Long RNA sequencing of human plasma exosomes reveals full coverage of diverse protein coding and long non-coding RNA

Sudipto K. Chakrabortty¹, Yasashi Tadigota¹, Michael Valentino¹, Robert Kitchen¹, Dalin Chan¹, Dominik Grimm¹, Lisa Bedford², Hidefumi Uchiyama³, Sunita Badola³ and Johan Skog¹

(1) Exosome Diagnostics, Inc. Cambridge, MA (2) Takeda Pharmaceuticals International Co. Translational and Biomarkers Research, Cambridge, MA (3) Takeda Pharmaceutical Co., Integrated Technology Research Laboratories, Fujisawa, Japan

Introduction

All tissues shed genetic material into biofluids such as blood, urine, CSF, and saliva. Liquid biopsies that target extracellular RNA enable monitoring of living cells, while those sampling cell-free DNA are limited to observing dead cells. Exosomes are a class of extracellular vesicles which are abundantly released from all cells into biofluids (Fig. 1). Exosomes represent the largest source of extracellular RNA and holds extraordinary potential for non-invasive biomarker discovery and diagnostics.

Growing research in oncology implicates exosomes in cell-to-cell communication and various patho-physiologies mediated via both protein and RNA. Although small RNAs have been extensively studied in exosomes, limited studies attempting to detect long RNAs have reported a relatively small proportion and poor transcript coverage. This has led many to conclude that exosomes only carry short fragments of mRNA and ncRNA.

We demonstrate a novel workflow optimized to perform RNAseq on long RNA from exosomes. Our results shed new light on the biology of exosomes and indicate that, in addition to well-recognized small RNA cargo, exosomes carry an abundance of long RNAs. This unique workflow enables biomarker discovery without the need for invasive tissue biopsies.

Why sequence long RNAs?

- Clinical biomarker space dominated by mRNA.
- No actionable microRNA based biomarkers have been identified yet.
- Interrogating long RNA cargo enables many approaches of biomarker discovery.

Workflow

2 ml plasma per sample
Clinical-grade exRNA isolation
Novel exoSeq long-RNA library workflow
Sequencing & analysis pipeline

Why sequence long RNAs?

- Clinical biomarker space dominated by mRNA.
- No actionable microRNA based biomarkers have been identified yet.
- Interrogating long RNA cargo enables many approaches of biomarker discovery.

Highly reproducible workflow

Average exosomal RNA correlation = 0.90

Uniform coverage of transcript

30% of genes have >80% coverage

Wide diversity of exosomal RNA cargo

Highly sensitive detection of RNA molecules

Top GO categories of exosomal mRNA

Conclusions

- Exosomes carry wide diversity of protein coding and long non-coding RNA.
- The workflow is highly reproducible.
- >75% of all protein coding genes are detected.
- Substantial proportion of transcripts have full coverage.
- The assay has a wide dynamic range with sensitivity down to 8 molecules.
- Optimal for measuring transcriptional activity, splice variants and fusion transcripts.
- Novel RNAseq workflow that can interrogate long RNA for biomarker discovery.