AN UNBIASED MULTI-OMICS APPROACH FOR THE DETECTION OF PANCREATIC CANCER BIOMARKERS UTILIZING ION-MOBILITY MASS SPECTROMETRY AND NANOPARTICLE-BASED PROTEOGRAPH TECHNOLOGY

Manoj Khadka, Kavya Swaminathan, John Blume, Joon Yong-Lee, Sai Ramaswamy, Yuya Kodama, Brian Young, Chinmay Belthangady, Manway Liu, Mi Yang, Preethi Prasad, Philip Ma, and Bruce Wilcox; **Corresponding Author:** bruce.wilcox@prognomiq.com

PrognomiQ Inc., 1900 Alameda de las Pulgas, Suite 100, San Mateo, CA, USA

INTRODUCTION

- Pancreatic cancer is the seventh leading cause of cancer-related death worldwide and the third leading cause of cancer-related death in the USA.[1]
- Pancreatic cancer typically results in poor disease prognosis at diagnosis since it usually goes undetected until the cancer has progressed to irreversible levels.
- Challenges in early detection have led to poor survival rates, highlighting the need for early diagnostic test development.[2]
- Although some tests are available, none have enough sensitivity and specificity for diagnostic early detection of cancer (CA19-9) and can require invasive biopsies.
- Early detection of pancreatic cancer may lead to better prognosis with improved patient outcomes.
- With recent technological advancements, it is now possible to perform systems biology level analysis leveraging a multi-omics approach to survey thousands of unique potential biomarker molecules from blood.
- Combining these orthogonal 'omics technologies is key to developing new early disease screening tests.

MULTI-OMICS WORKFLOW

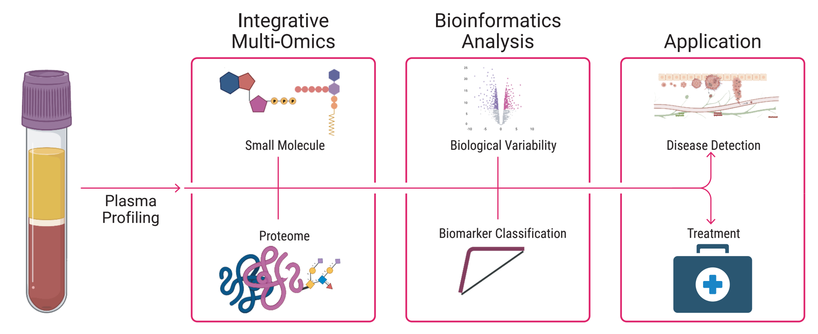

Integrative Multi-Omics
Bioinformatics Analysis
Application
Plasma Profiling
Small Molecule
Proteome
Biological Variability
Biomarker Classification
Disease Detection
Treatment

FIGURE 1

- In this case-control study, we measured the plasma proteome and lipidome of 196 human plasma subjects comprising 92 pancreatic cancer subjects and 104 healthy subjects (3 subjects were excluded from the analysis due to failed quality control metrics).
- Initial analysis of this study suggests that unique biological signatures of pancreatic cancer can be inferred using our multi-omics platform based on our findings of multiple statistically significant analyte differences between pancreatic cancer and healthy subjects.
- Future work includes building a predictive model for early cancer detection by combining analyte classes and leveraging additional 'omics data from blood.

EXPERIMENTAL METHOD

The study cohort comprised 196 human subjects (92 pancreatic cancer, 104 healthy) enrolled at various collection sites that were diagnosed as pancreatic cancer or enrolled as healthy controls. The subjects under cancer cohort were staged as 1, 2, 3 or 4, either prior or after enrollment. Subjects were part of the study "A Prospective Blood Sample Collection Study to Evaluate a Panel of Protein-based Biomarkers" which was approved by the Institutional Review Board(s) and all subjects provided written informed consent before samples were collected.

For the proteomics workflow, plasma samples were processed through the Proteograph using the standard five nanoparticle panel. Peptide samples were separated using the 60 samples per day method on the Evosep One LC system fitted with a traditional PepSep 8 cm C18 column. Samples were subjected to LC-MS/MS analysis on a timsTOF Pro 2 using Data Independent Acquisition mode with Parallel Accumulation-Serial Fragmentation (diaPASEF). The data were analyzed using DIA-NN implemented on Proteograph Analysis Suite (PAS).

For the lipidomics workflow, lipids were extracted from plasma using a 1:1 v/v butanol/methanol mixture. After centrifugation, 2 µL of clean lipid extract were injected into a Bruker Elute UPLC and separated on a Waters UPLC column. Mass spectrometry data was collected on a Bruker timsTOF Pro 2 under positive ionization mode utilizing DDA-PASEF. Data was preprocessed utilizing Bruker Compass Metaboscape 8.0.1 to detect, deconvolute and annotate lipid species. Lipids were identified based on exact mass and MS/MS information.

prognomiQ

www.prognomiq.com

METHODS

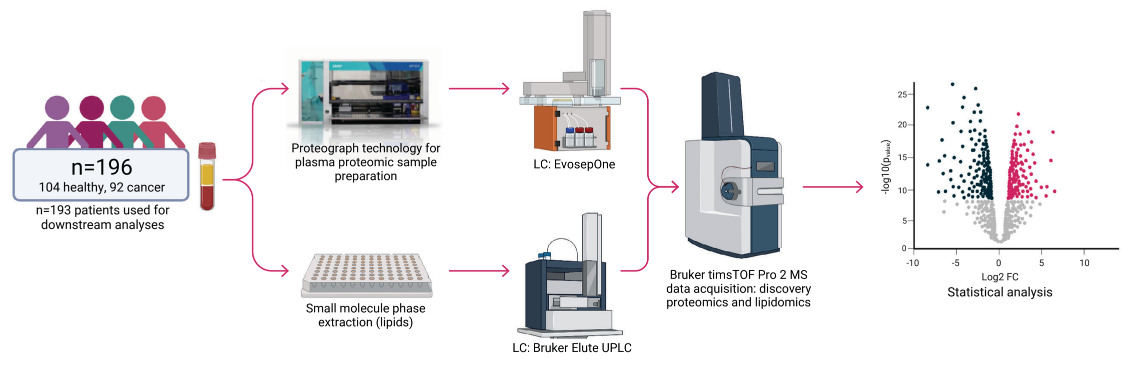

n=196
104 healthy, 92 cancer
n=193 patients used for downstream analyses
Proteograph technology for plasma proteomic sample preparation
LC: EvosepOne
Small molecule phase extraction (lipids)
LC: Bruker Elute UPLC
Bruker timsTOF Pro 2 MS data acquisition: discovery proteomics and lipidomics
Log2 FC
Statistical analysis

FIGURE 2

Experimental overview illustrating study design, sample processing, data acquisition, and downstream processing of proteomics and lipidomics data utilizing Seer Proteograph, Evosep, and Bruker platforms.

RESULTS: PROTEINS DETECTED ACROSS SAMPLES

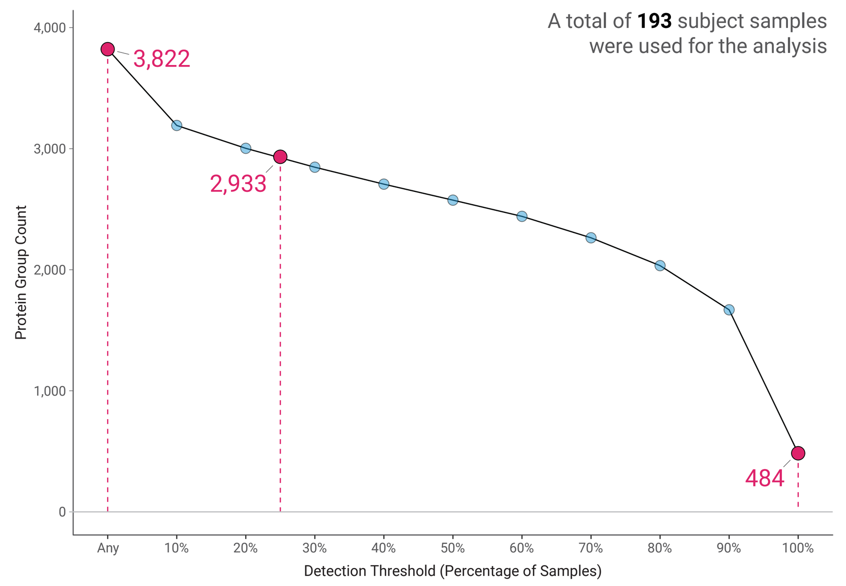

A total of 193 subject samples were used for the analysis
3,822
2,933
484
Protein Group Count
Detection Threshold (Percentage of Samples)

FIGURE 3A

In our proteomics workflow, we detected 3,822 protein groups across all 5 nanoparticles for the 193 pancreatic cancer and healthy subject samples using DIA-NN. We found 2,933 protein groups were identified in 25% of cohort, and 484 proteins were consistently identified in 100% of the cohort.

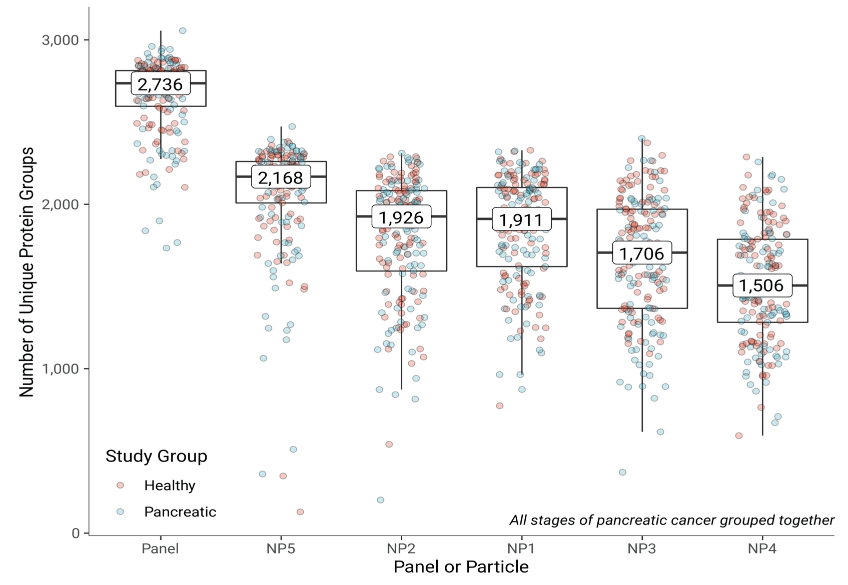

2,736
2,168
1,926
1,911
1,706
1,506
Number of Unique Protein Groups
Study Group
Healthy
Pancreatic
All stages of pancreatic cancer grouped together
Panel
NP5
NP2
NP1
NP3
NP4
Panel or Particle

FIGURE 3B

We detected a median of 2,736 protein groups across all 5 nanoparticles for the 193 subjects in this study, comparable to previously reported results where an average of 1664 proteins were detected.3 Nanoparticle 5 (NP5) provides the largest number and most diverse protein groups (2,168) amongst all the nanoparticles.

PROTEINS DETECTED IN PANCREATIC CANCER STUDY MATCHED TO HPPP DATABASE

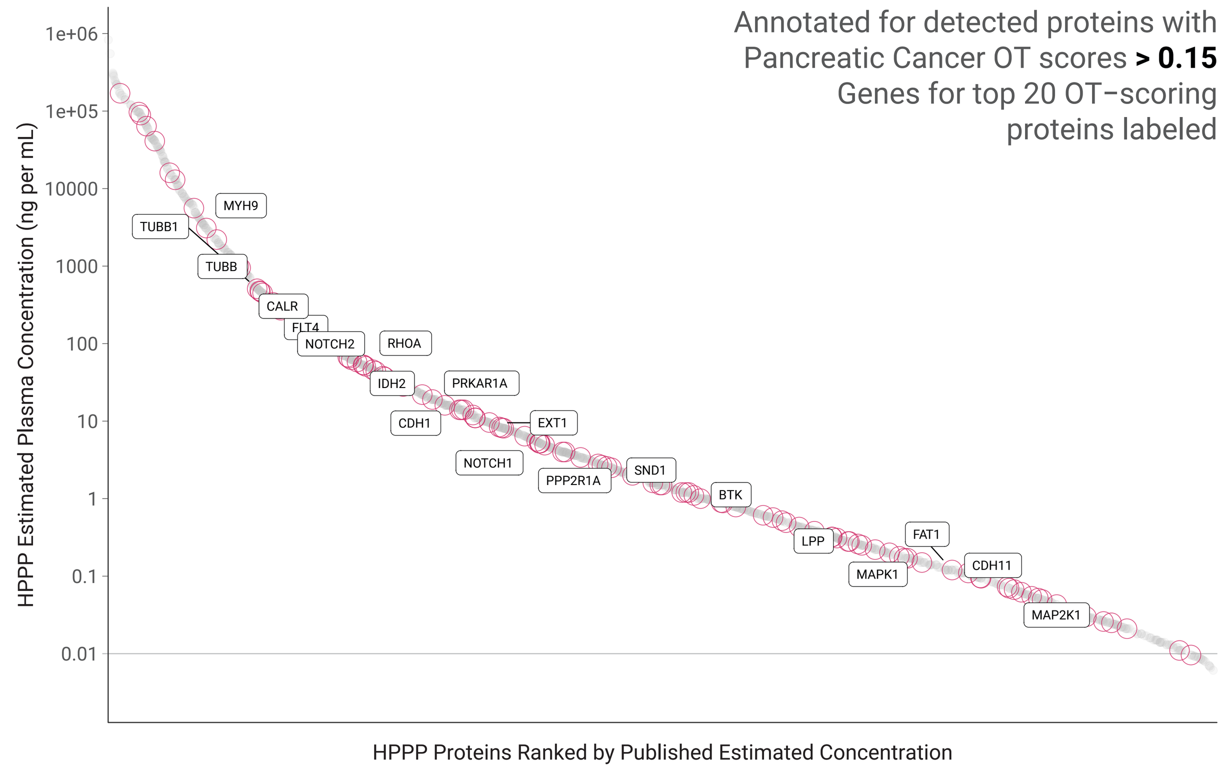

Annotated for detected proteins with Pancreatic Cancer OT scores > 0.15
Genes for top 20 OT–scoring proteins labeled
HPPP Estimated Plasma Concentration (ng per mL)
HPPP Proteins Ranked by Published Estimated Concentration

FIGURE 4

The 3,822 detected protein groups were mapped to the HPPP database. Identified proteins had concentrations ranging 8 orders of magnitude. All identified proteins are displayed above (grey) with proteins having significant pancreatic cancer OT scores <0.15 highlighted (pink).

VOLCANO PLOTS

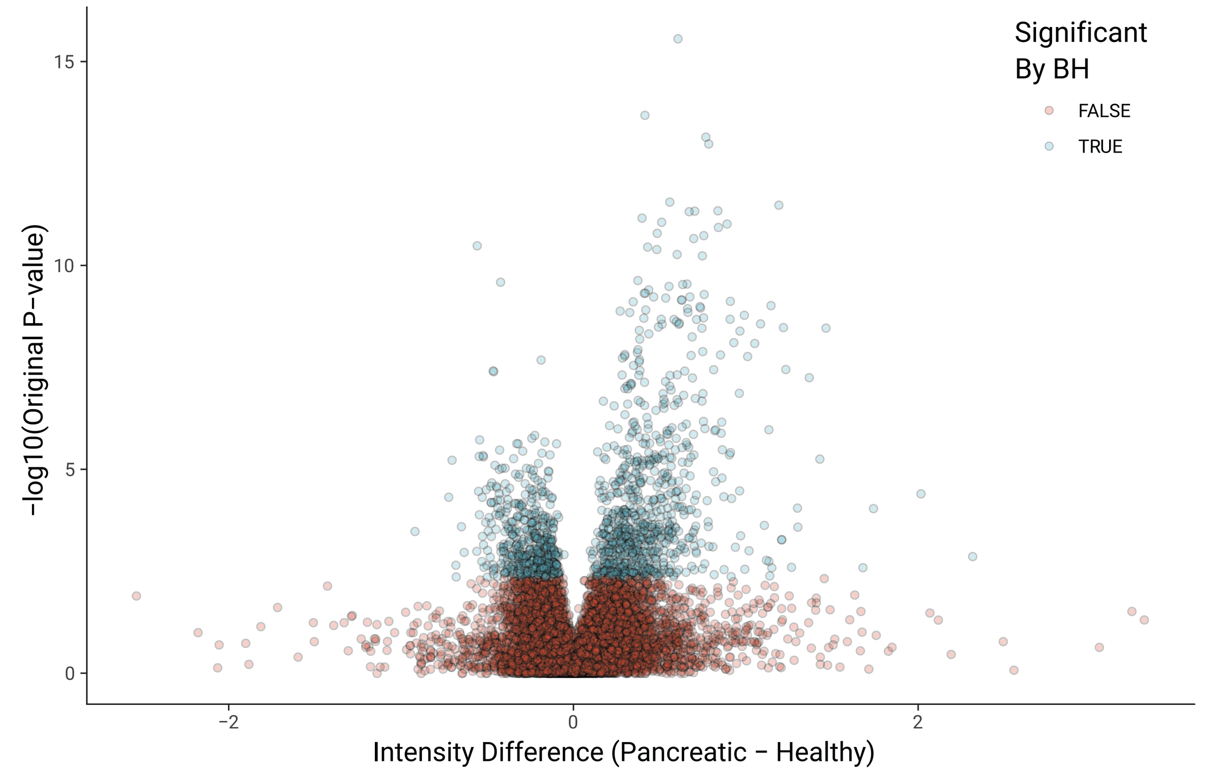

Significant By BH
FALSE
TRUE
−log10(Original P−value)
Intensity Difference (Pancreatic − Healthy)

FIGURE 5

The volcano plot indicates 124 out of a total of 3,822 protein groups were statistically significantly different between healthy and pancreatic cancer subjects calculated based on Wilcox test with Benjamini-Hochberg correction (p = 0.05). Significance testing using multiple-testing correction with 0.05 threshold using non-imputed data with at least 3 measures per class were used in this analysis.

VOLCANO PLOTS (CONT.)

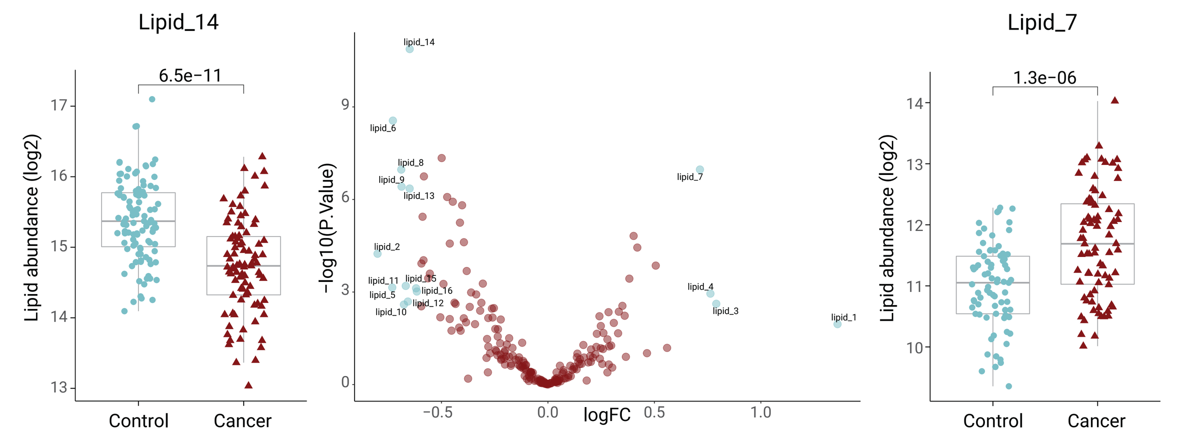

Lipid_14
6.5e−11
Lipid_7
1.3e−06
Lipid abundance (log2)
−log10(P.Value)
logFC
Control
Cancer

FIGURE 6

Volcano plot shows differential abundance of lipid species between normal and pancreatic cancer subjects calculated based on t-test and Benjamini-Hochberg correction. We found 16 of 259 lipids species to be significantly different (adjusted p-value<0.05) between healthy and pancreatic cancer subjects. Representative boxplots for two of the lipid species depict the differential abundance between healthy and cancer subjects.

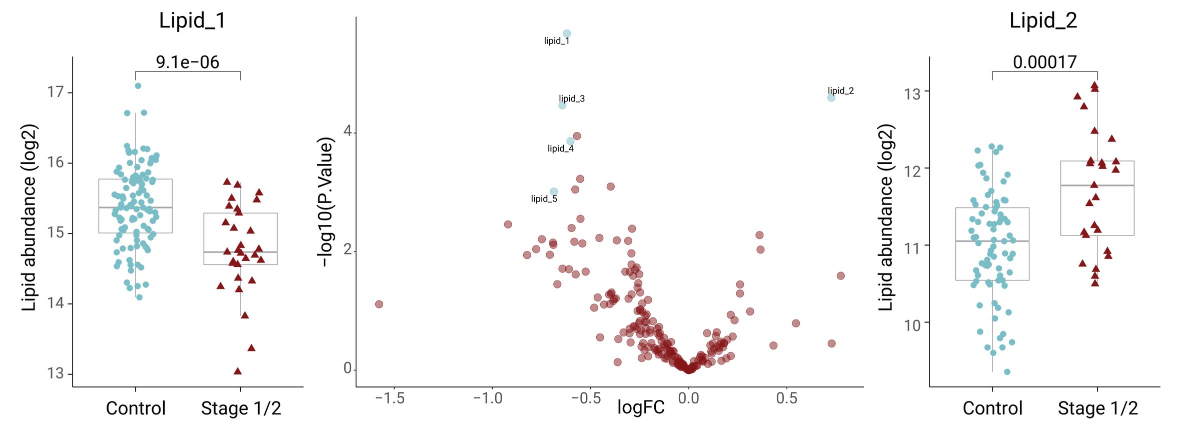

Lipid_1
9.1e−06
Lipid_2
0.00017
Lipid abundance (log2)
−log10(P.Value)
logFC
Control
Stage 1/2

FIGURE 7

Volcano plot shows differential abundance of lipid species between normal and pancreatic cancer (Stages 1 and 2) subjects calculated based on t-test and Benjamini-Hochberg correction. We found 5 of the 259 lipids species to be statistically significantly altered (adjusted p-value<0.05) between healthy and pancreatic cancer (stage 1 and 2) subjects. Representative boxplots for two of the lipid species depict the differential abundance between healthy and early-stage cancer subjects.

CONCLUSIONS

- Our multi-omics platform facilitated the evaluation of the global proteome and lipidome of our pancreatic cancer cohort and identified multiple putative biomarker candidates across analyte classes for early disease detection.
- Untargeted DIA proteomics data yielded 124 statistically significant proteins out of a total of 3,822.
- Untargeted lipidomics data showed 16 of 259 lipids species that were significantly different between healthy and pancreatic cancer subjects.
- We found 5 out of 259 lipid species between healthy and Stage 1 and 2 cancer subjects to be statistically significantly different which highlights the detection of a biological signal related to early stages of pancreatic cancer.
- Further work is focused on further combining complimentary analyte classes and integrating other analytes from genomics, transcriptomics, metabolomics, DNA methylomics, and glycomics data.

REFERENCES

[1] McGuigan, A. et. al. World J Gastroenterol. 24(43):4846-4861 (2018)
[2] Pashayan, N. et. al. Science. 368:589-590(2020)
[3] Blume J. et. al. Nat Commun. 11:3662(2020)
[4] Farrah T. et. al. Mol and Cell Prot. 10(9):1014(2011)
[5] Liu J-F. et. al. Biochem Pharmacol. 155:537-546(2018)
[6] Zhao, Z. et. al. J of Clin Onco. 25(19):2696-2701

ACKNOWLEDGMENTS

We would like to acknowledge Xuejun Peng and Heiko Neuweger from Bruker for their collaboration in data acquisition and data analysis.