Using Analytics to Improve Cancer Diagnosis and Therapy Selection

Developing and automating best-practice workflows that make analyzing, processing, and disseminating genomic data accessible to researchers and clinicians.

Introduction

The mission of the University of Melbourne Centre for Cancer Research (UMCCR) is to "drive innovation and implementation for clinical impact in cancer care." Led by Sean Grimmond, the UMCCR is working to accelerate improved outcomes for cancer patients through the use of genomic information in understanding cancer development and targeting therapies. The UMCCR collaborates with Victorian Comprehensive Cancer Centre (VCCC) Alliance partners to enable personalized cancer care programs across institutions.

A key part of this endeavor is ensuring that results from next-generation sequencing (NGS) data processing, analysis, and interpretation are broadly available to researchers and clinicians. This is the goal of the Genomics Platform Group at UMCCR, to improve "the scalability and reliability of sequencing workflows, better detect changes in cancer genomes, and make tumor data accessible in real time." To help achieve this objective, Dr. Oliver Hofmann, Associate Professor and Head of Bioinformatics, Clinical Pathology at UMCCR and the leader of the Genomics Platform Group, partnered with Illumina to gain early access to Illumina Connected Analytics.

The Search for Scalable Bioinformatics

When Dr. Hofmann and team began working with Illumina, they had a pipeline for high-throughput NGS data analysis in place. This solution reproducibly processed short-read sequencing data, but the team wanted to increase scalability by migrating from traditional high-performance computing (HPC) to a cloud-based platform. They started looking for alternative solutions.

Dr. Hofmann had a specific list of requirements for the new pipeline. It had to be an open-source solution, cloud-native, and easily accessed by the labs that would ultimately use the provided workflows to analyze and interpret data on their own. Long-term support would be essential for success, so a commercial product, rather than an academic one, seemed more viable. It needed to have fast bioinformatics to enable labs to go from genome to report in < 7 days. And, as a contributor to the Global Alliance for Genomics and Health (GA4GH), Dr. Hofmann wanted the solution to be responsive to GA4GH standards.

Finding Illumina Connected Analytics

The Genomics Platform Group began testing solutions to find one that would provide the needed support, speed, and reliability. When Dr. Hofmann spoke with the team at Illumina, he learned about a promising project in development, called Illumina Connected Analytics, that would meet his requirements.

Illumina Connected Analytics is a cloud-based system for storing, managing, and analyzing genomic data in a single environment. It is built on top of Amazon Web Services (AWS) to create a comprehensive data management pipeline for use globally, in a secure environment. Using Illumina Connected Analytics, users can access any public workflow and easily share their own workflows with others,

The Genomics Platform Group at UMCCR strives to improve the usability of genome sequencing data to improve cancer detection and treatment.

simplifying collaboration across labs. Because Illumina architected the platform to allow for robustness and extensibility, there was sufficient flexibility in the platform to allow for modifications to meet the team's unique needs. In short, it sounded exactly like what the Genomics Platform Group was seeking.

"Moving everything to one environment minimizes the processes that we have to vet for each new workflow and simplifies the accreditation process, which, in turn, supports setting up a clinical sequencing environment," states Dr. Hofmann. "It streamlines processes and enables us to leverage work that has already been done, so we are not constantly redoing the same proof work."

In addition, Illumina Connected Analytics offers the power of the DRAGEN™ Bio-IT Platform for accurate, ultrarapid secondary analysis of sequencing data. Using the DRAGEN pipelines, analysis can be completed in 6-7 hours. This speed is crucial; to turn around a curated report in 7 days, analysis needs to be completed within 48 hours. Data generated on a NovaSeq™ 6000 System is uploaded directly to Illumina Connected Analytics, where it can be analyzed directly after the sequencing run is complete. There are no intermediate data staging sites, speeding up the whole process.

Collaboration Produces Results

For the past several months, Dr. Hofmann and the Genomics Platform Group at UMCCR have been working with the Illumina team to optimize their workflows for Illumina Connected Analytics. Illumina is committed to supporting GA4GH standards, allowing the team to migrate existing workflows to the platform with minimal effort. The platform has provided the team with a commercially backed stable production environment, while allowing them to maintain control over deployed workflows. Dr. Hofmann describes this collaborative effort as "Terrific. I've worked on many academic/industry partnerships, but this is easily one of the best in terms of communication, hands-on support, and friendliness. Illumina has been responsive throughout all steps of the process, providing roadmaps, insights into design decisions and actual implementation support."

Going forward, the Genomics Platform Group will continue to work with GA4GH, ensuring that Illumina Connected Analytics is compliant with community-developed standards and allowing other researchers to run analysis on UMCCR data. Ultimately, the team wants to make the UMCCR-developed cancer analysis workflows available as turnkey solutions to the scientific community. The goal is to provide access to the information needed to understand what is driving a particular cancer and identify what that means for treatment.

Learn more

Illumina Connected Analytics, www.illumina.com/ConnectedAnalytics

Summary

Overview

The Genomics Platform Group at UMCCR strives to improve the scalability and reliability of sequencing workflows, enable better detection of changes in cancer genomes, and make tumor data accessible in real time.

Challenge

The Group needed a scalable, commercially supported, cloud-based environment for running and sharing workflows for analyzing and reporting genomics data.

Solution

The team chose to migrate current workflows and systems to Illumina Connected Analytics.

Benefits

Using Illumina Connected Analytics, the Genomics Platform Group has been able to digitally transform their sequencing workflows by optimizing and standardizing their analytic pipelines, streamlining the timeline from data output to curated report, and sharing workflows and processes across national and global collaborators all within a secure and compliant data environment.

> "We see immense promise in the simpler workflow management provided by Illumina Connected Analytics, all in one place instead of a hybrid HPC/cloud solution. The gains in turnaround time are impressive, and we look forward to moving our data portal and monitoring to Illumina Connected Analytics."

Illumina • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com
 Pub. No. 970-2020-010-A QB10596

illumina®