

REVIEW ARTICLE

Hypothesis testing and statistical analysis of microbiome



Yinglin Xia ^{a,b,**}, Jun Sun ^{b,*}

^a Division of Academic Internal Medicine and Geriatrics, Department of Medicine, University of Illinois at Chicago, Chicago, IL, USA

^b Division of Gastroenterology and Hepatology, Department of Medicine, University of Illinois at Chicago, Chicago, IL, USA

Received 10 April 2017; accepted 9 June 2017
Available online 23 June 2017

KEYWORDS

Bioinformatics;
Biostatistics;
Cancer;
Hypothesis testing;
IBD;
Microbiome;
Statistical methods
and models;
Vitamin D receptor

Abstract After the initiation of Human Microbiome Project in 2008, various biostatistic and bioinformatic tools for data analysis and computational methods have been developed and applied to microbiome studies. In this review and perspective, we discuss the research and statistical hypotheses in gut microbiome studies, focusing on mechanistic concepts that underlie the complex relationships among host, microbiome, and environment. We review the current available statistic tools and highlight recent progress of newly developed statistical methods and models. Given the current challenges and limitations in biostatistic approaches and tools, we discuss the future direction in developing statistical methods and models for the microbiome studies.

Copyright © 2017, Chongqing Medical University. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

The gut microbiome plays fundamental roles in the human health. It can be considered as a newly identified organ that interacts with other organs and influences the development

of disease.^{1,2} Human Microbiome Project (HMP) was initiated in 2008 by the National Institutes of Health Roadmap for Biomedical Research and constructed as a large, genome-scale community research project.³ The

* Corresponding author. Division of Gastroenterology and Hepatology, Department of Medicine, University of Illinois at Chicago, 840 S Wood Street, Room 704 CSB, MC716, Chicago, IL 60612, USA.

** Corresponding author. Division of Academic Internal Medicine and Geriatrics, Department of Medicine, University of Illinois at Chicago, 1747 W. Roosevelt Rd., Room 420, Chicago, IL 60618, USA.

E-mail addresses: yxia@uic.edu (Y. Xia), junsun7@uic.edu (J. Sun).

Peer review under responsibility of Chongqing Medical University.

HMP project needs data analysis, computational methods development, and the public availability of tools and data.⁴

Human gut microbiome study is to understand not only the microbiome community composition, but also the dynamic interactions among microbiome, host, environment, and disease intervention. The microbiome studies require a multi-disciplinary team effort, involving basic, translational, and clinical investigators. The next phase of research investigation of the gut microbiome should be guided by specific biological questions relevant to the clinical aspects and natural history of the disease, utilizing the full spectrum of 'omic' technologies, bioinformatic analysis, and experimental models.⁵ However, the significant roles of biostatisticians and bioinformatic and biostatistical methods in gut microbiome studies are underestimated, especially the appropriate use of biostatistic tests is largely ignored.

Here, we discuss the statistical hypothesis tests in both community composition and microbiome-host interactions. We review the utility of various statistical approaches for assessing the diversity of microbiome communities and analyzing and modeling the association between community composition of the microbiome and host. We summarize the current available statistic tools for microbiome studies. We highlight the recent progress in new statistical methods and models. In doing so, we provide specific examples of these methods and discuss how to appropriately apply them into microbiome study. In the meantime, we bring up the limitations and daunting challenges ahead of us that must be overcome in order to move the field forward. Furthermore, we discuss the development of statistical methods, the limits, and future direction.

Research and statistical hypotheses in human microbiome studies

In the current microbiome studies, there are mainly two themes: 1) to characterize the relationship between microbiome features and biological, genetic, clinical or experimental conditions; and 2) to identify potential biological and environmental factors that are associated with microbiome composition. The goal of these studies is to understand mechanisms of host genetic and environmental factors that shape our microbiome. Insights gained from the studies potentially contribute to the development of therapeutic strategies in modulating the microbiome composition in human diseases.^{6,7}

Dynamic interactions exist among environment, microbiome and host (Fig. 1). To study the complicated interactions among these factors, three general research hypotheses have been developed and used in the field: hypothesis 1 is to test the association between environment and host. There are no specific features for this hypothesis compared to other biomedical sciences. To test the hypothesis 1, we can use the standard statistical methods and models, which are commonly used in other biomedical sciences. For the microbiome studies, the focus is on the following research hypothesis 2 and 3:

The research hypothesis 2 is to test the association between microbiome and host: whether the composition of

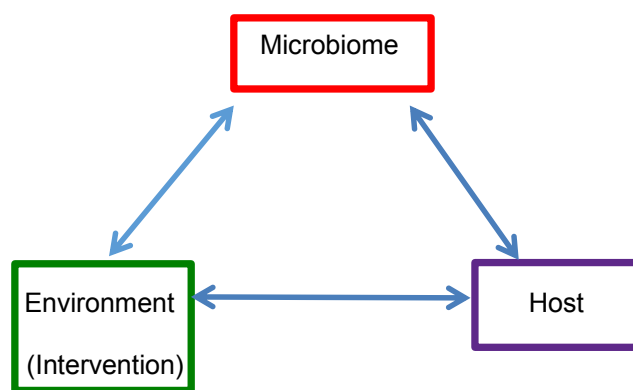


Figure 1 Dynamic Interactions among environment, microbiome and host for the research hypotheses in microbiome studies.

the microbiome or "dysbiotic" microbiome is linked to the health or disease of host. For example, in inflammatory bowel diseases (IBD) research,^{8,9} dysbiosis is associated with the progression of the diseases. Lack of vitamin D receptor (VDR) causes dysbiosis and changes the functions of the murine intestinal microbiome.¹⁰ Altered bacterial community is associated with different intestinal epithelial VDR status.¹¹

The research hypothesis 3 is to test whether microbiome is associated with environmental or biological covariates,¹² the impact of environmental factors on microbiome,¹³ or whether there is an effect of intervention on a specific microbiome composition (diversity) in health and disease. The examples include testing whether dietary interventions shape gut microbiota,^{8,14} the impact of a probiotic intervention on the composition of the human microbiota.¹⁵ The longitudinal studies have tested antibiotics and diet effects on gut microbial community structure,⁹ analyzed whether nutrition influence gut microbiome composition at the level of bacterial species,¹⁶ or hypothesized that antibiotic treatments affect the diversity of strains of gut bacteria.¹³ In a recent paper, Bokulich et al showed that antibiotic exposure and delivery mode alter bacterial diversity and delay microbiota maturation and infant diet affects diversity of intestinal microbiome.¹⁷

Statisticians usually develop their statistical hypotheses based on the research hypotheses. Based on the research hypotheses, the null statistical hypothesis is developed as "there is no difference of microbiome composition in health and disease (or experimental groups or genetic conditions)" or "there is no difference (change) of microbiome composition in different environmental factors (or intervention). Although these statistical hypotheses have the core theme that explores impacts of environmental or external factors (e.g. interventions) on composition and/or richness of microbiota, they could focus on various topics, including, alpha diversity (species diversity in each individual sample), bacterial richness, total number of unique operational taxonomic units (OTUs), phylogenetic diversity (the relative amount of diverse phylogenetic lineages), and species evenness in each sample.¹⁷

The statistical hypothesis could be alpha diversity. For example, for antibiotic studies, we can hypothesize that

Download English Version:

<https://daneshyari.com/en/article/5532632>

Download Persian Version:

<https://daneshyari.com/article/5532632>

[Daneshyari.com](https://daneshyari.com)