

國立東華大學應用數學系 專題演講

主講人：王澤毅博士

中央研究院資訊科學研究所

講題：Rank-based methods in metagenomic data analysis

時間：101年3月7日(星期三) 10:30-12:00

地點：理學院A324會議室

摘要

Metagenomics involves sampling, sequencing, and studying the genetic materials in microbial communities. Recent advances in next generation sequencing technologies have fueled an increase in both the number and scope of metagenomic projects, and large amount of data are being accumulated for analyzing microbiomes. Several statistical methods have been proposed for comparative analysis of taxonomic or functional compositions across microbial communities. However, the estimated abundances of taxonomic units or functional groups are highly noisy due to sampling biases and other systematic artifacts in metagenomic data processing, e.g., 16S rRNA chimeras, artificial replicates, and inaccurate binning.

We have developed a rank conversion scheme, MetaRank, to convert abundances into ranks for comparative analysis. MetaRank employs a series of statistical hypothesis tests to compare abundances within a microbial community and determine their ranks. Empirical tests on synthetic samples and real metagenomes confirm that MetaRank is able to diminish noise and facilitate the characterization of microbiomes. In this talk, I will introduce MetaRank and describe the strengths and weakness of rank-based methods in comparative metagenomics. Further, the statistical challenges in next generation sequencing data analysis are discussed.



※※※ 歡迎參加 ※※※ se1010307