MRGM: Mouse reference gut microbiome enabling comprehensive comparison of gut microbiome between mouse and human

Nayeon Kim[1,3], Chan Yeong Kim[1,3], Dongjin Park[2], Sang-Jun Ha[2] and Insuk Lee[1*]

[1] Department of Biotechnology, College of Life Science & Biotechnology, Yonsei University, Seoul 03722, Korea

[2] Department of Biochemistry, College of Life Science & Biotechnology, Yonsei University, Seoul 03722, Korea

[3] These authors contribute equally.

Gut microbiome is associated with many diseases and can be modulated by diet and drugs. Mouse may provide effective *in vivo* model for the study of diet and drug effect on human gut microbiome. For the microbiome profiling with high resolution and coverage, a comprehensive reference gut microbial genomes needs to be available. We constructed mouse reference gut microbiome (MRGM) which include newly assembled genomes from >850 metagenomes. There are several improvements in our mouse reference gut microbiome compared with previous databases. First, we included metagenome-assembled genomes (MAGs) from many samples with ultra-deep sequencing (>20 Gb), which was found to improve quality of MAGs. Second, we filtered the assembled genomes not only by conventional criteria; completeness and contamination, but also by genome chimerism. Third, we chose genomes with completeness ≥70% for the database and found that this threshold achieves an optimal trade-off between accuracy of reference genomes and coverage of sequencing reads with taxonomic classification. The MRGM provides over 46,267 non-redundant genomes by 70% completeness and 5% contamination, comprising 1,689 representative prokaryotic species. MRGM also provides catalog of 15.2 million coding genes along with their functional annotations.