ACS SyntheticBiology

Viewpoint

pubs.acs.org/synthbio

When It Rains, It Pores

Sriram Kosuri*,†,‡ and A. Michael Sismour*,†,‡

†Wyss Institute for Biologically Inspired Engineering, Harvard University, 3 Blackfan Circle, Fifth Floor, Boston, Massachusetts 02215, United States

‡Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, NRB 238, Boston, Massachusetts 02215, United States

Clive Brown of Oxford Nanopore stunned the crowd at the 2012 Advances in Genome Biology and Technology conference in Marco Island with the announcement of two new DNA sequencing instruments, based on protein nanopores, that will be available by the end of the year.[1] The talk detailed the release of two sequencing instruments, the GridIon and MinIon, that have the potential to dramatically change the landscape of DNA sequencing. This is the latest development in a series of breakthroughs over the past five years in DNA sequencing technologies that have brought down costs 10,000-fold.[2]

These advances are a direct result of the large private investments in technology development during this time. Dozens of sequencing methods are already commercially available or are in development, including pyrosequencing (Roche/454), sequencing-by-ligation (Life Technologies/SOLiD, Complete Genomics), sequencing-by-synthesis (Illumina/Intelligent Biosystems), FET sequencing (Life Technologies/Ion Torrent), microfluidic sequencing (GnuBio), electron microscopy sequencing (Halcyon Molecular), and optical single-molecule sequencing (Pacific Biosciences, Helicos). The widespread availability of cheap sequencing has fundamentally changed biological research, not only related to genome sequencing but also as a research tool to study diverse processes such as transcription, gene regulation, epigenetic modification, genome structure, differentiation, and evolution.

Oxford's technologies expand upon published work on using protein nanopores spanning a membrane to sequence DNA.[3] Briefly, a membrane that does not allow current to pass separates two compartments that are held at different voltages. A membrane protein known to form a small pore (nanopore), such as *Mycobacterium smegmatis* porin A, is inserted into the membrane, which allows a small measurable current.[4] Enzymes, such as phi29 DNA polymerase, thread individual DNA molecules through the nanopore in order to control the rate of DNA traversal across the membrane.[5] Changes in current caused by blockage of the pore by DNA bases can be correlated to the identity of those bases. In this manner they are able to obtain DNA sequences for long individual molecules of DNA. Oxford has spent several years screening appropriate pores and enzymes to control and detect nucleotides as they pass through the membrane, developing electronics to rapidly identify the sequence, and packaging the technology into reliable instrumentation.

If Oxford can deliver, they have several advantages over most other sequencing technologies on the market. First, electronic detection allows for fast scaling and miniaturization without the added complex optics inherent in many other systems. For example, current plans for the GridIon start at 2,000 nanopores per disposable cartridge and quickly move to 8,000 nanopores within a year. In addition, their announcement of the USB key form factor for the MinION demonstrates the ability to miniaturize the technology. Second, nanopore sequencing allows for very fast sequencing of DNA on the order of hundreds to thousands of bases per second per nanopore. For example, a cluster of 20 GridIons (8,000 nanopore version) can sequence a human genome in 15 min. Third, there is minimal sample preparation, which due to labor and consumables is a large component of overall sequencing costs. Fourth, single-molecule detection allows for reduction of PCR amplification biases and possibly detection of modified bases to study epigenetics. Although the quoted error rate (4%) is worse than that of Illumina and SOLiD, it is sufficient to achieve good mapping and significantly better than previously released single-molecule sequencers such as the PacBio RS. Fifth, read lengths of up to 100 kilobases can be achieved, paving the way for better *de novo* and metagenomic sequencing and an improved ability to detect structural variations and phasing in resequencing. Finally, real-time data acquisition allows for sequencing until reaching the desired level of coverage without prior optimizations.

While the latest announcement has generated a lot of excitement, the stated costs are still on par with current sequencing technologies. Reducing costs in the future will depend on how quickly Oxford can scale the technology. In addition, some assays such as RNA-Seq depend more on the number of reads than on read length. The announcement did not detail how easily the nanopores switch from one molecule to another, and thus it is unclear if these systems can compete with existing instruments on the number of reads per run. Finally, alignment algorithms and assembly methods that take advantage of longer read lengths and real-time data acquisition while accounting for dominant error types in the sequencing technology are presumably yet to be developed.

The impact of the Oxford's sequencers on synthetic biology might be modest. To date, next generation sequencing has largely been used indirectly by synthetic biologists, some of whom mine for novel or improved enzymatic function from the overflowing amount of sequence data now available due to next-generation sequencing. Since Oxford's sequencing costs are comparable to current technologies, it is unclear if they will appreciably accelerate the availability of metagenomic sequence, though long read lengths will certainly help. Perhaps the more pressing question for the synthetic biology community is whether such rapid progress will be seen in technologies that

Received: March 12, 2012
Published: April 20, 2012

   dx.doi.org/10.1021/sb300015f | *ACS Synth. Biol.* 2012, 1, 109–110

are more central, such as gene synthesis and genetic engineering techniques.

Energy, materials, chemicals, and medicine are all large markets that synthetic biology already affects and could be equally, if not more, lucrative than sequencing in the long run. For the synthetic biology community, reductions in the cost of gene synthesis, genetic engineering techniques, or other basic enabling technologies would allow expedited progress in the development of genes, pathways, and organisms tailored to specific functions. So why do not we see the same types of direct investments we see in sequencing technologies and computer chip manufacturing in gene synthesis or genetic engineering?

The problem may lie in the fact that enabling technologies such as gene synthesis and genetic engineering are used as a tool in research and development to help engineer novel organisms, which in turn generate products that can be sold to a broad consumers base. A company can take a single engineered organism and sell chemical, material, or energy products to a vast number of people. This is in contrast to sequencing technologies, where the investments were not driven by the research and development markets as much as the potential of genomic sequencing as a diagnostic tool and the ∼7 billion people that could benefit from it. In the microprocessor industry, which has experienced comparable gains, investments in fabrication technologies produce microchips that are directly sold in consumer products. Companies that invest in the enabling technologies for synthetic biology cannot directly monetize those investments by selling a product directly to consumers, as is the case with the microchip and DNA sequencing industries.

There have been many large investments in genetic engineering technologies. For example, Monsanto has heavily invested in technologies to better genetically engineer various plant species used in farming. Monsanto is able to monetize these gains by being a vertically integrated company that can sell products resulting from their enabling technologies to farmers around the world. However, this vertical integration simultaneously restricts access to these techniques. Another example is Sangamo Biosciences, whose 15-year investment in engineered nucleases allows for genetic engineering in arbitrary species. Sangamo is able to monetize on these investments and make them available to a larger community through licensing agreements with partner companies that allow for revenue sharing of downstream products.

So what can we expect from the pace of development for enabling technologies in synthetic biology? Reduction in the cost of synthetic gene products is widely considered to be one of the most important enabling technologies in synthetic biology. Tenuous estimates for the market for such synthetic gene products is $100 million/year in 2009[6,7] with current costs for synthetic genes from ∼$0.30−$0.80/base. A 1−2 orders of magnitude reduction in costs would allow gene synthesis to directly replace slow genetic engineering techniques and also allow for more rapid and extensive testing of synthetic designs. Assuming further investment in new technologies could reduce costs to such levels, how would a potential company recoup their investment? They could keep the prices relatively stable and obtain larger profits through lower production costs, albeit from this relatively small market. They could drop the price of gene synthesis and potentially capture the entire market (now $1−$10 million/year). While demand might expand and biology research would be positively impacted, there are no clear broad-based consumer markets akin to sequencing or microprocessor industries that can increase demand at the scale required to recoup such price reductions. Alternatively, they can leverage their ability to deliver large numbers of genetic constructs to develop more valuable engineered organisms either in-house (vertically integrated) or in partnerships (revenue sharing) that target much larger consumer markets. This approach is probably more lucrative and therefore more likely to attract necessary investment than trying to conquer the limited and now commoditized gene synthesis market. Unfortunately, in such cases, access to the technology must be limited in order to provide competitive advantage.

Thus, the dramatic technology developments in next-generation sequencing and its effects on the biological research community are going to be hard to replicate for enabling technologies for the synthetic biology research community. This, unfortunately, will likely remain the status quo until the enabling synthetic biology technologies themselves become products sold directly to ∼7 billion consumers.

AUTHOR INFORMATION

Corresponding Author

*E-mail: sri.kosuri@wyss.harvard.edu; msismour@genetics.med.harvard.edu.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Oxford Nanopore (**2012**) Oxford Nanopore introduces DNA 'strand sequencing' on the high-throughput GridION platform and presents MinION, a sequencer the size of a USB memory stick, http://www.nanoporetech.com/news/press-releases/view/39, accessed 2/11/2012.

(2) Wetterstrand K. A. DNA sequencing costs: Data from the NHGRI large-scale genome sequencing program, available at www.genome.gov/sequencingcosts, accessed 2/11/2012.

(3) Clarke, J., Wu, H.-C., Jayasinghe, L., Patel, A., Reid, S., and Bayley, H. (2009) Continuous base identification for single-molecule nanopore DNA sequencing. *Nat. Nanotechnol. 4*, 265−270.

(4) Derrington, I. M., Butler, T. Z., Collins, M. D., Manrao, E., Pavlenok, M., Niederweis, M., and Gundlach, J. H. (2010) Nanopore DNA sequencing with MspA. *Proc. Natl. Acad. Sci. U.S.A. 107*, 16060−16065.

(5) Lieberman, K. R., Cherf, G. M., Doody, M. J., Olasagasti, F., Kolodji, Y., and Akeson, M. (2010) Processive replication of single DNA molecules in a nanopore catalyzed by phi29 DNA polymerase. *J. Am. Chem. Soc. 132*, 17961−17972.

(6) Carlson, R. (2009) The changing economics of DNA synthesis. *Nat. Biotechnol. 27*, 1091−1094.

(7) Newcomb, J., Carlson, R., and Aldrich, S. C. (2007) *Genome Synthesis and Design Futures: Implications for the U.S. Economy*, Bio Economic Research Associates, Cambridge, MA.