**Title**
Global landscape of extracellular RNAs and their post-transcriptional modifications in human biofluids



**Author**
Koyano, Kikuye

**Publication Date**
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

# Global landscape of extracellular RNAs and their post-transcriptional modifications in human biofluids

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy in Bioinformatics

by

Kikuye Koyano

2021

ABSTRACT OF THE DISSERTATION

Global landscape of extracellular RNAs and their post-transcriptional modifications in human biofluids

by

Kikuye Koyano

Doctor of Philosophy in Bioinformatics

University of California, Los Angeles, 2021

Professor Xinshu Xiao, Chair

Coding and noncoding RNAs (ncRNAs) are essential molecules of life. While the roles of RNA have been extensively studied inside the cell, RNA has also been detected outside of the cell and stably expressed in body fluids. These extracellular RNAs (exRNAs) are packaged inside vesicles, lipoproteins, or in ribonucleoprotein complexes which protect them from degradation by ribonucleases. exRNAs are thought to be involved in cell-to-cell communication by sending messages via extracellular vesicles. Previous studies primarily focused on total expression of exRNAs, especially microRNAs whose presence in the extracellular space is now well established. Other aspects of the

transcriptome, such as RNA modifications, were rarely investigated in exRNA profiles. In addition to microRNAs, a plural of recent studies reported the detection of mRNA fragments in the extracellular space. However, many questions remain about the mechanisms that enable their stability, sorting, and function. In this dissertation, leveraging a large amount of public and in-house exRNA-seq data, we develop and apply bioinformatic tools to analyze the patterns of post-transcriptional modifications, mRNA fragments, and repeat-derived RNAs.

We developed miNTA, a highly accurate bioinformatic method to identify 3' end non-templated additions (NTAs) on microRNAs, a type of post-transcriptional modification. We observed enriched levels of 3' uridylation across 4 biofluids, consistent with previous findings. We demonstrated that 3' uridylation levels enabled improved segregation of biofluids relative to miRNA expression levels.

Next, we analyzed mRNA fragments, repeat-derived RNAs and another type of post-transcriptional modifications, RNA editing, in exRNA-seq data of 17 human biofluids. We observed enrichment of RNA binding protein (RBP) binding sites in mRNA fragments, supporting the hypothesis that RBPs play a role in selection and stability of mRNA fragments. Additionally, we reported hundreds of RNA editing sites across biofluids.

Expanding upon examples of RNA editing across biofluids, we probed RNA editing differences in Alzheimer's Disease and control individuals using plasma-derived exRNA-seq data. Editing levels of differential editing sites were elevated in Alzheimer's Disease patients, including three sites whose levels correlated with patient cognitive scores. These

findings support the potential to uncover useful information via RNA editing analysis in exRNA profiling.

The dissertation of Kikuye Koyano is approved.

Jingyi Li

Roy Wollman

Xianghong Jasmine Zhou

Xinshu Xiao, Committee Chair

University of California, Los Angeles

2021

To my family

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Professor Xinshu Xiao for her guidance and mentorship. Her passion for science, curiosity, visionary thinking, attention to detail, leadership, and work ethic have been inspirational and aspects I aim to emulate as a scientist. In addition to her breadth of scientific strengths, I am appreciative for her personal commitment to her students as scientists. Despite her ever-increasing responsibilities, Dr. Xiao always makes time for to discuss project challenges, questions, and ideas. I am forever grateful for all her support, undivided attention, and how she has helped me grow into a better scientist and person.

I would like to thank my committee members, Professor Jingyi Jessica Li, Professor Roy Wollman, and Professor Xianghong Jasmine Zhou, for their insightful comments, unique perspectives, detailed questions, and curiosity in each meeting.

Thank you to all past and present Xiao lab members who have consistently offered technical help, scientific suggestions, friendship, and comedic relief. You all have made the Xiao lab an enjoyable environment the last 6 years. Special thanks to Mudra Choudhury and Ashley Cass who were always willing lend an ear.

I would like to acknowledge Wendy Lee for giving me my first research opportunity at UC Santa Cruz in 2012. Over the last 9 years, she has consistently supported my goals and it was her example that first inspired me to obtain a Ph.D.

Next, I would like to express my gratitude to my family who have always supported me through the many ups and downs of life. I would like to give special thanks to my parents, Karen and Duane Koyano, for their love, encouragement, personal sacrifices, and support enabled me to pursue a Ph.D. I dedicate all my accomplishments to them. I would like to thank Kelly Koyano, our sisterhood holds a special place in my heart. Additionally, I would like to recognize my grandparents, who could not be here to witness this dissertation, but their love, perseverance through financial and racial struggles have contributed towards these achievements.

Thank you to all my friends who have supported me during this process. Thank you to my former soccer teammates at UC Santa Cruz who have always cheered me on, particularly Gloria Hernandez who I had the honor of also experiencing graduate school with at UCLA. I would also like to thank my colleagues in the bioinformatics program, specifically Kathy Burch for study sessions and laughs and to Artur Jaroszewicz, my climbing partner and retreat co-coordinator.

Last, but certainly not least, I would like to say a very special Kevin Nguyen who has supported me through Ph.D. application to defense. You are my trusted confidant and best friend.

This work was partially supported by Eureka Fellowship (Training Grant at UCLA, 2018-2019).

Chapter 2 is a version of the published open access article: **Koyano K**, Bahn JH, Xiao X. Extracellular microRNA 3' end modification across diverse body fluids.

Epigenetics 00, 1–16 (2020) 10.3389/FNCEL.2014.00314. K.K. and X.X. designed the study and wrote the paper. J.H.B conducted molecular and cellular experiments. K.K. conducted bioinformatics analysis. X.X. was the principal investigator.

Chapter 3 is currently being prepared as a manuscript for submission: Koyano K, Xiao X. “Extracellular RNA fragments and RNA editing in human biofluids.” K.K. and X.X. designed the study and wrote the paper. K.K. performed all described analyses. X.X. was the principal investigator.

Chapter 4 is currently being prepared as a manuscript for submission: Koyano K, Xiao X. “Extracellular RNA editing in plasma of Alzheimer’s Disease patients.” K.K. and X.X. designed the study and wrote the paper. K.K. performed all described analyses. X.X. was the principal investigator. I would like to thank Dr. Shusuke Toden and Dr. Arkaitz Ibarra for providing clinical metadata.

# VITA

| | |
|---|---|
| 2009 – 2013 | B.S. Biomolecular Engineering<br>University of California, Santa Cruz |
| 2013 | Intern, Protein Sciences<br>NGM Biopharmaceuticals, South San Francisco |
| 2014 – 2015 | Research Assistant<br>University of Southern California |
| 2019 | Teaching Assistant, PHYSCI125: Molecular Systems Biology,<br>University of California, Los Angeles |
| 2015 – 2021 | Graduate Student Researcher, Bioinformatics Interdepartmental Program, University of California, Los Angeles |

**Awards**

| | |
|---|---|
| 2018 | Eureka Scholarship (Integrative Biology and Physiology)<br>University of California, Los Angeles |
| 2013 | Highest Honors in the Major<br>UCSC Baskin School of Engineering |
| 2013 | Cum Laude<br>UCSC Baskin School of Engineering |
| 2011- 2013 | All-Academic Team Women's Soccer<br>The Association of Independents Division III |
| 2011 – 2013 | Scholar Athlete<br>University of California, Santa Cruz |

## Publications

**Koyano K,** Xiao X**.** Extracellular RNA editing in plasma of Alzheimer's Disease patients. In preparation.

**Koyano K,** Xiao X**.** Extracellular RNA fragments and RNA editing in human biofluids. In preparation.

Weitz SH, Quick-Cleveland J, Jacob JP, Barr I, Senturia R, **Koyano K**, Xiao X, Weiss S and Guo F. Fe(III) heme sets an activation threshold for processing distinct groups of pri-miRNAs in mammalian cells. bioRxiv (2020). 10.1101/2020.02.18.955294.

Tavanasefat H, Li F, **Koyano K**, Gourtani BK, Marty V, Mulpuri Y, Lee SH, Shin KH, Wong DTW, Xiao X, Spigelman I, Kim Y. Molecular consequences of fetal alcohol exposure on amniotic exosomal miRNAs with functional implications for stem cell potency and differentiation. PLoS One. 15, e0242276 (2020). 10.1371/journal.pone.0242276

**Koyano K**, Bahn JH, Xiao X. Extracellular microRNA 3' end modification across diverse body fluids. Epigenetics 16, 1000 (2020) 10.3389/FNCEL.2014.00314

He X*, Li F*, Bor B[#], **Koyano K**[#], Cen L, Xiao X, Shi W, Wong DTW. Human tRNA-Derived Small RNAs Modulate Host-Oral Microbial Interactions. Journal of Dental Research. 97, 1236 (2018). 10.1177/0022034518770605 (*Contributed equally, [#]Contributed equally)

Li F*, Kaczor-Urbanowicz KE*, Sun J, Majem B, Lo H, Kim Y, **Koyano K**, Rao SL, Kang SY, Mi Kim S, Kim KM, Kim S, Chia D, Elashoff D, Grogan TR, Xiao X, DTW Wong. Characterization of Human Salivary Extracellular RNA by Next-generation Sequencing. Clinical Chemistry. 64, 1085 (2018). 10.1373/clinchem.2017.285072 (*Contributed equally)

Kaczor-Urbanowicz KE, Kim Y, Li F, Galeev T, Kitchen RR, Gerstein M, **Koyano K**, Jeong S, Wang X, Elashoff D, Kang SY, Kim SM, Kim K, Kim S, Chia D, Xiao X, Rozowsky J, Wong DTW. Novel approaches for bioinformatic analysis of salivary RNA sequencing data for development. Bioinformatics. 34, 1 (2018). 10.1093/bioinformatics/btx504

# Chapter 1 : Introduction

RNAs are involved in a wide range of functions, from acting as messengers between DNA and protein to operating as gene regulators in important cellular processes (e.g. development, disease etc.)[1]. Although intracellular RNA has been extensively studied, abundant and diverse species of coding and non-coding RNAs have also been found stably expressed outside the cell, namely extracellular RNA (exRNA)[2–4]. exRNAs have been identified across diverse biofluids in mammals, nematodes, and plants[4–6].

exRNA in biofluids was unexpected since biofluids contain ribonucleases (RNases) outside of cells that degrade RNA molecules, such as foreign and viral RNAs. It is now established that exRNAs are stably present in biofluids through encapsulation in vesicles, lipoprotein, apoptotic bodies, or binding to ribonucleoprotein complexes[7–10]. While exRNAs were first thought to be non-functional and regarded as “junk” from cells, a large body of evidence has shown exRNAs to be functionally active in recipient cells and reflective of the cellular transcriptome of their cells of origin[11–13]. Cell-to-cell communication is hypothesized to be one of the main functions of exRNA[6].

Identification of exRNA unleashed a wave of studies focusing on its translational applications in identifying biomarkers via “liquid biopsies”, i.e. the non-invasive sampling of non-solid biological tissues (e.g. plasma, urine, etc.) in order to monitor diseases. Biomarkers are useful in areas of preventative medicine, drug-screening, and detecting

physiological responses. To use exRNA as biomarkers, it is important to understand their function and how they relate to diseases.

A majority of exRNA studies focus on expression levels, specifically of small non-coding RNAs called microRNAs (miRNAs). miRNAs are short 18-22nt ncRNA species. They play critical roles in gene regulation by targeting complementary mRNAs, modulating mRNA or protein levels. Exosomes containing miRNAs were shown to affect neighboring cells locally in the tumor microenvironment or in distal organs during cancer metastasis[12,14]. The presence of miRNAs in extracellular biofluids is well recognized. While most previous characterization of extracellular miRNAs focused on their overall expression levels, alternative sequence isoforms and modifications of miRNAs are known to exist but were rarely considered in the extracellular space. Non-templated additions on the 3' end of miRNA have been shown to impact targeting, stability, or turnover[15–17]. In chapter 2, we developed a highly accurate bioinformatic method, called miNTA, to identify 3' non-templated additions (NTAs) of miRNAs from small RNA-sequencing data.

The majority of extracellular RNA sequencing data is size-selected for small RNAs (17 - 35nt) as most existing studies focused on miRNAs. For example, the exRNA atlas is a data repository of the Extracellular RNA Communication Consortium (ERCC). This database currently only includes small RNA sequencing and qPCR-derived exRNA profiles[18]. Long RNA-seq in extracellular fluids is a neglected area overshadowed by miRNA profiling. Longer RNAs, such as lincRNAs and mRNAs, largely existing as fragments, have been found in exRNA fractions and importantly, shown to contain biologically relevant signals[19–22]. For example, mRNAs in low-density exRNA fractions

from mast cells correlate with levels of the corresponding intracellular mRNAs[23]. Additionally, long RNAs were shown to be differentially sorted into extracellular vesicles secreted by colorectal cancer cells[21]. Revealing the mechanisms that enable long exRNA stability, selection, and function is critical for understanding exRNA biology. Chapter 3 aims to address this significant gap by implementing bioinformatic methods to identify mRNA fragments from RNA-seq datasets across 17 biofluids. By identifying these RNA fragments, we revealed enrichment of binding sites from RNA binding proteins (RBPs).

Although several studies have examined the diverse exRNA molecules present in biofluids, repeat-derived RNAs and RNA modifications are two important aspects of RNA biology that have received relatively little attention thus far. RNAs generated by repeat elements have been shown to play roles in various aspects of gene regulation and RNA processing[24,25]. In addition to repeat-derived RNAs, we examined RNA editing in long exRNA-seq, a type of post-transcriptional RNA modification. The most common type of RNA editing is the conversion of the nucleotides adenosine to inosine (A-to-I), catalyzed by the ADAR family of proteins[26]. This conserved post-transcriptional modification affects coding and non-coding RNAs and plays important roles in immune response, neurodevelopment, RNA processing and disease[27–30]. While much is being learned about intracellular repeat-derived RNAs and RNA editing, little has been done to characterize these modifications in the extracellular space. In chapter 3 we examined repeat-derived RNAs and RNA editing from extracellular RNA isolated from 17 biofluids.

Alzheimer's Disease (AD) is a progressive neurodegenerative disease characterized by accumulation of amyloid plague in the brain and neuronal loss[31].

Overtime, symptoms of memory loss and difficulties in speech begin to manifest[32]. Due to the high prevalence of AD, affecting 6.2 million Americans over the age of 65[31], non-invasive biomarkers for early diagnosis are being heavily investigated[33]. RNA editing has been previously investigated for prognostic value and was shown to be decreased in post-mortem brains of AD individuals compared to controls[34,35]. However, exRNA editing profiles in AD has not been characterized. In Chapter 4 we explored the RNA editing profile of long exRNA from plasma of AD and control individuals. In this chapter, we reveal RNA editing dysregulation in plasma from AD individuals.

# Chapter 2 : Extracellular microRNA 3’ end modification across diverse body fluids

## 2.1 Abstract

microRNAs (miRNAs) are small non-coding RNAs that play critical roles in gene regulation. The presence of miRNAs in extracellular biofluids is increasingly recognized. However, most previous characterization of extracellular miRNAs focused on their overall expression levels. Alternative sequence isoforms and modifications of miRNAs were rarely considered in the extracellular space. Here, we developed a highly accurate bioinformatic method, called miNTA, to identify 3’ non-templated additions (NTAs) of miRNAs using small RNA-sequencing data. Using miNTA, we conducted an in-depth analysis of miRNA 3’ NTA profiles in 1047 extracellular RNA-sequencing data sets of 4 types of biofluids. This analysis identified hundreds of miRNAs with 3’ uridylation or adenylation, with the former being more prevalent. Among these miRNAs, up to 53% (22%) had an average 3’ uridylation (adenylation) level of at least 10% in a specific biofluid. Strikingly, we found that 3’ uridylation levels enabled segregation of different types of biofluids, more effectively than overall miRNA expression levels. This observation suggests that 3’ NTA levels *

* The work appearing in this chapter is published: **Koyano K**, Bahn JH, Xiao X. Extracellular microRNA 3’ end modification across diverse body fluids. Epigenetics 16, 1000 (2020) 10.3389/FNCEL.2014.00314

possess fluid-specific information relatively robust to batch effects. In addition, we observed that extracellular miRNAs with 3' uridylations are enriched in processes related to angiogenesis, apoptosis and inflammatory response, and this type of modification may stabilize base-pairing between miRNAs and their target genes. Together, our study provides a comprehensive landscape of miRNA NTAs in human biofluids, which paves way for further biomarker discoveries. The insights generated in our work built a foundation for future functional, mechanistic and translational discoveries.

## 2.2 Introduction

Recent studies revealed the existence of extracellular RNAs (exRNAs) in many types of biofluids [36,37]. exRNAs are mostly packaged in small extracellular vesicles, microvesicles or in complex with lipoproteins or ribonucleoproteins [18], which protect them from degradation by ribonucleases. exRNA expression could be highly cell type- or disease-specific [38,39], thus affording potential values as disease biomarkers [40]. Importantly, the functional roles of exRNAs are also starting to unfold [41–43]. For example, several studies reported the involvement of exRNAs in cell-to-cell communication in the local tumor microenvironment and during angiogenesis [9,41,44].

The most-often studied exRNAs are microRNAs (miRNAs), small 18-22nt noncoding RNAs that are potent regulators of mRNA and protein expression levels [15]. Most previous studies on extracellular miRNAs focused on interrogating their overall expression levels. Nonetheless, many miRNAs assume multiple sequence forms

resulted from alternative miRNA processing or post-transcriptional modification [45]. Specifically, a well-known class of post-transcriptional modification of miRNAs is non-templated addition (NTA) [46]. Two types of 3' miRNA NTAs have been reported [47,48], 3' adenylation catalyzed by GLD2 and 3' uridylation by the terminal uridyltransferase-4 and 7 (TUT4/TUT7). Both types of 3' NTAs may affect miRNA targeting, stability, or turnover [15–17].

Thus far, only a small number of studies examined miRNA NTAs in the extracellular space. For example, a study using cultured human B cells examined 3' NTAs of intracellular and extracellular exosomal miRNAs [49]. The authors observed that 3' NTAs of miRNAs showed distinct patterns in the two compartments, with 3' adenylation more enriched intracellularly and 3' uridylation overrepresented in exosomes. Another study examined global miRNA expression in blood cells, serum and exosomes [50]. They showed that 3' NTA patterns clustered in a blood-lineage specific manner and extracellular 3' NTAs were distinct from the intracellular profiles. These findings suggest that 3' NTA patterns of miRNAs may carry specific information that segregates extra- and intracellular miRNA profiles. The mechanisms underlying this specificity remain unclear.

In this study, we developed a new analysis pipeline, called miNTA, to identify NTAs of miRNAs in any small RNA-sequencing (RNA-seq) data set and applied it to 1047 extracellular samples derived from 4 types of biofluids. To our best knowledge, this is the largest study of extracellular miRNA NTA profiles in biofluids. Although many studies have examined NTAs of intracellular miRNAs, the bioinformatic pipelines

employed in most studies could be improved to enhance accuracy. Incorporating a number of stringency measures, our method achieves a low false discovery rate < 5%. Applied to the large number of biofluid samples, we observed 3' uridylation and adenylation as the two most prominent types of 3' NTAs in extracellular miRNAs, with the former being more prevalent than the latter. Our analysis showed that 5' NTAs are unlikely present in miRNAs, or extremely rare if exist at all. The levels of 3' NTAs varied widely across miRNAs. Importantly, 3' NTA levels can be used to segregate different types of biofluids, more effectively than miRNA expression levels. We also observed that extracellular miRNAs with 3' uridylations are enriched in processes related to angiogenesis, apoptosis and inflammatory response, and this type of modification may stabilize base-pairing between miRNAs and their target genes. Overall, our study provides global insights regarding 3' end modifications of miRNAs in extracellular fluids.

## 2.3 Results

### 2.3.1 miNTA: A bioinformatic pipeline to identify miRNA NTAs

To explore the diversity of NTAs in the extracellular space, we first developed a rigorous pipeline, called miNTA, to accurately identify NTAs of miRNAs in small RNA-seq data (Fig. 2.1A). While other methods to identify miRNA sequence variations exist[51–54], our method aims to improve the mapping strategies and reduce false positive modifications (Methods).

miRNA 3' adenylation and uridylation are known to be mediated by specific enzymes such as, but not limited to, GLD-2 and TUT4/TUT7, respectively[15]. To evaluate

if our pipeline detects biologically relevant NTAs, we performed double knockdown (KD) of the TUT4 and TUT7 enzymes in HEK293 cells, followed by small RNA-seq (Methods and Supplementary Fig. S2.1A). As expected, we observed reduced global 3' uridylation levels of miRNAs upon TUT4/7 KD relative to the controls (Fig. 2.1B). In addition, a concomitant increase of 3' adenylation levels was observed, consistent with previous literature[16,48]. Compared to two other popular methods, sRNAbench and Chimira[53,55], miNTA identified the highest fraction of 3' A and U, and correspondingly, lowest 3' G and C NTAs (Supplementary Fig. S2.1B, C). Since there are no known enzymes for 3' G and C NTAs on miRNAs in mammals, these NTAs are considered false positives. Similarly, we applied our method to two GLD2 KD datasets and observed an expected reduction of 3' adenylation compare to control cells (Supplementary Fig. S2.2) [47,56]. Thus, the results here demonstrate the superior performance of miNTA, which is further evaluated below.

### 2.3.2 Comprehensive catalog of miRNAs in the extracellular space

To investigate the landscape of extracellular miRNA 3' NTAs, we obtained small RNA-seq data sets (50 nt, strand-specific) of four bodily fluids of healthy subjects obtained from previous studies [57–61] (Fig. 2.2A). In total, we analyzed 1047 data sets, including 399 plasma samples, 163 samples of small extracellular vesicles isolated from plasma, 167 serum, 69 cerebral spinal fluid (CSF), and 249 urine samples. As a comparison, we also analyzed 297 intracellular data sets of seven types of human

peripheral blood cells (NK cells, B lymphocytes, cytotoxic T lymphocytes, T-helper cells, monocytes, neutrophils and erythrocytes) sorted from whole blood [50].

Although intracellular miRNAs have been studied extensively, the repertoire of extracellular miRNAs is still being explored. To create a comprehensive list of human miRNAs, as the first step of the pipeline, we ran miRDeep2 [62] on all extracellular and intracellular data sets to identify novel miRNAs. This procedure identified a total of 404 novel miRNAs, present in more than one sample (Fig. 2.2B). Interestingly, 85.3% of these novel miRNAs were detected in more than 10 samples, a slightly higher percentage than that (81.7%) of known miRNAs. Nonetheless, novel miRNAs had relatively lower expression levels than known miRNAs (Fig. 2.2C), likely explaining their absence in the miRBase annotation. Notably, certain novel miRNAs may have higher expression levels in the extracellular space, such as the example shown in Fig. 2.2D (derived from paired intra- and extracellular data sets [63], also see Supplementary Fig. S2.3). Henceforth, we include both annotated and novel miRNAs in the analysis.

### 2.3.3 miRNA NTAs in the extracellular fluids and evaluation of the NTA pipeline

Next, we examined NTA profiles identified in individual miRNA reads, without grouping reads per miRNA. On average across fluids, 5.2% or 0.3% of all miRNA reads had 3’ or 5’ end NTAs, respectively (Fig. 2.3A). For intracellular samples, 11.2% or 0.2% of total miRNA reads had 3’ or 5’ end NTAs, respectively (Supplementary Fig. S2.4A).

We next examined the nucleotide composition of the 3' NTAs across miRNA reads for each fluid. This analysis also allows us to evaluate the quality of our pipeline as miRNAs are expected to have predominantly two types of 3' NTAs (adenylation and uridylation), based on previous reports [47,49]. On average, >94% of all mono 3' end modifications in reads of extracellular samples were identified as either adenylation or uridylation, whereas 3' end cytosine or guanosine additions were each less than 3.4% (Fig. 2.3B). Similar results were also observed for intracellular groups (Supplementary Fig. S2.4B). If the observed 3' G or 3' C modifications were assumed to be false positives, then the false discovery rate (FDR) of our predicted mono 3' NTA-containing reads would be estimated to be <5% for exRNAs, and <2% for intracellular miRNAs. Note that these FDRs may be over-estimated since 3' G or C NTAs may exist, although no known enzymes have been reported for these modifications in mammals.

Our bioinformatic pipeline also allowed an investigation of 5' NTAs of miRNAs. In general, the prevalence of 5' NTAs were much lower than that of 3' NTAs (Fig. 2.3A). Despite this low level, a strong enrichment of cytosine among the 5' NTAs was observed (Supplementary Fig. S2.5A). A very low range of 5' Cs (average <0.028%) was observed for 78 miRNAs in Plasma from Lab 5 (Supplementary Fig. S2.5B). Since no known mechanisms exist to account for 5' C modification of miRNAs in mammals, the observed 5' C NTA may reflect technical rather than biological mechanisms. For example, the 5' adaptor used in small RNA library generation ended with a C nucleotide, which may have been read as the first base of the read as a type of sequencing error. Another report also observed a high proportion of 5' addition of C [54]. Although the 5' C may be an experimental artifact, the fact that this strong nucleotide

bias was observed despite the overall low prevalence of 5’ NTAs strongly supports the effectiveness of our pipeline.

### 2.3.4 miRNA 3’ uridylation is relatively more prevalent than 3’ adenylation in biofluids

Next, we examined the landscape of 3’ NTAs of miRNAs by grouping reads for each miRNA. In this analysis (and all others henceforth), we required ≥2 reads in each sample carrying the NTA nucleotide, with ≥5 samples satisfying this requirement in the same data set. When summarizing results for each fluid data set, we further required that ≥10 total reads mapped to the corresponding miRNA in each sample.

We observed that 32% (496/1541) or 26% (405/1541) of all detected miRNA species had 3’ uridylation or 3’ adenylation, respectively, considering all the biofluid data sets. Although the total numbers of modified miRNA species were similar, 3’ uridylation in reads mapped to miRNAs was significantly more frequent than 3’ adenylation, 58% and 27% among all types of 3’ NTAs, respectively (Fig. 2.3C). In contrast, the opposite trend was observed in intracellular samples, where a larger fraction of miRNAs had 3’ adenylations than uridylations (61% and 31%, respectively) (Fig. 2.3C). Overall, miRNA 3’ uridylation was more frequent in extracellular data sets (58%) compared to intracellular data sets (31%) (Wilcoxon rank sums test, p value = 0.003).

### 2.3.5 miRNAs exhibit a wide range of NTA levels in biofluids

The 3' uridylation levels varied greatly across miRNAs in each type of fluid (Fig. 2.4A). While many miRNAs had relatively low 3' U levels, a number of them had modest to high levels. For example, 39% (122/312) of 3' U-harboring miRNAs in plasma (Lab 1 data set) had ≥20% 3' U levels. A substantial fraction (23-53%) of 3' uridylated miRNAs had an average 3' U level of at least 10% in each specific type of biofluid (Fig. 2.4A). In particular, the hsa-let-7f-2-3p demonstrated a high level of 3' U (ranging 14-93%) across all extracellular data sets. miRNAs with ≥5% average 3' uridylation level across samples with the modification are listed in Supplementary Table S2.1. Note that the 3' U levels of the same miRNA between different data sets may not be comparable due to the distinct experimental protocols used to generate each data set. Among the three studies where at least 10 novel miRNAs were detected with 3' uridylation, two showed higher 3' U levels in novel miRNAs than known ones (Wilcoxon rank-based test, Supplementary Fig. S2.6).

Similarly, the 3' adenylation levels also varied across miRNAs (Fig. 2.4B). Consistent with the previous observation of lower 3' A levels compared to those of 3' U, only 14.7% (27/189) of 3' A-harboring miRNAs had a 3' A level ≥20% in the plasma (Lab 1) data set. About 6-22% of 3' adenylated miRNAs had an average 3' A level of at least 10% in each type of biofluid. Nonetheless, a small number of miRNAs had considerable levels of 3' adenylation (Supplementary Table S2.2). An example is hsa-miR-6513-3p that demonstrated a high level of 3' A (ranging 8-48%) across 3 extracellular data sets. No studies were detected with at least 10 novel 3' adenylated miRNAs.

### 2.3.6 NTAs often occur in miRNAs relevant to angiogenesis or signaling

To gain insights on the functional relevance of 3' uridylation of miRNAs, for each fluid in each data set, we performed Gene Ontology (GO) enrichment analysis on miRNAs that have an average 3' uridylation level ≥ 5% in at least 50% of samples. As background controls, miRNAs without 3' modifications in our data were chosen randomly by matching their expression levels with those that harbored 3' Us (Methods). Interestingly, the GO term "extracellular space" was significantly enriched (FDR < 0.05) for all data sets (Fig. 2.4C). Since this analysis controlled for expression, this observation supports the enrichment of 3' uridylation of miRNAs in the extracellular space. In addition, we observed terms related to angiogenesis, apoptosis, gene regulation and inflammatory response, indicating the involvement of 3' U-containing miRNAs in these processes. Since miRNAs in the let-7 miRNA family are known to be enriched with 3' uridylation [64], we repeated this analysis by excluding let-7 miRNAs. Similar enriched GO terms were observed, supporting the generality of the results for diverse miRNA species (Supplementary Fig. S2.7). For 3' adenylated miRNAs, an enrichment for angiogenesis-related terms was observed, but mostly from the urine data set (Lab 1) (Fig. 2.4D).

### 2.3.7 3' uridylation levels of miRNAs segregate biofluids robust to batch effects

Given the wide range of 3' uridylation levels in miRNAs, we asked whether this modification can help to segregate different types of fluids. tSNE clustering on 3'

uridylation levels of all expressed miRNAs showed that samples derived from similar fluid types tend to cluster together (Fig. 2.5A). Specifically, serum, urine and plasma samples each contained data sets generated by two different labs. For these fluids (especially serum and urine), we observed that samples of the same fluid type generated by different labs largely clustered together. This observation strongly suggests that 3' uridylation levels are informative in segregating fluid types. In addition, we observed that intracellular blood cell types descending from the lymphoid lineage, such as T lymphocytes (CD4+, CD8+), B lymphocytes (CD19+) and natural killer cells (CD56+), clustered separately from cell types of the myeloid lineage. Myeloid cell types such as erythrocytes (CD235a), neutrophils (CD15+), and monocytes (CD14+) also clustered separately from each other. Cell types clustering by blood lineage based on their 3'-end composition was previously reported [50]. Exosomal plasma samples did not cluster with plasma/serum samples, likely due to batch effects or the biological difference between exosomal and total extracellular miRNA contents. Segregation of samples based on 3' adenylation levels was not as effective as using 3' uridylation levels (Supplementary Fig. S2.8A), although serum and plasma samples from different labs did cluster together.

As a comparison, we also performed tSNE on normalized miRNA expression values (Methods). This analysis showed exacerbated batch effects where data generated from the same lab largely clustered together, instead of clustering by fluid types (Supplementary Fig. S2.8B). PCA analysis of these samples showed improved fluid segregation, but still confounded by batch effects to some degree (Supplementary

Fig. S2.8C). These results are consistent with observations made in previous studies where batch effects confounded the clustering of miRNA expression across samples [18].

### 2.3.8 Comparison of 3' uridylation of miRNAs between biofluids

These results suggest that different types of biofluids have distinct levels of 3' uridylation. Thus, 3' NTAs add another layer of information to distinguish fluid types. Nonetheless, since batch effects may not be completely absent, we avoided carrying out direct comparisons of the quantitative levels of NTAs across different studies. Instead, we performed differential modification analysis between fluids of the same study [65] (Methods).

For this analysis, we included miRNAs expressed in at least 20 samples in both fluids of a study. Overall, we identified 62 miRNAs with differential 3' uridylation levels (≥5% difference in modification levels, FDR <0.05) (Supplementary Table S2.2). Specifically, 25 miRNAs had differential 3' uridylation between urine and plasma (Lab 1), 11 between CSF and serum (Lab 3), 12 between urine and exosomes from plasma (Lab 4), and 29 between plasma and serum (Lab 5). Fig. 2.5B-D show three example miRNAs whose 3' uridylation levels were high in one fluid, but almost zero in another.

### 2.3.9 3' uridylation may increase miRNA base-pairing to its targets

Given the relatively high levels of 3' uridylation in extracellular miRNAs, we hypothesized that this modification may affect the base-pairing between miRNAs and

their targets. We focused on the top 20 unique miRNAs with highest average 3' uridylation levels (per fluid) across all samples (Methods). Using RNAhybrid [66], we observed that 3' uridylation tends to stabilize the interaction between miRNAs and their targets, to an extent much greater than 3' NTAs of G, C or A nucleotides (Fig. 2.6A, Chi-squared test $p < 0.05$). All miRNAs were significant against at least one of the background nucleotides. Target genes that paired with the 3' U of these miRNAs were enriched in a variety of processes including extracellular exosome, integral protein of plasma membrane and negative regulation of apoptosis (Fig. 2.6B). Thus, it is possible that 3' uridylation of extracellular miRNAs may affect miRNA targeting once taken up by recipient cells.

## 2.4 Discussion

We developed miNTA, a new bioinformatic pipeline, to identify 3' NTAs of miRNAs. With different types of data sets, we demonstrated that the pipeline yields accurate and sensitive predictions. Using miNTA, we analyzed the global landscape of 3' NTAs of extracellular miRNAs in >1000 biofluid samples, the largest study so far for extracellular miRNA modifications (to the best of our knowledge). We made a number of notable observations: (1) 3' uridylation levels of miRNAs are higher in the extracellular space than 3' adenylation levels, whereas the opposite was observed for intracellular miRNAs. (2) The level of 3' NTAs varied widely across miRNAs, with some miRNAs demonstrating nearly 100% 3' uridylation in certain biofluids. (3) 3' uridylation levels of miRNAs can inform segregation of different types of biofluids. In addition, such

segregation was more effective than that achieved by miRNA expression levels and largely robust to batch effects. (4) Extracellular miRNAs with 3' uridylations are enriched in functional categories related to angiogenesis, apoptosis and inflammatory response, and 3' uridylation may stabilize base-pairing between miRNAs and their target genes.

The effective segregation of biofluids by 3' uridylation levels of miRNAs indicates that this type of miRNA modification possesses fluid-specific signatures. Thus, miRNA modification levels could serve as biomarkers that are less susceptible to batch effects. This property is likely due to the fact that modification levels are normalized metrics relative to the total miRNA expression. Differences in uridylation levels across fluids may be due to different cell types contributing distinct miRNA species into the extracellular space. Together, these results show that 3' modification levels of miRNAs add an important layer of information to characterize extracellular RNA content.

The enrichment of 3' uridylated extracellular miRNAs in angiogenesis, apoptosis and inflammatory responses suggests that this modification may have important functional relevance. For example, angiogenesis is a tightly regulated process that requires endothelial cells to be in close communication with their environment [67]. Extracellular vesicles were previously shown to play a role in regulating angiogenesis [41,67]. Thus, 3' uridylation of miRNAs may be an important aspect contributing to this process. It should be noted that the functional enrichment analysis controlled for extracellular expression levels of 3' uridylated miRNAs. Thus, the enriched categories reflect functions that are particularly relevant to 3' uridylated miRNAs instead of extracellular miRNAs in general. We showed that 3' uridylation may strengthen miRNA

targeting. Interestingly, a previous study reported that 3' uridylation of miR-27a induced target repression of ‘non-canonical’ sites with only partial seed-match and extensive 3’ end pairing [16]. Future studies are needed to understand the functional relevance of 3’ uridylation in extracellular miRNAs.

Extracellular miRNAs have been heavily investigated to identify miRNAs associated with cancer, such as oncomiRs [68]. While most studies focused on canonical miRNA expression levels, miRNA modifications and sequence variants also have potential as diagnostic or prognostic biomarkers. For example, consideration of sequence variants of extracellular miRNAs in urine enabled improved segregation of control and prostate cancer patients than using canonical miRNA expression alone [69]. Similarly, expression of isomiR of miR-21-5p (3′ addition C) was significantly higher in serum from breast cancer patients compared to normal control samples [70]. In the same study, a multiple regression model including expression of miR-21-5p (3′ addition C), canonical miR-21-5p, and tRF-Lys (TTT) was able to discriminate stage 0 breast cancer from the control group. Thus, incorporating both known mature and miRNA variant levels and abundances may have valuable applications in translational research.

In summary, we presented an accurate pipeline to identify 3’ NTA patterns in extracellular miRNAs. Our large-scale analysis of data from different human biofluids supports that fluid-specific signatures exist in 3’ modifications of miRNAs. Our work extends the basis for future studies on the functional relevance of extracellular miRNA modifications and their values in biomarker discoveries.

## 2.5 Methods

### 2.5.1 Cell Culture

Human embryonic kidney cell line (HEK293T) was obtained from the ATCC. Cells were maintained in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum (FBS) with antibiotics at 37C in 5% CO2.

### 2.5.2 shTUT4 and shTUT7 knockdown

We used lentivirus-packaged short hairpin RNA (shRNA) to knock down TUT4 and TUT7. The shRNA sequences (purchased from IDT) were obtained from a previous study [48] and non-target shRNA control plasmids (SHC016) from a previous publication [71]. Co-transfection of pCMV-d8.91, pVSV-G and pLKO1 into HEK293T cells was performed using Lipofectamine 3000 (Thermo Fisher Scientific, Cat# L3000-008). Lentiviruses were collected from conditioned media 48hrs after transfection. Lentivirus-containing media was filtered using 0.45 μM PES syringe filter (VWR) and mixed with polybrene (8 μg/ml). Following 24hrs of infection, cells were incubated with puromycin (1 μg/ml) for 3-4 days. To make double knock down cell lines, second round of infected cells were incubated with hygromycin (200 μg/ml) for 3-4 days. Knockdown efficiency was evaluated by Western blot using TUT4 (Proteintech Inc, Cat# 18980-1-AP), TUT7 (Bethyl Laboratories cat# A305-089A) and beta-actin antibodies (Santa Cruz Biotech, Cat# sc-47778 HRP).

### 2.5.3 Extracellular and intracellular RNA isolation

Lentivirus-infected HEK293 cells were washed three times with PBS and the medium was switched to serum-free medium containing antibiotic-antimycotic (Thermo Fisher Scientific, Cat# 15240112). Following 24hrs incubation, the cell culture medium was collected and centrifuged at 2,000 g for 15 min at room temperature. To thoroughly remove cellular debris, the supernatant was centrifuged again at 12,000 g for 35 min at room temperature. Then the conditioned medium was used for RNA extraction with Trizol (Thermo Fisher Scientific, Cat# 10296028). Intracellular RNA was isolated using Direct-zol RNA mini prep kit (Zymo Research) from the same culture dish for extracellular RNA.

### 2.5.4 Small RNA library preparation

Small RNA sequencing libraries were generated using the NEBNext Small RNA library Prep kit and NEBNext multiplex oligos for Illumina according to the manufacturer's instructions (New England Biolabs, E7300). The final small RNA libraries were purified from 6% PAGE gel, and their concentrations were measured by Qubit fluorometric assay (Life Technologies). Libraries were sequenced on an Illumina HiSeq-3000 (50-bp single-end).

### 2.5.5 Small RNA-sequencing data processing

For each small RNA-seq data set, adapters and low-quality nucleotides were removed from raw fastq sequences using cutadapt (v.1.11) [72]. Following read mapping, we removed samples with <50,000 total reads mapped to miRNAs to ensure at least a modest sequencing coverage.

### 2.5.6 miNTA: read mapping

To enable comprehensive read mapping, miNTA includes a multi-step mapping strategy. miNTA rescues unmapped reads by performing 2 sequential rounds of trimming on the 3' or 5' end of unmapped reads and remapping them. First, all reads were aligned to the human genome (hg19) using Bowtie (version 1.1.2) [73], allowing up to one mismatch and retaining uniquely mapped reads only. This stringency aims to minimize ambiguous mapping results. Next, the unmapped reads were trimmed by 1 nucleotide at their 3' ends and realigned to the human genome according to the same requirements as described above. Unaligned reads from this step were trimmed again and aligned for another round. The above procedures were repeated using all original reads to carry out 5' trimming and identify 5' end NTAs. All remaining unmapped reads were restored to their original sequences and realigned after trimming 1 nucleotide each from the 3' and 5' end, respectively. Finally, the mapped reads (trimmed or untrimmed) were examined relative to the human genome reference to identify 3' and 5' NTAs.

### 2.5.7 miNTA: quality control (QC) procedures

*QC of mapped reads*

Incorrect mapping can lead to false positive predictions of NTAs. To ensure accurate read mapping, we realigned reads that mapped to miRNAs using BLAT [74] against the human genome. Those reads that did not yield consistent alignment results by BLAT and Bowtie were removed from further analysis.

*Canonical end positions of miRNAs*

For each miRNA in each sample, the canonical 3' end was defined as the position of the last nucleotide in the miRNA read sequence that matches the reference genome and supported by at least 50% of reads aligned to this miRNA. For 5'NTA analysis, the canonical 5' end of each miRNA was defined similarly for each sample. Subsequently, NTAs were identified relative to the canonical end positions for each miRNA.

*QC of modifications*

The predicted NTAs were further examined to eliminate those that may reflect genetic variants or technical artifacts. First, predicted NTAs that overlapped known SNPs or genetic mutations were removed [75–79]. Second, NTAs with a 100% modification level were removed as they may be due to mapping errors (similarly as in other studies of single-nucleotide variants [80,81]). Third, to minimize false positives due to likely sequencing errors, we removed reads with a PHRED score < 30 at the position corresponding to the NTA. Lastly, for NTAs of each miRNA, we required at least 2 reads

with the modification and the NTAs observed in at least 5 samples within the same fluid and data set group.

### 2.5.8 Comparison of detected 3'NTAs against Chimira and sRNAbench

As a proxy for error, we compared the proportion of identified 3'NTA that were Gs and Cs found in miNTA against sRNAbench (https://arn.ugr.es/srnatoolbox) and Chimira (version 1.5), two other methods known to identify 3'NTAs[53,55]. The comparison was performed on our six intracellular HEK293 samples (shControl and shTUT4 and shTUT7 double knockout, 3 biological replicates each). For each sample and each method, the number of 3'NTA Gs and Cs are counted and divided by the total number of all mono-3'NTA (A, U, C, G). All input fastq files had adapters previously trimmed.

### 2.5.9 Novel miRNA prediction

Novel miRNAs were predicted using miRDeep2 [62] by combining reads from all samples of the same fluid (or cell type) in the same study. To obtain high confidence predictions, we imposed the following criteria, similar to previous studies [82]: (1) a miRDeep2 score > 0; (2) ≥ 5 reads mapped to the passenger strand and ≥10 reads to the primary strand. For overlapping predictions, the one with the highest miRDeep2 score was chosen. The final set of putative novel miRNAs were combined with known miRNA annotations (miRBase V22) [83] for subsequent analyses.

### 2.5.10 Gene ontology enrichment analysis

Gene ontology (GO) terms for miRNAs were downloaded from http://geneontology.org/gene-associations/goa_human_rna.gaf.gz [84]. For each data set, miRNAs expressed in ≥ 50% of samples and with an average 3' uridylation level of ≥5% were included for the GO analysis. For each query miRNA, a control miRNA without 3' modifications was chosen randomly that matches the expression level of the query (± 20% relative to the query). GO terms present in the sets of query miRNAs and matched controls were collected, respectively. The process was repeated 10,000 times to construct 10,000 sets of control miRNAs, where each set has the same number of miRNAs as the query set. Query miRNAs that had less than 3 candidate controls were not included in this analysis. For each GO term associated with at least 2 query miRNAs, a Gaussian distribution was fit to the number of control miRNAs also associated with this term to calculate a p value. Significant GO terms were defined as those with FDR < 0.05.

### 2.5.11 miRNA read count normalization across data sets

To obtain miRNA expression levels, DESeq2 (version 1.14.1) [85] was used to normalize miRNA read counts across data sets. miRNAs associated with at least 10 reads in at least 50% of all samples were used to generate the DESeq2 scale factor for normalization.

### 2.5.12 tSNE and PCA clustering

miRNAs expressed with a minimum read count of 10 were included for clustering analysis. tSNE and PCA clustering were performed using the package Rtsne and prcomp, respectively. The expression was set to 0 in samples where a miRNA has no reads. Levels of NTAs or Log2 of the DEseq2 normalized expression was used in these analyses.

### 2.5.13 Differentially modified miRNA between fluids

Differential modification of miRNAs between two fluids was performed using the REDITs method [65]. miRNAs with an effect size ≥ 5% between fluids in expressed samples were included. Significant miRNAs were required to be expressed (read count ≥ 10) in at least 20 samples in both tested fluids with a FDR < 0.05.

### 2.5.14 miRNA target analysis

RNAhybrid (version 2.1.2) [66] with default settings was used to estimate the minimum free energy between miRNAs and putative target 3' UTR sequences. RNAhybrid p value was required to be <0.05 to call a significant minimum free energy binding site. Putative sites that base-paired with different types of 3' NTAs (A, U, C, G) were then examined. GO enrichment of target genes pairing with the 3' U NTAs was performed similarly as described above. For each miRNA, control genes were chosen from those putative target genes without 3' U-pairing and with matched gene length as

the query genes (±10%). Each query gene was required to have at least 10 controls to be included in the analysis.

## 2.6 Acknowledgements

We thank members of the Xiao laboratory for helpful discussions and comments on this work. We acknowledge the data production efforts and the subject donors for making available the data sets used in this study. The Extracellular Small RNA Profiles in Plasma, Urine and Saliva from College Athletes Study (dbGAP phs001258) was made possible by the NIH Common Fund Program on Extracellular RNA Communication Grant UH3 TR000891, Riddell and BRG Sports, the Flinn Foundation grant awards #1994 and #2307, the staff and athletes at Arizona State University, and TGen faculty and interns. Thanks to The Michael J Fox Foundation for supporting the collection of dbGAP phs00727, the subject donors to the program at Sun Health Research Institute. This work was supported in part by grants from the National Institute of Health (U01HG009417, R01AG056476 to X.X) and the Jonsson Comprehensive Cancer Center at UCLA. K.K. was supported by a Eureka Scholarship from the Department of Integrative Biology and Physiology at UCLA.

## 2.7 Figures

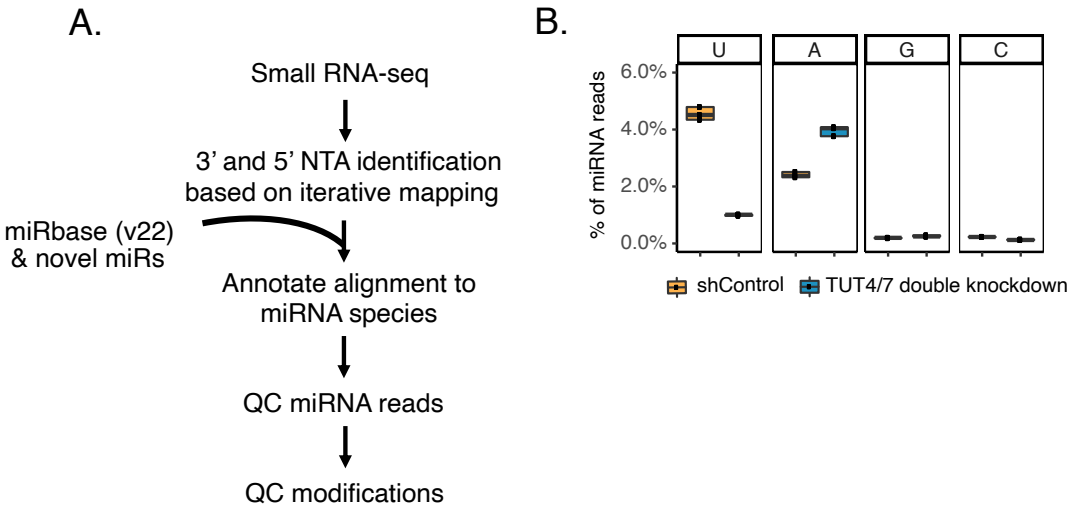



**Figure 2.1 Identification of non-templated additions (NTAs) of miRNAs.**

**A** miNTA, a bioinformatic pipeline to identify miRNA NTAs. Small RNA-seq reads were mapped to hg19 allowing up to one mismatch and retaining uniquely mapped reads only. An iterative mapping approach was applied to unmapped reads by sequentially trimming 1 nucleotide on the 3' and 5' ends before remapping. Each end was trimmed twice. All mapped reads were examined relative to the human genome reference to identify 3' and 5' NTAs. After annotating mapped reads to novel and known miRNAs, reads were passed through several quality control (QC) filters to remove likely false positives from mapping and sequencing errors. See Materials and Methods for more details. **B** Percentage of miRNA reads with 3' end non-templated mono-uridylation (U),

adenylation (A), guanidylation (G) and cytidylation (C) identified by miNTA using small RNA-seq data derived from control (shControl) and TUT4/7 double KD HEK293 cells.

A.

| Study Number | Fluids (N samples) | Reference |
|---|---|---|
| Lab 1 | Plasma (180); Urine (203) | Yeri A. et al. *Sci. Rep* (2017) |
| Lab 2 | Plasma Exosomes (50) | Yuan T. et al. *Sci. Rep* (2016) |
| Lab 3 | Serum (71); Cerebral Spinal Fluid (69) | Burgos K. et al. *PLoS One* (2014) |
| Lab 4 | Plasma Exosomes (113); Urine (46) | Ferrero G et al. *Oncotarget* (2019) |
| Lab 5 | Plasma (219); Serum (96) | Max K. et al. *PNAS* (2018) |
| Lab 6 | Intracellular Blood Cells (297) | Juzenas S. et al. *NAR* (2017) |

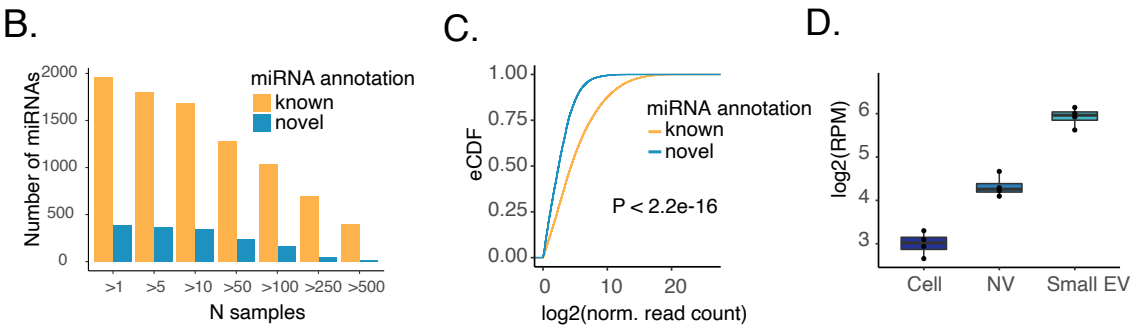



**Figure 2.2 Generation of a comprehensive catalog of extracellular miRNAs.**

**A** Extracellular and intracellular small RNA-seq datasets used in this study. **B** Number of known and novel miRNAs observed in greater than N samples (x axis) across all data sets in A. **C** Empirical cumulative distribution function (eCDF) of the abundance of known or novel miRNAs in all data sets in A. Normalized read counts were calculated using DESeq2 (Methods). P value was calculated via a two-sided Kolmogorov–Smirnov (KS) test. **D** Expression of a novel miRNA (chr7_40460) in whole-cell lysates (Cell), non-vesicle extracellular (NV) and small extracellular vesicle (Small EV) fractions isolated from Gli36 cells.

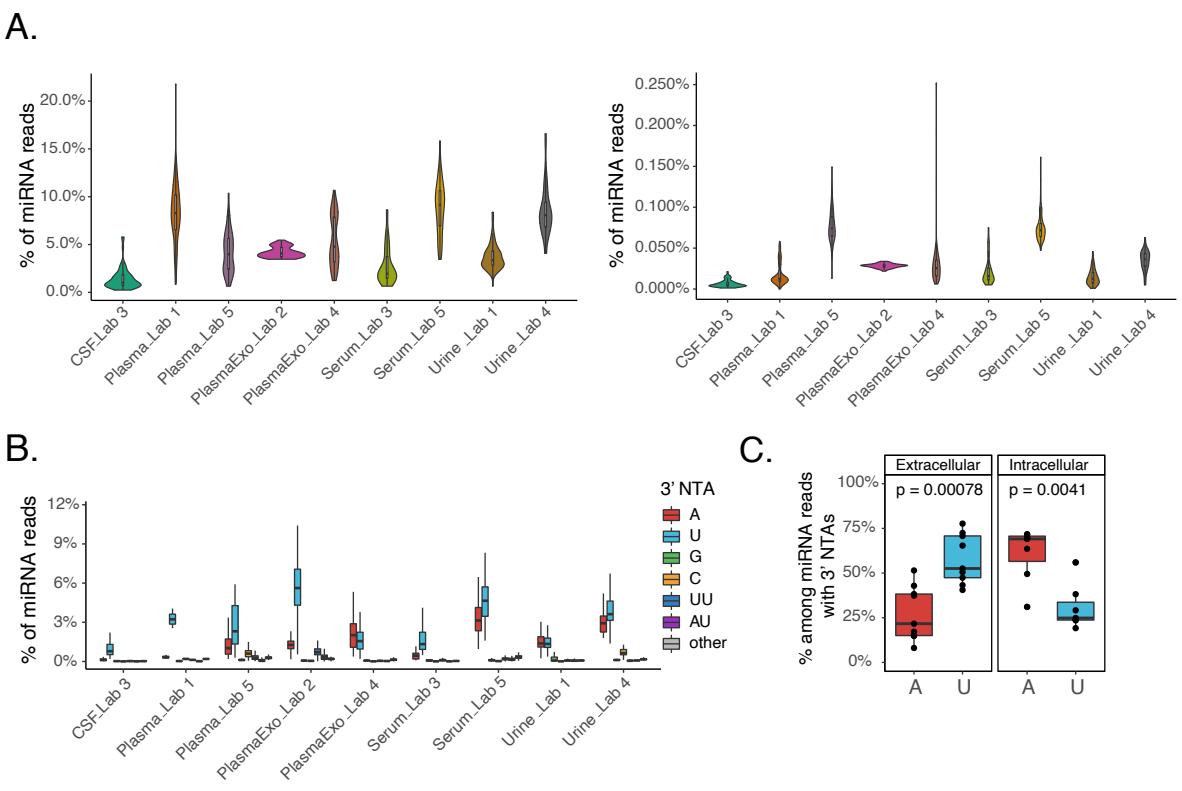



**Figure 2.3 NTA profiles of extracellular miRNAs across biofluids.**

**A** Percentage of miRNA reads with 3' (left) and 5' (right) NTAs across all samples in each data set. **B** Nucleotide composition of 3' NTAs in miRNA reads across all samples in each data set. **C** Average percentage of reads with 3' uridylation or adenylation among all miRNA reads with 3' NTAs. Each dot represents this average value for an extracellular fluid type or intracellular cell type in each study. P values were calculated via Wilcoxon rank-sum test.

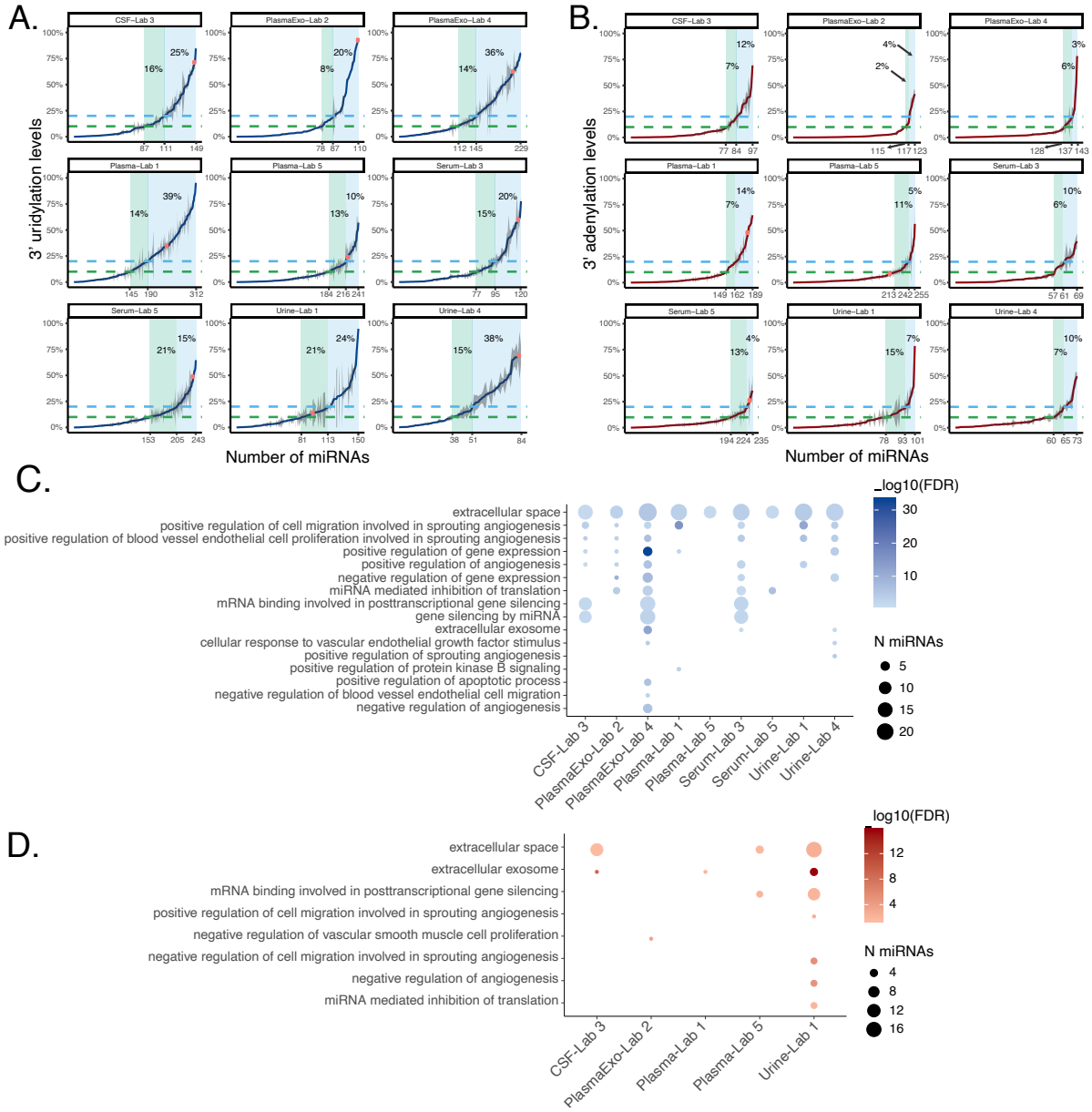



**Figure 2.4 Characteristics of miRNAs with 3' NTAs.**

**A** 3' uridylation levels of miRNA species in each data set. Blue curves represent the average 3' NTA levels and grey shades represent standard errors. The horizontal dashed green and blue lines represent 10% and 20% 3' U thresholds, respectively.

Shaded green and blue regions encompass miRNAs with 3' U levels between 10% - 20% and >20%, respectively, with the percentage of such miRNAs among all 3' U modified miRNAs shown. The miRNA hsa-let-7f-2-3p is highlighted as a red dot. **B** Similar to A but for miRNA 3' adenylation levels. The miRNA hsa-miR-6513-3p is highlighted as a red dot. **C** Gene ontology terms enriched among miRNAs with an average 3' uridylation level ≥ 5% (FDR < 0.05, Methods). The size of the dots reflect the number of miRNAs in each GO term. **D** Similar to C but for miRNA 3' adenylation.

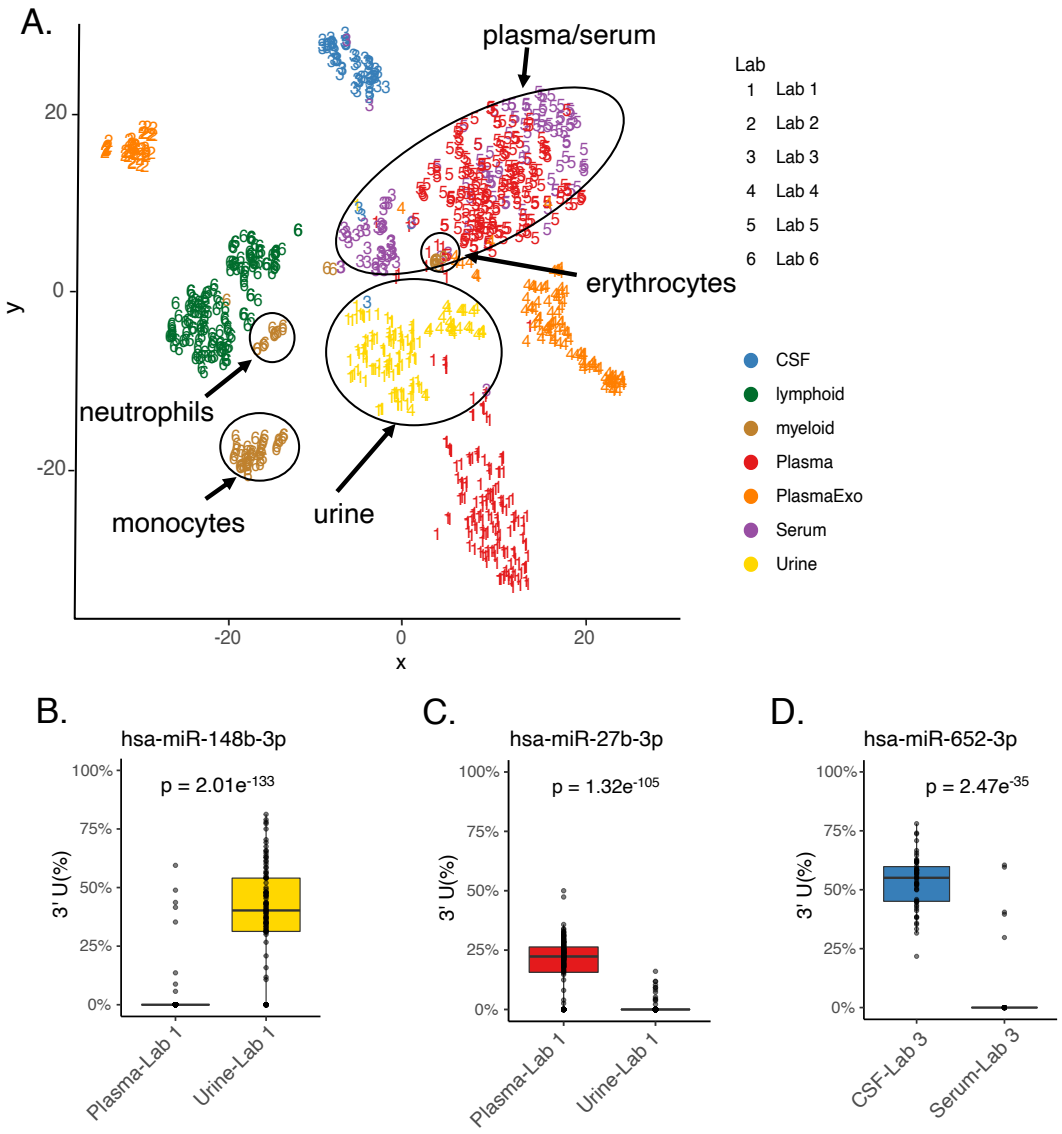



**Figure 2.5 Distinct miRNA 3' uridylation profiles across fluids.**

**A** tSNE clustering of samples using miRNA 3' uridylation levels. miRNAs expressed with a minimum read count of 10 were included for this analysis. **B-D** Example miRNAs observed with differential 3' uridylation levels between fluids of the same study. P values were calculated via REDITs (Methods).

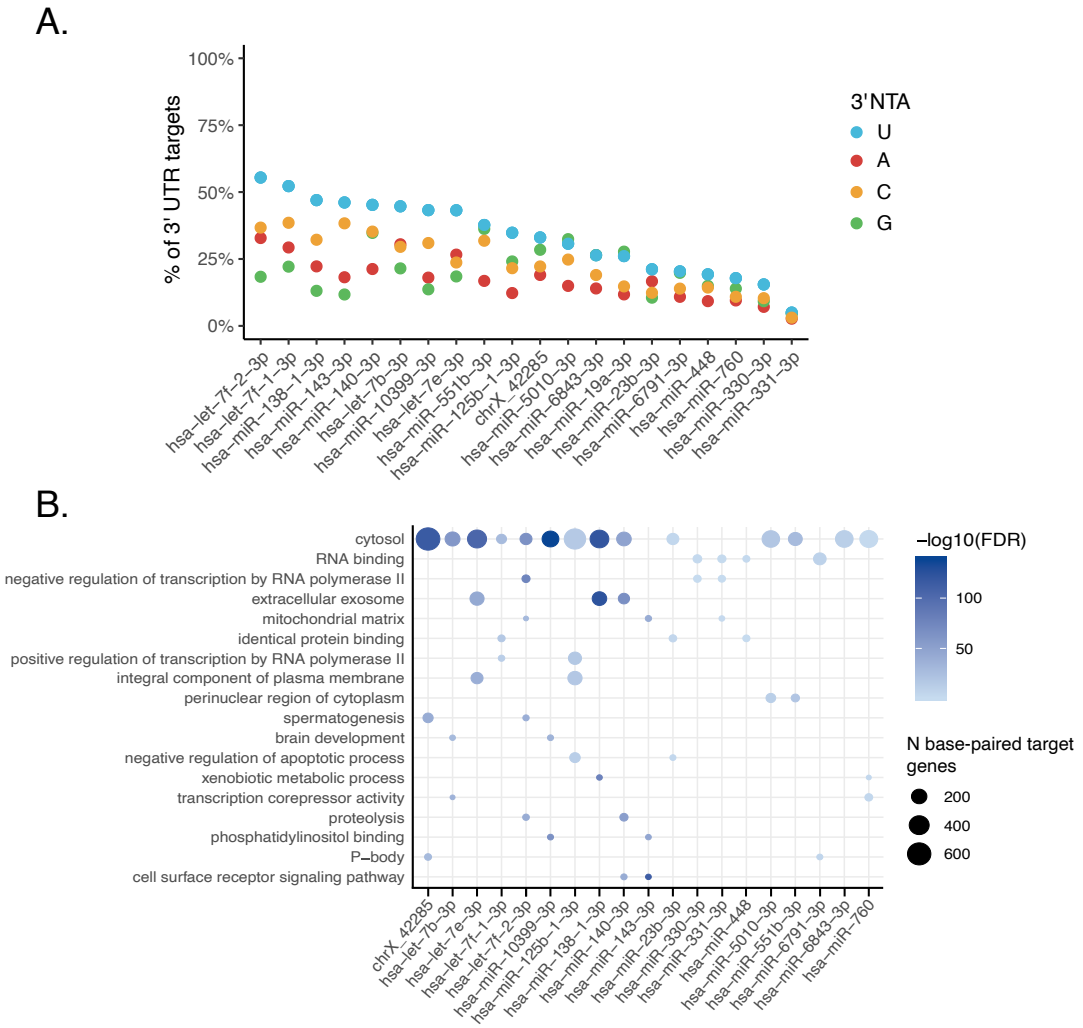
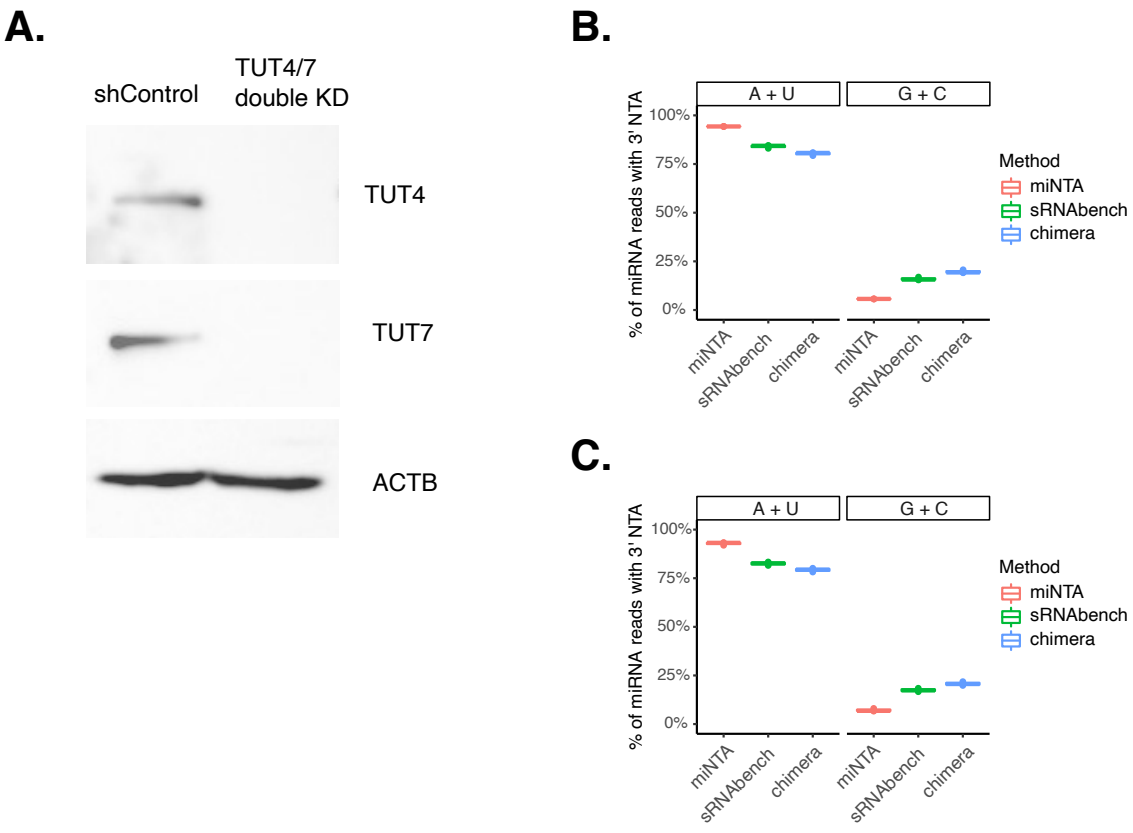



**Figure 2.6 3’ U base-pairs with predicted miRNA targets more often than other 3’ NTAs.**

**A** Percentage of predicted miRNA 3’ UTR targets that base-pair with 3’ U (or A, C, G) for the top 20 unique miRNAs with highest average 3’ uridylation levels (per fluid) across all samples (Methods). For all miRNAs, the number of 3’ Us base-paired with predicted miRNA targets is higher than at least one background nucleotide (A, C or G)

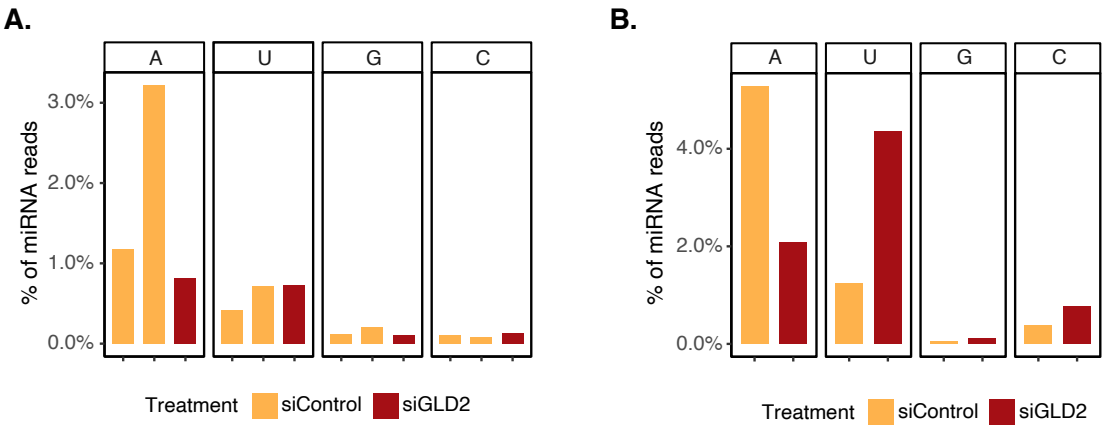
(Chi-squared test $p < 0.05$). **B** For each miRNA, the top 5 enriched GO terms among its target genes that base-paired with the 3' U were collected (Methods). Enriched terms observed in at least 2 data sets are shown (FDR < 0.05). The size of the dots represents the number (N) of base-paired target genes.

## 2.8 Supplementary Figures



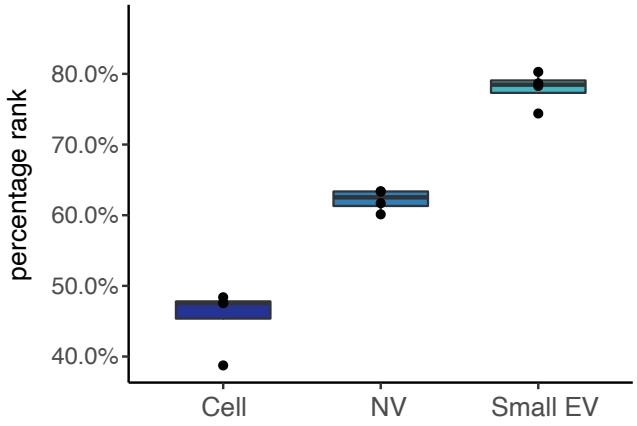

**Supplementary Figure S2.1 TUT4/7 double knockdown to validate miNTA pipeline.**

**A** Western blot of control (shControl) and TUT4/7 double KD in HEK293 cells. **B** Average percentage of reads with 3' uridylation and adenylation (A + U) or guanidinylation and cytidylation (G + C) among all miRNA reads with 3' NTAs. Methods were evaluated using shControl (N = 3) in HEK293 cells. Each dot represents a sample. **C** Similar to B, but in shTUT4 and shTUT7 double KD samples (N = 3) in HEK293 cells.



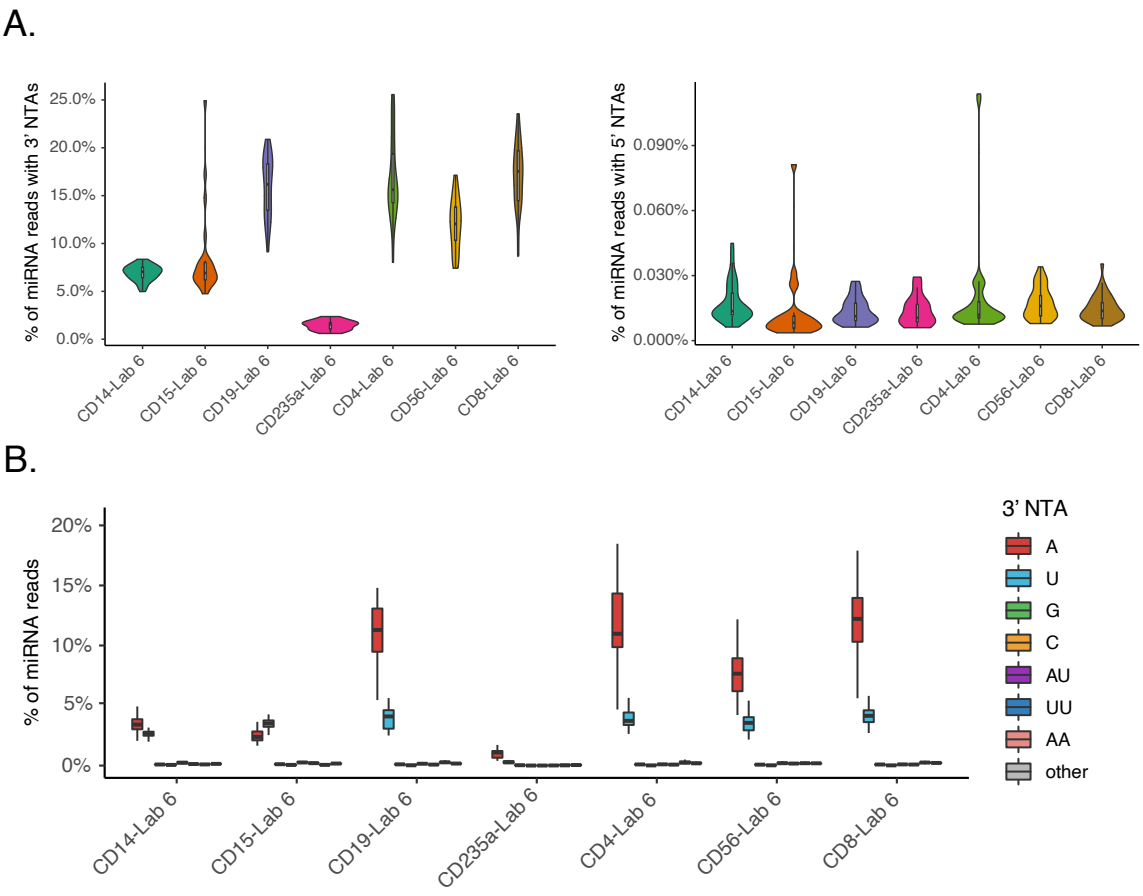

**Supplementary Figure S2.2 GLD2 knockdown to validate miNTA pipeline.**

**A** Western blot of control (shControl) and TUT4/7 double KD in HEK293 cells. **B** Average percentage of reads with 3' uridylation and adenylation (A + U) or guanidinylation and cytidylation (G + C) among all miRNA reads with 3' NTAs. Methods were evaluated using shControl (N = 3) in HEK293 cells. Each dot represents a sample. **C** Similar to B, but in shTUT4 and shTUT7 double KD samples (N = 3) in HEK293 cells.




**Supplementary Figure S2.3 A novel miRNA with higher expression levels in the extracellular space.**

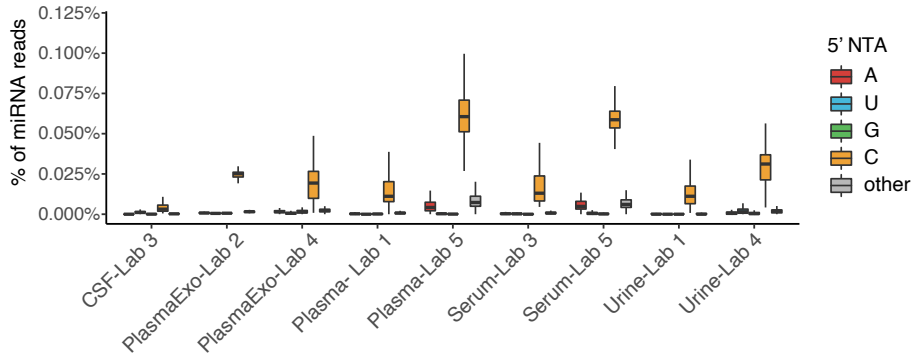
Percentage rank of a novel miRNA (chr7_40460) in whole-cell lysates (Cell), non-vesicle extracellular (NV) and small extracellular vesicle (small EV) fractions isolated from Gli36 cells.



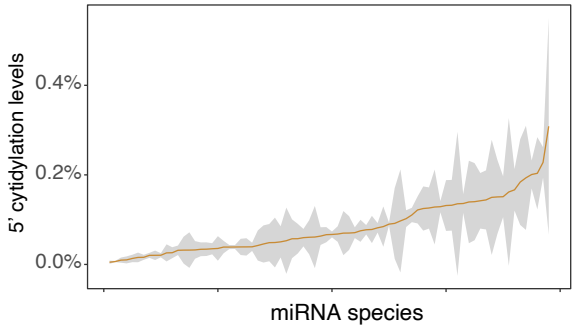

**Supplementary Figure S2.4 NTA profiles of intracellular miRNAs across human peripheral blood cells.**

**A** Percentage of miRNA reads with 3' (left) and 5' (right) NTA's in each dataset. **B** Nucleotide composition of 3' NTAs of miRNA reads across all samples in each intracellular blood-cell type.

A.

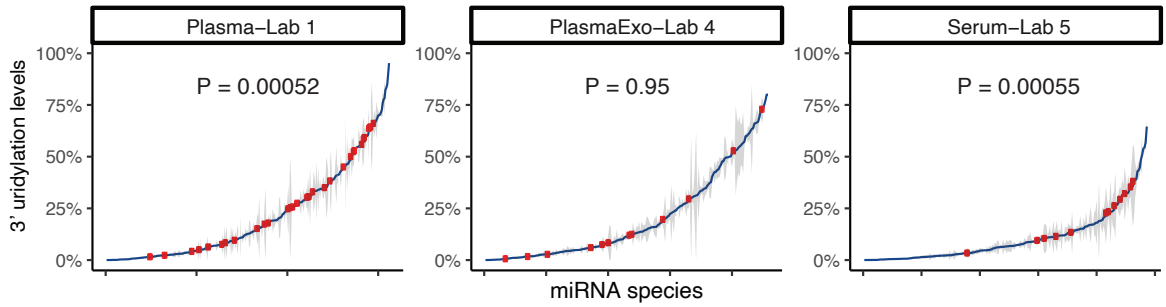



B.



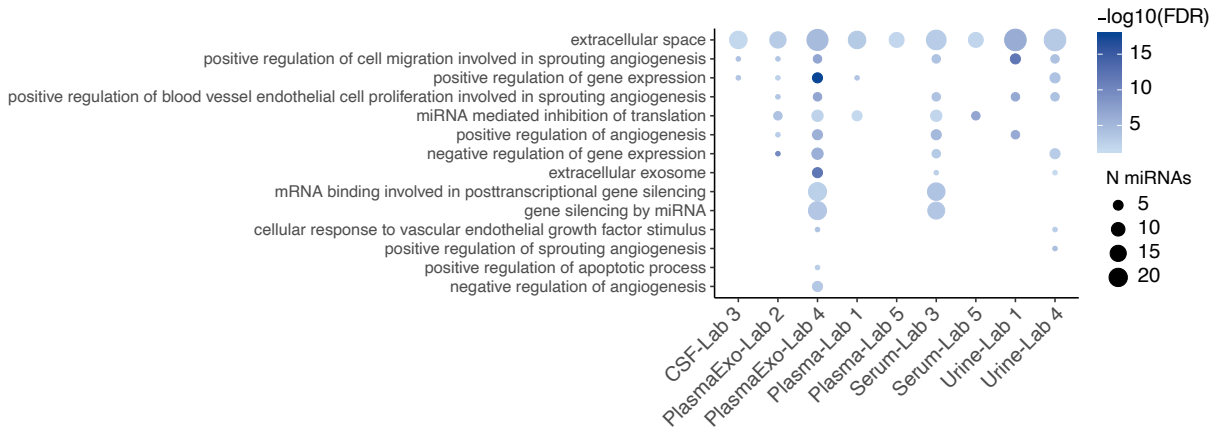

**Supplementary Figure S2.5 miRNA 5' NTAs across biofluids.**

**A** Nucleotide composition of 5' NTAs of miRNA reads across all samples in extracellular fluids from each study. **B** 5' cytidylation levels of miRNAs in Plasma-Lab 5. Orange curve represents the average 5' NTA levels and grey shades represent standard errors.



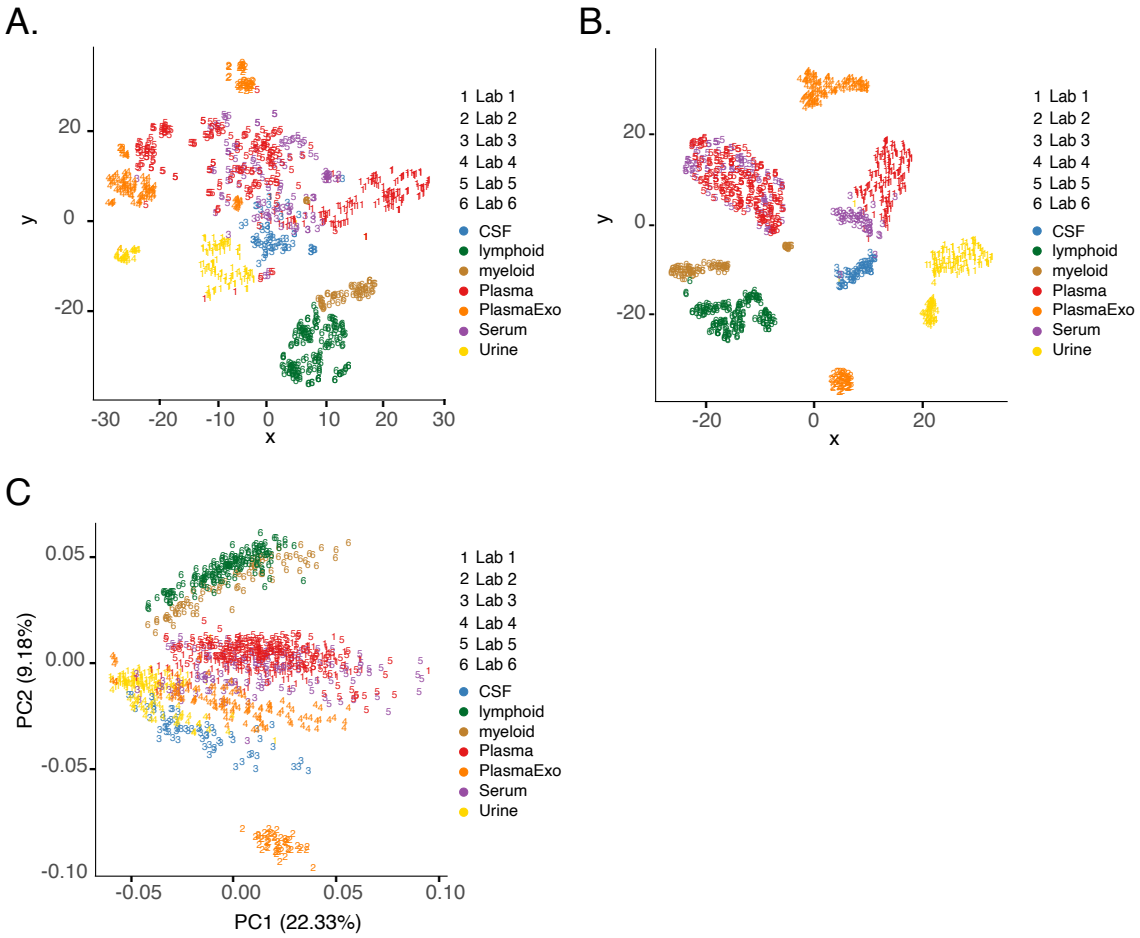

**Supplementary Figure S2.6 Comparison of 3' uridylation levels in novel and known miRNAs.**

3' uridylation levels of miRNAs in each data set with at least 10 novel miRNAs expressed (read count ≥ 10). Blue curves represent the average 3' NTA levels and grey shades represent standard errors. Red dots represent novel miRNAs. P values were calculated via a two-tailed Wilcoxon ranked sum test.




**Supplementary Figure S2.7 Gene ontology enrichment of 3' uridylated miRNAs excluding let-7 family.**

Gene ontology terms enriched among miRNAs, excluding the let-7 family, with an average 3' uridylation level ≥ 5% (FDR <0.05, Methods). The size of the dots reflects the number of miRNAs in each GO term.




**Supplementary Figure S2.8 Clustering samples from biofluids.**

A tSNE clustering of samples using miRNA 3' adenylation levels. miRNAs expressed with a minimum read count of 10 were included for this analysis. B Similar to A, but using normalized miRNA expression values (Methods). C Similar to A, but PCA clustering of samples.

## 2.9 Supplementary Tables

### Supplementary Table S2.1 miRNA 3' uridylation levels across biofluids.

| | miRNA | Fluid– data set | N samples (modified/expressed) | Average 3' U (%) | RPM |
|---|---|---|---|---|---|
| *1* | hsa–let–7f–2–3p | PlasmaExo–Lab 2 | 50/50 | 0.927 | 51.910 |
| *2* | hsa–let–7f–2–3p | CSF–Lab 3 | 19/22 | 0.713 | 11.936 |
| *3* | hsa–let–7f–2–3p | PlasmaExo–Lab 4 | 24/34 | 0.625 | 14.647 |
| *4* | hsa–let–7f–2–3p | Serum–Lab 3 | 10/13 | 0.597 | 14.423 |
| *5* | hsa–let–7f–2–3p | Serum–Lab 5 | 43/60 | 0.487 | 38.055 |
| *6* | hsa–let–7f–2–3p | Plasma–Lab 1 | 40/109 | 0.347 | 31.528 |
| *7* | hsa–let–7f–2–3p | Plasma–Lab 5 | 78/126 | 0.238 | 32.145 |
| *8* | hsa–miR–143–3p | PlasmaExo–Lab 2 | 50/50 | 0.899 | 4118.272 |
| *9* | hsa–miR–143–3p | Serum–Lab 5 | 65/67 | 0.544 | 3806.757 |
| *10* | hsa–miR–143–3p | Plasma–Lab 5 | 164/165 | 0.501 | 1802.856 |
| *11* | hsa–miR–143–3p | PlasmaExo–Lab 4 | 42/42 | 0.437 | 437.102 |
| *12* | hsa–miR–143–3p | Plasma–Lab 1 | 163/163 | 0.263 | 4153.692 |
| *13* | hsa–miR–143–3p | Urine–Lab 1 | 33/56 | 0.186 | 156.184 |
| *14* | hsa–miR–143–3p | CSF–Lab 3 | 51/51 | 0.124 | 5689.880 |
| *15* | hsa–miR–143–3p | Serum–Lab 3 | 42/42 | 0.090 | 8219.917 |
| *16* | hsa–miR–125b–1–3p | PlasmaExo–Lab 2 | 50/50 | 0.862 | 850.898 |
| *17* | hsa–miR–125b–1–3p | Urine–Lab 4 | 9/10 | 0.693 | 15.290 |
| *18* | hsa–miR–125b–1–3p | CSF–Lab 3 | 63/63 | 0.589 | 80.438 |
| *19* | hsa–miR–125b–1–3p | Serum–Lab 3 | 33/35 | 0.514 | 50.771 |
| *20* | hsa–miR–125b–1–3p | Urine–Lab 1 | 6/7 | 0.426 | 74.729 |
| *21* | hsa–miR–125b–1–3p | Plasma–Lab 5 | 15/19 | 0.385 | 11.137 |
| *22* | hsa–miR–125b–1–3p | Serum–Lab 5 | 9/12 | 0.282 | 16.000 |
| *23* | hsa–miR–125b–1–3p | Plasma–Lab 1 | 8/10 | 0.279 | 45.060 |
| *24* | hsa–miR–448 | CSF–Lab 3 | 59/60 | 0.848 | 119.525 |
| *25* | hsa–miR–760 | PlasmaExo–Lab 2 | 50/50 | 0.819 | 4788.726 |
| *26* | hsa–miR–760 | PlasmaExo–Lab 4 | 72/72 | 0.688 | 54.885 |
| *27* | hsa–miR–760 | Urine–Lab 4 | 5/6 | 0.674 | 11.583 |
| *28* | hsa–miR–760 | Plasma–Lab 1 | 98/109 | 0.667 | 21.832 |
| *29* | hsa–miR–760 | CSF–Lab 3 | 40/40 | 0.623 | 17.785 |
| *30* | hsa–miR–760 | Serum–Lab 3 | 13/13 | 0.513 | 13.600 |
| *31* | hsa–miR–760 | Serum–Lab 5 | 12/68 | 0.093 | 30.260 |
| *32* | hsa–miR–760 | Plasma–Lab 5 | 14/139 | 0.057 | 24.361 |
| *33* | hsa–miR–10399–3p | Plasma–Lab 1 | 130/136 | 0.777 | 51.312 |
| *34* | hsa–miR–10399–3p | Serum–Lab 3 | 22/23 | 0.626 | 18.774 |
| *35* | hsa–miR–10399–3p | PlasmaExo–Lab 4 | 61/62 | 0.477 | 48.842 |
| *36* | hsa–miR–10399–3p | Urine–Lab 4 | 10/10 | 0.457 | 8.540 |
| *37* | hsa–miR–10399–3p | Plasma–Lab 5 | 88/96 | 0.343 | 13.855 |

| | | | | | |
|---|---|---|---|---|---|
| *38* | hsa–miR–10399–3p | Serum–Lab 5 | 43/54 | 0.269 | 25.526 |
| *39* | hsa–let–7f–1–3p | Serum–Lab 3 | 7/7 | 0.776 | 14.443 |
| *40* | hsa–let–7f–1–3p | PlasmaExo–Lab 2 | 50/50 | 0.710 | 32.822 |
| *41* | hsa–let–7f–1–3p | CSF–Lab 3 | 21/23 | 0.701 | 7.978 |
| *42* | hsa–let–7f–1–3p | Plasma–Lab 1 | 39/46 | 0.687 | 11.741 |
| *43* | hsa–let–7f–1–3p | PlasmaExo–Lab 4 | 37/38 | 0.662 | 12.137 |
| *44* | hsa–let–7f–1–3p | Serum–Lab 5 | 26/33 | 0.390 | 17.067 |
| *45* | hsa–let–7f–1–3p | Plasma–Lab 5 | 63/79 | 0.338 | 14.306 |
| *46* | chrX_42285 | PlasmaExo–Lab 4 | 19/20 | 0.729 | 9.670 |
| *47* | chrX_42285 | Plasma–Lab 1 | 15/17 | 0.661 | 6.182 |
| *48* | hsa–miR–6791–3p | Plasma–Lab 1 | 25/26 | 0.725 | 6.438 |
| *49* | hsa–miR–331–3p | PlasmaExo–Lab 4 | 29/36 | 0.725 | 20.911 |
| *50* | hsa–miR–331–3p | PlasmaExo–Lab 2 | 42/49 | 0.723 | 10.090 |
| *51* | hsa–miR–331–3p | Serum–Lab 5 | 60/68 | 0.429 | 107.879 |
| *52* | hsa–miR–331–3p | Plasma–Lab 5 | 127/159 | 0.310 | 52.260 |
| *53* | hsa–miR–23b–3p | PlasmaExo–Lab 4 | 66/66 | 0.717 | 120.715 |
| *54* | hsa–miR–23b–3p | Plasma–Lab 1 | 138/146 | 0.712 | 313.392 |
| *55* | hsa–miR–23b–3p | Urine–Lab 1 | 24/38 | 0.449 | 37.311 |
| *56* | hsa–miR–23b–3p | Serum–Lab 5 | 27/28 | 0.447 | 571.996 |
| *57* | hsa–miR–23b–3p | Plasma–Lab 5 | 10/42 | 0.143 | 156.221 |
| *58* | hsa–miR–140–3p | Plasma–Lab 1 | 149/150 | 0.703 | 2966.460 |
| *59* | hsa–miR–140–3p | Serum–Lab 3 | 43/43 | 0.569 | 1585.847 |
| *60* | hsa–miR–140–3p | CSF–Lab 3 | 54/54 | 0.480 | 437.611 |
| *61* | hsa–miR–140–3p | Urine–Lab 1 | 102/104 | 0.368 | 63.066 |
| *62* | hsa–miR–140–3p | PlasmaExo–Lab 4 | 93/93 | 0.265 | 1747.963 |
| *63* | hsa–miR–330–3p | Plasma–Lab 1 | 127/134 | 0.701 | 74.216 |
| *64* | hsa–miR–330–3p | PlasmaExo–Lab 4 | 82/84 | 0.659 | 181.481 |
| *65* | hsa–miR–330–3p | Urine–Lab 4 | 12/14 | 0.653 | 11.636 |
| *66* | hsa–miR–330–3p | Serum–Lab 5 | 17/47 | 0.118 | 36.102 |
| *67* | hsa–miR–330–3p | Plasma–Lab 5 | 15/73 | 0.057 | 12.799 |
| *68* | hsa–let–7b–3p | Plasma–Lab 1 | 97/104 | 0.700 | 23.332 |
| *69* | hsa–let–7b–3p | Urine–Lab 4 | 8/9 | 0.672 | 8.922 |
| *70* | hsa–let–7b–3p | CSF–Lab 3 | 40/42 | 0.669 | 25.167 |
| *71* | hsa–let–7b–3p | Urine–Lab 1 | 15/21 | 0.570 | 18.367 |
| *72* | hsa–let–7b–3p | PlasmaExo–Lab 4 | 81/85 | 0.501 | 22.120 |
| *73* | hsa–let–7b–3p | Serum–Lab 3 | 26/27 | 0.494 | 32.096 |
| *74* | hsa–let–7b–3p | Serum–Lab 5 | 63/76 | 0.322 | 93.024 |
| *75* | hsa–let–7b–3p | Plasma–Lab 5 | 133/162 | 0.286 | 76.267 |
| *76* | hsa–miR–19a–3p | Urine–Lab 4 | 8/8 | 0.677 | 10.387 |

| | | | | | |
|---|---|---|---|---|---|
| *77* | hsa–miR–19a–3p | PlasmaExo–Lab 2 | 50/50 | 0.622 | 29.564 |
| *78* | hsa–miR–19a–3p | PlasmaExo–Lab 4 | 92/92 | 0.477 | 48.039 |
| *79* | hsa–miR–19a–3p | Urine–Lab 1 | 7/20 | 0.100 | 11.905 |
| *80* | hsa–miR–19a–3p | CSF–Lab 3 | 21/34 | 0.079 | 28.644 |
| *81* | hsa–miR–19a–3p | Serum–Lab 3 | 28/40 | 0.058 | 225.067 |
| *82* | hsa–miR–551b–3p | Plasma–Lab 1 | 8/9 | 0.676 | 4.844 |
| *83* | hsa–miR–551b–3p | CSF–Lab 3 | 12/12 | 0.494 | 9.942 |
| *84* | hsa–miR–551b–3p | Plasma–Lab 5 | 7/12 | 0.260 | 7.375 |
| *85* | hsa–miR–551b–3p | Serum–Lab 5 | 6/13 | 0.238 | 9.823 |
| *86* | hsa–miR–5010–3p | PlasmaExo–Lab 4 | 21/23 | 0.668 | 11.113 |
| *87* | hsa–miR–5010–3p | Plasma–Lab 1 | 36/45 | 0.599 | 9.033 |
| *88* | hsa–miR–5010–3p | Serum–Lab 5 | 10/14 | 0.507 | 8.407 |
| *89* | hsa–miR–5010–3p | Plasma–Lab 5 | 11/29 | 0.140 | 6.445 |
| *90* | hsa–let–7e–3p | Plasma–Lab 1 | 43/43 | 0.652 | 11.153 |
| *91* | hsa–let–7e–3p | CSF–Lab 3 | 8/11 | 0.146 | 7.800 |
| *92* | hsa–miR–6843–3p | Plasma–Lab 1 | 15/19 | 0.651 | 6.126 |
| *93* | hsa–miR–6843–3p | PlasmaExo–Lab 4 | 33/34 | 0.614 | 18.659 |
| *94* | hsa–miR–138–1–3p | PlasmaExo–Lab 2 | 50/50 | 0.646 | 53.416 |
| *95* | hsa–miR–4669 | PlasmaExo–Lab 4 | 24/24 | 0.636 | 21.008 |
| *96* | hsa–miR–4669 | Plasma–Lab 5 | 8/10 | 0.341 | 6.910 |
| *97* | hsa–miR–6798–3p | Plasma–Lab 1 | 9/9 | 0.635 | 6.700 |
| *98* | hsa–miR–6741–5p | PlasmaExo–Lab 4 | 30/31 | 0.607 | 8.806 |
| *99* | hsa–miR–652–5p | Plasma–Lab 1 | 6/7 | 0.606 | 8.171 |
| *100* | hsa–miR–432–5p | PlasmaExo–Lab 4 | 88/90 | 0.600 | 148.974 |
| *101* | hsa–miR–432–5p | Plasma–Lab 1 | 51/56 | 0.555 | 176.345 |
| *102* | hsa–miR–432–5p | Serum–Lab 5 | 37/45 | 0.434 | 48.616 |
| *103* | hsa–miR–432–5p | Plasma–Lab 5 | 27/52 | 0.304 | 20.944 |
| *104* | hsa–miR–432–5p | CSF–Lab 3 | 6/12 | 0.139 | 20.050 |
| *105* | hsa–miR–26b–3p | PlasmaExo–Lab 4 | 44/44 | 0.599 | 38.168 |
| *106* | hsa–miR–26b–3p | PlasmaExo–Lab 2 | 12/12 | 0.557 | 5.842 |
| *107* | hsa–miR–26b–3p | Serum–Lab 5 | 32/58 | 0.234 | 54.817 |
| *108* | hsa–miR–26b–3p | Plasma–Lab 1 | 117/129 | 0.221 | 64.847 |
| *109* | hsa–miR–26b–3p | CSF–Lab 3 | 31/35 | 0.182 | 12.183 |
| *110* | hsa–miR–26b–3p | Serum–Lab 3 | 9/18 | 0.089 | 28.522 |
| *111* | hsa–miR–676–3p | CSF–Lab 3 | 7/7 | 0.596 | 5.157 |
| *112* | hsa–miR–676–3p | PlasmaExo–Lab 2 | 50/50 | 0.546 | 662.862 |
| *113* | chr9_45534 | Plasma–Lab 1 | 19/20 | 0.593 | 9.235 |
| *114* | hsa–miR–125b–2–3p | CSF–Lab 3 | 10/11 | 0.591 | 392.700 |

| | | | | | |
|---|---|---|---|---|---|
| *115* | hsa–miR–125b–2–3p | Serum–Lab 3 | 29/30 | 0.586 | 252.207 |
| *116* | hsa–miR–125b–2–3p | PlasmaExo–Lab 4 | 20/22 | 0.497 | 8.223 |
| *117* | hsa–miR–125b–2–3p | Plasma–Lab 1 | 80/83 | 0.486 | 21.630 |
| *118* | hsa–miR–125b–2–3p | Urine–Lab 1 | 62/73 | 0.209 | 54.979 |
| *119* | hsa–miR–125b–2–3p | PlasmaExo–Lab 2 | 50/50 | 0.200 | 1959.074 |
| *120* | hsa–miR–125b–2–3p | Plasma–Lab 5 | 22/45 | 0.175 | 39.473 |
| *121* | hsa–miR–125b–2–3p | Serum–Lab 5 | 8/18 | 0.121 | 60.400 |
| *122* | chr11_4033 | Plasma–Lab 1 | 10/10 | 0.586 | 6.090 |
| *123* | hsa–miR–1301–3p | PlasmaExo–Lab 4 | 51/54 | 0.572 | 243.698 |
| *124* | hsa–miR–1301–3p | Plasma–Lab 1 | 119/124 | 0.558 | 108.293 |
| *125* | hsa–miR–1301–3p | Serum–Lab 5 | 26/33 | 0.338 | 62.015 |
| *126* | hsa–miR–1301–3p | Plasma–Lab 5 | 7/38 | 0.074 | 31.142 |
| *127* | hsa–miR–652–3p | Serum–Lab 5 | 61/70 | 0.567 | 604.350 |
| *128* | hsa–miR–652–3p | CSF–Lab 3 | 58/58 | 0.532 | 37.293 |
| *129* | hsa–miR–652–3p | Urine–Lab 1 | 13/16 | 0.501 | 20.644 |
| *130* | hsa–miR–652–3p | Urine–Lab 4 | 17/17 | 0.434 | 17.294 |
| *131* | hsa–miR–652–3p | PlasmaExo–Lab 2 | 39/39 | 0.390 | 154.736 |
| *132* | hsa–miR–652–3p | Plasma–Lab 1 | 90/134 | 0.383 | 393.093 |
| *133* | hsa–miR–652–3p | Plasma–Lab 5 | 124/157 | 0.343 | 275.629 |
| *134* | hsa–miR–652–3p | PlasmaExo–Lab 4 | 40/63 | 0.279 | 166.114 |
| *135* | hsa–miR–652–3p | Serum–Lab 3 | 5/32 | 0.072 | 350.800 |
| *136* | hsa–miR–4738–3p | PlasmaExo–Lab 4 | 5/7 | 0.567 | 2.829 |
| *137* | hsa–miR–3135a | PlasmaExo–Lab 4 | 7/7 | 0.560 | 5.200 |
| *138* | hsa–miR–3620–3p | Plasma–Lab 1 | 41/41 | 0.557 | 10.366 |
| *139* | hsa–miR–3620–3p | Serum–Lab 5 | 8/10 | 0.388 | 6.740 |
| *140* | hsa–miR–149–5p | Urine–Lab 1 | 65/68 | 0.556 | 77.794 |
| *141* | hsa–miR–149–5p | Urine–Lab 4 | 12/12 | 0.398 | 25.683 |
| *142* | hsa–miR–149–5p | CSF–Lab 3 | 34/43 | 0.100 | 49.758 |
| *143* | hsa–miR–6515–3p | PlasmaExo–Lab 4 | 6/6 | 0.553 | 4.083 |
| *144* | hsa–miR–6783–3p | Plasma–Lab 1 | 5/5 | 0.549 | 3.380 |
| *145* | hsa–miR–3661 | Plasma–Lab 1 | 13/16 | 0.548 | 7.100 |
| *146* | hsa–miR–4750–5p | PlasmaExo–Lab 4 | 5/7 | 0.548 | 5.214 |
| *147* | hsa–miR–4750–5p | Plasma–Lab 1 | 5/7 | 0.465 | 7.114 |
| *148* | hsa–miR–758–3p | Plasma–Lab 1 | 75/98 | 0.545 | 35.419 |
| *149* | hsa–miR–758–3p | Serum–Lab 5 | 30/43 | 0.320 | 18.693 |
| *150* | hsa–miR–758–3p | Plasma–Lab 5 | 8/11 | 0.291 | 14.473 |
| *151* | hsa–miR–758–3p | CSF–Lab 3 | 5/7 | 0.261 | 4.900 |
| *152* | hsa–miR–758–3p | PlasmaExo–Lab 4 | 24/29 | 0.178 | 20.755 |
| *153* | hsa–miR–200b–3p | Urine–Lab 1 | 110/110 | 0.531 | 609.075 |

| | | | | | |
|---|---|---|---|---|---|
| *154* | hsa–miR–200b–3p | Serum–Lab 3 | 7/7 | 0.439 | 66.371 |
| *155* | hsa–miR–200b–3p | Plasma–Lab 1 | 78/80 | 0.396 | 21.259 |
| *156* | hsa–miR–200b–3p | CSF–Lab 3 | 6/7 | 0.365 | 4.614 |
| *157* | hsa–miR–200b–3p | PlasmaExo–Lab 4 | 55/57 | 0.335 | 37.091 |
| *158* | hsa–miR–200b–3p | Urine–Lab 4 | 27/27 | 0.325 | 8790.785 |
| *159* | hsa–miR–4533 | PlasmaExo–Lab 4 | 22/23 | 0.526 | 8.965 |
| *160* | chr19_16571 | Plasma–Lab 1 | 18/18 | 0.524 | 9.750 |
| *161* | chr19_16571 | Serum–Lab 5 | 12/16 | 0.294 | 6.131 |
| *162* | hsa–miR–6777–3p | Plasma–Lab 1 | 8/9 | 0.522 | 10.456 |
| *163* | hsa–miR–6777–3p | Plasma–Lab 5 | 5/6 | 0.491 | 7.100 |
| *164* | hsa–miR–6772–3p | PlasmaExo–Lab 4 | 9/10 | 0.519 | 8.500 |
| *165* | hsa–miR–6772–3p | Plasma–Lab 1 | 90/97 | 0.262 | 26.203 |
| *166* | hsa–miR–221–3p | Urine–Lab 4 | 22/23 | 0.508 | 134.043 |
| *167* | hsa–miR–221–3p | PlasmaExo–Lab 4 | 52/55 | 0.496 | 4188.125 |
| *168* | hsa–miR–221–3p | Plasma–Lab 1 | 142/145 | 0.439 | 5087.316 |
| *169* | hsa–miR–221–3p | Urine–Lab 1 | 27/40 | 0.189 | 139.795 |
| *170* | hsa–miR–221–3p | Serum–Lab 5 | 9/12 | 0.129 | 12659.617 |
| *171* | hsa–miR–221–3p | Plasma–Lab 5 | 10/17 | 0.050 | 8210.259 |
| *172* | hsa–miR–1911–5p | CSF–Lab 3 | 41/42 | 0.491 | 297.374 |
| *173* | hsa–miR–3190–3p | PlasmaExo–Lab 4 | 7/10 | 0.490 | 6.590 |
| *174* | hsa–miR–421 | CSF–Lab 3 | 46/48 | 0.487 | 46.431 |
| *175* | hsa–miR–421 | Urine–Lab 1 | 30/46 | 0.349 | 26.254 |
| *176* | hsa–miR–421 | Serum–Lab 3 | 34/36 | 0.280 | 84.264 |
| *177* | hsa–miR–421 | PlasmaExo–Lab 4 | 30/60 | 0.127 | 20.763 |
| *178* | hsa–miR–421 | Plasma–Lab 1 | 33/106 | 0.126 | 209.107 |
| *179* | hsa–miR–421 | Serum–Lab 5 | 14/40 | 0.069 | 36.167 |
| *180* | hsa–miR–107 | Plasma–Lab 1 | 155/159 | 0.483 | 346.208 |
| *181* | hsa–miR–107 | PlasmaExo–Lab 4 | 32/32 | 0.417 | 28.328 |
| *182* | hsa–miR–107 | PlasmaExo–Lab 2 | 49/50 | 0.271 | 13.358 |
| *183* | hsa–miR–107 | CSF–Lab 3 | 54/55 | 0.263 | 49.175 |
| *184* | hsa–miR–107 | Urine–Lab 1 | 64/77 | 0.255 | 27.512 |
| *185* | hsa–miR–107 | Serum–Lab 3 | 20/26 | 0.220 | 37.423 |
| *186* | hsa–miR–107 | Serum–Lab 5 | 46/65 | 0.090 | 113.688 |
| *187* | hsa–miR–107 | Plasma–Lab 5 | 31/88 | 0.082 | 20.427 |
| *188* | hsa–miR–2110 | Urine–Lab 4 | 19/19 | 0.482 | 44.168 |
| *189* | hsa–miR–2110 | Plasma–Lab 1 | 13/24 | 0.333 | 32.892 |
| *190* | hsa–miR–3174 | Plasma–Lab 1 | 6/6 | 0.476 | 6.617 |
| *191* | hsa–miR–452–5p | CSF–Lab 3 | 10/12 | 0.470 | 11.700 |
| *192* | hsa–miR–452–5p | Serum–Lab 3 | 10/11 | 0.428 | 21.645 |

| | | | | | |
|---|---|---|---|---|---|
| *193* | hsa–miR–452–5p | PlasmaExo–Lab 4 | 26/30 | 0.239 | 22.290 |
| *194* | hsa–miR–196b–5p | Plasma–Lab 1 | 112/115 | 0.465 | 29.764 |
| *195* | hsa–miR–196b–5p | Urine–Lab 4 | 12/13 | 0.427 | 25.077 |
| *196* | hsa–miR–196b–5p | CSF–Lab 3 | 6/6 | 0.406 | 61.117 |
| *197* | hsa–miR–196b–5p | Urine–Lab 1 | 88/93 | 0.377 | 43.954 |
| *198* | hsa–miR–196b–5p | PlasmaExo–Lab 4 | 9/17 | 0.277 | 12.694 |
| *199* | hsa–miR–6819–3p | Plasma–Lab 1 | 125/125 | 0.462 | 45.350 |
| *200* | hsa–miR–6819–3p | Serum–Lab 5 | 14/20 | 0.281 | 7.805 |
| *201* | hsa–miR–199b–5p | PlasmaExo–Lab 4 | 30/32 | 0.460 | 14.128 |
| *202* | hsa–miR–199b–5p | Plasma–Lab 1 | 98/108 | 0.350 | 25.071 |
| *203* | hsa–miR–199b–5p | Serum–Lab 3 | 9/12 | 0.205 | 14.083 |
| *204* | hsa–miR–199b–5p | CSF–Lab 3 | 5/8 | 0.086 | 9.963 |
| *205* | hsa–miR–223–3p | Serum–Lab 3 | 29/30 | 0.453 | 854.660 |
| *206* | hsa–miR–223–3p | CSF–Lab 3 | 10/13 | 0.308 | 31.531 |
| *207* | hsa–miR–223–3p | Urine–Lab 4 | 6/6 | 0.288 | 73.933 |
| *208* | hsa–miR–223–3p | PlasmaExo–Lab 2 | 33/33 | 0.238 | 38.197 |
| *209* | hsa–miR–223–3p | Serum–Lab 5 | 54/54 | 0.189 | 24087.187 |
| *210* | hsa–miR–223–3p | Plasma–Lab 5 | 100/104 | 0.178 | 3054.075 |
| *211* | hsa–miR–223–3p | PlasmaExo–Lab 4 | 38/43 | 0.156 | 102.437 |
| *212* | hsa–miR–3188 | PlasmaExo–Lab 4 | 6/6 | 0.435 | 7.800 |
| *213* | hsa–miR–3188 | Plasma–Lab 1 | 11/12 | 0.412 | 7.167 |
| *214* | hsa–miR–342–3p | Plasma–Lab 1 | 101/107 | 0.435 | 130.684 |
| *215* | hsa–miR–342–3p | PlasmaExo–Lab 2 | 50/50 | 0.159 | 33.774 |
| *216* | hsa–miR–342–3p | PlasmaExo–Lab 4 | 28/46 | 0.113 | 25.602 |
| *217* | hsa–miR–491–5p | Plasma–Lab 1 | 13/22 | 0.431 | 13.286 |
| *218* | hsa–miR–491–5p | Serum–Lab 5 | 9/20 | 0.152 | 12.070 |
| *219* | hsa–miR–491–5p | CSF–Lab 3 | 6/8 | 0.084 | 12.988 |
| *220* | hsa–miR–187–5p | PlasmaExo–Lab 2 | 43/43 | 0.427 | 9.519 |
| *221* | hsa–miR–6842–3p | PlasmaExo–Lab 4 | 52/60 | 0.427 | 116.128 |
| *222* | hsa–miR–6842–3p | Plasma–Lab 1 | 151/156 | 0.341 | 114.016 |
| *223* | hsa–miR–6842–3p | CSF–Lab 3 | 12/15 | 0.233 | 6.373 |
| *224* | hsa–miR–6842–3p | Serum–Lab 5 | 17/29 | 0.206 | 11.110 |
| *225* | hsa–miR–6842–3p | Serum–Lab 3 | 13/16 | 0.197 | 30.200 |
| *226* | hsa–miR–6842–3p | Plasma–Lab 5 | 11/28 | 0.145 | 9.321 |
| *227* | hsa–miR–628–3p | PlasmaExo–Lab 4 | 23/32 | 0.426 | 31.378 |
| *228* | hsa–miR–628–3p | Plasma–Lab 1 | 87/101 | 0.159 | 33.787 |
| *229* | hsa–miR–132–3p | Urine–Lab 4 | 6/8 | 0.424 | 7.612 |
| *230* | hsa–miR–132–3p | PlasmaExo–Lab 4 | 23/25 | 0.404 | 11.404 |

| | | | | | |
|---|---|---|---|---|---|
| *231* | hsa–miR–132–3p | Urine–Lab 1 | 18/22 | 0.337 | 50.064 |
| *232* | hsa–miR–132–3p | Serum–Lab 3 | 17/20 | 0.297 | 29.115 |
| *233* | hsa–miR–132–3p | Plasma–Lab 1 | 107/120 | 0.275 | 35.443 |
| *234* | hsa–miR–132–3p | CSF–Lab 3 | 51/54 | 0.241 | 43.170 |
| *235* | hsa–miR–132–3p | Serum–Lab 5 | 54/68 | 0.091 | 168.004 |
| *236* | hsa–miR–132–3p | PlasmaExo–Lab 2 | 50/50 | 0.065 | 2427.680 |
| *237* | hsa–miR–132–3p | Plasma–Lab 5 | 115/169 | 0.062 | 104.042 |
| *238* | hsa–miR–3928–3p | Plasma–Lab 1 | 33/38 | 0.410 | 38.668 |
| *239* | hsa–miR–548l | Plasma–Lab 1 | 8/10 | 0.405 | 9.470 |
| *240* | hsa–miR–488–3p | CSF–Lab 3 | 27/27 | 0.404 | 17.289 |
| *241* | hsa–miR–944 | Plasma–Lab 5 | 5/6 | 0.400 | 4.700 |
| *242* | hsa–miR–425–3p | Plasma–Lab 1 | 147/150 | 0.389 | 133.875 |
| *243* | hsa–miR–425–3p | Serum–Lab 3 | 31/33 | 0.282 | 39.609 |
| *244* | hsa–miR–425–3p | CSF–Lab 3 | 18/21 | 0.235 | 16.114 |
| *245* | hsa–miR–425–3p | PlasmaExo–Lab 4 | 5/10 | 0.052 | 40.350 |
| *246* | hsa–miR–1227–3p | Plasma–Lab 5 | 7/9 | 0.389 | 5.889 |
| *247* | hsa–miR–148b–3p | Urine–Lab 1 | 74/89 | 0.386 | 44.230 |
| *248* | hsa–miR–148b–3p | Urine–Lab 4 | 27/27 | 0.197 | 407.604 |
| *249* | hsa–miR–148b–3p | Serum–Lab 5 | 53/64 | 0.171 | 1734.736 |
| *250* | hsa–miR–148b–3p | CSF–Lab 3 | 59/62 | 0.119 | 282.555 |
| *251* | hsa–miR–148b–3p | PlasmaExo–Lab 4 | 42/88 | 0.098 | 1557.841 |
| *252* | hsa–miR–148b–3p | Serum–Lab 3 | 40/45 | 0.079 | 731.418 |
| *253* | hsa–miR–127–3p | Urine–Lab 4 | 23/23 | 0.382 | 55.917 |
| *254* | hsa–miR–127–3p | PlasmaExo–Lab 4 | 98/99 | 0.271 | 1485.423 |
| *255* | hsa–miR–127–3p | Plasma–Lab 1 | 165/166 | 0.190 | 5711.132 |
| *256* | hsa–miR–127–3p | Serum–Lab 5 | 66/72 | 0.167 | 177.042 |
| *257* | hsa–miR–127–3p | CSF–Lab 3 | 64/64 | 0.147 | 3078.889 |
| *258* | hsa–miR–127–3p | Urine–Lab 1 | 42/58 | 0.147 | 157.738 |
| *259* | hsa–miR–127–3p | Serum–Lab 3 | 48/49 | 0.141 | 1308.204 |
| *260* | hsa–miR–127–3p | Plasma–Lab 5 | 102/148 | 0.135 | 49.958 |
| *261* | hsa–miR–6735–5p | Plasma–Lab 1 | 7/9 | 0.381 | 11.100 |
| *262* | hsa–miR–1343–3p | PlasmaExo–Lab 2 | 50/50 | 0.378 | 100.954 |
| *263* | hsa–miR–1343–3p | Plasma–Lab 1 | 11/34 | 0.066 | 10.041 |
| *264* | hsa–miR–1226–3p | PlasmaExo–Lab 4 | 9/12 | 0.376 | 4.942 |
| *265* | hsa–miR–1226–3p | Serum–Lab 5 | 25/30 | 0.350 | 12.980 |
| *266* | hsa–miR–1226–3p | Plasma–Lab 1 | 15/18 | 0.300 | 6.561 |
| *267* | hsa–miR–1226–3p | Plasma–Lab 5 | 9/17 | 0.169 | 8.506 |
| *268* | hsa–miR–152–3p | Urine–Lab 4 | 25/25 | 0.376 | 221.236 |
| *269* | hsa–miR–152–3p | CSF–Lab 3 | 6/7 | 0.244 | 70.771 |

| | | | | | |
|---|---|---|---|---|---|
| *270* | hsa–miR–152–3p | PlasmaExo–Lab 4 | 48/66 | 0.188 | 101.847 |
| *271* | hsa–miR–152–3p | PlasmaExo–Lab 2 | 50/50 | 0.083 | 414.958 |
| *272* | hsa–miR–152–3p | Serum–Lab 3 | 11/14 | 0.059 | 76.236 |
| *273* | hsa–miR–6513–3p | PlasmaExo–Lab 4 | 11/14 | 0.373 | 6.257 |
| *274* | hsa–miR–6513–3p | Plasma–Lab 1 | 22/26 | 0.279 | 7.565 |
| *275* | hsa–miR–6513–3p | Serum–Lab 5 | 8/18 | 0.133 | 7.739 |
| *276* | hsa–miR–5189–5p | Plasma–Lab 1 | 12/15 | 0.370 | 8.680 |
| *277* | hsa–miR–484 | PlasmaExo–Lab 4 | 43/43 | 0.366 | 863.951 |
| *278* | hsa–miR–484 | Urine–Lab 4 | 10/11 | 0.353 | 12.245 |
| *279* | hsa–miR–484 | Plasma–Lab 1 | 85/99 | 0.199 | 965.337 |
| *280* | hsa–miR–484 | Urine–Lab 1 | 24/35 | 0.182 | 23.351 |
| *281* | hsa–miR–484 | Serum–Lab 5 | 21/23 | 0.145 | 1555.013 |
| *282* | hsa–miR–484 | PlasmaExo–Lab 2 | 41/41 | 0.141 | 148.768 |
| *283* | hsa–miR–484 | CSF–Lab 3 | 20/28 | 0.111 | 16.150 |
| *284* | hsa–miR–92a–3p | Urine–Lab 1 | 110/110 | 0.362 | 738.359 |
| *285* | hsa–miR–92a–3p | Plasma–Lab 1 | 167/167 | 0.344 | 13930.833 |
| *286* | hsa–miR–92a–3p | CSF–Lab 3 | 62/64 | 0.199 | 91.897 |
| *287* | hsa–miR–92a–3p | Serum–Lab 3 | 49/50 | 0.135 | 882.814 |
| *288* | hsa–miR–92a–3p | Serum–Lab 5 | 74/77 | 0.102 | 1297.999 |
| *289* | hsa–miR–92a–3p | PlasmaExo–Lab 4 | 94/101 | 0.065 | 933.012 |
| *290* | hsa–miR–92a–3p | Urine–Lab 4 | 16/23 | 0.061 | 47.317 |
| *291* | hsa–miR–92a–3p | PlasmaExo–Lab 2 | 46/50 | 0.059 | 58.194 |
| *292* | hsa–miR–1273h–3p | PlasmaExo–Lab 4 | 30/31 | 0.359 | 13.039 |
| *293* | hsa–miR–1273h–3p | Serum–Lab 5 | 7/11 | 0.324 | 5.727 |
| *294* | hsa–miR–1273h–3p | Plasma–Lab 1 | 44/85 | 0.075 | 16.108 |
| *295* | chr3_29850 | Serum–Lab 5 | 5/9 | 0.354 | 5.722 |
| *296* | chr3_29850 | Plasma–Lab 1 | 120/127 | 0.275 | 42.520 |
| *297* | hsa–miR–187–3p | CSF–Lab 3 | 5/5 | 0.350 | 16.520 |
| *298* | hsa–miR–187–3p | PlasmaExo–Lab 2 | 6/7 | 0.054 | 37.000 |
| *299* | hsa–miR–203a–3p | Urine–Lab 4 | 27/27 | 0.349 | 4207.496 |
| *300* | hsa–miR–203a–3p | PlasmaExo–Lab 4 | 38/39 | 0.302 | 128.741 |
| *301* | hsa–miR–203a–3p | Urine–Lab 1 | 25/46 | 0.272 | 157.778 |
| *302* | hsa–miR–203a–3p | Plasma–Lab 5 | 32/91 | 0.071 | 21.842 |
| *303* | hsa–miR–203a–3p | Serum–Lab 5 | 20/49 | 0.061 | 32.857 |
| *304* | hsa–miR–203a–3p | Serum–Lab 3 | 5/20 | 0.061 | 236.375 |
| *305* | hsa–miR–1237–3p | Plasma–Lab 1 | 6/6 | 0.348 | 4.833 |
| *306* | hsa–miR–331–5p | PlasmaExo–Lab 4 | 25/27 | 0.348 | 7.748 |
| *307* | hsa–miR–331–5p | PlasmaExo–Lab 2 | 49/50 | 0.145 | 23.134 |
| *308* | hsa–miR–331–5p | Serum–Lab 5 | 30/51 | 0.064 | 27.598 |

| | | | | | |
|---|---|---|---|---|---|
| *309* | hsa–miR–331–5p | Plasma–Lab 5 | 84/151 | 0.060 | 35.903 |
| *310* | hsa–miR–331–5p | Plasma–Lab 1 | 9/37 | 0.051 | 12.268 |
| *311* | hsa–miR–370–3p | Plasma–Lab 1 | 129/131 | 0.345 | 104.206 |
| *312* | hsa–miR–370–3p | Serum–Lab 3 | 6/6 | 0.266 | 42.967 |
| *313* | hsa–miR–370–3p | PlasmaExo–Lab 4 | 84/87 | 0.168 | 564.694 |
| *314* | hsa–miR–370–3p | Plasma–Lab 5 | 63/95 | 0.165 | 33.501 |
| *315* | hsa–miR–370–3p | Serum–Lab 5 | 57/73 | 0.142 | 78.052 |
| *316* | hsa–miR–370–3p | CSF–Lab 3 | 12/19 | 0.087 | 15.084 |
| *317* | hsa–miR–328–3p | Urine–Lab 1 | 49/54 | 0.344 | 46.857 |
| *318* | hsa–miR–328–3p | Plasma–Lab 1 | 161/166 | 0.316 | 410.917 |
| *319* | hsa–miR–328–3p | Urine–Lab 4 | 11/13 | 0.264 | 14.862 |
| *320* | hsa–miR–328–3p | Serum–Lab 3 | 30/34 | 0.232 | 34.818 |
| *321* | hsa–miR–328–3p | CSF–Lab 3 | 51/55 | 0.180 | 67.922 |
| *322* | hsa–miR–328–3p | Plasma–Lab 5 | 140/171 | 0.164 | 128.891 |
| *323* | hsa–miR–328–3p | Serum–Lab 5 | 71/72 | 0.163 | 344.700 |
| *324* | hsa–miR–328–3p | PlasmaExo–Lab 4 | 100/101 | 0.137 | 541.950 |
| *325* | hsa–miR–3191–3p | PlasmaExo–Lab 4 | 8/9 | 0.343 | 7.167 |
| *326* | hsa–miR–376a–5p | Plasma–Lab 1 | 8/9 | 0.343 | 12.100 |
| *327* | hsa–miR–7–1–3p | Plasma–Lab 1 | 6/13 | 0.339 | 7.431 |
| *328* | hsa–miR–7–1–3p | PlasmaExo–Lab 2 | 21/35 | 0.242 | 13.729 |
| *329* | hsa–miR–7–1–3p | Plasma–Lab 5 | 13/59 | 0.124 | 23.336 |
| *330* | hsa–miR–7–1–3p | Serum–Lab 5 | 6/44 | 0.067 | 53.436 |
| *331* | hsa–miR–7706 | Plasma–Lab 1 | 127/139 | 0.337 | 52.328 |
| *332* | hsa–miR–7706 | Urine–Lab 1 | 12/19 | 0.302 | 14.663 |
| *333* | hsa–miR–7706 | PlasmaExo–Lab 4 | 51/55 | 0.204 | 20.773 |
| *334* | hsa–miR–7706 | Plasma–Lab 5 | 41/68 | 0.181 | 12.693 |
| *335* | hsa–miR–7706 | Serum–Lab 5 | 8/20 | 0.174 | 9.875 |
| *336* | hsa–miR–7706 | Serum–Lab 3 | 14/43 | 0.052 | 70.742 |
| *337* | hsa–miR–320a–3p | PlasmaExo–Lab 4 | 63/63 | 0.337 | 4792.944 |
| *338* | hsa–miR–320a–3p | Plasma–Lab 1 | 153/153 | 0.311 | 8569.275 |
| *339* | hsa–miR–320a–3p | Urine–Lab 1 | 53/53 | 0.234 | 397.553 |
| *340* | hsa–miR–320a–3p | CSF–Lab 3 | 64/64 | 0.190 | 967.467 |
| *341* | hsa–miR–320a–3p | Serum–Lab 3 | 50/50 | 0.139 | 5261.066 |
| *342* | hsa–miR–320a–3p | Serum–Lab 5 | 76/77 | 0.119 | 5823.710 |
| *343* | hsa–miR–320a–3p | Plasma–Lab 5 | 182/187 | 0.091 | 3038.949 |
| *344* | hsa–miR–320a–3p | PlasmaExo–Lab 2 | 50/50 | 0.084 | 28586.030 |
| *345* | hsa–miR–1976 | Plasma–Lab 1 | 67/79 | 0.335 | 17.441 |
| *346* | hsa–miR–1976 | PlasmaExo–Lab 4 | 47/70 | 0.172 | 13.714 |

| | | | | | |
|---|---|---|---|---|---|
| *347* | hsa–miR–1976 | Serum–Lab 5 | 19/48 | 0.088 | 21.231 |
| *348* | hsa–miR–18a–3p | PlasmaExo–Lab 4 | 46/54 | 0.331 | 13.235 |
| *349* | hsa–miR–18a–3p | Plasma–Lab 1 | 124/144 | 0.234 | 42.787 |
| *350* | hsa–miR–18a–3p | Serum–Lab 5 | 28/44 | 0.119 | 28.357 |
| *351* | hsa–miR–18a–3p | Plasma–Lab 5 | 72/140 | 0.059 | 43.877 |
| *352* | chr11_2103 | Serum–Lab 5 | 8/13 | 0.321 | 8.069 |
| *353* | hsa–miR–500a–3p | Urine–Lab 4 | 22/23 | 0.320 | 33.870 |
| *354* | hsa–miR–500a–3p | CSF–Lab 3 | 41/50 | 0.090 | 36.436 |
| *355* | hsa–miR–500a–3p | Plasma–Lab 1 | 135/159 | 0.069 | 193.188 |
| *356* | hsa–miR–500a–3p | Urine–Lab 1 | 84/107 | 0.069 | 162.335 |
| *357* | hsa–miR–500a–3p | PlasmaExo–Lab 4 | 12/74 | 0.059 | 17.193 |
| *358* | hsa–miR–29c–5p | CSF–Lab 3 | 13/14 | 0.318 | 15.600 |
| *359* | hsa–miR–29c–5p | Serum–Lab 5 | 9/22 | 0.081 | 30.882 |
| *360* | hsa–miR–769–3p | Serum–Lab 5 | 9/12 | 0.316 | 5.833 |
| *361* | hsa–miR–1307–3p | CSF–Lab 3 | 9/17 | 0.316 | 13.888 |
| *362* | hsa–miR–1307–3p | Urine–Lab 1 | 7/18 | 0.191 | 28.867 |
| *363* | hsa–miR–361–3p | PlasmaExo–Lab 4 | 39/39 | 0.313 | 365.554 |
| *364* | hsa–miR–671–3p | PlasmaExo–Lab 4 | 14/16 | 0.308 | 118.762 |
| *365* | hsa–miR–671–3p | Plasma–Lab 1 | 133/138 | 0.294 | 237.559 |
| *366* | hsa–miR–671–3p | Urine–Lab 1 | 14/28 | 0.151 | 31.154 |
| *367* | hsa–miR–671–3p | Plasma–Lab 5 | 5/19 | 0.149 | 9.737 |
| *368* | hsa–miR–671–3p | Serum–Lab 5 | 5/17 | 0.114 | 24.329 |
| *369* | chr6_38453 | Plasma–Lab 1 | 16/16 | 0.306 | 37.400 |
| *370* | chr16_12669 | Plasma–Lab 1 | 5/8 | 0.304 | 15.463 |
| *371* | hsa–miR–6793–3p | Plasma–Lab 5 | 5/9 | 0.303 | 4.956 |
| *372* | hsa–miR–4523 | Plasma–Lab 1 | 17/21 | 0.297 | 18.738 |
| *373* | hsa–miR–4523 | Serum–Lab 5 | 6/10 | 0.209 | 6.970 |
| *374* | chr20_23368 | PlasmaExo–Lab 4 | 45/48 | 0.296 | 67.579 |
| *375* | chr20_23368 | Plasma–Lab 1 | 144/148 | 0.254 | 211.926 |
| *376* | chr20_23368 | Serum–Lab 5 | 10/17 | 0.233 | 11.888 |
| *377* | chr20_23368 | Serum–Lab 3 | 8/16 | 0.072 | 30.381 |
| *378* | hsa–miR–6764–5p | PlasmaExo–Lab 4 | 9/10 | 0.295 | 6.220 |
| *379* | hsa–miR–1908–3p | Plasma–Lab 1 | 76/81 | 0.289 | 16.988 |
| *380* | hsa–miR–1908–3p | Serum–Lab 5 | 6/11 | 0.169 | 9.082 |
| *381* | hsa–miR–195–3p | Plasma–Lab 1 | 9/12 | 0.289 | 18.558 |
| *382* | hsa–miR–3615 | Plasma–Lab 1 | 58/65 | 0.286 | 297.682 |
| *383* | hsa–miR–3615 | Urine–Lab 4 | 9/11 | 0.146 | 18.655 |
| *384* | hsa–miR–3615 | Urine–Lab 1 | 13/32 | 0.100 | 19.859 |
| *385* | hsa–miR–3615 | Serum–Lab 5 | 51/55 | 0.078 | 447.775 |

| | | | | | |
|---|---|---|---|---|---|
| *386* | hsa–miR–23a–5p | Plasma–Lab 1 | 111/117 | 0.280 | 89.443 |
| *387* | hsa–miR–23a–5p | PlasmaExo–Lab 4 | 7/11 | 0.131 | 22.473 |
| *388* | hsa–miR–23a–5p | Serum–Lab 5 | 18/32 | 0.113 | 15.594 |
| *389* | chr18_16031 | Plasma–Lab 5 | 10/13 | 0.280 | 3.838 |
| *390* | chr18_16031 | Plasma–Lab 1 | 60/72 | 0.248 | 21.056 |
| *391* | chr18_16031 | Serum–Lab 3 | 13/20 | 0.111 | 14.695 |
| *392* | hsa–miR–106b–3p | Urine–Lab 4 | 12/13 | 0.274 | 34.554 |
| *393* | hsa–miR–106b–3p | PlasmaExo–Lab 2 | 22/23 | 0.114 | 74.474 |
| *394* | hsa–miR–106b–3p | Serum–Lab 5 | 39/54 | 0.063 | 549.869 |
| *395* | hsa–miR–486–3p | PlasmaExo–Lab 4 | 48/54 | 0.272 | 11.846 |
| *396* | hsa–miR–486–3p | Plasma–Lab 1 | 78/89 | 0.240 | 23.767 |
| *397* | hsa–miR–486–3p | Serum–Lab 3 | 9/13 | 0.152 | 14.708 |
| *398* | hsa–miR–486–3p | Serum–Lab 5 | 10/31 | 0.068 | 16.810 |
| *399* | hsa–miR–486–3p | Plasma–Lab 5 | 42/102 | 0.066 | 15.823 |
| *400* | hsa–miR–30a–3p | Plasma–Lab 5 | 116/130 | 0.269 | 65.398 |
| *401* | hsa–miR–30a–3p | Serum–Lab 5 | 49/53 | 0.265 | 71.679 |
| *402* | hsa–miR–30a–3p | PlasmaExo–Lab 4 | 18/28 | 0.210 | 17.375 |
| *403* | hsa–miR–30a–3p | Serum–Lab 3 | 16/37 | 0.066 | 228.024 |
| *404* | hsa–miR–30a–3p | CSF–Lab 3 | 27/52 | 0.058 | 663.937 |
| *405* | hsa–miR–126–5p | PlasmaExo–Lab 4 | 101/101 | 0.262 | 567.133 |
| *406* | hsa–miR–126–5p | PlasmaExo–Lab 2 | 50/50 | 0.187 | 46.886 |
| *407* | hsa–miR–126–5p | Serum–Lab 5 | 73/74 | 0.071 | 9265.234 |
| *408* | hsa–miR–339–3p | PlasmaExo–Lab 4 | 22/28 | 0.259 | 11.961 |
| *409* | hsa–miR–339–3p | Plasma–Lab 1 | 10/25 | 0.147 | 107.116 |
| *410* | hsa–miR–339–3p | Serum–Lab 5 | 41/52 | 0.111 | 116.488 |
| *411* | chr19_17473 | Plasma–Lab 1 | 81/92 | 0.257 | 28.096 |
| *412* | hsa–miR–1908–5p | Plasma–Lab 1 | 107/114 | 0.255 | 43.076 |
| *413* | hsa–miR–3120–3p | PlasmaExo–Lab 4 | 7/8 | 0.254 | 17.113 |
| *414* | hsa–miR–3120–3p | Serum–Lab 5 | 5/12 | 0.151 | 10.358 |
| *415* | hsa–miR–3120–3p | Plasma–Lab 1 | 58/82 | 0.109 | 21.270 |
| *416* | hsa–miR–193b–5p | Serum–Lab 3 | 5/8 | 0.254 | 17.387 |
| *417* | hsa–miR–193b–5p | Urine–Lab 4 | 6/15 | 0.060 | 15.680 |
| *418* | hsa–miR–7158–3p | CSF–Lab 3 | 8/8 | 0.249 | 10.113 |
| *419* | hsa–miR–3909 | Urine–Lab 4 | 5/6 | 0.244 | 7.350 |
| *420* | hsa–miR–665 | Serum–Lab 5 | 6/12 | 0.243 | 11.025 |
| *421* | hsa–miR–5189–3p | Plasma–Lab 1 | 12/18 | 0.240 | 9.633 |
| *422* | hsa–miR–1292–5p | PlasmaExo–Lab 4 | 27/34 | 0.240 | 10.385 |
| *423* | hsa–miR–1538 | Plasma–Lab 5 | 6/13 | 0.236 | 8.031 |

| | | | | | |
|---|---|---|---|---|---|
| *424* | hsa–miR–181d–5p | Serum–Lab 5 | 38/45 | 0.231 | 77.673 |
| *425* | hsa–miR–181d–5p | Plasma–Lab 1 | 87/97 | 0.155 | 178.716 |
| *426* | hsa–miR–181d–5p | Plasma–Lab 5 | 23/64 | 0.096 | 25.131 |
| *427* | hsa–miR–17–3p | Plasma–Lab 1 | 13/28 | 0.231 | 222.854 |
| *428* | hsa–miR–150–3p | PlasmaExo–Lab 4 | 32/58 | 0.231 | 14.997 |
| *429* | hsa–miR–150–3p | Serum–Lab 3 | 14/22 | 0.135 | 44.568 |
| *430* | hsa–miR–150–3p | Plasma–Lab 5 | 42/80 | 0.128 | 43.056 |
| *431* | hsa–miR–150–3p | Plasma–Lab 1 | 64/92 | 0.125 | 24.892 |
| *432* | hsa–miR–150–3p | Serum–Lab 5 | 17/39 | 0.076 | 56.756 |
| *433* | chr1_22324 | Serum–Lab 5 | 9/15 | 0.228 | 8.627 |
| *434* | chr1_22324 | Plasma–Lab 1 | 6/17 | 0.152 | 8.571 |
| *435* | chr1_22324 | Plasma–Lab 5 | 15/49 | 0.062 | 10.031 |
| *436* | hsa–let–7e–5p | Plasma–Lab 1 | 156/158 | 0.224 | 819.259 |
| *437* | hsa–let–7e–5p | Serum–Lab 5 | 65/71 | 0.182 | 107.593 |
| *438* | hsa–let–7e–5p | PlasmaExo–Lab 4 | 90/91 | 0.159 | 348.796 |
| *439* | hsa–let–7e–5p | Urine–Lab 1 | 92/97 | 0.144 | 507.657 |
| *440* | hsa–let–7e–5p | Urine–Lab 4 | 14/15 | 0.126 | 767.807 |
| *441* | hsa–let–7e–5p | CSF–Lab 3 | 28/41 | 0.109 | 209.256 |
| *442* | hsa–let–7e–5p | Serum–Lab 3 | 22/30 | 0.105 | 89.010 |
| *443* | hsa–let–7e–5p | Plasma–Lab 5 | 82/116 | 0.098 | 36.090 |
| *444* | hsa–miR–181c–3p | PlasmaExo–Lab 4 | 23/28 | 0.224 | 21.557 |
| *445* | hsa–miR–181c–3p | Plasma–Lab 5 | 39/78 | 0.142 | 19.024 |
| *446* | hsa–miR–181c–3p | Serum–Lab 5 | 25/48 | 0.117 | 39.498 |
| *447* | hsa–miR–377–5p | PlasmaExo–Lab 2 | 50/50 | 0.221 | 60.182 |
| *448* | hsa–miR–377–5p | PlasmaExo–Lab 4 | 7/9 | 0.160 | 6.078 |
| *449* | hsa–miR–218–2–3p | PlasmaExo–Lab 2 | 10/16 | 0.220 | 5.525 |
| *450* | hsa–miR–766–5p | PlasmaExo–Lab 4 | 23/27 | 0.214 | 28.341 |
| *451* | hsa–miR–27a–3p | Plasma–Lab 1 | 116/123 | 0.214 | 598.852 |
| *452* | hsa–miR–27a–3p | CSF–Lab 3 | 20/29 | 0.110 | 104.959 |
| *453* | hsa–miR–27a–3p | Serum–Lab 5 | 75/76 | 0.090 | 10099.964 |
| *454* | hsa–miR–27a–3p | Plasma–Lab 5 | 180/186 | 0.078 | 4156.726 |
| *455* | hsa–miR–27a–3p | Urine–Lab 4 | 21/22 | 0.074 | 448.964 |
| *456* | hsa–miR–27a–3p | PlasmaExo–Lab 4 | 101/101 | 0.069 | 848.982 |
| *457* | hsa–miR–27a–3p | Serum–Lab 3 | 28/41 | 0.053 | 547.400 |
| *458* | chr7_41677 | Serum–Lab 3 | 17/19 | 0.213 | 37.805 |
| *459* | chr7_41677 | Plasma–Lab 1 | 143/153 | 0.179 | 253.288 |
| *460* | chr7_41677 | PlasmaExo–Lab 4 | 49/56 | 0.121 | 298.075 |
| *461* | chr7_41677 | PlasmaExo–Lab 2 | 10/21 | 0.095 | 8.014 |
| *462* | chr7_41677 | Serum–Lab 5 | 30/52 | 0.095 | 37.717 |

| | | | | | |
|---|---|---|---|---|---|
| *463* | chr7_41677 | Plasma–Lab 5 | 14/31 | 0.091 | 28.165 |
| *464* | hsa–miR–2114–5p | CSF–Lab 3 | 7/8 | 0.211 | 5.825 |
| *465* | hsa–miR–1271–5p | CSF–Lab 3 | 9/10 | 0.210 | 11.100 |
| *466* | hsa–miR–1271–5p | Plasma–Lab 1 | 23/31 | 0.206 | 10.281 |
| *467* | hsa–miR–532–3p | Urine–Lab 1 | 26/32 | 0.205 | 27.659 |
| *468* | hsa–miR–532–3p | PlasmaExo–Lab 4 | 5/11 | 0.095 | 5.582 |
| *469* | hsa–miR–181b–5p | Urine–Lab 1 | 18/27 | 0.201 | 21.974 |
| *470* | hsa–miR–181b–5p | Plasma–Lab 1 | 116/138 | 0.192 | 29.920 |
| *471* | hsa–miR–181b–5p | Serum–Lab 3 | 10/13 | 0.178 | 24.923 |
| *472* | hsa–miR–181b–5p | PlasmaExo–Lab 2 | 50/50 | 0.174 | 20.942 |
| *473* | hsa–miR–181b–5p | Serum–Lab 5 | 32/52 | 0.103 | 32.442 |
| *474* | hsa–miR–181b–5p | Plasma–Lab 5 | 53/85 | 0.103 | 18.460 |
| *475* | hsa–miR–181b–5p | CSF–Lab 3 | 19/28 | 0.089 | 19.518 |
| *476* | hsa–miR–185–3p | Serum–Lab 3 | 30/33 | 0.198 | 31.761 |
| *477* | hsa–miR–185–3p | Plasma–Lab 1 | 152/158 | 0.175 | 149.645 |
| *478* | hsa–miR–185–3p | Plasma–Lab 5 | 160/179 | 0.169 | 68.554 |
| *479* | hsa–miR–185–3p | Serum–Lab 5 | 65/75 | 0.162 | 102.601 |
| *480* | hsa–miR–185–3p | CSF–Lab 3 | 6/9 | 0.147 | 8.644 |
| *481* | hsa–miR–185–3p | PlasmaExo–Lab 4 | 46/81 | 0.084 | 107.060 |
| *482* | hsa–miR–3605–3p | Plasma–Lab 1 | 62/85 | 0.195 | 28.527 |
| *483* | hsa–miR–3605–3p | PlasmaExo–Lab 4 | 21/66 | 0.145 | 15.226 |
| *484* | hsa–miR–212–3p | Serum–Lab 5 | 6/15 | 0.194 | 9.247 |
| *485* | hsa–miR–5589–3p | Serum–Lab 3 | 5/5 | 0.193 | 27.260 |
| *486* | hsa–miR–27b–3p | Plasma–Lab 1 | 130/164 | 0.193 | 6730.318 |
| *487* | hsa–miR–27b–3p | Serum–Lab 5 | 60/68 | 0.063 | 2552.149 |
| *488* | hsa–miR–4665–5p | Plasma–Lab 1 | 10/14 | 0.192 | 11.114 |
| *489* | hsa–miR–4665–5p | PlasmaExo–Lab 4 | 28/35 | 0.113 | 23.731 |
| *490* | hsa–miR–1843 | Serum–Lab 5 | 17/34 | 0.191 | 30.221 |
| *491* | hsa–miR–548j–5p | Plasma–Lab 1 | 80/101 | 0.190 | 27.223 |
| *492* | hsa–miR–374a–3p | Plasma–Lab 1 | 17/24 | 0.189 | 12.446 |
| *493* | hsa–miR–374a–3p | Urine–Lab 4 | 9/13 | 0.130 | 17.046 |
| *494* | hsa–miR–374a–3p | PlasmaExo–Lab 4 | 35/50 | 0.113 | 27.718 |
| *495* | hsa–miR–134–5p | PlasmaExo–Lab 4 | 98/100 | 0.182 | 577.529 |
| *496* | hsa–miR–134–5p | Plasma–Lab 1 | 157/160 | 0.175 | 327.274 |
| *497* | hsa–miR–134–5p | Serum–Lab 5 | 58/68 | 0.147 | 67.049 |
| *498* | hsa–miR–134–5p | Serum–Lab 3 | 28/35 | 0.130 | 65.934 |
| *499* | hsa–miR–134–5p | CSF–Lab 3 | 28/42 | 0.117 | 32.148 |
| *500* | hsa–miR–134–5p | Plasma–Lab 5 | 62/115 | 0.096 | 32.126 |
| *501* | hsa–miR–2277–5p | Serum–Lab 5 | 6/12 | 0.179 | 5.492 |

| | | | | | |
|---|---|---|---|---|---|
| *502* | hsa–miR–1249–3p | Plasma–Lab 1 | 13/23 | 0.179 | 11.170 |
| *503* | hsa–miR–99b–5p | Plasma–Lab 1 | 146/151 | 0.172 | 986.903 |
| *504* | hsa–miR–99b–5p | Serum–Lab 5 | 44/50 | 0.109 | 517.762 |
| *505* | hsa–miR–99b–5p | Urine–Lab 1 | 99/110 | 0.061 | 211.171 |
| *506* | hsa–miR–151a–3p | CSF–Lab 3 | 64/64 | 0.172 | 742.350 |
| *507* | hsa–miR–151a–3p | Urine–Lab 1 | 110/110 | 0.158 | 271.951 |
| *508* | hsa–miR–151a–3p | Serum–Lab 3 | 51/51 | 0.129 | 2469.657 |
| *509* | hsa–miR–151a–3p | Plasma–Lab 1 | 167/167 | 0.127 | 9829.999 |
| *510* | hsa–miR–151a–3p | Serum–Lab 5 | 75/77 | 0.120 | 417.825 |
| *511* | hsa–miR–151a–3p | Plasma–Lab 5 | 177/185 | 0.119 | 220.412 |
| *512* | hsa–miR–151a–3p | Urine–Lab 4 | 27/27 | 0.109 | 730.422 |
| *513* | hsa–miR–151a–3p | PlasmaExo–Lab 4 | 101/101 | 0.108 | 2866.488 |
| *514* | hsa–miR–151a–3p | PlasmaExo–Lab 2 | 25/25 | 0.073 | 829.812 |
| *515* | hsa–let–7i–5p | Serum–Lab 5 | 67/70 | 0.172 | 5764.631 |
| *516* | hsa–let–7i–5p | PlasmaExo–Lab 2 | 22/22 | 0.084 | 26963.055 |
| *517* | hsa–let–7i–5p | CSF–Lab 3 | 64/64 | 0.083 | 2288.272 |
| *518* | hsa–let–7i–5p | Plasma–Lab 1 | 164/164 | 0.068 | 16434.757 |
| *519* | hsa–let–7i–5p | PlasmaExo–Lab 4 | 39/45 | 0.065 | 34545.640 |
| *520* | hsa–let–7i–5p | Serum–Lab 3 | 32/33 | 0.062 | 3619.197 |
| *521* | hsa–let–7i–5p | Urine–Lab 1 | 109/110 | 0.061 | 1046.350 |
| *522* | hsa–let–7i–5p | Plasma–Lab 5 | 19/53 | 0.054 | 6840.147 |
| *523* | hsa–miR–138–2–3p | PlasmaExo–Lab 2 | 23/35 | 0.167 | 38.357 |
| *524* | hsa–miR–766–3p | Plasma–Lab 1 | 131/134 | 0.164 | 89.487 |
| *525* | hsa–miR–766–3p | Serum–Lab 3 | 6/10 | 0.105 | 25.470 |
| *526* | hsa–miR–766–3p | PlasmaExo–Lab 4 | 27/34 | 0.099 | 24.462 |
| *527* | hsa–miR–766–3p | Serum–Lab 5 | 42/65 | 0.098 | 47.283 |
| *528* | hsa–miR–1306–3p | PlasmaExo–Lab 4 | 6/13 | 0.161 | 6.746 |
| *529* | hsa–miR–1306–3p | Plasma–Lab 1 | 6/13 | 0.159 | 6.946 |
| *530* | hsa–miR–577 | Urine–Lab 1 | 24/32 | 0.161 | 19.337 |
| *531* | hsa–miR–30c–1–3p | PlasmaExo–Lab 4 | 9/16 | 0.160 | 8.369 |
| *532* | hsa–miR–4446–3p | PlasmaExo–Lab 4 | 67/76 | 0.159 | 107.429 |
| *533* | hsa–miR–4446–3p | Plasma–Lab 1 | 144/160 | 0.117 | 152.915 |
| *534* | hsa–miR–4446–3p | Plasma–Lab 5 | 8/18 | 0.110 | 21.378 |
| *535* | hsa–miR–4446–3p | Serum–Lab 3 | 7/12 | 0.106 | 18.900 |
| *536* | hsa–miR–4446–3p | Serum–Lab 5 | 21/37 | 0.099 | 22.014 |
| *537* | hsa–miR–28–3p | Plasma–Lab 5 | 88/98 | 0.158 | 150.942 |
| *538* | hsa–miR–28–3p | Serum–Lab 5 | 47/49 | 0.152 | 325.761 |
| *539* | hsa–miR–28–3p | Plasma–Lab 1 | 148/165 | 0.051 | 474.838 |

| | | | | | |
|---|---|---|---|---|---|
| *540* | hsa–miR–885–3p | PlasmaExo–Lab 2 | 50/50 | 0.158 | 127.600 |
| *541* | hsa–miR–885–3p | Plasma–Lab 5 | 7/17 | 0.113 | 11.271 |
| *542* | hsa–miR–25–3p | Urine–Lab 4 | 21/22 | 0.155 | 145.173 |
| *543* | hsa–miR–25–3p | PlasmaExo–Lab 4 | 42/42 | 0.096 | 2736.355 |
| *544* | hsa–miR–323a–5p | PlasmaExo–Lab 4 | 9/13 | 0.154 | 8.923 |
| *545* | hsa–miR–16–5p | Serum–Lab 5 | 38/53 | 0.152 | 60.187 |
| *546* | hsa–miR–16–5p | PlasmaExo–Lab 4 | 11/22 | 0.079 | 9.891 |
| *547* | hsa–miR–204–5p | PlasmaExo–Lab 4 | 5/11 | 0.151 | 7.245 |
| *548* | hsa–miR–204–5p | Urine–Lab 4 | 19/19 | 0.109 | 695.316 |
| *549* | hsa–miR–204–5p | Plasma–Lab 1 | 77/124 | 0.093 | 577.699 |
| *550* | hsa–miR–3200–3p | CSF–Lab 3 | 20/25 | 0.151 | 18.796 |
| *551* | hsa–miR–3200–3p | Plasma–Lab 5 | 72/125 | 0.130 | 32.320 |
| *552* | hsa–miR–3200–3p | Serum–Lab 5 | 15/34 | 0.124 | 16.421 |
| *553* | hsa–miR–3200–3p | Plasma–Lab 1 | 20/36 | 0.084 | 15.625 |
| *554* | hsa–miR–3200–3p | Serum–Lab 3 | 5/12 | 0.054 | 12.025 |
| *555* | hsa–miR–188–5p | Serum–Lab 5 | 33/51 | 0.147 | 50.918 |
| *556* | hsa–miR–188–5p | Plasma–Lab 5 | 72/126 | 0.102 | 34.003 |
| *557* | hsa–miR–188–5p | Urine–Lab 1 | 8/20 | 0.095 | 14.435 |
| *558* | hsa–miR–126–3p | Urine–Lab 1 | 7/13 | 0.147 | 20.823 |
| *559* | hsa–miR–126–3p | Plasma–Lab 1 | 162/166 | 0.141 | 2671.689 |
| *560* | hsa–miR–126–3p | PlasmaExo–Lab 2 | 50/50 | 0.139 | 1490.820 |
| *561* | hsa–miR–126–3p | Serum–Lab 3 | 43/47 | 0.107 | 394.532 |
| *562* | hsa–miR–126–3p | CSF–Lab 3 | 32/43 | 0.104 | 141.023 |
| *563* | hsa–miR–126–3p | Plasma–Lab 5 | 180/187 | 0.081 | 11605.440 |
| *564* | hsa–miR–126–3p | Serum–Lab 5 | 72/77 | 0.073 | 19866.526 |
| *565* | hsa–miR–584–5p | Urine–Lab 4 | 8/12 | 0.146 | 28.025 |
| *566* | hsa–miR–584–5p | Plasma–Lab 1 | 156/162 | 0.134 | 2101.617 |
| *567* | hsa–miR–584–5p | PlasmaExo–Lab 4 | 23/23 | 0.073 | 1957.852 |
| *568* | hsa–miR–584–5p | Urine–Lab 1 | 6/19 | 0.054 | 26.958 |
| *569* | hsa–miR–744–5p | Serum–Lab 5 | 76/77 | 0.144 | 549.996 |
| *570* | hsa–miR–744–5p | Plasma–Lab 5 | 44/68 | 0.078 | 198.534 |
| *571* | hsa–miR–744–5p | Plasma–Lab 1 | 166/167 | 0.072 | 4040.144 |
| *572* | hsa–miR–744–5p | PlasmaExo–Lab 4 | 78/81 | 0.063 | 1752.763 |
| *573* | hsa–miR–6741–3p | PlasmaExo–Lab 4 | 23/37 | 0.142 | 14.730 |
| *574* | hsa–miR–6741–3p | Plasma–Lab 1 | 69/95 | 0.120 | 22.886 |
| *575* | hsa–miR–6741–3p | Serum–Lab 5 | 7/18 | 0.112 | 8.244 |
| *576* | hsa–miR–423–3p | Plasma–Lab 1 | 167/167 | 0.141 | 12621.334 |
| *577* | hsa–miR–423–3p | PlasmaExo–Lab 4 | 72/72 | 0.124 | 8110.488 |
| *578* | hsa–miR–423–3p | Urine–Lab 1 | 109/110 | 0.115 | 829.458 |

| | | | | | |
|---|---|---|---|---|---|
| *579* | hsa–miR–423–3p | Urine–Lab 4 | 27/27 | 0.101 | 573.981 |
| *580* | hsa–miR–423–3p | Serum–Lab 5 | 64/66 | 0.089 | 1877.409 |
| *581* | hsa–miR–423–3p | Plasma–Lab 5 | 112/128 | 0.055 | 555.820 |
| *582* | hsa–miR–423–3p | Serum–Lab 3 | 49/51 | 0.053 | 1160.516 |
| *583* | hsa–miR–423–3p | CSF–Lab 3 | 63/64 | 0.051 | 266.772 |
| *584* | hsa–miR–147b–3p | Plasma–Lab 5 | 8/19 | 0.140 | 6.511 |
| *585* | hsa–miR–7849–3p | Plasma–Lab 1 | 19/28 | 0.138 | 22.279 |
| *586* | hsa–miR–146b–3p | PlasmaExo–Lab 4 | 8/13 | 0.138 | 7.177 |
| *587* | hsa–miR–146b–3p | Plasma–Lab 1 | 59/88 | 0.098 | 24.744 |
| *588* | hsa–miR–769–5p | Serum–Lab 5 | 22/22 | 0.136 | 159.559 |
| *589* | hsa–miR–769–5p | Plasma–Lab 1 | 165/166 | 0.109 | 1242.669 |
| *590* | hsa–miR–769–5p | Urine–Lab 1 | 65/96 | 0.083 | 127.610 |
| *591* | hsa–miR–769–5p | Plasma–Lab 5 | 43/69 | 0.078 | 97.620 |
| *592* | hsa–miR–769–5p | Serum–Lab 3 | 48/51 | 0.054 | 507.175 |
| *593* | hsa–miR–92b–3p | CSF–Lab 3 | 56/57 | 0.135 | 548.981 |
| *594* | hsa–miR–92b–3p | Plasma–Lab 1 | 8/11 | 0.129 | 315.864 |
| *595* | hsa–miR–92b–3p | Urine–Lab 1 | 87/96 | 0.118 | 258.889 |
| *596* | hsa–miR–92b–3p | Urine–Lab 4 | 9/14 | 0.085 | 24.493 |
| *597* | hsa–miR–92b–3p | Serum–Lab 3 | 10/17 | 0.058 | 184.424 |
| *598* | hsa–miR–92b–3p | Plasma–Lab 5 | 118/144 | 0.053 | 194.288 |
| *599* | hsa–miR–3177–3p | Plasma–Lab 1 | 59/78 | 0.134 | 17.209 |
| *600* | hsa–miR–3177–3p | Serum–Lab 5 | 5/13 | 0.086 | 8.569 |
| *601* | hsa–miR–127–5p | Serum–Lab 5 | 15/28 | 0.134 | 18.346 |
| *602* | hsa–miR–125a–5p | PlasmaExo–Lab 4 | 34/64 | 0.133 | 249.472 |
| *603* | hsa–miR–125a–5p | Urine–Lab 4 | 18/21 | 0.061 | 317.081 |
| *604* | hsa–miR–33a–3p | Serum–Lab 5 | 16/30 | 0.133 | 13.657 |
| *605* | hsa–miR–192–5p | PlasmaExo–Lab 4 | 37/37 | 0.129 | 595.692 |
| *606* | hsa–miR–192–5p | Serum–Lab 3 | 49/49 | 0.060 | 23435.431 |
| *607* | hsa–miR–4687–5p | Plasma–Lab 1 | 6/10 | 0.122 | 5.870 |
| *608* | hsa–miR–582–3p | Urine–Lab 1 | 5/21 | 0.120 | 14.095 |
| *609* | hsa–miR–424–3p | Urine–Lab 1 | 17/32 | 0.119 | 21.744 |
| *610* | hsa–miR–181b–3p | Plasma–Lab 5 | 7/21 | 0.118 | 10.171 |
| *611* | hsa–miR–3187–3p | PlasmaExo–Lab 4 | 22/36 | 0.116 | 8.547 |
| *612* | hsa–miR–200c–3p | Urine–Lab 1 | 18/37 | 0.115 | 17.468 |
| *613* | hsa–miR–200c–3p | PlasmaExo–Lab 4 | 49/74 | 0.080 | 69.384 |
| *614* | chr22_25085 | CSF–Lab 3 | 21/45 | 0.114 | 12.118 |
| *615* | hsa–miR–184 | CSF–Lab 3 | 13/19 | 0.114 | 33.758 |
| *616* | hsa–miR–184 | Urine–Lab 1 | 7/17 | 0.104 | 19.118 |
| *617* | hsa–miR–184 | Urine–Lab 4 | 10/18 | 0.098 | 44.106 |

| | | | | | |
|---|---|---|---|---|---|
| *618* | hsa–miR–34b–3p | CSF–Lab 3 | 55/57 | 0.111 | 149.005 |
| *619* | hsa–miR–10527–5p | Urine–Lab 4 | 8/13 | 0.110 | 11.831 |
| *620* | hsa–miR–10527–5p | PlasmaExo–Lab 4 | 32/75 | 0.061 | 20.312 |
| *621* | hsa–miR–299–3p | Serum–Lab 5 | 14/29 | 0.110 | 14.562 |
| *622* | hsa–miR–299–3p | Plasma–Lab 5 | 6/22 | 0.051 | 10.755 |
| *623* | hsa–miR–210–5p | Plasma–Lab 1 | 11/23 | 0.109 | 7.417 |
| *624* | hsa–miR–3605–5p | CSF–Lab 3 | 9/17 | 0.109 | 7.147 |
| *625* | hsa–miR–3605–5p | Serum–Lab 3 | 7/18 | 0.087 | 17.906 |
| *626* | hsa–miR–4690–3p | Plasma–Lab 5 | 5/14 | 0.103 | 6.607 |
| *627* | hsa–miR–93–3p | PlasmaExo–Lab 4 | 7/13 | 0.101 | 6.308 |
| *628* | hsa–miR–93–3p | Serum–Lab 5 | 33/62 | 0.075 | 66.560 |
| *629* | hsa–miR–1224–5p | PlasmaExo–Lab 4 | 12/21 | 0.100 | 8.114 |
| *630* | hsa–miR–3176 | Plasma–Lab 5 | 7/23 | 0.099 | 8.717 |
| *631* | hsa–miR–744–3p | PlasmaExo–Lab 2 | 8/28 | 0.096 | 15.654 |
| *632* | hsa–miR–4433b–3p | Plasma–Lab 1 | 73/93 | 0.096 | 76.530 |
| *633* | hsa–miR–4433b–3p | PlasmaExo–Lab 4 | 53/73 | 0.056 | 111.227 |
| *634* | hsa–miR–205–3p | Plasma–Lab 5 | 5/10 | 0.095 | 7.130 |
| *635* | hsa–miR–130b–3p | PlasmaExo–Lab 2 | 16/21 | 0.093 | 10.767 |
| *636* | hsa–miR–6813–5p | Plasma–Lab 1 | 8/16 | 0.093 | 8.450 |
| *637* | hsa–miR–150–5p | Plasma–Lab 5 | 175/182 | 0.093 | 1833.920 |
| *638* | hsa–miR–150–5p | Serum–Lab 5 | 71/75 | 0.086 | 1610.444 |
| *639* | hsa–miR–150–5p | Plasma–Lab 1 | 149/165 | 0.070 | 527.305 |
| *640* | hsa–miR–150–5p | PlasmaExo–Lab 2 | 33/50 | 0.069 | 17.450 |
| *641* | hsa–miR–150–5p | CSF–Lab 3 | 12/22 | 0.069 | 30.064 |
| *642* | hsa–miR–150–5p | PlasmaExo–Lab 4 | 88/101 | 0.054 | 229.451 |
| *643* | hsa–miR–340–3p | Plasma–Lab 1 | 94/118 | 0.091 | 90.207 |
| *644* | hsa–miR–340–3p | PlasmaExo–Lab 4 | 41/55 | 0.089 | 104.696 |
| *645* | hsa–miR–30b–3p | PlasmaExo–Lab 4 | 6/13 | 0.089 | 9.062 |
| *646* | hsa–miR–30b–3p | Plasma–Lab 1 | 26/61 | 0.063 | 13.595 |
| *647* | hsa–miR–501–3p | Urine–Lab 4 | 8/8 | 0.088 | 86.188 |
| *648* | hsa–miR–874–3p | Urine–Lab 1 | 7/21 | 0.088 | 18.986 |
| *649* | hsa–miR–146a–3p | Plasma–Lab 1 | 10/22 | 0.087 | 13.818 |
| *650* | hsa–miR–1306–5p | Serum–Lab 5 | 9/25 | 0.087 | 35.740 |
| *651* | hsa–miR–2355–3p | Plasma–Lab 5 | 28/56 | 0.087 | 34.079 |
| *652* | hsa–miR–2355–3p | Serum–Lab 5 | 13/30 | 0.051 | 65.880 |
| *653* | hsa–miR–483–3p | Plasma–Lab 5 | 10/27 | 0.086 | 13.667 |
| *654* | hsa–miR–3614–5p | Serum–Lab 3 | 6/12 | 0.085 | 16.450 |
| *655* | hsa–miR–5010–5p | Plasma–Lab 1 | 93/146 | 0.085 | 39.474 |

| | | | | | |
|---|---|---|---|---|---|
| *656* | hsa–miR–139–5p | PlasmaExo–Lab 4 | 35/36 | 0.085 | 312.697 |
| *657* | hsa–miR–627–5p | Serum–Lab 5 | 30/51 | 0.083 | 104.073 |
| *658* | hsa–miR–627–5p | Plasma–Lab 5 | 45/109 | 0.060 | 42.458 |
| *659* | hsa–miR–708–3p | CSF–Lab 3 | 14/25 | 0.083 | 15.208 |
| *660* | hsa–miR–6775–3p | Plasma–Lab 1 | 9/20 | 0.082 | 7.765 |
| *661* | hsa–miR–16–2–3p | Urine–Lab 4 | 14/23 | 0.082 | 24.104 |
| *662* | hsa–miR–671–5p | Plasma–Lab 5 | 5/12 | 0.082 | 13.983 |
| *663* | hsa–miR–671–5p | Serum–Lab 5 | 6/20 | 0.068 | 13.880 |
| *664* | hsa–miR–29a–3p | PlasmaExo–Lab 2 | 50/50 | 0.078 | 3519.200 |
| *665* | hsa–miR–29a–3p | PlasmaExo–Lab 4 | 70/76 | 0.076 | 78.122 |
| *666* | hsa–miR–29a–3p | Urine–Lab 1 | 17/31 | 0.075 | 54.155 |
| *667* | hsa–miR–29a–3p | CSF–Lab 3 | 33/35 | 0.064 | 352.517 |
| *668* | hsa–miR–29a–3p | Serum–Lab 3 | 26/35 | 0.054 | 341.220 |
| *669* | hsa–miR–25–5p | Plasma–Lab 1 | 79/124 | 0.077 | 37.504 |
| *670* | hsa–miR–486–5p | Serum–Lab 5 | 34/59 | 0.077 | 27.341 |
| *671* | hsa–miR–486–5p | Plasma–Lab 5 | 129/177 | 0.072 | 70.541 |
| *672* | hsa–miR–486–5p | Serum–Lab 3 | 51/51 | 0.058 | 3699.447 |
| *673* | hsa–miR–598–3p | Plasma–Lab 1 | 8/34 | 0.076 | 12.974 |
| *674* | hsa–miR–98–5p | Plasma–Lab 1 | 151/159 | 0.074 | 2723.174 |
| *675* | hsa–miR–98–5p | Serum–Lab 5 | 65/72 | 0.073 | 301.049 |
| *676* | hsa–miR–361–5p | Plasma–Lab 5 | 17/65 | 0.074 | 261.197 |
| *677* | hsa–miR–380–5p | PlasmaExo–Lab 4 | 6/12 | 0.073 | 7.625 |
| *678* | hsa–miR–190a–3p | Plasma–Lab 5 | 8/28 | 0.073 | 6.761 |
| *679* | hsa–let–7b–5p | Plasma–Lab 1 | 74/76 | 0.072 | 4363.834 |
| *680* | hsa–let–7b–5p | Urine–Lab 1 | 96/109 | 0.059 | 1030.483 |
| *681* | hsa–miR–205–5p | Urine–Lab 1 | 38/67 | 0.071 | 80.461 |
| *682* | hsa–miR–337–3p | Plasma–Lab 1 | 16/31 | 0.071 | 14.197 |
| *683* | hsa–miR–337–3p | Plasma–Lab 5 | 19/50 | 0.064 | 12.798 |
| *684* | hsa–miR–337–3p | Serum–Lab 5 | 18/51 | 0.052 | 29.369 |
| *685* | hsa–miR–431–5p | Serum–Lab 5 | 16/39 | 0.070 | 21.015 |
| *686* | hsa–miR–431–5p | Plasma–Lab 5 | 5/28 | 0.055 | 7.786 |
| *687* | hsa–miR–6847–5p | Plasma–Lab 1 | 6/18 | 0.070 | 7.072 |
| *688* | hsa–miR–194–5p | Serum–Lab 5 | 59/77 | 0.068 | 151.779 |
| *689* | hsa–miR–194–5p | Plasma–Lab 5 | 121/185 | 0.052 | 130.209 |
| *690* | hsa–miR–181a–3p | Plasma–Lab 1 | 138/157 | 0.067 | 150.768 |
| *691* | hsa–miR–4685–3p | Serum–Lab 5 | 7/20 | 0.067 | 10.790 |
| *692* | hsa–miR–4685–3p | Serum–Lab 3 | 6/17 | 0.060 | 27.341 |
| *693* | hsa–miR–363–3p | Urine–Lab 1 | 19/78 | 0.066 | 65.082 |
| *694* | hsa–miR–1294 | Plasma–Lab 1 | 10/26 | 0.066 | 12.531 |

| | | | | | |
|---|---|---|---|---|---|
| *695* | hsa–miR–195–5p | Serum–Lab 5 | 13/36 | 0.066 | 23.403 |
| *696* | hsa–miR–223–5p | Serum–Lab 5 | 16/27 | 0.064 | 130.867 |
| *697* | hsa–miR–100–5p | Urine–Lab 4 | 27/27 | 0.064 | 4901.448 |
| *698* | hsa–miR–15a–5p | PlasmaExo–Lab 4 | 13/77 | 0.063 | 21.399 |
| *699* | chr12_6918 | Plasma–Lab 1 | 8/19 | 0.063 | 7.984 |
| *700* | hsa–miR–454–5p | Plasma–Lab 5 | 5/27 | 0.063 | 10.611 |
| *701* | hsa–miR–193a–5p | Urine–Lab 4 | 14/25 | 0.060 | 35.396 |
| *702* | hsa–miR–193a–5p | Serum–Lab 3 | 5/10 | 0.060 | 28.860 |
| *703* | chr19_13717 | PlasmaExo–Lab 4 | 12/41 | 0.060 | 18.339 |
| *704* | hsa–miR–139–3p | PlasmaExo–Lab 4 | 52/62 | 0.060 | 187.208 |
| *705* | hsa–miR–30e–5p | Urine–Lab 1 | 108/110 | 0.058 | 656.218 |
| *706* | hsa–miR–30e–5p | Plasma–Lab 1 | 166/167 | 0.050 | 17172.875 |
| *707* | hsa–miR–2355–5p | Plasma–Lab 1 | 6/16 | 0.058 | 20.200 |
| *708* | hsa–miR–425–5p | PlasmaExo–Lab 2 | 17/17 | 0.057 | 219.594 |
| *709* | hsa–miR–27b–5p | PlasmaExo–Lab 4 | 9/17 | 0.057 | 17.541 |
| *710* | hsa–miR–27b–5p | Plasma–Lab 1 | 91/143 | 0.055 | 48.357 |
| *711* | hsa–miR–27b–5p | Serum–Lab 3 | 9/20 | 0.051 | 27.085 |
| *712* | hsa–miR–181a–2–3p | Plasma–Lab 5 | 50/99 | 0.057 | 83.442 |
| *713* | hsa–miR–30b–5p | Serum–Lab 5 | 66/73 | 0.057 | 748.392 |
| *714* | hsa–miR–4746–5p | Plasma–Lab 1 | 28/65 | 0.056 | 13.945 |
| *715* | hsa–miR–4746–5p | PlasmaExo–Lab 4 | 17/40 | 0.051 | 12.703 |
| *716* | hsa–miR–1307–5p | Serum–Lab 5 | 7/14 | 0.056 | 1616.386 |
| *717* | hsa–miR–23b–5p | PlasmaExo–Lab 4 | 22/44 | 0.054 | 20.334 |
| *718* | hsa–miR–485–5p | Plasma–Lab 1 | 66/123 | 0.052 | 47.296 |

## Supplementary Table S2.2 miRNA 3' adenylation levels across biofluids.

| | miRNA | Fluid– Data set | N samples (modified/expressed) | Average 3' A (%) | RPM |
|---|---|---|---|---|---|
| 1 | hsa–miR–944 | CSF–Lab 3 | 7/7 | 0.692 | 9.514 |
| 2 | hsa–miR–1237–3p | Plasma–Lab 5 | 5/6 | 0.564 | 3.300 |
| 3 | hsa–miR–1237–3p | Plasma–Lab 1 | 6/6 | 0.443 | 4.833 |
| 4 | hsa–miR–6775–3p | Plasma–Lab 1 | 18/20 | 0.547 | 7.765 |
| 5 | chr2_27505 | PlasmaExo–Lab 4 | 6/7 | 0.524 | 4.986 |
| 6 | hsa–miR–205–5p | Urine–Lab 4 | 22/22 | 0.494 | 295.800 |
| 7 | hsa–miR–205–5p | PlasmaExo–Lab 4 | 6/9 | 0.210 | 34.467 |
| 8 | hsa–miR–205–5p | Urine–Lab 1 | 31/67 | 0.075 | 80.461 |
| 9 | hsa–miR–205–5p | Plasma–Lab 1 | 10/20 | 0.068 | 28.255 |
| 10 | hsa–miR–7158–3p | CSF–Lab 3 | 8/8 | 0.490 | 10.113 |
| 11 | hsa–miR–6513–3p | Plasma–Lab 1 | 26/26 | 0.479 | 7.565 |
| 12 | hsa–miR–6513–3p | Serum–Lab 5 | 11/18 | 0.260 | 7.739 |
| 13 | hsa–miR–6513–3p | Plasma–Lab 5 | 5/16 | 0.084 | 10.387 |
| 14 | hsa–miR–141–3p | Urine–Lab 4 | 8/10 | 0.460 | 17.000 |
| 15 | hsa–miR–503–5p | Plasma–Lab 5 | 59/97 | 0.431 | 29.386 |
| 16 | hsa–miR–503–5p | PlasmaExo–Lab 2 | 10/11 | 0.417 | 7.182 |
| 17 | hsa–miR–503–5p | PlasmaExo–Lab 4 | 89/89 | 0.411 | 31.138 |
| 18 | hsa–miR–503–5p | Urine–Lab 4 | 7/7 | 0.248 | 12.700 |
| 19 | hsa–miR–6772–3p | Plasma–Lab 1 | 96/97 | 0.419 | 26.203 |
| 20 | hsa–miR–103a–2–5p | Plasma–Lab 5 | 43/73 | 0.413 | 16.807 |
| 21 | hsa–miR–103a–2–5p | Serum–Lab 5 | 11/24 | 0.344 | 16.408 |
| 22 | hsa–miR–99b–3p | Urine–Lab 1 | 11/16 | 0.395 | 9.581 |
| 23 | hsa–miR–99b–3p | Plasma–Lab 1 | 6/12 | 0.215 | 20.175 |
| 24 | hsa–miR–99b–3p | Urine–Lab 4 | 22/22 | 0.198 | 68.368 |
| 25 | hsa–miR–99b–3p | PlasmaExo–Lab 4 | 15/20 | 0.148 | 34.030 |
| 26 | hsa–miR–99b–3p | PlasmaExo–Lab 2 | 44/44 | 0.125 | 341.045 |
| 27 | hsa–miR–99b–3p | Serum–Lab 5 | 8/21 | 0.118 | 23.057 |
| 28 | hsa–miR–99b–3p | Plasma–Lab 5 | 9/24 | 0.116 | 17.438 |
| 29 | hsa–miR–140–5p | PlasmaExo–Lab 2 | 17/26 | 0.391 | 10.977 |
| 30 | hsa–miR–1911–3p | CSF–Lab 3 | 38/42 | 0.384 | 11.133 |
| 31 | hsa–miR–152–3p | Urine–Lab 4 | 25/25 | 0.379 | 221.236 |
| 32 | hsa–miR–152–3p | PlasmaExo–Lab 2 | 50/50 | 0.145 | 414.958 |
| 33 | hsa–miR–152–3p | PlasmaExo–Lab 4 | 45/66 | 0.128 | 101.847 |
| 34 | hsa–miR–183–5p | Serum–Lab 3 | 23/31 | 0.370 | 133.535 |
| 35 | hsa–miR–183–5p | Plasma–Lab 1 | 29/50 | 0.270 | 55.072 |
| 36 | hsa–miR–183–5p | CSF–Lab 3 | 6/17 | 0.108 | 56.282 |
| 37 | hsa–miR–183–5p | Plasma–Lab 5 | 13/112 | 0.050 | 153.373 |

| | | | | | |
|---|---|---|---|---|---|
| *38* | hsa–miR–324–3p | PlasmaExo–Lab 2 | 50/50 | 0.366 | 41.914 |
| *39* | hsa–miR–324–3p | Plasma–Lab 5 | 126/164 | 0.302 | 69.713 |
| *40* | hsa–miR–324–3p | Serum–Lab 5 | 51/64 | 0.231 | 64.717 |
| *41* | hsa–miR–324–3p | PlasmaExo–Lab 4 | 35/46 | 0.136 | 12.957 |
| *42* | hsa–miR–28–5p | Urine–Lab 1 | 55/58 | 0.363 | 23.652 |
| *43* | hsa–miR–28–5p | Plasma–Lab 1 | 153/158 | 0.180 | 389.855 |
| *44* | hsa–miR–28–5p | Serum–Lab 3 | 20/32 | 0.168 | 40.422 |
| *45* | hsa–miR–28–5p | CSF–Lab 3 | 21/38 | 0.153 | 18.361 |
| *46* | hsa–miR–28–5p | PlasmaExo–Lab 2 | 20/49 | 0.074 | 23.386 |
| *47* | hsa–miR–28–5p | Serum–Lab 5 | 15/71 | 0.054 | 159.197 |
| *48* | hsa–miR–320e | Plasma–Lab 5 | 16/19 | 0.352 | 58.758 |
| *49* | hsa–miR–320e | Serum–Lab 5 | 10/16 | 0.295 | 59.737 |
| *50* | hsa–miR–1226–3p | Plasma–Lab 1 | 16/18 | 0.329 | 6.561 |
| *51* | hsa–miR–1226–3p | Serum–Lab 5 | 21/30 | 0.183 | 12.980 |
| *52* | hsa–miR–1226–3p | Plasma–Lab 5 | 7/17 | 0.113 | 8.506 |
| *53* | hsa–miR–150–5p | CSF–Lab 3 | 21/22 | 0.327 | 30.064 |
| *54* | hsa–miR–150–5p | Serum–Lab 3 | 39/41 | 0.240 | 238.924 |
| *55* | hsa–miR–150–5p | Serum–Lab 5 | 75/75 | 0.186 | 1610.444 |
| *56* | hsa–miR–150–5p | Plasma–Lab 5 | 179/182 | 0.183 | 1833.920 |
| *57* | hsa–miR–150–5p | Plasma–Lab 1 | 157/165 | 0.151 | 527.305 |
| *58* | hsa–miR–3620–3p | Serum–Lab 5 | 7/10 | 0.324 | 6.740 |
| *59* | hsa–miR–3620–3p | Plasma–Lab 1 | 36/41 | 0.220 | 10.366 |
| *60* | hsa–miR–6842–3p | CSF–Lab 3 | 13/15 | 0.324 | 6.373 |
| *61* | hsa–miR–6842–3p | Serum–Lab 3 | 13/16 | 0.275 | 30.200 |
| *62* | hsa–miR–6842–3p | Plasma–Lab 1 | 145/156 | 0.151 | 114.016 |
| *63* | hsa–miR–6842–3p | Serum–Lab 5 | 9/29 | 0.056 | 11.110 |
| *64* | hsa–miR–29c–3p | CSF–Lab 3 | 21/25 | 0.309 | 108.464 |
| *65* | hsa–miR–29c–3p | Urine–Lab 1 | 41/48 | 0.307 | 45.554 |
| *66* | hsa–miR–29c–3p | Plasma–Lab 1 | 61/72 | 0.142 | 80.969 |
| *67* | hsa–miR–33a–5p | PlasmaExo–Lab 2 | 47/48 | 0.305 | 20.744 |
| *68* | hsa–miR–6819–3p | Serum–Lab 5 | 15/20 | 0.301 | 7.805 |
| *69* | hsa–miR–6819–3p | Plasma–Lab 1 | 122/125 | 0.287 | 45.350 |
| *70* | hsa–miR–1229–3p | Plasma–Lab 1 | 15/18 | 0.294 | 7.928 |
| *71* | hsa–miR–1229–3p | Serum–Lab 5 | 8/19 | 0.187 | 11.968 |
| *72* | hsa–miR–1227–3p | Plasma–Lab 5 | 7/9 | 0.293 | 5.889 |
| *73* | hsa–miR–636 | Plasma–Lab 1 | 28/31 | 0.292 | 24.484 |
| *74* | hsa–miR–636 | Plasma–Lab 5 | 7/58 | 0.058 | 11.066 |
| *75* | hsa–miR–125b–2–3p | Urine–Lab 4 | 5/5 | 0.263 | 52.160 |
| *76* | hsa–miR–125b–2–3p | PlasmaExo–Lab 4 | 15/22 | 0.162 | 8.223 |

| | | | | | |
|---|---|---|---|---|---|
| *77* | hsa–miR–532–5p | Serum–Lab 3 | 26/29 | 0.256 | 532.152 |
| *78* | hsa–miR–532–5p | Plasma–Lab 1 | 65/69 | 0.209 | 401.330 |
| *79* | hsa–miR–532–5p | Urine–Lab 1 | 95/104 | 0.095 | 124.988 |
| *80* | hsa–miR–532–5p | Plasma–Lab 5 | 166/178 | 0.060 | 422.170 |
| *81* | hsa–miR–28–3p | Plasma–Lab 5 | 81/98 | 0.246 | 150.942 |
| *82* | hsa–miR–28–3p | Serum–Lab 5 | 37/49 | 0.211 | 325.761 |
| *83* | hsa–miR–28–3p | Urine–Lab 4 | 24/24 | 0.132 | 214.171 |
| *84* | hsa–miR–28–3p | PlasmaExo–Lab 4 | 25/27 | 0.103 | 580.137 |
| *85* | hsa–miR–27a–3p | PlasmaExo–Lab 2 | 50/50 | 0.245 | 954.440 |
| *86* | hsa–miR–27a–3p | Urine–Lab 4 | 21/22 | 0.203 | 448.964 |
| *87* | hsa–miR–27a–3p | PlasmaExo–Lab 4 | 101/101 | 0.184 | 848.982 |
| *88* | hsa–miR–27a–3p | Serum–Lab 5 | 76/76 | 0.107 | 10099.964 |
| *89* | hsa–miR–27a–3p | Plasma–Lab 5 | 181/186 | 0.101 | 4156.726 |
| *90* | hsa–miR–3188 | PlasmaExo–Lab 4 | 6/6 | 0.241 | 7.800 |
| *91* | hsa–miR–3188 | Plasma–Lab 1 | 9/12 | 0.201 | 7.167 |
| *92* | hsa–miR–200a–5p | Urine–Lab 4 | 27/27 | 0.237 | 324.463 |
| *93* | hsa–miR–134–5p | CSF–Lab 3 | 40/42 | 0.236 | 32.148 |
| *94* | hsa–miR–134–5p | Serum–Lab 3 | 33/35 | 0.155 | 65.934 |
| *95* | hsa–miR–134–5p | Plasma–Lab 5 | 61/115 | 0.103 | 32.126 |
| *96* | hsa–miR–134–5p | Serum–Lab 5 | 43/68 | 0.094 | 67.049 |
| *97* | hsa–miR–134–5p | Plasma–Lab 1 | 145/160 | 0.081 | 327.274 |
| *98* | hsa–miR–3120–3p | Serum–Lab 5 | 7/12 | 0.235 | 10.358 |
| *99* | hsa–miR–3120–3p | PlasmaExo–Lab 4 | 6/8 | 0.125 | 17.113 |
| *100* | hsa–miR–483–5p | CSF–Lab 3 | 32/33 | 0.230 | 55.394 |
| *101* | hsa–miR–483–5p | Serum–Lab 3 | 24/26 | 0.205 | 58.338 |
| *102* | hsa–miR–483–5p | Plasma–Lab 1 | 11/18 | 0.169 | 23.122 |
| *103* | hsa–miR–483–5p | Serum–Lab 5 | 11/21 | 0.093 | 56.490 |
| *104* | hsa–miR–483–5p | Plasma–Lab 5 | 14/27 | 0.092 | 41.196 |
| *105* | hsa–miR–671–5p | Plasma–Lab 5 | 8/12 | 0.227 | 13.983 |
| *106* | hsa–miR–671–5p | Serum–Lab 5 | 11/20 | 0.123 | 13.880 |
| *107* | hsa–miR–361–5p | Urine–Lab 1 | 33/44 | 0.225 | 19.102 |
| *108* | hsa–miR–361–5p | Plasma–Lab 5 | 46/65 | 0.095 | 261.197 |
| *109* | hsa–miR–361–5p | Plasma–Lab 1 | 109/135 | 0.057 | 169.440 |
| *110* | chr11_4033 | Plasma–Lab 1 | 8/10 | 0.221 | 6.090 |
| *111* | hsa–miR–190b–5p | Serum–Lab 5 | 21/31 | 0.219 | 22.900 |
| *112* | hsa–miR–190b–5p | CSF–Lab 3 | 20/28 | 0.173 | 12.296 |
| *113* | hsa–miR–190b–5p | Serum–Lab 3 | 6/8 | 0.162 | 24.012 |
| *114* | hsa–miR–190b–5p | Plasma–Lab 5 | 27/65 | 0.110 | 12.568 |

| | | | | | |
|---|---|---|---|---|---|
| *115* | hsa–miR–190b–5p | Plasma–Lab 1 | 13/26 | 0.092 | 10.142 |
| *116* | hsa–miR–652–3p | CSF–Lab 3 | 58/58 | 0.216 | 37.293 |
| *117* | hsa–miR–652–3p | Plasma–Lab 5 | 122/157 | 0.197 | 275.629 |
| *118* | hsa–miR–652–3p | Serum–Lab 5 | 58/70 | 0.132 | 604.350 |
| *119* | hsa–miR–652–3p | Urine–Lab 1 | 8/16 | 0.098 | 20.644 |
| *120* | hsa–miR–652–3p | Urine–Lab 4 | 10/17 | 0.094 | 17.294 |
| *121* | hsa–miR–652–3p | Plasma–Lab 1 | 89/134 | 0.071 | 393.093 |
| *122* | hsa–miR–652–3p | PlasmaExo–Lab 2 | 39/39 | 0.059 | 154.736 |
| *123* | hsa–miR–187–3p | Urine–Lab 4 | 5/6 | 0.213 | 14.533 |
| *124* | hsa–miR–660–5p | Urine–Lab 1 | 92/101 | 0.212 | 90.197 |
| *125* | hsa–miR–660–5p | Plasma–Lab 1 | 115/150 | 0.066 | 78.795 |
| *126* | hsa–miR–660–5p | CSF–Lab 3 | 29/46 | 0.053 | 69.776 |
| *127* | hsa–miR–126–3p | Plasma–Lab 5 | 187/187 | 0.200 | 11605.440 |
| *128* | hsa–miR–126–3p | Serum–Lab 5 | 76/77 | 0.147 | 19866.526 |
| *129* | hsa–miR–126–3p | PlasmaExo–Lab 2 | 50/50 | 0.096 | 1490.820 |
| *130* | hsa–miR–126–3p | CSF–Lab 3 | 30/43 | 0.060 | 141.023 |
| *131* | hsa–miR–6747–3p | Plasma–Lab 5 | 7/13 | 0.196 | 4.592 |
| *132* | hsa–miR–4511 | Plasma–Lab 5 | 5/10 | 0.194 | 6.100 |
| *133* | hsa–miR–1908–3p | Plasma–Lab 1 | 69/81 | 0.193 | 16.988 |
| *134* | hsa–miR–543 | Plasma–Lab 1 | 35/44 | 0.189 | 23.793 |
| *135* | hsa–miR–543 | Serum–Lab 5 | 17/38 | 0.096 | 38.245 |
| *136* | hsa–miR–543 | Plasma–Lab 5 | 8/32 | 0.064 | 11.556 |
| *137* | hsa–miR–324–5p | CSF–Lab 3 | 21/25 | 0.187 | 11.816 |
| *138* | hsa–miR–324–5p | Plasma–Lab 5 | 24/59 | 0.062 | 91.203 |
| *139* | hsa–miR–324–5p | Serum–Lab 3 | 13/25 | 0.056 | 32.312 |
| *140* | hsa–miR–324–5p | Plasma–Lab 1 | 37/85 | 0.051 | 31.729 |
| *141* | hsa–let–7i–5p | Urine–Lab 1 | 110/110 | 0.182 | 1046.350 |
| *142* | hsa–let–7i–5p | CSF–Lab 3 | 64/64 | 0.152 | 2288.272 |
| *143* | hsa–let–7i–5p | Serum–Lab 3 | 33/33 | 0.146 | 3619.197 |
| *144* | hsa–let–7i–5p | Serum–Lab 5 | 61/70 | 0.144 | 5764.631 |
| *145* | hsa–let–7i–5p | Plasma–Lab 1 | 164/164 | 0.092 | 16434.757 |
| *146* | hsa–let–7i–5p | Plasma–Lab 5 | 49/53 | 0.084 | 6840.147 |
| *147* | hsa–let–7i–5p | Urine–Lab 4 | 27/27 | 0.079 | 4726.526 |
| *148* | hsa–miR–320d | Plasma–Lab 5 | 5/9 | 0.175 | 11.556 |
| *149* | hsa–miR–181a–2–3p | Plasma–Lab 5 | 64/99 | 0.174 | 83.442 |
| *150* | hsa–miR–181a–2–3p | Serum–Lab 5 | 10/33 | 0.114 | 63.333 |
| *151* | hsa–miR–122–5p | Urine–Lab 1 | 9/11 | 0.173 | 20.055 |
| *152* | hsa–miR–122–5p | Plasma–Lab 5 | 139/150 | 0.148 | 3976.959 |
| *153* | hsa–miR–122–5p | Plasma–Lab 1 | 143/156 | 0.148 | 474.339 |

| | | | | | |
|---|---|---|---|---|---|
| *154* | hsa–miR–122–5p | Serum–Lab 5 | 51/55 | 0.135 | 4445.575 |
| *155* | hsa–miR–122–5p | Serum–Lab 3 | 25/43 | 0.078 | 12625.272 |
| *156* | hsa–miR–122–5p | CSF–Lab 3 | 8/25 | 0.067 | 87.664 |
| *157* | hsa–miR–4690–3p | Plasma–Lab 5 | 7/14 | 0.171 | 6.607 |
| *158* | hsa–miR–486–5p | Serum–Lab 5 | 48/59 | 0.170 | 27.341 |
| *159* | hsa–miR–486–5p | Plasma–Lab 5 | 147/177 | 0.140 | 70.541 |
| *160* | hsa–miR–486–5p | Serum–Lab 3 | 50/51 | 0.128 | 3699.447 |
| *161* | hsa–miR–486–5p | CSF–Lab 3 | 31/42 | 0.089 | 70.095 |
| *162* | hsa–miR–486–5p | Plasma–Lab 1 | 155/160 | 0.053 | 2282.833 |
| *163* | hsa–miR–939–5p | Plasma–Lab 1 | 34/48 | 0.168 | 12.196 |
| *164* | hsa–miR–939–5p | Serum–Lab 5 | 6/13 | 0.139 | 5.108 |
| *165* | hsa–miR–939–5p | Plasma–Lab 5 | 6/20 | 0.094 | 7.705 |
| *166* | hsa–miR–99b–5p | Urine–Lab 1 | 109/110 | 0.164 | 211.171 |
| *167* | hsa–miR–99b–5p | Plasma–Lab 1 | 143/151 | 0.098 | 986.903 |
| *168* | hsa–miR–99b–5p | Serum–Lab 5 | 42/50 | 0.084 | 517.762 |
| *169* | hsa–miR–99b–5p | Plasma–Lab 5 | 32/66 | 0.057 | 451.845 |
| *170* | hsa–miR–339–5p | Plasma–Lab 1 | 81/101 | 0.162 | 27.419 |
| *171* | hsa–miR–339–5p | Serum–Lab 5 | 55/72 | 0.082 | 159.414 |
| *172* | hsa–miR–339–5p | PlasmaExo–Lab 4 | 48/73 | 0.072 | 56.318 |
| *173* | hsa–miR–339–5p | PlasmaExo–Lab 2 | 31/50 | 0.067 | 14.304 |
| *174* | hsa–miR–140–3p | Plasma–Lab 5 | 157/157 | 0.156 | 4224.594 |
| *175* | hsa–miR–140–3p | Serum–Lab 5 | 11/13 | 0.138 | 2891.723 |
| *176* | hsa–miR–223–5p | Urine–Lab 4 | 7/11 | 0.154 | 59.164 |
| *177* | hsa–miR–223–5p | PlasmaExo–Lab 2 | 40/43 | 0.097 | 33.128 |
| *178* | hsa–miR–223–5p | Serum–Lab 5 | 15/27 | 0.070 | 130.867 |
| *179* | hsa–miR–223–5p | PlasmaExo–Lab 4 | 32/34 | 0.061 | 227.915 |
| *180* | hsa–miR–223–5p | Plasma–Lab 5 | 18/60 | 0.056 | 39.223 |
| *181* | hsa–miR–99a–5p | Urine–Lab 4 | 13/13 | 0.154 | 4705.154 |
| *182* | hsa–miR–99a–5p | Urine–Lab 1 | 57/107 | 0.051 | 75.244 |
| *183* | chr11_2103 | Serum–Lab 5 | 5/13 | 0.151 | 8.069 |
| *184* | hsa–miR–4433b–5p | Serum–Lab 5 | 24/31 | 0.149 | 150.726 |
| *185* | hsa–miR–4433b–5p | Plasma–Lab 5 | 20/50 | 0.100 | 27.648 |
| *186* | hsa–miR–4433b–5p | Plasma–Lab 1 | 29/39 | 0.081 | 315.482 |
| *187* | hsa–miR–3200–3p | Serum–Lab 5 | 16/34 | 0.145 | 16.421 |
| *188* | hsa–miR–3200–3p | Plasma–Lab 5 | 68/125 | 0.116 | 32.320 |
| *189* | hsa–miR–3200–3p | CSF–Lab 3 | 13/25 | 0.060 | 18.796 |
| *190* | hsa–miR–217–5p | CSF–Lab 3 | 5/7 | 0.143 | 3.986 |
| *191* | hsa–miR–4672 | Plasma–Lab 5 | 5/13 | 0.140 | 4.915 |
| *192* | hsa–miR–576–5p | Serum–Lab 5 | 20/42 | 0.139 | 14.062 |

| | | | | | |
|---|---|---|---|---|---|
| *193* | hsa–miR–576–5p | Plasma–Lab 5 | 30/71 | 0.118 | 13.076 |
| *194* | hsa–miR–181c–3p | Serum–Lab 5 | 32/48 | 0.136 | 39.498 |
| *195* | hsa–miR–181c–3p | Plasma–Lab 5 | 27/78 | 0.066 | 19.024 |
| *196* | hsa–miR–887–3p | Plasma–Lab 5 | 6/17 | 0.136 | 6.729 |
| *197* | hsa–miR–3176 | Plasma–Lab 5 | 10/23 | 0.131 | 8.717 |
| *198* | hsa–miR–2277–3p | Serum–Lab 5 | 10/15 | 0.130 | 9.320 |
| *199* | hsa–miR–3615 | Urine–Lab 1 | 14/32 | 0.129 | 19.859 |
| *200* | hsa–miR–3615 | CSF–Lab 3 | 10/15 | 0.116 | 12.887 |
| *201* | hsa–miR–3615 | Plasma–Lab 1 | 53/65 | 0.083 | 297.682 |
| *202* | hsa–miR–3615 | Serum–Lab 3 | 7/15 | 0.057 | 144.233 |
| *203* | hsa–miR–92b–3p | Urine–Lab 1 | 89/96 | 0.125 | 258.889 |
| *204* | hsa–miR–92b–3p | Plasma–Lab 5 | 120/144 | 0.056 | 194.288 |
| *205* | hsa–miR–92b–3p | CSF–Lab 3 | 57/57 | 0.054 | 548.981 |
| *206* | hsa–miR–3661 | Plasma–Lab 1 | 9/16 | 0.123 | 7.100 |
| *207* | hsa–miR–345–5p | Serum–Lab 5 | 36/47 | 0.122 | 222.855 |
| *208* | hsa–miR–379–5p | Urine–Lab 4 | 5/8 | 0.121 | 13.075 |
| *209* | hsa–miR–95–3p | Serum–Lab 5 | 12/34 | 0.118 | 24.924 |
| *210* | hsa–miR–95–3p | Plasma–Lab 5 | 20/65 | 0.071 | 22.025 |
| *211* | hsa–let–7e–5p | Plasma–Lab 5 | 81/116 | 0.115 | 36.090 |
| *212* | hsa–let–7e–5p | Urine–Lab 1 | 91/97 | 0.100 | 507.657 |
| *213* | hsa–let–7e–5p | Serum–Lab 5 | 52/71 | 0.098 | 107.593 |
| *214* | hsa–let–7e–5p | Plasma–Lab 1 | 149/158 | 0.072 | 819.259 |
| *215* | hsa–let–7e–5p | CSF–Lab 3 | 31/41 | 0.066 | 209.256 |
| *216* | hsa–let–7e–5p | Urine–Lab 4 | 14/15 | 0.061 | 767.807 |
| *217* | hsa–miR–483–3p | Plasma–Lab 5 | 11/27 | 0.114 | 13.667 |
| *218* | hsa–miR–6726–3p | Serum–Lab 5 | 6/12 | 0.114 | 8.350 |
| *219* | hsa–miR–374a–5p | Serum–Lab 5 | 50/54 | 0.113 | 1299.131 |
| *220* | hsa–miR–26b–5p | Urine–Lab 1 | 106/110 | 0.110 | 451.201 |
| *221* | hsa–miR–26b–5p | CSF–Lab 3 | 49/55 | 0.056 | 376.329 |
| *222* | hsa–miR–126–5p | Plasma–Lab 5 | 179/180 | 0.108 | 5891.588 |
| *223* | hsa–miR–126–5p | Serum–Lab 5 | 74/74 | 0.101 | 9265.234 |
| *224* | hsa–miR–29c–5p | Plasma–Lab 5 | 22/60 | 0.107 | 25.413 |
| *225* | hsa–miR–29c–5p | Serum–Lab 5 | 8/22 | 0.060 | 30.882 |
| *226* | hsa–miR–582–3p | Plasma–Lab 5 | 10/34 | 0.107 | 15.703 |
| *227* | hsa–miR–100–5p | Urine–Lab 1 | 37/98 | 0.106 | 64.646 |
| *228* | hsa–miR–421 | Urine–Lab 1 | 21/46 | 0.106 | 26.254 |
| *229* | hsa–miR–421 | CSF–Lab 3 | 39/48 | 0.101 | 46.431 |
| *230* | hsa–miR–421 | Plasma–Lab 5 | 56/110 | 0.090 | 27.680 |

| | | | | | |
|---|---|---|---|---|---|
| *231* | hsa–miR–421 | Serum–Lab 3 | 29/36 | 0.074 | 84.264 |
| *232* | hsa–miR–421 | Serum–Lab 5 | 16/40 | 0.073 | 36.167 |
| *233* | hsa–miR–206 | Urine–Lab 4 | 15/21 | 0.105 | 28.081 |
| *234* | hsa–miR–589–3p | Plasma–Lab 1 | 53/80 | 0.105 | 20.676 |
| *235* | hsa–miR–374b–5p | Serum–Lab 5 | 30/40 | 0.104 | 476.295 |
| *236* | hsa–miR–33b–5p | Plasma–Lab 5 | 41/103 | 0.102 | 57.363 |
| *237* | hsa–miR–155–5p | Plasma–Lab 5 | 11/25 | 0.102 | 40.856 |
| *238* | hsa–miR–145–3p | Serum–Lab 5 | 32/55 | 0.101 | 88.013 |
| *239* | hsa–miR–5189–5p | Plasma–Lab 1 | 7/15 | 0.100 | 8.680 |
| *240* | hsa–miR–210–3p | Urine–Lab 4 | 18/27 | 0.099 | 37.726 |
| *241* | hsa–miR–210–3p | PlasmaExo–Lab 2 | 15/17 | 0.096 | 61.482 |
| *242* | hsa–miR–1296–5p | Serum–Lab 5 | 8/20 | 0.099 | 12.805 |
| *243* | hsa–miR–580–3p | Plasma–Lab 1 | 11/26 | 0.099 | 8.885 |
| *244* | hsa–miR–194–5p | Urine–Lab 4 | 19/25 | 0.096 | 37.208 |
| *245* | hsa–miR–194–5p | Serum–Lab 5 | 54/77 | 0.057 | 151.779 |
| *246* | hsa–miR–194–5p | Plasma–Lab 5 | 131/185 | 0.050 | 130.209 |
| *247* | hsa–miR–509–3–5p | Urine–Lab 4 | 5/11 | 0.095 | 34.300 |
| *248* | hsa–miR–200b–5p | Urine–Lab 4 | 26/27 | 0.094 | 227.044 |
| *249* | hsa–miR–185–3p | Plasma–Lab 5 | 113/179 | 0.094 | 68.554 |
| *250* | hsa–miR–487a–3p | Serum–Lab 5 | 7/17 | 0.094 | 14.094 |
| *251* | hsa–miR–195–5p | Serum–Lab 5 | 20/36 | 0.092 | 23.403 |
| *252* | hsa–miR–195–5p | Plasma–Lab 5 | 26/72 | 0.068 | 21.139 |
| *253* | hsa–miR–185–5p | Urine–Lab 4 | 21/27 | 0.092 | 71.407 |
| *254* | hsa–miR–185–5p | Serum–Lab 5 | 63/74 | 0.064 | 1169.324 |
| *255* | hsa–miR–328–3p | Urine–Lab 1 | 33/54 | 0.090 | 46.857 |
| *256* | hsa–miR–328–3p | Plasma–Lab 5 | 92/171 | 0.054 | 128.891 |
| *257* | hsa–miR–548e–3p | Serum–Lab 5 | 11/44 | 0.090 | 26.630 |
| *258* | hsa–miR–142–3p | Plasma–Lab 5 | 130/138 | 0.086 | 1704.154 |
| *259* | hsa–miR–142–3p | Serum–Lab 5 | 50/60 | 0.079 | 4908.595 |
| *260* | hsa–miR–3187–3p | PlasmaExo–Lab 4 | 20/36 | 0.086 | 8.547 |
| *261* | hsa–miR–125b–1–3p | CSF–Lab 3 | 56/63 | 0.085 | 80.438 |
| *262* | hsa–miR–125b–1–3p | Plasma–Lab 5 | 10/19 | 0.082 | 11.137 |
| *263* | hsa–miR–125b–1–3p | Serum–Lab 3 | 20/35 | 0.072 | 50.771 |
| *264* | hsa–let–7g–5p | Urine–Lab 1 | 101/110 | 0.085 | 253.794 |
| *265* | hsa–let–7g–5p | Serum–Lab 5 | 63/71 | 0.061 | 3462.300 |
| *266* | hsa–let–7g–5p | CSF–Lab 3 | 35/50 | 0.058 | 543.854 |
| *267* | hsa–miR–204–3p | Urine–Lab 4 | 26/27 | 0.083 | 240.778 |
| *268* | hsa–miR–98–5p | Urine–Lab 1 | 102/110 | 0.082 | 244.036 |
| *269* | hsa–miR–486–3p | Serum–Lab 5 | 12/31 | 0.082 | 16.810 |

| | | | | | |
|---|---|---|---|---|---|
| *270* | hsa–miR–486–3p | Plasma–Lab 1 | 55/89 | 0.079 | 23.767 |
| *271* | hsa–miR–486–3p | Serum–Lab 3 | 5/13 | 0.059 | 14.708 |
| *272* | hsa–miR–486–3p | PlasmaExo–Lab 4 | 23/54 | 0.054 | 11.846 |
| *273* | hsa–miR–486–3p | Plasma–Lab 5 | 37/102 | 0.053 | 15.823 |
| *274* | hsa–miR–1298–5p | PlasmaExo–Lab 2 | 50/50 | 0.081 | 93.028 |
| *275* | hsa–miR–181a–3p | Plasma–Lab 5 | 12/61 | 0.081 | 47.339 |
| *276* | hsa–miR–584–5p | Urine–Lab 4 | 7/12 | 0.080 | 28.025 |
| *277* | hsa–miR–629–5p | Serum–Lab 5 | 17/34 | 0.080 | 117.118 |
| *278* | hsa–miR–139–5p | Plasma–Lab 5 | 21/64 | 0.078 | 66.044 |
| *279* | hsa–miR–7–1–3p | Serum–Lab 5 | 24/44 | 0.078 | 53.436 |
| *280* | hsa–miR–6741–5p | PlasmaExo–Lab 4 | 11/31 | 0.078 | 8.806 |
| *281* | hsa–miR–299–3p | Serum–Lab 5 | 9/29 | 0.077 | 14.562 |
| *282* | hsa–miR–34c–3p | CSF–Lab 3 | 54/59 | 0.077 | 90.439 |
| *283* | hsa–miR–1307–3p | Urine–Lab 4 | 26/26 | 0.076 | 236.100 |
| *284* | hsa–miR–744–5p | Plasma–Lab 5 | 45/68 | 0.076 | 198.534 |
| *285* | hsa–miR–744–5p | Serum–Lab 5 | 71/77 | 0.072 | 549.996 |
| *286* | hsa–miR–378i | Serum–Lab 3 | 5/11 | 0.076 | 12.227 |
| *287* | hsa–miR–671–3p | PlasmaExo–Lab 4 | 10/16 | 0.075 | 118.762 |
| *288* | hsa–miR–25–3p | Urine–Lab 4 | 20/22 | 0.075 | 145.173 |
| *289* | hsa–miR–92a–3p | Urine–Lab 1 | 106/110 | 0.073 | 738.359 |
| *290* | hsa–miR–339–3p | Serum–Lab 5 | 39/52 | 0.073 | 116.488 |
| *291* | hsa–miR–339–3p | Plasma–Lab 5 | 32/54 | 0.053 | 88.280 |
| *292* | hsa–miR–10a–5p | Urine–Lab 1 | 39/72 | 0.073 | 13248.493 |
| *293* | hsa–miR–10a–5p | Urine–Lab 4 | 15/15 | 0.053 | 14753.553 |
| *294* | hsa–miR–3605–3p | Plasma–Lab 1 | 48/85 | 0.071 | 28.527 |
| *295* | hsa–miR–539–3p | Plasma–Lab 1 | 8/15 | 0.071 | 20.887 |
| *296* | chr22_25085 | Serum–Lab 5 | 23/45 | 0.071 | 34.280 |
| *297* | hsa–miR–1247–5p | Plasma–Lab 5 | 23/56 | 0.070 | 15.496 |
| *298* | hsa–miR–4746–5p | Plasma–Lab 1 | 31/65 | 0.069 | 13.945 |
| *299* | hsa–miR–1911–5p | CSF–Lab 3 | 10/42 | 0.069 | 297.374 |
| *300* | hsa–miR–146a–5p | Urine–Lab 1 | 34/74 | 0.069 | 44.524 |
| *301* | hsa–miR–146a–5p | CSF–Lab 3 | 61/63 | 0.062 | 491.554 |
| *302* | hsa–miR–139–3p | Plasma–Lab 5 | 10/43 | 0.067 | 23.051 |
| *303* | hsa–miR–15a–5p | Urine–Lab 1 | 9/21 | 0.066 | 16.838 |
| *304* | hsa–miR–654–3p | Serum–Lab 5 | 38/65 | 0.065 | 86.065 |
| *305* | hsa–miR–200c–3p | PlasmaExo–Lab 4 | 49/74 | 0.064 | 69.384 |
| *306* | hsa–miR–374a–3p | Plasma–Lab 1 | 10/24 | 0.064 | 12.446 |
| *307* | hsa–miR–187–5p | PlasmaExo–Lab 2 | 20/43 | 0.063 | 9.519 |
| *308* | hsa–miR–1250–5p | Serum–Lab 5 | 6/20 | 0.063 | 10.985 |

| | | | | | |
|---|---|---|---|---|---|
| *309* | hsa–miR–627–5p | Serum–Lab 5 | 24/51 | 0.062 | 104.073 |
| *310* | hsa–miR–27b–3p | Plasma–Lab 5 | 63/119 | 0.062 | 1392.807 |
| *311* | hsa–miR–487b–3p | Plasma–Lab 1 | 89/125 | 0.062 | 71.062 |
| *312* | hsa–miR–432–5p | Plasma–Lab 5 | 13/52 | 0.060 | 20.944 |
| *313* | hsa–miR–432–5p | Serum–Lab 5 | 16/45 | 0.058 | 48.616 |
| *314* | hsa–miR–378a–5p | Plasma–Lab 5 | 26/143 | 0.060 | 44.132 |
| *315* | hsa–miR–570–3p | Plasma–Lab 5 | 14/49 | 0.058 | 9.743 |
| *316* | hsa–miR–29a–3p | Serum–Lab 5 | 14/37 | 0.058 | 8320.273 |
| *317* | hsa–miR–361–3p | PlasmaExo–Lab 4 | 37/39 | 0.057 | 365.554 |
| *318* | hsa–miR–143–3p | PlasmaExo–Lab 4 | 26/42 | 0.056 | 437.102 |
| *319* | hsa–miR–766–3p | Serum–Lab 5 | 38/65 | 0.056 | 47.283 |
| *320* | hsa–let–7d–5p | Serum–Lab 5 | 40/46 | 0.056 | 771.611 |
| *321* | hsa–miR–6852–5p | Plasma–Lab 1 | 134/148 | 0.056 | 275.755 |
| *322* | hsa–miR–6852–5p | Serum–Lab 5 | 9/24 | 0.051 | 19.817 |
| *323* | hsa–miR–708–5p | PlasmaExo–Lab 2 | 28/38 | 0.056 | 54.747 |
| *324* | hsa–miR–127–3p | Urine–Lab 4 | 14/23 | 0.055 | 55.917 |
| *325* | hsa–miR–370–3p | Serum–Lab 5 | 42/73 | 0.055 | 78.052 |
| *326* | hsa–miR–3613–5p | Serum–Lab 5 | 28/54 | 0.055 | 54.367 |
| *327* | hsa–miR–455–5p | Plasma–Lab 5 | 6/21 | 0.054 | 16.971 |
| *328* | hsa–let–7c–5p | Serum–Lab 5 | 46/66 | 0.054 | 96.370 |
| *329* | hsa–let–7c–5p | Plasma–Lab 1 | 124/155 | 0.051 | 137.545 |
| *330* | hsa–miR–769–5p | Plasma–Lab 5 | 26/69 | 0.053 | 97.620 |
| *331* | hsa–miR–149–5p | Urine–Lab 4 | 6/12 | 0.053 | 25.683 |
| *332* | hsa–miR–30e–3p | Plasma–Lab 5 | 91/134 | 0.052 | 98.450 |
| *333* | hsa–miR–1249–3p | Plasma–Lab 1 | 5/23 | 0.052 | 11.170 |
| *334* | hsa–miR–143–5p | Serum–Lab 5 | 30/67 | 0.052 | 67.201 |
| *335* | hsa–miR–651–5p | Serum–Lab 5 | 16/44 | 0.052 | 31.582 |
| *336* | hsa–miR–335–3p | Plasma–Lab 5 | 18/57 | 0.052 | 19.414 |
| *337* | hsa–miR–150–3p | Serum–Lab 5 | 15/39 | 0.051 | 56.756 |

## Supplementary Table S2.3 Differentially 3' uridylated miRNAs in Plasma and Urine from Lab 1

| | | Plasma - Lab 1 | | | Urine - Lab 1 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | miRNA | Ave. 3' U (%) | Ave. RPM | sampleN | Ave. 3' U (%) | Ave. RPM | sampleN | 3'U Difference | beta_binomial.pvalue | FDR.BH |
| 1 | hsa–miR–107 | 48.0 | 345.2 | 160 | 26.9 | 27.1 | 84 | 21.2 | 2.144e–30 | 5.359e–30 |
| 2 | hsa–miR–148b–3p | 1.5 | 3948.0 | 168 | 38.0 | 42.3 | 102 | –36.5 | 8.038e–135 | 2.009e–133 |
| 3 | hsa–miR–140–3p | 70.3 | 2930.0 | 153 | 36.5 | 58.6 | 119 | 33.9 | 1.104e–75 | 9.197e–75 |
| 4 | hsa–miR–132–3p | 27.7 | 35.4 | 120 | 34.2 | 50.1 | 22 | –6.5 | 6.341e–06 | 7.206e–06 |
| 5 | hsa–miR–3615 | 28.5 | 294.7 | 66 | 10.1 | 19.7 | 33 | 18.4 | 3.067e–05 | 3.334e–05 |
| 6 | hsa–let–7e–5p | 22.2 | 810.2 | 160 | 14.3 | 507.3 | 115 | 7.9 | 2.744e–34 | 7.623e–34 |
| 7 | hsa–miR–27a–3p | 21.0 | 591.3 | 125 | 4.6 | 38.5 | 72 | 16.5 | 1.132e–18 | 1.768e–18 |
| 8 | hsa–miR–99b–5p | 17.5 | 984.1 | 152 | 6.1 | 195.9 | 130 | 11.3 | 6.455e–44 | 2.690e–43 |
| 9 | hsa–miR–143–3p | 26.1 | 4137.6 | 165 | 18.4 | 146.2 | 62 | 7.7 | 1.336e–42 | 4.773e–42 |
| 10 | hsa–miR–671–3p | 29.3 | 238.2 | 139 | 17.3 | 31.1 | 31 | 12.0 | 4.653e–24 | 9.694e–24 |
| 11 | hsa–miR–320a–3p | 30.9 | 8784.9 | 155 | 23.8 | 373.7 | 62 | 7.1 | 3.743e–04 | 3.743e–04 |
| 12 | hsa–miR–204–5p | 9.9 | 569.1 | 127 | 3.2 | 2364.1 | 129 | 6.7 | 1.946e–29 | 4.423e–29 |
| 13 | hsa–miR–23b–3p | 71.3 | 313.4 | 146 | 44.6 | 34.1 | 48 | 26.7 | 2.884e–23 | 5.150e–23 |
| 14 | hsa–miR–27b–3p | 19.0 | 6656.3 | 167 | 0.8 | 1996.2 | 126 | 18.1 | 1.055e–106 | 1.319e–105 |
| 15 | hsa–miR–221–3p | 43.8 | 5056.5 | 146 | 18.9 | 130.8 | 44 | 24.9 | 4.293e–19 | 7.156e–19 |
| 16 | hsa–miR–424–3p | 0.5 | 86.9 | 123 | 12.2 | 22.1 | 34 | –11.7 | 4.444e–36 | 1.389e–35 |
| 17 | hsa–miR–98–5p | 7.4 | 2674.8 | 162 | 2.0 | 233.1 | 130 | 5.4 | 1.772e–52 | 8.862e–52 |
| 18 | hsa–miR–125b–2–3p | 48.9 | 21.6 | 83 | 20.7 | 51.9 | 83 | 28.3 | 1.666e–23 | 3.203e–23 |
| 19 | hsa–let–7b–3p | 70.4 | 23.3 | 104 | 57.0 | 18.4 | 21 | 13.5 | 7.652e–05 | 7.970e–05 |
| 20 | hsa–miR–363–3p | 0.4 | 569.1 | 157 | 6.4 | 61.6 | 89 | –6.1 | 3.097e–68 | 1.936e–67 |
| 21 | hsa–miR–532–3p | 2.9 | 19.5 | 37 | 20.5 | 27.7 | 32 | –17.6 | 1.221e–10 | 1.796e–10 |
| 22 | hsa–miR–200b–3p | 39.7 | 21.3 | 80 | 53.1 | 569.7 | 130 | –13.4 | 3.211e–10 | 4.460e–10 |
| 23 | hsa–miR–196b–5p | 46.5 | 29.8 | 115 | 37.2 | 42.3 | 110 | 9.3 | 3.957e–09 | 4.710e–09 |
| 24 | hsa–miR–421 | 13.0 | 208.2 | 107 | 34.7 | 25.2 | 51 | –21.7 | 1.462e–09 | 1.828e–09 |
| 25 | hsa–miR–582–3p | 0.0 | 12.3 | 21 | 12.4 | 14.1 | 21 | –12.4 | 6.291e–10 | 8.278e–10 |

## Supplementary Table S2.4 Differentially 3' uridylated miRNAs in CSF and Serum from Lab 3

| | | CSF - Lab 3 | | | Serum - Lab 3 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | miRNA | Ave. 3' U (%) | Ave. RPM | sampleN | Ave. 3' U (%) | Ave. RPM | sampleN | 3'U Difference | beta_binomial.pvalue | FDR.BH |
| 1 | hsa–miR–125b–1–3p | 58.29 | 80.04 | 64 | 51.45 | 50.77 | 35 | 6.85 | 1.62e–02 | 1.62e–02 |
| 2 | hsa–miR–92a–3p | 19.98 | 91.90 | 64 | 13.55 | 866.60 | 51 | 6.43 | 2.31e–04 | 3.17e–04 |
| 3 | hsa–miR–484 | 12.36 | 16.15 | 28 | 3.40 | 334.78 | 38 | 8.96 | 1.46e–06 | 4.01e–06 |
| 4 | hsa–miR–23b–3p | 1.96 | 143.19 | 58 | 13.88 | 104.69 | 24 | –11.91 | 2.10e–13 | 1.15e–12 |
| 5 | hsa–miR–652–3p | 53.22 | 37.29 | 58 | 6.97 | 340.89 | 33 | 46.26 | 2.24e–36 | 2.47e–35 |
| 6 | hsa–miR–140–3p | 48.14 | 430.60 | 55 | 56.89 | 1585.85 | 43 | –8.75 | 4.69e–05 | 7.38e–05 |
| 7 | hsa–miR–27a–3p | 12.39 | 103.20 | 30 | 5.72 | 534.66 | 42 | 6.68 | 2.08e–03 | 2.54e–03 |
| 8 | hsa–let–7b–3p | 66.85 | 25.17 | 42 | 49.39 | 32.10 | 27 | 17.47 | 1.60e–02 | 1.62e–02 |
| 9 | hsa–miR–500a–3p | 9.26 | 36.44 | 50 | 3.27 | 102.92 | 42 | 5.99 | 2.65e–06 | 4.85e–06 |
| 10 | chr22_25085 | 11.39 | 12.12 | 45 | 2.94 | 176.28 | 51 | 8.45 | 1.27e–08 | 4.67e–08 |
| 11 | hsa–miR–421 | 48.66 | 46.43 | 48 | 28.41 | 84.26 | 36 | 20.25 | 1.91e–06 | 4.19e–06 |

## Supplementary Table S2.5 Differentially 3' uridylated miRNAs in PlasmaExo and Urine from Lab 4

| | | PlasmaExo - Lab 4 | | | Urine - Lab 4 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | miRNA | Ave. 3' U (%) | Ave. RPM | sampleN | Ave. 3' U (%) | Ave. RPM | sampleN | 3'U Difference | beta_binomial.pvalue | FDR.BH |
| 1 | hsa–miR–148b–3p | 9.28 | 1478.62 | 95 | 19.23 | 366.61 | 43 | –9.95 | 1.19e–15 | 2.85e–15 |
| 2 | hsa–miR–200c–3p | 8.49 | 67.79 | 76 | 2.90 | 4547.33 | 43 | 5.59 | 3.09e–14 | 5.30e–14 |
| 3 | hsa–miR–127–3p | 26.82 | 1393.64 | 108 | 35.47 | 52.30 | 34 | –8.66 | 3.55e–08 | 4.73e–08 |
| 4 | hsa–miR–152–3p | 19.32 | 101.85 | 66 | 35.49 | 198.06 | 38 | –16.17 | 4.09e–11 | 6.14e–11 |
| 5 | hsa–miR–125a–5p | 11.78 | 225.58 | 73 | 6.72 | 310.68 | 33 | 5.06 | 1.35e–04 | 1.48e–04 |
| 6 | hsa–miR–25–3p | 9.44 | 2689.79 | 44 | 15.45 | 150.08 | 32 | –6.01 | 5.27e–05 | 6.32e–05 |
| 7 | hsa–let–7a–5p | 5.15 | 83.65 | 42 | 0.06 | 65.29 | 32 | 5.09 | 5.88e–25 | 2.35e–24 |
| 8 | hsa–miR–16–2–3p | 4.04 | 847.99 | 103 | 9.64 | 24.10 | 23 | –5.59 | 4.06e–04 | 4.06e–04 |
| 9 | hsa–miR–500a–3p | 6.55 | 17.19 | 74 | 30.35 | 33.94 | 28 | –23.80 | 1.45e–14 | 2.90e–14 |
| 10 | hsa–miR–2110 | 0.21 | 92.32 | 42 | 44.66 | 45.84 | 26 | –44.45 | 9.59e–63 | 1.15e–61 |
| 11 | hsa–miR–375–3p | 1.12 | 119.28 | 67 | 6.35 | 3311.39 | 32 | –5.23 | 1.71e–20 | 5.14e–20 |
| 12 | hsa–miR–30a–3p | 22.16 | 17.16 | 29 | 0.11 | 2480.41 | 29 | 22.06 | 1.79e–46 | 1.07e–45 |

## Supplementary Table S2.6 Differentially 3' uridylated miRNAs in Plasma and Serum from Lab 5

| | | Plasma - Lab 5 | | | Serum - Lab 5 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | miRNA | Ave. 3' U (%) | Ave. RPM | sampleN | Ave. 3' U (%) | Ave. RPM | sampleN | 3'U Difference | beta_binomial.pvalue | FDR.BH |
| *1* | hsa–miR–148b–3p | 0.93 | 928.39 | 148 | 16.09 | 1672.78 | 75 | –15.16 | 2.77e–89 | 9.97e–88 |
| *2* | hsa–miR–484 | 2.07 | 2638.99 | 195 | 12.07 | 1489.12 | 32 | –10.00 | 1.00e–29 | 7.20e–29 |
| *3* | hsa–miR–744–5p | 8.26 | 194.37 | 78 | 14.54 | 546.11 | 90 | –6.28 | 1.32e–07 | 2.64e–07 |
| *4* | hsa–miR–16–5p | 0.51 | 56.59 | 139 | 14.76 | 61.87 | 58 | –14.26 | 1.87e–69 | 2.24e–68 |
| *5* | hsa–miR–425–5p | 0.85 | 1211.32 | 93 | 6.68 | 978.83 | 23 | –5.83 | 3.39e–21 | 1.36e–20 |
| *6* | hsa–miR–143–3p | 50.29 | 1855.33 | 183 | 55.49 | 4069.55 | 82 | –5.19 | 3.83e–02 | 4.86e–02 |
| *7* | hsa–miR–106b–3p | 1.19 | 1307.92 | 204 | 6.70 | 527.81 | 65 | –5.52 | 6.04e–38 | 5.43e–37 |
| *8* | hsa–miR–23b–3p | 13.37 | 163.89 | 50 | 41.22 | 527.97 | 32 | –27.85 | 1.90e–07 | 3.60e–07 |
| *9* | hsa–miR–92a–3p | 3.49 | 790.78 | 204 | 10.16 | 1307.37 | 91 | –6.66 | 1.35e–25 | 6.97e–25 |
| *10* | hsa–miR–33a–3p | 1.83 | 20.78 | 41 | 13.30 | 13.66 | 30 | –11.48 | 2.25e–08 | 5.06e–08 |
| *11* | hsa–miR–15b–3p | 0.42 | 154.82 | 80 | 5.99 | 135.84 | 34 | –5.57 | 9.53e–27 | 5.72e–26 |
| *12* | hsa–miR–7–1–3p | 12.23 | 28.14 | 60 | 6.44 | 58.27 | 46 | 5.79 | 7.88e–05 | 1.19e–04 |
| *13* | hsa–let–7f–2–3p | 23.94 | 32.63 | 129 | 50.29 | 41.48 | 62 | –26.36 | 6.56e–09 | 1.57e–08 |
| *14* | hsa–miR–652–3p | 33.75 | 273.43 | 169 | 55.34 | 604.04 | 83 | –21.59 | 5.78e–07 | 1.04e–06 |
| *15* | hsa–let–7i–5p | 5.04 | 6609.38 | 59 | 17.11 | 5645.85 | 82 | –12.07 | 1.08e–19 | 3.55e–19 |
| *16* | hsa–miR–18a–3p | 5.90 | 45.32 | 144 | 12.17 | 28.36 | 44 | –6.26 | 2.88e–04 | 3.99e–04 |
| *17* | hsa–let–7e–5p | 10.06 | 36.77 | 118 | 17.64 | 113.32 | 80 | –7.58 | 5.22e–08 | 1.11e–07 |
| *18* | hsa–miR–98–5p | 0.10 | 182.79 | 122 | 6.63 | 302.63 | 86 | –6.53 | 9.64e–79 | 1.73e–77 |
| *19* | hsa–miR–339–3p | 3.72 | 97.02 | 59 | 11.35 | 140.32 | 62 | –7.63 | 1.93e–06 | 3.16e–06 |
| *20* | hsa–miR–1976 | 2.77 | 33.21 | 149 | 9.73 | 22.46 | 49 | –6.96 | 6.22e–10 | 1.60e–09 |
| *21* | hsa–miR–26b–3p | 5.04 | 29.88 | 126 | 21.67 | 57.49 | 63 | –16.63 | 3.98e–21 | 1.43e–20 |
| *22* | hsa–miR–1306–5p | 0.22 | 33.00 | 67 | 8.36 | 41.98 | 27 | –8.14 | 4.28e–25 | 1.93e–24 |
| *23* | hsa–miR–1843 | 3.42 | 21.18 | 74 | 18.35 | 39.13 | 36 | –14.92 | 1.09e–10 | 3.01e–10 |
| *24* | hsa–miR–181d–5p | 10.05 | 25.13 | 64 | 22.10 | 81.50 | 51 | –12.05 | 9.77e–05 | 1.41e–04 |
| *25* | hsa–miR–330–3p | 5.91 | 12.80 | 73 | 11.38 | 39.65 | 49 | –5.47 | 2.11e–03 | 2.81e–03 |
| *26* | hsa–miR–99b–3p | 0.00 | 17.44 | 24 | 7.37 | 31.15 | 22 | –7.37 | 7.49e–12 | 2.25e–11 |
| *27* | hsa–miR–454–5p | 7.37 | 10.61 | 27 | 0.38 | 16.42 | 28 | 6.98 | 7.48e–07 | 1.28e–06 |
| *28* | hsa–miR–1301–3p | 10.54 | 32.74 | 39 | 31.04 | 72.08 | 38 | –20.50 | 7.92e–05 | 1.19e–04 |
| *29* | hsa–miR–23a–5p | 4.32 | 31.30 | 23 | 11.55 | 19.61 | 33 | –7.23 | 3.92e–02 | 4.86e–02 |

# Chapter 3 : Extracellular RNA fragments and RNA editing in human biofluids

## 3.1 Introduction

Extracellular RNAs (exRNAs) have been extensively studied across a variety of biofluids, cell lines, and disease phenotypes[18,37,86,87]. These molecules can be packaged in vesicles, lipoproteins, or within protein complexes, protecting them from RNases within biofluids[7,8]. While most initial studies focused on small non-coding RNA species such as miRNAs, RNA-seq of longer RNA species in extracellular fluids received less attention until recently. An increasing number of studies have now shown abundant and diverse RNAs generated from coding or long non-coding genes, which largely exist as fragments, in the extracellular space with potential functions[11,20,88–90]. For example, when mouse exosomal RNA was transferred to human mast cells, the recipient cells contained new mouse proteins, demonstrating that transferred exosomal mRNA can be translated[9]. Furthermore, extracellular mRNA fragments and long non-coding RNAs (lncRNAs) have been reported to be dysregulated in several cancer types[5,6]. Despite these exciting discoveries, many questions remain, such as mechanisms that enable long exRNA stability, exportation selectivity (i.e. sorting), and function.

RNA binding proteins (RBPs) are hypothesized to enable the sorting and stability of mRNA fragments[93]. For example, RBPs such as YBX1, SYNCRIP, and hnRNPA2B1 were shown to be key regulators in selecting and sorting miRNAs and other small non-

coding RNAs into exosomes[93–97]. However, for longer RNAs, such regulatory mechanisms remain uncharacterized. A recent study of exRNAs in plasma observed many mRNA fragments harboring RBP binding sites and full-length excised intron RNAs with stable predicted secondary structures, possibly contributing to their protection against plasma nucleases[22]. However, investigation of RBPs involved in regulating mRNA fragments across diverse biofluids has not been performed.

Although a number of studies have examined the diverse exRNA molecules present in biofluids, two important aspects received relatively little attention thus far: repeat-derived RNAs and RNA modifications. Retrotransposable elements are abundant repeat elements in eukaryotic genomes. Although most retrotranspons in the human genome lost their ability to retrotranspose, RNAs generated by these repreats have been shown to play roles in various aspects of gene regulation and RNA processing[24,25]. The expression levels of these RNAs may alter in disease conditions, potentially contributing to disease mechanisms. For example, a plural of recent studies reported induced expression of retrotransposons in neurodegenerative disorders[98–100]. Thus, the presence of repeat-derived RNA should be examined in biofluids, which will inform future biomarker studies.

Post-transcriptional RNA modifications are widespread and impact various aspects of RNA stability, function, and localization[101–104]. RNA editing is a post-transcriptional modification process, mediated by the ADAR protein, which converts the nucleotide adenosine to inosine in double-stranded RNA[26]. Since inosine is read as guanine by cellular machineries, RNA editing in coding mRNA regions can induce

protein changes not encoded in the genome. These RNA modifications often demonstrate dynamic variation across different cell types and physiological systems[49,105]. While RNA editing sites have been heavily studied intracellularly and are emerging as promising biomarker candidates, RNA editing sites in exRNAs have not been investigated systematically.

In this study, we characterized extracellular RNA fragments obtained by mRNA capture protocols across 17 unique biofluids from a previous study[106]. By assessing enriched exRNA fragments in transcripts, we observed enrichment of RBP binding site within these fragments compared to controls. In addition, we examined repeat elements across biofluids and identified elevated Alu elements in urine from bladder cancer patients compared to controls. We carried out systematic analysis of RNA editing events across biofluids and identified several recoding events, including one recoding site in the gene AZIN1 that has important functional implications in hepatocellular carcinoma[107]. Overall, our study provides global insights of mRNA fragments, repeat elements, and RNA editing in extracellular fluids.

## 3.2 Results

### 3.2.1. Identification of extracellular RNA peaks in biofluids

Long extracellular RNA (exRNA) sequencing has revealed large amounts of mRNA fragments. It has been hypothesized that these fragments may be protected by proteins bound to the RNA[108]. We aimed to characterize mRNA fragments in the extracellular space across biofluids and determine if these fragments are enriched with

RBP binding sites. To this end, we analyzed extracellular RNA-seq (exRNA-seq) datasets generated via mRNA capture from 17 biofluids by a previous study[106]. We excluded a few biofluids from the analysis, including saliva, stool, sputum, and sweat, as our study focuses on human RNA and data from these fluids are often confounded by bacterial RNA with some degree of homology to human sequences.

We used a peak-caller to identify preserved regions in the mRNA enriched with exRNA-seq reads. Briefly, to call continuous peaks along exons, paired end reads were first aligned to the human transcriptome. Then, the peak calling tool MACS2 was used to identify significant peaks (see Methods). Peaks from multiple transcripts of the same gene were merged. The average count of uniquely mapped and deduplicated reads of peaks within each fluid ranged from 19 (urine) to 417 (breast milk) (Supplementary Figure 3.1A), suggesting that most peaks have sufficient read coverage. Peaks with less than 4 uniquely mapped and deduplicated reads were removed to ensure modest coverage.

Across all biofluids, 96.4% and 3.3% of peaks overlap protein coding regions and pseudogenes, respectively, and less than 0.3% of peaks were mapped to other biotypes. Of the peaks located in protein coding regions, over 98.7% and 1.2% map to CDS regions and 3'UTRs, respectively. The high levels of exonic coverage and depletion of intronic and intergenic peaks is consistent with the usage of mRNA capture in library preparation[109] (Fig. 3.1A and Supplementary Fig. S3.1B). On average each biofluid had 9,635 peaks, ranging from 903 peaks in CSF to 28,382 in pancreatic cyst fluid (Supplemental Fig. S3.1C). The wide range in the number of peaks across biofluids

likely reflects the RNA diversity reported in previous publications[106]. Furthermore, we observed an average of 1.28 peaks per gene across fluids and an average peak length of 936nt per gene across biofluids. (Fig. 3.1B, Fig. 3.1C, and Supplemental Fig. S3.2).

Given the wide range of peak lengths in different genes, we investigated the types of genes with short or long peaks respectively. To this end, we performed gene ontology enrichment analysis. Genes with high peaks length (top 20%, average length 2319nt) were enriched for ribonucleoprotein complexes, immune, and signaling related terms (Fig. 3.2A). In contrast, genes with short peaks (bottom 20%, average 364nt) were enriched for nucleus related terms (Fig. 3.2B). Ribonucleoprotein complexes contain both RNA and RBPs, suggesting that genes with long peaks may be involved in RNA binding to enhance their stability or export.

### 3.2.2. Extracellular RNA peaks are enriched in RBP binding sites

Since genes with long peaks were enriched in ribonucleoprotein complexes-related GO terms, we asked if these peaks harbored RBP binding sites. To bioinformatically assess RBP binding sites across peaks, we overlapped our peaks with binding sites identified from eCLIP datasets of 120 RBPs processed by the ENCODE Consortium (derived from K562 cells). On average, a gene harbored 8 RBP binding sites within their peaks (Supplementary Fig. S3.3A). For comparison, we generated 100 sets of controls peaks within the same genes harboring the test peaks, matching gene feature and peak length (see Methods). If a control peak could not be drawn from the same gene due to the peak length being greater than 50% of the coding region, a

control gene with matched expression in K562 was selected (see Methods). Interestingly we observed that the exRNA peaks across all biofluids were more enriched with RBP binding sites than controls (Fig. 3.3A). Furthermore, for each RBP in each biofluid, we calculated the proportion of peaks harboring the RBP's binding sites. For example, exRNA peaks in bile was enriched with the binding sites of 71/120 RBPs compared to background control regions, while CSF was enriched with binding sites of 44 RBPs (Supplementary Fig. S3.3B).

To visualize the set of RBPs with the highest enrichment of their binding sites in the exRNA peaks, we selected highly enriched RBPs where more than 5% of peaks harbored the RBP's respective binding sites compared to the control sets (Fig. 3.3B). Several biofluids including urine, PPP, serum, ascites, PRP, bronchoalveolar fluid, synovial fluid, CSF, and PFP showed evident enrichment of LIN28B, UCHL5, PUM1, and ZNF800 binding sites.

### 3.2.3. UCHL5 binding sites enriched in exRNA peaks and protein expression in extracellular fractions

Among all tested RBPs, Ubiquitin C-terminal Hydrolase-L5, (UCHL5) had the largest enrichment relative to control peaks across multiple biofluids (Fig 3.3B). Specifically, 26.7% of all peaks (union of all biofluids) contained a UCHL5 binding site compared to 13.5% in controls (Supplementary Fig. S3.4A). Next, we examined whether genes with UCHL5 peaks were distinct between biofluids. We observed most genes harboring UCHL5 binding sites were shared across multiple biofluids

(Supplementary Fig. 3.4B). Pancreatic cyst fluid had the highest number of unique genes in this analysis, but this is likely due to increased library coverage and total number of mRNAs identified. Next, we examined the gene ontologies enriched among genes of UCHL5 targets. Several terms related to mRNA splicing and the spliceosome were enriched in most biofluids (Fig 3.4C).

To further confirm the interaction of UCHL5 with exRNA, we analyzed mass spectrometry data from a previous study of proteins in cells, large extracellular vesicles (lEV), small extracellular vesicles (sEV) and non-vesicles (NV) fractions[63]. Interestingly we observed UCHL5 protein counts in all fractions (Fig. 3.4D). In extracellular fractions, UCHL5 had more protein counts than 30% - 47% of all other 3053 proteins detected from mass spectrometry. In addition, of the 16 RBPs with at least 5% higher binding site enrichment compared to controls, 14 (87.5%) were identified in extracellular vesicles via mass spectrometry[110] (Fig. 3.4B). The enrichment of RBP binding sites and the presence of RBP proteins in extracellular fractions suggests they may play important roles in exRNA regulation.

### 3.2.4. CSF extracellular RNA peaks enriched in a unique set of RBPs compared to other biofluids

As shown in Fig. 4B, CSF has a unique RBP enrichment pattern compared to other biofluids. Among all RBPs, LIN28 had the largest enrichment relative to control peaks in CSF. We performed gene ontology on LIN28B targets in CSF and observed terms enriched for several neurological related categories such as response to

amphetamine, ryanodine receptors, and the rhodopsin signaling pathway (Fig. 3.3E). Ryanodine receptors are highly expressed in the central nervous system[111] and play roles in learning and memory[112,113]. In addition, rhodopsin is a known biomarker for retinal degradation and has been suggested as a potential biomarker for several neurological diseases[114]. Similarly, ZNF800 and DDX3X binding sites were highly enriched within CSF peaks and their targets were also enriched with neurological related terms (Supplementary Fig. S3.5A and S3.5B).

### 3.2.3. Identification of repeat elements across biofluids

Short interspersed nuclear elements (SINE) and endogenous retroviruses (ERVs) have been shown to play roles in gene regulation and disease[115–117]. However, their presence and potential roles in exRNAs have received little attention compared to coding and other noncoding RNAs. To investigate repeat elements in exRNA, we rescued reads that were unmapped to the transcriptome (which does not include standalone repeat genes) and mapped them to the human genome allowing up to 100 multiple mapping positions (see Methods). An average of 39,301 read pairs in each type of biofluid were rescued and annotated to repeat elements (Fig. 3.4A). Among these reads, most mapped to low complexity and simple repeat elements (67.1% to 19.5% respectively). The percent of reads mapping to SINE elements ranged from 1.1% in amniotic fluid to 23.3% in synovial fluid and no SINE elements were found in ascites and urine. (Fig. 3.4B). ERVs elements were lowly represented across biofluids, only contributing up to an average of 0.2% of repeat reads and 38 read counts. While all

healthy genomes have ERVS, some of which are functional, previous reports also observed low levels of ERVs in normal tissues and elevated expression in tumor tissues[118].

To gain further insight into SINE elements, we segregated the reads mapped to Alu lineages, AluS, AluJ, and AluY in each biofluid. Interestingly, we observed that the youngest Alu lineage, AluY, has less expression compared to older Alu lineages AluS and AluJ, ~30 million and ~65 million years old, respectively[119] (Fig. 3.4C). Since increased Alu expression have been shown to be dysregulated in cancers[116,120], we compared expression of Alu classes observed in urine samples from bladder cancer or control patients. Interestingly, we found that all Alu lineages were significantly enriched in the bladder cancer samples compared to controls (Fig. 3.4D). These findings support the presence of repeat elements in extracellular RNA across a variety of biofluids, the potential of which as diagnostic markers should be investigated in the future.

### 3.2.4. Global RNA editing across biofluids

Post transcriptional modification in exRNAs is a poorly studied field partly due to the low RNA concentration and coverage across genic bodies. The mRNA capture method used to generate the above dataset allows relatively high coverage of protein coding regions compared to that in previous studies, thus amenable to analysis of RNA modifications. Here, we focus on A-to-I RNA editing, a type of RNA modification with wide implications in disease mechanisms[27,121]. We applied our previously developed

methods[30] to identify known RNA editing sites cataloged in the REDIportal[122] database (see Methods).

The number of RNA editing sites identified in each fluid varied greatly, largely dependent on sequencing read coverage, with tears having the most (110) RNA editing sites and CSF having the fewest (2). (Fig. 3.5A). The number of detected edited sites correlated with read coverage (Supplementary Fig S3.6A). Although the overall percentage of reads that had an editing event is low (highest in serum, 0.028% reads) (Fig. 3.5B), the editing levels of some editing sites were relatively high in their respective samples (Fig. 3.5C). Interestingly, genes harboring RNA editing sites were enriched in GO terms related to vesicle biogenesis and lifecycle as well as a cytokine-mediated signaling pathway (Fig 3.5D). Furthermore, a total of 97 RNA editing sites were identified as recoding sites (that is, the A to G change corresponds to an amino acid change). Of the 97 recoding sites, 73 were edited in only one biofluid and 24 were edited in at least 2 biofluids (Fig 3.5E and Supplementary Fig. S3.6B). For example, the increased editing levels in AZIN1 (encoding antizyme inhibitor 1) was associated with hepatocellular carcinoma[107]. This recoding site (chr8:103841636) in AZIN1 was found edited in 7 biofluids with an editing level ranging from 2% to 11.5% in colostrum and breast milk, respectively. These findings suggest that RNA editing, and likely other post transcriptional modifications, are present in extracellular RNA and have the potential to be leveraged as biomarkers.

## 3.3 Discussion

In this study, we examined RBP binding site enrichment in mRNA fragments, expression of repeat elements, and RNA editing in exRNA transcriptomes from 17 diverse biofluids. The data were generated via mRNA capture, followed by sequencing. Accordingly, we carried out peak calling within the mRNA annotation (devoid of introns). This approach enabled us to detect thousands of peaks in coding regions at higher resolution than previous studies. For each mRNA, the average length of exRNA fragments in our data was ~1kb, which is in contrast to previous studies that reported shorter fragments (< 200 bases)[19,95,108]. Since most previous studies didn't use mRNA capture or peak calling in mRNAs, it is likely that their reported fragment length is an under-estimate.

Interestingly, the average peak length of ~1kb is consistent with previous findings that mRNAs of ~1kb in length can be transferred to cells[90]. It is unclear if the mRNA fragments in the biofluids contain functional start codons or regulatory sequences, such as internal ribosomal entry sites (IRES), necessary for translation [123,124]. Indeed, if mRNA fragments could be up-taken and translated in recipient cells, it is curious to speculate their potential function. For example, in cancer, such mRNA fragments may generate tumor-specific peptides to inhibit tumor cell proliferation or migration[125],[126]. Future studies should further investigate the potential function of the exRNA fragments.

We observed enrichment of RBP binding sites in mRNA fragments in all biofluids, suggesting that RBPs may play an important role in the selection and stability of extracellular RNAs. Intriguingly, biofluids with higher RBP binding site enrichment had

lower read coverage (Supplementary Fig. S3.7A and S3.4B). Since the percent of peaks overlapping RBP binding sites was calculated relative to the total number of peaks, biofluids with higher read coverage, and thus more tested peaks, may appear to have reduced enrichment. Another possibility is that RNA isolation from certain biofluids may have suffered from intracellular RNA contamination, explaining their apparently high RNA diversity (thus high total unique reads). In this case, the intracellular RNAs may not be enriched with RBP binding sites. On the other hand, enhanced enrichment of RBP binding sites in biofluids with lower read coverage indicates that the most abundant RNAs harbor RBP binding sites, possibly suggesting these RNAs rely on RBPs for stabilization or sorting into vesicles. Overall, our results support the existence of RBP binding site enrichment in exRNA fragments detected across biofluids.

A number of interesting RBPs were uncovered by our analyses. UCHL5 is one of the proteins with the strongest binding site enrichment across multiple biofluids. UCHL5 suppresses protein degeneration by deubiquitinating the distal subunit of the polyubiquitin chain[127,128]. In esophageal squamous cell carcinoma, hepatocellular carcinoma, epithelial ovarian cancer, and endometrial cancer, increased UCHL5 levels accelerated tumor growth and metastasis [129]. While UCHL5 may not be the most highly expressed protein in extracellular fluids, its binding site enrichment indicates that this protein may affect exRNA stability, export or transport. The functional relevance of this protein to exRNA profiles in the tumor microenvironment will be an interesting topic for future studies.

Although CSF was among the biofluids with low read coverage and low number of exRNA peaks, the patterns of RBP binding enrichment for exRNAs in CSF were very unique relative to other biofluids. Gene ontology of the exRNA binding targets of LIN28, ZNF800, and DDX3X in CSF revealed enrichment of ryanodine receptors and rhodopsin signaling pathways. These processes are both relevant to neurological function, consistent with the likely neurological origin of exRNAs in CSF.

Alu RNA abundance was shown to increase under cellular stresses, such as heat shock or viral infection and induced interferon response[117,119]. In addition, Alu RNA accumulation has been shown to induce the epithelial-to-mesenchymal transition of cells, a process crucial for metastasis and cell invasion[120]. Consistently, our results showed increased Alu levels in bladder cancer patients compared to controls. These findings suggests that Alu RNAs should be considered as potential biomarkers for cancer.

Lastly, we reported a comprehensive analysis of RNA editing in biofluids. Our finding suggests that it may be possible to detect RNA editing sites that may serve as biomarkers from non-blood-based biofluids. We were able to detect 97 recoding sites, including the well-known site in AZIN1 that is associated with hepatocellular carcinoma[107]. We observed that exRNA in tears had the highest number of edited sites and serum with the highest percent of edited reads among all biofluids. However, it is unknown to what extent cell- or tissue-specific RNA editing events may contribute to each biofluid's editome. Future studies examining tissue or cell specific RNA editing events in biofluids may help to deconvolute the contribution from different tissue types.

In summary, we analyzed a large number of exRNA-seq data derived from 17 biofluids. Our analysis revealed novel insights regarding RBP-exRNA binding, Alu RNA enrichment and RNA editing profiles across biofluids. This work builds the basis for future studies to examine the functional relevance of these exRNA characteristics, and their potentials to serve as biomarkers for human diseases.

## 3.4 Methods

### 3.4.1. Extracellular RNA-seq data processing

We downloaded exRNA-seq fastq files from the European Genome-phenome Archive (accession number EGAS00001003917[109]). For each data set, adapters and low-quality nucleotides were removed from raw fastq sequences using cutadapt[72] (v.1.11). We aligned reads to hg19 using hisat2[130] (version 2.1.0) with the parameters '--no-softclip --no-discordant --no-mixed -k 5'. Subsequently, uniquely mapped reads were extracted and duplicates were removed using samtools[131] (version 1.11).

### 3.4.3. exRNA-seq gene expression quantification

To assign reads to genes, mapped read coordinates were overlapped to Ensembl [132] gene annotations using BEDTools[133] intersect (version v2.26.0). Genes with less than 4 reads were removed. Furthermore, gene counts for each sample were normalized using a scale factor. To calculate the scale factor for each sample, the reads

were first mapped to ERCC spike-in sequences. Subsequently, the log2(counts) of the spike-in sequences were compared to those in the reference sample “Sweat_2” to generate a scale factor for each sample. Lastly, the sample’s scale factor was applied to the log2(counts) of all genes.

### 3.4.2. Peak calling

Since the RNA-seq libraries were generated following mRNA capture, we aimed to call continuous peaks along the mRNA. Thus, fastq samples were first aligned to the GRCh37 transcriptome (ensembl[132]) using hisat2[130] (version 2.1.0) with the parameters ‘--no-softclip --no-discordant --no-mixed -k 5’. Next, peak calling was performed on the mapped reads using MACS2[134] (version 2.2.7.1) with the parameters ‘--nomodel --gsize hs –broad’. Peaks with a fold change < 2 and an adjusted q-value < 0.05 were removed. To perform downstream analysis, the transcriptome coordinates of each peak were converted to genomic coordinates using the Bioconductor package GenomicRanges[135] (version 1.44.0) and the R package EnsDb.Hsapiens.v75[136] (version 2.99.0). To reduce redundant peaks from genes with multiple transcripts, peaks with overlapping coordinates were merged. Lastly, peaks with less than four reads from genome mapping described in the previous section were removed to ensure at least a modest sequencing coverage.

### 3.4.2. Overlapping peaks with eCLIP binding sites

RNA binding sites derived from eCLIP-seq of 120 RNA binding proteins in the K562 cell line were downloaded from the ENCODE consortium[137]. We overlapped exRNA peaks with eCLIP RBP binding peaks labeled as IDR (Irreproducible Discovery Rate) reflecting those both significantly enriched and reproducible in eCLIP biological replicates.

### 3.4.4. Enrichment of RBP binding sites

To identify if exRNA peaks were enriched with RBP binding sites, we generated a set of control peaks. For each exRNA peak in each sample, a control region of the same length was randomly picked from the same gene as the exRNA. In cases where a valid control region was not available in the same gene because the observed exRNA peak was longer than 50% of the gene's mRNA, another gene with matched expression level as the exRNA gene was used to draw a random peak. Since the eCLIP data were obtained from K562 cells, we used the gene expression values from this cell-line to define matched control genes (±10% relative to the query peak). Importantly, all control peaks were drawn from the transcriptome (i.e., not including introns). A total of 100 sets of random control peaks were constructed for each sample. For each control set, the proportion of peaks overlapping the RBP binding sites was calculated. This proportion of 100 control sets was fit with a Gaussian distribution to calculate a p value for the observed proportion of exRNA peaks overlapping RBP binding sites. Significantly enriched RBP binding sites were defined as those with FDR < 0.05.

### 3.4.5. Gene ontology enrichment

Gene ontology (GO) terms for genes (GRCh37 build) were downloaded from Ensembl[132]. For each enriched RBP, query genes were genes harboring peaks with the respective RBP's RNA binding sites. For each query gene, a control gene without the RNA binding site was randomly selected that matched the expression level and gene length of the query (± 10% relative to the query). Expression of genes from the respective biofluid was used to generate the random controls. GO terms present in the sets of query genes and matched controls were collected, respectively. A total of 10,000 sets of control genes were generated, where each set has the same number of genes as the query set. Query genes with less than five candidate controls were not included in this analysis. For each GO term associated with at least 3 query genes, a Gaussian distribution was fit to the number of control genes also associated with this term to calculate a p value. Significant GO terms were defined as those with FDR < 0.05.

GO analysis was performed similarly for genes harboring RNA editing sites. Random control genes were those without RNA editing sites that matched for gene length and expression (± 10% relative to the query).

### 3.4.6. Repetitive elements analysis

Reads unmapped to the transcriptome were mapped to the hg19 genome using hisat2[138] with the parameters '--no-softclip --no-discordant --no-mixed -k 100', allowing up to 100 multiple mapping positions. Read coordinates were overlapped with UCSC

RepeatMasker[139] using BEDTools[133] pairtobed and requiring strandedness. Each read was assigned to a repeat family and class. Gene expression counts for each repeat class was normalized by the sample's scale factor described above for gene expression. Differential expression of repeat classes between case and control groups was performed via a Mann Whitney U test.

### 3.4.8. Identification of RNA editing sites

Using our previously published methods[80,81,140,141] , we detected editing events within each sample recorded in the REDIportal v2[122] database. Editing sites that overlapped a variant listed in dbSNP (version 147) and with a minor allele frequency of > 1% were excluded from downstream analyses. To ensure sufficient coverage, we required RNA editing sites to have a minimum read coverage of five with two reads harboring the edited nucleotide (G).

## 3.5 Figures

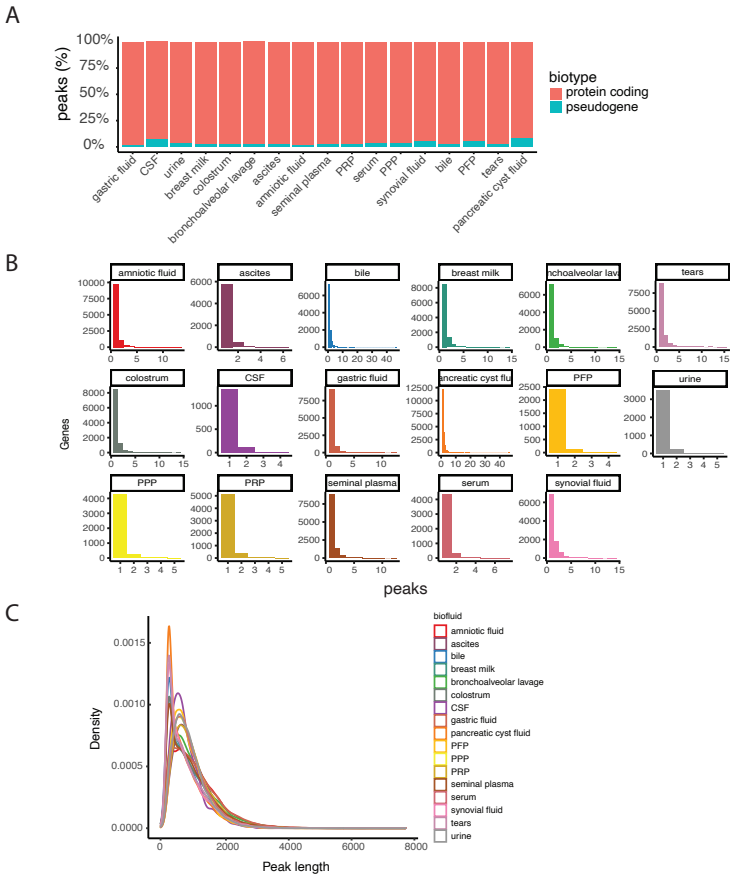



**Figure 3.1 Overview of exRNA peaks across biofluids.**

The following biofluids were studied: gastric fluid, cerebral spinal fluid (CSF), urine, breast milk, colostrum, bronchoalveolar lavage, ascites, amniotic fluid, seminal plasma, platelet rich plasma (PRP), platelet poor plasma (PPP), platelet free plasma (PPP), serum, synovial fluid, bile, tears, and pancreatic cyst fluid. **A** Percentage of extracellular peaks annotating to protein coding genes or pseudogenes in each biofluid. **B** Histogram of the number of peaks per protein coding gene in each biofluid. **C** Distribution of the peak lengths annotating to protein coding genes in each biofluid.

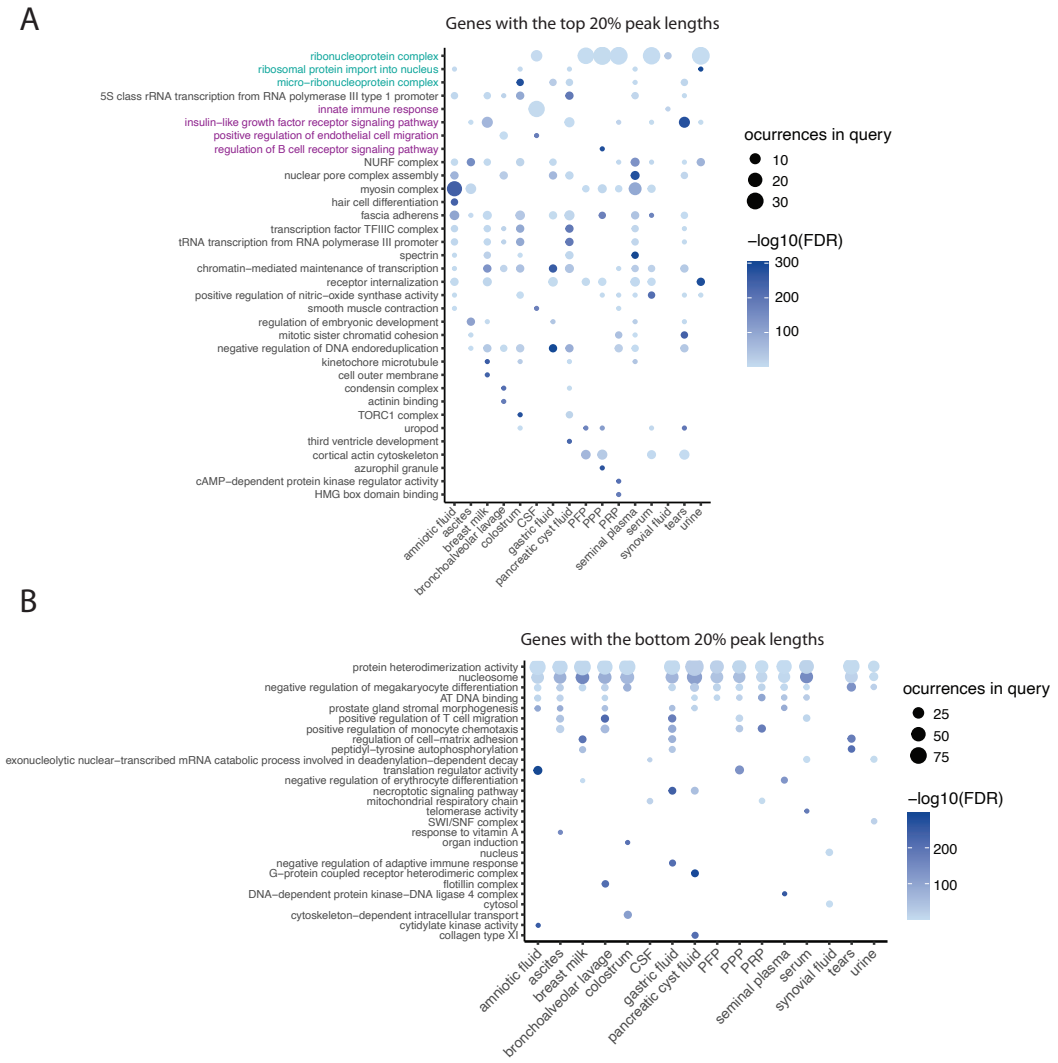



**Figure 3.2 Longer peaks are enriched in ribonucleoprotein complexes and immune related signaling terms.**

**A** Gene ontology enrichment of the genes with the top 20% longest peaks in each biofluid represented by color (-log10 transformed adjusted p value) and point size (occurrences in the query genes). Top 2 GO enriched terms for each biofluid are shown. Background genes match for ±10% gene length and expression in each respective

biofluid. Parent GO terms were removed (Methods). Turquoise text color indicates categories RNA binding and purple text color indicates immune or signaling related categories. **B** Similar to A, but using genes with the bottom 20% shortest peaks.

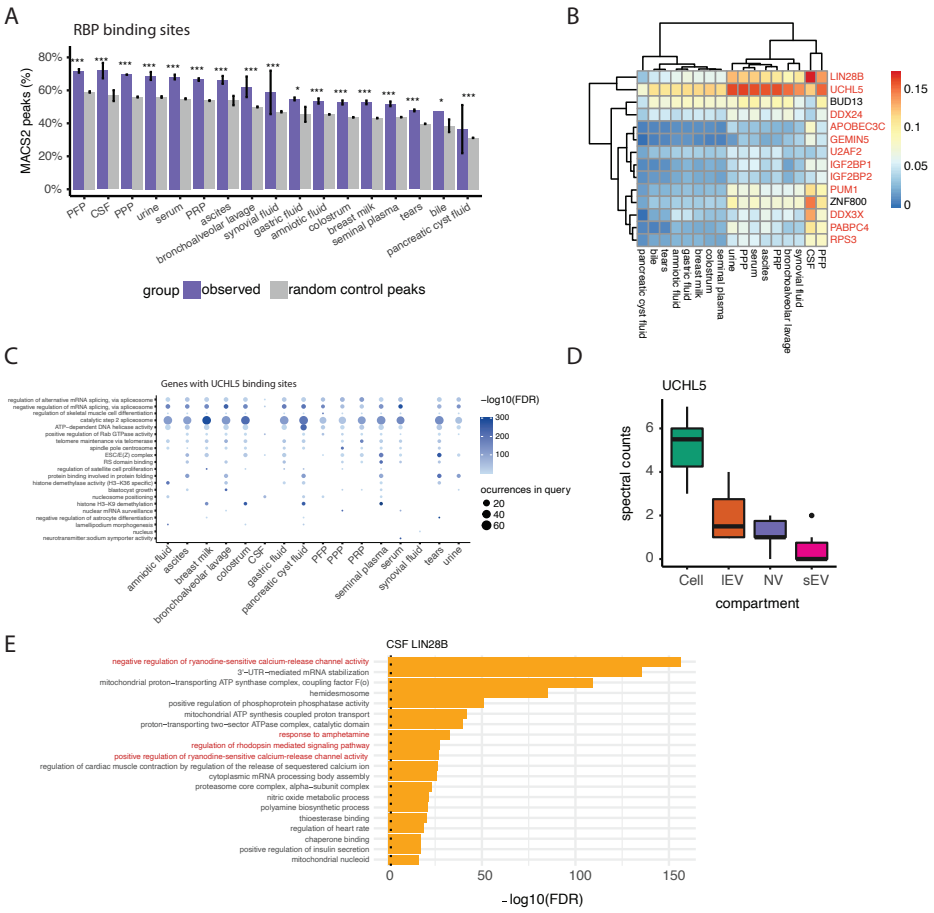



**Figure 3.3 exRNA peaks are enriched in RBP binding sites.**

**A** Percent of peaks overlapping binding sites from 120 RBP eCLIP experiments in K562 cell line compared to random control peaks in each biofluid. Control peaks match for gene, peak length, and gene feature (see methods). Significance of empirical p-values are shown as ns: $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. **B** Heatmap represents the set of RBPs with at least 5% change between the percent of observed peaks overlapping the respective RBP's binding sites compared to control peaks. The color

represents the percent of peaks overlapping the RBP binding sites. RBP names in red indicates those found in Vesiclepedia1 from mass spectrometry. **C** Gene ontology enrichment of the genes with peaks harboring UCHL5 binding sites in each biofluid represented by color (-log10 transformed adjusted p value) and point size (occurrences in the query genes). The top 2 significant terms are shown. Background genes without UCHL5 binding sites matching for ±10% gene length and expression in K562 cell line were used. Parent GO terms were also removed (Methods). **D** Spectral counts of UCHL5 in cell, large extracellular vesicles (lEV), non-vesicle (NV), and small extracellular vesicles (sEV) fractions from the DKO1 cell line2. **E** Gene ontology enrichment of the top 20 significant terms of the genes with peaks harboring LIN28 binding sites in CSF. Red text color indicates neurological related categories. Parent GO terms were removed (see methods). Black dotted line indicates FDR < 0.05

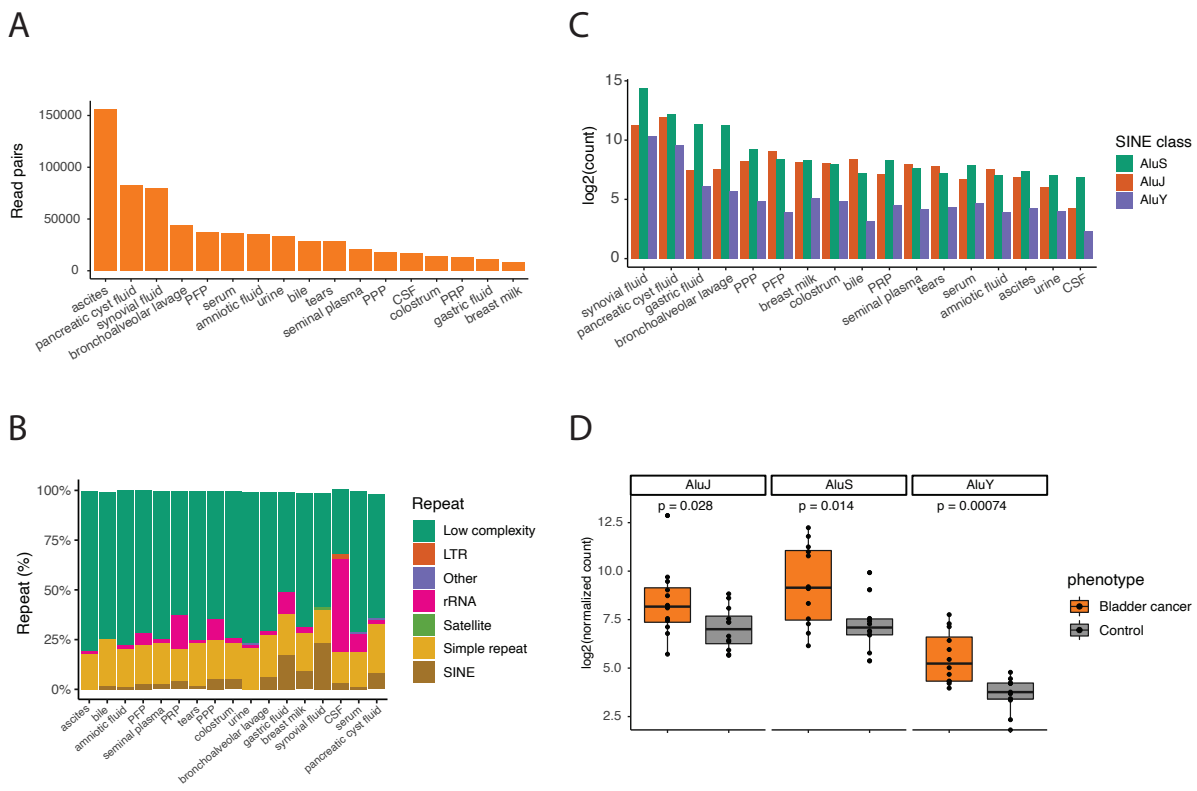



**Figure 3.4 Alu elements are expressed in biofluids and enriched in bladder cancer.**

**A** Number of rescued read pairs mapping to repeat elements from RepeatMasker **B** Percent of rescued reads mapping to repeat families. **C** The log2 normalized count of reads mapping to the Alu sub-families in each biofluid. **D** The log2 normalized count of reads mapping to Alu sub-family elements from urine of bladder cancer and controls patients. P-values were calculated via Wilcoxon rank-sum test.

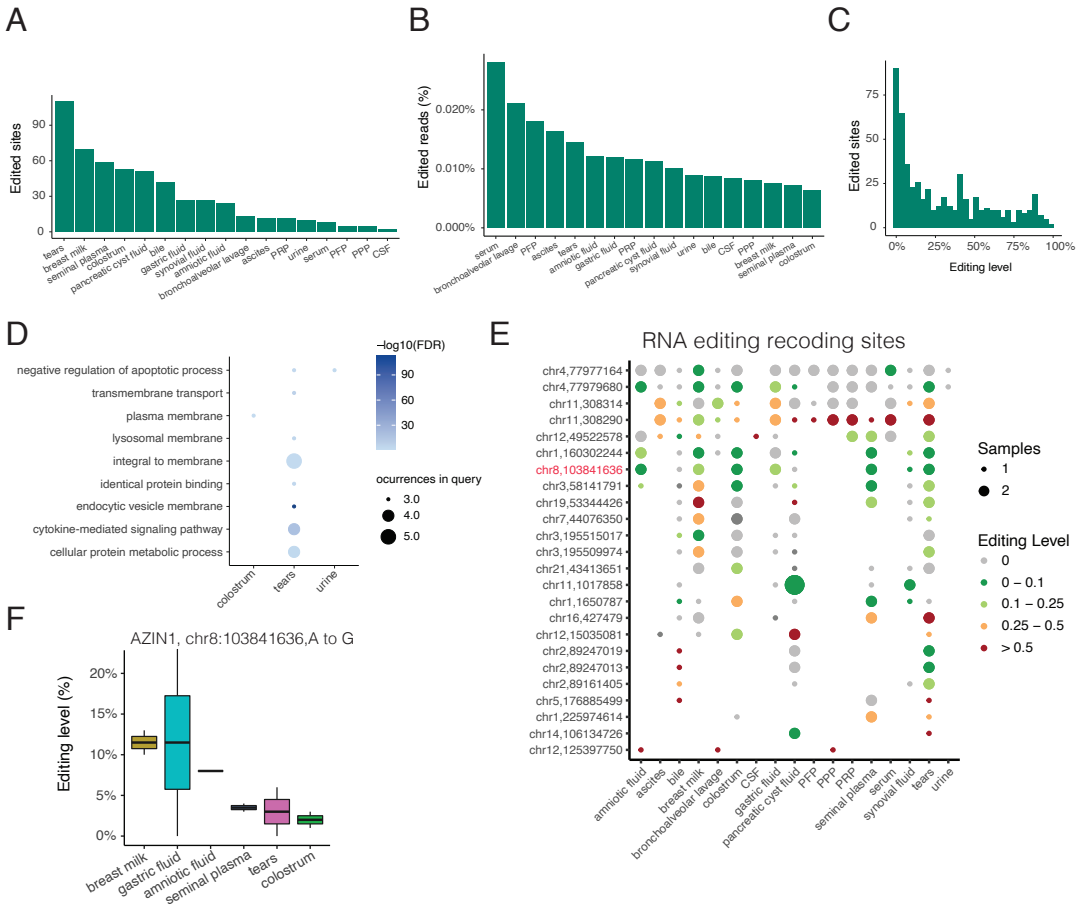
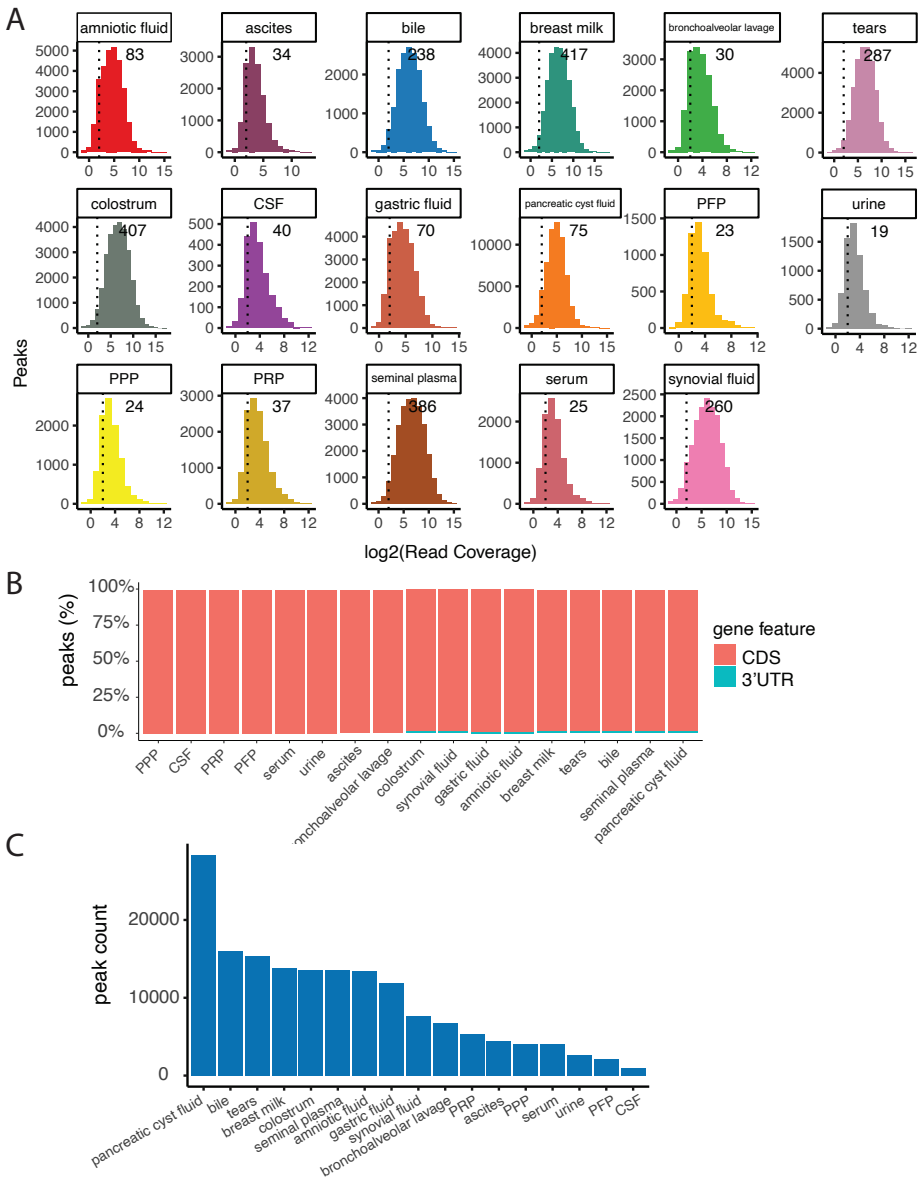



**Figure 3.5 RNA editing events identified across biofluids.**

**A** Histogram of the editing level of all edited sites identified in the REDIportal database4 across all biofluids. **B** The number of edited sites in each biofluid. **C** The total percent of edited reads in each biofluid. **D** Significant gene ontology terms for edited sites across biofluids represented by color (-log10 transformed adjusted p value) and point size (occurrences in the query genes). **E** Bubble plot of recoding sites edited in at least 2

biofluids represented by color (editing level group) and point size (number of samples covered in each biofluid). AZIN1 editing site is shown in red text. **F** Editing level of AZIN1 recoding site chr8:103841636,A to G in biofluids.

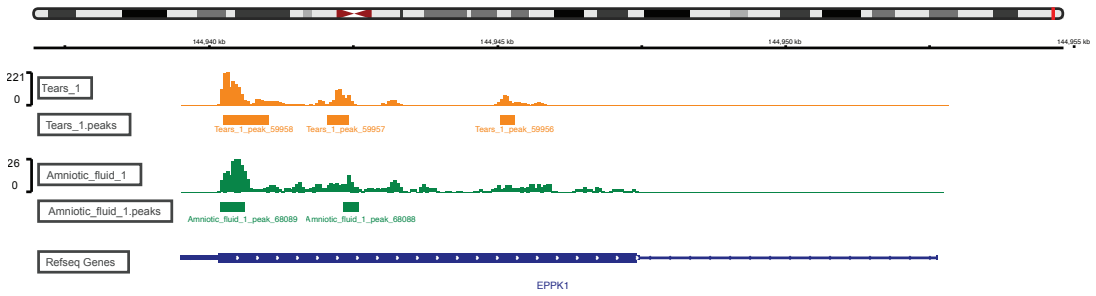
## 3.6 Supplementary Figures




**Supplementary Figure S3.1 Gene feature and coverage of exRNA peaks across biofluids.**

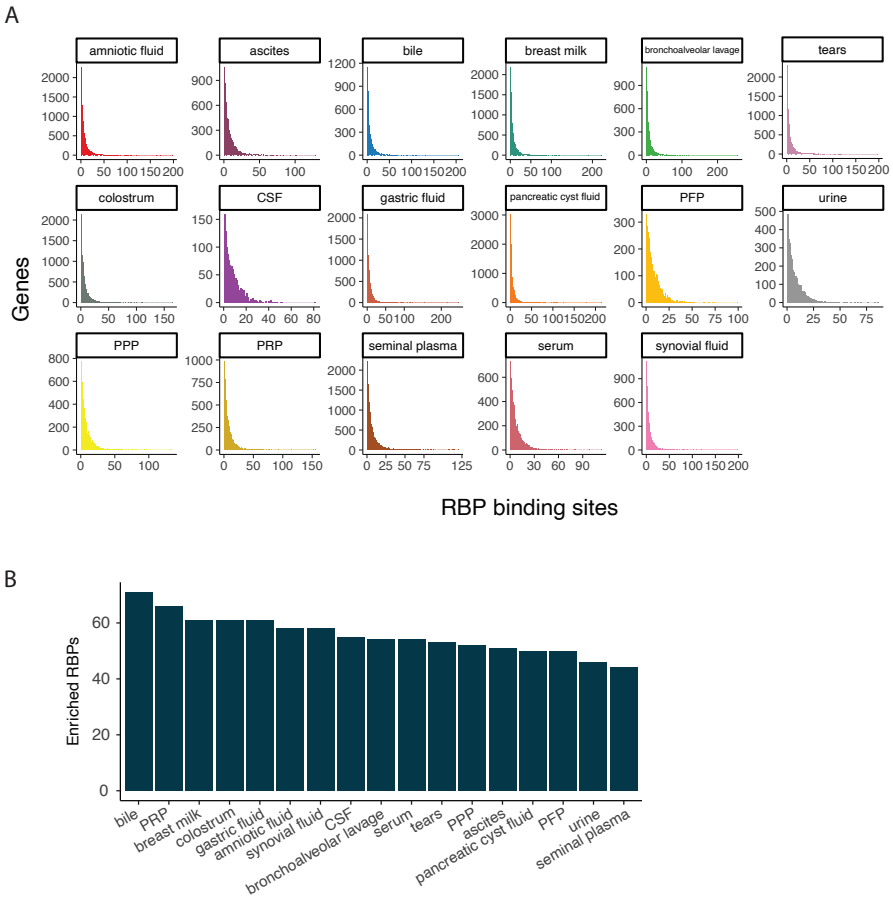
**A** Histogram of the log2 read coverage of peaks in each biofluid. Vertical dash line represents the log2(4) threshold to filter out lowly covered peaks. **B** Percent of peaks

annotating to coding sequence (CDS) or 3'UTR regions. **C** Number of peaks in each biofluid.



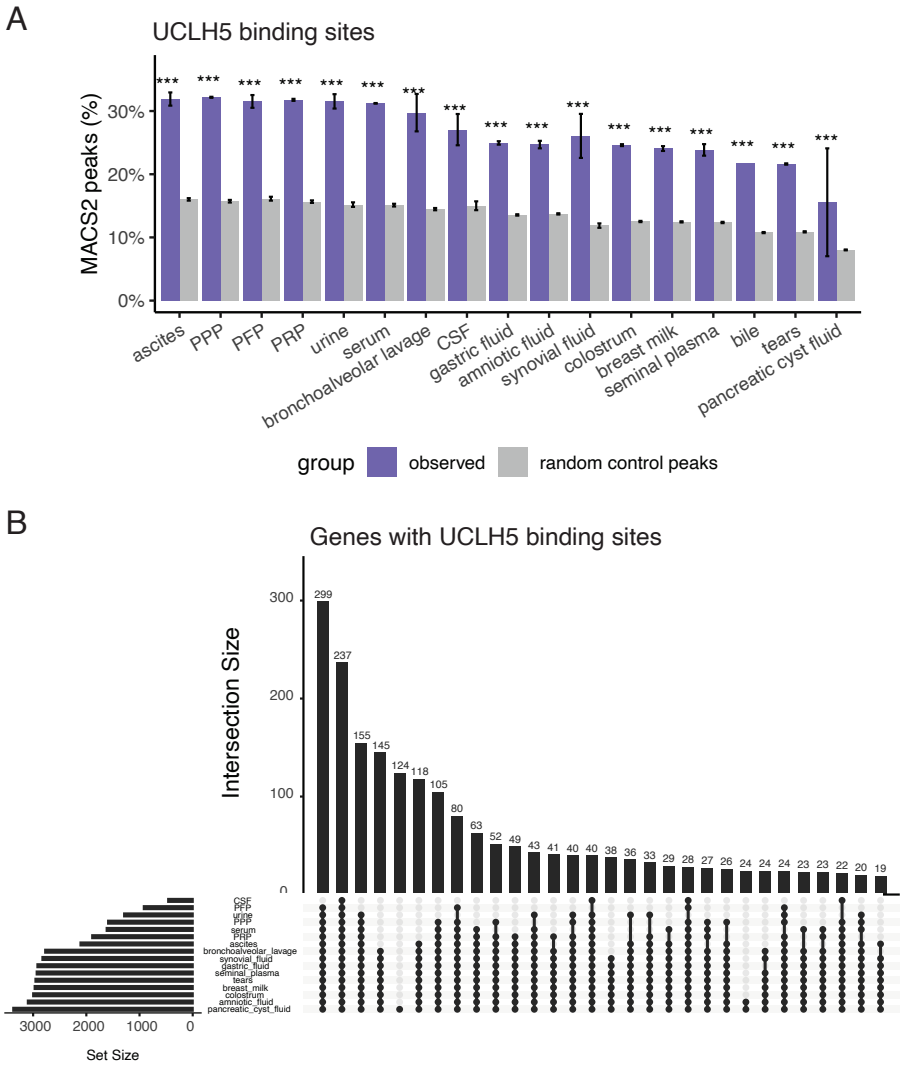

**Supplementary Figure S3.2 Example of peaks called on long extracellular RNA.**

Read density and peaks called in the gene EPPK1 for an Amniotic fluid and Tears sample.




**Supplementary Figure S3.3 Number of RBP binding sites in genes and number of enriched RBPs across biofluids.**

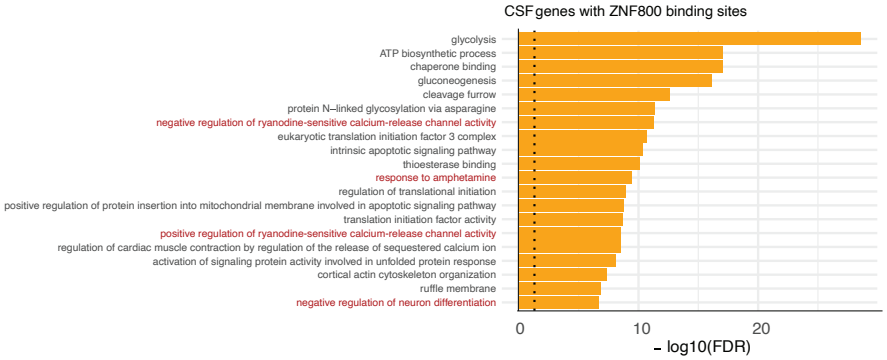
**A** Histogram of the number of RBP binding sites residing in peaks for protein coding genes across each biofluid. **B** The number of RBPs with binding site enrichment in exRNA peaks compared to matched controls (see Methods) across biofluids.



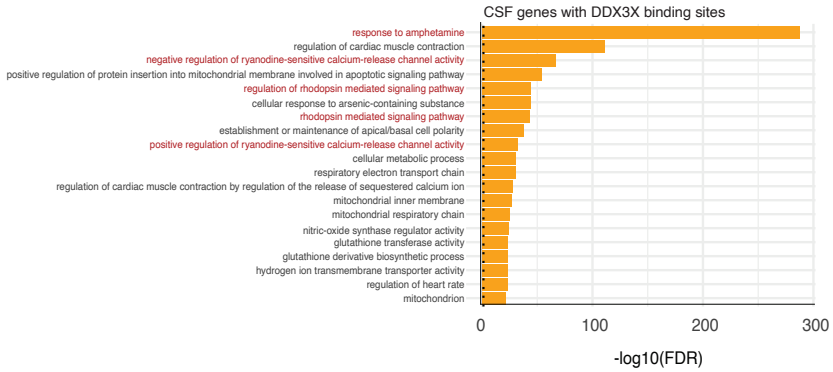

**Supplementary Figure S3.4 UCHL5 binding site enrichment in peaks across biofluids.**

**A** Percent of peaks overlapping UCHL5 binding sites from eCLIP experiments in K562 cell line compared to random control peaks in each biofluid. Control peaks match for gene, peak length, and gene feature (see methods). Significance of empirical p-values are shown. ns: p > 0.05, * p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001. **B** Histogram of the number of shared genes harboring UCHL5 binding sites across groups of biofluids.

A

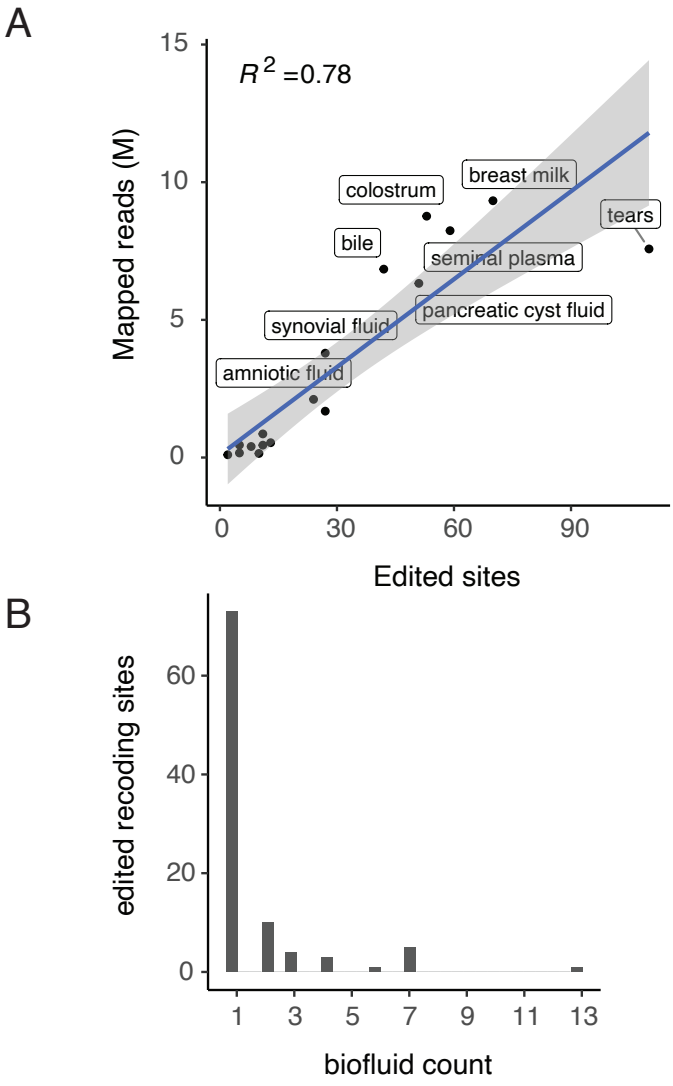



B



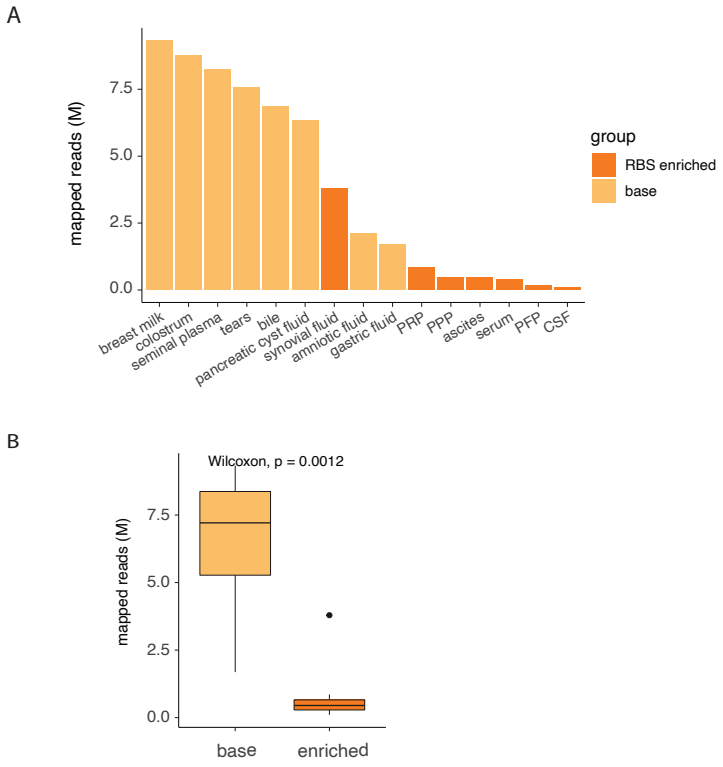

**Supplementary Figure S3.5 Genes enriched with ZNF800 and DDX3X binding sites are enriched in neurological related terms.**

**A** Gene ontology enrichment of the top 20 significant terms of the genes with peaks harboring ZNF800 binding sites in CSF. Red text color indicates neurological related categories. Parent GO terms were removed (see methods) **B** same as A, for DDX3X. Black dotted line indicates FDR < 0.05




**Supplementary Figure S3.6 RNA editing sites across biofluids.**

**A** Number of edited sites and the number of mapped reads (M) for each biofluid. **B** A histogram of the number of edited recoding site identified in multiple biofluids.




**Supplementary Figure S3.7 Biofluids with enhanced RBP enrichment have fewer mapped reads.**

**A** Number of mapped reads in the enhanced RBS enriched group (urine, PPP, serum, ascites, PRP, bronchoalveolar fluid, synovial fluid, CSF, and PFP) and the base group (breast milk, colostrum, seminal plasma, tears, bile, pancreatic cyst fluid, amniotic, and gastric fluid) **B** P-value between the two groups of mapped reads using the Wilcoxon rank-sum test

# Chapter 4 : Extracellular RNA editing in plasma of Alzheimer's Disease patients

## 4.1 Introduction

Alzheimer's Disease (AD) is the most common form of dementia, affecting 6.2 million Americans of age 65 and older and projected to grow to 13.8 million by 2060[31,32]. This progressive neurodegenerative disease is characterized by symptoms such as memory loss and language impairment[32]. Hallmark changes in the brain associated with AD include the accumulation of beta-amyloid plaques outside neurons and tau tangles inside neurons[32]. Importantly, these changes in the brain may begin many years before clinical symptoms emerge, revealing the need for early detection biomarkers.

Physical and neurological tests, tau levels in cerebrospinal fluid (CSF), or positron emission tomography (PET) brain imaging are used to diagnose the severity of cognitive impairment[32,142]. However, postmortem examination remains the gold standard for establishing AD pathology. Development of non-invasive tests associated with the molecular changes in the brain is paramount for diagnosis and developing effective therapies.

Extracellular RNAs (exRNAs) play important roles in cell-to-cell communication, they have also been heavily studied as non-invasive biomarkers[4,143]. exRNAs are found in various biofluids and stably packaged in exosomes, microvesicles, apoptotic bodies, high-density lipoproteins or associated with protein complexes[144]. Some exRNAs, such as microRNAs, mitochondrial RNAs, and mRNAs, have been reported to be

upregulated in plasma from AD patients[89,145–147]. While exRNA studies largely focused on gene expression levels, RNA modification patterns, such as adenosine-to-inosine (A-to-I) editing, are also detectable in exRNA-sequencing data. A-to-I editing is the most prevalent type of RNA editing in humans. A post-transcriptional modification process, A-to-I editing is essential for normal brain development, neuronal function, and immune response[28,101]. It is mediated by the ADAR protein, which converts the nucleotide adenosine to inosine in double-stranded RNAs[26]. Since inosine is read as guanine by cellular machineries, RNA editing in coding regions can induce protein changes not encoded in the genome (thus called recoding sites). Furthermore, RNA editing was shown to have reduced levels in post-mortem brains of AD individuals compared to controls[34,35]. Although several editing sites have recently been discovered in extracellular RNA[148–150], editing changes in exRNA from Alzheimer's disease individuals are not well understood.

Here we report dysregulated RNA editing events detected in plasma of 172 AD patients, relative to166 non-cognitive controls (NCIs) collected by a previous study[89]. We identified a set of differentially edited sites that show elevated editing in AD individuals. Several sites had editing levels associate with cognitive scores. Furthermore, genes harboring editing sites were enriched in toll-like receptor and innate immune signaling pathways. Overall, we provide insights regarding RNA editing in exRNAs of plasma of AD patients.

## 4.2 Results

### 4.2.1. Overview of whole-exome captured exRNA read annotations in plasma from AD and NCI samples

We first analyzed the exRNA expression profile of 338 plasma samples. These exRNA-seq data (75 base-paired-end strand specific) were obtained using whole-exome RNA capture from 172 AD and 166 NCI samples by a previous study[89]. Across AD and NCI samples, 90.3% and 92.4% of reads overlapped with protein-coding regions, respectively (Fig.4.1A). Additionally, samples with larger proportions of mapped reads overlapping rRNA or mitochondria rRNA genes may suggest reduced effectiveness of the whole-exome capture for those samples (Fig. 4.1A). Overall, the majority of reads were mapped to CDS (69.9% and 72.7%) and 3'UTRs (14.6% and 13.7%) for AD and NCI, respectively (Fig. 4.1B). The 8.9% and 6.9% of AD and NCI reads mapping to non-coding exons reflect reads mapping to non-coding genes. The high levels of CDS coverage and depletion of intronic and intergenic reads are consistent with the usage of whole-exome capture in library preparation[89].

### 4.2.2. Extracellular RNA editing events identified in plasma from AD and NCI samples

While RNA editing has been heavily studied intracellularly, it's abundance in extracellular fluids has not been well characterized. The mRNA capture method used to generate the above dataset allows relatively high coverage of protein coding regions compared to that in previous studies, thus amenable to analysis of RNA editing. We applied our previously developed methods[30] to identify known RNA editing sites

cataloged in the REDIportal[122] database (see Methods). Although the majority of reads were aligned to CDS regions (Fig 4.1A), most RNA editing sites were observed in 3'UTRs (60.2% and 65.6% in AD and NCI, respectively) (Fig 4.2A), consistent with previous reports[151]. While average editing level was modest (4.3%), we observed a number of editing sites with relatively high editing levels up to 95.4% (Fig 4.2B).

### 4.2.3. Differential RNA editing events identified in plasma from AD and NCI samples

To examine editing differences in AD and control patients, we performed differential editing analysis using REDITs[65] on all common sites (10% prevalence across all samples) with a minimum effect size of 2%. Briefly, REDITs implements a beta binomial model to identify differential editing sites and has been shown to reduce both false positives and false negatives. Using this method, we identified 28 differentially edited sites between the AD and control samples, of which 5 were recoding sites (Supplementary Table S4.1 and Supplementary Fig. S4.1). Among the 28 editing sites, FOXN3 was the only gene that harbored more than one differentially edited site. As shown in Supplementary Fig. S4.2, 3 differentially edited sites located in the 3'UTR of FOXN3 were observed, with elevated levels in AD samples compared to controls.

We next asked if there exists a hypo- or hyper-editing trend in exRNA editing of AD patients, since previous studies of RNA editing in brain samples of Alzheimer's Disease[152] or Autism Spectrum Disorders[141] showed hypo-editing. Interestingly, we observed that 21 of the 28 differentially edited sites had higher editing levels in AD

samples compared to controls (Fig 4.2C). Given the observed difference in RNA editing between AD and controls, we also calculated the editing activity of the 28 differentially edited sites represented by the RNA editing index (see Methods). We observed that AD samples had a higher editing index compared to NCIs (Fig 4.2D). These findings demonstrate RNA editing sites in exRNAs show differences between AD and control subjects.

### 4.2.3. Differential RNA editing levels correspond with cognitive scores

To further examine the relevance of RNA editing to AD, we asked if editing levels were correlated with cognitive metadata scores such as the Mini-Mental State Exam (MMSE) or Clinical Dementia Rating (CDR). Interestingly, we identified 3 differential RNA editing sites whose editing levels showed correlation with clinical data (Fig. 4.3A – 4.3C). For example, editing level of the recoding site chrX,153579950 was elevated in AD samples relative to NCI samples. Consistently, its editing level was lowest in the samples with normal MMSE scores (25+), but significantly elevated in samples with MMSE scores corresponding to mild (20-24) and severe (<12) dementia (Fig. 4.3A). Similar results were observed using the CDR scores. In addition, the other two editing sites showed the same trends, that is, higher editing in AD and higher editing in more severe AD patients.

### 4.2.4. Genes harboring differential RNA editing sites are enriched for immune signaling terms

We performed GO enrichment to understand the cellular processes of genes harboring differentially edited sites (Fig 4.4). Interestingly, we observed an enrichment of a variety of toll-like receptor signaling and other signaling pathways important to innate immune response. These results indicate that RNA editing could contribute to or be a consequence of immune signaling in AD.

## 4.3 Discussion

In this study, we examined RNA editing in 338 extracellular RNA samples from plasma of AD and control subjects. The data was generated via whole-exome capture, followed by sequencing, enabling detection of coding regions at higher resolution than previous studies. We identified 28 differentially edited sites in AD compared to controls, with an overall trend of hyperediting in AD. While previous studies reported lower levels of RNA editing in post-mortem brains in AD patients[34,35], upregulated editing in plasma likely reflect elevated editing levels in blood cells, which has been shown as the primary source of exRNAs in plasma[153]. Additionally, we did not find any differentially edited sites in brain specific genes.

RNA editing is known to have close relevance to innate immunity through altering double-stranded RNA (dsRNA) structures. Endogenous dsRNAs, if unmodified, may cause aberrant interferon and inflammatory responses[154]. A-to-I editing helps to eliminate immunogenic dsRNAs since the inosines in dsRNAs disrupt base-pairing

patterns and abolish dsRNA recognition by cytosolic sensor proteins. We observed that differential editing sites in exRNAs of AD were enriched in toll-like receptor signaling and innate immune response terms, which is in line with the known functional roles of RNA editing in innate immunity. However, it is unclear whether such editing changes are causal factors to AD pathogenesis, or the consequence of global transcriptome changes in AD.

Of the 28 differentially edited sites, FOXN3 was the only gene with more than 1 such site. The 3'UTR of FOXN3 harbored 3 differentially edited sites, all of which had higher editing levels in AD and nearly no editing in NCI controls. FOXN3 belongs to the forkhead box (FOX) protein family and functions as a transcription factor[155]. FOXN3 has been reported as a tumor suppressor, implicated in cell cycle regulation and tumorigenesis[155]. Since 3'UTRs are known to be hubs for miRNA targeting and regulation by RNA binding proteins (RBPs), RNA editing in the 3'UTR of FOXN3 may impact mRNA abundance by modifying the target sites of miRNAs or RBPs[156,157]. Future studies should investigate the functional consequences of these editing sites in FOXN3.

Among the 28 differentially edited sites, the editing levels of three sites were correlated with cognitive metadata such as CDR and MMSE scores. In addition, we identified 5 recoding sites, including a recoding site in AZIN1 which has important functional implications in hepatocellular carcinoma[107]. While no correlation of hepatocellular carcinoma and Alzheimer's disease currently exist[158], editing of AZIN1 may have other unknown functions. While our analysis shows elevated editing levels of

CCNI in plasma, reduced editing in the temporal lobe was previously observed of AD individuals compared to controls[34]. This difference is likely due to blood cells contributing a majority of the RNAs in plasma. Elevated RNA editing of a recoding site in CCNI was reported to produce edited tumour antigens that was presented on HLA and stimulated T-cell responses[159]. A body of evidence has shown an inverse association between AD and cancer prevalence[160]. Presentation of edited tumor antigens may elicit antitumor immune responses. Editing of CCNI may have multiple functions depending on its tissue type. RHOA is a member of the Rho family of small GTPases and is involved in regulating cell shape, attachment, and motility[161]. Increase levels of RHOA prior to neuronal death has been shown to contribute to neuronal apoptosis[162]. Interestingly, we observe increased editing in RHOA in AD patients, however the relationship of RHOA RNA editing and expression in AD plasma remains unknown. Overall, these findings suggest editing sites combined with expression or other molecular markers may add diagnostic value for Alzheimer's Disease and help understand underlying immune system changes.

In summary, we analyzed a large number of exRNA-seq data derived from 172 AD and 166 NCI samples and report increased editing levels of differentially edited sites in AD individuals. This work builds the basis for future studies to examine the functional relevance of exRNA editing and their potentials to serve as biomarkers for human diseases.

## 4.4 Methods

### 4.4.1. Extracellular RNA-seq data processing

We downloaded fastq files under the accession number PRJNA574438 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA574438/)[89]. For each RNA-seq data set, adapters and low-quality nucleotides were removed from raw fastq sequences using cutadapt[72] (v.1.11). We aligned reads to hg19 using hisat2[130] (version 2.1.0) with the parameters '--no-softclip --no-discordant --no-mixed -k 5'. Subsequently, uniquely mapped reads were extracted and duplicates were removed using samtools[131] (version 1.11).

### 4.4.2. ExRNA-seq gene expression quantification

To obtain gene expression values, mapped read coordinates were overlapped to Ensembl [132] gene annotations using BEDTools[133] intersect (version v2.26.0). Genes with less than 10 read counts were removed. TPM was then calculated for each gene.

### 4.4.3. Identification of RNA editing sites

Using our previously published methods[80,81,140,141] , we searched for editing events within each sample recorded in the REDIportal v2[122] database. Editing sites that overlapped a variant listed in dbSNP (version 147) and with a minor allele frequency of > 1% (i.e., relatively common SNPs) were excluded from downstream analyses. To

ensure sufficient coverage, we required RNA editing sites to have a minimum read coverage of 10 with 2 reads for the alternative allele.

#### 4.4.4. Differential RNA editing

To identify differentially edited sites between AD and controls, we first selected sites that were edited in at least 10% of all samples. Next, we performed differential editing using REDITs likelihood ratio test[65] on sites with an effect size of at least 2%. Differentially edited sites were FDR corrected.

#### 4.4.5. RNA editing index across biofluids

An RNA editing index was calculated for each sample. The RNA editing index was the average editing level of the 28 differentially edited sites.

#### 4.4.6. Gene ontology enrichment

Gene ontology (GO) terms for genes (GRCh37 build) were downloaded from Ensembl[132]. Query genes were genes harboring differential RNA editing sites. For each query gene, a control gene without the RNA editing site was randomly selected that matched the expression level and gene length and expression of the query (± 10% relative to the query). Expression of genes across all samples were used to generate the random controls. GO terms present in the sets of query genes and matched controls were collected, respectively. A total of 10,000 sets of control genes were generated,

where each set has the same number of genes as the query set. Query genes with less than five candidate controls were not included in this analysis. For each GO term associated with at least 3 query genes, a Gaussian distribution was fit to the number of control genes also associated with this term to calculate a p value. Significant GO terms were defined as those with FDR < 0.05.

## 4.5 Figures

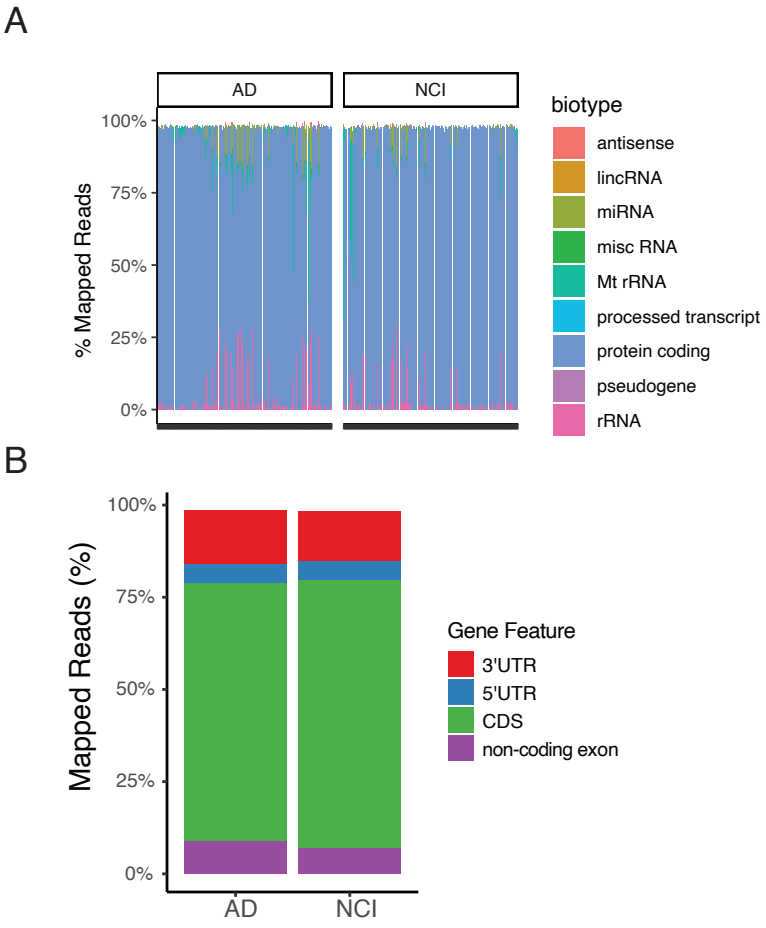



**Figure 4.1 Overview of whole-exome captured exRNA read annotations in plasma from AD and NCI samples.**

**A** Whole-exome captured exRNA in plasma from 172 Alzheimer's Disease (AD) and 166 non-cognitive control (NCI) samples obtained from a previous study1. A Percentage of mapped reads overlapping RNA biotypes in AD and NCI. **B** Percentage of mapped reads overlapping gene features in AD and NCI.

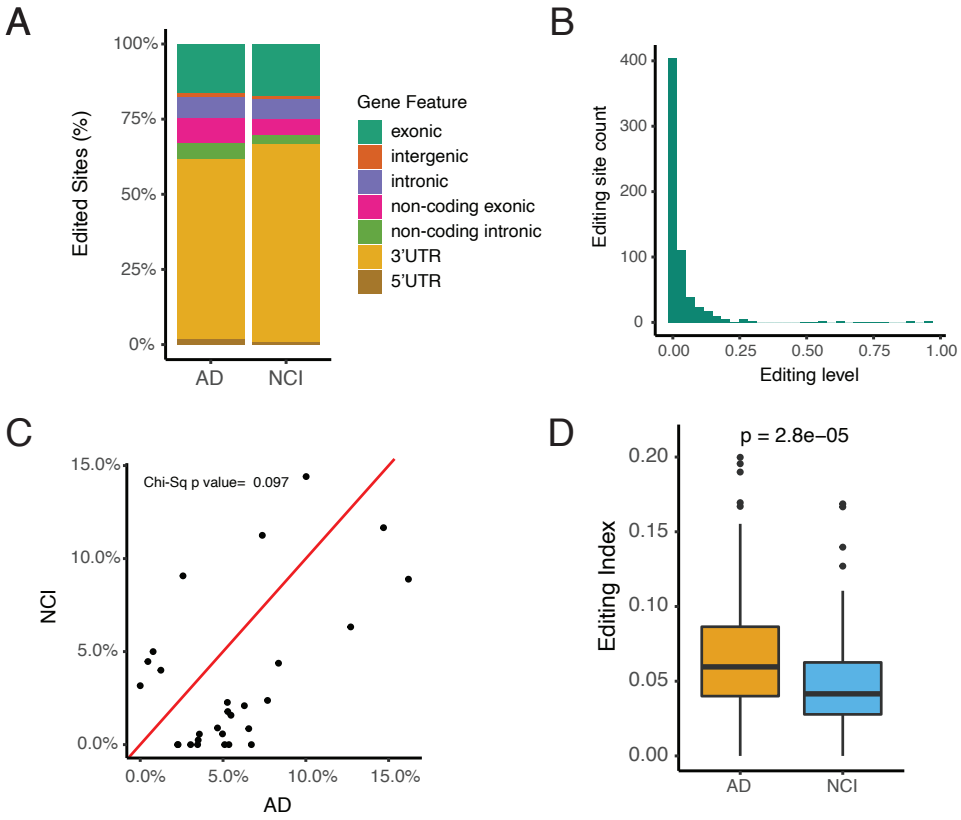



**Figure 4.2 Extracellular RNA editing events are elevated identified in plasma from AD.**

**A** Percentage of RNA editing sites overlapping gene features. **B** Histogram of the editing level of all edited sites identified from the REDIportal database across AD and NCI samples. **C** RNA editing levels of the 28 differentially edited sites in AD and NCI. **D** Sample RNA editing index of the differentially edited sites across AD and NCI samples. P values between AD and NCI were calculated via Wilcoxon rank-sum test.

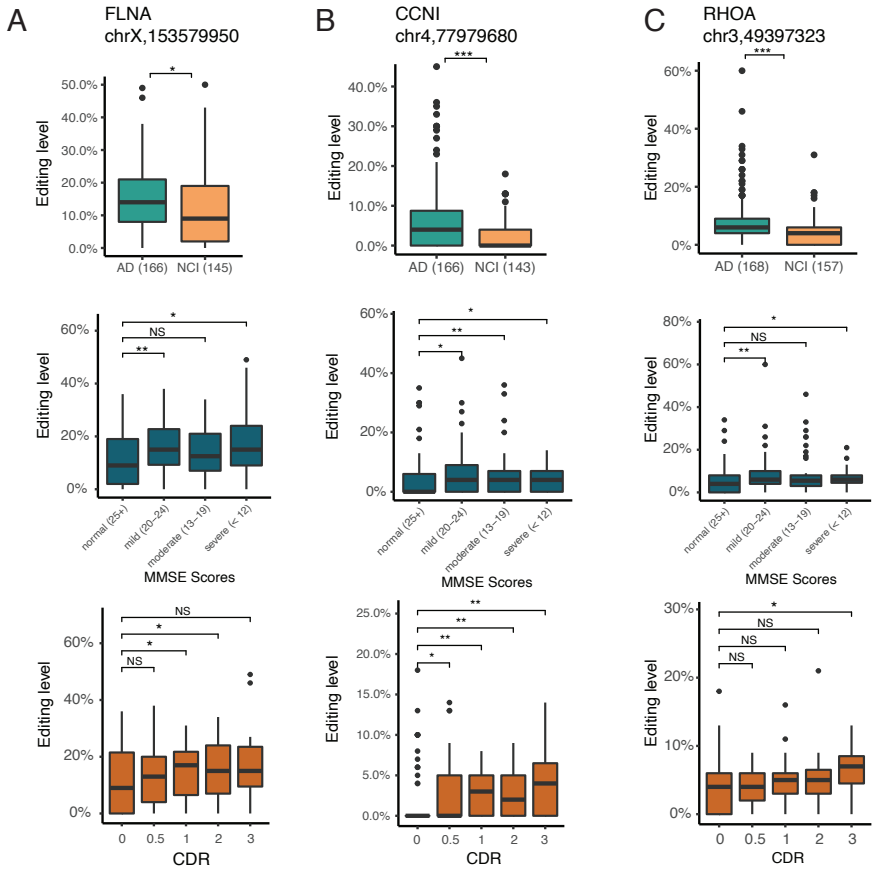



**Figure 4.3 Differential RNA editing levels correspond with cognitive scores.**

**A-C** Editing level of differentially edited sites with the number of samples expressed shown in parenthesis (upper panel). Editing level and binned Mini-Mental State Exam (MMSE) scores for dementia of each sample (middle panel). Editing level of differentially edited site and the Clinical dementia rating (CDR) scores for each sample ranging from 0 (normal) to 3 (severe dementia). Significance of p-values are shown as

NS: p > 0.05, * p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001. P-values were calculated via Wilcoxon rank-sum test.

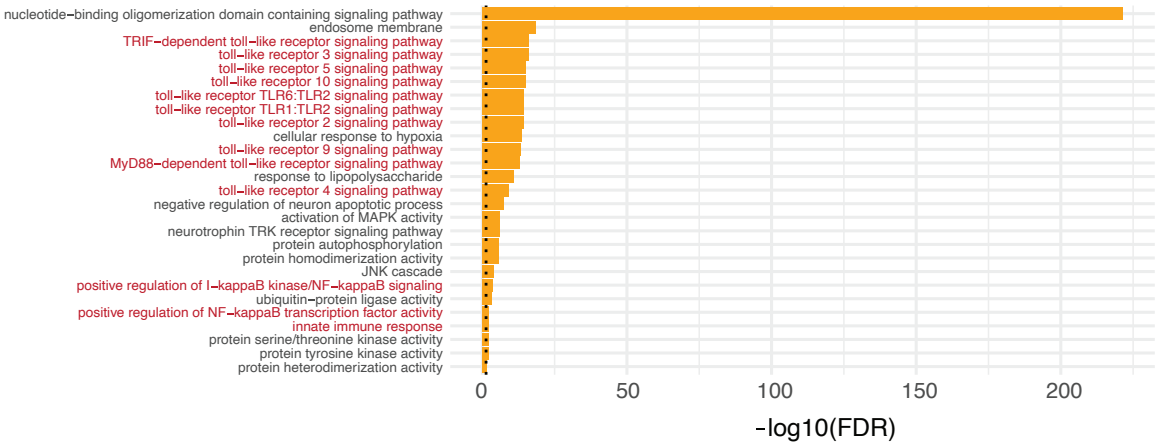
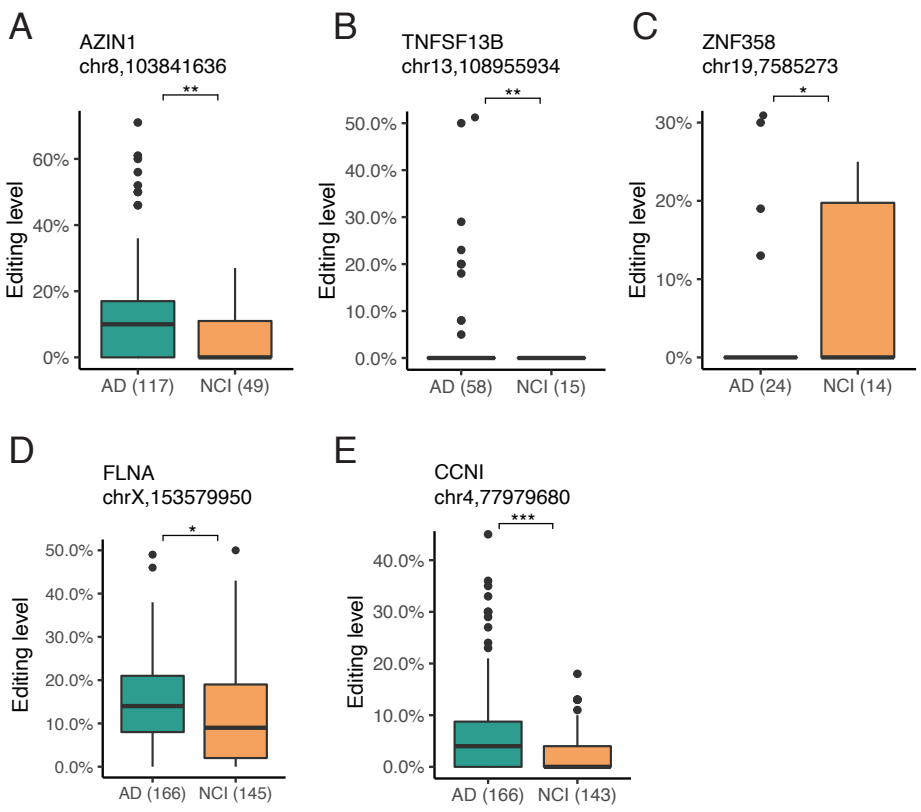



**Figure 4.4 Genes harboring differential RNA editing sites enrich for immune signaling terms.**

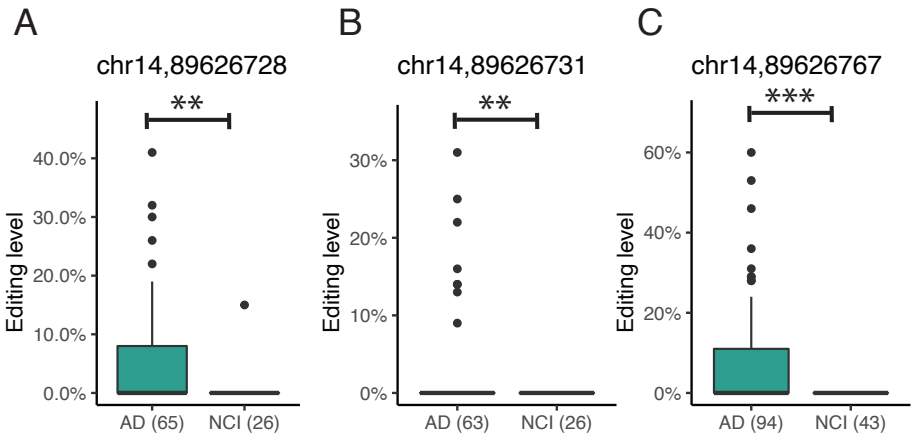
Gene ontology enrichment of the genes harboring differentially edited sites. Background genes match for ±10% gene length and expression. Parent GO terms were removed (Methods). Red text color indicates toll-like receptor and immune signaling categories.

## 4.6 Supplementary Figures




**Supplementary Figure S4.1 Differentially edited recoding sites in plasma from AD and control samples.**

**A-E** Differentially edited sites identified from REDITs on common sites in 10% of samples with a minimum effect size of 2% (Methods). Samples in AD or NCI are shown in parenthesis. Significance of FDR corrected p-values from REDITs are shown as NS: $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.




**Supplementary Figure S4.2 Differentially edited sites in the 3'UTR of FOXN3.**

**Supplementary Fig S4.2 Differentially edited sites in the 3'UTR of FOXN3**

**A-C** Differentially edited sites in the 3'UTR of FOXN3 identified from REDITs on common sites in 10% of samples with a minimum effect size of 2% (Meth- ods). Samples in AD or NCI are shown in parenthesis. Significance of FDR corrected p-values from REDITs are shown as NS: $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

## 4.7 Supplementary Tables

| | gene | RNA_region | name | AD_editing | NCI_editing | AD_samples | NCI_samples | FDR | recoding |
|---|---|---|---|---|---|---|---|---|---|
| 1 | RHOA | UTR3 | chr3,49397323 | 0.0835 | 0.0438 | 168 | 157 | 3.073e−13 | No |
| 2 | FOXN3 | UTR3 | chr14,89626767 | 0.0671 | 0.0000 | 94 | 43 | 3.073e−13 | No |
| 3 | CCNI | exonic | chr4,77979680 | 0.0629 | 0.0209 | 166 | 143 | 7.109e−11 | Yes |
| 4 | SELT | UTR3 | chr3,150345451 | 0.0655 | 0.0086 | 112 | 35 | 1.973e−06 | No |
| 5 | IFI30 | exonic | chr19,18284747 | 0.0526 | 0.0227 | 152 | 119 | 3.211e−06 | No |
| 6 | ARHGAP1 | UTR3 | chr11,46700510 | 0.0350 | 0.0025 | 82 | 53 | 1.400e−05 | No |
| 7 | EIF4EBP2 | UTR5 | chr10,72164154 | 0.0467 | 0.0090 | 51 | 48 | 4.676e−04 | No |
| 8 | FOXN3 | UTR3 | chr14,89626728 | 0.0497 | 0.0058 | 65 | 26 | 8.332e−04 | No |
| 9 | FOXN3 | UTR3 | chr14,89626731 | 0.0229 | 0.0000 | 63 | 26 | 1.320e−03 | No |
| 10 | CARD8 | UTR5 | chr19,48753070 | 0.0000 | 0.0317 | 34 | 12 | 1.320e−03 | No |
| 11 | TNFSF13B | exonic | chr13,108955934 | 0.0347 | 0.0000 | 58 | 15 | 6.847e−03 | Yes |
| 12 | AZIN1 | exonic | chr8,103841636 | 0.1271 | 0.0633 | 117 | 49 | 6.974e−03 | Yes |
| 13 | IKZF3 | UTR3 | chr17,37920136 | 0.0304 | 0.0000 | 23 | 15 | 6.974e−03 | No |
| 14 | SON | exonic | chr21,34923256 | 0.0535 | 0.0000 | 31 | 12 | 7.167e−03 | No |
| 15 | AB209315 | UTR3 | chr20,43707759 | 0.1002 | 0.1441 | 51 | 22 | 1.444e−02 | No |
| 16 | IRAK1 | UTR3 | chrX,153276566 | 0.0510 | 0.0000 | 30 | 14 | 1.444e−02 | No |
| 17 | UBE2K | intronic | chr4,39735046 | 0.0046 | 0.0447 | 41 | 15 | 2.303e−02 | No |
| 18 | VPS37B | ncRNA_intronic | chr12,123350793 | 0.0357 | 0.0057 | 77 | 23 | 2.350e−02 | No |
| 19 | C5AR1 | UTR3 | chr19,47824657 | 0.0078 | 0.0500 | 54 | 5 | 3.182e−02 | No |
| 20 | FLNA | exonic | chrX,153579950 | 0.1469 | 0.1166 | 166 | 145 | 3.413e−02 | Yes |
| 21 | CCDC97 | UTR3 | chr19,41828971 | 0.0529 | 0.0178 | 49 | 27 | 4.211e−02 | No |
| 22 | UBC | exonic | chr12,125396488 | 0.1620 | 0.0889 | 66 | 56 | 4.256e−02 | No |
| 23 | APOL6 | UTR3 | chr22,36056564 | 0.0548 | 0.0158 | 50 | 19 | 4.256e−02 | No |
| 24 | RNF125 | ncRNA_intronic | chr18,29651340 | 0.0226 | 0.0000 | 27 | 12 | 4.256e−02 | No |
| 25 | ZNF358 | exonic | chr19,7585273 | 0.0258 | 0.0907 | 24 | 14 | 4.256e−02 | Yes |
| 26 | FHDC1 | UTR3 | chr4,153898316 | 0.0737 | 0.1125 | 94 | 72 | 4.256e−02 | No |
| 27 | CTSS | UTR3 | chr1,150705221 | 0.0769 | 0.0238 | 58 | 16 | 4.256e−02 | No |
| 28 | SOD2 | UTR3 | chr6,160101805 | 0.0125 | 0.0400 | 93 | 17 | 4.273e−02 | No |

**Supplementary Table S4.1 Differentially edited sites between AD and control samples.**

Table of the 28 differentially edited sites identified from REDITS (Methods). The name of the editing site is the chromosome and the position. AD_editing and NCI_editing are the average editing level of samples with sufficient coverage.

# Chapter 5 : Concluding Remarks

Post-transcriptional modifications are known to be prevalent in intracellular RNAs, with functional roles in gene regulation and disease mechanisms[102]. However, they remain poorly studied in extracellular RNAs. In this work, we provided a comprehensive landscape of two post-transcriptional modifications, 3' end non-templated additions and RNA editing, in human biofluids. We also investigated the enrichment of RBP binding sites in mRNA fragments across biofluids, an aspect of extracellular RNA that has not been well studied.

In Chapter 2, we developed miNTA, a highly accurate bioinformatic method to identify 3' end non-templated additions. miNTA was applied to 1047 extracellular derived small RNA-sequencing datasets from 4 types of biofluids. We identified hundreds of miRNAs with 3' uridylation or adenylation, with the former being more prevalent. We found that 3' uridylation levels enabled segregation of different types of biofluids, more effectively than overall miRNA expression levels. This observation suggests that 3' NTA levels possess fluid-specific information relatively robust to batch effects. In addition, we observed that extracellular miRNAs with 3' uridylations are enriched in processes related to angiogenesis, apoptosis and inflammatory response. Lastly, we showed miRNA 3' uridylation may stabilize base-pairing between miRNAs and their target genes.

In Chapter 3 we examined RBP binding site enrichment in mRNA fragments, expression of repeat elements, and RNA editing in exRNA transcriptomes across 17 biofluids. We observed enrichment of RBP binding sites in mRNA fragments suggesting RBPs may play an important role in the selection and stability of exRNA. Specifically, we identified UCHL5 as one of the RBPs with the strongest binding site enrichment across multiple biofluids. Alu enrichment in urine was also observed in bladder cancer patients and could be considered as a potential biomarker for cancers. Furthermore, we reported RNA editing across biofluids and observed 97 recoding sites including editing in AZIN1, which is known to be associated with hepatocellular carcinoma.

In Chapter 4 we examined RNA editing in 338 extracellular RNA samples from plasma of AD and control subjects. We found 28 differentially edited sites with an overall trend of hyperediting in AD. Differentially edited sites in exRNA of AD were enriched in toll-like receptor signaling and innate immune response terms, which is in line with the known functional roles of RNA editing in innate immunity. Additionally, the editing level of three sites were correlated with cognitive metadata such as CDR and MMSE, suggesting RNA editing may add diagnostic value for AD.

In chapter 2 and in chapter 4 we observed post-transcriptional modification to be enriched in innate immune response and inflammatory terms. In the future, it would be interesting to investigate the roles of exRNA and these post-transcriptional modifications in immune-related systems and upon immune activation. While we address significant gap in knowledge, other important post-transcriptional modifications such as, but not limited to, RNA methylation (e.g. m6A, m1A, m5C), pseudouridylation, and 2'-OMe, have

yet to be extensively profiled in exRNAs and may reveal valuable insights on exRNA biology.

Identifying the cell or tissue origin of exRNA is an important but difficult question in the exRNA field. While several studies have implemented methods to deconvolute compartment-type or tumor contribution for exRNA[18,163], cell type contributions in different biofluids have not been studied. Both exRNA expression and RNA modifications may provide useful information for future deconvolution methods.

In summary, our studies expand the known repertoire of RNA modifications, RBP binding in extracellular RNA, and repeat-derived RNAs, which paves way for further biomarker discoveries. The insights generated in our work builds a foundation for future functional, mechanistic, and translational discoveries

# References


1. Morris, K. V. & Mattick, J. S. The rise of regulatory RNA. *Nat. Rev. Genet.* **15**, 423–437 (2014).

2. Fritz, J. V *et al.* Sources and Functions of Extracellular Small RNAs in Human Circulation. *Annu. Rev. Nutr.* **36**, 301–36 (2016).

3. X, C. *et al.* Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res.* **18**, 997–1006 (2008).

4. O'Brien, K., Breyne, K., Ughetto, S., Laurent, L. C. & Breakefield, X. O. RNA delivery by extracellular vesicles in mammalian cells and its applications. *Nature Reviews Molecular Cell Biology* 1–22 (2020). doi:10.1038/s41580-020-0251-y

5. HN, G. & MT, M. Examining the evidence for extracellular RNA function in mammals. *Nat. Rev. Genet.* **22**, 448–458 (2021).

6. Cai, Q. *et al.* Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes. *Science* **4142**, 1–9 (2018).

7. Vickers, K. C., Palmisano, B. T., Shoucri, B. M., Shamburek, R. D. & Remaley, A. T. MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. *Nat. Cell Biol.* **13**, 423–435 (2011).

8. Wang, K., Zhang, S., Weber, J., Baxter, D. & Galas, D. J. Export of microRNAs and microRNA-protective protein by mammalian cells. *Nucleic Acids Res.* **38**,

7248–7259 (2010).

9. Valadi, H. *et al.* Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **9**, 654–9 (2007).

10. J, S. *et al.* Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat. Cell Biol.* **10**, 1470–1476 (2008).

11. Wei, Z. *et al.* Coding and noncoding landscape of extracellular RNA released by human glioma stem cells. *Nat. Commun.* **8**, 1145 (2017).

12. Hoshino, A. *et al.* Tumour exosome integrins determine organotropic metastasis. *Nature* **527**, 329–335 (2015).

13. Pegtel, D. M. *et al.* Functional delivery of viral miRNAs via exosomes. *Proc. Natl. Acad. Sci. USA* **107**, 6328–6333 (2010).

14. Tkach, M. & Théry, C. Communication by Extracellular Vesicles: Where We Are and Where We Need to Go. *Cell* **164**, 1226–1232 (2016).

15. Gebert, L. F. R. & MacRae, I. J. Regulation of microRNA function in animals. *Nat. Rev. Mol. Cell Biol.* **20**, 21–37 (2019).

16. Yang, A. *et al.* 3′ Uridylation Confers miRNAs with Non-canonical Target Repertoires. *Mol. Cell* (2019). doi:10.1016/J.MOLCEL.2019.05.014

17. Kingston, E. & Bartel, D. Global analyses of the dynamics of mammalian microRNA metabolism. *Genome Res.* 1–40 (2019). doi:10.1101/gr.251421.119

18. Murillo, O. D. *et al.* exRNA Atlas Analysis Reveals Distinct Extracellular RNA Cargo Types and Their Carriers Present across Human Biofluids. *Cell* **177**, 463-477.e15 (2019).

19. Wei, Z. *et al.* Coding and noncoding landscape of extracellular RNA released by human glioma stem cells. *Nat. Commun.* **8**, 1145 (2017).

20. Batagov, A. O. & Kurochkin, I. V. Exosomes secreted by human cells transport largely mRNA fragments that are enriched in the 3′-untranslated regions. *Biol. Direct* **8**, 1–8 (2013).

21. Hinger, S. A. *et al.* Diverse Long RNAs Are Differentially Sorted into Extracellular Vesicles Secreted by Colorectal Cancer Cells. *Cell Rep.* **25**, 715-725.e4 (2018).

22. Yao, J., Wu, D. C., Nottingham, R. M. & Lambowitz, A. M. Identification of protein-protected mRNA fragments and structured excised intron RNAs in human plasma by TGIRT-seq peak calling. *Elife* **9**, 1–41 (2020).

23. Lässer, C. *et al.* Two distinct extracellular RNA signatures released by a single cell type identified by microarray and next-generation sequencing. *RNA Biol.* **14**, 58–72 (2017).

24. JY, L. *et al.* Genomic Repeats Categorize Genes with Distinct Functions for Orchestrated Regulation. *Cell Rep.* **30**, 3296-3311.e5 (2020).

25. Elbarbary, R. A., Lucas, B. A. & Maquat, L. E. Retrotransposons as regulators of gene expression. *Science* **351**, aac7247 (2016).

26. Nishikura, K. A-to-I editing of coding and non-coding RNAs by ADARs. *Nat. Rev.*

*Mol. Cell Biol.* **17**, 83–96 (2015).

27. Eisenberg, E. & Levanon, E. Y. A-to-I RNA editing - immune protector and transcriptome diversifier. *Nat. Rev. Genet.* **19**, 473–490 (2018).

28. M, B. & M, Ö. RNA Editing: A Contributor to Neuronal Dynamics in the Mammalian Brain. *Trends Genet.* **32**, 165–175 (2016).

29. A, G., D, V., D, M., MA, O. & LP, K. ADAR RNA editing in human disease; more to it than meets the I. *Hum. Genet.* **136**, 1265–1278 (2017).

30. Hsiao, Y.-H. E. *et al.* RNA editing in nascent RNA affects pre-mRNA splicing. *Genome Res.* **28**, 812–823 (2018).

31. P, L., S, A. & AI, B. The essential elements of Alzheimer's disease. *J. Biol. Chem.* **296**, (2021).

32. 2021 Alzheimer's disease facts and figures. *Alzheimers. Dement.* **17**, 327–406 (2021).

33. Paraskevaidi, M., Allsop, D., Karim, S., Martin, F. L. & Crean, S. Diagnostic Biomarkers for Alzheimer's Disease Using Non-Invasive Specimens. *J. Clin. Med.* **9**, (2020).

34. K, K. *et al.* Reduced levels of protein recoding by A-to-I RNA editing in Alzheimer's disease. *RNA* **22**, 290–302 (2016).

35. S, W., M, Y., P, K. & X, Z. ADeditome provides the genomic landscape of A-to-I RNA editing in Alzheimer's disease. *Brief. Bioinform.* **22**, (2021).

36. Godoy, P. M. *et al.* Large Differences in Small RNA Composition Between Human Biofluids. *Cell Rep.* **25**, 1346–1358 (2018).

37. Srinivasan, S. *et al.* Small RNA Sequencing across Diverse Biofluids Identifies Optimal Methods for exRNA Isolation. *Cell* **177**, 446-462.e16 (2019).

38. Mitchell, P. S. *et al.* Circulating microRNAs as stable blood-based markers for cancer detection. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 10513–8 (2008).

39. Danielson, K. M. *et al.* Plasma Circulating Extracellular RNAs in Left Ventricular Remodeling Post-Myocardial Infarction. *EBioMedicine* **32**, 172–181 (2018).

40. Schwarzenbach, H., Nishida, N., Calin, G. A. & Pantel, K. Clinical relevance of circulating cell-free microRNAs in cancer. *Nat. Rev. Clin. Oncol.* **11**, 145–156 (2014).

41. Lucero, R. *et al.* Glioma-Derived miRNA-Containing Extracellular Vesicles Induce Angiogenesis by Reprogramming Brain Endothelial Cells. *Cell Rep.* **30**, 2065-2074.e4 (2020).

42. Melo, S. A. *et al.* Cancer Exosomes Perform Cell-Independent MicroRNA Biogenesis and Promote Tumorigenesis. *Cancer Cell* **26**, 707–721 (2014).

43. Thomou, T. *et al.* Adipose-derived circulating miRNAs regulate gene expression in other tissues. *Nature* **542**, 450–455 (2017).

44. Gao, X. *et al.* Gliomas Interact with Non-glioma Brain Cells via Extracellular Vesicles. *Cell Rep.* **30**, 2489-2500.e5 (2020).

45. Ameres, S. L. & Zamore, P. D. Diversifying microRNA sequence and function. *Nat. Rev. Mol. Cell Biol.* **14**, 475–88 (2013).

46. Wyman, S. K. *et al.* Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. *Genome Res.* **21**, 1450–61 (2011).

47. Burroughs, A. M. *et al.* A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. *Genome Res.* **20**, 1398–410 (2010).

48. Thornton, J. E. *et al.* Selective microRNA uridylation by Zcchc6 (TUT7) and Zcchc11 (TUT4). *Nucleic Acids Res.* **42**, 11777–91 (2014).

49. Koppers-Lalic, D. *et al.* Nontemplated nucleotide additions distinguish the small RNA composition in cells from exosomes. *Cell Rep.* **8**, 1649–1658 (2014).

50. Juzenas, S. *et al.* A comprehensive, cell specific microRNA catalogue of human peripheral blood. *Nucleic Acids Res.* **45**, 9290–9301 (2017).

51. Wu, X. *et al.* sRNAnalyzer-a flexible and customizable small RNA sequencing data analysis pipeline. *Nucleic Acids Res.* **45**, 12140–12151 (2017).

52. Felipe Valter de Oliveira, L., Paula Christoff, A. & Margis, R. Sequence analysis isomiRID: a framework to identify microRNA isoforms. *Bioinforma. Discov. NOTE* **29**, 2521–2523 (2013).

53. Aparicio-Puerta, E. *et al.* sRNAbench and sRNAtoolbox 2019: intuitive fast small RNA profiling and differential expression. *Nucleic Acids Res.* **47**, W530–W535

(2019).

54. Kaushik, A., Saraf, S., Mukherjee, S. K. & Gupta, D. miRMOD: a tool for identification and analysis of 5’ and 3’ miRNA modifications in Next Generation Sequencing small RNA data. *PeerJ* **3**, e1332 (2015).

55. Vitsios, D. M. & Enright, A. J. Chimira: analysis of small RNA sequencing data and microRNA modifications. *Bioinformatics* **31**, 3365–7 (2015).

56. D’Ambrogio, A., Gu, W., Udagawa, T., Mello, C. C. & Richter, J. D. Specific miRNA Stabilization by Gld2-Catalyzed Monoadenylation. *Cell Rep.* **2**, 1537–1545 (2012).

57. Burgos, K. *et al.* Profiles of extracellular miRNA in cerebrospinal fluid and serum from patients with Alzheimer’s and Parkinson’s diseases correlate with disease status and features of pathology. *PLoS One* **9**, e94839 (2014).

58. Yuan, T. *et al.* Plasma extracellular RNA profiles in healthy and cancer patients. *Sci. Rep.* **6**, 19413 (2016).

59. Ferrero, G. *et al.* Small non-coding RNA profiling in human biofluids and surrogate tissues from healthy individuals: description of the diverse and most represented species. *Oncotarget* **9**, 3097–3111 (2018).

60. Max, K. E. A. *et al.* Human plasma and serum extracellular small RNA reference profiles and their clinical utility. *Proc. Natl. Acad. Sci. U. S. A.* 1–10 (2018). doi:10.1073/pnas.1714397115

61. Yeri, A. *et al.* Total Extracellular Small RNA Profiles from Plasma, Saliva, and

Urine of Healthy Subjects. *Sci. Rep.* **7**, 44061 (2017).

62. Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W. & Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* **40**, 37–52 (2012).

63. Jeppesen, D. K. *et al.* Reassessment of Exosome Composition. *Cell* **177**, 428-445.e18 (2019).

64. Heo, I. *et al.* Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. *Cell* **151**, 521–32 (2012).

65. Tran, S. S., Zhou, Q. & Xiao, X. Statistical inference of differential RNA editing sites from RNA-sequencing data by hierarchical modeling. *Bioinformatics* (2020). doi:10.1093/bioinformatics/btaa066

66. Krüger, J. & Rehmsmeier, M. RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Res.* **34**, W451-4 (2006).

67. Todorova, D., Simoncini, S., Lacroix, R., Sabatier, F. & Dignat-George, F. Extracellular vesicles in angiogenesis. *Circ. Res.* **120**, 1658–1673 (2017).

68. Esquela-Kerscher, A. & Slack, F. J. Oncomirs - microRNAs with a role in cancer. *Nat. Rev. Cancer* **6**, 259–69 (2006).

69. Koppers-Lalic, D. *et al.* Non-invasive prostate cancer detection by measuring miRNA variants (isomiRs) in urine extracellular vesicles. *Oncotarget* **7**, 22566–78 (2016).

70. Koi, Y. *et al.* Predicting the presence of breast cancer using circulating small RNAs, including those in the extracellular vesicles. *Cancer Sci.* **111**, 2104–2115 (2020).

71. Nostrand, E. L. Van *et al.* A Large-Scale Binding and Functional Map of Human RNA Binding Proteins. *bioRxiv* 1–74 (2018). doi:10.1101/179648

72. Compeau, P. E. C. *et al.* Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).

73. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).

74. Kent, W. J. BLAT--the BLAST-like alignment tool. *Genome Res.* **12**, 656–64 (2002).

75. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–11 (2001).

76. Cancer Genome Atlas Research Network *et al.* The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* **45**, 1113–20 (2013).

77. Wang, M. *et al.* The Mount Sinai cohort of large-scale genomic, transcriptomic and proteomic data in Alzheimer's disease. *Sci. data* **5**, 180185 (2018).

78. Karczewski, K. J. *et al.* The ExAC browser: displaying reference data information from over 60 000 exomes. *Nucleic Acids Res.* **45**, D840–D845 (2017).

79. Karczewski, K. J. *et al.* Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv* 531210 (2019). doi:https://doi.org/10.1101/531210

80. Lee, J., Ang, J. K. & Xiao, X. Analysis and design of RNA sequencing experiments for identifying RNA editing and other single-nucleotide variants. *RNA* **19**, 725–32 (2013).

81. Bahn, J. H. *et al.* Accurate identification of A-to-I RNA editing in human by transcriptome sequencing. *Genome Res.* **22**, 142–50 (2012).

82. de Rie, D. *et al.* An integrated expression atlas of miRNAs and their promoters in human and mouse. *Nat. Biotechnol.* **35**, 872–878 (2017).

83. Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* 1–8 (2018). doi:10.1093/nar/gky1141

84. Huntley, R. P. *et al.* Expanding the horizons of microRNA bioinformatics. *RNA* **44**, In press (2018).

85. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

86. Yagi, Y. *et al.* Next-generation sequencing-based small RNA profiling of cerebrospinal fluid exosomes. *Neurosci. Lett.* **636**, 48–57 (2017).

87. Godoy, P. M. *et al.* Large Differences in Small RNA Composition Between Human Biofluids. *Cell Rep.* **25**, 1346–1358 (2018).

88. Bahn, J. H. *et al.* The landscape of microRNA, Piwi-interacting RNA, and circular RNA in human saliva. *Clin. Chem.* **61**, 221–30 (2015).

89. Toden, S. *et al.* Noninvasive characterization of Alzheimer's disease by circulating, cell-free messenger RNA next-generation sequencinag. *Sci. Adv.* **6**, eabb1654 (2020).

90. Hinger, S. A. *et al.* Diverse Long RNAs Are Differentially Sorted into Extracellular Vesicles Secreted by Colorectal Cancer Cells. *Cell Rep.* **25**, 715–725 (2018).

91. Chen, M. *et al.* Transcriptome and long noncoding RNA sequencing of three extracellular vesicle subtypes released from the human colon cancer LIM1863 cell line. *Sci. Rep.* **6**, 38397 (2016).

92. Savelyeva, A. V. *et al.* Variety of RNAs in Peripheral Blood Cells, Plasma, and Plasma Fractions. *Biomed Res. Int.* **2017**, (2017).

93. Fabbiano, F. *et al.* RNA packaging into extracellular vesicles: An orchestra of RNA-binding proteins? *Journal of Extracellular Vesicles* **10**, (2020).

94. Shurtleff, M. J., Temoche-Diaz, M. M., Karfilis, K. V., Ri, S. & Schekman, R. Y-box protein 1 is required to sort microRNAs into exosomes in cells and in a cell-free reaction. *Elife* **5**, 1–23 (2016).

95. Shurtleff, M. J. *et al.* Broad role for YBX1 in defining the small noncoding RNA composition of exosomes. *Proc. Natl. Acad. Sci.* **114**, E8987–E8995 (2017).

96. Hobor, F. *et al.* A cryptic RNA-binding domain mediates Syncrip recognition and exosomal partitioning of miRNA targets. *Nat. Commun.* **9**, 831 (2018).

97. Villarroya-Beltri, C. *et al.* Sumoylated hnRNPA2B1 controls the sorting of miRNAs into exosomes through binding to specific motifs. *Nat. Commun.* **4**, 1–10 (2013).

98. DM, T. & SE, D. Aberrantly High Levels of Somatic LINE-1 Expression and Retrotransposition in Human Neurological Disorders. *Front. Genet.* **10**, (2020).

99. Saleh, A., Macia, A. & Muotri, A. R. Transposable Elements, Inflammation, and Neurological Disease. *Front. Neurol.* **10**, 894 (2019).

100. J, J.-H. *et al.* Whole-genome sequencing reveals principles of brain retrotransposition in neurodevelopmental disorders. *Cell Res.* **28**, 187–203 (2018).

101. Tan, M. H. *et al.* Dynamic landscape and regulation of RNA editing in mammals. *Nature* **550**, 249–254 (2017).

102. Frye, M., Jaffrey, S. R., Pan, T., Rechavi, G. & Suzuki, T. RNA modifications: what have we learned and where are we headed? *Nat. Rev. Genet.* **17**, 365–372 (2016).

103. Helm, M. & Motorin, Y. Detecting RNA modifications in the epitranscriptome: predict and validate. *Nat. Rev. Genet.* **18**, 275–291 (2017).

104. I, B. & T, K. Role of RNA modifications in cancer. *Nat. Rev. Cancer* **20**, 303–322 (2020).

105. Telonis, A. G. *et al.* Knowledge about the presence or absence of miRNA isoforms (isomiRs) can successfully discriminate amongst 32 TCGA cancer types. *Nucleic Acids Res.* **45**, 2973–2985 (2017).

106. Hulstaert, E. *et al.* Charting Extracellular Transcriptomes in The Human Biofluid RNA Atlas. *Cell Rep.* **33**, 108552 (2020).

107. Chen, L. *et al.* Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma. *Nat. Med.* **19**, 209–216 (2013).

108. Englert, J. A. *et al.* Whole blood RNA sequencing reveals a unique transcriptomic profile in patients with ARDS following hematopoietic stem cell transplantation. *Respir. Res.* **20**, 15 (2019).

109. Hulstaert, E. *et al.* Charting Extracellular Transcriptomes in The Human Biofluid RNA Atlas. *Cell Rep.* **33**, 108552 (2020).

110. M, P. *et al.* Vesiclepedia 2019: a compendium of RNA, proteins, lipids and metabolites in extracellular vesicles. *Nucleic Acids Res.* **47**, D516–D519 (2019).

111. F, M., M, F., H, A., K, W. & M, W. Developmental changes in expression of the three ryanodine receptor mRNAs in the mouse brain. *Neurosci. Lett.* **285**, 57–60 (2000).

112. F, B. *et al.* Loss of Ryanodine Receptor 2 impairs neuronal activity-dependent remodeling of dendritic spines and triggers compensatory neuronal hyperexcitability. *Cell Death Differ.* **27**, (2020).

113. X, L. *et al.* Role of leaky neuronal ryanodine receptors in stress-induced cognitive dysfunction. *Cell* **150**, 1055–1067 (2012).

114. Lenahan, C., Sanghavi, R., Huang, L. & Zhang, J. H. Rhodopsin: A Potential Biomarker for Neurodegenerative Diseases. *Front. Neurosci.* **14**, 326 (2020).

115. Deininger, P. Alu elements: Know the SINEs. *Genome Biol.* **12**, 1–12 (2011).

116. Kong, Y. *et al.* Transposable element expression in tumors is associated with immune infiltration and increased antigenicity. *Nat. Commun.* **10**, (2019).

117. Thomson, S. J. P. *et al.* The role of transposable elements in the regulation of IFN-λ1 gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 11564–11569 (2009).

118. Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61 (2015).

119. Bennett, E. A. *et al.* Active Alu retrotransposons in the human genome. *Genome Res.* **18**, 1875–83 (2008).

120. Di Ruocco, F. *et al.* Alu RNA accumulation induces epithelial-to-mesenchymal transition by modulating miR-566 and is associated with cancer progression. *Oncogene* **37**, 627–637 (2018).

121. Song, B., Shiromoto, Y., Minakuchi, M. & Nishikura, K. The role of RNA editing enzyme ADAR1 in human disease. *Wiley Interdiscip. Rev. RNA* e1665 (2021). doi:10.1002/wrna.1665

122. Mansi, L. *et al.* REDIportal: millions of novel A-to-I RNA editing events from thousands of RNAseq experiments. *Nucleic Acids Res.* **49**, D1012–D1019 (2021).

123. Zhao, J. *et al.* IRESbase: A Comprehensive Database of Experimentally Validated Internal Ribosome Entry Sites. *Genomics. Proteomics Bioinformatics* **18**, 129–139

(2020).

124. Chen, C.-K. *et al.* Structured elements drive extensive circular RNA translation. *Mol. Cell* **18**, 129–139 (2021).

125. Dersh, D., Hollý, J. & Yewdell, J. W. A few good peptides: MHC class I-based cancer immunosurveillance and immunoevasion. *Nat. Rev. Immunol.* **21**, 116–128 (2021).

126. Xie, M., Liu, D. & Yang, Y. Anti-cancer peptides: classification, mechanism of action, reconstruction and modification. *Open Biol.* **10**, 200004 (2020).

127. J, G. *et al.* Ubiquitin carboxyl-terminal hydrolase isozyme L5 inhibits human glioma cell migration and invasion via downregulating SNRPF. *Oncotarget* **8**, 113635–113649 (2017).

128. L, N. *et al.* Ubiquitin carboxyl-terminal hydrolase-L5 promotes TGFβ-1 signaling by de-ubiquitinating and stabilizing Smad2/Smad3 in pulmonary fibrosis. *Sci. Rep.* **6**, (2016).

129. D, L., Z, S., X, W. & L, O. Ubiquitin C-Terminal Hydrolase L5 (UCHL5) Accelerates the Growth of Endometrial Cancer via Activating the Wnt/β-Catenin Signaling Pathway. *Front. Oncol.* **10**, (2020).

130. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).

131. H, L. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics*

**25**, 2078–2079 (2009).

132. KL, H. *et al.* Ensembl 2021. *Nucleic Acids Res.* **49**, D884–D891 (2021).

133. AR, Q. & IM, H. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

134. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

135. M, L. *et al.* Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, (2013).

136. Rainer, J. EnsDb.Hsapiens.v75. *EnsDb.Hsapiens.v75: Ensembl based annotation package* R package version 2.99.0 (2017). Available at: http://bioconductor.org/packages/release/data/annotation/html/EnsDb.Hsapiens.v75.html. (Accessed: 15th September 2021)

137. ENCODE Project *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

138. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).

139. Haeussler, M. *et al.* The UCSC Genome Browser database: 2019 update. *Nucleic Acids Res.* **47**, D853–D858 (2019).

140. Chan, T. W. *et al.* RNA editing in cancer impacts mRNA abundance in immune response pathways. *Genome Biol.* **21**, 268 (2020).

141. Tran, S. S. *et al.* Widespread RNA editing dysregulation in brains from autistic individuals. *Nat. Neurosci.* **22**, 25–36 (2019).

142. Beach, T. G., Monsell, S. E., Phillips, L. E. & Kukull, W. Accuracy of the Clinical Diagnosis of Alzheimer Disease at National Institute on Aging Alzheimer's Disease Centers, 2005–2010. *J. Neuropathol. Exp. Neurol.* **71**, 266 (2012).

143. Li, K. *et al.* Advances, challenges, and opportunities in extracellular RNA biology: insights from the NIH exRNA Strategic Workshop. *JCI insight* **3**, (2018).

144. Lu, K.-C., Zhang, Y. & Song, E. Extracellular RNA: mechanisms of it's transporting into target cells. *ExRNA* **1**, 1–5 (2019).

145. D, S.-W. *et al.* Altered microRNAs related to synaptic function as potential plasma biomarkers for Alzheimer's disease. *Alzheimers. Res. Ther.* **11**, (2019).

146. Kim, K. M. *et al.* Mitochondrial RNA in Alzheimer's Disease Circulating Extracellular Vesicles. *Front. Cell Dev. Biol.* **8**, 1 (2020).

147. Z, Y. *et al.* Presymptomatic Increase of an Extracellular RNA in Blood Plasma Associates with the Development of Alzheimer's Disease. *Curr. Biol.* **30**, 1771-1782.e3 (2020).

148. Nigita, G. *et al.* Tissue and exosomal miRNA editing in Non-Small Cell Lung Cancer. *Sci. Rep.* **8**, 10222 (2018).

149. Zhu, Y. *et al.* Integrative analysis of long extracellular RNAs reveals a detection panel of noncoding RNAs for liver cancer. *Theranostics* **11**, 181 (2021).

150. Hosaka, T. *et al.* ADAR2-dependent A-to-I RNA editing in the extracellular linear and circular RNAs. *Neurosci. Res.* **147**, 48–57 (2019).

151. Zhang, Q. & Xiao, X. Genome sequence-independent identification of RNA editing sites. *Nat. Methods* **12**, 347–50 (2015).

152. Annese, A. *et al.* Whole transcriptome profiling of Late-Onset Alzheimer's Disease patients provides insights into the molecular changes involved in the disease. *Sci. Rep.* **8**, 4282 (2018).

153. Pritchard, C. C. *et al.* Blood Cell Origin of Circulating MicroRNAs: A Cautionary Note for Cancer Biomarker Studies. *Cancer Prev. Res.* **5**, 492–497 (2012).

154. NM, M. *et al.* The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. *Cell Rep.* **9**, 1482–1494 (2014).

155. Kong, X. *et al.* Recent Advances in Understanding FOXN3 in Breast Cancer, and Other Malignancies. *Front. Oncol.* **9**, 234 (2019).

156. Hentze, M. W., Castello, A., Schwarzl, T. & Preiss, T. A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* (2018). doi:10.1038/nrm.2017.130

157. JR, L., TA, Y. & JA, S. Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 9667–9672 (2007).

158. SW, L., HJ, C., CL, L. & KF, L. No correlation between Alzheimer's disease and risk of hepatocellular carcinoma in older people: an observation in Taiwan. *Geriatr. Gerontol. Int.* **14**, 231–232 (2014).

159. M, Z. *et al.* RNA editing derived epitopes function as cancer antigens to elicit immune responses. *Nat. Commun.* **9**, (2018).

160. Majd, S., Power, J. & Majd, Z. Alzheimer's Disease and Cancer: When Two Monsters Cannot Be Together. *Front. Neurosci.* **13**, (2019).

161. Stankiewicz, T. R. & Linseman, D. A. Rho family GTPases: key players in neuronal development, neuronal survival, and neurodegeneration. *Front. Cell. Neurosci.* **8**, (2014).

162. Y, K. *et al.* Involvement of RhoA and possible neuroprotective effect of fasudil, a Rho kinase inhibitor, in NMDA-induced neurotoxicity in the rat retina. *Brain Res.* **1018**, 111–118 (2004).

163. Shi, A. *et al.* Plasma-derived extracellular vesicle analysis and deconvolution enable prediction and tracking of melanoma checkpoint blockade outcome. *Sci. Adv.* **6**, 1–13 (2020).