

The Gut Microbiome Is Associated with Clinical Response to Anti-PD-1/PD-L1 Immunotherapy in Gastrointestinal Cancer



Zhi Peng¹, Siyuan Cheng¹, Yan Kou², Ziqi Wang¹, Rong Jin³, Han Hu², Xiaotian Zhang¹, Ji-fang Gong¹, Jian Li¹, Ming Lu¹, Xicheng Wang¹, Jun Zhou¹, Zhihao Lu¹, Quan Zhang², David T.W. Tzeng², Dongtao Bi², Yan Tan², and Lin Shen¹

ABSTRACT

We report on a comprehensive analysis of the gut microbiomes of patients with gastrointestinal (GI) cancer receiving anti-PD-1/PD-L1 treatment. The human gut microbiota has been associated with clinical responses to anti-PD-1/PD-L1 immunotherapy in melanoma, non-small cell lung cancer, and renal cell carcinoma. We aimed to investigate this association in GI cancers. We also identified bacterial taxa with patient stratification potential. We recruited 74 patients with advanced-stage GI cancer receiving anti-PD-1/PD-L1 treatment and collected their fecal samples prior to and during immunotherapy, along with clinical evaluations. Our 16S rRNA taxonomy survey on the fecal samples revealed an elevation of the *Prevotella/Bacteroides* ratio in patients, with a preferred response to anti-PD-1/PD-L1 treatment, and a particular subgroup of responders harboring a significantly higher abundance of *Prevotella*, *Ruminococcaceae*, and *Lachnospiraceae*.

The shotgun metagenomes of the same samples showed that patients exhibiting different responses had differential abundance of pathways related to nucleoside and nucleotide biosynthesis, lipid biosynthesis, sugar metabolism, and fermentation to short-chain fatty acids (SCFA). Gut bacteria that were capable of SCFA production, including *Eubacterium*, *Lactobacillus*, and *Streptococcus*, were positively associated with anti-PD-1/PD-L1 response across different GI cancer types. We further demonstrated that the identified bacterial taxa were predictive of patient stratification in both our cohort and melanoma patients from two previously published studies. Our results thus highlight the impact of gut microbiomes on anti-PD-1/PD-L1 outcomes, at least in a subset of patients with GI cancer, and suggest the potential of the microbiome as a marker for immune-checkpoint blockade responses.

See articles by Tomita et al., p. 1236, and Hakozaiki et al., p. 1243

Introduction

Gastrointestinal (GI) cancers are the most common cancers in the world. Despite improvements in therapeutic approaches in the last decade, the mortality of GI cancers is still among the highest, especially in gastric, esophageal, and colorectal cancers. The 5-year overall survival (OS) rate is only 20% to 30% for gastric cancer (1), whereas colorectal cancer-related deaths account for 16% of all deaths in the United States (2). Anti-PD-1/PD-L1 immune-checkpoint inhibitor (ICI) treatment has achieved a clinical breakthrough in many cancers, including advanced-stage melanoma, non-small cell lung cancer (NSCLC), and renal cell cancer (RCC; ref. 3). The FDA approved pembrolizumab to treat certain patients with recurrent locally or

metastasis gastric or gastroesophageal junction adenocarcinoma in 2017 (4). Many ongoing clinical trials aim to evaluate the safety and efficacy of anti-PD-1/PD-L1 immunotherapy in GI cancers (2). Only a small fraction of patients benefit from ICI therapy (10%–40%; refs. 5–7). The link between the human gut microbiota and clinical responses to anti-PD-1/PD-L1 treatment has begun to emerge in recent years. Accumulated evidence suggests a critical role of the gut microbiota in assisting ICI treatment (8, 9) in advanced melanoma, NSCLC, RCC, and urothelial cancer. In healthy individuals, the gut microbiota help maintain the gut barrier, along with the immune homeostasis, and may also promote anticancer immune surveillance through tumor antigenicity and adjuvanticity (10). The gut microbiota or its products can mimic tumor antigens and cause T cells to be primed locally before migrating to remote lymph nodes (11). It can also trigger systemic innate immune responses via pattern recognition receptors (PRR; ref. 12), which activate host responses against tumor cells.

Identification of bacteria taxa that directly or indirectly induce antitumor activities is critical to developing microbiome-based combinatory treatment that can improve the overall response rate of anti-PD-1/PD-L1 treatment. A few bacteria genus/species are found to be enriched in patients with preferred clinical outcomes, including *Akkermansia* [NSCLC, RCC (9) and hepatocellular carcinoma (13)], *Clostridiales* (melanoma; ref. 14), *Ruminococcaceae* [melanoma (14) and hepatocellular carcinoma (13)], *Faecalibacterium* (melanoma; ref. 14), *Bifidobacterium* (melanoma; ref. 15), *Collinsella* (melanoma; ref. 15), *Enterococcus* (melanoma; ref. 15), *Alistipes putredinis* (NSCLC; ref. 16), *Bifidobacterium longum* (NSCLC; ref. 16), and *Prevotella copri* (NSCLC; ref. 16). However, the list of gut bacteria clinically beneficial for ICI therapy is far from complete, and the prevalence of candidate taxa in responders and nonresponders often contradicts in different patient cohorts, even within the same cancer type.

¹Department of Gastrointestinal Oncology, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Peking University Cancer Hospital and Institute, Beijing, China. ²Xbiome, Shenzhen, China. ³Department of Immunology, School of Basic Medical Sciences, Peking University Health Science Center, Key Laboratory of Medical Immunology (Ministry of Health), Peking University, Beijing, China.

Note: Supplementary data for this article are available at Cancer Immunology Research Online (<http://cancerimmunolres.aacrjournals.org/>).

Z. Peng, S. Cheng, and Y. Kou contributed equally to this article.

Corresponding Authors: Lin Shen, Peking University Cancer Hospital and Institute, Fu-Cheng Road 52, Hai-Dian District, Beijing 100142, China. Phone: 86-10-88196561; E-mail: shenlin@bjmu.edu.cn; and Yan Tan, Scientific Research Building, Room 907, Tsinghua High-Tech Park, Nanshan District, Shenzhen, China. Phone: 86-075586530994; E-mail: yant@xbiome.com

Cancer Immunol Res 2020;8:1251-61

doi: 10.1158/2326-6066.CIR-19-1014

©2020 American Association for Cancer Research.

Here, we report on gut microbiomes from patients with GI cancer receiving anti-PD-1/PD-L1 therapy. The gut microbiotas from our cohort, regardless of their clinical responses, predominantly consisted of *Bacteroidetes* and *Firmicutes*, with a 2.5-fold elevation of the *Prevotella/Bacteroides* ratio in patients with favorable outcomes, and a subgroup of responders showed a higher abundance of *Prevotella*. Metagenomics analysis indicated that pathways related to nucleoside and nucleotide biosynthesis, lipid biosynthesis, sugar metabolism, and fermentation to short-chain fatty acids (SCFA) dictated the observed differences in the gut microbiota functionality among patients exhibiting different clinical outcomes. We further identified microbial biomarkers demonstrating reasonable performance on patient stratification within both our cohort and independent melanoma cohorts.

Materials and Methods

Patient cohort

Eighty-nine patients with stage III or stage IV GI cancer were enrolled in the study at the Beijing Cancer Hospital between February 2017 and January 2018 under Institutional Review Board (IRB)-approved protocols (2018KT66). Four patients not suitable for anti-PD-1/PD-L1 treatment thus were excluded. Eleven patients who received combined chemotherapy and anti-PD-1/PD-L1 immunotherapy were also excluded. The study was conducted in accordance with the Declaration of Helsinki. Signed voluntary informed consent was obtained from participating patients. All patients received or were scheduled to receive anti-PD-1/PD-L1 immune-checkpoint blockade therapy at the time being recruited. Fourteen patients also received anti-CTLA-4 in combination with the anti-PD-1/PD-L1 treatment. Patients received ICI therapy intravenously every 2 or 3 weeks until disease progression or intolerable toxicity. Responses to the therapy were assessed according to the patients' treatment plan following the Response Evaluation Criteria in Solid Tumors (RECIST) v1.1 standard. Clinical information including age, gender, diagnosis, microsatellite instability (MSI), allergy, and medication was collected from the medical records and is summarized in **Table 1**.

Fecal sample collection and DNA extraction

Patient fecal samples were collected at various time points throughout their immunotherapy treatment, as indicated in **Fig. 1**. Fecal samples were collected using the Wealthgene Fecal Microsolution TM Collection kit (catalog No. ML-001A, Wealthgene) and transported to the laboratory within 3 days. The Wealthgene kit is known to stabilize fecal samples under room temperature for up to 30 days. Fecal DNA was extracted according to the protocol provided by QIAamp PowerFecal DNA Kit Handbook (catalog No. 12830-50, Qiagen). The extracted DNA was stored at -20°C and transported to the sequencing facility on dry ice.

Bacterial 16S rRNA sequencing and taxonomic profiling

The extracted fecal DNA was used as the template for amplicon sequencing with barcoded primers (341F: 5'-CCTAYGGGRBGCAS-CAG-3') and 806R: 5'-GGACTACNNGGTATCTAAT-3') flanking the V3-V4 hypervariable region of the bacterial 16S rRNA gene. Approximately 200 ng DNA/sample was used for library construction. Sequencing was performed on the Illumina HiSeq 2500 platform (Novo Gene) according to the manufacturer's protocol on 250 bp \times 2 paired-end reads. Raw reads from all samples were quality filtered and denoised with DADA2 (17) to generate amplicon sequencing variants (ASV), which are essentially operational taxonomic units (OTU) with 100% identity. For ease of readability, we refer to these

Table 1. Patient statistics.

| Clinical factor | Total (n = 74) | R (n = 45) | NR (n = 29) | P value |
|-------------------------|-------------------|---------------|----------------|---------|
| Gender | | | | |
| Male | 53 | 31 | 22 | 0.6029 |
| Female | 21 | 14 | 7 | |
| Cancer | | | | |
| Colorectal cancer | 19 | 12 | 7 | 0.33 |
| Esophageal cancer | 14 | 8 | 6 | |
| Gastric cancer | 23 | 15 | 8 | |
| Others | 18 | 10 | 8 | |
| EBER | | | | |
| Positive | 6 | 4 | 2 | >0.9999 |
| Negative | 17 | 10 | 7 | |
| BMI | | | | |
| Normal | 39 | 25 | 14 | 0.4739 |
| Nonnormal | 33 | 18 | 15 | |
| Allergy history | | | | |
| Yes | 8 | 5 | 3 | >0.9999 |
| No | 66 | 40 | 26 | |
| MSI | | | | |
| MSI-H | 26 | 18 | 8 | 0.4305 |
| Non-MSI-H | 37 | 21 | 16 | |
| Prior treatment | | | | |
| ≥ 3 | 9 | 4 | 5 | 0.3105 |
| < 3 | 65 | 41 | 24 | |
| Age | | | | |
| ≤ 60 | 40 | 21 | 19 | 0.2373 |
| > 60 | 34 | 23 | 11 | |
| Treatment | | | | |
| Anti-PD-1 | 48 | 30 | 18 | 0.9415 |
| Anti-PD-1 + anti-CTLA-4 | 14 | 8 | 6 | |
| Anti-PD-L1 | 12 | 7 | 5 | |

Abbreviations: BMI, body mass index; MSI-H, MSI high; N, nonresponders; non-MSI-H, non-MSI high; R, responders.

ASVs as OTUs. We selected the reverse reads as input for DADA2 to ensure consistency when comparing with other studies. Taxonomy assignment of each unique OTU was performed with the pretrained Ribosomal Database Project classifiers (implemented in DADA2) using the SILVA (18) v132 reference database. OTUs with a total read count of less than 1,000 across all samples were considered belonging to rare microbes and were removed from all downstream analyses. The genus-level relative abundance was calculated by summing up the relative abundance of OTUs belonging to each genus on the phylogenetic tree. All the sequencing data were deposited at NCBI with accession number PRJNA615114.

Phylogenetic diversity of the gut microbiome

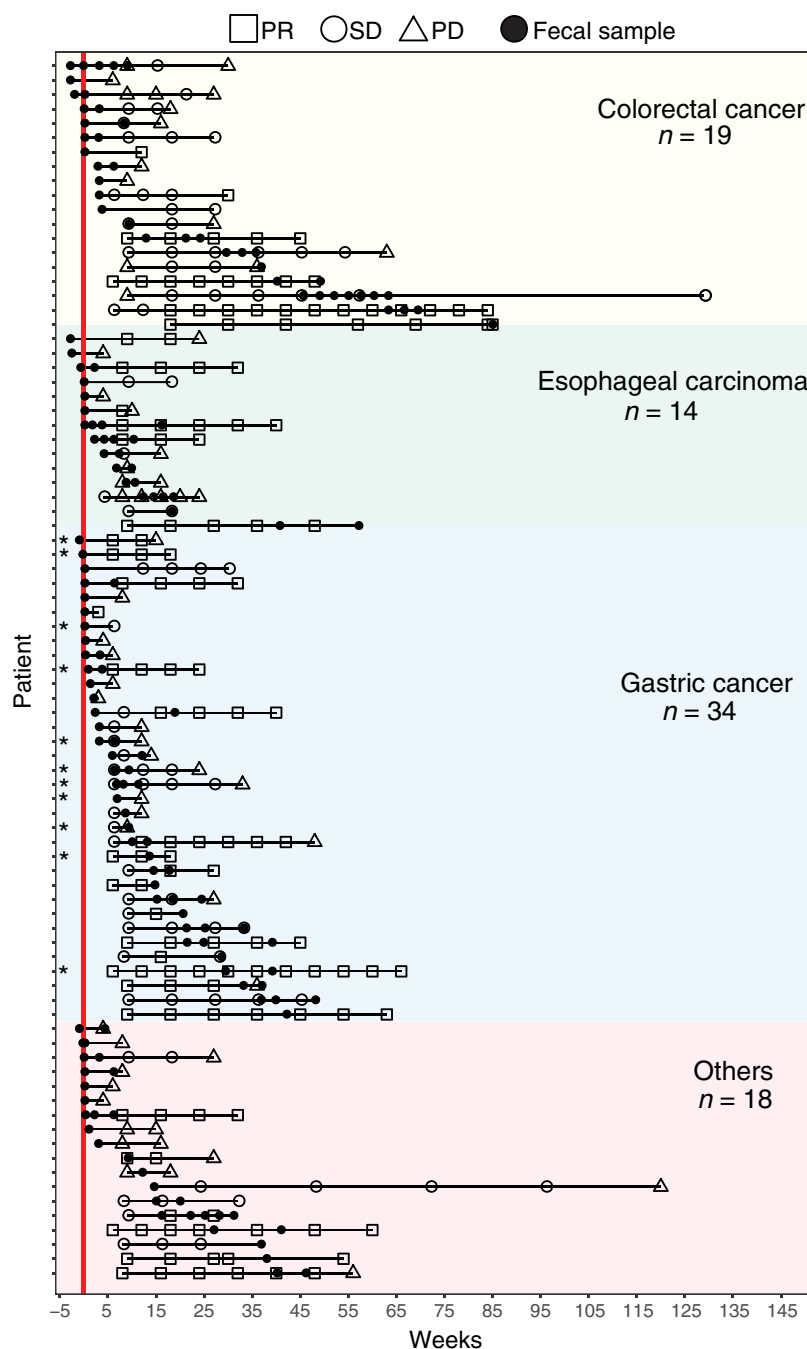
Alpha and beta diversity was computed using the R package phyloseq (19). For alpha diversity, the inverse Simpson index was calculated to represent the richness and evenness of each microbiome sample. To measure the distance between patients (beta diversity), principal coordinate analysis (PCoA) was performed based on pairwise Bray-Curtis dissimilarity. The statistical differences of alpha diversity across cancer types were calculated with the Wilcoxon test.

Identification of differentially abundant OTUs

We selected the first available fecal sample to represent the gut microbiome profile for each patient. Differentially abundant OTUs were identified using an omnibus method (20) that jointly tests the

Figure 1.

Overview of the clinical samples and collection time by cancer types. A total of 155 fecal samples from 85 patients with GI cancer were collected. Solid black dots indicate sample collection time relative to the start of anti-PD-1/L1 treatment; hollow shapes mark the weeks when tumor progression is evaluated. Patients who received chemotherapy along with anti-PD-1/PD-L1 treatment are marked with an asterisk. PD, progressive disease; PR, partial response; SD, stable disease.

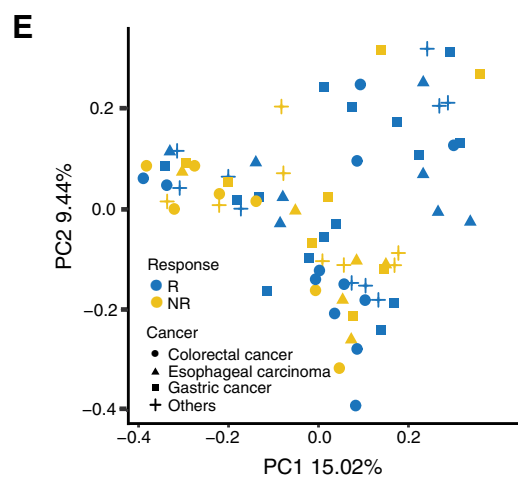
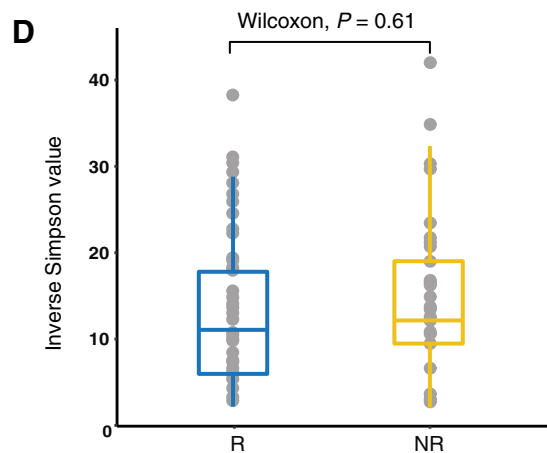
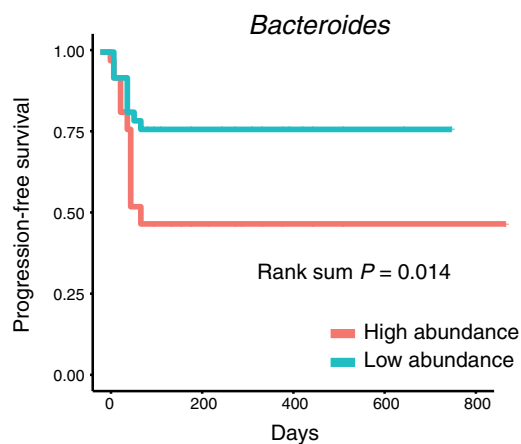
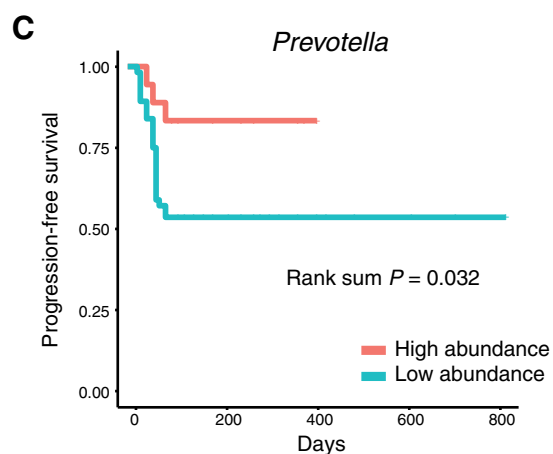
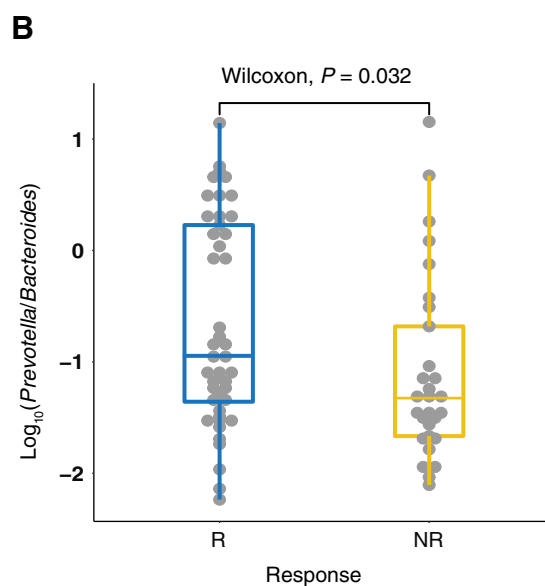
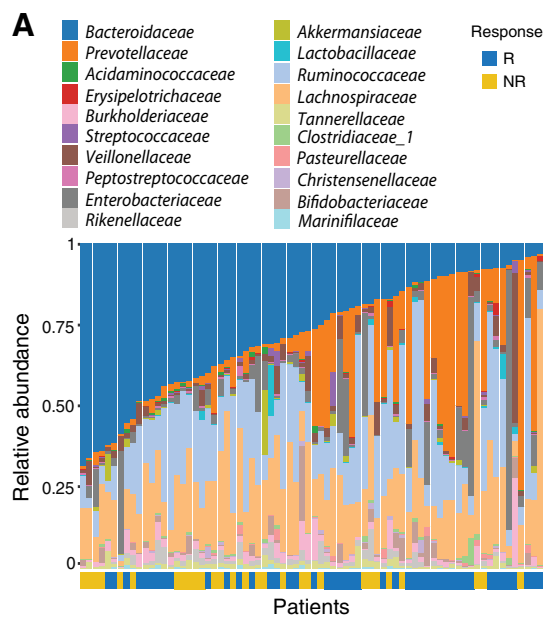


abundance, prevalence, and dispersion using a zero-inflated negative binomial regression model that was developed for microbiome data. For each taxon, the regression coefficient was tested using a likelihood ratio test, which follows asymptotically χ^2 distribution. The *P* value from such test essentially denoted whether taxa differed in any of the following: abundance, prevalence, or dispersion between responders and nonresponders. OTUs with *P* values less than 0.05 were reported. To remove the effect of any confounding factors, we used MaAslin (Multivariate Association with Linear models; ref. 21) to assess the association between bacterial abundance and gender, diagnosis, batch, body mass index (BMI), and MSI status.

Statistical analyses

The univariate survival analysis of the progression-free survival (PFS) was estimated using the Kaplan–Meier method and compared with the log-rank tests. We also performed Kolmogorov–Smirnov tests to examine the distribution of the *Prevotella/Bacteroides* ratio (defined by the ratio of summed relative abundance of OTUs assigned to *Bacteroides* and *Prevotella*, respectively) among our patients. Pearson coefficient was calculated to examine the correlation of the relative abundance of shared genera derived from 16S rRNA sequencing and the metagenomics sequencing methods (described below). All statistical analyses were performed in R.

Downloaded from <http://aacrjournals.org/cancerimmunolres/article-pdf/8/10/1251/2343638/1251.pdf> by guest on 28 August 2022



Metagenomic whole-genome shotgun sequencing

The whole-genome shotgun sequencing allows for species-level quantification of the gut microbiome and enables gene and functional profiling. We selected a subset of patients whose fecal samples were sufficient for the metagenomic analysis (Supplementary Table S1). Individual DNA libraries were constructed from each sample and loaded onto an Illumina NovoSeq 6000 (Novo Gene), yielding an average of 6 GB 150 bp × 2 pair-ended reads per sample. Raw reads were quality controlled (QCed) by KneadData (version 0.6.1), which integrated several QC tools such as FastQC (22) and Trimmomatic (23). After trimming the low-quality portion of reads and subsequently dropping reads less than 100 bp in length, we used bowtie2 (24) to map the filtered reads to the human genome (hg19) to remove host contaminants. The remaining high-quality reads were loaded to the MetaPhlan2 (ref. 25; version 2.2.0) pipeline for taxonomy profiling with default parameters.

Functional profiling of the metagenomes

Functional analyses were performed with HUMAnN2 (ref. 26; version 0.11.2) using the UniRef90 reference database. Briefly, the shotgun sequencing reads were mapped to the pangenomes of species identified by MetaPhlan2 (25) in previous steps. The majority of reads were mapped to these pangenomes harboring the coding sequences of proteins annotated in UniRef90 (27). Unmapped reads were translated to peptide sequences and mapped to UniRef90 by DIAMOND (28). The gene-level abundance was calculated as reads per kilobase units, as defined by HUMAnN2. The reads that failed to map to known species/genes were grouped as unclassified. The gene abundances were further annotated according to the KEGG Orthology (29) and MetaCyc (30) database, respectively.

Centered log-ratio transformation

Centered log-ratio (CLR) transformation addresses the constraints of microbial compositional data (31, 32). The CLR transformation $Z_{ij}^{(k)}$ of microbiome relative abundance $X_{ij}^{(k)}$ is defined as

$$Z_{ij}^{(k)} = \log \left(\frac{X_{ij}^{(k)}}{\left(\prod_{j=1}^p X_{ij}^{(k)} \right)^{1/p}} \right), \quad i = 1, \dots, n_r, \quad j = 1, \dots, p, \quad k = 1, 2 \quad (\text{A})$$

where i denotes samples, j denotes OTUs with at least one sequencing read, and k represents groups (responders vs. nonresponders). In order to avoid the zero-relative abundance in Equation (A), we replaced zero counts by a pseudo-count of 0.5 before performing the relative abundance normalization and CLR transformation. The CLR transformation maps composition data in the D-part Aitchison-simplex isometrically to a D-dimensional Euclidean vector subspace. Consequently, CLR transformation is not injective; the resulting covariance matrices are always singular.

Machine learning models and comparison with published melanoma data sets

The 16S rRNA sequencing data and metadata were retrieved from Matson and colleagues (15) and Gopalakrishnan and colleagues (14) as described in the original publications. For fair comparison, we processed the raw reads using the same workflow of our analysis. Both studies sequenced the V4 hypervariable region of the 16S rRNA gene; thus, we restricted the comparison to include only the reverse reads in our samples. Forty-two subjects (16 responders and 26 nonresponders) and 40 subjects (29 responders and 11 nonresponders) were kept from Matson and colleagues and Gopalakrishnan and colleagues, respectively, for subsequent comparison.

To predict the response status (responder/nonresponder) using the 16S-rRNA-derived microbiome profiles, we used the absolute abundance of 90 shared genera present in all three data sets as features for binary classification. We first trained several classification models [random forest, extra trees, support vector machine (SVM), elastic net, and k-nearest neighbors] on our GI data set using 5-fold cross-validation with exhaustive grid search to find optimal parameters for each model. For SVM, candidate values for the C penalty and kernel types were [0.1, 1.0, 5.0, 10.0, 50.0] and ["rbf," "poly," "sigmoid," "linear"]. Default parameters were used for other models. The train-validation splits were stratified with regard to the responder/nonresponder ratio. Within the validation folds, accuracy was used as the evaluation metrics to select the optimal parameters for each model. Two baseline models based on simple rules were included: (1) baseline-stratified makes random predictions with the probabilities based on the observed responder/nonresponder ratio and (2) baseline-prior always predicts the majority class from the training set. We also transformed the original genus-level features with polynomial transformation followed by univariate feature selection by ANOVA F-values to couple with SVM and elastic net classifiers. All classification models were implemented in the Scikit-learn Python package. We evaluated the classification models on independent holdout test sets constructed in various comparison scenarios.

Results

Our Chinese cohort includes 19 patients with colorectal cancer, 34 patients with gastric cancer, 14 patients with esophageal carcinoma, and 18 patients of other GI cancer types ($N = 85$; Table 1). The cohort was predominantly male, with a small proportion having allergy history. One hundred fifty-five samples were collected from these patients before and/or during anti-PD-1/PD-L1 treatment (Fig. 1; Supplementary Fig. S1). Eleven patients with gastric cancer who received chemotherapy in addition to the standard anti-PD-1/PD-L1 therapy were excluded from subsequent analysis. Fourteen patients who received anti-CTLA-4 treatment in combination with the anti-PD-1/PD-L1 therapy were included, as no significant associations

Figure 2.

The taxonomy composition of the gut microbiome is associated with clinical responses to anti-PD-1/PD-L1 treatment. **A**, Phylogenetic composition (from 16S rRNA sequencing) of common bacterial taxa at the family level, ordered by the most abundant taxa (*Bacteroidaceae*) across the cohort ($n = 74$). **B**, The ratio of the relative abundance of *Prevotella* and *Bacteroides* in responders and nonresponders. Wilcoxon test was performed on the \log_{10} -transformed ratio. The box represents the first (lower line) and third (upper line) quartiles as well as median (the line within the box). **C**, Kaplan-Meier survival analysis of patients with high versus low abundance of *Prevotella* (left) and *Bacteroides* (right). Eighteen and 56 patients showed high and low abundance of *Prevotella*, respectively; 37 and 37 patients exhibited high and low abundance of *Bacteroides*, respectively. P values were calculated with the log-rank sum test. **D**, Alpha diversity of patients with favorable and less favorable response to anti-PD-1/PD-L1 therapy. Inverse Simpson index was used to calculate alpha diversity with the relative abundance of OTUs from 16S rRNA sequencing, and the statistical difference was assessed with the Wilcoxon test. **E**, PCoA of fecal samples ($n = 74$, including 45 responders and 29 nonresponders) by response and diagnosis using Bray-Curtis dissimilarity. The x - and y -axes show the first and second principal coordinates, along with the percentage of variance explained on each dimension. NR, nonresponders; R, responders.

of the gut microbiome or response were observed with the choice of treatment (Table 1), and the downstream conclusions did not change when excluding these patients. Clinical responses including partial response (PR), stable disease (SD), and progressive disease (PD) were evaluated according to the RECIST 1.1 standard. We defined a patient as a responder if the patient achieved an objective response (PR/SD) lasting at least 3 months upon treatment start, or a nonresponder (PD) observed within 3 months of treatment start. The response rate was comparable with previous studies (Supplementary Table S2; ref. 33). No significant differences were found in age, drug, gender, EBER status, BMI, allergy history, MSI, or prior treatment between responders and nonresponders (Table 1; Supplementary Tables S2–S4).

Gut microbiome profile of patients with GI cancer on anti-PD-1 immunotherapy

We identified 3,852 OTUs at the 100% sequence similarity cutoff using the Silva database. We retained 319 OTUs for subsequent analyses after removing extremely low abundant OTUs (see Materials and Methods for details), as they are susceptible to random variation or false detection. The number of OTUs in each cancer type is shown in the Venn diagram in Supplementary Fig. S2. In brief, 306 of 319 OTUs were identified in all 74 patients. We found 308, 317, and 314 OTUs in all fecal samples of patients with colorectal cancer, esophageal carcinoma, and gastric cancer, respectively.

From 16S rRNA sequencing results, we observed a higher relative abundance of *Prevotellaceae* in the responder group, with a decreasing presence of *Bacteroidaceae* (Fig. 2A), at the family level. Indeed, the relative abundance of the *Bacteroides* genus was significantly lower in nonresponders (Supplementary Fig. S3). On the other hand, *Prevotellaceae* is predominantly represented by the genus *Prevotella_9* in our samples. Although its relative abundance is not significantly elevated in responders (Supplementary Fig. S3), the ratio of *Prevotella/Bacteroides* is significantly different between responders and nonresponders (Fig. 2B; Wilcoxon test, $P = 0.032$). We observed this trend in colorectal cancer and esophageal carcinoma but not in gastric cancer (Supplementary Fig. S4). We assigned all 74 patients into high and low abundance groups according to the median relative abundance of these two taxa (cutoff = 0.08 for *Prevotella* and 0.27 for *Bacteroides*). Survival analysis showed the high-*Prevotella* group was more likely to achieve PFS within 12 weeks of treatment start compared with those with low-*Prevotella* abundance ($P = 0.032$; Fig. 2C, left). On the contrary, the abundance of *Bacteroides* was negatively associated with PFS in this cohort ($P = 0.014$; Fig. 2C, right). *Akkermansia* and *Lactobacillus* are reported in previous studies to be positively associated with favorable outcomes of anti-PD-1 treatment in melanoma (9, 14). Our association analysis suggested that patients with a higher abundance of *Akkermansia* ($P = 0.031$), but not *Lactobacillus* ($P = 0.56$), were likely to benefit from anti-PD-1/PD-L1 therapy in our GI cancer cohort (Supplementary Fig. S5).

Phylogenetic diversity of the gut microbiome

Although a highly diverse gut microbiome has been associated with healthy state or preferred clinical response in several diseases (34–37), we found no significant difference in the alpha diversity between responders and nonresponders in our cohort (Fig. 2D; $P = 0.61$), regardless of cancer type (Supplementary Fig. S6). One melanoma study (14) reports decreased alpha diversity in a less favorable group of patients receiving anti-PD-1 immunotherapy. However, similar patterns have not been observed in other independent studies of melanoma and renal cell carcinoma (9, 15). In our study, a subgroup consisting of predominantly responders was clearly separated from

other patients (Fig. 2E). We further identified the observed variance was mainly driven by the differential abundance of *Prevotella* and was not associated with other clinical characteristics.

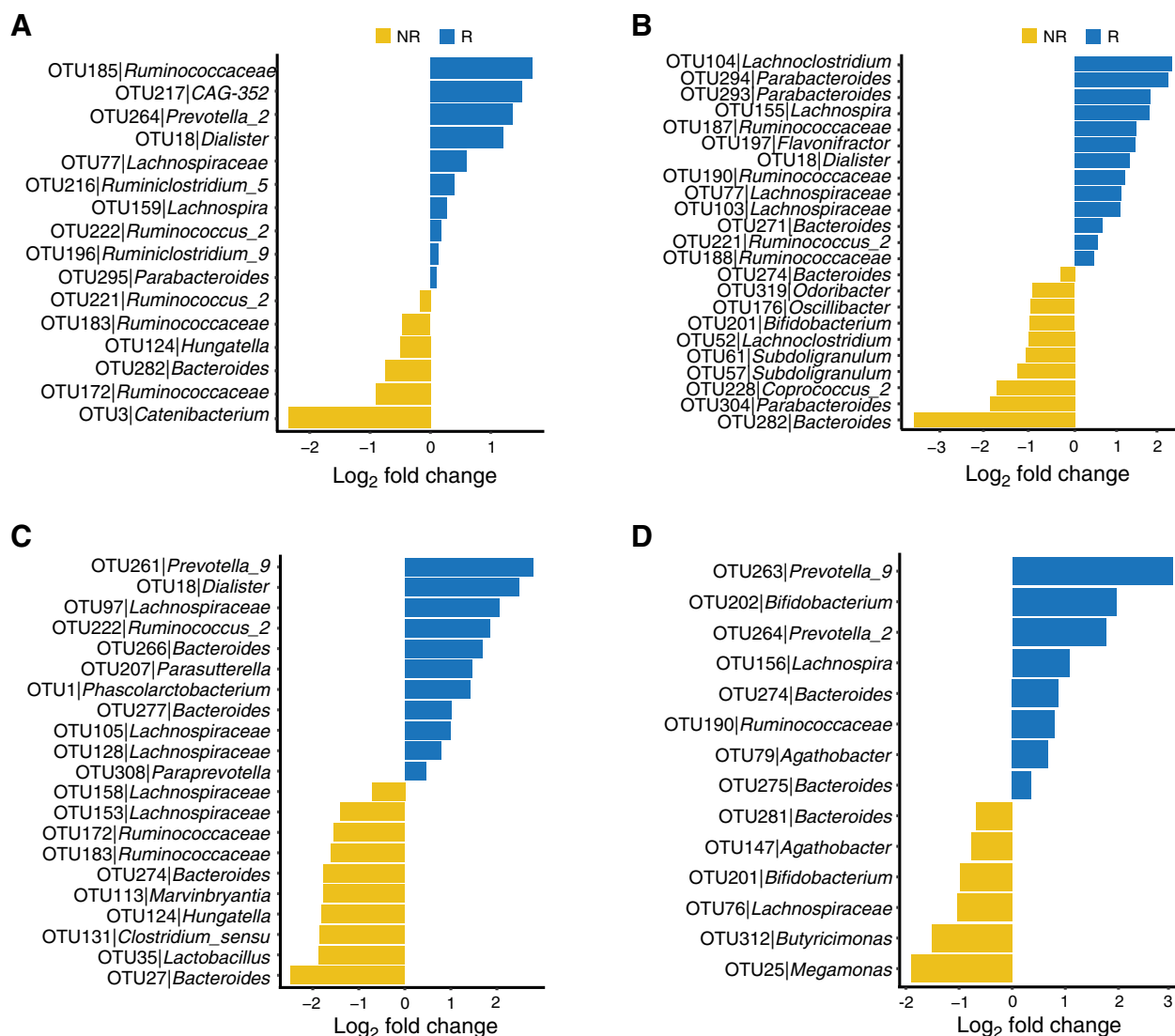
Differential abundance of bacterial taxa in patients with different clinical responses

An important aspect of studying the gut microbiome and its relationship with the human host is to derive biomarkers for diagnosis or prognostic prediction. Differential bacteria taxa between different response groups can identify clinically relevant microbial species as potential biomarkers. To remove confounding effects, we excluded any genus that was significantly associated with one or more factors, including age, gender, BMI, cancer types, MSI, and a few others, using MaAslin2 (ref. 21; Supplementary Fig. S7). We first performed differential analysis using DESeq2 (38), which failed to yield significantly different taxa between responders and nonresponders (Supplementary Table S5) due to its inability to count for excessive zeros in our microbiome data set. This problem was overcome by using a novel omnibus method that jointly tests the abundance, prevalence, and dispersion using a zero-inflated negative binomial regression model (20). Our differential analysis identified 10 OTUs enriched in responders and 6 OTUs enriched in nonresponders (Fig. 3A; Supplementary Table S6). OTUs belonging to *Ruminococcaceae*, *Prevotella*, and *Lachnospiraceae* were among the top enriched taxa in responders, whereas *Bacteroides*, *Catenibacterium*, and *Ruminococcaceae_NK4A214_group* were overrepresented in nonresponders. This is in line with previous studies in patients with melanoma, NSCLC, and RCC receiving anti-PD-1 immunotherapy reporting the diverse presence of *Ruminococcaceae* microbes between different response groups (9, 15, 39). Routy and colleagues observed certain *Prevotella* species enriched in responders of NSCLC and RCC (9), whereas Gopalakrishnan and colleagues reported a higher abundance of *Prevotella_histicola* in nonresponders of melanoma (14).

We also examined the differentially abundant genera within each GI cancer type (Supplementary Tables S7–S9). We identified 23 differential OTUs in colorectal cancer, among which *Lachnospiraceae*, *Parabacteroides*, *Lachnospiraceae*, *Ruminococcaceae*, *Flavonifractor* (*Eubacterium*), and *Dialister* were enriched in responders, and *Bacteroides*, *Parabacteroides*, *Coprococcus*, and *Subdoligranulum* were enriched in nonresponders (Fig. 3B). *Fusobacterium nucleatum* is reported to associate with colorectal cancer (40, 41). We did not observe its enrichment in either group, which could be due to the small sample size of our study. In esophageal carcinoma, a total of 21 differential OTUs showed significantly different abundance, with 11 enriched in responders and 10 enriched in nonresponders (Fig. 3C). In patients with gastric cancer, 8 OTUs were significantly enriched in responders, including *Prevotella*, *Bifidobacterium*, and *Lachnospiraceae*, and 6 OTUs were enriched in nonresponders (Fig. 3D), including *Megamonas*, *Butyrivimonas*, *Lachnospiraceae_UCG-001*, and *Agathobacter*. *Bacteroides* and *Ruminococcaceae* were enriched in both response groups according to different cancer types. These two taxa are among the most abundant gut microbes of our patients; the observed inconsistency, thus, was likely caused by different species belonging to the two taxa.

Metagenomics shotgun sequencing analysis of the gut microbiome

We performed metagenomics shotgun sequencing in 40 of the 74 patients (25 responders and 15 nonresponders; Supplementary Table S1) of whom the matching 16S rRNA sequencing data were available. Among the 17 differentially abundant OTUs identified by

**Figure 3.**

Differential abundance between response groups. Abundance in responders and nonresponders in all 74 patients (**A**), patients with colorectal cancer (**B**; $n = 19$), patients with esophageal carcinoma (**C**; $n = 14$), and patients with gastric cancer (**D**; $n = 23$). All differential OTUs were identified using the omnibus test, with $P < 0.05$. OTUs were ordered by the abundance in responders versus nonresponders. The length of the horizontal bars represents the log₂ fold change of each genus between responders and nonresponders. NR, nonresponders; R, responders.

16S rRNA-based analysis, OTUs belonging to *Prevotella* (Fig. 4A, top; Pearson correlation, $r = 0.89$) and *Bacteroides* (Fig. 4A, bottom; Pearson correlation, $r = 0.92$) demonstrated correlated abundance with those from the metagenomic sequencing data. This correlation confirmed the consistency between our 16S rRNA and metagenomics sequencing results. Using a centered log-norm transformation of the relative abundance, we identified, in the metagenomics sequencing data, significantly differential species from the *Eubacterium*, *Lactobacillus*, *Akkermansia*, *Catenibacterium* (also identified by 16S rRNA analysis; Fig. 3A), *Dialister*, and *Streptococcus* genera between responders and nonresponders (Supplementary Table S10; Supplementary Fig. S8; Mann-Whitney U test, $P < 0.05$). *Eubacterium* ferments fiber into SCFAs, including acetic acid, butyric acid, and propionic acid, and *Lactobacillus* and *Streptococcus* are the main contributors of lactic acid in the gut. A higher abundance of these genera in responders

may promote an intact intestinal environment (Fig. 4 with beneficial immune activity.

Differential bacterial pathways indicate potential mechanisms driving clinical responses

We used HUMAnN2 (26) to quantitatively identify the metabolic and biological pathways in our metagenomics data. Hierarchical clustering of differentially abundant (between responders and non-responders) pathways divided patients into two major groups (Fig. 4B), with response rates of 73% and 56%, respectively. The separation was primarily driven by pathways involved in nucleoside and nucleotide biosynthesis. Responders were enriched in pathways involved in fermentation to SCFAs, unsaturated fatty acid biosynthesis, vitamin and starch biosynthesis, whereas nonresponders showed higher read abundance for lipopolysaccharide biosynthesis, sugar

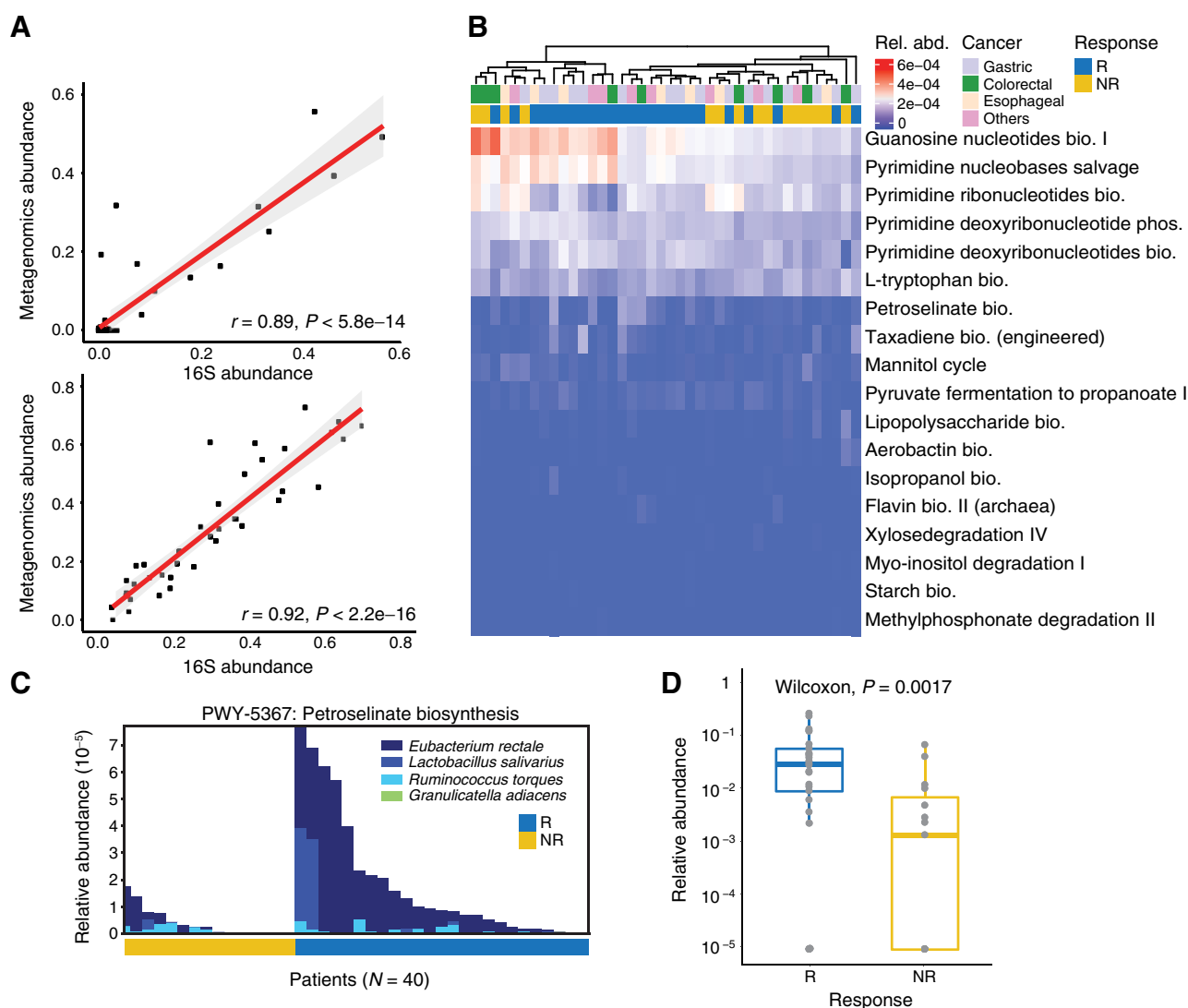


Figure 4.

Functional analysis of the gut microbiome with metagenomics sequencing data. **A**, Pearson correlation of *Prevotella* (top) and *Bacteroides* (bottom) abundance from 40 patients who had matching 16S rRNA and metagenomics sequencing samples. The diagram shows the regression line (in red) with 95% confidence intervals (the gray area). **B**, Functional analysis by metagenomics shotgun sequencing data. Hierarchical clustering of differentially abundant pathways was derived using Euclidean distance and complete linkage methods on log-normalized data. Color bars along the column dendrogram represent cancer types (top) and clinical responses (bottom). **C**, Contribution of bacterial species to petroselinate biosynthesis grouped by 25 responders and 15 nonresponders. **D**, Relative abundance of *Eubacterium rectale* in responders and nonresponders from metagenomics shotgun sequencing data. Statistical differences were assessed by the Wilcoxon test. The box represents the first (lower line) and third (upper line) quartiles as well as median (the line within the box). NR, nonresponders; R, responders.

degradation, and amino acid biosynthesis (Supplementary Table S11; Mann-Whitney U test, $P < 0.05$). The majority of the 18 differential pathways were contributed by a handful of species. For example, *Eubacterium rectale* dominated the petroselinate biosynthesis pathway among responders (Fig. 4C), and its abundance was significantly higher in responders (Fig. 4D). *Eubacterium rectale* is known for its ability to produce butyrate and has been associated with a fiber-rich diet (42). It is one of the few species that harbor the cBir1 antigen, which contributes to the gut-reactive T-cell repertoire (43).

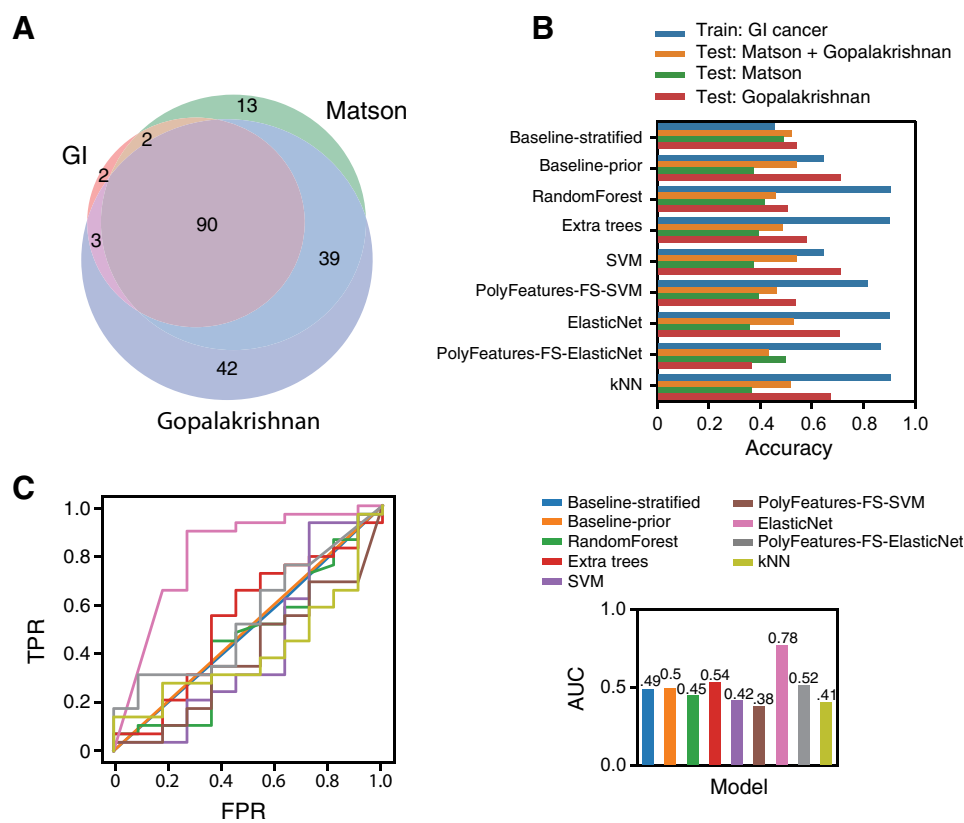
Patient stratification with the gut bacteria

A few pioneer studies have identified various sets of commensal bacteria enriched in patients with melanoma benefiting from the

anti-PD-1 immunotherapy (14, 15, 39). Nonetheless, which bacteria contribute to such differences are diverse according to the patient cohorts. We, thus, attempted to identify the signature bacterial taxa for responders in our cohort, followed by validation in independent melanoma cohorts. The assumption was that the commensal bacteria interplay with the host immune system in a systematic fashion, regardless of tumor types. To assess the patients' microbiome profiles immediately following anti-PD-1/PD-L1 treatment, we removed 13 patients with only pretreatment fecal samples. For each of the remaining 61 patients (Supplementary Table S1), the first posttreatment sample was used for all subsequent analysis (referred to as GI cancer). We obtained 16S sequencing data from two independent studies (refs. 14, 15; see Materials and

Figure 5.

Predicting clinical response with gut microbiome biomarkers and comparison with patients with melanoma. **A**, Genera identified in our cohort and two independent melanoma studies (14, 15) of patients receiving anti-PD-1 treatment. Fecal samples from 40 patients and 43 patients were included in the two studies. Numbers in the Venn diagram indicate the genera counts in each region. A total of 90 genera were observed in all three data sets. **B**, The accuracy of various machine learning models trained with taxonomy features of the patients with GI cancer. Blue bars represent cross-validation accuracy; orange, green, and red bars show the accuracy on test sets in melanoma patient cohorts. **C**, Receiver operating characteristic (ROC) curves of the prediction performance on 40 patients with melanoma from the Gopalakrishnan et al. study (14), using models trained with the patients with GI cancer. The bar plot shows the AUC (area under the curve) of each model. The x-axis represents the false positive rate [FPR, defined by false positive/(true negative + false positive)], and the y-axis represents the true positive rate [TPR, defined by true positive/(true positive + false negative)].



Methods for details). The relative abundance of OTUs (belonging to 90 genera) that were observed across all three data sets was kept to build classifiers (Fig. 5A). We built both parametric and nonparametric machine learning models to dissect the predictive power of these genera on patient responses (see Materials and Methods). All but SVM-based models trained with GI cancer achieved over 0.9 accuracy in a holdout test set (Fig. 5B; 5-fold nested cross-validation). We tested these GI cancer-trained models on patients in the two melanoma studies. As expected, the accuracy on either or combined data sets (mean accuracy = 0.49) was much less than the GI cancer test set (mean accuracy = 0.86), with noticeable better accuracy on patients in the Gopalakrishnan set (mean accuracy = 0.60; Fig. 5B, red bars).

In order to overcome unbalanced classes in the Gopalakrishnan set (29 responders and 11 nonresponders), we further generated two baseline models (see Materials and Methods). Only the ExtraTrees model (randomized decision trees on various subsamples of the training set) and elastic net-based models yielded better prediction than the baseline models. The elastic net model without polynomial features outperformed other nonparametric models [Fig. 5C; area under the curve (AUC) = 0.78], likely due to its ability to eliminate features through regularization in the training stage. Among the top 20 predictive features ranked by the elastic net model (Supplementary Table S12), the majority (15 of 20) belonged to the Firmicutes phylum. Ruminococcaceae, Lachnospiraceae, Bacteroides, and Catenibacterium harbored the most predictive features, indicating they might play a role in the response to anti-PD-1/PD-L1 therapy that is independent of cancer types, at least in a subset of our patients with GI cancer.

Discussion

We report a comprehensive analysis of the gut microbiomes of patients with GI cancer receiving anti-PD-1/PD-L1 treatment. We observed the ratio of the relative abundance of *Prevotella* and *Bacteroides* tends to be elevated in responders. The difference was primarily driven by the bimodal distribution of this ratio, which is a recurrent pattern in several large-scale population-based studies (44–46). Hence, the observed divergence of *Prevotella* versus *Bacteroides* prevalence may be an indicator of the enterotypes of our patients, possibly priming their immune system into different conditions.

Several bacteria taxa enriched in responders in our cohort are also reported in other anti-PD-1 clinical studies of other cancer types. For example, Routy and colleagues found that the relative abundance of *Akkermansia muciniphila* correlates with preferred clinical outcomes in patients with NSCLC and RCC (9). *Lactobacillus* is more abundant in metastatic melanoma responders (15). A metagenomics study on 39 patients with melanoma also identified the elevated presence of *Streptococcus* in responding patients (39). Immediate conclusions on the gut microbiome's role in anti-PD-1 therapy from these early findings are still controversial due to limited data availability. We plan to investigate the role of commensal bacteria in a larger patient cohort, coupled with host immune assays and mechanistic animal studies in the future.

Although emerging evidence points to the positive association between a diverse gut microbiota and preferred clinical outcomes of anti-PD-1 treatment, it is unclear if the observed effect is consistent across patient cohorts or cancer types. A study on a small cohort of Chinese patients with NSCLC receiving nivolumab shows a

significantly higher diversity of the gut microbiome in patients exhibiting PR or SD (16). Nevertheless, results from a handful of clinical studies on patients with advanced-stage melanoma show mixed findings (14, 15, 39). In our reanalysis of Matson and colleagues (15), we did not observe statistically significant differences in bacterial taxa relative abundance between responders and nonresponders. The variation in each study design, such as sample collection, inclusion/exclusion criteria, treatment regimen, diet, and population differences, as well as data analyses, may further complicate the interpretability of findings on microbiome compositions in different response groups. It is yet to be established whether a diverse gut microbiota merely reflects a patient's overall health prior to anti-PD-1/PD-L1 treatment, or that it harbors species that cooperate with the immune system to promote a positive response.

The mode of action of the cross-talk between gut microbes and the host immune system is yet to be elucidated. Three categories of possible mechanisms have been proposed (47): (i) through the T-cell responses induced by microbial antigens, (ii) through the engagement of pattern recognition receptors, and (iii) through small molecules produced by microbial metabolism. Our findings of the responder-enriched species that are capable of producing SCFAs (*Eubacterium rectale* and *Streptococcus*) may provide additional evidence on the proimmunotherapy effect of a subset of the commensal bacteria. SCFAs are known to inhibit inflammation through regulating Tregs. Although the exact molecular interplay remains to be determined, gut-derived SCFAs are observed to have a negative impact on colorectal carcinogenesis (48). Because SCFAs are the main fermentation products of dietary fiber, it is of particular interest to assess the impact of diet on gut microbiota and immune manipulation in patients with cancer.

We developed machine learning models with the goal to evaluate the usability of the gut microbiome as anti-PD-1/PD-L1 immunotherapy biomarkers. We further tested our models on patients with melanoma because (i) the small sample size of individual cancer types in our cohort prohibited the development of a robust, generalizable model; (ii) we reason that none of the possible mechanisms mentioned above seems cancer dependent; and (iii) the 16S RNA sequencing data were only available from patients with melanoma receiving anti-PD-1 therapy. In spite of the variation in study design, continent, diet, and even cancer types, some of our models achieved unexpected encouraging performance in these independent melanoma data sets, indicating that the gut microbiome affects host immune function in a systematic fashion, and it may serve as a repertoire for novel biomarkers of anti-PD-1/PD-L1 immunotherapy. Gharaibeh and colleagues (49) previously reported classifiers that were trained with species-level metagenomic profiles of patients with melanoma. They suggest that the prediction could be improved when incorporating KEGG orthologous features. In the same notion, we expect the performance of our models can be further enhanced by adding metagenome-derived features.

In 2017, the FDA approved the use of pembrolizumab in advanced solid tumors in patients with the MSI-high/DNA mismatch-repair-deficient tumors. We did not include MSI status in our models because our study was not designed to evaluate the response rate of patients with certain MSI status. The MSI status may introduce bias in the machine learning models. Both the ratio of MSI-high and the objective response rate of our patients with colorectal and noncolorectal cancer were consistent with published studies (33). It would be interesting to further evaluate the predictive performance of gut microbiome and MSI-based biomarkers or develop a combined marker set for potentially improved performance.

Our models were limited by the relatively small sample size ($n = 61$), coupled with sparse, high-dimensional features which hinders the prediction consistency. Besides, the imbalance between responders and nonresponders in our cohort imposes bias toward predicting the overrepresented class. It is of note that our primary results can only stand with the selected patient set, and it may suffer from the heterogeneity of where their tumor originated. Thus, an additional, independent validation cohort is needed, preferably with the same cancer type, to demonstrate the use of microbiome-based biomarkers in predicting anti-PD-1/PD-L1 response. We observed a subset of responders that exhibited clustering behavior across various analyses. Hence, they may represent a clinically relevant patient subgroup. Gopalakrishnan and colleagues identified similar subgrouping of responders when clustering patients with melanoma with their clusters-of-related-OTUs (crOTU) abundance (14). In either study, the key microbial drivers of subgrouping were different, and the exact mechanisms need further investigation in model systems.

In conclusion, we present data on gut microbiomes of patients with GI cancer receiving anti-PD-1/PD-L1 immunotherapy. In responders, we observed an elevated ratio of *Prevotella/Bacteroides*, along with enrichment of *Ruminococcaceae* and *Lachnospiraceae*. Commensal bacteria that produce SCFAs were overrepresented in patients with favorable outcomes, including *Eubacterium*, *Lactobacillus*, and *Streptococcus*. Using GI cancer-derived taxonomic features, we built machine learning classifiers with moderate performance on predicting the anti-PD-1/PD-L1 response of both patients with GI cancer and melanoma.

Disclosure of Potential Conflicts of Interest

H. Hu and Y. Tan report a patent for CN109680085A pending (Shenzhen Xbiome Biotech Co., Ltd.). No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Z. Peng: Conceptualization, resources, funding acquisition, investigation, methodology. **S. Cheng:** Resources, data curation, investigation, methodology. **Y. Kou:** Formal analysis, investigation, methodology, writing—original draft, writing—review and editing. **Z. Wang:** Validation, investigation. **R. Jin:** Validation, investigation. **H. Hu:** Investigation, methodology. **X. Zhang:** Validation, investigation. **J.-F. Gong:** Validation, investigation. **J. Li:** Validation, investigation. **M. Lu:** Validation, investigation. **X. Wang:** Validation, investigation. **J. Zhou:** Validation, investigation. **Z. Lu:** Validation, investigation. **Q. Zhang:** Formal analysis, visualization, methodology. **D.T.W. Tzeng:** Formal analysis, visualization. **D. Bi:** Data curation, methodology. **Y. Tan:** Conceptualization, supervision, investigation, writing—review and editing. **L. Shen:** Conceptualization, supervision, funding acquisition.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (nos. 2017YFC1308900 and 2017YFC0908400), the National Natural Science Foundation of China (no. 81602057), Clinical Medicine Plus X—Young Scholars Project of Peking University (PKU2019LCXQ020 and PKU2018LCXQ018), and Beijing Municipal Administration of Hospital's Youth Program (no. 20171102). The authors gratefully thank all the patients and their families for participating in the present study. The authors also thank Yiming Yin, Chao Yang, Chengsheng Zhu, and Zhenjiang Xu for discussion on the study design and interpreting results. They thank Mei Wang, Cong Zhang, and Ning Zhu for their help with the sequencing experiments. The authors appreciate Zichen Wang for his assistance in data analysis, along with Guanhua Ye for curating the metadata.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 20, 2019; revised March 27, 2020; accepted August 3, 2020; published first August 27, 2020.

References

- Niccolai E, Taddei A, Prisco D, Amedei A. Gastric cancer and the epoch of immunotherapy approaches. *World J Gastroenterol* 2015;21:5778–93.
- Bilgin B, Sendur MAN, Aknc MB, Dede DŞ, Yağın B. Targeting the PD-1 pathway: a new hope for gastrointestinal cancers. *Curr Med Res Opin* 2017;33:749–59.
- Hargadon KM, Johnson CE, Williams CJ. Immune checkpoint blockade therapy for cancer: an overview of FDA-approved immune checkpoint inhibitors. *Int Immunopharmacol* 2018;62:29–39.
- Marcus L, Lemery SJ, Keegan P, Pazdur R. FDA approval summary: pembrolizumab for the treatment of microsatellite instability-high solid tumors. *Clin Cancer Res* 2019;25:3753–8.
- Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N Engl J Med* 2015;373:1627–39.
- Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med* 2012;366:2443–54.
- Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med* 2015;373:1803–13.
- Chaput N, Lepage P, Coutzac C, Soularue E, Le Roux K, Monot C, et al. Baseline gut microbiota predicts clinical response and colitis in metastatic melanoma patients treated with ipilimumab. *Ann Oncol* 2017;28:1368–79.
- Routy B, Chatelier EL, Derosa L, Duong CPM, Alou MT, Daillère R, et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* 2018;359:91–7.
- Zitvogel L, Ayyoub M, Routy B, Kroemer G. Microbiome and anticancer immunosurveillance. *Cell* 2016;165:276–87.
- Tai N, Peng J, Liu F, Gulden E, Hu Y, Zhang X, et al. Microbial antigen mimics activate diabetogenic CD8 T cells in NOD mice. *J Exp Med* 2016;213:2129–46.
- Thaiss CA, Zmora N, Levy M, Elinav E. The microbiome and innate immunity. *Nature* 2016;535:65–74.
- Zheng Y, Wang T, Tu X, Huang Y, Zhang H, Tan D, et al. Gut microbiome affects the response to anti-PD-1 immunotherapy in patients with hepatocellular carcinoma. *J Immunother Cancer* 2019;7:193.
- Gopalakrishnan V, Spencer CN, Nezi L, Reuben A, Andrews MC, Karpinetz TV, et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* 2018;359:97–103.
- Matson V, Fessler J, Bao R, Chongsuwan T, Zha Y, Alegre M-L, et al. The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. *Science* 2018;359:104–8.
- Jin Y, Dong H, Xia L, Yang Y, Zhu Y, Shen Y, et al. The diversity of gut microbiome is associated with favorable responses to anti-PD-1 immunotherapy in Chinese non-small cell lung cancer patients. *J Thorac Oncol* 2019;14:1378–89.
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016;13:581.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013;41:D590–6.
- McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013;8:e61217.
- Chen J, King E, Deek R, Wei Z, Yu Y, Grill D, et al. An omnibus test for differential distribution analysis of microbiome sequencing data. *Bioinformatics* 2018;34:643–51.
- Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV, et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* 2012;13:R79.
- Babraham Bioinformatics-FastQC; [about 9 screens]. [cited 2018 Apr 16]. Available from: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–20.
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–9.
- Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, et al. MetaPhlan2 for enhanced metagenomic taxonomic profiling. *Nat Methods* 2015;12:902–3.
- Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods* 2018;15:962–8.
- Suzek BE, Wang Y, Huang H, McGarvey PB, Wu CH. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* 2015;31:926–32.
- Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 2015;12:59–60.
- Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 2016;44:D457–62.
- Caspi R, Billington R, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, et al. The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res* 2018;46:D633–9.
- Zhao N, Zhan X, Guthrie KA, Mitchell CM, Larson J. Generalized Hotelling's test for paired compositional data with application to human microbiome studies. *Genet Epidemiol* 2018;42:459–69.
- Cao Y, Lin W, Li H. Two-sample tests of high-dimensional means for compositional data. *Biometrika* 2018;105:115–32.
- Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med* 2015;372:2509–20.
- Stewart CJ, Ajami NJ, O'Brien JL, Hutchinson DS, Smith DP, Wong MC, et al. Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* 2018;562:583–8.
- Kang DW, Adams JB, Coleman DM, Pollard EL, Maldonado J, McDonough-Means S, et al. Long-term benefit of microbiota transfer therapy on autism symptoms and gut microbiota. *Sci Rep* 2019;9:5821.
- Kostic AD, Xavier RJ, Gevers D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* 2014;146:1489–99.
- Li Y, Wang H, Li X, Li H, Zhang Q, Zhou H, et al. Disordered intestinal microbes are associated with the activity of systemic lupus erythematosus. *Clin Sci* 2019;133:821–38.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.
- Frankel AE, Coughlin LA, Kim J, Froehlich TW, Xie Y, Frenkel EP, et al. Metagenomic shotgun sequencing and unbiased metabolomic profiling identify specific human gut microbiota and metabolites associated with immune checkpoint therapy efficacy in melanoma patients. *Neoplasia* 2017;19:848–55.
- Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, et al. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe* 2013;14:207–15.
- Kostic AD, Gevers D, Pedamallu CS, Michaud M, Duke F, Earl AM, et al. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res* 2012;22:292–8.
- Tanaka S, Yamamoto K, Yamada K, Furuya K, Uyeno Y. Relationship of enhanced butyrate production by colonic butyrate-producing bacteria to immunomodulatory effects in normal mice fed an insoluble fraction of *Brassica rapa* L. *Appl Environ Microbiol* 2016;82:2693–9.
- Shen C, Landers CJ, Derkowski C, Elson CO, Targan SR. Enhanced CBir1-specific innate and adaptive immune responses in Crohn's disease. *Inflamm Bowel Dis* 2008;14:1641–51.
- Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. *Nature* 2012;486:222–7.
- Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. *Nature* 2011;473:174–80.
- Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. *Science* 2011;334:105–8.
- Zitvogel L, Ma Y, Raoult D, Kroemer G, Gajewski TF. The microbiome in cancer immunotherapy: diagnostic tools and therapeutic strategies. *Science* 2018;359:1366–70.
- O'Keefe SJ, Li JV, Lahti L, Ou J, Carbonero F, Mohammed K, et al. Fat, fibre and cancer risk in African Americans and rural Africans. *Nat Commun* 2015;6:6342.
- Gharaibeh RZ, Jobin C. Microbiota and cancer immunotherapy: in search of microbial signals. *Gut* 2018;68:385–8.