genomeweb

BillionToOne Developing Coronavirus Dx Test Based on Sanger Sequencing

Apr 08, 2020 Julia Karow

NEW YORK – Prenatal and cancer molecular diagnostics company BillionToOne has come up with an unconventional way that could dramatically increase the number of available coronavirus diagnostic tests: a quantitative Sanger sequencing assay that harnesses the large existing base of capillary electrophoresis instruments in clinical laboratories.

In a manuscript posted on the company's website this week that will soon appear as a preprint on *BioRxiv*, BillionToOne researchers outlined their qSanger-COVID-19 test, which relies on sequencing a single amplicon from the viral genome along with an RNA spike-in control, followed by data analysis with the company's proprietary software.

In about two weeks, the Menlo Park, California-based firm plans to submit the test to the US Food and Drug Administration for Emergency Use Authorization, which would allow it to sell kits to almost any clinical laboratory that has a Sanger sequencing instrument available. This could potentially increase the coronavirus diagnostic test capacity by hundreds of thousands per day while circumventing some of the reagent supply issue experienced by providers of RT-PCR tests.

"For us, it's about getting this out there to as many people as possible," said Oguzhan Atay, BillionToOne's cofounder and CEO. "I hope that happens so that we can all go back to doing our core business as quickly as possible."

His company originally developed the patent-pending quantitative Sanger sequencing method for a noninvasive prenatal aneuploidy test, which the firm has not launched yet. Essentially, it relies on simultaneously PCR-amplifying and Sanger sequencing a target in a sample and a spike-in control. The sequence of the control is very similar to that of the target but has a frameshift, so the traces of the two sequences in the chromatogram initially overlap but are then offset by a few bases. From the difference in amplitude between the base peaks from the target and the spike-in control, an algorithm can calculate the amount of target in the sample.

"You add a spike-in that you know is at 100 genome copy equivalents to your sample, and if you see both peaks essentially at the same height, you know that your viral load is 100 genome copy equivalents," Atay said. While the sequencing protocol is easy to perform, the bioinformatics is a little tricky, especially near the limit of detection. "We have optimized our own machine learning algorithm to be able to detect those very small signals at the same limit of detection that qPCR does," he said.

A single amplicon from one of the viral genes and a single RNA spike-in control is sufficient for the assay. In their validation experiments, which used synthetic SARS-CoV-2 genomic RNA and AccuPlex SARS-CoV-2 reference material as a proxy for clinical samples, the researchers found the limit of detection to be about 10 to 20 GCEs.

The spike-in serves not only as a way to quantify the amount of virus in the sample but also as an internal control. "Normally, if there is no signal, you wouldn't know if it is due to a problem with amplification or extraction" or because no virus is present, Atay said, but with the spike-in, there will always be a signal if the PCR and Sanger reactions have worked. If there is viral RNA in the sample, the chromatogram will show mixed bases, but if there is no virus, the spike-in control will yield just one clean sequence.

The protocol also requires no RNA extraction. In their paper, the researchers showed that they could detect SARS-CoV-2 from viral particles suspended in transport media and directly added to the PCR master mix, with no prior RNA purification. This, Atay said, gets around the shortage of RNA extraction reagents that has limited qPCR-based coronavirus testing.

Also, he said, regular PCR instruments and capillary electrophoresis sequencers – the two instruments needed for the assay – are plentiful in clinical labs. Sanger sequencers are still widely used for fragment analysis and single-gene tests, and Atay estimated that US clinical labs probably still operate about 100 Applied Biosystems 3730xl Genetic Analyzers – Thermo Fisher Scientific's high-throughput CE instruments with either 48 or 96 capillaries – in total. Each 3730xl could in theory handle 3,840 patient samples per day in a largely automated format that requires little hands-on time. "You don't need to babysit it the way you have to babysit a qPCR machine," he said.

To get its COVID-19 assay into the hands of clinical labs, BillionToOne, in partnership with several contract manufacturers, has been developing a company-branded kit for the PCR amplification that it plans to submit for EUA. In the meantime, the company plans to ship research-use-only versions of its kit to research labs, starting as early as the end of this week. "The goal for us is to make this as easy as possible for other people to use," Atay said, so "clinical labs can simply add the sample into each of these wells, PCR amplify, and then Sanger sequence it."

The contract manufacturers should soon be able to produce kits for several 100,000 tests per day, he said, and capacity could be further increased by either working with a larger contract manufacturer or by moving away from the 96-well format and providing labs with reagents only that they need to pipette into plates themselves.

Atay said BillionToOne has almost all the data required for the EUA application in hand but still needs to replicate the results with the manufactured kits. It also wants to generate data on several different Sanger sequencing platforms to prove that they are equivalent to the 3730xl the company has been using internally, so they can be included in the EUA. It is doing so by sending out samples to sequencing service providers that operate other Sanger platforms.

Based on current reagent and manufacturing costs, Atay said BillionToOne will probably make the PCR amplification kits available at a price of around $15 per reaction, which will include the use of its analysis software. Customers can also opt to purchase their own reagents and use the firm's software for a fee. Sanger sequencing costs range between $2 and $6 per reaction, based on pricing from service providers, so the total reagent cost for a clinical lab to perform the qSanger COVID-19 test should be around $20 or less per sample, he said.

"I am delighted to see a new and exciting approach using Sanger sequencing," said Scott Tighe, technical director of the Vermont integrative genomics core lab at the University of Vermont Cancer Center. "It is without a doubt that combining Sanger sequencing with a no-extraction sample processing protocol would allow for the analysis of many more samples because of the throughput of the instrument. However, this is only part of the equation since sample setup, instrument run time, and final data interpretation also play a role slowing things down."

Tighe, who helped develop an RNA extraction-free RT-qPCR coronavirus assay that recently came out as a *BioRxiv* preprint, said that "if we as a scientific community can adopt a no-extraction protocol in general for RT-qPCR or a method such as qSanger, throughput would increase exponentially considering most centers have several instruments."

Capacity for the BillionToOne assay could be further increased, Atay suggested, if clinical labs could get their hands on the hundreds of Sanger sequencing instruments that are installed in research labs and core facilities. Tighe, for example, said that few clinical labs have high-throughput Sanger sequencers anymore and many of them send their sequencing out to a handful of vendors.

This could be done either by clinical labs renting or borrowing Sanger sequencers from research labs for the duration of the pandemic or by sending out PCR-amplified patient samples to research labs for sequencing. However, the latter would require relaxation of the current rules for diagnostic COVID-19 testing.

"We're saying that if that were allowed, that would create essentially millions of tests on top of maybe a few hundred thousand to a million that would come from clinical labs," Atay said, adding that his company has so far not seen any differences between Sanger results generated in its clinical lab and through research service providers. "If we had a distributed model where the safety-critical parts were happening within a clinical lab, and then [the sample] is sent out for sequencing, it would open up this capacity in a much faster way than it could otherwise be accomplished," he added. "But, this is a regulator question."

For BillionToOne, developing a coronavirus test was never part of the business plan. Last month, the firm raised $15 million in a Series A follow-on round to support the commercialization of its noninvasive prenatal Unity test and to fund R&D for cancer liquid biopsy assays and other work. "This is not our core business. It's not even something that would benefit us in the short term, mid term, or long term," Atay said. "We take away resources from all the other things that we're doing. But, we thought this was the right thing to do because we realized we have this [quantitative] Sanger sequencing technology that could be easily applied to this."