Unraveling Disease Mechanisms With The NextSeq® 500 System

Researchers at the Icahn School of Medicine at Mount Sinai are studying mtDNA, piRNA, and TCR repertoire diversity to unlock the secrets of cancer, infertility, and autoimmune disease.

Introduction

Extensive research is focused on how human genome aberrations contribute to the onset and progression of disease. Structural variations, gene mutations, deletions, and insertions have all been found to play a role. Yet up until recently, researchers have been limited in studying how genetic changes in the DNA or RNA driving specific biological subsystems might be involved in disease onset and progression.

An innovative research team at the Icahn School of Medicine at Mount Sinai led by Ravi Sachidanandam, PhD, is focused on developing new methods to study 3 of these important subsystems: mitochondrial DNA (mtDNA), T cell receptor (TCR) diversity, and the piwi-interacting RNA (piRNA) pathway. Each has been difficult to study in the past. mtDNA studies are challenging, because it is hard to parse out mtDNA from the whole genome for sequencing purposes to identify its potential role in disease processes. Sequencing the full TCR repertoire is time-consuming, limiting detailed study of the adaptive immune system, its role in autoimmune disease, and immune and vaccine response. Researchers have also faced challenges in deciphering role of the piRNA pathway in defending the germline against attacks from transposons that could impact fertility.

Working with Illumina sequencing systems for the past 5 years, Dr. Sachidanandam and his team invented a new method for removing biases in small RNA sequencing while studying piRNAs[1]. They also developed a method to sequence mtDNA with high resolution, without having to simultaneously sequence the nuclear genome[2], and are working on new, sequencing-based approaches for monitoring TCR response in blood. In 2015, his team installed a NextSeq 500 System in the Oncological Sciences department to expedite their research studies.

iCommunity spoke with Dr. Sachidanandam to learn how his team is using the NextSeq 500 System to develop novel sequencing methods.

Q: What sparked your interest in genomics?
Ravi Sachidanandam (RS): I was trained as a physicist and as I was thinking of what to do after my postdoc, I was reading a lot about genomics in newspapers and journals. I applied to the Cold Spring Harbor Laboratory and began working on a large-scale SNP project, which was the very first of its kind, to compile SNPs in the human genome. The database included > 300,000 SNPs. I eventually ended up working on several projects, which included a group that were amongst the first to generate data on the early Illumina sequencing systems, the Genome Analyzer® II.

The chair of genetics at Mount Sinai Hospital recruited me 5 years ago to build upon the strictly computational work that I had been doing at Cold Spring Harbor. When I made the move, they were thinking of starting a sequencing program and the responsibility for setting it up fell into my hands. I hired a technician and together we learned a lot about sequencing systems and workflows as we established the laboratory.

Q: How are your team's research goals intertwined with the development of new sequencing methodologies?
RS: It becomes a means to an end. We are investigating dynamic biological subsystems that researchers have struggled to analyze. We find that there is a continual loop where the computational results of our RNA and DNA sequencing efforts inform our wet-bench practices and vice versa. This flow of information has enabled us to develop innovative and powerful approaches, including new laboratory techniques, and computational algorithms and tools.

Our piRNA research is a good example, because it led to our first sequencing invention. piRNAs are the largest class of small, noncoding RNA molecules and play a role in gene silencing, specifically the silencing of transposons. They are present only in the germline. Alterations in the piRNA pathway can disrupt protection of the germline from transposon attack, potentially resulting in infertility.

Ravi Sachidanandam, PhD, is an Assistant Professor of Oncological Sciences at the Icahn School of Medicine at Mount Sinai.

For Research Use Only. Not for use in diagnostic procedures.

We accidentally discovered that the protocols used for small RNA sequencing are biased towards certain piRNA. We were able to invent a new method of sequencing small RNAs. The method works well for piRNAs and other small RNAs, including miRNAs.

Q: What other scientific questions are you trying to answer?
RS: There are 2 other biological subsystems that we are studying and trying to understand in greater depth. One of them is the genetics of mtDNA. Mitochondria are the engines of our cells, converting food into energy to fuel our muscles, brain, and other tissues. Sequencing mtDNA is difficult. In the past, we sequenced the whole genome and, in the process, picked up some mtDNA pieces. So while the role of mitochondria in disease might be widespread, we have not been able to decipher it because of our inability to sequence mtDNA and the genome separately.

The other subsystem we are studying is TCRs, which are part of our adaptive immune system. Each person's TCR repertoire possesses unique elements that are triggered by its response to viral and bacterial infections over their lifetime. It takes over when our immune system mistakes our own cells for a foreign entity and starts attacking them. TCRs also play a role in vaccine responses and how quickly the body clears an infection. We are developing new sequencing approaches to track these responses.

> "We find that there is a continual loop where the computational results of our RNA and DNA sequencing efforts inform our wet-bench practices and vice versa. This flow of information has enabled us to develop innovative and powerful approaches, including new laboratory techniques, and computational algorithms and tools."

Q: Can you talk about the Mseek technique that your team developed?
RS: The Mseek technique solved the mtDNA sequencing problem. In the past, when we sequenced cellular DNA, only a small amount of it was mtDNA and our resolution was poor. That is not surprising. The nuclear genome consists of 6 billion bases, while the mitochondrial genome is only 17,000 bases. Even though there are 1,000 mtDNA for every nuclear genome, total mtDNA adds up to only 17–20 million bases. So less than 1% of total DNA is mtDNA.

Anitha Jayaprakash, a student in the lab, invented Mseek, a method of isolating and sequencing just the mitochondrial DNA. The technique takes advantage of the fact that the mitochondrial genome is circular and the nuclear genome is linear. Mseek eliminates the linear nuclear DNA through an enzymatic process, leaving behind the circular mtDNA. From that point on, it is a simple process to sequence the mtDNA at very high resolution, about 1000× coverage.

Using this unique method, we screened large numbers of human cell lines and discovered the mtDNA diversity is widespread and stable. Since then, we have developed an entire program focused on understanding how the mitochondrial genome affects human development and metabolic disorders.

Now there is kind of a "gold rush" in this field because we can actually monitor and track mitochondria in different tissues using this technique and understand what is going on in various disorders.

> "The biggest advantages of the NextSeq 500 System are its throughput, optimal price point, and fast turnaround time. The installation was easy and we were able to start running projects on it immediately."

Q: Can you talk about the TCR technique that you are developing?
RS: We have developed a method of sequencing to monitor the TCR repertoire in people over time, potentially using it as a biomarker for disease. The key point is the use of universal primers for unbiased sampling of the repertoire, avoiding the biases that plague other approaches that are mostly based on PCR amplification. Using the technique we could learn how people react to different diseases or whether a vaccination is working. There are also potential implications for cancer immunogenomics.

Q: On which diseases is your research focusing?
RS: I am in the Oncological Sciences department, so clearly cancer is a focus. Within mtDNA, we are looking at what goes on in tumors and if there is an evolution of the mitochondrial genome as the tumor progresses. I also conduct research in autoimmune disease and collaborate with Dr. Terry Davies, a famous researcher of Graves' disease, an autoimmune thyroid disorder. We are trying to understand the role of the TCR repertoire in this disorder.

Q: Does your research support personalized medicine approaches to fighting disease?
RS: That is the goal, particularly of our TCR research. The idea is to be able to monitor disease through noninvasive techniques. Ultimately, we would like to be able to take a small blood sample and use our approach to figure out what the immune system is doing. It would enable us to tailor therapies to each person.

Q: How important are the tools that you use to the success of your research?
RS: The technologies and tools we use are extremely important. Sequencing is our main technique because we use it to infer things from tissue and fluid samples. Access to technology and a quick

For Research Use Only. Not for use in diagnostic procedures.

data turnaround are important for us because we develop techniques to address specific areas. For example, if I want to look at the TCR repertoire, I have to develop a method to monitor it in the blood. We spend a lot of time going back and forth—taking a sample, using a method to prepare and sequence it, looking at the sequencing data, and tweaking the method until we are sure that everything is working. Then we apply it to large sets of samples.

Q: How long have you been using the Illumina systems?
RS: I have been using Illumina systems since the Genome Analyzer II System was introduced in 2006. In this lab, we have used almost every Illumina sequencing system, including the MiSeq®, HiSeq®, and NextSeq Systems.

Q: How do the Illumina systems compare with other sequencing systems?
RS: The systems are good in terms of ease of use. The responsiveness of the Illumina technical support team has been important for us. We might be one of the heaviest users of this service because we are always trying pushing the boundaries of the sequencer, trying new things, wanting to customize this or that. Illumina Tech Support always tries to address our questions and help us find ways to perform what we are trying to do. That is a significant plus for Illumina, and as a customer I appreciate the added dedication.

> "Having our own NextSeq 500 System has let us generate data and get feedback in the lab quickly. It has eliminated our sequencing bottleneck."

Q: Why did you choose the NextSeq 500 System for your current study?
RS: The biggest advantages of the NextSeq System are its throughput, optimal price point, and fast turnaround time. The installation was easy and we were able to start running projects on it immediately.

The NextSeq System enables us to have real-time monitoring of our experiments so we can go back and forth between the lab, analysis, and sample preparation. We can quickly solve problems in the lab before we get too far. Before, we had to wait a few weeks for data and would lose track of what was going on in the lab. After 3 months working with the NextSeq 500 System, we have significantly shortened our debug cycle for the techniques we are developing.

Q: What do you think of its performance and data quality?
RS: Our yields exceeded the specified output significantly and the data quality has been comparable to what we used to get from the HiSeq System.

Before installing the NextSeq System, we were using a HiSeq System at a core lab and at various companies, or a MiSeq System. Having our own NextSeq 500 System has let us generate data and get feedback in the lab quickly. It has eliminated our sequencing bottleneck. Previously we would wait up to 8 weeks for data; now we get results back in a day. Data turnaround is not really a consideration in planning experiments anymore. Our work volume has increased, enabling us to pursue additional research directions.

Q: How many different sequencing applications are you performing with it?
RS: We perform mtDNA sequencing and the TCR repertoire analysis using our proprietary techniques. We are also performing mRNA-Seq, ChIP-Seq, whole-exome sequencing, amplicon sequencing, and some bacterial genome sequencing.

> "Previously we would wait up to 8 weeks for data; now we get results back in a day. Data turnaround is not really a consideration in planning experiments anymore."

Q: Do you use any Illumina library prep kits?
RS: For mRNA-Seq, we use the TruSeq® Stranded mRNA Library Prep Kit.

Q: What types of bioinformatics tools have you developed?
RS: We have developed all the bioinformatics tools we use in-house. We have several tools, resources, databases, and websites that allow others to analyze the data on our website. For example, Kismeth is a web-based tool for bisulfite sequencing analysis that enables researchers to analyze methylation states in plants and humans. Our mRNA-splicing resource, SpliceRack, is widely used in the field as are our tools for piRNA analyses.

Q: What are the next steps in your research?
RS: We will apply our tools to various problems. For example, we will be using the Mseek protocol to study mitochondrial genetics in a twins study, comparing mitochondrial genomes and seeing how they differ. Since twins come from the same mother, they should technically have identical mitochondrial genomes. We are testing that theory. We also want to compare the TCR repertoire in twins and see how they diverge over time.

References:

1. Jayaprakash AD, Jabado O, Brown BD, and Sachidanandam R. Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. Nucleic Acids Res. 2011; 39(21):e141. doi: 10.1093/nar/gkr693.
2. Jayaprakash AD, Benson EK, Gone S, et al. Stable heteroplasmy at the single-cell level is facilitated by intercellular exchange of mtDNA. Nucleic Acids Res 2015; 43(4): 2177–2187.

For Research Use Only. Not for use in diagnostic procedures.

Learn more about the Illumina systems and products mentioned in this article:

- NextSeq 500 System, www.illumina.com/systems/nextseq-sequencer.html
- MiSeq System, www.illumina.com/systems/miseq.html
- HiSeq System, www.illumina.com/systems/hiseq_2500_1500.html
- TruSeq Stranded mRNA Library Prep Kit, www.illumina.com/products/truseq_stranded_mrna_library_prep_kit.html

Illumina • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

For Research Use Only. Not for use in diagnostic procedures.

© 2015 Illumina, Inc. All rights reserved. Illumina, Genome Analyzer, HiSeq, MiSeq, NextSeq, TruSeq, and the pumpkin orange color are trademarks of Illumina, Inc. and/or its affiliate(s) in the U.S. and/or other countries. Pub. No. 770-2015-017 Current as of 23 June 2015

illumina®