For the first time in a patient, researchers use long-read genome sequencing

June 22 2017

A depiction of the double helical structure of DNA. Its four coding units (A, T, C, G) are color-coded in pink, orange, purple and yellow. Credit: NHGRI

When Ricky Ramon was 7, he went for a routine checkup. The pediatrician, who lingered over his heartbeat, sent him for a chest X-ray, which revealed a benign tumor in the top-left chamber of his heart. For Ramon, it was the beginning of a long series of medical appointments, procedures and surgeries that would span nearly two decades.

During this time, noncancerous tumors kept reappearing in Ramon's heart and throughout his body—in his pituitary gland, adrenal glands above his kidneys, nodules in his thyroid.

The trouble was, doctors couldn't diagnose his condition.

When Ramon was 18, doctors thought his symptoms were suggestive of Carney complex, a genetic condition caused by mutations in a gene called PRKAR1A. However, evaluation of Ramon's DNA revealed no disease-causing variations in this gene.

Now, eight years later, researchers at the Stanford University School of Medicine have used a next-generation technology—long-read sequencing—to secure a diagnosis for Ramon. It's the first time long-read, whole-genome sequencing has been used in a clinical setting, the researchers report in a paper to be published online June 22 in *Genetics in Medicine*.

Genome sequencing involves snipping DNA into pieces, reading the fragments, and then using a computer to patch the sequence together. DNA carries our genetic blueprint in a double-stranded string of molecular "letters" called nucleotides, or base pairs. The four types of nucleotides are each represented by a letter—C for cytosine and G for guanine, for example—and they form links across the two strands to hold DNA together.

'Illuminating a dark corner'

Current sequencing technologies cut DNA into "words" that are about 100 base-pairs, or letters, long, according to the study's senior author, Euan Ashley, DPhil, FRCP, professor of cardiovascular medicine, of genetics and of biomedical data science. Long-read sequencing, by comparison, cuts DNA into words that are thousands of letters long.

"This allows us to illuminate dark corners of the genome like never before," Ashley said.

"Technology is such a powerful force in medicine. It's mind-blowing that we are able to routinely sequence patients' genomes when just a few years ago this was unthinkable."

The study was conducted in collaboration with Pacific Biosciences, a biotechnology company in Menlo Park, California, that has pioneered a type of long-read sequencing. Lead authorship of the paper is shared by Jason Merker, MD, PhD, assistant professor of pathology and co-director of the Stanford Clinical Genomics Service, and Aaron Wenger, PhD, of Pacific Biosciences.

The type of long-read sequencing developed by the research team's collaborators at Pacific Biosciences can continuously spool long threads of DNA for letter-by-letter analysis, limiting the number of cuts needed.

"This is exciting," said Ashley, "because instead of having 100-base-pair 'words,' you now have 7,000- to 8,000-letter words."

Falling cost

Thanks to technological advances and increased efficiency, the cost of long-read sequencing has been falling dramatically. Ashley estimated the current cost of the sequencing used for this study at between $5,000 and

$6,000 per genome.

Though the cost of short-read sequencing is now below $1,000, according to Ashley, parts of the genome not accessible when cutting DNA into small fragments. Throughout the genome, series of repeated letters, such as GGCGGCGGC, can stretch for hundreds of base pairs. With only 100-letter words, it is impossible to know how long these stretches are, and the length can critically determine someone's predisposition to disease.

Additionally, some portions of the human genome are redundant, meaning there are multiple places a 100-base pair segment could potentially fit in, said Ashley. This makes it impossible to know where to place those segments when reassembling the genome. With longer words, that happens much less often.

Given these issues, 5 percent of the genome cannot be uniquely mapped, the researchers wrote. And any deletions or insertions longer than about 50 letters are too long to detect.

For patients with undiagnosed conditions, short-read sequencing can help doctors provide a diagnosis in about one-third of cases, said Ashley. But Ramon's case was not one of those.

The technique initially used to analyze Ramon's genes failed to identify a mutation in the gene responsible for Carney complex, though Ashley said co-author Tam Sneddon, DPhil, a clinical data scientist at Stanford Health Care who browsed through the database of Ramon's sequenced genome by hand, did notice something looked wrong. Ultimately, the long-read sequencing of Ramon's genome identified a deletion of about 2,200 base-pairs and confirmed that a diagnosis of Carney complex was indeed correct.

This work is an example of Stanford Medicine's focus on precision health, the goal of which is to anticipate and prevent disease in the healthy and precisely diagnose and treat disease in the ill.

An 'exceedingly rare' condition

Carney complex arises from mutations in the PRKAR1A gene, and is characterized by increased risk for several tumor types, particularly in the heart and hormone-producing glands, such as ovaries, testes, adrenal glands, pituitary gland and thyroid. According to the National Institutes of Health, fewer than 750 individuals with this condition have been identified.

The most common symptom is benign heart tumors, or myxomas. Open heart surgery is required to remove cardiac myxomas; by the time Ramon was 18 years old, he'd had three such surgeries. He is under consideration for a heart transplant, and having the correct diagnosis for his condition was important for the transplant team. Beyond the typical screening for a transplant, Ashley said the team needed to ensure there weren't other health issues that could be exacerbated by immune suppressants, which heart transplant patients must take to avoid rejection of the donated organ.

Though it helps his medical team to have a confirmed diagnosis of Carney complex, Ramon has found it disheartening to face the fact that he cannot escape his condition. "I was pretty sad," he said. "It took me a while to come to terms with the fact that I'll have this until the day I die."

He tries not to dwell on it, though. "Live one day at a time," he said. "The bad days are temporary storms, and they'll pass."

"His story is quite incredible," said Ashley, who said it was a privilege to be working on Ramon's team. "To have such a burden on such young

shoulders, and to decide whether or not he wants a transplant, requires incredible courage."

Because he couldn't wait any longer for a transplant, Ramon recently underwent his fourth surgery to remove three tumors in his heart. Joseph Woo, MD, professor and chair of cardiothoracic surgery at Stanford, performed the operation. "It is exceedingly rare to have tumors in the heart," said Ashley. "It was a particularly heroic operation." Though Ramon is still under consideration for a transplant, the need is less urgent now.

"I'm in good hands," Ramon said of the Stanford team. "I'm glad to be here."

A future in the clinic?

Ashley said he and many other doctors believe that long-read technology is part of the future of genomics.

"Now we get to see how to do it better," said Ashley. "If we can get the cost of long-read sequencing down to where it's accessible for everyone, I think it will be very useful."

Provided by Stanford University Medical Center