small RNA library prep





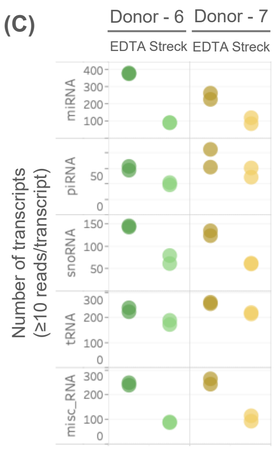


Figure 4. Low-bias Small RNA-seq metrics from cfRNA.
(A) NEBNext Low-bias Small RNA produces high quality sequencing metrics with high mapping rates, low rRNA content, and high small RNA recovery. **(B)** Small RNA transcripts detected with ≥10 reads, variability between donors collected with EDTA tubes are shown. **(C)** Transcript detection comparison between EDTA and Streck tube from the same donor. EDTA tubes produced higher numbers of small RNA transcripts with largest differences in miRNA and miscellaneous RNAs.

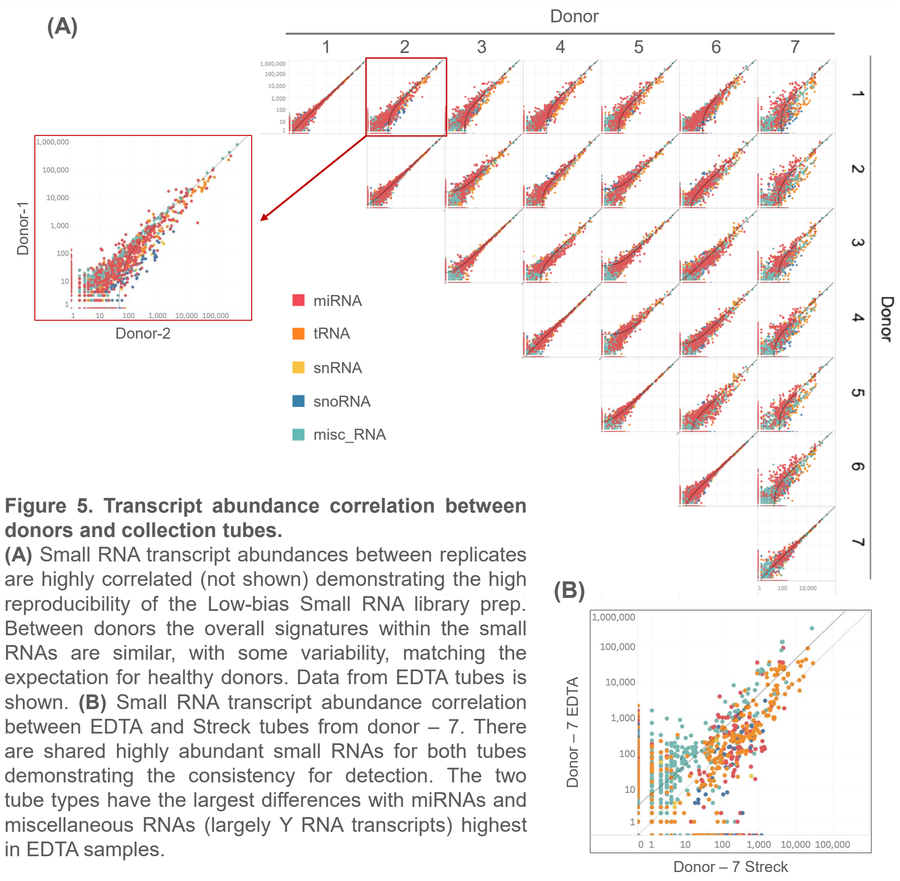


Figure 5. Transcript abundance correlation between donors and collection tubes.
(A) Small RNA transcript abundances between replicates are highly correlated (not shown) demonstrating the high reproducibility of the Low-bias Small RNA library prep. Between donors the overall signatures within the small RNAs are similar, with some variability, matching the expectation for healthy donors. Data from EDTA tubes is shown. **(B)** Small RNA transcript abundance correlation between EDTA and Streck tubes from donor – 7. There are shared highly abundant small RNAs for both tubes demonstrating the consistency for detection. The two tube types have the largest differences with miRNAs and miscellaneous RNAs (largely Y RNA transcripts) highest in EDTA samples.

Conclusions

- NEB’s Monarch solution for cell-free RNA extraction produces high quality cell-free RNA from plasma stabilized in different blood collection tubes.
- The NEBNext UltraExpress RNA Library Prep Kit produces high quality libraries from limited cfRNA input and enable target enrichment for transcripts’ profiling in plasma.
- The NEBNext Low-bias Small RNA Library Prep Kit produces high quality libraries from limited cfRNA inputs to allow profiling of small non-coding RNAs in plasma.
- Transcript abundance analysis shows variations between donors and blood collection tubes in both sequencing workflows.