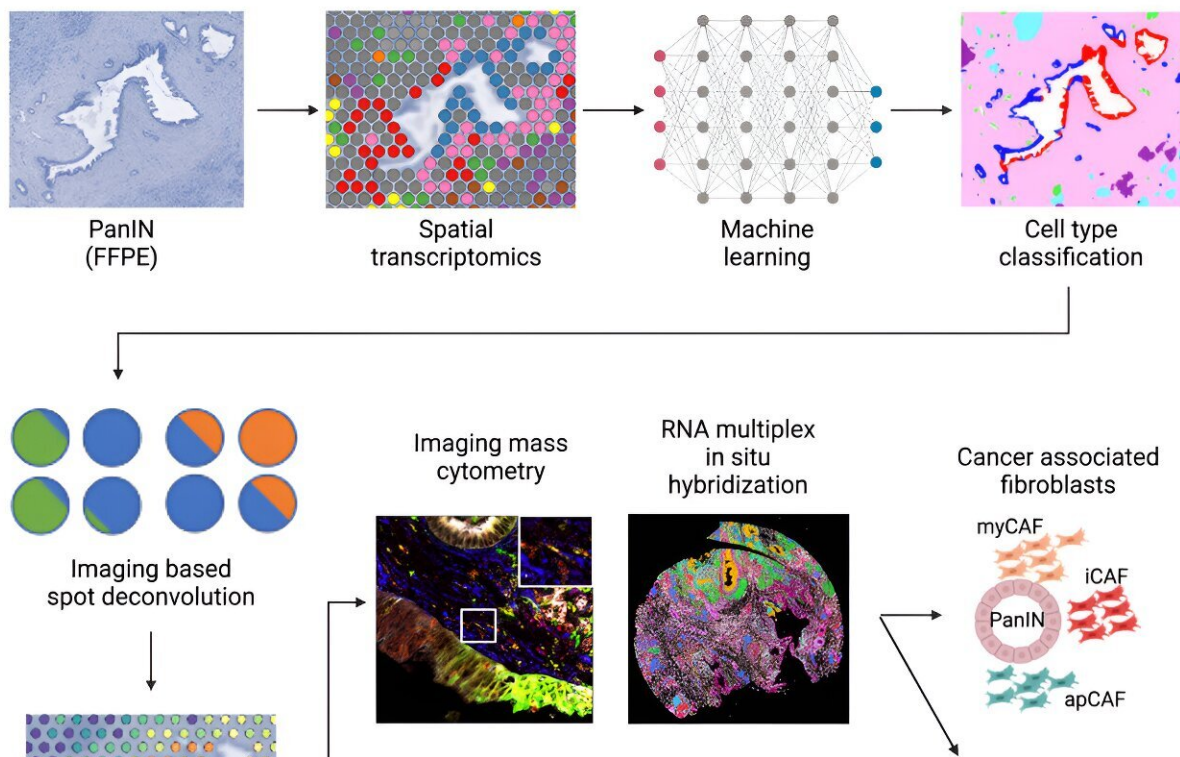# Identifying key markers in pancreatic cancer progression using a new analysis pipeline

August 29 2024

Novel computational pipeline developed by Johns Hopkins Kimmel Cancer Researchers to identify early molecular and cellular progression to pancreatic ductal adenocarcinoma. Credit: *Cell Systems* (2024). DOI: 10.1016/j.cels.2024.07.001

Using a new workflow that integrates spatial transcriptomics and

machine learning for imaging analysis and integration with single-cell datasets, researchers at the Johns Hopkins Kimmel Cancer Center have identified novel molecular and cellular markers in the development of one of the most aggressive, deadly pancreatic cancers: pancreatic ductal adenocarcinoma (PDAC).

PDAC arises from precancerous lesions in the pancreas. One of these lesion types, pancreatic intraepithelial neoplasias (PanINs), can appear in the pancreas years before they progress to invasive cancer. Because PanINs are so small, they cannot be detected by conventional clinical imaging tests.

Previous analysis methods, such as bulk sequencing and single-cell sequencing, can capture gene expression of cancer cells and other cell types in the tumor microenvironment. However, what is missing is the spatial relationships among these types of cells within and around tumors, says senior study co-author Luciane Kagohara, Ph.D., an assistant professor of oncology at the Johns Hopkins University School of Medicine.

Elana Fertig, Ph.D., professor of oncology, director of the oncology quantitative sciences division, and co-director of the Convergence Institute and the Single-Cell Training and Analysis Center at the Johns Hopkins University School of Medicine, was also a senior study co-author.

Using spatial transcriptomics—a technique used to measure and map gene expression in a tissue section—the researchers developed a three-way analysis pipeline to map changes in gene expression from nine patients in 14 PanINs, including five rare, high grade PanINs. The machine learning tools used for imaging analyses (CODA) and for the integration with PDAC single-cell datasets, through the innovative multiomics integration methods CoGAPS (coordinated gene activity in

pattern sets) and projectR, were developed in previous studies at Johns Hopkins.

The investigators have made their data and code available to other researchers to uncover further insights into [pancreatic cancer](link) and, through open-source tools, to adapt the analysis methods for spatial transcriptomics for biomedical research at large.

Results of their work, [published](link) Aug. 7 in *Cell Systems*, provide new insights into the [gene expression](link) and spatial distribution of different types of cells in the precancerous environment around PanINs. The findings are key to understanding how PanINs progress to PDAC, laying a foundation for future early detection of this and other types of pancreatic cancer.

The three-way analysis revealed that some key features of pancreatic cancer were present in PanINs.

"When we looked at the progression of normal to high grade PanIN lesions, we found that cell proliferation gradually increases while inflammatory signaling decreases, which can be important to understand the intrinsic low immunogenicity of these tumors," says Kagohara. This suggests that the cells in and around PanINs are already creating a more immunosuppressive environment before the invasive PDAC is fully developed, she says.

The findings also revealed spatial differences in the cells. "We found that cancer-associated fibroblasts, which play a major role in the biology of PDAC and their response to treatments, were already present at the premalignant stage, consistent with previous research in animal models of pancreatic cancer but not previously observed in humans," says Kagohara.

"PanINs are pretty small—less than 1 millimeter in size—so we were very surprised that even using very few cells, we were able to detect strong signatures from those lesions," she adds.

"That's very important in cancer because, for example, if we want to understand immune responses to treatments, we need to know which immune cells are closer to or further from the tumor. But we also need to understand if there are other cell types that are blocking these communications by forming a physical barrier or interacting with the neoplastic cells."

Beyond the main findings, the study highlights the power of computational tools to further understanding of human disease.

Spatial transcriptomics is a very powerful technology for generating data, says Kagohara, but the right computational tools are needed to analyze it. "With the right tools, you can enhance the results that you can extract from these novel technologies," she says.

To encourage more collaboration and analyses, the study's data is available on GEO (GSE254829), and the code is available on GitHub.

"As we generate more and more data on PanINs, the idea is to identify molecular signatures for the development of early detection tests," says Kagohara. "This is one of the first studies to generate a resource open to other researchers in the search for these early markers in PanINs so we can learn more about how the spatial distribution of molecular and cellular signatures affect the progression of pancreatic cancer and the response to treatments."

Other study co-authors include Alexander Bell, Jacob Mitchell, Ashley Kiemen, Melissa Lyman, Kohei Fujikura, Jae Lee, Eron Coyne, Sarah Shin, Sushma Nagaraj, Atul Deshpande, Pei-Hsun Wu, Dimitrios

Sidiropoulos, Rossin Erbe, James Chell, Lauren Ciotti, Jacquelyn Zimmerman, Denis Wirtz, Won Jin Ho, Neeha Zaidi, Elizabeth Thompson, Elizabeth Jaffee and Laura Wood of Johns Hopkins. Researchers from 10x Genomics of Pleasanton, California, contributed to the work.

More information: Alexander T.F. Bell et al, PanIN and CAF transitions in pancreatic carcinogenesis revealed with spatial data integration, *Cell Systems* (2024). DOI: 10.1016/j.cels.2024.07.001