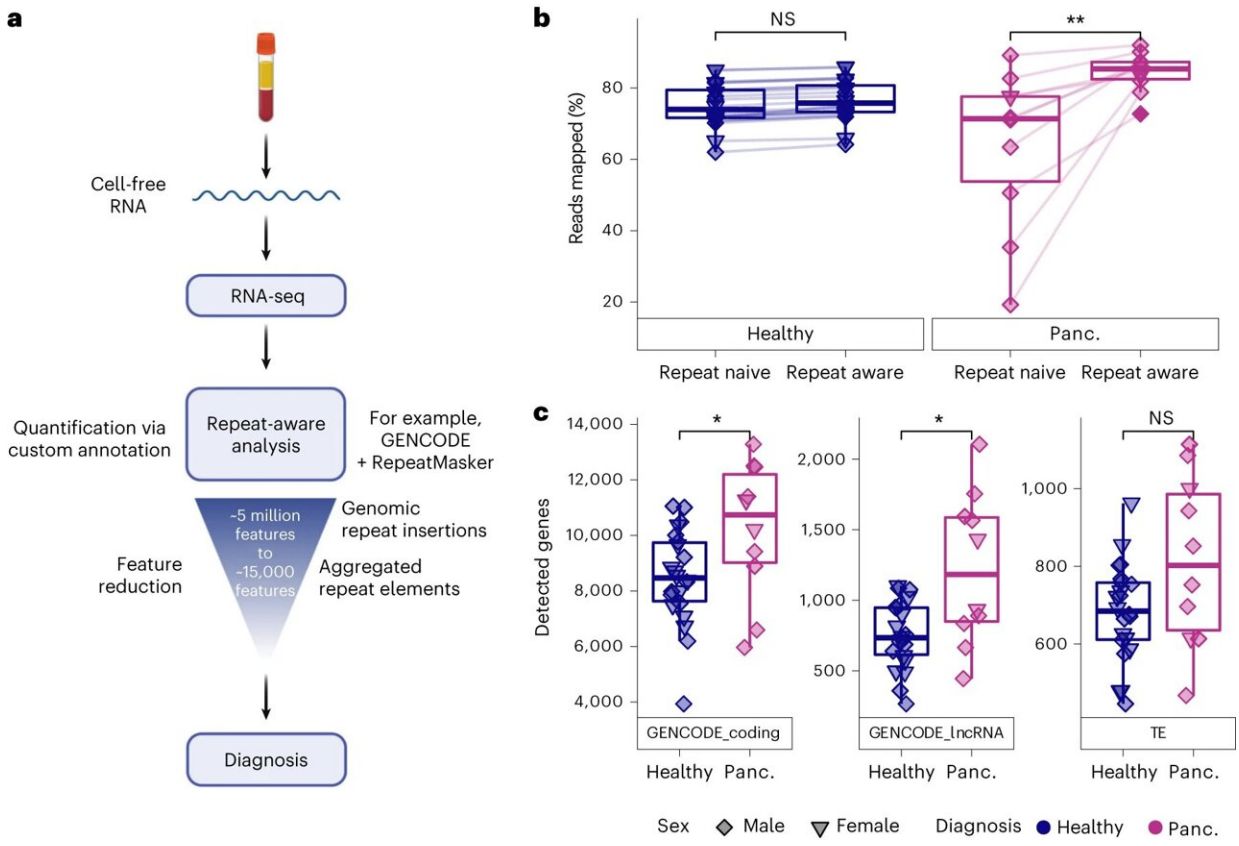
New blood test for noncoding RNA significantly improves cancer detection

August 31 2023



Cell-free RNA transcriptome profiling using repeat-aware COMPLETE-seq. **a**, Diagram of COMPLETE-seq RNA liquid-biopsy technology, highlighting the use of repeat-derived cell-free RNAs aggregated into a tractable feature set to enable diagnostic modeling. Created with BioRender.com. **b**, Comparison of mapping rates between use of a repeat-naive (GENCODE v.39) reference annotation (**$P = 0.0039$) and repeat-aware reference annotation (Wilcoxon, paired, two-sided). **c**, Comparison of gene detection distributions for each cohort across coding genes (GENCODE_coding; *$P = 0.043$), lncRNAs

(GENCODE_lncRNA; *$P = 0.035$) and TE subfamilies (Wilcoxon, two-sided). For the box plots, the center line represents the median, the box limits are upper and lower quartiles and whiskers represent 1.5× interquartile range. NS, not significant; panc., pancreatic cancer. Credit: *Nature Biomedical Engineering* (2023). DOI: 10.1038/s41551-023-01081-7

Cancer is most treatable in its early stages, so finding innovative and non-invasive methods to diagnose cancer early on is crucial for fighting the disease. Liquid biopsies, which require just a simple blood draw, are an emerging technology for non-invasively testing for cancer using DNA or RNA sequencing of a patient's blood.

Assistant Professor of Biomolecular Engineering Daniel Kim and his lab are developing more accurate and powerful liquid biopsy technologies that take advantage of signals from RNA "dark matter," an understudied area of the genome. Kim's new research shows that this genetic material is present in the blood of people with cancer and can be identified to diagnose specific cancer types such as pancreatic, lung, esophageal, and others early in the course of the disease.

Kim's lab developed an RNA liquid biopsy platform that detects both protein-coding RNA and RNA dark matter in the blood, and showed that this new approach significantly improves the performance of liquid biopsy for cancer diagnosis. This research was published today in the journal *Nature Biomedical Engineering*.

Focus on 'dark matter'

While most researchers and companies are pursuing DNA-based liquid biopsy for cancer diagnosis, Kim's approach is unique in its focus on RNA "dark matter," specifically noncoding and repetitive RNA.

Most of the three billion base pairs of DNA that make up the human genome are transcribed into RNA, and all of the RNA is collectively known as the transcriptome. The most commonly recognized function of RNA is to code for proteins in the body, but 75% of the human genome generates noncoding RNA that does not code for proteins.

A substantial portion of these noncoding RNAs are derived from repetitive elements, and these RNAs can travel out of the cell from which they originate and into the bloodstream. A healthy individual's blood typically would have very few of these repetitive noncoding RNAs. However, Kim's research has shown that even at the earliest stages of cancer, many of these repetitive RNAs are secreted out of cancer cells, making them potent biomarkers of early-stage disease.

RNA liquid biopsy technology developed by the Kim lab aims to detect cancer by sequencing "cell-free RNA" in a patient's blood to test for the presence of both protein-coding and repetitive noncoding RNA.

Kim's lab created a cell-free RNA sequencing and analysis platform called COMPLETE-seq to identify repetitive noncoding RNAs that are typically overlooked. After a patient's blood is drawn, this comprehensive approach analyzes the sample for all of the annotated areas of the transcriptome—the tens of thousands of RNAs that have already been well-documented— plus all of the five million noncoding repetitive elements that Kim's lab also focuses on.

"If you look at these different cancers, each has its own characteristic cell-free RNA profile, but a lot of these RNAs are coming from the millions of repeat elements that are found throughout the genome," Kim said.

"What we found was that when we trained machine learning models for cancer classification, the models perform better when you introduce

these repetitive cell-free RNAs as additional features. We see higher sensitivity in terms of detecting cancer, so we think that these repeat elements are actually providing a lot of rich cell-free RNA information that people previously hadn't looked for."

Improving tests

Other existing liquid biopsy tests have not been very sensitive for early stage cancer, with some tests missing up to 75% of stage I cancers, when the biological signal is low due to the small tumor size. Kim's paper shows that incorporating repetitive RNA into their liquid biopsy platform greatly increases the biological signal and boosts the performance of machine learning models tasked to identify cancer. As an example, using COMPLETE-seq improved performance to 91% sensitivity for identifying colorectal cancer.

"The value of our study is that we've now shown the potential of these repeat elements for diagnosing disease, so hopefully there'll be a lot of interest in leveraging repetitive RNAs to boost the sensitivity of these multi-cancer early detection tests," Kim said.

The research findings show that this technology can be used to identify a variety of cancer types. The lab initially focused on pancreatic cancer for this study, as there is an urgent clinical need for pancreatic cancer early detection, as late detection leads to worse outcomes for patients. Pancreatic cancer is also known to be driven by mutations in the KRAS gene, which is also a focus of Kim's lab.

After verifying findings in pancreatic cancer, the researchers also looked at a variety of other cancers, and plan to look at many more cancer types with additional samples across the progressive stages of cancer. The team is interested in collaborating with clinicians and companies to do this.

Kim's goal is to develop an RNA liquid biopsy test for multi-cancer early detection, using the rich information from repetitive RNAs to identify and diagnose disease with high sensitivity and specificity. Kim hopes his platform will not only diagnose cancer at the earliest stages but also help guide individualized, precise treatment strategies when the cancer is more treatable.

Moreover, his test could help to identify a recurrence of cancer, and also be used to study aging and to diagnose other types of diseases that alter the repetitive RNA landscape, such as Alzheimer's disease. He recently gave the UCSC Kraw Lecture on "Precision Health for All Through RNA" that laid out his vision for early detection and precise treatment of disease using RNA.

The researchers also used nanopore sequencing to read the cell-free RNAs floating in the blood, which allowed them to generate long-reads and determine the true length of these cell-free RNAs. Kim believes his lab is the first to use nanopore sequencing, a technique pioneered at UC Santa Cruz, for RNA liquid biopsies to diagnose cancer and to determine the full length of these cell-free RNAs.

Nanopore sequencing can be performed on a handheld device developed by Oxford Nanopore Technologies called the MinION. This holds promise for carrying out cancer screening in remote or resource-poor settings where larger, more expensive sequencers are not readily available.

"This study would not have been possible without the strong support of the American Cancer Society and all of its generous donors, leadership, staff, and volunteers, as well as all of the hard work of my Ph.D. student Roman Reggiardo (an NIH F99/K00 Fellow now at HHMI Investigator Howard Chang's lab at Stanford University) and all of our Kim lab members and collaborators," Kim said.

In addition to his role as an Assistant Professor in the Baskin School of Engineering associated with the Institute for the Biology of Stem Cells, the Genomics Institute, and the Center for Molecular Biology of RNA at UC Santa Cruz, Kim is also an Associate Member of the Canary Center at Stanford for Cancer Early Detection and a Research Scholar of the American Cancer Society.

More information: Daniel Kim et al, Profiling of repetitive RNA sequences in the blood plasma of patients with cancer, *Nature Biomedical Engineering* (2023). DOI: 10.1038/s41551-023-01081-7
www.nature.com/articles/s41551-023-01081-7

Provided by University of California - Santa Cruz