

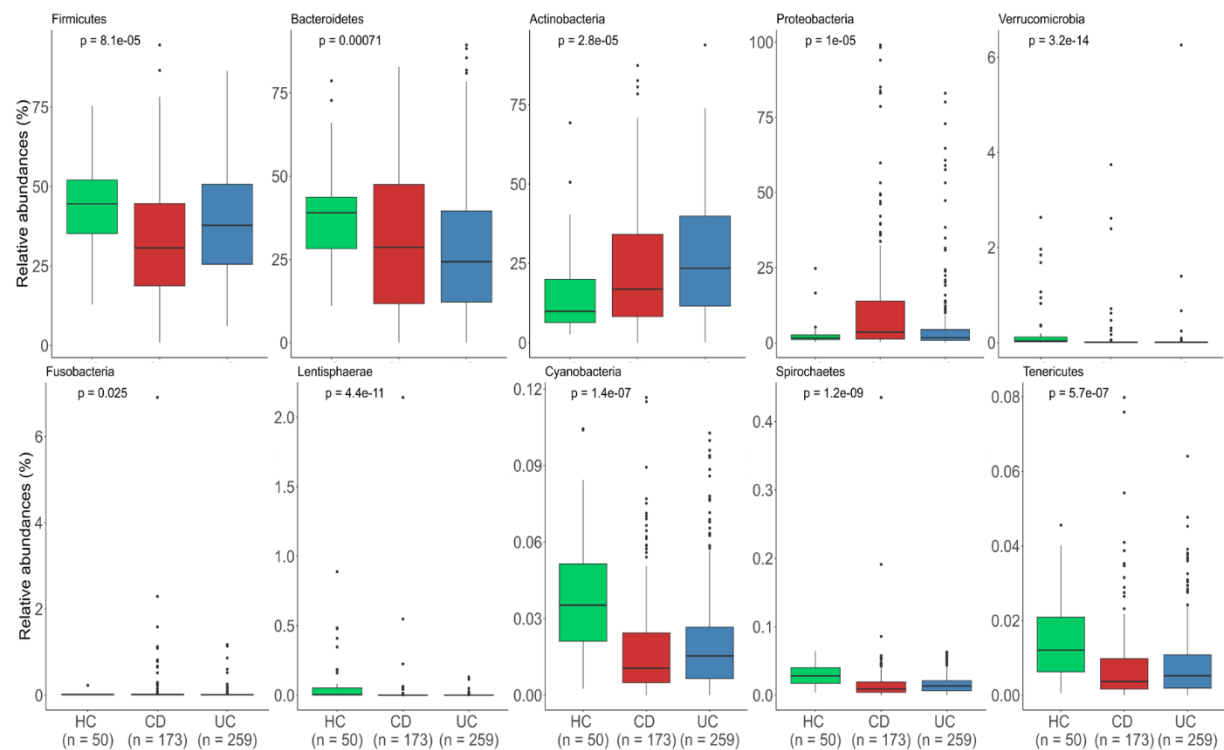
Supplementary file

Diagnosis of Crohn's Disease and Ulcerative Colitis Using the Microbiome

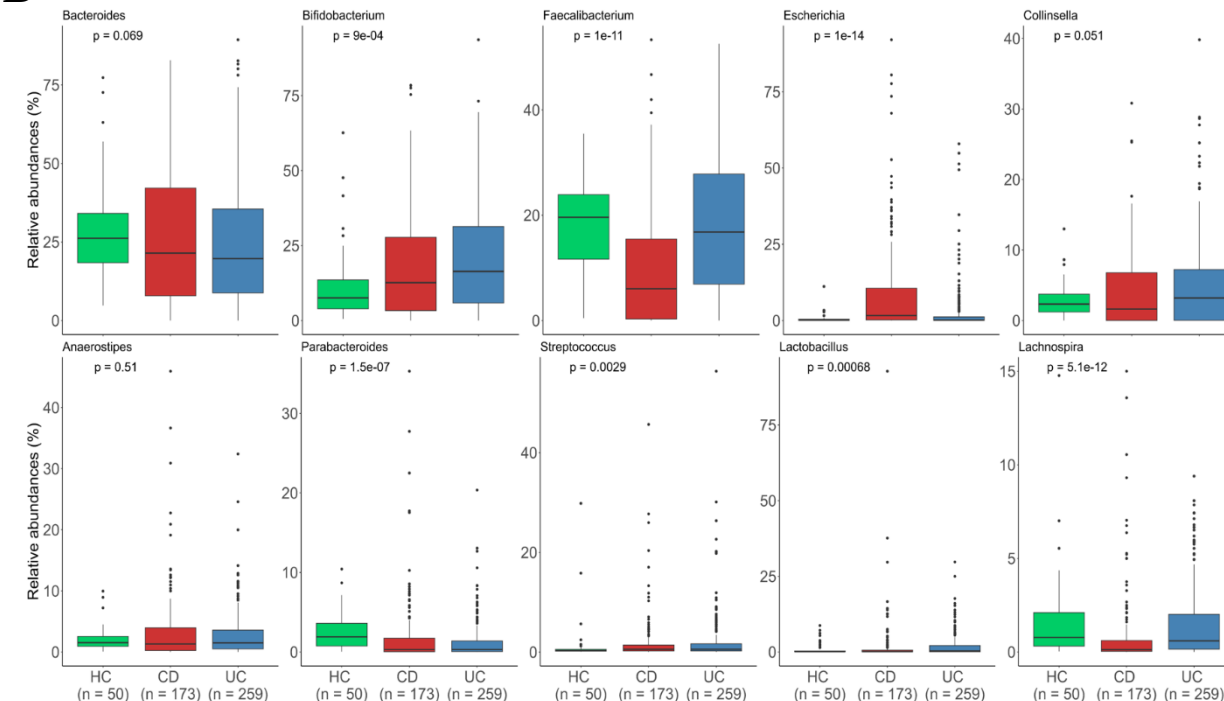
Da-Yeon Kang^{1,2*}, Jong-Lyul Park^{3*}, Min-Kyung Yeo^{4*},
Sang-Bum Kang⁵, Jin-Man Kim⁴, Ju Seok Kim^{6^}, and
Seon-Young Kim^{7^}

Fig. S1

A



B



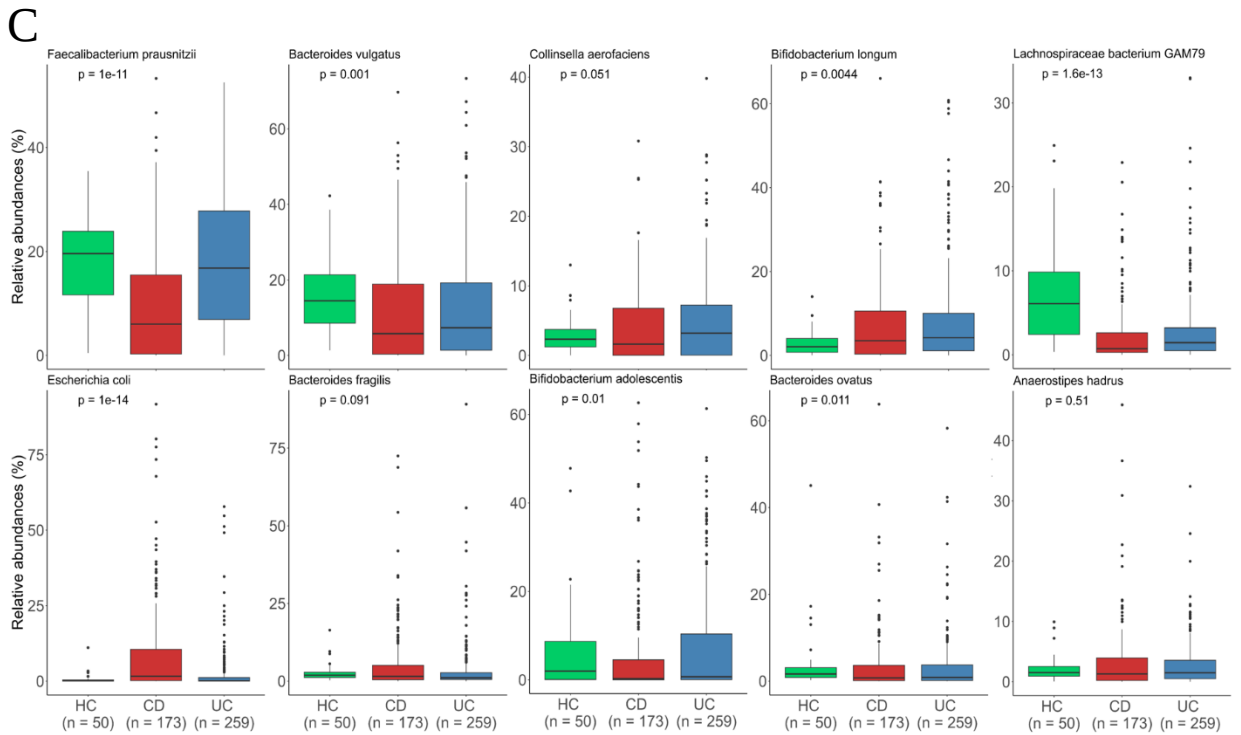
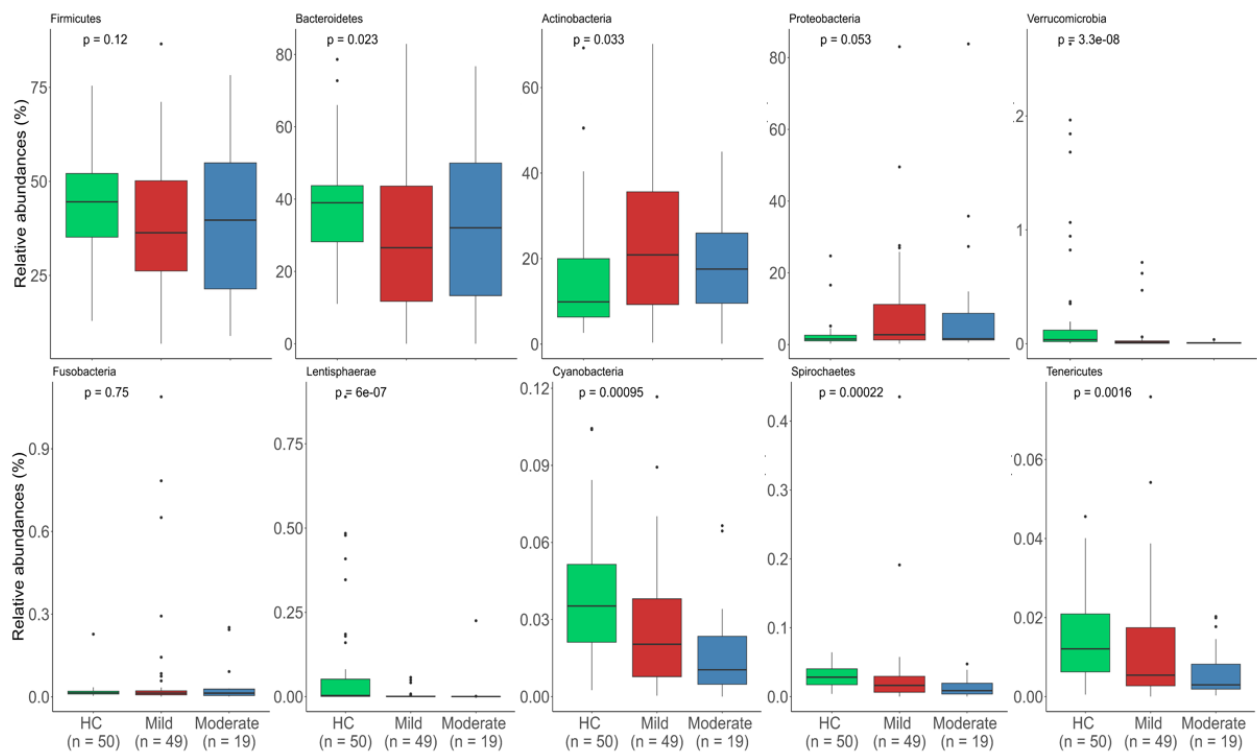


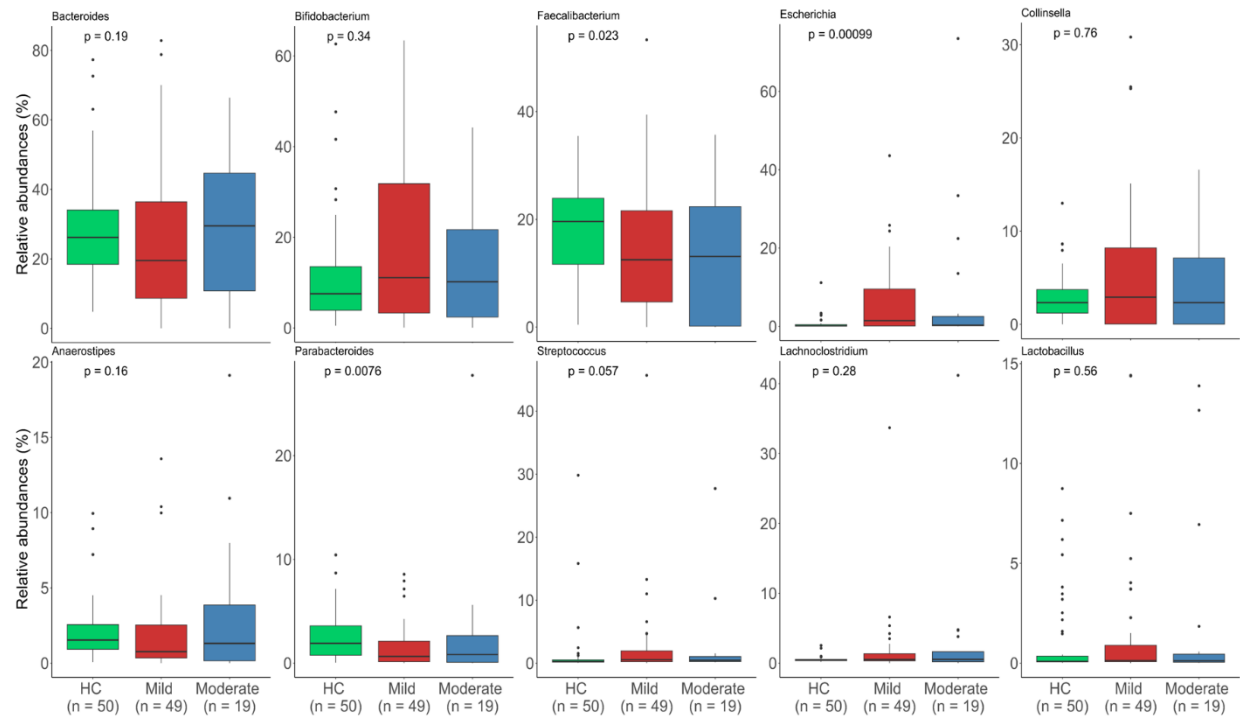
Fig. S1 Comparison of the relative abundance of microbiota among HC, CD, and UC in phylum (A), genus (B) and species (C) based on the Kruskal-Wallis test. Each figure is shown in the order of HC, CD and UC.

Fig. S2

A



B



C

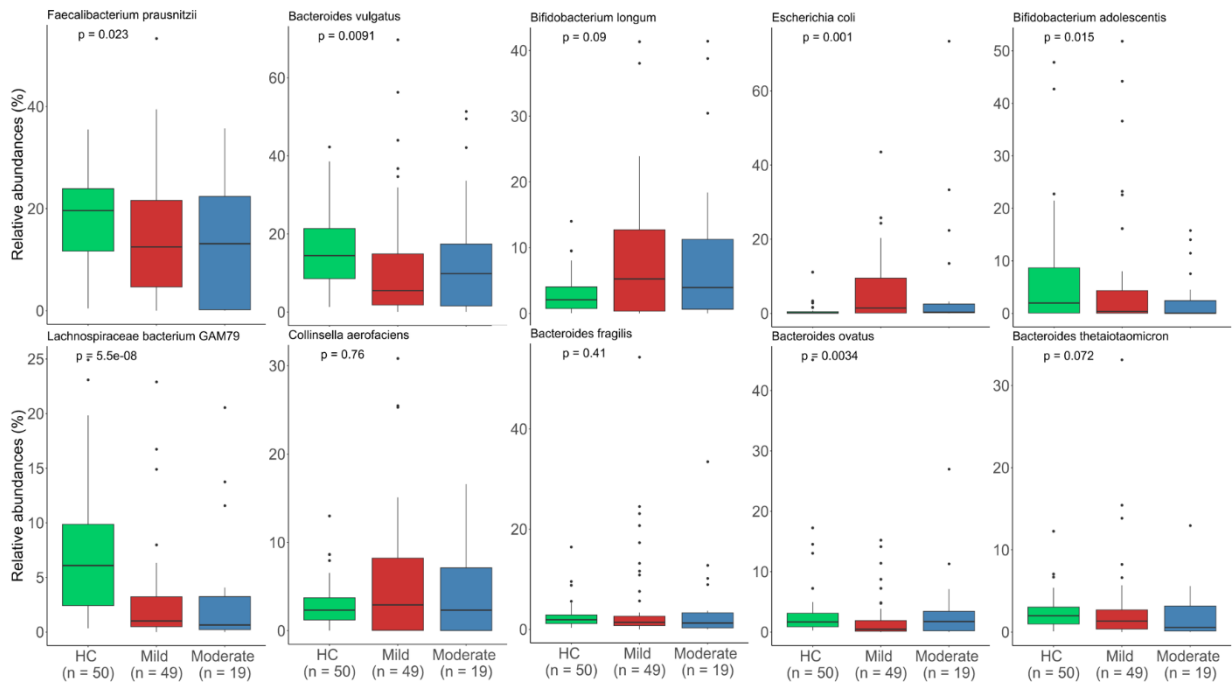
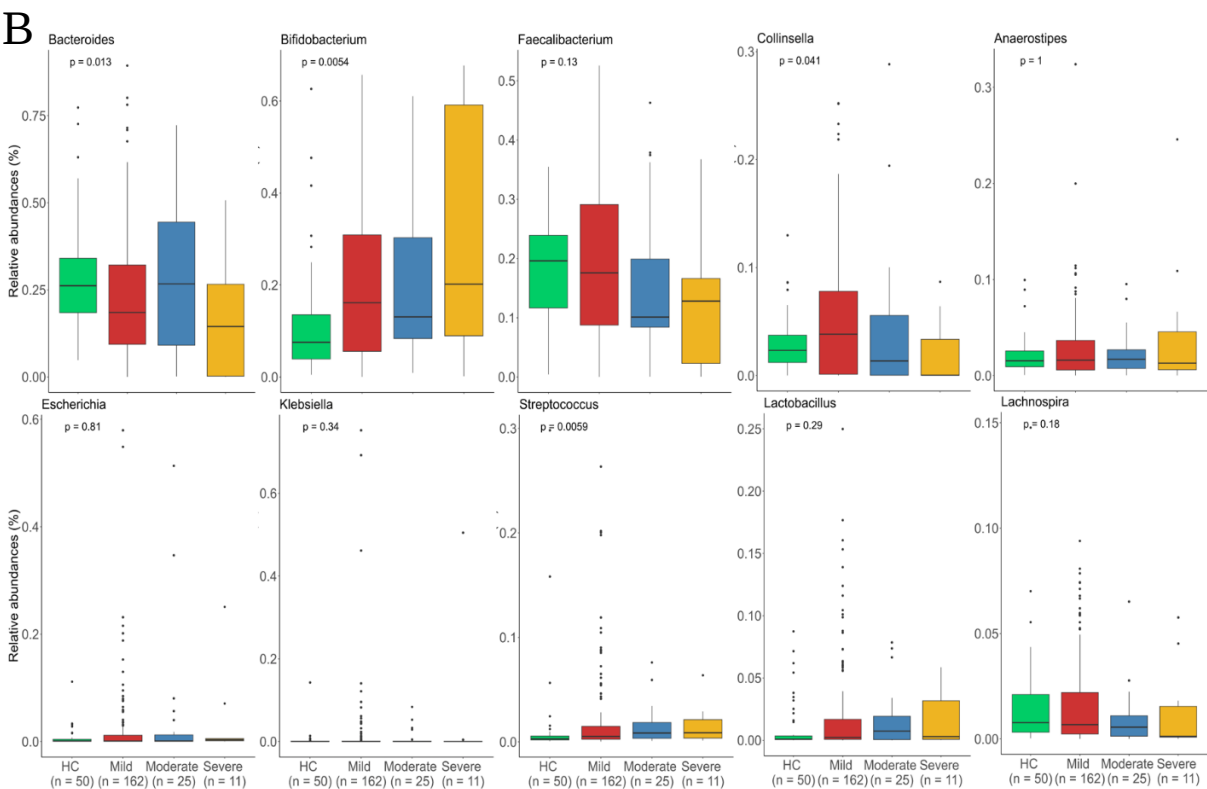
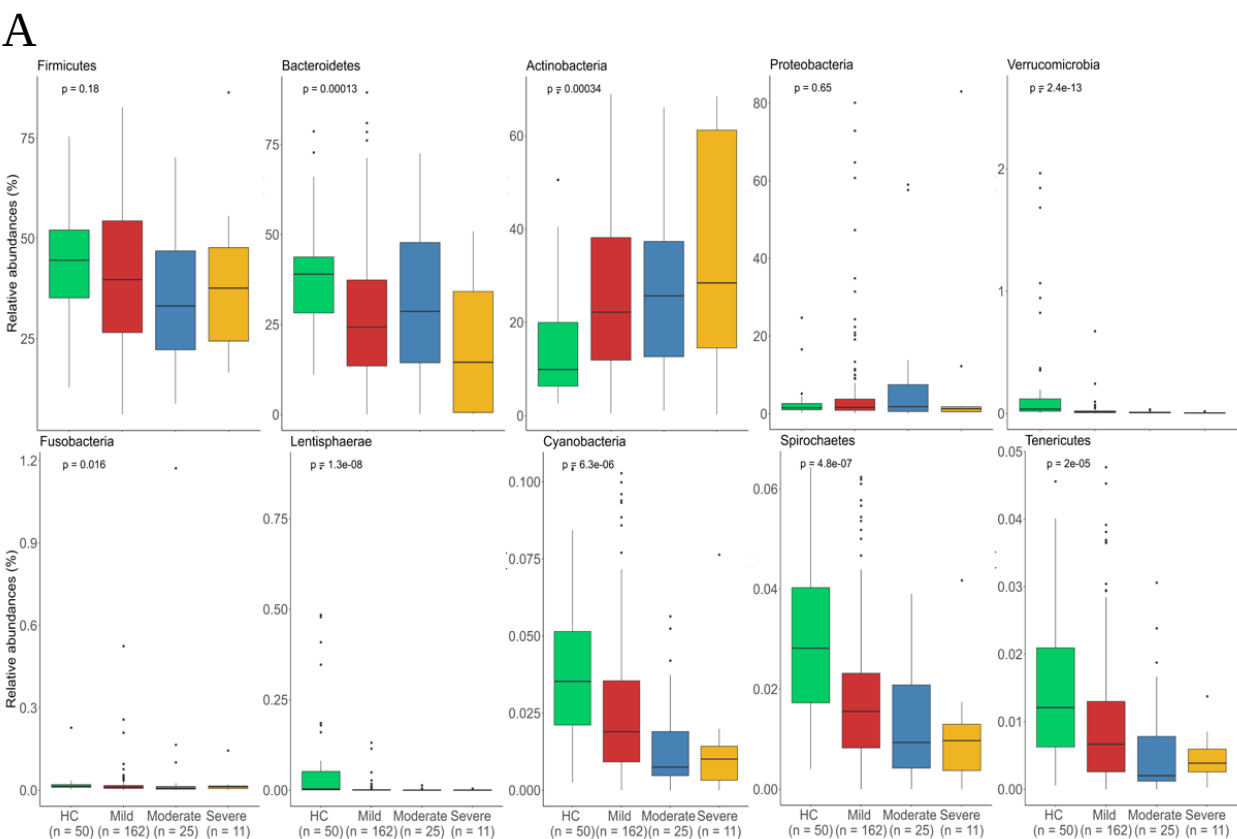


Fig. S2 Comparison of the relative abundance of microbiota among disease stages (HC, mild, and moderate) in CD patients at the phylum (A), genus (B) and species (C) level. Numbers are listed in the following order: HC, mild, and moderate.

Fig. S3



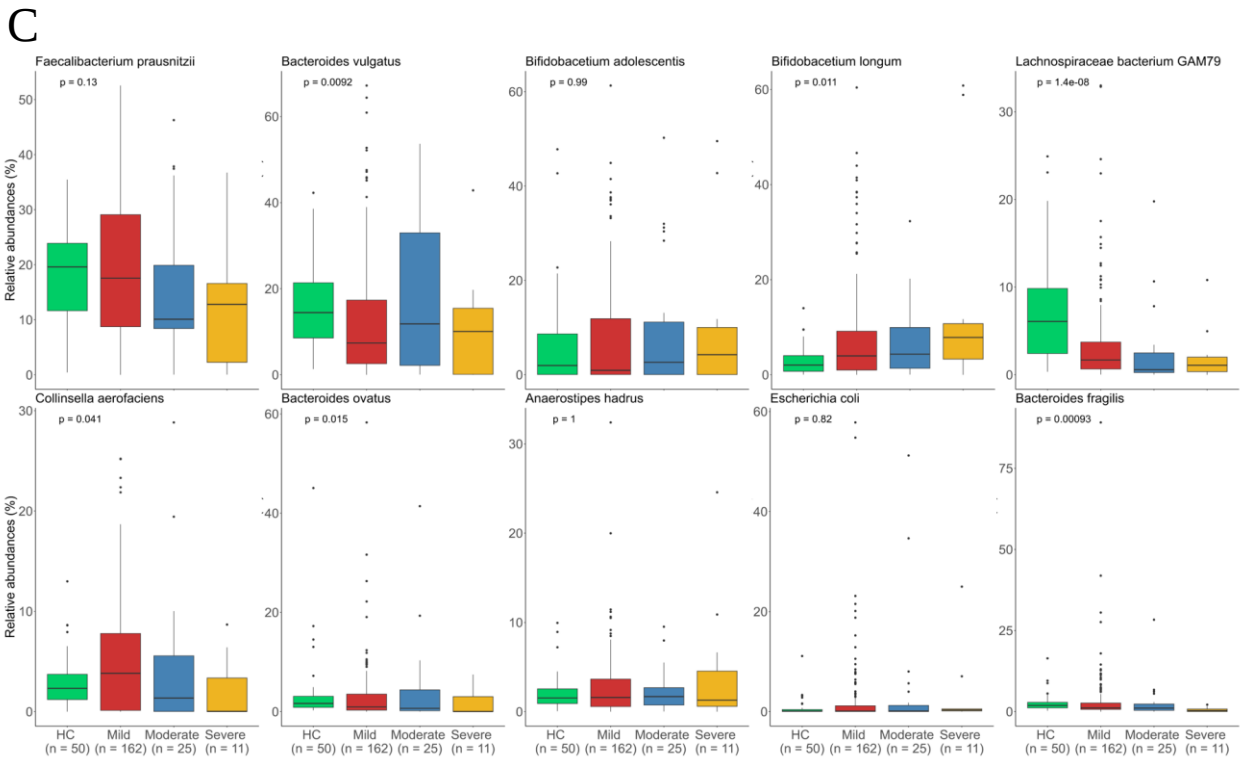


Fig. S3 Comparison of the relative abundance of microbiota between disease stages in UC patients at the phylum (A), genus, (B) and species (C) level. Numbers are listed in the following order: HC, mild, moderate, and severe.

Fig. S4

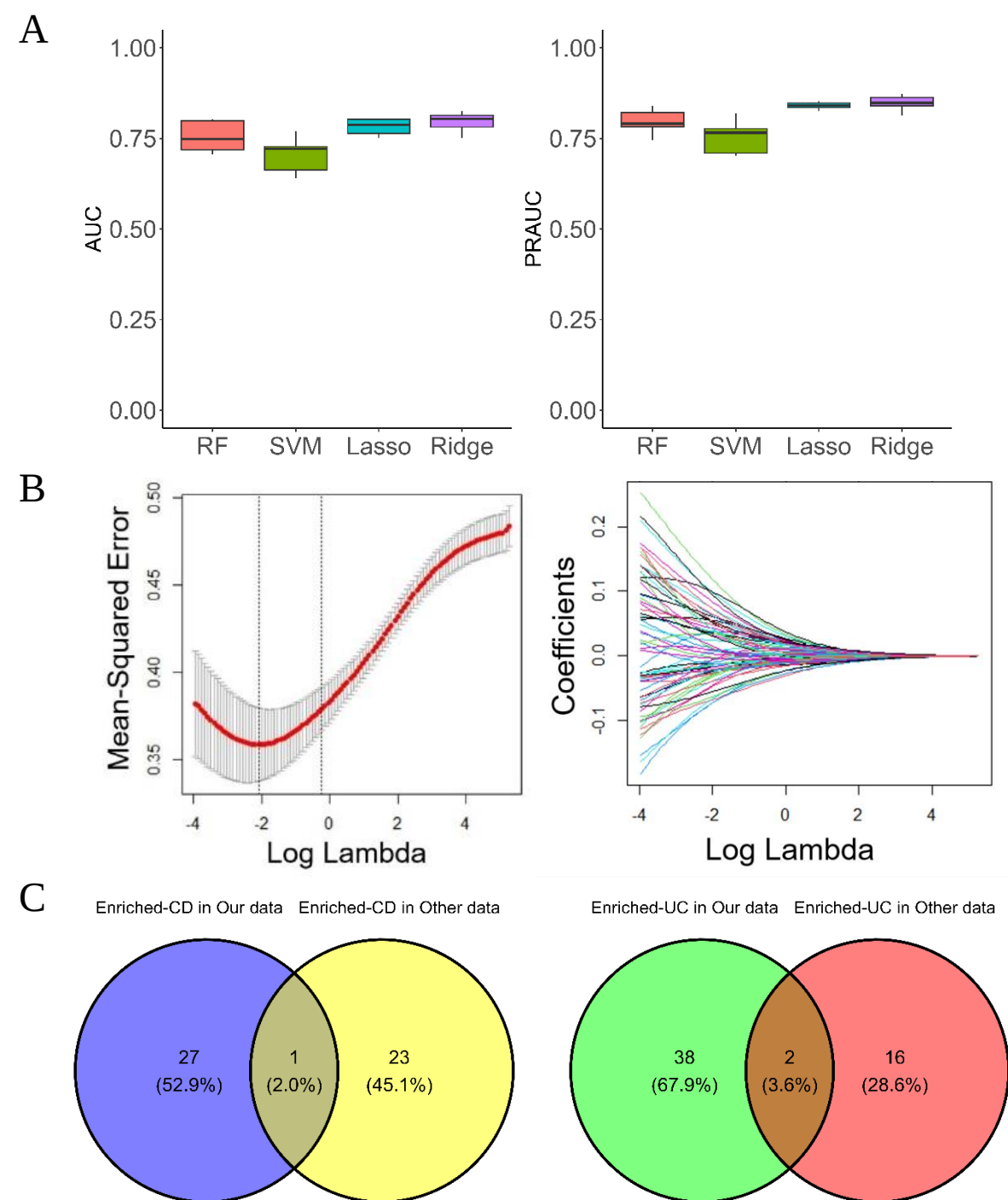


Fig. S4 (A) Comparison of four supervised machine learning models (RF, SVM, LASSO, and ridge) for the diagnosis of IBD subtypes of using differential abundance. (B) Parameter tuning of a supervised machine learning model for the classification of CD and UC individuals using gut microbiomes. A ten-fold cross-validation was performed to select the optimal lambda. The point where the mean square error (MSE) was minimized was designated as the best lambda ($\log [\lambda] = -2.088795$). (C) Venn diagram comparing our microbiome list with Clonney et al.[16]. RF, random forest; SVM, support vector machine; LASSO, least absolute shrinkage and selection operator.