

Innovation in clinical bioinformatics

Background

Genomic sequencing technologies are constantly evolving, and new technologies and bioinformatic methods regularly appear on the horizon.

Bioinformatics is the collection, classification, storage and analysis of biochemical and biological information using computers, including genetic or genomic codes. Clinical bioinformatics plays an important role in deciding a patient's genomic test result.

To enable patients to access to the best, fit-for-purpose genomic sequencing methods, diagnostic and bioinformatics experts must gain knowledge and experience in preparing for new horizon technologies.

Project summary

The objective: to apply state-of-the-art diagnostic approaches in clinical care and to support a common system for the management of clinical genomic data.

This project sought to support clinical bioinformatics activities that fostered collaboration, in order to share knowledge and expertise across the Melbourne Genomics members.

A shared bioinformatician role was created in 2018 to address important implementation opportunities – including the new human genome reference¹ in GenoVic² – and work across the Alliance for the benefit of all members.

As part of the shared bioinformatics initiative, four projects were completed:

- Evaluation of the new human reference genome (GRCh38)
- Investigation into approaches for validating bioinformatic analyses
- Development of portable pipeline guidelines and framework
- Evaluation, prioritisation and selection of a longitudinal quality control tool

The Melbourne Genomics Health Alliance members involved were: the Australian Genome Research Facility, the Peter MacCallum Cancer Centre, Monash Health, Murdoch Children's Research Institute (Victorian Clinical Genetics Services), The Royal Melbourne Hospital and Austin Health.

Impact

The new reference genome marked an important development which, when implemented into diagnostic tools, provides even more comprehensive analysis of the many thousands of gene changes in a person's genome.

Evaluation of the new human genome reference led to an earlier-than-anticipated implementation within GenoVic and incorporation into accredited testing in 2019, positioning Melbourne Genomics as a global leader in using the new reference in clinical genomic testing.

¹ The genome against which patients' genomic sequencing data is compared in the search for changes which may be disease-causing.

² See project summaries for GenoVic, the clinical genomics system built by and for the Melbourne Genomics members.

The Shared Bioinformatics Project led to decreased duplication of effort in the evaluation of the new reference genome and quality control tools. Two Alliance laboratories become early adopters of the new reference genome, improving the accuracy of genomic tests delivered in Victoria in general.

Increased knowledge-exchange led to an increase in clinical bioinformatic capabilities: in executing pipelines in the cloud and in potential implementation of large-scale longitudinal quality control processes.

This project also resulted in increased participation by Victoria's laboratories in international efforts in clinical bioinformatics. A Hackathon at the Broad Institute of Harvard and MIT – in which Melbourne Genomics participated – resolved critical failures and optimised WDL language³ for AWS⁴ cloud compatibility.

Lessons learnt

- Bioinformatic projects require resource commitment from member organisations, to ensure maximum benefit and to accommodate each laboratory environment.
- Project selection and scoping requires careful consideration to ensure timely completion and achievement of objectives.
- Participating laboratories favour open source products for bioinformatics solutions when there are no requirements for a complex user interface, integration, security or auditing. However, open source tools are often designed and tested in the research setting, and significant resourcing and development may be needed to meet diagnostic/clinical requirements.
- Laboratories require support to overcome challenges involved in moving away from legacy bioinformatics solutions, i.e. from local computing infrastructure to cloud computing. This includes development of augmented costing models and significant training that may not be readily available.

³ Workflow Definition Language.

⁴ Amazon Web Services, which is utilised by GenoVic. GenoVic's IT infrastructure is cloud-based, which ensures high availability and scalability. (See GenoVic project summaries.)