Cell Reports Medicine, Volume *4*

Supplemental information

Integrative multi-omics deciphers the spatial characteristics of host-gut microbiota interactions in Crohn's disease

Xiang Gao, Ruicong Sun, Na Jiao, Xiao Liang, Gengfeng Li, Han Gao, Xiaohan Wu, Muqing Yang, Chunqiu Chen, Xiaomin Sun, Liang Chen, Wei Wu, Yingzi Cong, Ruixin Zhu, Tiannan Guo, and Zhanju Liu

Supplemental Figures

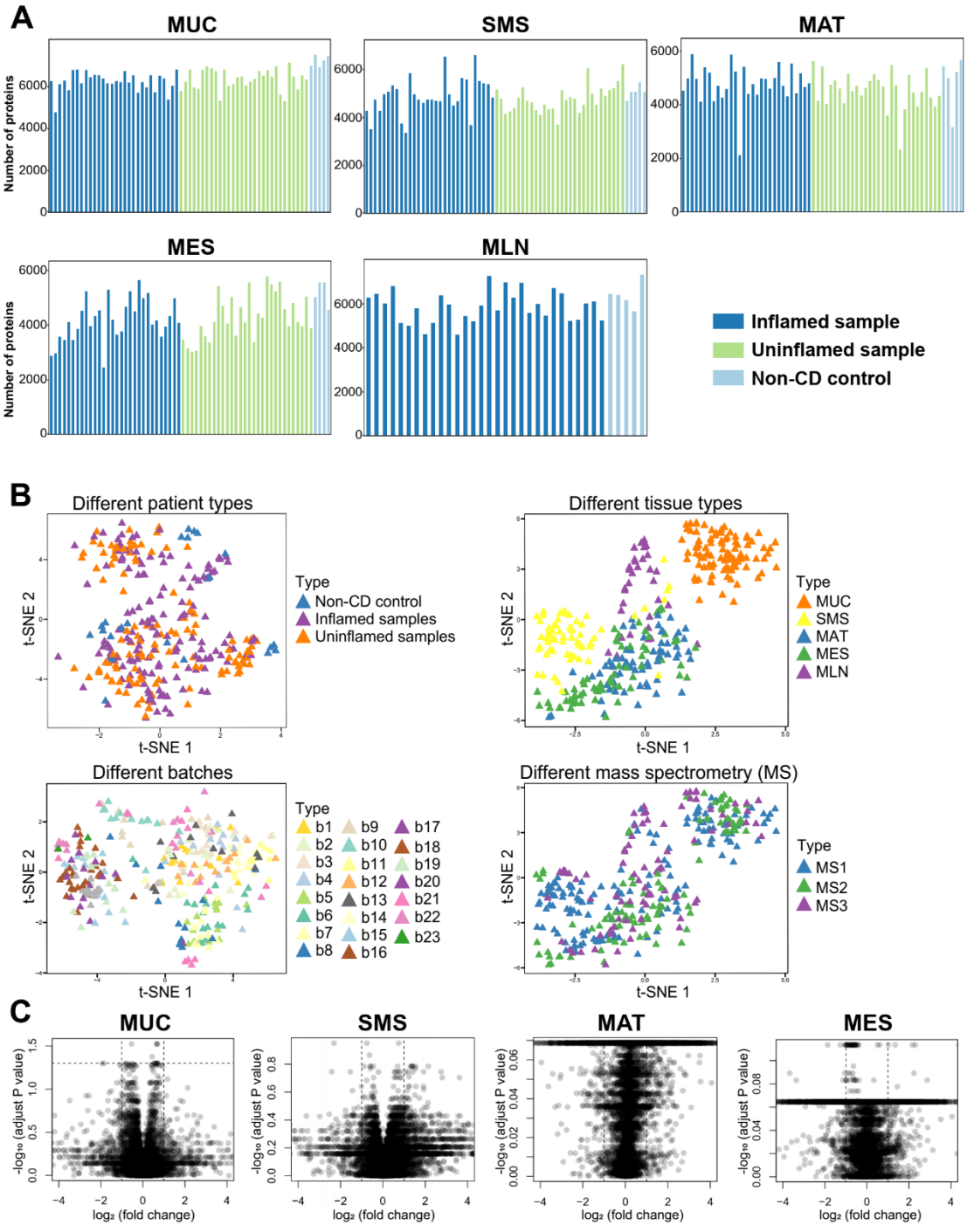

Figure S1. Quality control of proteomics data, Related to Figure 1

A. The bar plot shows the number of protein identifications in different tissues. Each column represents the number of the identified proteins in one sample of each tissue site. Dark blue represents the inflamed samples of CD patients. Light green represents the adjacent uninflamed samples of CD

patients. Light blue represents normal samples of NCs.

B. The t-SNE analysis of samples from different patients (top left), different tissue types (top right), different batches (bottom left), and different mass spectrometry analyses (bottom right).

C. The volcano plots compare the inflamed tissues with adjacent uninflamed tissues from CD patients, as indicated in the plot. Proteins with $\log_2$(FC) beyond 0.25 or below -0.25 with an adjusted p-value ≤ 0.05 were considered significantly differential expression.

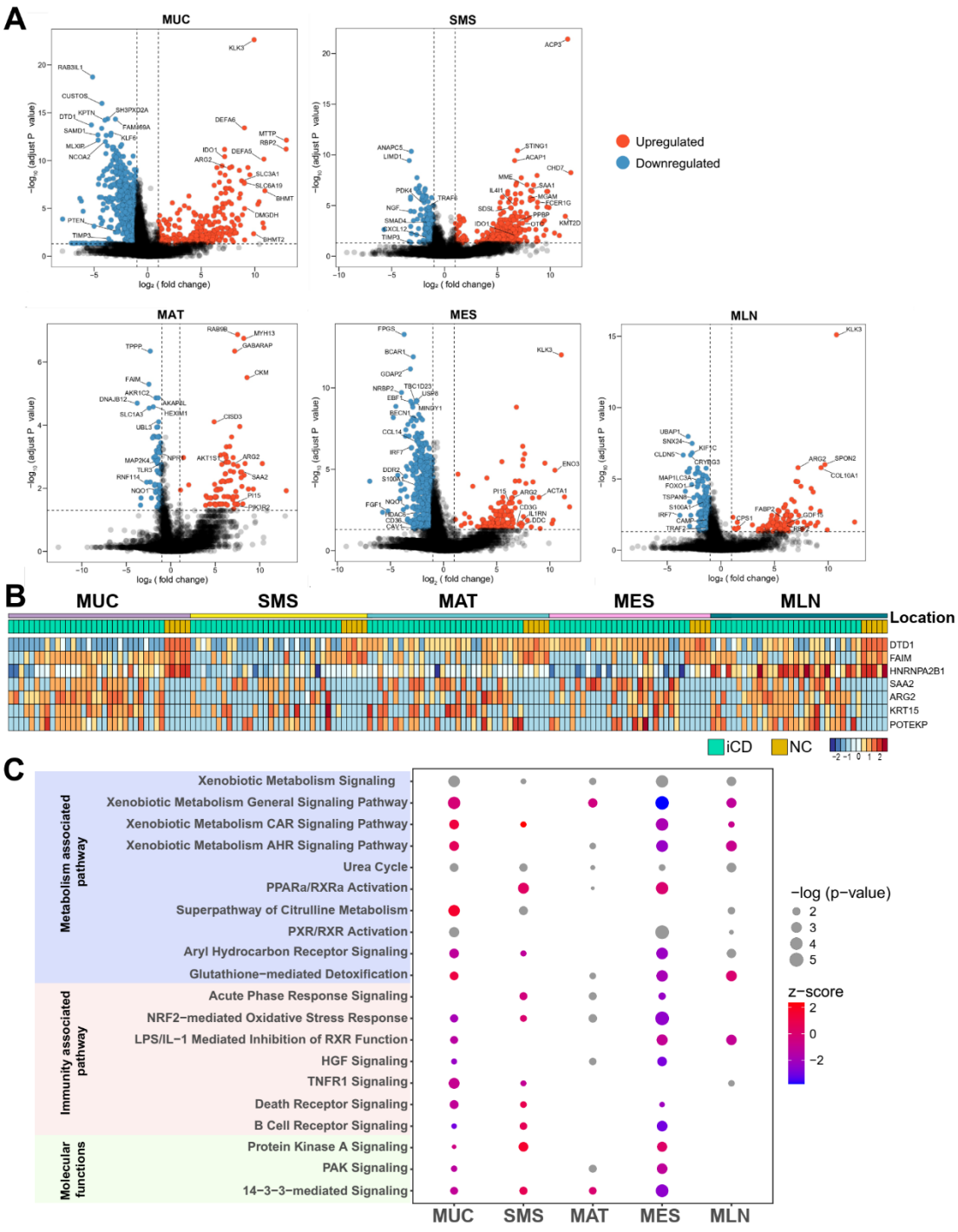

Figure S2. The landscape of dysregulated proteins in different tissues from CD patients and NCs, Related to Figure 2

A. The volcano plots compare inflamed tissues from CD patients with normal tissues from NCs, as indicated in the plot. Proteins with $\log_2$(FC) beyond 0.25 or below -0.25 with an adjusted p-value ≤ 0.05 were considered significantly differential expressions. Names of significantly down- (blue) and

up- (red) regulated proteins are shown in the plots.

B. The heatmap shows the dysregulated proteins in all five types of tissues (i.e., MUC, SMS, MAT, MES, and MLN) between CD patients and NCs. iCD, inflamed tissues of CD patients; NC, normal tissues of NCs.

C. The pathways are dysregulated across different types of tissues. Pathway analysis was performed using all dysregulated proteins in the specific tissues by IPA, and the most enriched pathways (adjusted p-value < 0.05) among the five inflamed tissues were shown. The size of the circles represents the $-\log_{10}$(p-value), and the color represents the Z score by IPA.

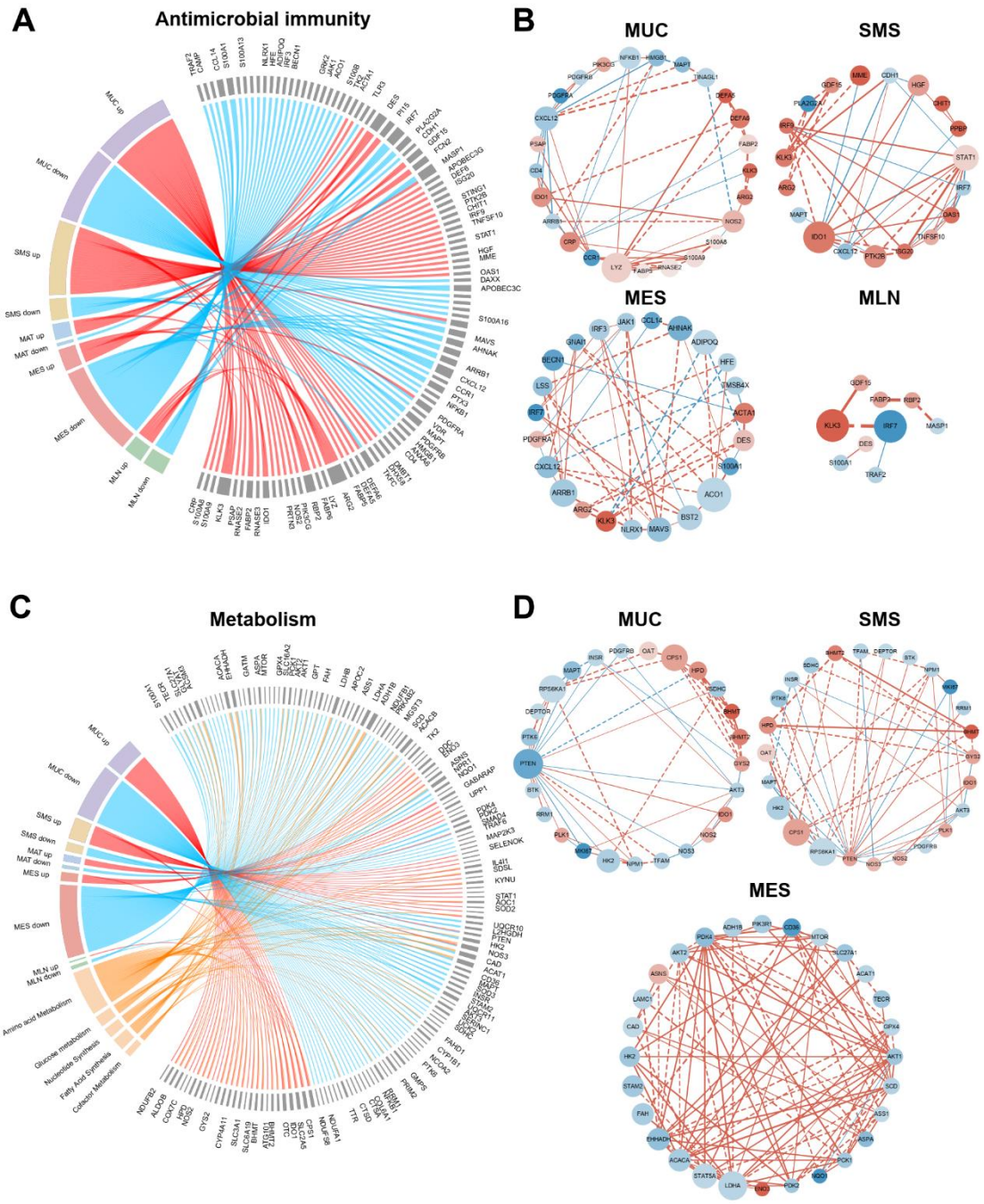

Figure S3. The landscape of dysregulated proteins involved in antimicrobial immunity and metabolism, Related to Figure 3

The chord diagrams show dysregulated proteins of antimicrobial immunity (A) and metabolism (C) in different tissues between CD patients and NCs. The length of the brick representing each protein corresponds to the sum of absolute $\log_2$(FC) in different types of tissues, and the length of the brick

representing each type of tissue corresponds to the sum of absolute $\log_2$(FC) in one or more proteins. The protein-protein interaction networks were generated from the dysregulated proteins of antimicrobial immunity (B) and metabolism (D) in different tissues. Red circles, upregulated proteins; blue circles, downregulated proteins; red lines, positive interaction; blue lines, negative interaction. The solid lines represent the interactions analyzed by the String database, and the dashed lines represent the interactions analyzed by our proteomics data. The circle size represents the absolute $\log_2$(FC) of each protein.

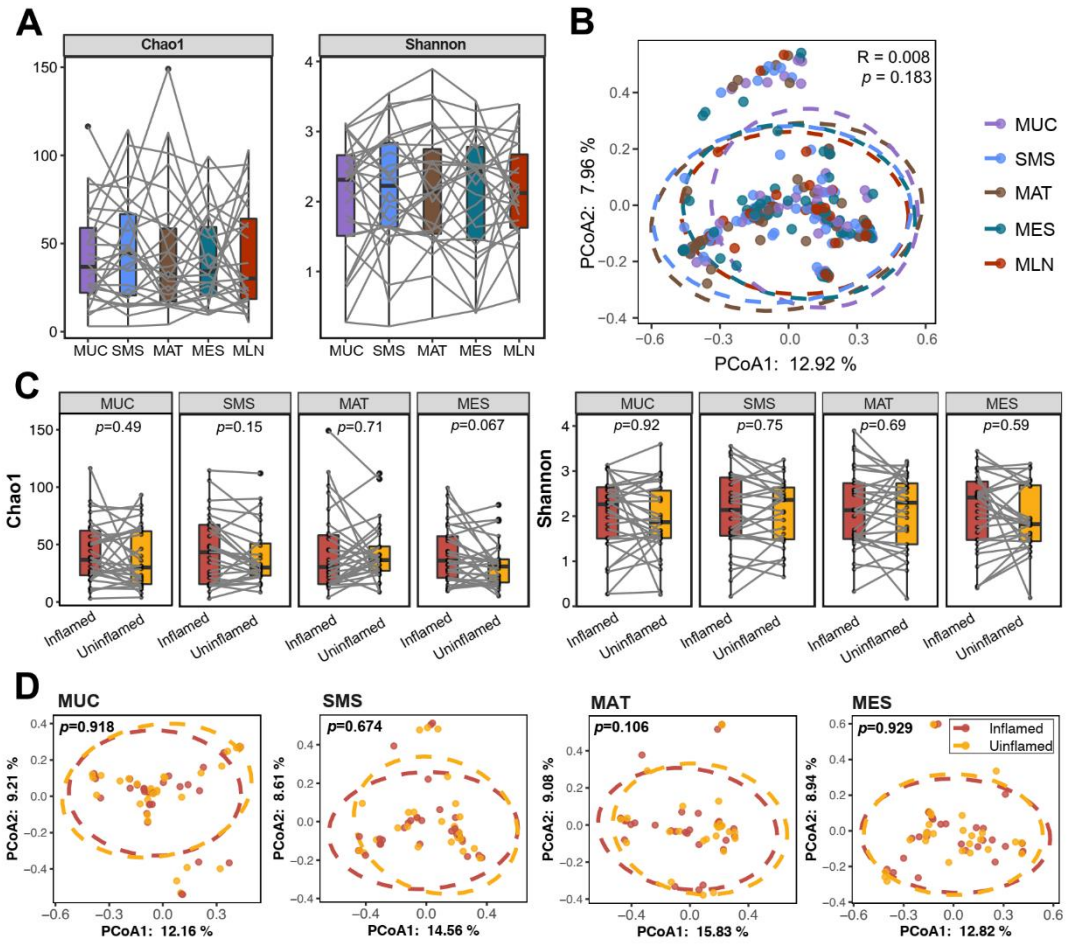

Figure S4. Microbial sequencing and compositional alterations across different tissues of CD patients, Related to Figure 4

A. Boxplots show the alpha diversity measured by Chao1 and Shannon index between inflamed (red) and adjacent uninflamed (orange) MUC, SMS, MAT, MES, and MLN of CD patients. All boxplots represent the 25th–75th percentile of the distribution, and the median is shown in the thick line in the middle of the box. The whiskers indicate the range of 1.5-fold IQR. The p values are calculated by paired Wilcoxon test.

B. Principal coordinate analysis of samples from inflamed tissues (red) and adjacent uninflamed tissues (orange) of CD patients based on the unweighted Unifrac distance in MUC, SMS, MAT, and MES,

respectively. The p values of beta diversity based on the unweighted Unifrac distance are calculated with PERMANOVA by 999 permutations.

C. Boxplots show the alpha diversity measured by Chao1 and Shannon index in inflamed MUC, SMS, MAT, MES, and MLN of CD patients. All boxplots represent the 25th–75th percentile of the distribution, and the median is shown in the thick line in the middle of the box. The whiskers indicate the range of 1.5-fold IQR. The p values were calculated by a two-tailed Wilcoxon test.

D. Principal coordinate analysis of samples from inflamed MUC, SMS, MAT, MES, and MLN of CD patients based on the unweighted Unifrac distance. The p values of beta diversity based on the unweighted Unifrac distance were calculated with PERMANOVA by 999 permutations. Tissue sites are labeled with different colors.

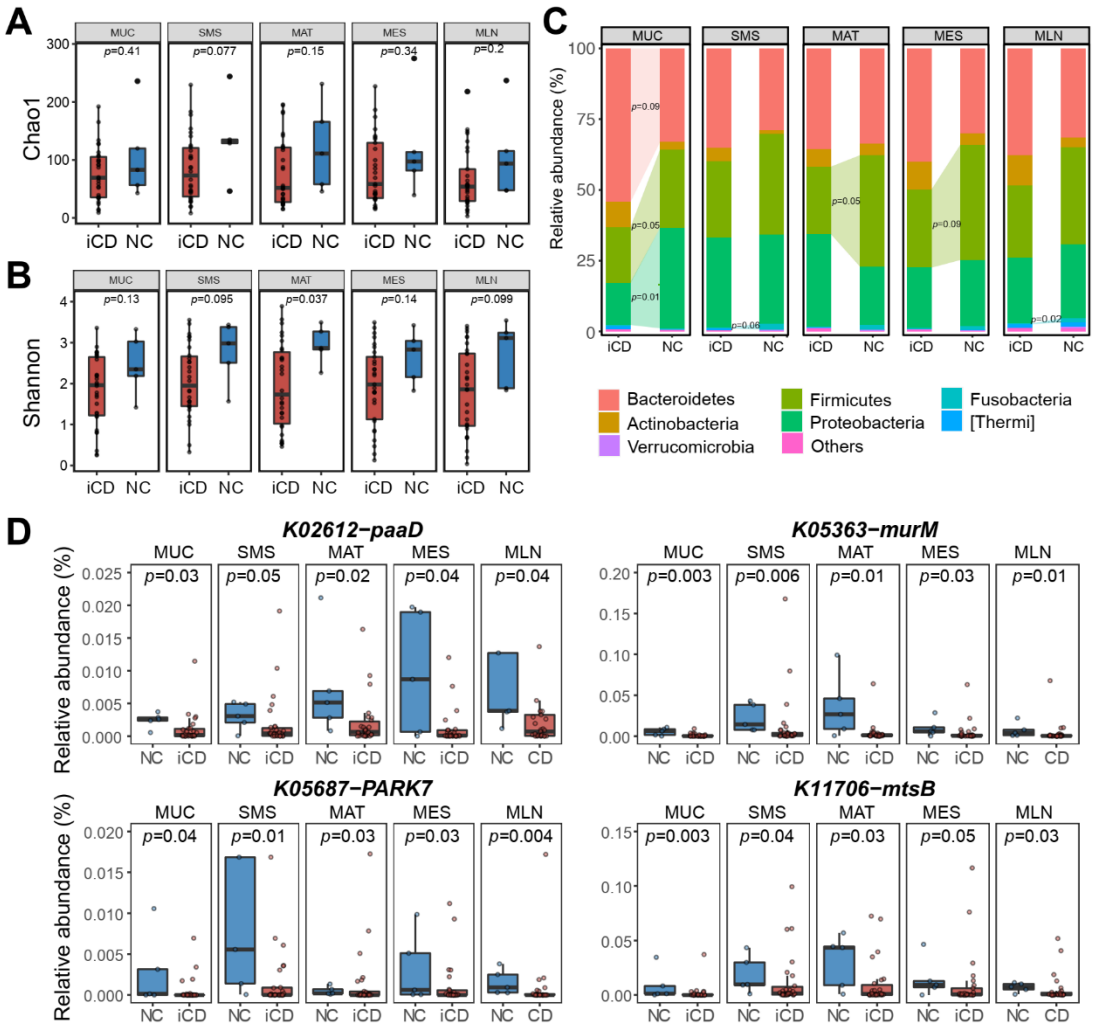

Figure S5. Microbial composition and functional alterations between inflamed tissues of CD patients and normal tissues of NCs, Related to Figure 4

A. Boxplots show the alpha diversity measured by Chao1 between inflamed tissues of CD patients and normal tissues of NCs. All boxplots represent the 25th–75th percentile of the distribution, and the median is shown in the thick line in the middle of the box. The whiskers indicate the range of 1.5-fold IQR, and the outliers are represented as dots. The p values were calculated by a two-tailed Wilcoxon test.

B. Boxplots show the alpha diversity measured by the Shannon index between inflamed tissues of CD patients and normal tissues of NCs. All boxplots represent the 25th–75th percentile of the

distribution, and the median is shown in the thick line in the middle of the box. The whiskers indicate the range of 1.5-fold IQR, and the outliers are represented as dots. The p values were calculated by a two-tailed Wilcoxon test.

C. Microbial compositions of inflamed tissues from CD patients (CD) and normal tissues from NCs at the phylum level. Only the abundant phyla are shown in the stacked bar plot, and the rare phyla are summed into others. p values were calculated by the two-tailed Wilcoxon test.

D. Boxplots show the relative abundance of differential KO genes between inflamed tissues of CD patients (red) and normal tissues of NCs (blue) in MUC, SMS, MAT, MES, and MLN, respectively. All boxplots represent the 25th–75th percentile of the distribution, and the median is shown in the thick line in the middle of the box. The whiskers indicate the range of 1.5-fold IQR, and the outliers are represented as dots. The p values were calculated by the two-tailed Wilcoxon test.

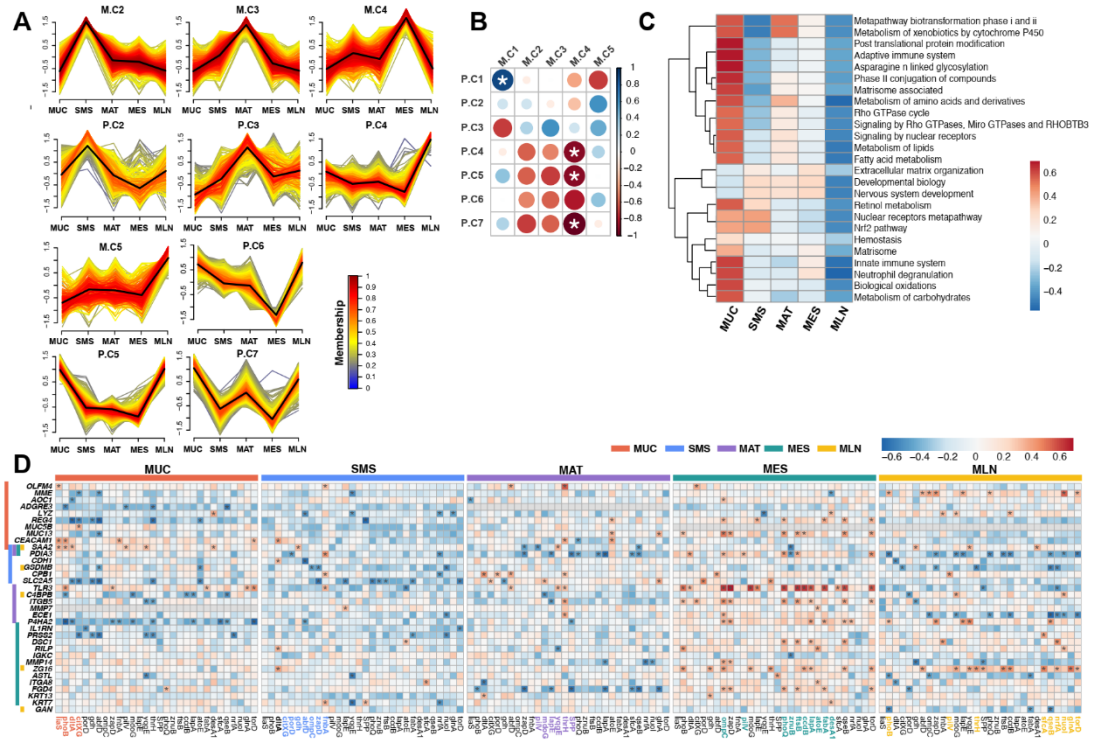

Figure S6. Correlations between host proteins and microbial functional genes, Related to Figure 6

A. The mfuzz clustering of differential host proteins and gut microbial genes across different tissues of CD patients. Membership scores indicate the degree to which the microbial gene or host protein belongs in each cluster. M.C., clusters based on microbial genes; P.C., clusters based on proteins.

B. The heatmap shows correlations by spearman rank correlation analysis between host protein clusters and microbial gene clusters. Circles labeled by stars represent significant correlations. P.C, protein clusters; M.C, microbial genes clusters.

C. Gene set variation enrichment analysis of host P.C1 across each tissue site.

D. The heatmap shows the spearman rank correlation between key proteins in P.C1 and microbial genes in M.C1. The star labeled in cells represents significant correlations (permutation p-value < 0.05), and the colors of cells indicate the degree of correlations. Tissue sites are labeled with different colors. MUC, red; SMS, blue; MAT, purple; MES, green; MLN, orange.

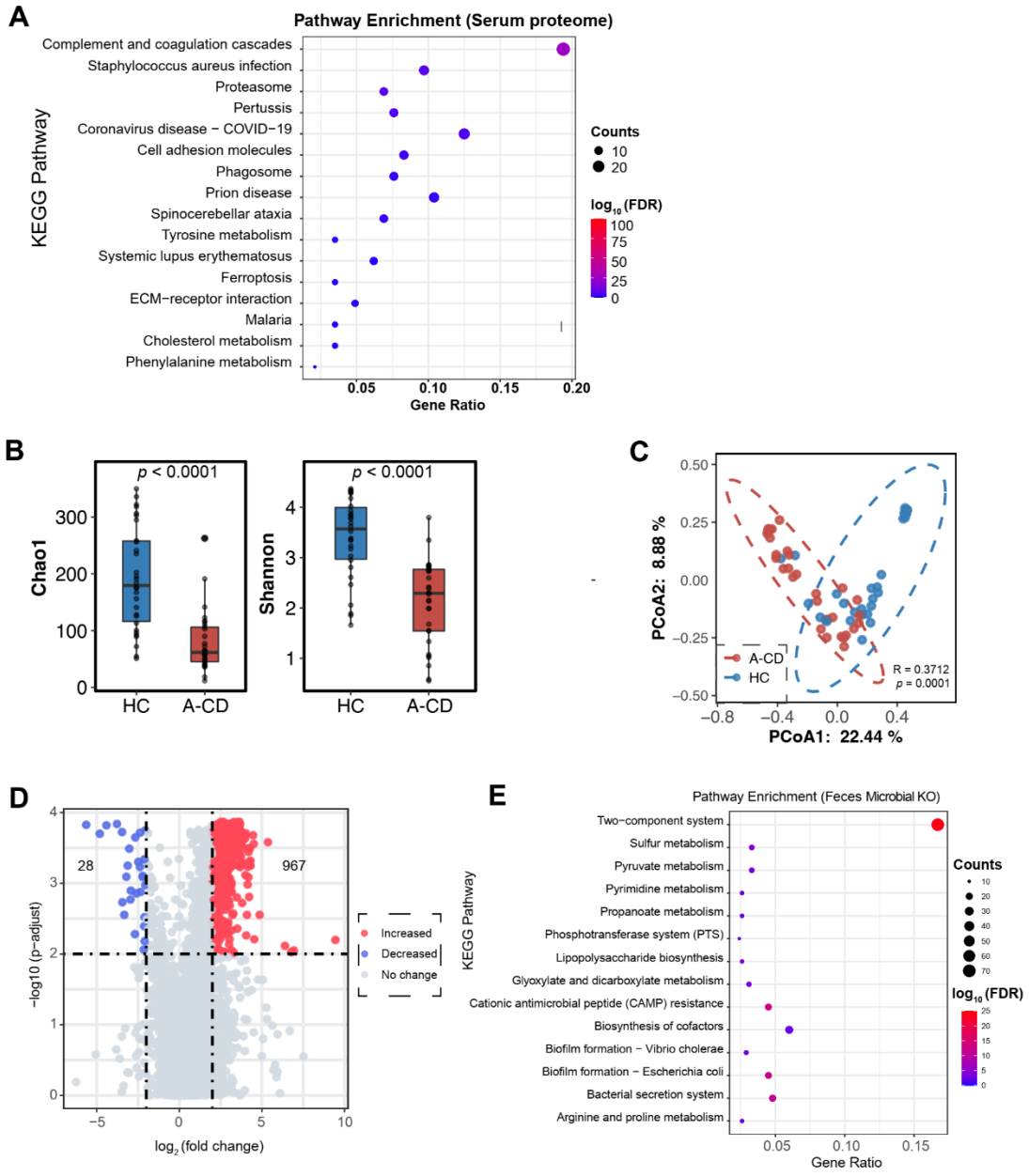

Figure S7. Alterations in host serum proteome and fecal microbiome of CD patients, Related to Figure 7

A. The KEGG pathway enrichment analysis of differential proteins identified in serum samples. The dot size corresponds to the enriched gene count, and the dot color represents the enrichment -$\log_{10}$(FDR).

B. Alterations in microbial alpha diversity were measured by Chao1 and Shannon index between healthy donors (HC, blue) and active CD patients (A-CD, red).

C. Alterations in microbial beta diversity were measured by the Bray-Curtis distance between healthy donors (HC) and active CD patients (A-CD). The p values were calculated with PERMANOVA by 999 permutations.

D. The volcano plot shows the differential microbial genes between healthy donors (HC) and active CD patients (A-CD).

E. The KEGG pathway enrichment analysis of differential microbial genes identified in fecal samples. The dot size corresponds to the enriched gene count, and the dot color represents the enrichment - $\log_{10}$(FDR).