Predictive power of different *Akkermansia* phylogroups in clinical response to PD-1 blockade against non-small cell lung cancer

Peixin Fan[1], Mi Ni[1], Yu Fan[1], Magdalena Ksiezarek[1], Gang Fang[1*]

[1]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

[*]Correspondence to Gang Fang, gang.fang@mssm.edu

Summary

Immune checkpoint blockade has emerged as a promising form of cancer therapy. However, only some patients respond to checkpoint inhibitors, while a significant proportion of patients do not, calling for the discovery of reliable biomarkers. Recent studies reported the importance of the gut microbiome in the clinical response to PD-1 blockade against advanced non-small cell lung cancer (NSCLC), highlighting Akkermansia muciniphila as a candidate biomarker. Motivated by the genomic and phenotypic differences across Akkermansia muciniphila strains and Akkermansia (Akk) phylogroups (AmIa, AmIb, AmII, AmIII and AmIV), we analyzed fecal metagenomic sequencing data from three publicly available NSCLC cohorts (n=425). Encouragingly, we found that patients' responses to PD-1 blockade are significantly different across different Akk phylogroups, highlighting the relatively stronger association between AmIa and positive response than the other phylogroups. We built a machine learning model based on Akk gene profiles, which improved the predictive power and shed light on a group of Akk genes that may be associated with the response to PD-1 blockade. In summary, our study underlines the benefits of high-resolution analysis of Akk genomes in the search for biomarkers that may improve the prediction of patients' responses to cancer immunotherapy.

Key words:

Non-small cell lung cancer (NSCLC), immunotherapy, *Akkermansia muciniphila*, phylogroup

Introduction

Lung cancer is the third most common type of cancer in the US and the leading cause of cancer-related mortality worldwide[1,2]. Non-small cell lung cancer (NSCLC) is the most prevalent form of lung cancer (84%). The emergence of immunotherapy, especially the immune checkpoint blockade, has revolutionized the cancer treatments with great promise, especially for NSCLC[3,4,5]. However, only some patients experience durable responses to immune checkpoint inhibitors (ICIs), while others do not respond. For example, although NSCLC has a relatively higher response rate to ICIs than other cancer types, only ~20% of NSCLC patients respond favorably to ICIs as monotherapy[6]. Therefore, it is urgent to understand critical factors that influence the response to immunotherapy and discover predictive biomarkers that may guide personalized medicine.

Several promising biomarkers have been discovered for predicting responses to immunotherapy and have been incorporated into clinical practice, including the expression of immune checkpoint proteins (e.g., PD-1 or PD-L1), pre-existing immune responses, tumor mutation burden, and tumor-associated proteins[6,7,8,9,10]. From a recent study, bronchoalveolar lavage fluid analysis of cellular profiles or secreted molecules can help identify patients with NSCLC who respond to ICIs[11]. In addition, recent studies reported an association between patients' gut microbiome and their responses to immunotherapies[12,13]. Animal model-based studies found that antitumor activity of ICIs was lower in microbiome-depleted mice such as germ-free mice and mice treated with antibiotics, compared to specific pathogen-free mice[14,15]. Patient cohort studies showed significant differences in gut microbiome composition between NSCLC, renal cell carcinoma, and melanoma patients who responded to anti-PD-1 therapy and those who did not[15,16,17,18]. Fecal microbiota transplantation (FMT) was shown to influence the

ICIs treatment of refractory melanoma and colorectal cancer in clinical trials[19,20], suggesting some specific gut microbes may serve as non-invasive predictive biomarkers.

For NSCLC, recent studies reported that *Akkermansia muciniphila*, a mucin-degrading gut commensal bacterium, was associated with better response to PD-1-based immunotherapy, and that it can enhance ICIs treatment in mice[21,22,23]. Functionally, *A. muciniphila* is associated with beneficial effects on host metabolism and homeostatic immunity[24,25]. It can also directly trigger intestinal adaptive immune responses through inducing T cell-dependent IgG1 and IgA[23]. Immunogenic activity of lipid (a diacyl phosphatidylethanolamine with two branched chains) and peptides (e.g., Amuc_1100 and P9) produced by *A. muciniphila* were observed *in vitro*[26,27]. Notably, several studies have demonstrated the genomic and phenotypic differences across *A. muciniphila* strains, which initiated further classification into *Akkermansia* (Akk) phylogroups (primarily AmIa, AmIb, AmII, AmIII and AmIV). Among these phylogroups, only AmIa and AmIb represent *A. muciniphila,* while the remaining are being actively characterized as potential new *Akkermansia* species[28,29,30,31]. Different Akk phylogroups were shown to have significant phenotypic differences that are associated with different host-microbe interactions: growth rates in mucin, binding to epithelial surfaces, and activation of innate immune responses[28,29,30,31,32]. A recent study also identified correlations between Akkermansia species and subspecies and human health outcomes[31]. Consistently, two different Akk strains showed different capacities to modulate differentiation of regulatory T cells (Tregs) in mice[15]. Based on these genetic and phenotypic variations among Akk strains, we hypothesized that NSCLC patients carrying different phylogroups of Akk in their gut microbiome have different responses to immunotherapy.

To test this hypothesis, we first determined the distribution of Akk phylogroups in NSCLC metagenomics datasets of 425 samples from three publicly available cohorts[15,22], and evaluated the predictive power of different Akk phylogroups for patients' responses to PD-1 blockade. Furthermore, we performed a rigorous analysis and highlighted a subset of genes in the Akk pangenome that are enriched in patient responders.

Results

Akk phylogroups across three cohorts of NSCLC patients with ICI treatment

We performed a meta-analysis across three publicly available NSCLC cohorts (*C1*: *Routy* et al. Discovery, n = 60; *C2*: *Routy* et al. Validation, n = 27; *C3 Derosa* et al. n = 338) that used fecal shotgun metagenomics to study the associations between Akk and response to PD-1 blockade[13,20]. All raw shotgun metagenomic sequencing data were processed consistently using the same bioinformatics pipeline (STAR Methods).

We first re-evaluated the associations between the presence/absence of Akk (Akk^+ or Akk^-, determined by MetaPhlAn[33], Methods) and response to PD-1 blockade against NSCLC. The sequencing yield (Gb) per sample did not have a significant difference between Akk^+ and Akk^- samples across three cohorts (P_{C1} = 0.98, P_{C2} = 0.53, P_{C3} = 0.18, Student's t test, Fig. S1). Consistent with previous studies, the Akk^+ group had higher (compared to the Akk^- group) proportion of progression-free patients [complete response (CR), partial response (PR), or stable disease (SD), Fig. S2A-C] and objective response (CR and PR, Supplementary Fig. 2D-F) in all cohorts. Noteworthy, in the *C2* cohort, all the 5 patients with CR or PR were characterized as Akk^+ group (P = 0.0009, Fig. S2E). We also evaluated the predictive power of Akk relative abundance, instead of Akk presence/absence, and found significant association

with better response (CR, PR, SD) (P = 0.025) in the *C1* cohort but not in *C2* or *C3*. These analyses supported the use of the presence/absence of Akk in the following analyses, which is consistent with a recent study that reported bacteria presence/absence is more robust than bacterial relative abundance in microbiome-based association studies[34]. Overall, NSCLC patients carrying Akk (presence vs. absence) had a higher response rate to PD-1 blockade, consistently across the three cohorts.

To distinguish Akk phylogroups in the meta-analysis, we first built a comprehensive Akk pangenome including diverse Akk strains (216 Akk isolate genomes from the NCBI database) that include five known Akk phylogroups (Fig. S3A; STAR Methods). Out of the 231 Akk^+ samples across the three cohorts, 24 (out of 30), 8 (out of 12) and 122 (out of 189) samples were assigned with an Akk phylogroup (Methods), respectively (Fig. 1A-B, Fig. S3B-D, C1: 10 AmIa, 10 AmIb, and 4 AmII; C2: 4 AmIa, 3 AmIb, and 1 AmII; C3: 58 AmIa, 37 AmIb, 21 AmII, and 6 AmIV). Consistent across the three cohorts, AmIa and AmIb are the two most prevalent Akk phylogroups, representing typical *A. muciniphila* species that are the most common Akk species found in the human microbiome[26], followed by AmII and AmIV.

Akk phylogroups have different predictive power for patient response to PD-1 blockade

We compared patients' response rate between NSCLC patients carrying different Akk phylogroups using a logistic regression model (with covariates; STAR Methods) and found that AmIa is associated with better response (than other phylogroups) to PD-1 blockade across all the three NSCLC cohorts (Fig. 2). Specifically, the rates of responders (CR, PR or SD) in the AmIa group were 80% in *C1* (P_{AmIa} = 0.037, Fig. 2A), 75% in *C2* (P_{AmIa} = 0.075, Fig. 2B), and 58.62% in *C3* (P_{AmIa} = 0.016, Fig. 2C), respectively. We also evaluated the logistic regression

models that included Akk phylogroup vs. Akk presence/absence ($Akk^{+/-}$, not stratifying phylogroups) to predict objective response (CR or PR) using Akaike information criterion (AIC)[35], which measures the relative prediction error between regression models (the smaller value the better model). Consistently across the three cohorts, the logistic regression models that included Akk phylogroup have better accuracy and AIC than the models that used Akk species: C1: AIC = 50.538 (Akk phylogroup) vs 52.907 ($Akk^{+/-}$); *C2*: AIC = -16.065 (Akk phylogroup) vs 20.375 ($Akk^{+/-}$); *C3*: AIC = 299.9 (Akk phylogroup) vs 305.19 ($Akk^{+/-}$). We further compared patient survival across different Akk phylogroups. In the *C1* cohort, 60% of patients carrying AmIa had progression-free survival (PFS) longer than 3 months (vs. 43.75% of the other patients, P_{AmIa} = 0.036, Fig. 2D). In the *C2* cohort, 75% of patients carrying AmIa had PFS longer than 3 months (vs. 33.33% of AmIb patients, P_{AmIa} = 0.0005, Fig. 2E). Consistently, in the C3 cohort, AmIa patients had longer overall survival (OS) than patients not carrying AmIa based on Cox proportional-Hazards regression analysis (P_{AmIa} = 0.0371, Fig. 2F). It is worth noting that both AmIa and AmIb are prevalent in the three NSCLC cohorts, yet AmIa is associated with better patients' response and survival than AmIb, consistently across the three cohorts.

Derosa et al. reported in the *C3* cohort that tumor microenvironment differed between Akk^{+} patients and Akk^{-} patients, highlighting several differentially expressed genes[22]. To evaluate if Akk phylogroups and patients' differential response and survival rate correlate with patients' tumor microenvironment, we examined the gene expression levels of 28 immune-related genes between patients with AmIa and AmIb phylogroups. Interestingly, 21 of the 28 genes ($P_{binomial}$ = 0.006) that were expressed higher in Akk^{+} patients than Akk^{-} patients showed greater expression level in patients carrying AmIa compared to those in AmIb, with 5 of them showing significantly difference (P_{adj} < 0.05, Fig. 2G). The *TDO2* gene, which is the interferon fingerprint and associated with checkpoint pathway, showed the largest difference between

AmIa and AmIb group (LogFC = 3.79, P_{adj} = 0.008). *CXCL10* (LogFC = 1.95, P_{adj} = 0.049), related to type II interferon signaling, *VCAM1* (LogFC = 1.94, P_{adj} = 0.015), related to leukocyte migration, *RB1* (LogFC = 1.69, P_{adj} = 0.033) related to tumor marker, and *CMKLR1* (LogFC = 1.62, P_{adj} = 0.048), related to dendritic cell and macrophage also had higher expression in AmIa than AmIb. These analyses suggest that the tumor microenvironment genes associated with Akk^{+} patients have more pronounced changes in their expression level in the patients carrying the AmIa phylogroup than the AmIb, consistent with their stronger anti-tumor immune response to ICI.

Taken together, these analyses suggest that the presence of AmIa phylogroup in NSCLC patients' gut microbiome is associated with better responses than the other Akk phylogroups to PD-1 blockage, survival and tumor microenvironment.

Akk gene-based prediction of response to anti-PD1 blockade

Building on the observations above that NSCLC patients carrying AmIa were associated with better response than the other phylogroups, we further hypothesized that AmIa genomes may be enriched for certain genes that are positively associated with the response to ICI. It is also worth noting that not all the patients carrying AmIa responded to PD-1 blockade while some patients with other Akk phylogroups did respond to PD-1 blockade. This heterogeneity within a phylogroup is not unexpected considering the genomic differences across strains within each phylogroup. Therefore, we further performed an Akk gene-level analysis to better understand the different predictive power between Akk phylogroups for patients' response to PD-1 blockade.

We built a machine-learning model (model I) with patients' Akk phylogroup, patients' metadata (covariates: age, gender, antibiotics, sequencing depth, which are available in all the three cohorts) to predict response (CR/PR/SD or PD/D). We compared the accuracy of model I with another model (model II) built with the addition of presence or absence of Akk genes. Considering the sample sizes of the three cohorts, we trained the two models using all the available Akk^+ samples with defined phylogroups in the combined *C1* and *C2* cohort (46 Akk^+ samples including 14 additional samples collected at a later time point from the cohorts), and the *C3* cohort (122 Akk^+ samples), respectively. For the combined cohort (*C1* and *C2*), the leave-one-out cross-validation (LOOCV) accuracy by the model I using random forest had an area under the receiver operating characteristic (AUROC) value of 0.76 (Fig. 3A), while the LOOCV of model II had an AUROC of 0.86 (Fig. 3B). For the *C3* cohort, the LOOCV of models I and II had AUROCs of 0.49 and 0.57 (with respect to the 54.92% responders in *C3*), respectively (Fig. 3C-D). The improvements in LOOCV accuracy of the training model with the addition of Akk gene feature (consistent between the combined *C1*/*C2* cohort and the *C3* cohort) indicate the association between certain Akk genes and patients' response to PD-1 blockade can provide better predictive power than Akk phylogroups itself.

To investigate which genes contribute significantly to improving the predictive power and their distribution across different phylogroups, we combined the 168 Akk^+ patients with defined Akk phylogroup across the three cohorts and constructed a model (Model III). The training model with patients' metadata, sequencing depth, and presence and absence matrix of all the 4,576 pan-Akk genes, selected using strict criteria to avoid ambiguous mapping of genes from other species to Akk (STAR Methods), reach the accuracy of 0.71 (Fig. 4B), which is higher than the accuracy solely based on Akk phylogroups (55.0%, Fig. 4A). 73 Akk genes (variable importance score (VIS) > 0.5, present > 1% of at least one Akk phylogroup) showed significantly

different between responders and non-responders ($P_{chi\text{-}square} < 0.05$, Table S1), with 35 enriched in the responders (Fig. 4C) and 38 enriched in the non-responders (Fig. 4D), respectively. Among the 35 Akk genes enriched in the responders, 22 genes had highest prevalence in AmIa, followed by 7 genes in AmIb, 3 genes in AmIV, 2 gene in AmII, and 1 gene had highest prevalence in both AmIa and AmII (Fig. 4C). These responder-associated genes include multiple carbohydrate-active enzymes (CAZy) encoding genes, such as glycosyltransferase family 1 protein (*CXU07_01840*), alpha-1,2-fucosyltransferase (*Amuc_2087*), and glycosyl transferase group 1 (*Amuc_2082*). The distribution of these CAZy encoding genes was much higher in AmIa followed by AmIb phylogroup and rarely identified in AmII or AmIV phylogroups. For the 38 Akk genes enriched in non-responders over responders, 18 genes had highest prevalence in AmIb, followed by 16 genes in AmII and 4 genes in AmIa (Fig. 4D). These genes encode several bacteriophage related proteins, such as bacteriophage tail assembly protein-like protein (*Amuc_1355*) and phage major capsid protein which had the highest prevalence in AmII, as well as phage portal protein, which was only detected in AmIb.

Besides, nonresponder-associated Akk genes also include certain CAZy encoding gene, such as glycosyltransferase 2-like domain-containing protein (CXU19_11355), which is frequently detected in AmII but not in other Akk phylogroups. These results suggest a link between carbohydrate utilization of gut *Akkermansia* and response to immunotherapy of hosts, likely reflecting variation in intestinal mucosal environments, particularly glycoside diversity between responders and nonresponders, which probably results from their distinct immune and metabolic status, that favor growth of different Akk phylogroups[36,37]. The patients dominated by AmI phylogroup may also benefit from bacterial glycoside fermentation end products, such as short-chain fatty acids, that enhance their response to PD-1 blockade[38,39]. AmIa and AmIb containing the most and second most genes enriched in responders is consistent with the

observation that showed patients with AmIa and AmIb are more likely (than the other phylogroups) to be responders. In contrast, AmIb and AmII harboring more genes enriched in non-responders, and particularly AmII had few genes enriched in responders, which is consistent with the observation that patients with AmII are more likely to be non-responders.

Several additional genes that had higher prevalence in responders also showed varied distribution across Akk phylogroups. For example, Amuc_1677, a gene enriched in responders, was reported to encode an enzyme involved in the (S)-acetolacetate to 2-methyl-2-oxovaleric acid step in the branched-chain amino acids to phosphatidylethanolamine pathway[27]. The frequency of Amuc_1677 across the Akk phylogroups is 51.85% in AmIa, 17.24% in AmIb, absent in AmII and AmIV. Among the patients with either AmIa or AmIb, those carrying Akk with an Amuc_1677 gene had a higher chance to be responders (Fig. 4E). Another gene enriched in responders, Amuc_2164 (glycoside hydrolase family 18, a predicted mucin binding chitinase) is only in strains from the AmI phylogroup (83% in AmIa, 99.14% in AmIb) but absent in the other phylogroup. These data indicate AmI strains, encode certain Akk genes that are associated with better response to PD-1 blockage, while AmII strains lack beneficial genes or might harbor certain genes that are negatively associated with patients' response to PD-1 blockade. It is worth noting these genes represent association but not causality, and their functionality in the context of responses to PD-1 blockage needs to be further explored in the future studies.

To summarize, we found that the model that incorporate Akk genes has enhanced performance in the prediction of patients' response to PD-1 blockade. Importantly, the gene-level association analysis (with patient response and with Akk phylogroups) helped elucidate the difference in the predictive power of different Akk phylogroups in clinical response to PD-1 blockade against NSCLC.

Discussion

In this study, we assessed the difference between Akk phylogroups with respect to their predictive capacity in responses to PD-1 blockade in NSCLC patients. Through the delineation of Akk phylogroups and an extensive meta-analysis across three previously published cohorts, we unveiled diverse Akk phylogroups among NSCLC's patients, highlighting the better power of the AmIa phylogroup, consistently across three independent cohorts, in predicting responses to immunotherapy compared to the other phylogroups. Also consistently, we observed stronger association between AmIa and patient survival and the RNA expression of patients' genes relevant to tumor microenvironment. Using a machine learning model predicting patients' response based on Akk gene profiles, we searched for the genes associated with AmIa's superior predictive capabilities, and highlighted a number of genes predominantly prevalent in the AmIa phylogroup and associated with patients' response to PD-1 blockade. Some of these genes are also present in other phylogroups, e.g. AmIb, but less prevalent compared to AmIa. This is consistent with the observation that some patients carrying AmIb are also responders. In addition, we found a few genes that are predominantly prevalent in non-responders and AmII, which are consistent with the worse response of patients carrying AmII. These revelations underscore the importance of discerning Akk phylogroups and Akk genes in disease association analysis, which may help the development of innovative microbiome-centric diagnostic tools and personalized medicine approaches for PD-1 blockage in NSCLC patients.

Our integration and analysis of the three cohorts also emphasized some limitations and the need for further validations through larger, independent studies, considering the complexity of NSCLC and patients' response to PD-1 blockage. First, although our analytical models used three co-variants shared across the three cohorts (age, gender, antibiotics), we think the

collection and incorporation of additional covariates is paramount in these analyses, given their potential to introduce confounding effects. For instance, cohorts C1 and C2 predominantly comprised patients who underwent chemotherapy before PD-1 blockade, whereas cohort C3 had a higher proportion of patients treated with radiotherapy before PD-1 blockade. These variations might explain the relatively better predictive power of AmIa in C1 and C2 than in C3, even though the three cohorts consistently highlight the importance of AmIa. Second, while this study was focused on the microbiome, we also recognize the substantial importance of host factors such as ECOG performance status, PD-L1 expression, tumor gene expression data, in predicting patients' response to PD-1 blockage[22]. Although ECOG performance status, PD-L1 expression, and tumor gene expression data are available in the C3 cohort, they were not part of the other two cohorts. Given the complexity of NSCLC and patients' response to PD-1 blockage, future development in this field necessitates studies with a holistic approach, integrating both microbiome and host gene expression data to decipher their relationships and their joint predictive power[23,40,41]. Besides, due to the limitation of annotated Akk gene based on our current knowledge, the functions of response-associated sequence clusters need to be further explored.

While Akk has persistently demonstrated a link to responses in NSCLC patients undergoing PD-1 blockade across varied cohorts, several other bacterial species like *Eubacterium hallii*, *Bifidobacterium adolescentis*, and *Parabacteroides distasonis* have also been reported, albeit with less consistency across cohorts[22,42,43]. This discrepancy might arise from the possibly more subtle yet legitimate associations these bacteria have with PD-1 blockade responses or due to the inherent patient heterogeneity observed across diverse cohorts. Indeed, the complexity of responses to PD-1 blockade in NSCLC is multifaceted, interweaving numerous factors from both the human host and the gut microbiome. While Akk is

a significant player in inducing intestinal adaptive immune responses during homeostasis, it's plausible that other bacterial species also play integral roles in this biological puzzle. With this rationale, we made an attempt to assess the predictive capabilities of additional bacterial species, probing their association with responses to PD-1 blockade (Supplementary Text). *Monoglobus pectinilyticus* exhibited a consistent positive association with responders across all the three NSCLC cohorts (Supplementary Text and Fig. S4), although future work is imperative to validate this finding on large and independent cohorts as well as the complementarity between Akk and additional bacterial species.

Beyond its association with responses to PD-1 blockade, Akk has also been implicated in a spectrum of other diseases, including obesity, type 2 diabetes, and inflammatory bowel disease, among others[44,45,46]. The insights from this study also suggest the use of similar analytical strategies to investigate the predictive capabilities of distinct Akk phylogroups in relation to these other diseases and their respective treatment outcomes. Given the recognition of Akk as a promising candidate in the development of next-generation probiotics, it is important to attain a nuanced understanding of the various Akk phylogroups, their specific genes, and genomic variations[47] across a range of diseases, which could unravel transformative insights and facilitate the advancement of innovative therapeutic interventions and proactive preventive measures.

Figure legend

Figure 1. NSCLC patients carrying different phylogroup of Akk varied in response to ICI treatment. (A) The Akk phylogroups carried by NSCLC patients from three cohorts. **(B)** The AmIa and AmIb phylogroup carried by NSCLC patients from three cohorts. The three cohorts are *Routy* Discovery cohort (n = 60), *Routy* Validation cohort (n = 27), and *Derosa* cohort (n = 338, *P* = 0.016).

Figure 2. **NSCLC patients carrying different phylogroup of Akk varied in response to ICI treatment. (A to C)** Correlations between phylogroup of Akk in the stool of NSCLC patients and response to PD-1 blockade. AmIa group had higher proportion of patients being CR, PR, or SD in *Routy* Discovery cohort (n = 60, *P* = 0.037) (**A**), *Routy* Validation cohort (n = 27, *P* = 0.075) (**B**), and *Derosa* cohort (n = 338, *P* = 0.016) (**C**). In *Routy* Validation cohort, AmIa group showed significantly higher proportion of patients being CR or PR ($P = 9.04\times10^{-8}$). CR; complete response. PR; partial response. SD; stable disease. PD; progressive disease. D; death. Data were analyzed using generalized linear model. (**D**-**F)** Correlations between phylogroup of Akk in stool of NSCLC patients and survival length. AmIa group had higher proportion of patients which had progression-free survival (PFS) longer than 3 months in *Routy* Discovery cohort (n = 60, *P* = 0.036) (**D**) and *Routy* Validation cohort (n = 27, *P* = 0.0005) (**E**), and had longer overall survival in *Derosa* cohort (**F**). Data were analyzed using generalized linear model. **G**, Bar graph showing most immune-related genes that were expressed at higher level in lung tumor of patients carrying Akk had higher expression level in AmIa group. Differential expression analysis was conducted using voom method. Akk-/un represents for the samples that are either

detected as Akk negative using MetaPhlAn 3 or detected as Akk positive in MetaPhlAN 3 but not in PanPhlan[*], $P < 0.05$; [#], $0.05 < P < 0.1$.

Figure 3. Akk gene-based model improved performance to predict response to PD-1 blockade. (A) The area under the receiver operating characteristic (AUROC) of the random forest training model to predict response using Akk phylogroup, patients' metadata, and sequencing depth in the combined C1 and C2 cohort (46 Akk^+ samples). **(B)** The AUROC of the random forest training model to predict response with Akk gene presence or absence information together with patients' metadata, and sequencing depth in the combined C1 and C2 cohort (46 Akk^+ samples). **(C)** The AUROC of the random forest training model to predict response using Akk phylogroup, patients' metadata, and sequencing depth in the C3 cohort (122 Akk^+ samples). **(D)** The AUROC of the random forest training model to predict response with Akk gene presence or absence information together with patients' metadata, and sequencing depth in the C3 cohort (122 Akk^+ samples).

Figure 4. Potential Akk genes that are associated with response to PD-1 blockade.

(A) The AUROC of the random forest training model to predict response using Akk phylogroup, patients' metadata, and sequencing depth in the combined C1, C2, and C3 cohort (168 Akk^+ samples). **(B)** The AUROC of the random forest training model to predict response with presence or absence information of Akk genes together with patients' metadata, and sequencing depth in the combined C1, C2, and C3 cohort (168 Akk^+ samples). (**C-D)** The heatmaps showing the distribution of response-associated Akk genes across Akk phylogroups. Only Akk genes with prevalence higher than 1% in at least one phylogroup were included in the heatmaps **(C)** The distribution of responder-associated Akk genes ($P_{\text{Chi-square}} < 0.05$) with

variable importance score (VIC) higher than 0.5 in the random forest model across different Akk phylogroups. Only Akk genes present in more than 1% of at least 1 phylogroup were included in the heatmap. **(D)** The distribution of nonresponder-associated Akk genes ($P_{\text{Chi-square}} < 0.05$) with VIC higher than 0.5 in the random forest model across different Akk phylogroups. Only Akk genes present in more than 1% of at least 1 phylogroup were included in the heatmap. **(e)** The bargraph showing the prevalence of responders and nonresponders of the patients in AmIa or AmIb that carry or not carry Amuc_1677 gene.

STAR Methods

Public populations of NSCLC patients

Raw gut metagenomic sequencing data of three populations (C1: *Routy* et al. Discovery, n = 60; C2: *Routy* et al. Validation, n = 27; C3: *Derosa* et al. n = 338) were downloaded from the Sequence Read Archive (SRA) under the BioProject accession PRJEB22863 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJEB22863) and PRJNA751792 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA751792), mainly from two recently published papers[15,22]. The raw tumor RNA sequencing data were downloaded from the SRA under accession the NCBI accession GSE182328 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE182328) from *Derosa* et al. (2022)[22].

Identification of Akk presence and Akk phylogroup in the stool metagenomes

The KneadData tool (v0.10.0) was used to perform quality control on raw metagenomic sequencing data. Briefly, adapters and low-quality reads were removed using Trimmomatic (v0.39), and repetitive sequences were trimmed using the TRF (v4.09.1). The human reads were detected by mapping remaining reads to the human genome (hg37) by bowtie2 (v.2.2.5),

and removed. *Akkermansia muciniphila* presence and relative abundances of bacterial taxa were determined and estimated by MetaPhlAn 3 (v3.0) with ChocoPhlAn 3 database[32]. PanPhlan pipeline (v3.0) was applied to profile *A. muciniphila* strains within stool metagenomes. A total of 216 available *A. muciniphila* isolate genomes (fasta file) were download from NCBI and were used to build a comprehensive pangenome with UniRef90 ID using PanPhlan pangenome exporter (https://github.com/SegataLab/PanPhlAn_pangenome_exporter). Due to high genetic diversity within *A. muciniphila* strains, this species has been suggested to be classified into four species-level phylogroups, and we considered the UniRef genes families-level pangenome as *Akkermansia* (Akk) pangenome. Then each metagenomic sample was mapped against the generated Akk custom pangenome (panphlan_map.py). The coverages of all gene positions were detected and extracted using samtools (v1.16), and then integrated to a gene family coverage profile for each sample. Pangenome-related Akk strains were predicted to exist if coverage curves display a satisfying plateau using the least stringent threshold of panphlan_profile.py (--min_coverage 1, --left_max 1.70, --right_min 0.30). Principal coordinates analysis (PCoA) of Jaccard distance was performed on UniRef90 ID presence and absence binary matrix of metagenomic samples using ggplot2 package in the R (v4.3) to distinguish phylogroup of Akk strains according to cluster with references.

Evaluation of correlations between Akk presence/absence, Akk phylogroup and response to ICI treatment

Generalized linear model was applied to evaluate the effects of Akk presence/absence or specific Akk phylogroup on the response. The patients were classified as responders (complete response (CR), partial response (PR), and stable disease (SD)) and non-responders (progressive disease (PD) and death (D)) according to the previous study[15]. Akk presence or Akk phylogroup with sequencing depth (number of raw reads) and patients' metadata (age,

gender, time point, ECOG performance status, PD-L1 expression if available) were included as independent variables. Responders, objective response (CR and PR), or survival lengths were served as dependent variables. Akaike information criterion (AIC) and accuracy were used to compare the quality of models.

Tumor RNA sequencing data analysis

To investigate relationships between the expression of immune-related genes in tumors and Akk phylogroup the patients were carrying, the differences in Akk-associated immune-related genes reported by previous study was analyzed using the published tumor RNA sequencing data[14]. The low-quality reads and adapter were removed from raw RNA sequencing data using Trim Galore (v0.6.6). The remaining reads were mapped to human GRCh38 cDNA reference transcriptome using STAR (v2.7.2a). The number of reads mapping to each gene was counted using featureCounts (v2.0.1). The differentially expressed genes (DEGs) between patients carrying AmIa (n = 3) and AmIb (n = 5) were analyzed in R using voom package (v3.38.14).

Construction and evaluation of the machine-learning model based on Akk genes profile

To determine whether Akk gene profiles could enhance the prediction of response to PD-1 blockade, the random forest models were built using R caret package (v 6.0-94). The patients from the three cohorts with Akk phylogroup assigned were included in this analysis. Additionally, to increase the sample size, an additional 14 fecal metagenomic samples from the C1 cohort collected after the 2nd (T1- 1 month), 4th (T2- 2 months) and 12th (T3- 6 months) injection of anti-PD-1/PD-L1 mAb and with Akk phylogroup assigned were included[22].

The training models were constructed with leave one out cross validation to distinguish responders and non-responders. The features used for model building consist of sequencing

depth (number of raw reads), patient metadata (age, gender, usage of antibiotics), and Akk profile (phylogroup, presence and absence of Akk only genes). The UniRef genes that only exist in *Akkermansia* were considered as Akk only genes. Briefly, the metadata for the full UniRef90 gene set was retrieved from the UniRef website (http://www.uniprot.org/uniref), which includes the organism/species information. All the Akk-related genes was extracted by using the Organism tag ("*Akkermansia*"). These Akk genes were classified into different categories based on the number of homologous. The genes that only existed in Akk were pulled out as Akk only genes. The models were optimized by tuning the mtry (number of randomly selected features to be sampled at each iteration) and the ntree (number of decision tree). The area under the receiver operating characteristic (AUROC) and accuracy were used to evaluate the performance of the training models. The variable importance scores were extracted for each feature in the models, and distribution of critical Akk genes with high score (> 0.5) with significant difference present in responders and non-responders ($P_{chi\text{-}square} < 0.05$) were determined across phylogroups.

References

1. Centers for Disease Control and Prevention. Lung Cancer Statistics. https://www.cdc.gov/cancer/lung/statistics/index.htm#:~:text=Lung%20cancer%20is%20the%20third,any%20other%20type%20of%20cancer. 2023.
2. Centers for Disease Control and Prevention. An Update on Cancer Deaths in the United States. https://www.cdc.gov/cancer/dcpc/research/update-on-cancer-deaths/index.htm.2020.
3. Zhou F, Qiao M, Zhou C. The cutting-edge progress of immune-checkpoint blockade in lung cancer. *Cell. Mol. Immunol.* 2021; **18**: 279-293.
4. Chiang AC, Herbst RS. Frontline immunotherapy for NSCLC—the tale of the tail. *Nat. Rev. Clin. Oncol.* 2020;**17**: 73-74.

5. MacManus M, Hegi-Johnson F. Overcoming immunotherapy resistance in NSCLC. *Lancet Oncol.* 2022; **23**: 191-193.

6. Doroshow DB, Bhalla S, Beasley MB, Sholl LM, Kerr KM, Gnjatic S, Wistuba II, Rimm DL, Tsao MS, Hirsch FR. PD-L1 as a biomarker of response to immune-checkpoint inhibitors. *Nat. Rev. Clin. Oncol.* 2021; **18**: 345-362.

7. Kambayashi Y, Fujimura T, Hidaka T, Aiba S. Biomarkers for predicting efficacies of anti-PD1 antibodies. *Front. Med. (Lausanne).* 2019*,* **6**: 174.

8. Liu Y, Altreuter J, Bodapati S, Cristea S, Wong CJ, Wu CJ, Michor F. Predicting patient outcomes after treatment with immune checkpoint blockade: A review of biomarkers derived from diverse data modalities. *Cell Genom.* 2024; **4**: 100444.

9. Lu S, Stein JE, Rimm DL, Wang DW, Bell JM, Johnson DB, Sosman JA, Schalper KA, Anders RA, Wang H et al. Comparison of Biomarker Modalities for Predicting Response to PD-1/PD-L1 Checkpoint Blockade: A Systematic Review and Meta-analysis. *JAMA Oncol.* 2019; **5**: 1195-1204.

10. Ren D, Hua Y, Yu B, Ye X, He Z, Li C, Wang J, Mo Y, Wei X, Chen Y et al. Predictive biomarkers and mechanisms underlying resistance to PD1/PD-L1 blockade cancer immunotherapy. *Mol. Cancer*. 2020; **19**: 31.

11. Masuhiro K, Tamiya M, Fujimoto K, Koyama S, Naito Y, Osa A, Hirai T, Suzuki H, Okamoto N, Shiroyama T et al. Bronchoalveolar lavage fluid reveals factors contributing to the efficacy of PD-1 blockade in lung cancer. JCI Insight. 2022; **7**: e157915.

12. Zhou CB, Zhou YL, Fang JY. Gut microbiota in cancer immune response and immunotherapy. *Trends Cancer.* 2021; **7**: 647-660 (2021).

13. Simpson RC, Shanahan ER, Batten M, Reijers ILM, Read M, Silva IP, Versluis JM, Ribeiro R, Angelatos AS, Tan J et al. Diet-driven microbial ecology underpins associations between cancer immunotherapy outcomes and the gut microbiome. *Nat. Med.* 2022; **28**: 2344-2352.

14. Iida N, Dzutsev A, Stewart CA, Smith L, Bouladoux N, Weingarten RA, Molina DA, Salcedo R, Back T, Cramer S et al. Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. *Science* 2013; **342**: 967-970.

15. Routy B, Le Chatelier E, Derosa L, Duong CPM, Alou MT, Daillère R, Fluckiger A, Messaoudene M, Rauber C, Roberti MP et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science.* 2018; **359**: 91-97.

16. Vernocchi P, Gili T, Conte F, Del Chierico F, Conta G, Miccheli A, Botticelli A, Paci P, Caldarelli G, Nuti M et al. Network analysis of gut microbiome and metabolome to discover microbiota-linked biomarkers in patients affected by non-small cell lung cancer. *Int. J. Mol. Sci.* 2020; **21**: 8730.

17. Deluce J, Maleki Vareki S, Fernandes R. The role of gut microbiome in immune modulation in metastatic renal cell carcinoma. *Ther. Adv. Med. Oncol.* 2022; **14**: 17588359221122714.

18. Gopalakrishnan V, Spencer CN, Nezi L, Reuben A, Andrews MC, Karpinets TV, Prieto PA, Vicente D, Hoffman K, Wei SC et al. Gut microbiome modulates response to anti–PD-1 immunotherapy in melanoma patients. *Science.* 2018; **359**: 97-103.

19. Routy B, Lenehan JG, Miller WH Jr, Jamal R, Messaoudene M, Daisley BA, Hes C, Al KF, Martinez-Gili L, Punčochář M et al. Fecal microbiota transplantation plus anti-PD-1 immunotherapy in advanced melanoma: a phase I trial. *Nat. Med.* 2023; **29**: 2121-2132.

20. Wang Y, Wiesnoski DH, Helmink BA, Gopalakrishnan V, Choi K, DuPont HL, Jiang ZD, Abu-Sbeih H, Sanchez CA, Chang CC et al. Fecal microbiota transplantation for refractory immune checkpoint inhibitor-associated colitis. *Nat. Med.* 2018; **24**: 1804-1808.

21. Zheng Y, Wang T, Tu X, Huang Y, Zhang H, Tan D, Jiang W, Cai S, Zhao P, Song R et al. Gut microbiome affects the response to anti-PD-1 immunotherapy in patients with hepatocellular carcinoma. *J. Immunother. Cancer.* 2019; **7**: 193.

22. Derosa L, Routy B, Thomas AM, Iebba V, Zalcman G, Friard S, Mazieres J, Audigier-Valette C, Moro-Sibilot D, Goldwasser F et al. Intestinal Akkermansia muciniphila predicts clinical response to PD-1 blockade in patients with advanced non-small-cell lung cancer. *Nat. Med.* 2022; **28**: 315-324.

23. Ansaldo E, Slayden LC, Ching KL, Koch MA, Wolf NK, Plichta DR, Brown EM, Graham DB, Xavier RJ, Moon JJ et al. *Akkermansia muciniphila* induces intestinal adaptive immune responses during homeostasis. *Science.* 2019; **364**: 1179-1184.

24. Yoon HS, Cho CH, Yun MS, Jang SJ, You HJ, Kim JH, Han D, Cha KH, Moon SH, Lee K, et al. *Akkermansia muciniphila* secretes a glucagon-like peptide-1-inducing protein that improves

glucose homeostasis and ameliorates metabolic disease in mice. *Nat. Microbiol.* 2021; **6**: 563-573.

25. Kim SM, Park S, Hwang SH, Lee EY, Kim JH, Lee GS, Lee G, Chang DH, Lee JG, Hwang J et al. Secreted *Akkermansia muciniphila* threonyl-tRNA synthetase functions to monitor and modulate immune homeostasis. *Cell Host Microbe.* 2023; **31**: 1021-1037.

26. Jian H, Liu Y, Wang X, Dong X, Zou X. *Akkermansia muciniphila* as a next-generation probiotic in modulating human metabolic homeostasis and disease progression: a role mediated by gut–liver–brain axes? *Int. J. Mol. Sci.* 2023; **24**, 3900.

27. Bae M, Cassilly CD, Liu X, Park SM, Tusi BK, Chen X, Kwon J, Filipčík P, Bolze AS, Liu Z et al. *Akkermansia muciniphila* phospholipid induces homeostatic immune responses. *Nature.* 2022; **608**: 168-173.

28. Guo X, Li S, Zhang J, Wu F, Li X, Wu D, Zhang M, Ou Z, Jie Z, Yan Q et al. Genome sequencing of 39 *Akkermansia muciniphila* isolates reveals its population structure, genomic and functional diversity, and global distribution in mammalian gut microbiotas. *BMC Genomics. 2017*; **18**: 1-2.

29. Becken B, Davey L, Middleton DR, Mueller KD, Sharma A, Holmes ZC, Dallow E, Remick B, Barton GM, David LA et al. Genotypic and phenotypic diversity among human isolates of *Akkermansia muciniphila*. *Mbio.* 2021; **12**: e00478-21.

30. Karcher N, Nigro E, Punčochář M, Blanco-Míguez A, Ciciani M, Manghi P, Zolfo M, Cumbo F, Manara S, Golzato D et al. Genomic diversity and ecology of human-associated *Akkermansia* species in the gut microbiome revealed by extensive metagenomic assembly. *Genome Biol.* 2021; **22**: 209.

31. Mueller KD, Panzetta ME, Davey L, McCann JR, Rawls JF, Flores GE, Valdivia RH. Pangenomic analysis identifies correlations between *Akkermansia* species and subspecies and human health outcomes. *Microbiome Research Reports.* 2024; **3**: 33.

32. Xing J, Li X, Sun Y, Zhao J, Miao S, Xiong Q, Zhang Y, Zhang G. Comparative genomic and functional analysis of *Akkermansia muciniphila* and closely related species. *Genes Genomics.* 2019; **41**, 1253-1264.

33. F. Beghini, L.J. McIver, A. Blanco-Míguez, L. Dubois, F. Asnicar, S. Maharjan, A. Mailyan, P. Manghi, M. Scholz, A.M. Thomas, M. Valles-Colomer, G. Weingart, Y. Zhang, M. Zolfo, C. Huttenhower,

E.A. Franzosa, N. Segata. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *Elife* **10**, e65088 (2021).

34. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, Mailyan A, Manghi P, Scholz M, Thomas AM et al. Host phenotype classification from human microbiome data is mainly driven by the presence of microbial taxa. *PLOS Comput. Biol.* 2022; **18**: e1010066.

35. Palmqvist S, Tideman P, Cullen N, Zetterberg H, Blennow K; Alzheimer's Disease Neuroimaging Initiative; Dage JL, Stomrud E, Janelidze S, Mattsson-Carlgren N et al. Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures. *Nat. Med.* 2021; **27**: 1034-1042.

36. Kudelka MR, Stowell SR, Cummings RD, Neish AS. Intestinal epithelial glycosylation in homeostasis and gut microbiota interactions in IBD. *Nat. Rev. Gastroenterol. Hepatol.* 2020; **17**: 597-617.

37. Moran AP, Gupta A, Joshi L. Sweet-talk: role of host glycosylation in bacterial pathogenesis of the gastrointestinal tract. *Gut.* 2011; **60**: 1412-1425.

38. Bergstrom K, Xia L. The barrier and beyond: Roles of intestinal mucus and mucin-type O-glycosylation in resistance and tolerance defense strategies guiding host-microbe symbiosis. *Gut Microbes*. 2022; **14**: 2052699.

39. Shuoker B, Pichler MJ, Jin C, Sakanaka H, Wu H, Gascueña AM, Liu J, Nielsen TS, Holgersson J, Nordberg Karlsson E et al. Sialidases and fucosidases of *Akkermansia muciniphila* are crucial for growth on mucin and nutrient sharing with mucus-associated gut bacteria. *Nat. Commun.* 2023; **14**: 1833.

40. Fidelle M, Rauber C, Alves Costa Silva C, Tian AL, Lahmar I, de La Varende AM, Zhao L, Thelemaque C, Lebhar I, Messaoudene M et al. A microbiota-modulated checkpoint directs immunosuppressive intestinal T cells into cancers. *Science.* 2023; **380**: eabo2296.

41. Park JS, Gazzaniga FS, Wu M, Luthens AK, Gillis J, Zheng W, LaFleur MW, Johnson SB, Morad G, Park EM et al. Targeting PD-L2-RGMb overcomes microbiome-related immunotherapy resistance. *Nature.* 2023; **617**: 377-385.

42. Sivan A, Corrales L, Hubert N, Williams JB, Aquino-Michaels K, Earley ZM, Benyamin FW, Lei YM, Jabri B, Alegre ML et al. Commensal Bifidobacterium promotes antitumor immunity and facilitates anti-PD-L1 efficacy. *Science.* 2015; **350**: 1084-1089.

43. Mao J, Wang D, Long J, Yang X, Lin J, Song Y, Xie F, Xun Z, Wang Y, Wang Y et al. Gut microbiome is associated with the clinical response to anti-PD-1 based immunotherapy in hepatobiliary cancers. *J. Immunother. Cancer.* 2021; **9**: e003334.

44. Depommier C, Everard A, Druart C, Plovier H, Van Hul M, Vieira-Silva S, Falony G, Raes J, Maiter D, Delzenne NM et al. Supplementation with *Akkermansia muciniphila* in overweight and obese human volunteers: a proof-of-concept exploratory study. *Nat. Med.* 2019; **25**: 1096-1103.

45. Forslund K, Hildebrand F, Nielsen T, Falony G, Le Chatelier E, Sunagawa S, Prifti E, Vieira-Silva S, Gudmundsdottir V, Pedersen HK et al. Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature.* 2015; **528**: 262-266.

46. Zhang T, Ji X, Lu G, Zhang F. The potential of *Akkermansia muciniphila* in inflammatory bowel disease. *Appl. Microbiol. Biotechnol.* 2021; **105**, 5785-5794.

47. Liu R, Zou Y, Wang WQ, Chen JH, Zhang L, Feng J, Yin JY, Mao XY, Li Q, Luo ZY et al. Gut microbial structural variation associates with immune checkpoint inhibitor response. *Nat. Commun.* 2023; **14**: 7421.

Acknowledgements

We thank Edward Mead, Yanchun Zhang and other members in the Fang Lab for their helpful feedback. This work was supported by R35 GM139655 (G.F.) from the National Institutes of Health. G.F. is a Hirschl Research Scholar by Irma T. Hirschl/Monique Weill-Caulier Trust, and a Nash Family Research Scholar. This work was also supported in part through the computational resources and staff expertise provided by the Department of Scientific Computing at the Icahn School of Medicine at Mount Sinai.

Author Contributions

P.F. performed all the data analyses. P.F. and G.F. designed the meta-analyses across the three cohorts. M.N., Y.F. and M.K. assisted in the Akk pangenome analysis. G.F. conceived and supervised the project. P.F. and G.F. wrote the manuscript with input and comments from all co-authors.

Competing Interests

The authors declare no competing interests.

Figure 1

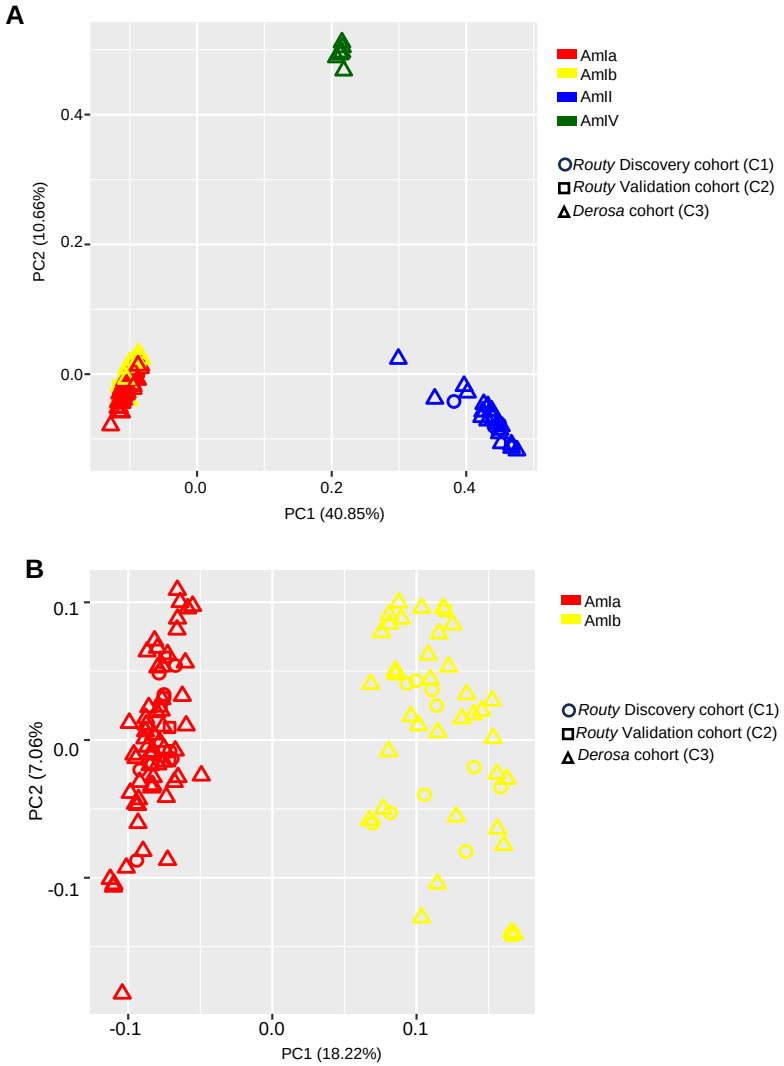

A
0.4
0.2
0.0
PC2 (10.66%)
0.0
0.2
0.4
PC1 (40.85%)
AmIa
AmIb
AmII
AmIV
Routy Discovery cohort (C1)
Routy Validation cohort (C2)
Derosa cohort (C3)
B
0.1
0.0
-0.1
PC2 (7.06%
-0.1
0.0
0.1
PC1 (18.22%)
AmIa
AmIb
Routy Discovery cohort (C1)
Routy Validation cohort (C2)
Derosa cohort (C3)

Figure 2

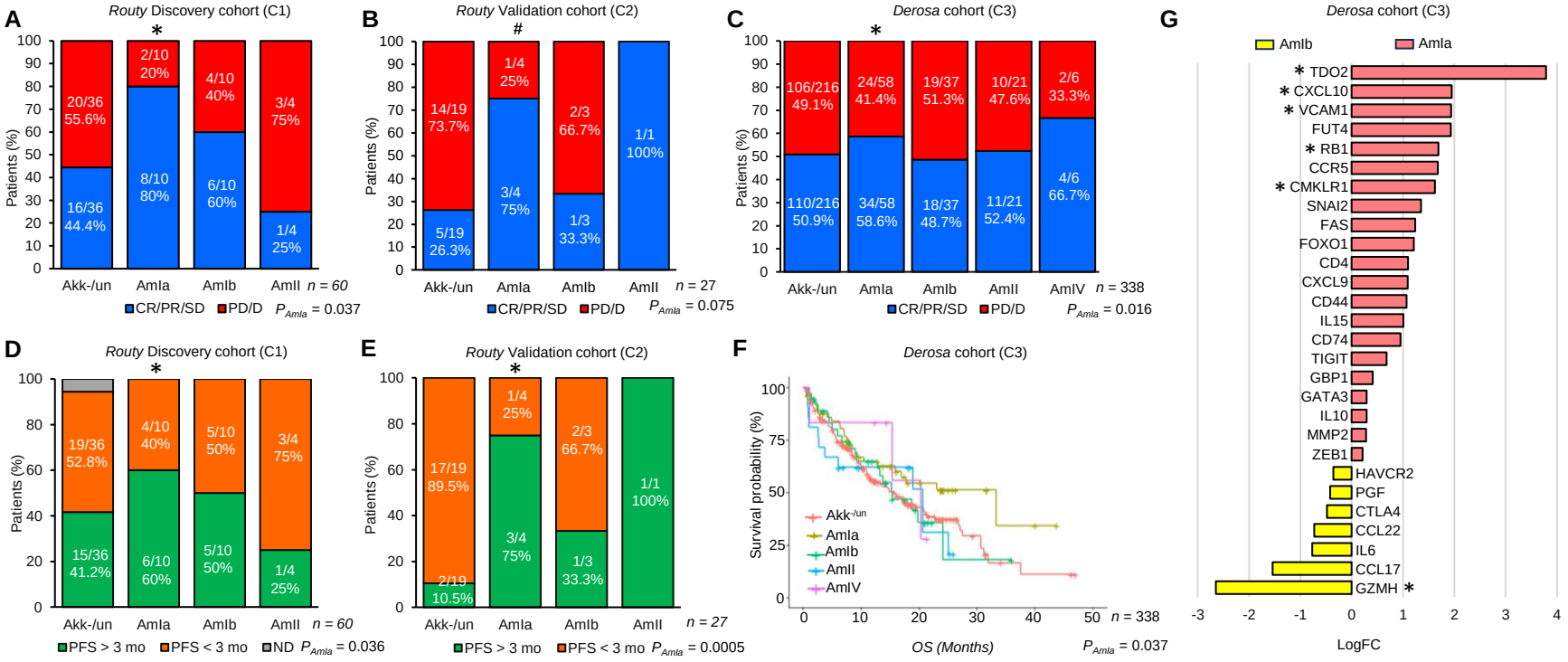

A
Routy Discovery cohort (C1)
Patients (%)
Akk-/un
Amla
Amlb
AmlI
n = 60
CR/PR/SD
PD/D
P_Amla = 0.037
B
Routy Validation cohort (C2)
n = 27
P_Amla = 0.075
C
Derosa cohort (C3)
AmlV
n = 338
P_Amla = 0.016
D
PFS > 3 mo
PFS < 3 mo
ND
P_Amla = 0.036
E
P_Amla = 0.0005
F
Survival probability (%)
OS (Months)
P_Amla = 0.037
G
Derosa cohort (C3)
LogFC
TDO2
CXCL10
VCAM1
FUT4
RB1
CCR5
CMKLR1
SNAI2
FAS
FOXO1
CD4
CXCL9
CD44
IL15
CD74
TIGIT
GBP1
GATA3
IL10
MMP2
ZEB1
HAVCR2
PGF
CTLA4
CCL22
IL6
CCL17
GZMH

Figure 3

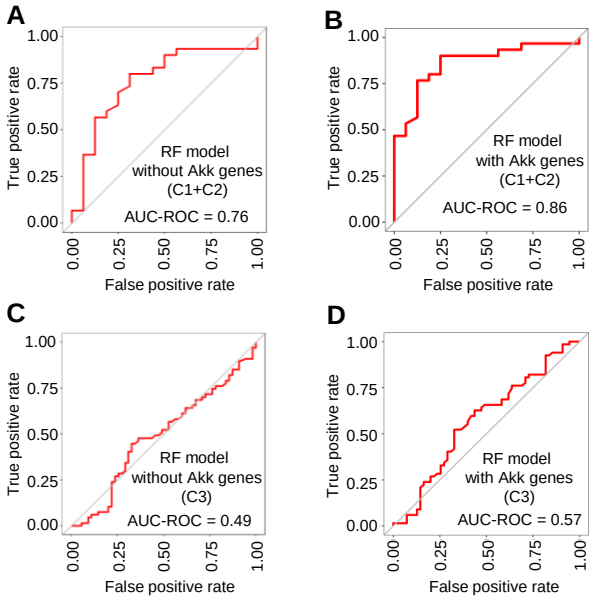
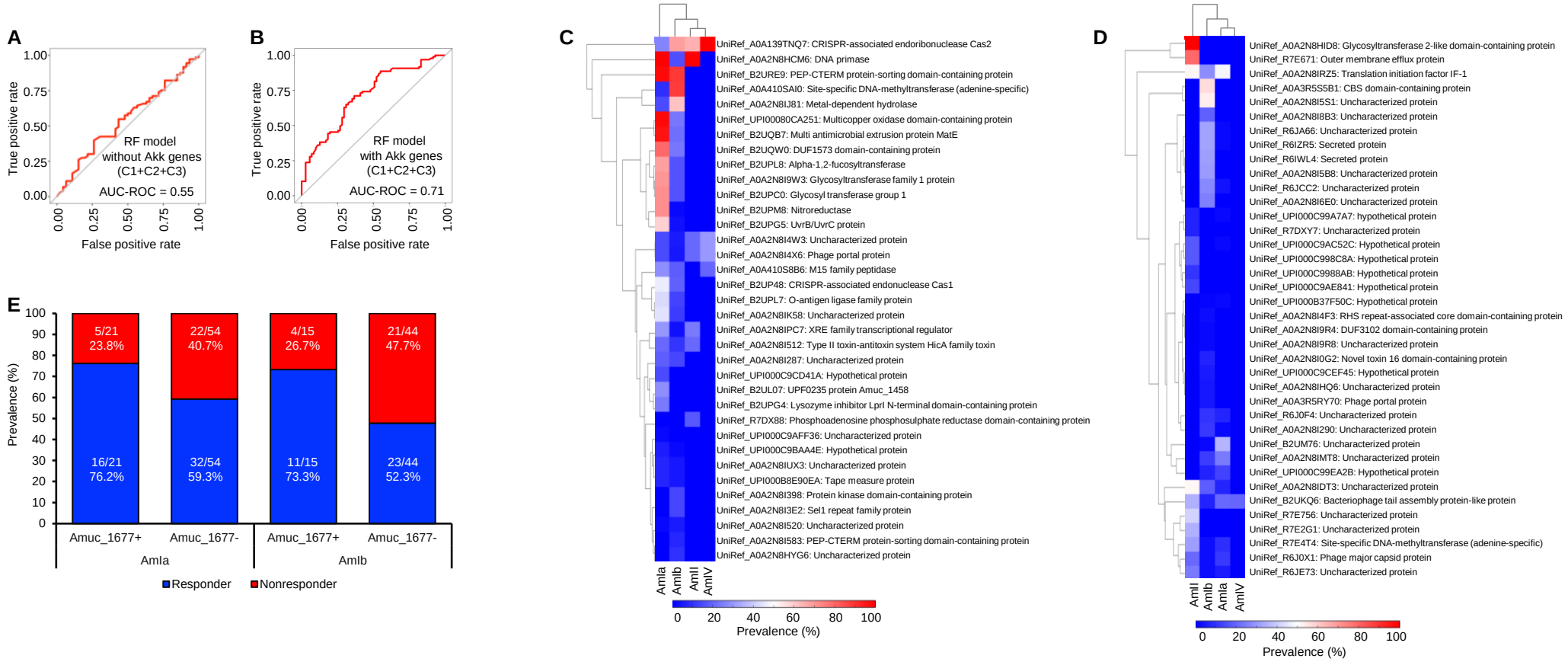

A
B
C
D
True positive rate
False positive rate
0.00
0.25
0.50
0.75
1.00
RF model
without Akk genes
(C1+C2)
AUC-ROC = 0.76
RF model
with Akk genes
(C1+C2)
AUC-ROC = 0.86
RF model
without Akk genes
(C3)
AUC-ROC = 0.49
RF model
with Akk genes
(C3)
AUC-ROC = 0.57

Figure 4


A
True positive rate
False positive rate
RF model
without Akk genes
(C1+C2+C3)
AUC-ROC = 0.55
B
True positive rate
False positive rate
RF model
with Akk genes
(C1+C2+C3)
AUC-ROC = 0.71
C
UniRef_A0A139TNQ7: CRISPR-associated endoribonuclease Cas2
UniRef_A0A2N8HCM6: DNA primase
UniRef_B2URE9: PEP-CTERM protein-sorting domain-containing protein
UniRef_A0A410SAI0: Site-specific DNA-methyltransferase (adenine-specific)
UniRef_A0A2N8IJ81: Metal-dependent hydrolase
UniRef_UPI00080CA251: Multicopper oxidase domain-containing protein
UniRef_B2UQB7: Multi antimicrobial extrusion protein MatE
UniRef_B2UQW0: DUF1573 domain-containing protein
UniRef_B2UPL8: Alpha-1,2-fucosyltransferase
UniRef_A0A2N8I9W3: Glycosyltransferase family 1 protein
UniRef_B2UPC0: Glycosyl transferase group 1
UniRef_B2UPM8: Nitroreductase
UniRef_B2UPG5: UvrB/UvrC protein
UniRef_A0A2N8I4W3: Uncharacterized protein
UniRef_A0A2N8I4X6: Phage portal protein
UniRef_A0A410S8B6: M15 family peptidase
UniRef_B2UP48: CRISPR-associated endonuclease Cas1
UniRef_B2UPL7: O-antigen ligase family protein
UniRef_A0A2N8IK58: Uncharacterized protein
UniRef_A0A2N8IPC7: XRE family transcriptional regulator
UniRef_A0A2N8I512: Type II toxin-antitoxin system HicA family toxin
UniRef_A0A2N8I287: Uncharacterized protein
UniRef_UPI000C9CD41A: Hypothetical protein
UniRef_B2UL07: UPF0235 protein Amuc_1458
UniRef_B2UPG4: Lysozyme inhibitor LprI N-terminal domain-containing protein
UniRef_R7DX88: Phosphoadenosine phosphosulphate reductase domain-containing protein
UniRef_UPI000C9AFF36: Uncharacterized protein
UniRef_UPI000C9BAA4E: Hypothetical protein
UniRef_A0A2N8IUX3: Uncharacterized protein
UniRef_UPI000B8E90EA: Tape measure protein
UniRef_A0A2N8I398: Protein kinase domain-containing protein
UniRef_A0A2N8I3E2: Sel1 repeat family protein
UniRef_A0A2N8I520: Uncharacterized protein
UniRef_A0A2N8I583: PEP-CTERM protein-sorting domain-containing protein
UniRef_A0A2N8HYG6: Uncharacterized protein
AmIa
AmIb
AmII
AmIV
Prevalence (%)
D
UniRef_A0A2N8HID8: Glycosyltransferase 2-like domain-containing protein
UniRef_R7E671: Outer membrane efflux protein
UniRef_A0A2N8IRZ5: Translation initiation factor IF-1
UniRef_A0A3R5S5B1: CBS domain-containing protein
UniRef_A0A2N8I5S1: Uncharacterized protein
UniRef_A0A2N8I8B3: Uncharacterized protein
UniRef_R6JA66: Uncharacterized protein
UniRef_R6IZR5: Secreted protein
UniRef_R6IWL4: Secreted protein
UniRef_A0A2N8I5B8: Uncharacterized protein
UniRef_R6JCC2: Uncharacterized protein
UniRef_A0A2N8I6E0: Uncharacterized protein
UniRef_UPI000C99A7A7: hypothetical protein
UniRef_R7DXY7: Uncharacterized protein
UniRef_UPI000C9AC52C: Hypothetical protein
UniRef_UPI000C998C8A: Hypothetical protein
UniRef_UPI000C9988AB: Hypothetical protein
UniRef_UPI000C9AE841: Hypothetical protein
UniRef_UPI000B37F50C: Hypothetical protein
UniRef_A0A2N8I4F3: RHS repeat-associated core domain-containing protein
UniRef_A0A2N8I9R4: DUF3102 domain-containing protein
UniRef_A0A2N8I9R8: Uncharacterized protein
UniRef_A0A2N8I0G2: Novel toxin 16 domain-containing protein
UniRef_UPI000C9CEF45: Hypothetical protein
UniRef_A0A2N8IHQ6: Uncharacterized protein
UniRef_A0A3R5RY70: Phage portal protein
UniRef_R6J0F4: Uncharacterized protein
UniRef_A0A2N8I290: Uncharacterized protein
UniRef_B2UM76: Uncharacterized protein
UniRef_A0A2N8IMT8: Uncharacterized protein
UniRef_UPI000C99EA2B: Hypothetical protein
UniRef_A0A2N8IDT3: Uncharacterized protein
UniRef_B2UKQ6: Bacteriophage tail assembly protein-like protein
UniRef_R7E756: Uncharacterized protein
UniRef_R7E2G1: Uncharacterized protein
UniRef_R7E4T4: Site-specific DNA-methyltransferase (adenine-specific)
UniRef_R6J0X1: Phage major capsid protein
UniRef_R6JE73: Uncharacterized protein
AmII
AmIb
AmIa
AmIV
Prevalence (%)
E
Prevalence (%)
5/21
23.8%
16/21
76.2%
22/54
40.7%
32/54
59.3%
4/15
26.7%
11/15
73.3%
21/44
47.7%
23/44
52.3%
Amuc_1677+
Amuc_1677-
AmIa
Amuc_1677+
Amuc_1677-
AmIb
Responder
Nonresponder