

LETTER TO EDITOR

Diagnostic and prognostic potential of the oral and gut microbiome for lung adenocarcinoma

Dear Editor,

Lung cancer is one of the most common cancers.¹ Although its primary risk factor is smoking,² 15–20% and >50% of male and female lung cancer patients, respectively, are nonsmokers,^{3,4} which indicates the involvement of other risk factors. Recently, alteration of the oral and gut microbiomes was observed in lung cancer patients,^{5,6} suggesting potential roles of microbiota in lung cancer. We investigated the oral and gut microbiomes of never-smoking lung adenocarcinoma patients to identify microbial signatures of lung cancer, which can be used for diagnosis and prognosis.

Oral and gut microbiomes were profiled from never-smoking lung adenocarcinoma patients (cancer, $n = 91$) and from age- and sex-matched never-smoking healthy controls (control, $n = 91$). Most patients were female

($n = 84, 92.3\%$) and had stage IV lung cancer ($n = 81, 89.0\%$; Table S1). Shannon's diversity index of the oral microbiome was significantly lower in patients, and the overall oral microbiome structure differed significantly between patients and controls (Figure 1A). The gut microbiome structure also differed between patients and controls, but the difference in Shannon diversity was insignificant (Figure S1A).

The patients were further categorized into treated (Treatment+, $n = 52$) and untreated groups (Treatment–, $n = 39$) based on, at the time of microbiome sampling, whether they had received one or more lines of chemotherapy or targeted drug therapy for lung cancer since their first lung cancer diagnosis. The oral microbiome diversity was reduced in the Treatment– group (Wilcoxon rank-sum, $p = 0.21$) and reduced more in the Treatment+ group

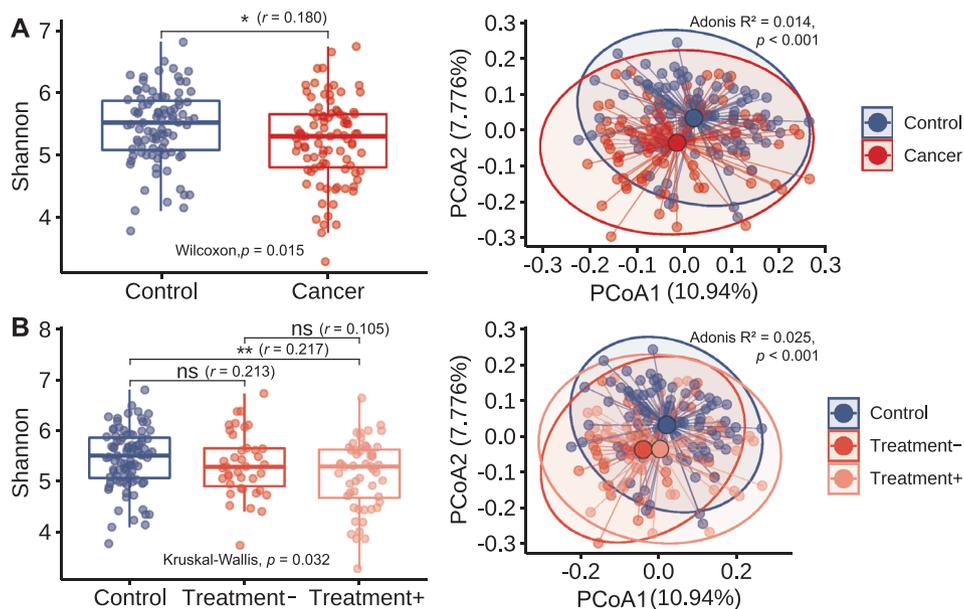


FIGURE 1 Oral microbiome diversity in healthy controls and lung cancer patients. Shannon's diversity index (left) and principal coordinate analysis plots (right) based on the Bray–Curtis distance. Samples were divided into (A) healthy controls and lung cancer patient groups, or (B) healthy controls, Treatment–, and Treatment+ groups. * $p < 0.05$; ** $p < 0.01$; ns, not significant; Wilcoxon rank-sum test, r means Wilcoxon effect size

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Clinical and Translational Medicine* published by John Wiley & Sons Australia, Ltd on behalf of Shanghai Institute of Clinical Bioinformatics

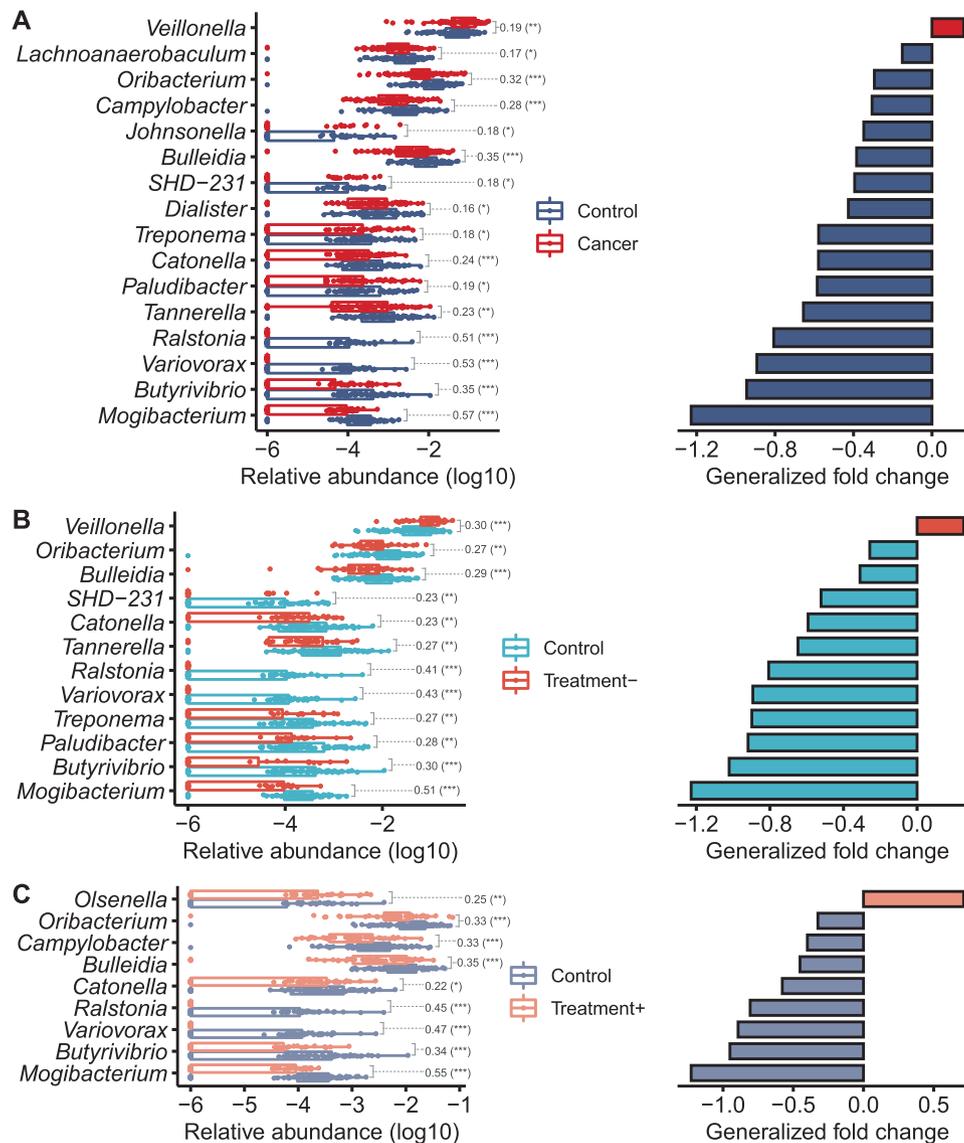


FIGURE 2 Oral bacteria genera that were differentially represented in the lung cancer patients and control groups. Log₁₀-transformed relative abundance (left) and generalized fold change (right) of differentially abundant genera among the different groups of subjects. Differentially abundant genera in healthy controls versus (A) all lung cancer patients, (B) Treatment– group, or (C) Treatment+ group were identified using a Wilcoxon rank-sum test ($q < 0.1$). * $q < 0.1$; ** $q < 0.05$; *** $q < 0.01$; the numbers between groups in the boxplot indicate the Wilcoxon effect size

(Wilcoxon rank-sum, $p = 0.01$; Figure 1B). No significant difference in the gut microbiome diversity was observed irrespective of treatment status (Figure S1B). Regardless of the oral and gut microbiomes, the samples in the Treatment– and Treatment+ groups were located closer to each other on the principal coordinate analysis plot than to the control samples (Figure 1B and Figure S1B).

We further compared the relative abundance of genera in the oral microbiome of the patients and controls. *Veillonella* was more abundant in the patients, whereas the abundance of 15 genera—*Mogibacterium*, *Butyrivibrio*, *Variovorax*, *Ralstonia*, *Catonella*, *Bulleidia*, and *Oribacterium*—decreased in the patients (Wilcoxon

rank-sum test, false discovery rate [FDR] < 0.1 ; Figure 2A). These seven genera were significantly less abundant in the patients regardless of treatment history (Figure 2B,C). By contrast, some taxa were associated with the treatment condition: *Olsenella* was more abundant in the Treatment+ patients, whereas *Veillonella* was more abundant in the Treatment– patients (Figure 2B,C).

Compared with oral bacteria, a relatively small number of gut bacterial taxa was associated with lung cancer. *Clostridium*, *Enterococcus*, and *Streptococcus* were enriched in patients, whereas *Faecalibacterium* was enriched in controls (Wilcoxon rank-sum test, FDR < 0.1 ; Figure S2A). A similar difference was observed between

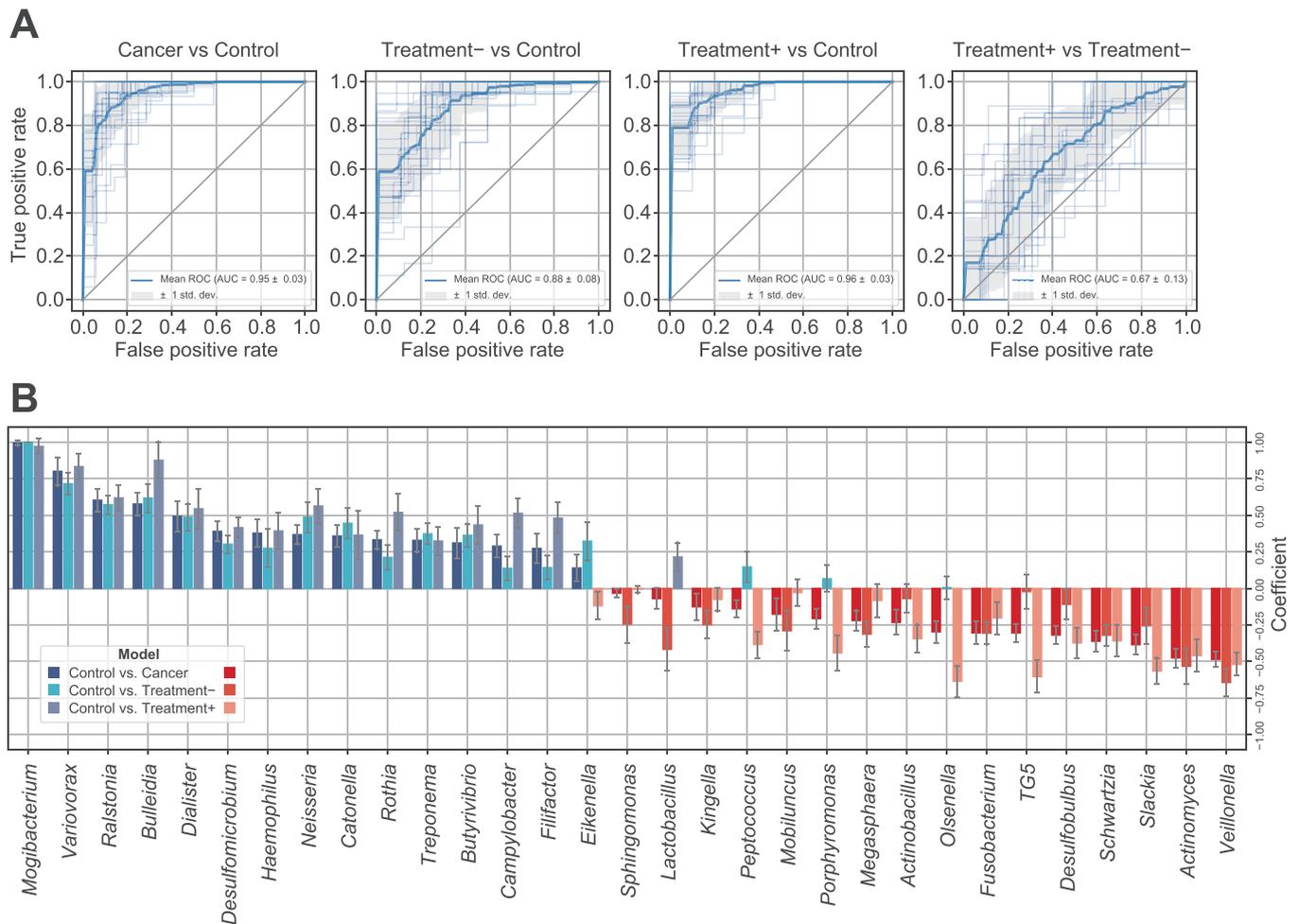


FIGURE 3 Lung cancer prediction from oral microbiome signatures. (A) Receiver operating characteristic (ROC) curve of the predictive models that distinguish the total cancer patients, Treatment+ group, Treatment- group, and control group from each other. The ROC curves of the test sets are plotted with semi-transparent lines, and the average curve is plotted with a solid line. The area representing one standard deviation from the average is colored in gray. The area under curve (AUC) is shown in the islet. (B) Normalized coefficients of the predictive models (Bayesian ridge). The genera for which the coefficients belonged to the highest and lowest 5th percentiles in the predictive model of cancer versus control are displayed. The higher abundance of a microbe with a positive coefficient increases the chance of predicting a healthy control

the Treatment+ and control groups, but not between the Treatment- and control groups (Figure S2B). Considering that microbes in the oral cavity can potentially migrate to the lungs,⁷ they may be more strongly associated with lung cancer than gut microbes.

Although the functions of lung cancer-associated microbes are not clearly understood, there are some references about them. *Veillonella* can activate cancer signaling pathways in airway epithelial cells,⁸ and the enrichment of *Enterococcus* and *Streptococcus* in the gut microbiome is linked to colorectal cancer.⁹ *Faecalibacterium* can modulate systemic immune responses.¹⁰ We further discussed the functional implications of these microbes in the Supporting Information.

We then evaluated whether the cancer status can be predicted from oral microbiome profiles. The patients

were successfully distinguished from the controls using a machine learning model, with an area under the receiver operating characteristic curve (AUC) of 0.95 in the five-fold cross-validation tests (Figure 3A and Figure S3). We further confirmed the performance of predictive models on an independent dataset including 41 patients with nonsmall cell lung carcinoma and 612 healthy controls (AUC = 0.88). The predictive model trained with the control and Treatment+ groups could distinguish cancer status (AUC = 0.96) better than the model trained with the control and Treatment- groups (AUC = 0.88; Figure 3A). The predictive performance on the independent dataset was also higher for the model trained with the Treatment+ group (AUC = 0.86) than that with the Treatment- group (AUC = 0.80). Although the prediction performance slightly varied, a similar set of microbial taxa

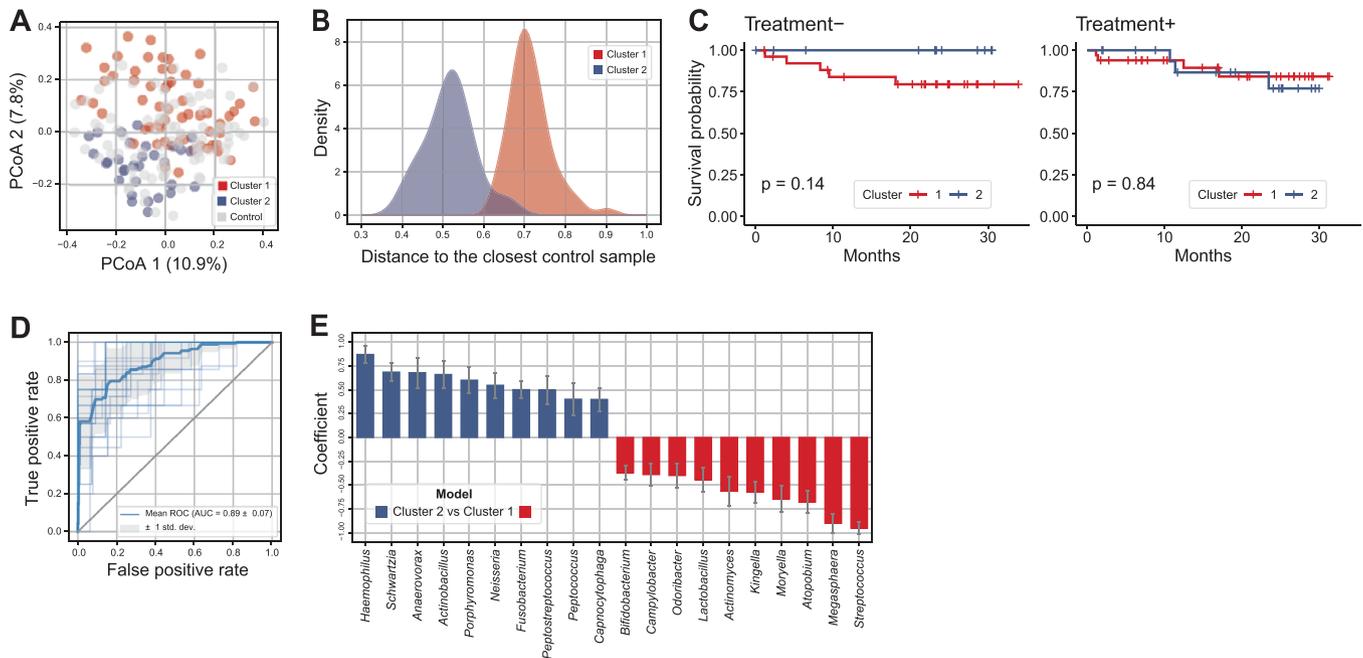


FIGURE 4 Prognostic impact of the oral microbiome profiles for lung cancer patients. (A) Principal coordinate analysis plot of the oral microbiome. Lung cancer patients were clustered into two groups based on Bray–Curtis distances. (B) Density plot of oral microbiome clusters based on the distance to the control sample with the highest similarity. (C) Survival curves for the two clusters of oral microbiome profiles. (D) Receiver operating characteristic curve of the predictive models that distinguish Clusters 1 and 2. (E) Normalized coefficients of the predictive models (Bayesian ridge)

was used in the two models (Figure 3B). Hence, the oral microbial taxa identified herein represent a general signature of lung cancer and can be used as diagnostic biomarkers.

Cancer status could also be predicted using gut microbiome profiles, but its predictive power was limited (AUC = 0.76; Figure S4A); the performance on the independent dataset was lower (AUC = 0.67). The gut microbiome of the Treatment+ group was slightly more informative in predicting the cancer status than that of the Treatment– group (Figure S4A). Similar genera were used in both models (Figure S4B), and their performances on the independent dataset were similar (AUC = 0.67, for both models, respectively).

To evaluate the prognostic impact of microbiome profiles, we clustered the patients into two groups by their oral microbiome similarity (Figure 4A). For the overall microbial structure, the first group (Cluster 1) was more dissimilar from the controls than the second group (Cluster 2; Figure 4B). In the Treatment– group, although statistical significance was not achieved, Cluster 2 patients had a trend toward longer survival than Cluster 1 patients ($p = 0.14$; Figure 4C). *Streptococcus* and *Megasphaera* were more abundant and *Haemophilus* was less abundant in the group with a worse survival outcome (Cluster 1; Figure 4D,E). However, no significant difference was observed in survival outcome among the Treatment+ patients.

Collectively, we found the alterations in the oral and gut microbiomes associated with lung adenocarcinoma and cancer treatment. The predictive models based on oral microbiome profiles could successfully distinguish the lung cancer patients from the healthy controls. These findings suggest the possibility of diagnosis of lung cancer, especially for nonsmokers, using oral microbiome profiles.

ACKNOWLEDGMENTS

This work was supported by the Main Research Program (grant number E0170600-05) of the Korea Food Research Institute (KFRI) funded by the Ministry of Science and ICT, and by the National Cancer Center (grant number 1910280).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ETHICS APPROVAL

This study was performed with approval from the National Cancer Center Institutional Review Board (approval number NCC2016-0208). All the participants provided written informed consent.

DATA AVAILABILITY STATEMENT

The oral and gut microbiome data of never-smoking patients with lung adenocarcinoma are available in the European Nucleotide Archive (<https://www.ebi.ac.uk/ena/>) under accession numbers PRJEB44168 and PRJEB44169, respectively.

AUTHOR CONTRIBUTIONS

Ji-Youn Han and Young-Do Nam conceived and designed the study. Kum Hui Hwang, Eun Jin Lim, and Ji-Youn Han collected the patients' samples and provided the clinical information. Mi Young Lim performed the experiments. Mi Young Lim and Seungpyo Hong analyzed the data and wrote the manuscript. All authors reviewed the manuscript. The authors read and approved the final manuscript.

Mi Young Lim¹ 
 Seungpyo Hong¹ 
 Kum Hui Hwang²
 Eun Jin Lim²
 Ji-Youn Han²
 Young-Do Nam¹ 

¹ Research Group of Healthcare, Korea Food Research Institute, Wanju-gun, Republic of Korea

² Center for Lung Cancer, National Cancer Center, Goyang-si, Republic of Korea

Correspondence

Young-Do Nam, Research Group of Healthcare, Korea Food Research Institute, 245 Nongsaengmyeong-ro, Iseo-myeon, Wanju-gun, Jeollabuk-do 55365, Republic of Korea.

Email: youngdo98@kfri.re.kr

Ji-Youn Han, Center for Lung Cancer, National Cancer Center, 323 Ilsan-ro, Ilsandong-gu, Goyang-si, Gyeonggi-do 10408, Republic of Korea.

Email: jymama@ncc.re.kr

Mi Young Lim and Seungpyo Hong contributed equally to this work.

ORCID

Mi Young Lim  <https://orcid.org/0000-0003-4881-1710>

Seungpyo Hong  <https://orcid.org/0000-0003-1273-0740>

Young-Do Nam  <https://orcid.org/0000-0002-2660-3569>

REFERENCES

1. International Agency for Research on Cancer. Global cancer observatory: cancer today. 2020. <https://gco.iarc.fr/today> Accessed December 16, 2020.
2. de Groot PM, Wu CC, Carter BW, Munden RF. The epidemiology of lung cancer. *Transl Lung Cancer Res.* 2018;7:220-233.
3. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin.* 2005;55:74-108.
4. Kurahara Y, Kawaguchi T, Tachibana K, et al. Small-cell lung cancer in never-smokers: a case series with information on family history of cancer and environmental tobacco smoke. *Clin Lung Cancer.* 2012;13:75-79.
5. Yan X, Yang M, Liu J, et al. Discovery and validation of potential bacterial biomarkers for lung cancer. *Am J Cancer Res.* 2015;5:3111-3122.
6. Zhuang H, Cheng L, Wang Y, et al. Dysbiosis of the gut microbiome in lung cancer. *Front Cell Infect Microbiol.* 2019;9:112.
7. Bassis CM, Erb-Downward JR, Dickson RP, et al. Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. *mBio.* 2015;6:e00037.
8. Tsay JJ, Wu BG, Badri MH, et al. Airway microbiota is associated with upregulation of the PI3K pathway in lung cancer. *Am J Respir Crit Care Med.* 2018;198:1188-1198.
9. Cheng Y, Ling Z, Li L. The intestinal microbiota and colorectal cancer. *Front Immunol.* 2020;11:615056.
10. Ferreira-Halder CV, Faria AVS, Andrade SS. Action and function of *Faecalibacterium prausnitzii* in health and disease. *Best Pract Res Clin Gastroenterol.* 2017;31:643-648.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.