

Aerobic bacterial group as an early-stage biomarker from fecal samples of patients with colorectal cancer without distant metastasis

Donghyoun Lee, Kung Ahn, Kyeongui Yun, Yunseok Oh, Young Suk Park, Yong Sung Kim, Jeong-An Gim, Seyoung Mun, Jae-Woo Mun, Kyudong Han and Yong Ju Ahn

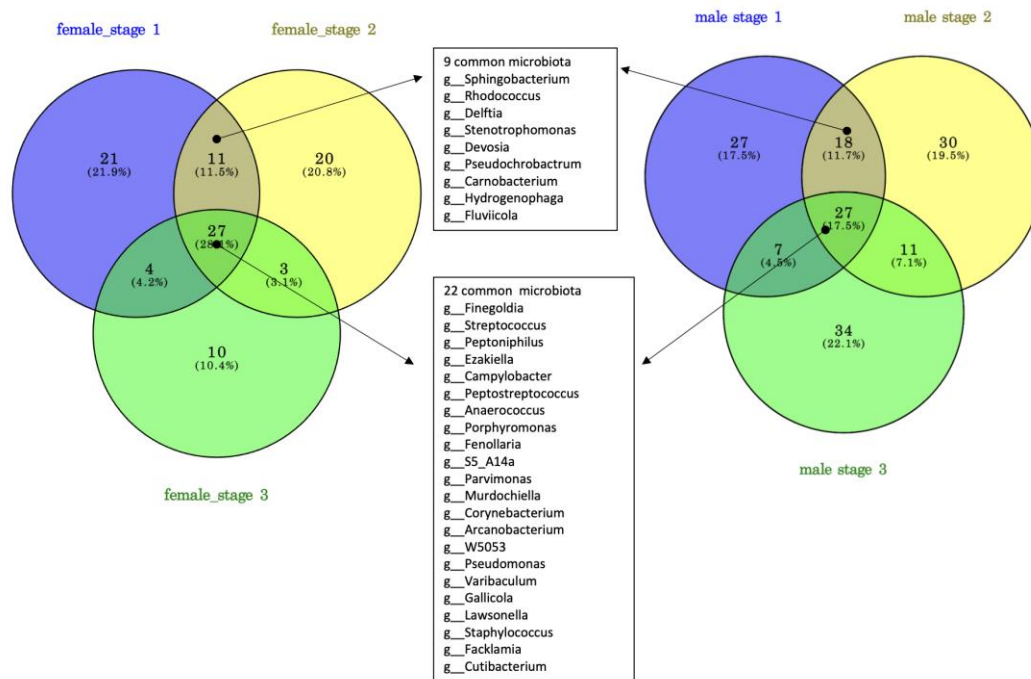


Figure S1. Venn diagram of microbiota genera meeting LDA score > 2 and $P < 0.05$ by cancer stage in males and females in colon cancer samples. The blue circle on the left Venn diagram indicates the female's stage 1, the yellow indicates stage 2, and the green indicates stage 3 genera. The Venn diagram on the right indicates comparison results by stage in males.

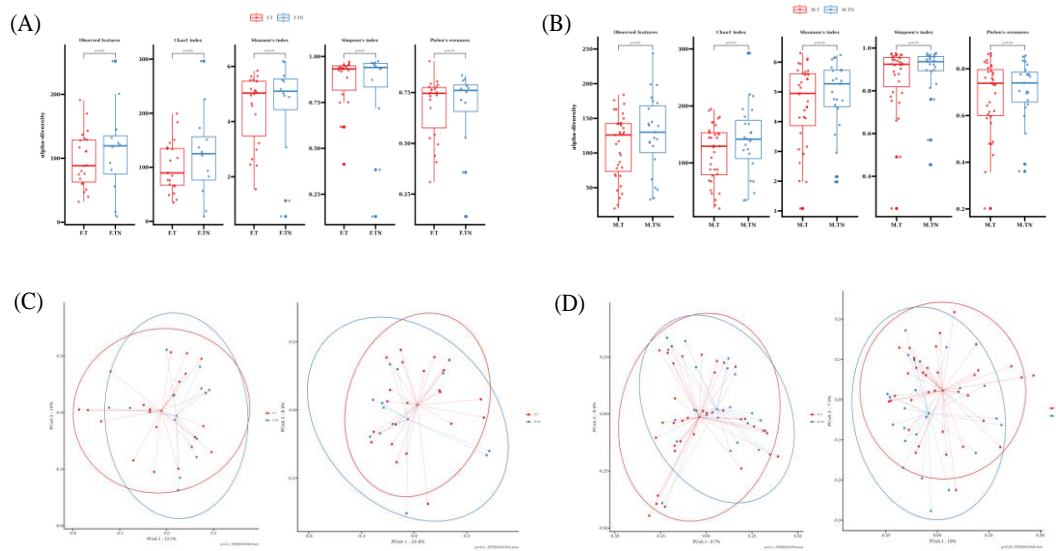


Figure S2. Diversity analysis results in the T group (without lymph node metastasis) and TN group (with lymph node metastasis) according to sex. (A) The alpha diversity box plot in the T and TN groups for each diversity index in the female sample, and (B) alpha diversity box plot in the male sample. The red box plot represents the T group, and the blue represents the TN group. (C) and (D) The beta diversity results. Red dots represent the T group; blue dots represent the TN group. There were no significantly different diversity results in the T and TN groups.

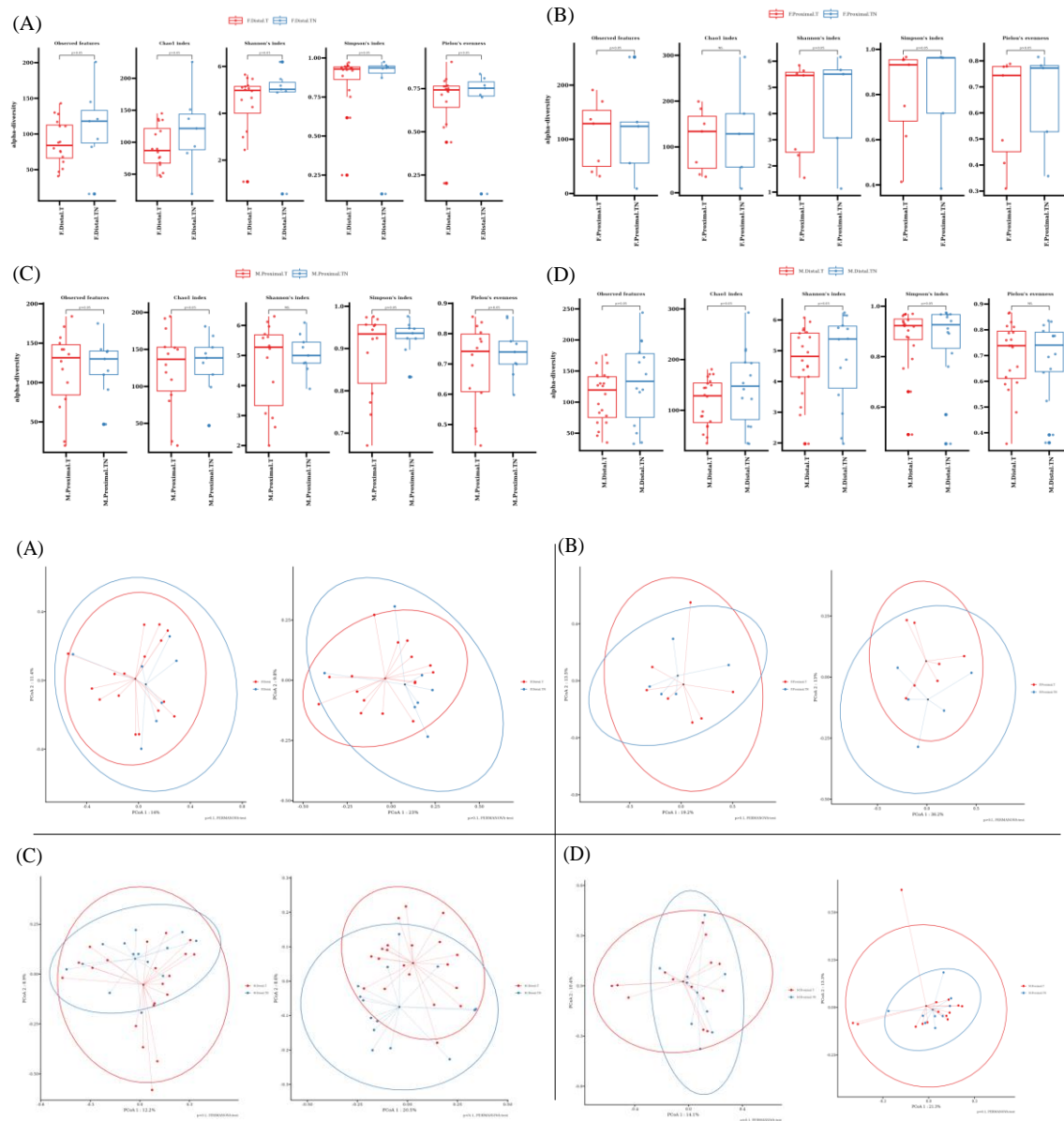


Figure S3. Diversity analysis results in the T group (without lymph node metastasis) and TN group (with lymph node metastasis) based on sex and cancer origin. (A) Alpha diversity boxplot in the T and TN groups for each diversity index in the female distal sample, and (B) alpha diversity boxplot in the T and TN groups for each diversity index in the female proximal sample. (C) The alpha diversity boxplot in the T and TN groups for each diversity index in the distal sample of males. (D) The alpha diversity boxplot in the T and TN groups for each diversity index in the proximal sample of males. The alpha diversity analysis showed no significant differences. In addition, the results in (D)-(G) are the T and TN groups for the female distal sample, female proximal sample, male distal sample, and male proximal sample, respectively. This PCoA plot was derived through the Bray-Curtis and unweighted unifrac method.

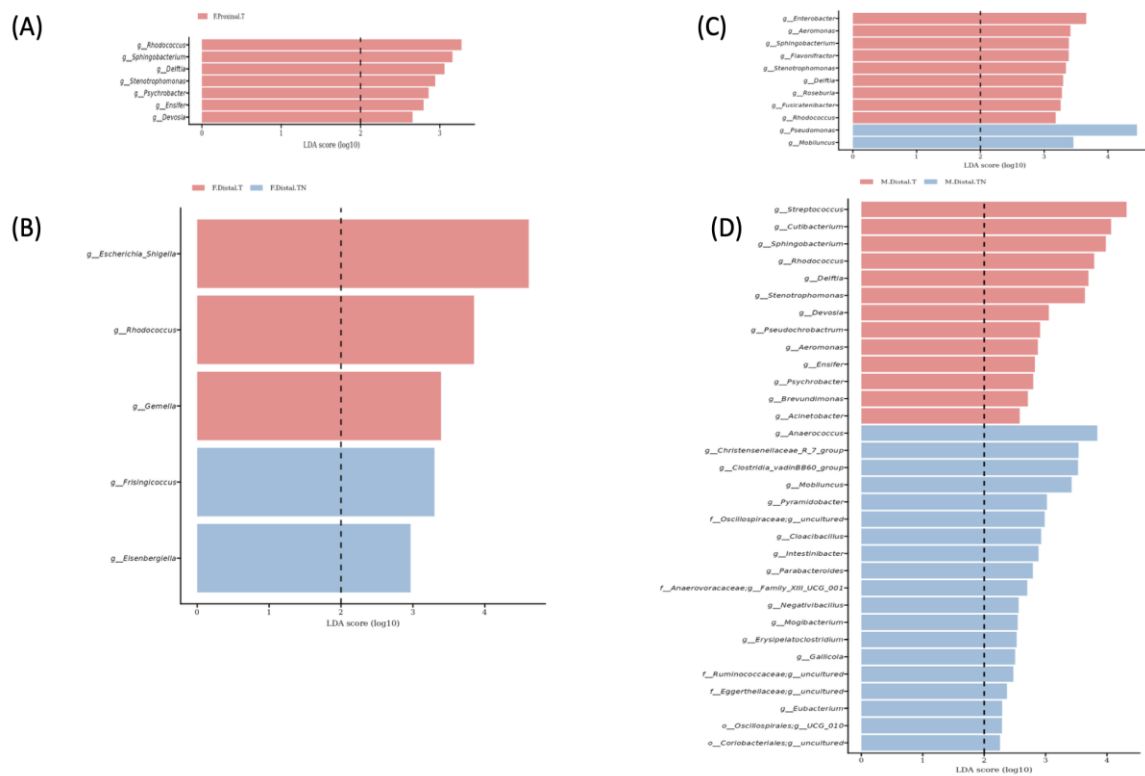


Figure S4. Lefse analysis results in the T group (without lymph node metastasis) and TN group (with lymph node metastasis) based on sex and cancer origin. The pink bars represent the T group, and the blue bars represent the TN group. (A) The result of lefse in the proximal part of a female. In the case of the TN group, it indicates no significant marker group. (B) indicates T in the proximal region in males and the TN group; (C) and (D) are the distal results in females and distal results in males, respectively.

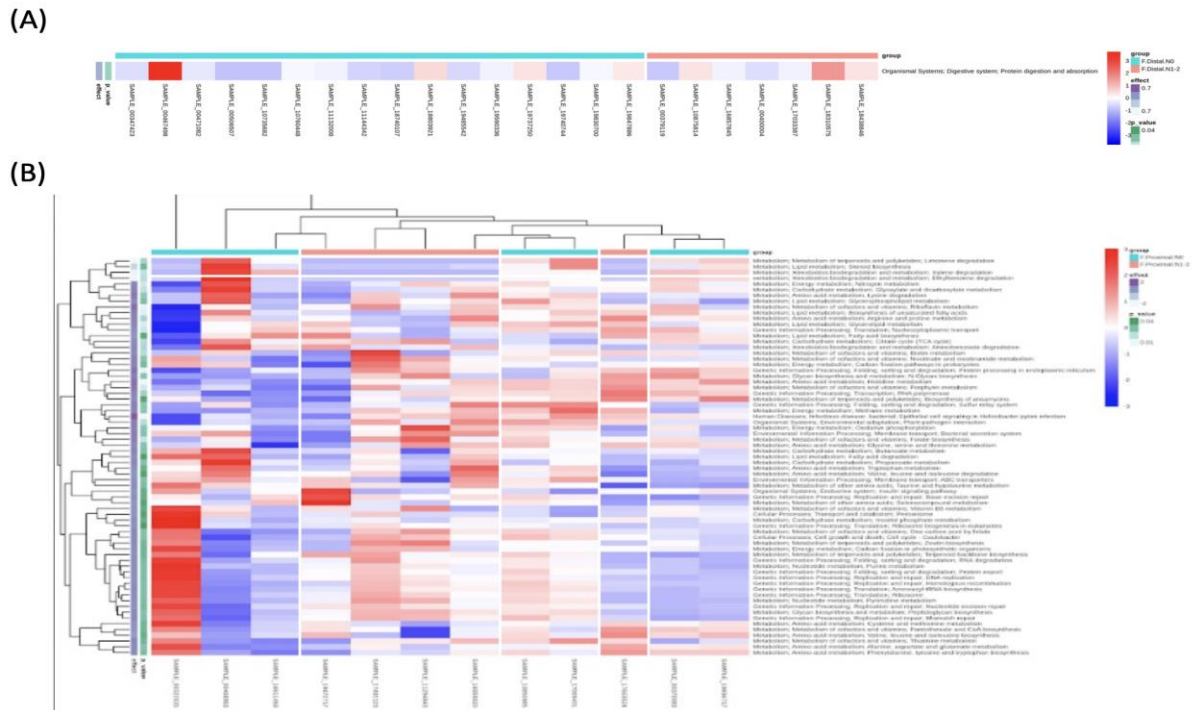


Figure S5. Functional analysis of microbiomes in groups stratified by colorectal cancer patient sex and, lymph node metastasis status, and cancer origin using PICRUSt 2.

The results of functional prediction analysis based on the origin (Proximal and Distal) of colorectal cancer within the gender-specific T and TN groups are shown. (A) Comparison between the female T group in the Distal origin (N0: no lymph node metastasis, represented by the light blue color group) and the TN group (N1-2: lymph node metastasis, represented by the pink color group). (B) Heatmap comparing these aspects in the Proximal origin.

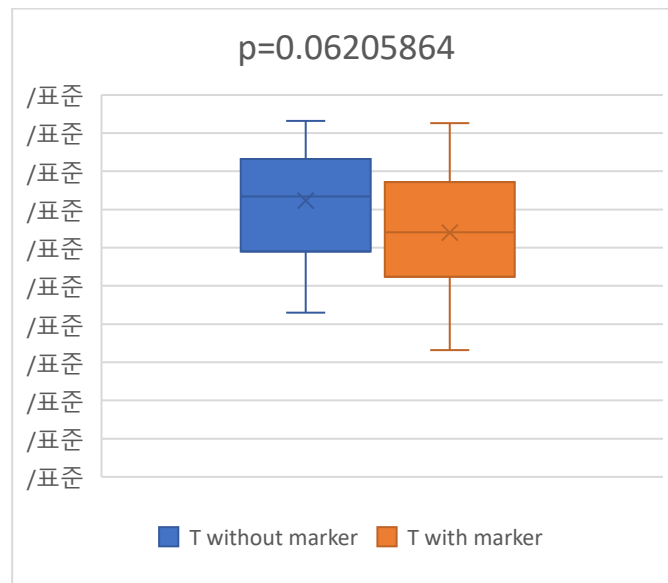


Figure S6. Comparison of lymphocyte ratios within the T group. Within the T group, the T without marker group refers to the group where 7 genera were not detected, while the T with marker group refers to the group where 7 genera were detected.

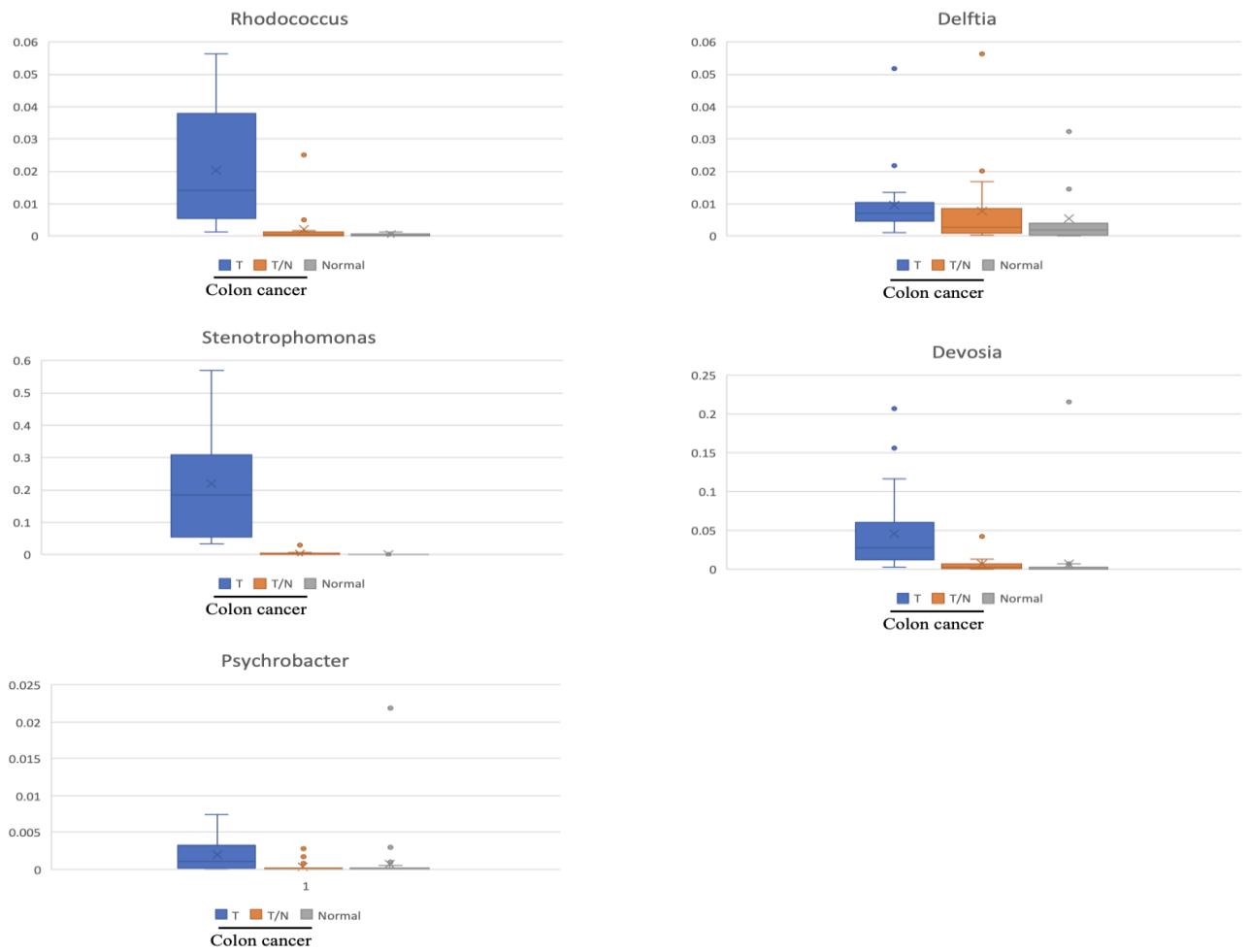


Figure S7. qPCR validation of candidate marker microbiome for early stages cancer detection. To validate through qPCR, the same samples were used with low-level primers. Validation was conducted using 54 samples from colorectal cancer (CRC) and 40 samples from normal individuals. The samples were validated using 54 randomly selected samples from the group and 40 normal samples from the control group, utilizing normal samples from the Korea Food Research Institute. In the case of *Ensifer* and *Sphingobacterium*, the primers did not work, and no detection was observed in quantitative, qualitative analysis

