GENOMICS

MinION takes center stage

One of the first genome sequences produced on a handheld nanopore sequencer shows the platform's potential as well as its challenges.

MinION (courtesy of Oxford Nanopore Technology)

Nicholas Loman of the University of Birmingham has been following the development of nanopore sequencing since Clive Brown, chief technology officer of Oxford Nanopore Technologies (ONT), announced plans for two nanopore sequencers in 2012.

Loman and others were particularly intrigued by the MinION (pronounced 'min-ion'), a handheld sequencer that could be connected to a laptop via a USB port and promised to deliver reads as long as 100 kilobases in real time. In November 2013, ONT started taking applicants for its MinION Access program, in which Loman and his PhD student Joshua Quick were participants.

The MinION consists of a membrane housing around 500 protein pores that connects two chambers filled with electrolytes and allows a current to flow through. As a DNA molecule passes through the pore, the current is blocked, producing a pattern that is characteristic for each base. The challenge is to measure the current with high enough spatial and temporal resolution to decode each base accurately. In June 2014, Loman and Quick posted a read from their first sample run on the MinION to figshare "simply to show that the instrument could do it," recalls Loman. "Some people thought it was a physical impossibility to sequence DNA with this method." But he acknowledges that the quality and accuracy of the read were not great.

Part of the problem was a suboptimal library preparation method—two enzymes need to be attached to the DNA to facilitate its processing through the pore. In what ONT

SYNTHETIC BIOLOGY

BACTERIAL RECALL

Bacterial populations get outfitted with stable analog genetic memory.

Synthetic biologists have recently begun to bestow the powers of memory on bacteria. One strategy creates permanent memories by coding the events sensed by a cell directly in its DNA, relying on irreversible recombinases to generate inversions that can be read as digital bits of information. Timothy Lu at the Massachusetts Institute of Technology and his graduate student Fahim Farzadfard are now claiming a greater share of DNA's enormous coding capacity by generating analog memory devices that can encode or 'write' a variety of sequence changes into nearly any genomic site.

Lu comes from an engineering background. "In our field, we know that you have to have computing plus memory to really make powerful systems," he says. Long-term memory could be used for environmental or health monitoring. Lu foresees it ultimately being used to study development or neuronal activity by recording cell signals *in vivo*.

Their new system synthetic cellular recorders integrating biological events (SCRIBE), began with a good question and a literature search. "He's very creative," says Lu about Farzadfard. "We talked about trying to do *in vivo* writing and he went out first to search for elements that would help us." In older papers, Farzadfard discovered retrons as a way to generate single-stranded DNA (ssDNA).

Retrons are bacterial genetic elements with obscure functions. "It is pretty amazing that no one still really knows what they do," says Lu. They consist of a reverse transcriptase, a short transcribed sequence that primes reverse transcription, and a template sequence for generating ssDNA. Farzadfard and Lu showed that expressing the cassette from an inducible promoter yields ssDNA in proportion to inducer levels. Moreover, a portion of the template can be replaced with other sequences.

To execute DNA writing, they designed a template homologous to a genomic target site but bearing mutations, and expressed viral beta recombinase in *Escherichia coli* to promote oligonucleotide-mediated recombination at that site. The result was a distribution

calls the R6 version of the library prep, a motor protein binds to one end, a hairpin connects the DNA strands and a second motor protein attaches to the hairpin. The purpose of these proteins is to slow the transition down enough for the individual bases to be recorded. The reads of highest accuracy are '2D reads', in which the template strand is sequenced, followed by the hairpin and the minus strand. Poor binding efficiency of the second motor protein to the hairpin reduced the number of 2D reads in the initial MinION sequencing runs.

In September, ONT presented the R7.3 library prep, with the key improvement of having the second motor protein already ligated to the hairpin. With the R7.3 kit in hand, it took Loman and Quick only two weeks to produce data on an *Escherichia coli* genome and upload it to the GigaDB sequence database (Quick *et al.*, 2014). Loman teamed up with Aaron Quinlan of the University of Virginia to write the program Poretools for quality control and downstream analysis (Loman and Quinlan, 2014). He hopes that their data will trigger the further development of tools for better alignment, variant calling and *de novo* assembly.

Loman also predicts that a mixture of different pores on a flowcell may improve the quality of the reads. Accuracy will also increase with the number of 2D reads. At present Loman sees an accuracy of 85% if both strands run through the pore correctly. If only one strand gets through, the accuracy drops to ~70%.

Loman deems ONT's handheld device as ideally suited for diagnostics or environmental monitoring that does not rely on very high accuracy. He is also leading efforts to field test the device.

Nicole Rusk

RESEARCH PAPERS

Quick, J. *et al*. A reference bacterial genome generated on the MinION™ portable single-molecule nanopore sequencer. *GigaScience* **3**, 22–28 (2014).

Loman, N.J. & Quinlan, A.R. Poretools: a toolkit for analyzing nanopore sequencing data. *Bioinformatics* **30**, 3399–3401 (2014).

of the mutations throughout the population. Modeling and simulations confirmed a straightforward relationship between the magnitude and duration of signal—analog properties—and mutation frequency in the population.

The researchers showed that memory can be written and rewritten at multiple loci in response to chemical or light induction and can be read out using antibiotic resistance, reporter expression or DNA sequencing. Writing can be independent from reading, for example in 'sample and hold' circuits, in which a second inducer is needed to trigger the reading of an initial inducer's output. Changes can be written into any DNA site to easily scale up the number of 'memory registers', so unlike other systems, SCRIBE is not limited by the number of orthogonal recombinases.

The intermediate frequency of beta-mediated recombination means that the population can record over long periods. In contrast, high-efficiency digital systems saturate quickly on receiving a signal. Some applications require the single-cell resolution of digital approaches, and Lu and Farzadfard are also working on higher-efficiency approaches.

In general, the system may need to be tested for each context to confirm the nature of the relationship between induction and writing. SCRIBE also requires taking a population census. "In a real-world environment, you may want to constrain the population in some way or have some way of normalizing the signal," says Lu.

SCRIBE is not currently used like a ticker tape that records the order of events; it is more like a closed lake at the base of a waterfall that can be sampled to discover what came over the falls. But the scientists are working on incorporating the logic of biological state machines, which have been used to record event sequences.

Lu is looking to other organisms, taking advantage of the system's modularity. Retrons function in other organisms, and writing systems such as CRISPRs (clustered regularly interspaced short palindromic repeats) may, in the future, make SCRIBE possible in any cell type.

Tal Nawy

RESEARCH PAPERS

Farzadfard, F. and Lu, T.K. Genomically encoded analog memory with precise *in vivo* DNA writing in living cell populations. *Science* **346**, 1256272 (2014).