

Metagenomics reveals an increased proportion of an *Escherichia coli*-dominated enterotype in elderly Chinese people

Jinyou LI^{1*}, Yue WU^{1*}, Yichen YANG¹, Lufang CHEN¹, Caihong HE¹, Shixian ZHOU¹, Shunmei HUANG¹, Xia ZHANG¹, Yuming WANG¹, Qifeng GUI¹, Haifeng LU², Qin ZHANG^{1✉}, Yunmei YANG^{1✉}

¹Zhejiang Key Laboratory for Diagnosis and Treatment of Physic-chemical and Aging-related Injuries, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China

²State Key Laboratory for Diagnosis and Treatment of Infectious Disease, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China

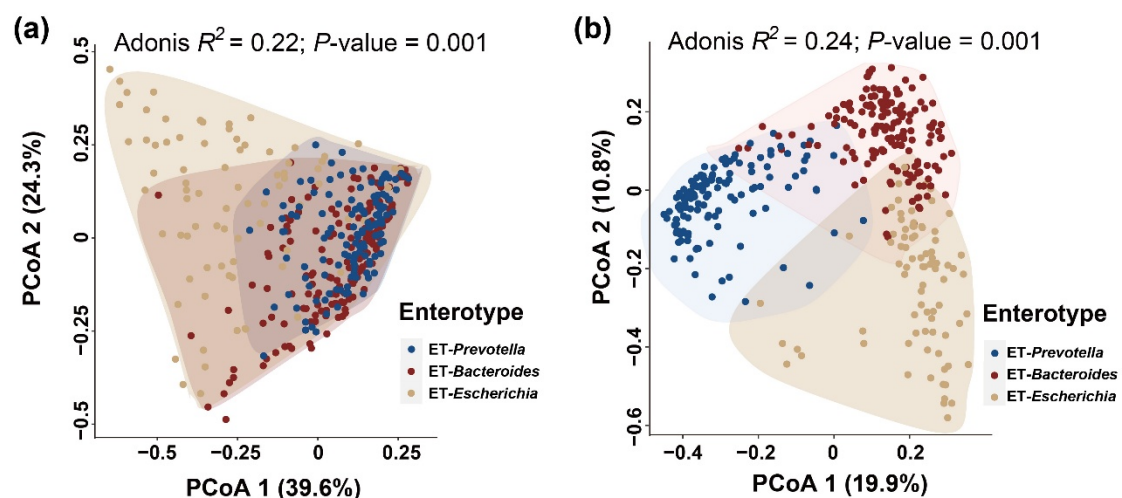


Fig. S1 Beta diversity differences support for the classification of three enterotype. Principal coordinate analysis (PCoA) based on Bray-Curtis dissimilarity metrics showed clustering of samples at the order level (a) and at the species level (b).

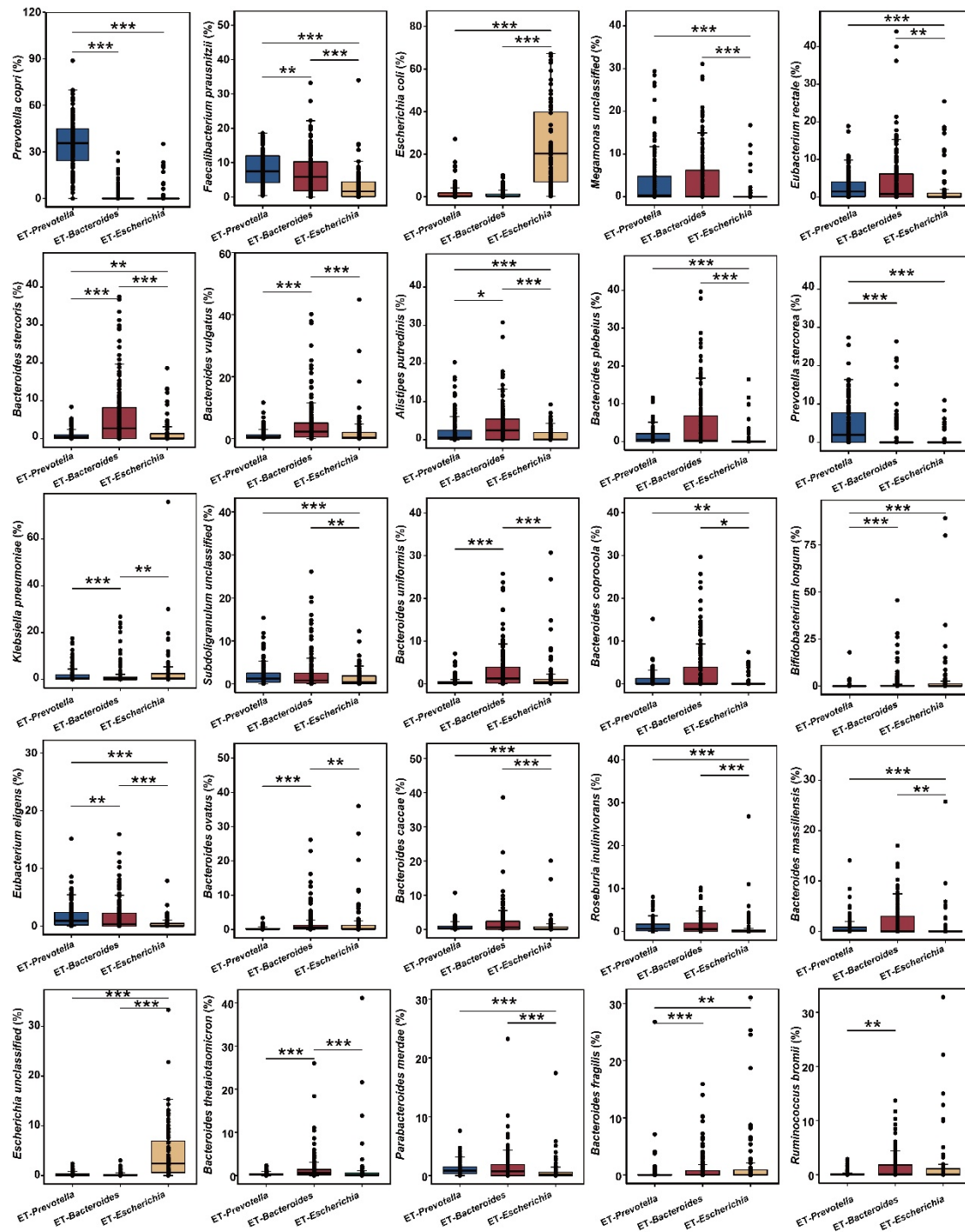


Fig. S2 Significant difference in the relative abundances of the top 25 most dominant species among the three enterotypes. The average relative abundance of these species was over 1% in at least one enterotype. The means, ranges and first and third quartiles are shown. Wilcoxon rank-sum test: ** $P < 0.01$; *** $P < 0.001$.

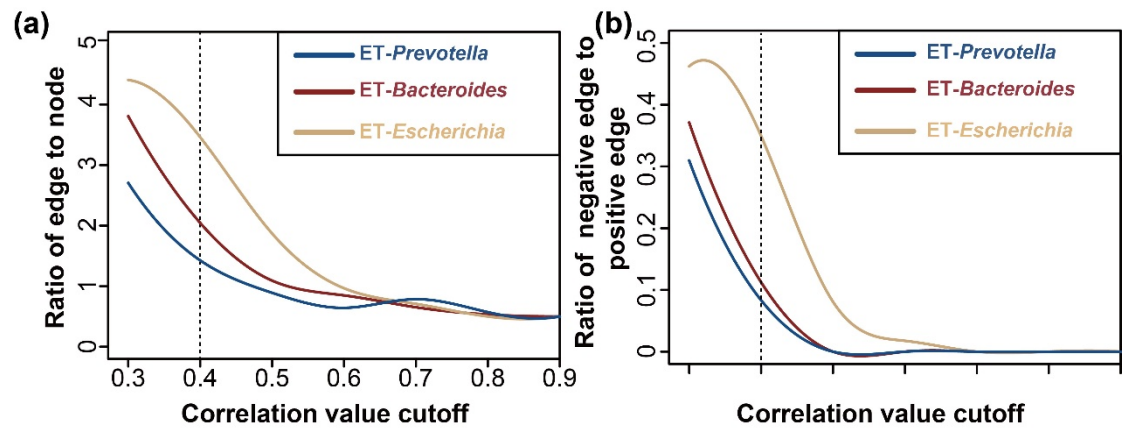


Fig. S3 Edges and nodes in each co-occurrence network. Ratio of edges to nodes (a) and ratio of negative edges to positive edges (b) in each co-occurrence network under different correlation value cutoffs from 0.3 to 0.9.

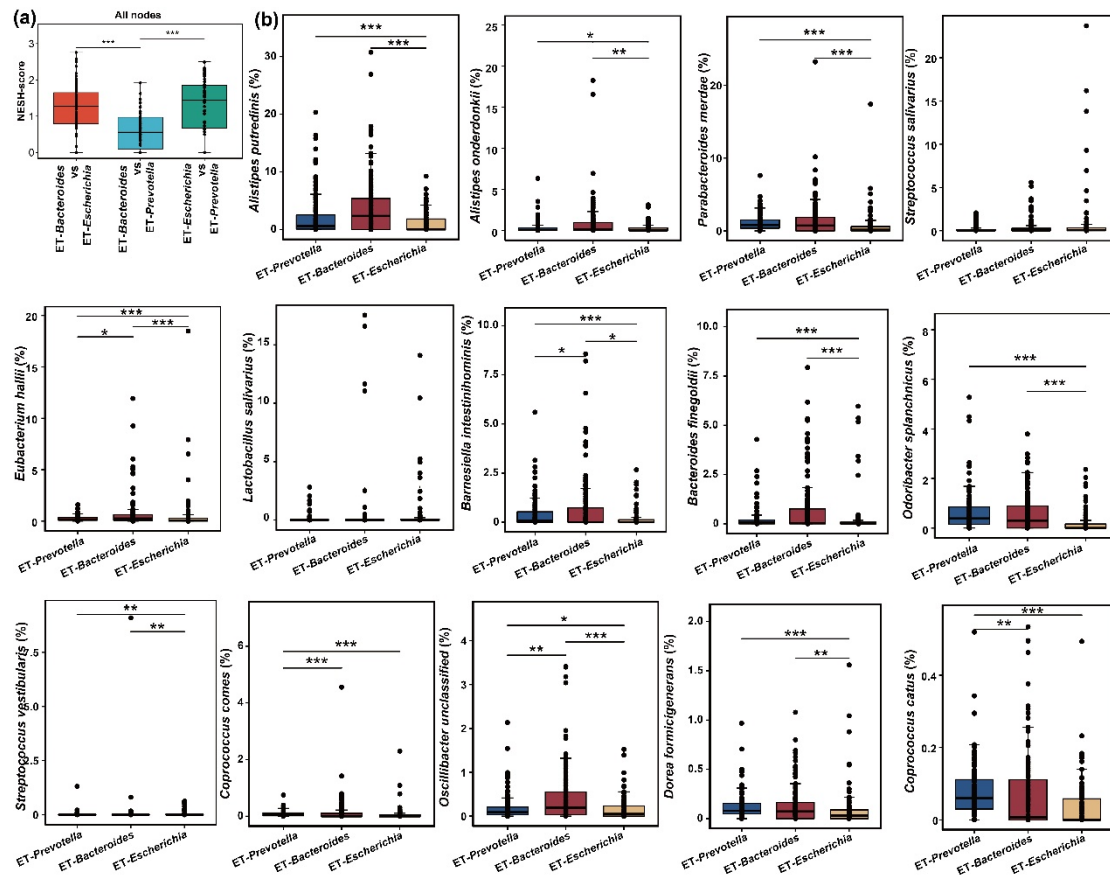


Fig. S4 Significant differences in the relative abundances of the selected biotic driver species among enterotypes. (a) Composition of NESH score values of nodes. The means, ranges, and first and third quartiles are shown. Wilcoxon rank-sum test: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. (b) Significant difference in the relative abundances of species identified by NetShift with NESH score values > 2 . The means, ranges, and first and third quartiles are shown. Wilcoxon rank-sum test: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

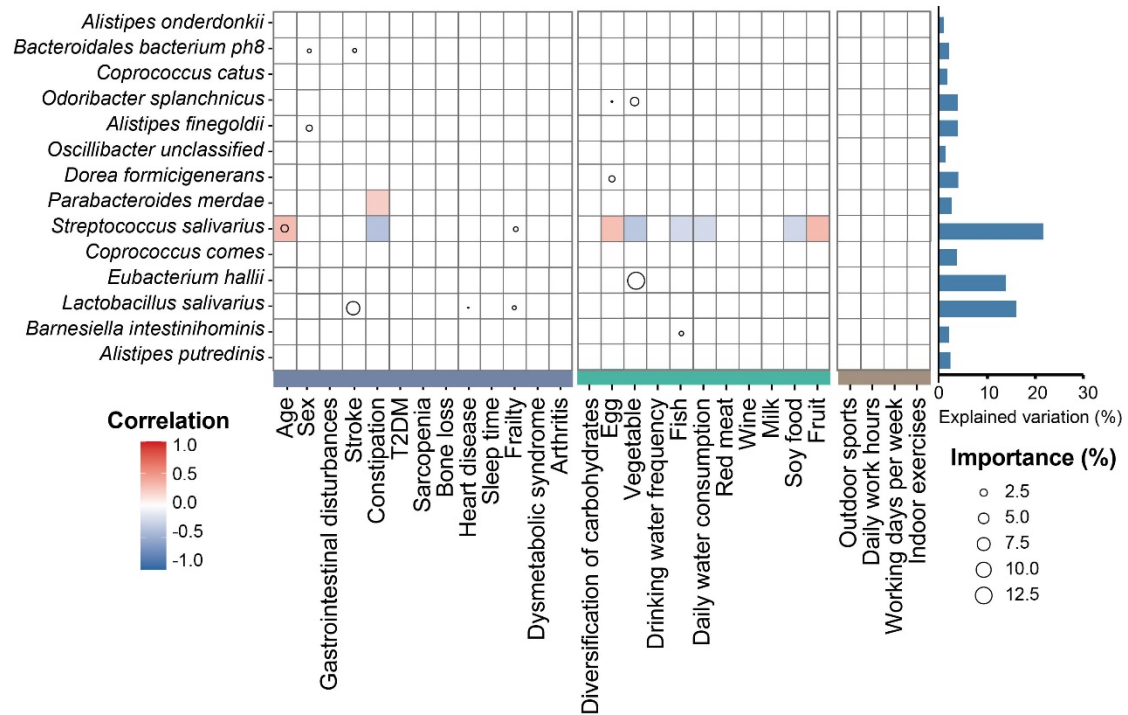


Fig. S5 Association between clinical and demographic variables and the relative abundances of the biotic driver species. The color bar shows correlation values, where red indicates a positive association, blue indicates a negative association, and only significant correlations are shown. Circle sizes represent the value of importance (%).

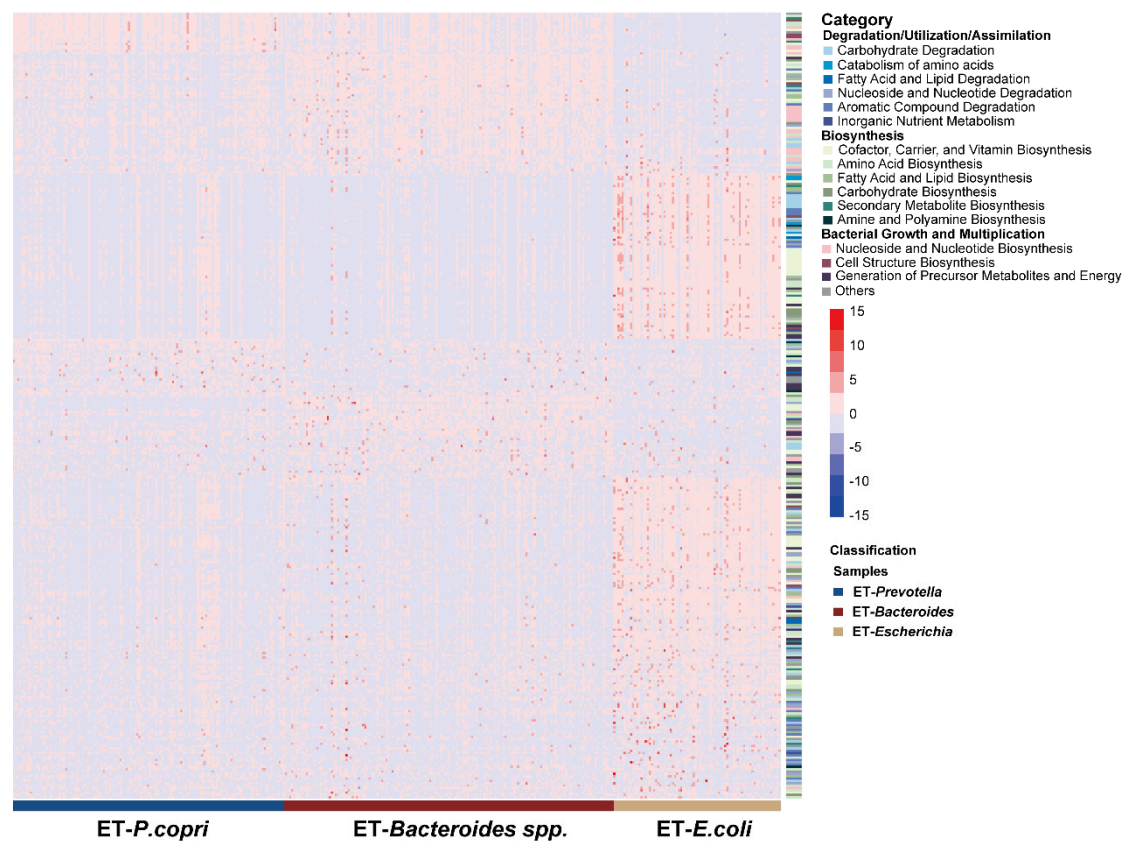


Fig. S6 Analysis of stool samples using shotgun metagenomics. The heatmap shows a clustering of functional pathways significantly differentially enriched in each enterotype classified by metabolic categories or functional modules calculated using the SPIEC-EASI (glasso) method.

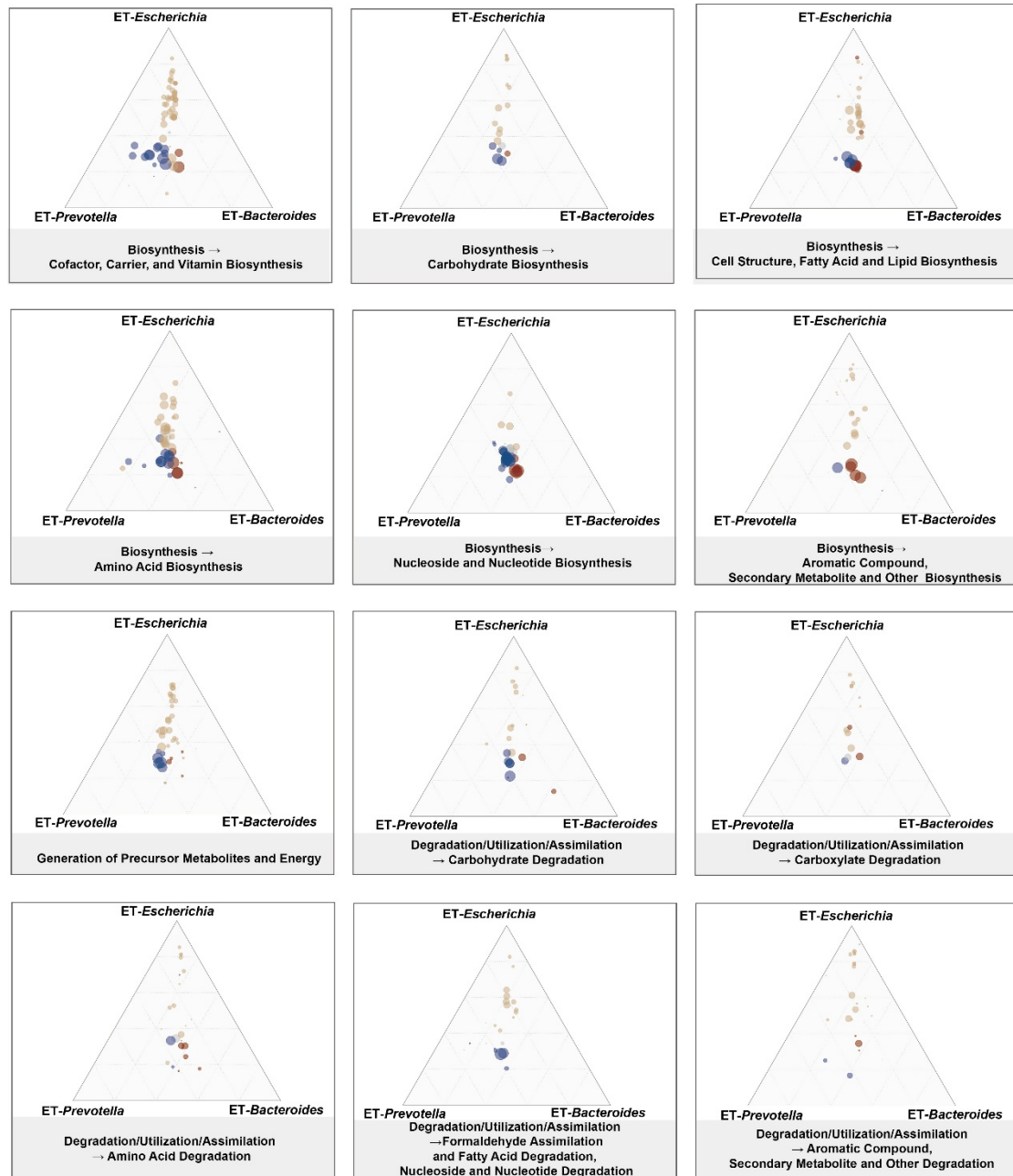


Fig. S7 Comparison of metabolic capacity between enterotypes. Dot sizes represent the mean relative abundances of pathways. Bacterial functions that were significantly enriched (FDR-corrected $P < 0.05$) in the comparison among enterotypes are indicated in color (red, *ET-Bacteroides*; gold, *ET-Escherichia*; blue, *ET-Prevotella*); pathways without significant differences in relative abundance are indicated in gray. Pathways from the same metabolic category are presented in each diagram. The closer the point is to the apex of the triangle, the higher the abundance of this pathway in the corresponding enterotype.