Sequencing Method Screens for All Known Bacterial Pathogens, Resistance Markers

Oct 24, 2018 | Madeleine Johnson

NEW YORK (GenomeWeb) – Using capture-based targeted sequencing, researchers at Columbia University have developed a method that can identify all known pathogenic bacteria from clinical samples as well as known resistance and virulence factors.

The technique reportedly lowers the limit of detection compared to unbiased high-throughput sequencing and results in a 1,000-fold increase in bacterial reads from blood samples.

The new method, called BacCapSeq, was developed in the lab of Ian Lipkin, and described in a paper published yesterday in *mBio*.

The Lipkin lab recently pioneered a similar capture-based method for viruses, called VirCapSeq, as previously reported. The new bacterial method "is the same in principle," said Lipkin.

"It gives you the same sort of enrichment that you get with VirCapSeq, but it also allows us to detect biomarkers for microbial resistance that appear very early after culture," he said.

BacCapSeq uses a probe set of approximately 4.2 million oligonucleotides. It detects 307 bacteria, including all those known to be pathogenic to humans as well as others that were selected based on a potential to become pathogenic in the future.

The probes for bacterial identification were chosen using the Pathosystems Resource Integration Center (PATRIC) database, and the Columbia team selected only one genome per bacterial species.

Importantly, the BacCapSeq method can also detect antimicrobial resistance following a short period in culture. The probe oligo set therefore also includes a subset targeting 2,169 known antimicrobial resistance genes and 30,178 known virulence factors. This subset was selected based on data in the Comprehensive Antibiotic Resistance Database (CARD) and the Virulence Factor Database (VFDB).

The group filtered the target sequence data set down to approximately 1 million genes and sent these on to Roche Sequencing Solutions where they were additionally filtered based on printing constraints. The final set of 4,220,566 oligos has an average length of 75 nucleotides, a mean melt temperature of 79° C, and average interprobe distance of 121 nucleotides.

For the resistance assessment, the bacteria is placed immediately in antibiotic directly out of the sample, and then after as little as three hours of culture the test can detect markers expressed that indicate the presence of resistance.

In the *mBio* study, BacCapSeq performance was assessed using human whole blood nucleic acid spiked with bacterial nucleic acids, whole blood spiked with bacterial cells, and blood culture samples obtained from a clinical microbiology lab.

The method resulted in up to a 1,000-fold increase in reads from blood samples, and a lower limit of detection compared to conventional high-throughput sequencing, with as few as 5 million reads generated per sample.

In comparison to commercially available highly-multiplexed PCR-based infectious disease detection panels, which can detect about 20 different pathogens, VirCapSeq can detect about a 15-fold greater number of pathogens in a single test.

On the other hand, metagenomic sequencing for unbiased infectious disease detection can potentially detect any organism present. However, Lipkin noted that getting data as rapidly, conveniently, and inexpensively as possible is important for diagnostics work, and metagenomics does not yet fit that bill.

"These capture methods allow you to get more sensitivity than you get with unbiased high-throughput sequencing, or metagenomic sequencing, because all the sequencing is invested in the targets that are of interest," he explained.

Beyond sensitivity, focusing using target capture also potentially reduces cost by bringing down the total number of reads needed. Lipkin highlighted recent work his lab has done in collaboration with the University of Geneva as well as with a group in Uganda using VirCapSeq to reduce reads needed for ID from about 300 million to 3 million.

Finally, the bioinformatics of targeted sequencing is more straightforward because results don't need to be compared to as large of a set of possible matches, and this will potentially save on time as well as on labor costs. Higher sensitivity can also increase throughput, and having more samples in a given run could also reduce costs.

"We're already planning to be able to run 96 samples at a time, and that is just a start," Lipkin said. Furthermore, he noted that faster ways to get the sequencing done are on the horizon, "Whether it is on the [Oxford] Nanopore system, or something similar."

Other groups use sequencing for infectious diseases detection. For example, Fry Laboratories is using 16S rRNA sequencing and 18S rRNA sequencing for detection. Still others are working out methods for single-cell sequencing, for example using droplet microfluidics, or methods for using whole-genome sequencing to determine resistance once a pathogen has been identified.

Lipkin thinks BacCapSeq can also be used as a discovery tool to develop better diagnostics. "Right now, we're doing this using sequencing, but we can use it for the discovery of biomarkers that will allow us to do things like consensus PCR, microarrays, and other approaches that will allow us to figure out what the appropriate antibiotic is to use in a given individual."

For viruses, with the exception of influenza and herpes, there aren't many drugs that can be used as treatment for infection. But for bacteria, BacCapSeq is "an important advance," Lipkin said, because correct ID and resistance information can tailor antibiotic choices. "It's like precision medicine, but for microbiology," he said.

The Columbia group has two commercialization plans, Lipkin said. First, it will distribute the probe set.

"We are in the process of validating VirCapSeq with the New York State Department of Health Wadsworth Center for use in CLIA labs," he said. The team plans to do the same with the BacCapSeq probe set.

The ultimate goal is to identify pathogens as well as resistance and virulence markers in three hours or less, he said.

"We're anticipating that in the next few years that is going to be possible," Lipkin said. At present, the turnaround time might be more in the 24- to 36-hour range, he said, which is at least much faster than the three- to five-day turnaround on culture-based drug sensitivity assays.

Columbia University will also be patenting the technique to get it out into the commercial marketplace. The method can be used for research purposes immediately, but in terms of use on patient samples, it will need to meet the bar for CLIA use. The group also intends to take the assay through the US Food and Drug Administration in partnership with an as-yet-to-be-determined commercial entity.

Filed Under Sequencing Clinical Sequencing Molecular Diagnostics Infectious Disease sepsis target capture targeted sequencing