Clinical Chemistry

Fragment Ends of Circulating Microbial DNA as Signatures for Pathogen Detection in Sepsis

Article:
Guangya Wang, W K Jacky Lam, Lowell Ling, Mary-Jane L Ma, Saravanan Ramakrishnan, Don C T Chan, Wing-Shan Lee, Suk Hang Cheng, Rebecca W Y Chan, Stephanie C Y Yu, Irene O L Tse, Wai Tat Wong, Peiyong Jiang, Rossa W K Chiu, K C Allen Chan, and Y M Dennis Lo.
Fragment Ends of Circulating Microbial DNA as Signatures for Pathogen Detection in Sepsis
Clin Chem 2023; 69(2): 189-201. https://doi.org/10.1093/clinchem/hvac197

Guest: Dr. Jacky Lam is an Assistant Professor in the Department of Chemical Pathology and the Assistant Dean for Research of the Faculty of Medicine of the Chinese University of Hong Kong.

Bob Barrett: This is a podcast from *Clinical Chemistry*, a production of the American Association for Clinical Chemistry. I'm Bob Barrett. Next-generation sequencing, or NGS, of human plasma samples can detect fragments of DNA from microbes, potentially providing a new tool to identify the infection-causing pathogen in patients with sepsis and other diseases. One limitation of NGS-based test is the inability to distinguish between the microbes responsible for causing disease and those introduced as contaminants during sample collection or in any stage of the testing process.

In order to provide reliable information that can guide appropriate treatment, test methods need to be able to identify which DNA comes from the disease-causing pathogens and which comes from microbes that just happen to be present at the sample. Some strategies have already been proposed but they run the risk of eliminating true pathogenic organisms, resulting in false negative results.

A new research article, appearing in the February 2023 issue of *Clinical Chemistry,* identifies a possible solution. A clinical laboratorian with expertise in cell-free DNA and its utility as a biomarker explains the potential benefits of NGS-based infectious disease testing, identifies the limitations of current methods, and describes a new strategy that may help overcome these issues.

In this podcast, we are pleased to be joined by one of the lead authors of this article. Dr. Jacky Lam is an Assistant Professor in the Department of Chemical Pathology and the Assistant Dean for Research of the Faculty of Medicine of the Chinese University of Hong Kong. His research focuses on the biological features of cell-free DNA and its potential applications in clinical diagnostics. So, Dr. Lam, how can next-generation sequencing-based approaches be applied to infectious disease testing?

Jacky Lam: Yes. Next-generation sequencing-based, sometimes they may be referred as liquid biopsy for infectious disease. So, what we actually do is we sequence the microbe cell-free DNA in the blood that is derived from the infection-causing microorganisms. So, these are the presence, these microbial DNA could surface biomarkers for the presence of the pathogens. And so, the good thing is that it is non-targeted approach versus those by PCR, so it could allow us to non-invasively identify a wide range of infections that is present in the body. So, it does offer an advantage when indeed many common pathogens are difficult to culture using the conventional way of the microbiology method.

So, this is one advantage. The other advantage is that some deep-seated infection, they may require some invasive biopsy of the infected tissue in order to confirm the infection. So, these next-generation sequencing-based approaches could offer an advantage while it could be used non-invasively for the detection of the infection.

Bob Barrett: So, what is the limitation of next-generation-based sequencing approaches for the diagnoses of infectious disease?

Jacky Lam: So, I think one major challenge, while we're talking about technical challenge, is the contamination. Because while the microbial cell-free DNA is actually in the background of some nuclear DNA from human, and as well there may be a background of the contaminants. So, where do these contaminants come from? They could come from the reagents. So, they could include the water that we use, the DNA extraction kits for the sequencing, and also from the laboratory environment.

So, sometimes those contaminants that come from the extraction kits, some researchers refer to it as the "kitome." So, that is the contamination from the reagent kits. Indeed, because of the presence of these contaminants, there would be some difficulty while you're trying to differentiate the microbial cell-free DNA that is generally derived from the infection-causing microorganism versus those DNA comes from the contaminants.

So, indeed the researchers have proposed some of several approaches to tackle this problem. So, first what they would use is they could the microbial DNA abundance--the level of the presence of these microbial DNA as a filtering parameter. Why this could be used? Because they are based on the assumption that those DNA come from the genuine pathogen would have a higher level than that of a contaminating microbe but there could be one problem, because if you're trying to use the abundance as a filtering parameter, so that

would potentially eliminate those genuine pathogens that are present at the low concentrations.

So, that is one way to tackle the problem of contamination. Another way is some researchers would concurrently prepare something called no-template control, that is NTC. So, what is an NTC, it's a new word subject of blind control. So, basically just some water and you add the DNA extraction kits, you add the sequencing reagents, and then you prepare the sequencing library, and then you subject this no-template control, NTC, to sequencing.

So, whatever being detected in the NTC would be regarded as contaminants but at the same time, one problem with these trying to tackle to issue of contamination is that it would mean possibly eliminate those clinically relevant microbes that could be present in the contaminants' background as well. So, these two approaches have the intrinsic problems. So that's why--that is one of the major technical challenges that is present in the current NGS space approaches for infectious disease.

Bob Barrett: Dr. Lam, you and your team have analyzed the end motif profile of microbial DNA fragments in blood samples from patients with sepsis, could you tell us a bit more about this end motif marker?

Jacky Lam: Certainly. Because in the current world where we are working on cell-free DNA and these cell-free DNA, not just the microbial cell-free DNA derived from the pathogens, indeed they exist in the circulation in the blood as fragments, so they are bits of fragments of DNA that is present in the circulation. So, interestingly, the fragmentation process is not random. While if we try to study the fragmentation profiles of cell-free DNA, they do give us some information on the tissue of origin of the cell-free DNA and as wel,l the pathophysiological process that is going on in the body.

So, our group has been active in the research on the biology of the fragmentation of cell-free DNA, which we describe as fragmentomics. And along the fragmentomics research, we have developed a number of so-called fragmentomics markers with diagnostic implications for diagnostics and also informing the tissue of origins. So, to name a few examples of these fragmentomic markers, so they would include fragment size, fragment ends, and end motif is one of these fragmentomics markers.

So, what is actually end motif? They refer to the few nucleotides at the ends of the cell-free DNA fragments, regardless of the site of the origin within the genome. So, I would like to give some example to the audience, so how it could be used. So, for example, using the hepatocellular

carcinoma, liver cancer model. So we have demonstrated a different, a distinct plasma DNA end motif profile of the HCC patients from the healthy control subjects and also the patients with chronic hepatitis B infection. So, indeed if we're trying to look at the end motif profile of the nuclear DNA, it give us some diagnostic information.

Moving back to what we have done in infectious disease. In the current work that we have published in *Clinical Chemistry*, so we have analyzed the end motif profile of the microbial cell-free DNA derived from the pathogens in a group of patients with sepsis. If we look at the end motif profile after normalization with consideration of the different microbial genomic context, so we have shown that the end motif profile of the microbial cell-free DNA indeed they resemble that of the nuclear cell-free DNA.

So, this is very interesting because this show that the microbial cell-free DNA was also subject to non-random fragmentation so similar to nucleus cell-free DNA that we have observed, so microbial cell-free DNA was also within reach with a certain pattern, and this pattern is that those microbial cell-free DNA will be reached with a CCN, two C at the end of the microbial cell-free DNA fragments.

So, this CCN is interesting because they are considered as some cleavage signatures of certain nucleases that played a role in the fragmentation of DNA in the blood and this enzyme is called DNASE1L3.

Bob Barrett: Well, finally Dr. Lam, how might end motif analysis address some of the limitations of next-generation sequencing-based infectious disease assays?

Jacky Lam: Yes. Exactly, this is what we are trying to do in the current work that is published because without finding of the end motif profiles of the microbial cell-free DNA being likely that is non-random and there is likely being negated by certain nucleases.

So, we postulate that because those contaminants, those contaminating microbial DNA introduced during the sequencing, during the sample preparation, they may not be subject to the same fragmentation process that is occurring in the blood, and because of the difference in the fragmentation process, so they may exhibit a different fragmentation, and therefore, end motif profile.

So, how do we prove it is to confirm this is that we have used the sequencing data from the no-template control that I have mentioned before. To reiterate, those no-template control, there is no patient sample but only the reagents and then we subject them to sequencing. In these no-template control, or

blind control, those microbial DNA detected must we would assume that there are contamination, there would be contaminating microbial DNA.

So, interestingly, when we're trying to compare the end motif, the fragmentation profile from the microbial cell-free DNA, from sepsis patient, versus those contaminating microbial DNA from NTC, from no-template control, and then we can see that they are different because those contaminating microbial DNA, they do not have a preferential end motif profile, unlike those from sepsis patient and reached with a certain CCN profile.

Just before, earlier on, I've mentioned that because NTC, in the no-template control. then we may detect some genuine pathogenic species. So to highlight in the current work, we have shown that because in the NTC, we could detect DNA sequences from the *Bacteroides* species, and also at the same time among the clinical samples of the sepsis patient, we could also detect the DNA sequences from the *Bacteroides* species.

As mentioned before, because if we try to use NTC as a filtering criteria, so we may erroneously eliminate the *Bacteroides* species because anything that is present in the NTC, they may be considered as a contamination. But if we're trying to interpret the end motif profile, indeed we can show that they are different end motif profile of the DNA from the *Bacteroides* species in the NTC versus from the genuine pathogens from the sepsis patients. So that would illustrate one advantage of using the end motif trying to differentiate the genuine pathogens versus the contaminations. And finally what we have done is we would like to further validate our finding and evaluate the performance of end motif analysis and by sequencing an additional 38 subjects without infection.

So, these subjects were also being in-hospital patients but they do not have any signs or symptoms of infection. So, we label this group as a non-infection group. So, we were trying to detect the microbial DNA sequences from these subjects and all these subjects would be considered as having the contaminating microbial cell-free DNA. And in that analysis, we are trying to compare the use of two parameters. So, the first is that we use the microbial DNA abundance as we mentioned earlier on, and second, we'll use what we call a microbial cell-free DNA end signature score.

So, this is one metric that we developed for the end motif analysis and also subsequent comparison. So, interestingly, when we're trying to evaluate the performance of these two parameters in the differentiation of the pathogenic microbes versus the contaminating microbes in the non-infection

group, so using that ROC analysis, we could show that the end signature score that is based on the end motif analysis, they could achieve a higher performance, a higher area under the curve score in the differentiation.

So, that basically we are trying to illustrate the use, the utility, of end motif analysis in the separation and the differentiation of the genuine pathogens versus the contaminating microbes.

Bob Barrett: That was Dr. Jacky Lam from The Chinese University of Hong Kong. He and his team proposed a new method to distinguish true infection-causing pathogens from other non-pathogenic microorganisms introduced as contaminants during specimen collection or analysis. His work was published in the February 2023 issue of *Clinical Chemistry*. I'm Bob Barrett. Thanks for listening.