

DiTing: A Pipeline to Infer and Compare Biogeochemical Pathways from Metagenomic and Metatranscriptomic Data

Chunxu Xue¹, heyu lin², Xiao-Yu Zhu¹, Jiwen Liu¹, Gary Rowley³, jonathan todd³, Meng Li⁴, and Xiao-Hua Zhang⁵

¹Affiliation not available

²The University of Melbourne School of Earth Sciences

³University of East Anglia School of Biological Sciences

⁴Shenzhen University

⁵Ocean University of China

November 18, 2020

Abstract

Metagenomics and metatranscriptomics are powerful tools to uncover key microbes and processes driving biogeochemical cycling in natural ecosystems. Currently available databases depicting metabolic functions from metagenomic/metatranscriptomic data are not dedicated to biogeochemical cycles. There are no databases encompass genes involved in the cycling of dimethylsulfoxide (DMS), an abundant organosulfur compound. Additionally, a recognized normalization mode to estimate and compare the relative abundance and environmental importance of pathways from metagenomic and metatranscriptomic data has not been available. These limitations impact the ability to accurately relate key microbial driven biogeochemical processes to differences in environmental conditions. Thus, an easy to use specialized tool that infers and visually compares the potential for biogeochemical processes, including DMS cycling, is urgently required. To solve these issues, we developed DiTing, a tool wrapper to infer and compare biogeochemical pathways among a set of given metagenomic or metatranscriptomic reads in one step, based on the KEGG (Kyoto Encyclopedia of Genes and Genomes) and a manually created DMS cycling gene database. Accurate and specific formulas for over 100 pathways were developed to calculate their relative abundance. Output reports detail the relative abundance of biogeochemically-relevant pathways in both text and graphical format. We applied DiTing to metagenomes from simulated data, hydrothermal vents and the Tara Ocean project. The DiTing outputs were consistent with genetic feature of genomes used in simulated benchmark data, and also demonstrated that the predicted functional profiles correlated strongly with changes in environmental conditions. DiTing can now be confidently applied to wider metagenomic and metatranscriptomic datasets.

Hosted file

DiTing_main_document.pdf available at <https://authorea.com/users/376840/articles/493696-diting-a-pipeline-to-infer-and-compare-biogeochemical-pathways-from-metagenomic-and-metatranscriptomic-data>



