

Research Use Statement (2,200 characters max, including spaces).

The main objective of the proposed research project is to provide an R software framework for processing, storing and analysis of the integrative Human Microbiome Project (iHMP) data. These data are publicly available via the iHMP data portal (<http://portal.hmpdacc.org/>) and contains the collection of a total of 16,531 samples from 4 major HMP studies: Multi-Omic Microbiome Study: Pregnancy Initiative (MOMS-PI), Type 2 diabetes mellitus (T2D), and Inflammatory Bowel Disease Multi-Omics data base (IBDMDB) (<https://www.hmpdacc.org/ihmp/>). This is the largest collection of publicly available longitudinal multi-omics studies collected to understand the microbial biomarkers of preterm pregnancy, diabetes, and inflammatory bowel disease. While basic clinical annotation information, such as pregnancy or IBD outcome, is available through the data portal, the majority of participants data important for the analysis of disease outcomes, such as age, socioeconomic status, and medical history requires dbGAP access. The goal of this project is to securely integrate these data into the R software framework enabling authorized users the in-depth analysis of -omics data. This project extends our previous work on the integration of microbiome data with the controlled access dbGAP data (<https://www.biorxiv.org/content/early/2018/08/29/299115>). This study will advance: (1) the accessibility of iHMP data; (2) increase researchers' interest in iHMP data; (3) advance understanding of the connection between multi-omics measurements and physiological factors that cause spontaneous pre-term birth, the progression of type 2 diabetes and IBD. This is the collaborative project, which will include the following investigators: (1) Virginia Commonwealth University: Ekaterina Smirnova, Assistant Professor, Department of Biostatistics; Mikhail Dozmorov, Assistant Professor, Department of Biostatistics, and Jennifer Fettweis, Assistant Professor, Department of Microbiology and Immunology; (2) CUNY School of Public Health: Levi Waldron, Associate Professor, Department of Epidemiology and Biostatistics; (3) Johns Hopkins University: Ni Zhao, Department of Biostatistics.

Non-Technical Summary (1,100 characters max, including spaces)

The human microbiome contributes to human well-being, disease progression, pregnancy outcomes. Several consortia provide open-access microbiome data; however, software tools for the analysis of it remain undeveloped. This project will develop an R package enabling the integrative analysis of the microbiome data securely integrated with clinical annotations available for authorized users.