IDENTIFICATION AND UTILITY OF DNA IN EXOSOMES

Paul Kurywchak

Follow this and additional works at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations

Part of the Cancer Biology Commons, Laboratory and Basic Science Research Commons, and the Molecular Genetics Commons

Recommended Citation
https://digitalcommons.library.tmc.edu/utgsbs_dissertations/841

This Dissertation (PhD) is brought to you for free and open access by the Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in UT GSBS Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact laurel.sanders@library.tmc.edu.
IDENTIFICATION AND UTILITY OF DNA IN EXOSOMES

A

DISSERTATION

Presented to the Faculty of

The University of Texas

MD Anderson Cancer Center UTHealth

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILISOPHY

by

Paul A. Kurywchak, B.S.
Houston, Texas

May 2018
Dedication

I’d first like to dedicate this dissertation to my parents, Dawn and Steve, who are responsible for shaping me into the person I am today. If it wasn’t for their commitment to my development and growth, none of this would have been possible. Regardless of what I’ve pursued in my life, they’ve always had confidence in me and trusted that I would do well. To my mom, thank you for being the person I know I can always go to for support, especially if there’s something I can’t find or there’s a last-minute project to finish. To my dad, thank you for being my role model, life coach, and never letting me buy something from an infomercial, no matter how badly I thought we needed it.

Second, I’d like to dedicate this dissertation to my better half, Brittany. Her encouragement and patience to put up with my persistently late nights is something I’m very thankful for. She has been my rock through the most strenuous periods of graduate school and I’m so glad I was lucky enough to find her. I will always cherish our many adventures at Memorial Park with the pups, that were so effective at clearing my mind, and I can’t wait to see what the future holds for us.

Third, I’d like to dedicate this dissertation to the rest of my family and my closest friends, who I also consider my family. Even though they were far away, I knew I could always count on them to be there for me and remind me that they always will be. My trips back home and elsewhere with them never failed to recharge my batteries and gave me the motivation I needed to keep working hard.

Lastly, I’d like to dedicate this dissertation to the memory of those that have valiantly battled cancer. Cancer has taken many people close to me, and I often reflect on their strength as a reminder of why research is so vitally important, to take time to appreciate the things that truly matter in life, and to never stop fighting to achieve my goals.
Acknowledgements

First and foremost, I must thank my advisor, Dr. Raghu Kalluri for accepting me into the Kalluri lab family and supporting me every step of the way. He has been an exceptional mentor in so many ways over the past five years, both inside and outside the laboratory. He gave me unrestricted freedom to explore and choose research questions that excite me, was always willing to engage in thought-provoking discussions to focus my efforts and stimulate new ideas, and knew precisely the right times to give me a push, even if it had been weeks since our last extended interaction. Regardless of where life takes me, I will always value my relationship with Raghu and consider him a primary contact for advice and support.

Second, thank you to Dr. Valerie LeBleu for showcasing her brilliance, fearlessness, and passion for science each and every day. I will likely never meet someone again that works as hard or as smart as she does. Not only was she a tremendous help to my development as a scientist, she was also an incredible role model for how to approach work of any kind. She taught me to approach my experiments with focus and purpose, and to never get discouraged when results don’t make sense or mistakes are made. Even while handling the workload of multiple people singlehandedly, she always found the time to answer my questions, help me troubleshoot experiments, and proofread my manuscripts.

Third, I need to thank Dr. Olga Volpert for her contributions to transforming my research from a collection of observations into manuscript drafts. She’s fantastic at processing and interpreting data on the first pass and understanding how it should fit in relation to other experiments. Furthermore, I greatly appreciate the patience she had with me and willingness to always step in and help out if she saw that I needed it.

Fourth, I have to acknowledge the support given to me by my fellow Kalluri lab members. Like Dr. LeBleu, Kalluri lab scientists have a work ethic that is infectious, and there are very few hours of the day someone from the lab is not working. In addition, the collective
expertise in the Kalluri lab are unparalleled, and I was fortunate enough to benefit from that over the years. Specifically, Dr. Pedro Correa de Sampaio, Dr. Sara Lovisa, Dr. Sushrut Kamerkar, and Dr. Eliot Fletcher were always people I knew I could count on to give me advice, or just lend an ear when I needed to decompress. Dr. Christoph Kahlert, Dr. Sonia Melo, and Kerri Wolf were also people I worked with closely and made this work possible. Furthermore, the lab wouldn’t be where it is if it weren’t for the amazing work from our technicians, lab manager, or administrative support staff.

Fifth, I must thank my collaborators outside of the laboratory that made my work possible. Chia-Chin Wu was a great colleague during my time as a graduate student and helped immensely with teaching me about how to interpret sequencing data. Natalia Maltsev and her team at the University of Chicago were also essential to the success of our sequencing efforts. They provided tremendous resources and support to make sure we got as much out of the data we could, and make it as impactful as possible.

Sixth, I am grateful for spending my last two years of my Ph.D. as a NIH T32 Training Grant fellow. My experiences in the program were very enriching and I thank Dr. Mien-Chie Hung for serving as director, and the rest of the associated staff for the support.

Lastly, I’d like to show my gratitude to my advisory committee, comprised of Dr. Valerie LeBleu, Dr. Florencia McAllister, Dr. Craig Logsdon, and Dr. Colin Dinney. They have been nothing but supportive, and their fresh perspectives on my research helped immensely. Not only did they offer their honest opinions about how to strengthen my data, but they also provided valuable insight and advice for how I should be preparing for my future and where they thought my strengths were best aligned, and is something I valued greatly.
Table of Contents

TITeL PAGE ............................................................................................................................. I
DEDICATION ........................................................................................................................... II
ACKNOWLEDGEMENTS ....................................................................................................... III
LIST OF FIGURES ................................................................................................................. VII
LIST OF TABLES .................................................................................................................. IX
ABSTRACT ........................................................................................................................... X

CHAPTER 1 ............................................................................................................................ 1

BACKGROUND AND SIGNIFICANCE ................................................................................. 1
FEATURES OF PANCREATIC AND BLADDER CANCERS AND EXISTING THERANOSTIC LIMITATIONS ........................................................................................................... 2

Pan-cancer statistics ............................................................................................................. 2
Pancreatic cancer statistics and treatment ............................................................................. 3
Pancreatic cancer risk screening .......................................................................................... 5
Biomarkers of pancreatic cancer .......................................................................................... 6
Genetics of pancreatic cancer ............................................................................................... 8
Bladder cancer statistics & treatment .................................................................................... 9
Bladder cancer risk screening .............................................................................................. 12
Biomarkers of bladder cancer .............................................................................................. 13
Genetics of bladder cancer ................................................................................................. 14

EXOSOMES AND THEIR DNA CONTENT ........................................................................... 15

Discovery of exosomes .......................................................................................................... 15
Biogenesis of exosomes ......................................................................................................... 17
Exosomal cargo and protein/RNA biomarker utility .............................................................. 19
Functions of exosomes in cancer .......................................................................................... 22
DNA in exosomes ................................................................................................................ 24
exosomalDNA as a biomarker ............................................................................................... 24
exosomalDNA in cancer ....................................................................................................... 26

DISSERTATION GOALS AND MAJOR FINDINGS .............................................................. 30

CHAPTER 2 ............................................................................................................................ 32

MATERIALS AND METHODS ......................................................................................... 32

EXPERIMENTS WITH PDAC CELL LINE AND SERUM/PLASMA SAMPLES .............................. 33

Cell culture .......................................................................................................................... 33
Exosome isolation ................................................................................................................. 33
Sucrose gradient ultracentrifugation .................................................................................... 35
Western blot ........................................................................................................................ 35
Transmission electron microscopy (TEM) ........................................................................... 36
DNase I treatment and DNA extraction .............................................................................. 36
DiO labeling of exosomes ..................................................................................................... 38
DNA library preparation and WGS ....................................................................................... 38
Alignment and variant analysis ............................................................................................. 39
Copy number analysis .......................................................................................................... 40
Healthy individual and PDAC patient serum samples (66) ................................................ 40

Digital PCR ........................................................................................................................... 41
Nuclei isolation & NV collection ........................................................................................... 42
PKH26 labeling of exosomes ............................................................................................... 43
Fluorescence microscopy ...................................................................................................... 43
PCR & Sanger sequencing .................................................................................................... 45
RT-qPCR ................................................................................................................................... 46
Experimental animals ........................................................................................................ 47
Statistical analysis ............................................................................................................ 47
MATERIALS AND METHODS FOR EXPERIMENTS WITH BC SAMPLES .................... 48
Patient information and specimen collection .............................................................. 48
Isolation of exosomes from urine and serum ................................................................. 48
Nanoparticle tracking analysis ....................................................................................... 50
Immunogold labeling and transmission electron microscopy (TEM) ......................... 50
Western blot analysis .................................................................................................... 50
DNA extraction ............................................................................................................... 51
Targeted PCR and Sanger sequencing .......................................................................... 51
DNA library preparation and whole exome sequencing .............................................. 52
Quality control, sequence alignment, and variant calling ........................................... 53
Enrichment and comparative analysis ........................................................................ 55
Statistical analysis ........................................................................................................ 56

CHAPTER 3 .................................................................................................................. 57
CHARACTERIZATION OF EXOSOMES AND EXODNA .............................................. 57
Characterization of PDAC cell line exosomes and exoDNA ....................................... 58
PDAC cell line exosomes share common features ....................................................... 58
Large fragments of dsDNA are predominantly enclosed within exosomes ................. 60
Whole genome sequencing reveals that exoDNA is reflective of cell DNA ............... 64
Characterization of healthy donor and PDAC patient serum exosomes and exoDNA ... 68
Characterization of urine and serum exosomes from healthy donors and BC patients 69
Urine exosomes have some distinct features from serum exosomes ........................ 69
BC patient urine samples contained elevated exosome and exoDNA concentrations .... 70
Urine and serum exosomes contain large fragments of genomic DNA ........................ 71

CHAPTER 4 .................................................................................................................. 73
UTILITY OF EXODNA AS A CANCER BIOMARKER .................................................. 73
Detection of PDAC-associated DNA mutations in human serum exosomes (66) ....... 74
Establishing thresholds for mutation detection by digital PCR (dPCR) ....................... 74
Novel and unique somatic variants in the DNA from urinary exosomes of BC patients 79
Urine exoDNA is suitable for PCR-amplification & Sanger sequencing .................... 79
WES of exoDNA can be negatively impacted by non-uniform whole-genome amplification .............. 80
Variant analysis revealed superior mutation capture in urinary exoDNA samples ....... 81
BC patient panels have a high proportion of mutations in untranslated regions (UTRs) 85
Gene network panels reconstruction through identification of driver gene mutations .... 86

CHAPTER 5 .................................................................................................................. 89
INVESTIGATION OF THE ORIGIN AND FUNCTION OF EXODNA ............................... 89
Assessing nuclei as sources of exosome-like particles ............................................... 90
Isolated nuclei produce exosome-like vesicles that contain genomic DNA ................. 90
Exosomes deliver transcriptionally-active gene fragments to recipient cells .............. 94

CHAPTER 6 .................................................................................................................. 98
DISCUSSION AND FUTURE DIRECTIONS ................................................................. 98
Discussion ...................................................................................................................... 99
Confirmation of the intraluminal localization of genome-spanning exoDNA ............... 99
exoDNA may originate from nucleus-derived vesicles and can be horizontally transferred 103
Conclusions ................................................................................................................ 106
Future Directions ....................................................................................................... 107

BIBLIOGRAPHY .......................................................................................................... 108
List of Figures

Figure 1. Cancer mortality statistics in the U.S...............................................................2
Figure 2. Pancreatic cancer statistics by race.................................................................4
Figure 3. The most commonly mutated genes in PDAC................................................8
Figure 4. Bladder cancer statistics by race.................................................................10
Figure 5. Release of exosomes containing gold-labeled transferrin.................................17
Figure 6. Schematic of exosome biogenesis..............................................................19
Figure 7. The advantage of circulating exosomes for capturing tumor mutations..........25
Figure 8. Current proposed mechanism for packaging of nucleic acids into exosomes......27
Figure 9. Routes of exogenous dsDNA delivery to other cells...................................28
Figure 10. Schematic of exosome and exoDNA isolation from cell culture supernatants......34
Figure 11. Quality control experiments to test DNase I and Qubit 3.0 performance........37
Figure 12. Schematic of the exosome isolation procedure for human urine samples........49
Figure 13. Bioinformatics workflow for interpreting whole exome sequencing data........55
Figure 14. Characterization of PDAC cell line exosomes...........................................58
Figure 15. PDAC cell line exosomes contain large intraluminal DNA fragments..............61
Figure 16. Visualization of DiO-stained exosomes in sucrose gradient preparations........63
Figure 17. Copy number analysis reveals that exoDNA spans all chromosomes..........64
Figure 18. Representative concentration and size distribution of exosomes from healthy donors (top) and PDAC patients (bottom)............................................................................68
Figure 19. Healthy donor (left) and patient (right) serum exoDNA concentrations........69
Figure 20. Characterization of exosome from urine and serum of healthy donors and BC patients.......................................................................................................................70
Figure 21. Comparative analysis of DNA from exosomes and matched tumor tissues........72
Figure 22. Definition of digital PCR parameters using cell line-derived exoDNA..........75
Figure 23. Representative 2D intensity scatter plot of wild-type and mutant amplicon........76

Figure 24. $KRAS^{G12D}$ and $TP53^{R273H}$ detection in human serum samples..................78

Figure 25. PCR & Sanger sequencing of BC hotspots in patient samples.................................80

Figure 26. Overlap in somatic variants identified in the urinary exoDNA, serum exoDNA and tumors of individual patients.................................................................84

Figure 27. Common network for driver genes shared between at least four patients..............86

Figure 28. Networks and driver genes in the urinary exoDNA and matched tumor DNA of patient 2..............................................................................................................87

Figure 29. Cell nuclei shed exosome-like vesicles in $ex \; vivo$ conditions.........................91

Figure 30. Characterization of NV DNA..............................................................................93

Figure 31. ExoDNA is taken up by cells $in \; vitro$ and $in \; vivo$ and can be utilized as a template for mRNA synthesis.................................................................95
List of Tables

Table 1. Pooled performance of commercially-available tests for BC surveillance………13
Table 2. Total reads and coverage of cell line WGS samples………………………….39
Table 3. Antibodies used for BC experiments…………………………………………….50
Table 4. Primer sets used for PCR and Sanger sequencing…………………………….52
Table 5. Whole exome sequencing total reads and coverage data……………………..53
Table 6. Total somatic variants, concordance, and contamination data………………53
Table 7. Total shared & unique somatic variants in driver genes in tumor samples……..81
Table 8. Top common somatic variants in driver genes in tumor samples……………….82
Table 9. Most commonly mutated driver genes in tumor samples…………………….82
Table 10. Top driver variants unique to urine samples………………………………….83
Table 11. Mutated miRNA binding sites in driver genes……………………………….85
Cancer-associated mortality has been declining for two decades but remains a significant public health problem, especially when patients initially present with advanced disease. Early detection methods have improved survival rates but remain unavailable for a majority of cancers due to a lack of sensitive biomarkers or numerous limitations associated with current diagnosis strategies. Approaches to develop “liquid biopsies” by detecting tumor cells or DNA in the blood have led to several breakthroughs and create the potential for non-invasive, routine assessment of diseases status. However, these biomarkers are rare and currently difficult to isolate, especially in the early stages of disease. Exosomes on the other hand, are phospholipid nanovesicles that reflect the molecular contents of their cell of origin and are abundantly present in bodily fluids. Due to this, exosomes have attracted considerable interest as a source for theranostic biomarkers, and may also be important participants of tumor progression. Here, we characterize exosomes and their DNA content (exoDNA) from human cell lines and bodily fluids to assess their utility as circulating biomarkers of disease. In addition, we explore the origins of exoDNA, as well as its delivery and transcriptional capacity in other cells. We found that exoDNA from cancer cell lines and serum and urine of patients with bladder cancer is encapsulated within exosomes and spans the entire genome. We also discovered unique somatic mutations in serum and urine exosomes that are absent in tumor tissue, and propose the use of urine exosomes for non-invasive biomarker detection in bladder cancer. Lastly, our results indicate that DNA-containing exosomes may be partly comprised of vesicles that evolve from the nuclear membrane and horizontally transfer DNA to recipient cells.
Chapter 1

Background and Significance
Features of pancreatic and bladder cancers and their existing theranostic limitations

Pan-cancer statistics

Cancer mortality rates in the United States were at their peak in 1991 at 215.1 deaths per 100,000 and have since fallen to 161.2 deaths per 100,000 in 2014, largely due to reductions in smoking rates and advances in screening and treatment strategies (Fig. 1A) (1). However, cancer still ranks only behind heart disease in overall causes of death, and is the leading cause of death in women aged 40-79 and men aged 45-79. Nearly 40% of all Americans have a lifetime probability of being diagnosed with invasive cancer, with an average of over 4,600 new diagnoses per day. In particular, men have a 20% higher incidence and 40% higher cancer-associated mortality than woman (1). Reductions in cancer incidence often require large shifts in wide-scale population behavior, such as decreased smoking rates, and until other healthy lifestyle changes are adopted en masse, these high rates of incidence are likely to persist. It’s estimated that over two million cancer-related deaths have been prevented in the United States since the cancer mortality peak in 1991, but further improvements for detection and treatment are still vitally needed, especially in cancers that have seen increases in annual mortality such as in the pancreas or liver (Fig. 1B) (1).

License number: 4304311331149.

Pancreatic cancer statistics and treatment

Advances in cancer detection and therapy as a whole have resulted in 40% or more fewer mortalities in the four major cancers (lung, breast, prostate, and colorectum) (1). However, pancreas cancer (PC) rates have not been reduced in the slightest despite improvements in detection and short-term survival due to therapy. 80% of PCs are classified as
pancreatic ductal adenocarcinomas (PDAC) and are known to be the deadliest of all cancers. Although relatively rare (ranked 12th in incidence), PC currently accounts for the 4th most deaths and is expected to move to 2nd by 2030 (2). As CT scanners became more available in the 1980s, it became possible to diagnose PC in the early stages when the disease was still localized and resectable by surgery (3). Even today though, less than 20% of patients are eligible for surgery at the time of diagnosis because 50-60% of them present with metastatic disease (Fig. 2A) (2). This is often the case because early symptoms are rarely observed, or are unspecific.


to PC like nausea, weight loss, back pain, and irregular bowel movements when they are observed (2). 5-year survival is extended to 17% for the patients that receive surgery, but has not improved in 20 years despite advancements in therapeutic strategies, and is usually due to liver metastases and recurrence (3-5). Concerns have also been raised about inducing
dissemination of tumor cells during surgery (6), but given the poor prognosis of PC in general, surgery is still likely to be the best path to potential cures for PC patients.

Gemcitabine was introduced in the late 1990s, which was reported to improve clinical response rates in patients with advanced PDAC to 23.8% from 4.8% when patients were given fluorouracil (5-FU). However, median survival was modestly extended from 4.4 months on 5-FU to 5.7 months on gemcitabine (7), but remains a first-line therapy for all patients with PDAC. New targeted therapies like elotinib and chemotherapy combinations like FOLFIRINOX (5-FU, Leucovorin, Oxaliplatin, and Irinotecan) and ABRAXANE (Nab-paclitaxel and gemcitabine) have since been established as first-line therapies, but still have only resulted in an overall 5-year survival increase of 5% to 8% in the last 20 years (Fig. 2B) (1-3, 8). Taken together, PDAC remains notoriously chemoresistant and it is obvious that much progress still needs to be made in the clinical management of PDAC.

Pancreatic cancer risk screening

Due to the low incidence of PC, it would not be ideal to do population-based screening unless a biomarker test was highly specific, affordable, and minimally invasive. There are currently surveillance programs available through clinical trials for high risk patients, but are few in number and limited to specialized cancer centers (9). The intent is to reduce PC mortality through treatment of potentially precancerous lesions such as intraepithelial neoplasms, papillary mucinous neoplasms, diabetes, or chronic pancreatitis. However, the underlying biology of these pathologies and the conditions that determine if they will undergo malignant transformation remain somewhat unclear, and PC can also be misdiagnosed as a precancerous lesion. Results from a single institution, retrospective study of 1656 patients in Asia determined that risk of developing PC was significantly higher in older patients at the onset of chronic
pancreatitis and 12-fold higher risk was observed in patients with a >60 pack-year smoking history, so these factors may be used as baseline features for determining inclusion in a surveillance clinical trial for progression to PC (10). Familial PC or other hereditary syndromes are currently the best predictors of PC onset, with risk increasing as more family members are affected (9). A study of 175 consecutive patients with PDAC at Memorial Sloan Kettering Cancer Center estimated that 5-10% of PDAC cases are familial (11). Although some factors are known to increase risk of PC, the approach of how to effectively treat these early lesions is still not clear. Unlike high-grade dysplasia in Barrett’s esophagus or germline BRCA mutations in breast or ovarian cancers where preemptive surgery can dramatically reduce cancer risk, PC does not have clear indications that suggests performing prophylactic removal of the pancreas would be worth the risk (9). Therefore, in addition to identifying specific biomarkers for early diagnosis, PC will likely only see significant improvements in survival when combined with novel treatment strategies, likely involving specific targeting of known driver mutations and/or improvements to immune checkpoint-based methods.

**Biomarkers of pancreatic cancer**

Carbohydrate antigen 19-9 (CA 19-9) is the most commonly used biomarker for PC, with a pooled sensitivity across multiple studies of 73% and specificity of 83% for detecting recurrent disease (12). Carcinoembryonic antigen is also a biomarker used in PC but has poorer performance than CA 19-9, with a sensitivity of 50% and a specificity of 65%, and has largely been replaced with CA 19-9. Both of these biomarkers are only recommended for use as recurrence biomarkers and not for initial diagnosis. CA 19-9 has been reported to be able to predict recurrence one to nine months before detection using imaging or other clinical indicators. However, in addition to the limited sensitivity and specificity and applicability to
recurrent disease only, CA 19-9 is limited due to its dependence on Lewis blood group antigens, which 5-10% of patients do not have (12, 13). Numerous preclinical biomarkers have been developed to replace CA 19-9, but most don’t proceed past the discovery phase of development due to the extensive validation and associated investment required to move a candidate biomarker in clinical trials (13). Some serum biomarkers have reached the last phases of development though and shown improvement over CA 19-9 alone. For example, Capello et al. recently showed in three independent validation datasets that combining CA 19-9 with TIMP1 and LRG1 in an enzyme-linked immunosorbent assay (ELISA) discriminated early-stage PDAC versus healthy controls with an average sensitivity of 76% and a specificity of 95% (14). In addition, work from Melo et al. showed that the cell surface proteoglycan glypican-1 (GPC-1) was specifically enriched on circulating exosomes and could distinguish patients with benign diseases and those with early stage PDAC with 100% sensitivity and specificity. GPC-1+ exosome concentrations in the blood correlated with tumor burden and was a negative predictor for patient survival (15). Similarly, Hu et al. also recently showed that serum exosomal GPC-1 mRNA is also a biomarker with perfect sensitivity and specificity for PC (16). GPC-1 was confirmed as a biomarker for PC in multiple validation cohorts in both of these studies, and is currently being evaluated by other researchers to further test its validity at the mRNA and protein level. Biomarkers with such perfect accuracy are unprecedented in the cancer field, but GPC-1 could be the breakthrough that’s so desperately needed for PC, along with understanding the underlying genetics. Until more is known about these promising new candidates, CA 19-9 and routine CT or PET-CT imaging will continue to be the standard approach for only monitoring recurrence in PC.
Genetics of pancreatic cancer

A major contributor to the emergence of cancer and its progression is genetic instability that leads to multistep mutation accumulation. Identifying functional mutations in master regulators of key cellular processes often provides insight to the disease’s aggressiveness and can be used for predicting patient response to therapy and outcomes. Relative to other cancers, the mutation rate in PDAC is low at 2.64 mutations per megabase (17, 18), with a small set of frequently mutated genes that have either been difficult to target therapeutically or had poor performance in clinical trials (Fig. 3) (19). KRAS is the primary driver gene that is mutated in PDAC, which has been reported to occur in 90-95% of cases (19). Furthermore, in tumors with wild-type KRAS, mutations in other members of the pathway are observed in 60% of cases, showing the central importance of this pathway in PDAC development (20).

Figure 3. The most commonly mutated genes in PDAC. Dark color represents the minimum reported mutation frequency and light color represents the maximum reported mutation frequency across different studies. Reproduced with open access permission from (Pihlak, R., J. M. J. Weaver, J. W. Valle, and M. G. McNamara. 2018. Advances in Molecular Profiling and Categorisation of Pancreatic Adenocarcinoma and the Implications for Therapy. Cancers 10: 17.) (20).
Due to the prevalence and importance of \textit{KRAS} mutations, several attempts have been made to target them in human clinical trials, but have failed to improve overall survival. However, recent studies have shown that \textit{KRAS} mutation is a prognostic indicator of poor patient outcomes (21), suggesting early detection of mutations in \textit{KRAS} could be used to qualify candidates for enrollment in surveillance clinical trials. Additionally, like in a majority of cancers, \textit{TP53} is also commonly mutated in PDAC and thought to drive tumor development in \textit{KRAS}-mutated, pre-malignant lesions (22). \textit{TP53} is currently being targeted in clinical trials for other cancers, but no PDAC trials have been started yet with these therapies. \textit{SMAD4} is the third gene that has been reported to be mutated in $>50\%$ of cases, but the role of these mutations is still unclear. In addition, there are conflicting reports about the presence of \textit{SMAD4} mutations being a prognostic indicator (23, 24), so more research will need to be done to determine its significance across broader contexts such as stage, age, smoking history, etc.

Large PDAC sequencing studies have been completed recently and were used to classify PDAC into subtypes based on single nucleotide polymorphism (SNP) patterns, large structural variants, copy number changes, and multi-platform approaches combining nucleic acid and protein analysis (19). The various classifications that have been generated will have to be further scrutinized across cohorts at multiple institutions, but may be able to discern which patients will respond to a given therapy and which would likely not respond or develop resistance.

\textit{Bladder cancer statistics & treatment}

Unlike pancreatic cancer, bladder cancer (BC) is common in men, ranking 4\textsuperscript{th} in incidence among all cancers, but is not nearly as lethal as PC (77\% 5-year survival for BC) because most patients often experience symptoms like haematuria (blood in the urine) and are diagnosed relatively early (1). However, 25\% of patients are initially diagnosed with muscle
invasive BC (MIBC) or metastatic disease, and there is no effective treatment for BC metastasis (25). Distribution of stage at diagnosis and mortality rates for BC have not changed in over a decade because active population screening is not performed, but given the incidence of the disease, is needed greatly (Fig. 4A&B) (26). Additionally, 50-70% of patients initially diagnosed with non-muscle invasive BC (NMIBC) have recurrent disease and 40-50% go on to develop MIBC, which is treated by performing a radical cystectomy (RC) (27, 28). Meta-analysis from Chen et al. concluded that both cancer-specific survival and overall survival is similar after RC in patients that present with MIBC and patients that progress to MIBC, indicating a need for biomarkers that define which NMIBC patients would benefit from earlier RC (28). The 10-year recurrence-free survival for patients that get a RC is relatively high at 86%, but falls to 34% when the disease has spread to lymph nodes, and patients with seemingly localized disease can later present with recurrence in the lymph nodes. Not to mention, RC

significantly impacts quality-of-life, so prevention and effective therapies that don’t require removal of the bladder would be the best-case scenario.

Transurethral resection of bladder tumor (TURBT) is first performed for all new BC cases to stage the disease, but BC arises from “field cancerization” of the entire urothelium, and the small specimens used for staging/diagnosis incompletely represent its molecular and cellular heterogeneity (29). White-light cystoscopy, a fairly invasive standard procedure is used to evaluate and monitor BC, yet its reliability for detection and diagnosis of early-stage BC is modest (30-33). Cystoscopy and cytology can be very accurate in the hands of a skilled urologist at a high volume academic center; however, most patients handled in smaller general practices are not diagnosed until the disease has progressed to MIBC. Misdiagnosis of BC leads to higher recurrence rates and can be lethal in the context of MIBC, due to the differences in the clinical management of MIBC and non-muscle invasive BC (NMIBC) (26). The high recurrence rates necessitate regular surveillance with cystoscopy and cytology, making BC the most expensive cancer on a lifetime per patient basis (26, 34, 35). Thus, more objective and sensitive strategies with improved prognostic capacity are urgently needed for BC monitoring.

In the case of a NMIBC diagnosis, endoscopic resection and instillation of the bladder with the Bacillus Calmette-Guerin vaccine is the first line of therapy, and is reportedly more effective than chemotherapy (26). BCG therapy has been shown to reduce progression by 37% and recurrence by 40% at 1 year, and 15% at 5 years (36, 37). Unfortunately, 50% of patients given BCG fail to respond, dramatically reducing survival (38). As mentioned above, RC alone is the standard approach for treating MIBC, but 50% of these patients later die due to previously unobserved metastases that develop (39). If metastasis has occurred to lymph nodes or other sites, adjuvant chemotherapy comprised of a cisplatin combination is given after RC, which like PC, only results in a favorable outcome for a small minority of patients (Fig. 4B) (26).
**Bladder cancer risk screening**

As mentioned previously, screening has reduced mortality rates for many cancers. BC seems favorable for screening since urine comes in direct contact with bladder tumors and is truly a non-invasive biomarker source. However, similarly to PC, BC is not recommended for population-based screening by most urological societies like the US Preventive Services Task Force because there are currently no suitable biomarkers for detecting early stage disease, which are likely the best chance to further improve BC outcomes (40, 41). Studies have been done using at-home dipstick tests for hemoglobin in the urine, but these suffered from pitfalls such as lacking appropriate control arms (41). For example, one study screened 1575 men, with 258 testing positive and 21 later being diagnosed with BC. Long-term follow-up of these subjects showed only 4.8% had MIBC, compared to 23.6% in non-screened patients diagnosed traditionally. After 14 years of follow-up, none of the screened subjects died of BC, showing that this screening method was highly misleading (42).

Also similar to PC, smoking is the top risk factor for BC development and increases risk 2.5 times compared to nonsmokers (43). Chronic exposure to aromatic amines and other carcinogens has also been linked to BC, but is challenging to study and is believed to account for only 2-8% of new cases (44). Sequencing studies have not yet identified signatures related to these associated risk factors, making screening not feasible based on exposure to these carcinogens alone (26). More recent non-randomized trials have shown that detection of blood or nuclear matrix protein number 22 (NMP22) in the urine may be used as screening markers, but to date, thorough evaluation of screening markers in general is very limited and requires more comprehensive investigation.
Biomarkers of bladder cancer

Urine is an attractive liquid biopsy candidate for BC due to direct contact with the tumor. In addition, urine may overcome the limitations imposed by the relative paucity of tissue specimens and better reflect the molecular heterogeneity of BC than small biopsies. There are currently six FDA-approved, commercial protein/DNA-based tests for BC detection and surveillance, but their sensitivity and specificity for recurrent disease (35-75% and 76-94%) do not dramatically exceed that of cystoscopy (49-93% and 47-96%) (Table 1) (35, 45).

<table>
<thead>
<tr>
<th>Test/Marker</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytology38</td>
<td>0.35 (0.13-0.75)</td>
<td>0.94 (0.85-1.00)</td>
</tr>
<tr>
<td>NMP22</td>
<td>0.61 (0.49-0.71)</td>
<td>0.84 (0.75-0.90)</td>
</tr>
<tr>
<td>NMP22 BladderChek</td>
<td>0.70 (0.40-0.89)</td>
<td>0.83 (0.75-0.89)</td>
</tr>
<tr>
<td>BTA TRAK</td>
<td>0.58 (0.46-0.69)</td>
<td>0.79 (0.72-0.85)</td>
</tr>
<tr>
<td>BTA STAT</td>
<td>0.60 (0.55-0.65)</td>
<td>0.76 (0.69-0.83)</td>
</tr>
<tr>
<td>uCytx</td>
<td>0.75 (0.64-0.83)</td>
<td>0.76 (0.70-0.81)</td>
</tr>
<tr>
<td>UroVysion FISH</td>
<td>0.55 (0.36-0.72)</td>
<td>0.80 (0.66-0.89)</td>
</tr>
<tr>
<td>Cxbladder Monitor39</td>
<td>0.93</td>
<td>NA</td>
</tr>
</tbody>
</table>

Cl, confidence interval.

Table 1. Pooled performance of commercially-available tests for BC surveillance.


The most promising test available, Cxbladder Monitor, hasn’t been FDA-approved but meets Clinical Improvement Act (CLIA) guidelines. This test combines patient data with gene expression signatures for 5 targets as an adjunct to cystoscopy to rule out recurrent BC and avoid invasive assessment methods. It can detect Ta disease with 86% sensitivity and higher stage disease with 95% sensitivity for recurrent disease, regardless of BCG treatment status (46). Further validation of this test may lead to replacing cystoscopy in particular settings (41).
Lastly, recent reports assessing exfoliated cells and cell-free DNA (cfDNA) in urine samples of BC patients yield data supporting further research of urinary cfDNA as a biomarker (47-52). For instance, a retrospective next-generation sequencing/digital droplet PCR study of serially collected NMIBC samples revealed that high levels of tumor DNA in the urine are predictive of disease progression (49). The source, mechanism, and kinetics of cfDNA release in benign and malignant settings remain the subject of intense investigation and debate; however, deep sequencing efforts suggest vanishingly low frequency of tumor-associated mutations in cfDNA (53-56). Additionally, deep sequencing of cfDNA is an arduous task due to low abundance in circulation and high fragmentation (150-200 bp average length), which is even higher in urine (10-200 bp) due to >100-fold increase in DNase I activity (57-62). Achieving high global sequencing coverage of these low quality/quantity samples necessitated the development of new technologies, however, the prevailing clinical use of cfDNA remains limited (54, 56). On the other hand, growing evidence from our laboratory and others points to circulating exosomes as a source of genomic DNA in cancer patients, which we postulated may also be true for urinary exosomes, and a new biomarker class for the detection and surveillance of cancer (63-71).

*Genetics of bladder cancer*

Due to the “field cancerization” that occurs in BC, which is classified by detecting mutations in adjacent tissues that appear histologically normal, BC has been associated with high intra-tumor heterogeneity because multiple genetic clones are often present (29). However, the extent of this heterogeneity differs in NMIBC versus MIBC, with MIBC harboring more mutations and chromosomal rearrangements and/or amplifications/deletions (26).
NMIBCs typically have stable genomes, with loss of chromosome 9 being the most common chromosomal alteration and present in 50% of tumors. Activating mutations in TERT (73-83%), FGFR3 (60-70%), and PIK3CA (16-25%) are the most common oncogene alterations and deleterious mutations/loss in CDKN2A (15-60%), KDM6A (12-60%), and STAG2 (32-36%) are the most common tumor suppressor alterations in NMIBCs. Based on the prevalence of activation mutations in the RAS-MAPK pathway in NMIBC, it’s thought that this pathway may be essential for the development of >80% of these tumors (26).

Unlike NMIBCs, MIBCs are often comprised of many subclones and contain genetic aberrations that are not highly observed across patients like TERT and FGFR3 in NMIBCs. For example, activating mutations in PIK3CA (9-20%), FGFR3 (5-20%), and HRAS (5-12%) are the most frequently mutated oncogenes in MIBC. Inactivating mutations in KMT2D (27%), ARIDIA (25%), and APC (6-16%) are the most frequent mutations in tumor suppressors. Like most solid cancers, MIBCs frequently show loss of function or amplification of the repressors that regulate the tumor suppressors TP53, PTEN, and RB1. In addition, although FGFR3 mutations are rarely observed in MIBCs, up to 40% of these tumors show upregulated expression of FGFR3 (26, 72). Like in PC, large-scale sequencing efforts are beginning to classify BC into different molecular subtypes. These classifications have shown prognostic importance and have also identified potentially actionable targets novel to BC. Furthermore, they may be useful to refer to if DNA-based biomarker tests are developed for BC.

**Exosomes and their DNA content**

*Discovery of exosomes*

In 1967, Peter Wolf identified that “platelet dust” containing Platelet Factor 3 (PF3 aka CD142) was responsible for coagulation occurring in platelet-free plasma (73), an observation
first reported by Chargaff and West in 1946 (74). Wolf used electron microscopy to characterize this platelet dust and discovered that it was rich in complex phospholipids needed for thrombin generation and was likely involved in blood clotting (73). This would be proved 22 years later when Sims et al. showed that an excessive bleeding disorder called Scott syndrome is attributed to a microparticle-formation defect in platelets (75). Subsequent work from Webber and Johnson, and Crawford, confirmed Wolf’s findings and further characterized platelet-dust as biologically active vesicles that are released into the circulation from platelets (76, 77).

Since Wolf’s discovery of platelet-dust, microparticles have been a topic of exponential interest and have led to debates regarding the definition of a microparticle. “Microparticles” is now a broad term that is synonymous with “extracellular vesicles”, and comprises the vesicle classes exosomes and microvesicles, which are formed through different processes. Microvesicles are generally described as vesicles that bud directly from the plasma membrane and are typically larger (100-1000nm) than exosomes (78). Exosomes however, are released from multivesicular bodies through the endosomal pathway and are typically 40-150nm in size (78, 79). Current methods for exosomes isolation often don’t distinguish exosomes from microvesicles of a similar size and specific markers for each population are under active investigation, but in this work “exosome” refers to vesicles that are < 220 nm in diameter.

The term “exosome” was first proposed to describe 40-1000nm vesicles released by normal and neoplastic cell lines (80), but the association of small vesicles with multivesicular bodies (MVBs) and the endosomal pathway in rat reticulocytes shown by Harding et al. (81) caused the nomenclature to shift to a more specific subset of vesicles. Harding & Stahl made this discovery when they were studying transferrin receptor turnover using gold-conjugated transferrin, which revealed localization to small vesicles released from multivesicular endosomes (later called multivesicular bodies) into the extracellular space (Fig. 5) (81).
Figure 5. Release of exosomes containing gold-labeled transferrin. The multivesicular body fuses with the plasma membrane and releases the exosomes contained within. Reproduced with permission from (Harding, C., J. Heuser, and P. Stahl. 1983. Receptor-mediated endocytosis of transferrin and recycling of the transferrin receptor in rat reticulocytes. The Journal of Cell Biology 97: 329-339.) (81). License number: 4306541129446.

Biogenesis of exosomes

Also studying transferrin receptor turnover on the plasma membrane (PM) of sheep reticulocytes, Pan et al. were the first to observe the steps of exosome biogenesis (82). First, inward budding of the PM results in clathrin-coated vesicles that become early endosomes. These endosomes undergo further inward budding that result in the creation of intraluminal vesicles (ILVs) containing cytosolic components. This causes the endosomal lumen to acidify and early endosomes mature into late endosomes and multivesicular bodies (MVBs). From here, the cholesterol and ceramide content of endosomal membrane at least partly determines if the endosomes will fuse with lysosomes, degrading its contents, or fuse with the plasma membrane,
releasing ILVs as exosomes (78). MVB membranes enriched with ceramide and cholesterol are thought to be fated for plasma membrane fusion and ILV release (83, 84).

Exosome packaging of cytosolic molecules occurs through two separate routes, ESCRT-dependent and ESCRT-independent mechanisms. The first requires the four endosomal sorting complexes required for transport (ESCRT 0, I, II, and III), which are individually comprised of 2-7 individual proteins and associate on the membrane of endosomes. Together, these complexes associate with other proteins like Alix and Rab GTPases to facilitate exosome packaging and biogenesis. Specifically, the ESCRT-0 heterodimer complex facilitates clustering of ubiquitinated proteins and/or receptors on the cell surface to package them into budding endosomes (85). Interestingly, when ubiquitination of endosomal proteins was blocked in a yeast model, ILVs no longer formed in endosomes (86). The ESCRT-I heterotetramer complex also helps gather ubiquitinated proteins, and also passes ubiquitinated proteins from ESCRT-0 to ESCRT-II for endosomal packaging. In addition, ESCRT-I recruits ESCRT-III to MVBs. As mentioned, the ESCRT-II heterotetramer is responsible for packaging ubiquitinated proteins into endosomes, and also associates with ESCRT-III to pinch budding ILVs off into MVBs (Fig. 6). Additionally, Vps4 facilitates ESCRT complex removal from membranes and recycling (78, 87, 88). The ESCRT-independent pathway is not as well understood, but is thought to exist because exosomes are still produced when all four ESCRT complexes are depleted (78). It seems to be dependent on sphingomyelinase and tetraspanins like CD63, which is a commonly used membrane protein found on exosomes. As mentioned, Rab GTPases are also involved in exosome biogenesis, and Rab27a and Rab27b seem to primary facilitators of MVB fusion to the plasma membrane. However, other Rab proteins are also involved and silencing of Rab27a or Rab27b still allows exosomes to be released, although in fewer numbers (89-91).
Endocytosis leads to the formation of early endosomes (EE), which become MVBs as ILVs form. MVBs can then degrade their cargo by fusing with lysosomes, or fuse with the plasma membrane and release ILVs as exosomes.


Exosomal cargo and protein/RNA biomarker utility

Exosomes carry heterogeneous cargo comprised of nucleic acids, proteins, and lipids that collectively reflect their cell of origin. Commonly found proteins used as markers for exosomes include tetraspanins (CD9, CD37, CD63, CD81, CD82), endosomal proteins (flotillin-1, Alix, Clathrin, Tsg101), and heat shock proteins (Hsp70, Hsp90). However, common exosome isolation procedures do not select for exosomes specifically, so it’s likely that microvesicles with a similar size/density also contain some of these protein markers (92). Other proteins are present in exosomes as well depending on the specific features of the cell that produces them. For example, exosomes from antigen-presenting cells are abundant in MHC
class II molecules and exosomes from immune cells in general contain proteins with immune function, which are absent in exosomes from non-immune cells (79). Nucleic acids (miRNA, lncRNA, circ.RNA, mRNA, rRNA, tRNA, ssDNA, and dsDNA) have also been found in exosomes, but generally seem to vary depending on their parental cell type (63, 93-97).

Human cell-derived exosomes were first found to carry nucleic acids when Ratajczak et al. reported the presence of mRNA in embryonic stem cell exosomes. This study also revealed that the mRNA in these exosomes could be translated to protein in recipient cells (98). Subsequently, Valadi et al. confirmed this finding in human and mouse mast cells as well, and expanded on it by showing that exosomes also contain small RNAs, including microRNAs. They determined that many of the identified mRNA transcripts were absent in the cytoplasm of the exosome-producing cells, suggesting its packaging into exosomes was an orchestrated process and not random (99).

In pancreatic cancer, a number of RNA/protein biomarkers have been proposed. Evaluation of exosome size and concentration from the serum of healthy donors and patients with benign disease or PDAC has also been done, but significant differences between the three groups have not been observed (15, 100). As mentioned previously, the most promising biomarker in general for PDAC, that is also associated with circulating exosomes is glypican-1. GPC-1 was found to be a reliable biomarker for distinguishing patients with PDAC versus benign diseases of the pancreas. It was also shown that GPC-1+ exosomes contained mutated KRAS G12D mRNA, whereas GPC-1- exosomes from the same sample did not, suggesting that GPC-1+ exosomes may be primarily derived from PDAC cells driven by mutant KRAS expression (15). In addition, using a novel, highly-sensitive biochip, Hu et al. also found that GPC-1 mRNA levels in exosomes distinguished PDAC patients with the same accuracy as the previous study (16). Other groups have challenged the efficacy of GPC-1 as a theranostic biomarker for PDAC (100, 101), so additional research will likely be needed until a majority
consensus among the field is reached. Also using a biochip, Liang et al. found that ephrin type-A receptor 2 (EphA2) on plasma exosomes could also identify PDAC with high accuracy from as little as 1 µl of plasma, and their assay outperformed ELISA (102). In addition, Costa-Silva et al. identified that migration inhibitory factor (MIF) in circulating exosomes may also be a predictor of PDAC liver metastasis (103). Lastly, studies have also shown that specific miRNAs in exosomes may also be PDAC biomarkers (104, 105). As with all biomarkers, these candidates will need to be further validated and evaluated in clinical trials, but they offer hope for a disease that desperately needs it.

Exosomes have also been explored as a source for biomarkers in bladder cancer. Unlike in pancreatic cancer, Liang et al. suggest that the concentration of exosomes in the urine distinguished BC patients from healthy controls with a sensitivity of 81.3% and specificity of 90%, although the sample number was relatively low for each group (BC: n=16, Healthy; n=8) (106). Assessing the content of urine exosomes, Lin et al. identified alpha-1 antitrypsin and H2B1K as potential diagnostic BC protein biomarkers, but the sensitivity was poor (alpha-1 antitrypsin: sensitivity 62%, sensitivity 97%, H2B1K: sensitivity 62%, specificity 92%), limiting their potential (107). In a small cohort of 8 MIBC patients and 3 healthy donors, Berrondo also used urine exosomes to suggest HOTAIR and four other long non-coding RNAs (lncRNAs) were enriched in MIBC urine exosomes, with 2-10 fold higher expression versus healthy donor urine exosomes by RNA sequencing and qRT-PCR (108). Lastly, Andreu et al. evaluated the protein and miRNA content of urine exosomes (16 NMIBC, 18 MIBC, and 9 healthy donors) and found that miR-146a correlated with NMIBC and was a negative predictor for relapse, and miR-375 correlated with MIBC (109). It’s interesting to note that urine offers advantages over blood for BC because of proximity to the tumor, and although the exosome biomarker studies that have been performed thus far for BC are not as robust as what has been done in PC, there is great potential for urine exosomes as a biomarker.
Exosomes are of particular interest in the biomarker field because they can be non-invasively and serially collected from every bodily fluid. Furthermore, because they contain an array of biomolecules, there is potential for simultaneous identification of multiple biomarkers to increase accuracy for distinguishing between healthy versus pathogenic states and different stages of disease. In addition, the “mixed bag” characteristic of exosomes increases the potential for identifying unique biomarkers that help determine responders to therapy and overall patient outcomes.

*Functions of exosomes in cancer*

In addition to their utility as biomarkers, exosomes have been thought to be important facilitators of intercellular communication through the diverse molecular cargo enclosed within or associated with their bi-layer, phospholipid membrane (65, 110-112). Exosomes have recently become a major focus in cellular crosstalk studies and have been implicated in the regulation of diverse physiological and pathological processes such as cellular homeostasis, embryonic development, angiogenesis, innate and adaptive immunity, cancer metastasis, and cardiovascular disease (113-120). They have also been associated with the spread of harmful prion proteins in neurodegenerative disorders, and as accessories to viral infection (121-123).

The functions of exosomes in PC development and progression have been explored and implicated in a wide range of processes including tumor microenvironment remodeling and metabolism, insulin resistance, immune modulation, and metastasis. Mu *et al.* demonstrated in a rat PDAC model that PDAC exosomes carry proteases and can degrade the ECM upon contact, stimulating tumor cell proliferation and metastasis (124). Additionally, Masamune *et al.* showed how PC exosomes interact with stellate cells to stimulate fibroblast activation and fibrogenesis to further promote tumor progression (125). Showing the systemic effects of PC exosomes,
Wang et al. reported that PC exosomes can interact with skeletal muscle cells and stimulate lipidosis and inhibit glucose intake. In addition, these exosomes could also inhibit insulin and PI3K/AKT signaling (126). Lastly, two groups have reported that PC exosomes can promote metastasis by priming the liver microenvironment. Costa-Silva et al. first showed that PC exosomes can be taken up by Kupffer cells and stimulate TGF-β and fibronectin production by hepatic stellate cells, which then lead to recruitment of bone marrow-derived macrophages to create an inflammatory microenvironment (103). Yu et al. subsequently followed up on this work and showed that exosomes from highly metastatic cells can increase the metastatic burden of cells that typically have weak metastatic capacity (127).

Bladder cancer exosomes have also been investigated as mediators of tumor progression, but most of the work in this context has been done in vitro. Yang et al. suggested that BC exosomes inhibited cell death in a dose- and time-dependent manner through their expression of Bcl-2 and Cyclin D1, and upregulation of the ERK pathway in recipient cells (128). Two groups further expanded on this by demonstrating exosomes from MIBC cell lines and the urine of patients with MIBC also contain biologically active cargo capable of inducing epithelial to mesenchymal transition (EMT) and cell migration in primary urothelial cells (129, 130). Lastly, Xue et al. also observed the induction of cell proliferation, migration, and invasion through EMT in vitro and in vivo when exosomes carrying the lncRNA UCA1 (urothelial cancer-associated 1) from a hypoxic MIBC cell line were delivered to cells with low expression of UCA1, and suggested it may also be a biomarker for BC because it could be detected in the serum of BC patients (131).
DNA in exosomes

Genomic DNA was first found in exosomes (exoDNA) from human and mouse glioblastoma cells by Balaj et al. They reported that the DNA was single stranded and c-Myc amplification in the cells led to increased concentrations of c-Myc RNA and DNA in matched exosomes. They also showed that exosomes from normal skin fibroblasts contained exponentially less DNA, suggesting exoDNA packaging may primarily be a feature of cancer cells, or at least cells with genomic amplifications (97). Subsequently, our laboratory discovered that exosomes from PDAC patient serum and cell lines also contained genomic DNA, but this DNA was primarily double-stranded and composed of large fragments >10kb in length. ExoDNA from these sources also spanned the entire genome and reflected the mutational status of the tumor, confirming that the exosomes collected from serum originated from the tumor. In addition, PCR products could not be amplified from the exosome-depleted serum, suggesting that perhaps circulating DNA is primarily enriched in vesicles and not freely floating, as previously believed (63). Further studies have confirmed our findings in exosomes from other cell types, biological fluids, and disease settings (64, 132-134). The capture of cancer-specific mutations in exoDNA means that circulating exosomes have utility as a non-invasive source for monitoring tumor genetics (63, 133).

exoDNA as a biomarker

A major contributor to the emergence of cancer and its progression is genetic instability that leads to multistep mutation accumulation. Identifying functional mutations in master regulators of key cellular processes often provides insight to the disease’s aggressiveness and can be used for predicting patient response to therapy and outcomes. Multiple successive tissue samples are required in order to study the dynamics of these genetic events, but repeated
biopsies are rarely performed due to the risk of complications. Obtaining initial biopsies can also be very challenging or nearly impossible due to anatomical location of a tumor or health status of the patient (135). In addition, intratumor heterogeneity and varying ratios of malignant cells to normal cells in different biopsy samples can lead to inconsistent results in downstream analysis, compromising personalized-medicine approaches. However, exosomes are thought to be produced by all cells, so there is an opportunity to potentially capture more tumor mutations from circulating exosomes than from small samplings of solid tissue (Fig. 7).

**Figure 7. The advantage of circulating exosomes for capturing tumor mutations.** As tumors grow, genomic instability and environmental stresses lead to the creation of new cellular clones with unique genotypes and expression profiles. Small biopsies samples are often not representative of this heterogeneity within an entire tumor, and the presence of multifocal primary tumors or metastases amplify the mutations that can be missed in a biopsy. Circulating cancer exosomes may overcome this limitation if they can be separated from healthy cell exosomes, or if genomic DNA is enriched primarily in cancer cell exosomes. *Reproduced with permission from* (Kalluri, R. 2016. The biology and function of exosomes in cancer. The Journal of Clinical Investigation 126: 1208-1215.) (65). License number: 4310990323535.
One study that provided initial evidence for this claim came from San Lucas et al., who performed whole genome, exome, and transcriptome sequencing from matched exosome and tissue biopsy samples from three patients with pancreaticobiliary cancers. In one case, there were 100-fold more mutations found in exoDNA versus matched tumor tissue DNA (133). Although encouraging, some caution is warranted because prior studies evaluating intratumor heterogeneity by sequencing 10 different regions of the same tumor did not observe such a drastic increase in total mutational burden (136, 137). However, there are many differences between these studies, such as sequencing and bioinformatics approaches, and tumor types evaluated, so it is clear that much more research using next-generation sequencing is needed to fully evaluate increased capture of mutations in circulating exosomes versus tumor tissue.

Evaluating the diagnostic potential of exoDNA versus circulating-free DNA (cfDNA), Allenson et al. concluded that more mutations in the driver gene KRAS could be identified in exoDNA versus matched cfDNA, however, 20% of healthy controls also tested positive for KRAS mutations, requiring further validation and careful consideration of determining allele frequency thresholds to limit false positive tests as much as possible (138).

**exoDNA in cancer**

There is emerging new research that has given some indication of why DNA is packaged into exosomes and what effects this can have on recipient cells. Work from Takahashi et al. suggested that nuclear genomic DNA is secreted via human diploid fibroblast exosomes to maintain cellular homeostasis. When exosome production was inhibited, DNA accumulated in the cytoplasm and led to a reactive oxygen species (ROS)-dependent DNA damage response (DDR) mediated by the cytosolic DNA sensor, STING (stimulator of interferon genes) (68). However, it should be noted that inhibition of exosomal secretion may have itself caused DDR
activation, preceding cytosolic DNA accumulation and suggests that this phenomenon may be dependent on cellular stress and not necessarily a common feature of healthy cells. In addition, work from Kitai et al. demonstrated that treatment of breast cancer cells with the topoisomerase I inhibitor topotecan significantly increases exoDNA production, leading to dendritic cell activation through the STING pathway (69). This further suggests exoDNA accumulation is likely a product of DDR and potentially other stress-induced signaling events (Fig. 8).

**Figure 8. Current proposed mechanism for packaging of nucleic acids into exosomes.**

Cellular stress may facilitate increased packaging on nucleic acids into intraluminal vesicles (ILVs) before release as exosomes. In particular, STING activation leading to ROS production and activation of DNA-damage response and/or nuclear breakdown during cell division may result in mislocalization of genomic DNA in the cytoplasm and packaging into ILVs.

Interestingly, research from Cai et al. has demonstrated that exoDNA can be delivered to recipient cells, localize to the nucleus, and may be transcribed (140). Additional work from Cai et al. also suggests exoDNA can promote cancer-associated phenotypes in vivo, when they showed delivery of chronic myeloid leukemia cell line exosomes containing the BCR/ABL fusion gene led to the onset of weight loss, splenomegaly, and neutrophilia in rats. BCR/ABL mRNA and protein was found in the neutrophils of the rats, and treatment with the mRNA synthesis inhibitor actinomycin D prevented these symptoms (141). There is still much to be learned mechanistically about the function of exoDNA in recipient cells, and other work from Lee et al. reported that the effects of HRAS gene transfer through exosomes are transient and do not result in tumorigenic conversion (142). Therefore, the field is developing a better understanding of how exosomes interact and deliver their contents to recipient cells, but there remains a limited understanding of the role of exoDNA in recipient cells and whether its function is context-specific (Fig. 9).
Figure 9. Routes of exogenous dsDNA delivery to other cells. Free dsDNA enters cells through 1) micropinocytosis and 2) endocytosis in a sequence-dependent manner (143-145). exoDNA can enter cells through 1) micropinocytosis, 2) endocytosis (multiple mechanisms), 3) phagocytosis, and 4) membrane fusion (146). Reproduced with permission from (Kurywchak, P., and R. Kalluri. 2017. An evolving function of DNA-containing exosomes in chemotherapy-induced immune response. Cell research 27: 722.) (147). License number: 4297870031325

Bakhoum et al. recently published that errors in chromosomal segregation lead to cytosolic DNA accumulation and subsequent STING activation, promoting tumor cell invasion and metastasis (148). The rate of metastasis formation decreased when chromosomal instability was suppressed, and metastatic nodules were found to have a significantly higher degree of chromosomal instability when compared to the primary tumor. Chromosomal instability was associated with high rates of micronuclei formation, which were structurally unstable and led to leakage of genomic DNA in the cytoplasm. In addition to STING activation, cytosolic DNA resulted in non-canonical NF-κB activation and increased nuclear localization of the transcription factor RelB (148). Although this study was not directly linked to exoDNA production, it supports previous work and suggests exoDNA is at least a consequence of chromosomal instability, if not an active contributor to a metastatic phenotype. Further research is needed to fully elucidate the mechanisms of exoDNA packaging and whether the production of cytosolic DNA and exoDNA has a specific function, or are merely consequences of cellular crisis. Regardless, there appears to be immediate utility for exoDNA and more broadly exosomes for non-invasive monitoring of cancer, and potentially other pathological conditions as well.
Dissertation goals and major findings

In this work, we aimed to further address intra- and extraluminal DNA localization in exosomes as well as its potential intra- and extracellular origins. Whole genome sequencing of DNase I-treated cell line exosomes collected using two different exosomes isolation methods confirmed that exoDNA spans the entire genome and accurately reflects the copy number profile of the exosome-producing cells. Furthermore, we used digital PCR to evaluate the clinical utility of exoDNA for identifying prominent driving mutations ($KRA S^{G12D}$ and $T P 5 3 ^ { R 2 7 3 H }$) in 48 serum samples from PDAC patients, 7 from intraductal papillary mucinous neoplasm (IPMN) patients, 9 chronic pancreatitis (CP) patients, and 114 healthy donors. This confirmed that exoDNA could be used to identify PC driving mutations and also revealed that these mutations can also be detected in a small number of seemingly healthy donors (66). Follow up of these healthy donors was not possible, but would be needed in further studies to determine if these mutations predicted development of PC.

Expanding the potential of exoDNA in BC, we evaluated the DNA content of DNase I-treated urinary exosomes from healthy donors and BC patients to determine the utility of exoDNA for identifying BC mutations. We found that urine samples from BC patients contained significantly more exosomes and more exoDNA. Whole exome sequencing of urinary exoDNA, matched serum exoDNA, tumor DNA, and peripheral blood mononuclear cell (PBMC) DNA (normal control) from five BC patients and subsequent comparative variant analysis revealed superior capture of somatic mutations, many distinctive driver gene variants, and a subset of miRNA-binding domain mutations in urinary exoDNA compared to serum exoDNA and tumor tissue DNA. Taken together, our results demonstrate the potential theranostic value of urinary exoDNA analysis for BC and offer a starting point to further refine such techniques and analyses to generate robust, rigorous, and reproducible assays for BC management.
In addition, due to the unknowns that remain in regard to how genomic DNA is packaged into exosomes, we investigated an alternative possibility that vesicle production occurs directly from the cell nuclei. We compare these nuclear vesicles (NVs) to exosomes and explore horizontal exoDNA transfer and expression in recipient cells. We demonstrate that isolated nuclei can generate small vesicles, which are similar to exosomes in size, encapsulated by a bi-layer membrane, and also contain significant amounts of nuclease-protected, intraluminal DNA. Thus, it is likely that a fraction of the DNA-containing exosomes may derive from nuclear vesicles (NVs), which are possibly taken up and/or fuse with endosomes or multivesicular bodies in the cytoplasm during exosome biogenesis. Finally, we show that exoDNA can be traced to recipient tissues and is transcribed within recipient cells. Taken together, this work provides further evidence that evaluation of exoDNA as a biomarker should be pursued heavily in future research efforts and also suggests that exoDNA may be shuttled directly from NVs.
Chapter 2

Materials and Methods
Materials and methods for experiments with PDAC cell line and serum/plasma samples

Cell culture

The following human cell lines were used: Panc-1 (American Type Culture Collection (ATCC)), T3M4 (Cell Bank, RIKEN BioResource centre), HPNE (ATCC), HMLE (ATCC), BJ fibroblasts (ATCC), MDA-MB-231 (ATCC), DOV13 (ATCC). The following murine cell lines were used: 4T1 (ATCC), 1493 (primary pancreas fibroblast cell line generated in our laboratory). All cell lines were regularly tested for mycoplasma contamination. Panc-1 and T3M4 cells were cultured in RMPI-1640 supplemented with 10% (v/v) FBS, 100 U ml\(^{-1}\) penicillin and 100 µg ml\(^{-1}\) streptomycin. HPNE cells were cultured in 70% (v/v) DMEM (glucose-free, 2 mM L-glutamine, 1.5 g L\(^{-1}\) sodium bicarbonate) supplemented with 25% (v/v) Medium M3 Base, 5% (v/v) FBS, 10 ng ml\(^{-1}\) human recombinant EGF, 5.5 mM D-glucose (1 g L\(^{-1}\)), 750 ng ml\(^{-1}\) puromycin, 100 U ml\(^{-1}\) penicillin, and 100 µg ml\(^{-1}\) streptomycin). HMLE cells were cultured in DMEM/F12 media supplemented with 5% (v/v) horse serum, 20 ng ml\(^{-1}\) EGF, 0.5 mg ml\(^{-1}\) hydrocortisone, 100 ng ml\(^{-1}\) cholera toxin, 10 µg ml\(^{-1}\) insulin, 100 U ml\(^{-1}\) penicillin, and 100 µg ml\(^{-1}\) streptomycin. BJ fibroblasts, MDA-MB-231, 4T1, and 1493 cells were cultured in DMEM supplemented with 10% (v/v) FBS, 100 U ml\(^{-1}\) penicillin and 100 µg ml\(^{-1}\) streptomycin. All cell lines were incubated in a humidified atmosphere at 37°C with 5% CO\(_2\) and 20% O\(_2\).

Exosome isolation

Exosomes were collected from the conditioned media (CM) as described previously (Fig. 10) (149). Briefly, 5 million cells were grown in 225 cm\(^2\) flasks overnight, washed with PBS, and then incubated in serum-free media for 48 hours. Serum samples from healthy donors or PDAC patients were thawed at 37°C prior to exosomes extraction. Next, the CM or serum
was sequentially centrifuged at 800g for 5 minutes and then at 2,000g for 10 minutes to remove cellular debris and large vesicles. The cleared supernatant was passed through a 0.2 μm pore filter and ultracentrifuged (UC) at 100,000g for 2 hours at 4°C. For isolation using PEG-based Total Exosome Isolation Reagent (Thermo), CM was centrifuged and filtered as above, then 0.5 volume of the reagent was added and incubated for 1 hour at 4°C. Samples were centrifuged at 10,000g for 1 hour at 4°C. For both methods, the supernatant was discarded and exosomes were resuspended in 200 μl PBS for downstream processing. For nanoparticle tracking analysis (NanoSight LM10), 10 μl aliquots of exosome suspension were diluted 1:100 in cell culture-grade H₂O and 30 second videos were captured in triplicate to determine size and concentration. For immunoblotting, exosomes were resuspended in 100 μl Triton X-100 lysis buffer (see below).
**Sucrose gradient ultracentrifugation**

Exosomes were resuspended in 2 ml of HEPES/sucrose stock solution (2.5 M sucrose, 20 mM HEPES/NaOH, pH 7.4), placed in a SW40 tube (Beckman), and overlaid with 9 ml linear sucrose gradient (2.0-0.25 M sucrose). For DiO/DNA extraction experiments, exosomes were resuspended in 1 ml of HEPES/Sucrose stock solution (0.25 M sucrose, 20 mM HEPES/NaOH, pH 7.4) and layered on top of a 10 ml linear sucrose gradient. Gradients were subjected to UC for 16 hours at 100,000 g at 4°C. Gradient fractions of 1 ml were collected from top to bottom and density of each fraction was measured with a refractometer. Fractions were diluted and washed in PBS by UC for 2 hours at 210,000 g at 4°C. Pellets were resuspended in Triton X-100 lysis buffer for western blot or in PBS for nanotracking analysis and DNA extraction.

**Western blot**

After UC, exosomes were resuspended in Triton X-100 lysis buffer (150 mM NaCl, 1% (v/v) Triton X-100, 10 mM Na₂HPO₄, 2 mM KH₂PO₄, pH 7.4, 50 mM 6-aminoheptanoic acid, 10 mM EDTA supplemented with protease inhibitors [Complete EDTA-free, Roche] added just before use) and were allowed to lyse in the same tubes for at least 1 hour on ice. Protein loading was normalized based on BCA protein quantification and samples were denatured at 95°C for 5 minutes in 1X Laemmli loading dye ([5x] 10% [w/v] SDS, 10 mM β-mercapto-ethanol, 20% [v/v] glycerol, 0.2 M Tris-HCl pH 6.8, 0.05 % [w/v] bromophenolblue). Exosome lysates were loaded onto 10% (v/v) polyacrylamide gels and separated at 50 mA for 1 hour using the X-Blot system (Thermo). Proteins were transferred to PVDF membranes (Immobilon P) by wet transfer for 2 hours at 80V. Membranes were blocked for 30 minutes in 5% BSA in TBST at room temperature and then incubated overnight at 4°C with primary antibodies at 1:500 dilution: anti-
Flotillin-1 (sc-25506, Santa Cruz Biotech), anti-CD9 (ab92726, Abcam). Blots were washed 3 times in TBS-T (10 minutes each) and incubated with 1:500 anti-Rabbit (A0545, Sigma) horseradish peroxidase (HRP)-conjugated secondary antibody for 1 hour at room temperature. Blots were washed 3 more times with TBST and then developed using chemiluminescence reagents (Pierce).

Transmission electron microscopy (TEM)

Fixed specimens were placed onto 400 µm mess carbon/formvar-coated grids and allowed to absorb for at least 1 minute. Grids were rinsed with PBS and placed in 2.5% glutaraldehyde in 0.1 M phosphate buffer for 15 minutes, rinsed in PBS and distilled water, allowed to dry, and stained for contrast with uranyl acetate. The samples were viewed with a Tecnai Bio Twin TEM (FEI) and images were taken with an AMT CCD camera (Advanced Microscopy Techniques).

DNase I treatment and DNA extraction

Isolated exosomes were treated with 25 U ml⁻¹ DNase I in 1X DNase I buffer and incubated for 30 minutes at 37°C. DNase I was then inhibited by incubation with 2mM EGTA (final concentration, pH 8.0) for 10 minutes at 65°C. DNA from cell culture exosomes was extracted using the DNeasy Blood and Tissue kit (Qiagen), according to the manufacturer’s instructions and eluted in 50 µl H₂O. DNA from serum exosomes was extracted using the QIAamp DNA Micro kit (Qiagen) according to the manufacturer’s instructions with the following modifications: exosomes lysis was performed at 56°C for 1 hour, and DNA was eluted with 25 µl of elution buffer. Concentration was measured using the Qubit 3.0 high sensitivity dsDNA kit. Assays to test Qubit reliability and DNase I performance under different
conditions are provided in Fig. 11A-C. Fragment size distribution was assessed using the High Sensitivity DNA Kit (Cat#5067-4626) on the Agilent Bioanalyzer 2100 according to the manufacturer’s instructions. The standard curve used for Bioanalyzer measurements is shown in Fig. 11D.

Figure 11. Quality control experiments to test DNase I and Qubit 3.0 performance.

Readings were performed in biological triplicates. A) Panc-1 DNA was serially diluted 1:2 four consecutive times to determine measurement accuracy. Concentration readings were nearly identical using broad-range sensitivity reagents ($r^2 = 0.996$) and high-sensitivity reagents ($r^2 = 0.986$). B) DNase I performance was tested in either PBS or C) H$_2$O using DNase I from two different suppliers, with and without reaction buffer. DNA concentration was measured using high sensitivity Qubit 3.0 reagents. D) Standard curve for Bioanalyzer measurements. Seconds on the x-axis correspond with fragment length in base pairs on the y-axis.

cyTOF of exosome-bound beads
Panc-1 and T3M4 cells were plated in two 225 cm\(^2\) flasks each containing 50 µM 5-Iodo-2’–deoxyuridine (IdU, Cat# I7125, Sigma), incubated overnight, and CM collected as above after switching to serum-free media for 24 hours. Exosomes were collected by UC and resuspended in 250 µl PBS. Exosomes were fixed by adding an equal volume of 3.2% PFA in PBS for 10 minutes at room temperature, washed by UC, and resuspended in 100 µl PBS. Fixed exosomes were permeabilized by adding 2 volumes of cold 100% methanol and incubating at -20°C for 20 minutes. The samples were washed once more by UC in PBS supplemented with 0.1% BSA for 2 hours and exosomes were resuspended in 45 µl 0.1% BSA in PBS. Final exosome preparations were incubated with anti-GAPDH-Er170 antibody (final concentration 0.02 mg ml\(^{-1}\) Cat# A00178, GenScript) and incubated with gentle rocking overnight at 4°C. Exosomes were washed by UC as above, resuspended in 200 µl PBS, and incubated with 5 µl Aldehyde/sulfate latex beads for at least 1 hour before cyTOF analysis.

**DiO labeling of exosomes**

Serum-free CM were collected, cleared of debris by sequential centrifugation as described above, and passed through a 0.2 µm pore filter. 5 µl Vybrant DiO (Invitrogen) was added to the cleared media (5 µl x 25 ml\(^{-1}\) CM) and incubated for 45 min at 37°C. Exosomes were then collected by UC for 2 hours, pellets resuspended in 1 ml HEPES stock solution, and layered onto a sucrose gradient. DiO in stock, serum-free RPMI-1640 was processed in the same way as CM and used as a negative control.

**DNA library preparation and WGS**

Libraries were prepared from purified DNA samples using the SureSelect QXT Whole Genome Library Prep kit, following the manufacturer’s instructions (Agilent, Version D0).
Prepared libraries for Panc-1, T3M4, 4T1, and DOV13 samples were submitted for 76 bp paired-end WGS on an Illumina HiSeq 2000. The median coverage for these samples was 8.2X (Table 2). Separately, resequencing for Panc-1 (only Panc-1 cell DNA and Panc-1 exoDNA +DNase I) and first-time sequencing for MDA-MB-231 samples was performed using the same library preparation method as above and 151 bp paired-end WGS on an Illumina NextSeq 500. Median coverage for these samples was 11.5X.

### Table 2. Total reads and coverage of cell line WGS samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Total Reads (million)</th>
<th>Mapping Rate (%)</th>
<th>Duplicate Rate Mapped Reads (%)</th>
<th>Mean Target Coverage</th>
<th>Median Target Coverage</th>
<th>8X% Targets</th>
<th>6X% Targets</th>
<th>4X% Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panc-1 Cells</td>
<td>274.46</td>
<td>94.04</td>
<td>6.28</td>
<td>6.26</td>
<td>5.71</td>
<td>19.99</td>
<td>45.56</td>
<td>76.8</td>
</tr>
<tr>
<td>Panc-1 Nuclei</td>
<td>617.93</td>
<td>94.38</td>
<td>7.58</td>
<td>14.35</td>
<td>13.37</td>
<td>87.12</td>
<td>93.48</td>
<td>96.32</td>
</tr>
<tr>
<td>Panc-1 Conditioned Media</td>
<td>419.45</td>
<td>94.35</td>
<td>6.99</td>
<td>9.02</td>
<td>8.23</td>
<td>52.64</td>
<td>73.5</td>
<td>87.6</td>
</tr>
<tr>
<td>Panc-1 exoDNA +DNase (UC)</td>
<td>160.48</td>
<td>93.96</td>
<td>6.92</td>
<td>3.61</td>
<td>3.14</td>
<td>2.06</td>
<td>8.79</td>
<td>32.94</td>
</tr>
<tr>
<td>T3M4 Cells</td>
<td>583.38</td>
<td>94.41</td>
<td>7.32</td>
<td>12.53</td>
<td>11.64</td>
<td>75.68</td>
<td>84.49</td>
<td>90.57</td>
</tr>
<tr>
<td>Panc-1 exoDNA +DNase (Isolation Reagent)</td>
<td>412.48</td>
<td>94.74</td>
<td>6.94</td>
<td>9.44</td>
<td>8.74</td>
<td>57.63</td>
<td>75.98</td>
<td>87.83</td>
</tr>
<tr>
<td>Panc-1 exoDNA (Isolation Reagent)</td>
<td>314</td>
<td>94.83</td>
<td>6.48</td>
<td>7.52</td>
<td>6.91</td>
<td>36.37</td>
<td>61.84</td>
<td>83.41</td>
</tr>
<tr>
<td>4T1 Cells</td>
<td>505.83</td>
<td>94.39</td>
<td>7.53</td>
<td>11.19</td>
<td>10.18</td>
<td>71.83</td>
<td>85.99</td>
<td>93.7</td>
</tr>
<tr>
<td>T3M4 Nuclei</td>
<td>188.14</td>
<td>95.29</td>
<td>5.21</td>
<td>4.69</td>
<td>4.2</td>
<td>7.58</td>
<td>21.97</td>
<td>54.25</td>
</tr>
<tr>
<td>T3M4 Conditioned Media</td>
<td>212.29</td>
<td>94.62</td>
<td>6.93</td>
<td>5.3</td>
<td>4.74</td>
<td>13.49</td>
<td>31.94</td>
<td>62.39</td>
</tr>
<tr>
<td>T3M4 exoDNA +DNase (UC)</td>
<td>99.91</td>
<td>92.98</td>
<td>6.5</td>
<td>2.75</td>
<td>2.21</td>
<td>1.45</td>
<td>5.21</td>
<td>18.73</td>
</tr>
<tr>
<td>T3M4 exoDNA (UC)</td>
<td>383.77</td>
<td>94.36</td>
<td>6.81</td>
<td>9.62</td>
<td>8.76</td>
<td>57.67</td>
<td>75.74</td>
<td>87.72</td>
</tr>
<tr>
<td>T3M4 exoDNA +DNase (Isolation Reagent)</td>
<td>138.26</td>
<td>94.3</td>
<td>6.56</td>
<td>3.78</td>
<td>3.24</td>
<td>4.62</td>
<td>13.65</td>
<td>36.61</td>
</tr>
<tr>
<td>T3M4 exoDNA Isolation Reagent</td>
<td>164.15</td>
<td>94.31</td>
<td>6.82</td>
<td>4.3</td>
<td>3.79</td>
<td>6.85</td>
<td>19.18</td>
<td>46.43</td>
</tr>
<tr>
<td>4T1 Cells</td>
<td>535.86</td>
<td>94.34</td>
<td>12.8</td>
<td>13.49</td>
<td>12.58</td>
<td>85.2</td>
<td>93.01</td>
<td>96.45</td>
</tr>
<tr>
<td>4T1 exoDNA UC</td>
<td>535.09</td>
<td>95.58</td>
<td>8.82</td>
<td>19.34</td>
<td>16.78</td>
<td>79.32</td>
<td>86.34</td>
<td>92.35</td>
</tr>
<tr>
<td>DOV13 Cells</td>
<td>528.94</td>
<td>95.57</td>
<td>6.6</td>
<td>12.08</td>
<td>11.73</td>
<td>83.32</td>
<td>91.63</td>
<td>95.73</td>
</tr>
<tr>
<td>DOV13 exoDNA +DNase UC</td>
<td>184.8</td>
<td>90.11</td>
<td>9.97</td>
<td>3.17</td>
<td>2.67</td>
<td>1.15</td>
<td>5.92</td>
<td>25.03</td>
</tr>
<tr>
<td>DOV13 exoDNA UC</td>
<td>530.86</td>
<td>94.26</td>
<td>7.26</td>
<td>11.38</td>
<td>11.2</td>
<td>75.56</td>
<td>85.37</td>
<td>92.24</td>
</tr>
</tbody>
</table>

**Alignment and variant analysis**

Sequencing reads were aligned to the hg19 reference genome using BWA-MEM (150). Aligned reads were further processed using the GATK tool kits for duplicate removal, base recalibration, and variant calling (151). The called variants were annotated using ANNOVAR based on UCSC known genes (152). Variants listed in dbSNP129, 1000 genome, and EPS 6500 were filtered out to remove potential germline variants. Protein function impact assessment of variants were performed using SIFT and Polyphane scores (152).
Copy number analysis

Copy number profiles from sequencing data were obtained by first deriving segments using Circular Binary Segmentation (CBS) (153) and then deriving log2 ratio scores with an in-house tool, exomeCN, which is a modified version of HMMcopy (154) tuned for our data. Gene-level copy number assessment was derived from copy number segments using the R package CNTools (155). The log^2 scores > 0.5 were considered copy number gains while log2 scores < -0.5 were considered copy number losses. The correlation of copy number profiles between cell and paired exosomes samples were calculated using Pearson and Spearman correlation. A compilation of known cancer genes: CancerGenes (156), Network of Cancer Genes (157), and the Sanger Cancer Gene Census in COSMIC (158) was used for cancer gene annotation.

Healthy individual and PDAC patient serum samples (66)

Serum samples were collected from 171 patients who underwent surgery major pancreatic resection due to underlying diagnosis at the Department of General, Visceral and Thoracic surgery, University Medical Center of Hamburg, Germany between 2003 and 2013. The diagnosis was confirmed upon histopathological analyses of the resected material. The study was approved by the Ethics Committee of the Chamber of Physicians in Hamburg, Germany (PVN1045). Written consent for using the samples for research purposes was obtained from all patients prior to surgery or blood drawing. Blood samples of patients with suspected PDAC were obtained from central venous catheter from each patient directly before surgery. After collection of the blood samples, serum was separated from the blood by centrifugation at 3400×g for 10 minutes and these serum samples were kept frozen at –80°C. Samples were chosen randomly from the pre-existing tumor bank unaware of the underlying disease at the
time of exosome analysis. Histopathological proven diagnoses (PDAC, IPMN, chronic pancreatitis and in a few cases autoimmune pancreatitis, common bile duct cancer, pancreatic cystadenoma, pancreatic neuroendocrine tumor, duodenal adenoma, and uterine sarcoma) was revealed after digital PCR analyses of serum-derived exosomal DNA. None of the patients with malignant disease received preoperative cancer related therapies.

Blood samples of healthy individuals were collected by MD Anderson Blood Bank. Serum samples were collected in Vacutainer Plus plastic serum tube (Becton Dickinson, Franklin Lakes, NJ) and stored at room temperature until the end of the day. Samples were then spun at 3000xg for 20 minutes and then stored at 4°C overnight. The serum was then extracted the next day and the samples were stored at -80°C.

Digital PCR (66)

We aimed to detail the experimental procedures following the guidelines proposed for reporting digital PCR data (159) and proceeded with analyses of our results keeping in mind considerations reported on sensitivity and specificity (160). To identify KRAS$^{G12D}$ and TP53$^{R273H}$ mutations in human sample exosomal DNA, digital PCR was performed using Taqman SNP Genotyping assays on the Quantstudio 3D system (Thermo Fisher) according to manufacturer's instructions. The probe assay ID used are AH6R5PI (KRAS G12D c35G>A (WT:C —> MUT:T) and AHWSLEX (TP53 R273H c.818G>A (WT: C —> MUT:T). For human samples, 6.5µl of DNA was added regardless of DNA concentration. Samples with DNA concentration greater than 8ng/µl were diluted 1:3 in elution buffer to ensure proper cluster separation during analysis. The functional abundance was calculated as below, and reported as a percentage using the formula below:
Functional Abundance (%) = Total number of mutant alleles / Total number of alleles (mutant and wildtype) – average false positive rate.

The analysis of the digital PCR data was performed with the manufacturer’s software (QuantStudio 3D Analysis Suite). The detection limit of the assay was determined using exosomal DNA extracted from cell lines. Exosomal DNA extracted from HMLE cell lines were used as wild-type DNA, while that extracted from Panc-1 was used as mutant DNA for KRAS (heterozygous G12D) and TP53 (homozygous R273H). False positive rate and threshold (intensity) limit for mutant alleles were determined using wild-type DNA at concentrations representative of that of samples (four replicates at 0.1, 0.2 and 0.3ng/µl for patients (totaling 0.65, 1.3, 1.95ng of DNA), and 1 and 2ng/µl for healthy donors (totaling 6.5 and 13ng of DNA). The average false positive rate was calculated from all replicates, and used to calculate the reported functional abundance. A titration of mutant exosomal DNA from 10% to 0.01% with a total of 0.5, 1, and 3ng DNA was performed, determining the functional abundance threshold at 0.25%. We also conducted a no template control (n = 1) and did not detect any mutant or wild-type alleles.

Nuclei isolation & NV collection

Cells were harvested with 0.05% trypsin, centrifuged for 3 minutes at 2000 rpm, resuspended in 5 ml PBS, and then centrifugation was repeated. The supernatant was discarded and cell pellets were resuspended in three volumes of ice-cold NP-40 lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.5% v/v NP-40 supplemented with protease inhibitors [Complete EDTA-free, Roche] added just before use) and isolated as described previously (161). Briefly, cells were aspirated in and out of a 1000 µl pipette 10 times,
incubated on a rotating wheel at 4°C for 10 minutes, and nuclei were pelleted by centrifugation 
(200g, 5 minutes, 4°C). The supernatant (cytosolic fraction) was removed with a pipette and 
then nuclei were resuspended in lysis buffer and centrifuged two more times to remove any 
remaining cytosolic components. Final nuclei pellets were resuspended in 1 ml serum-free 
RPMI-1640 media and counted using a Cellometer Mini (Nexcelom). Five million nuclei were 
placed in 25 ml isotonic buffer (serum-free RPMI-1640 media or PBS were used with the same 
results) in a 225 cm² flask and placed in a humidified incubator (37°C, 5% CO₂, 20% O₂). 
Nuclear vesicles (NVs) were collected at indicated time points following the protocol used for 
exosome isolation.

PKH26 labeling of exosomes
Exosomes were isolated as described above, resuspended in 1 ml PKH dye solution (4 µl 
PKH26 in 996 µl Diluent C [10 mM Tris-HCl, 0.1 mM EDTA, 1 mM DTT, 0.15% Triton X-
100, 200 µg ml⁻¹ BSA, 50% glycerol, pH 7.4]) and incubated for 5 minutes at room temperature. 
Two volumes of RPMI-1640 supplemented with 1% BSA was added to each sample and 
incubated for 5 minutes at room temperature. Volume was adjusted to a total of 11 ml prior to 
UC. The supernatant was discarded, exosomes were resuspended in 1 ml RPMI-1640 and then 
UC was repeated twice to wash away excess PKH26 dye. Exosomes were resuspended in 210 µl 
PBS and used for further analysis.

Fluorescence microscopy
For fluorescence microscopy, nuclei were incubated in NP-40 lysis buffer with 10 µg ml⁻¹ 
1,4,6-Diamidino-2-phenylindole DAPI, (10 min at room temperature), and washed twice with 
lysis buffer as described above. Nuclei were resuspended in 1ml serum-free RPMI-1640
containing 1 µM ER-Tracker Red dye (Invitrogen), incubated 30 min at 37°C, washed twice with RPMI-1640, and plated on round glass coverslips in a 12-well plate for 1 hour. The medium was aspirated and the settled nuclei were fixed to the coverslips by adding 4% paraformaldehyde (PFA, 2 min at 37°C). Coverslips were then mounted on glass slides and images were obtained with a Zeiss Observer 2.1 microscope (Zeiss Zen software, 2012 Blue edition).

For DNA tracking, Hoechst 33342 was added to the cell medium to a final 10 µg/ml and incubated for 30 minutes at 37°C. The cells were then plated onto glass coverslips in 6-well plates (1x10^5 cells/well). After overnight incubation, cells were washed with PBS and incubated with SYTO16-labeled genomic or exoDNA (5 ng in 500 µl serum-free media) for 3 hours. Cells were washed two times with PBS, incubated for additional 24 hours in serum-free media and fixed by adding 4% PFA to a final 1% for 10 minutes at room temperature. The cells were washed twice with PBS, and permeabilized with 0.1% (v/v) Triton X-100 in PBS for 5 min. After 2 more washes the cells were incubated for 20 minutes at room temperature in Phalloidin 647 solution (Invitrogen, 5 µl Phalloidin in PBS supplemented with 1% BSA). The cells were rinsed two more times and the coverslips were mounted on glass slides and allowed to dry at 4°C overnight. In a parallel experiment, exosomes were collected and subjected to fluorescent measurement to detect exoDNA. Five million Hoechst 33342-labeled or unlabeled cells per experimental condition were plated in 225 cm² flasks and incubated as above; 250 ng of SYTO16-labeled gDNA or exoDNA added was to the media, incubated for 3 hours, the media removed, cells washed twice in PBS, and incubated for another 24 hours in serum-free media. Exosomes were then collected by ultracentrifugation, resuspended in 100 µl PBS and treated with DNase I. Exosome concentration was assessed by NanoSight™, and then fluorescence was
measured at an of 355 nm excitation to detect Hoechst 33342 signal (genomic DNA) and at 488 nm to detect SYTO16 signal (exogenous DNA).

For exosome tracking in cell culture, $5 \times 10^4$ cells were plated on glass coverslips in 12-well plates and allowed to adhere overnight. Then the cells were washed with PBS and placed in 500 µl serum-free media. PKH-labeled exosomes were added ($1 \times 10^9$) to the cells and incubated for 3 hours. Cells were fixed and permeabilized as above, washed with PBS and stained with Phalloidin 488 solution (Invitrogen) (1% BSA in PBS, 5 µl Phalloidin 488) for 20 minutes at room temperature. After another PBS wash, coverslips were mounted on glass slides.

**PCR & Sanger sequencing**

One million 1493 cells were plated in 75 cm$^2$ flasks, allowed to adhere overnight, and treated with $1 \times 10^{10}$ DNase I-treated Panc-1 or T3M4 exosomes for 3 hours. The cells were immediately collected for DNA extraction, or washed and incubated for another 24 hours with fresh serum-free media before DNA isolation. In another experiment, the cells were treated for 3 hours with 1 µg Panc-1 or T3M4 exoDNA (either naked or packaged with lipofectamine 2000). After cell collection and DNA extraction, human TP53 exon 5-8 were PCR amplified as described previously (63), generating a 1564 bp product (forward 5’-TTCCTCTTCCTACAGTACTCC-3’, reverse 5’-CCTGCTTGCTACCGTCGTA-3’) under the following cycling conditions: 94°C for 1 minute, 2 cycles of 94°C for 10 seconds, 67°C for 30 seconds, 70°C for 30 seconds, 2 cycles of 94°C for 10 seconds, 64°C for 30 seconds, 70°C for 30 seconds; 2 cycles of 94°C for 10 seconds, 61°C for 30 seconds, and 70°C for 30 seconds; 35 cycles of 94°C for 10 seconds, 59°C for 30 seconds, 70°C for 30 seconds. Additionally, for PCR with animal tissue DNA, human KRAS exon 2 was PCR amplified, generating a 166 bp
product (forward 5’-AAGGCTGCTGAAAATGACTG-3’, reverse 5’-AGAATGGTCCTGCACCAGTAA-3’) under the same cycling conditions used above.

PCR products were purified using the QIAquick PCR purification kit (Qiagen), and Sanger sequencing was performed using BigDye terminator kit (v3.1, Life Technologies) according to the manufacturer’s instructions, with primers for TP53 Exon 5-8 (forward 5’-TCTTCCTACAGTACTCCCCT-3’, reverse 5’-GCTTGCTTACCTCGCTTAGT-3’). Sequencing products were separated on an ABI 3730 automated sequencer (Life Technologies).

**RT-qPCR**

One million 1493 cells were plated in 75 cm² flasks, allowed to adhere overnight, and treated with 1x10¹⁰ DNase I-treated Panc-1 or T3M4 exosomes for 3 hours. The cells were immediately collected for RNA extraction, or washed and incubated for another 24 hours with fresh serum-free media before RNA isolation using the RNeasy Isolation Kit (Qiagen). cDNA was synthesized according to the manufacturer’s instructions (High-Capacity cDNA Reverse Transcription Kit, Cat# 4368813, Life Technologies) and used for RT-qPCR (SuperScript III Platinum One-Step qRT-PCR Kit with ROX, Cat# 11745-500, Thermo) in a QuantStudio 7 Flex Real-Time PCR system, using primers for human wild-type KRAS (forward 5’-ACTTGTGGTAGTTGGAGCTGG-3’, reverse 5’-TTGGATCATATTCGTCCACAA-3’) and primers for mouse ß-actin (forward 5’- CATGTACGTTGCTATCCAGGC-3’, reverse 5’-CTCCTTAATGTCCACGCACGAT-3’) as an internal control under the following cycling conditions: 50°C for 3 minutes, 95°C for 5 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 30 seconds. Fold-change was calculated using the ΔΔCt method and statistical significance was determined using Welch’s t-test (unpaired, two-tailed, * ≤0.05, ** ≤0.005).
Experimental animals

Swiss Nude mice were purchased from the Experimental Radiation Oncology Breeding Core at MD Anderson and maintained and handled at the MD Anderson Cancer Center vivarium following the guidelines outlined by the National Institutes of Health and MD Anderson Animal Care and Use Committee (Protocol #12-13-11831).

Exosome tracking in vivo

Panc-1 exosomes were collected and labeled with PKH26 as described above. Labeled exosomes collected from the 48 hour-conditioned media from two 225 cm² flasks were injected intraperitoneally (IP) into nude mice and the animals were sacrificed 3 hours later. Livers and pancreata were dissected and frozen in OCT compound (Tissue-Tec) for confocal epifluorescence microscopy or flash frozen for DNA extraction (DNeasy Blood and Tissue Kit). Tissue sections were stained with DAPI before analysis to visualize cell nuclei. PCR was carried out with the primers for exon 5 of human KRAS: (forward, 5’-AAGGCCTGCTGAAAATGACTG-3’, reverse, 5’-AGAATGTCCTGCACCAGTAA-3’) and cycling conditions as described above.

Statistical analysis

Statistical analysis was performed using Prism 7 (GraphPad) and data are expressed as mean ± SD. Statistical significance was determined using Welch’s t-test (unpaired, two-tailed, * ≤ 0.05, ** ≤ 0.005). Exosome and exoDNA measurement experiments were done in biological duplicate or triplicate.
Materials and methods for experiments with BC samples

Patient information and specimen collection

All samples were preserved in the Bladder SPORE Tissue bank (University of Texas MD Anderson Cancer Center) and the collection and analysis was approved by the Institutional Review Board (Protocol PA15-0970). All patients provided informed consent and all samples were properly de-identified. Three patients with NMIBC, one Ta low-grade (P3), two T1 high grade (P1 and 2) and three patients with MIBC (P4-6) were initially entered into the study. Two NMIBC patients (P1 and 2) had progressed to MIBC at the time of sample collection enrolled in the study, prior to neo-adjuvant therapy. Blood and urine were collected prior to transurethral biopsy, except from two blood samples (P2 and P5) that were collected 32 and 15 days after the biopsy was performed. One patient (P1) received neo-adjuvant intravesical Bacillus Calmette-Guerin (BCG) therapy others were given 3-5 cycles of various chemotherapy regimens. For each biopsy, H&E staining and staging was performed by a blinded Bladder Core pathologist.

Human urines were purchased from Bioreclamation IVT (Baltimore, MD) and sera from healthy volunteers were from the University of Texas MD Anderson Cancer Center Blood Bank.

Isolation of exosomes from urine and serum

Urine samples (35 ml for healthy volunteers, 4 ml for BC patients) were centrifuged and supernatants placed on ice. Tamms-Horsfall glycoprotein pellets were dissolved in DTT (200 mg ml⁻¹) to liberate exosomes and centrifuged, the supernatants combined and remaining pellets were discarded. Supernatants were passed through 0.20 µm syringe filters (Corning, 431219) to remove large vesicles and debris prior to ultracentrifugation (UC) (200,000 x g, 3 hours). Resultant pellets were resuspended in PBS and a second UC wash step performed. Final exosome pellets were resuspended in 210 µl PBS and a 10 µl aliquot used for nanoparticle
analysis. Remaining exosomes were incubated with DNase I (25 U/ml, Promega, 9PIM610, Madison, WI) for 30 min at 37°C and the reaction was terminated by incubation with 1X DNase I stop solution (Promega) for 5 min at 65°C (Fig. 12). Exosomes from serum samples were collected as previously described (149). Briefly, patient sera (1 ml) were thawed on ice, adjusted to 11 ml with PBS, and passed through a 0.22 μm syringe filters. Exosomes were collected by UC and DNase I treated using the same procedure described for urine exosomes above.

Figure 12. Schematic of the exosome isolation procedure for human urine samples.
Nanoparticle tracking analysis

10 µl exosome suspensions were diluted 1:100 in cell culture grade H$_2$O and loaded via syringe pump onto a NanoSight™ (LM10, Malvern Instruments, United Kingdom). Tracking was performed at 25°C with the camera level set at 13-16 for urine samples and 12-13 for serum samples to perform readings at a similar brightness for each sample. Three 30 second videos per sample were used to determine the size range and concentration of the particles.

Immunogold labeling and transmission electron microscopy (TEM)

Serum and urine exosomes were subjected to a total of three UC (210,000 x g) washes in PBS, 2 x 3 hrs and then 1 x 14 hrs. Exosomes were resuspended in 50 µl PBS with 2.5% EM-grade glutaraldehyde. Immunogold labeling with anti-CD9 primary antibody (Table 3) and TEM were performed as described previously (149).

<table>
<thead>
<tr>
<th>Primary antibodies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antigen</strong></td>
</tr>
<tr>
<td>CD9</td>
</tr>
<tr>
<td>CD63</td>
</tr>
<tr>
<td>CD81</td>
</tr>
<tr>
<td>Ficolin-1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Secondary antibodies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ab Type</strong></td>
</tr>
<tr>
<td>IgG</td>
</tr>
<tr>
<td>IgG</td>
</tr>
</tbody>
</table>

Table 3. Antibodies used for BC experiments.

Western blot analysis

Exosome pellets from 35 ml urine and 1 ml serum from healthy donors were lysed for 1 hour on ice in 100 µl urea buffer (8M urea, 2.5% SDS, 5 µg ml$^{-1}$ leupeptin, 1µg ml$^{-1}$ pepstatin, 1mM phenylmethylsulphonyl fluoride, PMSF). Lysates were cleared by centrifugation (22,000 x g, 4°C, 15 minutes) and protein concentrations determined using a modified BCA assay. Each
sample was measured in duplicate using BSA standards in the urea lysis buffer. Lysates were denatured in 1X Laemmli loading buffer at 95°C (5 min) and 40 µg per lane were run on 4-12% Tris-Bis gel using 1X MES running buffer (Invitrogen, NP0002, Carlsbad, CA) at 150V for 1 hour and transferred to PVDF membrane (Millipore, IPVH00010, Burlington, MA) by wet transfer at 32V for 2 hrs. Membranes were blocked for 1 hour (room temperature) with 5% non-fat dry milk in TBS-T (1X TBS, 0.05% Tween-20) and incubated overnight with primary antibodies at 4°C in 2% milk in TBS-T (Table 3). Membranes were washed 4 x 15 min with TBS-T and incubated with 2" antibodies for 1 hour at room temperature in 2% milk in TBS-T (Table 3). For band visualization, membranes were washed 4 x 15 minutes with TBS-T and incubated with ECL western blotting substrate per manufacturer’s instructions.

**DNA extraction**

DNA from urine exosomes (35 ml urine for healthy donors, 4 ml for BC patients), serum exosomes (1 ml serum for both healthy donors and BC patients), PBMCs and tumor tissue (1 mg). DNA from PBMCs and tumor tissue was extracted using the DNeasy Blood and Tissue kit (Qiagen, 69506, Hilden, Germany) and with QIAamp MiniElute columns (Qiagen, 57414, Hilden, Germany) for exosomes samples, according to the manufacturer’s instructions. DNA concentration was measured using the Qubit 3.0 high sensitivity dsDNA kit (Thermo, Q32854, Waltham, MA) and fragment sizes assessed using the Bioanalyzer 2100 High Sensitivity DNA Kit (Agilent, 5067-4626, Santa Clara, CA). Samples were stored at -20°C.

**Targeted PCR and Sanger sequencing**

DNA from urinary exosomes, tumor tissue, and PBMCs was PCR-amplified in 25 µl volume (for primers see Table 4) using KAPA2G Robust Hot start DNA Polymerase (KAPA...
Biosystems, KK5522, Basel, Switzerland) following manufacturer’s instructions in a T100 Thermocycler (Bio-Rad, Hercules, CA). PCR products were separated on 1% agarose gel for 1 hour at 100V and purified using the Wizard SV Gel and PCR Clean-Up System (Promega, TB308, Madison, WI). Sanger sequencing was performed using the same primer sets, at the Sequencing and Microarray Facility (MD Anderson Cancer Center). The resulting sequences were aligned and probed for somatic variants with DNA STAR SeqMan Pro software (Version 12.3.1).

Table 4. Primer sets used for PCR and Sanger sequencing.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward – 5’</th>
<th>Reverse – 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>FGFR3-1</td>
<td>CATGTCTTTGCAGCCGAGA</td>
<td>GGCAGCTCAGAACTGTTAT</td>
</tr>
<tr>
<td>FGFR3-2</td>
<td>GTGACCCAGGACACAGTAT</td>
<td>TGGGCTAACAGGCCCTGA</td>
</tr>
<tr>
<td>FGFR3-3</td>
<td>CCTGAGGTCATCCAAGGCGCA</td>
<td>ACCTTGCTGCGATTCAG</td>
</tr>
<tr>
<td>HRAS-1</td>
<td>GGGCGAGGCTCACCTCTAT</td>
<td>CTGGGGCTGGCTGAGCA</td>
</tr>
<tr>
<td>HRAS-2</td>
<td>ACTGTTGGAATGCTCAGAGGA</td>
<td>AGAGGCTGGCCTGAGAATC</td>
</tr>
<tr>
<td>KDM6A</td>
<td>ACAACACAGCAGTTACTTTTCT</td>
<td>ATTCGCAAAGGCTGCCC</td>
</tr>
<tr>
<td>TP53-1</td>
<td>GGCACACTGGCTGCAAGTTT</td>
<td>TGCTGTCGCCGAGCATTT</td>
</tr>
<tr>
<td>TP53-2</td>
<td>AAGAAAGCCAGACGGAAACC</td>
<td>TCAACCATGTAAGGCTCCC</td>
</tr>
<tr>
<td>TP53-3</td>
<td>AACCCTCTCCGCCAGAGAC</td>
<td>CCAAGGCTGCTAGGCTCAC</td>
</tr>
<tr>
<td>TP53-4</td>
<td>TATGGAAGAATCGTAAGGGTGGA</td>
<td>ATCTTGGGCTGTGATATC</td>
</tr>
<tr>
<td>TP53-5</td>
<td>CGATGGAACTACTGACACCT</td>
<td>TCCCTAGCTGCTGCTTCT</td>
</tr>
<tr>
<td>TP53-6</td>
<td>GCTGCTCAACCATTGCTATCT</td>
<td>TACTCCCTGCGCTCAACAA</td>
</tr>
<tr>
<td>PIK3CA-1</td>
<td>CATCTGTAATCCGAGGGGA</td>
<td>AGCAGCTACGTCGACTCCAT</td>
</tr>
<tr>
<td>PIK3CA-2</td>
<td>ACATTGGAAGACCGCTGGCTCT</td>
<td>ATTCGTCTTGGCCTGCTG</td>
</tr>
<tr>
<td>TERT</td>
<td>AGTGGATTCGGGCGACAGA</td>
<td>CAGCGCTGCTGAAACTC</td>
</tr>
</tbody>
</table>

Table 4. Primer sets used for PCR and Sanger sequencing.

DNA library preparation and whole exome sequencing

Prior to library preparation, sera and urinary exoDNA were subjected to whole-genome amplification (WGA) using the REPLI-g Mini Kit (Qiagen, 150025, Hilden, Germany). Library preparation and whole exome sequencing was performed at the Sequencing & Microarray Facility (University of Texas MD Anderson Cancer Center). A total of 23 DNA samples (PBMC reference unavailable for P3) 12-15 µg per sample were submitted for sequencing. DNA capture/ library preparation was performed using the SureSelect Clinical Exome Kit V2 (Agilent, 5190-9501, Santa Clara, CA) followed by sequencing on an Illumina HiSeq 3000. Coverage and other sequencing quality metrics are provided in Table 5.
Table 5. Whole exome sequencing total reads and coverage data. Red numbers are below the quality threshold: Mapping rate (≥95%), Duplicate mapped reads (≤25%), Mean coverage (≥100X), Median coverage (≥50X)

Quality control, sequence alignment, and variant calling

Estimates of concordance and contamination for matched sample - normal (PBMC) pairs were performed using Conpair (162) for detection of sample swaps and cross-patient contamination in WES experiments (Table 6A-E).

Table 6A. Total somatic variants in each patient sample.
Identification of somatic mutations using WES data from PBMC DNA, tumor tissue DNA, and urine and serum exoDNA was performed according to the Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical
Oncology, and College of American Pathologists as described by Li et al., and Chang et al. (163, 164). Globus Genomics (165), a Galaxy-based platform that uses Amazon Web Services for scalable compute and storage resources was used for the reference genome alignment and GATK-based variant calling. Variant annotation was performed using ANNOVAR (166). Identified variations were annotated with additional information from the UniProt database (167, 168) and Cancer Gene Index. Somatic driver mutations with strong or potential clinical significance were designated if the somatic mutation was identified in a gene that was (a) annotated as an oncogene or tumor suppressor gene by UniProt keywords; (b) associated with bladder carcinogenesis (Cancer Gene Index); or (c) has an entry in the COSMIC database. An overview of this bioinformatics approach is also presented as a flowchart in Fig. 13.

**Figure 13. Bioinformatics workflow for interpreting whole exome sequencing data.**

*Enrichment and comparative analysis*

Analysis of somatic mutations with strong or potential clinical significance in BC patient NGS data was performed as follows. All genes containing somatic mutations were annotated using information from the Lynx Knowledge Base (LynxKB) (169). Enrichment analysis to
discover over-represented functional categories and molecular pathways in the identified gene sets was done using Lynx enrichment analysis tools and ToppGene (169, 170). Construction of molecular networks and pathways harboring somatic variations of strong clinical and diagnostic significance was also performed using the Lynx suite of tools (169). STRING 10 was used as an underlying global network for network-based gene prioritization (171). The identification of miRNAs potentially interacting with mutated UTRs was done using the information from SomaMIR 2.0 database and ToppGene (170, 172).

Comparative analysis of somatic mutations in individual patients was performed using customized analytical pipelines specifically developed in-house for this purpose. The somatic mutations identified in samples from the same patient were compared to establish which variations are unique to a particular sample and which are shared among two or more. Comparative analysis of somatic mutations between individual patients was performed using the additional customized analytical pipelines developed in-house. The results of analyses were visualized in part using InteractiVenn (173).

Statistical analysis

Data are expressed as mean ± SD, with the exception of Fig. 15C, which is expressed as mode ± SD. Statistical analysis was performed using GraphPad Prism version 7. Non-parametric unpaired, two-tailed Mann-Whitney tests were performed independently for all urine and serum datasets, but were represented as single graphs in Fig. 15B-C & E-F. The Holm-Sidak method was used to determine statistical significance, **P≤0.05.
Chapter 3

Characterization of exosomes and exoDNA

**Characterization of PDAC cell line exosomes and exoDNA**

*PDAC cell line exosomes share common features*

Exosome production from two pancreas cancer cell lines (Panc-1 and T3M4) was measured by NanoSight™. We observed an accumulation of exosomes maintained for 72 hours after serum withdrawal from sub-confluent cultures (**Fig. 1A**). Exosome concentration was not impacted by DNase I treatment (**Fig. 1B**) and nanotracking analysis also showed that exosomes had a nearly uniform exosome diameter, with a major peak at approximately 150nm for both cell lines (**Fig. 1C**). Both Panc-1 and T3M4 exosomes were positive for EV markers flotillin-1 and CD9 by western blot (**Fig. 1D & F**), as well as the tetraspanin markers CD9, CD63, and CD81 by flow cytometry analysis (**Fig. 1E**). Ultracentrifugation using a bottom-loaded sucrose gradient yielded homogeneously distributed fractions positive for the EV marker flotillin-1 (**Fig. 1F**), thus confirming uniform size distribution. Transmission electron microscopy (TEM) showed spherical structures in the 40-150nm diameter range with lipid bilayer membranes (**Fig. 1G**).
Figure 14. Characterization of PDAC cell line exosomes. A) Exosomes were collected from conditioned media at the indicated time periods and particle concentrations were measured by NanoSight™. B) Exosome quantification was performed after exosomes were in the presence or absence of DNase I and the DNase I inhibitor EGTA. C) Representative NanoSight™ plots for Panc-1 (left) and T3M4 exosomes (right). D) Western blots of exosomal lysates probed for Flotillin-1 (Flot-1) and CD9. E) Flow cytometry analysis of Panc-1 (top) and T3M4 (bottom) exosomes for common tetraspanin exosome markers. F) Contents of Panc-1 and T3M4 conditioned media ultracentrifugation pellets were separated by density using a sucrose gradient. Aliquots from 1 ml fractions were subjected to Western blotting and probed for the exosomal marker Flot-1. G) Transmission electron microscopy (TEM) of Panc-1 and T3M4 exosomes.
Large fragments of dsDNA are predominately enclosed within exosomes

To address intra- or extra-luminal localization of exoDNA, DNA was extracted from Panc-1 and T3M4 exosomes that were treated with DNase I at distinct steps relative to exosome lysis and with or without the ion chelating agent EGTA to inhibit DNase I activity. Quantitative measurements following DNA extraction showed that when DNase I was applied before exosome lysis, 65-75% of the total exosomal DNA in the untreated control group was recovered. In contrast, only 10% DNA was recovered when DNase I treatment was performed following exosome lysis (Fig. 15A&B). When exosomes were exposed to DNase I prior to lysis, in the absence of DNase I inhibitor EGTA, roughly 40% of the DNA remained intact, suggesting that intraluminal DNA became susceptible to DNase I digestion when exosomal membranes were lysed. This observation further supports the concept of exosomal membranes serving as protective barriers for intraluminal exoDNA. When exoDNA content was normalized to exosome numbers, T3M4 exosomes contained more DNA (base pairs/particle) than Panc-1 exosomes; however, this difference was lost after DNase I exposure (Fig. 15C).

To further characterize the DNA cargo in PDAC exosomes, Panc-1 and T3M4 cells were incubated with the thymidine analog 5-iodo -2’-deoxyuridine (IdU), which is incorporated into DNA during S phase of the cell cycle. Exosomes were isolated, permeabilized, and incubated with Er170 conjugated anti-GAPDH antibody (internal control). After subsequent binding to aldehyde sulfate latex beads, exosomes were analyzed using cytometry time of flight (CyTOF). Panc-1 exosomes showed the following distribution: 8.9% beads double-positive for both IdU and GAPDH, 10.6%, positive for GAPDH alone and 5.9% for IdU alone. For T3M4 exosomes, 27.3% of exosome-bound beads were double-positive for IdU and GAPDH, 26.7% for GAPDH only, and 4.2% for IdU only (Fig. 15D). These results suggest that in the exosomes population under consideration, a heterogeneous distribution of DNA-containing and GAPDH-positive populations is detected, with a fraction of exosomes that do not contain IdU labeled DNA. The
higher percentages of IdU-positive beads in T3M4 exosomes is consistent with the observed higher DNA content per particle (Fig. 15C). In addition, this concept of heterogeneity within exosome population was also observed by traditional flow cytometry when these cell line exosomes were probed for tetraspanins, showing non-uniform expression (Fig. 14E).

Fragment length assessment of exoDNA from untreated or DNase I-treated exosomes showed large (up to 10 kb detection limit) fragments in both conditions. In contrast, very little exoDNA remained when DNase I treatment was performed after exosome lysis (Fig. 15E).

15
Figure 15. PDAC cell line exosomes contain large intraluminal DNA fragments. A) Total Panc-1 and T3M4 exoDNA content was measured by Qubit. DNase I and EGTA (DNase I inhibitor) were added before (B) or after (A) exosome lysis. B) In the same experiment, percent DNA remaining after treatment was calculated and compared between conditions. C) In the same experiment, DNA content was normalized to exosome concentration (base pairs particle⁻¹). D) Cellular DNA was labeled with IdU, exosomes isolated, permeabilized, incubated with anti-GAPDH antibodies, bound to Aldehyde/sulfate latex beads, and analyzed by CyTOF. E) Exosomes were isolated, treated with DNase I and EGTA at indicated time points (B, before lysis; A, after exosome lysis), and exoDNA isolated and analyzed by capillary electrophoresis (Bioanalyzer Station). F) DiO-labeled exosomes from Panc-1 and HPNE cells, and control RPMI medium were layered on top-to-bottom sucrose gradient, separated by ultracentrifugation, 1 ml fractions collected, and exosome-containing fraction visualized using spectrophotometry at 488 nm wavelength. The density of exosome-containing and adjacent fractions was determined by refractometry. Subsequent DNA extraction confirmed the presence of exoDNA in the gradient layer with the highest fluorescence intensity (total DNA amount is shown).

Lastly, we evaluated the exoDNA in sucrose-fractionated exosome preparations. Ultracentrifugation of DiO-labeled exosomes from Panc-1 and HPNE (normal human epithelial-like pancreas ductal cells) layered on top of a sucrose gradient showed a distinctly visible, single
band (Fig. 16). A negative control (exosome-free RPMI medium) confirmed the specific labeling of exosome membranes with DiO dye (Fig. 16). The density of the DiO-positive and adjacent 1 ml gradient fractions was in the range typically associated with exosomes (1.11-1.20 g/ml). To assess DNA concentration, fluorescence intensity at 488nm excitation was measured in all gradient layers. Only one fraction at 1.15 g/ml (DiO-positive band) from the Panc-1 cell line generated an RFU measurement above the background (Fig. 15F). The presence of DNA in this fraction was confirmed after subsequent DNA extraction.

Figure 16. Visualization of DiO-stained exosomes in sucrose gradient preparations.
Whole genome sequencing reveals that exoDNA is reflective of cell DNA

DNA was isolated from human pancreas (Panc-1, T3M4), ovarian (DOV13), mammary (MDA-231), and murine mammary (4T1) cancer cell line cells and exosomes (as well as Panc-1 and TM4 nuclei, conditioned media, and PEG reagent-isolated exosomes (Fig. 17A-E). Libraries for WGS were prepared and run on an Illumina HiSeq 2500, resulting in a mean coverage of 8.6X across all samples (Table 2). Copy number and variant analysis demonstrated that DNA from both untreated and DNase I-treated exosomes (pre-lysis) closely corresponded to the DNA isolated from matched cells (Fig. 17A&B, Tables 1-6). These observations further support the claim that cells generate large populations of exosomes, which collectively contain genomic DNA that spans all chromosomes and is not enriched for specific loci.
Panc-1 exoDNA (Exosome Isolation Reagent)

Panc-1 exoDNA + DNase (Exosome Isolation Reagent)

Panc-1 Nuclei DNA

Panc-1 Conditioned Media DNA

T3M4 Cell DNA

T3M4 exoDNA (No DNase I)

T3M4 exoDNA (+DNase I)
T3M4 exoDNA (Exosome Isolation Reagent)

T3M4 exoDNA +DNase (Exosome Isolation Reagent)

T3M4 Nuclei DNA

T3M4 Conditioned Media DNA

C

MDA-MB-231 Cell DNA

MDA-MB-231 exoDNA +DNase UC
Figure 17. Copy number analysis reveals that exoDNA spans all chromosomes. ExoDNA from A) Panc-1, B) T3M4, C) MDA-MB-231, D) 4T1, and E) DOV13 cell lines was subjected to whole genome sequencing followed by copy and variant number analyses. Vertical shaded regions represent different chromosomes. DNase I treatment before exosome lysis was applied where indicated to confirm sequencing of only intraluminal DNA. Black indicates the raw profile and red represents the segmented (average) copy number profile on a log\(_2\) scale. The profiles show chromosomal gains (above the baseline) and losses (below the baseline).
Characterization of healthy donor and PDAC patient serum exosomes and exoDNA (66)

Serum exosomes from healthy donors (n = 114) and patients (PDAC, IPMN, CP, or others) (n = 76) were enriched using size exclusion filtration and ultracentrifugation and particle size and concentration was assessed using nanoparticle tracking analysis (NanoSightTM). This revealed that exosome preparations from both populations were mostly similar in size, as shown in a representative NanoSightTM plot for healthy and PDAC patient serum exosomes (Fig. 18).

Figure 18. Representative concentration and size distribution of exosomes from healthy donors (top) and PDAC patients (bottom).

Despite the low volume of some serum samples (as low as 150 µl), sufficient amounts of exoDNA were collected for downstream digital PCR (dPCR) experiments, with a circulating exosomal DNA concentration range of 0.102 – 1.35 ng/ml for patients (0.11 – 1.22 ng/µl for PDAC patients, 0.106–0.584 ng/µl for IPMN, 0.13– 1.35 ng/µl for CP, and 0.102–0.476 ng/µl for others), and 0.212 – 19.7 ng/µl for healthy subjects (Fig. 19).
Figure 19. Healthy donor (left) and patient (right) serum exoDNA concentrations.

Characterization of urine and serum exosomes from healthy donors and BC patients

*Urine exosomes have some distinct features from serum exosomes*

Nanotrack analysis revealed that in healthy donors, urinary exosomes are smaller than serum exosomes (mode diameter of $130.3 \pm 24.6$ nm vs. $177.4 \pm 28.2$ nm) and more exosomes were present in the sera compared to an equal volume of urine ($1.6 \times 10^{11} \pm 2.8 \times 10^{10}$ mL$^{-1}$ vs. $3.6 \times 10^{9} \pm 1.9 \times 10^{9}$ mL$^{-1}$, respectively, (Fig. 20A-C). Western blot analysis identified typical exosomal proteins in the exosomal lysates, including CD9, CD81, and flotillin-1 (FLOT-1) in the urinary (U) exosomes and CD9, CD63 and FLOT-1 in the serum (S) exosomes (Fig. 20D). TEM of exosomal preparations showed similar vesicular structures in the serum and urine (Fig. 20E). Immunogold TEM also confirmed CD9 presence on both urinary and serum exosomes (Fig. 20E). Together, this analysis revealed that although serum and urine exosomes share some markers typically found on all exosomes, they also had some unique characteristics, suggesting different tissues may have distinct exosome biogenesis and packaging mechanisms.
BC patient urine samples contained elevated exosome and exoDNA concentrations

Nanotracking analysis showed significantly higher exosome concentrations in the urine of BC patients (Fig. 20B). No significant difference in exosome size was observed (Fig. 20C). Quantitative assessment showed higher urinary exoDNA content in the BC patients compared to healthy donors when normalized per sample volume or per exosome (Fig. 20F&G).
**Figure 20. Characterization of exosomes from the urine and serum of healthy donors and BC patients.** A) Representative graphs for nanoparticle tracking analyses of exosomes from the urine and serum of a healthy donor. B) Total exosomes as determined by nanoparticle analysis, normalized to input volume (particles x ml\(^{-1}\)). C) Exosomes mode diameter as determined by nanoparticle analysis. D) Western blot analysis of exosome lysates from healthy human urine and serum probed for exosome markers CD9, CD63, CD81, and FLOT-1. E) TEM and CD9 immunogold staining of the urinary and serum exosomes from healthy volunteers. Scale bars = 100nm. F, G) exoDNA measurements in the urine and serum normalized per mL biological fluid F) or per particle G). Data are expressed as mean ± SD, with the exception of panel C, which is expressed as mode ± SD. Non-parametric Mann-Whitney (unpaired, two-tailed) tests were performed independently for all urine and serum datasets, but were represented as single graphs in panels B, C, E, & F. The Holm-Sidak method was used to determine statistical significance, **P ≤ 0.005. □ Healthy donors; ■ BC; ○ P1, • P2, ○ P3, • P4, • P5, ○ P6.

**Urine and serum exosomes contain large fragments of genomic DNA**

Capillary electrophoresis (Bioanalyzer\(^{TM}\)) was used to compare exoDNA fragment quality (average length) before and after healthy urine exosomes were treated with DNase I (Fig. 21A), as well as to compare fragment quality in the exosomal and tumor tissue DNA preparations (Fig. 21B-D). The quality was similar, suggesting that the majority of the DNA fragments are localized within the exosomal lumen and are shielded from enzymatic degradation (Fig. 21A). As expected, tumor tissue yielded high-quality DNA, with fragments ranging between 1,521 - 12,216 bp (Fig. 21B). Both urine and serum exoDNA presented with a similar DNA profile, with urinary exoDNA fragments ranging from 1,593 - 16,295 bp and serum exoDNA ranging from 1,508 - 29,640 bp (Fig. 21C-D).
Figure 21. Comparative analysis of DNA from exosomes and matched tumor tissues. A) DNA from healthy donor urine exosomes from a healthy volunteer which were untreated (left) or pre-treated with DNase I (right). B) DNA from tumor biopsies and C) DNase I-treated serum and D) urinary exosomes from BC patients and analyzed by capillary electrophoresis.
Chapter 4

Utility of exoDNA as a cancer biomarker

Detection of PDAC-associated DNA mutations in human serum exosomes (66)

Establishing thresholds for mutation detection by digital PCR (dPCR)

We previously reported the detection of a heterozygous KRAS mutation at codon 12 (KRAS<sup>G12D</sup> c35G>A) and a homozygous TP53 mutation at codon 273 (TP53<sup>R273H</sup> c.818G>A) using the genomic DNA from PDAC cell line, Panc-1, derived exosomes (63). To determine the detection limit of KRAS<sup>G12D</sup> and TP53<sup>R273H</sup> mutations in exosomes by digital PCR, we used exosomal DNA derived from Panc-1 cells as well as HMLE cells (wild-type for KRAS at codon 12 and wild-type for TP53 at codon 273). Exosomal DNA from HMLE cells (n = 3 replicates) indicated a 0 ± 0% rate of detection for KRAS<sup>G12D</sup> mutation and 0 ± 0.05% rate of detection for TP53<sup>R273H</sup> mutation, validating their wild-type status at both loci. In contrast, exosomal DNA from Panc-1 cells (n = 3 replicates) indicated a 59.62 ± 0.89% rate of detection for KRAS<sup>G12D</sup> mutation and 99.67 ± 0.13% rate of detection for TP53<sup>R273H</sup> mutation, validating their reported genotype for each mutation. Known amounts of HMLE exosomal DNA were used to determine the false positive detection rates upon digital PCR analyses for KRAS<sup>G12D</sup> and TP53<sup>R273H</sup> mutations. A stringent threshold of 8800 A. U. (arbitrary units, referring to the intensity of mutant allele for each mutation was used, reflecting an average false positive rate of 0.040% and 0.034% for KRAS<sup>G12D</sup> and TP53<sup>R273H</sup> mutation, respectively (Fig. 22A). Importantly, the wild-type HMLE exosomal DNA concentration range in these experiments reflects the range of exosomal DNA concentration obtained from serum samples of patients (vide infra). Whether relatively low or high concentrations of exosomal DNA were used, the average false positive rate remained consistent (Fig. 22A), supporting the reliability of the chosen limit set for defining false positives (< 8800 A. U.). To define the sensitivity of detection of KRAS<sup>G12D</sup> and TP53<sup>R273H</sup> mutations, we next performed a titration experiment, in which different amounts of Panc-1 mutant exosomal DNA were mixed with different amounts of wild-type HMLE
exosomal DNA. These mixtures reflected the ‘spiked-in’ mutation at an expected frequency of 0.05, 0.1, 0.25, 0.5, 1.0, 5.0 and 10% (Fig. 22B). The titration experiment revealed that the sensitivity of detection for both mutations by digital PCR was 0.25% (Fig. 22B), below which mutant alleles could not be detected reliably, especially at low concentrations (0.5 ng in Fig. 22B).

**Figure 22.** Definition of digital PCR parameters using cell line-derived exoDNA. A) Determination of the average false positive rate using HMLE exoDNA. B) Determination of the sensitivity threshold by measuring the relative functional abundance of Panc-1 exoDNA spiked in HMLE exoDNA. Defined percentages of spiked Panc-1 exoDNA with HMLE exoDNA were expressed as an expected percentage of spiked mutation. The sensitivity threshold was set to 0.25% to reliably detect either \(KRAS^{G12D}\) (left) or \(TP53^{R273H}\) (right) mutations. (66)

We reasoned that detection was limited by the low concentration of exosomal DNA, and decided that a functional abundance above 0.25% was a conservative and reliable threshold to
accurately define samples as ‘positive’ or ‘negative’ for the two mutations. Taking into consideration the respective false positive rates for each mutation, we defined a sample as ‘positive’ for a given mutation if 1) the intensity of mutant allele was higher than 8800 A.U. (Fig. 23) and 2) the functional abundance, with the average false positive rate deducted (as described in experimental procedures), was greater than 0.25%.
**Figure 23.** Representative 2D intensity scatter plot of wild-type and mutant amplicon for A) *KRAS*<sup>G12D</sup> and B) *TP53*<sup>R273H</sup>. PDAC patients (P16, P50), an IPMN patient (P30), a CP patient (P66), and control spiked-in reactions (0.5% mutant). Threshold for mutant was set to 8800. Yellow = no DNA; blue = mutant; red = wild-type; green = mutant and wild-type. (66)

Capture of *KRAS*<sup>G12D</sup> and *TP53*<sup>R273H</sup> mutations in human serum by dPCR.

Informed by the thresholds established by our control experiments, detection of *KRAS*<sup>G12D</sup> and *TP53*<sup>R273H</sup> mutations could be performed using digital PCR (Fig. 23). For PDAC patients, *KRAS*<sup>G12D</sup> and *TP53*<sup>R273H</sup> mutations were detected in 39.6% (19/48) and 4.2% (2/48) of the samples, respectively (Fig. 24A&B). The highest functional abundance observed was 47.45% (patient P8) and 0.25% (patients P27 and P65) for *KRAS*<sup>G12D</sup> and *TP53*<sup>R273H</sup> mutations, respectively (Fig. 24A). We found that 3 out of 7 IPMN patients harbored the *KRAS*<sup>G12D</sup> mutation with the highest observed functional abundance at 2.17%, and one of these patients also co-harbored the *TP53*<sup>R273H</sup> mutation (0.52% functional abundance) (Fig. 24A&B). In CP patients, *KRAS*<sup>G12D</sup> mutation was found in 5 out of 9 of the serum samples (highest functional abundance of 1.12%), but none were found to have the *TP53*<sup>R273H</sup> mutation (Fig. 24A&B). Five out of 12 (41.7%) patients diagnosed with other diseases (autoimmune pancreatitis, common bile duct cancer, pancreatic cystadenoma, pancreatic neuroendocrine tumor, duodenal adenoma, and uterine sarcoma) harbored the *KRAS*<sup>G12D</sup> mutation (highest functional abundance of 2.20%), and only 1 had the *TP53*<sup>R273H</sup> mutation (functional abundance of 0.33%). In healthy samples, the *KRAS*<sup>G12D</sup> mutation was observed in 2.6% (3/114) of individuals, with highest functional abundance of 3.47% (Fig. 24A), and none had the *TP53*<sup>R273H</sup> mutation.
Figure 24. \textit{KRAS}^{G12D} and \textit{TP53}^{R273H} detection in human serum samples. A) Functional abundance of \textit{KRAS}^{G12D} (top) and \textit{TP53}^{R273H} (bottom) mutations for the indicated groups. B) Percent distribution of \textit{KRAS}^{G12D} and \textit{TP53}^{R273H} mutations in the indicated groups with overlapping circles indicating individuals with both mutations. PDAC: n = 48, IPMN: n = 7, CP: n = 9, Others: n = 12, Healthy: n = 114. (66)

This study indicates that even small amounts of serum, as low as 150 µl, can be used to isolate exosomes and purify DNA to identify cancer specific mutations in humans. We identified, in 48 serum samples from PDAC patients, \textit{KRAS}^{G12D} mutation in 39.6% of the samples and \textit{TP53}^{R273H} mutation in 4.2% of them, leaving 27 samples without these 2 specific mutations. These results are relevant, as the prevalence of \textit{KRAS}^{G12D} mutation identified using PDAC tumor tissue is approximately 40–50% (174). Therefore, this exosomal DNA study likely captures most of the anticipated \textit{KRAS}^{G12D} mutations in PDAC patients. This also appears to be the case for \textit{TP53}^{R273H} mutation, with an estimated 7% of patients possibly presenting with a c.818G>A substitution reflecting the R273H mutation. Our study, using a large number of
healthy subjects to detect cancer associated mutations in liquid biopsy, leads to a cautionary note that identification of driver mutations may not signify presence of disease. On the other hand, patients with CP (at higher risk of developing PDAC lesions) appear to present with a greater rate of KRAS mutation in our study. Indeed, while identification of highly prevalent mutations in circulating exosomal nucleic acids by itself may not offer a specific diagnosis, this approach, when combined with imaging modalities and other diagnostic procedure, may enhance personalized care of high-risk individuals. Circulating exosomal nucleic acid analyses may allow us to further develop precision medicine techniques for tailoring personalized care, with early detection and cancer prevention in mind.

**Novel and unique somatic variants in the DNA from urinary exosomes of BC patients**

*Urine exoDNA is suitable for PCR-amplification & Sanger sequencing*

To determine whether urinary exoDNA can be used to identify hotspot mutations in BC, urinary exoDNA, matched tumor DNA, and PBMCs were PCR-amplified and subjected to Sanger sequencing. 15 primer sets targeting known BC hotspot regions in six genes (*TERT, FGFR3, PIK3CA, TP53, HRAS, KDM6A*) were utilized (Table 4) (50). Successful amplification of each target region was determined by gel electrophoresis and confirmed by Sanger sequencing. For each of the six queried targets, a positive result was observed with urinary exoDNA from at least one patient (Fig. 25A). Urine exoDNA from the patient 1 sample set showed the highest representation of queried genes, with positive PCR amplification for all 15 targets. Sanger sequencing performed for a subset of targets in PBMC, tumor, and urine exoDNA, resulted in base calls with distinct peaks for all samples, indicating high quality sequencing suitable for detecting mutations. A representative comparison of base calls from the patient 2 sample set in a mutational hotspot region of *TP53* is shown (Fig. 25B).
### Figure 25. PCR & Sanger sequencing of BC hotspots in patient samples.

**A)** Gel electrophoresis of PCR hotspot targets in urinary exoDNA samples. **B)** Representative Sanger sequencing alignment of *TP53* hotspot PCR product in PBMC, tumor, and urine exoDNA from patient 2.

*WES of exoDNA can be negatively impacted by non-uniform whole-genome amplification*

Whole exome sequencing (WES) data generated by Illumina HiSeq 3000 yielded a high mean target coverage (≥100X) for all but three samples using 76 bp paired-end reads. However,
the median coverage for most exoDNA samples was poor, indicating a subset of targets with a much higher coverage than the rest of the genome. This can be attributed to using whole-genome amplification (WGA) prior to library preparation for exoDNA samples, which is necessary to generate a DNA concentration suitable for WES, based on current technology requirements. Similar low-template WGA procedures have been known to create amplification bias and poor coverage, especially in samples with varying quality (175-177). This diminished coverage in some of our samples was also reflected in the total number of variants identified per sample, as well as in concordance and contamination data (Table 6A-E).

**Variant analysis revealed superior mutation capture in urinary exoDNA samples**

Comparison of the mutational profiles in tumor samples revealed a range of 107-152 total somatic driver gene variants in individual samples, with 5-12 variants shared between at least two patient samples (Table 7). Additionally, seven variants were shared among three of the patients, and one variant was common among four patients (P1/2/4 & 5, STK11/rs10415095, Table 8). *KLK10* and *IGF1R* were the two most commonly mutated somatic driver genes across all tumor samples (Table 9).

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>90</td>
<td>6</td>
<td>11</td>
<td>9</td>
<td>5</td>
<td>124</td>
</tr>
<tr>
<td>P2</td>
<td>6</td>
<td>109</td>
<td>12</td>
<td>10</td>
<td>7</td>
<td>144</td>
</tr>
<tr>
<td>P4</td>
<td>11</td>
<td>12</td>
<td>71</td>
<td>11</td>
<td>6</td>
<td>117</td>
</tr>
<tr>
<td>P5</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>110</td>
<td>5</td>
<td>152</td>
</tr>
<tr>
<td>P6</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>81</td>
<td>107</td>
</tr>
</tbody>
</table>

Table 7. Total shared & unique somatic variants in driver genes in tumor samples. The number of variants unique to each patient tumor sample are highlighted in blue. The total number of somatic driver variants in tumor samples from each patient are highlighted in yellow.
Table 8. Top common somatic variants in driver genes in tumor samples. A novel SNV in the 3’UTR of the STK11/LKB1 gene was the only somatic driver variant shared in four tumor samples. Variants were first ranked by prevalence across tumor samples, and then by prevalence in multiple samples from a single patient. CGI BC status indicates if this gene has been previously known to be associated with bladder cancer based on data from the Cancer Gene Index. U: Urine, T: Tumor, S: Serum.

<table>
<thead>
<tr>
<th>Gene</th>
<th>dbSNP ID</th>
<th>CGI BC?</th>
<th>RefSeq Func</th>
<th>P1</th>
<th>P2</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
</tr>
</thead>
<tbody>
<tr>
<td>STK11/LKB1</td>
<td>rs10415095</td>
<td>No</td>
<td>U, T</td>
<td>T</td>
<td>T</td>
<td>S, T</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>KLK10</td>
<td>rs11343599</td>
<td>No</td>
<td>U, T</td>
<td>T</td>
<td>-</td>
<td>U, S, T</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>IGF2</td>
<td>rs58312807</td>
<td>Yes</td>
<td>UTR3</td>
<td>-</td>
<td>T</td>
<td>T</td>
<td>U, S, T</td>
<td>-</td>
</tr>
<tr>
<td>AKT1</td>
<td>rs1130214</td>
<td>No</td>
<td>UTR5</td>
<td>U, T</td>
<td>T</td>
<td>S, T</td>
<td>-</td>
<td>T</td>
</tr>
<tr>
<td>PSCA</td>
<td>rs2976396</td>
<td>Yes</td>
<td>UTR3</td>
<td>-</td>
<td>U, T</td>
<td>T</td>
<td>-</td>
<td>U, T</td>
</tr>
<tr>
<td>PTK2</td>
<td>rs13258775</td>
<td>Yes</td>
<td>UTR5</td>
<td>-</td>
<td>T</td>
<td>T</td>
<td>U, T</td>
<td></td>
</tr>
<tr>
<td>PLEKHG2</td>
<td>rs251860</td>
<td>No</td>
<td>UTR3</td>
<td>U, T</td>
<td>T</td>
<td>T</td>
<td>U</td>
<td></td>
</tr>
<tr>
<td>ETV6</td>
<td>rs1051782</td>
<td>No</td>
<td>UTR3</td>
<td>-</td>
<td>T</td>
<td>-</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

Table 9. Most commonly mutated driver genes in tumor samples. Genes were ranked based on the total number of variants found in that gene across patient tumor samples. CGI BC status indicates if this gene has been previously known to be associated with bladder cancer based on data from the Cancer Gene Index.
Analysis of exoDNA samples revealed superior capture of BC tumor mutations in the urinary versus serum exosomes, as well as identification of *de novo* mutations that were not detectable in the matched tumor tissues. The genes with the most frequently identified variants unique to urine samples were ranked, including *KRAS* in three of five patients (Table 10).

<table>
<thead>
<tr>
<th>P1</th>
<th># Variants</th>
<th>P2</th>
<th># Variants</th>
<th>P4</th>
<th># Variants</th>
<th>P5</th>
<th># Variants</th>
<th>P6</th>
<th># Variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTSB</td>
<td>3</td>
<td>AHRR</td>
<td>4</td>
<td>CCL16</td>
<td>1</td>
<td>VHL</td>
<td>5</td>
<td>RNF213</td>
<td>3</td>
</tr>
<tr>
<td>ETV6</td>
<td>3</td>
<td>GAS7</td>
<td>3</td>
<td>DPH1</td>
<td>1</td>
<td>AKAP13</td>
<td>3</td>
<td>KRAS</td>
<td>2</td>
</tr>
<tr>
<td>KRAS</td>
<td>3</td>
<td>RNF213</td>
<td>3</td>
<td>ESR2</td>
<td>1</td>
<td>GPI</td>
<td>2</td>
<td>PCM1</td>
<td>2</td>
</tr>
<tr>
<td>LDLR</td>
<td>3</td>
<td>KRAS</td>
<td>2</td>
<td>GDF15</td>
<td>1</td>
<td>TEP1</td>
<td>2</td>
<td>VHL</td>
<td>2</td>
</tr>
<tr>
<td>PLG</td>
<td>3</td>
<td>CCL16</td>
<td>2</td>
<td>MT1B</td>
<td>1</td>
<td>FGF1</td>
<td>1</td>
<td>CDH1</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 10. Top driver variants unique to urine samples.** Genes were ranked based on the total number of variants found in that gene across patient urine samples. Shaded cells indicate that variants in that gene were identified in multiple patient urinary exoDNA samples.

When comparing somatic variants of cancer-associated genes (drivers) identified in tumor samples to those found in matched exoDNA, 22-74% of total variants were unique to tumor samples, 20-54% found exclusively in the urinary exoDNA, and 11-14% unique to serum samples (excluding those with poor median target coverage). Additionally, individual patient samples showed 7-21% overlap between urine and tumor, 5-6% between serum and tumor, and 5-8% between urine and serum. Lastly, in the two patients with the highest quality WES data (sets P2 and P5), 6% of somatic driver variants were shared between tumor, urine, and serum (Fig. 26). Patient 2 shared variants across the three samples in *ARHGAP35, ELAC2, ID4, KLK10, and NMT1*. Patient 5 shared variants across the three samples in *AKAP13, IGF2, KLK10, and SUFU*. 
Figure 26. Overlap in somatic variants identified in the urinary exoDNA, serum exoDNA and tumors of individual patients. A) Comparative analysis of somatic driver gene variants showed 22-74% were found in the tumor (T) DNA only, 7-54% were in the urinary (U) exoDNA samples only, and 3-17% were unique to serum (S) exoDNA samples. Up to 21% somatic driver variants were shared between the tumor and urinary exoDNA (T/U), 0-6% were shared between tumor and serum exoDNA (T/S), and 0-8% were common between the urine and serum samples (U/S). 0-6% were common between tumor, urine exoDNA, and serum exoDNA (T/U/S). B) Comparative analysis of all somatic variants showed a similar distribution with 26-70% of all variants unique to tumor samples, 11-50% unique to the urinary exoDNA, and 4-15% unique to serum exoDNA samples. Up to 20% were shared between tumor and urine samples, 0-5% between tumor and serum samples, 0-6% between urine and serum samples, and 0-4% between tumor, urine, and serum in individual patients.
**BC patient panels have a high proportion of mutations in untranslated regions (UTRs)**

Variant analysis of the WES data from tumor tissues and exoDNA using matched PBMC DNA as the reference sequence revealed a number of somatic variants in 3’ and 5’ untranslated regions (UTRs) of all samples, which were considerably more prevalent compared to the variations in the exonic regions. Non-coding sequence variants in 3’ and 5’ UTRs have recently been associated with high-penetrance hereditary disorders. Significant polymorphisms in the 5’ regions (178-181) and in the 3’UTRs were linked to glioma, colon, breast, and ovarian cancers (182-185). A recent study, which provides means of a unified analytic framework to prioritize such non-coding variants revealed over 130 potentially deleterious polymorphisms in breast and ovarian carcinoma (186). We found that six of the UTR variants localized in the miRNA binding domains of potential driver genes, suggesting that these mutations may interfere with miRNA binding to the gene transcripts and therefore prevent post-transcriptional regulation and promote cancer progression (Table 11).

<table>
<thead>
<tr>
<th>Gene</th>
<th>dbSNP ID</th>
<th>miRNA ID</th>
<th>Location</th>
<th>S</th>
<th>T</th>
<th>U</th>
<th>S</th>
<th># Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAX2</td>
<td>rs67035383</td>
<td>miR-185-5p</td>
<td>UTR3</td>
<td>N</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>KRAS</td>
<td>rs9266</td>
<td>miR-181-5p</td>
<td>UTR3</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>KRAS</td>
<td>rs712</td>
<td>miR-877-5p</td>
<td>UTR3</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>CSF1R</td>
<td>rs3828609</td>
<td>miR-155-5p</td>
<td>UTR3</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>PCM1</td>
<td>rs1057016</td>
<td>miR-599</td>
<td>UTR3</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>1</td>
<td>x</td>
</tr>
<tr>
<td>ADAM7</td>
<td>rs3173956</td>
<td>miR-382-3p</td>
<td>UTR3</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 11. Mutated miRNA binding sites in driver genes.** Six variants were found to be located in predicted miRNA binding domains, which disrupts post-transcriptional regulation of these drivers and may contribute to the progression of bladder cancer in these patients. Note the two KRAS mutations, one of which was identified in two urine exoDNA samples but absent in their matched tumor samples, and the other which was only identified in two tumor samples.
Gene network reconstruction through identification of driver gene mutations

Network reconstruction of driver genes with mutations shared across at least four patients revealed significant interactions between cancer-associated driver genes, primarily centered around the AKT1 pathway (Fig. 27).

Figure 27. Common network for driver genes shared between at least four patients.

Network reconstruction of mutated driver genes (Pathways in Cancer, KEGG) shows highly significant interaction of cancer-associated driver genes, with most abundant clustering around the AKT pathways. Major oncogenic nodes are shown in red. AKT pathway (black arrows), KRAS pathway (white arrows), Wnt pathway (closed triangles), and TGF-β pathway (open arrows).
In addition, the network analysis using matched urinary exoDNA and tumor samples from patient 2, for example, showed good overlap and a tight network centered around pathways with strong cancer relevance, which include oncogenes AKT1/2, BCL2, KRAS, MDM2, PDGFRB, AXIN1, IGF1R, tumor-suppressive LATS1 and BRCA1, and immunomodulatory IL4R (Fig. 28).
Figure 28. Networks and driver genes in the urinary exoDNA and matched tumor DNA of patient 2. A) Urinary exoDNA. B) Matched tumor sample. Note tight networks clustering around KRAS and AKT pathways, with mutated tumor suppressor pathways (TSC1, BRCA1) and altered TP53 pathway (MDM2). Note a significantly higher representation of mutated genes in the urinary exoDNA for this sample set.

Thus, we have demonstrated that genomic DNA can be found in exosomes of BC patients, but not those not of healthy donors using the same sample volumes. This urinary exoDNA can be used to identify cancer-specific mutational profiles, which partially match the profiles of parental tumor and represent the pathways and signatures characteristic for bladder cancer. Furthermore, our research suggests that urinary exoDNA is superior to serum exoDNA for mutational analysis of bladder cancer. Finally, urinary exoDNA contains subsets of mutations that are absent in the matched tumor specimens, likely due to the limited representation of the highly heterogeneous bladder tumor tissue in a small biopsy.
Chapter 5

Investigation of the origin and function of exoDNA
**Assessing nuclei as sources of exosome-like particles**

*Isolated nuclei produce exosome-like vesicles that contain genomic DNA*

Since exosomes contain genomic DNA, we investigated the possibility of vesicle generation directly from the cell nuclei. To observe the release of nuclear vesicles (NV), nuclei were isolated from Panc-1 and T3M4 cells using either a 0.5% Tween-20 or a 0.5% NP-40 lysis buffer. DNA and remaining endoplasmic reticulum (ER) were visualized by fluorescence microscopy using DAPI and ER Tracker Red, respectively, which showed that the NP-40 lysis buffer efficiently removed the ER from nuclei (Fig. 29A). This isolation method was then used for all subsequent experiments.

Nuclei were placed in tissue culture dishes in an isotonic buffer at physiological pH and incubated at 37°C (20% O₂, 5% CO₂) for 24, 48, and 72 hours to determine the kinetics of NV release. The number of NVs released into solution was determined by NanoSight™. Similar to exosomes, NVs were released over 48 hours of incubation; however, unlike exosomes, which are continuously released for at least 72 hours, NV abundance dramatically declined after 48 hours (Fig. 29B). The cause for this is unknown, but could be due to loss of vesicle integrity at some point between 48 and 72 hours. NanoSight™ measurements for both Panc-1 and T3M4 NVs showed diameters similar to those of exosomes (Fig. 29C). DiO staining followed by sucrose gradient separation showed that Panc-1 NVs float at the same density as Panc-1 exosomes (Fig. 29D). FACs analysis of Panc-1 and T3M4 NVs showed that unlike exosomes, NVs from both cell lines are negative for tetraspanin exosome markers (Fig. 29E). TEM of T3M4 EVs demonstrated relative structural integrity and bilayer membrane of the resultant particles. (Fig. 29F). TEM also confirmed that isolated nuclei were free of ER traces (Fig. 29F).
**Figure 29. Cell nuclei shed exosome-like vesicles in *ex vivo* conditions.** A) Nuclei were isolated using Tween-20 (left) or NP-40 (right) lysis buffers and stained with DNA (blue, DAPI) and endoplasmic reticulum (red, ER Tracker Red) fluorescent dyes. B) Isolated nuclei were incubated for the indicated time periods and NVs were isolated by ultracentrifugation and quantified by nanotracking analysis (Nanosight™). C) Representative nanotracking plots of vesicle preparations from the Panc-1 and T3M4 nuclei. D) Top-to-bottom sucrose density gradient was layered with DiO-labeled NVs from Panc-1 and HPNE cells, and control RPMI medium. Density and fluorescence intensity at 488 nm wavelength of NV-containing and adjacent fractions were determined to demonstrate the similarities between NVs and exosomes. E) Flow cytometry analysis of Panc-1 (top) and T3M4 (bottom) NVs for common tetraspanin exosome markers. F) Representative TEM of isolated T3M4 nuclei (left) and NVs (right).

Characterization of the DNA content of Panc-1 and T3M4 NVs also showed that the DNA associated with NVs is protected from DNase I. The NV DNA concentration reached its peak at 24 hours (**Fig. 30A**). Furthermore, fragment length profiling of intact and DNase I-treated NVs was similar to those found in matched exoDNA preparations (**Fig. 30B**).

Due to the endosomal origin of exosomes, it is equally possible that exoDNA originates from the cytoplasm, the nucleus, or is absorbed from the extracellular space. To differentiate between these sources, cellular DNA in live T3M4 was labeled with Hoechst 33342. The cells were then incubated in media, to which SYTO 16-labeled genomic or exosomal DNA was added and the DNA uptake observed by fluorescence microscopy. Exosomes were then collected and the uptake of foreign DNA material was examined. SYTO 16 signal was not detectable in the recipient cells by fluorescence microscopy, however cells incubated with isolated cellular or exoDNA presented with high rates of formation of micronuclei or nuclear protrusions, respectively, suggesting that following DNA uptake it was either segregated into...
micronuclei or retained in the cytoplasm causing nuclear rupture. Of note, exosomes from the cells incubated with exogenous genomic DNA showed a higher Hoechst 33342 signal (incorporation of endogenous nuclear DNA). In addition, incubation with exogenous DNA resulted in a slight increase of the exosomal SYTO16 levels compared to control. These observations suggest that most of the exosomal DNA originates from the nuclei of parental cells, however exogenous DNA may be incorporated as well. Furthermore, the presence of exogenous DNA led to increased packaging of endogenous nuclear DNA into exosomes, as shown by increased Hoechst 33342 signal in isolated exosomes after cells were treated with exogenous DNA (Fig. 30C).

30

![Graph](image)

**B**

**Panc-1**

**T3M4**

![Fluorescence spectra](image)

**C**

<table>
<thead>
<tr>
<th>Condition</th>
<th>No Treatment</th>
<th>+Hoechst only</th>
<th>+gDNA</th>
<th>+exoDNA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panc-1</strong></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
</tr>
<tr>
<td><strong>T3M4</strong></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
</tr>
</tbody>
</table>

**Fluorescence Units (RFU)**

- 355nm
- 488nm

*Note: RFU values are approximate and may vary depending on the specific experimental conditions.*
Figure 30. Characterization of NV DNA. A) NVs collected at different time points as in panel B were exposed to control buffer or DNase I solution. Following DNase I exposure, NVs were lysed and DNA isolated and quantified by Qubit. B) Fragment length of NV DNA was assessed by capillary electrophoresis (Bioanalyzer). C) Live T3M4 cells with Hoechst 33342-stained nuclei were incubated with serum-free medium alone, or 250ng SYTO-16 (488 nm emission peak)-stained T3M4 gDNA or exoDNA (5ng for microscopy samples). Fluorescence imaging was performed to assess exosome uptake and localization. Actin cytoskeleton was visualized by Phalloidin 647 staining. Collected exosomes from 5,000,000 cells incubated with 250ng DNA were DNase I treated and analyzed by fluorescence measurement at 350nm (Hoechst) and 488nm wavelengths (SYTO16) to evaluate the origin of internal exoDNA.

Exosomes deliver transcriptionally-active gene fragments to recipient cells

To evaluate the potential of exosomes for in vivo DNA transfer, nude mice were injected once i.p. with PKH-labeled Panc-1 exosomes that were collected from conditioned serum-free media that was incubated with cells in two 225 cm² flasks for 48 hours. Pancreas and liver tissues were collected 3 hours post injection, nuclei stained with DAPI, and examined for exosome uptake by fluorescence microscopy. To assess horizontal DNA transfer in these tissues, a human KRAS DNA fragment from exon 2 was PCR-amplified using human-specific primers. Fluorescence imaging detected PKH-labeled exosomes in the livers and, more abundantly, in the pancreata of recipient mice (Fig. 31A). PCR amplification of total DNA from these tissues suggested that the human-specific KRAS DNA sequences were transferred to the tissues of recipient mice (Fig. 31A).

To assess exosome interaction with recipient cells, Panc-1 exosomes were PKH26-labeled and added to the epithelial-like pancreas cells (HPNE), breast epithelial cells (HMLE),
or foreskin fibroblasts (BJ). Fluorescence microscopy demonstrated similar localization of exosomes in phalloidin-labeled cells, regardless of cell type (Fig. 31B). To assess DNA transfer, mouse pancreas fibroblasts (1493 cell line) were incubated with Panc-1 or T3M4 exosomes and the presence of exoDNA was assessed 3 hours after treatment or 24 hours later. PCR with human-specific primers showed the presence of exons 5-8 of human TP53 exosome-treated 1493 cells at both time points, but not in untreated control (Fig. 31C). Furthermore, in 1493 cells treated with Panc-1 exosomes, Sanger sequencing identified the KRAS G12D (G>A) mutation, which is specific to Panc-1 cells (Fig. 31D). RT-qPCR of cDNA generated from 1493 cells exposed to Panc-1 or T3M4 exosomes using primers that specifically amplify human KRAS transcript revealed that expression of the human-specific transcript only occurred in cells that were exposed to Panc-1 or T3M4 exosomes, and not in untreated cells (Fig. 31E). These results suggest that exoDNA from Panc-1 or T3M4 exosomes included human KRAS sequence, which were transcribed by the recipient cell machinery. To test whether this transfer of genetic material was through exosomal cargo (intraluminal DNA) and not the uptake of free DNA, 1493 cells were incubated with either purified Panc-1 or T3M4 exoDNA or exoDNA pre-packaged into cationic microvesicles (lipofectamine) prior to incubation. Human TP53 sequences were only detected when exoDNA was incubated with lipofectamine, suggesting that DNA uptake is dependent on the cationic lipid membranes of the carrier vesicles (Fig. 31F).
Figure 31. ExoDNA is taken up by cells *in vitro* and *in vivo* and can be utilized as a template for mRNA synthesis. A) Nude mice received i.p. injections of PKH-26 labeled Panc-1 exosomes. Exosome localization in liver and pancreas was detected by fluorescence microscopy and exoDNA was traced by PCR amplification of a 166 bp fragment of exon 2 of human *KRAS* gene (right). B) Cultured human and mouse epithelial cells and human fibroblasts were incubated with PKH-26 labeled exosomes (2000 exosomes cell⁻¹) for 3 hours and exosome
uptake examined by fluorescence microscopy. Cell nuclei were visualized with DAPI (blue) and
actin cytoskeleton highlighted by Phalloidin 488 staining (green). C) 1493 cells were treated
with Panc-1 or T3M4 exosomes and DNA was isolated 3 hours after incubation or after 24
hours. PCR was performed to amplify human-specific TP53 sequence. DNA from untreated
1493 cells was used as a negative control, and Panc-1 or T3M4 cellular DNA were used as
positive controls. D) DNA isolated from 1493 cell DNA after incubation with Panc-1 exosomes
at indicated time points was subjected to Sanger sequencing seeking human-specific KRAS
mutations transferred by exosomes. E) Human-specific KRAS mRNA in 1493 cells after
incubation with Panc-1/T3M4 exosomes was measured by qRT-PCR. Fold change was
calculated by normalizing to the internal reference sequence (mouse β-actin). F) Isolated Panc-1
or T3M4 exoDNA were pre-incubated with control buffer or Lipofectamine 2000 and added to
1493 cells. DNA isolated from treated 1493 cells was subjected to PCR with primers to amplify
human-specific TP53 sequence. DNA from untreated 1493 cells was used as a negative control.

Our study shows nuclear DNA can be packaged into exosomes, and this process may
include the formation of nuclear vesicles. ExoDNA is nuclease-resistant due to the protective
membrane and can be delivered on an organismal level, to distant tissues, where it can be
utilized as a template for gene expression and therefore may impinge on the phenotype of
recipient cells.
Chapter 6

Discussion and Future Directions
Discussion

Confirmation of the intraluminal localization of genome-spanning exoDNA

This study confirms the presence of genomic DNA in exosomes using an array of experimental strategies. Differential sensitivity to DNase I added pre- and post-lysis showed that indeed, exoDNA is intraluminal and protected by the exosomal membrane. Mass cytometry (cyTOF) analysis revealed the heterogeneity of the Panc-1 and T3M4 exosome populations (those that contain both GAPDH and IdU/DNA, GAPDH, but no IdU/DNA and IdU/DNA-containing exosomes negative for GAPDH). It is, however, not impossible that asymmetrical distribution of IdU between exosome populations may be due to the fact that IdU is incorporated in replicating DNA during S phase and the only exosomes that register in such assay as DNA (IdU) positive are the ones that were generated by cells in S phase. However, the heterogeneity of exosomal populations have been previously suggested by the work of others (187-190). For instance, exosomes released from the organoids formed by colon carcinoma cells could be separated into two distinct populations based on the expression of epithelial antigen A-33 vs. Ep-Cam (191). The use of beads instead of individual exosome analysis is also potentially confounding since multiple exosomes could bind to a single bead. However, the high sensitivity of cyTOF allows for detection of even weak positive signal, and the presence of beads positive for GAPDH alone implies that exosome populations are indeed heterogeneous. Further evidence of the intraluminal DNA localization in Panc-1 exosomes was provided by analysis of the DNA content in DiO-stained and DNase I-treated exosomes after separation on a sucrose gradient.

The results that were observed from cell line exosomes were also confirmed in BC patient serum and urine samples. Based on previously published work and the existing need for better biomarkers, BC was chosen as the ideal disease setting to study urinary exoDNA. Our study confirms the elevated exosome content in the urine of bladder cancer patients, which was
demonstrated previously by using an integrated double-ultrafiltration device conjugated to a nanochip (192). Additionally, we report for the first time that genomic DNA is contained within urine exosomes and there is high exoDNA content in DNase I-treated urinary exosomes from BC patients compared to healthy donors. Thus, our findings refute previously published work (193, 194), but may have broader implications if confirmed with additional research, since the presence of tumor DNA in urinary exosomes could be used in lieu of cystoscopy for accurate and sensitive detection of recurrent BC. In contrast, serum exosomes presented with detectable baseline levels of healthy exoDNA, which were not significantly altered in the patients with BC. Therefore, urine exosome and exoDNA concentration alone may prove to be a biomarker for BC when evaluated in larger patient cohorts.

ExoDNA is a suitable template for dPCR and next generation sequencing

Copy number analysis generated by WGS experiments showed a prodigious degree of overlap between exoDNA and cellular genomic DNA in all sequenced cell lines. Importantly, this remained true for exosomes treated with DNase I, and is another demonstration of intraluminal localization of exoDNA. The similarity of CNA profiles generated for exoDNA and cellular DNA was remarkable regardless of the method of exosome isolation (ultracentrifugation or commercially-available PEG-based precipitation reagent) and showed that as a collection, exosomes contain DNA that spans the entire genome in multiple cell lines from different tissues. However, correlation was impacted by low mean coverage, leading to lower confidence in calling true deletions. This suggests that deeper sequencing could improve the correlation rate and also allow for high-confidence comparative analysis of specific variants.

After this confirmation was done from cell line exoDNA, we decided to expand on previous research from our laboratory, where \( KRAS^{G12D} \) and \( TP53^{R273H} \) mutations were identified in PDAC patient serum exosomes (63). In this work, we screened a larger cohort of
banked healthy donor and patient serum exosomes for these mutations using the highly sensitive dPCR technique, which revealed similar clinical detection rates of these mutations in PDAC patients, suggesting that this may be a reliable non-invasive approach for monitoring tumor status based on mutant allele frequency. However, a small percentage of healthy donor serum exosome samples also contained the oncogenic $KRAS^{G12D}$ mutation. Since monitoring and follow-up sample collections from these patients was not done, it’s not possible to determine if these mutations were real and/or led to the development of PDAC or other cancer. For a test like this to be implemented clinically, these unknowns will need to be addressed before it can be used in an adjuvant setting to supplement other diagnostic methods.

Our results were also promising in the BC setting, suggesting that exosome or exoDNA concentrations alone may have utility as a biomarker. Since the amount of DNA in urinary exosomes of healthy donors is low, it is likely that exoDNA from the urine of BC patients is primarily derived from tumor cells. This could therefore result in a higher signal-to-noise ratio compared to that in cfDNA or in serum exoDNA. The poor representation of BC DNA in the serum exosomes suggests that bladder tumor cells may shed exosomes into the urine at a higher rate than into vasculature, possibly due to a higher proportion of tumor cells being directly exposed to the bladder lumen as opposed to blood vessels, and is supported by previous findings (49, 52). This leads to the notion that urine is a superior source for BC-related biomarkers. Therefore, our research suggests the collective sequencing of both tumor biopsy DNA and urine exoDNA reflects the genetic heterogeneity of the tumor more completely than a limited biopsy specimen alone. The unique subsets of mutations found specifically in urinary exoDNA argue in support of this concept.

UTR variants are often excluded from WES bioinformatics pipelines, which are generally focused on coding regions, but the significance of UTR mutations for mRNA and non-coding RNA regulation in cancer is increasingly appreciated (195-200). Indeed, despite the bias
towards the somatic driver variants, which were found predominantly in the UTRs and intronic regions, our computational analysis of mutational profiles in the exoDNA of BC patients using LynxKB software tools (169) implicated multiple cancer-associated pathways, including those specific for bladder cancer. Among them were pathways supporting cancer cells themselves as well as those more attributed to the tumor microenvironment (angiogenic factors, inflammatory chemokines and their cognate receptors).

The most affected cancer-driving nodes across the majority of the patients found in the urinary exoDNA and the tumor tissue DNA included AKT1-3, BCR, FOXO3, IGF2, KRAS and MTOR/RPTOR, all of which affect cancer cell proliferation, survival and metabolism (201-209). Frequent mutations in the SMO/WNT/FZD module found in three out of six patients suggests epithelial to mesenchymal transition may have occurred in the invasive BC tumors (210), or potentially a switch to the activated state of tumor stroma (211, 212). The significant overlap between the main driver nodes involved in cancer progression and the TME lend further supports to the validity of using urinary exoDNA as a non-invasive biomarker for BC and potentially other cancer types. The angiogenesis-related genes with mutations found both in the urinary exoDNA and in matched tumor samples included VEGFB, PDGFRA, and PDGFRB. In addition, significant mutational burden in the IL4R and IL6R genes could potentially augment an inflammatory microenvironment. Lastly, analysis of mutational profiles using the SomaMIR database identified mutations in the miRNA binding domains of multiple cancer-associated genes, including two in the 3’UTR of the KRAS gene.

Thus, we have demonstrated that genomic DNA can be found in significant amounts in exosomes of BC patients but not those of healthy donors. This urinary exoDNA can be used to identify cancer-specific mutational profiles, which partially match the profiles of parental tumors and represents the pathways and signatures characteristic for bladder cancer. Furthermore, our research suggests that urinary exoDNA is superior to serum exoDNA for
mutational analysis of bladder cancer. Finally, urinary exoDNA contains subsets of mutations that are absent in the matched tumor specimens, likely due to the limited representation of the highly heterogeneous bladder tumor tissue in a small biopsy.

*exoDNA may originate from nucleus-derived vesicles and can be horizontally transferred*

The origins of DNA in exosomes are still being debated. Research from others suggests that fragmented nuclear DNA is segregated into rupture-prone micronuclei, which consequently release genomic DNA fragments into cytoplasm (148). Direct rupture of the nuclear membrane due to physical pressure from surrounding environment during migration can also lead to cytoplasmic DNA localization (213-215). Another study demonstrated that packaging of cytoplasmic DNA fragments in exosomes is part of homeostatic response to bypass aberrant STING response (68). Taken together, these studies imply that genomic DNA is incorporated into exosomes via a cytoplasmic route that involves nuclear/micronuclear rupture. There is currently no evidence directly linking nuclear rupture with exoDNA content, however several studies demonstrate the possibility of the nuclear origin of EVs, including the presence of a nuclear protein high mobility group box protein (HMGB) 1 in the EVs derived from LPS-stimulated mouse macrophages or in the blood of human volunteers (216, 217). Rappa et al. report on the association between late endosomes and invaginations in the nuclear envelope, which allows for the transport of exosomal materials to the nuclei, via nuclear pores (218). It is possible that reverse transport (from the nuclei to exosomes) also takes place. Finally, time-lapse imaging revealed a new type of membrane protrusions (“bead-on-a-string”) as a means of EV formation from nuclear membrane blebs (apoptotpodia) in apoptotic cells (219). Our studies using isolated nuclei show that in addition to DNA leakage to the cytoplasm or loss of nuclear
membrane integrity in apoptotic cells (220-222), fragmented nuclear DNA may also be packaged into exosome-like nanosized vesicles that bud directly from the nuclear membrane.

Using isolated nuclei, we found that at least \textit{ex vivo}, exosome-like vesicular nanostructures of nuclear origin are generated. These NVs showed remarkable similarity to exosomes with regards to their diameter, shape, enclosing membrane and, most importantly, intraluminal localization of genomic DNA. The observed differences in the kinetics of particle formation could be attributed to diminished particle stability in the chosen experimental conditions. The cessation of NV production could be caused by energy requirements or caused by other limiting factors that are exhausted in isolated nuclei after 48 hours. Further experiments will determine whether and how NVs and their DNA content are incorporated into exosome biogenesis pathways. Such mechanisms could be NV fusion with early endosomes prior to or during formation of intraluminal vesicles (ILVs) and multivesicular bodies (MVBs), or through release of NV DNA into the cytoplasm and successive uptake by endosomes/ILVs/MVBs.

The analysis of differential staining of nuclear DNA in parent cells and of exogenous DNA in their microenvironment clearly demonstrate that most of the DNA found in exosomes originates from the nuclei of the donor cells and only a small fraction by endocytic uptake of free DNA. Taken together with FACs and cyTOF analysis, which implies heterogeneity of markers and DNA loading in exosomes, these data suggest that DNA-containing fractions of an exosome population could be derived from the NVs. This is the first indication of such a mechanism of exosome formation and further studies are needed to fully elucidate the mechanism(s) involved in exoDNA sorting and packaging. To date, a similar mechanism has been described exclusively for apoptotic bodies (223). It is currently accepted that exoDNA represents damaged DNA expunged from cells to prevent activation of repair mechanisms that induce cell cycle arrest and potentially activate survival mechanisms such as autophagy (115).

Indeed, upon exposure to genotoxic agents such as topotecan or doxorubicin, DNA accumulates
in the cytosol and subsequently in exosomes in a non-specific manner, since no enrichment for specific DNA loci has been observed. Attenuated exosome secretion leads to DDR activation via the STING pathway, which can be limited by overexpression of cytoplasmic DNase2a (69). Our experiments show horizontal transfer of exoDNA in vivo and in cell culture. Interestingly, our imaging analysis frequently showed exosomes overlying cell nuclei, suggesting a possibility of direct exosome fusion to nuclei. On the other hand, our results demonstrate the requirement for a positively charged lipid membrane, underscoring the critical role of exosomes in horizontal DNA transfer on an organismal level. Horizontal DNA transfer facilitated by exosomes has also been shown by others (224-226). For example, exosomes derived from c-HRAS-transformed rat intestinal epithelial cells (RAS-3) contained exoDNA encompassing HRAS sequence, which was transferred into immortalized rat fibroblasts and retained for up to 30 days (227). It is nevertheless unclear whether the resultant increase in fibroblast proliferation was a direct consequence of gene transfer or induced indirectly by other factors transferred by HRAS containing exosomes. Further analysis could clarify the precise route of internalization, localization, and functionality of exoDNA in recipient cells. Interestingly, non-nuclear transcription of cytoplasmic DNA by a membrane-associated RNA polymerase II has been reported, indicating that delivery of exoDNA to the nucleus may not be critical for use as a transcription template (228). Our initial analysis shows that exoDNA retains the ability to be utilized by host transcriptional machinery. Exosomes from the human cell lines Panc-1 and T3M4 were incubated with murine 1493 cells. RT-PCR analysis detected increased human-specific KRAS transcript 24 hours after treatment, suggesting de novo transcription of DNA fragments carried by exosomes and the possibility of altering the phenotype of recipient cells. Similar results have been published regarding BCR/ABL DNA sequences in exosomes from CML, which are transcribed in HEK293 cells and neutrophils (141, 229). Other studies suggest that exoDNA can activate anti-cancer immune response or promote a suitable
environment for metastatic outgrowth when a malignancy is present. In both cases, the activation of cytosolic DNA sensing pathway cGAS–STING (cyclic GMP-AMP synthase–stimulator of interferon genes) is implicated (69, 115, 140, 141, 230, 231).

Conclusions

In conclusion, this collective work advances the fields understanding of exosomes and their DNA content by demonstrating that exosomes from multiple cell lines and circulating bodily fluids in two different disease settings containing DNase I-protected genomic DNA. Most importantly, this is the first study demonstrating the presence of genomic DNA in urine exosomes, which are more abundant in the urine of patients with bladder cancer and span the entire genome. We have demonstrated the utility of this exoDNA for the non-invasive theranostic assessment and monitoring of both pancreatic and bladder cancer using next-generation amplification and sequencing methods. Therefore, in addition to allowing serial assessment of tumor genetics, exoDNA may also reveal additional driver mutations that could be crucial to an accurate prognosis and more appropriate clinical treatment strategy.

Lastly, we provide preliminary results that suggest exoDNA is derived from the nucleus of an exosome-producing cell, which may involve packaging of genomic DNA into nuclear membrane-derived vesicles that are similar to exosomes in size and protect DNA from DNase I treatment. It remains to be determined if these vesicles are released from cells as a distinct population of extracellular vesicles or fuse with multivesicular bodies before exosome release. Regardless, it is evident that exoDNA can be horizontally transferred to other cells and used a transcription template, but the functional importance of this remains poorly understood for now.
**Future directions**

Primary future directions in regards to exoDNA content will revolve around determining concentration and sensitivity threshold requirements for downstream assays, with the potential for the development of clinical tests. In addition, more research will be needed to directly compare the isolation difficulty and efficacy of exoDNA versus circulating tumor cells or circulating-free DNA as biomarkers for cancer and other diseases. Lastly, combinations of exosomal biomarkers and/or other clinical parameters will need to be evaluated to determine if superior specificity and sensitivity can be established to replace current practices. Much remains to be done in this space, but it is clear that exoDNA offers promise and unique advantages over other investigated biomarkers.

In regards to performing next-generation sequencing with exoDNA, although encouraging, novel approaches will be needed to refine and optimize the parameters of analysis, and incorporate larger cohorts of patients with samples from serial time points. Most notably, sequencing results impacted by low-template WGA should be improved either by using alternative WGA methodologies that amplify the input more uniformly, or by defining minimum sample/exoDNA requirements needed to eliminate the WGA entirely. Now, additional research and longitudinal prospective trials with larger sample sets are needed to meet the rigorous validation requirements for reliable biomarker development.

Lastly, much work remains to be done to understand the origin and function of exoDNA in vivo. Specifically, the presence of nuclear vesicles in living cells could be visualized in living cells by using the lentiviral vectors NLS-GFP or NLS-RFP (nuclear localization sequence) in combination with DNA & nuclear membrane tags/stains. Additionally, the functional role of exoDNA could be further explored by using sophisticated methods involving loading exosomes with reporter plasmids of delivering Cre+ exosomes to genetically engineered mice lacking Cre and determine if and in which tissues recombination takes place.
Bibliography


Stage Pancreatic Cancer. *JNCI: Journal of the National Cancer Institute* 109: djw266-djw266.


36. Sylvester, R. J., W. Oosterlinck, and A. P. M. van der Meijden. 2004. A SINGLE IMMEDIATE POSTOPERATIVE INSTILLATION OF CHEMOTHERAPY


Genomic Alterations in Liquid Biopsies from Patients with Bladder Cancer. *European urology* 70: 75-82.


Solary, J. C. Soria, F. Andre, and L. Lacroix. 2017. Whole exome sequencing for
determination of tumor mutation load in liquid biopsy from advanced cancer patients.

applications and technological advances of circulating tumor DNA in cancer monitoring.
Therapeutics and clinical risk management 13: 1363-1374.

Heitzer, E., P. Ulz, and J. B. Geigl. 2015. Circulating Tumor DNA as a Liquid Biopsy

(DNase) in the serum and urine of systemic lupus erythematosus (SLE)-prone
NZB/NZW mice by a new radial enzyme diffusion assay. Clinical and experimental

activity in human tissues and body fluids by a single radial enzyme-diffusion method.

Bryzgunova, O. E., and P. P. Laktionov. 2015. Extracellular Nucleic Acids in Urine:

Melkonyan, H. S., W. J. Feaver, E. Meyer, V. Scheinker, E. M. Shekhtman, Z. Xin, and

Su, Y. H., M. Wang, D. E. Brenner, A. Ng, H. Melkonyan, S. Umansky, S. Syngal, and
T. M. Block. 2004. Human urine contains small, 150 to 250 nucleotide-sized, soluble
DNA derived from the circulation and may be useful in the detection of colorectal


maintain cellular homeostasis by excreting harmful DNA from cells. *Nature communications* 8: 15287.


Extracellular Vesicles Causes Chronic Myeloid Leukemia in Immunodeficient Mice.

*PLOS ONE* 9: e105200.


155. Zhang, J. 2017. CNTools: Convert segment data into a region by sample matrix to allow for other high level computational analyses. *R package version 1.34.0.*


Lynx: a knowledge base and an analytical workbench for integrative medicine. *Nucleic acids research* 44: D882-887.


VITA

Paul Alexander Kurywchak was born the son of Dawn and Steve Kurywchak in Phoenix, Arizona. After receiving his diploma at McClintock High School, Tempe, Arizona in 2009, he attended Arizona State University in Tempe, Arizona. Paul graduated *magna cum laude* and received the degree of Bachelor of Science with a major in Molecular Biosciences and Biotechnology from ASU in May of 2013. In August of 2013, Paul entered The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences and joined the laboratory of Dr. Raghu Kalluri to pursue a Ph.D. in Cancer Biology.

Permanent address:

3125 S. Evergreen Rd.

Tempe, Arizona 85282