

Article

Plasma-Derived Extracellular Vesicles Convey Protein Signatures That Reflect Pathophysiology in Lung and Pancreatic Adenocarcinomas

Johannes F. Fahrmann ¹, Xiangying Mao ¹ , Ehsan Irajizad ², Hiroyuki Katayama ¹ , Michela Capello ¹, Ichidai Tanaka ^{1,3}, Taketo Kato ¹, Ignacio I. Wistuba ⁴, Anirban Maitra ^{4,5}, Edwin J. Ostrin ^{6,7} , Samir M. Hanash ^{1,4,8} and Jody Vykoukal ^{1,8,*} 

¹ Department of Clinical Cancer Prevention, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA; jffahrman@mdanderson.org (J.F.F.); xmao2@mdanderson.org (X.M.); hkatayama1@mdanderson.org (H.K.); michela.capello@gmail.com (M.C.); ichidai@med.nagoya-u.ac.jp (I.T.); tkato1@mdanderson.org (T.K.); shanash@mdanderson.org (S.M.H.)

² Department of Biostatistics, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA; eirajizad@mdanderson.org

³ Department of Respiratory Medicine, Nagoya University Graduate School of Medicine, Nagoya 466-8560, Japan

⁴ Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA; iiwistuba@mdanderson.org (I.I.W.); amaitra@mdanderson.org (A.M.)

⁵ Department of Pathology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA

⁶ Department of General Internal Medicine, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA; ejostrin@mdanderson.org

⁷ Department of Pulmonary Medicine, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA

⁸ The McCombs Institute for the Early Detection and Treatment of Cancer, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA

* Correspondence: jody@mdanderson.org; Tel.: +1-713-834-6095

Received: 7 April 2020; Accepted: 25 April 2020; Published: 2 May 2020



Abstract: Using a combination of mass-spectrometry and aptamer array-based proteomics, we characterized the protein features of circulating extracellular vesicles (EVs) in the context of lung (LUAD) and pancreatic ductal (PDAC) adenocarcinomas. We profiled EVs isolated from conditioned media of LUAD and PDAC cell lines to identify EV-associated protein cargoes released by these cancer cell types. Analysis of the resulting data identified LUAD and PDAC specific and pan-adenocarcinoma EV protein signatures. Bioinformatic analyses confirmed enrichment of proteins annotated to vesicle-associated processes and intracellular compartments, as well as representation of cancer hallmark functions and processes. Analysis of upstream regulator networks indicated significant enrichment of TP53, MYC, TGFB1 and KRAS-driven network effectors ($p = 1.69 \times 10^{-77}$ – 2.93×10^{-49}) manifest in the adenocarcinoma sEV protein cargoes. We extended these findings by profiling the proteome of EVs isolated from lung ($N = 15$) and pancreatic ductal ($N = 6$) adenocarcinoma patient plasmas obtained at time of diagnosis, along with EVs derived from matched healthy controls ($N = 21$). Exploration of these proteomic data revealed abundant protein features in the plasma EVs with capacity to distinguish LUAD and PDAC cases from controls, including features yielding higher performance in the plasma EV isolates relative to unfractionated plasmas.

Keywords: extracellular vesicles; exosomes; proteomics; liquid biopsy; lung cancer; pancreatic cancer; adenocarcinoma; biomarker discovery

1. Introduction

Extracellular vesicles (EVs) encompass a diverse group of lipid-bound nanoparticles, 50–1000 nm in diameter that are released by most cell types in normal as well as diseased states [1,2]. Exosomes are a distinct class of small extracellular vesicles (sEVs), approximately 50–150 nm-diameter that originate and traffic specifically via the cellular endosomal sorting apparatus [3]. Large extracellular vesicles (IEVs) include vesicle types commonly referred to as microvesicles, oncosomes and apoptotic bodies. Cells also release non-vesicle lipid structures including lipoproteins and other lipid-protein complexes and varied complexes of proteins, nucleic acids and other biomolecules. These vesicle and non-vesicle entities are actively being categorized and investigated and are currently of considerable research interest [4,5].

Emerging evidence indicates EVs function in cell to cell signaling, communication, transfer and exchange in both local and systemic environments [6,7]. The molecular cargoes of EVs have been shown promote cancer progression, invasion and metastasis, remodeling of the tumor microenvironment and angiogenesis [8,9]. We previously demonstrated that EVs disseminated from pancreatic cancer cells communicate a considerable repertoire of tumor antigens that are associated with host humoral immune response and circulating anti-tumor autoantibodies in pancreatic cancer patient plasmas [10].

EVs have been detected in blood, urine, fine needle aspirates, saliva, cerebrospinal fluid and ascites and are considered to convey signature features of the cells from which they originate [11]. Clinical assessment of EV contents from patient-derived biofluids offers promise of minimally invasive means for early detection of cancer and dynamically informing clinical decision-making. In addition, the assessment of EV constituents may yield improvements in assay classification performance, as relevant molecules could conceivably be enriched tumor-derived EVs compared to whole plasmas [12]. Accordingly, there is high interest in developing liquid biopsy assay potential of EVs as markers for early detection of disease, especially in the field of cancer [13–16]. EV-associated cancer marker types investigated include oncogenic micro- and mRNAs, mutated DNAs and tumor proteins [17–19].

Primary effort towards realizing liquid biopsy utility for EV cargoes has, to date, largely focused on EV-associated microRNAs. In lung cancer, a number of groups have reported diagnostic, prognostic and predictive microRNA markers for various lung cancer subtypes [20]. Exploration of EV-associated proteins in biofluids for lung cancer biomarker potential has been more limited and generally targeted or array based [20,21]. We recently demonstrated a feasible approach for exploration of plasma-derived EV protein signatures for detection of lung adenocarcinoma via untargeted mass spectrometry profiling of plasma-derived EVs from patients and matched cancer free controls [22].

Here, using a combination of mass-spectrometry and aptamer array-based proteomics [22,23], we further these findings and explore protein features conveyed by circulating EVs in the context of lung and pancreatic ductal adenocarcinomas. We profile sEVs isolated from conditioned media of a panel of lung and pancreatic adenocarcinoma cell lines and establish characteristic sEV-associated protein signatures of LUAD and PDAC. We then perform comparative profiling of plasma-derived sEVs and paired whole plasmas in the context of lung and pancreatic ductal adenocarcinomas to identify specific protein features with high performance in the circulating sEV compartment.

2. Results

2.1. Proteomic Characterization Extracellular Vesicles from LUAD and PDAC Cell Lines

In this study, we explored for protein signatures of lung and pancreatic adenocarcinoma that are communicated by extracellular vesicles. We first established profiles of sEV-associated proteins originating from cancer cells through proteomic characterization of sEVs derived from conditioned-media of lung (LUAD) and pancreatic ductal (PDAC) adenocarcinoma cell lines. sEVs were isolated by differential ultracentrifugation [24] of concentrated conditioned medias derived from a panel of ten cell lines (NCI-H23, NCI-H647, NCI-H1573, HCC4019, CFPAC-1, HPAF-II, SU.86.86, Panc 03.27, MIA PaCa-2 and PANC-1). Mass-spectrometry-based proteomic profiling was performed

on media-derived sEVs AND corresponding total cell lysates. Accession-level protein isoform data was summed to the parent gene name identifier and normalized according to Spectral Abundance (SA). Within the combined LUAD and PDAC sEV profiling data, 2678 sEV-associated protein features were identified in conditioned media of at least two cell lines in the panel with SA ≥ 2 . Within these protein IDs were 96 of the ExoCarta Top 100 proteins (<http://exocarta.org/>) most often identified in exosomes (Data S1).

Comparison of the mean SA distributions (Figure 1A) for sEV-associated proteins identified in the respective LUAD and PDAC cell lines indicates conserved features in addition to contrasting signatures of vesicle-associated proteins. Principal component analysis (PCA) of the integrated LUAD and PDAC cell line sEV profiling data (Figure 1B) indicates 44.2% of the data variation is accounted for by the first two principle components (PC1: 25.5%, PC2: 18.7%), providing a quantitative metric of extent to which differential, cancer type specific protein features are manifest in the cancer cell line conditioned media sEVs. Figure 1CD and Data S4 summarize cell line derived sEV features evidencing differential representation between the LUAD and PDAC cell lines (2-sided p -value < 0.05 and > 1.5 -fold mean \log_2 (SA) difference). Notably, features exclusive to either the lung or pancreatic cancer cell lines were observed in the cell line derived sEVs.

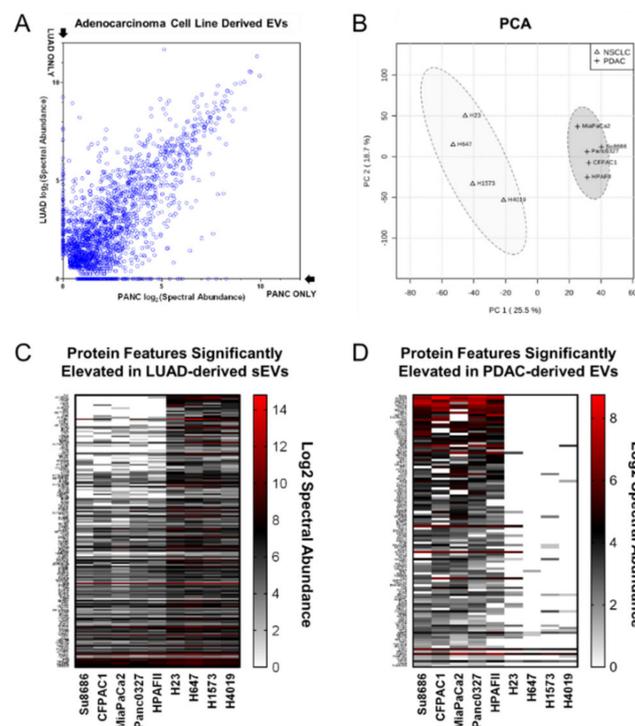


Figure 1. Proteomic profiling of lung and pancreatic ductal adenocarcinoma cell line derived small extracellular vesicles (sEVs). (A) Scatter-plot depicting \log_2 spectral abundance of proteins identified in sEVs from conditioned media of lung (Y-Axis) and pancreatic (X-Axis) adenocarcinoma cell lines. (B) Principal component analysis (PCA) of protein features identified in sEVs from conditioned media of lung and pancreatic adenocarcinoma cell lines. (C) Heatmap indicating statistically significantly (2-sided student T test $p < 0.05$) increased protein features in sEVs of lung versus pancreatic adenocarcinoma cell lines. (D) Heatmap indicating statistically significantly (2-sided student T test $p < 0.05$) increased protein features in EVs of pancreatic versus lung adenocarcinoma cell lines.

Search Tool for the Retrieval of Interacting Genes (STRING) database network [25] gene ontology (GO) enrichment analyses were performed for sEV protein features identified with spectral abundance values ≥ 8 in at least two of the adenocarcinoma panel cell lines (Table A1). These indicated highest-ranked statistical enrichment of GO Biological Process terms vesicle-mediated transport and establishment of localization (FDR = 6.28×10^{-84} – 6.73×10^{-73}); GO Molecular Function terms protein

binding and binding (FDR = 5.70×10^{-88} – 5.88×10^{-65}); and GO Cellular Component terms including cytosol, intracellular and vesicle, as well as organelle and endomembrane system (FDR = 2.42×10^{-154} – 1.81×10^{-64}). Ingenuity Upstream Regulator Analysis of the 2678 sEV-associated protein features identified with confidence in the LUAD and PDAC cell lines revealed (Table A2A, Data S2) the top statistically-significant mechanistic networks converged on established oncogenic mediators TP53 (Cellular tumor antigen p53), MYC (Myc proto-oncogene protein), TGFB1 (Transforming growth factor beta-1 proprotein), MYCN (N-myc proto-oncogene protein) and KRAS (GTPase KRas) ($p = 1.69 \times 10^{-77}$ – 2.93×10^{-49}). Ingenuity Pathway Analysis (IPA) Diseases and Functions analysis identified enrichment of annotations related to cancer, neoplasm or tumor; the top ten annotations, ranked by p -value ($p = 2.76 \times 10^{-133}$ – 1.03×10^{-118}), were all within the Cancer, Organismal Injury and Abnormalities diseases and functions category (Table A2B, Data S3). Among the top canonical pathways represented by sEV-associated protein cargoes were EIF2 signaling ($p = 6.31 \times 10^{-55}$) and mTOR signaling ($p = 1 \times 10^{-35}$) (Data S11). We also assessed the LUAD and PDAC cell line derived sEV protein dataset for specific representation cancer pathway components. The top five significantly enriched Ingenuity Canonical Pathways related to cancer signaling (Figure S1, Table A2C) were breast cancer regulation by stathmin1, ERK/MAPK signaling, role of tissue factor in cancer, FAK signaling and PI3K/AKT signaling ($p = 5.01 \times 10^{-32}$ – 2.00×10^{-17} ; pathway overlap = 47.4–35.6%). Non-small cell lung and pancreatic cancer signaling were also among the enriched pathways ($p = 7.24 \times 10^{-7}$ and 2.04×10^{-5} ; pathway overlap = 32.9% and 25.7%).

Biological relevance of the differentially represented protein features identified in the respective LUAD and PDAC cell line sEV isolates was also considered—174 features exhibited mean log₂ (SA) values > 1.5 fold higher in LUAD relative to PDAC cell line sEVs (Figure 1C, Data S4) and 115 were found to be > 1.5-fold higher in PDAC relative to LUAD cell line sEVs (Figure 1D, Data S4). STRING database network GO enrichment analyses of the sEV protein features identified with elevated spectral abundance in LUAD sEVs revealed statistical enrichment (Table A3) of protein annotations for GO Biological Process terms related to immune activation and effector processes as well as secretion, exocytosis and degranulation (FDR = 2.74×10^{-18} – 2.21×10^{-13}); GO Molecular Function terms encompassing small molecule and other metabolite binding as well as catalytic and other enzymatic activit(ies) (FDR = 3.10×10^{-20} – 3.56×10^{-14}); and GO Cellular Component terms comprising proteasome complex, secretory and vesicle compartments (FDR = 2.68×10^{-20} – 3.50×10^{-14}). STRING database network GO enrichment analyses of the sEV protein features identified with elevated spectral abundance in PDAC sEVs indicated statistical enrichment (Table A4) of protein annotations for GO Biological Process terms related to translation and cellular localization and transport (FDR = 1.72×10^{-10} – 5.36×10^{-10}); GO Molecular Function terms encompassing G-protein and nucleic acid binding (FDR = 1.23×10^{-9} – 1.51×10^{-5}); and GO Cellular Component terms comprising cytosol and cytoplasm as well as G-protein and ribonucleoprotein complex[es] (FDR = 7.57×10^{-17} – 8.30×10^{-7}).

2.2. Isolation of Extracellular Vesicles from LUAD and PDAC Patient Plasmas

In order to assess translational, cancer marker aspects of extracellular vesicle-associated protein cargoes, we profiled sEVs derived from retrospective plasma samples obtained from recently diagnosed, treatment naïve lung adenocarcinoma ($N = 15$) and pancreatic ductal adenocarcinoma ($N = 6$) cases as well as respective age and sex matched controls ($N = 21$) (Table A5). LUAD plasmas comprised Stage I, II and IV disease; all PDAC plasmas were from Stage IV disease, consistent with typical (>80%) initial clinical presentation of pancreatic cancer at advanced stage.

For in-depth proteomic exploration of sEVs from biofluids, vesicle recovery and purity yields are optimized based on a streamlined workflow comprising an initial 16,500× g centrifugation to deplete large-diameter EVs, followed by ultracentrifugation-based density flotation through a single-step density-overlay ($\rho = 1.14$ g/mL). In previous studies, we determined this approach facilitates efficient sample processing and throughput, while effectively depleting free, non-lipid bound protein backgrounds and enriching the sEV compartment for proteomic characterization [10]. Using this

density flotation approach, sEVs were isolated from LUAD and PDAC case and control plasma specimens according to standardized, uniform conditions (Figure 2A). An aliquot of unfractionated whole plasma from each cohort sample was also reserved to allow parallel profiling of paired samples to enable comparative analysis of feature representation in sEVs relative to intact plasma (Figure 2A). Nanoparticle-tracking analyses of representative samples confirmed enrichment of sEVs in the harvest fractions with an average nanoparticle diameter of 74 ± 8 nm and 98 ± 18 nm for cases and controls, respectively (Figure 2B). In concordance with the nanoparticle-tracking data, transmission electron microscopy (TEM) of representative samples from the low-density harvest fraction confirmed recovery of vesicles of size and morphology typical of exosomes and sEVs (Figure 2C), this was further indicated by positive detection by immunoblotting of representative samples from the harvest fraction with standard exosome markers CD63 and TSG101 (Figure S2).

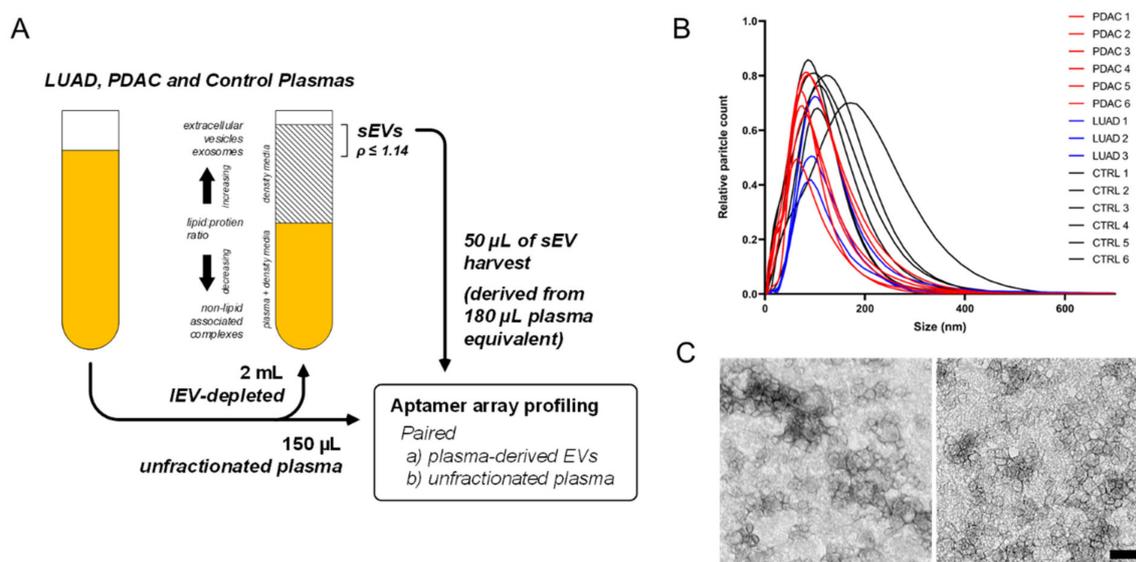


Figure 2. Isolation and physical characterization of plasma-derived sEVs. **(A)** Schematic illustrating density-based fractionation approach for isolation of high-purity small extracellular vesicles (sEVs) from plasma and other biofluids. **(B)** Size distribution plot of plasma-derived sEV harvests based on nanoparticle tracking analysis. **(C)** Representative transmission electron micrographs for isolated plasma sEVs; left is case, right is control; scale bar indicates 100 nm.

2.3. Profiling of Extracellular Vesicles from LUAD and PDAC Patient Plasmas

We performed affinity-based proteomic profiling of patient plasma-derived sEVs using multiplex aptamer-arrays (SOMAscan, SomaLogic, Inc., Boulder, CO, USA) to investigate representation of adenocarcinoma protein features uncovered in our initial proteomic characterization of LUAD and PDAC cell line sEVs. We additionally explored the potential for these sEV-associated features to distinguish cancer cases from controls in plasma. Aptamer-array assays were performed on individual samples of plasma sEVs ($N = 42$) and the corresponding unfractionated plasma paired-aliquot ($N = 42$) to assess expression levels of 1,305 targets encompassing a diverse set of molecular functions, including cancer [23,26,27].

The distribution of proteins among LUAD and PDAC cases and controls in plasma-derived sEVs and unfractionated plasmas as analyzed by aptamer array is summarized in Figure S3. Classifier performances of individual proteins analyzed in the respective compartments were calculated using T test and Receiver Operating Characteristic (ROC) area under the curve (AUC) values to evaluate capacity for distinguishing cases from matched controls. The LUAD and PDAC cohorts were analyzed separately and also as a combined, pan-adenocarcinoma case cohort. A total of 108 and 397 proteins were identified in plasma-derived sEVs that yielded statistically significant (2-sided Wilcoxon rank sum test $p < 0.05$) ROC AUCs for distinguishing cases from controls in the respective LUAD and PDAC

plasma cohorts, whereas 328 and 272 significant proteins were identified in unfractionated plasma in the respective LUAD and PDAC cohorts (Data S5 and S6). *p*-values following adjustment for multiple hypothesis testing (*q*-values) are also provided in Data S5 and S6.

Next, we focused on protein features that yielded ROC AUCs ≥ 0.7 for distinguishing cases from controls. In the LUAD cohort, 37 plasma-derived sEV protein features exhibited individual ROC AUCs ≥ 0.7 with capacity to distinguish cases from controls, whereas 117 proteins exhibited ROC AUCs ≥ 0.7 in unfractionated plasma (Figure 3A, Data S5); 34 of the 37 proteins with ROC AUCs ≥ 0.7 in the sEV compartment also exhibited higher AUC performance in the sEV compartment compared to unfractionated plasma (Figure 3A; Data S5). IPA of the 37 plasma-derived sEV protein features with ROC AUCs ≥ 0.70 revealed NKX3-1 ($p = 1.4 \times 10^{-9}$) as a top causal network, Cancer, Cell Death and Survival as a top disease function ($p = 1.57 \times 10^{-8}$) and AHR, VEGF, EGF, IRF1 and STAT3 as predicted significant upstream regulators (Data S12).

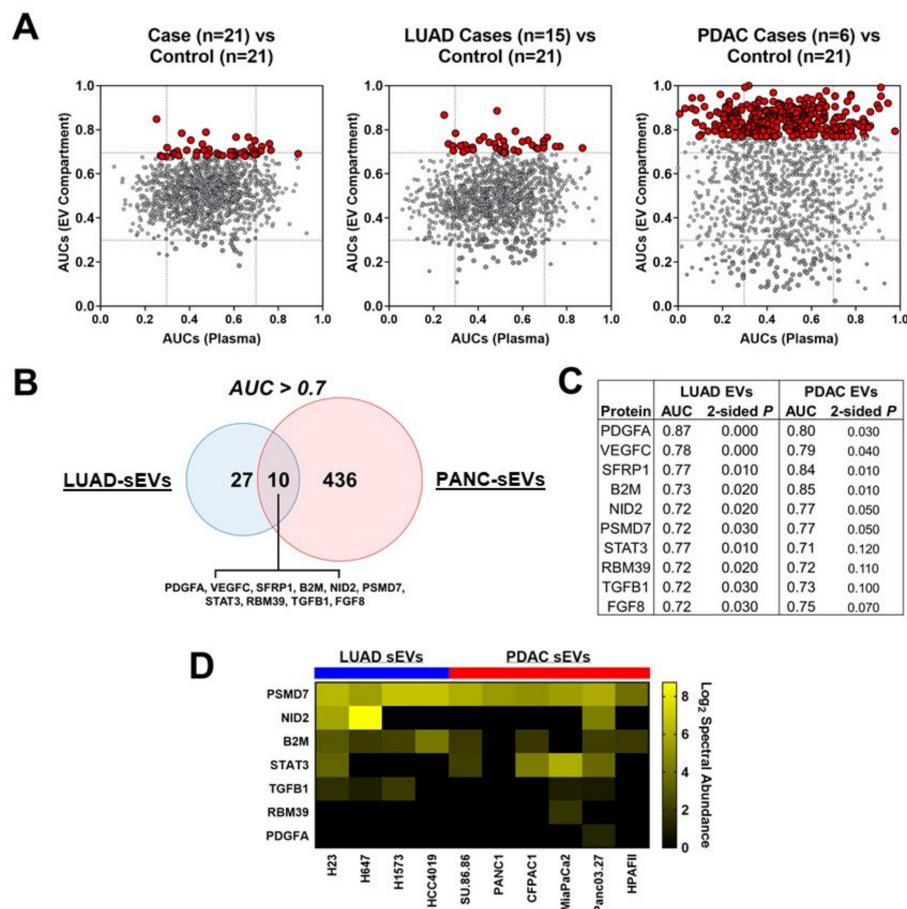


Figure 3. Profiling of adenocarcinoma patient plasma-derived sEVs. (A) Scatter plot illustrating Receiver Operating Characteristic (ROC) area under the curve (AUC) of aptamer array protein features identified in the sEV compartment (Y-axis) and matched unfractionated total plasma (X-axis) comparing all cases ($N = 21$) versus controls ($N = 21$), lung adenocarcinomas (LUAD) cases ($N = 15$) versus controls ($N = 21$) and pancreatic ductal (PDAC) adenocarcinoma cases ($N = 6$) versus controls ($N = 21$). Red nodes indicate protein features that were statistically significant (Wilcoxon rank sum test $p < 0.05$) in the sEV compartment. (B) Venn diagram for sEV-derived protein features with AUC point estimates > 0.7 in distinguishing LUAD and PDAC cases from cancer-free controls. (C) Table summarizing AUC and *p*-values for the 10 overlapping protein features that exhibited AUC point estimates > 0.7 for distinguishing both LUAD and PDAC cases from cancer-free controls. (D) Heatmap indicating cell line sEV protein log₂ (Spectral Abundance) for overlapping features common between LUAD and PDAC cases (C) that were also identified (7/10) in the adenocarcinoma cell line-derived sEVs.

In the PDAC cohort, 446 plasma-derived sEV protein features yielded individual ROC AUCs ≥ 0.7 for delineating case from matched-controls whereas 273 proteins were identified with individual ROC AUCs ≥ 0.7 in unfractionated plasma (Figure 3A, Data S6); 413 of the 446 proteins with ROC AUCs ≥ 0.7 in the sEV compartment also exhibited higher AUC performance in the sEV compartment compared to unfractionated plasma (Figure 3A; Data S6). IPA of the 446 plasma-derived sEV protein features with ROC AUCs ≥ 0.70 indicated *YAP1* as a top causal network, Cancer, Organismal Injury and Abnormalities as a top disease function and KRAS and TP53 as top predicted statistically significant upstream regulators (Data S13).

We also considered adenocarcinoma plasma sEV features ($N = 10$) that exhibited ROC AUCs ≥ 0.7 classifier performance for distinguishing LUAD as well as PDAC cases from controls (Figure 3B). Six of these sEV-associated proteins (PDGFA, VEGFC, SFRP1, B2M, NID2 and PSMD7) yielded statistically significant (2-sided Wilcoxon rank sum test $p \leq 0.05$) ROC AUCs for distinguishing both LUAD and PDAC cases from controls (Figure 3C). Intersection of these high classifier performance plasma sEV protein features with the LUAD and PDAC cell line sEV data showed concordant representation in the in vitro and in vivo profiles, particularly for sEV-associated PSMD7, B2M and STAT3 (Figure 3D).

2.4. Extracellular Vesicle Associated Protein Cargoes Reflect Cancer Pathophysiology and Present a Diverse Intercellular Milieu

To further explore the extent to which the identified plasma-derived sEV-associated protein features are associated with cancer, we overlapped protein targets with ROC AUCs ≥ 0.7 in the sEV compartment with those proteins identified in the initial in-depth proteomic profiling of LUAD and PDAC cancer-cell line derived sEVs. For LUAD, 25 of the 37 plasma-derived sEV protein features that exhibited ROC AUCs ≥ 0.7 were also identified in the LUAD cancer cell line sEVs. Ingenuity Upstream Regulator Analysis of the 25 high-classifier performance plasma sEV-associated proteins also identified in the cell line sEVs revealed the top upstream gene, RNA and protein regulators included AHR, VEGF, EGF, IL6, IRF1 and STAT3 ($p = 1.22 \times 10^{-7}$ – 3.45×10^{-6}), consistent with plasma sEV representation of cancer-associated network mediators (Table A6A, Data S9).

For PDAC, 257 of the 446 plasma-derived sEV protein features with ROC AUCs ≥ 0.7 were also represented in PDAC cancer cell line sEVs. Ingenuity Upstream Regulator Analysis of these 257 high-classifier performance plasma sEV-associated protein features that were concordant with the cell line sEVs revealed the top upstream gene, RNA and protein regulators included TNF, IFNG, TGFBI, IL1B, KRAS, HGF and TP53 ($p = 2.45 \times 10^{-29}$ – 2.96×10^{-18}), indicating patent manifestation of oncogenic protein effectors in the plasma sEV compartment. (Table A6B, Data S10).

We also investigated for biological relevance of these overlapping features by performing subcellular localization analysis based on the COMPARTMENTS localization evidence database scores [28], filtering for genes confidently assigned to at least one of the 11 subcellular localizations (confidence score ≥ 2)—nucleus, cytosol, cytoskeleton, peroxisome, lysosome, endoplasmic reticulum, Golgi apparatus, plasma membrane, endosome, extracellular spaces and mitochondrion (Data S7 and S8). Compartments were then ranked based on the number of genes with high confidence (Figure 4A,B). Interestingly, subcellular localization analysis of these features common between plasma sEVs and cancer cell line-derived sEVs for LUAD and PDAC indicated conspicuous representation of proteins annotated as localized to the nucleus and mitochondria (Figure 4A,B), suggesting that altered protein localization can coordinately reflect disease status that is captured and conveyed by EV profiles.

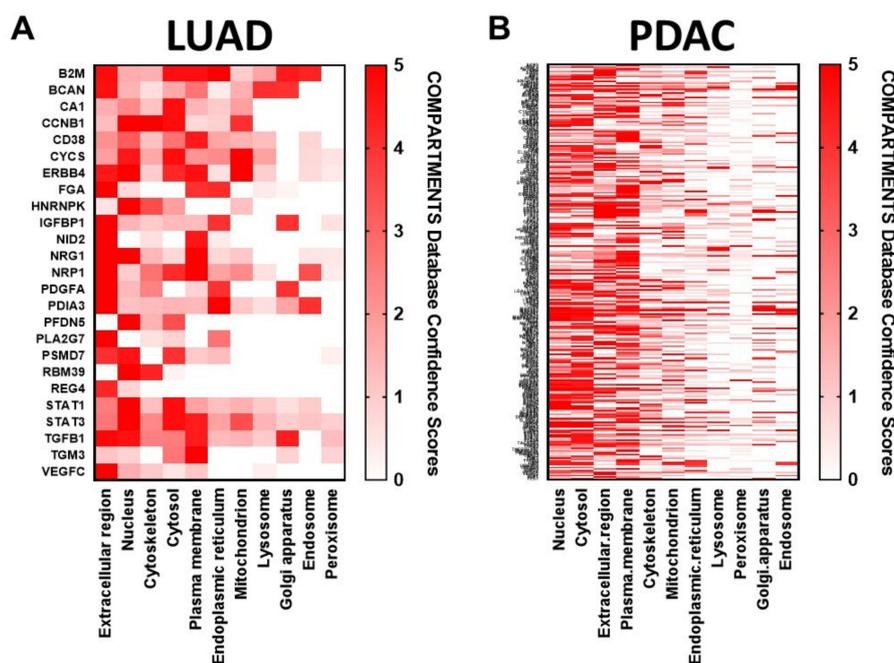


Figure 4. COMPARTMENTS subcellular localization analysis for (A) LUAD and (B) PDAC plasma-derived sEV features exhibiting ROC AUC classifier performance of >0.7 that were also identified in adenocarcinoma cell line-derived sEVs. The red shading indicates the COMPARTMENTS integrated confidence score for localization to each of eleven indicated subcellular compartments. A full list of the genes together with their compartment confidence scores is provided in Data S7 and S8.

3. Discussion

Through comprehensive proteomic profiling of extracellular vesicles derived from cell line conditioned medias and patient plasmas, we have identified protein signatures of lung and pancreatic ductal adenocarcinomas that are conveyed by cancer cell-disseminated extracellular vesicles. These protein features provide circulating evidence of wide-ranging oncogenic reprogramming and functional alteration in cancer cells. Aptamer array-based profiling of sEVs enriched from LUAD patient plasmas enabled enumeration of 37 protein features with $AUC \geq 0.7$ case:control classifier performance; 34 (92%) of these indicated improved performance in the sEV compartment compared to in intact plasma. Correspondingly, profiling of sEVs enriched from PDAC patient plasmas revealed 446 protein features with $AUC \geq 0.7$; 413 (93%) of these showed improved performance in the sEV compartment. The number of high classifier performance plasma sEV-associated features identified in the PDAC cohort likely derives from the advanced disease stage of these patients, consistent with the typical clinical encounter of pancreatic cancer.

We note that a novel component of this study in relation to other studies, which have similarly evaluated utility of EV-associated protein cargo for early detection or prognostication [29,30], is the direct comparison protein features in plasma-derived EVs versus unfractionated plasma from the same individual for utility in distinguishing case from control. This approach enables us to address a fundamental question and conclude that there are indeed EV-associated features that exhibited improved classifier performance specifically in the plasma sEV compartment.

Interestingly, the plasma sEV-associated proteins (Figure 3B,C) identified in plasma with $AUC \geq 0.7$ case:control classifier performance in both the LUAD and PDAC cohorts denote a pan-adenocarcinoma sEV signature and suggest intriguing implications regarding markers that yield insight into cancer-host interchange within the tumor microenvironment. For example, PDGF (Platelet-derived growth factor subunit A) has been shown to be involved in recruitment and activation of cancer-associated fibroblasts in lung adenocarcinoma [31] and to also exert similar effect on pancreatic stellate cells and fibroblasts in PDAC tumors [32,33]. Tumor derived VEGFC (Vascular endothelial growth factor C) has been

demonstrated to support metastatic expansion of primary tumors into the lymphatic vasculature in many tumor types, challenging the traditional concept of a purely passive role of the lymphatic vessels in cancer metastasis [34–36]. PSMD7 (26S proteasome non-ATPase regulatory subunit 7) has been reported as overexpressed in a variety of tumors and is hypothesized to be involved in maintenance of proteasome function in cancer cells through its roles in mediating endosomal trafficking and protein recycling [37,38]. This is especially intriguing in the context of our previous discovery of another member of the ubiquitin-proteasome pathway [39], HUWE1, as a plasma EV-associated marker of lung adenocarcinoma with good performance for distinguishing cases from controls [22]. NID2 (Nidogen-2) is a basement membrane protein primarily produced by mesenchymal cells [40] that has been linked to mesenchymal/de-differentiated phenotypes in breast cancer and melanoma [41]. Elevated serum NID2 has been reported in cases of ovarian and esophageal cancers [42]. In esophageal, lung, bladder and oral cancers, *NID2* methylation and reduced tissue expression are observed [43], suggesting biomarker context of NID2 may diverge in tumor tissue and EV compartments. SFRP1 (Secreted frizzled-related protein 1) is a glycoprotein modulator of Wnt-signaling that has been observed to play both tumor suppressor and oncogenic roles in a number of human cancers and to also be subject to epigenetic regulation via DNA methylation or microRNA transcriptional silencing [44]. A recent study demonstrated functional relevance of exosome-associated SFRP family proteins and that human lung cancer cells take up SFRP-containing exosomes, whereby SFRPs co-localize with β -catenin in both the cytoplasm and nucleus to mediate β -catenin/TCF-mediated transcription of Wnt target genes known to effect cancer stem cell properties [45]. B2M (Beta-2-microglobulin) is a component of the class I major histocompatibility complex (MHC-I) and is a crucial factor required for MHC-I assembly and maintenance of stable surface presentation of antigens to immune system effectors [46]. We previously found MHCs and presentation pathway proteins to be enriched on the surface of PDAC cell derived EVs [10] and there are multiple lines of evidence that tumor cells employ EVs to transfer MHCs loaded with tumor-derived peptides to antigen presenting cells, thereby activating anti-tumor CD8 T-cell response [47]. Interestingly, inactivation of *B2M* in cancer cells has been associated with downregulation of the MHC-I complex, abnormal immune surveillance that contributes to cancer development and attenuated responses to anti-PD-1/anti-PD-L1 immunotherapies [48], highlighting the potential for cancer EV cargoes to report on tumor immune status. Notably, serum-derived EV-associated LBP (Lipopolysaccharide-binding protein) has previously been shown to be elevated in PDAC cases as compared to healthy controls. Consistent with this study, we also observed that LBP is elevated in EVs of PDAC cases as compared to controls, thus providing independent validation [29].

The cell line and plasma sEV-associated proteins uncovered in this study are also known effectors of cancer-associated mechanistic regulators including transcription factors TP53 and MYC; signal transducer KRAS; cell growth and proliferation cytokines TGF β 1, HGF, VEGF and EGF; and proinflammatory and/or immunoregulatory cytokines TNF, IL1 β , IL6 and IFN γ , all with established, fundamental roles in cancer etiology. In addition, evaluation of the identified plasma sEV-associated proteins regarding subcellular compartment of origin revealed prominent representation, interestingly, of proteins manifest in the adenocarcinoma cancer cell line and patient plasma-derived sEVs annotated to nuclear and mitochondrial compartments, in addition to the expected extracellular and plasma membrane cellular compartments. This is indicative of sEV conveyance of multifaceted signatures of cancer and broad utility for interrogation of cancer intercellular dynamics.

Exploration of cancer associated extracellular vesicles has revealed diverse repertoires of proteins, nucleic acids, lipids and metabolites that reflect the altered molecular expression, sorting, trafficking and fates that underlie cancer pathophysiology [4]. The promise of liquid biopsy interrogation of EV multidimensional cargoes for real time reporting of tumor dynamics would redefine clinical management of cancer [49]. This is especially relevant in the emerging era of anti-cancer immunotherapy in which understating cancer immune-infiltrate interaction and varying tumor phenotype would guide personalized therapies and predict treatment outcomes [10].

As clinical oncology is increasingly guided by molecular insight and deployment of targeted precision therapies, approaches for improved characterization of tumor molecular features are ever more essential. Liquid biopsy based on assay of tumor-derived information disseminated into the peripheral blood or other biofluids complements traditional tissue biopsy and offers advantages of being less invasive, allowing for serial interrogation and potential for overcoming sampling error by integrating information across cancer cell subpopulations and immune and stromal infiltrates within the tumor milieu. Liquid biopsy research efforts to date have concentrated primarily on circulating tumor cells (CTCs), cell-free DNA (cfDNA), circulating tumor DNA (ctDNA) and cell free tumor RNA, with EVs being of more recent interest [50]. It is likely that no single mode of deriving tumor molecular information will be sufficient arrive at an optimal biomarker panel for the desired clinical application. Therefore, it is essential to evaluate the strengths and complementary contributions of various approaches to liquid biopsy [49,51].

The feasibility of using ctDNA for tracking and monitoring tumor dynamics, drug response and therapy resistance has been demonstrated through numerous studies [52]. Nevertheless, practical challenges remain as cfDNA is generally highly fragmented and the total amount of ctDNA can be as low as 0.01% of the total cfDNA in circulation [52]. Strategies to improve this by enriching for ctDNA associated with circulating tumor-derived EVs have been developed [18]. While ctDNA can yield cancer genotype profiles, it offers limited utility for discerning expression level dynamics that underlie phenotypic plasticity and give rise to tumor stemness, progression or therapeutic escape. The potential of circulating tumor cells in liquid biopsy has been widely investigated in scientific and clinical studies [50]. CTCs provide direct tumor information including real-time phenotypic profiles that can inform treatment selection and monitor response. CTCs are, however, exceedingly rare in blood samples and are outnumbered by normal leukocytes by 10^6 -fold or more. A clinical blood sample may contain as few as 5–10 CTCs and these cells may be heterogeneous, necessitating that analyses be performed at the single-cell level [50].

EVs provide an additional source of material for liquid biopsy based profiling of tumors. Similar to CTCs, their molecular cargoes can report tumor genotype and phenotype. Mechanistically, their constitutive expression and physical features seem to yield relatively straightforward conveyance of abundant tumor EVs into blood and biofluids compared to the relatively complex process of CTC extravasation into the peripheral circulation. This makes tumor EVs and their associated cargoes an appealing target for liquid biopsy. Nevertheless, challenges remain, for example in identifying surface protein profiles that would enable capture and enrichment of tumor-derived EVs from uninformative, circulating background populations. Herein, we have identified cancer-associated EV-protein features that may offer utility for early interception of disease. Our findings provide a basis for inclusion of identified EV-resident protein targets on validation studies that encompass other biomarker candidates [51,53,54].

4. Materials and Methods

4.1. Blood Samples

Whole blood samples were collected at MD Anderson Cancer Center (MDACC) through informed consent following institutional review board (IRB) approval (PA14-0552). Healthy control samples were obtained from volunteers in the clinic waiting rooms and for the most part, are relatives of the patients. Plasma was prepared from EDTA-treated whole blood by two successive room temperature centrifugation steps for 12 min at $1200\times g$, without braking and subsequently stored in $-80\text{ }^{\circ}\text{C}$ until use.

4.2. Exosome Isolation

Cell line sEVs were purified from cell-conditioned media obtained from LUAD and PDAC cell lines (NCI-H23, NCI-H647, NCI-H1573, HCC4019, CFPAC-1, HPAF-II, SU.86.86, Panc 03.27, MIA PaCa-2 and PANC-1) by differential ultracentrifugation after the basic protocol of Théry, et al. [24].

Briefly, cells were cultured for 48 h in serum-free media and subjected to sequential centrifugation steps of $800\times g$ and $2000\times g$. The resulting supernatant was next filtered through a $0.22\ \mu\text{m}$ PES vacuum filter (Corning) and concentrated ~ 50 -fold by tangential flow filtration (TFF) using a 100 kDa MWCO membrane (Vivaflow 200R, Sartorius AG, Göttingen, Germany). The concentrated TFF retentate was then subjected to ultracentrifugation at $100,000\times g$ for 2 h in 45Ti fixed angle rotor (Beckman Coulter Inc., Brea, CA, USA). The supernatant was removed and PBS added to the pellet for an overnight washing step, again at $100,000\times g$. The resultant sEV pellet was resuspended in PBS and harvested for further analyses.

For the isolation of plasma sEVs we applied a density gradient flotation approach. Cell debris and large EVs were depleted by centrifugation at $2000\times g$ for 20 min followed by $16,500\times g$ for 30 min; the resulting supernatant was additionally filtered through a pre-wetted $0.22\ \mu\text{m}$ vacuum filter (Steriflip SCGP00525, MilliporeSigma, Burlington, MA, USA). The IEV-depleted plasma was mixed with OptiPrep iodixanol solution (MilliporeSigma, St. Louis, MO, USA) to a final density of 1.18 g/mL. This was loaded into the bottom of a polycarbonate ultracentrifuge tube (Seton Scientific Inc., Petaluma, CA, USA) and overlaid with 2–3 mL of 1.14 g/mL iodixanol/PBS (phosphate buffered saline) solution to form a single-step density fractionation gradient. Ultracentrifugation was performed for at $100,000\times g$ for 16 h at $8\ ^\circ\text{C}$. Vesicles were collected in a single fraction from the top of the tube, proceeding downward to recover 0.6 mL of overlaid gradient volume.

4.3. Transmission Electron Microscopy

Extracellular vesicle aliquots were fixed in 2% paraformaldehyde; 5 μL of EV suspension was then applied to each formvar/carbon-coated 200 mesh nickel grid and allowed to adsorb for 20 min. The grids were washed twice on 100 μL drops of PBS, followed by eight washes with deionized water. Uranyl acetate (2%) was used as a counterstain; after 1 min of staining, excess uranyl acetate was blotted from the grid edge with Whatman No. 1 filter paper and the grids were air-dried. EM grids and reagents were from EMS (Hatfield, PA, USA). Imaging was performed using a JEM1010 TEM (JEOL, Peabody, MA, USA) at an accelerating voltage of 80 kV. Digital images were taken with AMT imaging software (Advanced Microscopy Techniques Corp, Danvers, MA, USA).

4.4. Particle Size Distribution and Quantification

EV yields were quantified via Brownian diffusion size analyses using ZetaView Nanoparticle-tracking analysis (NTA) instrumentation (Particle Metrix, Meerbusch, Germany). Sample aliquots were diluted 10^2 – 10^6 -fold to achieve optimal concentration for analysis; 1.0 mL of diluted sample was used for each analysis. Light scattering of individual particles in solution was digitally recorded, particle trajectory and displacement were automatically analyzed by image analysis tracking software and the particle-size distribution was determined from the observed Brownian motion of individual particles according to the Stokes-Einstein relationship.

4.5. Protein Quantitation, SDS-PAGE and Western Blot Assay

Protein quantification was performed using Pierce BCA (bicinchoninic acid) protein assay (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's recommended microplate assay procedure. Absorbance was measured with a SpectraMax M5 multi-mode microplate reader using SoftMax Pro data acquisition and analysis software (MolecularDevices, San Jose, CA, USA). Vesicle isolates were denatured in $4\times$ Laemmli sample buffer at $100\ ^\circ\text{C}$ for 10 min. Proteins were separated using 4–15% sodium dodecyl sulfate polyacrylamide gel electrophoresis in Tris/Glycine/SDS running buffer and transferred to Immun-Blot PVDF (polyvinylidene difluoride) membrane (all reagents and supplies from Bio-Rad, Hercules, CA, USA). Immunoblotting was performed with the following primary antibodies: CD9 (EXOAB-CD9A-1, System Biosciences, Palo Alto, CA, USA); TSG101 (EXOAB-TSG101-1, System Biosciences). Blots were washed and incubated with appropriate HRP-conjugated secondary antibodies (Amersham ECL, Cytiva, Marlborough, MA, USA) and detected

using Pierce ECL western blotting substrate (Thermo Fisher Scientific) with chemiluminescence optimized autoradiography film.

4.6. Aptamer Array

The SOMAscan assay is arrayed from SOMAmer (slow off-rate modified DNA aptamer) reagents with affinity optimized for individual protein targets. Unfractionated plasma (150 μ L) was prepared for SOMAscan assay using SomaLogic Plasma Diluent and Assay Buffer. Aliquots of plasma-derived EVs from each sample were normalized to 180 μ L plasma input equivalent and prepared for SOMAscan assay by diluting solubilization buffer containing 120 mM NaCl, 5 mM KCl, 5 mM MgCl₂, 40 mM HEPES pH 7.5, 0.05% Tween20, 1% DDM (*w/v*) and 0.5% sodium deoxycholate (*w/v*) and incubated for 15 min with mild agitation, followed by centrifugation at 14,000 \times *g* for 5 min. Final volume sample volume was standardized to 120 μ L. SOMAscan assays were performed by SomaLogic. Briefly, samples were incubated with a pools of SOMAmer reagents for equilibration binding followed by two sequential bead-based immobilization and washing steps to eliminate unbound and non-specifically bound proteins and unbound SOMAmer reagents. The remaining SOMAmer reagents were isolated and quantified simultaneously on a custom Agilent hybridization array such that the measured quantity of each SOMAmer is proportional to the corresponding target protein concentration in the original sample.

4.7. Mass Spectrometry Analysis

For cell line derived EVs, protein digestion and identification by LC-MS/MS (Liquid Chromatography-tandem Mass Spectrometry) was performed based on our established protocol [55]. Intact protein separation was performed using a UPLC (Ultra Performance Liquid Chromatography) system (Waters Corporation, Milford, MA, USA) with reversed-phase column 4.6 mm \times 150 mm (Column Technology Inc., Fremont, CA, USA); eluted proteins were subjected to in-solution trypsin digestion followed by LC-MS using a NanoAcquity UPLC system equipped with a Waters Symmetry C18 nanoAcquity trap-column (180 μ m \times 20 mm, 5 μ m) and a C18 analytical column (75 μ m \times 150 mm, 1.8 μ m, Column Technology Inc.) coupled in-line with a Waters SYNAPT G2-Si mass spectrometer. LC-HDMSE (Liquid Chromatography-data independent High Definition Mass Spectrometry) data were acquired in resolution mode with SYNAPT G2-Si using Waters Masslynx (version 4.1, SCN 851). The capillary voltage was set to 2.80 kV, sampling cone voltage to 30 V, source offset to 30 V and source temperature to 100 $^{\circ}$ C. Mobility utilized high-purity N₂ as the drift gas in the IMS TriWave cell. Pressures in the helium cell, Trap cell, IMS TriWave cell and Transfer cell were 4.50 mbar, 2.47 \times 10⁻² mbar, 2.90 mbar and 2.53 \times 10⁻³ mbar, respectively. IMS wave velocity was 600 m/s, helium cell DC 50 V, Trap DC bias 45 V, IMS TriWave DC bias V and IMS wave delay 1000 μ s. The mass spectrometer was operated in V-mode with a typical resolving power of at least 20,000. All analyses were performed using positive mode ESI using a NanoLockSpray source. The lock mass channel was sampled every 60 s. The mass spectrometer was calibrated with a [Glu1] fibrinopeptide solution (300 fmol/ μ L) delivered through the reference sprayer of the NanoLockSpray source. Accurate mass LC-HDMSE data were collected in an alternating, low energy (MS) and high energy (MSE) mode of acquisition with mass scan range from *m/z* 50 to 1800. The spectral acquisition time in each mode was 1.0 s with a 0.1-s inter-scan delay. In low energy HDMS (High Definition Mass Spectrometry) mode, data were collected at constant collision energy of 2 eV in both Trap cell and Transfer cell. In high energy HDMSE mode, the collision energy was ramped from 25 to 55 eV in the Transfer cell only. The RF applied to the quadrupole mass analyzer was adjusted such that ions from *m/z* 300 to 2000 were efficiently transmitted, ensuring that any ions observed in the LC-HDMSE data less than *m/z* 300 arose from dissociations in the Transfer collision cell. The acquired LC-HDMSE data were processed and searched against protein knowledge database (Uniprot) through ProteinLynx Global Server (PLGS, Waters Corporation) with 4% false discovery rate (FDR). Spectral counts were normalized to total spectral abundance—each identified

peptide count was divided by the total count for each analysis and rescaled using a constant factor of 50,000.

4.8. Statistical and Bioinformatics Analyses

Raw assay data were log₂-transformed, followed by median and 95% confidence interval calculation using bootstrap method as a resampling technique. Raw data were z-score transformed within each group LUAD and PDAC cases and matched healthy controls and ROC AUC was determined for 1305 proteins. *P* values were computed using two-sided Wilcoxon rank-sum test to compare cases and controls and the values were subjected to multiple testing correction to determine *q*-values. Corresponding plots were generated using ggplot2 (version 3.2.1) in the R software environment (version 3.6.1, The R Foundation, <https://www.r-project.org>).

For the proteins that were identified both in protein array with high performance (AUC ≥ 0.7) and in previous cell line exosome proteomics (MS2 ≥ 5 in at least one cell line), the MS2 counts in total exosome extract (TEE), cargo and surface were normalized to the total MS2 count in each cell line. The normalized count (per 10,000 total count) were represented in heatmaps using gplots (version 3.0.1.1, The Comprehensive R Archive Network, <https://github.com/talgalili/gplots>) in the R software environment (version 3.6.1) for visualization.

Protein subcellular localizations were determined using COMPARTMENTS database (<https://compartments.jensenlab.org/Search>). The confidence levels of protein targets associated with 9 major subcellular localizations were first extracted and the heatmaps were generated using gplots (version 3.0.1.1) in the R software environment (version 3.6.1). Unsupervised hierarchical clustering was performed by Ward's method to aggregate the compartments with similar patterns of protein localization.

To identify the canonical pathways that were enriched for the genes with high performance in lung and pancreatic adenocarcinoma respectively, Ingenuity Pathway Analysis (IPA) and Ingenuity Upstream Regulator Analysis was used to further explore feature relationships, independent of established canonical pathways according to the associations contained within the Ingenuity Knowledge Base.

5. Conclusions

Through comprehensive proteomic analyses of cell line and plasma-derived extracellular vesicles, we have investigated the repertoire of proteins carried by sEVs in lung and pancreatic adenocarcinomas. We demonstrate that sEVs indeed harbor cancer cell disseminated disease signatures that may not be detected by or that would complement those attained from conventional profiling of total plasma. These findings demonstrate value in continued efforts to query the plasma sEV proteome towards discovery of novel biomarkers for the diagnosis and clinical management of cancer.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-6694/12/5/1147/s1>, Figure S1: Top five Ingenuity Canonical Cancer Signaling Pathways with significant representation in LUAD and PDAC cell line EVs, Figure S2: Characterization of plasma-derived sEVs, Figure S3: Distribution of proteins among LUAD and PDAC cases and controls in plasma-derived sEVs and unfractionated plasmas as analyzed by aptamer array, Data S1–S13: The MS proteomics data generated and analyzed during this study have been deposited to the ProteomeXchange Consortium via the EMBL-EBI PRIDE partner repository with the dataset identifier PXDxxxxxx. Other relevant data supporting the findings of this study including IPA project files are available within the Article and Supplementary Files or are available from the authors upon reasonable request.

Author Contributions: Conceptualization, J.F.F., E.J.O., S.M.H. and J.V.; methodology, H.K. and J.V.; formal analysis, J.F.F., X.M., E.I. and J.V.; investigation, H.K., M.C., I.T. and J.V.; resources, I.I.W., A.M. and S.M.H.; data curation, J.F.F., X.M., E.I., M.C., I.T., T.K. and J.V.; writing—original draft preparation, J.F.F. and J.V.; writing—review and editing, X.M., E.I., H.K. and S.M.H.; project administration, S.M.H. and J.V.; funding acquisition, I.I.W., A.M. and S.M.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Institutes of Health (NIH) through the University of Texas Lung Specialized Programs of Research Excellence (SPORE) grant P50CA70907 [I.I.W.]; NIH grants 5-R01-CA218230-02, 1-P50-CA221707-01A1, 5-U01-CA196403-05, 5-U01-CA200468-04 and 1-U24-CA224020-01 [A.M.]; and The Cancer Prevention and Research Institute of Texas grant RP180505 [S.H.]. The UTMDACC Institutional Tissue Bank (ITB) and High Resolution Electron Microscopy Facility (HREMF) are supported by NIH

Cancer Center Support Grant (CCSG) P30-CA016672; confocal microscopy in the BSRB Microscopy Facility at UTMDACC is supported by NIH shared instrumentation grant (1-S10-OD024976-01).

Acknowledgments: The authors acknowledge Kenneth Dunner, Jr. (UTDMACC) for providing valuable technical assistance with electron microscopy sample preparation and image acquisition.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. STRING gene ontology (GO) enrichment analysis of sEV protein features identified in two or more adenocarcinoma panel (LUAD and PDAC) cell lines with ≥ 8 SA. Top ten statistically enriched Biological Process, Molecular Function and Cellular Component GO terms.

Pathway ID	Pathway Description	Observed Gene Count	False Discovery Rate
A. GO enrichment, Biological Process: LUAD and PDAC Cell Line sEVs.			
GO:0016192	vesicle-mediated transport	476	6.28×10^{-84}
GO:0051234	establishment of localization	790	6.73×10^{-73}
GO:0006810	transport	769	1.52×10^{-70}
GO:0051179	localization	896	2.09×10^{-70}
GO:0009987	cellular process	1735	5.45×10^{-70}
GO:0071840	cellular component organization or biogenesis	903	9.03×10^{-69}
GO:0016043	cellular component organization	871	1.01×10^{-64}
GO:0006887	exocytosis	272	5.14×10^{-61}
GO:0045055	regulated exocytosis	251	3.56×10^{-58}
GO:0032940	secretion by cell	293	2.31×10^{-55}
B. GO enrichment, Molecular Function: LUAD and PDAC Cell Line sEVs.			
GO:0005515	protein binding	1092	5.70×10^{-88}
GO:0005488	binding	1517	5.88×10^{-65}
GO:1901265	nucleoside phosphate binding	464	6.80×10^{-55}
GO:0000166	nucleotide binding	463	1.08×10^{-54}
GO:0043168	anion binding	539	3.26×10^{-53}
GO:0097367	carbohydrate derivative binding	467	5.88×10^{-53}
GO:0036094	small molecule binding	503	1.40×10^{-51}
GO:0017076	purine nucleotide binding	420	5.50×10^{-51}
GO:0035639	purine ribonucleoside triphosphate binding	409	1.11×10^{-50}
GO:0032555	purine ribonucleotide binding	415	7.85×10^{-50}
C. GO enrichment, Cellular Component: LUAD and PDAC Cell Line sEVs.			
GO:0005829	cytosol	1058	2.42×10^{-154}
GO:0005737	cytoplasm	1639	1.56×10^{-151}
GO:0044444	cytoplasmic part	1489	7.89×10^{-150}
GO:0044424	intracellular part	1759	3.34×10^{-110}
GO:0005622	intracellular	1772	2.75×10^{-107}
GO:0044422	organelle part	1334	1.14×10^{-87}
GO:0044446	intracellular organelle part	1308	8.44×10^{-86}
GO:0032991	protein-containing complex	864	3.87×10^{-78}
GO:0044464	cell part	1836	7.39×10^{-76}
GO:0031982	vesicle	542	1.65×10^{-74}

Table A2. Ingenuity Upstream Regulator, Diseases and Functions and Pathway Analyses of sEV protein features identified in two or more adenocarcinoma panel (LUAD and PDAC) cell lines with SA ≥ 2 .

A. Top-ranked statistically enriched Upstream Regulators.			
Upstream Regulator	Molecule Type	p-Value	Mechanistic Network
TP53	transcription regulator	1.69×10^{-77}	1105 (22)
MYC	transcription regulator	1.13×10^{-70}	1312 (20)
MAPT	other	1.26×10^{-63}	1145 (18)
TGFB1	growth factor	5.91×10^{-56}	1169 (20)
MYCN	transcription regulator	1.07×10^{-50}	1063 (18)
KRAS	enzyme	2.93×10^{-49}	1119 (19)
B. Top ten statistically enriched IPA Diseases and Functions annotations.			
Category	Diseases or Functions Annotation	p-Value	#Molecules
Cancer, Organismal Injury & Abnormalities	Nonhematologic malignant neoplasm	2.76×10^{-133}	2431
Cancer, Organismal Injury & Abnormalities	Non-hematological solid tumor	4.36×10^{-133}	2436
Cancer, Organismal Injury & Abnormalities	Epithelial neoplasm	2.79×10^{-131}	2410
Cancer, Organismal Injury & Abnormalities	Carcinoma	4.72×10^{-131}	2407
Cancer, Organismal Injury & Abnormalities	Non-melanoma solid tumor	4.72×10^{-131}	2421
Cancer, Organismal Injury & Abnormalities	Tumorigenesis of tissue	4.25×10^{-130}	2414
Cancer, Organismal Injury & Abnormalities	Cancer	3.96×10^{-123}	2457
Cancer, Organismal Injury & Abnormalities	Malignant solid tumor	2.43×10^{-120}	2445
Cancer, Organismal Injury & Abnormalities	Extracranial solid tumor	9.08×10^{-119}	2442
Cancer, Organismal Injury & Abnormalities	Solid tumor	1.03×10^{-118}	2452
C. Top ten statistically enriched IPA Cancer Signaling Canonical Pathways.			
Ingenuity Canonical Pathway	p-Value	Ratio	#Molecules
Breast Cancer Regulation by Stathmin1	5.01×10^{-32}	0.440	88/200
ERK/MAPK Signaling	5.01×10^{-22}	0.378	73/193
Role of Tissue Factor in Cancer	6.31×10^{-19}	0.436	51/117
FAK Signaling	1.26×10^{-18}	0.474	45/95
PI3K/AKT Signaling	2.00×10^{-17}	0.356	62/174
Molecular Mechanisms of Cancer	1.58×10^{-16}	0.261	102/391
Telomerase Signaling	3.16×10^{-13}	0.383	41/107
Glioma Invasiveness Signaling	3.98×10^{-13}	0.446	37/74
BAG2 Signaling Pathway	1.58×10^{-12}	0.558	24/43
PTEN Signaling	3.98×10^{-11}	0.333	42/126

Table A3. GO enrichment analysis of 174 sEV protein features identified with ≥ 1.5 -fold higher log₂(SA) in LUAD vs PDAC panel cell line sEVs. Top ten statistically enriched GO terms.

Pathway ID	Pathway Description	Observed Gene Count	False Discovery Rate
A. GO enrichment analysis, Biological Process: ≥ 1.5 -fold higher log ₂ (SA) in LUAD vs. PDAC sEVs.			
GO:0016192	vesicle-mediated transport	57	2.74×10^{-18}
GO:0002443	leukocyte mediated immunity	35	2.79×10^{-16}
GO:0045055	regulated exocytosis	36	3.43×10^{-16}
GO:0043312	neutrophil degranulation	31	4.72×10^{-16}
GO:0002252	immune effector process	40	6.09×10^{-16}
GO:0032940	secretion by cell	40	1.07×10^{-15}
GO:0002366	leukocyte activation involved in immune response	33	1.30×10^{-15}
GO:0002274	myeloid leukocyte activation	32	1.40×10^{-15}
GO:0045321	leukocyte activation	36	1.36×10^{-13}
GO:0001775	cell activation	38	2.21×10^{-13}
B. GO enrichment analysis, Molecular Function: ≥ 1.5 -fold higher log ₂ (SA) in LUAD vs. PDAC sEVs.			
GO:0004298	threonine-type endopeptidase activity	15	3.10×10^{-20}
GO:0036094	small molecule binding	64	7.39×10^{-17}
GO:1901265	nucleoside phosphate binding	59	7.39×10^{-17}
GO:0003924	GTPase activity	25	1.80×10^{-16}
GO:0000166	nucleotide binding	58	2.40×10^{-16}
GO:0005525	GTP binding	26	3.67×10^{-15}
GO:0003824	catalytic activity	94	4.39×10^{-15}
GO:0097367	carbohydrate derivative binding	56	8.68×10^{-15}
GO:0016787	hydrolase activity	59	2.24×10^{-14}
GO:0032553	ribonucleotide binding	51	3.56×10^{-14}

Table A3. Cont.

Pathway ID	Pathway Description	Observed Gene Count	False Discovery Rate
C. GO enrichment analysis, Cellular Component: ≥ 1.5 -fold higher $\log_2(\text{SA})$ in LUAD vs. PDAC sEVs.			
GO:0005839	proteasome core complex	15	2.68×10^{-20}
GO:0005829	cytosol	93	1.48×10^{-17}
GO:0030141	secretory granule	39	3.32×10^{-17}
GO:0099503	secretory vesicle	41	4.69×10^{-17}
GO:0044433	cytoplasmic vesicle part	49	1.04×10^{-16}
GO:0000502	proteasome complex	16	1.15×10^{-16}
GO:0031410	cytoplasmic vesicle	60	1.18×10^{-16}
GO:0031982	vesicle	61	1.18×10^{-16}
GO:0044444	cytoplasmic part	126	1.92×10^{-16}
GO:0060205	cytoplasmic vesicle lumen	24	3.50×10^{-14}

Table A4. GO enrichment analysis of 115 sEV protein features identified with ≥ 1.5 -fold higher $\log_2(\text{SA})$ in PDAC vs LUAD panel cell line sEVs. Top ten statistically enriched GO terms.

Pathway ID	Pathway Description	Observed Gene Count	False Discovery Rate
A. GO enrichment analysis, Biological Process: ≥ 1.5 -fold higher $\log_2(\text{SA})$ in PDAC vs. LUAD sEVs.			
GO:0006401	RNA catabolic process	17	1.72×10^{-10}
GO:0006412	translation	20	1.72×10^{-10}
GO:0043604	amide biosynthetic process	22	1.79×10^{-10}
GO:0006413	translational initiation	14	1.82×10^{-10}
GO:0000184	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	13	3.02×10^{-10}
GO:0000956	nuclear-transcribed mRNA catabolic process	15	3.35×10^{-10}
GO:0006614	SRP-dependent cotranslational protein targeting to membrane	12	3.35×10^{-10}
GO:0051641	cellular localization	42	3.35×10^{-10}
GO:0006518	peptide metabolic process	21	4.88×10^{-10}
GO:0046907	intracellular transport	33	5.36×10^{-10}
B. GO enrichment analysis, Molecular Function: ≥ 1.5 -fold higher $\log_2(\text{SA})$ in PDAC vs. LUAD sEVs.			
GO:0031683	G-protein beta/gamma-subunit complex binding	8	1.23×10^{-9}
GO:0003735	structural constituent of ribosome	11	4.10×10^{-7}
GO:0005198	structural molecule activity	20	4.80×10^{-7}
GO:0061608	nuclear import signal receptor activity	5	1.89×10^{-6}
GO:0005048	signal sequence binding	7	2.27×10^{-6}
GO:0044877	protein-containing complex binding	22	3.48×10^{-6}
GO:0003723	RNA binding	20	7.75×10^{-6}
GO:0017076	purine nucleotide binding	30	1.16×10^{-5}
GO:0031821	G protein-coupled serotonin receptor binding	4	1.51×10^{-5}
GO:0035639	purine ribonucleoside triphosphate binding	29	1.51×10^{-5}
C. GO enrichment analysis, Cellular Component: ≥ 1.5 -fold higher $\log_2(\text{SA})$ in PDAC vs. LUAD sEVs.			
GO:0005829	cytosol	75	7.57×10^{-17}
GO:0032991	protein-containing complex	72	7.86×10^{-16}
GO:0044444	cytoplasmic part	95	7.79×10^{-13}
GO:0005737	cytoplasm	102	7.82×10^{-12}
GO:0005622	intracellular	111	3.33×10^{-10}
GO:0044424	intracellular part	110	3.78×10^{-10}
GO:0005834	heterotrimeric G-protein complex	8	2.04×10^{-9}
GO:1990904	ribonucleoprotein complex	22	3.95×10^{-8}
GO:0044464	cell part	113	6.82×10^{-7}
GO:0019897	extrinsic component of plasma membrane	10	8.30×10^{-7}

Table A5. Clinicopathological features of lung and pancreatic ductal adenocarcinoma case and control plasma cohorts.

Feature	Cases, LUAD & PDAC		LUAD Cases		PDAC Cases		Controls	
	N	%	N	%	N	%	N	%
N	21		15		6		21	
Sex								
Male	9	42.9%	5	33.3%	4	66.7%	9	42.9%
Female	12	57.1%	10	66.7%	2	33.3%	12	57.1%
Age, y	69 ± 12		70 ± 12		65 ± 3		65 ± 4	
Smoking								
Never			5	33.3%		-		-
Former			6	40.0%		-		-
Current			4	26.7%		-		-
Stage								
I	8	38.1%	8	53.3%				-
II	7	33.3%	7	46.7%				-
IV	6	28.6%			6	100%		-

Table A6. Ingenuity Upstream Regulator Analysis of sEV protein features with ROC AUC ≥ 0.7 classifier performance also identified in adenocarcinoma panel cell lines. Top-ranked statistically enriched gene, RNA and protein Upstream Regulators.

Upstream Regulator	Molecule Type	p-Value	Mechanistic Network
A. Lung adenocarcinoma (LUAD) plasma-derived sEVs			
AHR	ligand-dependent nuclear receptor	1.22×10^{-7}	
Vegf	group	1.14×10^{-6}	
EGF	growth factor	1.38×10^{-6}	13 (7)
IL6	cytokine	1.58×10^{-6}	13 (10)
IRF1	transcription regulator	1.93×10^{-6}	
STAT3	transcription regulator	3.45×10^{-6}	13 (8)
TAF4	transcription regulator	3.91×10^{-6}	
TP73	transcription regulator	5.19×10^{-6}	
B. Pancreatic ductal adenocarcinoma (PDAC) plasma-derived sEVs			
TNF	cytokine	2.45×10^{-29}	134 (13)
APP	other	3.50×10^{-25}	154 (18)
INFG	cytokine	5.92×10^{-24}	143 (14)
TGFB1	growth factor	9.62×10^{-23}	156 (19)
IL1B	cytokine	2.97×10^{-20}	134 (13)
KRAS	enzyme	1.63×10^{-19}	162 (22)
HGF	growth factor	2.26×10^{-18}	153 (19)
TP53	transcription regulator	2.96×10^{-18}	133 (16)

References

1. Tkach, M.; Kowal, J.; Théry, C. Why the need and how to approach the functional diversity of extracellular vesicles. *Phil. Trans. R. Soc. B* **2018**, *373*, e20160479. [[CrossRef](#)] [[PubMed](#)]
2. Théry, C.; Witwer, K.W.; Aikawa, E.; Alcaraz, M.J.; Anderson, J.D.; Andriantsitohaina, R.; Antoniou, A.; Arab, T.; Archer, F.; Atkin-Smith, G.K.; et al. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): A position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *J. Extracell. Vesicles* **2018**, *7*, e1535750. [[CrossRef](#)] [[PubMed](#)]
3. Witwer, K.W.; Théry, C. Extracellular vesicles or exosomes? On primacy, precision and popularity influencing a choice of nomenclature. *J. Extracell. Vesicles* **2019**, *8*, e1648167. [[CrossRef](#)] [[PubMed](#)]

4. Van Niel, G.; D'Angelo, G.; Raposo, G. Shedding light on the cell biology of extracellular vesicles. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 213–228. [[CrossRef](#)]
5. Jeppesen, D.K.; Fenix, A.M.; Franklin, J.L.; Higginbotham, J.N.; Zhang, Q.; Zimmerman, L.J.; Liebler, D.C.; Ping, J.; Liu, Q.; Evans, R.; et al. Reassessment of Exosome Composition. *Cell* **2019**, *177*, 428–445. [[CrossRef](#)] [[PubMed](#)]
6. Mathieu, M.; Martin-Jaular, L.; Lavieu, G.; Théry, C. Specificities of secretion and uptake of exosomes and other extracellular vesicles for cell-to-cell communication. *Nat. Cell Biol.* **2019**, *21*, 9–17. [[CrossRef](#)]
7. Daßler-Plenker, J.; Küttner, V.; Egeblad, M. Communication in tiny packages: Exosomes as means of tumor-stroma communication. *Biochim. Biophys. Acta BBA-Rev. Cancer* **2020**, *1873*, e188340. [[CrossRef](#)]
8. Tkach, M.; Théry, C. Communication by Extracellular Vesicles: Where We Are and Where We Need to Go. *Cell* **2016**, *164*, 1226–1232. [[CrossRef](#)]
9. Becker, A.; Thakur, B.K.; Weiss, J.M.; Kim, H.S.; Peinado, H.; Lyden, D. Extracellular Vesicles in Cancer: Cell-to-Cell Mediators of Metastasis. *Cancer Cell* **2016**, *30*, 836–848. [[CrossRef](#)]
10. Capello, M.; Vykoukal, J.V.; Katayama, H.; Bantis, L.E.; Wang, H.; Kundnani, D.L.; Aguilar-Bonavides, C.; Aguilar, M.; Tripathi, S.C.; Dhillon, D.S.; et al. Exosomes harbor B cell targets in pancreatic adenocarcinoma and exert decoy function against complement-mediated cytotoxicity. *Nat. Commun.* **2019**, *10*, e254. [[CrossRef](#)]
11. Murillo, O.D.; Thistlethwaite, W.; Rozowsky, J.; Subramanian, S.L.; Lucero, R.; Shah, N.; Jackson, A.R.; Srinivasan, S.; Chung, A.; Laurent, C.D.; et al. exRNA Atlas Analysis Reveals Distinct Extracellular RNA Cargo Types and Their Carriers Present across Human Biofluids. *Cell* **2019**, *177*, 463–477. [[CrossRef](#)] [[PubMed](#)]
12. Rontogianni, S.; Synadaki, E.; Li, B.; Liefwaard, M.C.; Lips, E.H.; Wesseling, J.; Wu, W.; Altelaar, M. Proteomic profiling of extracellular vesicles allows for human breast cancer subtyping. *Commun. Biol.* **2019**, *2*, 1–13. [[CrossRef](#)] [[PubMed](#)]
13. Jansen, F.H.; Krijgsveld, J.; van Rijswijk, A.; van den Bemd, G.-J.; van den Berg, M.S.; van Weerden, W.M.; Willemsen, R.; Dekker, L.J.; Luider, T.M.; Jenster, G. Exosomal Secretion of Cytoplasmic Prostate Cancer Xenograft-derived Proteins. *Mol. Cell. Proteom.* **2009**, *8*, 1192–1205. [[CrossRef](#)] [[PubMed](#)]
14. Minciocchi, V.R.; Zijlstra, A.; Rubin, M.A.; Vizio, D.D. Extracellular vesicles for liquid biopsy in prostate cancer: Where are we and where are we headed? *Prostate Cancer Prostatic Dis.* **2017**, *20*, 251–258. [[CrossRef](#)] [[PubMed](#)]
15. Willms, E.; Cabañas, C.; Mäger, I.; Wood, M.J.A.; Vader, P. Extracellular Vesicle Heterogeneity: Subpopulations, Isolation Techniques and Diverse Functions in Cancer Progression. *Front. Immunol.* **2018**, *9*, e738. [[CrossRef](#)]
16. Kosaka, N.; Yoshioka, Y.; Fujita, Y.; Ochiya, T. Versatile roles of extracellular vesicles in cancer. *J. Clin. Investig.* **2016**, *126*, 1163–1172. [[CrossRef](#)]
17. 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. Quantitative and stoichiometric analysis of the microRNA content of exosomes. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 14888–14893. [[CrossRef](#)]
18. Castillo, J.; Bernard, V.; San Lucas, F.A.; Allenson, K.; Capello, M.; Kim, D.U.; Gascoyne, P.; Mulu, F.C.; Stephens, B.M.; Huang, J.; et al. Surfaceome profiling enables isolation of cancer-specific exosomal cargo in liquid biopsies from pancreatic cancer patients. *Ann. Oncol.* **2018**, *29*, 223–229. [[CrossRef](#)]
19. Carvalho, A.S.; Baeta, H.; Silva, B.C.; Moraes, M.C.S.; Bodo, C.; Beck, H.C.; Rodriguez, M.S.; Saraswat, M.; Pandey, A.; Matthiesen, R. Extra-cellular vesicles carry proteome of cancer hallmarks. *Front. Biosci. Landmark Ed.* **2020**, *25*, 398–436.
20. Chen, R.; Xu, X.; Qian, Z.; Zhang, C.; Niu, Y.; Wang, Z.; Sun, J.; Zhang, X.; Yu, Y. The biological functions and clinical applications of exosomes in lung cancer. *Cell. Mol. Life Sci.* **2019**, *76*, 4613–4633. [[CrossRef](#)]
21. Sandfeld-Paulsen, B.; Aggerholm-Pedersen, N.; Bæk, R.; Jakobsen, K.R.; Meldgaard, P.; Folkersen, B.H.; Rasmussen, T.R.; Varming, K.; Jørgensen, M.M.; Sorensen, B.S. Exosomal proteins as prognostic biomarkers in non-small cell lung cancer. *Mol. Oncol.* **2016**, *10*, 1595–1602. [[CrossRef](#)] [[PubMed](#)]
22. Vykoukal, J.; Sun, N.; Aguilar-Bonavides, C.; Katayama, H.; Tanaka, I.; Fahrman, J.F.; Capello, M.; Fujimoto, J.; Aguilar, M.; Wistuba, I.I.; et al. Plasma-derived extracellular vesicle proteins as a source of biomarkers for lung adenocarcinoma. *Oncotarget* **2017**, *8*, 95466–95480. [[CrossRef](#)] [[PubMed](#)]
23. Webber, J.; Stone, T.C.; Katilius, E.; Smith, B.C.; Gordon, B.; Mason, M.D.; Tabi, Z.; Brewis, I.A.; Clayton, A. Proteomics Analysis of Cancer Exosomes Using a Novel Modified Aptamer-based Array (SOMAscan) Platform. *Mol. Cell. Proteom.* **2014**, *13*, 1050–1064. [[CrossRef](#)] [[PubMed](#)]

24. Théry, C.; Amigorena, S.; Raposo, G.; Clayton, A. Isolation and Characterization of Exosomes from Cell Culture Supernatants and Biological Fluids. In *Current Protocols in Cell Biology*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2001; ISBN 978-0-471-14303-1.
25. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [[CrossRef](#)]
26. Gold, L.; Ayers, D.; Bertino, J.; Bock, C.; Bock, A.; Brody, E.N.; Carter, J.; Dalby, A.B.; Eaton, B.E.; Fitzwater, T.; et al. Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery. *PLoS ONE* **2010**, *5*, e15004. [[CrossRef](#)]
27. Welton, J.L.; Brennan, P.; Gurney, M.; Webber, J.P.; Spary, L.K.; Carton, D.G.; Falcón-Pérez, J.M.; Walton, S.P.; Mason, M.D.; Tabi, Z.; et al. Proteomics analysis of vesicles isolated from plasma and urine of prostate cancer patients using a multiplex, aptamer-based protein array. *J. Extracell. Vesicles* **2016**, *5*, e31209. [[CrossRef](#)]
28. Binder, J.X.; Pletscher-Frankild, S.; Tsafo, K.; Stolte, C.; O'Donoghue, S.I.; Schneider, R.; Jensen, L.J. COMPARTMENTS: Unification and visualization of protein subcellular localization evidence. *Database* **2014**, *2014*. [[CrossRef](#)]
29. Jiao, Y.-J.; Jin, D.-D.; Jiang, F.; Liu, J.-X.; Qu, L.-S.; Ni, W.-K.; Liu, Z.-X.; Lu, C.-H.; Ni, R.-Z.; Zhu, J.; et al. Characterization and proteomic profiling of pancreatic cancer-derived serum exosomes. *J. Cell. Biochem.* **2019**, *120*, 988–999. [[CrossRef](#)]
30. Wang, N.; Song, X.; Liu, L.; Niu, L.; Wang, X.; Song, X.; Xie, L. Circulating exosomes contain protein biomarkers of metastatic non-small-cell lung cancer. *Cancer Sci.* **2018**, *109*, 1701–1709. [[CrossRef](#)]
31. Safi, A.; Sadmi, M.; Martinet, N.; Menard, O.; Vaillant, P.; Gallati, H.; Hosang, M.; Martinet, Y. Presence of Elevated Levels of Platelet-Derived Growth Factor (PDGF) in Lung Adenocarcinoma Pleural Effusions. *Chest* **1992**, *102*, 204–207. [[CrossRef](#)]
32. Lunardi, S.; Muschel, R.J.; Brunner, T.B. The stromal compartments in pancreatic cancer: Are there any therapeutic targets? *Cancer Lett.* **2014**, *343*, 147–155. [[CrossRef](#)] [[PubMed](#)]
33. Hu, C.; Yang, J.; Su, H.-Y.; Waldron, R.T.; Zhi, M.; Li, L.; Xia, Q.; Pandol, S.J.; Lugea, A. Yes-Associated Protein 1 Plays Major Roles in Pancreatic Stellate Cell Activation and Fibroinflammatory Responses. *Front. Physiol.* **2019**, *10*, e1467. [[CrossRef](#)] [[PubMed](#)]
34. Karaman, S.; Detmar, M. Mechanisms of lymphatic metastasis. *J. Clin. Investig.* **2014**, *124*, 922–928. [[CrossRef](#)] [[PubMed](#)]
35. Stacker, S.A.; Williams, S.P.; Karnezis, T.; Shayan, R.; Fox, S.B.; Achen, M.G. Lymphangiogenesis and lymphatic vessel remodelling in cancer. *Nat. Rev. Cancer* **2014**, *14*, 159–172. [[CrossRef](#)] [[PubMed](#)]
36. Ma, Q.; Dieterich, L.C.; Ikenberg, K.; Bachmann, S.B.; Mangana, J.; Proulx, S.T.; Amann, V.C.; Levesque, M.P.; Dummer, R.; Baluk, P.; et al. Unexpected contribution of lymphatic vessels to promotion of distant metastatic tumor spread. *Sci. Adv.* **2018**, *4*, eaat4758. [[CrossRef](#)] [[PubMed](#)]
37. Popeda, M.; Stokowy, T.; Bednarz-Knoll, N.; Jurek, A.; Niemira, M.; Bielska, A.; Kretowski, A.; Kalinowski, L.; Szade, J.; Markiewicz, A.; et al. NF-kappa B Signaling-Related Signatures Are Connected with the Mesenchymal Phenotype of Circulating Tumor Cells in Non-Metastatic Breast Cancer. *Cancers* **2019**, *11*, e1961. [[CrossRef](#)]
38. Shi, K.; Zhang, J.; Zhao, R.; Yang, L.; Guo, D. PSMD7 downregulation induces apoptosis and suppresses tumorigenesis of esophageal squamous cell carcinoma via the mTOR/p70S6K pathway. *FEBS Open Bio.* **2018**, *8*, 533–543. [[CrossRef](#)]
39. Hegde, A.N.; Upadhyay, S.C. The ubiquitin–proteasome pathway in health and disease of the nervous system. *Trends Neurosci.* **2007**, *30*, 587–595. [[CrossRef](#)]
40. Böse, K.; Nischt, R.; Page, A.; Bader, B.L.; Paulsson, M.; Smyth, N. Loss of Nidogen-1 and -2 Results in Syndactyly and Changes in Limb Development. *J. Biol. Chem.* **2006**, *281*, 39620–39629. [[CrossRef](#)] [[PubMed](#)]
41. Klink, D.J.; Torang, A. An unsupervised feature extraction and selection strategy for identifying epithelial-mesenchymal transition state metrics in breast cancer and melanoma. *BioRxiv* **2019**. [[CrossRef](#)]
42. Chai, A.W.Y.; Cheung, A.K.L.; Dai, W.; Ko, J.M.Y.; Lee, N.P.Y.; Chan, K.T.; Law, S.Y.-K.; Lung, M.L. Elevated levels of serum nidogen-2 in esophageal squamous cell carcinoma. *Cancer Biomark.* **2018**, *21*, 583–590. [[CrossRef](#)] [[PubMed](#)]

43. Chai, A.W.Y.; Cheung, A.K.L.; Dai, W.; Ko, J.M.Y.; Ip, J.C.Y.; Chan, K.W.; Kwong, D.L.-W.; Ng, W.T.; Lee, A.W.M.; Ngan, R.K.C.; et al. Metastasis-suppressing NID2, an epigenetically-silenced gene, in the pathogenesis of nasopharyngeal carcinoma and esophageal squamous cell carcinoma. *Oncotarget* **2016**, *7*, 78859–78871. [[CrossRef](#)] [[PubMed](#)]
44. Baharudin, R.; Tieng, F.Y.F.; Lee, L.-H.; Ab Mutalib, N.S. Epigenetics of SFRP1: The Dual Roles in Human Cancers. *Cancers* **2020**, *12*, e445. [[CrossRef](#)] [[PubMed](#)]
45. Liang, C.-J.; Wang, Z.-W.; Chang, Y.-W.; Lee, K.-C.; Lin, W.-H.; Lee, J.-L. SFRPs Are Biphasic Modulators of Wnt-Signaling-Elicited Cancer Stem Cell Properties beyond Extracellular Control. *Cell Rep.* **2019**, *28*, 1511–1525. [[CrossRef](#)] [[PubMed](#)]
46. Sade-Feldman, M.; Jiao, Y.J.; Chen, J.H.; Rooney, M.S.; Barzily-Rokni, M.; Eliane, J.-P.; Bjorgaard, S.L.; Hammond, M.R.; Vitzthum, H.; Blackmon, S.M.; et al. Resistance to checkpoint blockade therapy through inactivation of antigen presentation. *Nat. Commun.* **2017**, *8*, 1–11. [[CrossRef](#)] [[PubMed](#)]
47. Zeng, F.; Morelli, A.E. Extracellular vesicle-mediated MHC cross-dressing in immune homeostasis, transplantation, infectious diseases and cancer. *Semin. Immunopathol.* **2018**, *40*, 477–490. [[CrossRef](#)] [[PubMed](#)]
48. Pereira, C.; Gimenez-Xavier, P.; Pros, E.; Pajares, M.J.; Moro, M.; Gomez, A.; Navarro, A.; Condom, E.; Moran, S.; Gomez-Lopez, G.; et al. Genomic Profiling of Patient-Derived Xenografts for Lung Cancer Identifies B2M Inactivation Impairing Immunorecognition. *Clin. Cancer Res.* **2017**, *23*, 3203–3213. [[CrossRef](#)]
49. Cohen, J.D.; Li, L.; Wang, Y.; Thoburn, C.; Afsari, B.; Danilova, L.; Douville, C.; Javed, A.A.; Wong, F.; Mattox, A.; et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science* **2018**, *359*, 926–930. [[CrossRef](#)]
50. Kelley, S.O.; Pantel, K. A New Era in Liquid Biopsy: From Genotype to Phenotype. *Clin. Chem.* **2020**, *66*, 89–96. [[CrossRef](#)]
51. Fahrman, J.F.; Bantis, L.E.; Capello, M.; Scelo, G.; Dennison, J.B.; Patel, N.; Murage, E.; Vykoukal, J.; Kundnani, D.L.; Foretova, L.; et al. A Plasma-Derived Protein-Metabolite Multiplexed Panel for Early-Stage Pancreatic Cancer. *JNCI J. Natl. Cancer Inst.* **2019**, *111*, 372–379. [[CrossRef](#)]
52. Elazezy, M.; Joosse, S.A. Techniques of using circulating tumor DNA as a liquid biopsy component in cancer management. *Comput. Struct. Biotechnol. J.* **2018**, *16*, 370–378. [[CrossRef](#)] [[PubMed](#)]
53. Capello, M.; Bantis, L.E.; Scelo, G.; Zhao, Y.; Li, P.; Dhillon, D.S.; Patel, N.J.; Kundnani, D.L.; Wang, H.; Abbruzzese, J.L.; et al. Sequential Validation of Blood-Based Protein Biomarker Candidates for Early-Stage Pancreatic Cancer. *JNCI J. Natl. Cancer Inst.* **2017**, *109*. [[CrossRef](#)] [[PubMed](#)]
54. Guida, F.; Sun, N.; Bantis, L.E.; Muller, D.C.; Li, P.; Taguchi, A.; Dhillon, D.; Kundnani, D.L.; Patel, N.J.; Yan, Q.; et al. Assessment of Lung Cancer Risk on the Basis of a Biomarker Panel of Circulating Proteins. *JAMA Oncol.* **2018**, *4*, e182078. [[PubMed](#)]
55. Wang, H.; Hanash, S. Intact-Protein Analysis System for Discovery of Serum-Based Disease Biomarkers. In *Serum/Plasma Proteomics*; Simpson, R.J., Greening, D.W., Eds.; Humana Press: Totowa, NJ, USA, 2011; Volume 728, pp. 69–85. ISBN 978-1-61779-067-6.

