Kenneth Rainin Foundation
2023 Innovations Symposium
July 17 – July 18, 2023 | Palace Hotel, San Francisco

Strain-resolved inference of microbial gene content in fecal microbiota transplantation to treat ulcerative colitis

Byron J. Smith[1,2], Katherine S. Pollard[1,2,3]
[1]The Gladstone Institutes, Data Science and Biotechnology [2]Chan Zuckerberg Biohub
[3]University of California, San Francisco, Department of Epidemiology and Biostatistics

Background: Although Fecal microbiota transplant (FMT) has been used as an experimental treatment for inflammatory bowel disease (IBD), in particular ulcerative colitis (UC), patient and donor selection remains a challenge. One reason for this is that distinct strains of the same bacterial species can differ in relevant traits, such as resistance to antimicrobial therapies, metabolism of pharmaceutical products, and interactions with the host immune system.

Hypothesis: We hypothesize that the variable gene content among strains present in FMT treatments from different donors can contribute to the unpredictable response of UC patients to the therapy.

Methods: To explore this hypothesis, we developed a novel approach for microbial gene content reconstruction that harnesses metagenomic sequencing and pangenome profiling on multiple samples, and identifies robust correlations in sequence abundance in order to confidently assign genes to individual genomes. Importantly, by incorporating strain tracking based on single-nucleotide polymorphisms, we expand this approach to also capture variable gene content. We validated our method using metagenomic data from a realistic synthetic community, and then applied it to a large, publicly available collection of metagenomes from IBD patients and healthy controls, as well as a trial of FMT for UC.

Results: Our analysis revealed extensive gene content variation across hundreds of species and thousands of strains. Specifically, we found that in *Escherichia coli*—which is often at elevated abundance in IBD patients—variable genes are enriched in functions such as motility, cell surface components, and defense mechanisms. We also identified distinct gene content in conspecific strains from two FMT donors and used functional annotations to predict the impacts of these differences on bacterial physiology.

Conclusions: Our approach enhances the accuracy and resolution of gene content determination for individual strains in metagenomic data.

Impact on our knowledge of IBD: These findings underscore the importance of considering variation in bacterial strains in the rational design of microbial therapies for IBD and motivate further exploration of strain diversity across patients and donors.