

PacBio Shares Early-Access Customer Experiences, New Applications for Sequel II

Mar 07, 2019 | [Monica Heger](#)

SAN FRANCISCO (GenomeWeb) – Pacific Biosciences' Sequel II instrument is now in the hands of five early-access customers, who over the last month ran a total of 58 SMRT cells, company officials revealed last week.

Jonas Korlach, PacBio's chief scientific officer, said in an interview at the Advances in Genome Biology and Technology meeting last week that the new system will launch commercially in the second quarter and that the company will also have trade-in programs for existing Sequel customers wanting to upgrade. The firm will release the expected performance of the system, including average yield, read length, and accuracy, closer to its commercial launch. "We'll take the results from the external sites and improve and tweak a few things," Korlach said.

He also noted that the company aims for a mid-2019 launch of a targeted sequencing protocol that makes use of CRISPR/Cas9 technology and does not rely on amplification. PacBio researchers and collaborators described such a protocol in a 2017 BioRxiv preprint and company scientists collaborated with the Parkinson's Institute and Clinical Center in Sunnyvale, California to target and sequence through pathogenic long repeat expansions. Korlach noted, though, that the original protocol was "too cumbersome," taking between four and five days, and that researchers have since simplified it to between one and one and a half days.

The Sequel II's SMRT chips have 8 million zero-mode waveguides — the wells on the SMRT chip where DNA molecules are analyzed — eight times the number of the earlier instrument. The ZMW increase is expected to result in an approximately eightfold increase in throughput per SMRT cell.

In a presentation at AGBT, Marty Badgett, senior director of product management at PacBio, said that the five early-access customers have run a total of 58 SMRT cells so far, 31 with PacBio's circular consensus sequencing technique and 27 with the standard continuous long read (CLR) sequencing protocol.

Average yield per CCS SMRT cell was 250.4 Gb with 16.7 Gb of unique CCS yield. Average read length was 62.5 kb, with insert sizes between 10 and 13 kb. For the CLR application, early-access customers ran 27 SMRT cells with an average per-cell yield of 67.4 Gb and an average read length of 20.5 kb.

Control samples from PacBio accounted for about 20 percent of the runs, while customers' own samples made up the remaining 80 percent.

Jeremy Schmutz, a faculty investigator at the HudsonAlpha Institute for Biotechnology and an early-access customer of the Sequel II, said in an interview that the increase in throughput will make certain applications feasible that weren't previously possible using the original Sequel due to cost and time constraints.

For instance, he said he is interested in using the instrument to do de novo assembly of plant genomes. While the assembly quality itself is not much different using data from the new instrument, the "major difference is that the throughput is much higher and the cost is projected to be lower," he said, which would allow for projects to be scaled up. "We should look forward to seeing many more de novo genomes," he said.

Thus far, his group has tested out plant de novo sequencing, human genome sequencing using the CCS protocol, and cDNA sequencing. On average, his team is getting around 70 Gb of yield from the long-read protocol and 300 to 330 Gb for CCS libraries, he said.

He estimated that the cost per base on Sequel II would be four to five times less than on the original Sequel, but still around four to five times more than on an Illumina instrument. "The thing I'm most excited about is really cataloguing true human variation," he said. "You can identify the structural variants in a genome and have much higher confidence in the structural variant calls than with Illumina and [you] can also address things like repeats," he said.

Similarly, for plant genomics studies, while resequencing with Illumina yields a lot of information about plant diversity and SNPs, much variation remains undetected since the data is mapped to a reference. "But plants tend to be hypervariable," he said, so "when you view [the data] through the lens of a reference, you avoid portions of the genome." By contrast, long-read sequencing will lead to the "ability to start to build up more comprehensive pan-genome references in plants, so that when we do association studies, we have reference blocks for those plants," he said, which will enable linking of traits and variants.

Also at the conference, Jason Underwood, a principal scientist at PacBio, described a collaboration with Dolomite Bio, a UK startup that has developed a single-cell instrument called Nadia that is compatible with the Drop-seq technology, and with researchers at the University of California, San Francisco, to do single-cell cDNA sequencing on the Sequel II.

In one experiment, the researchers analyzed both chimpanzee and human cells, generating between 5 million and 6 million CCS reads from each sample type.

Underwood noted that the team found evidence they had indeed captured full-length cDNAs, including molecules at the expected size that showed the "hallmark" poly-A tail.

While single-cell sequencing with long reads yields fewer genes per cell than with short reads, Underwood said that the technique would offer a different type of information, such as full-length transcripts and the ability to identify splice junctions and transcriptional start sites.

This reprint has been modified from the original online version to omit pricing information for the instrument.