Early Pancreatic Cancer Detection Using Extracellular Vesicle DNA Methylation Signatures in Blood

Sheng Zhang[1], Hongzhang He[1,2]*, Siyang Zheng[1]*

[1]Carnegie Mellon University, Biomedical Engineering, Pittsburgh, PA, USA. [2]Captis Diagnostics, Pittsburgh, PA, USA (contact: Siyang Zheng, siyangz@andrew.cmu.edu and Hongzhang He, hongzhah@andrew.cmu.edu)

In 2020, there were 495,773 new cases and 466,003 deaths for pancreatic cancer worldwide. Furthermore, pancreatic cancer is emerging to be the leading cause of cancer related deaths in the U.S. by the year 2030, surpassing lung and colon cancer.

The majority of pancreatic cancer patients (~80%) are diagnosed at the most advanced (metastatic) stage, for which the 5-year survival is 3%. Dramatically, studies show that if it can be detected at an early stage when surgical removal is feasible, the 5-year survival rate significantly increased to 37%. This data highlights the crucial importance of early detection for pancreatic cancer treatment with curative intent. Unfortunately, currently there are no reliable techniques for routine screening to identify pancreatic cancer at early stages. In this Abstract, we used a machine learning based feature selection procedure to extract an optimal set of DNA methylation markers to differentiate cancer from normal pancreatic tissue and normal blood. Then, we validated the performance of those DNA methylation markers to detect pancreatic cancer in a cohort of pancreatic cancer patients.

Firstly, we collected a large amount of genome-wide methylation microarray data (targeting over 450,000 methylation sites) of tissue samples from several publicly available databases, including samples of solid pancreatic tumors (n = 384), noncancerous pancreatic tissues (n = 126) and healthy whole blood (n = 453). The raw microarray data were preprocessed using a pipeline implemented in R, which removes inaccurate measurements and corrects for bias caused by technical issues. A two-stage feature selection procedure was designed and implemented in Python, which makes use of various filters and a multivariate embedded method, support vector machine recursive feature elimination (SVM-RFE)(*1*). The features were evaluated using two metrics, classification performance and stability, based on a tradeoff of which we selected a set of differentially methylated loci for discriminating the samples. The set of prospective markers includes 129 hypermethylated and 142 hypomethylated tumor-specific loci, 8 hypermethylated and 8 hypomethylated pancreatic tissue specific loci.

To validate the performance of those pancreatic cancer-specific methylation markers, as a feasibility study, we isolated extracellular vesicles from cancer cells and plasma of cancer patients using lipid nanoprobe (LNP)(*2*). Extracellular vesicles (EVs) are lipid-bilayer-enclosed vesicles of sub-micrometer size that are secreted by virtually all cell types. In a pilot study encompassing 11 pancreatic cancer patients (Stage I-IV) and age-matched 10 healthy donors (HD), we employed quantitative methylation-specific polymerase chain reaction assays (qMSP)(*3*) to examine the promoter methylation status of the genes within plasma EV DNA. The receiver operating characteristic (ROC) curves were employed to assess the performance of a combined promoter methylation signature in EV DNA for the detection of pancreatic cancer. Our findings indicate that the promoter methylation of genes can be specific to pancreatic cancer, suggesting the efficient packaging of methylation information from genomic DNA into EVs.

References: 1. I. Guyon, J. Weston, S. Barnhill, V. Vapnik, *Machine Learning* **46**, 389-422 (2002). 2. Y. Wan *et al.*, *Nature Biomedical Engineering* **1**, 0058 (2017). 3. V. J. Bailey *et al.*, *Clinical Chemistry* **56**, 1022-1025 (2010).