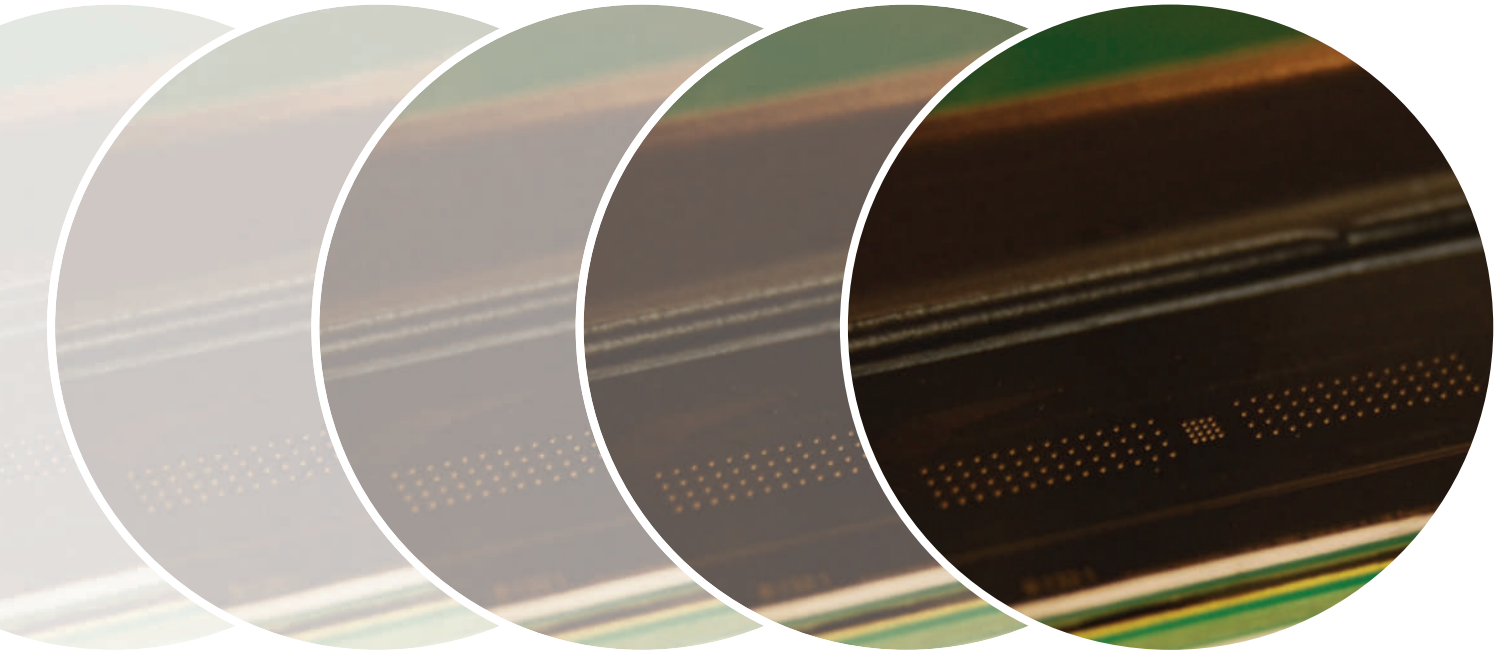


STANDARD AND PORES

Could the next generation of genetic sequencing machines be built from a collection of miniscule holes? **Katharine Sanderson** reports.

DNA sequencing is a technology on the move. In April, 454 Life Science, based in Branford, Connecticut, sequenced the entire genome of James Watson in two months for less than US$1 million[1]. In this issue, Illumina, based in San Diego, California, reports the sequence of a human genome obtained for a quarter of that price and in eight weeks[2]. Companies are positioning themselves aggressively to go further, faster and cheaper.

Many consider the ideal technology to be 'single molecule' sequencing, which reads from individual DNA fragments without the need for amplification and the risk of introducing errors. Pacific Biosciences, based in Menlo Park, California, has placed itself centre stage, promising to deliver such a service by watching enzymes build DNA base by fluorescently tagged base. But the single-molecule technology that the US National Human Genome Research Institute (NHGRI) in Bethesda, Maryland, has invested most in is nanopore sequencing, in which DNA is read as it threads through a tiny hole. The technique has received $40 million of a total of $68 million spent in the institute's drive to generate human genomes for $1,000. $4.2 million of that went to Hagan Bayley, a chemical biologist at the University of Oxford, UK, to back research that forms the basis of Oxford Nanopore Technologies, the company he founded, and the one that is closest to making a working nanopore sequencer.

Jeffrey Schloss, NHGRI programme director of technology development, says that nanopore sequencing is the only method the institute has supported so far that has the potential to sequence DNA directly from cells without amplification, modification or use of expensive reagents such as fluorescent tags. Oxford Nanopore Technologies's chief executive Gordon Sanghera says that he would like his technology to "dominate the world, ultimately". But Sanghera faces stiff competition. Pacific Biosciences, and Complete Genomics in Mountain View, California, are just two of the companies that have announced their ambition to become the chief provider of genetic sequencing. There is still scepticism in the scientific community about whether nanopore sequencing can deliver, says Schloss, and there is a simple reason: "Pacific Biosciences and Complete Genomics have both sequenced some DNA. Nanopores have not."

Oxford Nanopore Technologies's 128-pore chips (top) build on work by Hagan Bayley (above).

One of the first suggestions that nanopores could form the basis for DNA sequencing came in 1996, when a team led by Daniel Branton, a biophysicist at Harvard University, showed that the presence of DNA could be detected as it passed through a pore by the interruption in the flow of ions through the aperture[3].

The pores, made from a ring of seven α-haemolysin membrane proteins, are the same as those that the infectious bacterium *Staphylococcus aureus* pushes into the membranes of other cells in order to create damaging holes. Branton's result suggested that the identity of each of the four bases traversing the hole might be revealed by distinctive changes in ion flow, which can be read as an electrical signal.

From small beginnings

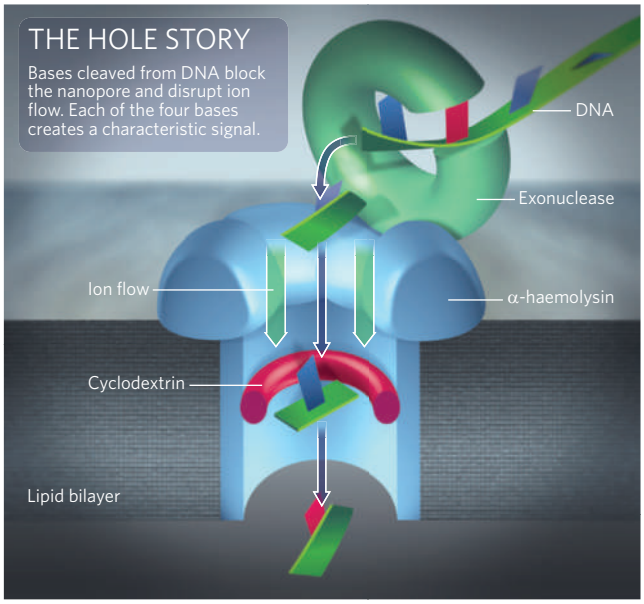
Bayley and Sanghera founded the company in 2005 to develop nanopores as sensor systems for DNA and other molecules, but the company quickly decided to focus on DNA sequencing. Bayley provided 20 years of experience studying nanopores and Sanghera, who had previously worked for Abbott Laboratories, the business know-how. Of the $35 million that has been raised to finance the company, all from private and institutional investors, $20 million came in a financing round in March this year.

In 2006, Bayley showed that the distinction between bases could be made when each was held in place in the nanopore for long enough[4].

"The breakthrough is that one free nucleotide gives a distinguishable signal," says Tim Harris, from the applied physics and instrumentation group at the Howard Hughes Medical Institute's Janelia Farm Research Campus in Ashburn, Virginia.

DNA cannot, for now, be run continuously through the nanopore, partly because of the need to hold each base in the pore long enough to disrupt the flow of ions. So, to do their sequence detection, Bayley's group has used genetic engineering and chemistry to make two alterations. At the pore's mouth, the team placed an exonuclease, an enzyme that grabs the ends of a DNA molecule from a solution running over the top. The enzyme then severs each base and directs it into the hole (see graphic). At the other end of the pore, the group inserted a cyclodextrin plug, a ring-shaped molecule that narrows the neck. The passing bases have to squeeze through this plug and, as they do so, a phosphate group on the nucleotide briefly binds the cyclodextrin and blocks the pore. Because the bases are different sizes, they sit within the cyclodextrin for different lengths of time, and fill it to different extents, giving characteristic readouts for each base.

"The advantage of this technique is, first of all, it's a single-molecule technique, so you don't have to amplify or clone your DNA," says Bayley. There are no fluorescent tags and, in theory, minimal sample preparation. "Also you're directly sequencing the genomic DNA, so, in principle, as well as just getting the four bases you should be able to get modified bases," Bayley adds. Oxford Nanopore Technologies says it has unpublished data showing that the system can better discriminate between the four bases and detect 5-methylcytosine, a chemically altered version of cytosine that is commonly involved in gene regulation.



In May this year, the company decided that its technology had advanced far enough to announce its intention to develop a next-generation sequencing system. The company had also been quietly vacuuming up the intellectual-property rights from some of the leading nanopore research teams, signing licensing deals with leaders in the field such as Branton, and David Deamer and Mark Akeson at the University of California, Santa Cruz. "They're eliminating their competition," says Harold Swerdlow, head of sequencing technology at the Wellcome Trust Sanger Institute in Cambridge, UK.

Key questions

The part of the project that the company is reluctant to talk about is the bit that everyone most wants to know: how this will be scaled up into a working, multichannel sequencer. How many working pores could be used in parallel, and how quickly would it sequence a DNA strand? And crucially, when will sequencing data be made available?

Early prototypes in the company's lab look far from complete. A ten-square-centimetre chip, capable of holding 128 pores that will sequence different DNA fragments, sprouts plastic tubing that delivers the samples and naked wiring that connects to an electronics box. But those at the company are tight-lipped about the details of the final product, how it might work and when. They say that they do not want to oversell themselves by making a

ACGT spells hype

When Complete Genomics, based in Mountain View, California, announced in early October that it would "offer complete human genomes for $5,000", Jim Hudson wondered if it could live up to its claim. So when Hudson, the president and co-founder of the HudsonAlpha Institute for Biotechnology in Huntsville, Alabama, met company representatives at a meeting later that week, he offered them an envelope stuffed with several thousand dollars cash in exchange for his sequence. "I said, 'I want one'," Hudson says.

The representatives declined the envelope, explaining that the company isn't actually taking orders at that price — which was promised by 2009. Hudson and many other scientists are still left wondering whether the company can live up to its promises.

A gap has opened up between the claims made by companies offering the next generation of cheap, ultrafast DNA sequencing and the data to back them up. The $1,000 genome is expected to attract biotechnology and pharmaceutical firms to a sequencing market that has so far been limited to academic centres. Sequencing companies wanting a share of this new market are walking a delicate balance between gaining investor interest and avoiding damaging their credibility with unsupportable claims.

In December 2005, Helicos Biosciences of Cambridge, Massachusetts, announced that it had sequenced the genome of a bacteriophage using a technology that could potentially sequence entire human genomes in one day. The company then raised $40 million in venture capital and nearly $46 million in its May 2007 initial public offering. But problems with its machines then caused long delays while competitors announced that they had sequenced human genomes. Helicos still seems far from achieveing this and has sold few machines, causing its share price to drop from a high of more than $17 this January to a low of 60 cents last month. Chief scientific officer Patrice Milos says the company has ironed out its problems and will present new data at the Advances in Genome Biology and Technology (AGBT) meeting in February 2009.

Helicos has also been hurt by the debut of another big name in this field. Pacific Biosciences of Menlo Park, California, has been developing a new sequencing technology based on DNA polymerase, an enzyme that copies DNA, and made a splash by presenting data at the AGBT meeting in Marco Island, Florida, this February. There, the company showed that it had sequenced pieces of DNA about 150-base-pairs long — a tiny step towards its eventual goal of sequencing entire human genomes in less than 10 minutes, but a tangible piece of data that has built up the company's credibility among

specific prediction that they will do *X* in *Y* time, and then disappointing or surpassing those expectations. "There's a danger for a company like this to come out too soon," Sanghera says. "It's a very difficult commercial strategy." (see 'ACGT spells hype').

"There's danger for a company like this to come out too soon." — Gordon Sanghera

Swerdlow is talking with all of the new companies. "It's quite difficult to decide who's telling the truth," he says, "It's all hearsay to some extent." He remains optimistic but unconvinced about Oxford Nanopore Technologies. His concern is whether the reagents needed to run a sequence might break down the biological pore in some way.

"I do think that there is some scepticism about direct nanopore sequencing," says Barrett Bready, chief executive of sequencing start-up NABsys in Providence, Rhode Island. He says this scepticism is based on the inherent difficulty of the problem. "The four bases actually differ by only a few atoms. These differences must be detected in the face of noise from various sources."

NABsys, formed in 2004 by Xinsheng Sean Ling, a physicist at Brown University in Providence, is also pursuing nanopore sequencing, but seems to be further from a working machine than its Oxford rival. In 2007, Ling and John Oliver, another NABsys scientist, received two NHGRI sequencing grants worth $1.32 million in total. The method is based on a silicon chip dotted with synthetic nanopores. Through these pores pass 100,000-base-long fragments of genomic DNA that have six-base-long probes attached to them at intervals. The method uses a library of probes, each having a different, but known, sequence. As the DNA passes through the pore, the points at which a probe is attached can be detected from the current in the chip. The time gaps between those current readings allows the location of the probes to be determined. Once lots of fragments are probed in this way, a picture of the entire genome can be put together from these sequences. But Bayley is dubious. "You can engineer [proteins] with angstrom precision, which you simply can't do with a pore in plastic or silicon nitride at this point." And Harris says that NABsys's sample preparation, which involves reengineering the DNA, is clunky. "This seems like an improbably gymnastic sample process for something that has to be fast, and essentially free."

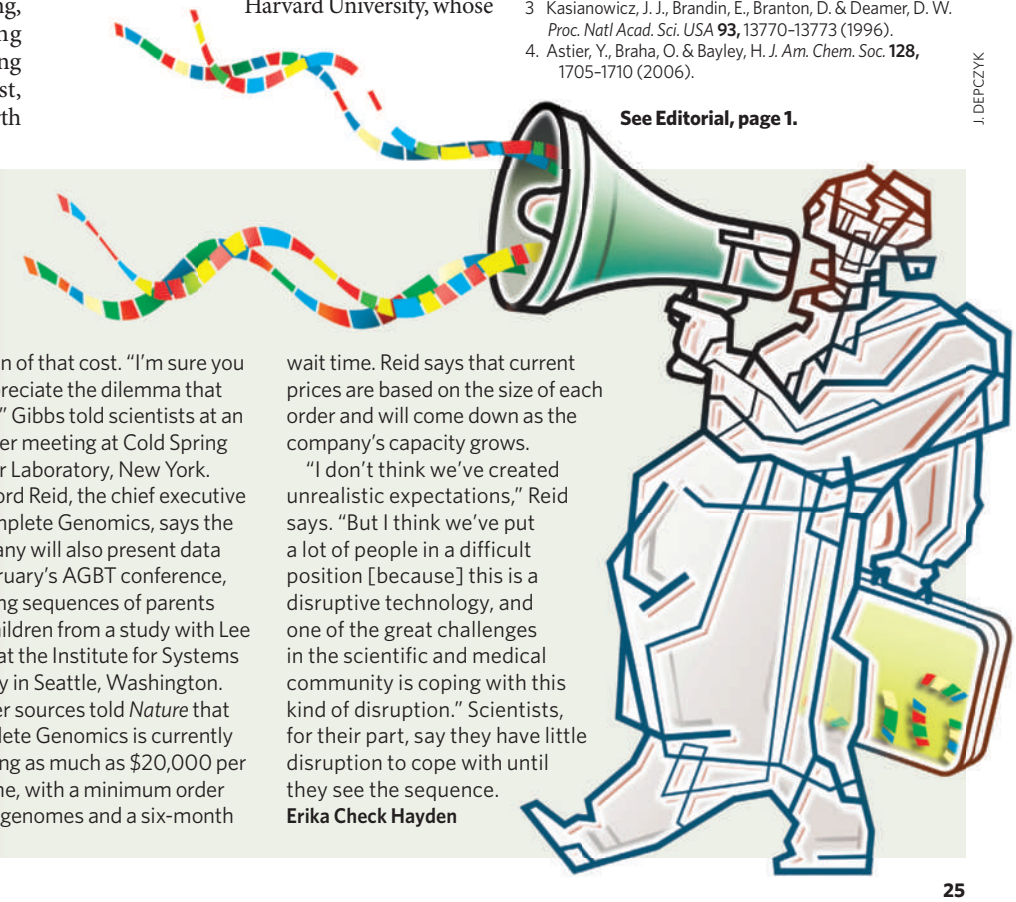
George Church, a molecular geneticist at Harvard University, whose work has also been licensed by Oxford Nanopore Technologies, thinks that the sequencing race will be won by whichever company has the lowest instrument cost and the highest throughput per instrument. Sequencing methods that rely on a digital camera to record colour changes from fluoresently tagged bases — such as Pacific Biosciences' technology — are winning that race over nanopores, he says. "Digital cameras are capable of collecting millions of bits of information at close to the maximum data-flow rate that a PC can handle." The cost of these cameras has dropped because of huge consumer use, says Church. "It does not seem to be a similar case for massively parallel ion-channel monitors."

Schloss says that the NHGRI views nanopore sequencing as a long-term goal. "We expected, when we launched the programme in 2004, that it might well take ten years to achieve the goal of using nanopores for sequencing DNA." Sanghera has no such reservations. "Our products are going to be so good that we're just going to let the technical data speak for itself. All things will flow from that." ■

Katharine Sanderson is a reporter for *Nature* based in London.

1. Wheeler, D. A. *et al. Nature* **452,** 872–876 (2008).
2. Bentley, D. R. *et al. Nature* **456**, 53–59 (2008).
3 Kasianowicz, J. J., Brandin, E., Branton, D. & Deamer, D. W. *Proc. Natl Acad. Sci. USA* **93,** 13770–13773 (1996).
4. Astier, Y., Braha, O. & Bayley, H. *J. Am. Chem. Soc.* **128,** 1705–1710 (2006).

See Editorial, page 1.



J. DEPCZYK

scientists — and helped it to raise $100 million in venture capital as of July 2008.

Complete Genomics' surprise entrance was more audacious than anything the field has seen so far. The company pledged to cut sequencing costs to $5,000 in six to nine months, but hasn't offered any data, annoying some potential customers. Richard Gibbs, director of the Baylor College of Medicine genome sequencing centre in Houston, Texas, says he had just convinced a private donor to fund a large cancer genome sequencing study when Complete Genomics' announcement hit the press, prompting the donor to ask why Gibbs needed $350,000 per genome when a new company was sequencing genomes at a fraction of that cost. "I'm sure you all appreciate the dilemma that poses," Gibbs told scientists at an October meeting at Cold Spring Harbor Laboratory, New York.

Clifford Reid, the chief executive of Complete Genomics, says the company will also present data at February's AGBT conference, showing sequences of parents and children from a study with Lee Hood at the Institute for Systems Biology in Seattle, Washington.

Other sources told *Nature* that Complete Genomics is currently charging as much as $20,000 per genome, with a minimum order of five genomes and a six-month wait time. Reid says that current prices are based on the size of each order and will come down as the company's capacity grows.

"I don't think we've created unrealistic expectations," Reid says. "But I think we've put a lot of people in a difficult position [because] this is a disruptive technology, and one of the great challenges in the scientific and medical community is coping with this kind of disruption." Scientists, for their part, say they have little disruption to cope with until they see the sequence.

Erika Check Hayden