#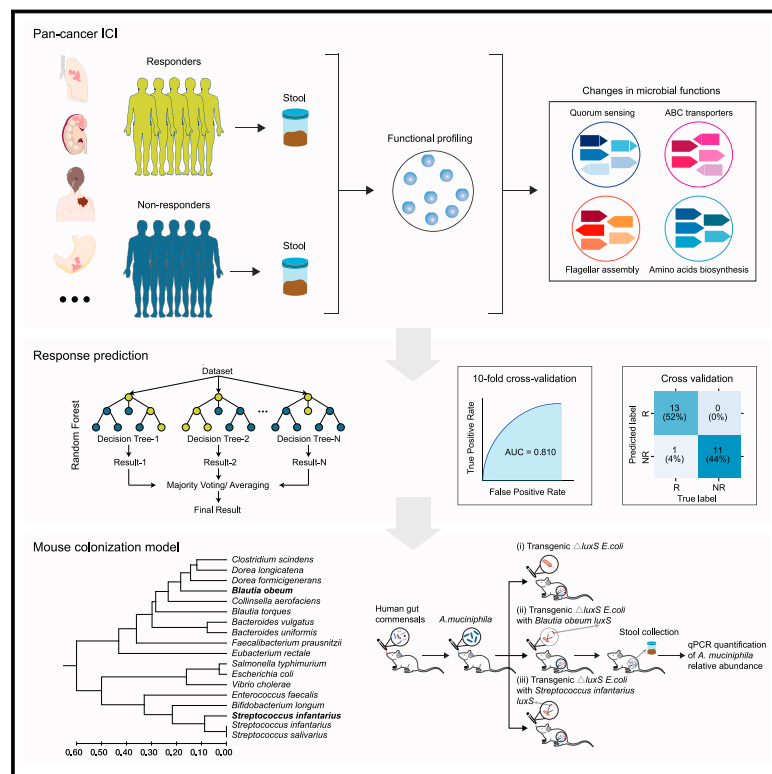 Enriched pathways in gut microbiome predict response to immune checkpoint inhibitor treatment across demographic regions and various cancer types

Graphical abstract

Authors

Xunhui Cai, Jennifer Y. Cho, Lijun Chen, ..., Jianjun Zhang, Ansel Hsiao, Tian Xia

Correspondence

jzhang20@mdanderson.org (J.Z.),
ansel.hsiao@ucr.edu (A.H.),
tianxia@hust.edu.cn (T.X.)

In brief

Microbiology; Bioinformatics; Cancer

Highlights

- Four pathways are enriched in ICI responders vs. non-responders across 12 datasets

- RF model using gene markers achieves an AUC of 0.810 in ICI response prediction

- AI-2 promotes *A. muciniphila* colonization via community-dependent mechanisms

CellPress

Article

Enriched pathways in gut microbiome predict response to immune checkpoint inhibitor treatment across demographic regions and various cancer types

Xunhui Cai,[1] Jennifer Y. Cho,[2,3] Lijun Chen,[1] Yufeng Liu,[4] Fenghu Ji,[4] Katia Salgado,[2] Siyi Ge,[2] Dehua Yang,[5] Hui Yu,[6,7] Jianbo Shao,[6] P. Andrew Futreal,[8] Boris Sepesi,[9] Don Gibbons,[10] Yaobing Chen,[1] Guoping Wang,[1] Chao Cheng,[11] Meng Wu,[12] Jianjun Zhang,[8,10,*] Ansel Hsiao,[2,*] and Tian Xia[1,4,13,*]

[1]Institute of Pathology, Tongji Hospital, Department of Pathology, School of Basic Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[2]Department of Microbiology and Plant Pathology, University of California, Riverside, Riverside, CA, USA
[3]Department of Biochemistry, University of California, Riverside, Riverside, CA, USA
[4]School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, China
[5]The National Center for Drug Screening, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China
[6]Clinical Laboratory, Wuhan Children's Hospital, Wuhan, China
[7]Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[8]Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
[9]Department of Thoracic and Cardiovascular Surgery, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
[10]Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
[11]Department of Medicine, Baylor College of Medicine, Houston, TX, USA
[12]Department of Immunology, Harvard Medical School, Boston, MA, USA
[13]Lead contact
*Correspondence: jzhang20@mdanderson.org (J.Z.), ansel.hsiao@ucr.edu (A.H.), tianxia@hust.edu.cn (T.X.)
https://doi.org/10.1016/j.isci.2025.112162

SUMMARY

Understanding the effect of gut microbiota function on immune checkpoint inhibitor (ICI) responses is urgently needed. Here, we integrated 821 fecal metagenomes from 12 datasets to identify differentially abundant genes and construct random forest models to predict ICI response. Gene markers demonstrated excellent predictive performance, with an average area under the curve (AUC) of 0.810. Pathway analyses revealed that quorum sensing (QS), ABC transporters, flagellar assembly, and amino acid biosynthesis pathways were enriched between responders (R) and non-responders (NRs) across 12 datasets. Furthermore, *luxS*, *manA*, *fliC*, and *trpB* exhibited consistent changes between R and NR across 12 datasets. Follow-up microbiota transplant experiments showed that inter-species signaling by different QS autoinducer-2 (AI-2) molecules (synthesized by *luxS*) can act on overall community function to promote the colonization of *Akkermansia muciniphila*, which is associated with superior ICI responses. Together, our data highlight the role of gut microbiota function in modulating the microbiome and antitumor immunity.

INTRODUCTION

Cancer immunotherapy is considered the "fifth pillar" of cancer therapy, joining the ranks of surgery, chemotherapy, radiation, and targeted therapy.[1] The principle of immunotherapy is to boost or restore the ability of the immune system to detect and destroy cancer cells by overcoming mechanisms by which tumors evade and suppress the immune response.[2] Immune checkpoint inhibitors (ICIs) targeting cytotoxic T-lymphocyte-associated protein 4 (CTLA-4)[3] and the programmed cell death protein 1 (PD-1)[4] have revolutionized the therapeutic landscape of many cancer types, but responses to ICIs are often heterogeneous. A recently recognized important host factor that could contribute to inter-patient heterogeneity is the differential composition of the patients' gut microbiota and microbiome,

which has been shown to affect antitumor immunity and therapeutic efficacy of ICIs.[5-8]

Interactions between the host immune system and microbiome have been recently scrutinized to identify immunogenic bacterial species that may impact responses to ICI therapy through facilitating recruitment and activation of CD8[+] T cells, identifying several potential bacterial determinants.[5-10] Gopalakrishnan et al.[5] found higher abundance of *Faecalibacterium prausnitzii* in responders (R) compared to non-responders (NRs) with metastatic melanoma treated with anti-PD-1, while, in a different cohort of patients with metastatic melanoma, Matson et al.[6] reported that response to anti-PD-1 therapy was associated with an increased abundance of a consortium of eight species driven by *Bifidobacterium longum*. In another study of metastatic melanoma patients, Frankel et al.[8] observed
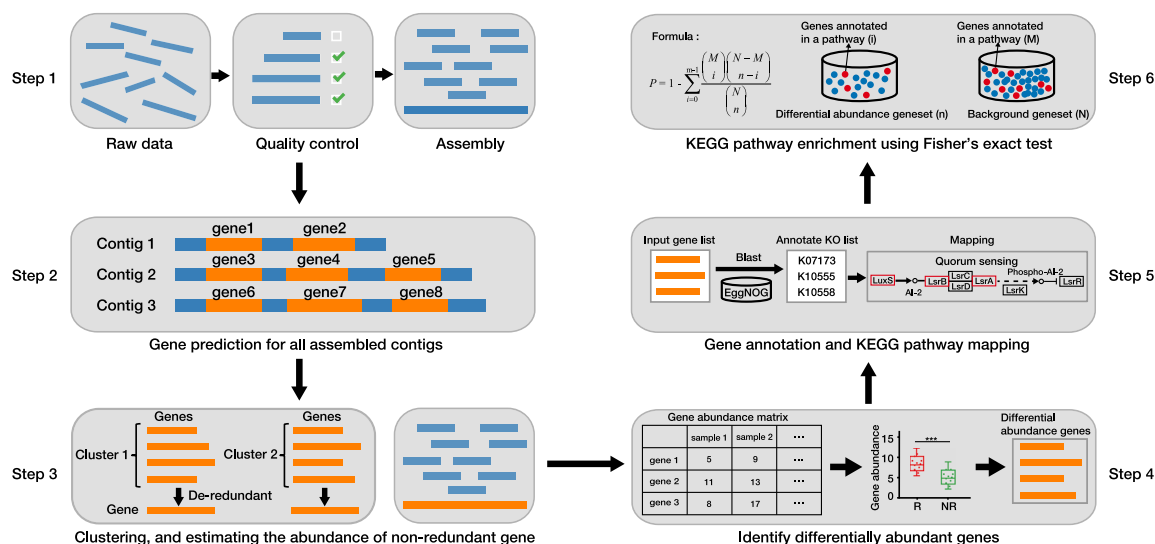
Figure 1. Schematic diagram illustrating the pipeline for identifying enriched pathways in metagenomic shotgun sequencing data
Step 1, raw sequence reads are quality-filtered by removing adapters, low-quality reads (shorter than 20 bp), and human sequences. Step 2, qualified reads are assembled into longer contigs, followed by gene prediction in the assembled contigs to overcome challenges posed by repeat sequences. Step 3, genes are clustered to form reference gene catalogs that are used to map reads from each sample and generate a gene abundance matrix. Step 4, the Wilcoxon rank-sum test is used to identify genes with significantly different abundances between the R and NR groups. Step 5, reference gene catalogs and differentially abundant gene sets are mapped to the eggNOG database for functional annotation. Step 6, Fisher's exact test is used to identify significantly enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways in a cohort, with correction for the false discovery rate (FDR).

that responders to anti-PD-1 therapy were enriched for *Dorea formicigenerans*. In non-small cell lung cancer (NSCLC), Routy et al.[7] observed an increased relative abundance of *Akkermansia muciniphila* in anti-PD-1 R compared to NR.

Species-centric analysis to identify consistent microbial biomarkers for predicting ICI efficacy is limited by rapid evolution at the strain level and the enormous inter-individual heterogeneity in human gut microbiomes.[11] Instead, the functional capacity of the microbiota may be a better biomarker for tumor response to ICI therapy. However, a comprehensive analysis of the functional capacity of the gut microbiome across different ICI cohorts is lacking. To address this, we first integrated fecal shotgun metagenome data collected prior to ICI treatment from 12 human datasets across various cancer types, geographic locations, and multiple checkpoint inhibitors. Then, genes were predicted from metagenomes of ICI R and NR, and differential analysis was conducted to identify the gene markers that distinguish between the two groups. To evaluate the performance of gene markers in predicting ICI response, classifiers were constructed using six popular machine learning methods based on differentially abundant genes. Furthermore, we applied a pathway-centric computational pipeline to identify the consistent changes in pathways and genes between R and NR across studies. Using a mouse colonization model, we examined the role of *luxS* (which synthesizes quorum sensing [QS] autoinducer-2 [AI-2] molecules) in driving the fitness or abundance of immunogenic bacterium *A. muciniphila*. Our study not only revealed a comprehensive pan-cancer intestinal functional profile in ICI R and NRs across multiple nations but also contributed to the accurate prediction of ICI response using microbial gene markers.

RESULTS

Characterization of shotgun metagenome data from patients receiving ICI therapy

We collected fecal metagenome data from 12 public datasets with available information on response to ICIs and associated metadata (Table S1). Data from samples collected after the initiation of ICI therapy were excluded to focus on microbial features potentially predictive of ICI response. Upon sample filtering, 821 metagenomes were retained. R and NR categorization was obtained from the metadata of each dataset. Metagenomes from 3 cohorts (Derosa et al.,[12] McCulloch et al.,[13] and Gopalakrishnan et al[5]) were obtained from patients with anti-PD-1 alone, whereas metagenomes from 5 cohorts (Matson et al.,[6] Lee et al.,[10] Peng et al.,[14] Routy et al.,[7] and Gunjur et al[15]) were obtained from patients undergoing anti-PD-1, -PD-L1, -CTLA-4, or combination immunotherapy. In total, 448 R and 373 NR patients were included. These datasets comprised patients from 6 countries (China, France, the Netherlands, the UK, the USA, and Australia), with 6 different tumor types (melanoma, gastrointestinal cancer [GI], NSCLC, renal cell carcinoma [RCC], neuroendocrine neoplasms [NENs], and upper gastrointestinal and biliary cancers [UGB]).

Construction of the ICI response prediction model based on differentially abundant genes

To identify genes with the potential for predicting ICI response across multiple geographical locations and tumor types, we first conducted a differential abundance analysis on the combined dataset consisting of 821 fecal metagenomes from 12 datasets (Figure 1, step 1–4). In total, 55,292 genes (representing 1.18%
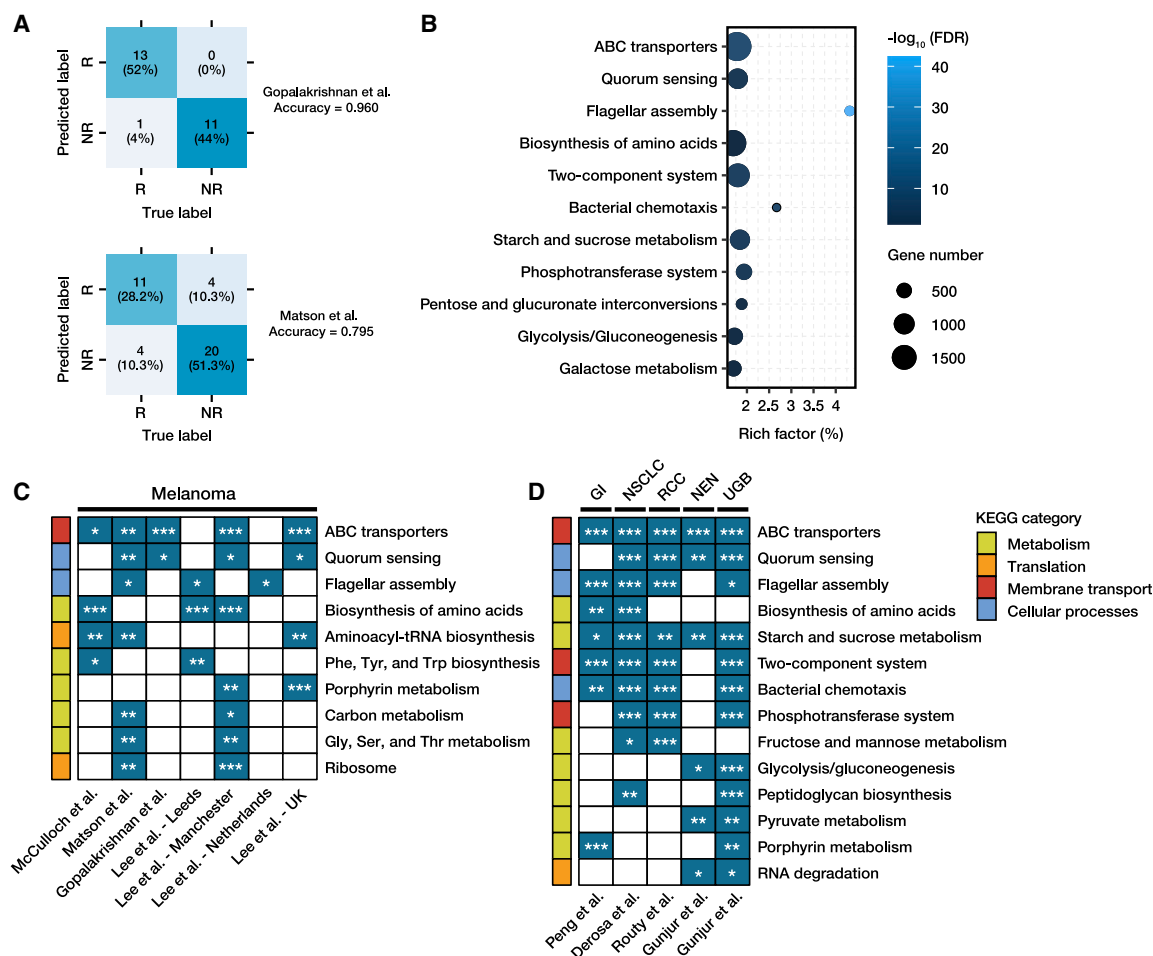
Figure 2. Microbial functional alterations between the R and NR groups across cohorts

(A) Confusion matrix evaluations of prediction models with 55,292 features in Gopalakrishnan et al.'s and Matson et al.'s cohorts.

(B) Pathway enrichment analysis based on genes with significantly different abundance in the combined dataset. The y axis indicates the pathway name, and the x axis indicates rich score. Rich score represents the percentage of the number of differentially abundant genes annotated in a pathway relative to the total number of genes annotated in that pathway. A higher rich score indicates a greater degree of enrichment. The bubble size indicates the number of genes. The color bar indicates the FDRs.

(C) Enriched pathways of differentially abundant genes between R and NR when pooling 7 melanoma datasets with anti-PD-1 therapy. The heatmap displays only pathways that are enriched in at least two of the seven independent datasets. The significance was determined using Fisher's exact test, and p values were adjusted by FDR method. *FDR < 0.05; **FDR < 0.01; ***FDR < 0.001.

(D) Enriched pathways between R and NR when pooling five other tumor-type datasets with anti-PD-1 therapy. The heatmap displays only pathways that are enriched in at least two of the five independent datasets. The significance was determined using Fisher's exact test, with p values adjusted for FDR. *FDR < 0.05; **FDR < 0.01; ***FDR < 0.001. GI, gastrointestinal cancer; NSCLC, non-small cell lung cancer; RCC, renal cell carcinoma; NEN, neuroendocrine neoplasms; UGB, upper gastrointestinal and biliary cancers.

of all characterized genes) were identified with significantly differential abundance ($p < 0.05$, Wilcoxon rank-sum test, |log2(fold change)| ≥ 1) between R and NR groups. Utilizing these differentially abundant genes, we established 6 popular machine learning classification models by combining 10 datasets with anti-PD-1 therapy (see STAR Methods). The cross-cohort validation showed that the random forest (RF) classifier achieved the highest accuracy, precision, and F1-score when evaluated on the unseen Gopalakrishnan et al. and Matson et al. cohorts, suggesting that the RF classifier model is the optimal choice for predicting ICI response across various studies (Figure 2A; Table S2).

To identify the functional characteristics of these genes, we further performed EggNOG annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis (Figure 1, step 5–6). Among the 55,292 genes, 41,467 were functionally annotated and we identified 11 pathways that showed statistically significant enrichment (Fisher's exact test, false discovery rate [FDR] < 0.05) in the combined dataset (Figure 2B). The 11 enriched pathways included QS, ABC transporters, flagellar assembly, biosynthesis of amino acids, two-component system, bacterial chemotaxis, starch and sucrose metabolism, phosphotransferase system, pentose and glucuronate

A

B
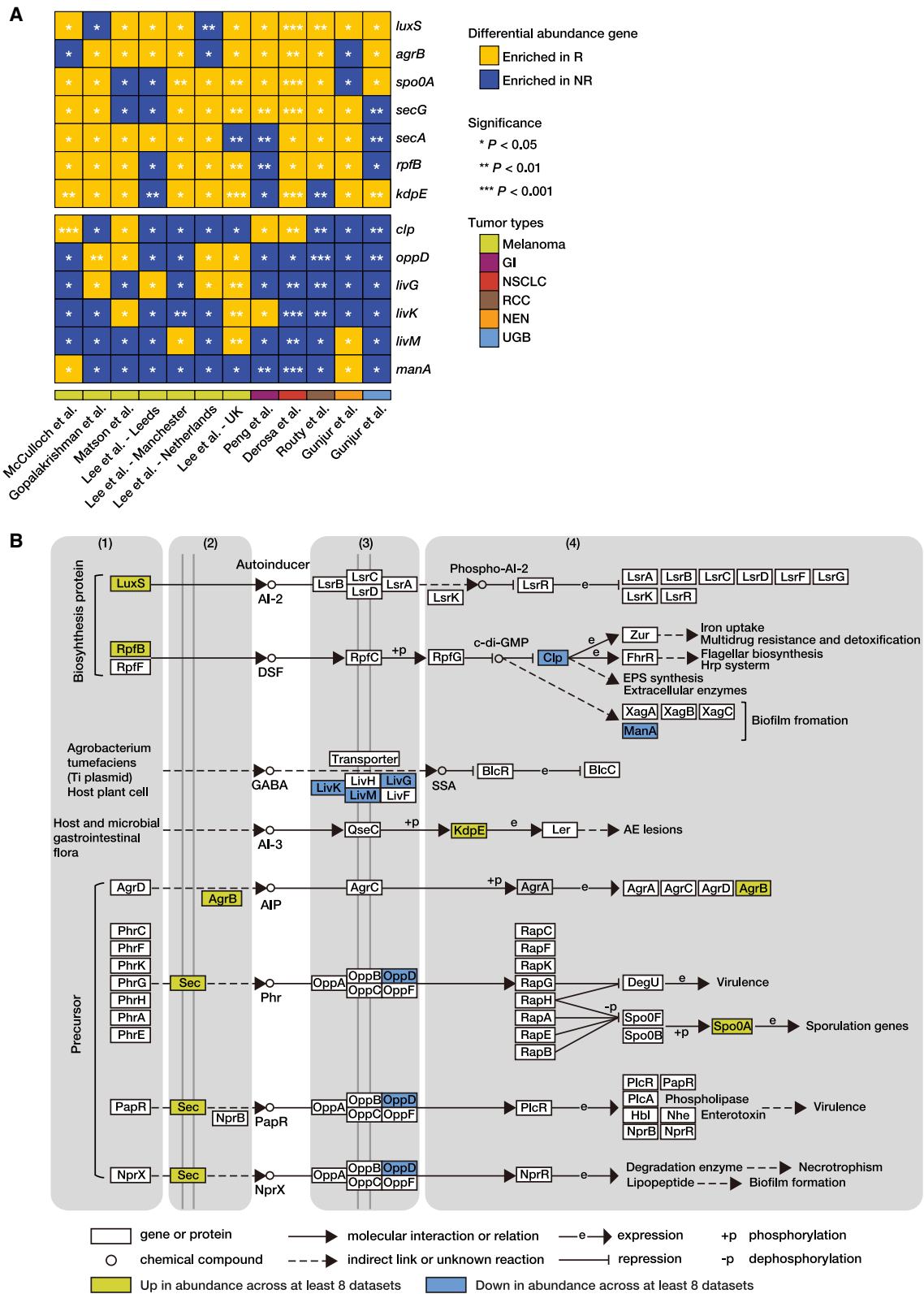
(legend on next page)

interconversions, glycolysis/gluconeogenesis, and galactose metabolism. Next, we constructed 10-fold cross-validation RF models using the differentially abundant genes belonging to these 11 pathways (5,761 genes) as selected features, which achieved an average area under the curve (AUC) of 0.763 for predicting R and NR in the discovery cohort. The accuracy of cross-cohort validation was 0.880 in the Gopalakrishnan et al. cohort and 0.744 in the Matson et al. cohort (Figure S1A). These data highlight that the 11 enriched pathways of the gut microbiome represent potential biomarkers for accurately predicting ICI response phenotype.

Uniform computational analysis across cohorts identifies consistently enriched KEGG pathways

To test whether the identified enriched pathways in the combined dataset are common and robust across studies, we applied pathway analysis pipeline (Figure 1) to 7 independent metastatic melanoma datasets, including those from Matson et al. ($n = 39$, 15 R, 24 NRs), Gopalakrishnan et al. ($n = 25$, 14 R, 11 NRs), McCulloch et al. ($n = 94$, 49 R, 45 NRs), and Lee et al. (United Kingdom: $n = 55$, 33 R, 22 NRs; Netherlands: $n = 55$, 33 R, 22 NRs; Manchester: $n = 25$, 13 R, 12 NRs; Leeds: $n = 18$, 10 R, 8 NRs) cohorts (Table S1). Given that pathways achieving significant enrichment (FDR < 0.05) in several individual datasets are more compelling, we conducted meta-analysis to combine the pathways enriched in two or more datasets. We identified a total of 10 candidate pathways that have statistically significant differential enrichment between R and NR groups (Figure 2C). Among these pathways, ABC transporters demonstrated significant differential enrichment in the highest number of datasets, followed by QS, flagellar assembly, biosynthesis of amino acids, aminoacyl-tRNA biosynthesis, phenylalanine, tyrosine and tryptophan biosynthesis, porphyrin metabolism, carbon metabolism, glycine, serine and threonine metabolism, and ribosome. These terms were further categorized into four major processes: metabolism, translation, membrane transport, and cellular processes.

To understand whether these pathways are specific to melanoma or tumor agnostic, we applied our analysis pipeline (Figure 1) in five other tumor-type datasets treated with ICI therapy, including Peng and colleagues in GI ($n = 40$, 25 R, 15 NRs), Derosa and colleagues in NSCLC ($n = 338$, 177 R, 161 NRs), Routy and colleagues in RCC ($n = 62$, 42 R, 20 NRs), and Gunjur and colleagues in NEN ($n = 32$, 22 R, 10 NRs) and UGB ($n = 38$, 15 R, 23 NRs). We identified enriched pathways separately in each of the five cohorts and then combined the results with meta-analysis. A total of 14 candidate pathways were found to be statistically enriched between R and NR in at least two cohorts, including QS, ABC trans-porters, flagellar assembly, biosynthesis of amino acids, starch and sucrose metabolism, two-component system, bacterial chemotaxis, phosphotransferase system, fructose and mannose metabolism, glycolysis/gluconeogenesis, peptidoglycan biosynthesis, pyruvate metabolism, porphyrin metabolism, and RNA degradation (Figure 2D).

Taken together, these results showed that ABC transporters, QS, flagellar assembly, and biosynthesis of amino acids pathways are significantly enriched between R and NR across different cancer types, geographic locations, and various checkpoint inhibitors. Of note, we observed that ABC transporters and QS pathways consistently co-enriched in eight of the twelve datasets (Figures 2C and 2D), suggesting an amalgamation of the ABC transporters and QS pathways. In addition, we constructed 10-fold cross-validation RF models using the differentially abundant genes belonging to these four pathways (3,522 genes) as input features, which achieved an average AUC of 0.750 for predicting R and NR in the discovery cohort. The accuracy of cross-cohort validation was 0.800 in the Gopalakrishnan et al. cohort and 0.718 in the Matson et al. cohort (Figure S1B). Considering that these four pathways have a profound impact on the composition of the gut microbiota[16–19] and perform well in predicting ICI response, subsequent analyses focused on the genes with consistent changes from these pathways.

Bacterial QS genes showed uniform changes between R and NR across multi-datasets

Gut bacteria use QS to efficiently synchronize behaviors across populations in response to changes in population density and community species composition by the signaling molecules called autoinducers (AIs).[20] In the QS process, ABC transporters are primarily involved in the transmembrane transport of AI molecules, such as AI-2 and peptide signals.[21,22] Next, we examined differential abundance of QS genes across all 12 datasets and then combined the results with meta-analysis. In total, 13 differentially abundant genes were identified with consistent changes, including 7 genes that were enriched in R and 6 genes that were enriched in NR (Figure 3A). Specifically, the abundance of *luxS* showed the most consistent enrichment in the R group ($p < 0.05$, log2(fold change) ≥ 1 in 10 of the 12 datasets), followed by *agrB*, *spo0A*, *secG*, *secA*, *rpfB*, and *kdpE*. In contrast, *manA*, an extracellular mannanase that influences biofilm maturation,[23] exhibited the most consistent enrichment in NR ($p < 0.05$, log2(fold change) ≤ -1 in 10 of the 12 cohorts), followed by *livM*, *livK*, *livG*, *oppD*, and *clp*. A KEGG map (map02024) demonstrates that the QS process can be divided into at least 4 steps: (1) production of AIs by the bacterial cell; (2) release of AIs, either actively or passively, into the surrounding

Figure 3. Alterations in quorum sensing genes between R and NR in 12 ICI datasets
(A) Heatmap displays quorum sensing genes with differential abundance between R and NR when pooling 12 datasets with anti-PD-1 therapy across cancer types. The heatmap displays log2-normalized abundance fold changes in each dataset. To standardize across different datasets, we further performed a *Z* score transformation on the log-transformed values. Yellow and blue colors represent genes with significant differential abundance ($p < 0.05$, |log2(fold change)| ≥ 1) between R and NR in at least eight of the twelve independent datasets. Significance was determined using the Wilcoxon rank-sum test. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. GI, gastrointestinal cancer; NSCLC, non-small cell lung cancer; RCC, renal cell carcinoma; NEN, neuroendocrine neoplasms; UGB, upper gastrointestinal and biliary cancers.
(B) Quorum sensing pathway as presented in KEGG. Genes enriched in R and NR are marked in yellow and blue, respectively.
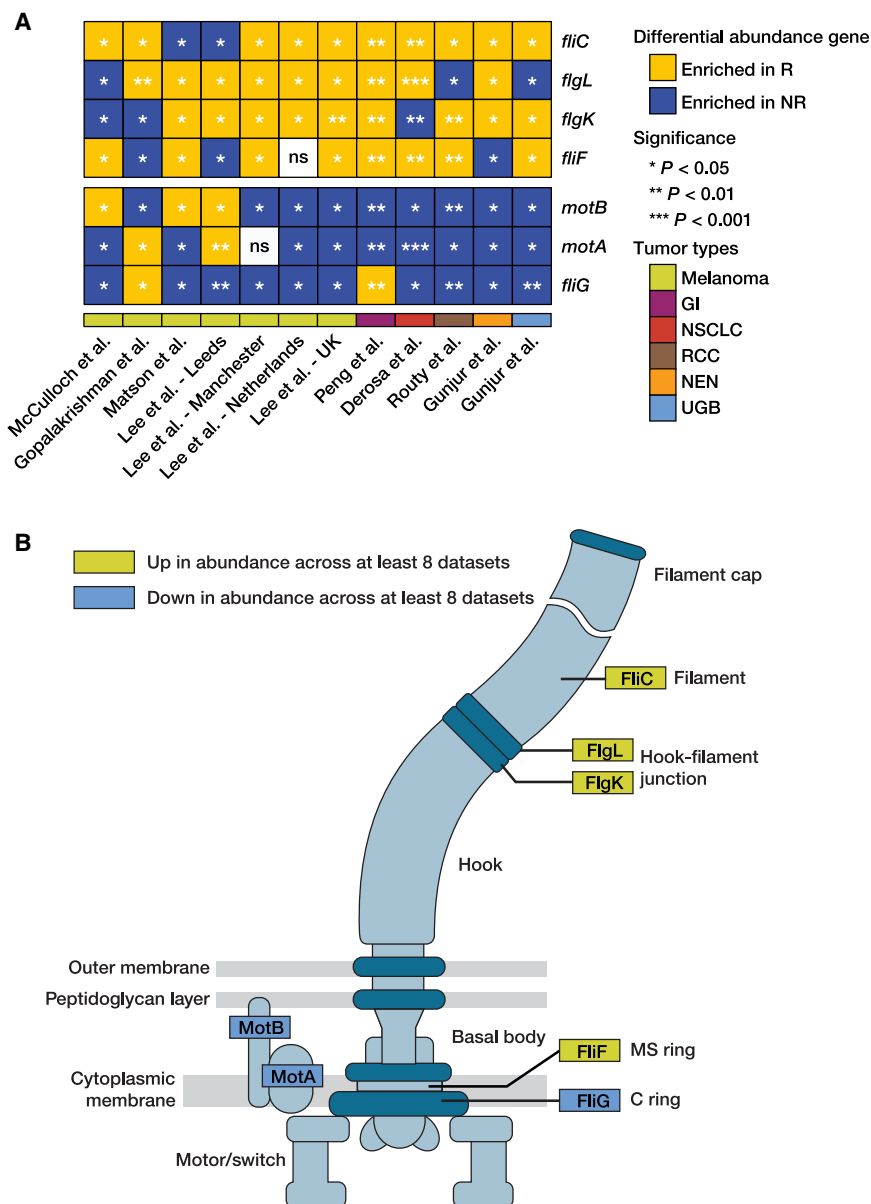
Figure 4. Alterations in flagellar assembly genes between R and NR in 12 ICI datasets

(A) Flagellar assembly genes exhibit differential abundance between R and NR when pooling 12 datasets with anti-PD-1 therapy across cancer types. The heatmap displays log2-normalized abundance fold changes in each dataset. To standardize across different datasets, we further performed a Z score transformation on the log-transformed values. Yellow and blue colors represent genes with significant differential abundance ($p < 0.05$, $|log2(fold\ change)| \geq 1$) between R and NR in at least eight of the twelve independent datasets. Significance was determined using the Wilcoxon rank-sum test. ns, $p > 0.05$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. GI, gastrointestinal cancer; NSCLC, non-small cell lung cancer; RCC, renal cell carcinoma; NEN, neuroendocrine neoplasms; UGB, upper gastrointestinal and biliary cancers.

(B) Flagellar assembly pathway as presented in KEGG. Genes enriched in R and NR are marked in yellow and blue, respectively.
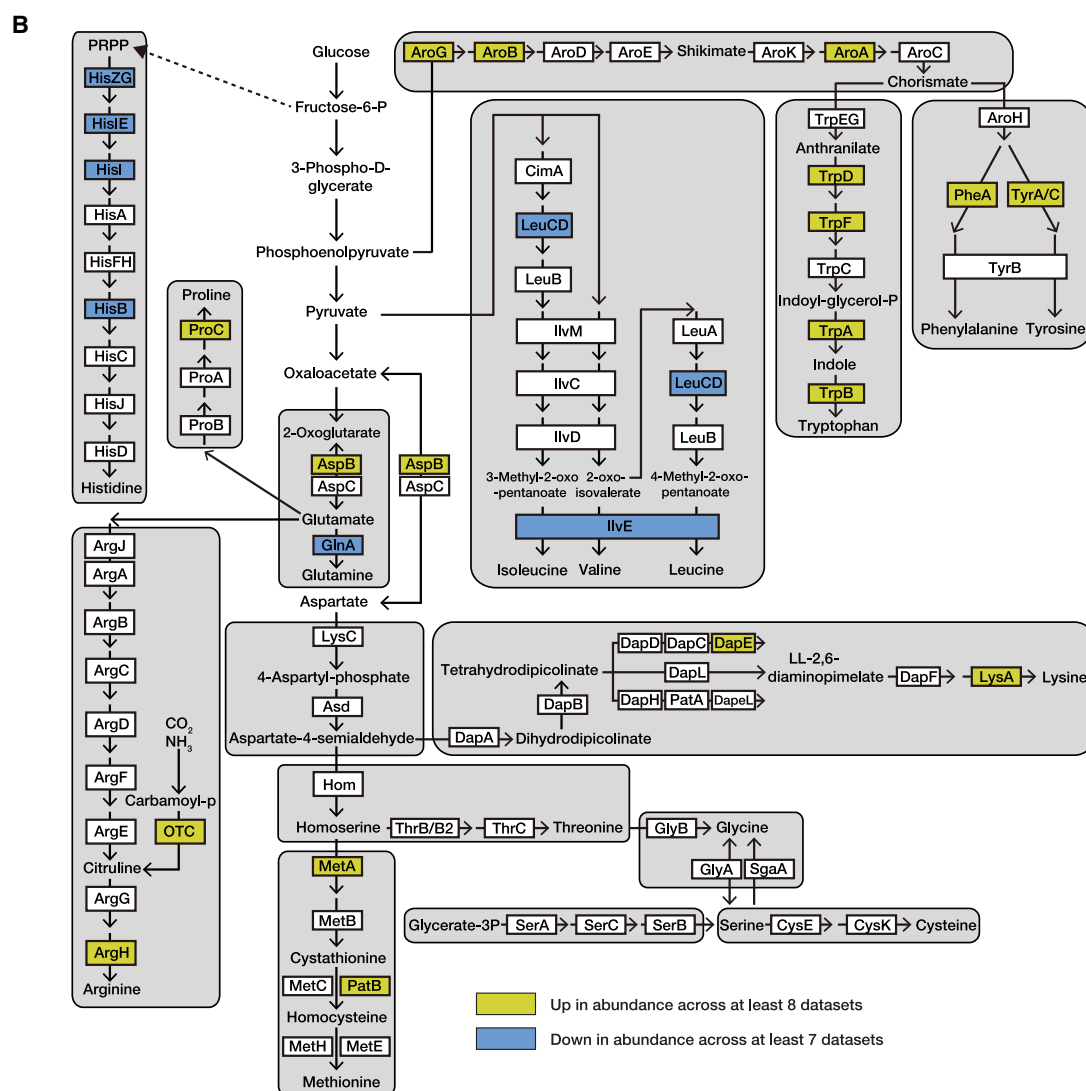
tion factor that regulates numerous cellular processes, including iron uptake and flagellar biosynthesis.[28]

Bacterial flagellar genes showed consistent changes between R and NR across cohorts

The primary function of the flagellum is motility, and the expression and synthesis of more than 30 proteins require a significant amount of energy from the bacterium.[29] We next examined whether any gut bacterial flagellar assembly genes had significantly differential abundance between the R and NR groups across 12 datasets. In total, seven differentially abundant genes were identified with consistent changes, including *fliC*, *flgL*, *flgK*, and *fliF* that were enriched in R and *fliG*, *motA*, and *motB* that were enriched in NR (Figure 4A). A KEGG map (map02040) demonstrates that the bacterial flagellum can be divided into three substructures: the filament, the hook, and the basal body (Figure 4B). The enriched R genes include *flgL* and *flgK*, which are required to form the junction between the hook and the filament, as well as flagellin *fliC*, the basic subunit of filament, which has been extensively examined for its ability to activate host immunity as an antitumor agent.[30] In addition, FliF is the protein comprising the MS ring of the bacterial flagellar basal body, which is the base for the assembly of flagellar axial structures.[31] In contrast, all three enriched NR genes are known to be involved in motor rotation. FliG, along with FliM and FliN, forms the rotor of the bacterial flagellar motor, while MotA and MotB form a transmembrane proton channel, serving as the stator of the bacterial flagellar motor to couple proton flow with torque generation.[32]

environment; (3) recognition of AIs by specific receptors; and (4) leading to changes in gene regulation once they exceed a threshold concentration (Figure 3B). Interestingly, we found that the enriched R genes were mainly associated with the synthesis and export of AIs, while the enriched NR genes were mainly related to the recognition and import of AIs (Figure 3B). Of note, *luxS* was involved in the synthesis of the signal molecule AI-2, which can mediate communication within and between many bacterial species, regulate the composition of the intestinal flora, and correct imbalances of the intestinal flora.[24,25] *spo0A* plays a crucial role in initiating sporulation, which is one of the protective mechanisms that some bacteria employ to respond to environmental changes, such as nutritional starvation and extreme growth conditions.[26] KdpE is a response regulator that activates expression of *ler*,[27] and Clp is a global transcrip-

A

B

(legend on next page)

Alteration of amino acid biosynthesis genes between R and NR across multi-datasets

Gut microorganisms can degrade undigested luminal proteins and peptides into amino acids. Although these amino acids cannot be absorbed by colon epithelium to a large extent, they are precursors of many end products of metabolism, such as short-chain fatty acids[33] and indole-3-aldehyde,[34] which are considered to be beneficial for ICI treatment. We next investigated genes exhibiting differential abundance associated with the biosynthesis of amino acids across all 12 datasets and then integrated the results through meta-analysis. In total, we identified 29 differentially abundant amino acid-associated genes ($p < 0.05$, |log2(fold change)| $\geq$ 1 in at least 8 datasets) with consistent changes, 17 genes enriched in the R group and 12 genes enriched in the NR group (Figure 5A). To distinguish the subset of amino acids that exhibited uniform changes, we examined the amino acid biosynthesis metabolic map (map01230) and overlaid the differential abundance data on it (Figure 5B). Using this map, we identified that the 17 genes enriched in the R group were associated with the biosynthesis of aromatic amino acid (*trpD*, *trpF*, *trpA*, *trpB*, *pheA*, *tyrA2*, *aroG*, *aroB*, and *aroA*), arginine (*argH* and *OTC*), lysine (*lysA* and *dapE*), methionine (*metA* and *patB*), proline (*proC*), and glutamine (*aspB*). In contrast, the 12 genes enriched in the NR group were related to the biosynthesis of branched-chain amino acids (*leuD* and *ilvE*), histidine (*hisI*, *hisG*, *hisB*, *hisIE*, and *hisZ*), glutamine (*glnA*), and carbohydrate metabolism (*gpmI*, *acnA*, *gapA*, and *prsA*).

Evaluate the expression of differentially abundant genes in an independent validation cohort

None of these ICI studies on the gut microbiota metagenome have incorporated metatranscriptomic data into their analysis,[5–7,10,12–14] which are crucial for understanding actual expression levels of metagenomic functions in the microbial community. To confirm the robustness of our findings across the 12 metagenome datasets, we validated the enriched pathways and the expression of differentially abundant genes in Peters et al.[35] cohort, which performed metagenomic shotgun sequencing on fecal samples from 27 melanoma patients and metatranscriptome sequencing on a subset of 17. Specifically, we first applied the pathway analytic pipeline (Figure 1) to the metagenomic samples (n = 27, 15 R, 12 NRs) to characterize the overall enriched functional pathways. We identified a total of 12 significantly enriched pathways between R and NR groups (FDR < 0.05), including QS, flagellar assembly, arginine biosynthesis, valine, leucine and isoleucine biosynthesis, bacterial chemotaxis, two-component system, aminoacyl-tRNA biosynthesis, porphyrin and chlorophyll metabolism, RNA polymerase,

microbial metabolism in diverse environments, methane metabolism, and 2-Oxocarboxylic acid metabolism (Figure 6A). Of note, the QS, flagellar assembly, and two biosynthesis of amino acid pathways were all significantly enriched, confirming the robustness of these microbial functions between R and NR groups. Next, we examined the expression level of these biomarker genes identified across the 12 metagenome datasets using the metatranscriptome samples (n = 17, 8 R, 9 NRs). For the QS genes, the expression levels of *luxS*, *agrB*, *spo0A*, *secG*, *secA*, *rpfB*, and *kdpE* were also significantly up-regulated in the R group compared to the NR group ($p < 0.05$, log2(fold change) $\geq$ 1, Figure 6B). In contrast, 4 of the 6 genes enriched in NR, including *livK*, *livG*, *livM*, and *manA*, were also significantly down-regulated between R and NR groups ($p < 0.05$, log2(fold change) $\leq$ −1, Figure 6B). For the flagellar assembly genes, *fliC*, *fliF*, *flgL*, *flgK*, *motA*, *motB*, and *fliG* were all significantly highly represented in the R group compared to the NR group ($p < 0.05$, log2(fold change) $\geq$ 1, Figure 6C). For the amino acid biosynthesis genes, 16 of the 17 genes identified as enriched in R were also up-regulated in the metatranscriptome data ($p < 0.05$, log2(fold change) $\geq$ 1, except for *trpA*), and 7 of the 12 genes (including *leuD*, *ilvE*, *gpmI*, *acnA*, *gapA*, *prsA*, and *glnA*) enriched in NR were also down-regulated ($p < 0.05$, log2(fold change) $\leq$ −1, Figure S2). Overall, these results suggest that the majority of significantly differentially abundant genes in the microbial community exhibit consistent trends of change at both abundance and expression levels.

AI-2 was associated with colonization of immunogenic strains in mouse model

In the earlier analysis, we found that the AI-2 synthesis gene *luxS* was consistently enriched in the R group across ten of the twelve metagenome datasets, which was further validated in one metatranscriptome dataset. Given that AI-2 can regulate the composition of the gut microbiota,[25,36] this prompted us to further investigate its potential role in modulating the composition of immunogenic bacteria species. Recently, *A. muciniphila* was reported to mediate ICI response in several human cancer trials and mouse models.[7,10,12,37–39] Therefore, we will examine the role of AI-2 in driving the fitness or abundance of *A. muciniphila*. We first examined the genome of *A. muciniphila* for putative homologs to known AI-2-related pathways. We mapped *A. muciniphila* against the KEGG database to identify putative QS pathways. Strikingly, no *luxS*, *lsr*, or *luxPQ* homologs were identified. *A. muciniphila* also does not encode a Com or Agr autoinducer peptide (AIP) sensing/regulation system. These findings suggested that the relationship of *A. muciniphila* to AI-2 may be an indirect one working via the effect of bacterial signaling on the rest of the gut commensal community. To

Figure 5. Alterations in biosynthesis of amino acid genes between R and NR in 12 ICI datasets

(A) Differentially abundant biosynthesis of amino acid genes between R and NR was identified by pooling 12 datasets using anti-PD-1 therapy across cancer types. The heatmap displays log2-normalized abundance fold changes in each dataset. To standardize across different datasets, we further performed a Z score transformation on the log-transformed values. Yellow and blue colors represent genes with significant differential abundance ($p < 0.05$, |log2(fold change)| $\geq$ 1) between R and NR in at least seven of the twelve independent datasets. Significance was determined using the Wilcoxon rank-sum test. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. GI, gastrointestinal cancer; NSCLC, non-small cell lung cancer; RCC, renal cell carcinoma; NEN, neuroendocrine neoplasms; UGB, upper gastrointestinal and biliary cancers.

(B) Biosynthesis of amino acids pathway as presented in KEGG. Genes enriched in R and NR are marked in yellow and blue, respectively.
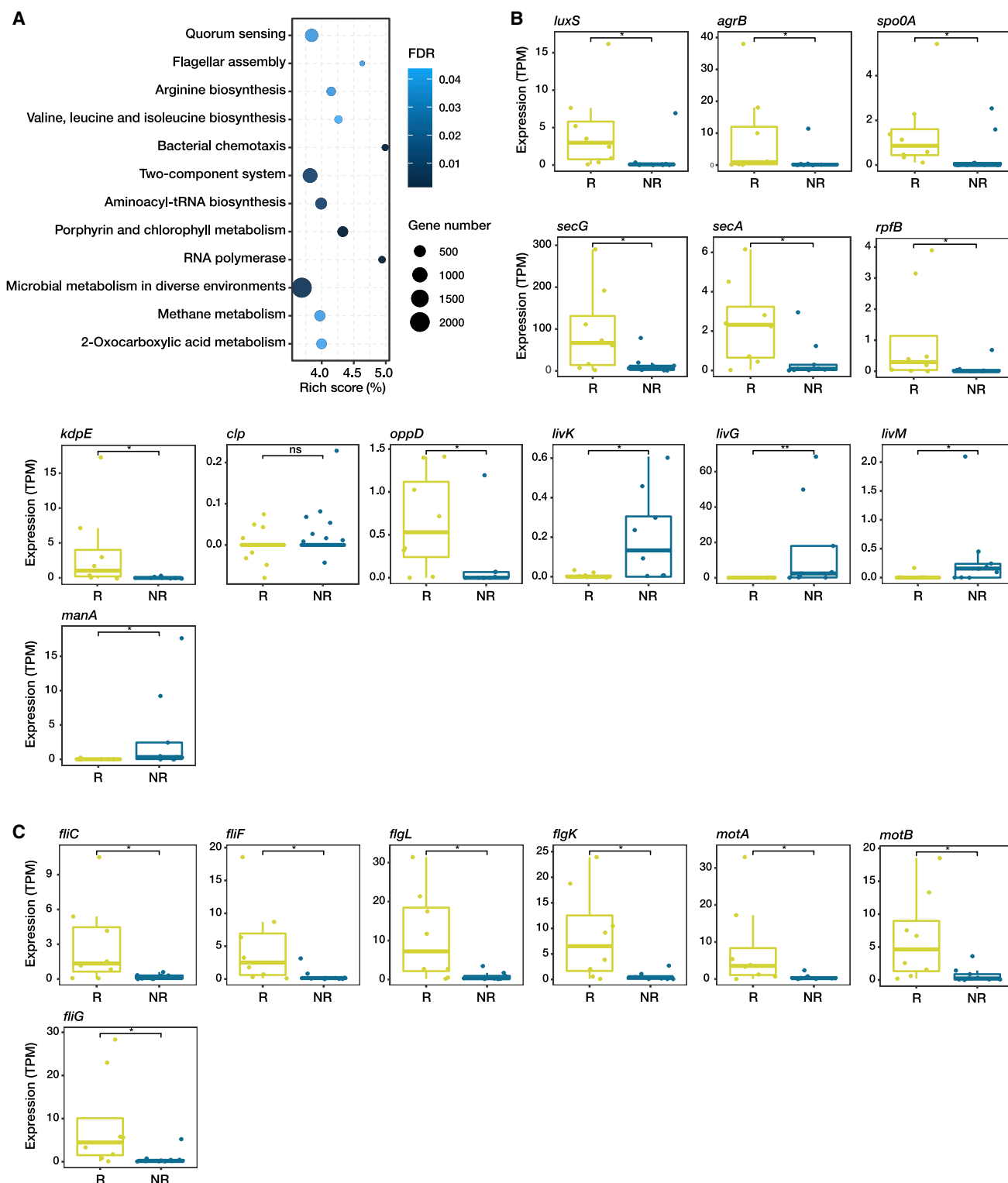
Figure 6. Evaluation of the enriched microbial pathways and the expression of biomarker genes in an independent cohort

(A) Pathway enrichment analysis based on genes with significantly different abundance in Peters et al.'s cohorts. The y axis indicates the pathway name, and the x axis indicates rich score. Rich score represents the percentage of the number of differentially abundant genes annotated in a pathway relative to the total number

(legend continued on next page)

confirm this, we used a suckling mouse colonization model[40] where the putative immunogenic taxa *A. muciniphila* was co-colonized with transgenic Δ*luxS E. coli* modified to express the AI-2 synthase enzymes from two human gut commensals, *Blautia obeum* (*B. obeum*) and *Streptococcus infantarius* (*S. infantarius*). These two commensals have genomes that contain *luxS* homologs of high sequence variability (33.3% amino acid identity, 33.6%–34.2% amino acid identify with *E. coli* LuxS, Figure 7A). The LuxS enzymes of these different species were codon optimized for expression in *E. coli* and stably integrated in single copy into the *E. coli* chromosome and driven by the same constitutive promoter. Constitutive expression of these different AI-2 synthase enzymes drove consistent AI-2 signal levels in the filtered supernatants (Figure 7B) when measured using the *Vibrio* BB170 bioassay,[41] suggesting that these enzymes produce AI-2 signals able to mediate cross-species communication.

Next, we designed an experiment to explore the association of AI-2 molecules with immune-modulatory species in a mouse colonization model (Figure 7C). Three groups of mice received the immunomodulatory strain *A. muciniphila*, a defined human gut commensal community, and either of (1) Δ*luxS E. coli*, *E. coli* expressing the LuxS enzyme of two different human gut commensals, (2) *B. obeum*, or (3) *S. infantarius*. The background model microbiome was based on healthy US human gut microbiomes and included *Bifidobacterium longum*, *Dorea formicigenerans*, and *Faecalibacterium prausnitzii* (Table S3). qPCR with universal 16S primers showed similar overall microbial load and no statistically significant differences between AI-2 signal types in either compartment (Figure 7D). We found using quantitative PCR with *A. muciniphila*-specific primers that the presence of *E. coli* expressing AI-2 from *B. obeum* was associated with a significantly increased abundance of *A. muciniphila* in the small intestine and large intestine compared to Δ*luxS E. coli* (Figure 7E). To confirm that these effects were mediated via changes in community behavior as opposed to direct effects on *A. muciniphila*, we co-colonized *A. muciniphila* with either Δ*luxS E. coli* or *E. coli* expressing the *B. obeum luxS*, without the human gut commensal community (Figure 7F). As expected, we observed no statistically significant signal-dependent colonization phenotype. Taken together, these data show that different LuxS signal products have varying community-dependent effects on the distribution of gut microbial communities, including those with immune-modulatory potential.

DISCUSSION

The role of the human microbiome on cancer progression and response to cancer therapy has recently been under intense investigation. Several studies have suggested that gut microbiota can modulate response to ICIs across several cancer types.[5,12,14,37,39] Gut microbes may impact antitumor immunity

via a number of mechanisms, including interaction of microbial components or products with antigen-presenting cells (APCs), which can help prime an adaptive immune response; induction of cytokine production by APCs or lymphocytes; and the effects of microbial metabolites.[42] Notably, there was only modest overlap in these checkpoint blockade responsive microbiome signatures across cohorts, though some phylogenetic commonalities exist among identified bacterial taxa in studies utilizing different checkpoint inhibitors.[10] At the species level, there is enormous inter-individual heterogeneity in gut microbiomes, which has hindered efforts to define a core microbiome shared between healthy individuals. Instead, the functional capacity of the microbiota, as defined by abundances of genes involved in metabolic pathways, has been advanced as a better metric for defining a core healthy microbiota.[43,44] Indeed, the KEGG pathways were more robustly represented across studies and showed good performance for predicting ICI response.

In this study, our pathway-centric method identified significantly enriched ABC transporters, QS, flagellar assembly, and biosynthesis of amino acid pathways between R and NR across 12 independent datasets with varying populations, cancer types, and multiple checkpoint inhibitors. ABC transporters are members of a transport system superfamily, with representatives in all phyla, from prokaryotes to humans.[45] In gut microenvironments, bacteria use ABC transporters to uptake a large variety of nutrients, biosynthetic precursors, trace metals, and vitamins and to export sterols, lipids, drugs, and a large variety of primary and secondary metabolites.[16] Previous research has emphasized the importance of bacterial ABC transporter in secreting products to trigger host immune response,[46] but it remains unclear whether bacterial ABC transporter affects the clinical response to ICI therapy.

Many bacteria use flagellar motors for motility.[47] Our results indicate that genes related to flagellar assembly, such as *fliC*, *flgL*, *flgK*, and *fliF*, were enriched in the R group, while *fliG*, *motA*, and *motB* were enriched in the NR group. Flagellin (FliC) is known to elicit strong immune response via Toll-like receptor 5 (TLR5).[48,49] In a phase 2 open-label clinical trial (NCT03637803), the synergistic effects of the oral probiotic MRx0518 (*Enterococcus gallinarum* in capsule form) in combination with pembrolizumab (anti-PD-1) were investigated in patients with NSCLC, RCC, bladder cancer, or melanoma. While the results have not yet been published, MRx0518 mainly relies on free flagellin to activate the TLR5 and nuclear factor κB signaling pathways and exert antitumor effects.[50] In addition, the role of other flagellar assembly genes in the context of ICI response warrants further investigation in future studies.

The gut microbiota has the capacity to facilitate the *de novo* synthesis of essential amino acids, which are implicated in amino acid homeostasis in the host.[51] Data from the Gopalakrishnan et al. study indicated that anabolic functions, notably amino acid biosynthesis, were more prominent in R, potentially

of genes annotated in that pathway. A higher rich score indicates greater enrichment. The bubble size indicates the number of genes. The color bar indicates the FDRs.

(B) Expression of the 13 quorum sensing genes that were significantly differentially abundant across 12 metagenome datasets. TPM, transcripts per million. ns, $p > 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. Wilcoxon rank-sum test.

(C) Expression levels of *fliC*, *fliF*, *flgL*, *flgK*, *motA*, *motB*, and *fliG*. TPM, transcripts per million. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. Wilcoxon rank-sum test.
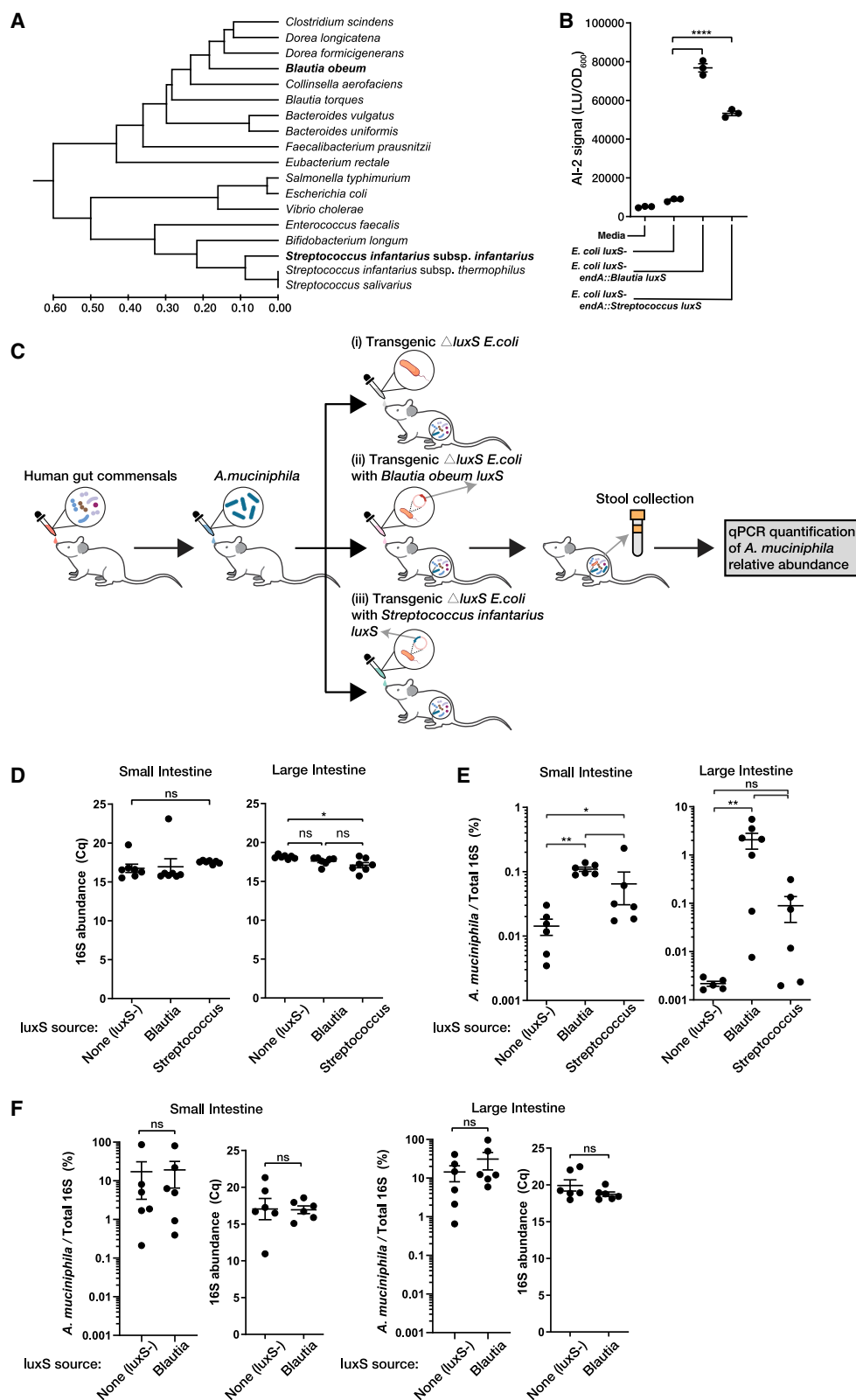
enhancing host immunity,[52] while catabolic functions were predominant in NR.[5] Consistent with these findings, our results indicated that amino acid biosynthesis pathways, including aromatic, arginine, lysine, methionine, glutamine, and proline, were enriched in the R group. In addition, we found that the carbohydrate metabolism and biosynthesis of the branched-chain amino acid and histidine pathways were enriched in the NR group. These findings highlight the potential role of amino acid biosynthesis in the context of ICI response.

QS process enables bacteria in a population to synchronize their actions and engage in group behaviors. The most commonly studied QS signals belong to one of the following three categories: acylated homoserine lactones (AI-1), used by gram-negative bacteria; peptide signals, used by gram-positive bacteria; and AI-2, used by both gram-negative and gram-positive bacteria.[53] Of these systems, AI-2 is the only one thought to be an inter-species signaling system, capable of mediating communication within and across many bacterial species. Despite the abundance of AI-2 signaling homologs in the gut microbiome, very few studies have attempted to define the commensal community and resulting host effects of AI-2 manipulation. Recent findings reveal that the AI-2 concentration is increased in tumors compared to the surrounding normal tissue in human colorectal cancer.[54] Moreover, AI-2 level was associated with increased CD3[+] T cell numbers and decreased CD4/CD8 ratio within colorectal cancer tissues[54]; infiltration of CD8[+] T cells within cancer tissues has been demonstrated to contribute to ICI response in a variety of cancers.[55–58] Our results indicate that the AI-2 synthesis gene *luxS* and peptide export genes, such as *agrB*, *secG*, and *secA*, were enriched in the R group, suggesting a potential inter-species communication within the gut microbiota of patients responding to ICI therapy. Conversely, genes related to the recognition and import of peptide signals, such as *livM*, *livK*, *livG*, and *oppD*, were enriched in the NR group, indicating a predominance of intra-species communication within the gut microbiota of NR group patients. The individual QS gene regulons associated with these signals will require additional investigation, including targeted culture isolation approaches to obtain pure bacterial populations that may be QS responsive, to elucidate.

Our studies show a potential role of *A. muciniphila* in cancer immunotherapy response. Of note, Mager et al.[59] identify inosine as a metabolite produced by *A. muciniphila* and *Bifidobacterium pseudolongum* that markedly increases the level of intratumoral and splenic CD4[+] and CD8[+] T cells and further enhances the efficacy of ICI therapy in a mouse model. In addition, the outer membrane protein Amuc of *A. muciniphila* strengthens the antitumor immune response in subcutaneous melanoma and colorectal tumor-bearing mice by activating the Toll-like receptor 2 signaling pathway.[60] However, no relevant literature regarding the effect of QS products on *A. muciniphila* has been published. Our results show that the provision of exogenous AI-2 from the common human gut commensal *B. obeum*, but not *Streptococcus* species, promotes *A. muciniphila* colonization in the small and large intestine of animals in a defined human gut microbiota context. This suggests that microbial AI-2 may be able to affect the gut microbiome to favor the expansion of immunomodulatory species and that separate AI-2 products of the AI-2 synthase enzyme LuxS can drive divergent community phenotypes. Therefore, these data support a mechanistic model where AI-2 action on the overall microbiota promotes accumulation of *A. muciniphila*, which can increase ICI therapy efficacy and antitumor immune response via inosine and Amuc production in the gut microenvironment. Furthermore, Zargar et al.[61] demonstrated that AI-2 from nonpathogenic bacteria affects host immune function by stimulating the expression of inflammatory cytokine interleukin-8, suggesting that AI-2 directly mediates interkingdom communication between the microbiome and the host in the gastrointestinal tract. In conclusion, our findings suggest that AI-2 is likely a vital regulatory molecule for gut microbiota response to cancer ICI therapy.

Limitations of the study

Although our data further advance the understanding and impact of gut bacterial genes or functional pathways on the response of tumor patients to ICI treatments, it is important to acknowledge several limitations of our animal modeling. As numerous gut microbes do not have readable accessible genetic manipulation systems available, genetic manipulation of AI-2 signaling pathways is unfortunately infeasible. Future studies are warranted to validate the impact of bacteria QS signals on ICI efficacy in animal models. Furthermore, the mechanisms of different QS AI-2 molecules and other bacterial AIs (including AI-1 and AIP signals) on the colonization of *A. muciniphila* remain to be investigated.

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Tian Xia (tianxia@hust.edu.cn).

Materials availability
Unique strains and reagents generated in this study are available from the lead contact with a completed Materials Transfer Agreement.

Figure 7. Commensal-encoded AI-2 quorum sensing signals promote *A. muciniphila* colonization
(A) Amino acid similarity of LuxS sequences from selected human gut-associated microbes.
(B) AI-2 signal levels of cell-free supernatants from indicated bacterial strains. ****, $p < 0.0001$. Mann-Whitney U test. Error bars represent mean ± SEM.
(C) Schematic of mouse experiments.
(D) Overall 16S levels measured by quantitative PCR. ns, $p > 0.05$; *, $p < 0.05$. Mann-Whitney U test. Error bars represent mean ± SEM.
(E) Colonization of *A. muciniphila* with defined model community and *E. coli* expressing AI-2 signals from indicated bacteria. ns, $p > 0.05$; *, $p < 0.05$; **, $p < 0.01$. Mann-Whitney U test. Error bars represent mean ± SEM.
(F) Colonization of *A. muciniphila* and *E. coli* expressing AI-2 signals from indicated bacteria, without the defined model community. ns, $p > 0.05$. Mann-Whitney U test. Error bars represent mean ± SEM.

Data and code availability

- This paper analyses existing, publicly available data. The accession numbers for the datasets are listed in the key resources table.
- All original code has been deposited at https://github.com/Xunhui-Cai/RandomForestClassifier and is publicly available as of the date of publication.
- All other requests: any additional information required to reanalyze the data reported will be shared by the lead contact upon request.

AUTHOR CONTRIBUTIONS

X.C., Y.L., F.J., M.W., J.Z., and T.X. conceived the study. X.C., L.C., Y.L., and F.J. generated the data and performed the analyses. A.H., J.Y.C., K.S., and S.G. performed the validation experiments. D.Y., H.Y., C.C., M.W., G.W., and T.X. supervised the research. T.X., C.C., L.C., P.A.F., B.S., D.G., J.S., Y.C., and J.Z. interpreted the clinical correlation. X.C., A.H., T.X., and J.Z. wrote the manuscript, with contributions from all authors. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used ChatGPT 4.0 in order to correct grammar and spelling. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Animals
  - Bacterial strains
  - Public datasets inclusion and metadata acquisition
- METHOD DETAILS
  - Metagenomic datasets
  - Workflow of obtaining non-redundant metagenomic gene catalogue
  - Analysis of genes with abundance difference
  - KEGG pathway enrichment
  - Integrated analysis for identification of differentially abundant genes
  - Machine learning methods
  - Shotgun metatranscriptome data analysis
  - Phylogenetic analysis
  - Bacterial growth conditions
  - Generation of AI-2 expression strains
  - BB170 AI-2 bioassay
  - Infant mouse colonization
  - Quantitative PCR analysis of *A. muciniphila* abundance
- QUANTIFICATIONS AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2025.112162.

REFERENCES

1. Oiseth, S.J., and Aziz, M.S. (2017). Cancer immunotherapy: a brief review of the history, possibilities, and challenges ahead. J. Cancer Metastasis Treat. *3*, 250–261. https://doi.org/10.20517/2394-4722.2017.41.

2. Disis, M.L. (2014). Mechanism of action of immunotherapy. Semin. Oncol. *41*, S3–S13. https://doi.org/10.1053/j.seminoncol.2014.09.004.

3. Halpert, M.M., Konduri, V., Liang, D., Chen, Y., Wing, J.B., Paust, S., Levitt, J.M., and Decker, W.K. (2016). Dendritic cell-secreted cytotoxic T-lymphocyte-associated protein-4 regulates the T-cell response by downmodulating bystander surface B7. Stem Cell. Dev. *25*, 774–787. https://doi.org/10.1089/scd.2016.0009.

4. van der Kooij, M.K., Joosse, A., Speetjens, F.M., Hospers, G.A.P., Bisschop, C., de Groot, J.W.B., Koornstra, R., Blank, C.U., and Kapiteijn, E. (2017). Anti-PD1 treatment in metastatic uveal melanoma in the Netherlands. Acta Oncol. *56*, 101–103. https://doi.org/10.1080/0284186X.2016.1260773.

5. Gopalakrishnan, V., Spencer, C.N., Nezi, L., Reuben, A., Andrews, M.C., Karpinets, T.V., Prieto, P.A., Vicente, D., Hoffman, K., Wei, S.C., et al. (2018). Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. Science *359*, 97–103. https://doi.org/10.1126/science.aan4236.

6. Matson, V., Fessler, J., Bao, R., Chongsuwat, T., Zha, Y., Alegre, M.-L., Luke, J.J., and Gajewski, T.F. (2018). The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. Science *359*, 104–108. https://doi.org/10.1126/science.aao3290.

7. Routy, B., Le Chatelier, E., Derosa, L., Duong, C.P.M., Alou, M.T., Daillère, R., Fluckiger, A., Messaoudene, M., Rauber, C., Roberti, M.P., et al. (2018). Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. Science *359*, 91–97. https://doi.org/10.1126/science.aan3706.

8. Frankel, A.E., Coughlin, L.A., Kim, J., Froehlich, T.W., Xie, Y., Frenkel, E.P., and Koh, A.Y. (2017). Metagenomic shotgun sequencing and unbiased metabolomic profiling identify specific human gut microbiota and metabolites associated with immune checkpoint therapy efficacy in melanoma patients. Neoplasia *19*, 848–855. https://doi.org/10.1016/j.neo.2017.08.004.

9. Gharaibeh, R.Z., and Jobin, C. (2019). Microbiota and cancer immunotherapy: in search of microbial signals. Gut *68*, 385–388. https://doi.org/10.1136/gutjnl-2018-317220.

10. Lee, K.A., Thomas, A.M., Bolte, L.A., Björk, J.R., de Ruijter, L.K., Armanini, F., Asnicar, F., Blanco-Miguez, A., Board, R., Calbet-Llopart, N., et al. (2022). Cross-cohort gut microbiome associations with immune checkpoint inhibitor response in advanced melanoma. Nat. Med. *28*, 535–544. https://doi.org/10.1038/s41591-022-01695-5.

11. Thursby, E., and Juge, N. (2017). Introduction to the human gut microbiota. Biochem. J. *474*, 1823–1836. https://doi.org/10.1042/BCJ20160510.

12. Derosa, L., Routy, B., Thomas, A.M., Iebba, V., Zalcman, G., Friard, S., Mazieres, J., Audigier-Valette, C., Moro-Sibilot, D., Goldwasser, F., et al. (2022). Intestinal *Akkermansia muciniphila* predicts clinical response to

PD-1 blockade in patients with advanced non-small-cell lung cancer. Nat. Med. *28*, 315–324. https://doi.org/10.1038/s41591-021-01655-5.

13. McCulloch, J.A., Davar, D., Rodrigues, R.R., Badger, J.H., Fang, J.R., Cole, A.M., Balaji, A.K., Vetizou, M., Prescott, S.M., Fernandes, M.R., et al. (2022). Intestinal microbiota signatures of clinical response and immune-related adverse events in melanoma patients treated with anti-PD-1. Nat. Med. *28*, 545–556. https://doi.org/10.1038/s41591-022-01698-2.

14. Peng, Z., Cheng, S., Kou, Y., Wang, Z., Jin, R., Hu, H., Zhang, X., Gong, J.-F., Li, J., Lu, M., et al. (2020). The gut microbiome is associated with clinical response to anti–PD-1/PD-L1 immunotherapy in gastrointestinal cancer. Cancer Immunol. Res. *8*, 1251–1261. https://doi.org/10.1158/2326-6066.CIR-19-1014.

15. Gunjur, A., Shao, Y., Rozday, T., Klein, O., Mu, A., Haak, B.W., Markman, B., Kee, D., Carlino, M.S., Underhill, C., et al. (2024). A gut microbial signature for combination immune checkpoint blockade across cancer types. Nat. Med. *30*, 797–809. https://doi.org/10.1038/s41591-024-02823-z.

16. Rees, D.C., Johnson, E., and Lewinson, O. (2009). ABC transporters: the power to change. Nat. Rev. Mol. Cell Biol. *10*, 218–227. https://doi.org/10.1038/nrm2646.

17. Akahoshi, D.T., and Bevins, C.L. (2022). Flagella at the host-microbe interface: Key functions intersect with redundant responses. Front. Immunol. *13*, 828758. https://doi.org/10.3389/fimmu.2022.828758.

18. Wu, L., Tang, Z., Chen, H., Ren, Z., Ding, Q., Liang, K., and Sun, Z. (2021). Mutual interaction between gut microbiota and protein/amino acid metabolism for host mucosal immunity and health. Anim. Nutr. *7*, 11–16. https://doi.org/10.1016/j.aninu.2020.11.003.

19. Galloway, W.R.J.D., Hodgkinson, J.T., Bowden, S.D., Welch, M., and Spring, D.R. (2011). Quorum sensing in Gram-negative bacteria: small-molecule modulation of AHL and AI-2 quorum sensing pathways. Chem. Rev. *111*, 28–67. https://doi.org/10.1021/cr100109t.

20. Li, Q., Ren, Y., and Fu, X. (2019). Inter-kingdom signaling between gut microbiota and their host. Cell. Mol. Life Sci. *76*, 2383–2389. https://doi.org/10.1007/s00018-019-03076-7.

21. Slamti, L., and Lereclus, D. (2019). The oligopeptide ABC-importers are essential communication channels in Gram-positive bacteria. Res. Microbiol. *170*, 338–344. https://doi.org/10.1016/j.resmic.2019.07.004.

22. Taga, M.E., Miller, S.T., and Bassler, B.L. (2003). Lsr-mediated transport and processing of AI-2 in *Salmonella typhimurium*. Mol. Microbiol. *50*, 1411–1427. https://doi.org/10.1046/j.1365-2958.2003.03781.x.

23. Amos, M.R., Sanchez-Contreras, M., Jackson, R.W., Muñoz-Berbel, X., Ciche, T.A., Yang, G., Cooper, R.M., and Waterfield, N.R. (2011). Influence of the *Photorhabdus luminescens* phosphomannose isomerase gene, *manA*, on mannose utilization, exopolysaccharide structure, and biofilm formation. Appl. Environ. Microbiol. *77*, 776–785. https://doi.org/10.1128/aem.02326-10.

24. Hsiao, A., Ahmed, A.M.S., Subramanian, S., Griffin, N.W., Drewry, L.L., Petri, W.A., Haque, R., Ahmed, T., and Gordon, J.I. (2014). Members of the human gut microbiota involved in recovery from *Vibrio cholerae* infection. Nature *515*, 423–426. https://doi.org/10.1038/nature13738.

25. Ji, Y.C., Sun, Q., Fu, C.Y., She, X., Liu, X.C., He, Y., Ai, Q., Li, L.Q., and Wang, Z.L. (2021). Exogenous autoinducer-2 rescues intestinal dysbiosis and intestinal inflammation in a neonatal mouse necrotizing enterocolitis model. Front. Cell. Infect. Microbiol. *11*, 694395. https://doi.org/10.3389/fcimb.2021.694395.

26. Piggot, P.J., and Hilbert, D.W. (2004). Sporulation of *Bacillus subtilis*. Curr. Opin. Microbiol. *7*, 579–586. https://doi.org/10.1016/j.mib.2004.10.001.

27. Hughes, D.T., Clarke, M.B., Yamamoto, K., Rasko, D.A., and Sperandio, V. (2009). The QseC adrenergic signaling cascade in enterohemorrhagic *E. coli* (EHEC). PLoS Pathog. *5*, e1000553. https://doi.org/10.1371/journal.ppat.1000553.

28. Xu, K., Lin, L., Shen, D., Chou, S.-H., and Qian, G. (2021). Clp is a "busy" transcription factor in the bacterial warrior, *Lysobacter enzymogenes*.

Comput. Struct. Biotechnol. J. *19*, 3564–3572. https://doi.org/10.1016/j.csbj.2021.06.020.

29. Macnab, R.M. (2003). How bacteria assemble flagella. Annu. Rev. Microbiol. *57*, 77–100. https://doi.org/10.1146/annurev.micro.57.030502.090832.

30. Hajam, I.A., Dar, P.A., Shahnawaz, I., Jaume, J.C., and Lee, J.H. (2017). Bacterial flagellin—a potent immunomodulatory agent. Exp. Mol. Med. *49*, e373. https://doi.org/10.1038/emm.2017.172.

31. Komatsu, H., Hayashi, F., Sasa, M., Shikata, K., Yamaguchi, S., Namba, K., and Oosawa, K. (2016). Genetic analysis of revertants isolated from the rod-fragile *fliF* mutant of *Salmonella*. Biophys. Physicobiol. *13*, 13–25. https://doi.org/10.2142/biophysico.13.0_13.

32. Morimoto, Y.V., and Minamino, T. (2014). Structure and function of the bi-directional bacterial flagellar motor. Biomolecules *4*, 217–234. https://doi.org/10.3390/biom4010217.

33. Luu, M., Riester, Z., Baldrich, A., Reichardt, N., Yuille, S., Busetti, A., Klein, M., Wempe, A., Leister, H., Raifer, H., et al. (2021). Microbial short-chain fatty acids modulate CD8$^+$ T cell responses and improve adoptive immunotherapy for cancer. Nat. Commun. *12*, 4077. https://doi.org/10.1038/s41467-021-24331-1.

34. Bender, M.J., McPherson, A.C., Phelps, C.M., Pandey, S.P., Laughlin, C.R., Shapira, J.H., Medina Sanchez, L., Rana, M., Richie, T.G., Mims, T.S., et al. (2023). Dietary tryptophan metabolite released by intratumoral *Lactobacillus reuteri* facilitates immune checkpoint inhibitor treatment. Cell *186*, 1846–1862. https://doi.org/10.1016/j.cell.2023.03.011.

35. Peters, B.A., Wilson, M., Moran, U., Pavlick, A., Izsak, A., Wechter, T., Weber, J.S., Osman, I., and Ahn, J. (2019). Relating the gut metagenome and metatranscriptome to immunotherapy responses in melanoma patients. Genome Med. *11*, 61. https://doi.org/10.1186/s13073-019-0672-4.

36. Thompson, J.A., Oliveira, R.A., Djukovic, A., Ubeda, C., and Xavier, K.B. (2015). Manipulation of the quorum sensing signal AI-2 affects the antibiotic-treated gut microbiota. Cell Rep. *10*, 1861–1871. https://doi.org/10.1016/j.celrep.2015.02.049.

37. Derosa, L., Routy, B., Fidelle, M., Iebba, V., Alla, L., Pasolli, E., Segata, N., Desnoyer, A., Pietrantonio, F., Ferrere, G., et al. (2020). Gut bacteria composition drives primary resistance to cancer immunotherapy in renal cell carcinoma patients. Eur. Urol. *78*, 195–206. https://doi.org/10.1016/j.eururo.2020.04.044.

38. Salgia, N.J., Bergerot, P.G., Maia, M.C., Dizman, N., Hsu, J., Gillece, J.D., Folkerts, M., Reining, L., Trent, J., Highlander, S.K., and Pal, S.K. (2020). Stool microbiome profiling of patients with metastatic renal cell carcinoma receiving anti–PD-1 immune checkpoint inhibitors. Eur. Urol. *78*, 498–502. https://doi.org/10.1016/j.eururo.2020.07.011.

39. Zheng, Y., Wang, T., Tu, X., Huang, Y., Zhang, H., Tan, D., Jiang, W., Cai, S., Zhao, P., Song, R., et al. (2019). Gut microbiome affects the response to anti-PD-1 immunotherapy in patients with hepatocellular carcinoma. J. Immunother. Cancer *7*, 193–197. https://doi.org/10.1186/s40425-019-0650-9.

40. Alavi, S., and Hsiao, A. (2020). Protocol for microbiome transplantation in suckling mice during *vibrio cholerae* infection to study commensal-pathogen interactions. STAR Protoc. *1*, 100200. https://doi.org/10.1016/j.xpro.2020.100200.

41. Bassler, B.L., Greenberg, E.P., and Stevens, A.M. (1997). Cross-species induction of luminescence in the quorum-sensing bacterium *Vibrio harveyi*. J. Bacteriol. *179*, 4043–4045. https://doi.org/10.1128/jb.179.12.4043-4045.1997.

42. Helmink, B.A., Khan, M.A.W., Hermann, A., Gopalakrishnan, V., and Wargo, J.A. (2019). The microbiome, cancer, and cancer therapy. Nat. Med. *25*, 377–388. https://doi.org/10.1038/s41591-019-0377-7.

43. Integrative HMP iHMP Research Network Consortium (2014). The Integrative Human Microbiome Project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease. Cell Host Microbe *16*, 276–289. https://doi.org/10.1016/j.chom.2014.08.014.

44. Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J.H., Chinwalla, A.T., Creasy, H.H., Earl, A.M., FitzGerald, M.G., and Fulton, R.S. (2012). Structure, function and diversity of the healthy human microbiome. Nature 486, 207–214. https://doi.org/10.1038/nature11234.

45. Polgar, O., and Bates, S.E. (2005). ABC transporters in the balance: is there a role in multidrug resistance? Biochem. Soc. Trans. 33, 241–245. https://doi.org/10.1042/BST0330241.

46. Dasgupta, A., Sureka, K., Mitra, D., Saha, B., Sanyal, S., Das, A.K., Chakrabarti, P., Jackson, M., Gicquel, B., Kundu, M., and Basu, J. (2010). An oligopeptide transporter of *Mycobacterium tuberculosis* regulates cytokine release and apoptosis of infected macrophages. PLoS One 5, e12225. https://doi.org/10.1371/journal.pone.0012225.

47. Akahoshi, D.T., and Bevins, C.L. (2022). Flagella at the host-microbe interface: key functions intersect with redundant responses. Front. Immunol. 13, 828758. https://doi.org/10.3389/fimmu.2022.828758.

48. Griffin, M.E., and Hang, H.C. (2022). Microbial mechanisms to improve immune checkpoint blockade responsiveness. Neoplasia 31, 100818. https://doi.org/10.1016/j.neo.2022.100818.

49. Khim, K., Bang, Y.J., Puth, S., Choi, Y., Lee, Y.S., Jeong, K., Lee, S.E., and Rhee, J.H. (2021). Deimmunization of flagellin for repeated administration as a vaccine adjuvant. npj Vaccines 6, 116. https://doi.org/10.1038/s41541-021-00379-4.

50. Lauté-Caly, D.L., Raftis, E.J., Cowie, P., Hennessy, E., Holt, A., Panzica, D.A., Sparre, C., Minter, B., Stroobach, E., and Mulder, I.E. (2019). The flagellin of candidate live biotherapeutic *Enterococcus gallinarum* MRx0518 is a potent immunostimulant. Sci. Rep. 9, 801. https://doi.org/10.1038/s41598-018-36926-8.

51. Collins, S.M., Surette, M., and Bercik, P. (2012). The interplay between the intestinal microbiota and the brain. Nat. Rev. Microbiol. 10, 735–742. https://doi.org/10.1038/nrmicro2876.

52. Blacher, E., Levy, M., Tatirovsky, E., and Elinav, E. (2017). Microbiome-modulated metabolites at the interface of host immunity. J. Immunol. 198, 572–580. https://doi.org/10.4049/jimmunol.1601247.

53. Jimenez, A.G., and Sperandio, V. (2019). Quorum sensing and the gut microbiome. In Quorum Sensing, G. Tommonaro, ed. (Elsevier), pp. 151–169. https://doi.org/10.1016/B978-0-12-814905-8.00006-X.

54. Li, Q., Peng, W., Wu, J., Wang, X., Ren, Y., Li, H., Peng, Y., Tang, X., and Fu, X. (2019). Autoinducer-2 of gut microbiota, a potential novel marker for human colorectal cancer, is associated with the activation of TNFSF9 signaling in macrophages. OncoImmunology 8, e1626192. https://doi.org/10.1080/2162402X.2019.1626192.

55. Banchereau, R., Chitre, A.S., Scherl, A., Wu, T.D., Patil, N.S., de Almeida, P., Kadel III, E.E., Madireddi, S., Au-Yeung, A., Takahashi, C., et al. (2021). Intratumoral CD103+ CD8+ T cells predict response to PD-L1 blockade. J. Immunother. Cancer 9, e002231. https://doi.org/10.1136/jitc-2020-002231.

56. Bassez, A., Vos, H., Van Dyck, L., Floris, G., Arijs, I., Desmedt, C., Boeckx, B., Vanden Bempt, M., Nevelsteen, I., Lambein, K., et al. (2021). A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. Nat. Med. 27, 820–832. https://doi.org/10.1038/s41591-021-01323-8.

57. Hsu, C.-L., Ou, D.-L., Bai, L.-Y., Chen, C.-W., Lin, L., Huang, S.-F., Cheng, A.-L., Jeng, Y.-M., and Hsu, C. (2021). Exploring markers of exhausted CD8 T cells to predict response to immune checkpoint inhibitor therapy for hepatocellular carcinoma. Liver Cancer 10, 346–359. https://doi.org/10.1159/000515305.

58. Li, C., Phoon, Y.P., Karlinsey, K., Tian, Y.F., Thapaliya, S., Thongkum, A., Qu, L., Matz, A.J., Cameron, M., Cameron, C., et al. (2022). A high OXPHOS CD8 T cell subset is predictive of immunotherapy resistance in melanoma patients. J. Exp. Med. 219, e20202084. https://doi.org/10.1084/jem.20202084.

59. Mager, L.F., Burkhard, R., Pett, N., Cooke, N.C.A., Brown, K., Ramay, H., Paik, S., Stagg, J., Groves, R.A., Gallo, M., et al. (2020). Microbiome-derived inosine modulates response to checkpoint inhibitor immunotherapy. Science 369, 1481–1489. https://doi.org/10.1126/science.abc3421.

60. Shi, L., Sheng, J., Chen, G., Zhu, P., Shi, C., Li, B., Park, C., Wang, J., Zhang, B., Liu, Z., and Yang, X. (2020). Combining IL-2-based immunotherapy with commensal probiotics produces enhanced antitumor immune response and tumor clearance. J. Immunother. Cancer 8, e000973. https://doi.org/10.1136/jitc-2020-000973.

61. Zargar, A., Quan, D.N., Carter, K.K., Guo, M., Sintim, H.O., Payne, G.F., and Bentley, W.E. (2015). Bacterial secretions of nonpathogenic *Escherichia coli* elicit inflammatory pathways: a closer investigation of interkingdom signaling. mBio 6, e00025. https://doi.org/10.1128/mBio.00025-15.

62. Uritskiy, G.V., DiRuggiero, J., and Taylor, J. (2018). MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome 6, 158. https://doi.org/10.1186/s40168-018-0541-1.

63. Zhu, W., Lomsadze, A., and Borodovsky, M. (2010). *Ab initio* gene identification in metagenomic sequences. Nucleic Acids Res. 38, e132. https://doi.org/10.1093/nar/gkq275.

64. Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P.A. (2017). metaSPAdes: a new versatile metagenomic assembler. Genome Res. 27, 824–834. https://doi.org/10.1101/gr.213959.116.

65. Shen, W., Le, S., Li, Y., and Hu, F. (2016). SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PLoS One 11, e0163962. https://doi.org/10.1371/journal.pone.0163962.

66. Mikheenko, A., Saveliev, V., and Gurevich, A. (2016). MetaQUAST: evaluation of metagenome assemblies. Bioinformatics 32, 1088–1090. https://doi.org/10.1093/bioinformatics/btv697.

67. Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28, 3150–3152. https://doi.org/10.1093/bioinformatics/bts565.

68. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 14, 417–419. https://doi.org/10.1038/nmeth.4197.

69. Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering, C., and Bork, P. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. Mol. Biol. Evol. 34, 2115–2122. https://doi.org/10.1093/molbev/msx148.

70. Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., Rattei, T., Mende, D.R., Sunagawa, S., Kuhn, M., et al. (2016). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res. 44, D286–D293. https://doi.org/10.1093/nar/gkv1248.

71. Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60. https://doi.org/10.1038/nmeth.3176.

72. Ma, S., Shungin, D., Mallick, H., Schirmer, M., Nguyen, L.H., Kolde, R., Franzosa, E., Vlamakis, H., Xavier, R., and Huttenhower, C. (2022). Population structure discovery in meta-analyzed microbial communities and inflammatory bowel disease using MMUPHin. Genome Biol. 23, 208. https://doi.org/10.1186/s13059-022-02753-4.

73. Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35, 1547–1549. https://doi.org/10.1093/molbev/msy096.

74. Bassler, B.L., Wright, M., and Silverman, M.R. (1994). Multiple signalling systems controlling expression of luminescence in *Vibrio harveyi*: sequence and function of genes encoding a second sensory pathway. Mol. Microbiol. 13, 273–286. https://doi.org/10.1111/j.1365-2958.1994.tb00422.x.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Vibrio harveyi AI-2 bioassay strain	Bassler et al.[74]	BB170
luxS- E. coli strain BW30045	*E. coli* Genetic Resources at Yale. CGSC, The Coli Genetic Stock Center	CGSC#8227
E. coli expressing *B. obeum luxS*	This paper	BW30045_RO_AI-2
Streptococcus infantarius subsp. *infantarius*	ATCC	ATCC BAA-102
Bifidobacterium longum subsp. *longum*	DSMZ	DSM 20219
Collinsella aerofaciens	ATCC	ATCC 25986
Dorea formicigenerans	ATCC	ATCC 27755
Eubacterium rectale	ATCC	ATCC 33656
Faecalibacterium prausnitzii	DSMZ	DSM 17677
Bacteroides vulgatus	ATCC	ATCC 8482
Blautia obeum	ATCC	ATCC 29174
Clostridium scindens	ATCC	ATCC 35704
Akkermanisa muciniphila	ATCC	ATCC BAA-835
Deposited data		
Metagenomic data	Gopalakrishnan et al.[5]; Matson et al.[6]; McCulloch et al.[13]; Lee et al.[10]; Peng et al.[14]; Derosa et al.[12]; Routy et al.[7]; Gunjur et al.[15]; Peters et al.[35]	European Nucleotide Archive (ENA): PRJEB22893; PRJNA399742; PRJNA762360; PRJEB43119; PRJNA615114; PRJNA751792; PRJEB22863; PRJEB49516; PRJNA541981
Metatranscriptome sequencing data	Peters et al.[35]	ENA: PRJNA541981
Source code	This study	https://github.com/Xunhui-Cai/RandomForestClassifier
Experimental models: Organisms/strains		
Mouse: CD-1	Charles River Laboratories	N/A
Oligonucleotides		
E.coli-F primer: CCAAAACAGCTTTCGCTACGTTGCTGGCTCGTTTTAACACGGAGTAAGTGTTAGAAAAATTCATCCAGCA	This study	N/A
E.coli-R primer: GGTTGTACGCGTGGGGTAGGGGTTAACAAAAGAATCCCGCTAGTGTAGGCGGGCAGTGAAAGGAAGGCC	This study	N/A
A. muciniphila-F primer: CAGCACGTGAAGGTGGGGA	This study	N/A
A. muciniphila-R primer: TTACCGCGGCTGCTGGCAC	This study	N/A
16S F PCR Primer forward: 5′-CTCCTACGGGAGGCAGCAG-3′	IDT	N/A
16S R PCR Primer reverse: 5′-TTACCGCGGCTGCTGGCAC-3′	IDT	N/A
Recombinant DNA		
B. obeum codon optimized LuxS placed downstream of the $P_{Ltet-O-1}$ constitutive promoter sequence derived from the plasmid vector pZE21 vector (pMK_*B. obeum_luxS*)	This paper	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Python	Python Software Foundation	https://www.python.org/
R v3.6	R Development Core Team	https://cran.r-project.org/
metaSPAdes v3.13.0	Nurk et al.[64]	https://github.com/ablab/spades
MetaWRAP v1.0	Uritskiy et al.[62]	https://github.com/bxlab/metaWRAP
eggNOG-mapper	Huerta-Cepas et al.[69]	http://eggnog-mapper.embl.de
DIAMOND	Buchfink et al.[71]	http://ab.inf.uni-tuebingen.de/software/diamond
MetaGeneMark v.3.38	Zhu et al.[63]	http://exon.gatech.edu/GeneMark/license_download.cgi
SeqKit v.0.10.1	Shen et al.[65]	https://bioinf.shenwei.me/seqkit/
MetaQUAST v5.2.0	Mikheenko et al.[66]	http://quast.sourceforge.net/metaquast
CD-HIT v4.7	Fu et al.[67]	https://github.com/weizhongli/cdhit
Salmon	Patro et al.[68]	https://github.com/COMBINE-lab/Salmon
MEGA X	Kumar et al.[73]	https://www.megasoftware.net
MMUPHin	Ma et al.[72]	https://www.bioconductor.org/packages/release/bioc/html/MMUPHin.html

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Animals

4-day-old suckling CD-1 mice were purchased from Charles River Laboratories. No differentiation was observed between sexes and animal data were pooled by sex where applicable. Animals in the study were treated and housed under specific-pathogen-free or germfree conditions. All animal protocols were approved by University of California, Riverside's Institutional Animal Care and Use Committee (protocol number: 20210031).

Bacterial strains

A list of bacterial strains used in this work can be found in the key resources table and Table S3. Bacterial culture methods are detailed in the method details section.

Public datasets inclusion and metadata acquisition

The raw metagenomic data were obtained from the European Nucleotide Archive (ENA) using the ENA identifier or Sequence Read Archive (SRA) using SRA identifier. Metadata was retrieved from the original publication. A total of 821 samples were included in the analysis after discarding samples collected following the start of any treatments. The response phenotype was evaluated based on the Response Evaluation Criteria in Solid Tumors 1.1 (RECIST 1.1) as described in the original studies. Patients' characteristics are described in Table S1.

METHOD DETAILS

Metagenomic datasets

This study uses public metagenomic datasets. These cohorts were obtained from Gopalakrishnan et al.[5] (PRJEB22893); Matson et al.[6] (PRJNA399742); Routy et al.[7] (PRJEB22863); Lee et al.[10] (PRJEB43119); Derosa et al.[12] (PRJNA751792); McCulloch et al.[13] (PRJNA762360); Peng et al.[14] (PRJNA615114); Gunjur et al.[15] (PRJEB49516) and Peters et al.[35] (PRJNA541981).

Workflow of obtaining non-redundant metagenomic gene catalogue

Raw metagenomic DNA sequencing datasets for clinical trial samples were downloaded from NCBI (Table S1) and then filtered and assembled using metaWRAP software.[62] Next, gene prediction was performed on the assembled contigs using MetaGeneMark.[63] The predicted genes from all samples were clustered to generate a non-redundant gene catalogue. Detailed descriptions of quality control, contig assembly and gene prediction are provided below:

Quality control

The MetaWRAP-Read_qc module was used to trim the raw sequencing reads for low-quality bases and adapter sequences, and removed human contamination using default parameters.[62]

Assembly

High-quality reads of each sample were assembled into longer contigs using metaSPAdes.[64] Assembly parameter K-mer size was set to 21, 33 and 55. K-mers whose frequency is less than or equal to 1 were dropped to reduce errors and improve assembly accuracy. Contigs shorter than 1,000 bp were removed using SeqKit v0.10.1.[65] MetaQUAST v5.2.0 was then run on the remaining contigs to compute genome statistics of the metagenomics assemblies,[66] including the number and maximal length of contigs, total assembly length, N50 values, L50 values, and the number of misassembled contigs (Table S4). Finally, misassembled contigs, as identified by MetaQUAST, were discarded to ensure high-quality of the assembly.

Gene prediction

The assembly results from each sample were integrated into a fasta file, and genes were predicted using MetaGeneMark v3.38.[63] The predicted gene sequences and their corresponding amino acid sequences were then obtained.

Redundant gene removal

Redundant protein sequences were clustered using CD-HIT v4.7.[67] Any two genes with amino acid sequence identity greater than 98% and covering more than 90% were clustered together. The representative sequence of each cluster was extracted into a de-duplicated protein sequence file. A de-duplicated gene sequence file was then extracted based on the de-duplicated protein sequence file by gene-protein mapping. Furthermore, cluster representative genes shorter than 100 bp were discarded to generate the final non-redundant gene catalogue.

Analysis of genes with abundance difference
Gene abundance estimation

Salmon was used to count quality-controlled reads against the non-redundant gene catalogue.[68] Gene abundance was normalized using the transcripts per million (TPM) method for each sample. The gene abundance data for each sample within a cohort were subsequently combined into a gene abundance matrix.

Differentially abundant genes assessment

A list of genes with significantly differential abundance between R and NR was obtained by performing Wilcoxon rank-sum test with $P < 0.05$ and $|\log_2 (\text{fold change})| \geq 1$.

KEGG pathway enrichment

Non-redundant genes were annotated with KEGG orthology (KO) clusters using eggNOG-mapper v1[69] based on EggNOG orthology data.[70] Sequence searches were performed using DIAMOND with default settings.[71] The annotated differentially abundant genes for each dataset are provided in Table S5. Of note, ABC transporters, QS, flagellar assembly, and biosynthesis of amino acid KOs were obtained from KEGG BRITE 'ko00001.keg' (list of KO genes with pathway maps) under the 'Quorum sensing', 'ABC transporters', 'Flagellar assembly', and 'Biosynthesis of amino acids' categories. Based on the EggNOG annotation file, the following four numbers were tallied: M, N, i, and n, where M is the total number of genes annotated in a pathway term, N is the total number of genes with KO annotation, i is the number of genes annotated with a pathway term in the differentially abundant geneset, and n is the total number of genes in the differential abundance geneset. Notably, if a gene had multiple pathway assignments, each assignment was considered as a separate gene. Pathway enrichment analysis was further performed by Fisher's exact test with the four numbers as follows:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}$$

Finally, P values were corrected for multiple hypothesis testing using the Benjamini-Hochberg procedure in R, and a false discovery rate < 0.05 was defined as the significance threshold. Scatter plots of enriched KEGG pathway were generated using the R package ggplot2.

Integrated analysis for identification of differentially abundant genes

Upon integrating the different datasets, batch effects were first removed using MMUPHin analysis.[72] Two-sided Wilcoxon rank-sum tests were used for determining differentially abundant genes. The fold change in gene abundance was calculated as the mean difference between R and NR groups. Genes with $P < 0.05$ and $|\log2 (\text{fold change})| \geq 1$ were considered significantly differentially abundant across all 12 datasets.

Machine learning methods

We divided the 12 ICI datasets into a discovery cohort (10 datasets, 755 samples) and a validation cohort (2 datasets, including 39 samples from Matson et al. and 25 samples from Gopalakrishnan et al.). For the discovery cohort, six popular machine learning algorithms were applied with the scikit-learn v.1.1.3 package in Python (https://www.python.org/) using 10-fold cross-validation for discriminating responders and non-responders. Features used for model construction included all differentially abundant genes (55,292 features), genes from 11 enriched pathways (5,761 features), and genes from 4 enriched pathways (3,522 features). Notably, the optimal 55,292-feature RF model was constructed with n_estimators set to 293 and max_depth set to 35. The optimal

5,761-feature RF model was constructed with n_estimators set to 265, max_depth set to 49, and max_features set to 69. The optimal 3,522-feature RF model was constructed with n_estimators set to 320, min_samples_split set to 24, and max_features set to 2. For the validation cohort, the optimal RF classification models trained on the discovery cohort were applied to the Matson et al. and Gopalakrishnan et al. datasets for external evaluation. Finally, accuracy, precision, recall, and F1-score were used to evaluate the generalizability of the models across different contexts.

Shotgun metatranscriptome data analysis

The analysis process of the shotgun metatranscriptome samples is generally similar to that of the shotgun metagenome samples. First, we performed quality control on the fecal RNA sequencing reads using MetaWRAP.[62] Then, we obtained a non-redundant gene sequence set by analyzing the shotgun metagenome samples through our standard pipeline. Furthermore, we aligned the quality-controlled RNA sequencing reads to the non-redundant gene set to generate a gene expression matrix. Finally, genes that were differentially expressed between R and NR were identified using Wilcoxon rank-sum test with $P < 0.05$ and $|log2 (fold change)| \geq 1$. Box plots were generated using the R package ggplot2.

Phylogenetic analysis

LuxS sequences were aligned, and a phylogenetic tree was generated using MEGA version X.[73] Sequences were aligned and analyzed by ClustalW using the default setting. The phylogenetic tree was generated using the aligned result by UPGMA Tree using the default setting.

Bacterial growth conditions

Strains incorporated into the defined human gut community are listed in Table S3. All strains were grown in LYH-BHI liquid media (BHI supplemented to 5 g/L yeast extract, 5 mg/L hemin, 1 mg/mL cellobiose, 1 mg/mL maltose, and 0.5 mg/mL cysteine-HCl). In order to optimize the growth of A. muciniphila, LYH-BHI media was supplemented with 3% of porcine mucin. All cultures were placed inside a static incubator at 37°C for 48 hours under anaerobic conditions (5% H_2, 20% CO_2, balance N_2), and then subcultured at 1:100 in fresh LYH-BHI media for 48 hours prior to use. All E. coli strains containing luxS expression constructs were grown aerobically in Luria broth medium (LB media) with agitation at 37°C overnight and subcultured at 1:100 in fresh LB media overnight prior to use.

Generation of AI-2 expression strains

E. coli strains overexpressing the luxS enzymes of S. infantarius or B. obeum were constructed using an E. coli luxS mutant strain (BW30045). The B. obeum luxS coding region (from genome position 33,305-33,784) and the S. infantarius luxS coding region (from genome position 414,476-414,994) were codon-optimized for E. coli, placed downstream of the $P_{Ltet-O-1}$ constitutive promoter sequence derived from the plasmid vector pZE21, and the resulting construct cloned into vector pMK using the GeneArt Subcloning & Express Cloning Service (ThermoFisher). These expression constructs were then amplified and inserted into the endA gene of the E. coli genome using primers (forward: 5'-CCAAAACAGCTTTCGCTACGTT
GCTGGCTCGTTTTAACACGGAGTAAGTGTTAGAAAAATTCATCCAGCA-3', reverse: 5'-GGTT
GTACGCGTGGGGTAGGGGTTAACAAAAAGAATCCCGCTAGTGTAGGCGGGCAGTGAAAGGAAGGCC-3').

BB170 AI-2 bioassay

Cultures of signal-producing E. coli strains or luxS- parental strains were grown overnight in LB, and subcultured 1:100 into 12 mL of LB and grown with agitation at 37°C until $OD_{600} \approx 0.2$. They were then centrifuged, and the supernatant was filtered using 0.22 μm filters. AI-2 activity in the resulting cell-free supernatants was then assessed using the BB170 bioassay.[74] Briefly, overnight cultures of reporter strain BB170 in LM medium were diluted at 1:1,000 in AB medium and 10 μL of cell-free supernatant or heat-treated cell-free supernatant were then added to 90 μL of BB170 dilution. The luminescence and OD_{600} of each sample were measured immediately and after ~3 hours growth at 30°C with agitation.

Infant mouse colonization

Four-day-old CD-1 suckling mice were purchased from Charles River Laboratories. To deplete the gut microbiota, the suckling mice were fasted for 1.5 hours and treated with approximately 1 mg/g body weight of streptomycin by gavage with a 30-gauge plastic tubing. Pups were then placed back with a lactating dam. After 24 hours, the mice were separated into three groups ($n = 7$), and each mouse received a maximum gavage volume of 50 μL of a defined bacterial mixture. The bacterial mixture for gavage consisted of a total microbial mass of 300 μL of $OD_{600} = 0.4$ per inoculum, divided into (i) 45% microbial community, 20% E. coli expressing luxS from either Streptococcus, B. obeum, or an empty expression vector, and 35% of A. muciniphila or (ii) equal quantities of E. coli strains and A. muciniphila. Gavaged mice were placed in at 30°C for 16 hours and then sacrificed. The small and large intestines were dissected, separated, and homogenized in 5 mL LYH-BHI followed by DNA isolation.

Quantitative PCR analysis of A. muciniphila abundance

To determine the abundance of A. muciniphila relative to the bacterial load, DNA extracts were isolated from both the small and large intestines of colonized suckling mice. 500 μL of intestinal homogenates were mixed with 500 μL of 0.1 mm Zirconia beads (BioSpec),

210 μL of 20% SDS, 500 μL of phenol:chloroform:isoamyl-alcohol (24:24:1, Fisher Scientific), and then lysed using a bead beater at 2,400 RPM for one minute. Next, qPCR assays were performed using primers specific to *A. muciniphila* (Forward: 5'-CAGCACGT GAAGGTGGGGA-3', reverse: 5'-CCTTGCGGTTGGCTTCAGATC-3') and a primer set targeting the total bacterial load (Forward: 5'-CTCCTACGGGAGGCAGCAG-3', reverse: 5'-TTACCGCGGCTGCTGGCAC-3'). Each reaction was performed in triplicates and consisted of 12.5 μL of iQ SYBR Green Supermix (BIO-RAD2, Hercules, CA), 0.25 μL of forward and reverse primers at 10 μm concentration, 10 μL nuclease free water and 2 μL of genomic DNA (500 ng/μL).

QUANTIFICATIONS AND STATISTICAL ANALYSIS

All *in vitro* experiments were independently conducted for at least three times, and the quantitative data were presented as mean ± standard deviation. Statistical analyses were performed using GraphPad Prism Software (v9). Statistical details of the experiment can be found in the figures and figure legends. Wilcoxon test were performed to determine the statistical differences of the bacterial genes, and the differences were considered to be statistically significant when $P < 0.05$. Pathway enrichment was analyzed using Fisher's exact test. P values were corrected for multiple hypothesis testing via the Benjamini-Hochberg procedure in R, and a false discovery rate (FDR) < 0.05 was set as the significance threshold.