Proteomic profiling of tumor-derived extracellular vesicles enables early lung cancer detection

Kelley Van Vaerenberghe, Margie Kinnersley, Tre Blohm, Tia Seibold, Rachel Short-Miller, Sean Lodmell, Amanda Mast, Riley Kemp, Claire Seibold, and Katie Havranek

FYR, Missoula, MT USA

#7084

Introduction

- Lung Cancer is the leading cause of cancer-related death in the U.S., with deaths exceeding those of breast, colorectal, and prostate cancers combined[1].
- Non-small cell lung cancer (NSCLC), the most prevalent type, has a 5-year survival rate of just 25%[2].
  - NSCLC usually presents with no symptoms until later stages after the cancer has spread, making early detection rare.
- Early detection can boost NSCLC survival rates by 36%, yet the recommended screening method, low-dose CT, lacks specificity, provides no comprehensive prognostic information and is inaccessible to many patients[3,4].
- Liquid biopsies based on cell free DNA (cfDNA) have emerged as convenient, cost-effective alternatives to traditional screening but often lack sensitivity and specificity for early-stage cancers[5,6].
- Blood plasma extracellular vesicles (EVs), which carry heterogeneous protein, nucleic acid, and metabolite cargos derived from various cell types including cancer cells, provide a richer diagnostic analyte base than cfDNA.
- Here we present a novel tumor-derived EV (TDEV) enrichment technique called SPARCs™, applied to plasma from early-stage NSCLC patients.
- Biomarkers across stages are likely to change, and results indicate SPARCs™ can detect tumor progression from circulating EVs.

Objective

The objective of this study was to identify RNA and protein biomarkers in enriched TDEVs from cancer vs. healthy donors to advance early screening capabilities for NSCLC. Proteomic results are highlighted here.

Materials and Methods

Plasma was processed from whole blood collected in Streck Cell-Free BCT preservative (Streck, La Vista, Nebraska) and stored at -80ºC. For EV isolation, plasma was thawed, re-spun to clear debris and subjected to ion-exchange chromatography. Purified EVs were then characterized in concordance with the Minimal Information for Studies of Extracellular Vesicles (MISEV) 2023 guidelines with respect to particle concentration and size (Zetaview nanoparticle tracking system, ParticleMetrix, Ammersee, Germany), and presence/absence of category 1, 2 and 3 protein markers (Jess automated western blot system, Biotechne, Minneapolis, MN). EVs were incubated with Tumor SPARCs™ to enrich for tumor-derived EVs. Purified RNA was used to generate bulk RNAseq libraries and sequenced on an Element Biosciences AVITI system (San Diego, CA). SPARC-enriched EVs were subjected to digestion and subsequent LC-MS/MS on the Orbitrap Astral Instrument (Thermo Fisher Scientific, Waltham, MA) at Cedars Sinai Precision Biomarker Laboratories (Beverly Hills, CA) using data independent acquisition. 10-fold cross-validation was used for Machine Learning model training, and the best model was selected based on AUC. Candidate biomarkers were selected based on leading model AUC.

EV Characterization

- EV extractions yielded an average of 1.2x10[11] particles /mL plasma with a mean diameter of 155.3 nm (Fig. 1).
- There was no significant difference in concentration or size between NSCLC and healthy donor particles.

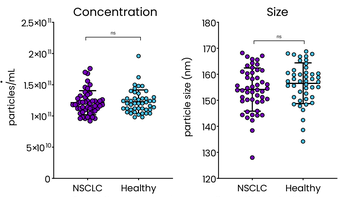

Figure. 1. Nanoparticle tracking analysis, Kruskal–Wallis test p<0.05.

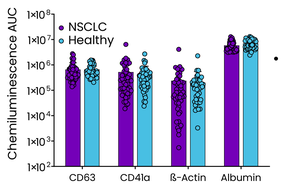

Figure. 2. Automated Western blot analysis.

- Particle preparations contained MISEV marker proteins CD63 (category 1a), CD41a (category 1b), β-Actin (category 2b), and Albumin (category 3a), confirming the presence of extracellular vesicles (Fig. 2).

Study Design

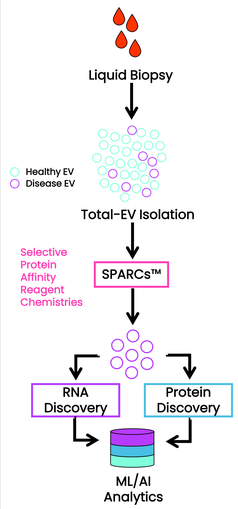

Table 1. Samples used in this study

	NSCLC	Healthy
N	49	45
Age		
<51	4	5
51-60	11	13
61-70	21	23
71-80	12	4
81-90	1	0
>90	0	0
Sex		
Male	28	26
Female	21	19
Race/Ethnicity		
Caucasian	49	37
Asian	0	2
Black	0	6
Staging		
Stage 1 (I, IA1, IA2, IA3, IB)	25	-
Stage 2 (IIA, IIB)	24	-
Subtype		
Adenocarcinoma (LUAD)	37	-
Squamous Cell Carcinoma (LUSC)	10	-
Adenosquamous	2	-
Specimen Providers	2	2
Smoking Status		
Unknown	-	8
Never used	17	23
Previous use	22	9
Current smoker	10	5

Lung Cancer EV Proteomic Signature

- LC-MS/MS proteomics identified a total of 4176 unique proteins across all samples.
- Differential expression analysis comparing NSCLC and healthy controls identified 454 differentially expressed proteins (DEPs), with 250 proteins significantly enriched and 204 proteins significantly depleted (Fig. 3).
- Top DEPs represent both novel markers as well as features known to be relevant in NSCLC.

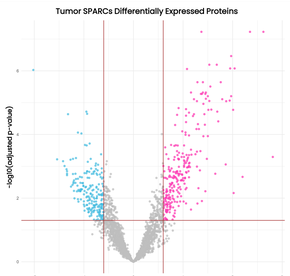

Figure 3. Volcano plot with differentially expressed proteins enriched by Tumor SPARCs. Cutoff thresholds at p_{adj} 0.05 and log_2 fold-change ±0.6.

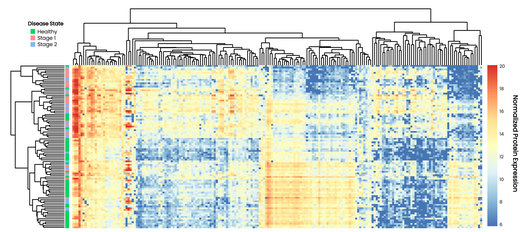

Figure 4. Heatmap hierarchical clustering of normalized protein expression for the top 200 DEPs from Tumor SPARCs enrichments comparing NSCLC vs. healthy patients. The top 200 DEPs were chosen by sorting by absolute log2 FoldChange. Each row represents a single individual and each column represents a single DEP.

Figure 5. Significant results of GSEA for Gene Ontology and KEGG utilizing the full protein list ranked by Log2FC between NSCLC and Healthy groups. GO categories used include Biological Process (BP), Cellular Component (CC), and Molecular Function (MF).

- Identification of enriched KEGG and GO terms between NSCLC and healthy groups (Fig. 5) identified increased translation and defense response while showing decreased metabolic processing and mitochondrial function, results consistent with disease progression.

Unique Signatures of Stage 1 and Stage 2

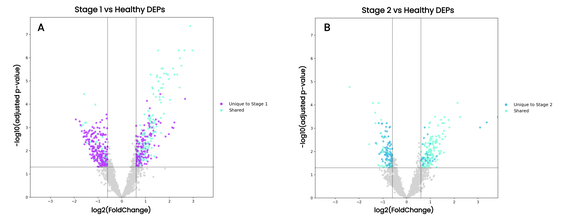

Figure 6. Stage specific comparisons of TDEV protein relative abundances. (A) Volcano plot of Stage I vs Healthy. (B) Volcano plot of Stage II vs Healthy. Cutoff thresholds at p_{adj}≤0.05 and a log_2 fold-change of ± 0.6.

- Differential expression analysis comparing Stage 1 NSCLC vs healthy controls identified a total of 559 DEPs (Fig. 6A).
- Differential expression analysis comparing Stage 2 NSCLC vs healthy controls identified a total of 319 DEPs (Fig. 6B).

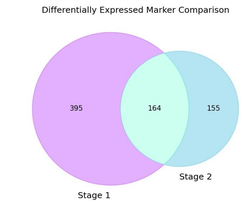

Fig. 7. Venn Diagram depiction of Stage 1 vs 2 DEP comparison overlap.

- Identification of shared vs unique DEPs across stages highlights more differences than similarities between stages. Of the 714 unique DEPs identified by the Stage specific comparison, 395 are unique to Stage 1, 155 are unique to Stage 2, and 164 are shared among both stages (Fig. 7).

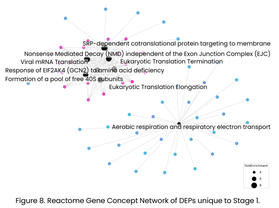

Figure 8. Reactome Gene Concept Network of DEPs unique to Stage 1.

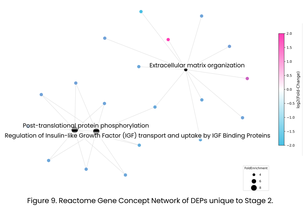

Figure 9. Reactome Gene Concept Network of DEPs unique to Stage 2.

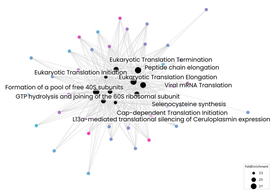

Figure 10. Reactome Gene Concept Network of DEPs shared between Stage 1 and 2.

- Reactome Gene Concept Networks for stage-specific DEPs identify unique signatures. Metabolic changes in Stage 1 (Fig. 8A) and changes in insulin growth factor signaling and ECM dynamics in Stage 2 (Fig. 9), coupled with a shared change in ribosome components and function (Fig. 10) are indicative of disease progression.

- Published NSCLC markers known to increase or decrease through cancer progression based on tissue data are captured by SPARCs (Fig. 11), highlighting the possibility of liquid biopsy-based progression monitoring.

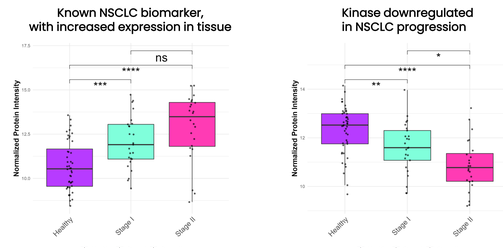

Figure 11. Box plots of biomarkers known to change during NSCLC progression. p ≤ 0.05 (*), p ≤ 0.01 (**), p ≤ 0.001 (***), and p ≤ 0.0001 (****).

Biomarker Discovery

- Supervised machine-learning (ML) algorithms were used to detect differences between NSCLC and Healthy samples.
- 10-fold cross-validation was used to calculate performance metrics (Table 2).
- The best model for protein was selected based on AUC (Table 2, Fig. 12) and was based on information from 9 genes.

Table 2. Performance metrics for leading model: NSCLC vs. Healthy

	Protein	RNA
AUC	0.95	0.98
Combined Sensitivity	0.90	0.98
Stage I Sensitivity	0.92	0.96
Stage 2 Sensitivity	0.88	1
Specificity	0.87	0.98
No of Features	9	24

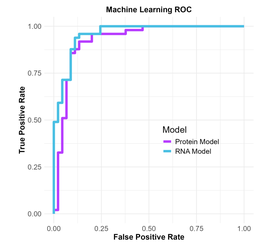

Figure 12. ROC curve for leading RNA and Protein models.

- The 9 features used for the protein model are cancer associated and represent strong biomarker candidates that can distinguish Healthy and NSCLC samples (Fig. 13).

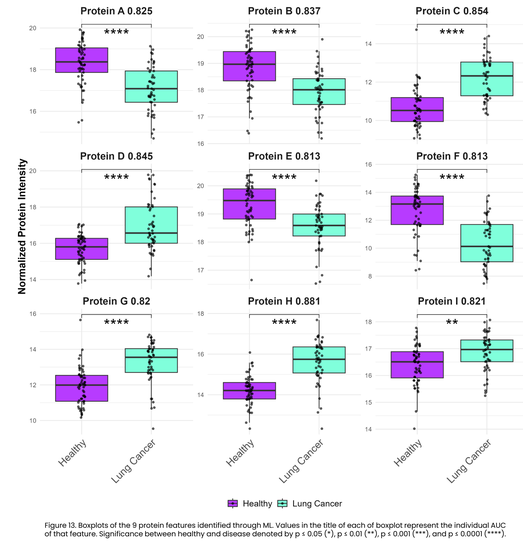

Figure 13. Boxplots of the 9 protein features identified through ML. Values in the title of each of boxplot represent the individual AUC of that feature. Significance between healthy and disease denoted by p ≤ 0.05 (*), p ≤ 0.01 (**), p ≤ 0.001 (***), and p ≤ 0.0001 (****).

Conclusions and Future Directions

- SPARCs enrichment produces molecular signatures that are unique to Stage 1 and 2 lung cancers, demonstrating the sensitivity of the platform.
- DEPs shared by both Stage 1 and stage 2 have a strong enrichment of markers and mechanisms related to the immune response (neutrophils), cytoplasmic translation, and ribosomes, as well as a strong dysregulation of metabolic processing and mitochondrial function.
- Biomarkers across cancer stages are likely to change, and our technology highlights this tumor progression, demonstrating the promising ability of liquid biopsies with TDEVs to augment information gathered from ctDNA.
- Future studies will include blood plasma from stage 3 and 4 NSCLC samples to help identify markers of tumor progression.

References

1. *NCI Cancer Stat Facts: Common Cancer Sites.* Available from: https://seer.cancer.gov/statfacts/html/common.html.
2. Society, A.C., *Cancer Facts & Figures 2024*. Atlanta: American Cancer Society, 2024.
3. Babar I., M.P., Anjum F., *Lung Cancer Screening*. StatPearls. 2024, Treasure Island, FL: StatPearls Publishing.
4. Poon, C., et al., *Why is the screening rate in lung cancer still low? A seven-country analysis of the factors affecting adoption.* Front Pub. Health, 2023. **11**: p. 1264342.
5. Connal, S., et al., *Liquid biopsies: the future of cancer early detection.* J Transl Med, 2023. **21**(1): p. 118.
6. Stejskal, P., et al., *Circulating tumor nucleic acids: biology, release mechanisms, and clinical relevance.* Mol Cancer, 2023. **22**(1): p. 15.