MetaPrism: A Toolkit for Joint Taxa/Gene Analysis of Metagenomic Sequencing Data

CURRENT STATUS: POSTED

Jiwoong Kim
University of Texas Southwestern Medical Center at Dallas

Shuang Jiang
Southern Methodist University

Guanghua Xiao
University of Texas Southwestern Medical Center at Dallas

Yang Xie
University of Texas Southwestern Medical Center at Dallas

Dajiang Liu
Pennsylvania State University

Qiwei Li
University of Texas at Dallas

Andrew Koh
University of Texas Southwestern Medical Center at Dallas

Xiaowei Zhan
University of Texas Southwestern Medical Center

Xiaowei.Zhan@Utsouthwestern.edu

Corresponding Author
ORCiD: https://orcid.org/0000-0002-6249-7193

DOI: 10.21203/rs.2.22868/v1

SUBJECT AREAS
General Microbiology

KEYWORDS
metagenomics sequence analysis, joint analysis, microbiome biomarker
Abstract

Background In microbiome research, metagenomic sequencing generates enormous amounts of data. These data are typically classified into taxa for taxonomy analysis, or into genes for functional analysis. However, a joint analysis where the reads are classified into taxa-specific genes is often overlooked.

Result To enable the analysis of this biologically meaningful feature, we developed a novel bioinformatic toolkit, MetaPrism, which can analyze sequence reads for a set of joint taxa/gene analyses: 1) classify sequence reads and estimate the abundances for taxa-specific genes; 2) tabularize and visualize taxa-specific gene abundances; 3) compare the abundances between groups, and 4) build prediction models for clinical outcome. We illustrated these functions using a published microbiome metagenomics dataset from patients treated with immune checkpoint inhibitor therapy and showed the joint features can serve as potential biomarkers to predict therapeutic responses.

Conclusions MetaPrism is a toolkit for joint taxa and gene analysis. It offers biological insights on the taxa-specific genes on top of the taxa-alone or gene-alone analysis. MetaPrism is open source software and freely available at https://github.com/jiwoongbio/MetaPrism. The example script to reproduce the manuscript is also provided in the above code repository.

Full-text

Due to technical limitations, full-text HTML conversion of this manuscript could not be completed. However, the manuscript can be downloaded and accessed as a PDF.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

Supplementary.pdf