Extracellular vesicle and particles small RNA sequencing Analytical Validation
Icahn School of Medicine at Mount Sinai (ISMMS) CIMAC
Date: November 22, 2022

Version 1.0

Performance Lab:

Edgar Gonzalez-Kozlova, PhD, Assistant Professor
Navneet Dogra, PhD, Assistant Professor
Seunghee Kim-Schultz, PhD, Associate Director
Sacha Gnjatic, PhD, Professor
1470 Madison Avenue, New York, NY 10029

This assay uses liquid biopsy inputs (plasma/serum (1-5mL) or urine (10-25mL)) to isolate extracellular vesicles and particles (EVPs) for RNA sequencing using size exclusion chromatography (SEC). The minimum RNA requirement post isolation is 1-0.25ng of nucleic acid material. We disclose that this technology has being approved and currently in use for R01 (R01AI172899) and R21 (R21AG078848) due it's extraordinary high reproducibility (R^2~0.8/0.9 across batch/assay and technology).

The table below lists the nature of initial results produced (primary output), quality control (QC) steps and analyses, and eventual output.

Assay type	Primary assay outputs	Pre-processing/ Normalization /QC	Initial analyses	Derived data outputs
EVP small RNAseq	~20,000+ gene mappings or 700,000+ transcript mappings based on 50 million 100pb reads.	EVP isolation using SEC and QC by Nanochip and Biochip. RNA isolation QC.	EVP-RNAseq in-house and Excerpt pipelines. Differential expression analysis using regression modeling strategies.	Coding and non-coding gene expression by gene biotype or read cluster/contig.

Extracellular Vesicles RNA Characterization	
(o) Ultra-low assay inputs	Minimal input of 1-0.25ng of RNA from EVPs. EVs are purified from a liquid biopsy of either (a) 1-3mL of plasma/serum or (b) 10-25mL urine. (Fig. 1)
(i) Accuracy	The EVP-RNAseq offers bulk-cell level detection of ~20,000 genes and ~700,000 transcripts with reproducibility values above 90% for technical and biofluids and 70% cross EVP isolation technologies, without background noise due usage of positive and negative sequencing controls. (See Fig. 3-6).
(ii) Precision	The precision of intra-assay/cross-technology stands with R^2:0.77/Rho: 0.9. The precision of intra-batch and technical replicates stands with R^2:0.9/Rho:0.9. (Figs. 3-6)
(iii) Analytical sensitivity	We have experimentally confirmed sensitivity and linearity of measurements in direct comparison with bioanalyzer and sequencing quality control using in-house and published pipelines (R^2:0.9). Gene counts detection of rare transcripts with confidence up to 5 counts per transcript for libraries of 40 million reads. (Figs. 3-6)
(iv) Analytical specificity including interfering substances	The 3 quality controls steps in our SOPs help us have high quality, low to zero impurities in the RNA before sequencing. (Figs. 3-6). Canonical EVP markers (CD9, CD63, and CD81) are used to verify isolation integrity.

(vii) Standardization, harmonization, reproducibility, and ruggedness.	We have developed numerous QC steps to improve standardization, including extensive review of the distribution, size, and density profiles of our isolated EVPs. These can be safely shipped to collection sites, and to facilitate and standardize RNA extraction procedures. These can be applied to whole blood, tissues for centralized analysis. All batch samples have positive and negative controls to maximize reproducibility and avoid environmental RNA contaminations. Positive controls use include UHRR (universal human reference RNA and miR163s). Negative controls involve a blank sample for environmental RNA. (When RNA is amplified in the negative control samples, then we subtract matching reads from all other samples).
(viii) Establishment of appropriate quality control and improvement procedures	Positive and negative control samples with expected RNA distributions and markers are used to validate novel RNA transcripts, in addition to the multiple other controls included below. Positive controls use include UHRR (universal human reference RNA and miR163s). Negative controls involve a blank sample for environmental RNA. (When RNA is amplified in the negative control samples, then we subtract matching reads from all other samples). Computational adjustment methods (generalized mixed effect linear models) are used to model potential technical artifacts and relevant sample information. (Figs. 3-6)

1. Purpose of assay

Minimally invasive monitoring through liquid biopsies is a critical strategy to diagnose and decide treatment directions in heterogeneous human populations. Here, we provide a robust and reproducible platform for surveying the RNA contents of extracellular vesicles and particles (EVPs) of 50-250nm size. Currently literature shows that these EVPs are enriched in cancer signatures of potential prognostic and diagnostic value. This is a broadly applicable approach that provides valuable information about the cellular and tumor secretome associated with cancer mechanisms and therapeutic responses. In the context of clinical trials, the aim is to potentially identify correlative extracellular vesicle and particles (EVPs) RNA markers associated with response and disease. EVP-RNAs have been associated with influencing changes in immune infiltrating and tumor cell phenotypes, including angiogenesis, metastasis, immune response, and drug resistance[1,2]. Some of the key challenges for this technology is to increase the reproducibility and capture rate of EVPs and their separation from other extracellular components, allowing their recovery from less abundant fluid samples. The ISMMS-CIMAC has established a robust EVP-small RNAseq platform addressing these challenges using cutting-edge equipment, protocols, and computational approaches.

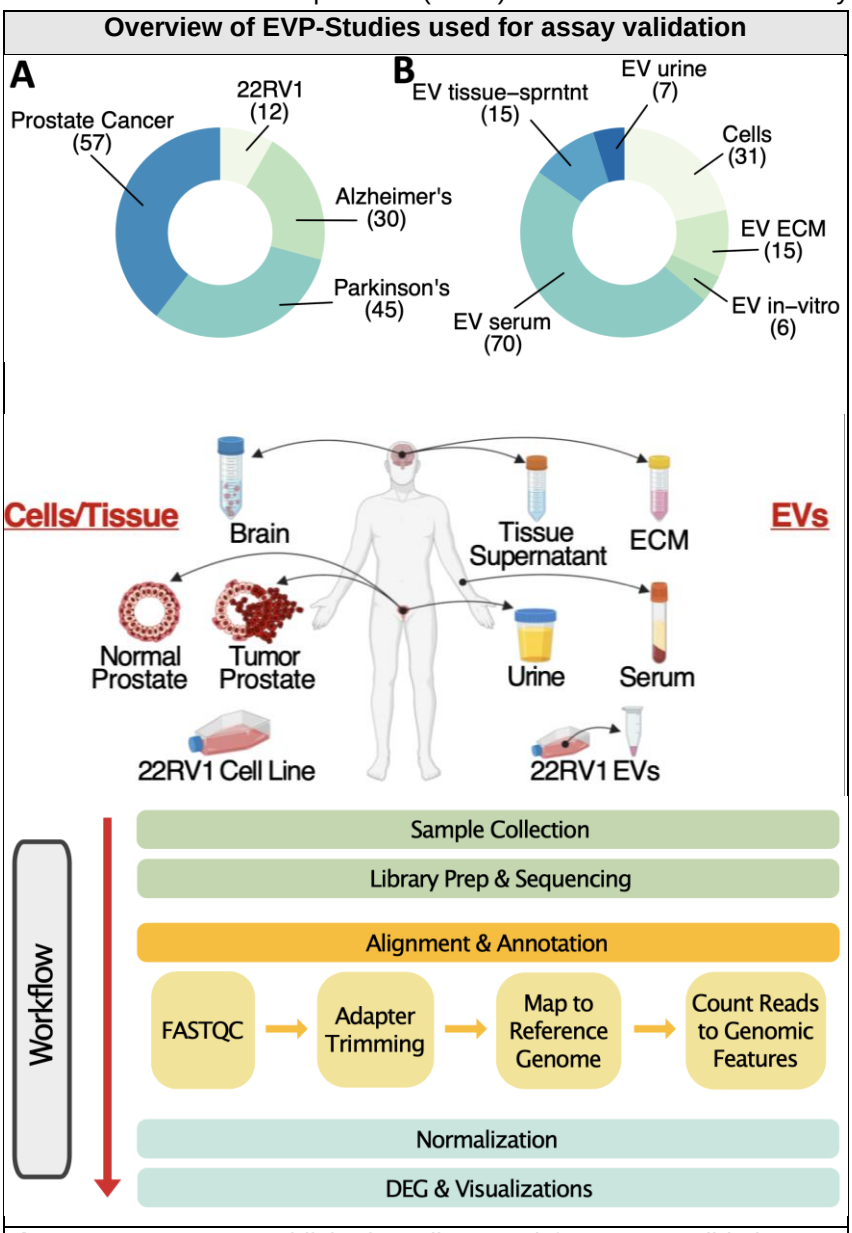


Figure 1. EVP-EVP-Published studies used for assay validation. **(A)** Sample's cellular origin of patient derived data and cell lines used for assay validation. **(B).** Tissue type derived samples divided into EVs or cells. **(C).** Overview of patient derived sample types used for validation. Computational workflow of downstream data processing.

EVP-small RNAseq[3] uses a combination of ultra-centrifugation and size exclusion chromatography to separate EVPs from other undesired components of fluids such as free DNA/RNA and cellular debris. Commercially available small RNA kits such as SMARTer or RNAeasy allow the quick recovery of all small RNA molecules and expand the read length to at least 100bps from fragments as small as 15bps, allowing their sequencing with standard RNAseq.

Removal of ribosomal RNAs is also possible as an optional step when rRNAs are not of interest. This strategy maintains low costs per sample and minimize experimental variability. Good quality samples can be sequenced at resolutions of 50 million reads. The usage of sequencing positive and negatives controls allows to perform strict quality control and filter out sequencing errors and environmental RNA noise associated with sample and library prep. Positive controls use include UHRR (universal human reference RNA and miR163s). Negative controls involve a blank sample for environmental RNA. (When RNA is amplified in the negative control samples, then we subtract matching reads from all other samples). Computational adjustment methods (generalized mixed effect linear models) are used to model potential technical artifacts and relevant sample information. EVP-

The reproducible and comprehensive platform for isolating EVPs, EVP-RNAs and computational tools for analysis have been established through an industry partnership between the ISMMS and IBM completed in 2020. Currently, the EVP pipeline/work frame is led by Drs. Dogra and Gonzalez-Kozlova. It includes optimized sample processing SOPs, reagents for library prep, isolation columns and antibodies for validation of EVP surface markers, as well as a supercomputer-based processing and analysis pipeline. The sequencing platform used by this pipeline is part of the Genetics and Genomics

department core. Additional sample processing facilities include the Human Immune Monitoring Center (HIMC) led by Drs. Gnjatic and Kim-Schulze. Thus, this work frame allows for in-depth cutting-edge correlative studies of the secretome composed by EVPs and their RNA cargo from tissues (tumor or adjacent normal) and biofluids (urine and blood) (Figure 1).

2. Assays performance characteristics based on Mt Sinai CIMAC's internal validation

2A. General materials and methods for assay characterization and validation.

Assay validation was performed using: (A) multiple cell lines commercially available (22RV1, 22RV1RD, HCC, MCF, BT) as well as (B) bloodserum and plasma collected from healthy, liver and prostate cancer patients/donors at Mt. Sinai and processed as described[4]. The isolation of EVs can be done from either, plasma, serum, or urine using standard kits and columns for size exclusion chromatography. Thus, the panel of reagents and datasets used for validation can be found in Table 1.

Table 1

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
CD81	Abcam	ab239687
Calnexin	Abcam	ab22595
Gold conjugated secondary antibody	Abcam	ab105285
Flotillin (anti-Flot-1)	Abcam	ab133497
Biological Samples		
Early liver cancer patient serum samples	ISMMS: Department of Pathology	IRB completed
Prostate cancer patient serum samples	ISMMS: Department of Urology	IRB completed
Human live and preserved Alzheimer's brains	ISMMS: Department of Neuroscience	Ongoing study
Human prostate cancer cell line 22RV1	American Type Culture Collection	ATCC CRL-2505
Chemicals, Peptides, and Recombinant Proteins		
RPMI 1640 cell culture medium	GIBCO	
Bovine serum albumin in phosphate buffer saline	Sigma Aldrich	
3% Glutaraldehyde	Sigma Aldrich	
Osmium tetraoxide	Sigma Aldrich	
Acid-phenol	Invitrogen	Cat# 4478545
Chloroform	Invitrogen	Cat# 4478545
Lysis Buffer (2% SDS/1X protease inhibitor/0.1M Ambic)	Invitrogen	Cat# 4478545
Critical Commercial Assays		
Total EV RNA and Protein Isolation Kit	Invitrogen	Cat# 4478545
Deposited Data		
Raw and normalized 22RV1 gene and protein counts	Dogra Lab.	GSE123736
Human reference genome Ensembl GRCh38.p13	Ensembl	https://useast.ensembl.org/Homo_sapiens/Info/Annotation
Human breast cancer cell lines EV and cell proteomic counts	Hurwitz et al., 2016	PMID: 27894104
Human colon cancer cell lines EV and cell gene counts	Hinger et al., 2018	PMID: 30332650
Human normal and cancer-associated fibroblasts EV and cell gene counts	Herrera et al., 2018	PMID: 30075793

Human reticulocyte derived EV protein counts	Diaz-Varela et al., 2018	PMID: 30232403
Human glioblastoma cell derived EV gene counts	Wei et al., 2017	PMID: 29074968
Software and Algorithms		
QIAGEN Ingenuity Pathway Analysis	QIAGEN	https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-ipa/?cmpid=QDI_GA_IPA&gclid=CjwKCAjwgr6TBhAGEiwA3aVuIS7UzgcUf8gTC8ad6uXMnBwZEeEb_lhG88SStmpLMASvAUl0tbI9tBoCJeAQAvD_BwE
FeatureCounts read summarization function	Subread	http://subread.sourceforge.net/
bowtie aligner (version 2.5.4b)	bowtie	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
R version 4.1.1	R Core Team	https://www.R-project.org/

2B. Reagents controls and calibrators

We have developed and optimized a purification EV consisting of a combination of differential-centrifugation and size exclusion chromatography. This procedure is a significant improvement of the ultra-low input protocol published by Dr. Dogra's group in 2019. This is the backbone of all our EV studies, which offers the advantage of streamlining sample processing, minimizing reagent batch effects, reducing potential errors of reagent addition, and facilitating the implementation of computational tools for accurate automated data QC and analysis, and high dimensional meta-analyses across various internal and published EVP-RNA data sets. Fig. 2. summarizes the necessary steps towards isolating EVPs and other small components such as microvesicles and free nucleic acids or proteins. First, the tissue, culture, or biofluid is dissociated from cells or other large components. Then, differential centrifugation is applied to concentrate all small particles <1um. Next, size exclusion chromatography is used to separate extracellular vesicles from other small components. The resulting fractions can be pooled or maintained separately at -80C for extended periods of time due the remarkable resistance of EVPs to heat and mechanical damage[1,8].

To overcome the greatest challenge of EVP-RNAs (low input samples) we have developed a comprehensive and critical quality control measurements using Pico and small RNA kits for bioanalyzer (Table 2). The final minimal input for sequencing stands between 1-0.25ng of RNA which can be obtained from a liquid biopsy 1-3mL of plasma/serum, 10-25mL urine or from solid tissue digestion recovered from supernatants (rescue from other experiments). If samples pass these QC metrics and contain sufficient material (1-0.25 ng RNA), they proceed to the genetics and genomics core sequencing facilities. Further, transmission electron microscopy has been used regularly by our group to additional quality controlling the particle sizes and abundance in various biofluids (blood and urine) (Fig.3 A-D). Our published results emphasize that the distribution of EVs ranges from 50 to 250nm, peaking at ~100 (Fig.3. E)[2,3,4,5]. Finally, another control commonly used in EV studies is the presence of the transmembrane protein CD9 (canonical tetraspanin) in the surface of EVs (Fig. 3F) and the absence of MIgG (common false positive signal). This type of quantitative analysis allows to estimate the abundance of EV particles, shape, and distribution across various tissues and biofluids[6,7].

Extracellular Vesicle Isolation by size exclusion chromatography

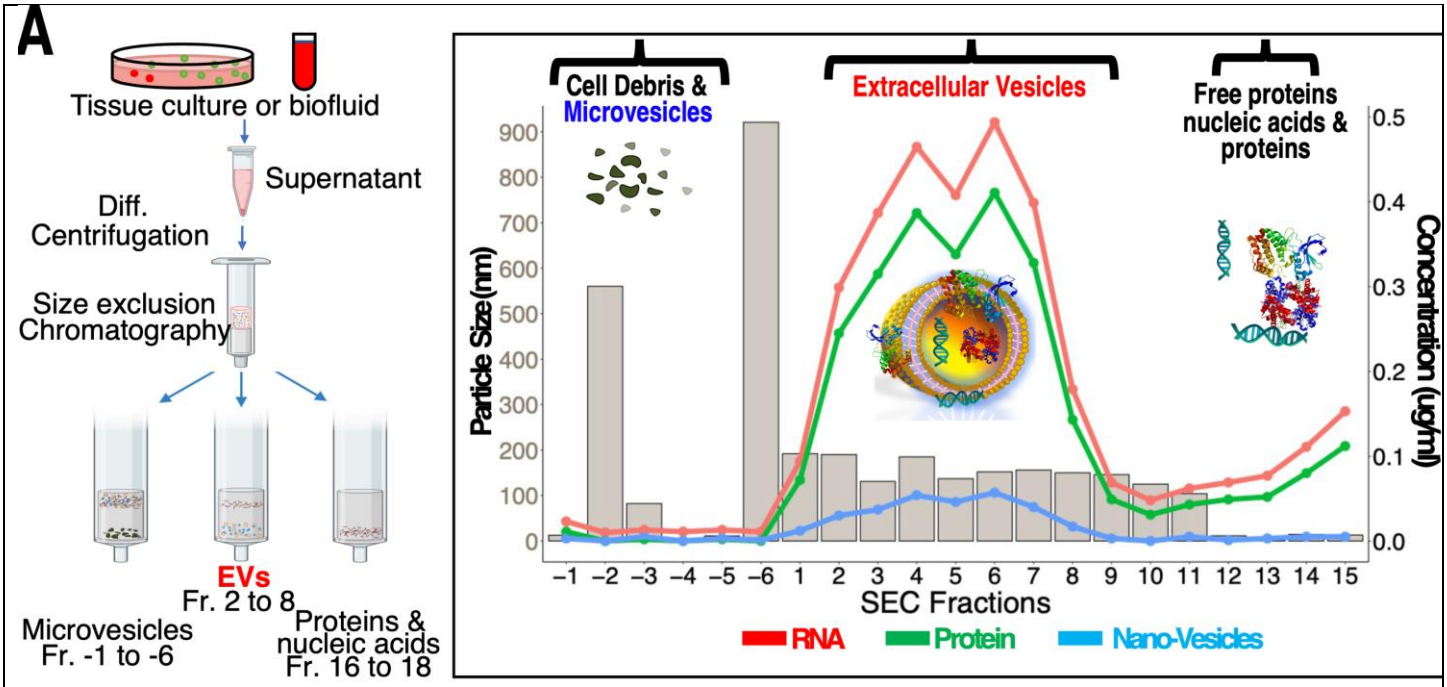


Figure 2. EV isolation schematics and fraction characterization. Left. Illustration summarizes critical steps for EV isolation. Right. Combined Bar-Line plot showing SEC fractions in the x-axis, y-axis left particle size (nm) (grey bar plots) and, y-axis right concentration (ug/mL) for RNAs, Proteins and Nano-Vesicles (<250nm).

EVP-

Extracellular Vesicle Transmission Electron Microscopy Validation in prostate cancer patient samples

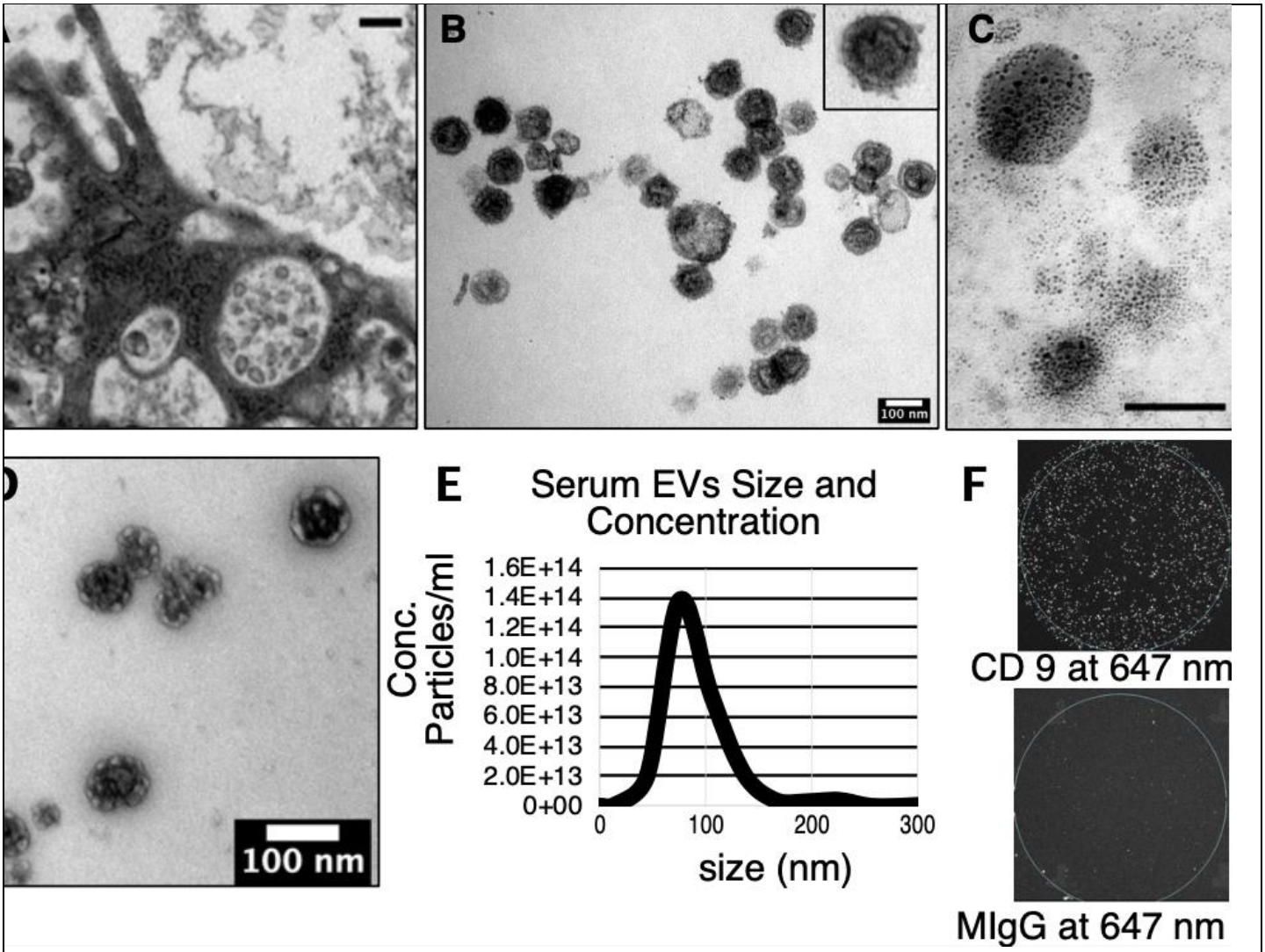


Figure 3. EV's size and concentration. A-D. TEM image of UC isolated EVs (scale bar 100nm). **E.** A size and concentration analyses of serum EVs show distribution of vesicles (50-150 nm). **F.** immuno-fluorescence co-localization analyses of serum EVs confirmed the presence of tetraspanin CD9 on vesicle surface. Optical image of immuno-fluorescence co-localization analyses shows EVs specificity to CD9 but not to mouse IgG.

Table 2. Sample processing quality control

Sample Processing	Data acquisition	Data processing and analysis
Strict SOPs and detailed sample processing logs	Commercially available reagents and columns for SEC.	Automated and manual quality control of sample, isolation and alignment.
EVP-Isolation QC	TEM & Western blot	Semi-Manual analysis of visual and quantitative inspection of isolated EVs.
Optimized Pico & nano kits	Bioanalyzer	Semi-automated analysis of RNA concentrations.
Sequencing QC	FastQC and alignment logs (Bowtie & STAR)	Automated analysis pipeline includes metrics for alignment and multimapping

2C. Validation of detection accuracy relative to gene biotype composition

For validation of small RNAs EVPs and cellular accuracy of composition relative to genomic annotation of transcripts based on their origin, a series of 6 cell lines were processed in parallel our computational pipeline. Eleven major cellular compartments of origin sections were identified using KEGG and GO databases (Fig. 4 A&B). As described by our[2], EV small RNA composition is different from cellular small RNAs (Fig. 4 B). A linear regression using restricted cubic splines shows a correlation of 80% and 77% for the composition of small RNAs from EVs and cells, respectively (Fig. 4 C & D).

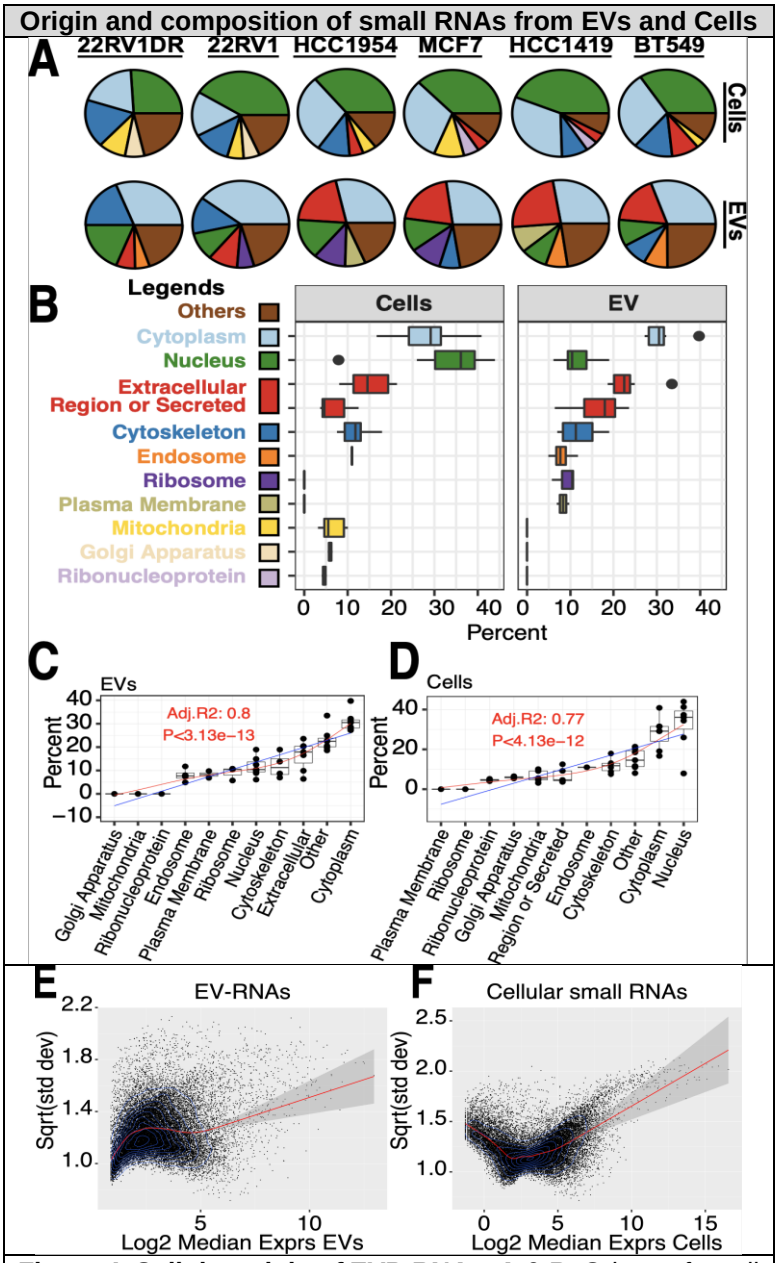


Figure 4. Cellular origin of EVP-RNAs. A & B. Origen of small RNAs showed in pie and boxplots plots separated by either cellular or isolated EV samples **C & D.** Linear regression using cubic splines (x~rcs(y,5)) with 5 splines, for RNA origin source separated by either cellular or isolated EV samples. X-axis indicates the RNA-origin and y-axis the percent of transcripts. **E & F.** Dispersion plot showing in x-axis the log2 median average expression for EVs and Cells, respectively.

This validation shows that heterogeneity in the source of origins of small RNAs represents about 20% of the total variation detected. Further, we show that the reproducibility between cell lines of small RNAs origin using linear models is around 80%. It's important to mention that these cell lines correspond to different sequencing sites and EVP-isolation protocols. Our published results show an agreement of >90% between human samples for RNA origin of transcripts in liver cancer[3]. Interestingly, these results highlight the diversity of small RNA molecules packed in EVPs and their association to cellular small RNAs.

2D. Validation of assay sensitivity

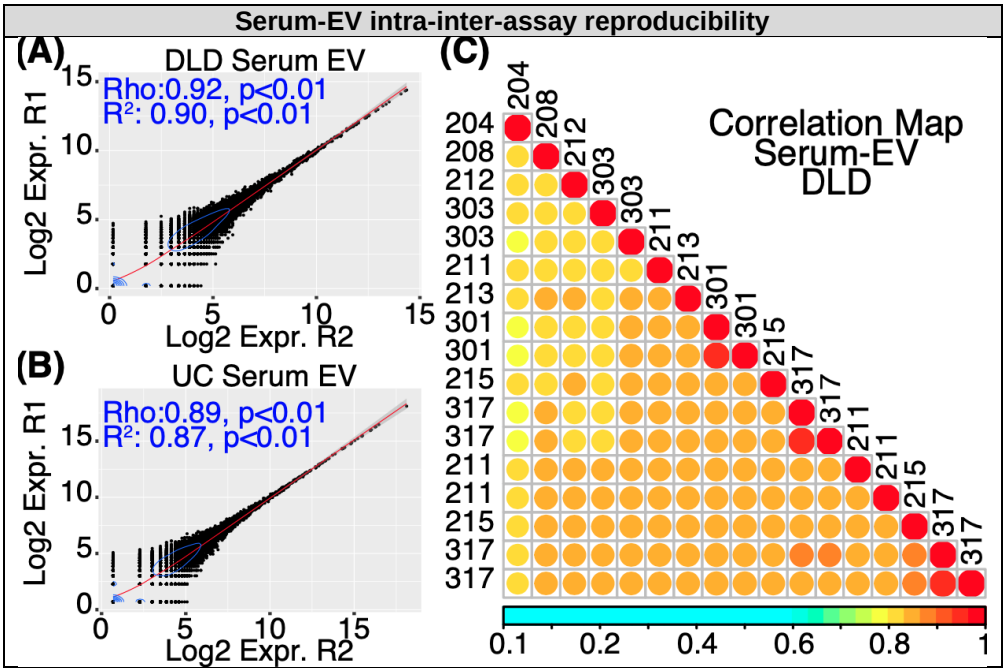
For validation of sensitivity, the samples were analyzed by combined bioanalyzer and unsupervised contig clustering. These results show that EVPs are enriched for small RNAs primarily < 100nt while cells have a different distribution of RNAs with a median transcript length of 700nt. These published results[4,5] also show that small EVP-RNA sequencing can detect approximately ~20,000 genes with more than 5 counts per gene and ~500,000 unique transcripts including intronic and non-coding regions classically referred as 'black matter'. The differences in RNA expression are further emphasized in the distribution of normalized counts (Fig.4 E & F). Cellular RNAs have a square root variance of 1.25 with smaller and higher variance of genes outside the IQR area, with either very low or high expression. However, EVP-RNAs were shown to have a more consistent expression of lowly expressed genes. This phenomenon is partially attributed to more duplicated sequencing in EVP-RNAs compared to cellular small RNAs[3,4].

2E. Validation of intra- and inter-assay reproducibility

To evaluate overall inter- and intra-assay variability in reagents, sample processing and instrument performance[3,4,5], we used our liver and prostate cancer cohorts (n=20) and three technologies SEC, UC and nanoDLD[7] with 3-5 technical and inter assay replicates processed in 3 different days and corresponding to 3 tissues (serum, urine, and biopsy) through a Mount Sinai – IBM collaboration. Variations in the number of detected genes and transcripts, genes enriched in either EVs or Cells and association with disease were

identified. These results allowed us to refine the EVP-RNA detection and quantification pipeline at Mount Sinai, allowing to quantify expression at 3 different levels: (1) Gene expression, (2) Transcript and, (3) Read cluster. We also show the agreement in recovered EVP-RNAs between different technical and biological variables cross assay or technology is very high for ultra-low input samples (1-0.25ng RNA) ~0.75 Spearman/Pearson correlation (Fig. 5C). Technical replicates across different isolation batches, library preparations and biofluids show a correlation of RNA contents >90% agreement (Fig. 5 A & B), while biological variability remains >60% agreement (Fig. 5C). Also, these results highlight that patient/disease specific EVP-RNA signatures represent about 30% of the total EVP-RNAs extracted, making it an outstanding tool for discovery of biomarkers.

2F. Optimization of sample storage conditions to ensure data consistency of stored samples



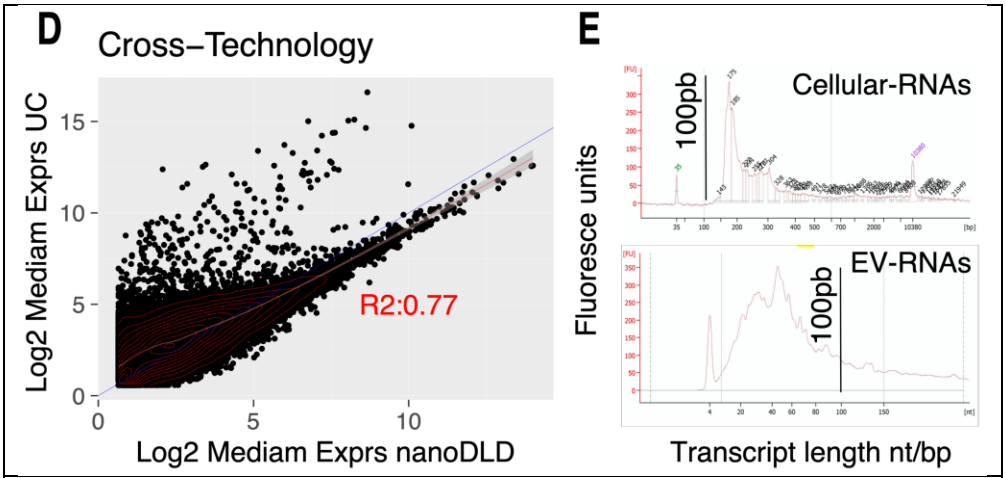


Figure 5. Correlation between biological and technical replicates. We performed EV sequencing using two different assays UC and DLD on serum samples (n=5). **A & B.** Shows the correlation between same technical replicates (isolation of EVs) for DLD and UC. **C.** Correlation matrix between technical (library preps/technology) and biological replicates for DLD and UC, showing a Pearson/Spearman correlation of min >0.7 and max >0.9.

Samples processed for EVP analysis are incredibly flexible, these can be recovered from any fresh or frozen samples (-80), live tissues, biofluids, and supernatants from other assays such as single cell RNAseq. However, the tissue with lowest recovery rate is FFPE, requiring larger volumes of samples.

This has the advantage of mitigating the need for complicated acquisition of samples and the lack of instrument availability. We have performed a range of studies to optimize short term and long-term storage conditions for EVs and currently we have the capabilities of storing them up to 1 week at 4°C and >2 years at -80°C with minimal degradation of content quality[4], anecdotally EV samples older than 50 years have been reported to be successfully recovered.

2G. Validation of inter-instrument (technology) reproducibility

We have confirmed the reproducibility of data generated using two independent instruments at two different institutions (MSSM and IBM). EVP-RNAs were isolated by UC or nanoDLD, both instruments using patient isolated samples. After isolation of particles, we used our processing and QC pipeline to extract RNA from these samples and sequenced the sample independently. Post data normalization we correlated the expression of EVP-RNAs and obtain a coefficient of 0.77, highlighting high reproducibility and best practices of our methods even cross institutions. This also emphasize that these vesicles have highly consistent RNA content based on the biofluid and tissue used for extraction which suggests a highly controlled packing mechanism[1,3,5,8]. Overall, throughout validation and our published research, we have stablished a precise, efficient, and accurate for EVP studies of potential diagnostic and therapeutic use.

Finally, to show the reproducibility between biological samples, we to show the correlation between UC and SEC in Plasma-EVPs from Alzheimer's patients and brain derived ECM-EVPs tissue (Fig. 6). In the preliminary data from a recently funded R21 study by NIA where Drs. Dogra (PI) and Gonzalez-Kozlova (CO-I) collaborate, we obtained serum and brain tissue for isolation of EVPs. Here, we compared the most common method UC/Density gradient with our SEC approach and show a correlation of 0.9 Spearman and 0.8 R^2 in serum (Fig.6B) and ECM (Fig.6C). These results, emphasize and validate that our approach is ideal for isolation of EVPs with extremely high reproducibility in the area of liquid biopsies. These results also show the differences of EVPs compared to cellular debris or cell free nucleic acids (Fig.6A). However, the proposed assay specifically targets EVPs due their relevance in cancer and immune-monitoring.

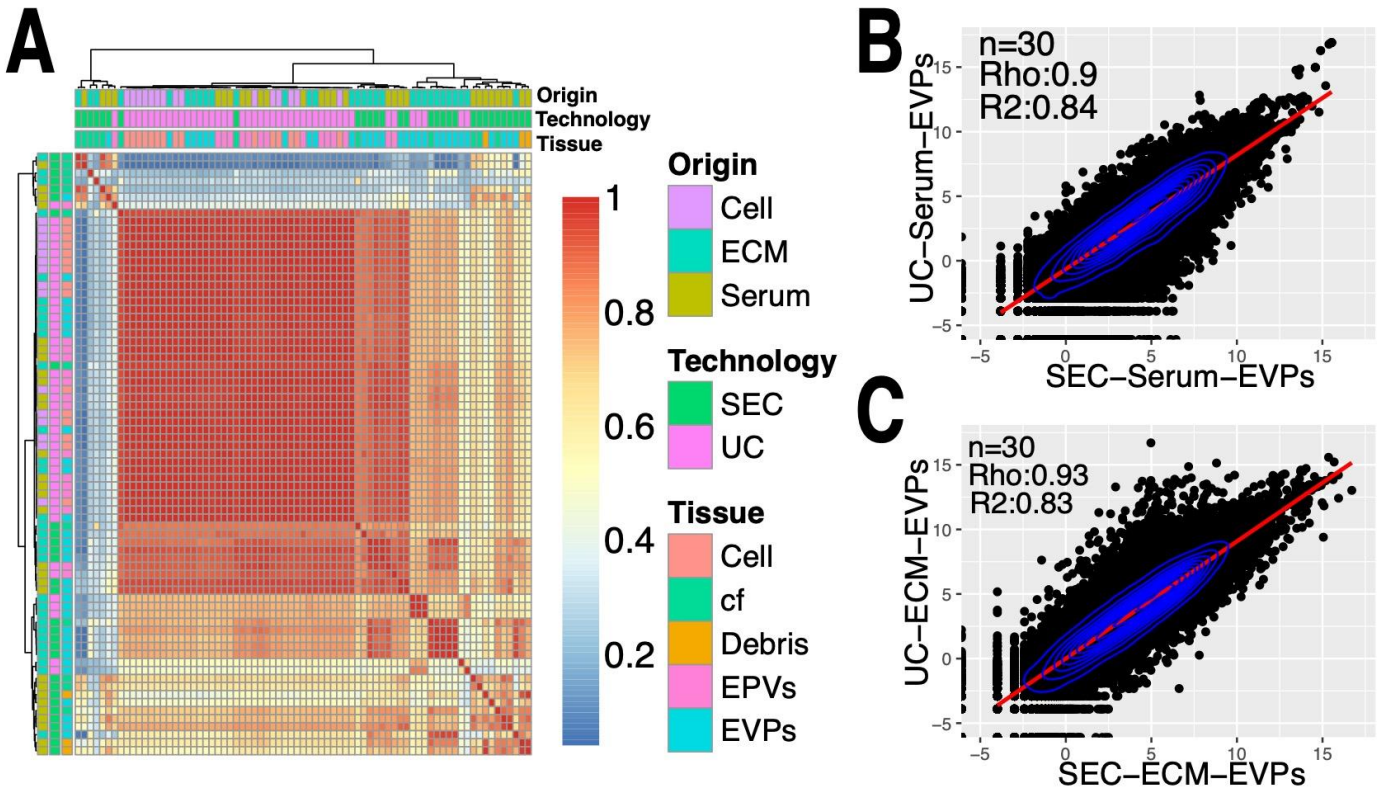


Fig. 6. A. Shows the Pearson correlation coefficient between individual Alzheimer samples for Serum and brain derived tissues (ECM-extracellular matrix and cells). The lowest correlation was between cell free nucleic acids and other tissues. **B.** Scatterplot showing the correlation values between genes for 30 EVP samples (15 UC and 15 SEC) from serum, correlation values shown in the figure. **C.** Correlation scatterplot for 30 EVP samples (15 UC and 15 SEC) from extracellular matrix derived EVPs. Correlation between EVPs from the most common UC/Density gradient method compared to SEC show a reproducibility of ~80-90%.

2H. Criteria and metrics for defining significant changes (e.g., between timepoints, between responders and non-responders).

The criteria for identifying significant molecular changes in a study are tied to the study experimental design. We strongly encourage use baseline and first treatment timepoint to identify signatures associated with outcomes or treatment arms.

We have published in peer reviewed high impact journals molecular diagnostic and prognostic signatures of EVP-RNAs. We describe below relevant study details:

In a pioneering study we integrated MRI and H&E histology slides with electron microscopy (EM) to describe EVs in prostate cancer (PCa). First, MRI confirmed the presence of tumor lesion in the prostate **(Fig. 7B)**. Second, post-biopsy H&E confirmed the distribution of epithelial carcinoma entirely absent in the normal tissue (**Fig. 7A&C**). Then, we captured secretion of vesicles into the lumen of prostate ductal epithelium from both tumor and normal tissue (**Fig.7D&E**). Prostate tumors shed 2.6X more EVs quantified as multivesicular bodies (MVBs) compared to normal cells (~400 MVBs measured in 3 different tissue sections) **(Fig.7E)**. *While* elevated levels of MVBs and EVs have been previously observed in multiple cancers[1,5,8] direct electron microscopic visualization and quantification of secretory vesicles had not been described. Thus, showing that **Prostate tumor tissues are enriched in EVP-RNAs.**

Pathology of tumor and EVs

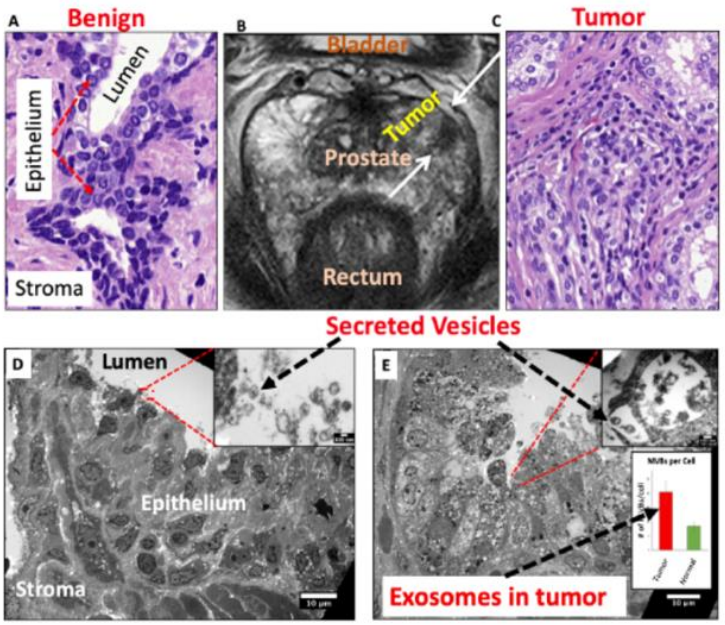


Figure 7. Microscopic analyses of a prostate cancer patient in his 70s (Gleason score 9) normal (A) and tumor (C) biopsy tissue. (B) MRI (T2WI) showing tumor lesion. Arrows indicate lesion location. Transmission EM analyses of normal (D) and tumor € tissue (Inset. vesicle secretion in the luminal area of prostate tissue, scale bar is 100nm).

Tumor specific EVP-RNAs can be identified across different biofluids and tissues (Fig. 8). We isolated tissue matched serum-EVs, confirmed presence of canonical markers (CD9, CD63, and CD81), physical EV properties and compared the differential expression of EVP-RNAs between urine and serum EVs (Fig. 8A). This approach reveals EVP-RNAs measured in parallel across biofluids and tissue. Further, we identified 186 differentially expressed EVP-RNAs between normal and tumor tissue with high average expression across patients[3,4,5] (Fig. 8B). Thus, showing that multiple tissue comparisons are reproducible across technologies (nanoDLD-IBM nanochip and Ultra-Centrifugation) and capable of revealing potential EVP-RNA biomarkers.

We applied a similar strategy[3] in HCC and identified non-coding EVP-RNAs as potential early predictors of disease in a discovery cohort (n=15), and identified small EVP-RNA (<100pb length) clusters (smRCs). These EVP-RNA markers were highly distinguishing between HCC and chronic liver disease (CLD) **(Fig.9A)**. Then, we trained regression models using these smRCs with our discovery cohort and tested their capacity as early diagnostic markers in an independent cohort (n=206) using RT-qPCR **(Fig.9B).** These were compared to serum glycoprotein AFP, a well-known standard in HCC identification. Our results indicate that EVP-RNAs are more precise markers than AFP and that using AFP+smRCs have even greater efficiency for early detection[3,4,5] **(Fig. 9B)**.

Overlapping prostate cancer signatures across tissues

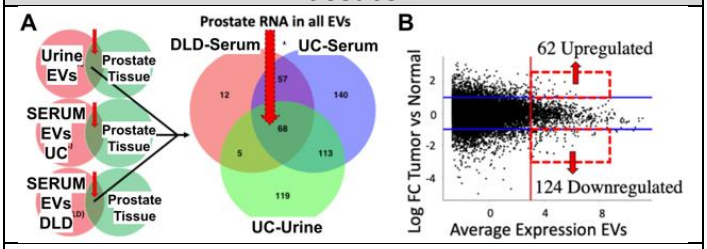


Figure 8. Overlap between prostate tissue RNA and serum/urine EVs RNA. Final analyses present a set of common 68 EVP-RNAs present in prostate tissue and serum/urine (C) A set of 186 discovered RNA (62 up- and 124 down-regulated genes in prostate tumor tissue vs adj. normal.

Finally, we observed upregulation of many known prostate-cancer associated genes in matched biopsies before and after prostatectomy. Importantly, these were upregulated in cancer patients only **(Fig.10A&B)**. These genes, including major tumor suppressors NKX3.1, BRCA1, MDX4, were highly expressed in the cancer patient exosomes along with genes involved in ROS, P53 pathways and cytokine production[1,4,5,8] (**Fig.10A&B**). The over expression of the tumor suppressor genes and inflammatory cytokines in EVP-RNAs can be explained by following mechanisms: **1)** The disposal/delivery mechanism from exosomes may be playing a key role in the dissemination of tumor-associated (oncogenes or tumor suppressors) RNA. **2**) The overexpression of cytokines in EVs may stem from local immune response. Remarkably, these results support the hypothesis that upregulated cancer circulating EVs play a key role in the dissemination of tumor-associated RNA and may be viable target for liquid biopsy biomarkers. In addition, these results and other studies suggest EVP-RNAs can characterize tumor and immune component transcriptomes[1,5,8,9,10].

Non-coding RNAs and small read clusters validation with RT-qPCR

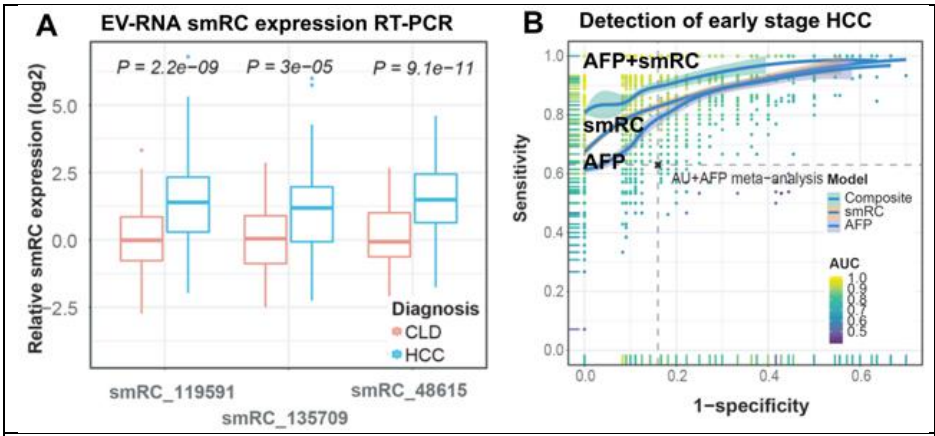


Figure 9. Hepatocellular carcinoma early detection using non-coding EVP-RNAs. A. Expression of 3 serum-EVP-RNA clusters named small RNA clusters (smRCs) **B.** An independent validation cohort with 209 patient samples comparing early detection of HCC using as predictors AFP, smRCs and AFP+smRCs, where smRCS or EVP-RNAs outperform AFP, while AFP+smRCs showed to the be best overall method of diagnostics.

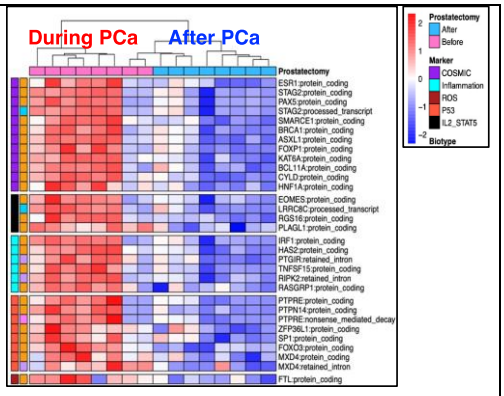


Figure 10. Detection of EVP-RNA in matching patients during and after prostate cancer (undetectable PSA ~3 months after prostatectomy). Y-Axis color coded notes show expression before and after prostatectomy (PCa)

2I. Inclusion of individual samples for analysis.

The proposed technology is ideal for comparing changes between baseline – timepoint and associations with outcomes. While, an accurate experimental design is ideal for obtaining a clear EVP-RNA signature associated with response or treatment, individual patient samples can be processed to obtain quantifications of the most abundant RNA species. Many small and non-coding RNAs including miRNAs, are well known to correlated with specific cancer phenotypes and descriptive databases are available (exAtlasRNA, miRbase). The expression of EVP-RNAs from an individual sample can be ranked based on their relative expression and top 5% expression can be evaluated for potential known biomarkers of outcomes.

3. References

1. Kalluri R, LeBleu VS. The biology, function, and biomedical applications of exosomes. Science. 2020 Feb 7;367(6478):eaau6977. doi: 10.1126/science.aau6977. PMID: 32029601; PMCID: PMC7717626.

2. Chen T, Gonzalez-Kozlova E, Soleymani T, La Salvia S, Kyprianou N, Sahoo S, Tewari A, Cordon-Cardo C, Stolovitzky G, Dogra N. Extracellular Vesicles Carry Distinct Proteo-Transcriptomic Signatures That are Different from Their Cancer Cell of Origin. IScience 2022. https://doi.org/10.1101/2021.09.20.460963

3. Von Felden J, ..., Gonzalez-Kozlova E, ... Tewari A, Stolovitzky G, Losic B, Villanueva A. Unannotated small RNA clusters associated with circulating extracellular vesicles detect early-stage liver cancer. Gut 2021, PMID: 34321221.

4. Dogra N, Ahsen ME, Gonzalez-Kozlova E, … Sahoo S, Tewari A, Cordon-Cardo C, Losic B, Stolovitzky GA. exRNA Signatures in Extracellular Vesicles and their Tumor-Lineage from Prostate Cancer. MedRxiv 2021. doi:10.1101/2020.09.28.20190

5. Gaglani S, Gonzalez-Kozlova E, Lundon DJ, Tewari AK, Dogra N, Kyprianou N. Exosomes as A Next-Generation Diagnostic and Therapeutic Tool in Prostate Cancer. Int J Mol Sci 2021. PMID: 34576294.

6. Charoenviriyakul C, Takahashi Y, Nishikawa M, Takakura Y. Preservation of exosomes at room temperature using lyophilization. Int J Pharm. 2018 Dec 20;553(1-2):1-7. doi: 10.1016/j.ijpharm.2018.10.032. Epub 2018 Oct 11. Erratum in: Int J Pharm. 2019 Mar 25;559:427-428. PMID: 30316791.

7. Smith JT, Wunsch BH, Dogra N, Ahsen ME, Lee K, Yadav KK, Weil R, Pereira MA, Patel JV, Duch EA, Papalia JM, Lofaro MF, Gupta M, Tewari AK, Cordon-Cardo C, Stolovitzky G, Gifford SM. Integrated nanoscale deterministic lateral

displacement arrays for separation of extracellular vesicles from clinically-relevant volumes of biological samples. Lab Chip. PMID: 30468237.

8. Gonzalez-Kozlova E. Computational methods for precision oncology - Chapter: Molecular Profiling of Liquid Biopsies. Springer. https://doi.org/10.1007/978-3-030-91836-1

9. Zhang P, Wu X, Gardashova G, Yang Y, Zhang Y, Xu L, Zeng Y. Molecular and functional extracellular vesicle analysis using nanopatterned microchips monitors tumor progression and metastasis. Sci Transl Med. 2020 Jun 10;12(547):eaaz2878. doi: 10.1126/scitranslmed.aaz2878. PMID: 32522804; PMCID: PMC8024111.

10. Poggio M, Hu T, Pai CC, Chu B, Belair CD, Chang A, Montabana E, Lang UE, Fu Q, Fong L, Blelloch R. Suppression of Exosomal PD-L1 Induces Systemic Anti-tumor Immunity and Memory. Cell. 2019 Apr 4;177(2):414-427.e13. doi: 10.1016/j.cell.2019.02.016. PMID: 30951669; PMCID: PMC6499401.