Next-Generation Sequencing Provides a High-Quality, Efficient, Affordable Replacement for Sanger Sequencing

For Bryan Paeper, who has sequenced thousands of exomes, the MiSeq® small-scale next-generation sequencer speeds up and simplifies operations.

The Northwest Genomics Center at the University of Washington was created in 2009 using funds from a $25 million "Grand Opportunity" (GO) grant from the American Reinvestment and Recovery Act (ARRA) administered by the National Heart, Lung, and Blood Institute (NHLBI). Dr. Deborah Nickerson, who pioneered exome sequencing, created and still directs the facility. At the time of the grant, the new genome center was tasked with sequencing 4,000 exomes in two years. The ultimate goal was to identify genetic connections to disease. Having successfully accomplished this initial charter, the lab now plans to sequence small custom targets for several groups around the world.

As a research consultant at the Northwest Genomics Center, Bryan Paeper leads the team responsible for generating thousands of sequences of data each year. With the increasing demand for sequencing projects, he is looking at the MiSeq system to speed up and simplify operations so that he can meet growing needs.

Q: How will you integrate the MiSeq system into your sequencing process?

Bryan Paeper (BP): Our main use for the MiSeq will be to QC libraries before they are sequenced on the HiSeq® system. With the MiSeq, we can QC 96 or 384 multiplexed libraries and know which ones are good to go on a HiSeq in just one day. That is huge for us. Before, we had to put them onto a HiSeq for five days to figure out which ones should continue and which ones shouldn't. The cost savings is key because the five days of putting the libraries on a HiSeq costs more than putting them on a MiSeq.

Q: What other projects are amenable for the MiSeq system?

BP: The MiSeq system is basically going to replace Sanger sequencing for us. We'll get more efficiency for the smaller target projects. Sanger sequencing is costly and takes a lot of time. Because the MiSeq is so quick and relatively easy to use, it will make a big difference.

"MiSeq not only saves on cost for reagents, but also in time and personnel. The difference is huge."

Bryan Paeper leads a team of five at the Northwest Genomics Center where they generate sequence data for the Nickerson lab at the University of Washington and for collaborators worldwide.

Q: How much Sanger sequencing was done in the lab?

BP: Sanger was the only sequencing method used in this lab for the last 10 years. Up until two years ago we had an ongoing grant with NHLBI to do exome sequencing. People would send their target region and all their samples to us. We would design all of the probes to do the PCR and the sequencing. It ran nonstop.

Q: How will the MiSeq system improve the projects previously done on Sanger sequencers?

BP: It's going to be greatly simplified. I personally like Illumina's system for the small targets. I like doing direct PCR instead of capture enrichment. With direct PCR on small samples you're basically getting 100% enrichment and you only sequence what you want to sequence. In contrast, with a capture you can get anywhere from 30–70% enrichment and end up with a reasonably large percentage of data that's not in the target that you're after. For the really small targets, Illumina's design is a lot more efficient.

Q: How else do you see the MiSeq system improving upon Sanger sequencing?

BP: With Sanger sequencing, we used to PCR amplify and sequence each exon across the entire exome in individual wells. This could require preparation, PCR, and sequencing of hundreds and hundreds of 96- or 384-well plates. It was weeks and weeks of work. With the Illumina platform, you prep the library in one tube and put that one sample onto a machine for a day.

Q: How will using the MiSeq system impact your costs?

BP: Sanger sequencing is more expensive. Even though you literally PCR amplify one stretch of DNA and you only have to generate one sequence to be sure that it's right, it is still more costly than having to obtain a certain fold coverage to get all the gaps closed with the short reads on a MiSeq. MiSeq not only saves on cost for reagents, but also in time and personnel. The difference is huge.

Q: How does using the MiSeq system compare to your current methods?

BP: Using the MiSeq system is certainly far, far easier. It's one hole with an arrow pointing at it that says "put sample here." There's no clustering. There's no cleaning the flow cell to get the optics to work properly. There's no getting it loaded in the proper orientation on the machine. There's only one sample so you can't have any problems there.

Q: How does data from the MiSeq compare to what you are used to seeing?

BP: The data quality from the MiSeq is much higher, all the way out for 150 bases. Everything was above Q30.

Q: How does the data compare to what your collaborators are used to seeing?

BP: Some groups want us to run samples and give them back raw data. We are willing to do that, but it definitely comes with associated costs on their end that they're not always prepared for. We've had groups that had no idea of the volume of data they were going to get, let alone what to do with it once they had it. A lot of people have never used any of the new downstream analysis tools. It is not as simple as putting it back into Sequencher. Sequencing is not the same paradigm it used to be.

"MiSeq does exactly what it's supposed to. It works in the specified amount of time. It produces the specified data quality. I've actually been really impressed."

Q: How do you see the MiSeq system changing the research questions people are asking?

BP: MiSeq will introduce more flexibility. You can ask one question and if you don't necessarily get the answer that you're looking for, you can rephrase your question and do it again in a very short turnaround time, which you cannot do now. Some of those larger projects took weeks to months to get in place, get a target designed, and get samples in house. Then it was weeks to months to actually get the work done, do the analysis, and feed it back to the investigator. At that point if they didn't find what they're looking for, the odds of them trying again or starting over were almost nil. I think researchers will start asking better questions. The more you can put into the question and the design of your experiment up front, the better answers you're going to get to your questions in the end.

Q: How would you summarize the primary benefits of the MiSeq system?

BP: Quick and easy. For our purposes, the MiSeq system is going to save us time, save us a fair amount of money, and give us better information about what we're moving forward with as far as samples and libraries go, and that's just in the QC realm. The other is the ability to save money for the people that want to run samples with us, which usually means instead of spending less money, they run more samples. MiSeq does exactly what it's supposed to. It works in the specified amount of time. It produces the specified data quality. I've actually been really impressed.

Learn more about the MiSeq system at
www.illumina.com/miseq

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

FOR RESEARCH USE ONLY

Pub. No. 770-2011-039 Current as of 16 January 2012

illumina®

AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
TCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTT
CGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACC
AACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAA
AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
GATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGAGC
CGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAGTA