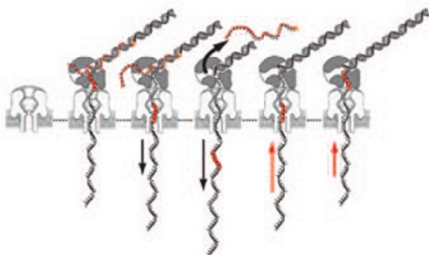
GENOMICS

A nanopore workhorse

Voltage-driven and enzyme-regulated control allows precise ratcheting of a DNA strand in and out of a nanopore.



Nanopore sequencing by ratcheting of a DNA strand, bound to a blocking oligomer and a polymerase, through a nanopore; arrows denote ratcheting direction. Reprinted from *Nature Biotechnology*.

For over two decades nanopores and their potential for sequencing have intrigued innovative scientists such as David Deamer from the University of California, Santa Cruz and Daniel Branton and George Church from Harvard University, who were among the very first to embrace the idea of nanopore sequencing technology. Initially they were met with profound skepticism, but over the years the doubts gave way to interest, culminating in the recent announcement of a commercial nanopore sequencer by Oxford Nanopore Technologies.

The details of the inner workings of the Oxford Nanopore instrument have yet to be revealed, but their promise of cheap, fast and long single-molecule sequence reads is tantalizing. So much so that it would be easy to forget that this development benefited from others that provided basic knowledge of the interaction between DNA and protein nanopores, such as studies similar to a recently published one by Mark Akeson at the University of California, Santa Cruz.

The basic principle of nanopore sequencing is simple: a DNA molecule traverses a pore in response to an applied voltage and in doing so blocks the ionic current in a sequence-specific way. However, challenges remain, as Akeson points out, "one needs to efficiently

GENE EXPRESSION

IN-DEPTH FUNCTIONAL DISSECTION OF ENHANCERS

Massively parallel reporter assays enable the functional dissection of enhancer sequences in cells and *in vivo*.

Biologists excel at cataloging natural genetic sequence variation and are improving ways to associate this variation with meaningful function. What is missing, according to Jay Shendure of the University of Washington, Seattle, are empirical ways to test for functional sequence variation on a large scale.

Two recent reports help fill this need by using deep sequencing to monitor the activity of many synthetic enhancer variants that drive expression of uniquely tagged RNA reporters. Researchers in the Shendure laboratory originally used the idea to interrogate the function of residues in core promoters *in vitro*. In one of the recent reports, a team led by Shendure, Len Pennacchio at Lawrence Berkeley National Laboratory and Nadav Ahituv at the University of California San Francisco modified the assay to accommodate longer enhancers and *in vivo* testing (Patwardhan *et al.*, 2012).

The group first engineered enhancer variants by stitching together oligo mixtures 'doped' with random substitutions using an overlapping PCR strategy. They tested variants with changes at 2–3% of positions of two human and one mouse enhancer up to 620 base pairs long. They placed each enhancer haplotype in front of a reporter with a sequence tag in a plasmid vector and sequenced the library to characterize each enhancer and associate it with a tag.

The resulting mixtures were incredibly complex, representing every possible variant at every position. They injected each enhancer library as a single shot into mouse tail veins and sequenced tags from mouse liver RNA. "People are always saying, 'this mutation caused a 1.5-fold effect, a twofold effect [on the likelihood of observing a specific trait],' but what does that mean functionally?" asks Shendure. Their data allowed them to directly model quantitative effects of sequence variation on gene regulation.

capture a DNA molecule at the entrance of the pore, and the molecule must go through the pore with enough temporal and spatial control to allow recording of single nucleotide blockages. After the sequencing of one strand is completed, the next one must be captured promptly."

Akeson's team tackled these challenges with α-hemolysin, a well-characterized protein. The team found that a complex of phi29 DNA polymerase, single-stranded DNA and a short oligomer that hybridizes to the DNA and blocks DNA synthesis is stable and enzymatically inactive in solution. When the electric field around the nanopore captures the complex, the DNA strand is pulled into the pore and this displaces the blocking oligomer triggers synthesis by the enzyme. As the DNA is replicated, it is pulled out of the nanopore again. Notably, both the voltage-driven zipper and the replication-driven ratchet proceed in distinct steps that are slow enough to be recorded and represent the transition of one nucleotide at a time.

Akeson and his colleagues carefully looked at the error profile of this setup: they detected insertions when the enzyme stepped backward causing some bases to be read twice, and deletions when the incorporation of a base was too rapid for accurate recording. The cumulative error rate is about 25%, but both types of errors can be substantially reduced, according to Akeson, by using an enzyme that does not step back and by improving the temporal resolution to less than one millisecond per event.

Long-term improvements needed to transition from experiments in academic laboratories to a commercial sequencing platform require parallelization, stability of the pore and extensive computational work—something Oxford Nanopore has done exceptionally well, according to Akeson, who consults for the company.

In 6 to 12 months we will know if they are ready for prime time or if more groundwork with the likes of α-hemolysin needs to be done.

Nicole Rusk

RESEARCH PAPERS

Cherf, G.M. *et al.* Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision. *Nat. Biotechnol.* advance online publication (14 February 2012).

In the second report, Tarjei Mikkelsen at the Broad Institute and colleagues designed enhancer sequence variants on custom microarrays (Melnikov *et al.*, 2012). This limited the number and length of the elements but provided better control over the composition of the libraries. They transfected plasmid pools into human cells and sequenced RNA and DNA tags, using the ratio to determine expression associated with the enhancers.

Results from one experiment for the two enhancers they tested agreed with those from painstaking experiments carried out over decades. The group tested linear, nonlinear and thermodynamic quantitative models of sequence-activity relationships and found that relatively simple linear models describe the elements surprisingly well. Their use of inducible enhancers enabled the testing of multiple cell states in the same system.

The two studies had in common several interesting conclusions: the researchers uncovered new binding sites in well-characterized enhancers and found that most sequence changes had small but significant effects on gene expression. They found no obvious signs of strong interactions between sequence positions in either study.

The assays can be used to test the functional effects of variants from candidate disease-associated regulatory regions implicated in genetic studies and to better engineer cell state–specific enhancers for synthetic biology or gene therapy applications. Mikkelsen points out that their approach helped to optimize enhancer characteristics that could be difficult to address with traditional directed-evolution approaches. "Mutations that increase the activity of an enhancer in the target cell state will often decrease its specificity for that state. We can assay large numbers of mutations in every relevant cell state and then combine the data *in silico* to find the optimal trade-off," he says.

Shendure sees a larger recent trend of applying mutagenesis to massively parallel functional assays. "I see a [new] field bubbling up," he says.

Tal Nawy

RESEARCH PAPERS

Melnikov, A. *et al.* Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat. Biotechnol.* **30**, 271–277 (2012).

Patwardhan, R.P. *et al.* Massively parallel functional dissection of mammalian enhancers *in vivo*. *Nat. Biotechnol.* **30**, 265–270 (2012).