

Cloud computing method greatly increases gene analysis

September 8 2010

Researchers at the Johns Hopkins Bloomberg School of Public Health have developed new software that greatly improves the speed at which scientists can analyze RNA sequencing data. RNA sequencing is used to compare differences in gene expression to identify those genes that switched on or off when, for instance, a particular disease is present. However, sequencing instruments can produce billions of sequences per day, which can be time-consuming and costly to analyze.

The software, known as Myrna, uses "[cloud computing](#)," an Internet-based method of sharing computer resources. Faster, cost-effective analysis of [gene expression](#) could be a valuable tool in understanding the genetic causes of disease. The findings are published in the current edition of the journal [Genome Biology](#). The Myrna software is available for free download at <http://bowtie-bio.sf.net/myrna>.

Cloud computing bundles together the processing power of the individual computers using the Internet. A number of firms with large computing centers including, Amazon and Microsoft, rent unused computers over the Internet for a fee.

"Cloud computing makes economic sense because cloud vendors are very efficient at running and maintaining huge collections of computers. Researchers struggling to keep pace with their sequencing instruments can use the cloud to scale up their analyses while avoiding the headaches associated with building and running their own computer center," said lead author, Ben Langmead, a research associate in the Bloomberg

School's Department of Biostatistics. "With Myrna, we tried to make it easy for researchers doing RNA sequencing to reap these benefits."

To test Myrna, Langmead and colleagues Kasper Hansen, PhD, a postdoctoral fellow, and Jeffrey T. Leek, PhD, senior author of the study and assistant professor in the Department of Biostatistics, used the software to process a large collection of publicly available RNA sequencing data. Processing time and storage space were rented from Amazon Web Services. According to the study, Myrna calculated differential expression from 1.1 billion RNA sequencing reads in less than 2 hours at cost of about \$66.

"Biological data in many experiments—from brain images to genomic sequences—can now be generated so quickly that it often takes many computers working simultaneously to perform statistical analyses," said Leeks. "The cloud computing approach we developed for Myrna is one way that statisticians can quickly build different models to find the relevant patterns in sequencing data and connect them to different diseases. Although Myrna is designed to analyze next-generation sequencing reads, the idea of combining cloud computing with statistical modeling may also be useful for other experiments that generate massive amounts of data."

Provided by Johns Hopkins University Bloomberg School of Public Health

Citation: Cloud computing method greatly increases gene analysis (2010, September 8) retrieved 9 April 2024 from <https://phys.org/news/2010-09-cloud-method-greatly-gene-analysis.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--