Guest Editorial

BRJ
Biomedical Research Journal

Next Generation Sequencing in Healthcare

Tania Fernandez

Founder and CEO, DreamCatcher Ventures, 1288 Columbus Avenue, #133, San Francisco, CA 94133, USA

The year 2013 was an eventful year witnessing revolutionary discoveries in the world of extraordinary medical advances and healthcare technology. We saw a spate of new and promising discoveries ranging from detecting lung cancer with a cough, pancreatic cancer with accurate faster, cheaper paper diagnostics, to the possibility of using the Human Immunodeficiency Virus to treat genetic disorders in children (Radcliffe, 2013). Emerging newer technologies have fuelled the momentum of the 'genomic revolution'. The completion of sequencing of the human genome project in 2003 (National Human Genome Research Institute, 2010), translated into the rise of the 'omics' era creating mega scientific data. It also gave rise to a breed of genomic companies that focused on application of the emerging technologies, particularly in medical science.

Deciphering the Genetic Code: DNA Sequencing Update

The ability to sequence DNA represented a breakthrough milestone in DNA research. The Sanger sequencing method, developed by the Nobel laureate Frederick Sanger, became the most widely used DNA sequencing method. Sequencing is a way of 'reading' DNA molecules, two complementary strands coiled together to form the double helix. The entire human genome contains about 3.1 billion molecular base pairs per set of chromosomes in a cell.

The story of genome sequencing was not a single 'eureka' moment characterizing the Archimedes discovery, but a compelling story of unbridled passion and continuous advancements pushing the frontiers of genomic technology into interdisciplinary amalgamation of science and technology including advanced materials, nanotechnology, biology, chemistry, enzymology, modeling, and mega data understanding. The story of genome sequencing is one of war and price. The Human Genome Project was one of the costliest 'contests' ever held, a multibillion-dollar government-led effort in the Reagan era, to sequence the entire human genome

Key words: DNA sequencing, human genome project, genomics research.
***Corresponding Author:** Tania Fernandez, Founder and CEO, DreamCatcher Ventures, 1288 Columbus Avenue, #133, San Francisco, USA.
Email: tania.fernandez@dreamcatcher.ventures

which was nearly beaten to the punch by a private company called Celera. The race to sequence the genome became the subject of The Genome War (Barbujani, 2004).

In 2001, the cost to sequence an entire human genome was USD 100 million (National Human Genome Research Institute, 2014). Since then, the cost has moved swiftly downwards and indicated USD 1 million around 2007 when the genome of Nobel laureate Professor James Watson was sequenced. The price has continued on its downward curve, falling to about USD 3,000–5,000 in 2013, although specialized sequencing for cancer patients often costs more. While researchers, companies and investors still argue and agonize on costs of sequencing the human genome, there is no disputing that the pricing has rapidly plummeted downwards.

Genome analysis is one of the fastest-emerging fields in the world, with the recent pricing close to passing USD 1,000, a milestone (Herper, 2014), and continuing to decelerate. The race is on to reach a price of USD 100 per complete genome. Cutting-edge 'next-generation sequencing', allows a greater throughput by parallelizing the sequencing process and producing millions of sequences concurrently. So-called third generation sequencing methods have since supplemented the second generation sequencing methods enabling a greater throughput while at the same time reducing the time to result and costs. Third generation sequencing involves real-time sequencing of single DNA molecules without needing to amplify DNA using PCR. Chip-based sequencing eliminates the need for expensive reagents and uses relatively inexpensive equipment, further lowering costs significantly, with increasing sequencing throughput speeds. Current gene sequencing technologies frequently require working with short snippets of DNA. These must be processed by large sequencers in a laboratory, and may take days to completion. 'Nanopore technology', the revolutionary advancement that the world stays tuned to will accelerate the genomic revolution. The excitement of nanopore DNA sequencing is in creation of 'tricorder-like' devices for detecting pathogens or diagnosing genetic disorders rapidly and on-the-spot, and may result in 'Point-of-Care' diagnostics for patients.

Besides, the human genome, one of the most compelling genetic mapping project is unraveling the genetic code of various cancers. Sequencing of cancer genomes allows scientists and doctors to discover gene mutations that contribute to cancer, potentially leading to better detection methods and treatments. The diagnosis and personalized treatment of cancer patients play a major role because the underlying cancer-causing mutations can vary greatly between tumors of different tissue and cell types, and between individuals with the same tumor. These genetic characteristics may have a huge

impact on the efficacy of anti-cancer drugs. Information gleaned through whole-genome sequencing has reclassified and stratified cancers based on the genetic makeup rather than only TNM classification of cancers and the location in the body. The system has resulted in a paradigm shift in cancer therapy for oncologists, and concurrent evaluation of the potential benefits of personalized cancer therapy. Besides, cost-effective approaches to whole exome (coding regions of the genome) sequencing available since 2009 (Maher, 2009), has been useful in predisposition studies in several cancer types (Jones *et al*., 2009) leading to 'Predictive Diagnosis' indicating critical importance of individual genomic constitution as a high risk factor.

'Deep Digital Sequencing' developed at The Genome Institute, Washington, USA, is commonly used to examine mutations in patients' tumor tissue samples, repeated 1000 times or more, generating frequency of the mutation. The data indicates evolution of cancer cells and molecular pathology of progression of the cancer (Maher, 2012). As cancer evolves, tumors acquire new mutations retaining the original cluster of mutations resulting in converting the normal cell to a malignant cell. The authors suggested that drugs targeted to genetic changes that occur early in the course of cancer may be more effective. On the other hand, drugs targeted to mutations observed exclusively in later-evolving cancer cells, may not have much effect on the disease and may not kill all the tumor cells. Sequencing is revolutionizing medical science and has the potential to serve as a powerful and cost-effective diagnostic tool in the management of cancer.

The Genomics Opportunity

Next-generation sequencing technologies have been and continue to be deployed in clinical laboratories, enabling rapid 'Bench to Bedside' transformations in 'Molecular Medicine'. As a reference point it is worthwhile to remember that the first complete cancer genome sequenced was that of acute myeloid leukemia (AML) cells, a severe form of cancer that initiates in the bone marrow. The AML genome was sequenced by creating single-read libraries from several micrograms of DNA, as replicate library-preps. Each of these was sequenced on 98 runs to generate 6 billion single-end 32bp sequencing reads. At 3 days per lane (a year on a Genome Analyser II, Illumina) and approximately USD 640 per lane in consumables, culminated in a final cost of USD 500,000. Today the same genome can be analyzed using 500 ng of DNA in a PCR-free library prep and run on one HiSeq X Ten (Illumina) to generate 375M PE1 50bp reads in 3 days for USD 1000. The masked and hidden costs will need to be added. However, the trend is clear, with whole genome sequencing 100 times faster and 500 times cheaper, thus making it useful for clinical

analysis (Core Genomics, 2014).

The advances in software development for molecular sequence analysis makes it feasible to analyze the vast terabytes of data generated by sequencing the genome, as 'big data' continue to aid the genomics revolution. The science has resulted in several startups flourishing with a torrent of venture capital dollars poured into the powerful 'genome interpretation' and 'data analytical' space. This in turn has led to understanding the clinical significance of genomic data to doctors and patients, thus affecting the most valuable stakeholder in the cycle, the patient.

The DNA sequencing market has expanded consistently to 18% a year and is expected to reach nearly USD 7 billion in 2016 (BCC Research, 2012) with cost price making it affordable, available and accessible to the ultimate consumers.

Genomics Research: Controversies

Study reported by Roberts *et al*. (2012) on 53,666 identical twins in cancer registries from the United States, Sweden, Finland, Denmark and Norway, to gauge the predictive capacity of personal genome sequencing clearly emphasizes that prediction will remain probabilistic and not deterministic as behavior, environment and random events may often tip the ability to be predictive with certainty (Kolata, 2012). Thus, today sequencing is frequently used to better understand the mechanistic aspects of diseases and preempt better therapies.

The need of the hour is to build databases of known disease-causing genetic mutations, robust sequencing and interpretation methodologies to validate cause-and-effect relationships between genes, behavior, environment and disease. Repeat and reproducible mutation data post sequencing large and varied populations will be reflected in routine clinical applications of the robust genomic information.

Promising Better Health

The promise of personal genomics is here to stay, with a major role in better health with 'Personalized Medicine', risk information for several diseases including cancers, diabetes and heart diseases; potential of individuals to metabolize drugs and the need for drugs to be personalized, not only for patient treatment but for identification of variants in carrier status in several diseases. The testing should be available and affordable to all. The personal genomic data will make it feasible for each individual to actively participate in shaping their health profiles and bettering health score cards. The promise of personalized medicine will have to be weighed against the challenges posed by the technological, financial and ethical limitations.

It is a game with many stakeholders: scientists, entrepreneurs, doctors, policy makers, investors and the most important stakeholder – the patient. We must tread with caution and rule with ethics.

REFERENCES

Barbujani G. The genome war: How Craig Venter tried to capture the code of life and save the world. *N Engl J Med* 2004;351:304.

BCC Research. DNA sequencing: Emerging technologies and applications. March 2012. Available at: http://www.bccresearch.com/market-research/biotechnology/dna-sequencing-technologies-applications-bio045d.html (accessed on 31st July 2014).

Core Genomics. BlogSpot. June 2014. Available at http://core-genomics.blogspot.in (accessed on 31st July 2014).

Herper M. The $1,000 genome arrives – for real, this time. *Forbes.com* 2014, January 14. Available at: http://www.forbes.com/sites/matthewherper/2014/01/14/the-1000-genome-arrives-for-real-this-time/ (accessed on 31st July 2014).

Jones S, Hruban RH, Kamiyama M, Borges M, Zhang X, Parsons DW, *et al*. Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. *Science* 2009;324:217.

Kolata G. Study says DNA's power to predict illness is limited. *New York Times* 2012, April 2 (accessed on 31st July 2014).

Maher B. Exome sequencing takes centre stage in cancer profiling. *Nature* 2009;459:146147.

Maher B. A bend in the river for cancer genomics. *Nature News Blog* 2012, April 4. Available at: http:///blogs.nature.com/news/2012/04/a-bend-in-the-river-for-cancer-genomics-at-aacr.html (accessed on 31st July 2014).

National Human Genome Research Institute. The human genome project completion. *National Institutes of Health* 2010. Available at: http://www.genome.gov/11006943 (accessed on 31st July 2014).

National Human Genome Research Institute. DNA sequencing costs. *National Institutes of Health* 2014. Available at: http://www.genome.gov/sequencingcosts (accessed on 31st July 2014).

Radcliffe S. 10 top medical and technological innovations of 2013. *Healthline News* 2013, December 5, 2013 (accessed on 31st July 2014).

Roberts NJ, Vogelstein JT, Parmigiani G, Kinzler KW, Vogelstein B, Velculescu VE. The predictive capacity of personal genome sequencing. *Sci Tranl Med* 2012;4:133ra58.