Gut microbiota signatures are associated with toxicity to combined CTLA-4 and PD-1 blockade

Jennifer Wargo (jwargo@mdanderson.org)
MD Anderson https://orcid.org/0000-0003-3438-7576

Miles Andrews
Monash University https://orcid.org/0000-0003-1231-8641

Connie Duong
Gustave Roussy Cancer Campus https://orcid.org/0000-0002-9067-8557

Vancheswaran Gopalakrishnan
University of Texas MD Anderson Cancer Center

Valerio Iebba
Institut Gustave Roussy https://orcid.org/0000-0002-0716-306X

Wei-Shen Chen
University of South Florida Morsani College of Medicine

Lisa Derosa
Institut Gustave Roussy, 94800 Villejuif, France; Departement de Médicine Oncologique, Gustave Roussy Cancer Campus, 94800 Villejuif, France. https://orcid.org/0000-0003-0527-2964

Bertrand Routy
Gustave Roussy Cancer Center

Gladys Ferrere
Gustave Roussy Cancer Center

Aurélie Fluckiger
Gustave Roussy Cancer Center

Maria Roberti
Deutsches Krebsforschungszentrum DKFZ https://orcid.org/0000-0001-5788-6840

Paule Opolon
Gustave Roussy Cancer Center

Whijae Roh
Broad Institute https://orcid.org/0000-0002-7395-9939

Christine Spencer
Parker Institute for Cancer Immunotherapy

Irina Fernandez Curbelo
University of Texas MD Anderson Cancer Center

Luis Vence
MD Anderson Cancer Center https://orcid.org/0000-0003-3441-4095

Alexandre Reuben
The University of Texas MD Anderson Cancer Center https://orcid.org/0000-0003-4510-0382
Zachary Cooper
AstraZeneca https://orcid.org/0000-0003-1059-0940
Peter Prieto
University of Rochester Medical Center
M. A. Wadud Khan
University of Texas M.D. Anderson Cancer Center https://orcid.org/0000-0003-0596-862X
Alexander Lazar
The University of Texas MD Anderson Cancer Center https://orcid.org/0000-0002-6395-4499
Michael Tetzlaff
University of California San Francisco
Courtney Hudgens
The University of Texas MD Anderson Cancer Center https://orcid.org/0000-0001-8312-7485
Pierre-Olivier Gaudreau
MD Anderson Cancer Center, Department of Surgical Oncology
Luigi Nezi
Istituto Europeo di Oncologia https://orcid.org/0000-0002-4670-7656
Didier Raoult
Aix Marseille Université Méditerranée Infection Foundation https://orcid.org/0000-0002-0633-5974
Lauren Haydu
The University of Texas MD Anderson Cancer Center
Hussein Tawbi
University of Texas MD Anderson Cancer Center https://orcid.org/0000-0003-1942-851X
Patrick Hwu
The University of Texas MD Anderson Cancer Center
Wen-Jen Hwu
University of Texas MD Anderson Cancer Center
Rodabe Amaria
University of Texas M.D. Anderson Cancer Center https://orcid.org/0000-0001-6876-278X
Elizabeth Burton
MD Anderson Cancer Center, Department of Surgical Oncology
Scott Woodman
The University of Texas MD Anderson Cancer Center
Adi Diab
University of Texas M.D. Anderson Cancer Center
Sapna Patel
MD Anderson Cancer Center
Isabella Glitza

University of Texas M.D. Anderson Cancer Center

Jianhua Zhang

The University of Texas MD Anderson Cancer Center https://orcid.org/0000-0001-5412-9860

Nadim Ajami

The University of Texas MD Anderson Cancer Center

Joseph Petrosino

Baylor College of Medicine

Robert Jenq

Adult Bone Marrow Transplantation Service, Department of Medicine, Memorial Sloan Kettering Cancer Center

Michael Davies

MD Anderson https://orcid.org/0000-0002-0977-0912

Jeffrey Gershenwald

MD Anderson

Padmanee Sharma

The University of Texas MD Anderson Cancer Center https://orcid.org/0000-0003-4658-055X

James Allison

The University of Texas MD Anderson Cancer Center

Andrew Futreal

The University of Texas MD Anderson Cancer Center https://orcid.org/0000-0001-8663-2671

Laurence Zitvogel

Inserm U1015, Gustave Roussy https://orcid.org/0000-0003-1596-0998

Maryam TIDJANI ALOU

CNRS

Satoru Yonekura

Inserm U1015, Gustave Roussy https://orcid.org/0000-0003-4047-3016

Alexandria Cogdill

The University of Texas MD Anderson Cancer Center

Reetakshi Arora

The University of Texas MD Anderson Cancer Center

Latasha Little

The University of Texas MD Anderson Cancer Center

Curtis Gumbs

The University of Texas MD Anderson Cancer Center

Khalida Wani

M. D. Anderson Cancer Center

Margaret Callahan

Memorial Sloan Kettering Cancer Center https://orcid.org/0000-0002-9087-0012

Mathew Adamow

Memorial Sloan Kettering Cancer Center
Michael Postow
MSKCC
Charlotte Ariyan
Memorial Sloan-Kettering Cancer Center
Julie Gardner
MD Anderson
Jennifer McQuade
Department of Melanoma Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX https://orcid.org/0000-0002-2393-2172
Michael Wong
The University of Texas MD Anderson Cancer Center
Li Zhao
The University of Texas MD Anderson Cancer Center

Article

Keywords: cancer, immunotherapy, checkpoint inhibitors, genomics, whole exome sequencing, lymphocyte infiltrate, microbiome, immune-related toxicity, PD-1, CTLA-4, B. intestinalis

Posted Date: December 7th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-119925/v1

Version of Record: A version of this preprint was published at Nature Medicine on July 8th, 2021. See the published version at https://doi.org/10.1038/s41591-021-01406-6.

Abstract

Treatment with combined immune checkpoint blockade (CICB) targeting CTLA-4 and PD-1 is associated with clinical benefit across tumor types, but a high rate of immune-related adverse events (irAE). Insights into biomarkers and mechanisms of response and toxicity to CICB are needed. To address this, we profiled the blood, tumor and gut microbiome of 77 advanced melanoma patients treated with CICB, with a high rate of any ≥Grade 3 irAEs (49%) with parallel studies in pre-clinical models. Tumor-associated immune and genomic biomarkers of response to CICB were similar to those identified for ICB monotherapy, and toxicity from CICB was associated with a more diverse peripheral T cell repertoire. Profiling of gut microbiota demonstrated a significantly higher abundance of Bacteroides intestinalis in patients with toxicity, with upregulation of mucosal IL-1b in patient samples of colitis and in pre-clinical models. Together, these data offer potential new therapeutic angles for targeting toxicity to CICB.

Main Text

Treatment with combined immune checkpoint blockade (CICB) targeting CTLA-4 and PD-1 is associated with clinical benefit across tumor types, but a high rate of immune-related adverse events (irAE). Insights into biomarkers and mechanisms of response and toxicity to CICB are needed. To address this, we profiled the blood, tumor and gut microbiome of 77 advanced melanoma patients treated with CICB, with a high rate of any ≥Grade 3 irAEs (49%) with parallel studies in pre-clinical models. Tumor-associated immune and genomic biomarkers of response to CICB were similar to those identified for ICB monotherapy, and toxicity from CICB was associated with a more diverse peripheral T cell repertoire. Profiling of gut microbiota demonstrated a significantly higher abundance of *Bacteroides intestinalis* in patients with toxicity, with upregulation of mucosal IL-1b in patient samples of colitis and in pre-clinical models. Together, these data offer potential new therapeutic angles for targeting toxicity to CICB.

Treatment with CICB is associated with high rates of objective responses[1], however a substantial proportion of patients experience immune-related adverse events (irAE)[2,3]. Interestingly, rates of clinical response and irAE appear to be linked[4,5], though distinct mechanisms behind therapeutic toxicity are incompletely understood. Robust biomarkers of response to CICB are currently lacking, and it is likely that up to 40% of unselected melanoma patients treated with CICB would be expected to respond to PD-1 blockade alone, and thus could potentially be spared the increased risk of severe irAE associated with this regimen[1,6,7].

To help address this issue, we studied biomarkers of response and toxicity to CICB in a cohort of 77 patients with advanced predominantly cutaneous-type melanoma receiving CICB, either on clinical trials or as standard of care therapy (**Fig. 1a**; **Supplementary Data Table 1**). The majority of patients had stage IV disease (n=65, 84%), and were naïve to prior systemic therapy (n=57, 74%) (**Supplementary Data Table 1**). In this cohort, the rate of any grade irAE was high (n=72, 93.5%) and nearly half of patients (49%)

experienced severe (≥ Grade 3) irAEs (**Supplementary Data Table 2**), which is in line with other published series[1,3,8-10]. Progression-free survival was markedly greater in patients with cutaneous-type melanomas than mucosal or uveal melanomas (**Extended Data Fig. 1a**.), hence response-based analyses were restricted to the cutaneous-type cohort within which there was a trend towards a coupling of response and toxicity (p=0.10, Chi-square test). Despite heterogeneity in terms of melanoma subtype, disease stage, and extent of prior treatment, we expected biomarkers of toxicity to be widely applicable and thus did not restrict our cohort under consideration for this outcome.

We first performed whole-exome sequencing in available pre-treatment tumor samples (**Supplementary Data Table 3**) to assess the association of total mutational burden (TMB) with response to CICB (n=26). We observed higher TMB in responders (R, n=20) compared with non-responders (NR, n=6) to CICB overall (**Fig. 1b**, p=0.20), consistent with findings from prior studies[11-14]. The sample size was admittedly limited, however there appeared to be 2 subsets within the responder population; one with a high TMB for whom anti-PD-1 monotherapy may have been sufficient; and one with a lower TMB in whom CICB was likely more necessary. Qualitative assessment of the mutational landscape in this cohort did not reveal significant differences in the frequency of mutations in common melanoma driver, IFN-γ-pathway and antigen processing pathway gene sets in R vs NR (**Extended Data Fig. 1b**). Similarly, no significant differences were noted in neoantigen burden between R and NR (**Fig. 1c,d**; **Extended Data Fig. 1c**). However, tumor samples were only available for a relatively small proportion of patients, limiting the ability to draw strong conclusions from this data.

We next assessed the association between copy number loss and response to CICB based on findings from our group and others suggesting that a high burden of copy number loss was associated with resistance to treatment with sequential checkpoint blockade targeting CTLA-4 and PD-1[15]. In the current cohort, we observed that NR to CICB had a significantly higher burden of copy number loss compared with R (p=0.04; **Fig. 1e**). Resistance to CICB was principally associated with copy number loss affecting chromosomes 5, 10 and 15 (**Extended Data Fig. 1d,e**). Several genes previously implicated in resistance to immune checkpoint blockade monotherapies appeared either exclusively (*CD74*) or disproportionately (*PDIA3*, *B2M*, *PTEN*) affected by copy number losses in NR tumors (**Fig. 1f**, **Extended Data Fig. 1f**), suggesting potential immuno-genomic mechanisms of resistance to CICB[15-19].

Next, we assessed the density of CD8+ T cells in baseline tumor biopsies of R versus NR to CICB given prior studies highlighting the prognostic significance of the density and distribution of CD8+ T cells in response to ICB monotherapy[17,20]. In these analyses, a higher density of CD8+ T cells was observed in tumors of R compared with NR, (n=19 R, n=6 NR, p=0.052, one-sided; **Fig. 1g**), We also assessed the T cell

repertoire via TCR sequencing in baseline tumor samples of R versus NR to CICB, and noted that T cell repertoire entropy was higher in R to CICB (p=0.058; **Extended Data Fig. 1g**).

Given the growing evidence regarding the role of the gut microbiota in response to checkpoint blockade monotherapy[21-23], we next assessed the association of gut microbiota signatures with response to CICB. Importantly, we assessed profiles in patients on CICB and performed studies in pre-clinical models for cross-species validation of putative microbial contributors to response and/or toxicity. To do this, we first profiled baseline fecal microbiome samples in patients on CICB using 16S rRNA gene sequencing (n=54; **Supplementary Data Table 3, Extended Data Fig. 2a**). Assignment of taxonomy at the species level was done by aligning the representative centroid sequence of each operational taxonomic unit (OTU) against the NCBI 16S reference database with BLAST. Species names were only considered for OTUs that had an unambiguous annotation, (**Supplementary Data Table 4**), while all other OTUs were treated as unclassified at the species level. The taxa Firmicutes (phylum) and Clostridiales (order) were previously shown to be response-associated in the context of PD-1 blockade[21] but displayed similar abundance between R and NR patients to CICB (p=0.39, p=0.38, respectively; **Extended Data Fig. 2b,c**). In this cohort we did not observe any significant differences in alpha diversity in R versus NR (**Extended Data Fig. 2d**) which is in contrast to our previous findings in patients on anti-PD-1 monotherapy, which may be related to the limited sample size but might also suggest that gut microbiome diversity is less critical in the setting of treatment with CICB. We next focused on novel candidate discovery by identifying potential candidate taxa associated with response by assessing compositional differences between R and NR using LEfSe (**Fig. 2a**) and pairwise comparisons (**Extended Data Fig. 3a**), and identified several differentially-enriched bacterial taxa in R, including *Bacteroides stercoris, Parabacteroides distasonis,* and *Fournierella massiliensis* (p=0.03, p=0.04, p=0.008, respectively by Mann-Whitney tests). Taxa enriched in NR included *Klebsiella aerogenes* and *Lactobacillus rogosae*, among others (p=0.04, p=0.02 respectively; **Supplementary Data Table 5**). Importantly, given the limitations of assigning taxonomy at the species level using 16S rRNA gene sequencing data, we sought to verify our results using whole metagenomic shotgun sequencing (WMS) which was performed on a subset of patients (n=27 R, n=11 NR). Employing Bray-Curtis dissimilarities, the Procrustes analysis revealed that taxonomic annotations at the species level by 16S and WMS methods produced similar clustering patterns (Mantel: r=0.65, p=0.001; **Extended Data Figure 3b**), suggesting a high degree of concordance between the species identification by either method. These studies also confirmed previously identified associations with *B. stercoris* (p=0.07) and *P. distasonis* (p=0.024) (**Fig. 2b**). Furthermore, we also observed a strong correlation between species abundances estimated using 16S versus WMS (Spearman's rho=0.934, p=2.2e-16; **Extended Data Fig. 3c**). In order to account for the effect of known clinical predictors of response, we next performed multivariable logistic regression using species-level abundances from WMS of our top bacterial candidates (**Extended Data Fig. 3d**) and continued to observe a positive association with response (*B. stercoris*, adjusted OR=4.76, 95% C.I.= 0.93 – 36.5; **Supplementary Data Table 6**).

We next examined candidate taxa associated with response to CICB in pre-clinical syngeneic transplantable tumor models (**Extended Data Fig. 4a**). In these studies, treatment with CICB was associated with tumor regression in two tumor models compared with control treated mice (**Fig. 2c**). We then interrogated fecal microbial composition in these models to determine if this was associated with therapeutic benefit. We performed a supervised analysis (Partial Least Squares Discriminant Analysis; PLS-DA) to explore the variance in the microbiome composition at T0 comparing mice eventually tumor-bearing versus tumor-free mice at T2 (sacrifice after 2 systemic injections of CICB), and observed a clear discrimination between the two groups (p=0.001; **Fig. 2d**). The relative contribution of each bacterial species abundance at T0 to the observed group separation was next assessed using the PLS-DA-derived variable importance (VIP) score, revealing that *Parabacteroides distasonis* was predictive of response to CICB (**Fig. 2e**), which was also observed in melanoma patients (**Fig. 2a**). Importantly, the relative abundance of *P. distasonis* was significantly higher in mice that became eventually tumor-free post-CICB (**Fig. 2f)** and was negatively correlated with tumor size (**Extended Data Fig. 4b**). Together, these data identify associations between distinct commensal species such as *P. distasonis* and beneficial tumor responses to CICB, with some overlap noted between taxa identified in murine tumor models and in our patient cohort.

Following assessment of gut microbiota and efficacy in mice, we next assessed the relationship between the gut microbiota and toxicity (colitis and ileitis) in our murine models. Consistent with previous studies utilizing immunocompetent mice, CICB did not trigger overt colitis-associated manifestations such as diarrhea or weight loss. However, histologic abnormalities of the gut epithelium and lamina propria pathognomonic of sub-clinical toxicity such as villi shortening in ilea, crypt or mucosal attritions, and inflammatory infiltrates could be scored and analyzed in ilea and colons. These inflammatory changes were then assessed with regard to the microbial composition via 16S rRNA gene sequencing of fecal samples at T0. In these studies, treatment with CICB was associated with subclinical ileitis in both MCA and RET tumor models (MCA p=0.042, RET p=0.024; **Fig. 3a,b**). Given known evidence that commensal microbes can induce colitis via IL-1β[24], we treated mice concurrently with an IL-1R antagonist (anakinra) which is approved for the treatment of patients with rheumatoid arthritis. Treatment with CICB and concurrent IL-1R antagonism was associated with significantly less inflammation (MCA p=0.0094, RET p=0.0009; **Fig. 3a,b**). Pre-treatment with oral antibiotics was also associated with reduced ileitis in MCA205 animals, however the magnitude of reduced inflammation was modest in comparison (MCA p=0.032, RET p=0.25; **Fig. 3a,b**). Colonic inflammation (as opposed to ileitis) was only observed in the RET model in tumor-bearing mice (p=0.018; **Fig. 3c**), and principal coordinate analyses of fecal microbial beta diversity demonstrated a clear association between the bacterial repertoire at day 7 post-tumor inoculation (T0) and subsequent subclinical colitis in this model (**Fig. 3d, Extended Data Fig. 4c**).

To confirm that intestinal inflammation associated with CICB was associated with increased IL-1B, we performed transcriptional profiling of the ilea and colons of tumor-bearing mice 24 hours after the first systemic administration of CICB. The development of intestinal inflammation (diagnosed by pathological scoring) was accompanied by a prompt and selective transcriptional upregulation of *Il1b*, but not *Tnfa* or *Il6*, and only in the presence of intact gut microflora (MCA p=0.0032**,** RET p=0.0016; **Fig. 3e,f**). Importantly, pharmacological inhibition of IL-1R using anakinra mitigated CICB-induced intestinal inflammation in both tumor models (MCA p=0.0094, RET p=0.0009; **Fig. 3a,b**) without compromising the anti-tumor efficacy of CICB (p<0.0001; **Fig. 3g**). We validated these findings in a cohort of patients who developed colitis (grade 3 or 4) while on immune checkpoint blockade. To do this, expression levels of *IL1B*, *IL17*, and *TNF* were assessed in colonic biopsies from affected patients and were compared to normal colon tissue samples from patients undergoing surgery (**Supplementary Data Table 7**). We observed higher expression of *IL1B* and *IL17* in colitis versus normal tissue but did not observe differences in *TNF* expression (*IL1B* p=0.042*, IL17* p=0.041; **Fig. 3h-j**).

We next assessed associations between gut microbial signatures and toxicity to CICB in our patient cohort. No significant differences were noted in the overall alpha diversity of the gut microbiome in patients who did or did not develop grade ≥3 irAEs (**Extended Data Fig. 5a**). However, discovery analysis revealed that several bacterial taxa were differentially enriched in baseline gut microbiome samples of patients who did develop grade ≥3 irAEs versus those who did not, including *Bacteroides intestinalis* and *Intestinibacter bartlettii* (p=0.009, p=0.009 by Mann-Whitney tests, respectively; **Supplementary Data Table 8**, **Extended Data Figure 5b**), albeit with a small fold-change for the latter. Taxa enriched in patients who did not develop grade ≥3 irAEs were also identified, including *Anaerotignum lactatifermentans* and *Dorea formicigenerans* by LEfSe and pairwise comparisons (p=0.016 and p=0.06 respectively; **Fig. 4a**, **Supplementary Data Table 8**, **Extended Data Figure 5b**). We confirmed the concordance of species level datasets from 16S and WMS methods by Procrustes analysis (Mantel: r=0.665, p=0.001; **Extended Data Figure 5c**). In addition, confirmation of bacterial candidates was done using WMS (**Fig. 4b**, **Extended Data Figure 5d**) and a strong positive correlation was observed between abundance of *B. intestinalis* quantified using 16S versus WMS (Spearman's rho=0.62, p=4.2e-6, **Extended Data Figure 5e**, **Supplementary Data Table 8**). Interestingly, associations were noted between gut microbiota signatures and peripheral immune profiles of patients developing grade ≥3 irAEs versus grade <3 irAEs (**Fig. 4c**). As before, associations were maintained after adjusting for clinical predictors in a multivariable logistic regression model using species-level abundances estimated from WMS (*B. intestinalis*: adjusted-OR= 4.54 (95% C.I.= 1.06 – 24.7)*; D. formicigenerans*: adjusted-OR= 0.35 (95% C.I.= 0.082 – 1.35); **Supplementary Data Table 6**).

We then examined associations between CICB-associated toxicity and systemic immune profiles in available clinical samples to identify signatures in the peripheral blood for patients at higher risk of irAEs

using multi-parameter flow cytometry phenotyping of peripheral blood mononuclear cells (PBMCs) in patients at baseline and on-treatment. In these studies, we observed a higher proliferative index in effector and central memory CD8+ T lymphocytes at early on-treatment time points in patients who developed grade ≥3 irAEs (p=0.0044, p=0.013, respectively, n=14; **Extended Data Fig. 6a,b**), potentially suggesting that accelerated expansion of cytotoxic T cells could contribute to immune-related toxicity. We then assessed the expression of CD28 and CD27 in T cell subsets of these patients (gating, **Extended Data Fig. 6c**), as these markers are known to be progressively down-regulated in antigen-experienced T cells adopting a distinct "aged" functional state[25,26]. In these analyses, we observed significantly lower expression of surface CD28 and CD27 on circulating CD4+ and CD8+ effector T lymphocytes, respectively, of patients who did not develop severe irAEs in a discovery cohort (CD27 in CD4 Teff, p=0.0022; CD28 in CD4 Teff, p=0.014; CD27 in CD8 Teff, p=0.072; CD28 in CD8 Teff, p=0.04; **Extended Data Fig. 6d,e left panels**). We assessed associations between these markers and toxicity in a second cohort of melanoma patients on CICB (**Supplementary Data Table 9**) and modest trends were noted **(Extended Data Fig. 6d,e right panels**), although methods of subset enumeration were non-identical between these cohorts. TCR sequencing of PBMCs in our own cohort revealed a significantly higher diversity of the baseline T cell repertoire in patients who subsequently experienced high-grade irAEs (p=0.028, n=24; **Extended Data Fig. 6f**), in line with previous reports on checkpoint blockade monotherapy[27,28]. Polyclonal expansion of T cell clones from baseline to on-treatment was also observed, with patients experiencing grade ≥3 toxicity on CICB having expansion of ≥55 circulating $CD8^+$ T cell clones compared to those with grade <3 irAE (p=0.22, **Extended Data Fig. 6g**), as observed with ICB monotherapy targeting CTLA-4[28]. Together this data suggests that peripheral immune signatures may be present that are associated with a higher risk of developing toxicity, though these clearly need optimization and validation in additional and larger cohorts.

Following this, we sought to establish potentially causal relationships between intestinal enrichment of the distinct commensals identified in our studies and intestinal inflammation during CICB. To do this, we gavaged mice with three different strains of *B. intestinalis* following gut decontamination with antibiotic treatment (ATB). Increased ileal damage was observed *in B. intestinalis*-colonized mice compared with mice allowed to spontaneously repopulate their intestinal microflora following ATB (p=0.0021; **Fig. 4d**). Gavage with *B. intestinalis* was also associated with induction of ileal *Il1b* transcription (p=0.0025; **Fig. 4e**). Interestingly, CICB facilitated the over-representation of *B. intestinalis* over other *Bacteroides* species such as *B. fragilis* or *B. uniformis,* as assessed by quantitative PCR in the feces of mice following four systemic injections of CICB (p=0.0068; **Fig. 4f,g).** To further confirm the impact of CICB in favoring ileal colonization with *B. intestinalis* and thereby promoting ileal *Il1b* transcription, we utilized the avatar mouse model system[29] in which FMT was performed in ATB-treated mice (**Fig. 4h**), using fecal material from three healthy human donors harboring either low or high endogenous levels of *B. intestinalis* (**Fig. 4i**). Three weeks post-FMT, the colonization of the recipient mouse intestines with *B. intestinalis* corresponded with their donor and was thus significantly higher in feces of mice receiving *B. intestinalis*-

high donor FMT than mice receiving *B. intestinalis*-low donor FMT (p<0.0001; **Fig. 4i**). In this model system, we again observed *B. intestinalis*-associated induction of ileal *Il1b* expression after systemic injections of CICB (p=0.0287; **Fig. 4j**).

Taken together, these studies build on prior findings in immune checkpoint monotherapies to identify novel biomarkers of response and irAE in the context of CICB, to which unique features may apply. Many of the predictive factors for checkpoint blockade monotherapy are also confirmed for CICB response and resistance (including TMB, $CD8^+$ T cell density, and burden of copy number loss) and toxicity (more diverse TCR repertoire and more naive T cell phenotype). Our results were limited in part by small numbers of evaluable samples for some analyses, including matched fecal microbiome and peripheral blood lymphocyte phenotyping. Due to this, we were limited in our ability to generate comprehensive predictive models of response and toxicity incorporating all known and novel clinical/biomarker parameters and as such, are unable to estimate the relative contribution of any single factor to clinical outcomes. Accordingly, additional studies of larger cohorts of ICB-treated patients will be needed to address the complexities of how the multiple tumor, microenvironment, host and microbiome factors interact to influence ICB outcome. Nonetheless, new findings pertaining to CICB were unveiled in this melanoma cohort and confirmed in mouse models of melanoma and sarcoma. First, the intestinal microbiota paves the way to the CICB-induced toxicity, as shown by the capacity of ATB to mitigate ileitis in mouse cancer models, and the significance of the microbiome profiles to the taxonomic level of species in our cohort. Secondly, CICB appears to favor the dominance of prevalent commensal bacteria associated with clinical benefit (such as *P. distasonis*) or intestinal toxicity (such as *B. intestinalis*) in melanoma patients and RET mouse melanomas. Thirdly, these parallel human and mouse studies highlight a contribution of commensals to not only subclinical colitis, but also to ileal damage associated with increased transcription of *IL1B*, that could be at least partially prevented by an IL-1R antagonist. Importantly, these findings are corroborated by additional work in human cohorts[30], and murine models[31], and these additional studies support approaches that will abrogate toxicity to combined immune checkpoint blockade while preserving (or even enhancing) therapeutic response. Together, insights from these studies could inform new strategies in biomarkers of response and toxicity to CICB, as well as new therapeutic targets to potentially abrogate toxicity beyond TNFaR blockade[32].

Methods

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jennifer A. Wargo, (jwargo@mdanderson.org).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Patient cohort

Patients with advanced (stage III/IV) melanoma treated at The University of Texas MD Anderson Cancer Center between 01/23/2014 and 08/31/2017 who received at least one dose of ipilimumab in combination with a PD-1 checkpoint blockade agent (either nivolumab or pembrolizumab) as combination immune checkpoint blockade (CICB) were identified from detailed retrospective and prospective review of clinic records (**Supplementary Data Table 1**). CICB treatment was provided as part of clinical trial or expanded access program protocols (NCT01844505, NCT02186249, NCT02089685, NCT01621490, NCT02519322, NCT02320058) or as standard of care therapy. Due to known differences in underlying biology and immunotherapy responses between melanoma subtypes, only cutaneous melanomas were included for analysis of response as this was the dominant subset. All subtypes (i.e.: cutaneous, mucosal and uveal) were included for toxicity analyses. To enable translational analyses, patients without available biospecimens relevant to the CICB treatment period, or for whom insufficient data were available to determine radiographic responses and toxicity outcomes were excluded.

Mice

All mice experiments were approved by the local institutional board and performed in accordance with government and institutional guidelines and regulations. Female C57Bl/6 and BALB/c were purchased from Harlan (France) and Janvier (France), respectively. Mice were utilized between 8 and 16 weeks of age. All mouse experiments were performed at Gustave Roussy Cancer Campus and mice were housed in specific pathogen-free conditions.

Cell lines

MCA205 and RET melanoma (a transgene-enforced expression of the *Ret* proto-oncogene under the control of the metallothionein-1 promoter driving spontaneous melanomagenesis, kindly provided by Professor Viktor Umansky) (syngeneic from C57BL/6J mice) were cultured at 37°C under 5% CO_2 in RPMI-1640 medium supplemented with 10% heat-inactivated fetal bovine serum (FBS), 1% penicillin/streptomycin, 2 mM L-glutamine and 1% of sodium pyruvate and non-essential amino acids (all from Gibco-Invitrogen), referred herein as complete RPMI medium. Cell lines were regularly tested for mycoplasma contamination and were not used after 10 passages.

METHOD DETAILS

Clinical assessments and biospecimens

Response assessments. Clinical response annotation was performed independently by at least two clinical investigators per patient (MCA, PAP, HT). Treatment responses were defined using the best overall response (BOR) according to RECIST 1.1 criteria[33] comparing tumor burden on restaging imaging performed at standard disease re-assessment time points studies with baseline (pre-treatment) studies. Longitudinal restaging scans were evaluated throughout the period of treatment until the initiation of a subsequent line of therapy or last known follow-up date. Imaging modality was matched whenever possible, favoring contrast-enhanced CT of the chest, abdomen and pelvis, contrast-enhanced MRI or CT brain, and imaging of the neck or extremities as indicated by known sites of disease. Patients were classified as "responders" (R) if they achieved objective complete response (CR; 100% reduction in tumor burden) or partial response (PR; ≥30% reduction in tumor burden) attributable to CICB. Patients were classified as non-responders if they achieved a BOR of progressive disease (PD; ≥20% increase in disease burden) or stable disease (SD; not meeting criteria for CR/PR/PD) (**Supplementary Data Table 2**). Mice were defined as responders (R) if their tumors either regressed or were stable during treatment, or as non-responders (NR) when tumors increased in size over two consecutive measurements.

Toxicity assessments. Immune-related adverse events (irAE) was scored according to the NCI Common Terminology Criteria for Adverse Events (CTCAE) 4.0 criteria and immune-relatedness to CICB therapy ("possible", "probable", "definite" association) assigned by consensus opinion of at least two independent clinical investigators (MCA, HT, WSC). Binary toxicity classification was based on whether patients experienced any grade 3 or higher irAE versus less than grade 3 irAE (**Supplementary Data Table 2**).

Biospecimen collections. Available pre- and on-treatment tumor and peripheral blood samples were identified by querying institutional research biospecimen holdings and, when necessary, archival pathology holdings from diagnostic specimens. Tumor biopsies were obtained as punch, core needle, or excisional biopsies and preserved as snap-frozen (for RNA/DNA extraction) or formalin-fixed paraffin-embedded (FFPE; for immunohistochemistry or DNA extraction) specimens. Peripheral blood samples underwent density-gradient centrifugation to isolate peripheral blood mononuclear cells (PBMC) prior to cryopreservation until required for germline DNA extraction or flow cytometry. Biospecimens were retrieved, collected and analyzed under UT MD Anderson Cancer Center Institutional Review Board-approved protocols in accordance with the Declaration of Helsinki. Fecal samples were obtained on an outpatient basis after detailed in-person explanation and instruction by a treating clinician to facilitate stool capture free of water/urine using a single-use toilet insert. Fecal samples were stabilized immediately using the OMNIgene-GUT Kit (DNA Genotek Inc, Ottawa, Canada) according to the manufacturer's recommendations, involving contact only with a provided sterile spatula. Stabilized fecal

samples were returned in person or by mail to a central laboratory at The University of Texas MD Anderson Cancer Center within 30 days of collection and stored at -80ºC immediately upon receipt. For sequencing, samples were shipped as-is and in bulk to the Alkek Center for Metagenomics and Microbiome Research at Baylor College of Medicine. Patient-level sample utilization is as shown in **Supplementary Data Table 3**.

Colon biopsies from a cohort of ICB-treated patients taken at the time of clinical grade 3-4 colitis, and from a separate cohort of non-ICB-treated patients without gut symptoms were identified from systematic chart review, as summarized in **Supplementary Data Table 7**. Archival FFPE material was retrieved and freshly cut sections used to extract RNA for downstream estimation of inflammatory cytokine expression by qPCR.

Genomic analyses

Whole exome sequencing analysis. Whole-exome sequencing (WES) was performed using the same protocol as previously described[15]. A total of 26 pre-treatment samples were included (19R, 7NR). DNA was extracted from tumor samples after pathological assessment and confirmation of tumor content. Matched peripheral blood leukocytes were collected as germline DNA control. The initial genomic DNA input into the shearing step was 750 ng. End repair, A-base addition, adapter ligation using forked Illumina paired-end adapters, and library enrichment polymerase chain reaction (PCR) was performed using the KAPA Hyper Prep Kit (#KK8504) followed by solid-phase reverse immobilization bead cleanup and cluster generation. Library construction was performed per the manufacturer's instructions. Target enrichment was performed using the Agilent SureSelectXT Target Enrichment (#5190-8646) protocol as per the manufacturer's instructions, using 650-750 ng of prepared libraries. Enriched libraries were normalized to equal concentrations using an Eppendorf Mastercycler EP Gradient instrument, pooled to equimolar amounts on the Agilent Bravo B platform and quantified using the KAPA LibraryQuantification Kit (#KK4824). Pooled libraries were adjusted to 2 nM, denatured with 0.2 M NaOH, diluted using Illumina hybridization buffer, and underwent cluster amplification using HiSeq v3 cluster chemistry and the Illumina Multiplexing Sequencing Primer Kit as per manufacturer's instructions. Pools were then sequenced on an Illumina HiSeq 2000/2500 v3 system using 76 bp paired-end reads, and analyzed using RTA v.1.13 or later. The mean coverage for exome data was 221´ in tumors and 100´ in germ line. Aligned BAM (hg19) files were then processed using Picard and GATK software to identify duplication, realignment and recalibration. Somatic point mutations were identified using MuTect (v1.1.4) and small insertions/deletions identified using Pindel (v0.2.4). Additional post-calling filters were then applied, including: (a) total read count in tumor sample > 30, (b) total read count in matched normal sample > 10, (c) VAF (Variant Allele Frequency) in tumor sample > 0.05, (d) VAF in matched normal sample < 0.01, and (e) SNVs reported in dbSNP129 and 1000 Genomes Project were removed.

Copy number alteration analysis. Copy number alteration analysis was performed as previously described[15]. Essentially, Sequenza (v2.1.2) algorithm was applied to the aligned BAM data to obtain the $\log_2$ copy number ratio (tumor/normal) for each tumor sample. Using R package "CNTools" (v1.24.0), copy number gain ($\log_2$ copy ratios > $\log_2 1.5$) and loss ($\log_2$ copy ratios < $-\log_2 1.5$) at the gene level were identified. The burden of copy number gain or loss was defined as the total number of genes with copy number gain or loss per sample. To define recurrent CNA, R package "cghMCR" (v1.26.0) was applied to the calculated $\log_2$ copy ratios (tumor/normal) to identify genomic regions of recurrent CNAs (minimum common regions, MCRs). To identify genes preferentially lost or gained in responders versus non-responders, Fisher's exact test was performed at each gene location, and statistical significance was defined by FDR adjusted $p<0.05$. Genes with CNA in less than 3 samples were excluded.

Neoantigen prediction. Non-synonymous exonic mutations (NSEM) from WES were reviewed and all possible 8- to 12-mer peptides encompassing NSEM were used for neoantigen prediction and compared with wild type peptides. HLA of each case was predicted using PHLAT[34]. Binding affinity was evaluated, taking into account patient HLA, by the NetMHCpan (v2.8) algorithm[35,36]. Candidate peptides with a predicted IC50<500 nM were considered HLA-binding.

Immune analyses

Flow cytometry - MDACC. Peripheral blood mononuclear cells (PBMCs) obtained from the study patients were analyzed by members of the MD Anderson Immunotherapy Platform. Pre-treatment and post-treatment blood samples were drawn for immunophenotypic analysis of PBMCs. PBMC samples were available from 20 patients, including 10 patients with ≥Grade 3 irAE, and 10 patients with <Grade 3 irAE. Multiparametric flow cytometry analysis of PBMCs was performed using fluorescently conjugated monoclonal antibodies across several panels: CD4 AF532 (SK3, eBioscience), CD3 PerCP-Cy5.5 (UCHT1, Biolegend) CD8 AF700 (RPA-T8, BD Biosciences), CD127 BV711 (HIL-7R-M21, BD Biosciences), ICOS PE-Cy7 (ISA-3, eBioscience), PD-1 BV650 (EH12.1 BD Biosciences) and FOXP3 PE-e610 (PCH101; eBioscience); CD3 PE-CF594, CD4 Pe-Cy5.5, CD8 AF532, CD45RA BV650 (HI100, Biolegend), CCR7 BV785 (G043H7, Biolegend) CD27 PeCy5 (0323, eBioscience), CD28 APC-e780 (CD28.2 eBioscience), PD-1 BV650 (EH12.1 BD Biosciences), EOMES e660 (WD1928, eBioscience), and TBET BV605 (4B10 Biolegend). Live/Dead fixable yellow stain was obtained from Thermo Fisher Scientific. Samples were run using an LSR Fortessa (BD Biosciences) and analyzed using the FlowJo software program. After appropriate forward/side scatter and live single cell gating, we determined the frequency of total CD3+ T cells, CD8+ T cells (CD3+CD8+) and CD4+ T cells (CD3+CD4+). Among the CD4, CD4+ effector T cells (CD4+FOXP3-) and CD4+ regulatory T cells (CD4+FOXP3+CD127-/low). PD-1 and ICOS expression were evaluated on these populations. CD45RA and CCR7 expression on CD4 and CD8 T cells was used to

define naïve, T central memory (TCM), T effector memory (TEM) and effector T (Teff) sub-populations. PD-1, CD28, CD27, EOMES and TBET expression was evaluated in each of these compartments.

Flow cytometry – MSKCC. Peripheral blood mononuclear cells (PBMC) were isolated and cryopreserved from patient whole blood samples. Flow cytometry was performed in the Immune Monitoring Facility at Memorial Sloan Kettering Cancer Center (MSKCC) to examine T cell phenotypic markers. Human peripheral blood mononuclear cells (PBMC) samples were thawed and stained with a fixable viability stain (FVS510, BD Biosciences) and a cocktail of antibodies to the following surface markers: CD45RA-BUV395 (BD, HI100), CD4-BUV496 (BD, SK3), ICOS-BUV563 (BD, DX29), CD25-BUV615 (BD, 2A3), TIM-3-BUV661 (BD, 7D3), CD27-BUV737 (BD, L128), CD8-BUV805 (BD, SK1), CD57-BV421 (BD, NK-1), CXCR5-BV480 (BD, RF8B2), CD14-BV570 (BioLegend, M5E2), CD19-BV570 (BioLegend, HIB19), CCR4-BV605 (BioLegend, L291H4), CCR7-SB645 (eBioscience, 3D12) HLA-DR-BV711 (BD, G46-6), CD3-BV750 (BD, SK7), CD28-BV786 (BD, CD28.2), PD-1-BB515 (BD, MIH4), CD127-BB700 (BD, HIL-7R-M21), CD38-BB790 (BD, HIT2), TIGIT-PE (eBioscience, MBSA43), and GITR-PE-Cy7 (eBioscience, eBioAITR), in the presence of Brilliant Stain Buffer Plus (BD). Cells were next fixed and permeabilized with the FoxP3/Ki-67 Fixation/Permeabilization Concentrate and Diluent (eBioscience), and subsequently stained intracellularly with LAG-3-BB660 (BD, T47-530), Ki-67-AlexaFluor700 (BD, B56), FoxP3-PE-Cy5.5 (eBioscience, PCH101), CTLA-4-PE-Cy5 (BD, BNI3), Eomes-PE-eFluor610 (eBioscience, WD1928), T-bet-APC (eBioscience, ebio4B10), Granzyme B-APC-Fire750 (BioLegend, QA16A02), in the presence of Brilliant Stain Buffer Plus (BD). Stained cells were acquired on a BD Biosciences FACSymphony and analyzed using FlowJo software (FlowJo, LLC).

Immunohistochemistry. A hematoxylin & eosin (H&E) stained slide from each FFPE tumor sample was obtained to confirm the presence of tumor. Heavily pigmented samples were pretreated with melanin bleaching by low concentration hydrogen peroxide. The selected antibody panel included programmed death-ligand 1 (PD-L1) clone E1L3N (1:100, Cell Signaling Technology), PD-1 clone EPR4877 (1:250, Epitomics), CD3 polyclonal (1:100, DAKO), CD4 clone 4B12 (1:80, Leica Biosystems), CD8 clone C8/144B (1:25, Thermo Scientific), FOXP3 clone 206D (1:50, BioLegend) and Granzyme B clone 11F1 (ready to use, Leica Microsystems). IHC staining of a limited antibody panel was performed using a Leica Bond Max automated stainer (Leica Biosystems, Buffalo Grove, IL). The IHC reaction was performed using Leica Bond Polymer Refine detection kit (Leica Biosystems) and diaminobenzidine (DAB) was used as chromogen. Counterstaining was with hematoxylin. All IHC slides were scanned using an Aperio AT Turbo (Leica Biosystems) prior to all downstream IHC analyses. Using the Aperio Image Toolbox analysis software (Leica Biosystems), average values for each markers from five randomly-selected 1mm^2 areas within the tumor region were selected for digital analysis as previously described[37]. PD-L1 expression was evaluated by H-score, which evaluates the percentage of positive cells (0 to 100) and the intensity of staining (0 to 3+), with a total score ranging from 0 to 300. The remaining markers were scored as density of cells.

TCR Sequencing. DNA was extracted from available FFPE tumor tissues (19R, 6NR) and PBMC (15 patients with ≥Grade 3 irAE, and 12 patients with <Grade 3 irAE) using the QIAamp DNA FFPE Tissue Kit (Qiagen). Next generation TCR sequencing of CDR3 variable regions was performed using the ImmunoSeq hsTCRB kit (Adaptive Biotechnologies) followed by sequencing on a MiSeq 150´ (Illumina) and analysis using the ImmunoSeq^{TM} Analyzer software v3.0 (Adaptive Biotechnologies), considering only samples for which a minimum of 1000 unique templates were detected. Clonality is an index inversely correlated with TCR diversity and was measured as 1-(entropy)/$\log_2$(# of productive unique sequences). Preferential clonal expansion was defined as the number of T cell clones significantly expanded in post-treatment compared to pre-treatment blood samples.

Murine models

Antibiotic treatments. Mice were treated with an antibiotic solution (ATB) containing ampicillin (1 mg/ml), streptomycin (5 mg/ml), and colistin (1 mg/ml) (Sigma-Aldrich), with or without the addition of vancomycin (0.25 mg/ml) via the drinking water. Solutions and bottles were replaced 3 times and once weekly, respectively. Antibiotic activity was confirmed by cultivating fecal pellets resuspended in BHI+15% glycerol at 0.1 g/ml on COS (Columbia Agar with 5% Sheep Blood) plates for 48 h at 37°C in aerobic and anaerobic conditions. The duration of ATB treatments was slightly different based on the experimental settings. In brief, mice were treated for 2 weeks prior to tumor implantation and continuously throughout the experiment in MCA205 and RET experiments, whilst in experiments where FMT were used, ATB treatment was administered for 3 days prior to fecal microbiota transfer.

Tumor challenge and treatment. Flanks of mice were subcutaneously (s.c.) injected with 0.8×10^6 MCA205 or 0.5×10^6 RET cells. Treatment commenced when tumors reached 20 to 30 mm^2. Mice were injected intraperitoneally (i.p) every three days with anti-PD-1 mAb (250 µg/mouse; clone RMP1-14, 6 injections in MCA205, 5 injections in RET) and/or anti-CTLA-4 mAb (100 µg/mouse, clone 9D9, 5 injections in both MCA205 and RET) with or without anti-IL-1R antagonist (anakinra, 500 µg/mouse, injected i.p. three times per week) or respective isotype controls as indicated in figures. All mAbs for *in vivo* use were obtained from BioXcell (West Lebanon, NH, USA), using the recommended isotype control mAbs except anakinra (Swedish Orphan Biovitrum, Sweden).

Fecal microbiota transfer experiments. After 3 days of ATB treatment, fecal microbiota transfer (FMT) was performed using samples from healthy volunteers whose fecal shot gun sequencing analyses revealed the presence or absence of *B. intestinalis*. Frozen fecal samples were thawed and thoroughly vortexed. Large particulate material was allowed to settle by gravity. 200 µL of supernatant was administered in a single dose by oral gavage. An additional 100 µL was topically applied onto the fur of

each animal. Two weeks after FMT, C57BL/6J mice were inoculated with 1 x 10^5 RET tumor cells in 100 µL PBS were injected subcutaneously. CICB began 7 days after tumor inoculation and mice were sacrificed at 24 hrs post 1st administration of anti-CTLA4+anti-PD1 ip to harvest ilea and perform qPCR for several gene products (IL-1b, TNFa, IL-6, IL-17).

Gut colonization with dedicated commensal species. *Bacteroides intestinalis* CSURP836 (provided by Institut hospitalo-universitaire Méditerranée Infection, Marseille, France; isolated from a human sample), *B. intestinalis* from everImmune (isolated from stools of a lung cancer patient prior to immunotherapy) and *B. intestinalis* (isolated from a mouse sample) were cultured on COS plates in anaerobic conditions using anaerobic generators (Biomerieux) at 37°C for 24 - 72 hours. Suspensions of 10^9 CFU/mL were obtained using a spectrophotometer (Eppendorf) at an optical density of 1 measured at 600 nm. Oral gavages of 10^9 CFU in 100 µL were administered 24 hours prior to antibody treatment and with each antibody treatment. Bacteria were verified using a Matrix-Assisted Laser Desorption/Ionization Time of Flight (MALDI-TOF) mass spectrometer (Microflex LT analyser, Bruker Daltonics, Germany).

Cytokine quantification. Stool samples were collected and stored at -80°C until further processing. Samples were thawed and re-suspended (at 100 mg/mL) in PBS containing 0.1% Tween 20. After a 20 min incubation with shaking at room temperature, samples were centrifuged for 10 min at 12,000 rpm and supernatants were harvested and stored at -20°C until analysis. Lipocalin-2 levels were measured using the mouse Lipocalin-2/NGAL DuoSet ELISA kit (R&D Systems, Minneapolis, MN) following the manufacturer's instructions.

Immunohistochemistry. Gut tissue was preserved in either formalin fixed paraffin embedded (FFPE) or optimum cutting temperature compound (OCT). At mouse sacrifice the ileum and colon were removed, washed in PBS, cut longitudinally, rolled and fixed in 4% PFA overnight at 4°C or, in some experiments for 2 hours at room temperature. Tissue was then either paraffin-embedded with a Tissue-Tek® VIP® 6 Vacuum Infiltration Processor (Sakura) or rehydrated in 15% sucrose for 1h followed by 30% sucrose overnight, OCT embedded (Sakura) and snap frozen. Longitudinal sections were counterstained with hematoxylin, eosin & safran stain (H&E).

Histological assessment of gut tissue for toxicity. A scoring system was developed with a pathologist (P.O.). *Ileum:* Inflammatory foci, appearance of the submucosa, length of villi, and the thickness of lamina propria were scored for each section. The score was defined as: 0 = normal, 1 = focal and minor lesions; 2 = diffuse and minor lesions; 3 = diffuse, minor and major lesions; 4 = major lesions with areas containing

only connective tissue. *Colon*: Inflammatory infiltrate, defined as either physiological (0), low (1), moderate (2) and high (3) levels were scored.

immune gene expression by real-time quantitative PCR analysis. RNA was extracted using the RNeasy Mini Kit (Qiagen) and reverse transcribed into cDNA using SuperScript III Reverse Transcriptase and the RNaseOUT™ Recombinant Ribonuclease Inhibitor (Life Technologies) using random primers (Promega, Wisconsin, United States) and the Deoxynucleoside Triphosphate Set, PCR grade (Roche, Basel, Switzerland). Gene expression was analyzed by real-time quantitative PCR (RT-qPCR) using the TaqMan method with TaqMan® Gene Expression Assays and Taqman Universal Master Mix II (Invitrogen) according to the manufacturer's instructions on the 7500 Fast Real Time PCR system (Applied Biosystems). Expression was normalized to the expression of the housekeeping gene of β-2 microglobulin by means of the 2 −ΔCt method. The following primers were used (all from TaqMan® Gene Expression Assay, ThermoFisher): *B2m* (Mm00437762_m1), *Il1b* (Mm00434228_m1), *Il6* (Mm00446190_m1), *Tnf* (Mm00443258_m1), *IL1B* (Hs01555410_m1), *B2M (*Hs00187842_m1), *IL17A (*Hs00174383_m1), *TNF* (Hs00174128_m1)

Microbiome studies

Patient fecal sample collection. Baseline stool samples were collected using the OMNIgene GUT kit (DNA Genotek, Ottawa, Canada). A total of 54 stool samples were subject to bacterial 16S rRNA gene sequencing, including a cutaneous/unknown primary cohort (29R, 11NR; 24 with ≥Gr3 irAE, 16 with <Gr3 irAE), and for toxicity analyses only, a mucosal cohort (3 with and 5 without ≥Gr3 irAE) and a uveal melanoma cohort (2 with and 4 without ≥Gr3 irAE. Within this cohort, a number of samples obtained early after initiation of CICB were included as surrogate baseline samples, as our parallel study on longitudinal samples collected from patients undergoing immune checkpoint blockade monotherapies showed no significant change in fecal microbiota early after treatment initiation[21].

Human fecal bacterial DNA extraction Preparation and sequencing of the human fecal samples was performed in collaboration with the Alkek Center for Metagenomics and Microbiome Research (CMMR), Baylor College of Medicine using methods adapted from the NIH-Human Microbiome Project[38,39]. Extended details of the analytical pipeline have been reported previously[21]. Briefly, bacterial genomic DNA extracted using the MO BIO PowerSoil DNA Isolation Kit (MO BIO Laboratories, USA) underwent PCR amplification of the 16S rRNA gene V4 region (2 x 250 bp) and was sequenced using the MiSeq platform (Illumina, Inc, San Diego, CA).

Processing of 16S rRNA gene sequences. Quality filtered sequences with >97% identity were clustered into Operational Taxonomic Units (OTUs) and classified phylogenetically against the NCBI 16S ribosomal RNA sequence database (release date September 1, 2018) using the NCBI-BLAST+ package 2.8.1. 2018). The pipeline involves the following steps:

1. The **fastq_mergepairs** command within VSEARCH[40] was used to merge paired-end reads, with a maximum of 10 mismatches to create consensus sequences, followed by dereplication using the **derep_fulllength** command, sorting by decreasing length (**sortbylength** command; 200 to 350 bp), and sorting by decreasing cluster size of representative sequences (**sortbysize** command, minimum 2)
2. OTU clustering, selection, and exclusion of chimeras (97%) was done using the **cluster_otus** command through the UPARSE[41] algorithm within USEARCH[42]
3. Representative OTU sequences were then classified using the NCBI 16S database with BLAST (Basic Local Alignment Search Tool). This step was done in R using the blastn UNIX executable and served as the database against which the original merged reads were mapped. At the species level, only OTUs with an unambiguous assignment were classified, whereas all others were annotated as 'unclassified' (**Supplementary Data Table 4**)
4. Next the **usearch_global** command was used to query the database of merged reads for high identity hits using the previously generated representative OTU sequences as reference. The identity threshold used for this step was 0.97. The mapped OTUs were converted into an OTU table using a series of python scripts summarized in uc2otutab.py
5. Microbiome indices to estimate of alpha and beta diversity was calculated in QIIME[43]. In order to estimate the phylogenetic distances among OTUs, sequences were first aligned by the PyNAST[44] method using the **py** command. **filter_alignment.py** was then used to filter the sequence alignment by removing the highly variable regions
6. Next, the **py** script was used to create the phylogenetic tree from multiple sequence alignment and the beta_diversity.py script was used to estimate beta diversity using Bray-Curtis dissimilarity, Weighted and Unweighted UniFrac distance matrices[45].
7. In order to estimate alpha diversity, the OTU table was first rarefied using the single_rarefaction.py command in QIIME. The rarefaction cutoff used was the total read count for the sample with the least number of reads. The alpha_diversity.py script in QIIME was then used to estimate alpha diversity

Whole metagenome shotgun sequencing (WMS). DNA extracted for 16S rRNA gene sequencing was also used for WMS to minimize biases introduced in the extraction process. The sequencing was done at CosmosID where samples were quantified using Qubit4 and individual sequencing libraries were prepared using proprietary methods. Pooled libraries were sequenced on the Illumina NextSeq 550 platform in a

300-cycle run. Raw FASTQ files were made available through CosmosID's client portal and annotated taxonomically using metaphlan2[46] following exclusion of host reads with kneaddata.

Statistical assessment of microbial biomarkers using LEfSe. The LEfSe method was used to compare abundances of all bacterial clades according to response (i.e.: between R versus NR) and by occurrence of toxicity (i.e.: between patients with ≥Grade 3 irAE versus those with <Grade 3 irAE) using the Kruskal-Wallis test (statistical significance was defined as $p<0.05$)[47]. Bacterial taxa with differential abundance between study groups were used as input for the linear discriminant analysis (LDA) to calculate an effect size. LEfSe analysis for murine taxa was performed with Mothur v1.39.5.

Mouse fecal sample collection, DNA extraction and microbiota characterization. At least two longitudinal stool samples were collected from mice (n=71) and stored at -80°C until DNA extraction. Preparation and sequencing of mouse fecal samples was performed at IHU Méditerranée Infection, Marseille, France. Briefly, DNA was extracted using two protocols. The first protocol consisted of physical and chemical lysis, using glass powder and proteinase K respectively, then processing using the Macherey-Nagel DNA Tissue extraction kit (Duren, Germany)[48]. The second protocol was identical to the first protocol, with the addition of glycoprotein lysis and de-glycosylation steps[49]. The resulting DNA was sequenced, targeting the V3–V4 regions of the 16S rRNA gene as previously described[50]. Raw FASTQ files were analyzed with Mothur pipeline v.1.39.5 for quality check and filtering (sequencing errors, chimerae) on a Workstation DELL T7910 (Round Rock, Texas, United States). Raw reads (15512959 in total, on average 125104 per sample) were filtered (6342281 in total, on average 51147 per sample) and clustered into Operational Taxonomic Units (OTUs), followed by elimination of low-populated OTUs (till 5 reads) and by de novo OTU picking at 97% pair-wise identity using standardized parameters and SILVA rDNA Database v.1.19 for alignment. In all, considering RET and MCA samples, 427 bacterial taxa were identified using a prevalence threshold of ≥20% (i.e. present in at least 20% of samples). Sample coverage was computed with Mothur and resulted to be on average higher than 99% for all samples, thus meaning a suitable normalization procedure for subsequent analyses. Bioinformatic and statistical analyses on recognized OTUs were performed with Python v.2.7.11. The most representative and abundant read within each OTU (as evidenced in the previous step with Mothur v.1.39.5) underwent a nucleotide Blast using the National Center for Biotechnology Information (NCBI) Blast software (ncbi-blast-2.3.0) and the latest NCBI 16S Microbial Database accessed at the end of April 2019 (ftp://ftp.ncbi.nlm.nih.gov/blast/db/). A matrix of bacterial relative abundances was built at each taxonomic level (phylum, class, order, family, genus, species) for subsequent multivariate statistical analyses.

Mouse microbiota and OTU-level analyses. For mouse experiments, raw data were firstly normalized then standardized using QuantileTransformer and StandardScaler methods from Sci-Kit learn package v0.20.3. Normalization using the output_distribution='normal' option transforms each variable to a strictly Gaussian-shaped distribution, whilst the standardization results in each normalized variable having a mean of zero and variance of one. These two steps of normalization followed by standardization ensure the proper comparison of variables with different dynamic ranges, such as bacterial relative abundances, tumor size, or colonic infiltrate score. Measurements of α diversity (within sample diversity) such as observed_otus and Shannon index, were calculated at OTU level using the SciKit-learn package v.0.4.1. Exploratory analysis of β-diversity (between sample diversity) was calculated using the Bray-Curtis measure of dissimilarity calculated with Mothur and represented in Principal Coordinate Analyses (PCoA), while for Hierarchical Clustering Analysis (HCA) 'Bray-Curtis' metrics and 'complete linkage' method were implemented using custom scripts (Python v.2.7.11). We implemented Partial Least Square Discriminant Analysis (PLS-DA) and the subsequent Variable Importance Plot (VIP) as a supervised analysis wherein the VIP values (order of magnitude) are used to identify the most discriminant bacterial species among tumor-bearing and tumor-free mice, and among the different timepoints (T0, T2, T5). As depicted in Figure 2e, bar thickness reports the fold ratio (FR) value of the mean relative abundances for each species among the two cohorts whilst not applicable (NA) refers to comparisons with a group with zero relative abundance. An absent border indicates mean relative abundance of zero in the compared cohort(s). In order to compare the microbiota taxa with gene expression datasets or tumor size and colonic toxicity, a multivariate statistical Spearman (or Pearson for mouse data) correlation analysis (and related P values) was performed with custom Python scripts. Mann-Whitney U and Kruskal-Wallis tests were employed to assess significance for pairwise or multiple comparisons, respectively, considering a p-value <0.05 as significant.

Pairwise comparisons of relative abundances between taxa identified within patient samples were performed using Mann-Whitney tests followed by bootstrapping with 1000 permutations. Only taxa that were present in at least 40% of all samples were considered. Rarefaction limits for the calculation of alpha diversity were set based on the least number of reads in all fecal samples. Taxonomic alpha-diversity of patient samples was estimated using the Inverse Simpson Index calculated as D= (pi is the proportion of the total species S that is comprised by the species i)[51], and additional diversity metrics as indicated in figures. Correlations between relative abundance of candidate taxa and peripheral immune markers were estimated using Spearman's rho. ANalysis Of SIMilarity (ANOSIM, which represents the difference of datasets' centroids) or, when indicated, Pearson correlation coefficient, were computed with Python 2.7.11.

Quantification of bacteria in fecal samples by qPCR. Genomic DNA was extracted from fecal samples using the QIAamp DNA Stool Mini Kit (Qiagen) following the manufacturer's instructions. Targeted qPCR systems were applied using either TaqMan technology (for systems targeting All Bacteria domain) or SYBR Green for different *Bacteroides* species. The following primers and probes were used:

Target	PCR system	Primers and probes	Oligo sequence	Reference
All bacteria	TaqMan	Forward	CGGTGAATACGTTCCCGG	52,53
		Reverse	TACGGCTACCTTGTTACGACTT	
		Probe	6 FAM -CTT GTA CAC ACC GCC CGT C-MGB	
B. intestinalis	SYBR Green	Forward	AGCATGACCTAGCAATAGGTTG	54
		Reverse	ACGCATCCCCATCGATTAT	
B. uniformis	SYBR Green	Forward	TCTTCCGCATGGTAGAACTATTA	55
		Reverse	ACCGTGTCTCAGTTCCAATGTG	
B. fragilis	SYBR Green	Forward	TGATTCCGCATGGTTTCATT	54
		Reverse	CGACCCATAGAGCCTTCATC	

Statistical analyses. Data analyses and representations were performed either with the R software (http://www.R-project.org/), Microsoft Excel (Microsoft Co., 436 Redmont, WA, US) or Prism 5 (GraphPad, San Diego, CA, USA). Patient cohort survival curves were generated using the R package "survival"[56]. Between-group comparisons of patient cohort genomic and immune parameters were performed using unpaired Mann-Whitney U tests or Fisher's exact test in the case of low-sample dichotomous variables, taking $p<0.05$ as statistically significant. All comparisons were two-sided unless a strong *a priori* hypothesis warranted a one-sided approach (indicated where appropriate). Permutation testing was performed by randomly permuting sample labels for a total of 1000 iterations. Multivariable logistic regression models were built using the best subsets approach to adjust for the effect of clinical prognostic variables. Separate models were built for response and toxicity outcomes and for each model, bacterial candidates identified during the taxonomic discovery phase were considered primary predictors. Abundances estimated from WMS were used as input. All patients were categorized as high or low for a bacterial candidate based on the median relative abundance. We allowed a maximum of two other clinical covariates (given constraints of event rates) from among age at entry, sex, *BRAF* mutation status (wild type vs mutant), AJCC stage (Stage III and IV vs Stage I and II), baseline LDH (high vs low), and melanoma subtype (uveal/mucosal vs cutaneous).

In murine studies, statistical analyses gathering more than two groups were performed using ANOVA followed by pairwise comparisons with Bonferroni adjustments. Differential enrichment analyses in murine studies were corrected for multiple hypothesis testing using FDR at 10% two-stage Benjamini-Hochberg. ANOSIM and PLS-DA p-values were automatically calculated after 999 permutations.

Otherwise, for two groups, statistical analyses were performed using the unpaired t-test. Outliers within a given distribution were tested using Grubbs' test (https://graphpad.com/quickcalcs/Grubbs1.cfm) with a threshold at $p<0.05$. All tumor growth curves were analyzed using software developed in Professor Guido Kroemer's laboratory and information about statistical analyses can be found at this following link: https://kroemerlab.shinyapps.io/TumGrowth/ [57]. Briefly, for longitudinal analyses, original tumor measurements were log transformed before statistical testing. When complete regressions of tumors were observed, zeros were imputed by the minimum value divided by 2. An automatic outlier detection at $p<0.1$ was retained, both for the longitudinal analyses and the Kaplan-Meier curves. Survival curves were estimated using the Cox regression and the multiple testing was taken account using the Bonferroni adjustment. p-values were two-sided with 95% confidence intervals and were considered significant when $p<0.05$. Symbol significance: *$p<0.05$, **$p<0.01$, ***$p<0.001$.

References

1 Larkin, J. *et al.* Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. *N Engl J Med* **373**, 23-34, doi:10.1056/NEJMoa1504030 (2015).

2 Hammers, H. J. *et al.* Safety and Efficacy of Nivolumab in Combination With Ipilimumab in Metastatic Renal Cell Carcinoma: The CheckMate 016 Study. *J Clin Oncol* **35**, 3851-3858, doi:10.1200/JCO.2016.72.1985 (2017).

3 Sznol, M. *et al.* Pooled Analysis Safety Profile of Nivolumab and Ipilimumab Combination Therapy in Patients With Advanced Melanoma. *J Clin Oncol* **35**, 3815-3822, doi:10.1200/JCO.2016.72.1167 (2017).

4 Attia, P. *et al.* Autoimmunity correlates with tumor regression in patients with metastatic melanoma treated with anti-cytotoxic T-lymphocyte antigen-4. *J Clin Oncol* **23**, 6043-6053, doi:10.1200/JCO.2005.06.205 (2005).

5 Das, S. & Johnson, D. B. Immune-related adverse events and anti-tumor efficacy of immune checkpoint inhibitors. *J Immunother Cancer* **7**, 306, doi:10.1186/s40425-019-0805-8 (2019).

6 Robert, C. *et al.* Nivolumab in previously untreated melanoma without BRAF mutation. *N Engl J Med* **372**, 320-330, doi:10.1056/NEJMoa1412082 (2015).

7 Robert, C. *et al.* Pembrolizumab versus Ipilimumab in Advanced Melanoma. *N Engl J Med* **372**, 2521-2532, doi:10.1056/NEJMoa1503093 (2015).

8 Wolchok, J. D. *et al.* Nivolumab plus ipilimumab in advanced melanoma. *N Engl J Med* **369**, 122-133, doi:10.1056/NEJMoa1302369 (2013).

9 Postow, M. A. *et al.* Nivolumab and ipilimumab versus ipilimumab in untreated melanoma. *N Engl J Med* **372**, 2006-2017, doi:10.1056/NEJMoa1414428 (2015).

10 D'Angelo, S. P. *et al.* Efficacy and Safety of Nivolumab Alone or in Combination With Ipilimumab in Patients With Mucosal Melanoma: A Pooled Analysis. *J Clin Oncol* **35**, 226-235, doi:10.1200/JCO.2016.67.9258 (2017).

11 Hellmann, M. D. *et al.* Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden. *N Engl J Med* **378**, 2093-2104, doi:10.1056/NEJMoa1801946 (2018).

12 Hugo, W. *et al.* Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell* **165**, 35-44, doi:10.1016/j.cell.2016.02.065 (2016).

13 Snyder, A. *et al.* Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N Engl J Med* **371**, 2189-2199, doi:10.1056/NEJMoa1406498 (2014).

14 Van Allen, E. M. *et al.* Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207-211, doi:10.1126/science.aad0095 (2015).

15 Roh, W. *et al.* Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Sci Transl Med* **9**, doi:10.1126/scitranslmed.aah3560 (2017).

16 Ekmekcioglu, S. *et al.* Inflammatory Marker Testing Identifies CD74 Expression in Melanoma Tumor Cells, and Its Expression Associates with Favorable Survival for Stage III Melanoma. *Clin Cancer Res* **22**, 3016-3024, doi:10.1158/1078-0432.CCR-15-2226 (2016).

17 Peng, W. *et al.* Loss of PTEN Promotes Resistance to T Cell-Mediated Immunotherapy. *Cancer Discov* **6**, 202-216, doi:10.1158/2159-8290.CD-15-0283 (2016).

18 Tanese, K. *et al.* Cell Surface CD74-MIF Interactions Drive Melanoma Survival in Response to Interferon-gamma. *J Invest Dermatol* **135**, 2775-2784, doi:10.1038/jid.2015.204 (2015).

19 Zaretsky, J. M. *et al.* Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma. *N Engl J Med* **375**, 819-829, doi:10.1056/NEJMoa1604958 (2016).

20 Tumeh, P. C. *et al.* PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568-571, doi:10.1038/nature13954 (2014).

21 Gopalakrishnan, V. *et al.* Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* **359**, 97-103, doi:10.1126/science.aan4236 (2018).

22 Matson, V. *et al.* The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. *Science* **359**, 104-108, doi:10.1126/science.aao3290 (2018).

23 Routy, B. *et al.* Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* **359**, 91-97, doi:10.1126/science.aan3706 (2018).

24 Seo, S. U. *et al.* Distinct Commensals Induce Interleukin-1beta via NLRP3 Inflammasome in Inflammatory Monocytes to Promote Intestinal Inflammation in Response to Injury. *Immunity* **42**, 744-755, doi:10.1016/j.immuni.2015.03.004 (2015).

25 Moro-Garcia, M. A., Alonso-Arias, R. & Lopez-Larrea, C. Molecular mechanisms involved in the aging of the T-cell immune response. *Curr Genomics* **13**, 589-602, doi:10.2174/138920212803759749 (2012).

26 Chen, Y., Gorelik, G. J., Strickland, F. M. & Richardson, B. C. Decreased ERK and JNK signaling contribute to gene overexpression in "senescent" CD4+CD28- T cells through epigenetic mechanisms. *J Leukoc Biol* **87**, 137-145, doi:10.1189/jlb.0809562 (2010).

27 Oh, D. Y. *et al.* Immune Toxicities Elicted by CTLA-4 Blockade in Cancer Patients Are Associated with Early Diversification of the T-cell Repertoire. *Cancer Res* **77**, 1322-1330, doi:10.1158/0008-5472.CAN-16-2324 (2017).

28 Subudhi, S. K. *et al.* Clonal expansion of CD8 T cells in the systemic circulation precedes development of ipilimumab-induced toxicities. *Proc Natl Acad Sci U S A* **113**, 11919-11924, doi:10.1073/pnas.1611421113 (2016).

29 Routy, B. *et al.* The gut microbiota influences anticancer immunosurveillance and general health. *Nat Rev Clin Oncol* **15**, 382-396, doi:10.1038/s41571-018-0006-2 (2018).

30 Hailemichael, Y., et al. *IL6 blockade abrogates colitis and promotes response to checkpoint blockade* (2020).*Submitted*

31 Zhou, Y., et al. *Intestinal toxicity to CTLA-4 blockade driven by IL-6 and myeloid infiltration* (2020). *Submitted*

32 Perez-Ruiz, E. *et al.* Prophylactic TNF blockade uncouples efficacy and toxicity in dual CTLA-4 and PD-1 immunotherapy. *Nature* **569**, 428-432, doi:10.1038/s41586-019-1162-y (2019).

33 Eisenhauer, E. A. *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* **45**, 228-247, doi:10.1016/j.ejca.2008.10.026 (2009).

34 Bai, Y., Ni, M., Cooper, B., Wei, Y. & Fury, W. Inference of high resolution HLA types using genome-wide RNA or DNA sequencing reads. *BMC Genomics* **15**, 325, doi:10.1186/1471-2164-15-325 (2014).

35 Nielsen, M. *et al.* NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PLoS One* **2**, e796, doi:10.1371/journal.pone.0000796 (2007).

36 Hoof, I. *et al.* NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics* **61**, 1-13, doi:10.1007/s00251-008-0341-z (2009).

37 Chen, P. L. *et al.* Analysis of Immune Signatures in Longitudinal Tumor Samples Yields Insight into Biomarkers of Response and Mechanisms of Resistance to Immune Checkpoint Blockade. *Cancer Discov* **6**, 827-837, doi:10.1158/2159-8290.CD-15-1545 (2016).

38 Human Microbiome Project, C. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207-214, doi:10.1038/nature11234 (2012).

39 Human Microbiome Project, C. A framework for human microbiome research. *Nature* **486**, 215-221, doi:10.1038/nature11209 (2012).

40 Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahe, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584, doi:10.7717/peerj.2584 (2016).

41 Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* **10**, 996-998, doi:10.1038/nmeth.2604 (2013).

42 Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460-2461, doi:10.1093/bioinformatics/btq461 (2010).

43 Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335-336, doi:10.1038/nmeth.f.303 (2010).

44 Caporaso, J. G. *et al.* PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **26**, 266-267, doi:10.1093/bioinformatics/btp636 (2010).

45 Lozupone, C. & Knight, R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71**, 8228-8235, doi:10.1128/AEM.71.12.8228-8235.2005 (2005).

46 Segata, N. *et al.* Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods* **9**, 811-814, doi:10.1038/nmeth.2066 (2012).

47 Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol* **12**, R60, doi:10.1186/gb-2011-12-6-r60 (2011).

48 Dridi, B., Henry, M., El Khechine, A., Raoult, D. & Drancourt, M. High prevalence of Methanobrevibacter smithii and Methanosphaera stadtmanae detected in the human gut using an improved DNA detection protocol. *PLoS One* **4**, e7063, doi:10.1371/journal.pone.0007063 (2009).

49 Angelakis, E. *et al.* Glycans affect DNA extraction and induce substantial differences in gut metagenomic studies. *Sci Rep* **6**, 26276, doi:10.1038/srep26276 (2016).

50 Million, M. *et al.* Increased Gut Redox and Depletion of Anaerobic and Methanogenic Prokaryotes in Severe Acute Malnutrition. *Sci Rep* **6**, 26051, doi:10.1038/srep26051 (2016).

51 Morgan, X. C. & Huttenhower, C. Chapter 12: Human microbiome analysis. *PLoS Comput Biol* **8**, e1002808, doi:10.1371/journal.pcbi.1002808 (2012).

52 Furet, J. P. *et al.* Comparative assessment of human and farm animal faecal microbiota using real-time quantitative PCR. *FEMS Microbiol Ecol* **68**, 351-362, doi:10.1111/j.1574-6941.2009.00671.x (2009).

53 Suzuki, M. T., Taylor, L. T. & DeLong, E. F. Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Appl Environ Microbiol* **66**, 4605-4614, doi:10.1128/aem.66.11.4605-4614.2000 (2000).

54 Odamaki, T. *et al.* Distribution of different species of the Bacteroides fragilis group in individuals with Japanese cedar pollinosis. *Appl Environ Microbiol* **74**, 6814-6817, doi:10.1128/AEM.01106-08 (2008).

55 Tong, J., Liu, C., Summanen, P., Xu, H. & Finegold, S. M. Application of quantitative real-time PCR for rapid identification of Bacteroides fragilis group and related organisms in human wound samples. *Anaerobe* **17**, 64-68, doi:10.1016/j.anaerobe.2011.03.004 (2011).

56 Therneau, T. M. & Grambsch, P. M. *Modeling Survival Data: Extending the Cox Model.*, (Springer, 2000).

57 Enot, D. P., Vacchelli, E., Jacquelot, N., Zitvogel, L. & Kroemer, G. TumGrowth: An open-access web tool for the statistical analysis of tumor growth curves. *Oncoimmunology* **7**, e1462431, doi:10.1080/2162402X.2018.1462431 (2018).

Declarations

ACKNOWLEDGEMENTS

This research was supported by the generous philanthropic contributions to The University of Texas MD Anderson Cancer Center Moon Shots Program™ from the Lyda Hill Foundation and utilized platform assistance from the Cancer Genomics Laboratory and Immunotherapy Platform; from the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation, and the AIM at Melanoma Foundation. Additional support was provided to PAF from the Cancer Prevention Research Institute of Texas and Welch Foundation. MAD receives research support from the National Institutes of Health and National Cancer Institute (NIH 2T32CA009666-24 and NIH/NCI 1R01CA187076-02). MCA is supported by a National Health and Medical Research Council of Australia CJ Martin Early Career Fellowship (#1148680). WSC was supported by a National Institutes of Health T32 Training Grant (T32CA163185). APC is supported by the CPRIT Research Training Program (RP170067), the Fulbright France Commission Franco-Americainé and the John J Kopchick Foundation. AR is supported by the Kimberley Clark Foundation Award for Scientific Achievement provided by MD Anderson's Odyssey Fellowship Program. P.O. Gaudreau was supported by the *Fonds de Recherche Québec–Santé's* (FRQS) Resident Physician Health

Research Career Training Program (32667). The authors would like to acknowledge the assistance of the animal facility team at Gustave Roussy. LZ is funded by grants from EU H2020 ONCOBIOME, Ligue contre le Cancer (équipe labelisée); Agence Nationale de la Recherche (ANR) – Projets blancs; ANR under the frame of E-Rare-2, the ERA-Net for Research on Rare Diseases; Association pour la recherche sur le cancer (ARC); Cancéropôle Ile-de-France; Chancellerie des universités de Paris (Legs Poix), Fondation de France; Fondation pour la Recherche Médicale (FRM); a donation by Elior; Fondation Carrefour; Institut National du Cancer (INCa); Inserm (HTE); ANR germanofrench; LabEx Immuno-Oncology; the French Ministry of Health PIA2, RHU Torino Lumière (ANR-16-RHUS-0008); the Swiss Bridge Foundation; the Seerave and Carrefour Foundation; the SIRIC Stratified Oncology Cell DNA Repair and Tumor Immune Elimination (SOCRATE). This work was supported by the French Government under the « Investissements d'avenir » (Investments for the Future) program managed by the Agence Nationale de la Recherche (ANR, fr: National Agency for Research), (reference: Méditerranée Infection 10-IAHU-03). LD is supported by 'Parcours d'excellence en cancérologie - Fondation Philanthropia'. This work was supported by Région Provence Alpes Côte d'Azur and European funding FEDER PRIMI.

AUTHOR CONTRIBUTIONS

Conceptualization: LZ, JAW, PAF. Investigation: MCA, CD, WSC, VI, VG, LD, WR, ZAC, BR, APC, GF, IFC, LV, AR, AF, MPR, MTA, PO, JB, CS, PP, MAWK, AJL, CWH, MA, POG, MTT, HAT. Provision/Acquisition of Data and Materials: CS, ZAC, LEH, HAT, PH, WJH, RNA, APC, MKC, MAP, CEA, EMB, SEW, MAD, AD, SPP, ICG, JEG, NJA, JP, RJ, PS, JPA, JAW. Formal Analysis: MCA, WSC, CD, VG, VI, MAWK, MA, NJA, LZ. Data Curation: MCA, WSC, VG, MAWK, VI. Writing – Original Draft: MCA, WSC, LZ. Writing – Review & Editing: MCA, CD, VG, APC, VI, LZ, JAW. Writing – Review and Approval of Final Manuscript: all authors. Visualization: MCA, VG, CD, MAWK. Supervision: LZ, JAW, PAF. Funding Acquisition: JAW, LZ, PAF.

DATA AND SOFTWARE AVAILABILITY

WES data from this study are available from the European Genome-Phenome Archive under study accession EGAS00001003857. Human fecal 16S rRNA gene sequencing and WMS data are being deposited into the EGA. Murine fecal 16S rRNA gene sequence reads from this study have been submitted to the NCBI under the Bioproject ID PRJNA484225. All tumor growth curves were analyzed using software developed in Professor Guido Kroemer's laboratory and information about statistical analyses can be found at this following link: https://kroemerlab.shinyapps.io/TumGrowth/.

COMPETING INTEREST DECLARATION

MCA reports advisory board participation and honoraria from Merck Sharp and Dohme, outside the submitted work. VG is a co-inventor on US patent (PCT/US17/53,717) relating to the microbiome. VG and ZAC are currently employees of AstraZeneca. RNA reports research funding from, Bristol-Myers Squibb, Merck and Genentech, all outside the submitted work. HAT reports personal fees from Novartis, grants from Merck and Celgene, and grants and personal fees from BMS and Genentech, all outside of the submitted work. MAD has been a consultant to Roche/Genentech, Array, Novartis, BMS, GSK, Sanofi-Aventis, Vaccinex and Apexigen, and he has been the PI of research grants to MD Anderson by Roche/Genentech, GSK, Sanofi-Aventis, Merck, Myriad, and Oncothyreon. WJH reports research grants from Merck, Bristol-Myers Squibb, MedImmune, GlaxoSmithKline and has served on an advisory board for Merck, all outside the submitted work. JEG reports advisory board participation with Merck, Regeneron, Bristol-Myers Squibb, Novartis, Syndax, and Castle Biosciences. AJL reports personal fees from BMS, Novartis, Genentech/Roche, and Merck; personal fees and non-financial support from ArcherDX and Beta-Cat; grants and non-financial support from Medimmune/AstraZeneca and Sanofi; grants, personal fees and non-financial support from Janssen, all outside the submitted work. MTT reports personal fees from Myriad Genetics, Seattle Genetics and Novartis, all outside the submitted work. MAP reports honoraria from BMS and Merck, consulting fees from BMS, Merck, Array BioPharma, Novartis, Incyte, NewLink Genetics, Aduro, Eisai, and institutional support from RGenix, Infinity, BMS, Merck, Array BioPharma, Novartis, AstraZeneca. SPP reports insitutional support for clinical trial from InxMed. JLM reports honoraria from Roche, BMS, Merck. R.R.J. has consulted for Karius, Merck, Microbiome DX, and Prolacta, and is on the scientific advisory boards of Kaleido, LISCure, Maat Pharma, and Seres, and has received patent royalties licensed to Seres. PS reports consulting, advisory roles, and/or stocks/ownership for Achelois, Adaptive Biotechnologies, Apricity Health, BioAlta, BioNTech, Codiak Biosciences, Constellation, Dragonfly Therapeutics, Forty-Seven Inc., Hummingbird, ImaginAb, Infinity Pharma, Jounce Therapeutics, Lave Therapeutics, Lytix Biopharma, Marker Therapeutics, Oncolytics, Phenomics, and Polaris: and owns a patent licensed to Jounce Therapeutics. JPA reports consulting, advisory roles, and/or stocks/ownership for Achelois, Adaptive Biotechnologies, Apricity Health, BioAlta, BioNTech, Codiak Biosciences, Constellation, Dragonfly Therapeutics, Forty-Seven Inc., Hummingbird, ImaginAb, Jounce Therapeutics, Lave Therapeutics, Lytix Biopharma, Marker Therapeutics, Phenomics, and Polaris: and owns a patent licensed to Jounce Therapeutics. BR reports advisory board membership for Vedanta and research funding Vedanta, Davoltera, Kaleido. VG, CNS, AR and JAW are co-inventors on US patent (PCT/US17/53,717) relating to the microbiome. JAW, VG, MCA, LZ and VI are co-inventors on a provisional US patent relating to the microbiome, relevant to the current work. LZ is the main founder of EverImmune, a biotech company devoted to the use of commensal bacteria for the treatment of cancers, is in the Board of administrators of Transgene, and in the SAB of EpiVax, Lytix Biopharma and received research contracts from Kaleido, BMS, Incyte, Transgene, MERUS and GSK. JAW reports speaker fees from Imedex, Dava Oncology, Omniprex, Illumina, Gilead, MedImmune and Bristol-Myers Squibb; consultant/advisor roles or advisory board membership for Roche-Genentech, Novartis, Astra-Zeneca, GlaxoSmithKline, Bristol-Myers Squibb, Merck/MSD, Biothera Pharma, and Microbiome DX; and receives clinical trial support from GlaxoSmithKline, Roche-Genentech, Bristol-Myers Squibb, and Novartis, all outside the current work.

Description Of Supplementary Files

SUPPLEMENTARY DATA TABLES

Supplementary Data Table 1: Patient characteristics.

Supplementary Data Table 2: Clinical outcomes.

Supplementary Data Table 3: Biospecimen use overview.

Supplementary Data Table 4: Human fecal microbial OTU table.

Supplementary Data Table 5: Pairwise comparison of human fecal bacterial abundances by response.

Supplementary Data Table 6: Multivariable adjustment of human fecal microbial candidates from whole metagenomic sequencing.

Supplementary Data Table 7: Characteristics of patients in colonic cytokine analysis sample cohort.

Supplementary Data Table 8: Pairwise comparison of human fecal bacterial abundances by toxicity.

Supplementary Data Table 9: Characteristics of patients in MSKCC sample cohort.

EXTENDED DATA FIGURE LEGENDS

Extended Data Figure 1: Cohort description and tumor intrinsic genomic parameters.

(a) Kaplan-Meier curve of progression-free survival in the patient cohort stratified by melanoma subtype (n=77, n=63 cutaneous/unknown primary, n=8 mucosal, n=6 uveal). (**b**) Landscape of non-synonymous variants (NSV) identified by whole-exome sequencing (n=26 tumors) affecting selected genes recurrently mutated in melanoma, IFN-signaling genes and antigen processing/presentation genes. (**c**) Differences in counts of total predicted neoantigens, and all binding neoantigens in patients grouped by best overall response (R=responder (blue), n=20; NR=non-responder (red), n=6; all p>0.05, Mann-Whitney test). (**d**) Genome-wide SGOL scores and (**e**) barplot of the number of genes affected by copy number losses aggregated by chromosome, demonstrating dominant copy number loss burden within chromosomes 5, 10 and 15. (**f**) Copy number loss-affected genes located on chromosome 10 include a broad variety of

functional classes. (**g**) Entropy of pre-treatment intratumoral T cell receptor (TCR) repertoires comparing R (n=19) versus NR (n=6) repertoires (p=0.058, Mann-Whitney test).

Extended Data Figure 2: Fecal microbiome composition and diversity at baseline in CICB-treated patients.

(**a**) Stacked bar plot depicting 16S microbial composition of each analyzed fecal sample from the cutaneous and unknown primary cohort at the order level (n=39). (**b,c**) Comparison of group-wise abundances of Firmicutes (b) and Clostridiales (c) by response outcome in the cutaneous/unknown primary cohort (n=39). (**d**) Inverse Simpson alpha diversity of the fecal microbiome grouped by response in CICB-treated patients with cutaneous or unknown primary melanomas (n=39) taken at baseline (p=0.68, Mann-Whitney test; R=responder, NR=non-responder).

Extended Data Figure 3: Microbial associations with CICB response are confirmed by whole metagenomic sequencing.

(**a**) Volcano plot of pairwise comparisons of bacterial taxa (at all levels) identified from 16S sequencing (n=40) dichotomized by response to CICB using Mann-Whitney tests applied to 1000 permutations of differential bacterial abundance. (**b**) Procrustes analysis demonstrating high concordance between taxonomic identification using either 16S or WMS methods within the response cohort (Mantel: r=0.650, p=0.001). (**c**) A strong positive correlation was observed between abundance of *Bacteroides stercoris* quantified using 16S versus WMS (Spearman's rho=0.934 p=2.2e-16). (**d**) Confirmation of bacterial candidate associations with response using WMS.

Extended Data Figure 4: Associations between prevalent bacterial taxa and tumor response in murine models.

(**a**) Experimental setting for murine studies shown in Fig. 2 and 3. Treatment of established transplantable tumors (MCA205 sarcoma or RET melanoma) by intraperitoneal (i.p.) administrations of CICB and feces collection at three time points for 16S rRNA gene sequencing. Feces collection time points: T0=before treatment initiation (Day 0), T2=48 hours after 2 treatments (Day 5), T5=48 hours after 5 treatments (Day 14). In studies utilizing antibiotic (ATB) treatment, ATB was commenced 14 days prior to tumor inoculation and continued throughout. (**b**) Pearson correlation between the relative abundance of *Parabacteroides distasonis* (at T0, T2, and T5) and standardized tumor size at T5 in MCA205 and RET tumor-bearing mice. (**c**) Heatmap of Spearman correlations between the most prevalent (>20%) bacterial species identified in mouse feces at different time points (T0, T2, T5) from RET tumor-bearing mice and colon inflammatory infiltrates. Data are derived from combined discovery and validation cohort animals.

Red represents a positive correlation, while blue represents a negative correlation with colonic infiltrate score. Following FDR adjustment, no significant correlations were observed.

Extended Data Figure 5: Microbial associations with immune-related toxicity are confirmed by whole metagenomic sequencing.

(**a**) Inverse Simpson alpha diversity from 16S sequencing of baseline fecal microbiota in CICB-treated patients (n=54) was not associated with subsequent development of high-grade immune-related adverse events (irAE). Mann-Whitney test. (**b**) Volcano plot of pairwise comparisons of bacterial taxa (at all levels) dichotomized by experience of high grade (≥Grade 3) immune-related adverse events (n=54 patients) using Mann-Whitney tests applied to 1000 permutations of differential bacterial abundance. (**c**) Procrustes analysis demonstrating high concordance between taxonomic identification using either 16S or WMS methods (Mantel: r=0.665, *p*=0.001). (**d**) Confirmation of bacterial candidate associations with toxicity using WMS (≥Gr3 irAE: n=25 Yes, n=21 No). (**e**) A strong positive correlation was observed between abundance of *Bacteroides intestinalis* quantified using 16S versus WMS (Spearman's rho=0.62, *p*=4.2e-6).

Extended Data Figure 6: Immune markers of CICB toxicity.

(**a**,**b**) Comparison of Ki67+ cells within CD8+ T effectors (Teff; a) and T central memory (TCM; b) cells in early on-treatment blood samples between patients with available blood samples (n=14) grouped according to high-grade irAE (Mann-Whitney test: *p* values as shown). (**c**) Gating strategy for key CD4/8+ T cell populations. (**d**,**e**) Percentage of CD28+ cells within CD4+ Teff (c) and CD27+ cells within CD8+ Teff (d) measured at baseline in this patient cohort (MDACC; left panels) and a separate cohort of patients treated with CICB at Memorial Sloan-Kettering Cancer Center (MSKCC; right panels). Data are grouped by experience of high-grade irAE (Mann-Whitney test: *p* values as shown). (**f**) Boxplot depicting a higher diversity of the peripheral T cell repertoire as measured by TCR Vβ sequencing in patients experiencing high-grade irAE (n=24, Mann-Whitney test; *p* value as shown). (**g**) Boxplot showing the number of significantly expanded T cell clones (pre- to on-treatment) detected by TCR sequencing of the peripheral blood immune repertoire, grouped by presence or absence of high-grade irAE (n=16, Mann-Whitney test: *p* value as shown).

Figures

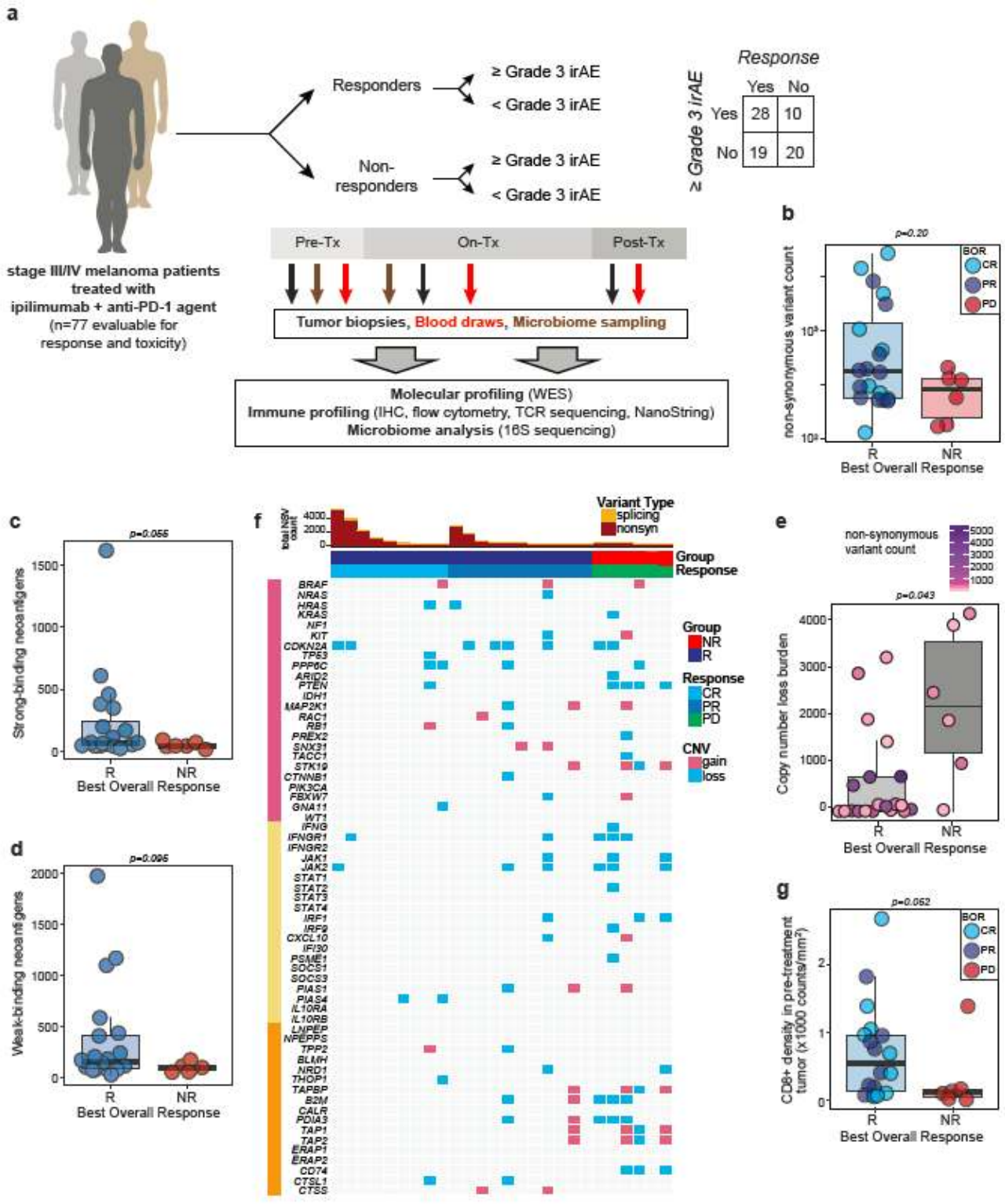

Figure 1

Molecular and immune predictors of response. (a) Cohort of patients with advanced melanoma (n=77) evaluated for clinical outcomes and correlative biospecimen analyses prior to and following initiation of combined anti-CTLA-4 and anti-PD-1 blockade. (b) Non-synonymous variant (NSV) count in pre-treatment tumor samples (n=26) grouped by binarized best overall response (BOR; R=responder, n=20, NR=non-responder, n=6). Specific objective responses are indicated by color of each data point (p=0.20, Mann-

Whitney test). (c,d) Strong and weak binding neoantigen predictions grouped by BOR as in (b). (e) Copy number loss burden (affected genomic regions) in pre-treatment tumor samples (n=26) grouped by binarized best overall response (p<0.05, Mann-Whitney test). Tumor mutation burden is indicated for each sample by color (f) Oncomap of copy number alterations affecting genes belonging to three groups: commonly mutated or copy-number altered in melanoma (pink), IFN-γ related signaling (yellow) and antigen processing (orange). (g) Density of infiltrating CD8+ cells (counts/mm2) in pre-treatment tumors by singlet stain immunohistochemistry grouped by binarized response (n=19 R, n=6 NR; p=0.052, one-sided Mann-Whitney test).

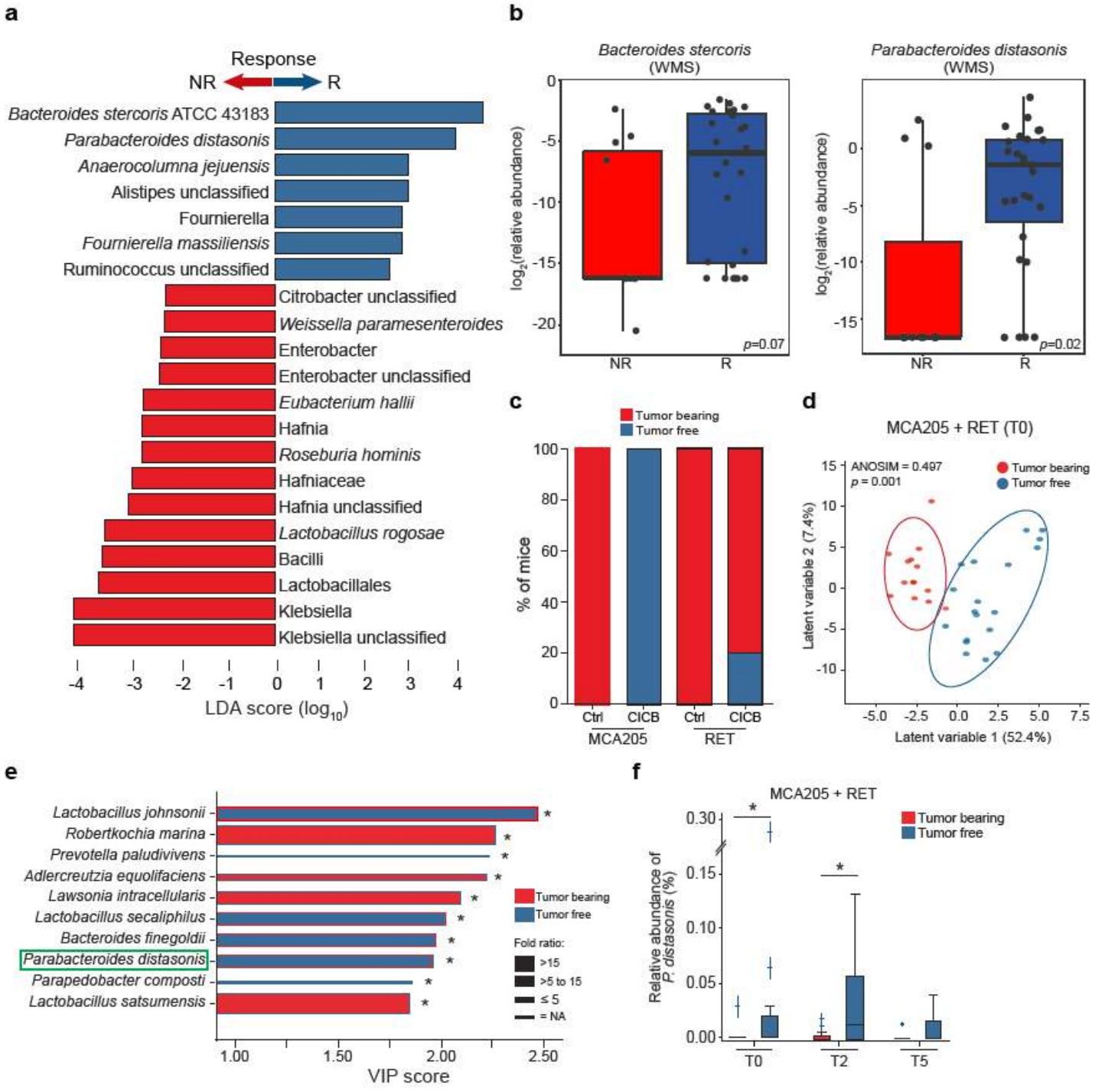

Figure 2

Gut microbial associations with CICB response. (a) LDA-score plot of bacterial taxa significantly enriched in patients with either R (n=29) or NR (n=11) to CICB from the cutaneous and unknown primary cohort (n=40). LDA=linear discriminant analysis. $p<0.05$. (b) abundance of candidate response taxa determined by WMS compared between response groups (total n=38 cutaneous melanoma patients: n=27 R, n=11 NR. Mann-Whitney test p-values as shown). (c) Percentages of tumor-free versus tumor-bearing mice after 4 intraperitoneal administrations of anti-PD-1 Ab + anti-CTLA-4 mAb (CICB) or isotype control mAbs (Ctrl) used to treat day 7 established MCA205 or RET tumors (n=24 mice/group, pooled data from two experiments). (d) Partial Least Squares Discriminant Analysis (PLS-DA) plot of the variance in beta-diversity at T0, between CICB treated mice which were eventually tumor-free or tumor-bearing at sacrifice, in both tumor models combined. LV, latent variable. ANOSIM defines the separation of the groups; the p value defines the significance of such separation after 999 permutations of the samples. (e) Variable importance (VIP) score barplot highlighting bacterial species present at T0 significantly enriched in the group defined by the bar color (highest mean relative abundance) compared to the group defined by the border color (lowest mean relative abundance), indicating mice that were eventually tumor-free versus tumor-bearing following CICB treatment (RET and MCA205 models). An absent border indicates mean relative abundance of zero in the compared cohort(s). The green box highlights a species in common with patient data. Mann-Whitney test: *$p<0.05$, **$p<0.01$, ***$p<0.001$, ns=not significant. Bar thickness reports the fold ratio value of the mean relative abundances for each species among the two cohorts. N/A=not applicable. (f) Relative abundance of Parabacteroides distasonis (at T0, T2, and T5) in tumor free and tumor bearers over time.

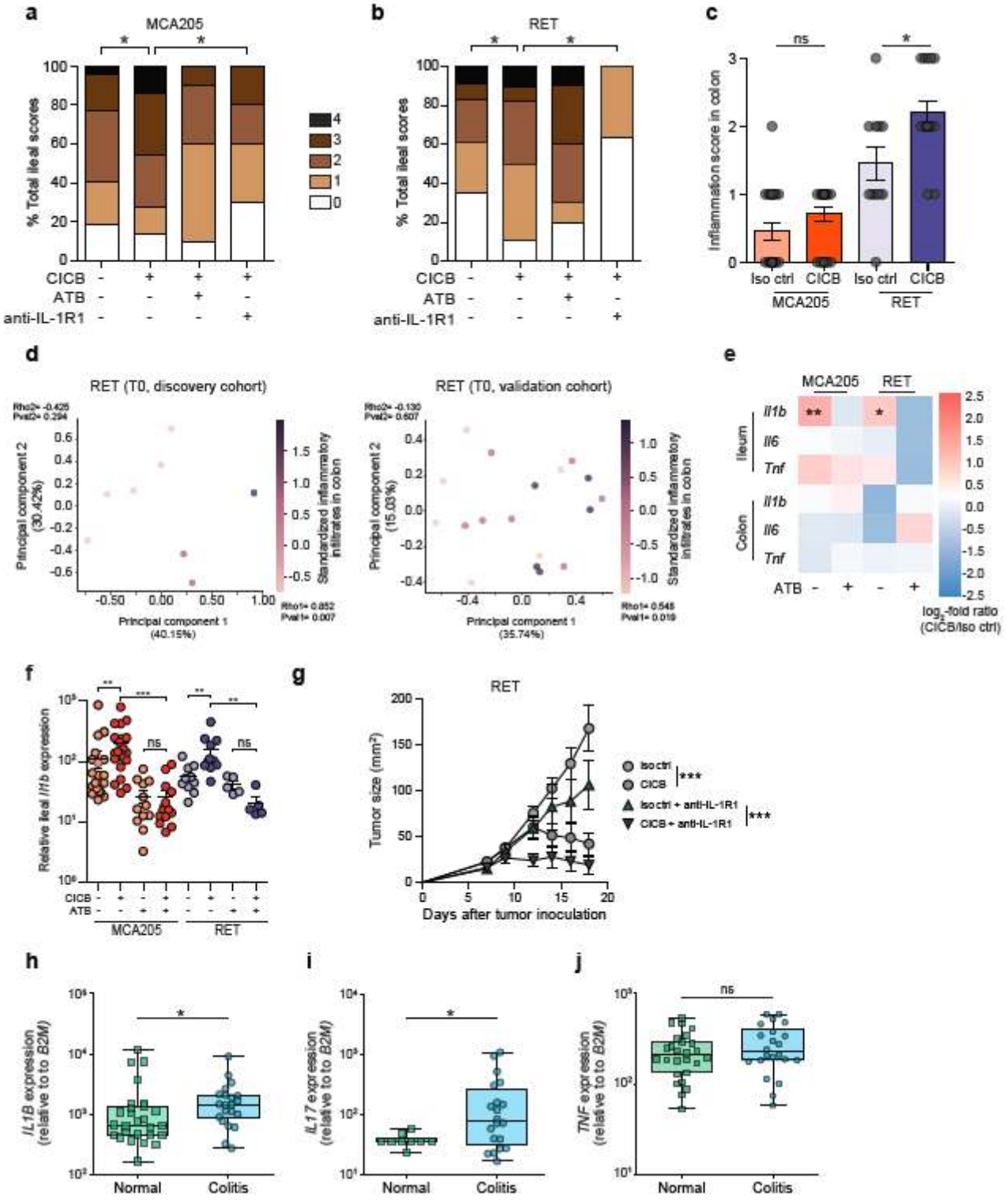

Figure 3

Role of gut microbiota and ileal IL-1β in CICB-induced intestinal inflammation in tumor bearing mice. (a-b) Scoring (range 0-4) of H&E-stained inflammatory infiltrates and pathological lesions of the ilea in MCA205 (a) or RET (b) tumor-bearing mice treated with isotype control or CICB, ± antibiotics (ATB) or with the IL-1R1 antagonist, anakinra, 24 hours after at least one CICB injection (n=9-22/group). Student's t-test. (c) Scoring of inflammatory colonic lesions in MCA205 and RET tumor-bearing mice analogous to that

shown in (a,b). (d) Beta diversity ordination (Bray-Curtis dissimilarity) of the fecal microbiota assessed by sequencing of 16S rRNA gene amplicons colored according to score of colonic inflammatory infiltrate in RET tumor-bearing mice. Intensity of purple indicates increasing inflammatory infiltrate score in the discovery (left panel) and validation (right panel) cohorts. Bacterial relative abundances and colonic inflammatory infiltrate were both normalized and standardized before correlation analysis. Pearson correlation and associated p values comparing each principal component with inflammatory infiltrate are indicated. (e) Heat map of log2-fold change of pro-inflammatory immune gene expression (CICB-treated vs isotype) in ilea and colons of MCA205 and RET tumor-bearing mice ± antibiotics. n=10-22 mice/group. Mann-Whitney test: *p<0.05, **p<0.01, ***p<0.001. (f) Relative ileal Il1b expression in tumor-bearing mice treated with isotype/CICB ± antibiotics. n=10-22 mice/group. Mann-Whitney test: *p<0.05, **p<0.01, ***p<0.001. (g) Tumor growth kinetics of RET melanoma in mice treated with CICB ± IL-1R1 antagonist anakinra. Shown are mean ± SEM tumor sizes from a representative experiment of two yielding similar results, comprising 6 mice/group. Anova test: ***p<0.001. (h-j) Quantitative PCR measurement of the relative IL1B (a), IL17 (b), TNF (c) expression in colon samples from melanoma patients experiencing immune-related colitis following ICB comparing areas of active inflammation (colitis) with areas of normal colonic tissue (‘normal’; intra-patient or cancer-free controls). Additional details are provided in Supplementary Data Table 7.

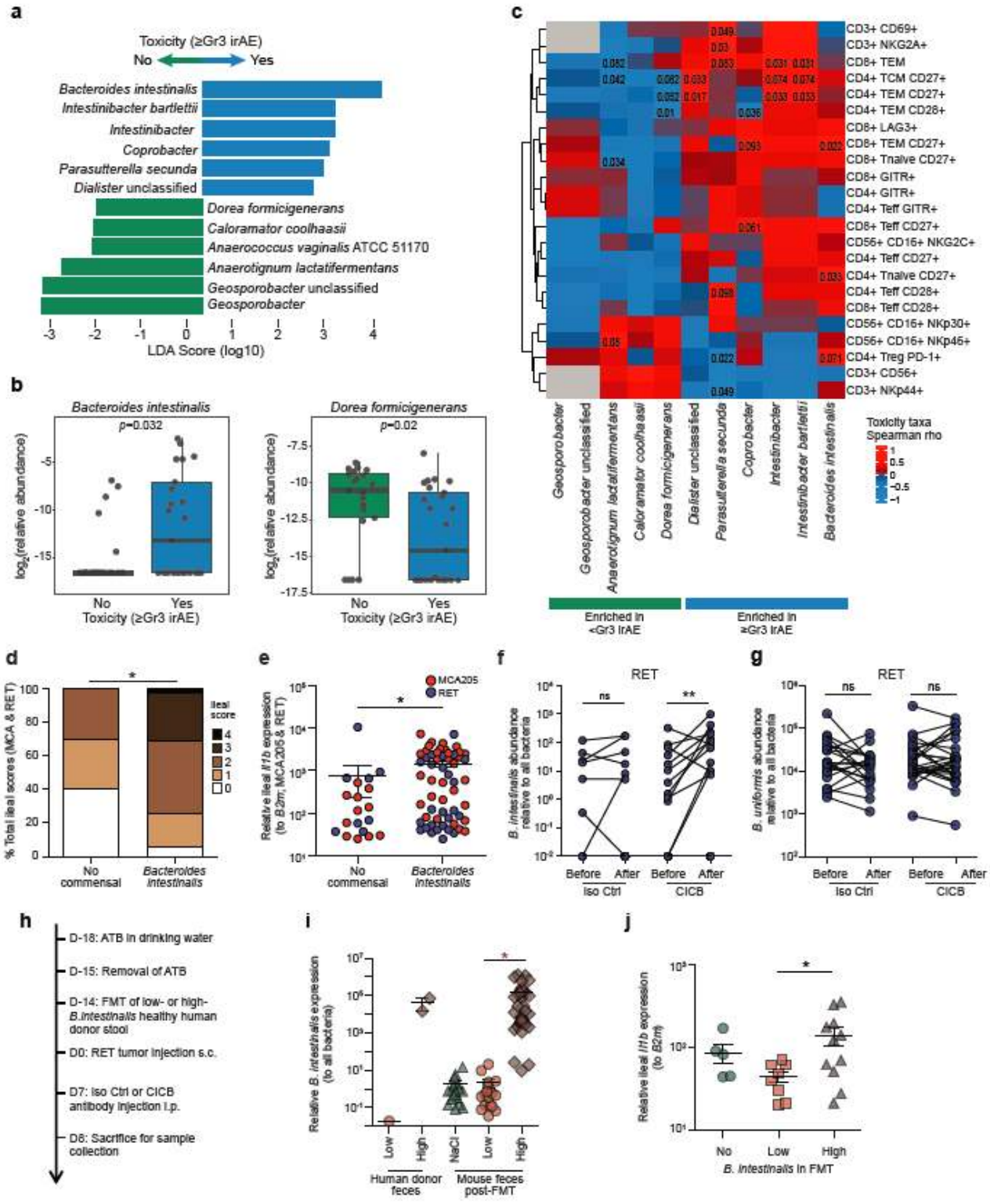

Figure 4

B. intestinalis is associated with intestinal IL-1β and colitis in the melanoma cohort. (a) LDA-score plot from LEfSe analysis of bacterial taxa significantly associated with development of, or freedom from, high-grade (≥Grade 3, n=29) immune-related adverse events (irAE) in all patients with available fecal samples (n=54). p<0.05. (b) Abundance of candidate taxa by WMS (total n=46 patients: n=25 ≥Grade 3, n=21 <Grade 3). Mann-Whitney p-values as shown. (c) Heatmap of correlation (Spearman's rho) between

key toxicity-associated or non-toxicity-associated bacterial taxa and circulating immune subsets quantitated by multiparameter flow cytometry of baseline blood samples (n=13). TEff=T effector cells, TCM=T central memory, TEM=T effector memory, Treg=regulatory T cell. (d,e) Total ileal toxicity scores (d) and relative Il1b expression (e) across MCA205 and RET tumor models showing higher toxicity (d) and higher Il1b expression (e) in animals treated with antibiotic microbiota ablation and subsequently colonized with B. intestinalis by gavage versus spontaneous recolonization (no administered commensal). Ileal toxicity was assessed at day 9 post-tumor inoculation in MCA205 (n=5-32/group, red dots) and RET (n=5-26/group, blue dots) combined. Data represent a pool of two individual experiments using three different strains of B. intestinalis, at 48 hrs post-oral gavage. For ileal toxicity scoring, mice were classified according to low (score 0 or 1) vs high toxicity (score 2, 3 or 4) and compared by Chi-square test: *p<0.05. Ileal Il1b expression was analyzed using Mann-Whitney test: *p<0.05. (f,g) qPCR quantification of the relative abundance of Bacteroides intestinalis (f) or non-implicated B. uniformis (g) in feces of mice treated with isotype versus CICB, before and after therapy. Data are shown in a paired manner (lines link samples from individual mice). n=21-26 mice/group. Wilcoxon signed-rank test: **p<0.01. (h) Experimental schema for RET model mice receiving FMT from human donors of differing B. intestinalis abundance and subsequent treatment with CICB or control. (i) Differing B. intestinalis content in human donor feces (“Low” vs “High”) and murine colonization following FMT or sham FMT (NaCl) was confirmed by qPCR. Mann-Whitney test: *p<0.05. (j) Mice receiving FMT from B. intestinalis “high” donor feces displayed higher expression of Il1b measured by qPCR of ileal tissue sampled 24 hours after administration of a single dose of CICB. Mann Whitney test: *p<0.05.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigureED1.pdf
- FigureED2.pdf
- FigureED3.pdf
- FigureED4.pdf
- FigureED5.pdf
- FigureED6.pdf
- SupplementaryDataTables123679.docx
- SupplementaryDataTable4OTUinfo.csv
- SupplementaryDataTable5MWresponse.txt
- SupplementaryDataTable8MWtoxicity.txt