HMP DACC Releases New Collection of Data

Baltimore, Md. — April 17, 2014.

Researchers with the Data Analysis and Coordination Center (DACC) for the NIH Human Microbiome Project (HMP) are releasing a significant amount of new sequence data generated by the HMP sequencing centers and Demonstration Projects. The data was produced through the Human Microbiome Project (HMP), a large-scale multi-institutional research effort, funded by the National Institutes of Health Common Fund. The mission of the HMP has been to provide baseline datasets, tools, protocols and resources to help scientists better understand the microbial communities of the human body and their effects on health and disease. The DACC is located at the Institute for Genome Sciences, University of Maryland School of Medicine. The DACC was established to collect, process, analyze, and disseminate the vast quantity of HMP sequence and analysis data. The DACC is also charged with providing software tools and computing resources to the research community to maximize their ability to use the HMP data. Sequencing was performed by HMP Demonstration Projects and sequencing centers including Baylor College of Medicine, Broad Institute, J. Craig Venter Institute, and Washington University. This is the latest in a series of data release milestones from the HMP.

This release includes whole metagenome sequencing reads from ~800 healthy human cohort samples (beyond the 764 previously released), both raw and trimmed 16S reads from ~5,000 healthy human cohort samples (beyond the 5,034 previously released), and associated metadata. Also included are raw whole metagenome sequencing reads and raw 16S reads from fourteen Demonstration Projects focused on specific disease conditions. With this new release, the total amount of data available on the DACC resource has surpassed 14 terabytes. The DACC data browser has been reorganized to facilitate access to the new data. This data is also being mirrored on Amazon S3 in their public datasets collection (http://aws.amazon.com/datasets/1903160021374413). This will enable users to leverage free data

transfers between Amazon S3 and Amazon EC2 in order to analyze the data in Amazon EC2. So far, five terabytes of DACC data has been made available on Amazon S3.

"HMP data has been extremely valuable to researchers worldwide as they explore the impact of the microbiome on human health and disease" said Owen White, PhD, Principal Investigator for the DACC and Associate Director, Bioinformatics at the Institute for Genome Sciences (IGS) at the University of Maryland School of Medicine. "Making data accessible in efficient and user-friendly retrieval systems is vitally important to the medical and scientific communities. The HMP DACC is meeting this need for HMP data."

DACC website: http://www.hmpdacc.org/
The DACC research and website has been funded by NIH grant U01-HG004866.

####

About the University of Maryland School of Medicine
Established in 1807, the University of Maryland School of Medicine is the first public medical school in the United States and the first to institute a residency-training program. The School of Medicine was the founding school of the University of Maryland and today is an integral part of the 11-campus University System of Maryland. On the University of Maryland's Baltimore campus, the School of Medicine serves as the anchor for a large academic health center which aims to provide the best medical education, conduct the most innovative biomedical research and provide the best patient care and community service to Maryland and beyond. www.medschool.umaryland.edu

About the Institute for Genome Sciences
The Institute for Genome Sciences (IGS) is a research center within the University of Maryland School of Medicine. Comprised of an interdisciplinary team of investigators, the Institute uses the powerful tools of genomics and bioinformatics to understand genome function in health and disease, to study molecular and cellular networks in a variety of model systems, and to generate data and bioinformatics resources of value to the international scientific community.
www.igs.umaryland.edu