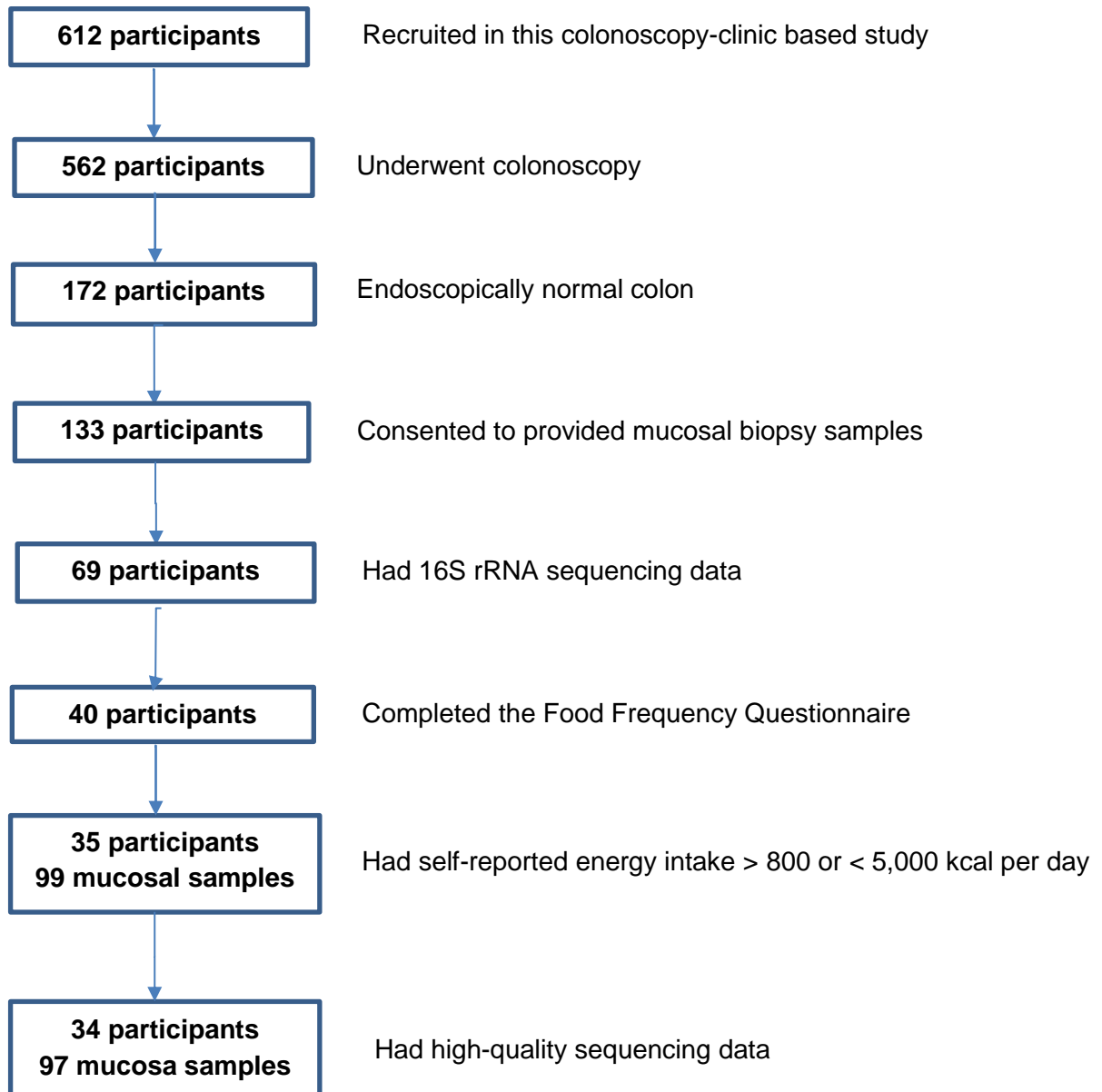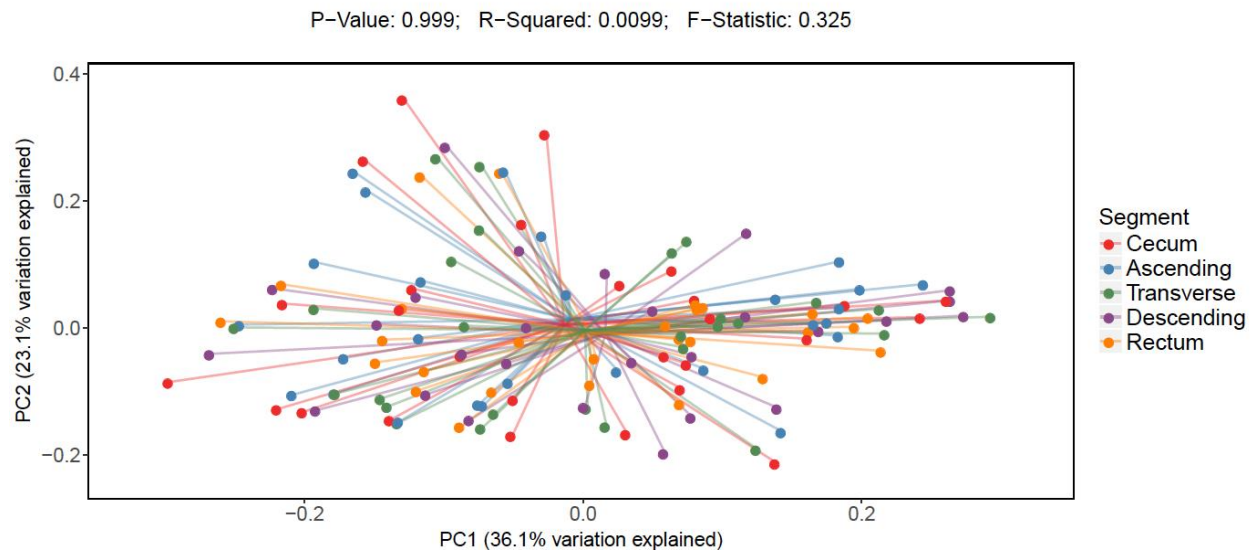**Dietary quality and colonic mucosa-associated gut microbiome in humans**

Yanhong Liu, Nadim J. Ajami, Hashem B. El-Serag, Clark Hair, David Y. Graham, Donna L. White, Liang Chen, Zhensheng Wang, Sarah Plew, Jennifer Kramer, Rhonda Cole, Ruben Hernaez, Jason Hou, Nisreen Husain, Maria E. Jarbrink-Sehgal, Fasiha Kanwal, Gyanprakash Ketwaroo, Yamini Natarajan, Rajesh Shah, Maria Velez, Niharika Mallepally, Joseph F. Petrosino, Li Jiao
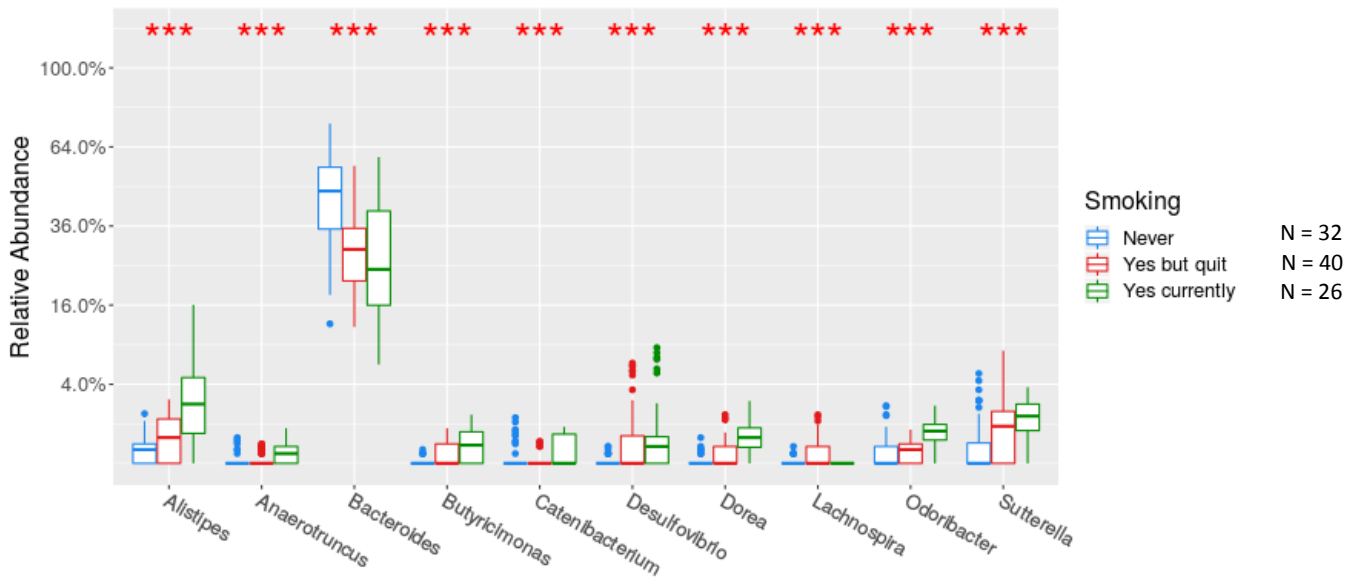
Supplemental Material

| 612 participants | Recruited in this colonoscopy-clinic based study |

↓

| 562 participants | Underwent colonoscopy |

↓

| 172 participants | Endoscopically normal colon |

↓

| 133 participants | Consented to provided mucosal biopsy samples |

↓

| 69 participants | Had 16S rRNA sequencing data |

↓

| 40 participants | Completed the Food Frequency Questionnaire |

↓

| 35 participants
99 mucosal samples | Had self-reported energy intake > 800 or < 5,000 kcal per day |

↓

| 34 participants
97 mucosa samples | Had high-quality sequencing data |

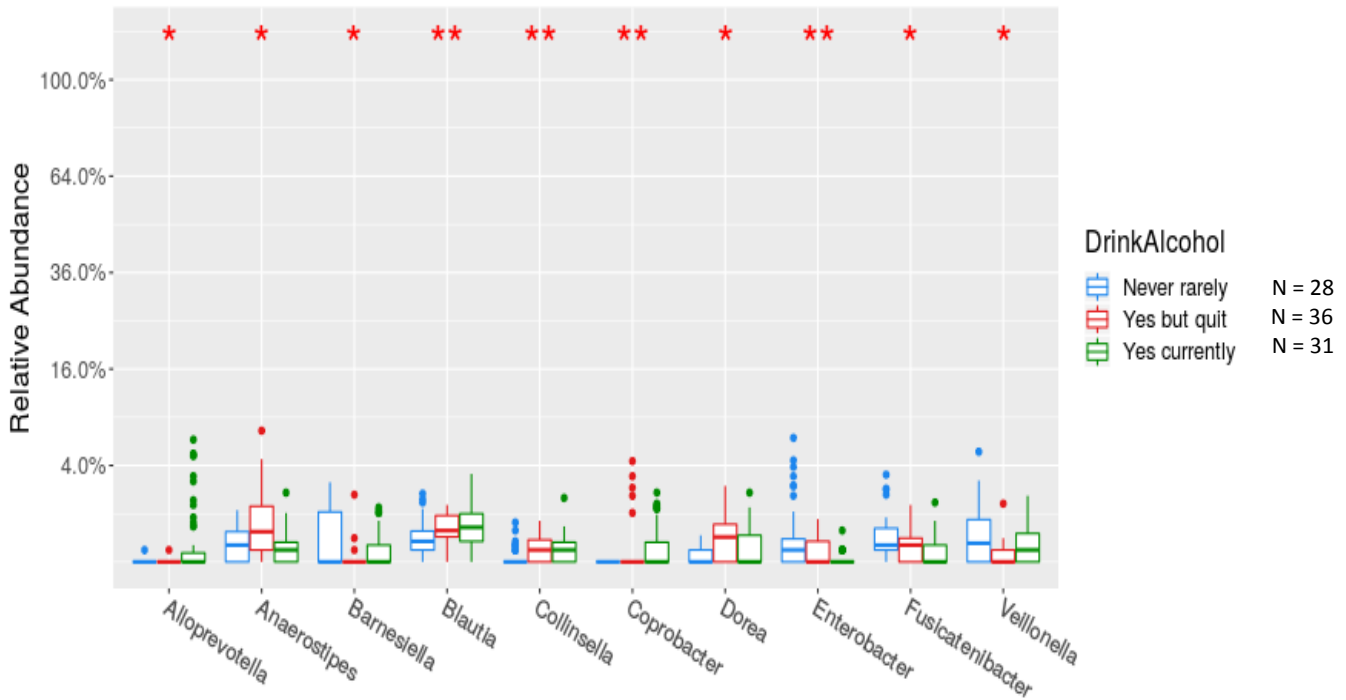**Supplemental figure 1**: Flow chart of sample selection in our study

Supplemental Figure 2. Beta-diversities of the mucosa-associated gut bacterial communities (based on colonic biopsies collected from five segments of the colon of 27 individuals).
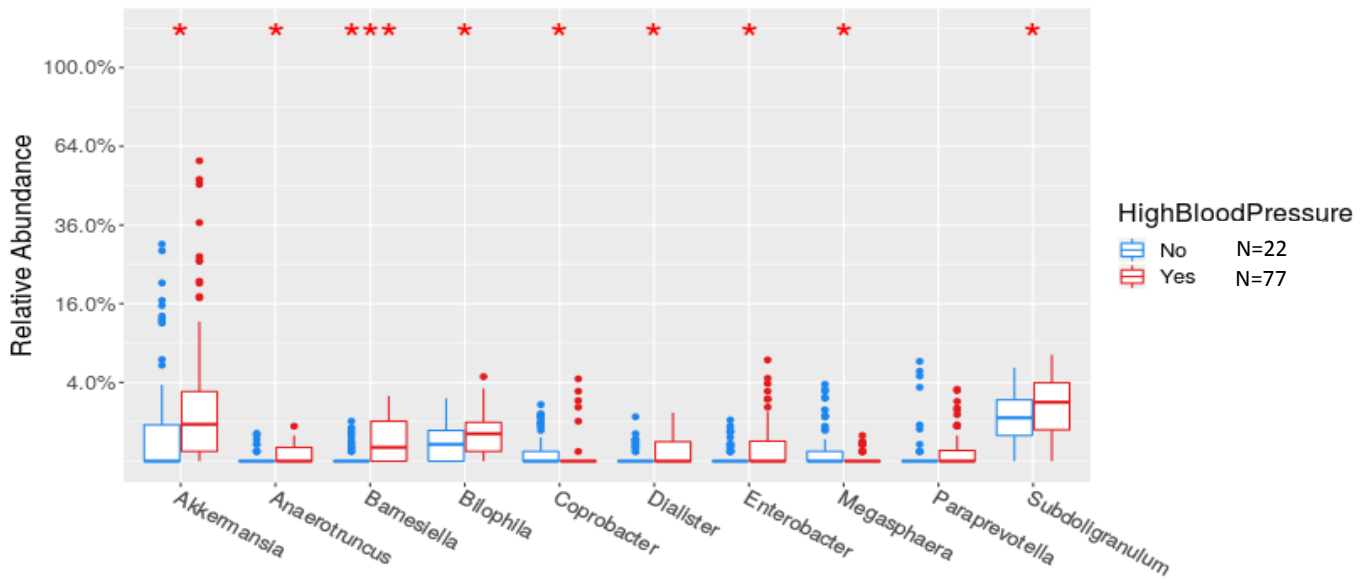
Supplemental Material
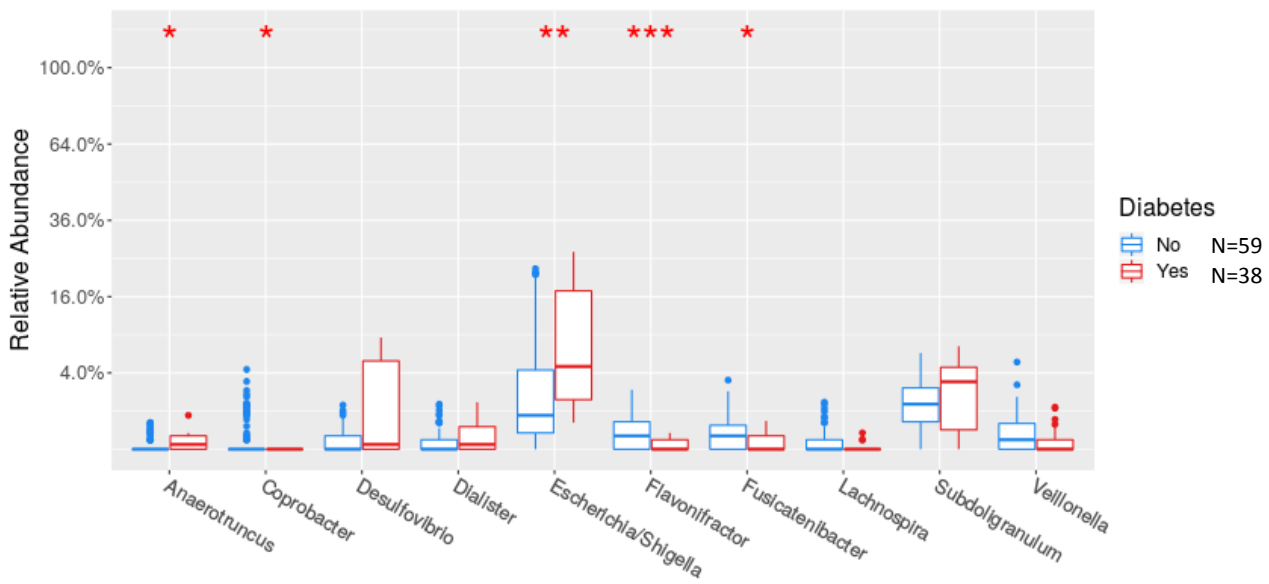

Supplemental figure 3 – A
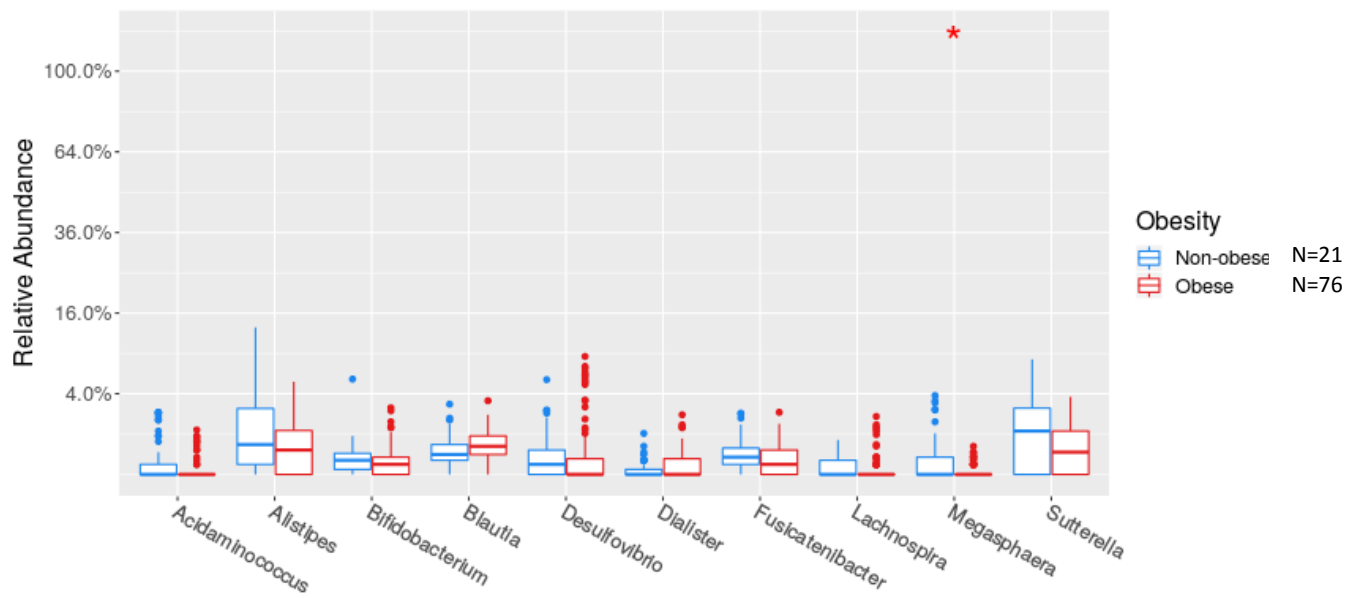

Supplemental figure 3 – B

Supplemental Material

Supplemental figure 3 – C
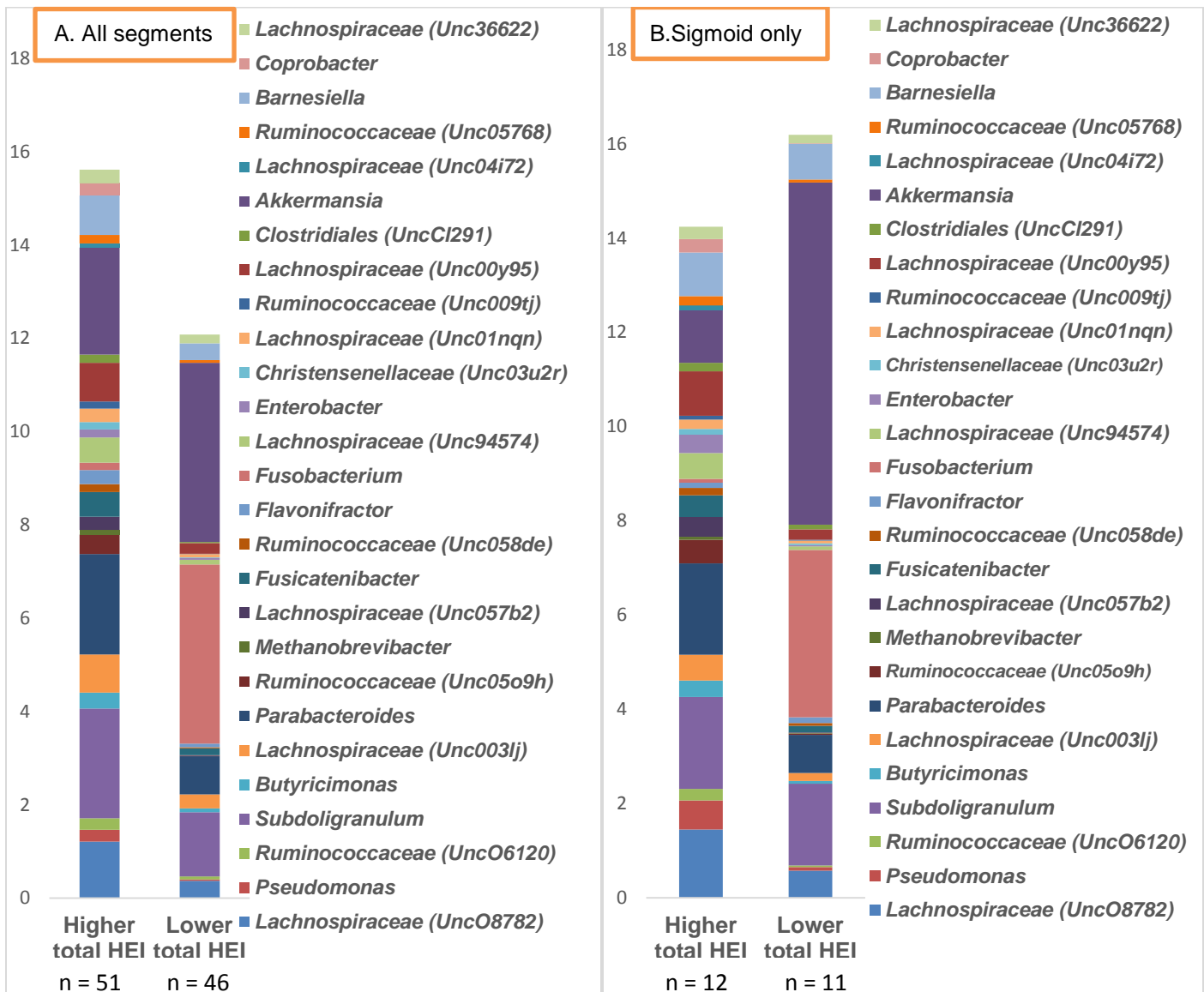

Supplemental figure 3 – D

Supplemental figure 3 – E

Supplemental Figure 3. Relative abundance (%) of mostly differed bacterial genera (rank by *P* values) by smoking status (A), alcohol use (B), hypertension (C), diabetes (D), and obesity (E).

*FDR *P* value < 0.05; ** FDR *P* value < 0.005; *** FDR *P* value < 0.0005
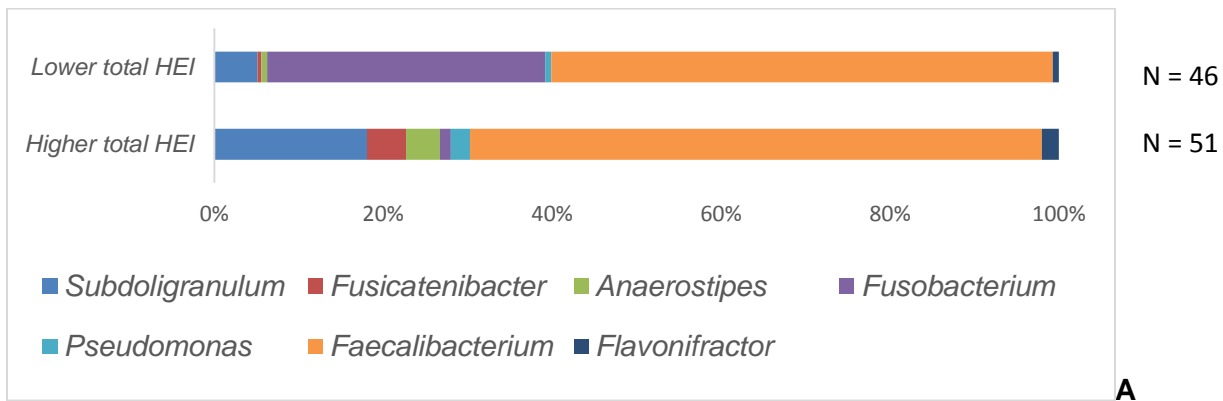
Supplemental Material

Supplemental Figure 4. Stacked bar chart of the relative abundance (Y axis) of the bacterial genera according higher vs. lower total HEI score using the sequencing data from all segments (97 biopsies from 34 participants) (panel A) and from sigmoid only (23 biopsies from 23 participants). Only the genera significantly differed in all segments (FDR-adjusted *P* values < 0.05) were shown in the figure. Bacteria were ordered by *P* values. The bottom Lachnospiraceae (UncO8782) had the smallest adjusted *P* value and the Lachnospiraceae (UncO8782) had the largest *P* value.
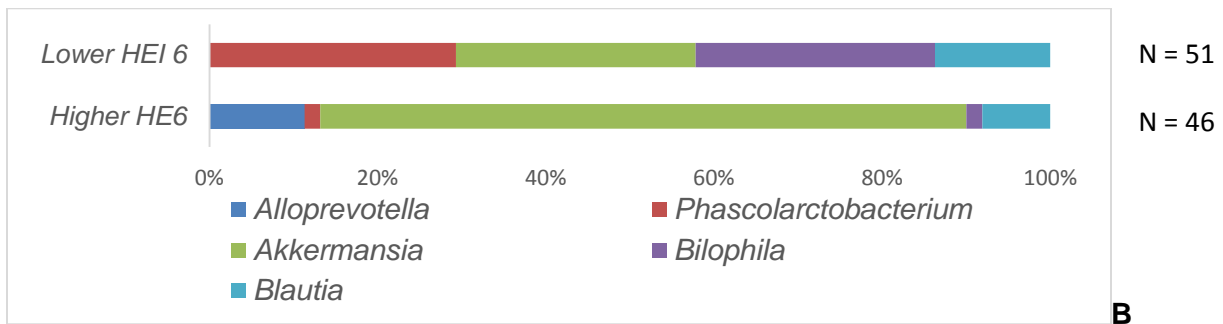
Y axis: relative abundance (%)

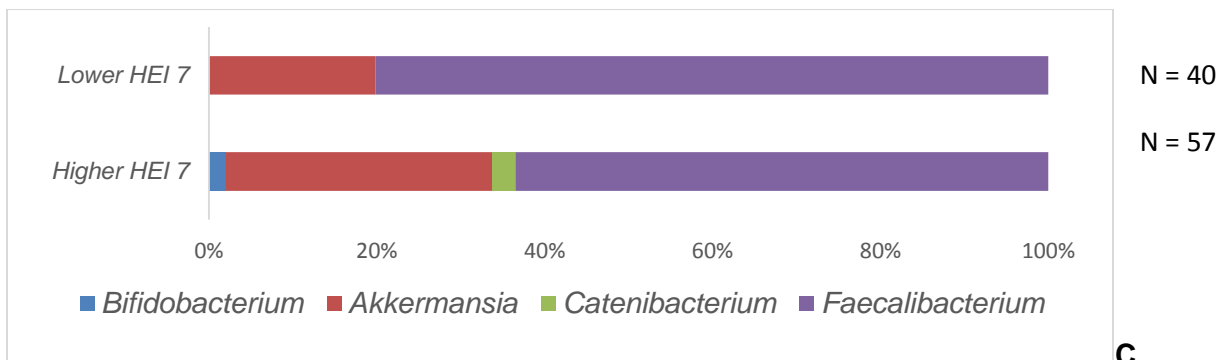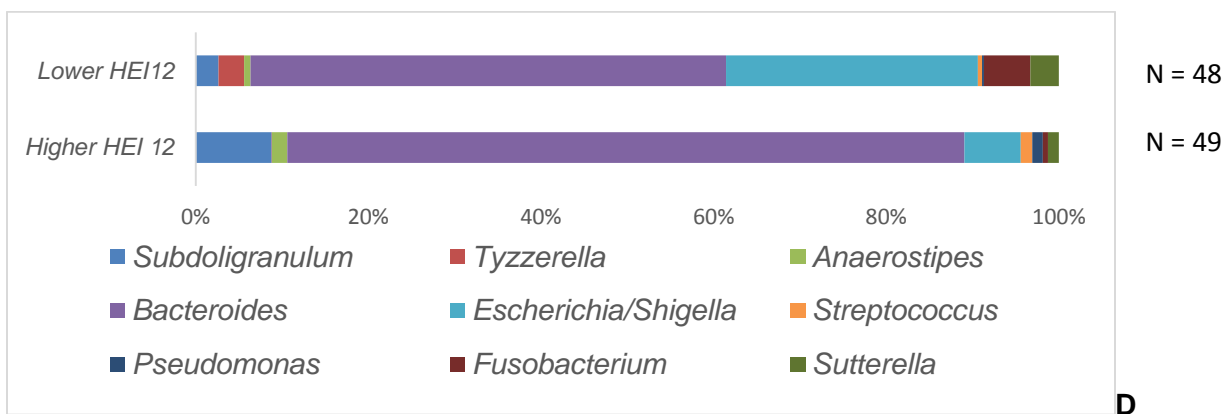Abbreviations: HEI: healthy eating index; Unc: unclassified.

Supplemental Figure 5. Stacked bar chart of the relative abundance (%, X axis) of the bacterial genera according higher *vs.* lower total HEI score (panel A), whole grains (panel B), milk products and soy beverages (panel C), and SoFFAS (panel D) using the sequencing data from the single sequencing batch (56 biopsies

from 13 participants). Only the genera significantly differed by dietary quality was shown (FDR-adjusted *P* values < 0.05) were shown in this figure. Unclassified genera were not shown in the figure. Bacteria were ordered by *P* values. The bacteria on the left side had the smallest adjusted *P* value and the bacterium on the right side had the largest *P* value.

X axis: relative abundance (%).

Abbreviations: HEI: healthy eating index; SoFAAS: Calories from solid fats, alcoholic beverages, and added sugars