

UC San Diego

UC San Diego Previously Published Works

Title

Small RNA Sequencing across Diverse Biofluids Identifies Optimal Methods for exRNA Isolation

Permalink

<https://escholarship.org/uc/item/2dc3s5wf>

Journal

Cell, 177(2)

ISSN

0092-8674

Authors

Srinivasan, Srimeenakshi
Yeri, Ashish
Cheah, Pike See
et al.

Publication Date

2019-04-01

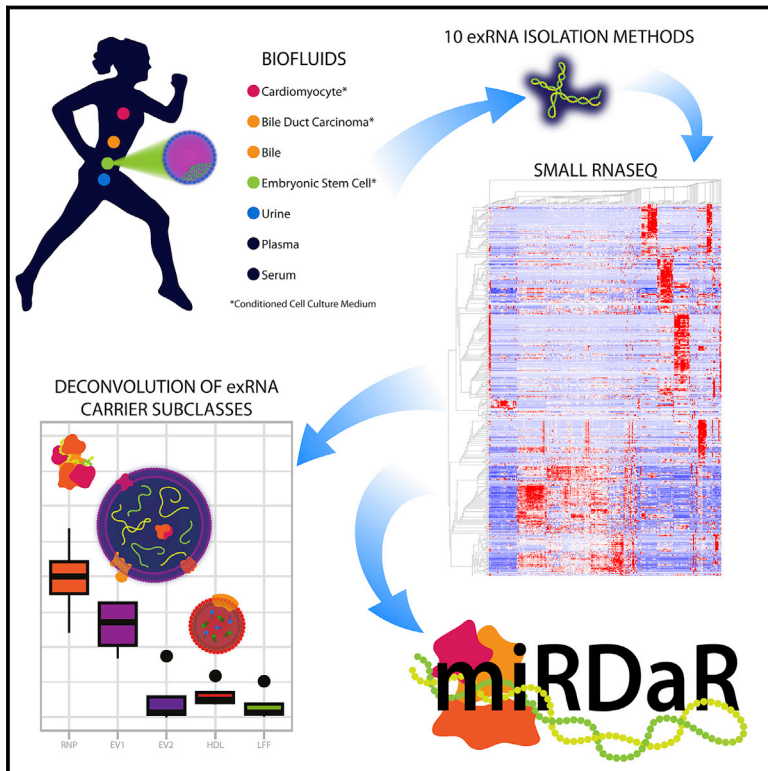
DOI

10.1016/j.cell.2019.03.024

Peer reviewed

Small RNA Sequencing across Diverse Biofluids Identifies Optimal Methods for exRNA Isolation

Graphical Abstract



Authors

Srimeenakshi Srinivasan, Ashish Yeri, Pike See Cheah, ..., Roger P. Alexander, Saumya Das, Louise C. Laurent

Correspondence

llaurent@ucsd.edu

In Brief

A systematic comparison of 10 extracellular RNA isolation methods across 5 biofluids will aid researchers in selecting optimal approaches for individual studies with the overall goal of enhancing reliability and reproducibility for a rapidly growing field.

Highlights

- exRNA sequencing complexity and reproducibility varies across isolation methods
- Deconvolution shows differential access to exRNA carriers by different methods
- Performance of exRNA isolation methods vary across biofluids and RNA species
- miRDaR enables customized selection of optimal exRNA isolation methods



Small RNA Sequencing across Diverse Biofluids Identifies Optimal Methods for exRNA Isolation

Srimeenakshi Srinivasan,¹ Ashish Yeri,² Pike See Cheah,^{3,4,17} Allen Chung,^{5,17} Kirsty Danielson,^{2,17} Peter De Hoff,^{1,17} Justyna Filant,^{6,17} Clara D. Laurent,^{1,2,17} Lucie D. Laurent,^{1,17} Rogan Magee,^{7,17} Courtney Moeller,^{1,17} Venkatesh L. Murthy,^{8,17} Parham Nejad,^{9,17} Anu Paul,^{9,17} Isidore Rigoutsos,^{7,17} Rodosthenis Rodosthenous,^{2,17} Ravi V. Shah,^{2,17} Bridget Simonson,^{2,17} Cuong To,^{1,17} David Wong,^{5,17} Irene K. Yan,^{10,17} Xuan Zhang,^{3,17} Leonora Balaj,^{3,11,18} Xandra O. Breakefield,^{3,18} George Daaboul,^{12,18} Roopali Gandhi,^{9,18} Jodi Lapidus,^{13,18} Eric Londin,^{7,18} Tushar Patel,^{10,18} Robert L. Raffai,^{5,18} Anil K. Sood,^{6,14,15,18} Roger P. Alexander,^{16,18} Saumya Das,^{2,18} and Louise C. Laurent^{1,19,*}

¹Department of Obstetrics, Gynecology, and Reproductive Sciences and Sanford Consortium for Regenerative Medicine, University of California, San Diego, La Jolla, CA, USA

²Cardiovascular Research Center, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

³Neurology and Radiology Services and Program in Neuroscience, Harvard Medical School, Massachusetts General Hospital, Boston, MA, USA

⁴Department of Human Anatomy, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Malaysia

⁵Department of Surgery, University of California, San Francisco and VA Medical Center San Francisco, San Francisco, CA 94121, USA

⁶Department of Gynecologic Oncology and Reproductive Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

⁷Computational Medicine Center, Sidney Kimmel College of Medicine, Thomas Jefferson University, Philadelphia, PA, USA

⁸Department of Medicine, Division of Cardiovascular Medicine, Frankel Cardiovascular Center, University of Michigan, Ann Arbor, MI, USA

⁹Department of Neurology, Ann Romney Center for Neurologic Diseases, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

¹⁰Mayo Clinic Florida, Jacksonville, FL, USA

¹¹Department of Neurosurgery, Harvard Medical School, Massachusetts General Hospital, Boston, MA, USA

¹²NanoView Biosciences, Boston, MA, USA

¹³Oregon Health Sciences University, Portland, OR, USA

¹⁴Center for RNA Interference and Non-Coding RNAs, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

¹⁵Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

¹⁶Pacific Northwest Research Institute, Seattle, WA 98122, USA

¹⁷These authors contributed equally

¹⁸Senior author

¹⁹Lead Contact

*Correspondence: l Laurent@ucsd.edu

<https://doi.org/10.1016/j.cell.2019.03.024>

SUMMARY

Poor reproducibility within and across studies arising from lack of knowledge regarding the performance of extracellular RNA (exRNA) isolation methods has hindered progress in the exRNA field. A systematic comparison of 10 exRNA isolation methods across 5 biofluids revealed marked differences in the complexity and reproducibility of the resulting small RNA-seq profiles. The relative efficiency with which each method accessed different exRNA carrier subclasses was determined by estimating the proportions of extracellular vesicle (EV)-, ribonucleoprotein (RNP)-, and high-density lipoprotein (HDL)-specific miRNA signatures in each profile. An interactive web-based application (miRDaR) was developed to help investigators select the optimal exRNA isolation method for their studies. miRDaR provides comparative statistics for all expressed miRNAs or a selected subset of miRNAs in the desired biofluid for each exRNA isolation method and returns a ranked list of exRNA isolation methods prioritized by complexity, expression level, and repro-

ducibility. These results will improve reproducibility and stimulate further progress in exRNA biomarker development.

INTRODUCTION

The discovery of extracellular RNAs (exRNAs) in biofluids has sparked considerable interest in their biological functions and potential clinical applications. Diagnostic, prognostic, and therapeutic exRNA biomarkers have been reported for a variety of diseases. However, few have been validated across studies, likely due to methodological differences among studies, including differences exRNA isolation and measurement techniques. A comprehensive understanding of the variability associated with these procedures is therefore a prerequisite for improving the reproducibility of exRNA biomarker studies.

Multiple sources of heterogeneity influence the exRNA profiles obtained from biofluid samples. Most biofluids contain exRNAs derived from multiple cell types and even multiple tissues. Therefore, biological variables, such as the age, fitness, and state of individual organs may impact exRNA profiles. In addition, exRNAs can be associated with different carrier subclasses, including ribonucleoprotein complexes (RNPs), lipoprotein complexes (LPPs),



and extracellular vesicles (EVs) (Skog et al., 2008; Turchinovich et al., 2011; Valadi et al., 2007; Vickers et al., 2011), which carry different protein and RNA cargo (Karimi et al., 2018; Lässer et al., 2017) and may be present in different proportions based on environmental and epigenetic factors. exRNA isolation methods that differentially extract exRNAs from the various carrier subclasses and measurement techniques with different sequence biases may also be sources of variability. Systematic investigations to identify differences among exRNA isolation and measurement techniques would address this variability and improve the reproducibility of studies in this field. This collaborative study involving six laboratories in phase 1 of the Extracellular RNA Communication Consortium (ERCC1) focused on exRNA isolation, while other ERCC1 studies compared exRNA profiling methods (Giraldez et al., 2018; Yeri et al., 2018).

Gene expression profiling of different tissues using RNA sequencing (RNA-seq) has allowed investigators to examine cellular processes in health and disease in unprecedented detail. Although standard RNA-seq methods perform poorly on exRNA samples, small RNA-seq has been successfully used to profile exRNA isolated using several methods from different biofluids (Burgos et al., 2013; Cheng et al., 2014; Freedman et al., 2016; Li et al., 2015, 2018; Shah et al., 2017; Tang et al., 2017; Williams et al., 2013; Yeri et al., 2017). Small RNA-seq thus offers the possibility of unbiased discovery of disease-specific exRNA biomarkers. Recent studies demonstrating the generally high intra- and interlab reproducibility of commonly used small RNA-seq methods on exRNA and other low-input samples (Giraldez et al., 2018; Yeri et al., 2018), as well as the popularity of small RNA-seq for exRNA biomarker studies, led us to focus on small RNA-seq as the primary readout for this study.

The utility of exRNA profiling for biomarker discovery or for mechanistic studies is critically dependent on the notion that the effects of biological variables will not be obscured by variability from technical factors, such as those arising from exRNA isolation or analysis. Another key assumption is that candidate exRNA biomarkers will be efficiently isolated and measured by the methods employed, without significant variance across replicates or among test sites. Here, we report results from a comprehensive and rigorous study encompassing many of the available exRNA isolation methods on standardized samples of five biofluids across multiple test sites. Our studies reveal that exRNA carrier subclasses are preferentially enriched by specific isolation methods. We further characterize the success rate and reproducibility of each exRNA isolation method for each biofluid and provide an interactive web-based tool that investigators can use to select the optimal methodology for their study. We hope that our findings will inform future studies in exRNA biology, aid in improving the rigor and reproducibility of exRNA analyses, and smooth the path to discovery and application of disease-specific exRNAs as clinically relevant biomarkers.

RESULTS

Biofluid Samples

Standardized biofluid samples were used for exRNA isolation experiments (Figure S1). To avoid potential artifacts arising from interindividual differences, pooled samples of plasma, serum,

and urine from 10 female or 10 male individuals were included. Bile was collected from three donors with suspected cholangiocarcinoma (P1–P3) and one healthy adult donor (P4). Cell culture conditioned medium (CCCM) was collected from three cell types in three labs: human embryonic stem cells (hESCs, Lab5), primary neonatal rat ventricular myocytes (NRVMs, Lab2), and human cholangiocarcinoma cells (KMBCs, Lab6). Standardized serum, plasma, and urine samples were collected by Lab5 and consisted of one female pool, one female individual, one male pool, and one male individual sample each.

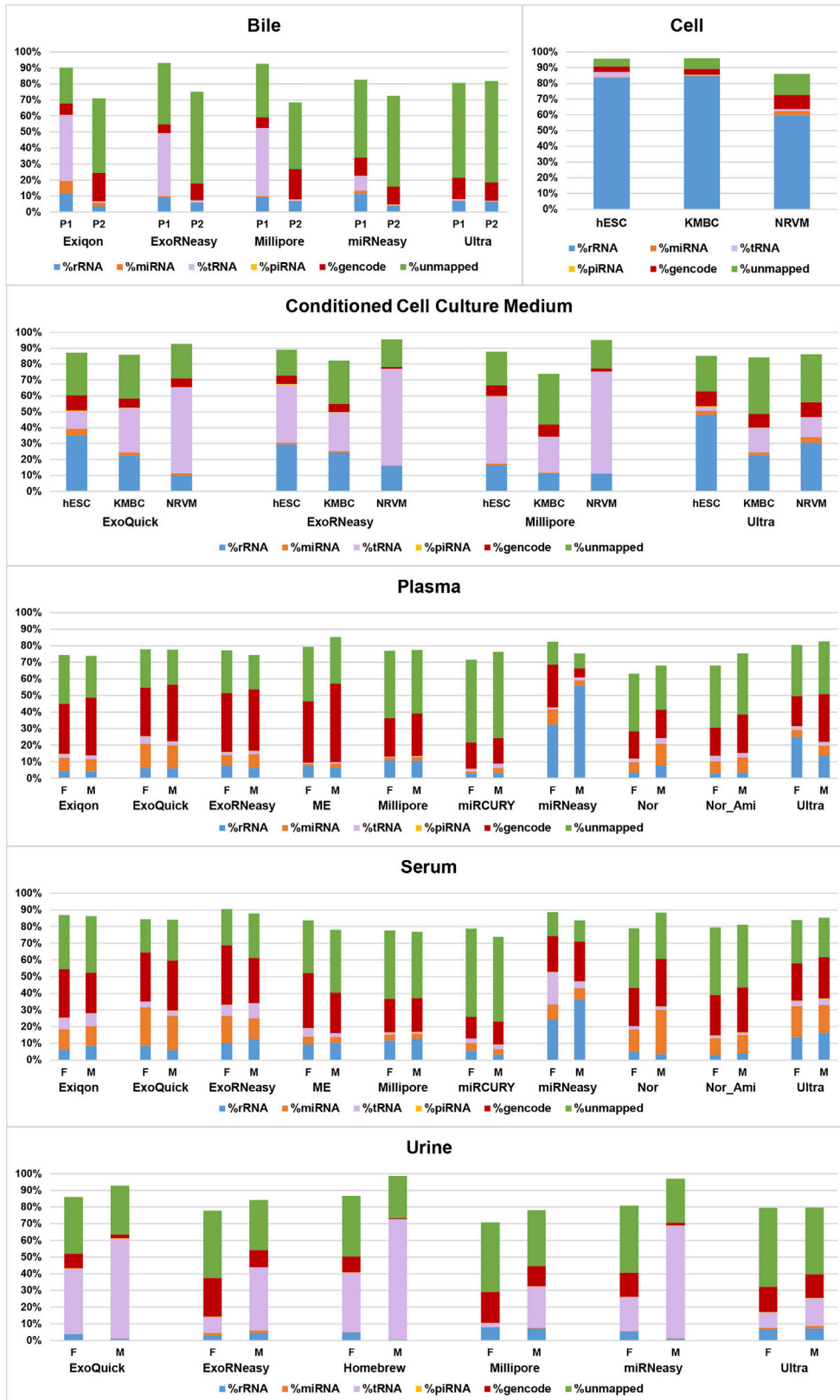
exRNA Isolation Methods

Multiple exRNA isolation methods were tested for each biofluid. Selection of the methods was based on compatibility with the biofluid type and volume per manufacturers' protocols. Methods designed to isolate total exRNA include miRNeasy micro (QIAGEN, referred to as "miRNeasy"), miRCURY Biofluids (Exiqon, "Exiqon"), and plasma/serum exosome purification and exRNA isolation kit with/without RNA concentration using an Amicon 3K filter (Norgen Biotek, "Nor_Ami" and "Nor," respectively). Methods with pre-enrichment of exRNA carriers were: precipitation—ExoQuick+Seramir (SBI, "ExoQuick") and miRCURY Exosome (Exiqon, "miRCURY"), sequential membrane filtration (Millipore, "Millipore"), ultracentrifugation ("Ultra"), and affinity purification—ExoRNeasy (QIAGEN; "ExoRNeasy") and ME (New England Peptide; "ME"). miRNeasy was used to isolate RNA from the material enriched by the ExoRNeasy, Millipore, Ultra, and ME methods.

exRNA isolation was performed in triplicate at each participating lab (Figure S1). For bile (200 μ L sample input volume), exRNA isolation was performed in one lab using five methods: ExoRNeasy, miRNeasy, Exiqon, Ultra, and Millipore. exRNA isolation from CCCM (4 mL input) was performed in three labs using four methods: ExoRNeasy, Ultra, Millipore, and ExoQuick. For serum and plasma (500 μ L input), three exRNA isolation methods (ExoRNeasy, miRNeasy, Exiqon) were tested in four (plasma) or three (serum) labs, six methods (Nor, Nor_Ami, Ultra, Millipore, and ExoQuick) were tested in two labs, and one method (miRCURY) was tested in one lab. For urine (500 μ L input), exRNA isolation was performed in one lab using six methods: ExoQuick, ExoRNeasy, miRNeasy, Homebrew, Millipore, miRNeasy, and Ultra.

RNA Size Distribution and Yield Vary According to Biofluid Type and exRNA Isolation Method

The RNA size distributions ascertained using the bioanalyzer RNA pico chip differed among biofluids (except for plasma and serum), exRNA isolation methods, and (for CCCM) source cell lines (Figure S2). All serum and plasma samples displayed a bimodal distribution, with shorter (<200 nt) RNAs predominating in ExoQuick and ExoRNeasy samples, longer (>200 nt) RNAs predominating in miRNeasy samples, and both peaks being well represented in Exiqon, Nor, and Nor_Ami samples. There was greater diversity in the RNAs size distributions for CCCM, with marked differences among both exRNA isolation methods and source cell lines but good reproducibility across labs. The overall yield of exRNA was higher for hESCs and NRVMs compared to KMBCs. The Millipore and ExoRNeasy methods



(legend on next page)

favored isolation of shorter RNAs, while ExoQuick was more efficient at isolating longer RNAs, especially for hESC and NRVM. Ultra showed the most variable results, with mostly shorter RNAs in the samples from Lab2, both short RNAs and full-length 18S and 28S rRNAs from Lab5, and poor overall yield for Lab6. In urine, the RNA size distributions were similar, but the yields varied across methods, with ExoQuick, Ultra, and Millipore having lower yields than ExoRNeasy, miRNeasy and Homebrew. We measured exRNA concentrations using RiboGreen (ThermoFisher), but the results were variable and often showed undetectable levels of RNA, even for samples in which RNA was clearly detected by Bioanalyzer, qRT-PCR, and small RNA-seq. The use of RiboGreen for quantification of isolated exRNA is therefore not recommended by the investigators.

qRT-PCR Identifies Sources of Variability in exRNA Isolation and Quantification

Quantification of three miRNAs was performed: let-7a-5p (since let-7 miRNAs were reported to be selectively secreted in exosomes; [Beltrami et al., 2017](#); [Chevillet et al., 2014](#); [Ohshima et al., 2010](#)), miR-16-5p (reported to be associated with AGO1- and AGO2-containing RNPs in plasma; [Turchinovich and Burwinkel, 2012](#)), and miR-223-3p (also associated with AGO-containing RNPs; [Turchinovich and Burwinkel, 2012](#)), with each lab analyzing the exRNA samples isolated in that lab in triplicate.

qRT-PCR data were analyzed using ANOVA, with the variability calculated as the sum of squares divided by the total error ([Table S1](#)). For bile, exRNA isolation was performed in a single lab, and the largest source of variability was sample donor (P1–P4), followed closely by exRNA isolation method, with variation among the three miRNAs being markedly lower. For CCCM, lab was the largest source of variability, distantly followed by target miRNA, exRNA isolation method, and source cell line. The serum and plasma qRT-PCR data were analyzed together in two ways. First, we included the variables biofluid type (serum versus plasma), exRNA isolation method, target miRNA, and lab; subsequently, we added sex and sample (individual versus pooled) as additional covariates. For both analyses, the largest source of variability was target miRNA, followed by lab and exRNA isolation method. For the first analysis, biofluid type contributed little to the variability in the data, and for the second analysis, biofluid type, sex, and sample made negligible contributions. For urine, exRNA isolation was performed in one lab, and the largest source of variability was isolation method, followed by sex of the biofluid donor, with target miRNA and sample making negligible contributions. The residual for urine was large (76.8%), indicating that most of the observed variability was not explained by known variables.

These results showed that when examining qRT-PCR data on a small number of miRNAs, the exRNA isolation method was a substantial source of variability for all biofluids except CCCM. In all multi-lab experiments, the lab variable contributed significantly to variability. On detailed inspection of the data, we noted

that the two labs that used the same qRT-PCR machine produced the most similar results. This observation suggested that the qRT-PCR process, rather than the exRNA isolation process, might have been the major source of lab-to-lab variability.

For CCCM and serum/plasma, target miRNA was also an important source of variability, suggesting that the three miRNAs assayed may be carried in different compartments in these biofluids. However, we felt that rigorous assessment of exRNA carrier subclasses would require assessment of a larger number of miRNAs and thus proceeded to perform small RNA-seq on the exRNA samples. Since the lab-to-lab variability in qRT-PCR results might have arisen during exRNA measurement rather than exRNA isolation, we performed small RNA-seq library preparation and sequencing for all exRNA samples in one laboratory (Lab5).

Overall Small RNA-Seq Results

A subset of exRNA samples were selected for small RNA sequencing. For plasma, serum, and CCCM, all three replicate exRNA isolations from one lab and one replicate from each of the other labs were sequenced. Since sample type did not contribute to variability in the qRT-PCR data from plasma and serum, we focused on the pooled samples for these biofluids. For bile, samples from two out of four donors were analyzed, as insufficient material remained from the other two cases after the qRT-PCR experiments for small RNA-seq. For urine, all samples from the female and male pools, one female individual, and one male individual were analyzed. The NEBNext small RNA-seq library preparation kit was used for library construction.

Metadata and summary small RNA-seq library statistics are provided in [Table S2.1](#). The initial mapped miRNA dataset ([Table S2.2](#)) was processed through a series of filtering and normalization steps ([Table S2.3–11](#)), and the number of libraries passing each step was tracked ([Table S2.12](#)).

Distribution of RNA Biotypes Varies by Biofluid, Donor/Cell Line, and exRNA Isolation Method

We calculated the percent of reads mapped to these RNA biotypes: rRNA, miRNA, tRNA, piwi-interacting RNA (piRNA), other GENCODE sequences, and unmapped reads ([Figure 1](#)) and compared the percent of each RNA biotype across methods within each biofluid type ([Table S2.13–17](#)). For bile, the %miRNA was significantly higher for Exiqon compared to all other methods, and %tRNA was significantly higher for P1 than P2. The distribution of RNA biotypes was similar among all cellular RNA samples, with a predominance of rRNA sequences. For CCCM, the %tRNA was significantly lower for Ultra than the other methods, and the %miRNA was higher for ExoQuick and Ultra than ExoRNeasy and Millipore. The distributions of RNA biotypes were similar between plasma and serum, with the miRNeasy and Ultra samples displaying significantly higher %rRNA compared to other methods and ME, Millipore, and miRCURY showing significantly lower %miRNA compared to Exiqon, ExoQuick, ExoRNeasy, miRNeasy,

Figure 1. Distribution of RNA Biotypes

Distributions of rRNA, miRNA, tRNA, piRNA, other GENCODE transcripts and unmapped reads are shown for libraries in the “miRNA_AllBiofluid_TMR” dataset. Values are averaged across all pass-filter replicate libraries.

See also [Figures S1](#) and [S2](#) and [Tables S1](#) and [S2](#).

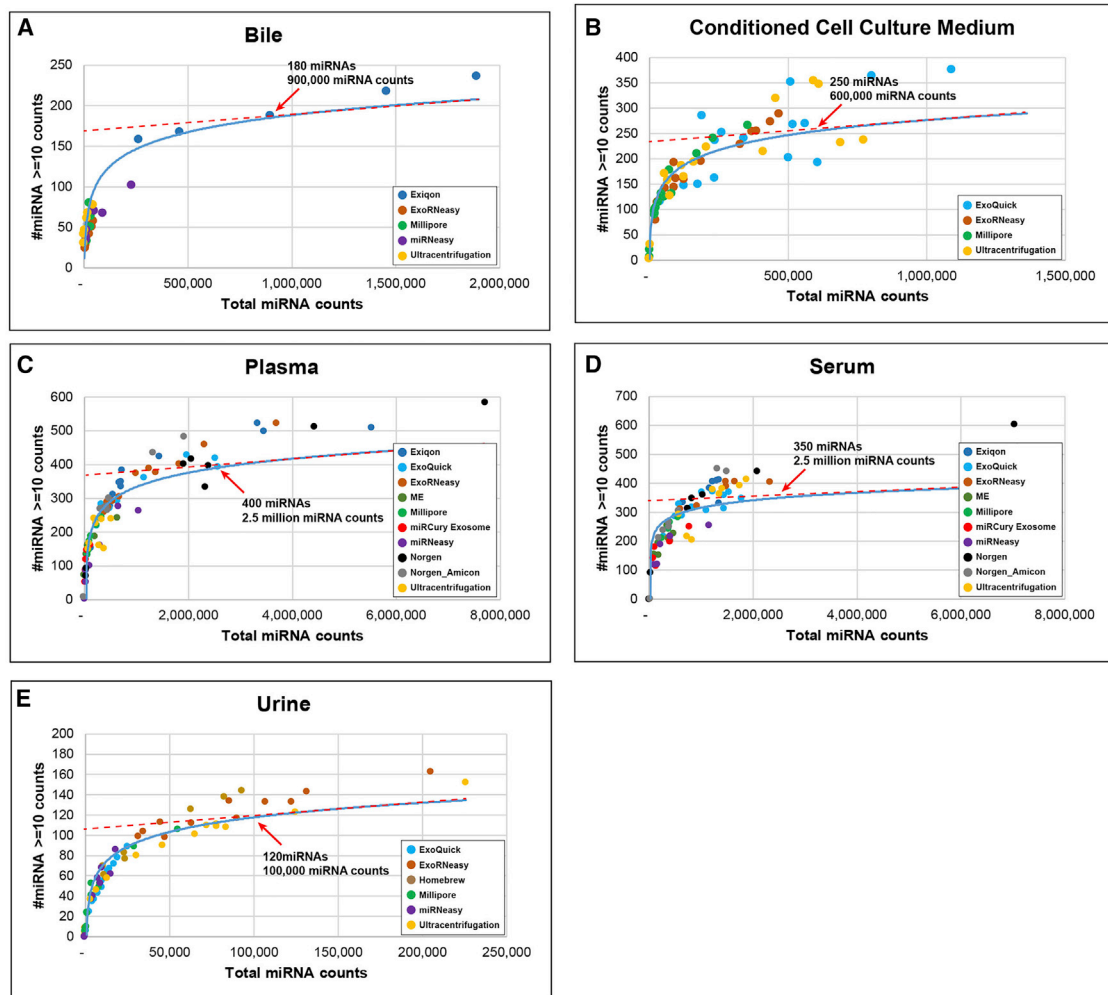


Figure 2. Scatterplots Illustrating Relationship between Small RNA Sequencing Library Complexity and Sequencing Depth

(A–E) Complexity for each library (y axis) was calculated as the number of miRNAs present at ten or more raw read counts in the “miRNA_AllBiofluid_RawData” dataset. Sequencing depth (x axis) was measured as total miRNA counts. Data for libraries from the different biofluids are shown: bile (A), CCCM (B), plasma (C), serum (D), and urine (E). Libraries prepared from samples isolated using each method are color coded according to the legends. For each graph, the best-fit curve is shown in blue, and the slope of the flatter portion of the log curve is indicated with a dashed red line. The estimated point of diminishing returns for each graph is indicated by the red arrow.

See also [Table S2](#).

Nor, and Ultra. The urine libraries overall had a high %tRNA, with the ExoRNeasy and Ultra methods yielding the highest (but still quite low) %miRNA.

miRNA Results from Small RNA-Seq miRNA Complexity Correlates with miRNA Read Depth

For every exRNA isolation method and every biofluid, the miRNA complexity (number of miRNAs with ≥ 10 raw counts) positively correlated with total miRNA read count (Figure 2). For each biofluid, the complexity began to plateau at a characteristic miRNA read depth, with the number of detected miRNAs at this plateau differing among biofluids: 180 detected miRNAs at 900,000 total miRNA counts for bile, 250 miRNAs at 600,000 total miRNA counts for CCCM, 400 miRNAs at 2.5 million total miRNA counts for plasma, 350 miRNAs at 2.5 million

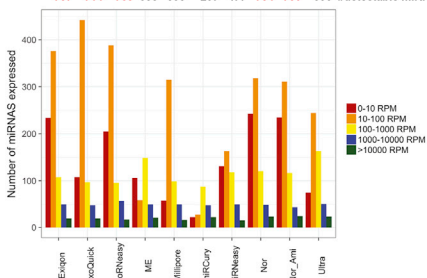
total miRNA counts for serum, and 120 miRNAs at 100,000 total miRNA counts for urine. For bile, only the Exiqon libraries reached this plateau.

Assessment of the Performance of exRNA Isolation Methods by Using miRNA Complexity, Expression, and Reproducibility Metrics

An ideal exRNA isolation method would display high yield and reproducibility for miRNAs from the desired exRNA carrier subclass(es) in the selected biofluid. To assess the relative performance of the tested exRNA isolation methods, we scaled and filtered the data to account for differences in sequencing depth among libraries and then calculated five metrics (Table S3): complexity (number of different miRNAs detected), mean number of detected miRNAs, mean percentage of replicates in which a given miRNA was detected, percent coefficient of

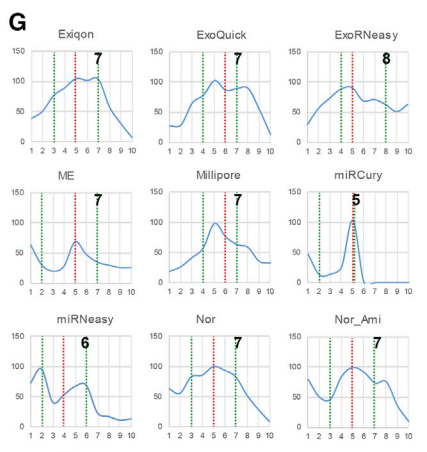
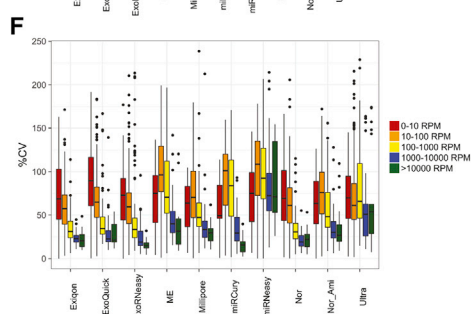
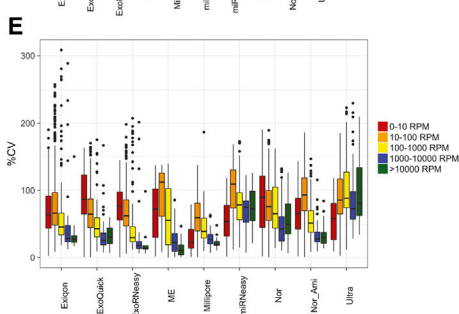
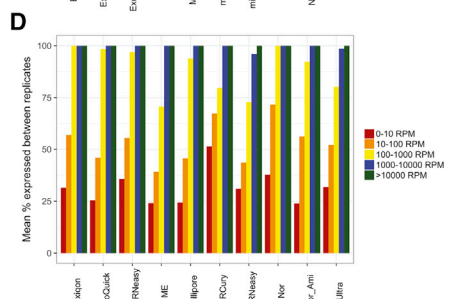
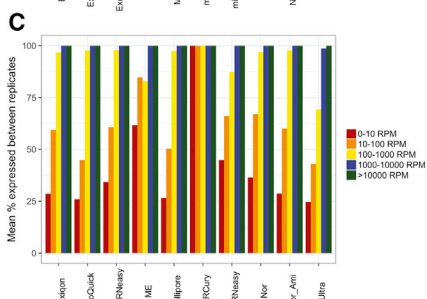
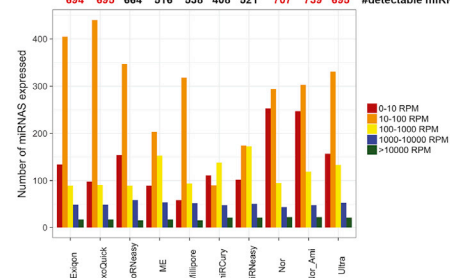
A Plasma

12	12	12	8	8	8	8	8	8	8	# total samples
11	10	11	2	5	1	3	6	7	6	# pass-filter samples
787	714	763	385	538	207	477	754	730	556	#detectable miRNAs

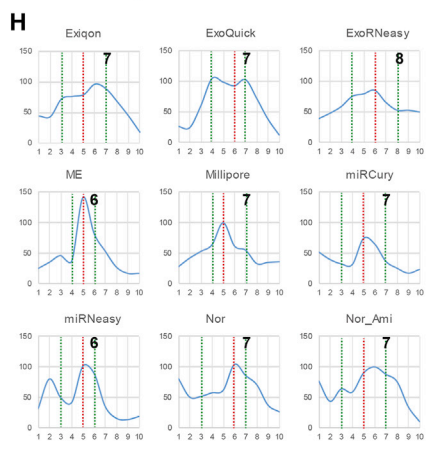


B Serum

12	12	12	8	8	8	8	8	8	8	# total samples
7	12	7	7	6	3	5	5	7	8	# pass-filter samples
694	695	664	516	538	408	521	707	739	695	#detectable miRNAs



Rank	Method	75% IQS	Complexity
1	ExoRNeasy	8	763
2	Exiqon	7	787
3	Nor	7	754
4	Nor_Ami	7	730
5	ExoQuick	7	714
6	Millipore	7	538
7	ME	7	385
8	Ultra	6	556
9	miRNeasy	6	477
10	miRCury	5	207



Rank	Method	75% IQS	Complexity
1	ExoRNeasy	8	664
2	Nor_Ami	7	739
3	Nor	7	707
4	ExoQuick	7	695
5	Exiqon	7	694
6	Millipore	7	538
7	miRCury	7	408
8	Ultra	6	695
9	miRNeasy	6	521
10	ME	6	516

(legend on next page)

variance (%CV) across replicates, and a combined expression/reproducibility score, which we term the integrated quality score (IQS) (Figures 3 and S3). Since the number of miRNAs detected and the probability of it being expressed consistently across technical replicates is dependent on the average expression level, results were displayed according to mean miRNA expression level (expressed in reads per million [RPM] miRNA counts) for each biofluid and isolation method (Figures 3A–3F and S3A–S3O). For each biofluid/isolation method combination, the IQS for each miRNA is the sum of the mean expression quantile and the %CV quantile across exRNA isolation kits. The scores for both the mean expression quantile and the %CV quantile range from 1–5: 1 for lowest mean expression or highest %CV and 5 for highest mean expression or lowest %CV across all isolation methods that had measurable expression for a given miRNA. The IQS values were computed for miRNAs that had measurable expression in at least four methods. (Table S3). Thus, a higher IQS for a given isolation method indicates better analytical performance due to higher expression and higher reproducibility. By comparing the distribution of the IQS values among exRNA isolation methods (Figures 3G, 3H, and S3P–S3T), we can easily select the optimal method for isolation for either a targeted list of miRNAs or for all miRNA detected for the biofluid of interest. The ideal isolation method would have a marked left skew, indicating a higher proportion of miRNAs with higher expression and reproducibility. Calculation of the 75th percentile IQS (IQS75) as an indicator of the left skewness in combination with the complexity of each isolation method enables ranking of the performance each of the tested exRNA methods.

Using these metrics, the findings for plasma and serum were similar, with ME, Millipore, miRCURY, miRNeasy, and Ultra displaying markedly lower complexity compared to the other methods (Figures 3A and 3B). For most methods, the number of expressed miRNAs decreased with increasing miRNA expression level, with the exception of the 0–10 RPM expression window, which generally contained fewer miRNAs compared to the 10–100 window (likely due to removal of miRNAs with <5 samples expressing at least 10 RPM during filtering; Table S2.6). The percent of miRNAs expressed among replicates was positively correlated with average expression level, except for the miRCURY kit in plasma, which showed 100% for all expression levels because there was only one pass-filter sample (Figures 3C and 3D). For Exiqon, ExoQuick, ExoRNeasy, and Nor,

the %CV decreased with increasing expression, as expected (Figures 3E and 3F). ME, Millipore, miRCURY, Nor_Ami, and Ultra displayed a lower %CV in the 0–10 RPM window compared to the 10–100 RPM window, likely due to a higher number of zero values for low-expressed miRNAs. We also observed that the miRNeasy and Ultra data for both plasma and serum, as well as the Nor data for plasma, did not display the expected decrease in %CV with increasing miRNA expression, indicating suboptimal reproducibility for these methods. For each method, we plotted the distribution of IQS values and compiled a table ranking the methods first by descending IQS75 and then by descending complexity (Figures 3G and 3H). In the table, the highest IQS75 value (IQS75max) and all IQS75max-1 values were highlighted in green, and the highest complexity value and complexity values within 10% of the highest values were highlighted in red. Methods that had both a highlighted IQS value and a highlighted complexity value were bolded. ExoRNeasy ranked highest for both plasma and serum, with the highest IQS75 (Figures 3G and 3H), largely due to high reproducibility (Figures 3I and 3J). The order of the next four methods (Exiqon, ExoQuick, Nor, and Nor_Ami) differed slightly between plasma and serum, but all four had the same IQS75 values and comparable complexities. Based on these results, it would seem that the optimal exRNA isolation method for plasma and serum studies would be ExoRNeasy; however, the potential biases that different methods have for specific exRNA carrier subclasses should be kept in mind, as discussed in detail below.

For bile, Exiqon clearly ranked highest, with a high IQS75 (Figure S3P) and a complexity that was far higher than the other three methods (Figure S3A). However, none of the bile methods showed a consistent decrease in variability with increasing expression level (Figure S3K), suggesting that even Exiqon was suboptimal. For CCCM, the relative performance of the different methods varied among cell lines (Figures S3B–S3D, S3G–S3I, S3L–S3N, and S3Q–S3S). For hESCs, all four methods had comparable IQS75 values and complexities, but only ExoQuick and ExoRNeasy performed well for KMBC, and ExoRNeasy was the only method with good performance for NRVM. Thus, only ExoRNeasy performed well across all three cell lines (Figures S3Q–S3S). For urine, none of the methods had both high complexity and reproducibility. For example, ExoQuick had the highest IQS75 but low complexity, while ExoRNeasy showed high complexity but a relatively low IQS75 (Figure S3T).

Figure 3. Complexity and Reproducibility of exRNA Isolation Methods for Plasma and Serum

(A, C, E, G) Plasma.

(B, D, F, H) Serum

For (A)–(F), values are given for each RNA isolation method for each expression window.

(A and B) Average number of miRNAs expressed.

(C and D) Mean percentage of replicates in which the miRNAs with the indicated mean expression level were detected.

(E and F) Boxplots indicating the %CV as a function of mean expression level. For plasma, no CV data are shown for the miRCURY exosome kit, as only one sample passed filter. The box indicates the median and interquartile range (IQR), the whiskers indicate the 25th percentile (Q1) – 1.5 × IQR and the 75th percentile (Q3) + 1.5 × IQR, and the individual dots indicate outliers.

(G and H) Plots indicating the distribution of IQSs across individual miRNAs comprised of the sum of the mean expression quantile (1 for lowest expression, 5 for highest expression) and %CV quantile (1 for highest %CV, 5 for lowest %CV) for each exRNA isolation method. The 25th, 50th, and 75th percentile IQS values are indicated by the vertical dashed lines.

Table: the 75th percentile IQS (IQS75) is highlighted in green if it is the top value (IQS75max) or IQS75max-1, the complexity is highlighted in red if it is within 10% of the highest value among all methods, and methods that meet both of these criteria are bolded; the methods are sorted first by IQS75 and then by complexity. See also Figures S3 and S4 and Tables S2 and S3.

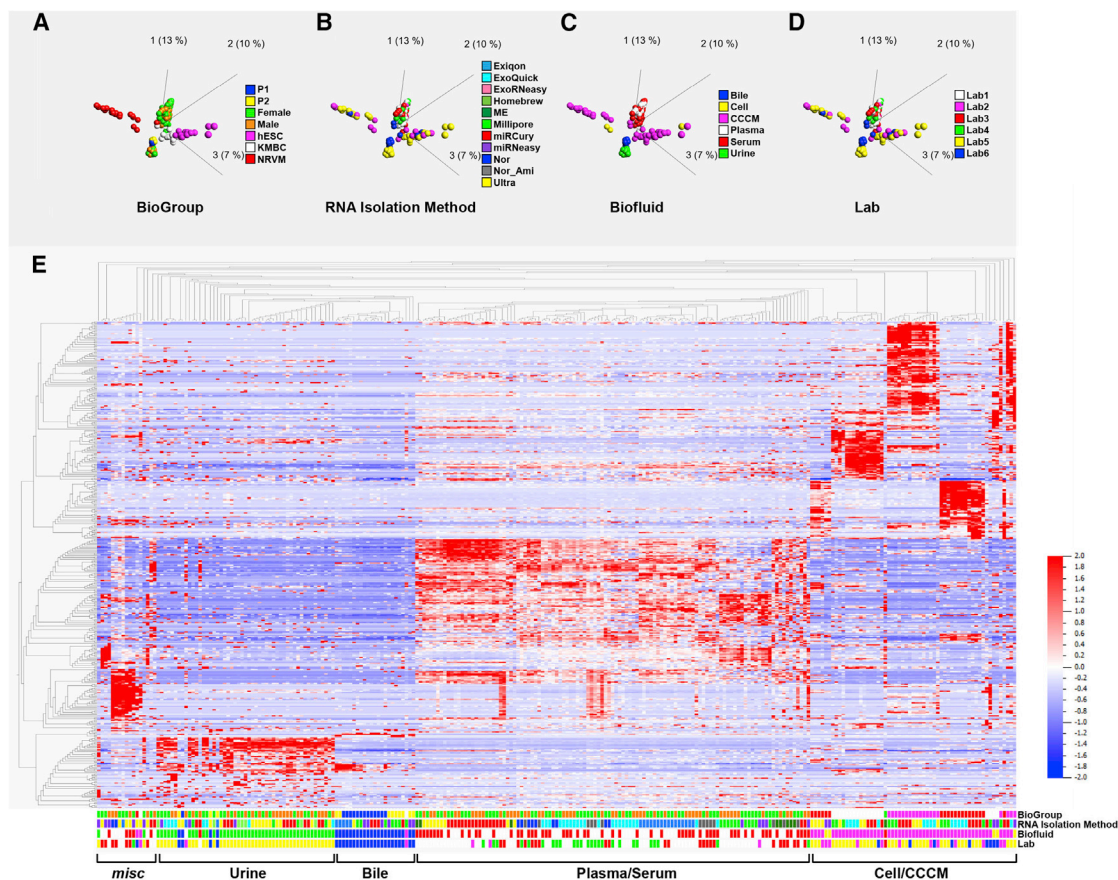


Figure 4. PCA and Hierarchical Clustering Analysis of miRNA Data from “miRNA_AllBiofluid_Scaled_Filtered” Dataset

(A–D) PCA plots with samples color coded by biological group (P1 and P2 for bile; hESC, KMBC, and NRVM cell lines for CCCM; and female [F] and male [M] for plasma and serum) (A), exRNA isolation method (B), biofluid type (C), and lab (D).

(E) Heatmap showing biclustering of miRNAs and samples by Euclidean distance with average linkage. The biological group, exRNA isolation method, biofluid, and lab for each sample are color coded below the heatmap using the same color scheme as in (A)–(D).

See also [Table S2](#).

To make these results easily accessible to investigators, we developed an interactive web-accessible application, miRDar (miRNA detection- and reproducibility-based selection of exRNA isolation methods). The investigator specifies the biofluid and the set of miRNAs of interest either by selecting a predefined set (consisting of all detected miRNAs or miRNAs observed to be enriched in specific carrier subclasses [see below]) or by building a set from individual miRNAs in a pull-down list. miRDar will return graphs showing for each exRNA isolation method: the mean miRNA expression level, the mean %CV, the mean expression quantile and mean %CV quantile, the distribution of IQS values for the constituent miRNAs, and a table ranking the methods by IQS75 and complexity (Figure S4). miRDar will also provide a table listing for each miRNA and each isolation method—a comprehensive set of data quality metrics.

From these results, we see that a combination of IQS75 and complexity enables quantitative and systematic comparison of the overall quality of the miRNA data obtained from exRNA samples isolated using each of the tested methods. However, these metrics do not reveal whether different exRNA isolation methods

produce similar miRNA profiles or if differences may arise from biases of the methods for particular exRNA carrier subclasses. Thus, we proceeded to address these questions directly.

miRNA Profiles Differ Substantially among exRNA Isolation Methods for Bile, Plasma, and Serum, but Not for CCCM or Urine

Unsupervised clustering of the complete scaled and filtered dataset (Table S2.7) showed that samples clustered primarily by biofluid, except for plasma and serum, which clustered together (Figure 4). Bile and plasma/serum subclustered according to exRNA isolation method, while CCCM subclustered by cell line. There was no apparent clustering by lab. To further explore sources of variability in the miRNA profiles for each biofluid, we analyzed the biofluids separately.

For bile (Table S2.8), samples clustered predominantly by donor (P1 versus P2) and then by isolation method (Figure S5A1 and 2). The only reproducible profile across replicates was seen for Exiqon (Figure S5A3), consistent with the low miRNA read depth and complexity for the other methods (Figure 2). Taken together with the results from the quality metrics, we conclude

that for bile, the only acceptable isolation method was Exiqon, but that the complexity even for Exiqon was low. Thus, further optimization of exRNA isolation from bile is warranted.

For cell and CCCM (Table S2.9), the samples clustered most strongly according to the source cell line (Figure S5B). The group of miRNAs expressed specifically in the cell and CCCM samples from the hESC line included the known pluripotency-associated miRNAs encoded by a large miRNA cluster on chromosome 19 (Laurent et al., 2008). Within each cell line, samples clustered weakly by exRNA isolation method, but not by lab. Given the similarities in the miRNA profiles produced by the different exRNA isolation methods, we conclude that the choice of method for CCCM studies should be driven by library data quality, which favors ExoRNeasy.

Plasma and serum samples did not separate by biofluid type (plasma versus serum) or BioGroup (female versus male) (Figure 5 and Table S2.10). Differential expression analysis of the plasma/serum data as a whole or stratified by exRNA isolation method confirmed that there were no miRNAs that were significantly ($q < 0.01$) differentially expressed based on biofluid (plasma versus serum) or BioGroup (female versus male). Biclustering showed that plasma/serum samples clustered by exRNA isolation method to form three main groups, and miRNAs formed four large clusters (Figure 5). The Cluster1 miRNAs were most highly expressed in Millipore and miRNeasy samples. The Cluster3 miRNAs were highest in Nor and Nor_Ami samples, and the Cluster4 miRNAs were highest in ExoRNeasy, Ultra, and ME samples. The Cluster2 miRNAs, as well as the other other isolation methods, were not as easily categorized. The ExoQuick samples were distributed between two groups: one with high expression of a subset of Cluster3 miRNAs and the other with high expression of Cluster4 miRNAs. These two groups of ExoQuick samples did not differ by donor sex, lab, or biofluid. The Exiqon samples contained high levels of a subset of Cluster1 and a subset of Cluster3 miRNAs. Cluster2 miRNAs were not clearly associated with any isolation method. Taken together, these results suggested that the Cluster1–4 extracellular miRNAs may be associated with distinct physical carriers, which are differentially accessed by the different exRNA isolation methods—a concept that was further explored through deconvolution analysis, as described below.

Compared to the other biofluids, urine (Table S2.11) displayed a biofluid-specific extracellular miRNA profile (Figure 4), but focusing on urine alone (“miRNA_urine_Scaled_Filtered”), we did not observe clustering by BioGroup (female versus male donor) or exRNA isolation method (Figure S5C). These results suggest that the choice of exRNA isolation method for urine should be driven by the overall small RNA-seq data-quality metrics, which indicated that ExoQuick was superior to the other tested methods.

Deconvolution Analysis Indicates that Different exRNA Carrier Subclasses Carry Distinct Sets of miRNAs from Plasma and Serum and Are Differentially Purified by Different exRNA Isolation Methods

To further explore the possibility that distinct sets of extracellular miRNAs are associated with different physical carriers, which are differentially purified by different exRNA isolation methods, we profiled putative carrier subclasses (EVs, AGO2-associated

RISC, and HDL) isolated by immunoprecipitation or sequential density fractionation and then used the resulting carrier subclass-specific signatures to deconvolute the miRNA profiles from each exRNA isolation method.

To isolate EVs and AGO2-associated RNA-induced silencing complex (RISC), we performed immunoprecipitations using antibodies raised against CD63, CD81, and CD9 (EV markers) and AGO2 on the female serum pool sample. A male plasma sample was separated into high-density lipoprotein (HDL), low-density lipoprotein (LDL), very low-density lipoprotein (VLDL), chylomicron, and lipoprotein-free (LFF) fractions by sequential density ultracentrifugation. RNA was isolated from the immunoprecipitated material and lipoprotein fractions and subjected to small RNA-seq. The LDL, VLDL, and chylomicron samples yielded too few counts (<300 raw miRNA counts) to be analyzed. Because the principal-component analysis (PCA) plot of the CD63+, CD81+, CD9+, HDL, and LFF fractions showed substantial overlap of the CD81 and CD9 profiles (Figure S6A) and because immunocapture/fluorescent staining of plasma EVs using antibodies against CD63, CD81, and CD9 suggested that there are two major populations of canonical EVs—CD63-/CD81+/CD9+ and CD63+/CD81+/CD9+ (Figure S6B)—the data from the CD81 and CD9 fractions were combined.

To identify miRNAs that were highly specific to the CD63, combined CD81/CD9, AGO2, HDL, and LFF carrier subclasses, the top 10 miRNAs based on the highest fold change with respect to the next highest subclass were selected as the signature miRNAs for each subclass to give a total of 50 miRNAs (Table S4.1–2). We then performed deconvolution analysis to estimate the relative proportions of each of the carrier subclasses in our plasma and exRNA profiles. To confirm the validity of the deconvolution method, we examined the samples from the immunoprecipitation and LPP fractionation experiments and established that each sample was comprised of the expected carrier subclass (Figure 5F1 and Table S4.3). Next, we deconvoluted the profiles for each of the exRNA isolation methods (Figures 5F2 and S6H–S6J and Table S4.4). We observed that the isolation methods could be grouped according to their deconvolution patterns and that these groupings mirrored those seen in the unsupervised clustering analysis (Figure 5E). ExoRNeasy, ME, and Ultra samples had similar patterns, mostly accounted for by the CD63 (>50%) and CD81/CD9 (~25%) signatures. The Exiqon, ExoQuick, miRCURY, Nor, and Nor_Ami samples displayed a mixed composition consisting of mostly of AGO2 (~50%) and CD63 (35%). The AGO2 signature was strongest in the Millipore samples. The miRNeasy samples showed the broadest representation of subclasses, with ~25% of their profiles attributable to each of the AGO2, CD63, HDL, and LFF subclasses.

These results suggest that the choice of exRNA isolation method should take into consideration the desired exRNA carrier subclass(es). To specifically isolate EV-associated miRNAs in plasma or serum, one should use ExoRNeasy, which has the best quality metrics among the EV-selective methods. However, to access both EV- and AGO2-associated miRNAs, ExoQuick for plasma or Norgen for serum are the best choices. We incorporated these findings into miRDaR, allowing investigators to identify the optimal exRNA isolation method for each exRNA carrier subclass for plasma and serum.

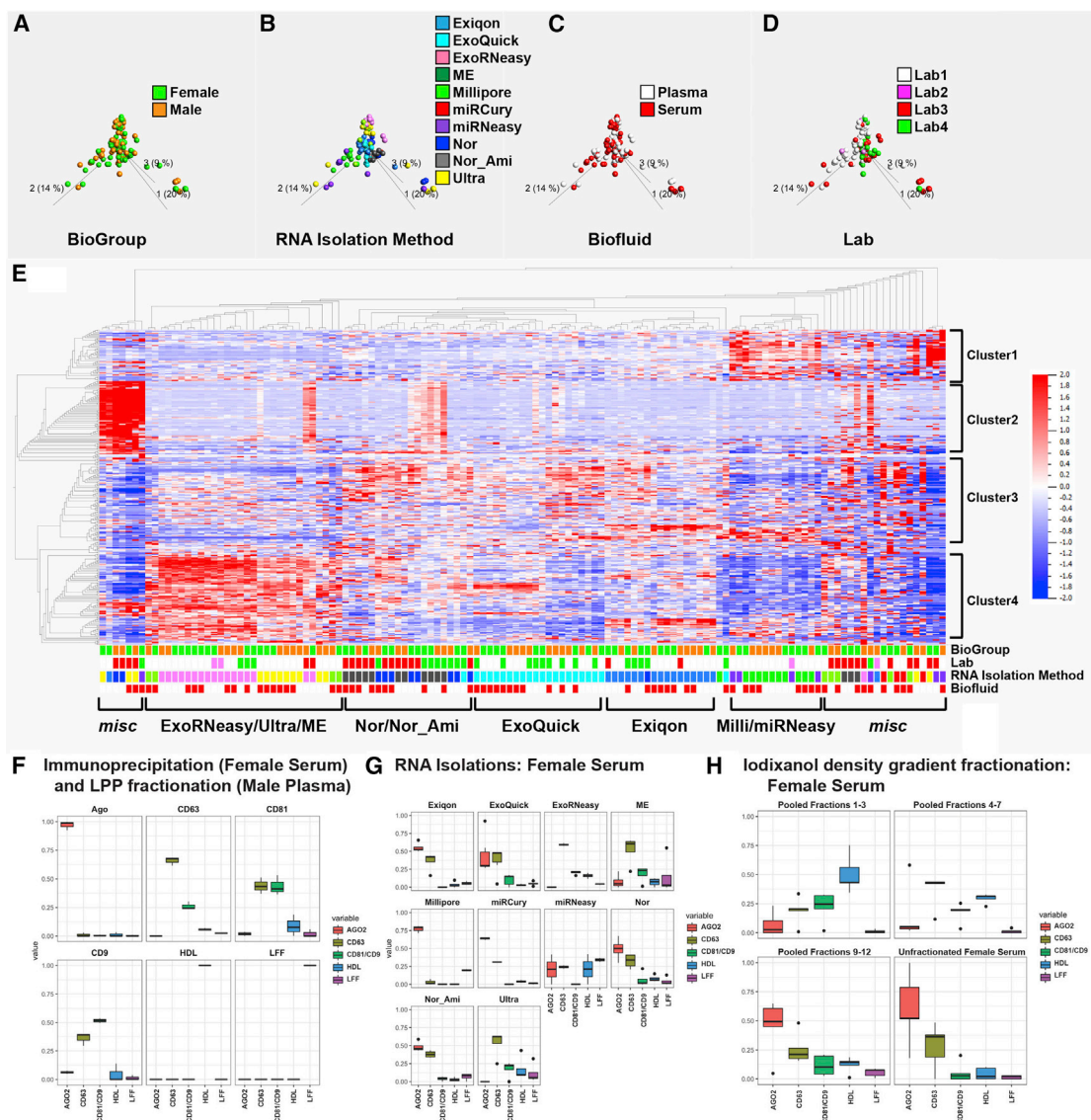


Figure 5. Analysis of Plasma and Serum miRNA Data

(A–D) PCA plots with samples color coded by biological group (female and male, A), exRNA isolation method (B), biofluid type (C), and lab number (D).

(E) Heatmap showing biclustering of miRNAs and samples. The biological group, biofluid, lab, and exRNA isolation method for each sample is color coded below the heatmap using the same color schemes as in (A)–(D). Groups 1–4, indicated to the right of the heatmap, are sets of miRNAs preferentially isolated by specific exRNA isolation methods.

(F) Deconvolution results for exRNA samples extracted from purified carrier subclasses. Box-and-whisker plots show proportions of CD63+, CD81+, CD9+, AGO2+, HDL, and LFF fractions for each sample.

(G) Deconvolution results for exRNA samples from the tested exRNA isolation methods. ExRNA samples isolated from the female serum pool using the indicated exRNA isolation methods were analyzed.

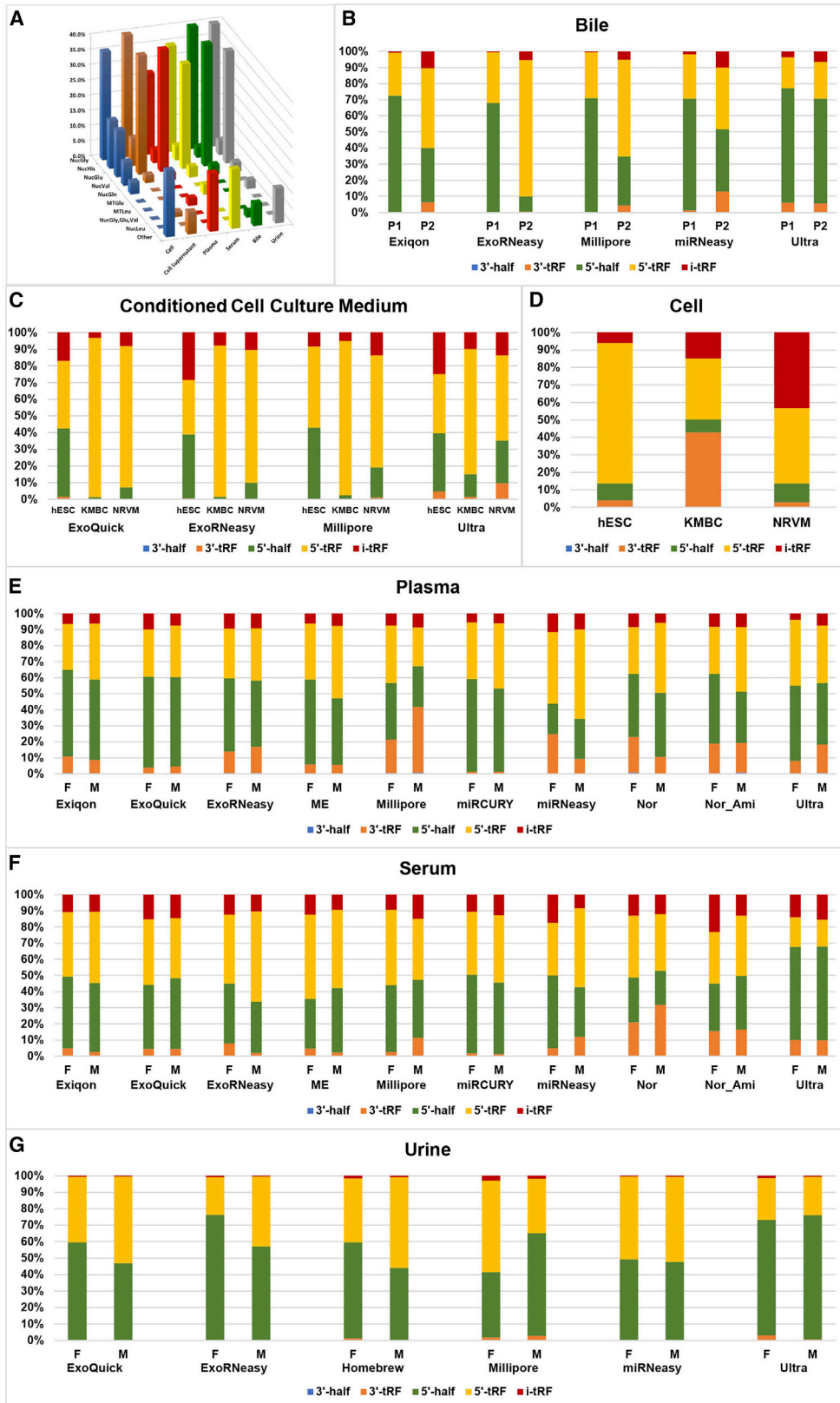
(H) Deconvolution results for exRNA samples isolated from iodixanol density gradients. ExRNA samples isolated from the female serum pool before and after fractionation on an iodixanol gradient were analyzed.

On the boxplots, the box indicates the median and IQR, the whiskers indicate $Q1 - 1.5 \times IQR$ and $Q3 + 1.5 \times IQR$, and the individual dots indicate outliers. See also Figures S5 and S6 and Tables S2 and S4.

Deconvolution Analysis of Plasma and Serum Fractionated by Density Gradient Ultracentrifugation

We explored the representation of CD63-, CD81/CD9-, AGO2-, HDL-, and LFF-associated miRNAs in plasma and serum fractions obtained by cushioned-density gradient ultracentrifugation

(C-DGUC) on iodixanol gradients (Li et al., 2018a). We fractionated plasma and serum samples from five female and five male donors by C-DGUC, collecting 12 fractions with densities from 1.028 to 1.259 g/mL (Figures S6C–S6E). Nanoparticle tracking analysis showed similar particle sizes for all fractions, with an



(legend on next page)

average mean size of 87.6 nm and average mode of 75.8 nm and a bimodal distribution of particles across the fractions, with one peak at fraction 5 and a second peak at fraction 12 (Figure S6F). Based on this distribution, we pooled fractions 1–3, 4–7, and 9–12 (fraction 8 was in the area of overlap between the two peaks). The density of the Fxn9_12 pool (1.11–1.26 g/mL) encompassed the ranges of the low-density (LD, 1.09–1.21 g/mL) and high-density (HD, 1.24–1.31 g/mL) fractions analyzed in two studies using CCCM from dendritic cells (Kowal et al., 2016; Lässer et al., 2017). In contrast to these previous reports, which showed a bimodal distribution of canonical EV markers within this higher density range, the peak levels of Flotillin 1 and CD9 in our experiments were seen in Fxn4_7 (1.05–1.08 g/mL; Figure S6E). We detected APOA1 and AGO2 in the 9–12 fraction pools, which is consistent with prior studies reporting the densities of HDL (1.06–1.21; Gotto and Jackson, 1977) and AGO-associated RNPs (Höck et al., 2007; in which AGO+ RNPs were found in a trimodal distribution spanning 1.05 to 1.3 g/mL). Deconvolution analysis of small RNA-seq data generated from these pools, and from unfractionated input plasma and serum, revealed marked differences in the proportion of the exRNA carrier subclasses among fraction pools and between the fractionated and unfractionated samples (Figures 5H and S6L–S6O and Table S4.6). The unfractionated samples showed roughly equal proportions of the AGO2 and CD63 subclasses and very low contributions from the CD81/CD9, HDL, or LFF subclasses. Fxn1_3 had a predominantly HDL signature, while the Fxn4_7 profile was largely accounted for by the CD63 and CD81/CD9 signatures (consistent with the highest levels of Flotillin 1 and CD9 protein by western blot; Figure S6G), with a smaller contribution from HDL. ~50% of Fxn9_12 was attributable to the AGO2 subclass (consistent with the highest level of AGO2 protein by western blot; Figure S6G), with the rest of profile accounted for by roughly equal levels of the CD63, CD81/CD9, and HDL subclasses. We note that the deconvolution reveals the proportion of each sample accounted for by the constituent subclasses rather than the distribution of each subclass across the samples. Therefore, even though a larger proportion of Fxn1_3 than of Fxn9_12 was attributable to HDL, this does not indicate that there was more HDL-associated miRNA in Fxn1_3 than Fxn9_12 in absolute terms. In fact, given the markedly higher numbers of raw miRNA read counts and levels of APOA1 obtained from Fxn9_12 compared to Fxn1_3 and Fxn4_7 (Table S4.6), we would conclude that HDL was predominantly found in Fxn9_12.

tRNA Results from Small RNA-Seq

The reads for each library were mapped to the “tRNA space” using the MINTmap pipeline (Loher et al., 2017, 2018) (tRNA_AllBiofluid_RawData; Table S5.1) and processed through a series of filtering and normalization steps to produce the tRNA_All

Biofluid_Processed (Table S5.2) dataset. The upper bound of the small RNA-seq library size selection procedure was 180 bp (corresponding to a ≤ 60 bp insert size). Therefore, we expected to obtain data on tRNA fragments (tRFs), but not full-length tRNAs. Biofluid type and source cell line had marked effects on the distribution of tRF types (Figure 6 and Table S5.3); the impact of exRNA isolation method was limited to an increased fraction of 3' tRF in the plasma Millipore samples and a lower fraction in the plasma miRCURY samples. The tRF distributions were similar between plasma and serum (with 5' halves and 5' tRFs being more abundant than 3' tRF and internal tRF [i-tRF]), except for a higher abundance of 3' tRFs in plasma. Bile and urine samples were almost completely comprised of 5' halves and 5' tRFs. The CCCM tRNA sequences, particularly those for KMBC, were predominantly made up of 5' tRFs, while the tRF distributions in cells were characterized by higher proportions of 3' tRFs (especially for KMBC) and i-tRFs (especially for NRVM). 3' halves were rare in all samples.

Comparing the tRNA amino acid distributions among sample types, we found that all biofluids had similar profiles, with 30%–40% being represented by both nGly and nGlu and 2%–5% by nHis and nVal (Figure 6 and Table S5.4). The prevalence of nGly and nGlu is likely the result of stability due to dimerization (Tosar et al., 2018). In contrast to the miRNA data, there was poor correlation between the cell and CCCM samples for tRNA data (Figures 6, S7A, and S7B and Table S5.4). For the cell samples, each cell line showed a distinct profile. However, there were two major clusters of CCCM samples: one cluster contained nearly all of the KMBC samples, along with the NRVM_Ultra and hESC_Ultra samples, and was dominated by the nGlu isodecoder; the other cluster contained all of the non-Ultra hESC and NRVM samples and was dominated by the nGly isodecoder. Also in contrast to the miRNA data, where clear clustering of plasma/serum samples by exRNA isolation method was seen, there was only weak clustering of ExoRNeasy samples for plasma and Nor and Ultra samples for serum in the tRNA data (Figures S7C and S7D). tRNA amino acid distributions for all of the bile samples were very similar (Figure S7E). For urine, there were two clusters, one with roughly equal fractions of nGly and nGlu and containing the ExoRNeasy and Ultra samples and the other with mostly nGly and containing the ExoQuick, Homebrew, Millipore, and miRNeasy samples (Figure S7F). These results suggest that bile contains one, CCCM and urine contain two, and plasma and serum contain several tRNA carrier subclasses.

The size distributions of the tRFs were similar across biofluids, cell lines, and exRNA isolation methods, ranging between 30–32 nt for cells, 30 nt for hESC CCCM, 21–31 nt for KMBC and NRVM CCCM, 29–34 nt for P1 and 32–34 for P2 in bile, 31–36 nt for plasma, 31–34 nt for serum, and 30–31 nt for urine (Figures

Figure 6. tRNA Analysis

(A) Distribution of tRNA amino acid distributions. Bar plots represent the percentage of tRF reads mapping to each amino acid. The top five amino acids are plotted for each biofluid type. The remaining amino acids were combined into the “other” category. The most abundant amino acids differ between the biofluids. (B–G) Distribution of tRNA fragment types (bile, B; CCCM, C; cell, D; plasma, E; serum, F; urine, G). The distributions of 3' half, 3' tRF, 5' half, 5' tRF, and i-tRF are shown for each library. The y axis is the percentage of tRF type relative to total reads present within the tRNA space only and not relative to all sequence reads generated. Values are averaged across all pass-filter replicate libraries.

See also Figure S7 and Table S5.

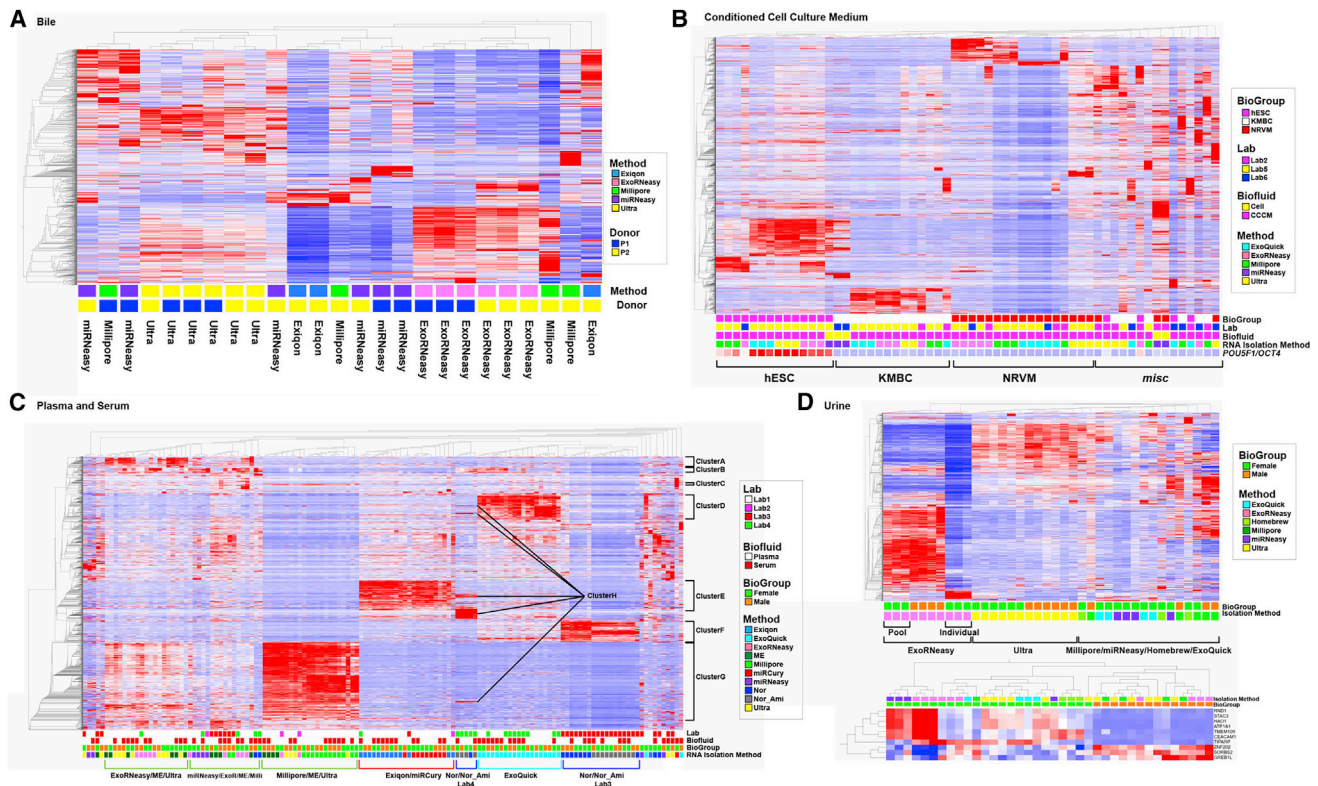


Figure 7. Hierarchical Clustering Analysis of mRNA Data

Each heatmap shows biclustering of mRNAs and samples.

(A) Hierarchical clustering analysis of mRNA data from “mRNA_bile_filtsc” dataset. The donor and exRNA isolation method for each sample are color coded below the heatmap.

(B) Hierarchical clustering analysis of miRNA data from “mRNA_CellCCCM_filtsc” dataset. The sample type (cell or CCCM), cell line, and exRNA isolation method for each sample are color coded below the heatmap.

(C) Hierarchical clustering analysis of mRNA data from “mRNA_plasmaserum_filtsc” dataset. The biological group, biofluid, lab, and exRNA isolation method for each sample are color coded below the heatmap. Groups 1–4, indicated to the right of the heatmap, are sets of mRNAs that are preferentially isolated by specific exRNA isolation methods.

(D) Hierarchical clustering analysis of mRNA data from “mRNA_urine_filtsc” dataset. Upper panel: heatmap showing biclustering of mRNAs and samples. The biological group and exRNA isolation method for each sample are color coded below the heatmap. Lower panel: heatmap showing hierarchical clustering of mRNAs that are differentially expressed between the female and male urine samples ($q \leq 0.01$).

See also [Tables S6](#) and [S7](#).

[S7G–S7L](#)). Overall, the exRNA isolation method had a weaker effect on the tRNA profile than the miRNA and mRNA profiles.

mRNA Fragment Results from Small RNA-Seq

The small RNA-seq data were also mapped to mRNA sequences and sequentially processed ([Tables S6](#) and [S7.1–6](#)). The results from the mRNA fragment analysis differed substantially from the miRNA analysis. As noted above, the bile samples clustered predominantly by donor (P1 versus P2) in the miRNA dataset, while the exRNA isolation method drove the clustering in the mRNA dataset ([Figure 7A](#)). Moreover, the Exiqon method yielded the most reproducible bile miRNA profile, while Ultra and ExoRNeasy produced the most consistent mRNA profiles.

As with the miRNA dataset, the CCCM samples clustered most strongly by source cell line in the mRNA dataset ([Figure 7B](#)). However, in contrast to the miRNA dataset, the exRNA isolation method had a strong effect on the mRNA fragment profile within

each cell line. For hESC and KMBC, ExoQuick, ExoRNeasy, and Ultra had similar profiles, and the Millipore samples formed a separate cluster. For NRVM, on the other hand, ExoRNeasy and Millipore clustered together, and the ExoQuick and Ultra samples formed two additional clusters. We note that a known pluripotency-associated mRNA, *POU5F1/OCT4*, was consistently and specifically expressed in the cell and CCCM samples from the hESC line. Approximately 25% of the samples did not cluster according to cell lines or exRNA isolation method (the miscellaneous samples in [Figure 7B](#)). Most of these miscellaneous samples were isolated in Lab2 and Lab6, suggesting that lab-to-lab variability may have a significant impact on the reproducibility of isolation of extracellular mRNA fragments from CCCM. Given the strong effects of the source cell line, exRNA isolation method, and lab on the CCCM profiles, we suggest that these variables should be held constant for a given experiment when possible.

As in the miRNA dataset, plasma and serum samples clustered most strongly by exRNA isolation method in the mRNA fragment data (Figure 7C), but with a different clustering pattern. Instead of the distinct clustering of ExoRNeasy, Ultra, and ME apart from miRNeasy and Millipore seen in the miRNA dataset, mRNA fragment analysis showed these five methods forming one large cluster with three subclusters that were not explained by lab, biofluid, biogroup, or exRNA isolation method (Figure 7C). Also, the mRNA data did not cluster the Nor, Nor_Ami, Exiqon, and ExoQuick samples together but rather clustered the Exiqon and miRCURY samples together, separated the ExoQuick samples into their own cluster, and clustered the Nor and Nor_Ami samples but subclustered them according to lab. From the perspective of the mRNA sequences, there were eight distinct mRNA clusters (ClusterA–ClusterH; Figure 7C). We performed functional enrichment analyses on each of these clusters and found that ClusterB was enriched in blood- and platelet-associated mRNAs, as well as mitochondrial mRNAs, while there was extensive overlap in the tissue-, subcellular-localization-, and biological-function-related terms for the other clusters (Table S7.7). We note that the enrichment of blood- and platelet-associated mRNAs in ClusterB could not have been due to hemolysis, as replicate exRNA isolations from the same standardized samples showed different levels of these transcripts.

Unlike in the urine miRNA results, which did not cluster samples by BioGroup or exRNA isolation method, the urine mRNA results showed distinct clustering according to both of these variables (Figure 7D). The strongest effect came from the exRNA isolation method, with the ExoRNeasy samples forming two tightly grouped clusters, and the Ultra samples clustering slightly more loosely. The ExoRNeasy clusters demonstrate the potential effects of interindividual variability on exRNA results, as we see that the female pooled samples cluster with the male pooled and male individual samples, while the female individual samples formed a separate cluster with a distinct mRNA profile. Within the ExoRNeasy and Ultra clusters, we noted separation of the female and male urine samples (Figure 7D). Differential expression analysis revealed 10 mRNAs with significantly different expression in female and male samples (Figure 7D, lower panel), none of which are encoded on the X or Y chromosomes or code for commonly known sex-specific proteins. Two of the differentially expressed mRNAs have kidney-related functions (mutations in GREB1L are associated with congenital kidney malformations, including renal agenesis in mice and humans; De Tomasi et al., 2017; Sanna-Cherchi et al., 2017; a mutation in HAO1 is associated with hyperoxaluria; Frishberg et al., 2014). We then compared our results with a recent publication reporting on differentially expressed transcripts between female and male kidneys in human and mouse (Si et al., 2009). In this previous report, Ceacam1 was found to be expressed 4.6-fold higher in healthy mouse female (compared to male) kidneys, which is concordant with our results. However, Tiparp was expressed 1.5-fold higher in male healthy mouse kidneys, and SORBS2 was expressed 1.5-fold higher in female healthy human glomeruli, which are in the opposite direction as in our data.

An overall consideration of the miRNA, tRNA, and mRNA fragment results suggests that the mechanisms underlying the loading of different RNA biotypes into exRNA carrier subclasses

may be quite different, and thus, the optimal method for exRNA isolation will differ depending on the targeted RNA biotype.

DISCUSSION

This study has produced key findings that will guide future exRNA biomarker and biology studies. By rigorously comparing different exRNA biotypes in a variety of biofluids from different individuals, as well as CCCM from three different cell types, we were able to identify major sources of technical and biological sources of variability and construct a set of three “best practices” for exRNA profiling studies.

First, for miRNA studies, small RNA-seq libraries should be sequenced deep enough that a plateau in complexity is reached. Our results suggest that each biofluid has a characteristic maximal complexity, which is reached at approximately the same target miRNA read depth regardless of the exRNA isolation method used: 180 detected miRNAs with 0.9 million miRNA reads for bile, 250 miRNAs with 0.6 million miRNA read for CCCM, 400 miRNAs with 2.5 million miRNA reads for plasma, 300 miRNAs with 2.5 million miRNA reads for serum, and 120 miRNAs with 100,000 miRNA reads for urine. If it is not possible to sequence all of the libraries in a dataset to the ideal depth, it is important to computationally correct for differences in read depth among the constituent libraries either by downsampling the number of miRNA reads per library to the same value or by filtering out the low-expressed miRNAs that would not be detectable in the lower miRNA read-depth libraries. The latter approach may lead to more robust validation, as shown by our previous studies that suggest improved cross-platform validation for miRNAs above a certain expression level (Yeri et al., 2018).

Second, given the marked differences among exRNA isolation methods in yield, reproducibility, and relative efficiency of accessing the exRNAs associated with different exRNA carrier subclasses, the optimal exRNA isolation method for a given study will depend on the efficiency and reproducibility with which the miRNA(s) of interest are isolated. To help investigators identify the optimal methods for their studies, we have developed an interactive web-based application, miRDaR, which will return quality metrics for each exRNA isolation method for any set of miRNAs entered by the user for a given biofluid.

Third, given the strong effects of biofluid, cell line, exRNA isolation method, and RNA biotype on exRNA profiles, these variables should be held constant within any given study. Moreover, integrative or comparative analysis among studies for which these variables differ should be performed with caution. The one pair of biofluids that would likely yield similar results is plasma and serum, given the similarity in exRNA profiles between plasma and serum throughout our study. If it is necessary to compare results between two exRNA isolation methods, it is preferable to use methods that access the different carrier subclasses with similar efficiencies and produce a similar profile for the target RNA biotype. For example, for miRNAs in plasma and serum, ExoRNeasy, ME, and Ultra represent a group of exRNA isolation methods that preferentially isolate EV-associated miRNAs, while Exiqon, ExoQuick, miRCURY, Nor, and Nor_Ami represent a second group of exRNA isolation methods that isolate both EV- and AGO2+ RNP-associated miRNAs

(Figures 5E, 5G, and S7H–S7K). The methods within each of these two groups produce similar miRNA profiles and thus are reasonably good choices if a switch of exRNA isolation method is necessary. The Millipore, miRNeasy, and DGUC fractions each have distinct patterns of access to the carrier subclasses and overall miRNA profiles, and thus, we cannot recommend switching from any of them to any other exRNA isolation method. A promising deconvolution-based strategy to correct for carrier subclass heterogeneity and thus enable cross-study analyses is presented in another of the ERCC1 flagship articles in this issue (Murillo et al., 2019).

We recognize that there are emerging carrier subclasses—e.g., exomeres (Zhang et al., 2018)—that were not included in our deconvolution analysis. We expect that exRNA profiling data will be generated on these and other as-yet undiscovered carrier subclasses over time and plan to incorporate them into our deconvolution approach to improve our understanding of the composition of different biofluids and the exRNA samples produced by different exRNA isolation methods. We also expect that additional methods for exRNA isolation will be developed over time and that the data from this study can be used to evaluate the exRNA produced by these new methods and assess their relative efficiency and reproducibility.

This study was limited to scalable exRNA isolation methods for which exRNA extraction from the tested biofluid was supported by the manufacturer. We also note that differences in rotor type and duration of centrifugation among participating laboratories for the Ultra experiments could have contributed to the high variability seen among the Ultra samples, as suggested by previous publications (Cvjetkovic et al., 2014; Jeppesen et al., 2014). Therefore, this study should not be seen as a definitive assessment of the performance of the Ultra method. Also, exRNA samples were not analyzed using alternative small RNA-seq methods (such as those using randomized adaptor sequences or template switching) or long RNA-seq. Future studies will be necessary to determine the impact of these variables on exRNA profiling results.

Over the past 10 years, the perceived value of using exRNAs in biofluids for diagnosis, prognosis, and monitoring therapeutic intervention in a variety of diseases has expanded exponentially. Given differences in technical and biological variables among studies, it has been hard to reach consensus about the clinical utility of specific biomarkers. This is especially difficult for biomarkers that consist of differences in levels of specific exRNAs rather than detection of disease-associated mutations (Skog et al., 2008) and splice variants (Antoury et al., 2018), which may be more distinctive and tolerant of technical variation. The study described in this report was designed to quantify the effects of major sources of variability and provide a set of best practices to improve the rigor and reproducibility of exRNA biomarker studies going forward. To promote the application of our findings to future studies, we have developed an interactive web-accessible application, which we call miRDaR. miRDaR extracts and displays relevant data from this study to assist investigators with selection of the optimal exRNA isolation methods for their studies. To enable investigators to compare results from their laboratories with our data, aliquots of the pooled plasma, serum, and urine samples used in this study are available upon request

through the ERCC Virtual BioRepository (<https://genboree.org/vbr-hub/>).

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Biofluid samples from human subjects
 - Cell Cultures
- METHOD DETAILS
 - Biofluid collection
 - exRNA Isolation
 - qRT-PCR
 - Immunoprecipitation of exRNA carriers
 - Fractionation of plasma serum using Cushioned Density Gradient Ultracentrifugation
 - Lipoprotein Fractionation from plasma
 - Western Blot
 - Single Particle Interferometric Reflectance Imaging Sensing (NanoView) analysis
 - Small RNA sequencing library preparation
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Analysis of qRT-PCR data
 - Analysis of small RNA Sequencing data
 - miRNA Data Processing
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cell.2019.03.024>.

ACKNOWLEDGMENTS

This publication is part of the NIH Extracellular RNA Communication Consortium paper package and was supported by the NIH Common Fund's exRNA Communication Program. We would like to thank Aileen Fernando, Fabian Flores, and the UCSD Clinical and Translational Research Institute for assisting us with collection of the standardized plasma and serum samples used in this study. We would like to thank Thomas Touboul for assisting us with collection of hESC supernatant for this study. This study was supported by the Extracellular RNA Communication Consortium funded by the NIH Common Fund (NIH UH3TR000906 [S.S., P.D.H., C.D.L., L.D.L., C.M., C.T., L.C.L.], NIH U19179563 [P.S.C., X.Z., L.B., X.O.B.], NIH UH3TR000901 [K.D., R.V.S., B.S., S.D., J.L.], NIH UH3TR000943 [A.K.S.], NIH UH3TR000890 [P.N., A.P., R.G.]), NIH UH3TR000884 [I.K.Y., T.P.], NIH U19CA179512 [R.L.R.], NIH U54DA036134 [R.P.A.]) and other funding sources (NIH P01069246 [P.S.C., X.Z., L.B., X.O.B.], NIH R01HL122547 [K.D., R.V.S., B.S., S.D., J.L.], NIH R35CA209904 and CA217685, the American Cancer Society Research Professor Award, and the Frank McGraw Memorial Chair in Cancer Research [A.K.S.], NIH K23 HL127099 [R.V.S.], NIH R01HL133575 [R.L.R.], NIH T32 HD007203 [S.S.], NIH/NHLBI HL141424 [I.R.], NIH/NCI CA195204 [E.L., I.R., R.M.], Thomas Jefferson University Institutional Funds [E.L., I.R., R.M.]).

AUTHOR CONTRIBUTIONS

Conceptualization, L.C.L., S.D., A.K.S., T.P., R.G., X.O.B., L.B., R.L.R.; Formal Analysis, L.C.L., A.Y., V.L.M., R.V.S., J.L., I.R., E.L., C.T.; Investigation, S.S.,

P.S.C., A.C., K.D., P.D.H., J.F., C.D.L., R.M., C.M., V.L.M., P.N., A.P., I.R., R.R., R.V.S., B.S., C.T., D.W., I.K.Y., X.Z., L.B., G.D., E.L., J.L., R.P.A. Data Curation, S.S., L.C.L.; Writing—Original Draft, S.S., V.L.M., R.V.S., E.L., R.L.R., L.C.L.; Writing—Review & Editing, S.S., L.B., X.O.B., R.G., E.L., T.P., R.L.R., A.K.S., S.D., L.C.L.; Visualization, L.D.L., R.P.A., L.C.L.; Supervision, L.B., X.O.B., R.G., E.L., T.P., R.L.R., A.K.S., S.D., L.C.L.; Project Administration, S.S., L.C.L.; Funding Acquisition, X.O.B., R.G., T.P., R.L.R., A.K.S., S.D., L.C.L.

DECLARATION OF INTERESTS

S.D. is a founding member of Dyrnamix, which did not play any role in this study and has a patent on extracellular RNA biomarkers for cardiac remodeling. Over the past 12 months, R.V.S. has received funds from Amgen (scientific advisory board), Myokardia (consulting), and Best Doctors (consulting). R.V.S. is a co-inventor on a patent for ex-RNAs signatures of cardiac remodeling. R.G. has been employed by Sanofi Genzyme since September 2017, working on a multiple sclerosis treatment. A.K.S. has the following interests: Advisory board member for Kiyatec and Merck; research funding from M-Trap; stock holder for Biopath; patents for EGFL6 antibodies and siRNA delivery systems.

Received: August 3, 2018

Revised: December 29, 2018

Accepted: March 11, 2019

Published: April 4, 2019

REFERENCES

- Antoury, L., Hu, N., Balaj, L., Das, S., Georghiou, S., Darras, B., Clark, T., Breakefield, X.O., and Wheeler, T.M. (2018). Analysis of extracellular mRNA in human urine reveals splice variant biomarkers of muscular dystrophies. *Nat. Commun.* **9**, 3906.
- Beltrami, C., Besnier, M., Shantikumar, S., Shearn, A.I., Rajakaruna, C., Laftah, A., Sessa, F., Spinetti, G., Petretto, E., Angelini, G.D., and Emanuelli, C. (2017). Human Pericardial Fluid Contains Exosomes Enriched with Cardiovascular-Expressed MicroRNAs and Promotes Therapeutic Angiogenesis. *Mol. Ther.* **25**, 679–693.
- Burgos, K.L., Javaherian, A., Bompreszi, R., Ghaffari, L., Rhodes, S., Courtright, A., Tembe, W., Kim, S., Metpally, R., and Van Keuren-Jensen, K. (2013). Identification of extracellular miRNA in human cerebrospinal fluid by next-generation sequencing. *RNA* **19**, 712–722.
- Cheng, L., Sharples, R.A., Scicluna, B.J., and Hill, A.F. (2014). Exosomes provide a protective and enriched source of miRNA for biomarker profiling compared to intracellular and cell-free blood. *J. Extracell. Vesicles* **3**.
- Chevillet, J.R., Kang, Q., Ruf, I.K., Briggs, H.A., Vojtech, L.N., Hughes, S.M., Cheng, H.H., Arroyo, J.D., Meredith, E.K., Gallichotte, E.N., et al. (2014). Quantitative and stoichiometric analysis of the microRNA content of exosomes. *Proc. Natl. Acad. Sci. USA* **111**, 14888–14893.
- Cvjetkovic, A., Lötvall, J., and Lässer, C. (2014). The influence of rotor type and centrifugation time on the yield and purity of extracellular vesicles. *J. Extracell. Vesicles* **3**.
- De Tomasi, L., David, P., Humbert, C., Silbermann, F., Arrondel, C., Tores, F., Fouquet, S., Desgrange, A., Niel, O., Bole-Feysot, C., et al. (2017). Mutations in GREB1L Cause Bilateral Kidney Agenesis in Humans and Mice. *Am. J. Hum. Genet.* **101**, 803–814.
- Freedman, J.E., Gerstein, M., Mick, E., Rozowsky, J., Levy, D., Kitchen, R., Das, S., Shah, R., Danielson, K., Beaulieu, L., et al. (2016). Diverse human extracellular RNAs are widely detected in human plasma. *Nat. Commun.* **7**, 11106.
- Frishberg, Y., Zeharia, A., Lyakhovetsky, R., Bargal, R., and Belostotsky, R. (2014). Mutations in HAO1 encoding glycolate oxidase cause isolated glycolic aciduria. *J. Med. Genet.* **51**, 526–529.
- Giraldez, M.D., Spengler, R.M., Etheridge, A., Godoy, P.M., Barczak, A.J., Srinivasan, S., De Hoff, P.L., Tanriverdi, K., Courtright, A., Lu, S., et al. (2018). Comprehensive multi-center assessment of small RNA-seq methods for quantitative miRNA profiling. *Nat. Biotechnol.* **36**, 746–757.
- Gotto, A.M., and Jackson, R.L. (1977). Structure of the Plasma Lipoproteins — A Review. In *Atherosclerosis IV*, G. Schettler, Y. Goto, Y. Hata, and G. Klose, eds. (Springer).
- Heberle, H., Meirelles, G.V., da Silva, F.R., Telles, G.P., and Minghim, R. (2015). InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* **16**, 169.
- Höck, J., Weinmann, L., Ender, C., Rüdell, S., Kremmer, E., Raabe, M., Urlaub, H., and Meister, G. (2007). Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO Rep.* **8**, 1052–1060.
- Jeppesen, D.K., Hvam, M.L., Primdahl-Bengtson, B., Boysen, A.T., Whitehead, B., Dyrskjot, L., Orntoft, T.F., Howard, K.A., and Ostfeld, M.S. (2014). Comparative analysis of discrete exosome fractions obtained by differential centrifugation. *J. Extracell. Vesicles* **3**, 25011.
- Karimi, N., Cvjetkovic, A., Jang, S.C., Crescitelli, R., Hosseinpour Feizi, M.A., Nieuwland, R., Lötvall, J., and Lässer, C. (2018). Detailed analysis of the plasma extracellular vesicle proteome after separation from lipoproteins. *Cell. Mol. Life Sci.* **75**, 2873–2886.
- Kowal, J., Arras, G., Colombo, M., Jouve, M., Morath, J.P., Primdahl-Bengtson, B., Dingli, F., Loew, D., Tkach, M., and Théry, C. (2016). Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. *Proc. Natl. Acad. Sci. USA* **113**, E968–E977.
- Lässer, C., Shelke, G.V., Yeri, A., Kim, D.K., Crescitelli, R., Raimondo, S., Sjöstrand, M., Gho, Y.S., Van Keuren Jensen, K., and Lötvall, J. (2017). Two distinct extracellular RNA signatures released by a single cell type identified by microarray and next-generation sequencing. *RNA Biol.* **14**, 58–72.
- Laurent, L.C., Chen, J., Ulitsky, I., Mueller, F.J., Lu, C., Shamir, R., Fan, J.B., and Löring, J.F. (2008). Comprehensive microRNA profiling reveals a unique human embryonic stem cell signature dominated by a single seed sequence. *Stem Cells* **26**, 1506–1516.
- Li, X., Mauro, M., and Williams, Z. (2015). Comparison of plasma extracellular RNA isolation kits reveals kit-dependent biases. *Biotechniques* **59**, 13–17.
- Li, F., Kaczor-Urbanowicz, K.E., Sun, J., Majem, B., Lo, H.C., Kim, Y., Koyano, K., Rao, S.L., Kang, S.Y., Kim, S.M., et al. (2018). Characterization of Human Salivary Extracellular RNA by Next-generation Sequencing. *Clin. Chem.* **64**, 1085–1095.
- Li, K., Wong, D.K., Hong, K.Y., and Raffai, R.L. (2018a). Cushioned-Density Gradient Ultracentrifugation (C-DGUC): A Refined and High Performance Method for the Isolation, Characterization, and Use of Exosomes. *Methods Mol. Biol.* **1740**, 69–83.
- Li, K., Wong, D.K., Luk, F.S., Kim, R.Y., and Raffai, R.L. (2018b). Isolation of Plasma Lipoproteins as a Source of Extracellular RNA. *Methods Mol. Biol.* **1740**, 139–153.
- Loher, P., Telonis, A.G., and Rigoutsos, I. (2017). MINTmap: fast and exhaustive profiling of nuclear and mitochondrial tRNA fragments from short RNA-seq data. *Sci. Rep.* **7**, 41184.
- Loher, P., Telonis, A.G., and Rigoutsos, I. (2018). Accurate Profiling and Quantification of tRNA Fragments from RNA-Seq Data: A Vade Mecum for MINTmap. *Methods Mol. Biol.* **1680**, 237–255.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**.
- Melman, Y.F., Shah, R., Danielson, K., Xiao, J., Simonson, B., Barth, A., Chakir, K., Lewis, G.D., Lavender, Z., Truong, Q.A., et al. (2015). Circulating MicroRNA-30d Is Associated With Response to Cardiac Resynchronization Therapy in Heart Failure and Regulates Cardiomyocyte Apoptosis: A Translational Pilot Study. *Circulation* **131**, 2202–2216.
- Murillo, O., Thistlethwaite, W., Rozowsky, J., Subramanian, S.L., Lucero, R., Shah, N., Jackson, A.R., Srinivasan, S., Chung, A., Laurent, C.D., et al. (2019). ExRNA Atlas Analysis Reveals Distinct Extracellular RNA Cargo Types and Their Carriers Present across Human Biofluids. *Cell* **177**, this issue, 463–477.

- Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn, M., and Alizadeh, A.A. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* *12*, 453–457.
- Ohshima, K., Inoue, K., Fujiwara, A., Hatakeyama, K., Kanto, K., Watanabe, Y., Muramatsu, K., Fukuda, Y., Ogura, S., Yamaguchi, K., and Mochizuki, T. (2010). Let-7 microRNA family is selectively secreted into the extracellular environment via exosomes in a metastatic gastric cancer cell line. *PLoS ONE* *5*, e13247.
- Pliatsika, V., Loher, P., Telonis, A.G., and Rigoutsos, I. (2016). MINTbase: a framework for the interactive exploration of mitochondrial and nuclear tRNA fragments. *Bioinformatics* *32*, 2481–2489.
- Pliatsika, V., Loher, P., Magee, R., Telonis, A.G., Londin, E., Shigematsu, M., Kirino, Y., and Rigoutsos, I. (2018). MINTbase v2.0: a comprehensive database for tRNA-derived fragments that includes nuclear and mitochondrial fragments from all The Cancer Genome Atlas projects. *Nucleic Acids Res.* *46* (D1), D152–D159.
- Sanna-Cherchi, S., Khan, K., Westland, R., Krithivasan, P., Fievet, L., Rasouly, H.M., Ionita-Laza, I., Capone, V.P., Fasel, D.A., Kiryluk, K., et al. (2017). Exome-wide Association Study Identifies GREB1L Mutations in Congenital Kidney Malformations. *Am. J. Hum. Genet.* *101*, 1034.
- Shah, R., Yeri, A., Das, A., Courtright-Lim, A., Ziegler, O., Gervino, E., Ocel, J., Quintero-Pinzon, P., Wooster, L., Bailey, C.S., et al. (2017). Small RNA-seq during acute maximal exercise reveal RNAs involved in vascular inflammation and cardiometabolic health: brief report. *Am. J. Physiol. Heart Circ. Physiol.* *313*, H1162–H1167.
- Si, H., Banga, R.S., Kapitsinou, P., Ramaiah, M., Lawrence, J., Kambhampati, G., Gruenwald, A., Bottinger, E., Glicklich, D., Tellis, V., et al. (2009). Human and murine kidneys show gender- and species-specific gene expression differences in response to injury. *PLoS ONE* *4*, e4802.
- Skog, J., Würdinger, T., van Rijn, S., Meijer, D.H., Gainche, L., Sena-Esteves, M., Curry, W.T., Jr., Carter, B.S., Krichevsky, A.M., and Breakefield, X.O. (2008). Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat. Cell Biol.* *10*, 1470–1476.
- Tang, Y.T., Huang, Y.Y., Zheng, L., Qin, S.H., Xu, X.P., An, T.X., Xu, Y., Wu, Y.S., Hu, X.M., Ping, B.H., and Wang, Q. (2017). Comparison of isolation methods of exosomes and exosomal RNA from cell culture medium and serum. *Int. J. Mol. Med.* *40*, 834–844.
- Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* *282*, 1145–1147.
- Tosar, J.P., Gámbaro, F., Darré, L., Pantano, S., Westhof, E., and Cayota, A. (2018). Dimerization confers increased stability to nucleases in 5' halves from glycine and glutamic acid tRNAs. *Nucleic Acids Res.* *46*, 9081–9093.
- Turchinovich, A., and Burwinkel, B. (2012). Distinct AGO1 and AGO2 associated miRNA profiles in human cells and blood plasma. *RNA Biol.* *9*, 1066–1075.
- Turchinovich, A., Weiz, L., Langheinz, A., and Burwinkel, B. (2011). Characterization of extracellular circulating microRNA. *Nucleic Acids Res.* *39*, 7223–7233.
- Valadi, H., Ekström, K., Bossios, A., Sjöstrand, M., Lee, J.J., and Lötvall, J.O. (2007). Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* *9*, 654–659.
- Vickers, K.C., Palmisano, B.T., Shoucri, B.M., Shamburek, R.D., and Remaley, A.T. (2011). MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. *Nat. Cell Biol.* *13*, 423–433.
- Williams, Z., Ben-Dov, I.Z., Elias, R., Mihailovic, A., Brown, M., Rosenwaks, Z., and Tuschl, T. (2013). Comprehensive profiling of circulating microRNA via small RNA sequencing of cDNA libraries reveals biomarker potential and limitations. *Proc. Natl. Acad. Sci. USA* *110*, 4255–4260.
- Yano, H., Maruiwa, M., Iemura, A., Mizoguchi, A., and Kojiro, M. (1992). Establishment and characterization of a new human extrahepatic bile duct carcinoma cell line (KMBC). *Cancer* *69*, 1664–1673.
- Yeri, A., Courtright, A., Reiman, R., Carlson, E., Beecroft, T., Janss, A., Siniard, A., Richholt, R., Balak, C., Rozowsky, J., et al. (2017). Total Extracellular Small RNA Profiles from Plasma, Saliva, and Urine of Healthy Subjects. *Sci. Rep.* *7*, 44061.
- Yeri, A., Courtright, A., Danielson, K., Hutchins, E., Alsop, E., Carlson, E., Hsieh, M., Ziegler, O., Das, A., Shah, R.V., et al. (2018). Evaluation of commercially available small RNAseq library preparation kits using low input RNA. *BMC Genomics* *19*, 331.
- Zhang, H., Freitas, D., Kim, H.S., Fabijanic, K., Li, Z., Chen, H., Mark, M.T., Molina, H., Martin, A.B., Bojmar, L., et al. (2018). Identification of distinct nanoparticles and subsets of extracellular vesicles by asymmetric flow field-flow fractionation. *Nat. Cell Biol.* *20*, 332–343.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-CD63 Antibody (for immunoprecipitation)	BD PharMingen	556019; RRID:AB_396297
Anti-CD81 Antibody (for immunoprecipitation)	Santa Cruz Biotechnology	sc-166029; RRID:AB_2275892
Anti-CD9 Antibody (for immunoprecipitation)	Santa Cruz Biotechnology	sc-13118; RRID:AB_627213
Anti-AGO2 Antibody (for immunoprecipitation)	Abcam	ab57113; RRID:AB_2230916
Anti-Flotillin1 Antibody (for western blot)	Cell Signaling Technology	18634; RRID:AB_2773040
Anti-CD9 Antibody (for western blot)	Abcam	ab92726; RRID:AB_10561589
Anti-APOA1 Antibody (for western blot)	Santa Cruz Biotechnology	sc-376818
Anti-AGO2 Antibody (for western blot)	Abnova	H00027161-A01; RRID:AB_1200191
Rabbit anti-mouse IgG HRP (for western blot)	Santa Cruz Biotechnology	sc-358914; RRID:AB_10915700
Goat anti-rabbit IgG HRP (for western blot)	ThermoFisher	31460; RRID:AB_228341
Biological Samples		
Primary Neonatal Rat Ventricular Myocytes	Prepared in house	Melman et al., 2015
Chemicals, Peptides, and Recombinant Proteins		
mTESR1 medium	Stem Cell Technologies	85850
Growth Factor Reduced Matrigel	Corning	354230
Dulbecco's Modified Eagle Medium	ThermoFisher	11965092
Antibiotic-Antimycotic	ThermoFisher	15240062
Protease inhibitor cocktail	EMD Millipore	539134
EZ-Link Sulfo-NHS-LC-Biotin reagent	ThermoFisher	21327
Dynabeads MyOne Streptavidin T1	Invitrogen	65601
20X PBS	ThermoFisher	28348
Qiazol Lysis Reagent	QIAGEN	79306
Iodixanol	Sigma-Aldrich	D1556-250ML
Critical Commercial Assays		
miRNeasy Micro kit	QIAGEN	217084
ExoRNeasy Midi	QIAGEN	77044
ExoRNeasy Maxi kit	QIAGEN	77064
ExoQuick plasma Prep and Exosome Precipitation kit	System Biosciences	EXOQ5TM-1
ExoQuick Serum tissue culture kit	System Biosciences	RA800-TC1
SeraMir Exosome RNA Purification Column Kit	System Biosciences	RA808A-1
Plasma/serum circulating and exosomal RNA Purification mini Kit	Norgen BioTek	51000
ME Kit	New England Peptide	ME-010
miRCURY Exosome Isolation kit	Exiqon	200101
miRCURY Biofluids RNA Isolation kit	Exiqon	300112/300113
miRNeasy mini kit	QIAGEN	217004
RNA 6000 Nano Pico Kit	Agilent Technologies	5067-1513
ExoView Tetraspanin Chip	NanoView Biosciences	EV-TC-TTS-01
ExoView Tetraspanin Labeling Antibodies	NanoView Biosciences	EV-TC-AB-01
NEBNext Small RNA Library Prep Set for Illumina (Multiplex Compatible)	New England Biolabs	E7330L
DNA Clean & Concentrator	Zymo Research	D4013
Iet-7a-5p TaqMan assay	ThermoFisher	Assay ID 000377, catalog #4427975

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
miR-16-5p TaqMan assay	ThermoFisher	Assay ID 000391, catalog #4427975
miR-223-3p TaqMan assay	ThermoFisher	Assay ID 002295, catalog #4427975
Deposited Data		
qRT-PCR data	See Table S1	See Table S1
Small RNAseq data, miRNA	See Table S2	See Table S2
Small RNAseq data, tRNA	See Table S5	See Table S5
Small RNAseq data, mRNA	See Tables S6 and S7	See Tables S6 and S7
Processed Small RNAseq data	GEO	GEO: GSE123865 and GEO: GSE123864
Experimental Models: Cell Lines		
WA09 Human Embryonic Stem Cells	Wisconsin Alumni Research Foundation	Thomson et al., 1998
KMBC cholangiocarcinoma cell line	From Dr. Masamichi Kojiro (Kurume University) via Drs. Gregory J Gores and Nicholas F LaRusso (Mayo Clinic)	Yano et al., 1992
Software and Algorithms		
exceRpt Small RNA-Seq Pipeline for exRNA Profiling	Genboree Bioinformatics	http://genboree.org/site/exrna_toolset/
Omics Explorer 3.3	Qlucore	https://qlucore.com/
R 3.5	R Foundation for Statistical Computing	https://cran.r-project.org
MINTmap	Loher et al., 2017, 2018	N/A
DESeq2	Love et al., 2014	N/A
CIBERSORT	Newman et al., 2015	N/A

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Louise C. Laurent (l Laurent@ucsd.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS**Biofluid samples from human subjects**

All human biofluid samples were collected with written consent from donors ≥ 18 years of age under an IRB protocols approved by the Human Research Protections Programs at UCSD (for plasma, serum, and urine) and Mayo Clinic Florida (for bile). Biofluid and RNA samples were labeled with study identifiers; no personally identifiable information was shared among participating laboratories. Raw data will be deposited in a controlled-access database (dbGAP).

Cell Cultures

WA09 human embryonic stem cells (hESC) were cultured in mTeSR1 medium (Stem Cell Technologies, # 85850) on standard tissue culture dishes coated with Growth Factor Reduced Matrigel at a 1:200 dilution (Corning, # 354230) in a humidified incubator at 37°C and 5%CO₂. Primary neonatal rat cardiomyocytes (NRVM) were isolated from 1 day old postnatal Sprague-Dawley pups using mechanical and collagenase-based enzymatic disruption as previously described ([Melman et al., 2015](#)) and cultured in Dulbecco's Modified Eagle Medium (DMEM, ThermoFisher, #11965092) on Primaria 60x15 mm dishes (Corning, #353802) in a humidified incubator at 37°C and 5%CO₂. KMBC cells were cultured in DMEM (ThermoFisher, #11965092) with 10% FBS and 1% Antibiotic-Antimycotic (ThermoFisher, #15240062) on standard tissue culture dishes coated with collagen (50 μ g/mL) in a humidified incubator at 37°C and 5%CO₂.

METHOD DETAILS**Biofluid collection**

Serum and plasma were collected from healthy adult donors, 10 female and 10 male. Briefly, blood was collected using 19 gauge needles (MONOJECT ANGEL WING Blood Collection Set 19G x 3/4," ITEM# 79027, MFG#8881225281, Moore Medical) and 60 mL syringes from a peripheral vein and transferred to 50 mL conical tubes. For serum, after 45-60 min at room temperature to allow clotting to occur, the blood was centrifuged at 2000 xg for 20 min and the supernatant was transferred to fresh tubes. For plasma, the 60 mL syringes were pre-filled with 440 μ l 0.5M K₂EDTA to prevent coagulation. The uncoagulated blood was then

transferred to 50 mL conical tubes and centrifuged at 2000 xg for 20 min to remove cells and cell debris. The clear supernatant was transferred to fresh tubes. The resulting cell-free serum and plasma were pooled using equal volumes from each donor to make four pools: female serum, male serum, female plasma, and male plasma. These pools were split into 1.5 mL aliquots and stored at -80°C . In addition, serum and plasma from one female and one male donor were also aliquoted and stored. Aliquots of these standardized samples were distributed to each of three (for serum) or four (for plasma) labs. Three “Core” exRNA isolation methods (ExoRNeasy (QIAGEN), miRNeasy (QIAGEN), and miRCURY Biofluids (Exiqon)) were tested in all labs, and 7 methods (Exosomal RNA Purification Kit (Norgen Biotek) w/o Amicon 3K, Exosomal RNA Purification Kit (Norgen Biotek) w/ Amicon 3K, Ultra + miRNeasy, Millipore + miRNeasy, ME Kit (New England Peptide) + miRNeasy, ExoQuick+Seramir (SBI), miRCURY Exosome Isolation + miRCURY Biofluids (Exiqon)) were tested in a subset of labs (Figure S1). exRNA isolation was performed in triplicate for each method in each lab. An input volume of 500 μL and elution volume of 30 μL were used for each plasma and serum exRNA isolation

CCCM was collected from three cell types in three different labs: human embryonic stem cells (hESCs at Lab5); primary rat cardiomyocytes (NRVM at Lab2); and cholangiocarcinoma cells (KMBC at Lab6). Cells were seeded at 1 million cells/mL and allowed to attach for 24 hours. The media was then replaced with serum-free (hESCs and NRVM cells, for which the growth media do not contain serum) or EV-cleared media (for KBMC) and the supernatant was collected after 24 hours. Immediately after collection of the CCCM, cells and cell debris were removed by centrifugation at 2,000 xg for 10 min. The cell-free supernatants were each split into 12 aliquots and stored at -80°C . The aliquots for each cell line were divided into three sets. Each lab retained one set and sent the other two sets to the other two labs, and then each of the three labs performed exRNA isolation experiments in triplicate for the supernatants from each cell type using four methods: ExoRNeasy; Ultra; Millipore; and ExoQuick (Figure S1). Upon thawing of the CCCMs, Lab5 proceeded directly to exRNA isolation, while Lab2 and Lab6 passed the thawed supernatants through 0.22 μm filters prior to exRNA isolation. An input volume of 4 mL and elution volume of 30 μL were used for each CCCM exRNA isolation

Bile was collected from three patients with suspected cholangiocarcinoma (P1-3) and one healthy adult donor (P4). Cells and cell debris were removed by centrifugation at 3000 xg for 10 min at 4°C and the cell-free bile was stored at -80°C in 1 mL aliquots. One lab performed exRNA isolation experiments in triplicate for the bile from each donor using five methods: ExoRNeasy; miRNeasy; Exiqon; Ultra; and Millipore (Figure S1). An input volume of 200 μL and elution volume of 30 μL were used for each bile exRNA isolation.

Urine was collected from healthy adult donors, 10 female and 10 male (the same donors as for plasma and serum). Cells and cell debris were removed by centrifugation at 2,000 xg for 10 min, and the resulting cell-free urine was pooled using equal volumes from each donor to make two pools: female urine pool and male urine pool. These pools were split into 1.5 mL aliquots and stored at -80°C . One lab performed exRNA isolation experiments in triplicate for the urine from each donor using five methods: ExoQuick; ExoRNeasy; Homebrew; Millipore; and Ultra (Figure S1). An input volume of 500 μL and elution volume of 30 μL were used for each urine exRNA isolation.

exRNA Isolation

Commercially available exRNA isolation kits were used in addition to ultracentrifugation followed by miRNeasy RNA extraction from the pelleted material and a “homebrew” exRNA isolation method. 500 μL of serum/plasma/urine or 200 μL of bile or 4 mL of CCCM was used as the starting material for each RNA isolation. Total RNA was isolated from the biofluids using the manufacturer’s recommended protocol for commercial kits with the following modifications.

QIAGEN miRNeasy Micro kit (catalog # 217084) – 5x volumes of the QIAzol Lysis Reagent was added to the biofluids and incubated for 5 min. To this equal volume of chloroform was added, incubated for 3 min and centrifuged for 15 min at 12,000 xg at 4°C . The RNA in the aqueous phase was precipitated by adding 1.5 x volumes of 100% ethanol and then loaded on to MinElute spin column and centrifuged at 1,000 xg for 15 s. The columns were then washed with 700 μL Buffer RWT, 500 μL Buffer RPE and 500 μL 80% ethanol consecutively by centrifuging for 15 s at $\geq 8000 xg$. After a final drying spin at full speed for 5 min, RNA was eluted in 14 or 30 μL RNase-free water directly to the center of the spin column membrane and centrifuging for 1 min at 100 xg followed by 1 min at full speed.

QIAGEN ExoRNeasy Midi kit (catalog # 77044 – serum/plasma/urine/bile) or QIAGEN ExoRNeasy Maxi kit (catalog # 77064 – CCCM) – The biofluids were mixed with an equal volume of XBP buffer, loaded onto the exoEasy spin column and centrifuged for 1 min at 500 xg . The column was then washed with 800 μL (serum/plasma/urine/bile) /10 mL (CCCM) of XWP buffer at 5,000 xg for 5 min. To extract RNA, 700 μL Qiazol was added directly to the column and centrifuged for 5 min at 5,000 xg . The lysed samples were then incubated with 90 μL of chloroform and centrifuged at 12,000 xg for 15 min at 4°C . The RNA in the aqueous phase was precipitated with 2x volumes of 100% ethanol, loaded onto a minElute spin column (QIAGEN, part of miRNeasy micro kit – catalog # 217084) and centrifuged at 1000 xg for 15 s. The spin columns were then washed once with 700 μL RWT and twice with 500 μL RPE buffers. After a drying spin for 5 min at full speed, the RNA was eluted with 14 or 30 μL water with a slow 1 min spin at 100 xg followed by a spin at full speed.

System Biosciences ExoQuick plasma Prep and Exosome Precipitation kit (catalog # EXOQ5TM-1 – serum/plasma/urine) or ExoQuick Seramir tissue culture kit (catalog # RA800-TC1) –125 μL ExoQuick Exosome Precipitation Solution was added to 500 μL of serum/plasma/urine and incubated for 30 min at 4°C . Before the addition of the precipitation solution, plasma was incubated for 5 min with 5 μL thrombin (500U/mL) and centrifuged at 10,000 xg for 5 min. The fibrin free supernatant was then treated with the precipitation solution and centrifuged at 1,500 xg for 30 min. The resulting pellet was resuspended in 50 μL sterile PBS. RNA

was then isolated using SeraMir Exosome RNA Purification Column Kit (catalog # RA808A-1). The resuspended pellet was incubated with 350 μ L Lysis Buffer for 5 min. After addition of 200 μ L of 100% ethanol, the mixture was loaded onto the spin column and centrifuged for 1 min at 13,000 rpm. The columns were washed twice with 400 μ L Wash Buffer and centrifuged at 13,000 rpm. The columns are spin for an additional 2 min to dry the column. The RNA is eluted in 30 μ L of water with a slow 2 min spin at 2000 rpm followed by a 1 min spin at 13,000 rpm. For CCCMs, 800 μ L of ExoQuick-TC was added to 4 mL of CCCM, incubated at 4°C overnight and centrifuged at 13000 rpm for 2 min. The RNA was then isolated from the pellet using the SeraMir Exosome RNA Purification Column kit.

Millipore – For serum/plasma/urine/bile the filter device (AU-0.5 filter – EMD Millipore catalog # UFC501024) was equilibrated by centrifuging with 500 μ L of PBS for 10 min at 14,000 x g. Next, 500 μ L of sample (200 μ L for bile) was applied to the AU-0.5 filter and centrifuged for 30 min at 14,000 x g. 500 μ L of PBS was added to the filter and again centrifuged for 30 min at 14,000 x g. The concentrated extracellular vesicles were collected by placing the filter upside down in a fresh microcentrifuge tube and centrifuging for 2 min at 2000 x g. For CCCMs the AU-15 millipore filter (EMD Millipore catalog # UFC901024) was washed with 2 mL of PBS for 10 min at 3,500 – 4,000 xg in a swinging bucket rotor. 4 mL of the precleared CCCM was added to the filter and centrifuged at 3,500 – 4,000 xg in a swinging bucket rotor for 30-50 min until 200 – 500 μ L of sample remained in the filter. The RNA was then isolated using the miRNeasy micro protocol described above.

Plasma/serum circulating and exosomal RNA Purification mini Kit (Norgen BioTek catalog # 51000) with and without Amicon 3K (catalog # UFC500396) – 500 μ L of biofluid was incubated with 100 μ L warmed PS Solution A and 900 μ L warmed PS Solution B (containing 2-Mercaptoethanol) for 10 min at 60°C. To this 1.5 mL of 100% ethanol was added and centrifuged for 30 s at 100 xg. To the pellet 750 μ L PS Solution C was added and incubated for 10 min at 60°C. 750 μ L 100% Ethanol was added, loaded onto the filter column and centrifuged for 1 min at 16000 xg. The column was washed thrice with 400 μ L Wash Solution for 1 min at 16000 xg and then centrifuged once again for 3 min to dry the membrane. RNA was eluted in 30 μ L water with a slow spin for 2 min at 300 xg followed by a spin for 3 min at 16000 xg. For samples that were processed further with Amicon filtration, the eluted RNA was diluted with 320 μ L water, loaded onto the Amicon filter (3K) and spun for 8 min at 14,000 xg. The RNA is then collected by inverting the column and centrifuging for 2 min at 8,000 xg

New England Peptide ME Kit (catalog # ME-010) - Protease inhibitor cocktail (5 μ L) (EMD Millipore catalog # 539134) was added to 500 μ L of serum/plasma diluted with equal volume of PBS. After centrifugation for 7 min at 10,000 x g at room temperature to remove debris, 20 μ L reconstituted Vn96 peptide was added to the supernatant and incubated at room temperature for 30 min on a rotator. Extracellular vesicles were precipitated by centrifuging for 7 min at 10,000 xg. The pellet was washed with 1 mL PBS and 5 μ L protease inhibitor cocktail twice by centrifuging for 7 min at 10,000 xg. The washed pellet was then lysed with 700 μ L of Qiazol and subjected to RNA isolation using miRNeasy micro kit as described above.

miRCURY Exosome Isolation kit (catalog # 200101) + miRCURY Biofluids RNA Isolation kit (catalog # 300112/300113) – Extracellular vesicles were precipitated from 500 μ L of serum or plasma by incubating with 200 μ L of Precipitation Buffer A for 60 min at 4°C. The samples were centrifuged at 1500 xg for 30 min, and the pellet was resuspended in 270 μ L of Resuspension buffer. RNA was isolated from either the whole biofluid or the resuspended extracellular vesicles using the miRCURY RNA isolation kit. Resuspended solution was incubated for 10 min with 90 μ L of Lysis Buffer solution and then with 30 μ L of Protein Precipitation solution for 1 min. The samples were then centrifuged at 11000 xg for 3 min, clear supernatant was transferred to a new collection tube and 400 μ L of isopropanol was added. After that, the samples were loaded onto microRNA Spin columns, centrifuged at 11000 xg, washed with the provided wash solutions and finally the RNA was eluted in 30 μ L of water. The elution was carried out with a slow 1 min spin at 100 xg followed by a spin at 11,000 xg.

Homebrew – For this method, 500 μ L of a heated 2x lysis buffer with 0.1 M Tris, 12.5 mM EDTA, 0.15 M NaCl, 6.8% SDS and 0.67 mg of Proteinase K was added to 500 μ L urine and incubated for 10 min at 60°C. To this, 1.2x volumes of GITC(4 M)/ β -Mercaptoethanol (0.1 M) buffer, 1:100 volume of 3M NaOAc and an equal volume of Acid Phenol was added and incubated for 5 min at 60°C. Following this, equal volume of Chloroform:Isoamyl alcohol was added and centrifuged at 12,000 xg for 15 min at 4°C. The RNA in the aqueous phase was precipitated by adding 1.5x volumes of 100% ethanol and loaded onto the QIAGEN minElute spin columns. The rest of the protocol was identical to the miRNeasy micro kit method.

Ultracentrifugation - For the Ultra method, the centrifuge time, speed and model varied between the labs. serum/plasma/bile/urine were brought upto 3 mL with PBS. The samples were centrifuged at 100,000 xg by all labs except Lab 6, where the samples were centrifuged at 120,000 xg. Labs 5 and 6 centrifuged the samples for 70 min, labs 1 and 3 for 90 min and lab 2 did the spin overnight (16h). The models of centrifuges and rotors used were as follows – Lab 1 (Optima-Maxx, MLA-55), Lab 2 (Beckman L8-M, SW-41 ti), Lab 3 (Optima-Maxx-TL, TLA-100.3), Lab 6 (Optima-L100 XP, 70Ti), Lab 5 (Optima-Maxx-XP, MLS-50). The pellet was washed with 3 mL PBS and centrifuged once again. The washed pellet was then lysed with 700 μ L Qiazol and subjected to RNA isolation using miRNeasy micro kit as described above.

qRT-PCR

qRT-PCR was performed using TaqMan primer-probes designed to quantify: let-7a-5p (ThermoFisher, Assay ID 000377, catalog #4427975); miR-16-5p (ThermoFisher, Assay ID 000391, catalog #4427975); and miR-223-3p (ThermoFisher, Assay ID 002295, catalog #4427975). For each sample and each miRNA target, qRT-PCR reactions were performed in triplicate and averaged.

Immunoprecipitation of exRNA carriers

Antibody biotinylation: Antibodies raised against CD63, CD81, CD9, and AGO2 were used. Sodium azide was removed from antibody stocks using the Zeba spin desalting column (7K MWCO, 0.5 ml, Thermo Fisher Scientific, Cat#89882). Antibodies were then biotinylated using the EZ-Link Sulfo-NHS-LC-Biotin reagent (ThermoFisher, Cat#21327), following manufacturer's protocol. Briefly, 10 mM biotin solution was prepared by dissolving 1 mg of no-weight Sulfo-NHS-LC-Biotin in 180 μ L ultrapure water (purified by Milli-Q Biocel System). Appropriate volume of biotin was added to antibody in order to gain about 20-fold excess biotin-to-antibody molar ratios. The mixture was incubated at room temperature for 2 hr. The biotinylated antibody was then filtered using another desalting column and the final concentration of the biotinylated antibody was measured using a NanoDrop UV spectrophotometer (ThermoFisher) based on absorption at 280 nm.

Magnetic bead preparation: Dynabeads MyOne Streptavidin T1 (Invitrogen, Cat#65601) suspension was transferred to 2.0 mL microcentrifuge tube and placed on the DynaMag-2 magnetic rack followed by aspiration of supernatant. The tube was removed from the magnetic rack and washed with 0.01% Tween-20. Washing step was repeated twice. For blocking purpose, the beads were washed 3 times in PBS containing 0.1% BSA prior to use.

Immunoprecipitation: The immunoprecipitation procedure was performed by incubating the serum with antibody conjugated beads. Briefly, serum from non-pregnant females was thawed and diluted 1:1 with double filtered 1X PBS (Pierce™ 20X PBS, ThermoFisher, Cat#28348). Every 1,000 μ L of serum was invert-mixed with 6 μ g biotinylated antibody for 20 min at RT on a HulaMixer Sample Mixer (ThermoFisher) at 10 rpm. Then, 390 μ L of washed Dynabeads was added to the mixture and invert-mixed for 25 min at RT on a Hula mixer at 10 rpm. The mixture was then washed three times with 0.1% BSA and subjected to RNA extraction.

RNA extraction from Dynabeads: RNA was extracted using the miRNeasy mini kit (QIAGEN, Cat#217004) following manufacturer's protocol. In brief, the Dynabeads were subjected to phenol/chloroform extraction step for RNA extraction using Qiazol Lysis Reagent (QIAGEN, Cat#79306) followed by chloroform. The aqueous phase was used as input into the miRNeasy procedure and the RNA was eluted in 14 μ L of nuclease-free water. To avoid contamination with genomic DNA, the RNA samples were also treated with deoxyribonuclease I (DNase I, Invitrogen). The quality of RNA was assessed by using the RNA 6000 Nano Pico Kit (Agilent Technologies, Cat#5067-1513) and the Bioanalyzer 2100 (Agilent Technologies). The eluted RNA was dried down using a speedvac, and used as input into the small RNaseq library preparation process. Small RNaseq libraries were generated and size selected as described above.

Fractionation of plasma serum using Cushioned Density Gradient Ultracentrifugation

A volume of 0.8 mL of each serum or plasma sample was individually mixed with 38 mL of PBS, placed into an ultracentrifuge tube, and underlaid with 2 mL 60% iodixanol (OptiPrep, Sigma-Aldrich) as recently described (Li et al., 2018a). The tubes were spun at 100,000 xg for 2 hours at 4°C in a Type 50.2 Ti rotor. The bottom 3 mL (2 mL iodixanol cushion + 1 mL supernatant) was removed, mixed, and underlaid under a step gradient of iodixanol (5%–10%–20% iodixanol diluted in 0.25 M sucrose, 1 mM EDTA, and 10 mM Tris-HCl, pH 7.4). This was spun at 100,000 xg for 18 hours at 4°C. Twelve 1 mL fractions were then collected, starting from the top of the gradient. Nanoparticle Tracking Analysis (NanoSight) was performed on the fractions. Fractions 1-3, 4-7, and 9-12 were pooled and RNA was extracted using miRNeasy from 500 μ L of each pool, as well as 500 μ L of unfractionated serum from each donor, and small RNaseq libraries were generated and size selected as described above. 37.5 μ L of each pool was also used for western blot (see below). The refractive index of the iodixanol fractions were measured using the RBC-6000 Refractometer (LAXCO), and the obtained values converted to density based on a standard curve derived from measurements on 10%, 20%, 40%, and 60% iodixanol solutions.

Lipoprotein Fractionation from plasma

Lipoproteins were fractionated from cell-free human plasma by sequential density ultracentrifugation as previously described (Li et al., 2018b). Male plasma was cleared of cells by centrifugation at 3,000 xg for 10 min. The cell-free plasma was centrifuged at 52,000 rpm for 16 h at 8°C in a TLA 100.3 rotor in an Optima TL Ultracentrifuge (Beckman instruments, Fullerton, CA) and the top 20% containing chylomicrons was removed. The remaining material was adjusted to 1.021 g/mL with KBr, and after centrifugation for another 16 h using the same parameters, the top fraction containing VLDL was collected. The remainder was adjusted to a density of 1.063 g/mL with KBr and centrifuged for an additional 16 h as described above, and the top fraction containing IDL/LDL was removed. Finally, the remaining material was adjusted to 1.21 g/mL with KBr and centrifuged at 16 hr as described above, and the top fraction containing HDL was collected. All lipoprotein preparations were extensively dialyzed against PBS.

Western Blot

Western blot was performed on the pooled iodixanol fractions from C-DGUC of the female and male plasma and serum fractions and the input plasma and serum samples using the primary antibodies raised against Flotillin-1, CD9, APOA-I, and AGO2. 37.5 μ L of the pooled iodixanol fractions from C-DGUC of the female and male plasma and serum fractions and 1 μ L of the input plasma and serum samples were resolved on a 10% Tris-glycine polyacrylamide gel and transferred onto PVDF. The membrane was blocked with 5% non-fat milk for 1 hour at room temperature. Primary antibodies were diluted 1:500 in 1% milk and blots were probed overnight at 4°C. Blots were washed four times for five minutes with 0.1% PBST and incubated with either anti-mouse IgG HRP (Santa Cruz) or anti-rabbit IgG HRP (ThermoFisher) at a dilution of 1:1000 for 1 hour at room temperature. Membrane were then washed four times for five minutes with 0.1% PBST and rinsed with PBS. Blots were incubated with Amersham ECL Prime (GE Life Sciences) and imaged using

the ImageQuant LAS 4000. Nanoparticle tracking analysis was performed on an LM14 NanoSight instrument (Malvern). The samples were diluted with PBS to create 1 mL solutions containing 10^8 - 10^9 nanoparticles. To achieve this concentration range, samples were typically diluted between 1:50 to 1:400. Data were collected as a mean reading of three videos of one minute in length with parameters being set at a camera level of 13 and detection threshold of 3.

Single Particle Interferometric Reflectance Imaging Sensing (NanoView) analysis

Plasma samples were purified using qEV columns (IZON) and fractions 5-8 were combined. 35 μ L of the combined fractions were incubated on the ExoView Tetraspanin Chip (NanoView Biosciences, EV-TC-TTS-01) placed in a sealed 24 well plate for 16 hours at room temperature. The ExoView Tetraspanin Chips have antibodies against CD81, CD63, CD9, and Mouse IgG1 Isotype control in triplicate. After incubation, the ExoView chips were washed on an orbital shaker once with PBST (0.05% Tween-20) for 3 min, and then 3 times in PBS for 3 min. Then chips were then incubated with ExoView Tetraspanin Labeling Antibodies (NanoView Biosciences, EV-TC-AB-01) that consist of (anti-CD81 Alexa-555, anti-CD63 Alexa 488, and anti-CD9 Alexa-647). The antibodies were diluted 1:6000 in PBST with 2% BSA. The chips were incubated with 250 μ L of the labeling solution for 2 hours at room temperature without shaking. The chips were then washed once in PBST for 3 min, 3 times in PBS for 3 min, rinsed in filtered DI water, and dried. The chips were then imaged using the ExoView R100 reader with the ExoScan 2.5.5 acquisition software. The data were then analyzed using the ExoViewer 2.5.0 software package.

Small RNA sequencing library preparation

Small RNA sequencing libraries were constructed using the NEB Next small RNA library kit according to the manufacturer's protocol except for the following modifications. All reactions were conducted at 1/5th the recommended volume and the adapters were diluted to 1/6th the supplied concentration, with 18 cycles of PCR. For the RNA prepared from the cell, CCCM, plasma, and serum samples, 1.2 μ L of each RNA sample was used for small RNaseq library construction; for the bile and urine samples, 4 μ L of the purified RNA was dried down using a speedvac without heat, resuspended in 1.2 μ L water, and used as input for small RNaseq library preparation. The library product was then cleaned using a Zymo DNA clean and concentrate kit (Zymo Research, D4013). The libraries were then pooled (up to 48 samples) based on picogreen concentration measurements and proportion of the desired PCR product and adaptor dimers as observed on a Fragment Analyzer high sensitivity DNA array (Advanced Analytical). The pooled libraries were then size selected to remove adaptor dimers using the pippin prep HT. The lower limit of the size selection was set to 115 or 125 to remove the adaptor dimers. The upper limit was either 160/180 for plasma, serum, urine and bile or 150 for cells and CCCMs. The size selected libraries were then sequenced on an Illumina HiSeq 4000 as 50 cycle single end reads.

The resulting initial dataset was analyzed to determine the total input read (TIR) count and the total miRNA read (TMiR) count. Libraries that yielded < 10,000 TIRs were considered to have failed library preparation. Libraries that yielded \geq 10,000 TIRs and < 100,000 TMiRs were handled in one of three ways:

1. Libraries for which the percent TMiR \geq 0.08% and for which there was sufficient material (small RNA library or isolated exRNA sample) to obtain 100,000 TMiRs were re-run, and the data from the two runs were combined.
2. Libraries in which the percent TMiR (TMiR/TIR x 100) < 0.08% were not prepared again, as it would require > 12.5 million TIR to obtain 100,000 TMiRs.
3. Libraries for which the percent TMiR \geq 0.08% and for which there was insufficient material (small RNA library or isolated exRNA sample) to obtain 100,000 TMiRs were not re-run.

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of qRT-PCR data

For the bile, CCCM, and urine experiments, the qRT-PCR replicates were averaged prior to analysis (to focus on non-technical sources of variability). A linear model was estimated Ct modeled as a function of non-coding RNA target, exRNA isolation method, and laboratory site as independent variables. A type III sum of squares for each parameter was estimated, with the percent variability for each independent variable calculated as the ratio of the sum of squares for that variable divided by the model sum of squares. For plasma and serum experiments a repeated-measures ANOVA was used modeling Ct as a function of non-coding RNA target, sample number, exRNA isolation method, laboratory site, and sex. R 3.5 was used for analysis (R Foundation for Statistical Computing, Vienna, Austria).

Analysis of small RNA Sequencing data

Small RNaseq library quality control and calculation of distribution of RNA biotypes

For samples that were sequenced twice, the data from the two runs were combined prior to analysis. Libraries for which the number of total input reads (TIR) < 10,000 were considered to have failed library preparation. Libraries with total mapped reads (TMR) < 100,000 failed the TMR filter and were removed. We then calculated the percent of reads (compared to successfully clipped reads) that were mapped into these categories: rRNA; miRNA; tRNA; piRNA; other gencode sequences; and unmapped reads (Figure 1). For

the bile libraries, the fractions of each of the five RNA biotypes were compared between donor P1 and donor P2 (two-tailed t tests paired by kit type, [Table S2.13](#)) and among exRNA isolation methods (two-tailed t tests paired by kit type, [Table S2.13](#)). For the CCCM libraries, the fractions of each of the five RNA biotypes were compared among exRNA isolation methods (two-tailed t tests paired by kit type, [Table S2.14](#)). For plasma and serum, the fractions of each of the five RNA biotypes were compared among exRNA isolation methods (two-tailed t tests paired by kit type, [Table S2.15–16](#)). For urine, the fractions of each of the five RNA biotypes were compared among exRNA isolation methods (two-tailed t tests paired by kit type, [Table S2.17](#)).

miRNA Data Processing

For samples that were sequenced twice, the data from the two runs were combined prior to mapping. Data were mapped using the ExceRpt small RNA sequencing data analysis pipeline on the Genboree Workbench (http://genboree.org/site/exrna_toolset/). Mapping parameters included a minimum read length of 15 nucleotides, with 0 mismatches allowed.

In order to evaluate the quality, reproducibility, and complexity of the miRNA data from the libraries constructed from the exRNA samples isolated using each kit, the initial mapped miRNA dataset ([Table S2.2](#)) was processed through a series of filtering and normalization steps ([Table S2.3–11](#)), the specifics of which are described below. For each combination of sample type and kit, the number of libraries that passed each filtering step was tracked ([Table S2.12](#)). We observed that the number of failed samples and the reasons that samples were filtered out varied widely between biofluid type and exRNA isolation method. For bile, the number of Total Mapped Reads (TMRs) were ample ($\geq 590,722$) for all libraries, but even using a cutoff of 5,000 Total miRNA Reads (TmiRs), 2 out of 6 of the ExoRNeasy and Ultra libraries were excluded due to low fractions of miRNA reads (see below, [Figure 1](#)). For CCCM, 1 Millipore and 1 Ultra library had fewer than 100,000 TMRs, and 1 each of the Millipore, Ultra, and ExoQuick libraries had fewer than 5,000 TmiRs. Nearly all of the plasma and serum libraries had at least 100,000 TMRs. Using a TmiR cutoff of 100,000, high proportions ($> 25\%$) of the plasma libraries from the ME, miRCURY, Millipore, and miRNeasy did not pass filter, but all of the methods except miRCURY had high pass rates for serum. Among the biofluids, urine had the lowest pass filter rates, with about a third of the libraries from the Homebrew, Millipore, and miRNeasy samples having fewer than 100,000 TMRs, and $> 25\%$ of the remaining ExoQuick, Millipore, and miRNeasy libraries failing a TmiR cutoff of 5,000 ([Table S2.12](#)).

The initial set of miRNA data for all libraries was called the “miRNA_AllBiofluid_RawData” dataset ([Table S2.2](#)). Libraries for which the number of total input reads (TIR) $< 10,000$ were considered to have failed library preparation. Libraries with TMR $< 100,000$ failed the TMR filter and were removed to yield the “miRNA_AllBiofluid_TMR” dataset ([Table S2.3](#)). The libraries were further subjected to a quality assessment in which plasma and serum libraries with $< 100,000$ total miRNA reads (TmiRs) were removed, and bile, cell, CCCM, and urine libraries with $< 5,000$ TmiRs were removed, to yield the “miRNA_AllBiofluid_TmiR” dataset ([Table S2.4](#)); the threshold was set lower for the bile, cell, CCCM, and urine libraries, as insufficient numbers of these libraries would pass the higher filter to allow for further analysis. The data were then normalized by linearly scaling each component library to a total miRNA read count of 1 million, such that the resulting counts were expressed as Reads Per Million Scaled miRNA Reads (RPMSmiR). The miRNAs in this “miRNA_AllBiofluid_Scaled” ([Table S2.5](#)) were then filtered to remove miRNAs with fewer than 3 bile, CCCM, or urine samples, or fewer than 5 plasma or serum samples that contained at least 10 or 100 RPMSmiR to yield the “miRNA_AllBiofluid_Scaled_Filtered_RPMSmiR10” and “miRNA_AllBiofluid_Scaled_Filtered_RPMSmiR100” ([Table S2.6–7](#)) datasets. For each biofluid type, the data for the relevant libraries were extracted from the “miRNA_AllBiofluid_Scaled_RPMSmiR100” dataset and the miRNAs were filtered as follows: 1) For the bile libraries, miRNAs with fewer than 3 samples that contained at least 100 RPMSmiR were removed to yield the “miRNA_bile_Scaled_Filtered” dataset ([Table S2.8](#)); 2) For the CCCM libraries, miRNAs with fewer than 3 samples that contained at least 100 RPMSmiR were removed to yield the “miRNA_cellCCCM_Scaled_Filtered” dataset ([Table S2.9](#)); 3) For the plasma and serum libraries, miRNAs with fewer than 5 serum or plasma samples that contained at least 100 RPMSmiR were removed to yield the “miRNA_plasmaserum_Scaled_Filtered” dataset ([Table S2.10](#)); For the urine libraries, miRNAs with fewer than 3 samples that contained at least 100 RPMSmiR were removed to yield the “miRNA_urine_Scaled_Filtered” dataset ([Table S2.11](#)).

miRDaR metrics calculations

From “miRNA_AllBiofluid_Scaled_Filtered_RPMSmiR10” ([Table S3](#)), for each biofluid/exRNA isolation method combination, we counted the number of libraries that passed our quality filters, and computed the percentage of pass-filter libraries compared to all sequenced libraries in that category ([Figures 3, S3, and S4](#)). For each miRNA in each biofluid/exRNA isolation method combination, the mean, range, and interquartile range of expression (in rpm), the standard deviation and %CV, the expression window (Mean_cut, 1:0–10 rpm, 2:10–100 rpm, 3:100–1000 rpm, 4:1,000–10,000 rpm, 5: $>10,000$ rpm), the quantiles of mean expression and %CV values, and the integrated quality score (IQS = sum of mean expression quantile and %CV quantile values) were calculated. Quantile values will be calculated only for miRNAs that were detectable for ≥ 4 methods for CCCM and ≥ 5 methods for all other biofluids. For investigator requests including fewer than 5 miRNAs, the mean IQS value for each method will be computed and will be used to rank the methods. For requests including ≥ 5 miRNAs, the 75th percentile IQS value and complexity will be computed and used to rank the methods (e.g., [Figures 3G–3H and S3P–S3T](#)).

mRNA Fragment Data processing

Data were mapped using the ExceRpt small RNA sequencing data analysis pipeline on the Genboree Workbench (<http://genboree.org/theCommons/projects/exrna-tools-may2014/wiki/Small%20RNA-seq%20Pipeline>). Mapping parameters included a minimum read length of 15 nucleotides, with 0 mismatches allowed. A detailed description of the data filtering and normalization procedures are provided in the [Supplemental Information](#).

miRNA and mRNA Fragment Data Analysis

Pairwise and Multi-group differential expression analyses were performed in the Qlucore Omics Explorer 3.3 using the “Two-Group” and “Multigroup” comparison tools. Principal Component Analysis, Hierarchical Clustering, and data visualization were also performed using the Qlucore Omics Explorer 3.3. Venn diagrams were constructed using InteractiVenn.(Heberle et al., 2015)

tRNA Data processing and analysis

tRF profiles from all samples, including NRVM, were processed using an established tRF pipeline. Briefly, short RNA-seq data were processed for quality control and adaptor trimming with cutadapt.(Martin, 2011) Reads were mapped to the tRFs using the MINTmap pipeline.(Loher et al., 2017, 2018) MINTmap enforces a perfect match to the genome without insertions or deletions while also reporting whether a fragment belongs unambiguously to tRNA space or can be found elsewhere in the human genome.(Pliatsika et al., 2018; Pliatsika et al., 2016) We pooled all tRF sequences that showed normalized expression ≥ 1.0 Reads Per Million (RPMMt) in at least one library and used them in our analyses.

miRNA deconvolution analysis

The objective of the deconvolution of miRNA expression of the plasma, serum and the vesicle fractionated Optiprep samples sequenced was to estimate the proportion of the immunoprecipitations (AGO2, CD9, CD63, CD81, and HDL) present in each of these samples. There were two major steps in the deconvolution analysis: A) To construct a signature miRNA set using differentially expressed miRNAs characteristic of each of the immunoprecipitations. B) Using the signature miRNA set to estimate proportions of these immunoprecipitations in each sample using a support vector regression model. Only female, non-pregnant samples that had at least 10,000 input reads and more than 15 miRNAs expressed with at least 10 read counts were included in the analyses. The samples from the various sources (plasma, serum, Iodixanol and the immunoprecipitations) were normalized separately using the median ratio method using DESeq2 (Love et al., 2014).

The signature miRNAs for the pulldown samples were selected by compiling the top 10 miRNAs for each pulldown (computed as the 10 miRNAs with the highest fold-change difference between their expression in a given pulldown compared to the pulldown with the next highest level of expression) (Table S4.2). Therefore, for the 5 pulldowns, 50 miRNAs were selected as the signature miRNAs. Deconvolution analysis was performed using the CIBERSORT (Newman et al., 2015) package that employs a linear support vector regression model to estimate proportions. The major advantages of Cibersort’s SVR-regression is that there is in-built variable selection to use the most appropriate miRNAs to estimate the proportions and an empirical p value determines the significance of the deconvolution regression.

DATA AND SOFTWARE AVAILABILITY

miRDaR (miRNA Detection- and Reproducibility-based selection of exRNA isolation methods) web application: <https://exrna.shinyapps.io/mirdar/>. Small RNA sequencing data mapped to miRNA, tRNA, and gencode annotations (raw and processed) are provided in Tables S1, S2, S3, S4, S5, S6, and S7. The accession number for the processed small RNA sequencing data reported in this paper is GEO: GSE123865 and GEO: GSE123864. The raw small RNA sequencing reads have been deposited in dbGAP. All tRF profiles can be interactively explored at <https://cm.jefferson.edu/MINTbase/>.

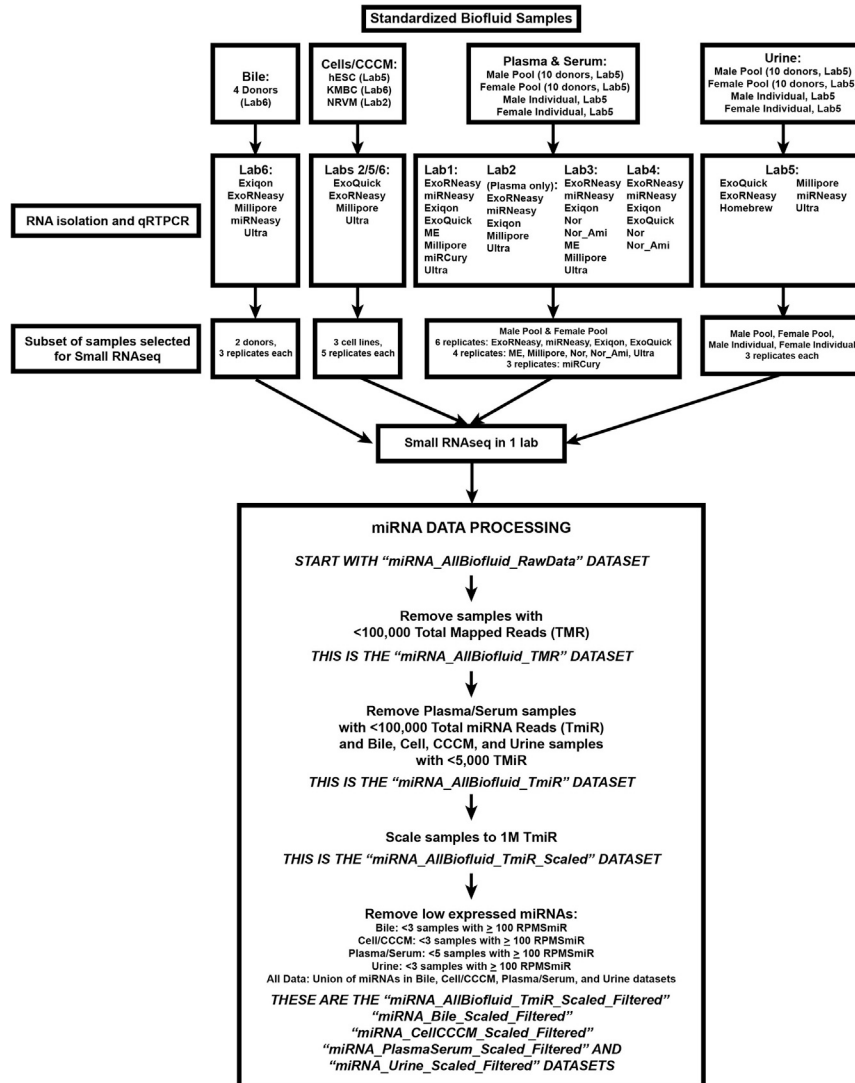


Figure S1. Flowchart Indicating Overall Project Plan, Related to Figure 1

For each biofluid, the exRNA isolation methods used and the number of laboratories that carried out the exRNA isolation the qRT-PCR experiments are indicated. miRNA data processing plan from the small RNaseq data is also shown.

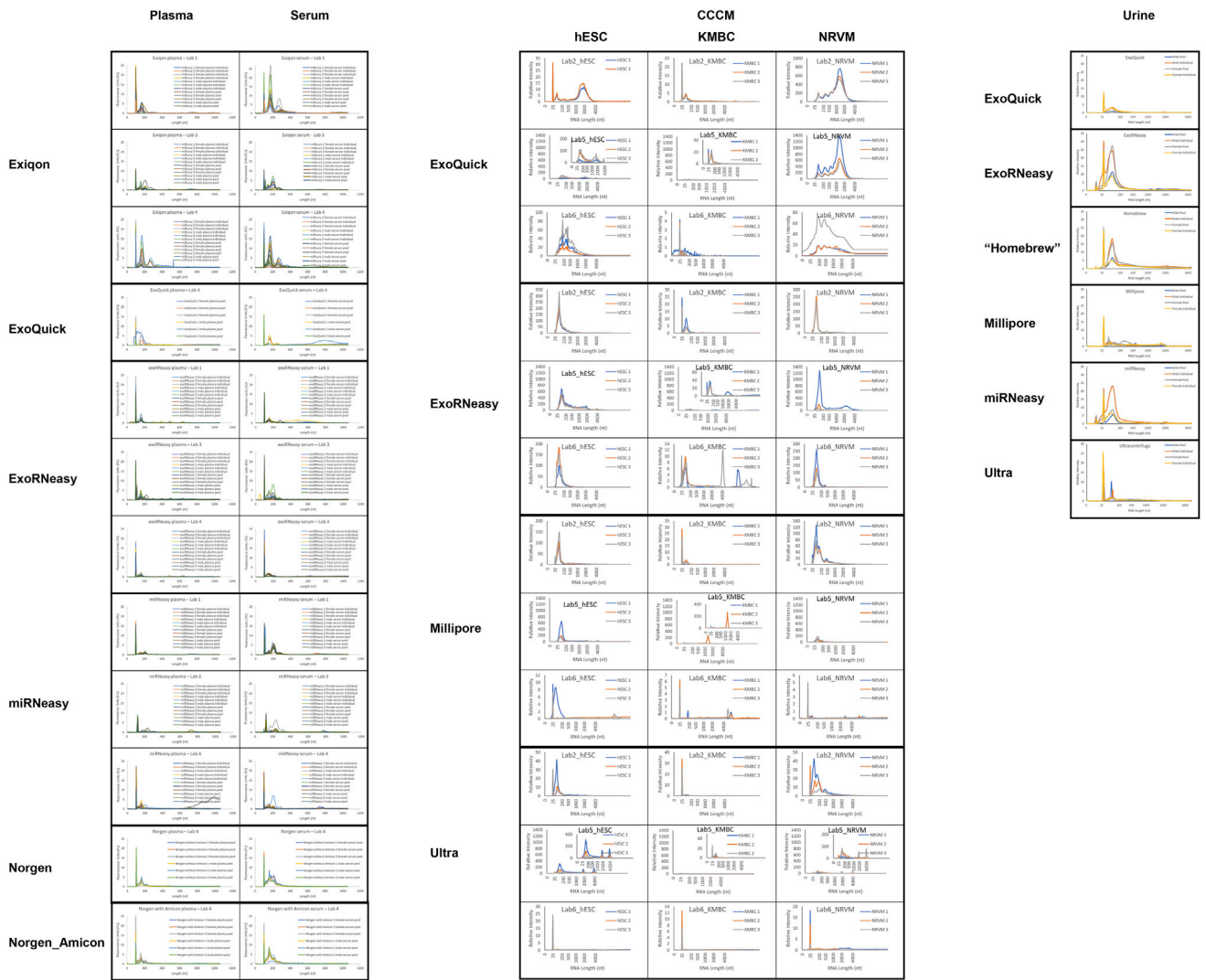


Figure S2. Bioanalyzer Traces for Extracellular RNA Samples, Related to Figure 1
 Each isolated RNA sample was run on a Bioanalyzer RNA Pico Chip (Agilent).

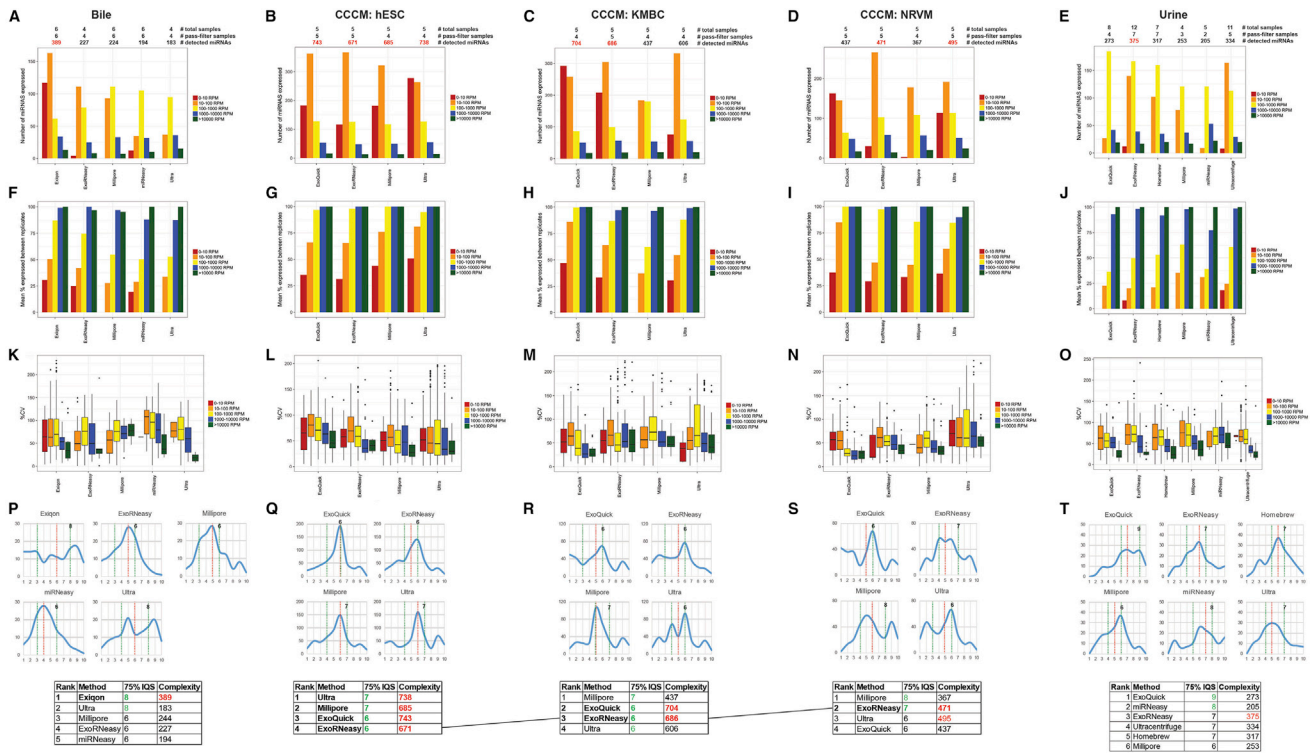


Figure S3. Complexity and Reproducibility of exRNA Isolation Methods for Bile, CCCM, Plasma, Serum, and Urine, Related to Figure 3 (A, F, K, P) Bile, (B-D, G-I, L-N, Q-S) CCCM, and (E, J, O, T) urine. For A-O, values are given for each RNA isolation method across different expression levels. A-E. Average number of miRNAs expressed. F-J. Mean % of replicates in which the miRNAs with the indicated mean expression level were detected. K-O. Boxplots indicating the %CV as a function of mean expression level. For plasma, no CV data are shown for the miRCURY Exosome kit, as only 1 sample passed filter. The box indicates the median and interquartile range (IQR), the whiskers indicate Q1-1.5*IQR and Q3+1.5*IQR, and the individual dots indicate outliers. P-T. Plots indicating the distribution of individual miRNA quality scores, comprised of the sum of the mean expression quantile (1 for lowest expression; 5 for highest expression) and %CV quantile (1 for highest %CV, 5 for lowest %CV), for each exRNA isolation method. The 25th, 50th, and 75th percentiles are indicated by the vertical dashed lines. In the table, the 75th percentile IQS (IQS75) is given and highlighted in green if it is the top value (IQS75max) or IQS75max-1, the complexity is highlighted in red if it is within 10% of the highest value among all methods, and methods that meet both of these criteria are bolded; the methods are sorted first by 75th percentile IQS and then by complexity.

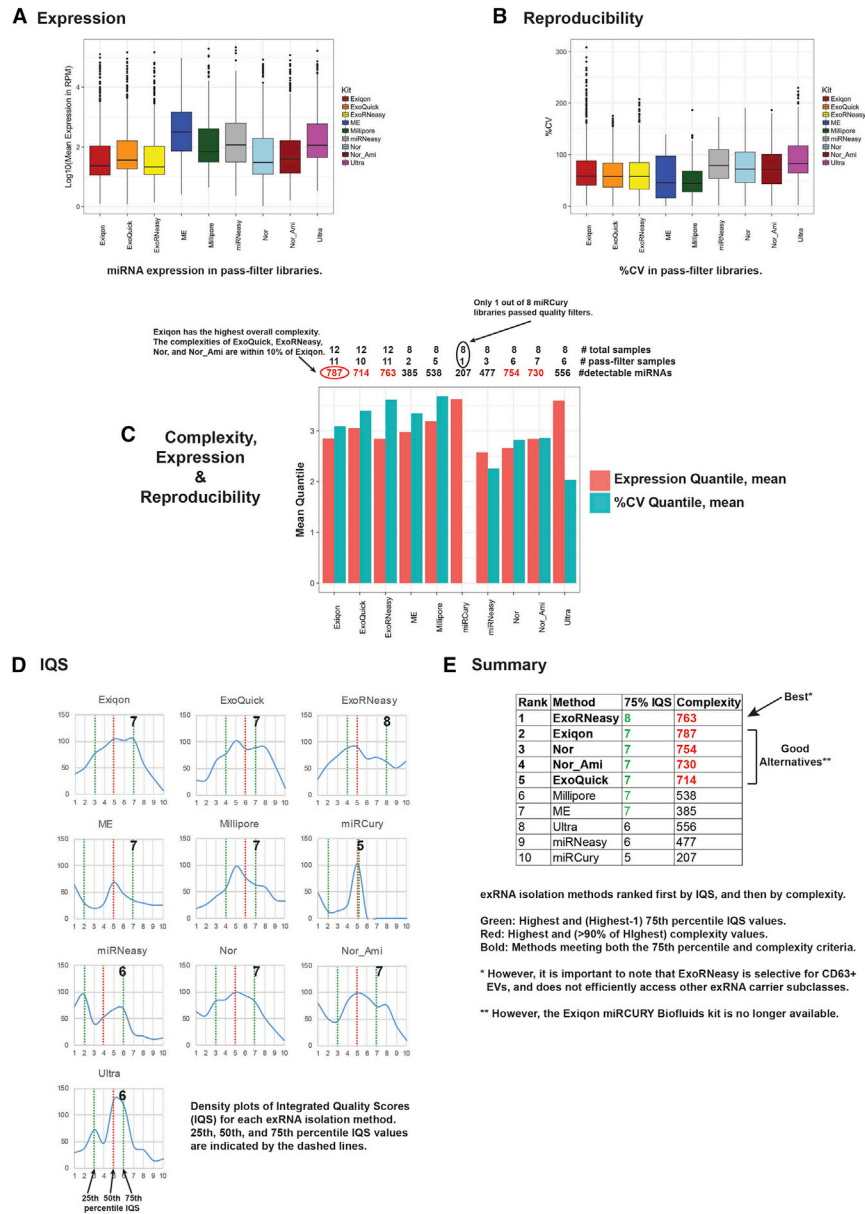


Figure S4. Example of the miRDaR results that will be returned for queries including 5 or more miRNAs, Related to Figure 3

The results shown are for all miRNAs detected from plasma. A. Boxplot displaying the mean miRNA expression level for each exRNA isolation method. B. Boxplot displaying the mean %CV for each method. C. Barchart showing the mean expression quantile and mean %CV quantile for each method. D. Density plots of the IQS values for the constituent miRNAs for each method. The 25th, 50th, and 75th percentile IQS values are marked with the dashed lines. E. Table ranking the exRNA isolation methods by 75th percentile IQS and complexity (computed as the number of miRNAs with non-zero IQS values). The 75th percentile IQS is highlighted in green if it is the top value (n) or n-1, the complexity is highlighted in red if it is within 10% of the highest value among all methods, and methods that meet both of these criteria are bolded. On the boxplots, the box indicates the median and interquartile range (IQR), the whiskers indicate Q1-1.5*IQR and Q3+1.5*IQR, and the individual dots indicate outliers.

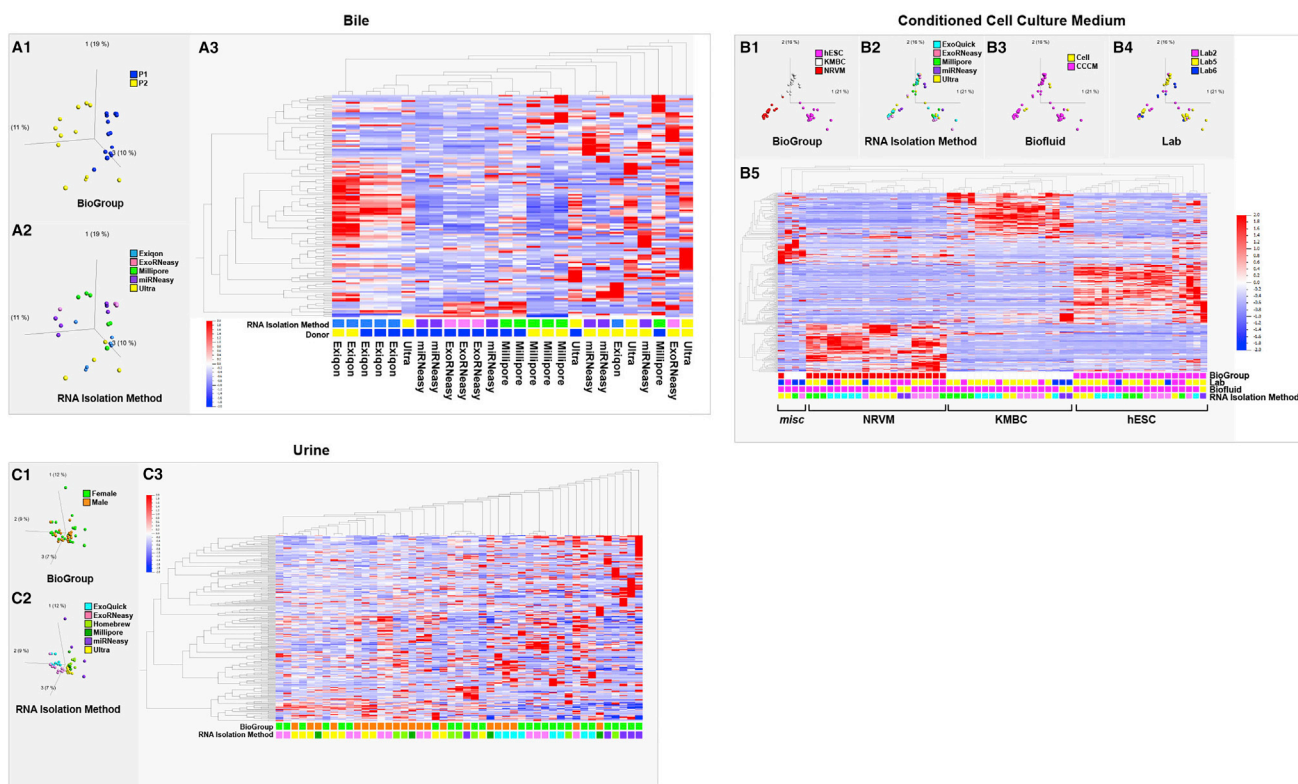


Figure S5. PCA and hierarchical clustering analysis of miRNA data for Bile, CCCM, and Urine, Related to Figure 5

A. Analysis of “Bile_Scaled_Filtered” dataset. A1. PCA plot with samples color coded by donor. A2. PCA plot with samples color coded by exRNA isolation method. A3. Heatmap showing hierarchical clustering results (biclustering of miRNAs and samples was performed). The donor and exRNA isolation method for each sample are color coded below the heatmap. B. Analysis of “Cell_CCCM_Scaled_Filtered” dataset. PCA plots with samples color coded by source cell line (B1), exRNA isolation method (B2), biofluid type (B3), and Lab number (B4). B5. Heatmap showing hierarchical clustering results (biclustering of miRNAs and samples was performed). The sample type (cell or cell Supernatant), cell line, and exRNA isolation method for each sample are color coded below the heatmap using the same color schemes as in Panels B1-B4. C. Analysis of “Urine_Scaled_Filtered” dataset. PCA plots with samples color coded by BioGroup (Female and Male, (C1) and exRNA isolation method (C2). C3. Heatmap showing hierarchical clustering results (biclustering of miRNAs and samples was performed). The BioGroup and exRNA isolation method for each sample are color coded below the heatmap using the same color schemes as in Panels C1-C2.

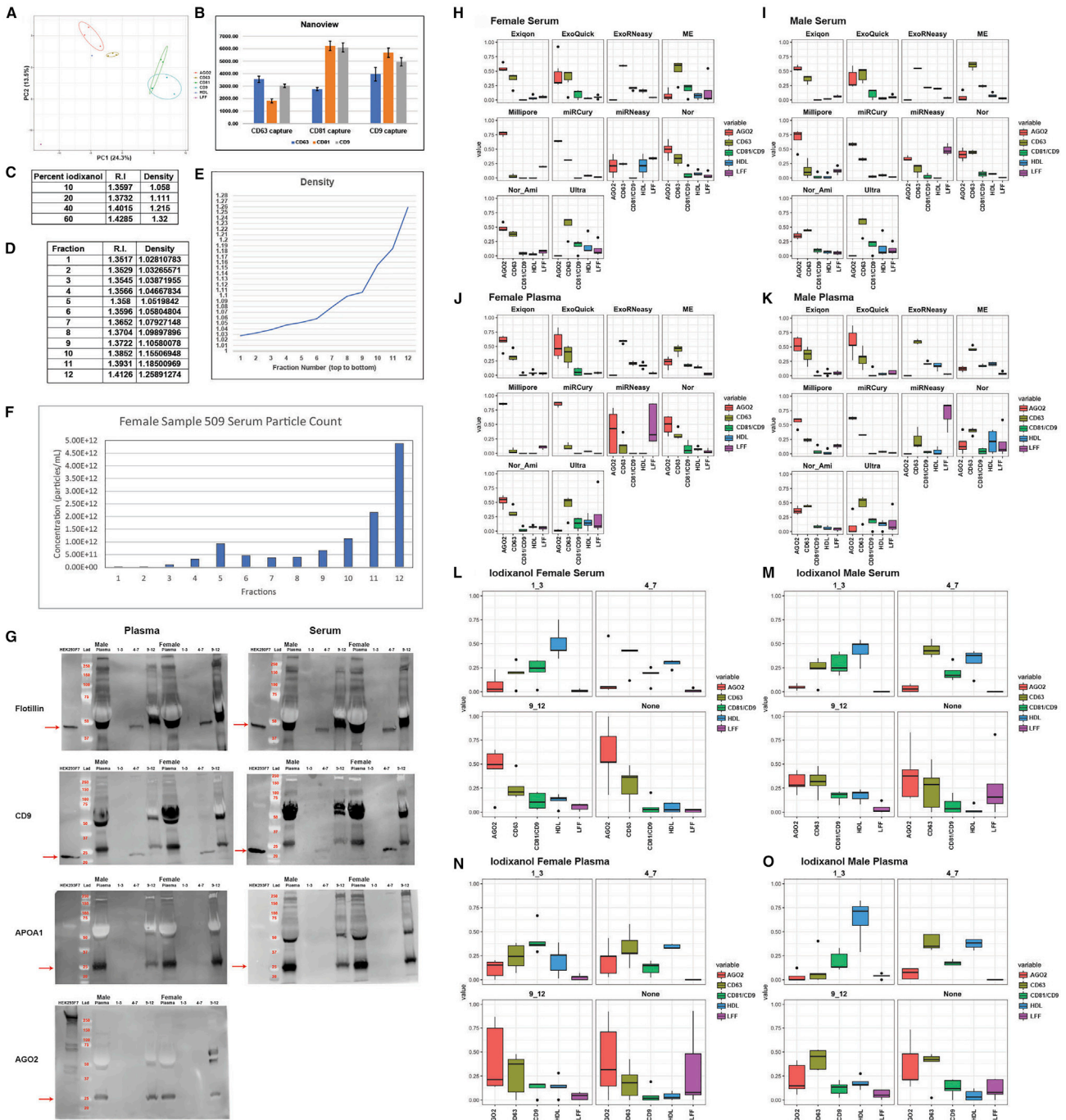


Figure S6. Characteristics of samples used in deconvolution analysis, Related to Figure 5

A. PCA plot of miRNA profiles of AGO2, CD63, CD81, CD9 immunoprecipitations and HDL and LFF fractions. B. Bar graph of Nanoview experiment on EVs from pooled human plasma, captured using antibodies to the targets indicated on the x axis, and then immunostained using fluorescently labeled antibodies raised against the targets indicated in the legend. C-E. Measured densities of iodixanol density gradient fractions. Based on a standard curve derived from refractive index (RI) measurements on samples with known concentrations of iodixanol (C), the density of each fraction of an iodixanol gradient was calculated (D, E). F. Particle concentration across fractions from iodixanol density gradient ultracentrifugation. G. Western blots using antibodies raised against flotillin, CD9, APOA1, and AGO2 for fractions 1-3, 4-7, and 9-12 from iodixanol density gradient ultracentrifugation. Data are shown for Female plasma and serum samples for subject 510. H-O. Proportions of CD63+, CD81+, CD9+, and AGO2+ extracellular miRNA carriers plasma and serum exRNA samples isolated using a variety of approaches. ExRNA samples isolated from Female serum Pool using the indicated exRNA isolation methods (H). ExRNA samples isolated from Male serum Pool using the indicated exRNA isolation methods (I). ExRNA samples isolated from Female plasma Pool using the indicated exRNA isolation methods (J). ExRNA samples isolated from Male plasma Pool using the indicated exRNA isolation methods (K).

(legend continued on next page)

indicated exRNA isolation methods (J). ExRNA samples isolated from Male plasma Pool using the indicated exRNA isolation methods (K). ExRNA samples isolated from Female serum Pool before and after fractionation on an iodixanol/Optiprep gradient (L). ExRNA samples isolated from Male serum Pool before and after fractionation on an iodixanol/Optiprep gradient (M). ExRNA samples isolated from Female plasma Pool before and after fractionation on an iodixanol/Optiprep gradient (N). ExRNA samples isolated from Male plasma Pool before and after fractionation on an iodixanol/Optiprep gradient (O). On the boxplots, the box indicates the median and interquartile range (IQR), the whiskers indicate $Q1-1.5*IQR$ and $Q3+1.5*IQR$, and the individual dots indicate outliers.

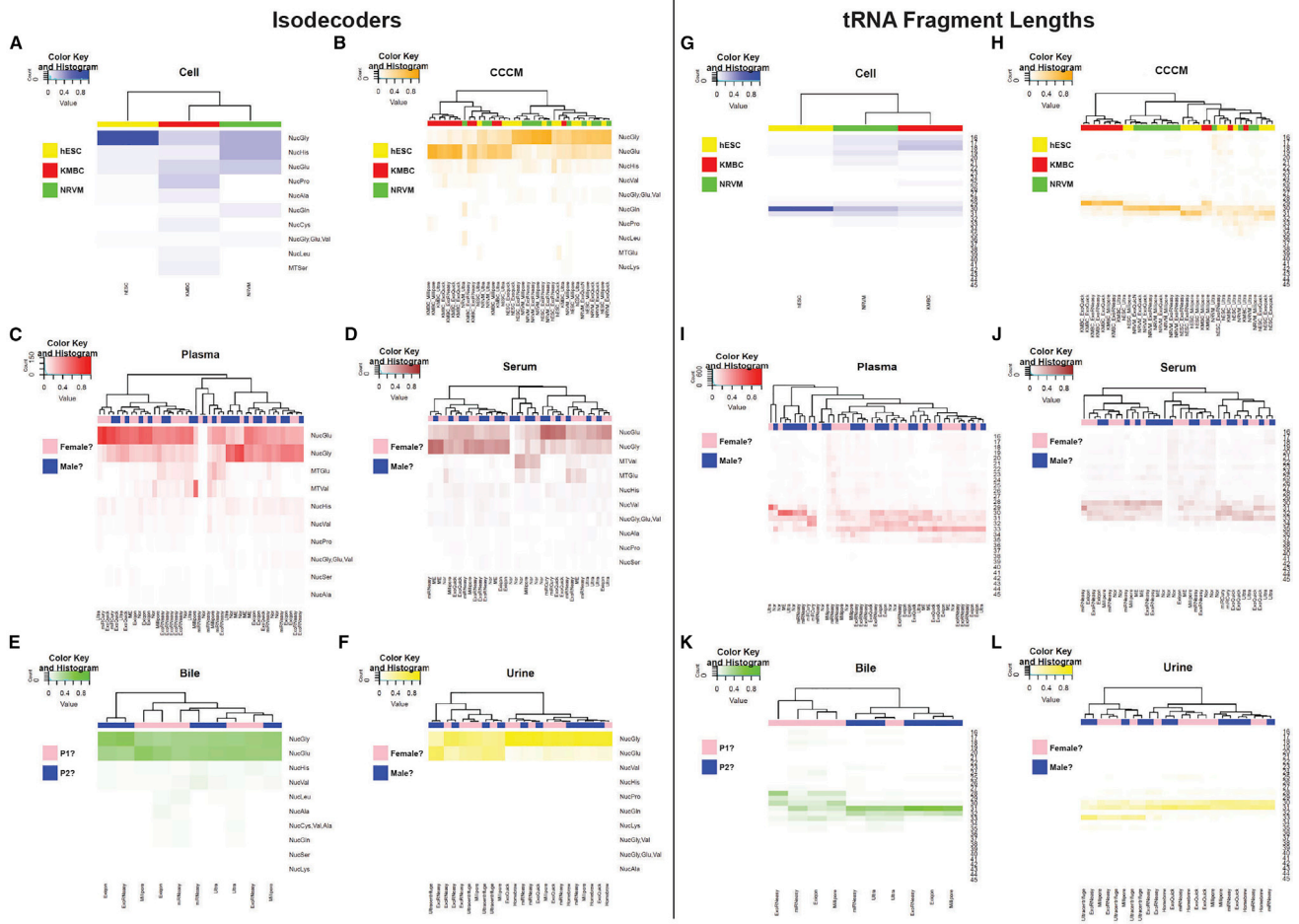


Figure S7. Detailed tRNA Analysis, Related to Figure 6

A-F. Heatmaps representing tRNA amino acid distributions for: cells (A); CCCM (B); plasma (C); serum (D); bile (E); urine (F). G-L. Heatmaps representing tRNA fragment length distributions for: cells (G); CCCM (H); plasma (I); serum (J); bile (K); urine (L).