141. MaxBin: An Automated Metagenomic Binning and Reassembly Algorithm

Steven W. Singer1,2*(SWSinger@lbl.gov), Yu-Wei Wu1,3, Jeffrey A. Kimbrel1,3, Stephanie Eichorst1,3, Yung-Hsu-Tang1,4, Janet Jansson1,2,5, Susannah G Tringe6, Blake A. Simmons1,5

1Joint BioEnergy Institute, Emeryville, CA; 2Earth Sciences and 3Physical Biosciences Divisions, Lawrence Berkeley National Laboratory, Berkeley, CA; 4San Francisco State University, San Francisco, CA; 5Biological Sciences Division, Pacific Northwest National Laboratory; 6Joint Genome Institute, Walnut Creek, CA; 7Biological and Materials Science Center, Sandia National Laboratories, Livermore, CA

http://downloads.jbei.org/data/MaxBin.html.

Project Goals: The Microbial Communities group at the Joint BioEnergy Institute translates activities found in the environment to applications in the biorefinery. The group develops bioinformatic tools to investigate microbial consortia that deconstruct plant biomass. These tools have broad application in understanding the metabolic potential of natural and engineered microbial communities.

Recovering individual genomes from metagenomic datasets by binning assembled sequence provides access to the metabolic potential of uncultivated microbial populations that may have important roles in natural and engineered ecosystems. We have developed a binning algorithm, MaxBin, which automates the binning of assembled metagenomic scaffolds using an expectation-maximization algorithm and is capable of binning individual metagenomes using tetranucleotide frequencies and read coverage of assembled sequences. MaxBin 2.0 was developed to bin genomes from multiple samples using differential read coverage and was compared to CONCOCT and MetaBAT, two recently described metagenomic binning algorithms that recover genomes using tetranucleotide frequencies and differential coverage. MaxBin also incorporates a reassembly function, which automatically retrieves the paired-end reads associated with the assembled sequences of binned genomes and reassembles them. MaxBin has been applied to recover genomes from microbial communities relevant to biomass deconstruction, including thermophilic microbial communities grown on biomass and halophilic communities found along a salinity gradient in San Francisco Bay.

Publications

1. “MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation- maximization algorithm” Wu, Y.W.; Tang, Y.H.; Tringe, S.G.; Simmons, B.A.; Singer, S.W. Microbiome, 2014, 2, 26.

This work was performed as part of the DOE Joint BioEnergy Institute (http://www.jbei.org) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE- AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy