The need for standards in human microbiome research

Cara N. Wilder, PhD, Juan Lopera, PhD, Dev Mittar, PhD, Carol Horton, MS, Briana Benton, BS

The human body harbors a series of diverse, dynamic microbial communities comprising bacteria, archaea, viruses, and eukaryotes. These communities, collectively termed the human microbiome, play a vital role in human health and disease, particularly with regard to physiology and development, immunomodulation, metabolic regulation, and protection against pathogenic strains.[1,2] By understanding how the human microbiome develops and changes over time with respect to lifestyle and environmental changes or various disease states, and how this in turn affects individual biology, we are offered a powerful tool for personalized healthcare and precision medicine.

Early studies on the human microbiome have been predominantly dependent on microbial cultivation; however, this approach is limited as many microorganisms cannot be cultured *in vitro*.[1] Thus, culture-independent methods for the analysis of microbial communities have been sought out. Accelerated advancements in sequencing have facilitated the use of a metagenomics-based approach for profiling whole microbial genomes directly from their natural environment. This technology has provided further understanding on the genetic potential of microbial communities, how the human microbiome may evolve over time with regard to species abundance and community composition, and how dysbiosis could potentially affect an individual's predisposition to disease.

Though metagenomic studies have offered a wealth of information on the human microbiome, the complexities associated with commonly used methods have posed significant challenges toward assay standardization. Here, bias can be introduced at every stage of a metagenomics workflow, from sample collection, DNA extraction, amplification, library preparation, sequencing, to data analysis.[1,3] Consequently, this bias can obscure the true composition of a microbial community, leading to inaccurate analyses and incorrect conclusions.

Take DNA extraction, for example. The quantity and quality of DNA extracted from Gram-negative versus Gram-positive bacteria frequently vary due to the inherent differences between their cell wall compositions. Therefore, one of the very first steps in the workflow can itself lead to biases and variations in microbiome data. When additional steps are added to the process, the inherent differences and complexities within each of these steps further compound the number of biases introduced throughout the workflow. This then leads to inaccurate analyses and incorrect conclusions. Optimization at each step of the analytical process through standardized methods and controls then becomes essential to the validity and reproducibility of scientific outcomes.

One of the primary challenges hindering assay standardization is the limited availability of credible reference materials. Here, the use of mock microbial communities as controls can help identify issues, determine error rates, and normalize
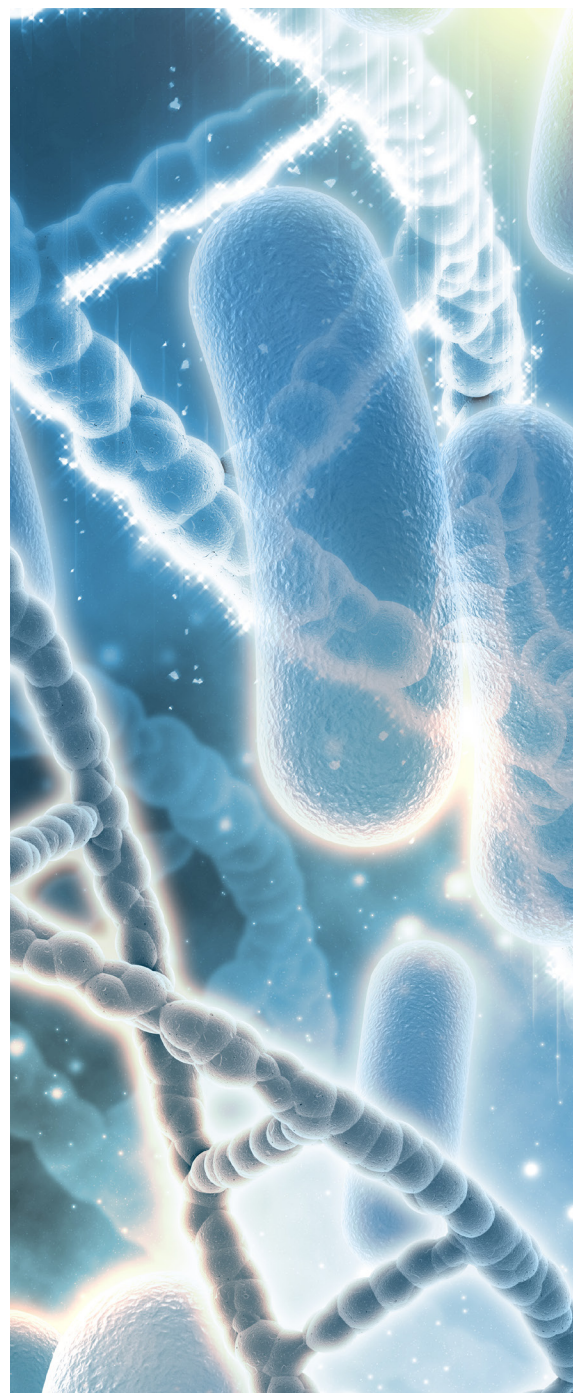
sources of assay bias during sample processing and analysis, in turn improving result interpretation. These mixed communities are typically defined in composition, represent an appropriate level of diversity, and may include a variety of genera sourced on or within the human body, each exhibiting different, relevant phenotypic and genotypic attributes.

However, given the diversity of microorganisms, it is not possible to make a single standard that can represent any and every species present in the human microbiome. Rather, the composition of a standard would need to vary depending on its application as a daily run control or in assay development and optimization. Therefore, to meet the diverse needs within microbiome research, ATCC worked with key opinion leaders in the field to develop a continually growing portfolio of ATCC® Microbiome Standards. For each mock microbial community, strains were carefully selected based on their clinical relevance as well as by the presence of specific phenotypic and genotypic attributes ranging from cell wall properties to genome complexity and size. The standards were then prepared as genomic nucleic acid or whole cell mixtures from fully sequenced, characterized, and authenticated cultures from ATCC. These products are available with even or staggered abundance as well as varying levels of complexity ranging from 3 to 20 strains per sample.

ATCC® Microbiome Standards enable the optimization of metagenomics workflows and microbiome research applications, providing reliable comparative data while improving assay consistency. Further, they support different aspects of microbiome research such as pathogen detection; environmental testing; research on the mycobiome and virome; and site-specific studies on the oral, gut, vaginal, and skin microbiomes. The inclusion of these mock community controls throughout a metagenomics study is essential for the identification of potential biases and can help further elucidate the impact of assay variation on the profiles of microbial communities obtained from microbiome samples.

To further enhance the use of these microbiome standards, ATCC has collaborated with One Codex to combine the power of physical laboratory standards with the leading bioinformatics platform for microbial genomics and metagenomics. Through this collaboration, ATCC has worked in conjunction with One Codex to develop an easy-to-use data analysis module for the ATCC® Microbiome Standards that includes pre-loaded metadata; the ability to analyze both shotgun and 16S rRNA data; automated quality scores assessing true positives, false positives, and relative abundance; and data management, storage, sharing, and graphing capabilities. The combination of the standards and data analysis platform provides an ideal tool for standardizing data from a wide range of sources as well as generating consensus among various microbiome applications and analyses.

Overall, a metagenomics-based approach for profiling the human microbiome has greatly enhanced our knowledge of microbial communities and how these communities affect human health and disease. However, because biases can be introduced throughout the metagenomics workflow, the true species abundance and community composition can be obscured. It is only with authenticated standards and optimized workflows that scientists will be able to achieve a clear understanding of how the human microbiome changes and develops over time, ultimately relating back to a deeper knowledge of human physiology. ATCC® Microbiome Standards combined with the One Codex data analysis platform offer a unique solution for assay standardization and eliminating bias during data analysis. Together, these tools enable you to optimize your metagenomics research applications with confidence and improve the consistency and reproducibility of your data run after run.

References

1. Committee on Metagenomics: Challenges and Functional Applications. *The New Science of Metagenomics: Revealing the secrets of our Microbial Planet*. The National Academies Press, Washington, DC, 2007.
2. Wang W L, *et al*. Application of metagenomics in the human gut microbiome. *World J Gastroenterol* 21(3): 803-814, 2015.
3. Brooks JP, *et al*. The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiol* 15: 66, 2015.