Multi-cohort, multi-regional metagenomic association study of inflammatory bowel disease patients and non-IBD controls

Sungho Lee[1] and Insuk Lee[1]

Presenting Author: Sungho Lee (shlee354@yonsei.ac.kr) / Correspondence: Insuk Lee (insuklee@yonsei.ac.kr)
[1]Department of Biotechnology, College of Life Science and Biotecnnology, Yonsei University, Seoul, Korea

Abstract

Inflammatory bowel disease (IBD), which includes Crohn's disease (CD) and ulcerative colitis (UC), is a group of chronic inflammatory conditions of the gastrointestinal tract. The incidence and prevalence of IBD have been high mainly in Western countries such as North America, northern and western Europe. In the last two decades, incidences of IBD in Asia and eastern Europe have also risen steadily due to westernized diet and changes in other environmental factors(Kaplan, G. G., 2015). IBD is one of the representative diseases associated with gut microbial dysbiosis. Common changes in the gut microbiome in IBD patients include an increase in facultative anaerobes and a decrease in obligately anaerobic producers of short-chain fatty acids (SCFAs)(Morgan, X. C. et al., 2012). Besides functional dysbiosis of the gut microbiome, multiple factors including host genetics and environmental conditions such as diet also affect immunological responses and inflammation in the intestine; the complex interplay between these factors must be considered. Taking this into account, we performed a multi-cohort metagenomic association study of IBD, utilizing metagenomic shotgun sequencing data of 1,002 cross-sectional stool samples collected from published studies. Sequencing data were pre-processed and batch corrected with compositionality-aware methods. Clean reads were then aligned to a comprehensive gut microbial genome database HRGM(Kim, CY. et al., In Press) to establish taxonomic profiles of all samples. Batch effects and differences among multi-regional cohorts were assessed, and differently abundant microbial species were then identified.

Results

Figure 1. UMAP visualization of initial batch effect of 1,002 metagenomic samples.

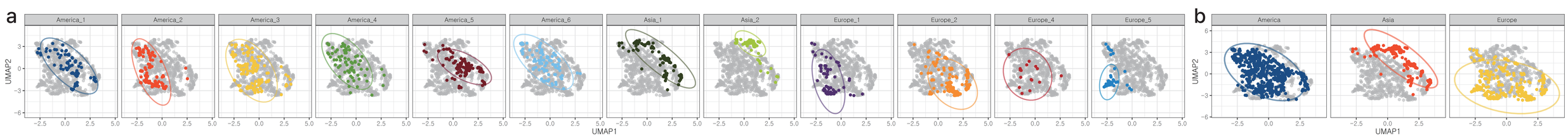

Figure 2. UMAP visualization of batch information and disease annotation after batch correction procedure.

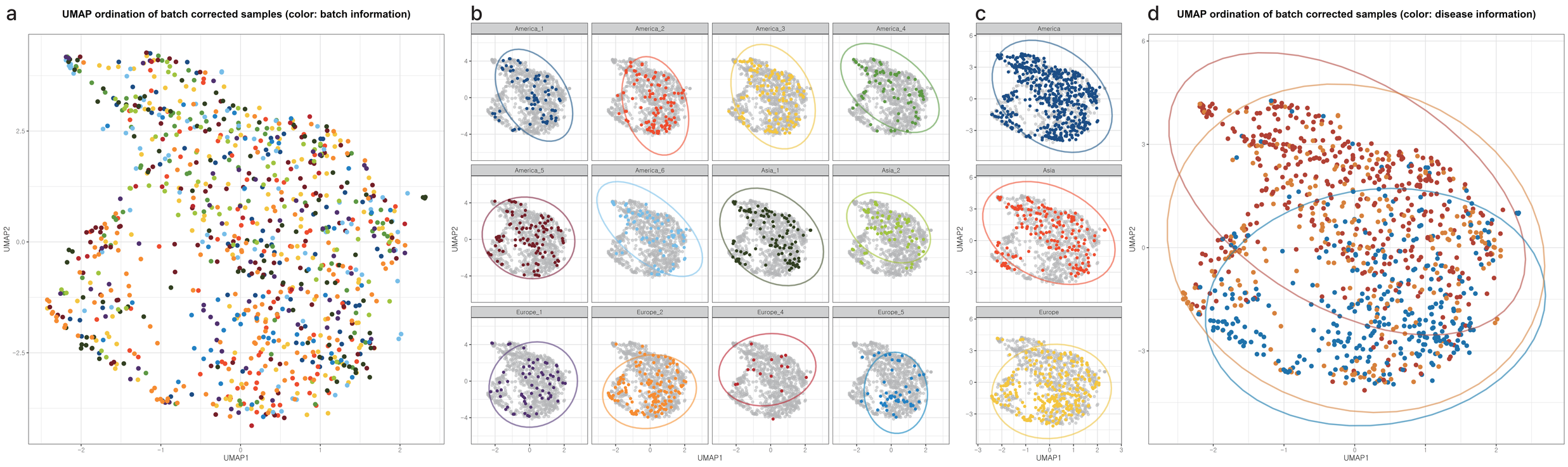

Figure 3. Differentially abundant species between CD/nonIBD and UC/nonIBD.

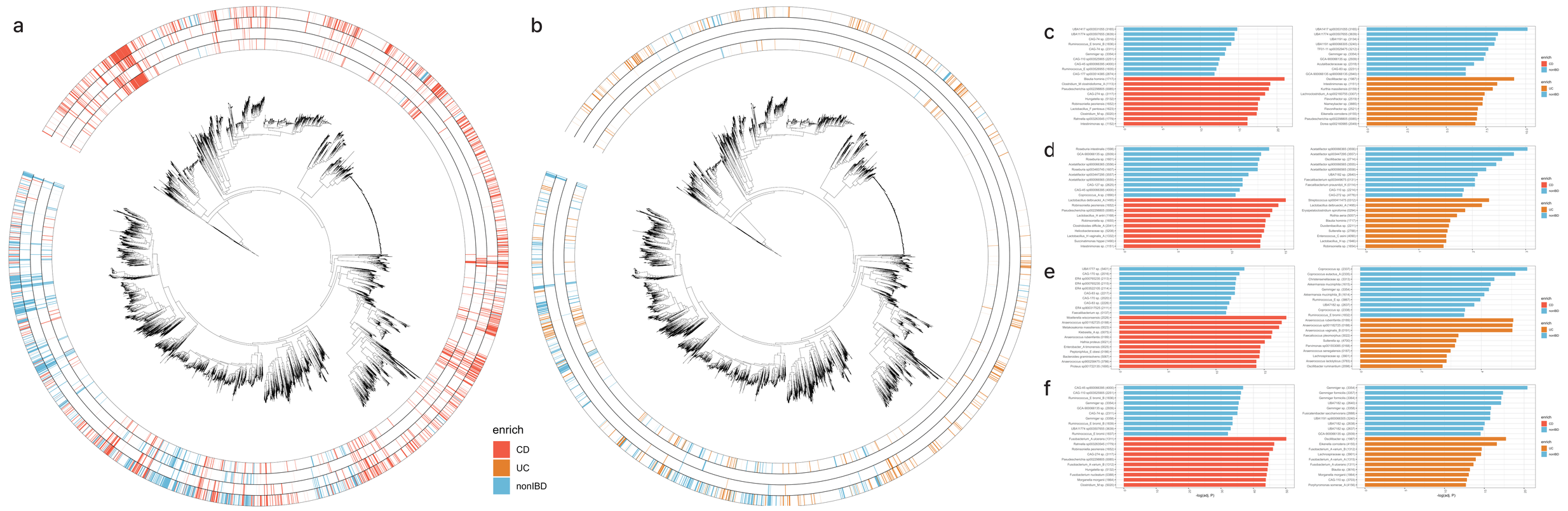

Figure 1. UMAP visualization of samples before batch correction. (a) Cohort-specific or (b) country-specific datapoints were highlighted and 95% confidence ellipses were shown.
Figure 2. UMAP visualization of samples after batch correction. (a) Overview of all 1,002 datapoints, (b) Cohort-specific, (c) country-specific, and (d) class-specific datapoints with 95% confidence ellipses.
Figure 3. Differentially abundant (DA) taxa. DA taxa between (a) CD vs. nonIBD and (b) UC vs. nonIBD were visualized on a phylogenetic tree of bacterial species included in HRGM database. Top 10 DA taxa for CD, UC, and nonIBD in (c) American cohort, (d) Asian cohort, (e) European cohort, and (f) all 1,002 samples were shown (Wilcoxon rank sum test, Bonferroni correction).

Discussion

Heterogeneity in 1,002 metagenomic samples from 12 datasets (Fig. 1) were handled by batch correction with removeBatchEffect function in R (Ritchie, ME. et al., 2015) (Fig. 2a, 2b, 2c), while preserving separation between CD, UC, and nonIBD samples (PERMANOVA p = 0.001) (Fig 2d). Differential abundance analysis of microbial species (Fig. 3) showed an increased abundance of beneficial microbes, such as Gemmiger spp., Akkermansia municiphila, and Faecalibacterium prausnitzii, in non-IBD individuals, and an increase of facultative anaerobes including Fusobacterium in CD and UC patients.

Further Studies

We found a meaningful discordance of gut microbial composition between samples from different geographical regions, which is probably due to differences in lifestyle factors and genetics. These diversities will need to be further analyzed within each regional cohort. Differentially abundant species found in this study will be experimentally validated after functional curation. We aim to discover novel gut microbial taxa associated with IBD which can be utilized as a biomarker for diagnosis and microbial therapy.