Research Paper

Predicting cancer immunotherapy response from gut microbiomes using machine learning models

Hai Liang[1], Jay-Hyun Jo[1], Zhiwei Zhang[2], Margaret A. MacGibeny[1,3], Jungmin Han[1], Diana M. Proctor[4], Monica E. Taylor[1], You Che[1], Paul Juneau[5,6], Andrea B. Apolo[7], John A. McCulloch[8], Diwakar Davar[9], Hassane M. Zarour[9], Amiran K. Dzutsev[10], Isaac Brownell[1,11], Giorgio Trinchieri[10], James L. Gulley[11] and Heidi H. Kong[1,10]

[1]Dermatology Branch, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD 20892, USA

[2]Biostatistics Branch, Division of Cancer Treatment and Diagnostics, National Cancer Institute, NIH, Bethesda, MD 20892, USA

[3]Department of Medical Education, West Virginia University, Morgantown, WV 26506, USA

[4]Translational and Functional Genomics Branch, National Human Genome Research Institute, NIH, Bethesda, MD 20892, USA

[5]NIH Library, Division of Library Services, Office of Research Services, NIH, Bethesda, MD 20892, USA

[6]Zimmerman Associates Inc., Fairfax, VA 22030, USA

[7]Genitourinary Malignancies Branch, Center for Cancer Research, NCI, NIH, Bethesda, MD 20892, USA

[8]Genetics and Microbiome Core, Laboratory of Integrative Cancer Immunology, Center for Cancer Research, NCI, NIH, Bethesda, MD 20892, USA

[9]Department of Medicine and UPMC Hillman Cancer Center University of Pittsburgh, Pittsburgh, PA 15213, USA

[10]Laboratory of Integrative Cancer Immunology, Center for Cancer Research, NCI, NIH, Bethesda, MD 20892, USA

[11]Center for Immuno-Oncology, Center for Cancer Research, NCI, NIH, Bethesda, MD 20892, USA

Correspondence to: Heidi H. Kong, **email**: konghe@mail.nih.gov

ABSTRACT

Cancer immunotherapy has significantly improved patient survival. Yet, half of patients do not respond to immunotherapy. Gut microbiomes have been linked to clinical responsiveness of melanoma patients on immunotherapies; however, different taxa have been associated with response status with implicated taxa inconsistent between studies. We used a tumor-agnostic approach to find common gut microbiome features of response among immunotherapy patients with different advanced stage cancers. A combined meta-analysis of 16S rRNA gene sequencing data from our mixed tumor cohort and three published immunotherapy gut microbiome datasets from different melanoma patient cohorts found certain gut bacterial taxa correlated with immunotherapy response status regardless of tumor type. Using multivariate *selbal* analysis, we identified two separate groups of bacterial genera associated with responders versus non-responders. Statistical models of gut microbiome community features showed robust prediction accuracy of immunotherapy response in amplicon sequencing datasets and in cross-sequencing platform validation with shotgun metagenomic datasets. Results suggest baseline gut microbiome features may be predictive of clinical outcomes in oncology patients on immunotherapies, and some of these features may be generalizable across different tumor types, patient cohorts, and sequencing platforms. Findings demonstrate how machine learning models can reveal microbiome-immunotherapy interactions that may ultimately improve cancer patient outcomes.

INTRODUCTION

In the last decade, the use of cancer immunotherapy targeting immune checkpoint inhibitors (ICIs) to boost T cell mediated cancer cell clearance has significantly improved cancer patient survival [1]. Commonly used ICIs include monoclonal antibodies targeting programmed cell death protein (PD-1) and its ligand (PD-L1) or CTL antigen 4 protein (CTLA-4) [2]. ICI combination therapy is also widely used to increase the effectiveness of other cancer treatments [3]. Although immunotherapy can significantly improve treatment outcomes amongst different cancer types as compared to other treatment modalities [4], approximately half of patients do not respond to immunotherapy [3, 5, 6]. To optimize treatment outcomes, current efforts are aimed at elucidating the internal or external features of patients or tumors that correlate with immunotherapy responsiveness [7, 8].

Increasing evidence has emerged that gut microbial communities help shape the host immune system [9–11]. Several studies have suggested specific gut bacteria can influence immunotherapy outcomes by modulating the immune responses of patients with metastatic melanoma, non–small cell lung cancer, and renal cell cancer [12–33]. Treatment responders generally exhibit increased gut microbial community diversity and are enriched in certain bacterial taxa including *Akkermansia* and *Bifidobacterium* [16, 19]. Prior studies have primarily focused on the impact of individual bacterial taxa on immunotherapy outcomes in melanoma patients. It is currently unclear whether the identified response signals are generalizable across tumor types. Furthermore, findings may not be reproducible across different patient cohorts due to geographical variation in the microbiome [34] as well as differences in sequencing platform or analysis methodology. Here, we took a tumor-agnostic approach to identify microbiome features associated with immunotherapy response from a discovery cohort of patients with nine different advanced stage cancers. To uncover common immunotherapy response signals regardless of tumor type, we conducted a combined meta-analysis integrating the discovery cohort data with three previously published 16S rRNA gene sequencing datasets from melanoma patients. Using the combined dataset, we trained and validated models with machine learning algorithms to predict patients' clinical responses, followed by cross-sequencing-platform validation using shotgun metagenomic sequencing data.

RESULTS

Specific bacterial taxa associated with immunotherapy response in advanced stage cancer patients

We sought to investigate the association between patients' baseline gut microbiomes and clinical outcomes of immunotherapy. The discovery cohort consisted of 16 patients with late-stage solid tumors enrolled in different immunotherapy trials at the National Cancer Institute (NCI). Immunotherapy clinical response rate in the cohort was 38% (6 responders and 10 non-responders) (Supplementary Table 1). Despite a male predominance in the cohort (12 of 16 male patients), patient demographics did not differ significantly between responders (83% male; mean age 58.5 ± 7.8 years) and non-responders (70% male; mean age 57.1 ± 15.5 years) (two-tailed Student's *t*-test, $p = 0.8140$). However, antibiotic use 30 days prior to immunotherapy was more common among non-responders than responders (Fisher's exact test, $p = 0.2335$); although not statistically significant due to a small sample size, our data align with prior observations of a negative association between recent antibiotic use and immunotherapy outcomes [35, 36].

Sequencing data from 16 patient stool samples collected prior to immunotherapy encompassed 1583 amplicon sequencing variants (ASVs, single DNA sequences). Taxonomic profiles at the genus level showed high relative abundances of *Bacteroides* and firmicutes across all samples (Figure 1A, Supplementary Tables 2–4). Neither alpha nor beta diversity, calculated at the ASV level, differed between responder and non-responder patients.

Ten taxa, including 5 within the family Lachnospiraceae and 4 within Coriobacteriaceae, exhibited significant differences in the mean relative abundance between responders versus non-responders (Figure 1B, $p < 0.05$, unadjusted, Supplementary Table 5). Consistent with previous studies, 6 genera -- including *Akkermansia*, *Parabacteroides* and *Prevotella,* which have been reported as differentially abundant between responders and non-responders, were also found to differ in this cohort though this difference was not statistically significant (Supplementary Figure 1). To examine relationships among individual genera within the responders vs. non-responders, we performed a network analysis which highlighted that bacterial taxa are deeply interconnected in the responder network but less connected in the non-responder network (Supplementary Figure 2), suggesting a need to examine gut immunotherapy microbial predictors at a more global level.

Gut microbial compositions distinguish responders from non-responders in the combined dataset

To expand the potential generalizability of cancer immunotherapy gut microbiome analyses, we validated our findings using three additional published immunotherapy microbiome datasets from melanoma patients that used a similar sequencing strategy (16S rRNA, V4 region) and comparable response status reporting: Matson et al., [16], Gopalakrishnan et al., [14], and Peters et al., [18]

(Supplementary Table 6). The combined dataset was comprised of 128 patient samples (70 responders and 58 non-responders), yielding a total of 5654 unique ASVs (Supplementary Figure 3, Supplementary Tables 7–9). While Chao1 richness was statistically significantly higher in responders as compared to non-responders ($p =$ 0.0041, Wilcoxon rank-sum test), Shannon diversity did not differ significantly ($p = 0.2200$, Wilcoxon rank-sum test) (Figure 2A). All 4 datasets were evenly dispersed in the first and second dimensions of NMDS, suggesting that these datasets can be analyzed together due to a lack of study-specific clustering (Supplementary Figure 4).
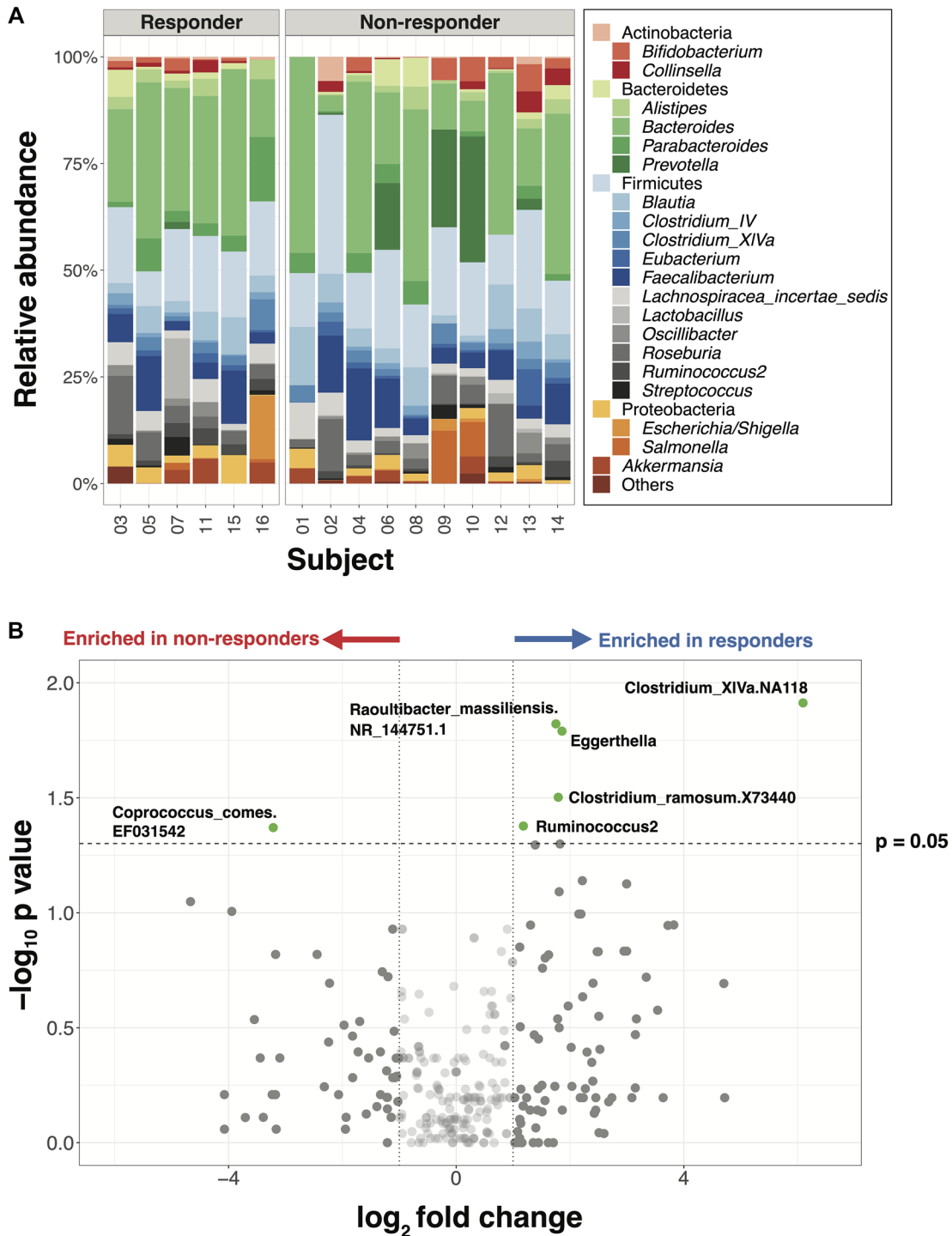

Figure 1: Major gut bacterial taxa of responders and non-responders from the NCI cohort. (**A**) Bar plot of phylogenetic composition of the top 25 bacterial taxa at the genus level, grouped by the response types ($n = 16$). (**B**) Volcano plot with all taxa signals from species to phylum levels. Taxa signals with greater than 2-fold change and statistically significant differences between responders and non-responders are highlighted in green (Wilcoxon rank-sum test, $p < 0.05$, unadjusted).

Analysis of the combined dataset showed statistically significant community differences between the microbiomes of responders and non-responders from phylum to genus levels (Figure 2B). Unsupervised hierarchical clustering of the ASV table based on Bray-Curtis distance at the phylum level clustered patients into two groups with significantly different response rates (cluster1: 0.67, cluster 2: 0.41, p = 0.0030, 2-sided proportion z-test) (Figure 2C). Patients
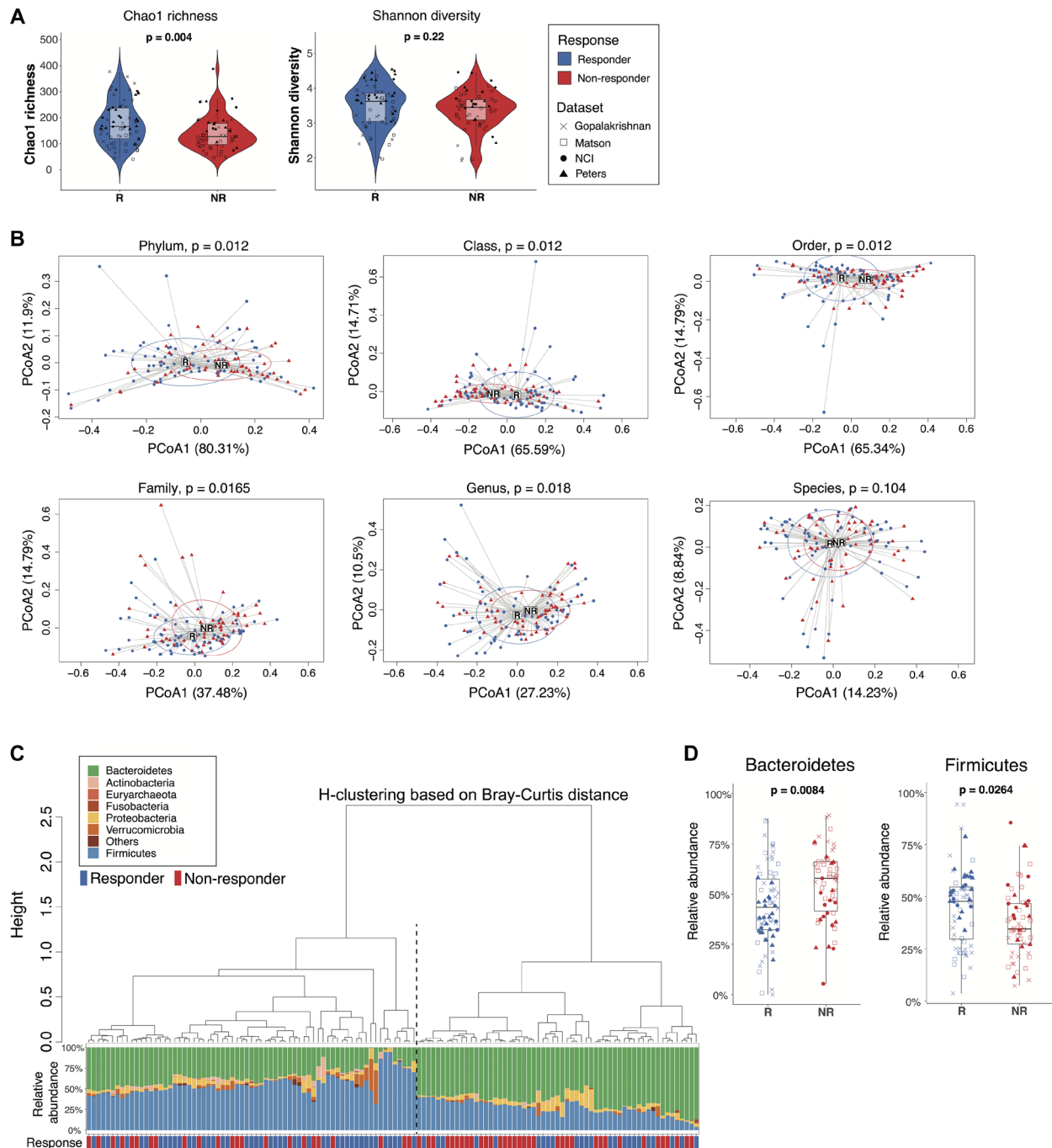

Figure 2: Comparisons of gut microbiome between responders and non-responders from the combined dataset. (A) Alpha diversity of the gut microbiome from responders and non-responders was compared by Wilcoxon rank-sum test (unadjusted p values). Left: Chao1 richness. Right: Shannon diversity. Boxes represent the first and third quartiles. Upper and lower whiskers extend from the box hinge to the largest/smallest value no further than 1.5*IQR. (B) Betadisper permutest plots with Bray-Curtis distance from phylum to species level (top left to bottom right) showing the centroid points and distribution of responder and non-responder sample groups. P values were adjusted with FDR correction. (C) Agglomerative hierarchical clustering of all patient samples using Ward's method with Bray-Curtis distances at the phylum level. Stacked bar plot shows the relative abundances of bacterial phyla for individual patients. Black dotted line separates cluster 1 (higher response rate) from cluster 2 (lower response rate) (p = 0.003, 2-sided proportion z-test). (D) Boxplot of selected taxa with differential relative abundance between responders and non-responders (Wilcoxon rank-sum test, p < 0.05, FDR-corrected).

in cluster 1 with higher relative abundances of Firmicutes had higher response rates, while patients in cluster 2 were enriched in Bacteroidetes and had lower response rates. Hierarchical clustering at the genus level showed a similar pattern of two clusters of patients with different response rates though not statistically significantly different (cluster1: 0.44, cluster 2: 0.62, p = 0.0520, 2-sided proportion z-test) (Supplementary Figure 5). To further evaluate the robustness of these findings, we used Unifrac distance analysis to account for the phylogenetic relationships between taxa. At the genus level, unweighted Unifrac distance, which takes into consideration only the presence or absence of taxa, failed to distinguish responders from non-responders by hierarchical clustering (p = 0.0830, 2-sided proportion z-test) (Supplementary Figure 6A). However, weighted-Unifrac, which accounts for relative abundance differences, showed two significantly different hierarchical clusters (p = 0.0150, 2-sided proportion z-test) (Supplementary Figure 6B).

Because Bacteroidetes and Firmicutes have previously been associated with immunotherapy response rates [12, 13, 19], we performed a targeted analysis of individual phyla from the combined dataset, which demonstrated a statistically significantly higher relative abundance of Bacteroidetes in non-responders and Firmicutes in responders (Figure 2D). A logistic model predicting responder status based on the relative abundances of Bacteroidetes and Firmicutes yielded an ROC curve with an estimated AUC value of 0.67, suggesting moderate prediction accuracy based on these two phylum level signals alone (Supplementary Figure 7). Thus, microbiome analyses of the combined dataset distinguished responders from non-responders and provided insight into taxa associated with these differences.

Variable selection identifies microbial signatures most closely linked to response status

To look beyond individual taxa and more deeply examine groups of taxa (microbial signatures) associated with response status, we utilized *selbal*, an algorithm-based modeling and variable selection method to identify microbial signatures that distinguish responders from non-responders [37]. Of the 20 *selbal*-selected genera, 11 were enriched in responders, while the remaining 9 genera were overrepresented in non-responders (Figure 3A). When analyzed by univariate tests, 7 of the 20 selected genera showed statistically significant differences in relative abundance between responders and non-responders based on unadjusted p values but were not statistically significant after FDR correction (Supplementary Figure 8, Wilcoxon rank-sum test), suggesting that *selbal* provides the advantage of identifying certain taxa that may collectively have associations with response status but would not otherwise be identified through univariate analyses. Among these 7 genera, *Bacteroides, Bilophila*

and *Blautia,* which were overrepresented in non-responders, have been reported in prior studies as enriched in the gut microbiomes of non-responders or patients with shorter progression-free survival (PFS) [14, 18]. *Barnesiella*, Lachnospiraceae_NA2674, *Subdoligranulum*, and *Aestuariispira* were enriched among responders. Consistent with prior studies, we also identified *Bifidobacterium* and *Prevotella* as important predictors of response status using *selbal*, though not statistically significant in univariate analyses [16, 18]. Hierarchical clustering based on Bray-Curtis distances calculated using *selbal*-selected genera clustered patients into two groups with significantly different response rates (cluster1: 0.46, cluster 2: 0.67, p = 0.0200, 2-sided proportion z-test) (Figure 3B). A small group of non-responders within cluster 2 had higher relative abundance of *Prevotella* (Figure 3B), suggesting a potential association between *Prevotella* and poorer outcomes.

Gut microbiome feature modeling predicts clinical response status

We next sought to develop models using microbiome features as clinical response predictors. We included as predictors relative abundances of taxa (species to phylum levels) after removing rare microbes as well as 10 alpha diversity indexes selected to comprehensively represent the community diversity characteristics. The majority of these indexes were weakly correlated with each other based on a Spearman correlation study (Supplementary Table 10, Supplementary Figure 9). Samples from the four combined datasets were randomly divided into a training set (n = 88) and a validation set (n = 40). To identify the model with the highest prediction accuracy, we tested a large number of diverse machine learning algorithms during model training, including penalized logistic regression with ridge, lasso or elastic net (EN, alpha = 0.5), regression tree (RT), random forest (RF), neural network (NN), support vector machine (SVM), as well as the SuperLearner (SL) ensemble-based algorithm. Prediction accuracy of all models was measured and compared across taxonomic levels using the ROC curve and AUC value. Across all models, AUC values generally increased at higher taxonomic ranks (family and above) (Figure 4, Supplementary Table 11). The inclusion of alpha diversity indexes as predictive features led to a modest, though not statistically significant increase (0.042, two-sided P = 0.062, Supplementary Figure 10) in the average AUC (across algorithms and taxa levels). The best prediction accuracy was attained at the class rank with alpha diversity indexes using SVM (AUC = 0.737). Without diversity indexes, ridge regression at the phylum level yielded the best prediction (AUC = 0.717).

Next, we evaluated the accuracy of the prediction equations derived from the training set for predicting tumor response in the validation samples. The AUC

estimates in this external validation tended to be lower than those from the internal cross-validation but remained higher than 0.5 (AUC for a purely random assignment) in most cases (Figure 4, Supplementary Table 12). AUC estimates were generally higher when alpha diversity indexes were added to taxa signals as additional features for prediction, but the difference was not statistically significant. The highest AUC estimates were 0.71 and 0.73 for predictions made with or without alpha diversity indexes, respectively. Interestingly, the models generally performed better at lower taxonomic levels (genus/species) in the validation dataset rather than at the phylum or class level.

Cross-platform validation confirms predictive capacity of the models across sequencing platforms

Because shotgun metagenomic sequencing is an additional method used for analyzing gut microbiomes, we further investigated whether the prediction equations generated with the training set of 16S rRNA V4 sequencing data would predict responses in two independent shotgun metagenomic sequencing datasets from melanoma patients: Frankel et al., [13] (19 responders and 20 non-responders) and McCulloch et al., [38] (19 responders and 8 non-responders). To
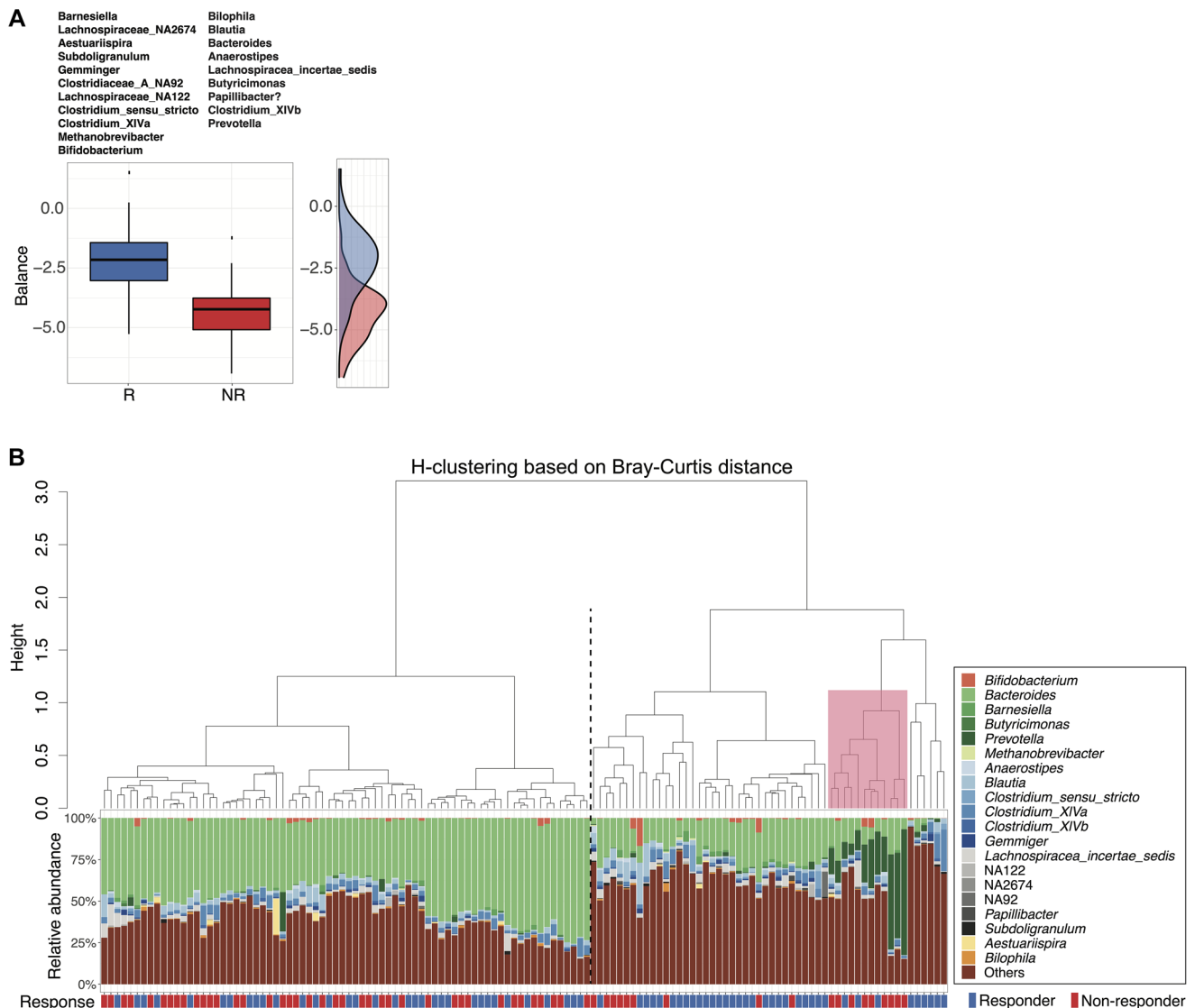

Figure 3: Microbial signatures associated with response status as determined by *selbal* variable selection. (**A**) Boxplots represent the distribution of balance scores in responders versus non-responders. The balance reflects the compositional difference between the two groups of genera selected by *selbal*. Specific genera enriched in each group are listed above the boxplots. A density plot of balance scores for each group is shown on the right. (**B**) Agglomerative hierarchical clustering of all patient samples using Ward's method with Bray-Curtis distances calculated from *selbal*-selected genera. Stacked bar plot shows the relative abundances of the selected genera for individual patients. Black dotted line separates cluster 1 (lower response rate) from cluster 2 (higher response) (*p* = 0.02, 2-sided proportion z-test). Pink box highlights a small group (primarily non-responders) within cluster 2 with higher relative abundance of *Prevotella*.

generate comparable taxonomy variables as the 16S rRNA amplicon sequencing data, we included relative abundances of family to phylum level taxa as features without alpha diversity indexes (Supplementary Table 13). We tested our prediction equations on the two shotgun metagenomic datasets both individually and jointly. AUC estimates were generally higher with McCulloch et al., data, but the difference in prediction accuracy was not statistically significant between datasets (Supplementary Figure 11, Supplementary Table 14). The best prediction accuracy was attained at the order level with the RF algorithm (AUC = 0.73) when testing with McCulloch et al., dataset. The highest AUC estimates for the Frankel et al., dataset and the combined datasets were 0.67 and 0.63, respectively, revealing a weaker yet still meaningful prediction accuracy. Thus, statistical models utilizing gut microbiome community features were robust to a different sequencing platform.

DISCUSSION

In this study, we identified baseline microbiome features associated with immunotherapy response based on a combined analysis of three previously published 16S rRNA gene sequencing datasets from melanoma patients along with our additional cohort of advanced stage solid tumor patients enrolled in NCI immunotherapy trials. By using a tumor-agnostic approach for the discovery cohort and by analyzing these data in combination with larger melanoma datasets, we uncovered microbiome response signals that may be generalizable across tumor types;

however, additional investigation with large-scale cohorts of different tumor types will be needed to confirm these findings. In the meta-analysis, we found that responder microbiomes more clearly separated from non-responder microbiomes at higher taxonomic levels. Whereas previous studies have primarily focused on individual taxa associated with immunotherapy response, we examined more complex community-based interactions using *selbal* analysis and uncovered differentially abundant groups of taxa (i.e., microbial signatures) that associated with clinical response status. To predict immunotherapy response, we developed and validated statistical models using taxonomic variables and community diversity indexes as predictive features. These models had good prediction accuracy in 16S training and validation datasets and, importantly, maintained predictive capacity across sequencing platforms when verified separately on shotgun metagenomic sequencing datasets.

Although several previous studies have identified potential microbiome-based biomarkers for cancer immunotherapy response, prior investigations have primarily focused on melanoma patients and largely inconsistent taxa. Different studies have reported higher relative abundances of *Faecalibacterium*, *Bifidobacterium* or *Akkermansia* in responders [12–14, 16, 18, 19], and *Bacteroides* or Bacteroidales in non-responders [12, 14, 18], yet, no clear consensus signals have emerged. In our meta-analysis, we accounted for multiple variables that could contribute to the inconsistencies observed across studies by geographically restricting our analyses to US-based patient cohorts with consistent responder/non-
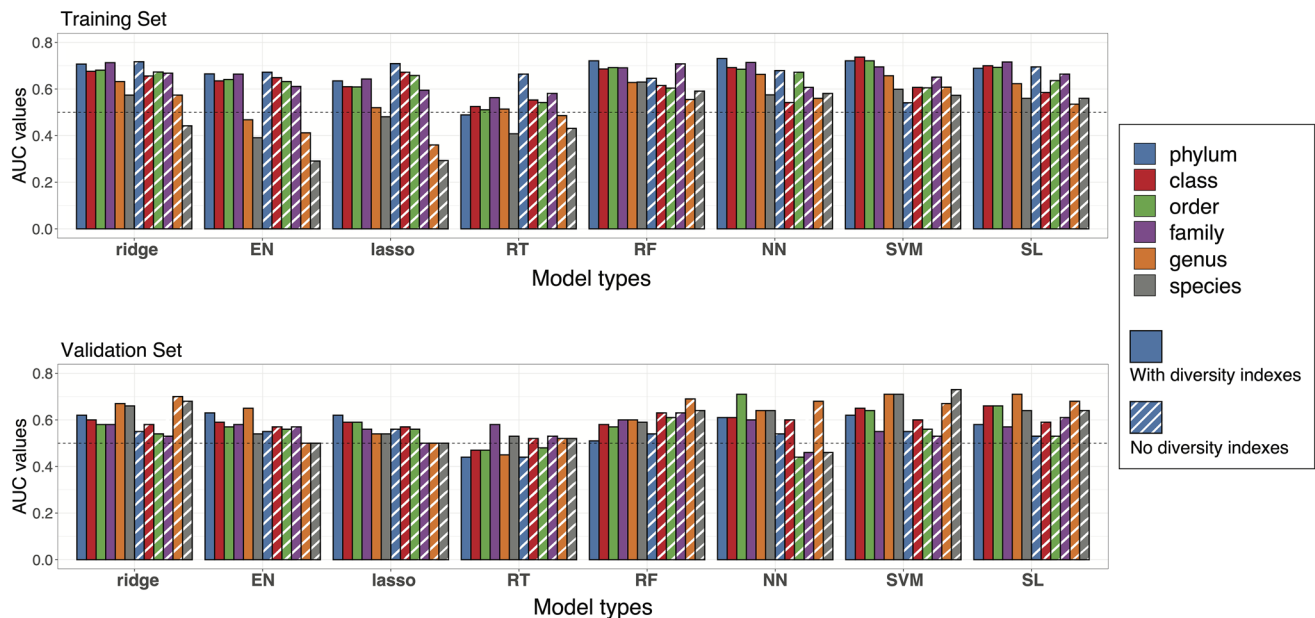

Figure 4: Prediction accuracies of statistical models. Performance of the models at different taxa levels with or without diversity indexes were evaluated by estimated AUC values of ROC curves. Bar plots of AUC values are reported with colors representing different taxa levels and stripes representing the models without diversity indexes. Upper panel: AUC values from the training set in cross validation. Lower panel: AUC values from model testing in the validation set.

responder classification; by including only datasets that used comparable sequencing techniques (16S V4 region); and by re-analyzing previously published datasets using one standardized pipeline. When we initially analyzed the NCI cohort alone, we took an agnostic approach to include multiple tumor types to find common microbiome features. A few bacterial genera were identified as significantly enriched among responders (e.g., *Ruminococcus2* and *Eggerthella)*, and some previously reported genera (*Prevotella* and *Akkermansia*) were also differentially abundant between responders and non-responders but not statistically significant, possibly due to the small sample size [14, 19]. In the meta-analysis, we further validated the identified common microbiome features with larger datasets from melanoma patients. Responder microbiomes separated more distinctly from non-responder microbiomes when analyzed at higher taxonomic levels, indicating that taxonomic rank may be an important variable to consider when searching for microbiome features consistently associated with response status. The lack of significant separation between responders and non-responders at lower taxonomic levels could reflect a prevalence of low abundance species or genera that are not adequately resolved by the sequencing methodology. This could also result from heterogeneity in the patient population and inter-individual variation between subjects, which may be more apparent at lower taxonomic ranks. By using unsupervised hierarchical clustering at the phylum level, we observed a higher response rate in patients enriched in Firmicutes and a lower response rate in patients enriched in Bacteroidetes. Interestingly, recent systemic antibiotic usage has been associated with an increase in the Bacteroidetes/Firmicutes ratio and linked to poorer immunotherapy outcomes [35, 36, 39]. These findings align with the microbiome features associated with non-responders in our analyses.

While individual taxa may strongly influence biological outcomes, we sought to identify more complex microbial community interactions that might improve prediction of immunotherapy response. Thus, we analyzed our combined dataset using *selbal*, a newer method suitable for finding differentially abundant groups of taxa (microbial signatures) in compositional data [37, 40]. *Selbal* analyses identified two groups of genera associated with responders or non-responders, which included several genera discussed in previous publications (i.e., *Bacteroides*, *Bifidobacterium*, *Clostridium*_XIVa, *Blautia* and *Gemmiger*). Importantly, most of the *selbal*-selected taxa were not initially identified using univariate analyses, highlighting the relevance of multivariate analyses in identifying taxa groups, which may collectively associate with response status through direct or functional interactions.

To predict immunotherapy response status, we developed and validated statistical models using global microbiome features and major taxonomic variables.

Using 16S data, our training and validation sets demonstrated a favorable prediction accuracy with the highest AUC value around 0.75, which is equal to or higher than the prediction accuracy reported for several previously published microbiome-based models of immunotherapy response [41–43]. Our models achieved this level of prediction accuracy based solely on the inclusion of taxonomic features and diversity indexes, and they outperformed some previous models that included both taxonomic features and metagenome-derived functional features (e.g., Kyoto Encylopedia of Genes and Genomes/KEGG ortholog differential abundances) [41, 43]. Model prediction accuracy varied substantially across taxonomic levels. Interestingly, taxonomic rank appeared to have a greater effect on prediction accuracy than the chosen machine learning algorithm, indicating that taxonomy level may be an important driver in model performance. Because the inclusion of functional features generally increased model performance in prior studies, the prediction capacity of our models may be even further enhanced by the future inclusion of functional features. Strikingly, our models were additionally validated on microbiome data generated using a shotgun metagenomic sequencing platform with one of two shotgun datasets yielding a robust prediction accuracy. Multiple factors including the technical differences of distinct sequencing platforms and batch effects from different experiments likely affected the effectiveness of the models. However, the cross-platform applicability of our models is encouraging and suggests the possibility of broader generalizability across disparate datasets.

Our study serves as an example of how small patient cohorts can be appropriately integrated into a larger body of published data to elucidate biologically meaningful conclusions. Prior efforts to optimize consistency between studies have focused on matching potential confounding variables between cases and controls at the study design stage [44]. This remains an ideal approach yet may preclude the use of publicly available sequencing data. Based on our meta-analysis, there are several important parameters to consider in future studies to effectively combine distinct datasets for the consistent identification of gut microbiome features associated with disease states or treatment response. These factors include careful selection and inclusion of samples from the following: patients with equivalent classification of disease state or treatment response; patient cohorts from comparable geographic regions; patients with appropriately annotated clinical metadata; datasets generated using equivalent sequencing regions/technologies; and datasets processed and re-analyzed using a single standardized pipeline. Our meta-analysis provides a model for combining distinct datasets in a controlled manner; this methodology may enable biological findings from large-scale analyses to be extrapolated and applied to size-limited patient cohorts in future studies.

Our study is limited by the sample size. While our findings suggest that certain immunotherapy response signals may be generalizable across different tumor types, this must be confirmed in larger patient cohorts. Large-scale investigations are needed to distinguish potential tumor-agnostic response signals from those that may be associated with specific tumor types. An additional limitation of this study is the resolution of our sequencing data when analyzing at lower taxonomy levels. We utilized an ASV clustering method at species and genus levels, and the bacterial variables generated by this method cannot be applied to other sequencing methods. This problem could be rectified by using long read 16S sequencing methodologies, which would likely improve classification accuracy at lower taxonomic levels [45, 46]. The performance of our models may be potentially enhanced by the inclusion of additional clinical variables such as antibiotic usage, body mass index, transcriptomic or proteomic data. The KEGG pathways or meta-transcriptomics features from gut microbiome have been analyzed by several research groups to identify the different patterns between the microbiome of responders and non-responders [18, 27, 29, 42, 43]. Thus, additional microbiome features may be beneficial for future model development.

In conclusion, analyses of our cohort and the combined microbiome dataset have provided a robust assessment of immunotherapy patients' gut microbiomes. The development of reliable models provides additional opportunity to distinguish and predict immunotherapy responders from non-responders. However, the interactions between key microbial taxa and host immunity still need to be elucidated. Ultimately, this research will assist in identifying microbial biomarkers or novel therapeutic targets to improve immunotherapy outcomes and the overall survival of cancer patients.

MATERIALS AND METHODS

NCI patient cohort and sample collection

Adult oncology patients who were enrolled in National Cancer Institute cancer immunotherapy clinical trials (NCT01772004, NCT02517398, NCT02496208) at the National Institutes of Health Clinical Center in Bethesda, MD USA were recruited to participate in gut microbiome sample collection through IRB-approved protocols (NCT00001506, NCT02471352) from 2014 to 2018. Any subject willing to participate and provide stool samples were eligible. All patients enrolled in this study provided written informed consent. Demographic, medical treatment and clinical information are summarized in Supplementary Table 1. Radiographic response to therapy while receiving immunotherapy was determined by treating investigators and assessed using response evaluation criteria in solid tumors (RECIST) v1.1.

Subjects classified by RECIST criteria as having complete response (CR), partial response (PR), or stable disease (SD) for at least 6 months were considered responders. Subjects with progressive disease (PD) or stable disease for less than 6 months were considered non-responders. Response status reflects the best overall response from the start of treatment until disease progression or the end of the study. Mean age between responders and non-responders was compared using Student's t-test. Subjects provided stool samples prior to starting immunotherapy trials. Samples were immediately stored in −80°C freezer until processing.

DNA extraction, PCR amplification, and sequencing

Stool DNA was extracted following the PowerSoil DNA Isolation Kit protocol (Qiagen Cat no. 47014). Final DNA concentration was measured by Qubit BR DNA Kit. Extracted DNA was used for PCR amplification of the 16S rRNA gene V4 region with 515F/806R primer pair (515F: 5′-GTGCCAGCAGCCGCGGTAA-3′, 806R: 5′-GGACTACCAGGGTATCTAAT-3′), which contains adapters for Illumina MiSeq sequencing and sample-specific barcodes. PCR reactions were performed in 30 cycles. PCR products were pooled in equimolar amounts before purification with AMPure xp magnetic beads and then quantified with Qubit BR DNA Kit. Sequencing was performed on the NGS MiSeq platform (Illumina, Inc, San Diego, CA, USA) at NIH Intramural Sequencing Center or NIH Laboratory of Integrative Cancer Immunology, Microbiome and Genetics Core with reagent, sequencing, and collection controls used to test for background and collection contamination. 2 × 250 bp paired end MiSeq run was used to yield forward and reverse reads with close to full overlap.

Sequencing read analyses and taxonomy profiling

Raw sequencing reads were analyzed using the DADA2 pipeline (version 1.16.0) in R (version 4.0.5) to generate exact amplicon sequencing variants (ASV) based on the guidelines at https://benjjneb.github.io/dada2/tutorial.html [47]. Briefly, raw sequences were processed to remove barcodes and quality filtered with default settings (maxN = 0, maxEE = c(2,2), truncQ = 2, rm.phix = TRUE). Based on the quality scores and for the purpose of maximum merging, forward and reverse reads were truncated to be 180 bp and 139 bp and followed by primer trimming. Dereplicating, merging, and chimera removal were processed with default settings. After processing, a seqtab table was generated with rows corresponding to samples and columns corresponding to ASVs.

For taxonomy profiling, ASVs from the seqtab table were aligned against the RefSeq-RDP reference

database (RefSeq-RDP16S_v2_May2018.fa.gz) to generate the taxonomic table assigned to the species level [48]. ASVs which failed to assign at species and genus levels were clustered by sequence similarity to species/genus-level operational taxonomic units (OTU) by kmer package (version 1.1.2) in R (version 3.6.1). Thresholds for clustering were set to 0.99 (for species-level OTUs) and 0.97 (for genus-level OTUs). Output OTUs and representative sequences were combined with assigned species or genera to form the final taxonomic table. For taxonomic tables at family or upper levels, un-assigned ASVs were removed from downstream analyses.

Microbiome community analyses

Read counts from the taxonomic table were converted into relative abundances by dividing the count of each taxon by the sum of read counts within each sample. Alpha diversity indexes were calculated with vegan package (version 2.5.7) or microbiome package (version 1.8.0), and Bray-Curtis dissimilarity distances were calculated via vegan package in R. Both alpha and beta diversity calculations were based on ASV-level relative abundances. Bar-plots and boxplots to visualize the microbiome community difference between responders and non-responders were generated with genus-level relative abundances. To reduce the number of rare microbes and multi-testing, only the taxa with more than 0.1% relative abundance in at least 10% of samples at each taxonomy level from species to phylum were included in the volcano plot. Wilcoxon rank-sum tests with unadjusted as well as false discovery rate (FDR)-corrected p-values were used to determine the taxa with significantly different relative abundance ($p < 0.05$) between responder and non-responder samples.

Literature review and published dataset collection

PubMed was searched for published literature relating to gut microbiome-immunotherapy interactions to find microbiome sequencing data, particularly datasets of 16S rRNA gene sequencing of the V4 region. Based on consideration of sequencing region/techniques and availability of sequencing datasets and patient metadata, 3 eligible studies from Matson et al., [16], Gopalakrishnan et al., [14], and Peters et al., [18] were selected for statistical model development. Sequencing datasets were downloaded from the Sequence Read Archive (SRA) with the accession numbers PRJNA399742 for Matson et al., PRJEB22894 for Gopalakrishnan et al., and PRJNA541981 for Peters et al., Patients' response information was obtained from the study GitHub repository (Matson dataset) (https://github.com/cribioinfo/sci2017_analysis), the SRA data library website (Peters dataset), or kindly received from the authors (Gopalakrishnan dataset).

Response status of pooled samples

Response status after immunotherapy was consistently assigned for patients from selected datasets. Responder and non-responder groups were classified according to RECIST 1.1 based on available metadata. For NCI and Gopalakrishnan datasets, patients were classified as responders if they achieved CR, PR, or SD with PFS no less than 6 months. Patients with PD, or SD with PFS less than 6 months were considered non-responders. For Matson et al., in which PFS information is not available, patients with SD were treated as non-responders according to the publication [16]. Patients with PFS no less than 6 months were considered responders in Peters dataset, as only PFS information is reported. In total there are 128 samples, including 70 responders and 58 non-responders.

Combining sequencing data and taxonomy profiling

Sequencing data were analyzed by DADA2 pipeline as described above. According to the quality scores and pipeline for NCI dataset, forward and reverse reads were truncated to be 140 bp before merging for samples from Matson et al., and Peters et al., Reads for Gopalakrishnan et al., samples were merged already and subsequently truncated to be 254 bp. Dereplicating, merging, and chimera removal were processed with default settings. Seqtab tables with rows corresponding to samples and columns corresponding to ASVs from each DADA2 run were combined into a final seqtab table for taxonomy classification. All ASVs were trimmed to 252 bp and merged by ASV sequence in R during table combining. 5654 unique ASVs were generated after combining all samples from 4 datasets.

ASVs were aligned against the RefSeq-RDP reference database (RefSeq-RDP16S_v2_May2018.fa.gz) to generate the taxonomic table as described above [48]. ASVs that failed to assign at species and genus levels were clustered by the sequence similarity to species/genus-level OTU signals as described above. Output OTU signals and representative sequences were combined with assigned species or genera to form the final taxonomic table. For taxonomic tables at family or upper levels, un-assigned ASVs were removed from downstream analyses.

Diversity calculation

Alpha diversity indexes were calculated with vegan and microbiome package in R with ASV-level relative abundances. For data standardization (maximum method), raw reads were standardized with decostand function from vegan package (method = "max"). Chao1 richness index was calculated with read counts rarefied to 7378 reads per sample, which is the lowest sequencing depth among all samples. Wilcoxon rank-sum tests were performed

to determine the diversity indexes with statistically significant differences ($p < 0.05$) between responders and non-responders.

Beta diversity was assessed based on ASV-level relative abundances using the Bray-Curtis dissimilarity distance calculated via vegan package in R. Non-metric multi-dimensional scaling (NMDS) plot was used to visualize the distance between samples from different datasets.

Statistical analyses

Permutational multivariate analyses of variance (PERMANOVA) test was performed to study the group difference between responder and non-responder samples by considering both the difference of mean values and within-group variation (dispersion) [49]. Only signals with more than 0.1% relative abundance in at least 10% of samples were included at each taxonomy level from species to phylum (final number of taxa features from species to phylum levels are 169, 83, 25, 15, 13 and 7, respectively). The test was performed with adonis function (vegan package) based on Bray-Curtis distance matrices calculated at each taxonomy level. Betadisper plots were created to visualize the centroids of each group and the distribution of data points in the principal coordinates-derived Euclidean space.

Hierarchical agglomerative clustering (i.e., Ward's method) was applied to Bray-Curtis distance matrices or Unifrac distances calculated with filtered information at each taxonomy level with agnes function (cluster package, version 2.1.0). Percentage of responders from 2 major patient clusters were compared using 2-sided proportion z-test. A p-value of 0.05 was set as the cutoff of statistical significance. *P*-values were adjusted using FDR correction except when specified in the text to be unadjusted.

Relative abundance values of Firmicutes and Bacteroidetes were fit into a generalized linear model by glm function in R. Response type was used as the predictor factor with parameter "family = binomal". A receiver operating curve (ROC) reflecting model results was generated by rocit function from ROCit package (version 2.1.1) using the empirical method.

Selection of taxonomy signals

Key genera associated with response types were studied using *selbal* package (version 0.1.0). A forward selection process with 5-fold cross-validation was performed to optimize the taxa variables and maximize the balance score between the two response groups. We used filtered relative abundances at genus level to minimize the number of tests. Variables selected in the global balance table were reported as key genera reflecting microbiome differences between responder and non-responder samples. Boxplot with balance scores was generated from *selbal* analysis

output. Boxplots for each selected genera were created in R (version 3.6.1). Wilcoxon rank-sum tests were performed to determine the genera with statistically significant difference ($p < 0.05$) between responders and non-responders using unadjusted and FDR-corrected *p*-values.

Tumor response prediction using machine learning algorithms

The SuperLearner (SL) package in R was used to make predictions based on the SL and individual algorithms (penalized logistic regression with ridge, lasso or elastic net (EN), as well as regression tree (RT), random forest (RF), neural network (NN), support vector machine (SVM)) via wrapper functions for the glmnet, rpart, randomForest, nnet and svm packages. SL yields an optimized weighted average of the aforementioned algorithms.

The algorithms were evaluated using the training set of 88 patients and internally validated through 10-fold cross validation. The resulting prediction equations were externally validated using the validation set of 40 patients and separate shotgun metagenomic sequencing datasets. A nonparametric bootstrap procedure was used to obtain standard errors for area under the curve (AUC) estimates. For internal cross-validation, the bootstrap procedure was to repeatedly sample 88 patients with replacement from the training set of 88 patients and compute an AUC estimate for each bootstrap sample. In external validations, the prediction equations from the training data were considered fixed, and bootstrap sampling was performed on the validation set of 40 patients or the shotgun metagenomic sequencing dataset.

Shotgun metagenomic sequencing data analyses

PubMed was searched for publications with available shotgun metagenomic sequencing datasets. We analyzed metagenomic sequencing data reported in the study by Frankel et al., from the SRA with accession number SRP115355 [13] and by McCulloch et al., [38] whose data were made available with accession number PRJNA762360. Reads mapping to human genome database (humanGRCh38) were identified and removed using Bowtie2 and samtools programs. Prinseq was used for trimming of low-quality reads and seqtk further split the output files into forward and reverse reads which went through MetaPhlAn3 pipeline for taxonomic profiling [50]. Taxonomy information was compiled for downstream analyses.

Abbreviations

ICIs: immune checkpoint inhibitors; PD-1: programmed cell death protein 1; PD-L1: programmed death-ligand 1; CTLA-4: CTL antigen 4 protein; NCI: National Cancer Institute; ASV: amplicon sequencing variant; NMDS: non-metric multi-dimensional scaling;

ROC: receiver operating curve; PFS: progression-free survival; PERMANOVA: permutational multivariate analyses of variance; SL: SuperLearner; EN: elastic net; RT: regression tree; RF: random forest; NN: neural network; SVM: support vector machine; AUC: area under the curve; KEGG: Kyoto Encyclopedia of Genes and Genomes; OTU: operational taxonomic units.

Author contributions

H.H.K., J.L.G., G. T. and I.B. conceived the study and procured funding. H.H.K. designed and supervised the study. M.T., A.B.A, I.B., J.L.G. and H.H.K performed data and sample collection. J.-H.J. contributed to sample processing. H.L., J.-H.J., Z.Z., M.M., J.H. and Y.C. performed data analyses. H.L., J.-H.J., Z.Z., M.M., J.H., D.P., Y.C., and P.J. contributed to data interpretation. J.A.M., D.D., H.M.Z, A.K.D. and G.T. contributed sequencing data. H.L. and H.H.K. drafted initial manuscript. All authors reviewed the results and approved the final version of the manuscript.

Availability of data and materials

The sequencing data generated in this study are available under NCBI Bioproject PRJNA797954. Supplementary Tables 2 to 14 are also available at https://github.com/skinmicrobiome/gut-microbiome-ML.

CONFLICTS OF INTEREST

Authors have no conflicts of interest to declare.

REFERENCES

1. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer. 2012; 12:252–64. https://doi.org/10.1038/nrc3239. [PubMed]

2. Topalian SL, Drake CG, Pardoll DM. Immune checkpoint blockade: a common denominator approach to cancer therapy. Cancer Cell. 2015; 27:450–61. https://doi.org/10.1016/j.ccell.2015.03.001. [PubMed]

3. Wolchok JD, Chiarion-Sileni V, Gonzalez R, Rutkowski P, Grob JJ, Cowey CL, Lao CD, Wagstaff J, Schadendorf D, Ferrucci PF, Smylie M, Dummer R, Hill A, et al. Overall Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. N Engl J Med. 2017; 377:1345–56. https://doi.org/10.1056/NEJMoa1709684. [PubMed]

4. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, Gonzalez R, Robert C, Schadendorf D, Hassel JC, Akerley W, van den Eertwegh AJ, Lutzky J, et al. Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med. 2010; 363:711–23. https://doi.org/10.1056/NEJMoa1003466. [PubMed]

5. Larkin J, Chiarion-Sileni V, Gonzalez R, Grob JJ, Rutkowski P, Lao CD, Cowey CL, Schadendorf D, Wagstaff J, Dummer R, Ferrucci PF, Smylie M, Hogg D, et al. Five-Year Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. N Engl J Med. 2019; 381:1535–46. https://doi.org/10.1056/NEJMoa1910836. [PubMed]

6. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, Powderly JD, Carvajal RD, Sosman JA, Atkins MB, Leming PD, Spigel DR, Antonia SJ, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. N Engl J Med. 2012; 366:2443–54. https://doi.org/10.1056/NEJMoa1200690. [PubMed]

7. Keenan TE, Burke KP, Van Allen EM. Genomic correlates of response to immune checkpoint blockade. Nat Med. 2019; 25:389–402. https://doi.org/10.1038/s41591-019-0382-x. [PubMed]

8. Pawelec G. Immune correlates of clinical outcome in melanoma. Immunology. 2018; 153:415–22. https://doi.org/10.1111/imm.12870. [PubMed]

9. Hooper LV, Littman DR, Macpherson AJ. Interactions between the microbiota and the immune system. Science. 2012; 336:1268–73. https://doi.org/10.1126/science.1223490. [PubMed]

10. Pouncey AL, Scott AJ, Alexander JL, Marchesi J, Kinross J. Gut microbiota, chemotherapy and the host: the influence of the gut microbiota on cancer treatment. Ecancermedicalscience. 2018; 12:868. https://doi.org/10.3332/ecancer.2018.868. [PubMed]

11. 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. https://doi.org/10.1126/science.aar6918. [PubMed]

12. Chaput N, Lepage P, Coutzac C, Soularue E, Le Roux K, Monot C, Boselli L, Routier E, Cassard L, Collins M, Vaysse T, Marthey L, Eggermont A, et al. Baseline gut microbiota predicts clinical response and colitis in metastatic melanoma patients treated with ipilimumab. Ann Oncol. 2017; 28:1368–79. https://doi.org/10.1093/annonc/mdx108. [PubMed]

13. Frankel AE, Coughlin LA, Kim J, Froehlich TW, Xie Y, Frenkel EP, Koh AY. 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. https://doi.org/10.1016/j.neo.2017.08.004. [PubMed]

14. Gopalakrishnan V, Spencer CN, Nezi L, Reuben A, Andrews MC, Karpinets TV, Prieto PA, Vicente D, Hoffman K, Wei SC, Cogdill AP, Zhao L, Hudgens CW, et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. Science. 2018; 359:97–103. https://doi.org/10.1126/science.aan4236. [PubMed]

15. Hakozaki T, Richard C, Elkrief A, Hosomi Y, Benlaïfaoui M, Mimpen I, Terrisse S, Derosa L, Zitvogel L, Routy B, Okuma Y. The Gut Microbiome Associates with Immune Checkpoint Inhibition Outcomes in Patients with Advanced Non-Small Cell Lung Cancer. Cancer Immunol Res. 2020; 8:1243–50. https://doi.org/10.1158/2326-6066.CIR-20-0196. [PubMed]

16. Matson V, Fessler J, Bao R, Chongsuwat T, Zha Y, Alegre ML, Luke JJ, Gajewski TF. The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. Science. 2018; 359:104–8. https://doi.org/10.1126/science.aao3290. [PubMed]

17. Peng Z, Cheng S, Kou Y, Wang Z, Jin R, Hu H, Zhang X, Gong JF, Li J, Lu M, Wang X, Zhou J, Lu Z, et al. The Gut Microbiome Is Associated with Clinical Response to Anti-PD-1/PD-L1 Immunotherapy in Gastrointestinal Cancer. Cancer Immunol Res. 2020; 8:1251–61. https://doi.org/10.1158/2326-6066.CIR-19-1014. [PubMed]

18. Peters BA, Wilson M, Moran U, Pavlick A, Izsak A, Wechter T, Weber JS, Osman I, Ahn J. Relating the gut metagenome and metatranscriptome to immunotherapy responses in melanoma patients. Genome Med. 2019; 11:61. https://doi.org/10.1186/s13073-019-0672-4. [PubMed]

19. Routy B, Le Chatelier E, Derosa L, Duong CPM, Alou MT, Daillère R, Fluckiger A, Messaoudene M, Rauber C, Roberti MP, Fidelle M, Flament C, Poirier-Colame V, et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. Science. 2018; 359:91–97. https://doi.org/10.1126/science.aan3706. [PubMed]

20. Vétizou M, Pitt JM, Daillère R, Lepage P, Waldschmitt N, Flament C, Rusakiewicz S, Routy B, Roberti MP, Duong CP, Poirier-Colame V, Roux A, Becharef S, et al. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. Science. 2015; 350:1079–84. https://doi.org/10.1126/science.aad1329. [PubMed]

21. Jin Y, Dong H, Xia L, Yang Y, Zhu Y, Shen Y, Zheng H, Yao C, Wang Y, Lu S. The Diversity of Gut Microbiome is Associated With Favorable Responses to Anti-Programmed Death 1 Immunotherapy in Chinese Patients With NSCLC. J Thorac Oncol. 2019; 14:1378–89. https://doi.org/10.1016/j.jtho.2019.04.007. [PubMed]

22. Song P, Yang D, Wang H, Cui X, Si X, Zhang X, Zhang L. Relationship between intestinal flora structure and metabolite analysis and immunotherapy efficacy in Chinese NSCLC patients. Thorac Cancer. 2020; 11:1621–32. https://doi.org/10.1111/1759-7714.13442. [PubMed]

23. Zheng Y, Wang T, Tu X, Huang Y, Zhang H, Tan D, Jiang W, Cai S, Zhao P, Song R, Li P, Qin N, Fang W. Gut microbiome affects the response to anti-PD-1 immunotherapy in patients with hepatocellular carcinoma. J Immunother Cancer. 2019; 7:193. https://doi.org/10.1186/s40425-019-0650-9. [PubMed]

24. Derosa L, Routy B, Fidelle M, Iebba V, Alla L, Pasolli E, Segata N, Desnoyer A, Pietrantonio F, Ferrere G, Fahrner JE, Le Chatelier E, Pons N, et al. Gut Bacteria Composition Drives Primary Resistance to Cancer Immunotherapy in Renal Cell Carcinoma Patients. Eur Urol. 2020; 78:195–206. https://doi.org/10.1016/j.eururo.2020.04.044. [PubMed]

25. Cvetkovic L, Régis C, Richard C, Derosa L, Leblond A, Malo J, Messaoudene M, Desilets A, Belkaid W, Elkrief A, Routy B, Juneau D. Physiologic colonic uptake of ^{18}F-FDG on PET/CT is associated with clinical response and gut microbiome composition in patients with advanced non-small cell lung cancer treated with immune checkpoint inhibitors. Eur J Nucl Med Mol Imaging. 2021; 48:1550–59. https://doi.org/10.1007/s00259-020-05081-6. [PubMed]

26. Salgia NJ, Bergerot PG, Maia MC, Dizman N, Hsu J, Gillece JD, Folkerts M, Reining L, Trent J, Highlander SK, Pal SK. Stool Microbiome Profiling of Patients with Metastatic Renal Cell Carcinoma Receiving Anti-PD-1 Immune Checkpoint Inhibitors. Eur Urol. 2020; 78:498–502. https://doi.org/10.1016/j.eururo.2020.07.011. [PubMed]

27. Wind TT, Gacesa R, Vich Vila A, de Haan JJ, Jalving M, Weersma RK, Hospers GAP. Gut microbial species and metabolic pathways associated with response to treatment with immune checkpoint inhibitors in metastatic melanoma. Melanoma Res. 2020; 30:235–46. https://doi.org/10.1097/CMR.0000000000000656. [PubMed]

28. Li L, Ye J. Characterization of gut microbiota in patients with primary hepatocellular carcinoma received immune checkpoint inhibitors: A Chinese population-based study. Medicine (Baltimore). 2020; 99:e21788. https://doi.org/10.1097/MD.0000000000021788. [PubMed]

29. Sakurai T, De Velasco MA, Sakai K, Nagai T, Nishiyama H, Hashimoto K, Uemura H, Kawakami H, Nakagawa K, Ogata H, Nishio K, Kudo M. Integrative analysis of gut microbiome and host transcriptomes reveals associations between treatment outcomes and immunotherapy-induced colitis. Mol Oncol. 2022; 16:1493–507. https://doi.org/10.1002/1878-0261.13062. [PubMed]

30. Oster P, Vaillant L, Riva E, McMillan B, Begka C, Truntzer C, Richard C, Leblond MM, Messaoudene M, Machremi E, Limagne E, Ghiringhelli F, Routy B, et al. *Helicobacter pylori* infection has a detrimental impact on the efficacy of cancer immunotherapies. Gut. 2022; 71:457–66. https://doi.org/10.1136/gutjnl-2020-323392. [PubMed]

31. Yin H, Yang L, Peng G, Yang K, Mi Y, Hu X, Hao X, Jiao Y, Wang X, Wang Y. The commensal consortium of the gut microbiome is associated with favorable responses to anti-programmed death protein 1 (PD-1) therapy in thoracic neoplasms. Cancer Biol Med. 2021; 18:1040–52. https://doi.org/10.20892/j.issn.2095-3941.2020.0450. [PubMed]

32. Shaikh FY, White JR, Gills JJ, Hakozaki T, Richard C, Routy B, Okuma Y, Usyk M, Pandey A, Weber JS, Ahn J, Lipson EJ, Naidoo J, et al. A Uniform Computational Approach Improved on Existing Pipelines to Reveal Microbiome Biomarkers of Nonresponse to Immune Checkpoint Inhibitors. Clin Cancer Res. 2021; 27:2571–83. https://doi.org/10.1158/1078-0432.CCR-20-4834. [PubMed]

33. Shen YC, Lee PC, Kuo YL, Wu WK, Chen CC, Lei CH, Yeh CP, Hsu C, Hsu CH, Lin ZZ, Shao YY, Lu LC, Liu TH, et al. An Exploratory Study for the Association of Gut Microbiome with Efficacy of Immune Checkpoint Inhibitor in Patients with Hepatocellular Carcinoma. J Hepatocell Carcinoma. 2021; 8:809–22. https://doi.org/10.2147/JHC.S315696. [PubMed]

34. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, Heath AC, Warner B, Reeder J, et al. Human gut microbiome viewed across age and geography. Nature. 2012; 486:222–27. https://doi.org/10.1038/nature11053. [PubMed]

35. Derosa L, Hellmann MD, Spaziano M, Halpenny D, Fidelle M, Rizvi H, Long N, Plodkowski AJ, Arbour KC, Chaft JE, Rouche JA, Zitvogel L, Zalcman G, et al. Negative association of antibiotics on clinical activity of immune checkpoint inhibitors in patients with advanced renal cell and non-small-cell lung cancer. Ann Oncol. 2018; 29:1437–44. https://doi.org/10.1093/annonc/mdy103. [PubMed]

36. Huemer F, Rinnerthaler G, Westphal T, Hackl H, Hutarew G, Gampenrieder SP, Weiss L, Greil R. Impact of antibiotic treatment on immune-checkpoint blockade efficacy in advanced non-squamous non-small cell lung cancer. Oncotarget. 2018; 9:16512–20. https://doi.org/10.18632/oncotarget.24751. [PubMed]

37. Rivera-Pinto J, Egozcue JJ, Pawlowsky-Glahn V, Paredes R, Noguera-Julian M, Calle ML. Balances: a New Perspective for Microbiome Analysis. mSystems. 2018; 3:e00053–18. https://doi.org/10.1128/mSystems.00053-18. [PubMed]

38. McCulloch JA, Davar D, Rodrigues RR, Badger JH, Fang JR, Cole AM, Balaji AK, Vetizou M, Prescott SM, Fernandes MR, Costa RGF, Yuan W, Salcedo R, et al. Intestinal microbiota signatures of clinical response and immune-related adverse events in melanoma patients treated with anti-PD-1. Nat Med. 2022; 28:545–56. https://doi.org/10.1038/s41591-022-01698-2. [PubMed]

39. Panda S, El khader I, Casellas F, López Vivancos J, García Cors M, Santiago A, Cuenca S, Guarner F, Manichanh C. Short-term effect of antibiotics on human gut microbiota. PLoS One. 2014; 9:e95476. https://doi.org/10.1371/journal.pone.0095476. [PubMed]

40. Susin A, Wang Y, Lê Cao KA, Calle ML. Variable selection in microbiome compositional data analysis. NAR Genom Bioinform. 2020; 2:lqaa029. https://doi.org/10.1093/nargab/lqaa029. [PubMed]

41. Gharaibeh RZ, Jobin C. Microbiota and cancer immunotherapy: in search of microbial signals. Gut. 2019; 68:385–88. https://doi.org/10.1136/gutjnl-2018-317220. [PubMed]

42. Limeta A, Ji B, Levin M, Gatto F, Nielsen J. Meta-analysis of the gut microbiota in predicting response to cancer immunotherapy in metastatic melanoma. JCI Insight. 2020; 5:e140940. https://doi.org/10.1172/jci.insight.140940. [PubMed]

43. Heshiki Y, Vazquez-Uribe R, Li J, Ni Y, Quainoo S, Imamovic L, Li J, Sørensen M, Chow BKC, Weiss GJ, Xu A, Sommer MOA, Panagiotou G. Predictable modulation of cancer treatment outcomes by the gut microbiota. Microbiome. 2020; 8:28. https://doi.org/10.1186/s40168-020-00811-2. [PubMed]

44. Vujkovic-Cvijin I, Sklar J, Jiang L, Natarajan L, Knight R, Belkaid Y. Host variables confound gut microbiota studies of human disease. Nature. 2020; 587:448–54. https://doi.org/10.1038/s41586-020-2881-9. [PubMed]

45. Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, Leopold SR, Hanson BM, Agresta HO, Gerstein M, Sodergren E, Weinstock GM. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat Commun. 2019; 10:5029. https://doi.org/10.1038/s41467-019-13036-1. [PubMed]

46. Jeong J, Yun K, Mun S, Chung WH, Choi SY, Nam YD, Lim MY, Hong CP, Park C, Ahn YJ, Han K. The effect of taxonomic classification by full-length 16S rRNA sequencing with a synthetic long-read technology. Sci Rep. 2021; 11:1727. https://doi.org/10.1038/s41598-020-80826-9. [PubMed]

47. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016; 13:581–83. https://doi.org/10.1038/nmeth.3869. [PubMed]

48. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 2014; 42:D633–42. https://doi.org/10.1093/nar/gkt1244. [PubMed]

49. Anderson MJ, Ellingsen KE, McArdle BH. Multivariate dispersion as a measure of beta diversity. Ecol Lett. 2006; 9:683–93. https://doi.org/10.1111/j.1461-0248.2006.00926.x. [PubMed]

50. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, Mailyan A, Manghi P, Scholz M, Thomas AM, Valles-Colomer M, Weingart G, Zhang Y, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife. 2021; 10:e65088. https://doi.org/10.7554/eLife.65088. [PubMed]