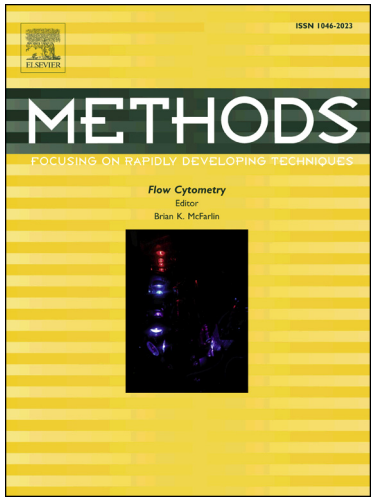
Accepted Manuscript



Functional Association Prediction by Community Profiling

Dazhi Jiao, Wontack Han, Yuzhen Ye

PII: S1046-2023(17)30027-0
DOI: http://dx.doi.org/10.1016/j.ymeth.2017.04.018
Reference: YMETH 4194

To appear in: *Methods*

Received Date: 18 January 2017
Revised Date: 31 March 2017
Accepted Date: 20 April 2017

Please cite this article as: D. Jiao, W. Han, Y. Ye, Functional Association Prediction by Community Profiling, *Methods* (2017), doi: http://dx.doi.org/10.1016/j.ymeth.2017.04.018

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Functional Association Prediction by Community Profiling

Dazhi Jiao[a,b], Wontack Han[a,b], Yuzhen Ye[a,*]

[a] *150 S. Woodlawn Ave, Indiana University, Bloomington, IN 47405*
[b] *These authors contributed equally to this work*

Abstract

Recent years have witnessed unprecedented accumulation of DNA sequences and therefore protein sequences (predicted from DNA sequences), due to the advances of sequencing technology. One of the major sources of the hypothetical proteins is the metagenomics research. Current annotation of metagenomes (collections of short metagenomic sequences or assemblies) relies on similarity searches against known gene/protein families, based on which functional profiles of microbial communities can be built. This practice, however, leaves out the hypothetical proteins, which may outnumber the known proteins for many microbial communities. On the other hand, we may ask: what can we gain from the large number of metagenomes made available by the metagenomic studies, for the annotation of metagenomic sequences as well as functional annotation of hypothetical proteins in general? Here we propose a community profiling approach for predicting functional associations between proteins: two proteins are predicted to be associated if they share similar presence and absence profiles (called community profiles) across microbial communities. Community profiling is conceptually similar to the phylogenetic profiling approach to functional prediction, however with fundamental differences. We tested different profile construction methods, the selection of reference metagenomes, and correlation metrics, among others, to optimize the performance of this new approach. We demonstrated that the community profiling approach alone slightly outperforms

*Corresponding author: yye@indiana.edu

دریافت فوری ← متن کامل مقاله

ISIArticles
مرجع مقالات تخصصی ایران

✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
✓ امکان دانلود نسخه ترجمه شده مقالات
✓ پذیرش سفارش ترجمه تخصصی
✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
✓ دانلود فوری مقاله پس از پرداخت آنلاین
✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات