GenomeNext achieves unprecedented throughput of 1,000 genomes analyzed per day

March 18 2015

GenomeNext, LLC, a leader in genomic data management and integrated analysis, announced today that, through the "Intel Heads In The Clouds Challenge on Amazon Web Services (AWS)" with support from JHC Technology, and in conjunction with Nationwide Children's Hospital, has benchmarked whole genome sequencing analysis at an unprecedented 1,000 genomes per day.

GenomeNext and Nationwide Children's were challenged to perform analysis of the complete 1000 Genomes Consortium's sequencing dataset in one week. This dataset consists of raw genomic sequence data from 2,504 individuals sampled across 26 different populations.

All 5,008 samples (2,504 whole genomes & 2,504 exomes) were analyzed on GenomeNext's Platform, leveraging its proprietary genomic sequence analysis technology developed by Nationwide Children's, operating on the AWS Cloud powered by Intel Xeon processors.

The entire analysis was completed in one week, with as many as 1,000 genome samples being completed per day, generating close to 100TB of processed result files. The results of the analysis will be published in the coming months, and Dr. Peter White, director of Nationwide Children's Biomedical Genomics Core, and his team found there was a high-degree of correlation with the original analysis performed by the Consortium, with additional variants potentially discovered during the analysis

performed utilizing GenomeNext's Platform.

"The successful completion of this proof-of-concept not only sets a groundbreaking timeframe for the analysis of a massive quantity of genomic data, but demonstrates the utility of the GenomeNext solution, eliminating the sequence analysis computational bottlenecks, enabling researchers and clinicians to keep pace with processing the magnitude of genomic data analysis required for population-scale genomics" said James Hirmas, CEO of GenomeNext.

GenomeNext believes this is the fastest, most accurate and reproducible analysis of a dataset of this magnitude. The goal is to integrate the new knowledge into GenomeNext's analysis pipeline to assist researchers and clinicians in identifying rare pathogenic variants. Access to a growing and diverse repository of DNA sequence data, including the ability to integrate and analyze the data is critical to accelerating the promise of precision medicine. "We believe this dataset will become invaluable to genomic medicine as a whole and further catalyze the identification of disease causing genetic variants," Hirmas said.

"At Nationwide Children's we have a strategic goal to introduce genomic medicine into multiple domains of pediatric research and healthcare," Dr. White said. "Rapid diagnosis of genetic disease can be critical in newborns, so our initial focus was to create an analysis pipeline that was extremely fast, but didn't sacrifice clinical diagnostic standards of reproducibility and accuracy. Having achieved that, we discovered that a secondary benefit of this analysis pipeline was that it could be adapted for population scale genomic analysis."

"Intel is proud of our work with GenomeNext and AWS to deliver these remarkable performance results with Nationwide Children's Hospital," said Ketan Paranjape, General Manager Life Sciences at Intel. "As we look to bringing precision medicine mainstream we will need to continue

to work on solutions based on a converged infrastructure of compute, storage and networking products as well as optimized software packages that will enable population scale genomic analysis."

"Analysis and bioinformatics pipelines need to be able to keep up with the new genomic data being generated as sequencing moves closer to clinical practice," said Matt Wood, General Manager, Data Science at Amazon Web Services, Inc. "It's great to see that important projects such as this are able to take advantage of the scale and flexibility of the cloud, advancing and improving the delivery of care for patients worldwide."

An article describing GenomeNext's novel parallelization strategy for human genome analysis was recently published in *Genome Biology* (genomebiology.com/2015/16/1/6/). Through implementation of this peer reviewed analysis method, the GenomeNext platform fully automates the analytical process, enabling an efficient, accurate and deterministic execution of all steps in the entire human genome analysis workflow.

The Software as a Service (SaaS) is offered through the company's website, www.genomenext.com/ . Alternative configurations, including private cloud and custom pipelines are considered upon request.

More information: "Churchill: an ultra-fast, deterministic, highly scalable and balanced parallelization strategy for the discovery of human genetic variation in clinical and population-scale genomics." *Genome Biology* 2015, 16:6 DOI: 10.1186/s13059-014-0577-x

Provided by Nationwide Children's Hospital