

## New DNA dataset is potent, accessible tool

## April 14 2015

Scientists focused on producing biofuels more efficiently have a new powerful dataset to help them study the DNA of microbes that fuel bioconversion and other processes.

In a paper published in *Nature Scientific Data*, researchers from the Department of Energy's Oak Ridge National Laboratory, North Carolina State University and LanzaTech describe methods and results for sequencing the Clostridium autoethanogenum bacterium. These and other microorganisms play important roles in biofuels, agriculture, food production, the environment, health and disease.

Armed with this knowledge, researchers can modify and optimize the <u>microbes</u> to convert waste into fuel or chemicals. The new *Nature* journal is dedicated to making scientifically valuable datasets more accessible, and as a result, ORNL scientists envision additional advances over the next few years.

"The publication of this genome dataset comprised of high-quality results can serve as a benchmark to assist others in developing new sequencing technologies and computational algorithms," said Steve Brown, a co-author and member of ORNL's Biosciences Division.

This particular microbe, classified as complex because of its more than 4.3 million base pairs of DNA, posed a significant challenge to scientists who used five sequencing technologies to achieve their results. Brown also noted that the dataset represents three generations of sequencing technologies and contains six types of data from four next-generation



sequencing platforms.

"These data are comprehensive and there are few similar examples," Brown said. "We see the data being used in a variety of ways, as a benchmark for new <u>sequencing technologies</u> and testing new bioinformatics algorithms for microbial genome sequences."

The team's publication builds on work published last year in *Biotechnology for Biofuels*, in which researchers announced they had sequenced the entire genome of the Clostridium autoethanogenum bacteria. In this paper, researchers included the raw sequence data generated from a PacBio instrument that produces valuable long DNA reads. The previously unpublished data from N.C. State are now accessible via the new journal.

Brown noted that researchers around the world will now be able to study Clostridium autoethanogenum DNA methylation motifs - naturally occurring chemical modifications - in their efforts to understand and exploit them for genetic engineering.

Provided by Oak Ridge National Laboratory

Citation: New DNA dataset is potent, accessible tool (2015, April 14) retrieved 27 April 2024 from <u>https://phys.org/news/2015-04-dna-dataset-potent-accessible-tool.html</u>

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.