

World Leader in Genomic Research, J. Craig Venter Institute, Selects Infobright to Help Unlock the Secrets of the Oceans

Overview

The J. Craig Venter Institute (JCVI), formed in October 2006 through the merger of several affiliated and legacy scientific research organizations, is a world leader in genomic research. Its president and founder Dr. Venter is best known for the pioneering research he and his team did in decoding the first draft of the human genome.

Since its founding, JCVI has been engaged in some of the most fruitful and exciting research in the biological sciences. One of the key initiatives is to unlock the secrets of the oceans by sampling, sequencing and analyzing the DNA of the microorganisms living in there. The Sorcerer II Global Ocean Sampling Expedition has already uncovered more than 40 million new genes and thousands of new protein families from organisms found in sea water. The team uses a metagenomic approach in which genomic analysis (the analysis of all DNA in an organism) is conducted on the entire community of microbes within the sample rather than isolating and culturing each individual microbial species.

As part of the Expedition, JCVI researchers have traveled around the world, collecting samples from such diverse locations as deep sea ocean vents, the air above New York City, soils in Panamanian jungles and the human mouth and gut. These samples are then analyzed against a database that contains previously decoded microbial genomes. Through this process, many new organisms are being identified, and their genomic information has become part of the ever-expanding database of known organisms. JCVI researchers are using this information to better understand the evolution of these organisms, and how they are related to each other. This information can also lead to advances in human health, climate and environmental remediation, agricultural improvements, and in other important areas.

Process and Tools

Not surprisingly, genomic sequencing and analysis requires capturing, storing and analyzing large data sets. To make the data accessible to researchers both within JCVI and the scientific community at large, JCVI developed APIS, the Automated Phylogenomic Inference System. APIS is updated as the genomes of new organisms are completed.



INDUSTRY Scientific Research

ORGANIZATION

J. Craig Venter Institute

SOLUTION

Scientific research database for metagenomics

RESULTS

Queries run 10 times faster, and the data compression is 17:1.

“While it is still early in our use of Infobright, it has proven to be very efficient and easy to deploy.

We don’t need to do a lot of manipulation of the system to get the performance we need, and the compression and performance are big advantages over alternative databases.”

Michael Heaney
JCVI
IT

CASE STUDY

APIS includes a number of database instances. Two public databases, previously stored in MySQL, were moved to the Infobright® analytic database to accommodate the growing number of users who wanted access to the data.

Why Infobright?

Infobright, designed for fast queries against large data volumes, was an ideal choice for JCVI as it provided:

- **17:1 data compression:** One database was 433GB in MySQL, and only 25GB of space in Infobright. The second database was compressed from 112GB of data in MySQL to 7GB in Infobright.
- **Query speed improvements of 10x**
- **Low administrative burden on IT:** No indexes, no need to change schemas as the data moved from MySQL to Infobright, easy to manage.
- **Low cost:** Today, Infobright is supporting an increasing number of researchers using a standard server, and running in a virtual machine with 2 CPUs and 8 GB RAM. Its high rate of compression means a significant decrease in the amount of storage needed to support the growing data volume.

Infobright technology is used by many organizations to enable greater insight from large volumes of data. Its use by JCVI for this important research is particularly exciting and meaningful for Infobright's employees and community.

About Infobright

Infobright, the open source analytic database company, delivers a high performance database that is the preferred choice for analytic applications and data marts. Fast and easy to implement, with operational simplicity and low cost, Infobright is being used by enterprises, SaaS and software companies in online businesses, telecommunications, financial services and other industries to provide rapid access to critical business data.

For more information, please visit www.infobright.com or join our open source community at www.infobright.org.

CONTACT US

Corporate Headquarters:

47 Colborne Street, Suite 403
 Toronto, Ontario M5E1P8 Canada
 Tel. 416 596 2483
 Toll Free 877 596 2483
 Fax 416 352 5674

Sales :

Toll Free 877 596 2483 x225
 416 596 2483 x225
info@infobright.com

European Office:

Block 4, Suite 204
 Europa Business Centre
 Harcourt Road
 Dublin 2 Ireland
 Tel: +353 87 743 71 07

www.infobright.com

www.infobright.org

Twitter: [Infobright](#)

YouTube: [infobrightdb](#)