Accepted Manuscript

High-Resolution Characterization of the Human Microbiome

Cecilia Noecker, Colin P. McNally, Alexander Eng, Elhanan Borenstein

PII: S1931-5244(16)30127-X

DOI: 10.1016/j.trsl.2016.07.012

Reference: TRSL 1077

To appear in: *Translational Research*

Received Date: 29 April 2016

Revised Date: 12 July 2016

Accepted Date: 15 July 2016

Please cite this article as: Noecker C, McNally CP, Eng A, Borenstein E, High-Resolution Characterization of the Human Microbiome, *Translational Research* (2016), doi: 10.1016/j.trsl.2016.07.012.

High-Resolution Characterization of the Human Microbiome

Cecilia Noecker[1,*], Colin P McNally [1,*], Alexander Eng [1,*], Elhanan Borenstein [1,2,3,†]

[1] Department of Genome Sciences, University of Washington, Seattle, WA
[2] Department of Computer Science and Engineering, University of Washington, Seattle, WA
[3] Santa Fe Institute, Santa Fe, NM
* These authors contributed equally to this work (order determined randomly)
[†] Corresponding author: EB, elbo@uw.edu

Abstract

The human microbiome plays an important and increasingly recognized role in human health. Studies of the microbiome typically employ targeted sequencing of the 16S rRNA gene, whole metagenome shotgun sequencing, or other meta-omic technologies to characterize the microbiome's composition, activity, and dynamics. Processing, analyzing, and interpreting these data involve numerous computational tools that aim to filter, cluster, annotate, and quantify the obtained data and ultimately provide an accurate and interpretable profile of the microbiome's taxonomy, functional capacity, and behavior. These tools, however, are often limited in resolution and accuracy and may fail to capture many biologically- and clinically-relevant microbiome features, such as strain-level variation or nuanced functional response to perturbation. Over the past few years, extensive efforts have been invested toward addressing these challenges and developing novel computational methods for accurate and high-resolution characterization of microbiome data. These methods aim to quantify strain-level composition and variation, detect and characterize rare microbiome species, link specific genes to individual taxa, and more accurately characterize the functional capacity and dynamics of the microbiome. These methods and the ability to produce detailed and precise microbiome information are clearly essential for informing microbiome-based personalized therapies. In this review, we survey these methods, highlighting the challenges each method sets out to address and briefly describing methodological approaches.

Introduction

Recent marked advances in sequencing technologies have been followed by an explosion of studies utilizing these technologies to explore a wide range of microbial communities, including those that inhabit the human body. Such studies apply targeted sequencing of the 16S rRNA gene as well as whole metagenome shotgun sequencing to characterize the human microbiome in numerous settings. Analyses of these sequencing data commonly use an assortment of clustering, binning, annotation, and assembly algorithms to ultimately profile the composition of species in each sample, the set of genes they collectively encode, or the genome sequence of specific member species (Figure 1). Combined, these efforts to map the human microbiome in health and