Speed reading

As competition hots up for the $10m Archon Genomics X PRIZE, it turns out that physicists may have the leading edge in the quest to sequence the human genome quickly and cheaply, as **Philip Ball** explains

Philip Ball is a science writer and journalist based in London, UK, e-mail p.ball@btinternet.com

Which is harder: building a manned spacecraft, or sequencing a genome? The last time the US-based X PRIZE Foundation put a $10m bounty on the table, it was for exploring outer space: for privately building a spaceship to fly three passengers to an altitude of 100 km. But the same sum is currently on offer for a triumph in inner space. The X PRIZE Foundation, which offers whopping cash incentives for "radical breakthroughs" in technology that will benefit humanity, seeks to reward the first privately funded company that can accurately sequence 100 human genomes in 10 days, at a cost of less than $10 000 per genome.

According to the foundation, the Archon Genomics X PRIZE, launched in 2006, aims to stimulate an innovation that would revolutionize medicine by ushering in a day "when you and your doctor sit down to review a copy of your own personal genome". That is probably too simplistic a picture – after all, we still do not know what most human genes do, nor how they interact in gene-related diseases. Nevertheless, if the extraordinary improvements in speed and cost of genome sequencing – stimulated by the Human Genome Project (HGP) – continue, they will surely have a profound impact on our understanding of genetic biology, and consequently not just on healthcare but on areas ranging from microbiology to anthropology.

And it is looking increasingly likely that basic physics and chemistry will be central to such advances. To read the sequence of bases in a strand of DNA faster and more easily, techniques entirely different from those used for the HGP are needed. These may rely on an ability to manipulate DNA and enzymes at the level of single molecules, and will therefore involve nanotechnology. What is more, one of the most promising of the recent candidate methods uses the material that claimed this year's Nobel Prize for Physics: graphene (sheets of carbon one atom thick).

Techniques to read the sequence of bases in a strand of DNA faster and more easily may rely on an ability to manipulate it at the level of single molecules, and will therefore involve nanotechnology

Threading the needle

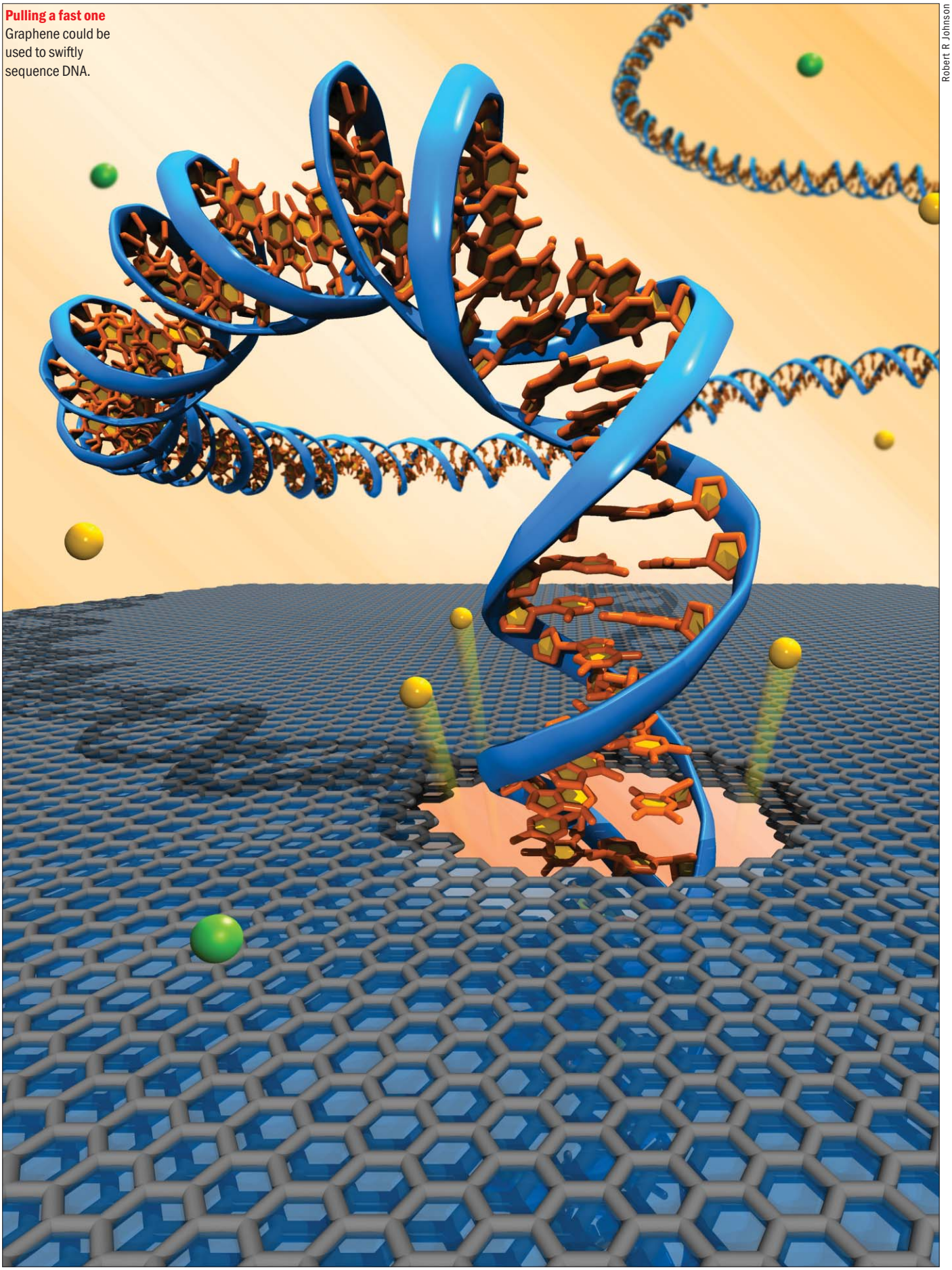
The first draft of the entire human genome sequence, announced by the HGP in 2001, was obtained using a complicated procedure: the DNA is chopped up into fragments by enzymes, "amplified" by replication, chemically labelled with fluorescent dyes, and then separated by the technique of gel electrophoresis to read out the sequence of nucleotide bases. During the HGP, this approach was streamlined and automated to speed it up, and in the past few years it has been supplemented by new, even faster, techniques.

But for all their ingenuity, these methods still provide a very cumbersome, roundabout way of getting at information that exists in the molecule as a simple linear sequence of four "letters" (the chemical bases, denoted A, T, G and C, on each nucleotide building block). How much better it would be if we could just zoom in on a single strand of DNA and "see" these letters one by one.

In the mid-1990s chemist David Deamer of the University of California, Santa Cruz was one of several researchers who realized that there might be a way to do just that. Deamer imagined threading a DNA strand through a pore so small that the molecule almost filled it. In a salt solution, pores will let dissolved ions pass through, and these tiny ion currents can be measured using the "patch-clamp" technique developed in the late 1970s. But if a DNA base partially blocks the pore, this reduces the ion current. And because the chemical structure of each DNA base is slightly different, their pore-blocking effects may be distinguishable: in other words, the ion current can vary depending on which of the four bases is blocking the pore neck at any moment. If the DNA strand is pulled steadily through the pore by an electric field, step changes in the ion current should reveal the sequence of bases.

In 1996 Deamer teamed up with Dan Branton of Harvard University to demonstrate a proof of principle of this technique. They showed that short strands of both DNA and RNA containing just one kind of nucleotide could be drawn through the nanopore in the centre of a bacterial membrane protein called α-hemolysin (α-HL), a hollow mushroom-shaped molecule with a narrow cylindrical neck that can be anchored in a lipid membrane layer. They saw that the ion current was reduced when the DNA and RNA strands passed through, and that the longer the strand, the longer the blockade. "With further improvements," the researchers boldly claimed at that time, "the method could in principle provide direct, high-speed detection of the sequence of bases in single molecules of DNA or RNA." Yet it was so radical an approach to DNA sequencing that few geneticists took any notice. "Many thought it was crazy," says Branton.

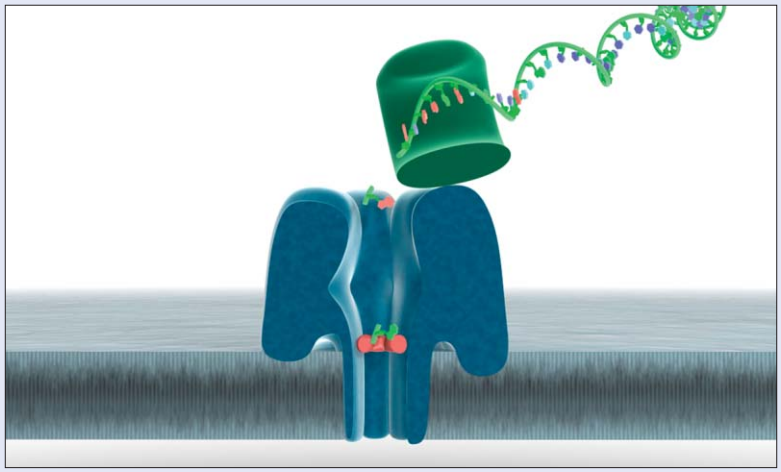
At much the same time, British chemist Hagan Bayley, then at Texas A&M University and now at Oxford University, was also exploring the use of α-HL as a sensor of substances lodged within the pore. In 1999 his group reported that when the pore was equipped with an "adapter" – the ring-shaped sugar molecule cyclodextrin, which fitted snugly into the opening – changes in the ion current could be used to discern various small organic molecules. Although Bayley was aware of Deamer's notion of nanopore genome sequencing, he says that it was only in 2004, when the US National



Pulling a fast one
Graphene could be used to swiftly sequence DNA.

Robert R Johnson

1 DNA muncher



Oxford Nanopore Technologies

A DNA strand is grasped by an exonuclease protein, which chops off individual bases one by one. These then feed into the mushroom-shaped α-HL molecule, which sits within an electrically resistant lipid bilayer, and through the even smaller cyclodextrin molecule attached within it. Different DNA bases produce characteristic disruptions in the flow of ions in the salt solution passing through the nanopore. Electrodes above and below the lipid bilayer both drive and detect the change in this tiny ion current, enabling the DNA sequence to be identified.

Institutes of Health (NIH) announced an initiative for developing technologies that could offer human genome sequences for $1000 apiece, that he decided to take up the challenge. With an NIH grant, in 2005 Bayley and his collaborators showed that DNA bases can be distinguished by their different attenuation of ion currents when a DNA strand is threaded through α-HL.

Speed limit

But, inevitably, DNA sequencing with α-HL is not this simple. For one thing, the narrow neck of α-HL is about 5 nm long, much longer than the 4 Å height of a single base pair on DNA, which means that about 10–15 bases contribute to pore blocking at any moment. In addition, the electrophoresis tended to whip the DNA strand through at such a rate – about 1–10 μs per base – that it was impossible to collect enough signal to reliably identify each base. In short, much more control was needed of both the pore itself and the movement of DNA through it.

The passage of DNA can be slowed by attaching beads to it to create drag or by lowering the temperature of the ion solution. But one of the most promising approaches uses nature's own speed police: so-called processive DNA enzymes. These are proteins that bind to DNA and step along the strand one base at a time performing tasks such as repair, cutting or chemical modification. If such an enzyme, selected to inspect each base in turn but not actually do anything to it, is tethered at the mouth of the pore, the DNA cannot go through any faster than it is passed on by the enzyme. At typically a few milliseconds per base, the DNA is slowed by about three orders of magnitude. To impose a DNA speed limit, says Branton, "the use of processive DNA enzymes at or near the pore mouth is the best option I know that is available today".

Bayley has devised a more drastic way to control the speed, which could also solve the problem of having several bases lodged in the pore at any moment. Instead of threading the DNA through, he suggests that it might be snipped up by a DNA-munching enzyme called an exonuclease, which cuts nucleotides off the end one by one. If the exonuclease is tied sufficiently close to the pore entrance by a chemical link, the free nucleotides will have nowhere to go but down into the pore (figure 1). In 2006 Bayley's group showed that an α-HL pore fitted with an adapter (a chemical derivative of cyclodextrin) could distinguish each of the four isolated nucleotides from one another by their blockage effects on the ion current.

In 2005, shortly before making this discovery, Bayley co-founded Oxford Nanopore Technologies, a company that aims to develop his fundamental research in nanopore sensing for DNA sequencing and other applications. While the company is exploring all options, it seems that this "exosequencing" approach may be the one closest to commercialization (figure 2), although it is cautious about releasing any details yet.

Another boon of Bayley's α-HL sensor is that the adapter in the pore neck can create enough sensitivity not just to distinguish nucleotides, but also to spot chemical modifications to them. Despite what genome-sequencing efforts sometimes imply, the message in your genes does not depend solely on their sequence. This is because genetic information is overwritten by chemical modifications to DNA, called epigenetic tags, that may alter the way a gene is expressed. In particular, cytosine (C) bases are sometimes altered by the addition of a methyl or hydroxylmethyl group, so that in effect there are not four but six key "letters" (bases and modified bases) in the "epigenome". Bayley and colleagues have shown that both exosequencing and sequencing of threaded strands can in principle identify all six letters of the epigenome. Since some researchers consider that sequencing the human epigenome will be at least as important as sequencing the primary genome, this is an immensely promising advantage.

Last August a team led by Jens Gundlach at the University of Washington – a physicist who combines an interest in gravitational waves with DNA sequencing – added a new pore to the game that has potential advantages over α-HL. The researchers found that an engineered form of the membrane protein from *Mycobacterium smegmatis* can also be used to distinguish all four nucleotides, with the advantage that the "sensing region" in the narrow neck of the pore is only about 5 Å long, and so should not suffer from α-HL's defect of potentially reading many bases at once.

Chips with holes

The problem with a protein in a lipid membrane is that it is not the most robust of systems to develop into a device. That is one reason why some research groups have considered making artificial nanopores from inorganic materials such as silicon nitride, carved, for example, with an electron beam. Another attraction of this approach is that it could offer new read-out mechanisms based on semiconductor microelectronics. For example, one might measure the tunnelling current flowing *across* an insulating pore using two electrodes embedded on opposite sides of the opening: basically the same kind of current that is used for atomic-resolution imaging in scanning tunnelling microscopy.

Calculations imply that objects placed between the electrodes may alter this current in a way that depends on their exact chemical nature. Indeed, early last year Stuart Lindsay and colleagues at Arizona State University reported experiments showing that all four DNA bases can be distinguished in this way.

Despite the daunting technical challenges, a European team led by Radomir Zikic at the Institute of Physics in Belgrade, Serbia, and involving other groups in Switzerland, Ireland and Israel, is attempting to develop a device that uses this detection principle. They envisage pulling DNA through a silicon-nitride nanopore while monitoring the current between two carbon-nanotube electrodes – tiny electrical contacts made of rolled-up graphene.

So far, the researchers have managed to deposit nanotubes on the substrate membrane and drill through the assembly with an electron beam. But to be able to discriminate between different bases using the tunnelling current, the ends of the nanotubes need to be very close together and the hole itself must be no more than a nanometre or two across; if the pore is too wide, then the DNA bases can rotate, making the data harder to analyse. Aleksandra Radenovic at the Swiss Federal Polytechnic Institute in Lausanne (EPFL) says that the researchers are considering adding chemical groups to the nanotube ends to bind the bases and hold them steady. And to slow down the DNA's speed, they are aiming to hold the strand in an optical trap. "It's a difficult route, and might not work," she admits. But the degree of control over the molecular motion that could be achieved in such a device would make it valuable for fundamental studies in molecular biology, such as interactions between DNA and proteins, even if it does not prove the best method for fast sequencing.

One of the challenges with solid-state nanopores is that the sheets in which they are punched are typically rather thick, so that again many nucleotides occupy the pore simultaneously. Another problem is that it is hard to make pores with a very precisely defined and reproducible size and shape. Bayley has recently teamed up with physicist Cees Dekker at the Delft University of Technology in the Netherlands to combine the benefits of the organic and inorganic approaches by capturing α-HL inside an inorganic nanopore in silicon nitride. First, the researchers attach an α-HL molecule to the end of a double-helical strand of DNA, which serves here simply as a pull-rope. They use electrophoeresis to pull the DNA through a pore that is too narrow to allow the protein through, so that it gets lodged in the opening. The DNA is then snipped off.

Graphene to the rescue?

But now there is a new option that is generating huge excitement. Earlier this year, three research groups independently reported that they had drawn DNA through a nanopore drilled into a sheet of graphene (figure 3): a sheet of pure carbon in which the atoms are linked in a flat array of hexagons. First isolated in 2003 by Andre Geim and Konstantin Novoselov from the University of Manchester in the UK, who shared this year's Nobel Prize for Physics for their efforts, graphene has been hailed as one of the most exciting new materials of our age: highly electrically conducting, ultrathin, strong and inexpensive. "It's an incredible material," says Bayley. The effective thickness of single-layer graphene of just 6 Å means that only a single base will reside inside a pore when threaded with DNA.

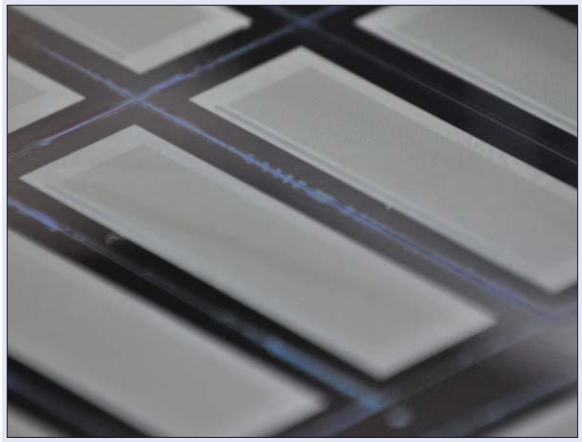
Threading of a graphene nanopore by DNA, and the consequent blockage of an ion current, was reported at much the same time by Dekker's team at Delft, a group at the University of Pennsylvania, and by Branton and colleagues at Harvard. Although none of the groups has yet demonstrated that the ion-current blockades can distinguish the DNA bases, they feel that graphene may prove to be pivotal to the technology. "I believe that graphene may indeed be a game changer," says Marija Drndic of the Pennsylvania team.

Graphene's high conductivity also suggests new ways to probe the nature of DNA in the pore. Bayley is sceptical that measuring the transverse tunnelling current will be a good method, because it is likely to be acutely sensitive to uncontrollable factors such as the precise nucleotide orientation. But he thinks that sequence information might be teased out of the perturbation of one of the exotic 2D electronic-transport properties of graphene. What is more, he says, the pores could be better defined than those drilled in other inorganic materials: for example, the edges of the pore could be tidied up by chemical modification. It might even be feasible to use chemical methods to build precise holes in the first place, for example by excising a precise number of hexagons from the graphene sheet – which is exactly what Dekker's group is now working towards.

Who dares, wins

If any of these techniques is going to become a commercial reality, says Bayley, speed may be at least as important as cost. That is one reason why there is a big push to develop highly parallel systems in which hundreds or even thousands of pores are all reading bits of DNA at the same time. As single-molecule methods have higher error rates than the traditional labelling

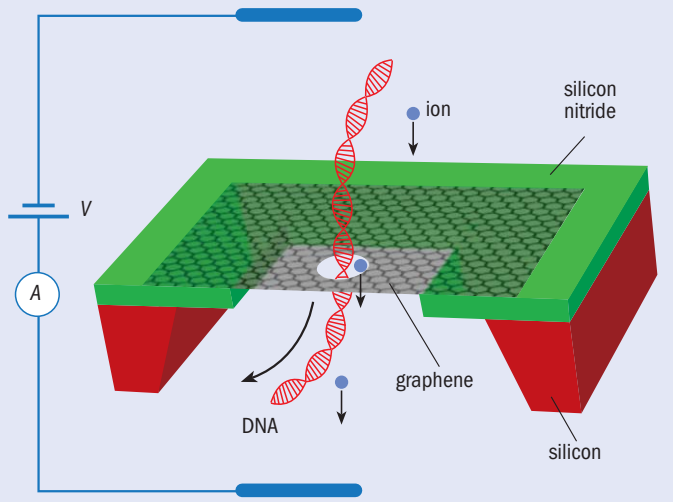
2 Nanopore device



This silicon chip developed by Oxford Nanopore Technologies contains an array of microwells, each of which is an individually addressable electronic channel. When used in a device, each microwell is filled with salt solution and capped with a lipid bilayer, into which a nanopore–enzyme complex is inserted, as shown in figure 1. The individual nanopores could act as real-time electronic sensors for DNA.

Earlier this year, three research groups independently reported that they had drawn DNA through a nanopore drilled into a sheet of graphene

3 Graphene nanopore



Graphene may provide an exciting alternative to the protein methods that have so far dominated DNA sequencing. Here, DNA is drawn through a single sheet of graphene, and the resulting change in ion current is detected by an ammeter.

techniques, each base in a genome may have to be read in perhaps 10 or 20 separate runs to weed out errors. Oxford Nanopore can now routinely produce solid-state chips containing many wells that each have a single α-HL molecule housed in a lipid membrane. Bayley foresees a time when a viral genome might be sequenced in barely any more time than it takes to push the button and set the process in motion, while bacteria might take a few seconds and the more complex eukaryotes perhaps tens of minutes. That could be immensely valuable for instant diagnosis of pathogenic infections, for instance.

Curiously, it seems it is precisely because so much else is at stake that the $10m bait of the Archon Genomics X PRIZE does not make many researchers drool. "I don't even really know what it's all about", says Bayley, although he admits that Oxford Nanopore might dip a toe into the race once it has got a working system (the company is not yet at liberty to say when that will be). And others working on the technology, such as Dekker and Branton, say they have neither the resources nor the patience needed for the laborious job of turning innovative discoveries into reliable commercial devices. "As basic researchers we do not think making this a competitive race will make us or most other research laboratories more productive", says Branton. "A $10m prize is simply not what motivates us."

The first to stump up the $1000 registration fee to enter the X PRIZE race, in November 2009, was a biotechnology company in Taiwan that is developing a single-molecule fluorescence method. It has now been joined by several others, but the technology is moving so fast that there are rumours that the X PRIZE Foundation may be about to shift the goalposts. "The existing methods could probably almost get there already," says Bayley. But he, like many others working on candidate technologies for ultrafast sequencing, feels that the real prize would not be a pot of gold, but the buzz of having transformed genomics into a discipline that can really start to explain how we are put together. ■