

GENOMICS IN AUSTRALIA

Andrew Gilbert*

General Manager, Bioplatforms Australia Limited, Macquarie University, NSW 2109

***Corresponding author: agilbert@bioplatforms.com**

Introduction

It is widely believed that the 21st century is the age of biology. While combustion, electricity and nuclear power defined scientific advance in the last century (1), genome research marks the advent of a new era where biotechnology is shaping our lives and economy as fundamentally as electricity did in the 20th century.

Rapid advancements in genomics technology have given scientists revolutionary abilities in comprehensively analysing biochemical pathways, cells and organisms at the molecular level. In the first instance, Sanger sequencing provided a fundamental tool for the routine analysis of individual genes, chromosomes and, ultimately, full genomes. Driven by the goal of deciphering the human genome, automation and process parallelisation allowed scientists to generate unprecedented levels of data and reach new frontiers in biological discovery. The Human Genome Project epitomises these advances, however, the enormous cost of large-scale sequencing initiatives meant such projects were prohibitive for most laboratories.

From 2005, the introduction of next-generation platforms is making genomics technology more accessible. Three principal platforms released by Roche, Illumina and Applied Biosystems all allow nucleotide sequencing with a throughput rate more than 2000 times that achieved by 'traditional' Sanger technology, with corresponding increases in quality and simultaneous decreases in cost (2). As a result, complete genomes are routinely sequenced and analysed. Indeed, there are abundant examples of individual organisms being sequenced as well as complete families analysed at the genome-wide level. The technology is even tackling ancient DNA, with sequencing undertaken on the nuclear genome of a 28,000-year-old mammoth and Neanderthal DNA (3).

Current Genomics Infrastructure Framework in Australia

Bioplatforms Australia was established in 2007 to ensure strategic management of Commonwealth funds for life science research infrastructure. With the aim of maximising accessibility, promoting collaboration and driving integration, Australia's genomic capabilities have been integrated into a forward-looking network of genomics providers.

Genomics Australia, the new consortium that now defines Australia's genomic capabilities, combines the pre-existing Australian Genome Research Facility (AGRF) with more specialised transcriptomics and epigenetics laboratories. This framework ensures a more holistic genomics sector that can better plan and prioritise strategic infrastructure investments. It also allows more effective support from

Bioplatforms Australia in providing industry-wide advocacy and pursuing initiatives that benefit the sector as a whole. One such example involves the specific engagement of the National Health and Medical Research Council (NHMRC), Commonwealth Scientific and Industrial Research Organisation (CSIRO) and Australian Research Council (ARC). These leading research organisations present many opportunities for leveraging the nation's genomics capacity in a way that supports and promotes meritorious research.

High-Throughput Sequencing Capability

For more than a decade, the AGRF served Australian researchers with leading-edge capabilities in genome sequencing. It made some of the first acquisitions of next-generation sequencing technology in Australia when it commissioned both Roche GS FLX (454) and Illumina GAI (Solexa) high-throughput systems. The Applied Biosystems SOLiD system was also added to its arsenal to ensure the breadth of scientific need was supported. AGRF's technology suite now offers more rapid and comprehensive tools for genomic analysis than ever before available in Australia.

Complementary Technology

Sanger and next-generation sequencing capabilities are complemented by substantial Australian capacity in transcriptomics and gene expression. The Ramaciotti Centre for Gene Function Analysis (Professor Ian Dawes) at the University of New South Wales and the Expression Genomics Laboratory (Professor Sean Grimmond) at the Institute for Molecular Bioscience, University of Queensland, are both members of Genomics Australia and prominent laboratories in these fields. Both undertake leading-edge research programs and provide technology services to the broader scientific community.

Epigenomics, the genome-wide analysis of epigenetic events such as DNA methylation and chromatin remodelling affecting gene expression and control, offers another facet of investigation and discovery. Professor Frances Shannon (University of Canberra) and Professor Liz Dennis (CSIRO Division of Plant Industry), through the Australian Cancer Research Foundation Biomolecular Research Facility at the Australian National University and in collaboration with the Centre for Comparative Genomics at Murdoch University, have developed analytical methodologies and bioinformatic pipelines enabling detailed investigation of epigenetic events within cell development. These advances do much to enhance discovery efforts and have been deployed to the bioscience community through new services offered by the AGRF.

Bioinformatics – Opportunities with Challenges

Next-generation sequencing is only possible because of modern computing power allowing compilation of sequence data, mining and analysis, together with advanced modelling.

Bioinformatics continues to pose considerable challenges in the Australian setting. A lack of capacity and a highly dispersed workforce has meant Australia has not made the same footholds as the US and Canada in establishing bioinformatics as a discrete science.

During a series of workshops directed at establishing a bioinformatics strategy for Australian Bioscience, EMBL Australia invited the Director of the UK European Bioinformatics Institute, Professor Ewan Birney, to contribute to an analysis of Australian bioinformatics capabilities (4).

Professor Birney noted that most major genomics projects are 'analysis bottlenecked'. That is, too little investment occurs in analytical capability relative to data generation, especially since the explosion of data generated by next-generation sequencing. In Europe, bioinformatics investment in most genomics grants equates to 50-60% of the funds and reaches 80% in some instances. This is very different to Australia, where bioinformatics grant proposals often 'miss the target' with grant review panels, leaving this 'analysis bottleneck' largely unresolved.

Bioplatforms Australia investments are also directed to developing high-level bioinformatics capabilities in Australia and, again, is acting to promote better collaboration and integration of bioinformatics among 'omics research in Australia.

Australian Genomics Projects

To strengthen Australia's record in genome science and improve our scientific competitiveness in these areas, genomic projects are now receiving new priority with some exciting research efforts under way.

Cancer Research

In 2008, Australia made a substantial commitment to international efforts to understand the genetic causes of cancer with its involvement in the International Cancer Genome Consortium (5). NHMRC committed \$27.5 million, with the University of Queensland, NSW Cancer Council, Silicon Graphics and Applied Biosystems boosting funds to over \$40 million.

Australia now has a team of genomics and cancer experts providing full genomic analysis of pancreatic and ovarian cancer. Australia's commitment to the consortium represents a major stride in Australia's global genomics engagement.

Australia's First Genome Project

An ambitious project to sequence the entire genome of *Acropora millepora*, a branching coral found in Great Barrier Reef, will be the first animal genome to be fully sequenced and assembled in Australia. This staghorn coral has a surprisingly large genome with around



20,000 genes – similar to humans. It is a monumental project because of its potential to address the impact of environment stress on Australia's coral reef system, but also because the evolutionary significance of corals may contribute to the idea of ancestral genetic complexity (6) and reveal significant insights into the molecular basis of many human genetic disorders.

The project is a joint undertaking between the AGRF and researchers from the ARC Centre of Excellence for Coral Reef Studies (CoECRS) and represents another example of Australia's growing genomics capacity and expertise.

Wine – Extending Genomics to the Systems Biology Frontier

The Australian Wine Research Institute (AWRI), based in South Australia, has also instigated a pioneering project. An interdisciplinary team led by Paul Chambers, Research Manager for Biosciences at the AWRI, is seeking to gain new insights into a winemaking yeast, *Saccharomyces cerevisiae*. The project will also offer a valuable demonstration of a systems biology approach – a new view of biology that does not investigate individual genes or proteins, but rather, the behaviour and relationships of all the elements of a particular biological system (7).

AWRI's wine yeast project involves collecting and integrating data sets generated from 'omics investigations of wine yeast in a model fermentation. The team is seeking to more precisely engineer the role of yeast in the winemaking process within the complexity of a whole system. Besides new and valuable knowledge for commercial winemaking and other yeast-dependent industries, the work is pioneering, given that it breaks with the tradition of a reductionist approach. Furthermore, it will offer practical experience in tackling the significant challenge of integrating mountains of 'omics data to create useful models of predictive benefit.

Future Genomics Projects

Professor Lee Hood states "the foundation of systems biology is thinking of biology as an informational science. There are two types of biological information: the digital information embedded in the genome and the environmental information that impinges on it" (8).

The limited availability of high-quality quantitative data remains a major bottleneck for mathematical modelling in biology and medicine. Generating such data (e.g., genomes, quantitative proteomics and metabolomics) is more costly and time-consuming than conventional experiments, making it virtually impossible for small research teams (9).

In a new genomics initiative, Bioplatforms Australia will add to Australia's scientific wealth by purchasing dormant instrument time and using it to create 'omic datasets of organisms and specimens of scientific and commercial significance. Specimens will be chosen in light of national priorities for innovation and take account of market demand from meritorious researchers and commercial research efforts. Wheat, barley, cancer types and unique Australian fauna and flora are likely candidates for the in-depth 'omic analysis proposed. The resulting datasets will then be made accessible to the research community

for any number of research questions across all scientific disciplines.

This strategy will both leverage and expand the capabilities of Australian 'omics experts. In particular, it should help to address the bottleneck in analytical bioinformatics highlighted by Professor Birney. Access to large datasets, with funding for bioinformaticians and the infrastructure required for advanced computational biology, will boost Australia's bioinformatics sector and such development can be built around strategic projects. This capability will then offer ongoing benefit to the broader Australian genomics community.

Summary

Revolutions in genomics technology have been transforming biology and it is clear that genomics has not yet exhausted its potential. Australia is picking up its pace in developing a strong genomics sector that will continue to advance as its infrastructure base grows and our bioinformatics sector matures.

References

1. Venter, C., and Cohen, D. (2004) *New Perspectives Quarterly* **Fall**, 73-77
2. Ideker, I., Galitski, T., and Hood, L. (2001) *Annu. Rev. Genomics Hum. Genet.* **2**, 343-372
3. Schuster, S. (2008) *Nature Methods* **5**, 16-18
4. EMBL Australia, Bioinformatics for the Future, March 2009, http://www.emblaustralia.org/site/DefaultSite/filesystem/documents/EMBL_Bioinformatics_Final.pdf
5. International Cancer Genome Consortium <http://www.icgc.org/>
6. Dean, T., (2009) *Australian Life Scientist* **November/December**, 38-41
7. Dean, T., (2009) *Australian Life Scientist* **November/December**, 9-12
8. Professor Lee Hood, Director, Institute for Systems Biology, Seattle, USA. Quoted from a symposium at the Plant and Animal Genomes Conference, San Diego, 13 January 2009
9. European Science Foundation (2008) *Advancing Systems Biology for Medical Applications* **December**, 7

