

Democratizing sequencing

The DNA sequencing method that earned Fred Sanger the 1980 Nobel Prize was, with a few critical technical improvements, the very same method that allowed the undertaking and completion of the human genome project. This summer, sequencing methods made the scientific news again with the publication in *Nature* and *Science* of two novel techniques. These works are the tip of an iceberg of activity aimed at developing the next generation of sequencing technologies to provide affordable routine genome sequencing. With commercial companies making it a business objective and agencies such as the US National Human Genome Research Institute making it a funding priority, the goal of ultimately reducing the cost of sequencing from the present \$10 million to \$1,000 per human genome is aggressive but not beyond reach.

The ability to easily sequence mammalian genomes would be tremendously enabling in many areas of research: from improving genetic studies in humans, to dissecting the cellular circuitry involved in cancer, to increasing the power of comparative genomics for annotation of the human genome. Reducing the cost to \$1,000 per genome could also make it a medical tool for making prognoses and predictions of treatment efficacy—a pharmacogenomics revolution of healthcare.

The two recent papers describing successful implementations of new sequencing principles clearly indicate that a substantial reduction in cost can already be achieved in the near future. The company 454 Life Sciences reports on the development of an integrated platform, using treasures of engineering and ‘sequencing-by-synthesis’ (p646). The platform, available as an instrument that is already selling to sequencing facilities, provides a 100-fold throughput improvement over the conventional Sanger technique and corresponding cost savings that make the company say they are ‘democratizing sequencing’. The independent approach taken by the Church group at Harvard University uses the technique known as ‘sequencing-by-ligation’ to achieve a ninefold decrease in cost (*Science*; published online 4 August 2005; doi: 10.1126/science.1117389). The gains in cost and throughput achieved by both groups rely principally on eliminating bacterial transformation from the process and operating in very small volumes in a highly parallel fashion, allowing millions of reactions to take place simultaneously. There is still room for improvement—in read length, accuracy and ability to correctly discriminate stretches of the same

base—but these technologies clearly constitute a first quantum leap in cost reduction.

The next wave of technologies now in the works is aimed at sequencing single molecules. Eliminating the template amplification, which is still a prerequisite in present methods, would reduce cost, time and amount of starting material while eliminating amplification bias and dephasing. Among the notable efforts are those of the UK-based company Solexa, which tackled the problem by optimizing novel chemistry for a base-by-base sequencing reaction on a very high-density solid support. Also, the Webb group at Cornell University is developing, in collaboration with Nanofluidics, the implementation of an optical tool called ‘zero-mode waveguide’, which restricts fluorescence observation to a region as small as the catalytic site of an immobilized polymerase, allowing visualization of fluorescent nucleotide incorporation in real time.

One step further removed from a practical application, but promising extreme efficiency gains, are the radically new technologies that do not involve DNA synthesis. The prominent theme in that futuristic arena is inferring a sequence by measuring the distinct variations of current that each nucleotide causes as a stretched molecule of DNA passes through a nanopore.

So, even if we are not close to the \$1,000 genome yet, the technology is on the move. But is society ready for this? The short answer is no. How exactly should genomic information be interpreted and used? Someone’s genome sequence, determined to answer a specific question, comes with a flurry of unrelated information that the individual may not want to know, such as susceptibility to diseases for which there is no treatment. Questions related to confidentiality and consent also remain partially unanswered, and a legislative framework will be required to prevent misuse and improper disclosure of information. Research in these societal implications is actively supported by initiatives such as the US Ethical, Legal and Social Issues (ELSI) program, which operates with a portion of the human genome project budget.

The scientists who develop and use novel sequencing technologies will be instrumental in the ensuing debates, providing input on important elements such as predictive value of the data and norms for quality control. Their commitment to this debate, and to education of the public and the legislators is crucial. Democratizing sequencing is not only an issue of cost.