Clinical Chemistry

Pan-Cancer Detection and Typing by Mining Patterns in Large Genome-Wide Cell-Free DNA Sequencing Datasets

Article:
Huiwen Che, Tatjana Jatsenko, Liesbeth Lenaerts, Luc Dehaspe, Leen Vancoillie, Nathalie Brison, Ilse Parijs, Kris Van Den Bogaert, Daniela Fischerova, Ruben Heremans, Chiara Landolfo, Antonia Carla Testa, Adriaan Vanderstichele, Lore Liekens, Valentina Pomella, Agnieszka Wozniak, Christophe Dooms, Els Wauters, Sigrid Hatse, Kevin Punie, Patrick Neven, Hans Wildiers, Sabine Tejpar, Diether Lambrechts, An Coosemans, Dirk Timmerman, Peter Vandenberghe, Frédéric Amant, and Joris Robert Vermeesch.
Pan-Cancer Detection and Typing by Mining Patterns in Large Genome-Wide Cell-Free DNA Sequencing Datasets
Clin Chem 2022;68(9): 1164–76. https://doi.org/10.1093/clinchem/hvac095

Guest: Dr. Joris Vermeesch is Professor of Molecular Cytogenetics and Genome Research and Director of the Genomics Core Facility at the Catholic University of Leuven in Belgium.

Bob Barrett: This is a podcast from *Clinical Chemistry,* sponsored by the Department of Laboratory Medicine at Boston Children's Hospital. I am Bob Barrett.

Analysis of cell-free DNA is a promising non-invasive biomarker in liquid biopsy for cancer management. In 2019, a paper reported that cancer-specific copy number aberrations were indicative of early tumors and had potential as a cancer screening tool. The authors of that study have followed up their work by data mining the patterns of cell-free DNA by shallow whole-genome sequencing data sets from patients with cancer, in hope that could improve cancer detection. That paper, titled "Pan-Cancer Detection and Typing by Mining Patterns in Large Genome-Wide Cell-Free DNA Sequencing Datasets" appears in the September 2022 issue of *Clinical Chemistry*.

The senior author of that paper is Dr. Joris Vermeesch. He is Professor of Molecular Cytogenetics and Genome Research and director of the Genomics Core Facility at the Catholic University of Leuven in Belgium. We are pleased to have Dr. Vermeesch as our guest in this podcast. First of all, Doctor, liquid biopsy to monitor cancers is a hot topic, as is machine learning. Your study bridges these two areas. So, what makes this study special?

Joris Vermeesch: Okay, thanks for the question. I think the liquid biopsy field is indeed exploding in the realization that tumors shed DNA in the bloodstream, opened up new avenues for monitoring and diagnosing, screening cancer. We actually observed that whereas majority of the field is focusing on targeted analysis of tumor specific variants, we noticed that if a tumor sheds DNA, you get actually a pattern of the fragmentation of the specific tumor cells in the blood, and these patterns are tumor

specific. And so, by doing just a genome-wide analysis, we can deduce those patterns.

What makes our study unique is the fact that the observation that if you have large amount of data from normal individuals, non-cancerous individuals, and people with cancer, that we can based on the pattern recognition. So, we use the machine learning to do pattern recognition, which allows us, not only which of the individuals have cancer and don't have a cancer, but even more we can actually map the profiles and deduce which cancer the person is carrying based on the fragments of DNA in the blood. So, I think there's more or less unique, or the uniqueness of our study is one, we use genome-wide analysis whereas the majority of the studies use target analysis. And secondly, we use the fact that we had large amount of data, large amount of data sets from different individuals to observe these patterns. If you would look at individual samples, you cannot see that immediately.

Bob Barrett: What were some of the unique results that you and your co-workers found in this study?

Joris Vermeesch: I started a little bit with a joke, which was not in the publication. But we first observed that we could deduce disease state, let's say over an aberrant status in an individual by using samples, this low-pass sequencing data of cfDNA, circulating, free-floating DNA, of individuals who are pregnant and comparing those with the fragmentome patterns of those who were not pregnant. Why do I take that example? Well, we could actually see who was pregnant and who's not. So, the approach we have is a great pregnancy test, although obviously of no use because anyone who is three months pregnant would know, or majority of people at least know, that they're pregnant. So, it's of no avail. But the point I want to make is that the pregnancy itself is not a disease state. The only difference is that there's a placenta. The placenta is chromosomally or genomically perfectly normal, but the fragments that are derived from that can be picked up.

There was one part, but obviously not the major part of this particular paper, where we zoom into cancer. So, we went down by looking at the sensitivity and specificity of -- or the ability to detect. First, we looked into lymphomas, leukemia, so blood cancers. And there, we had a very, very high accuracy. Meaning, sensitivity and specificity over 95%, even in the low grade leukemias and lymphomas. That may be surprising, definitely the high accuracy is very impressive and was much higher than any existing methodology at this moment, at least genome-wide methodologies, where you don't have a priori knowledge about the mutations in the leukemias or lymphomas. The current gold standard is based on copy number variations, which are present in the majority

of the cancers, but some cancers do not have any chromosomal abnormalities. And then we have no means of, just by looking at the genome-wide analysis profile, whether or not there is a cancer.

But by now, looking at this fragment or patterns, we can, with the sensitivity and specificity higher than 95%. If we look at the solid tumors, so we looked at the series of solid tumors, there's friability. Some tumors have a very high sensitivity and specificity even we had a grade 1, grade 2, grade 3, and grade 4 tumors. Obviously, the higher the grade is easier it is to detect because the tumors are usually bigger. There are more cells dying. There's more DNA entering the bloodstream. There are more fragments in the blood and those can be more easily be detected. But even in the lower grade tumors, we have quite a high sensitivity and specificity, with ovarian tumors amongst the highest. Breast tumors may be amongst the least visible in the spectrum, but nevertheless, also there we have quite a good pickup rate.

Bob Barrett: So, it sounds like this is much more than just a very complex pregnancy test.

Joris Vermeesch: Yeah, absolutely. Also, maybe I could, coming back to the pregnancies, I can say that we do detect by this methodology, cancers in pregnancy occasionally. I'm heading the laboratory doing the non-invasive prenatal testing here in Belgium, or one of the labs, and actually this methodology picks up cancer in pregnancy, which occasionally occur. Usually, the women are not aware because fatigue, which just comes with the cancer, is also part of pregnancy, and therefore, symptoms are not necessarily recognized. I think it's an added value for the pregnant woman. Obviously, it's a secondary finding and the women are not waiting for it, but in retrospect they're always happy if they get a diagnosis early, because it improves the cancer treatment successes.

Bob Barrett: And it brings us up to clinical outcomes, which are most important in such research. How can this study affect patients?

Joris Vermeesch: I showed the example where it has immediate effect because there are different non-invasive prenatal tests available, but one test, which is often used as low-pass genome-wide sequencing of circulating fee floating DNA. And there, the test is just another way of analyzing the data and could be implemented directly. And that's actually something we are exploring at the moment. For the other applications, I think in cancer, because that's what the test is aiming for, we are looking for ways to use it or we're actually doing multiple follow-up studies to explore for which cancer types this type of analysis can be useful. The ultimate vision is that all of us can use such a screen, say every year, to detect early cancers

before you have any symptoms. That I think will happen somewhere in our lifetime. If it's not this particular test, I think other tests based on this liquid biopsy will be able to do so.

So, I think that's the ultimate aim. However, it's clear that the risk in implementing such a test right away is that we would have too low a specificity and therefore drive too many people into the oncology units to follow up where that may not be needed. So, we really need to understand the dynamics of this test better before we can do that, but I think that's where we want to go to.

Bob Barrett: Well, let's look ahead then. What are the next steps in your research? Where do you go from here?

Joris Vermeesch: Okay. So, that's the ultimate goal, but starting right away with the population screen I think is challenging or is maybe not the right time yet. I must say that we do have a pilot study where we sequenced thousand individuals between 60 and 80 years old. That's actually the time period in our life when you have the highest chance of cancer. So, where we use this, we're now looking at this data to see whether we can actually improve cancer detection in this age group. So, that's a population where there is high risk for cancer. These are populations where we can start exploring the value of our assay.

We are also looking for other situations where there's high risk for cancer. But what we're also trying to do is trying to understand how we can use this test in monitoring of cancer treatment. So, once a cancer is removed or people are treated, some people respond better than others to treatment, and by just taking a blood draw and analyzing those patterns, we can see whether there's a reaction or there's no reaction. For some cancers, it is very valuable because then you can do probably a faster change of treatment. So, that's something we are exploring.

And the second aspect we're exploring is minimal residual disease or relapse detection. Basically, when cancers are treated, we hope everything, all the cells of the cancer, are gone. But as we know, some cancer types, large number of individuals relapsed. Example is pancreatic cancer, where 50% relapse after two years of treatment to three years of treatment. So, the question is, can liquid biopsy or the methodologies we've been developing be used for earlier relapse detection before the cancer really manifests itself in a physical way or in a visible way using certain scans? This is the way we are now progressing to explore the value of this methodology.

And last but not least, the approach does not stop with cancer detection. Any cell type in the body is affected by degradation. It might be detected in blood serum. So, we are exploring the use of this methodology in other diseases. And actually, we have another publication upcoming where we show that we can detect immunological problems in individuals where we can actually deduce whether a person has an immune disorder or not.

Bob Barrett: That was Dr. Joris Vermeesch, Professor of Molecular Cytogenetics and Genome Research and Director of the Genomics Core Facility at the Catholic University of Leuven. He has been our guest in this podcast on "Pan-Cancer Detection and Typing by Mining Patterns in Large Genome-Wide Cell-Free DNA Sequencing Datasets." That paper appears in the September 2022 issue of *Clinical Chemistry*. I am Bob Barrett. Thanks for listening.