ABSTRACT

In lung cancer, enrichment of the lower airway microbiota with oral commensals commonly occurs, and ex vivo models support that some of these bacteria can trigger host transcriptomic signatures associated with carcinogenesis. Here, we show that this lower airway dysbiotic signature was more prevalent in the stage IIIB–IV tumor–node–metastasis lung cancer group and is associated with poor prognosis, as shown by decreased survival among subjects with early-stage disease (I–IIIA) and worse tumor progression as measured by RECIST scores among subjects with stage IIIB–IV disease. In addition, this lower airway microbiota signature was associated with upregulation of the IL17, PI3K, MAPK, and ERK pathways in airway transcriptome, and we identified Veillonella parvula as the most abundant taxon driving this association. In a KP lung cancer model, lower airway dysbiosis with V. parvula led to decreased survival, increased tumor burden, IL17 inflammatory phenotype, and activation of checkpoint inhibitor markers.

SIGNIFICANCE: Multiple lines of investigation have shown that the gut microbiota affects host immune response to immunotherapy in cancer. Here, we support that the local airway microbiota modulates the host immune tone in lung cancer, affecting tumor progression and prognosis.

See related commentary by Zitvogel and Kroemer, p. 224.
INTRODUCTION

Lung cancer has remained the leading cause of cancer death worldwide. In this past year alone, lung cancer occurred in approximately 2.1 million patients and was responsible for 1.7 million deaths (1). Targeting certain somatic mutations has improved survival, but this is applicable to only ~30% of subjects with lung adenocarcinoma (2, 3). More recently, immunotherapy that targets inhibitory checkpoint molecules, such as programmed death 1 (PD-1), has been shown to affect the responses of T cells to neoantigens and improve survival in lung cancer (4–8). However, 40% to 60% of patients will not respond to or will develop resistance to immunotherapy (7).

Recent investigations have identified gut microbiota signatures that are associated with augmenting antitumor immunity and responding to PD-1 blockade in murine models and in prospective analyses of immunotherapy-responsive cancer cohorts (9–11). For example, modulation of the microbiota in germ-free mice can enhance antitumor immunity and augment effects of checkpoint blockade (12, 13). Matson and colleagues found that in patients with melanoma, anti–PD-1 treatment responders had a higher abundance of B. longum, C. aerofaciens, and E. faecium compared with nonresponders (11). Gopalakrishnan and colleagues demonstrated that patients with higher bacterial diversity and increased relative abundance of Ruminococcaceae in the gut had enhanced systemic and antitumor immune responses (10). Routy and colleagues identified that the relative abundance of A. muciniphila was associated with a favorable clinical response to immunotherapy (9). Although most investigations have focused on the gut microbiome, no human studies have studied the lower airway microbiota and lung cancer prognosis despite growing evidence supporting the role of the lung microbiota in lower airway inflammation (14–16).

Our understanding of the role of lung microbiota in health and disease is rapidly evolving with evidence that some phenotypic characteristics of the local lung immune tone appear to be more closely correlated to the lung microbiome than to the gut microbiome (14). Culture-independent techniques show that the lower airways of normal individuals commonly harbor oral bacteria such as Prevotella and Veillonella (15, 17–19). Our group has described that lower airway dysbiosis is associated with increased host inflammatory tone in the lung of healthy individuals (15, 19). This same lower airway dysbiotic signature was found to differentiate between subjects with lung cancer and subjects with benign lung nodules (16). Importantly, we have shown in humans and in ex vivo experimental models that this dysbiotic signature likely triggers transcriptomic signatures (PI3K and MAPK) previously reported in non–small cell lung cancer (NSCLC; refs. 16, 20), including the p53 mutation pathway (21). In order to explore the clinical implications of the lower airway microbiota in lung cancer, we utilized a prospective human cohort and a preclinical model to identify lower airway dysbiotic signatures that may affect the prognosis in this disease.

RESULTS

Lung Cancer Cohort

Between March 2013 and October 2018, we recruited 148 subjects with lung nodules from the NYU Lung Cancer Biomarker Center who underwent clinical bronchoscopy for diagnostic purposes and in whom lower airway brushes were obtained for research (Supplementary Fig. S1). Fifteen subjects had non–lung primary tumors (metastasis), 12 had benign lung nodules, and 38 subjects had other nonmalignant diagnosis and were excluded. The remaining 83 subjects had a final diagnosis of lung cancer and were included for this project. Among these subjects, all had microbiome 16S rRNA gene-sequencing data, 70/83 had transcriptomic data, and 75/83 had greater than six months of follow-up clinical data. Supplementary Table S1 describes the demographics and clinical characteristics of this cohort: 91% were current or former smokers with a mean history of 46 pack-years. Eighty-nine percent had a diagnosis of NSCLC, of which 65% had adenocarcinoma and 49% were found to have stage IIIB–IV. The median survival was 2.1 years; 54% received chemotherapy, 30% received radiotherapy, 24% received surgery, and 14% received immunotherapy. All biospecimens were obtained prior to treatment. Using the Cox proportional hazards model, we determined that surgical treatment and stage IIIB–IV were significantly associated with overall survival (OS; Supplementary Table S2).

Microbiomic Signatures Associated with Stage and Prognosis

In addition to lower airway brushings, we obtained buccal brushes and bronchoscope background control samples that were included in the 16S rRNA gene-sequencing analysis. As compared with background controls, the bacterial load was ~10 times higher in lower airway brushing samples and ~10,000 times higher in the upper airways (buccal; P < 0.001; Supplementary Fig. S2). Alpha diversity based on the Shannon index showed greater diversity among lower airway samples than upper airway and background control samples (P < 0.001; Supplementary Fig. S3A). Principal coordinate analysis (PCoA) based on Bray–Curtis dissimilarity index showed significant compositional differences across sample types (Fig. 1A; PERMANOVA P < 0.001). Across lower airway samples, there were also compositional differences between small cell lung cancer (SCLC) and NSCLC (PERMANOVA P = 0.01). Among NSCLC samples, there were no statistically significant differences in α-diversity and β-diversity between squamous cell carcinoma and adenocarcinoma.

We then evaluated microbial differences in lower airway samples based on the clinical NSCLC stage, grouped as I–IIIA and IIIB–IV of tumor–node–metastasis (TNM) classification. The selection of this cutoff point for TNM classification allowed for dichotomized analyses, and we support it based on prior prognosis/survival data and cancer management guidelines related to surgical management of NSCLC (22–25). Alpha diversity was similar across staging groups of NSCLC (Supplementary Fig. S3B for comparison across individual stages and Supplementary Fig. S3C for two-group comparisons of stages IIIB–IV vs. I–IIIA). Compositional differences between the I–IIIA and IIIB–IV lung cancer groups were noted based on β-diversity analysis (Fig. 1B, left; P = 0.005), where stage IIIB–IV lung cancer was compositionally more similar to buccal samples than I–IIIA stage lung cancer samples (Fig. 1B, right). Compositional differences comparing all individual stages (I–IV) were also noted based...
Lung Microbiome and Lung Cancer Prognosis

Figure 1. Lung microbiota in lung cancer and cancer survival. A, PCoA of airway samples shows a difference in β-diversity (P = 0.01, PERMANOVA) between small cell lung cancer and NSCLC (n = 83). B, Among patients with NSCLC (n = 74), PCoA shows a difference in β-diversity (P = 0.005, PERMANOVA) between stages IIIB–IV and I–IIIA NSCLC (left); lower airway microbiota of stages IIIB–IV were more similar to buccal microbiota than lower airway microbiota of stages I–IIIA (right; P < 0.0001, Bray–Curtis distance). C, Left, PCoA based on cancer stage and survival at six months and one year shows difference in β-diversity (P < 0.05, PERMANOVA). C, Right, lower airway microbiota in lung cancer and worse survival at six months or one year were more similar to buccal microbiota than with better survival in both stages IIIB–IV (n = 36) and I–IIIA (n = 37) groups (P < 0.05, Bray–Curtis distance). D, Stage IIIB–IV lung cancer was associated with having a higher proportion of subjects whose lower airway microbiota were classified as enriched with oral taxa (supraglottic predominant taxa, SPT) vs. background taxa (background predominant taxa, BPT), P = 0.006. E, Enrichment of the lower airway with pneumotypeP was associated with better survival in stage I–IIIA cancer than enrichment with pneumotypeP, P < 0.05; there was no difference in stage IIIB–IV cancer. F, Bray–Curtis dissimilarity index between lower airway and buccal samples was inversely associated with delta RECIST score for stage IIIB–IV NSCLC measured at 6 to 12 months (Spearman r = −0.48, P = 0.03).
on β-diversity analysis (Supplementary Fig. S4A; P = 0.047), where lower airway samples from more advanced stages had a greater similarity to buccal samples than lower airway samples from earlier-stage subjects (Supplementary Fig. S4B). Microbiome regression-based kernel association test (MiRKAT) analysis showed that differences noted in the microbial community profiles between stage I–IIIA and IIIB–IV NSCLC were not due to differences in location of the samples. Interestingly, subanalysis on patient samples where tumor PD-L1 expression was available (n = 39) shows that subjects with high PD-L1 expression (≥80%, n = 12) had a lower airway microbiota with greater similarity to upper airway microbiota versus the disease of similarity found among patients with lower tumor PD-L1 expression (0%, n = 16; 1%–79%, n = 11; P < 0.05; Supplementary Fig. S5).

Compositional differences based on six-month and one-year survival were also identified in β-diversity analysis (Fig. 1C, left), where samples from subjects with decreased survival were associated with greater compositional similarity to buccal samples than samples from subjects with better outcomes (Fig. 1C, right). Shannon index showed decreased α-diversity among samples from subjects with <6-month survival in both stages I–IIIA and IIIB–IV, but this difference was not statistically significant at one year (Supplementary Fig. S6). Multivariate PERMANOVA analysis demonstrated that the association between microbial community composition and six-month/one-year mortality was independent of TNM staging (Supplementary Fig. S7). No statistically significant differences were noted in α- or β-diversity analyses of buccal microbiota between subjects with different stages or mortality.

DESeq analyses were then performed to evaluate for taxonomic differential enrichment between SCLC and NSCLC and between the I–IIIA and IIIB–IV groups of NSCLC (Supplementary Fig. S8A). Importantly, lower airway samples from patients in the IIIB–IV stage group were enriched with many operational taxonomic units (OTU), which annotated to the genera Moraxella, Fusobacterium, Pseudomonas, and Haemophilus, and were decreased in abundance of Actinomyces (Supplementary Fig. S8B; Supplementary Table S1). Using a mixed-effect model that adjusts for sample location, we report the top 20 OTUs ranked by their absolute coefficients estimates as having a differential abundance between the stage I–IIIA and IIIB–IV groups (Supplementary Table S3). Once again, stage IIIB–IV lung cancer was enriched with OTUs recognized as oral commensals, such as Haemophilus, Fusobacterium, Gemella, Prevotella, and Granulicatella.

Among stage I–IIIA and IIIB–IV subgroups, multiple OTUs were differentially enriched when worse versus better survival groups were compared (both at 6 and 12 months). Several of the OTUs annotated to the genera Veillonella, Prevotella, and Streptococcus were found to be enriched in samples from subjects with worse prognosis (Supplementary Fig. S9A–S9D; Supplementary Tables S2–S5). In order to further explore for taxonomic associations with mortality while considering TNM staging, we constructed β-diversity biplots that allowed for colocation of lower airway samples and taxa driving the spatial distribution. Using a multivariate analysis adjusted by TNM stage, Supplementary Fig. S10 shows that poor prognosis was associated with enrichment of the lower airway microbiota with oral commensals (such as Streptococcus, Prevotella, and Veillonella). When analysis was repeated only considering the lower airway samples with closest proximity to the cancer, similar results were found (Supplementary Fig. S11).

Using a mixed-effect model adjusted by smoking status, stage (I–IIIA/IIIB–IV), and treatment type, we identified top OTUs associated with OS. Supplementary Table S4 reports the top 20 OTUs ranked by absolute coefficient estimates associated with OS. Poor prognosis was associated with enrichment with OTUs recognized as oral commensals that belong to the genera Prevotella, Streptococcus, Lactobacillus, and Gemella.

Utilizing a Dirichlet multinomial mixture (DMM) model, we established that samples can be divided into two clusters: cluster one consists of all the upper airway samples and ~60% of lower airway samples and cluster two consists of all the bronchoscope background control samples and ~40% of the lower airway samples (Supplementary Fig. S12A and S12B). Thus, similar to previously published data (15), our cohort consists of one cluster of lower airway samples enriched with background predominant taxa (BPT), such as Flavobacterium and Pseudomonas, whereas the second cluster was enriched with supraglottic predominant taxa (SPT), such as Veillonella, Streptococcus, Prevotella, and Haemophilus (Supplementary Fig. S12C; Supplementary Table S6). Supplementary Table S5 shows that we did not identify statistically significant differences in demographic or clinical characteristics, other than stage IV TNM staging (P < 0.05), between subjects with a lower airway microbiota that clustered as BPT versus SPT. Applying decontam (26) approach to these data, an analytic pipeline that accounts for taxa most likely to be contaminants, we identified Flavobacterium as a background contaminant (also most prevalent and abundant OTU in background controls) and oral commensals, such as Veillonella and Streptococcus, as most representatives of lower airway microbiota (Supplementary Fig. S13).

We then used the DMM grouping to evaluate whether the prevalence of SPT/BPT was different among stage I–IIIA and IIIB–IV NSCLC and/or associated with prognosis. The percentage of SPT was higher in lower airway samples from subjects from the stage IIIB–IV NSCLC group compared with lower airway samples from the stage I–IIIA NSCLC group (Fig. 1D; P = 0.006). Importantly, the Kaplan–Meier survival analysis shows that among subjects with stage I–IIIA NSCLC, the SPT pneumotype was associated with worse survival than the BPT pneumotype (Fig. 1E, P = 0.047). In stage IIIB–IV NSCLC, there were no statistically significant differences in survival between the SPT and BPT pneumotypes, although the overall mortality was much worse, with a median survival of less than one year as found in the above analysis. To further evaluate microbial signatures associated with treatment response, we analyzed a subset of patients with stage IIIB–IV NSCLC (thus nonsurgical) with available longitudinal imaging which allowed us to calculate the Response Evaluation Criteria in Solid Tumors (RECIST; ref. 27). Correlation analysis between delta RECIST score and β-diversity dissimilarity between upper and lower airways showed a significant inverse correlation (Fig. 1F; Spearman r = −0.48, P = 0.03). Thus, although overall mortality was not associated with pneumotype categorization in the stage IIIB–IV group, having a positive delta RECIST score, indicating tumor progression, was associated with having
a lower airway microbiota more similar to that of upper airways. Taxonomic differences between a dichotomized RECIST score showed lower airway samples from patients with tumor progression (RECIST = progressive disease or stable disease) were enriched with Veillonella, Streptococcus, Prevotella, and Rothia when compared with lower airway samples from patients with tumor regression (RECIST = complete response or partial response; Supplementary Fig. S14 and Supplementary Table S7).

Transcriptomic Signatures Associated with Stage, Prognosis, and Microbiota

After quality control, RNA-sequencing (RNA-seq) data were obtained on 70 lower airway samples from 70 subjects with NSCLC. We then compared global transcriptomic differences between stage I–IIIA and IIIB–IV NSCLC with PCoA based on the Bray-Curtis dissimilarity index. In contrast to microbiota data, there were no statistically significant differences in β-diversity between these two groups. DESeq analysis showed that there were only 20 genes differentially regulated in stage IIIB–IV compared with stage I–IIIA NSCLC (Supplementary Fig. S15; Supplementary Table S8). Similarly, very few transcripts were found differentially expressed when comparing better versus worse outcomes at six-month and one-year survival (Supplementary Table S8).

We then used DESeq to compare transcriptomic signatures associated with a distinct lower airway microbiota based on DMM and found that there were 209 genes upregulated and 88 genes downregulated in airway brushes based on DMM and found that there were 209 genes upregulated and 88 genes downregulated in airway brushes of subjects with SPT lower airway microbiota versus BPT lower airway microbiota (Fig. 2A; Supplementary Table S9, FDR < 0.25). Subanalysis of the transcriptomic data among stage I–IIIA and IIIB–IV NSCLC showed the most significant differences for SPT versus BPT within stage I–IIIA NSCLC. Functional enrichment analysis [Ingenuity Pathway Analysis (IPA)] of differentially expressed genes between SPT and BPT (all samples or stage I–IIIA NSCLC samples) showed that SPT was associated with upregulation of the following top canonical pathways: p53 mutation, PI3K/PTEN, ERK, and IL6/IL8 (Fig. 2B).

Multiomic Analysis

To better characterize host/microbe interaction in lung cancer, we used a multiomic analytic framework that evaluates for associations between co-occurring taxa and host RNA transcriptome signatures. We estimated co-occurrence probabilities between taxa and the host transcripts adding the probability ranks for the taxa being associated with stage I–IIIA or IIIB–IV lung cancer using MMvec (27, 28). Based on the predicted microbe–transcript co-occurrences, there were two distinct clusters of taxa (Fig. 2C; interactive figure available at https://segalmicrobiomelab.github.io). The first cluster consisted of SPT-associated taxa (belonging to the genera Veillonella, Prevotella, and Streptococcus) that had a high probability of being observed in subjects with stage I–IIIA NSCLC; however, it is important to note that many of the high abundant genera in this cluster (stages I–IIIA) likely represent background taxa as identified by decontam (Supplementary Fig. S13) and not true lower airway taxa. Among SPT-associated taxa, a Veillonella taxon (OTU#85419) had the highest relative abundance and a high probability of being found in subjects with stage IIIB–IV lung cancer. This taxon was also highly associated with cell adhesion molecules, IL17, cytokines and growth factors, chemokine signaling pathway, TNF, JAK–STAT, and PI3K–AKT signaling pathway (Supplementary Table S10). Using BLAST (28), the sequence of this OTU most closely aligned with Veillonella parvula.

Lung Dysbiosis Preclinical Model

To evaluate the causal effects of lower airway dysbiosis on lung cancer progression, we tested the effects of lower airway dysbiosis induced by Veillonella parvula in a preclinical lung cancer model (KP mice; Fig. 3A). We selected this bacterium because we have found it to be a good marker for SPT, it was consistently associated with NSCLC (16), and it was the taxa with the highest relative abundance identified in our multiomic analysis as associated with stage IIIB–IV and transcriptomic signatures. Of note, lower airway dysbiosis induced by other oral commensals, such as Streptococcus mitis and Prevotella melaninogenica, also led to increased lower airway inflammation but at a lesser degree than V. parvula (Supplementary Figs. S16 and S17A–S17B). Thus, as a proof of concept, we chose Veillonella parvula as our lower airway dysbiosis model for the KP lung cancer mice.

Dysbiosis was induced once KP seeding was determined. Induction of lower airway dysbiosis with V. parvula in wild-type (WT) mice did not affect the mice’s survival or weight gain. In contrast, within KP lung cancer mice, exposure to dysbiosis (KP + Dys) led to decreased survival, weight loss, and increased tumor burden (Fig. 3A and B; Supplementary Fig. S18A and S18B). The experiment was repeated at an early sac time point (three weeks post induction of dysbiosis) to evaluate the immune response to dysbiosis with host transcriptomics, T-cell profiling, and cytokine measurements. PCoA of host transcriptomics showed clear differences between the four experimental conditions, where dysbiosis led to greater compositional changes than lung cancer alone (Supplementary Fig. S19A). Characterization of immune cell subsets inferred from bulk transcriptomics (CIBERSORT) identified clear clustering by condition where lower airway dysbiosis led to an increase in Th1 cells and activation of dendritic cells (Supplementary Fig. S19B). IPA showed that dysbiosis led to upregulation of the PI3K/AKT, ERK/MAPK, IL17A, IL6/IL8, and inflammasome pathways (Fig. 3C). Comparisons between transcriptomic signatures induced by lower airway dysbiosis in the NSCLC mouse model and those identified in SPT among subjects with NSCLC showed concordant signatures related to IL17 signaling, chemokine, Toll-like receptor, PD-L1 signaling, and PI3K–AKT signaling, among others (Supplementary Fig. S20A and S20B). Although there are notable differences between transcriptomic signatures in human and mouse data, these findings provide a promising direction for follow-up. Lastly, lung dysbiosis induced by V. parvula led to the recruitment of Th17 cells, with increased levels of IL17 production, increased expression of PD-1+ T cells, and recruitment of neutrophils (Fig. 3D; Supplementary
Figure 2. Airway transcriptome in NSCLC based on lung microbiota. Comparisons between microbiome and host transcriptomic signatures were conducted using samples where paired matched data were available (n = 70). A, Volcano plot of differentially expressed genes (FDR < 0.25) between pneumotypePRT vs. pneumotypePNS in all, stage I–IIIA only, or stage IIIB–IV only lower airway samples. B, Unsupervised hierarchical heat map of canonical pathway analysis based on IPA (RRID-SCR_008653) using the airway transcriptome of all subjects and those with stage I–IIIA disease comparing pneumotypePRT vs. pneumotypePNS groups. Subanalysis using samples from patients with stage IIIB–IV disease is not presented given the paucity of differentially expressed genes between the groups. Orange shows upregulation of pathway, and blue shows downregulation of pathway. C, Network analysis based on conditional co-occurrence probability of microbiome and transcriptome data; microbiome nodes (circles) are colored red for stage IIIB–IV lung cancer and green for stage I–IIIA lung cancer (based on a gradient) and sized by relative abundance. Edges connect microbiome nodes to pathway nodes, and edge width is based on their conditional probability.

Fig. S21). Spatial analysis with IHC targeting CD4+ T cells, CD8+ T cells, and neutrophils shows that the increase of these inflammatory cells in response to dysbiosis occurred predominantly in tumor-spared lung tissue (Fig. 3E; Supplementary Fig. S22A). Interestingly, in the tumor there was a decrease in CD4+ T cells associated with lower airway dysbiosis.

To further assess the functional importance of dysbiotic-induced IL17 activation in lung tumorigenesis, dysbiotic-KP mice were treated with monoclonal antibodies against IL17 or isotype antibody control for two weeks after tumor initiation (Fig. 4A). Tumor luminescence data showed that IL17 blockade led to a decrease in tumor burden over the second week.
compared with isotype control (P = 0.0059; Fig. 4B). Immune profiling evaluated at day 14 after IL17 blockade showed that treatment with anti-IL17 antibodies was associated with decreased RORγt+ CD4+ T cells and neutrophils, and a non-statistically significant trend toward lower IL17+ CD4+ and IL17+ TCRγδ+ T cells (Fig. 4C). Histologic assessment with IHC shows that IL17 blockade led to a decrease in CD4+ T cells, CD8+ T cells, and neutrophils in the spared nontumor lung tissue but not in the tumor itself (Fig. 4D; Supplementary Fig. S22B). Overall, these data suggest that lower airway dysbiosis in lung cancer and cancer survival. A, Experimental condition and Kaplan–Meier survival showing decreased survival in mice with lung cancer and dysbiosis (LC + Dys, n = 22) compared with LC (n = 20) alone (P < 0.001). Dys did not affect mouse survival in WT control (n = 10 for each group). B, Quantitative data of tumor burden (measured as lumens prior to death or sacrifice normalized to baseline lumens) showing that LC + Dys mice had increased tumor burden (P < 0.05, n = 5 for each experimental condition).

C, IPA was used to identify dysregulated transcriptomic pathways. D, Immune profiling of lung tissue by FACS and cytokine measurement demonstrates that lower airway dysbiosis induces Th17 and PD-1 T-cell phenotype in the lung. E, IHC analysis comparing LC and LC + dys shows increase in CD4+ T cells, CD8+ T cells, and neutrophils in the non-tumor region after dysbiosis. Minimal difference in immune response was seen within the tumor itself [n = 4 (LC) vs. n = 8 (LC + dys) mice/group; each dot represents different regions analyzed color coded by mice].

Figure 3. Preclinical model of lung dysbiosis in lung cancer and cancer survival. A, Experimental condition and Kaplan–Meier survival showing decreased survival in mice with lung cancer and dysbiosis (LC + Dys, n = 22) compared with LC (n = 20) alone (P < 0.001). Dys did not affect mouse survival in WT control (n = 10 for each group). B, Quantitative data of tumor burden (measured as lumens prior to death or sacrifice normalized to baseline lumens) showing that LC + Dys mice had increased tumor burden (P < 0.05, n = 5 for each experimental condition). C, IPA was used to identify dysregulated transcriptomic pathways. D, Immune profiling of lung tissue by FACS and cytokine measurement demonstrates that lower airway dysbiosis induces Th17 and PD-1 T-cell phenotype in the lung. E, IHC analysis comparing LC and LC + dys shows increase in CD4+ T cells, CD8+ T cells, and neutrophils in the non-tumor region after dysbiosis. Minimal difference in immune response was seen within the tumor itself [n = 4 (LC) vs. n = 8 (LC + dys) mice/group; each dot represents different regions analyzed color coded by mice].
Dysbiosis contributes to a tumor-inflammatory microenvironment characterized by an increase in the Th1 and Th17 phenotypes, activation of dendritic cells with potential antigen presentation capacity, and an increase in checkpoint inhibitor markers within the surrounding lung tissue.

**DISCUSSION**

The lower airway microbiota, whether in health or disease state, are mostly affected by aspiration of oral secretions, and the lower airway microbial products are in constant interaction with the host immune system (15, 19, 29–31). In this study, we demonstrate that a lower airway dysbiotic signature present in patients with lung cancer affects tumor progression and clinical prognosis, likely due to alteration in stage I–IIIA immune tone promoting inflammation and checkpoint inhibition. First, patients with stage I–IIIA NSCLC are more likely to have enrichment of the lower airway microbiota with oral commensals compared with patients with stage I–IIIA disease. In addition, this dysbiotic signature was associated with: (i) worse outcome at six months and one year (for both stage I–IIIA and IIIB–IV groups); (ii) OS in the stage I–IIIA group; and (iii) tumor progression in stage IIIB–IV disease. Our preclinical data using an NSCLC mouse support a model in which aspiration of oral commensals (identified in our human cohort) affects the lower airway inflammatory tone and promotes tumor cell proliferation. Dysbiosis in these mice led to upregulation of the ERK/MAPK, IL1, IL6, and inflammasome signaling pathways. Immune profiling showed that lung dysbiosis led to a substantial increase in Th17 and PD-1+ cells. Previous preclinical models of cancer have shown the association between lung dysbiosis and lung inflammation but have limited human microbiome data to support clinical relevance (especially considering that the human and murine microbiota differ; refs. 32–35). Our data identified that enrichment of the lower airway microbiota with human oral commensals, such as Veillonella, contributes to a local protumor immune tone leading to progression of NSCLC, suggesting that micro-aspiration and/or impaired airway clearance likely affect the pathogenesis of this disease (36). Several lines of investigation have shown that increased inflammation and decreased immune surveillance, characterized...
by IL17 tone and checkpoint inhibition, are associated with poor prognosis in NSCLC. Increased local and systemic IL17 (37, 38), systemic IL6 (39), and higher neutrophil-to-T cell ratio (40) are associated with a poor prognosis in lung cancer. PD-L1, the ligand for PD-1, is induced in nonlymphoid cells and tumor cells under inflammatory conditions triggered by several cytokines, such as IFNγ and pathogen-associated molecular patterns (PAMP; refs. 41–43). In addition, many signaling molecules (e.g., NFkB, MAPK, PI3K, mTOR, and JAK/STAT) that affect proliferation, apoptosis, and cell survival induce PD-L1 expression (44, 45). In a transgenic mouse model expressing a conditional IL17A allele and a conditional KRASG12D, increased IL17 production was associated with accelerated lung tumor growth, decreased responsiveness to checkpoint inhibition, and decreased survival (46).

In many cancer models (breast cancer, gastric carcinoma, and lung cancer), inflammasome activation, through IL1β signaling, leads to an inflammatory response characterized by decreased antitumor immune surveillance (47–49). In the current investigation, we show that the increase in IL17 inflammatory tone triggered by lower airway dysbiosis can be blunted by anti-IL17 blocking antibodies, which seemed to lead to a decrease in the tumor burden. More experiments are obviously needed to further characterize the phenotype of inflammatory profile in the tumor and surrounding tissue, to understand the molecular mechanisms by which lower airway inflammatory cells respond to lower airway dysbiosis, and to better characterize how these factors affect tumor burden and survival. However, the above-discussed investigations and the data presented in the current paper support the balance between Th17 inflammation and immune surveillance affects NSCLC pathogenesis, and, thus, future investigations are warranted to explore the role of IL17 blockade in this disease.

Immune-checkpoint molecules, such as PD-1, mediate the response of T cells to neoantigens and are now first-line therapy for advanced NSCLC (4–8). However, 40% to 60% of patients will not benefit from these therapies, and existing biomarkers (e.g., expression of PD-1 ligand) have limited capacity to predict efficacy (7, 50). Different gut microbiota signatures have been identified as associated with augmented antitumor immune surveillance (47–49). In the current investigation, we show that the increase in IL17 inflammatory tone triggered by lower airway dysbiosis can be blunted by anti-IL17 blocking antibodies, which seemed to lead to a decrease in the tumor burden. More experiments are obviously needed to further characterize the phenotype of inflammatory profile in the tumor and surrounding tissue, to understand the molecular mechanisms by which lower airway inflammatory cells respond to lower airway dysbiosis, and to better characterize how these factors affect tumor burden and survival. However, the above-discussed investigations and the data presented in the current paper support the balance between Th17 inflammation and immune surveillance affects NSCLC pathogenesis, and, thus, future investigations are warranted to explore the role of IL17 blockade in this disease.

Immune-checkpoint molecules, such as PD-1, mediate the response of T cells to neoantigens and are now first-line therapy for advanced NSCLC (4–8). However, 40% to 60% of patients will not benefit from these therapies, and existing biomarkers (e.g., expression of PD-1 ligand) have limited capacity to predict efficacy (7, 50). Different gut microbiota signatures have been identified as associated with augmented antitumor immune surveillance and a PD-1 blockade response (9–11). In the gut, higher α-diversity and enrichment of Ruminococcaceae were associated with a favorable response to anti-PD-1 treatment in patients with melanoma (10, 51), and modulation of the microbiota in germ-free mice can enhance antitumor immunity and augment effects of checkpoint blockade (12, 13). In germ-free or antibiotic-treated mice, lung adenocarcinoma (KRAS mutation/p53 deletion) development is decreased compared with specific pathogen-free mice (32). In this model, lung microbiota activate IL1β and IL23 cytokines from myeloid cells and induce IL17-producing γδ T cells. Thus, although most studies have focused on the effect of the gut microbiome on cancer development and progression, there is increasing evidence to suggest that the local lung microbiota play a pivotal role in lung cancer pathogenesis and treatment. Multiple lines of investigation have shown that the lower airway microbiota are major determinants of the airway immune tone in health and many disease states. For example, recent preclinical models have shown that lower airway mucosal inflammation is primarily associated with the composition of the lower airway microbiota rather than the composition of the gut or upper airway microbiota (14).

In humans, we have shown that pneumotypeSPT is associated with increased local inflammatory cells and the Th17 phenotype (15, 52), and the lower airway microbial metabolism can be modulated by, for example, chronic macrolide therapy, leading to release of microbial metabolites with anti-inflammatory effects (53, 54). Anaerobes are commonly found in the lower airways and can survive oxygen stress by forming multilocular complexes within the hypoxic environment present in biofilms (55, 56). Short-chain fatty acids produced by fermentation, such as butyrate, are one energy source for anaerobes (57), and we have shown that their presence in the lower airways is higher in pneumotypeSPT and regulates IFNγ and IL17A production in CD4+/CD8+ lymphocytes (58). In NSCLC, we recently demonstrated that pneumotypeSPT is associated with several inflammatory cancer-related pathways, such as ERK/MAPK and PI3K/AKT (16), that can lead to chronic inflammation, altered Treg/Th17 balance (59–61), augmented Th17 differentiation (62, 63), and induction of PD-L1 expression (44, 45). Our current findings expand the above observations by demonstrating that a dysbiotic signature characterized by enrichment of the lower airway microbiota with oral commensals can contribute to the progression of disease.

Among the limitations pertinent to this study, we should point out that there is a significant degree of disease heterogeneity and the appropriate subanalyses could be explored only with a much larger cohort. For example, we decided to focus on NSCLC because there were few cases of small cell lung cancer. Further, within NSCLC there were several pathologic subtypes, driver mutation status, PD-L1 status, etc. The small subsample size prevents us from conducting the appropriate subanalysis. However, our analysis and models were stratified and adjusted by staging (dichotomized as I–IIIA and IIIB–IV stage groups and adjusted by individual TNM stages), which is a very significant covariate associated with prognosis and treatment modality. Interestingly, we found a few host transcriptomic signatures associated with a disease stage whereas there were many more transcriptomic signatures associated with lower airway microbiota subtype (SPT/BPT). It is possible that the histologic heterogeneity within NSCLC will affect these results, and a larger cohort may allow to control for this. Other potential confounders related to patients’ clinical condition, such as swallowing and deglutition problems, cannot be fully accounted for in the current cohort but may have significant impact on our results. Given our finding of the enrichment of the lower airway microbiota with oral commensal as associated with prognosis, future investigations that include evaluation of swallowing functions are warranted. Low biomass samples are subjected to contamination with background DNA (coming from the reagents, bronchoscopy, or sequencing noise; refs. 64, 65). To address concerns regarding DNA contamination during sample collection and preparation, we applied decontam (26) analysis and showed that Flavobacterium, a taxon identified in the multiomic analysis and dominant of BPT, is likely a background contaminant. This is consistent with prior data showing no growth from lower airway...
samples characterized as BPT (16). We therefore induced airway dysbiosis in our mouse model with Veillonella and compared it with PBS (which 16S rRNA gene sequencing composition most resembles BPT) rather than a separate bacterium as control. Our investigation supports the hypothesis that the lower airway microbiota contribute to a local protumor immunity; however, we did not investigate the systemic inflammatory response in this model. Further support for the relevance of this mechanism will need to focus on blocking the immune response to the microbial exposure in the setting of lung cancer and evaluating the effects of induced lower airway dysbiosis during immunotherapy. In the current investigation, we did not explore the association between lung microbiota and response to immunotherapy because this treatment was applied in a relatively small fraction of patients (16%) and the vast majority of the samples were collected before this therapy became standard of care. Also, although we identified a taxonomic signature associated with inflammatory tone and prognosis in lung cancer, we cannot determine the molecular signatures present in the microbial community that may be responsible for this association. Future investigations that exploit novel functional microbiomic approaches (e.g., metagenome, metatranscriptome, and metabolome) should focus on molecular markers with significant immunomodulatory activity. In our preclinical model, we tested whether Veillonella parvula was sufficient to induce lower airway inflammation and worsening of tumor progression. Other oral commensals, when present in the lower airways, may also be contributing to this process and may need to be further evaluated as key components of lower airway dysbiosis in isolation or in complex microbial communities. Although the lower airway microbiota were associated with staging and survival, other dysbiotic signatures in other mucosae could also have significant associations. Even though we did not identify significant microbiota signatures in the buccal samples, future investigations should include gut samples as well to establish the relative role of the microbiota of different mucosae niches to the pathogenesis of lung cancer. Finally, further validation of the results presented here will require a second cohort where sampling approach and design are customized to overcome some of the limitations here described.

This study has broad clinical implications regarding lung cancer pathogenesis and treatment response. Identification of lower airway dysbiotic signatures associated with lung cancer prognosis may be important to personalize approaches for lung cancer treatment and prognosis. Fecal microbiota transplant, a strategy with proven efficacy in difficult-to-treat Clostridium difficile infection and inflammatory bowel disease (66, 67), can influence the susceptibility to anti–PD-1 cancer immunotherapy (9, 10), and its clinical impact is now being tested in humans within ongoing clinical trials. Despite the evidence that the local microbiota affect the local inflammatory tone of the lung, there are no human trials aiming to modify the lung microbiome in the setting of malignancy. The data presented here suggest that lower airway dysbiosis induced by microaspiration of oral commensals affects lung tumorigenesis by promoting an IL17-driven inflammatory phenotype, a pathway amenable for targeted therapy that may have potential implications in this disease. A better understanding of the microbial host interaction in the lower airways will be needed to uncover how the lung cancer–associated microbiota could be modulated to affect prognosis and response to immunotheapy.

METHODS

Subjects

All subjects signed written informed consent to participate in this study that was approved by the Institutional Review Board of New York University. Participants included patients who had suspicious nodules on chest imaging and who underwent clinical bronchoscopy. Lung cancer subtype, somatic mutation, and stage were recorded after histopathologic confirmation. We excluded subjects with a prior history of cancer or recent (less than one month) antibiotic use. RECIST (27) score was analyzed at the 6- to 12-month time point after diagnosis of lung cancer, where these data were most consistently available.

Bronchoscopic Procedure

Both background and supraglottic (buccal) samples were obtained prior to the procedure as previously described (16). The background samples were obtained by passing sterile saline through the suctioning channel of the bronchoscope prior to the procedure. For this project, we obtained multiple lower airway samples from different locations, including 82 from the right mainstem, 59 from the airways leading to the lung cancer lesion (involved segments), and 69 from the airways spared of disease on the contralateral lung. A detailed description of the number of samples and the analyses performed is provided in Supplementary Table S6.

Bacterial 16S rRNA-Encoding Gene Sequencing

High-throughput sequencing of bacterial 16S rRNA-encoding gene amplicons (V4 region; ref. 68) was performed. Reagent control samples and mock mixed microbial DNA were sequenced and analyzed in parallel (Supplementary Fig. S23). The obtained 16S rRNA gene sequences were analyzed with the Quantitative Insights Into Microbial Ecology (QIIME; RRID:SCR_008249) 1.9.1 package (69). OTUs were not removed from upstream analysis. PERMANOVA testing was used to compare the compositional differences of groups. A prevalence-based method using the R package decontam (v1.6.0; ref. 26) was used to identify potential contaminants from the sequencing data sets. In this process, all reads from background bronchoscope control samples were identified as negative controls and thus possible sources of contaminants. No OTU was removed from the analyses performed and data from the 16S microbiome for this manuscript are available (Sequence Read Archive, RRID:SCR_001370: #PRJNA592147).

Sample clustering of meta-communities was based on Dirichlet multinomial mixtures (DMM) modeling (70).

Transcriptome of Bronchial Epithelial Cells

RNA-seq was performed on bronchial epithelial cells obtained by airway brushing, as described (71–73), using the Hi-Seq/Illumina platform at the NYU Langone Genomic Technology Center (data available at Sequence Read Archive: # PRJNA600487), KEGG (74, 75) annotation was summarized at levels 1 to 3. Genes with a false discovery rate (FDR)–corrected adjusted P value <0.25 were considered significantly differentiated, unless otherwise specified. Pathway analysis using differentially regulated genes (FDR < 0.25) was done using IPA (RRID:SCR_008653; QIAGEN Inc.; ref. 76). Gene set enrichment analysis was performed with differential genes (FDR < 0.25) for data set comparison, R package gsease v1.4.1 (77).
**Experimental Mouse Model**

The mice utilized in this experiment were five-week-old females at the time of use. The strain was B6(Cg)-Tyrc-2J/J mice purchased via vendor (Jackson Laboratory; cat. #000058). The mice were kept in Skirball Animal Facility and were kept under controlled conditions with cycles of 12-hour daylight and 12-hour darkness. Mice were euthanized by carbon dioxide asphyxiation followed by cardiac puncture. Blood, skin swabs, oral swabs, lung lavage, lung tissue, humerus bone marrow, cecum, terminal ileum, and fecal pellets were collected for study. The Institutional Animal Care and Use Committee of the New York University School of Medicine approved all procedures, and experiments were carried out following their guidelines (IACUC# s16-00032).

**KP Model Lung Cancer**

The KP model of lung cancer histopathologically resembles that of human cancers and has been used to study translational models of lung cancer in mice (78). The KP model of lung cancer is based on KRASGLS12Dp;Trp53R16+/-NSCLC models require induction by use of replication-deficient adenoviruses expression Cre (Ad-Cre) to induce transient Cre expression in the lungs of mice. Once tumor burden is increased in the mice, the lungs are harvested and the KP lung cancer cells grown in cell culture (79). Cell culture lines of KP lung cancer cells are grown in DMEM 10% FBS plus gentamicin under aerobic conditions (Bactron 300, Shel Lab) and then stored in 20% glycerol tryptic soy broth at −80°C. These bacteria were grown in anaerobic conditions (Bactron 300, Shel Lab) and then stored in 20% glycerol tryptic soy broth at −80°C. The colonies were collected from the plate and resuspended in 1 mL of sterile PBS. Nitrogen (N2) in an anaerobic chamber for 24 to 48 hours. The colonies were harvested from the cell culture when 90% confluent. The goal was to grow cells to 3,000,000 KP cells/mL (or 150,000 cells/50 μL). To detect in vivo luminescence, images were acquired using the IVIS spectrum (PerkinElmer) after intraperitoneal injection of Luciferin (Promega). We then proceeded to intratracheal inoculation of KP cells. The mice were anesthetized utilizing isoflurane until sedated. The mice were then placed on an intubation platform, and with blunt forceps, their tongue was gently pulled ventrally until the pharynx was exposed (78). Then, an Esol SafeCell catheter (Esol International Inc.; cat. # 26746) was introduced through the vocal cords of the mice, and a 50-μL inoculum of lung cancer (1.5 × 10^5 KP cells) was placed into the catheter. The mice were then removed from the intubation platform to recover from anesthesia on a heat pad.

**Creation of Veillonella Parvula Inoculum**

The following human oral commensals were obtained: Veillonella parvula, Prevotella melanogenumica, and Streptococcus mitis (ATCC). These bacteria were grown in anaerobic conditions (Bactron 300, Shel Labs) and then stored in 20% glycerol tryptic soy broth at 80°C. To prepare oral commensal challenge, bacteria strains were thawed and streaked on anaerobic PRAS-Brucella Blood agar plates (Anaerobe Systems). The plates were incubated at 37°C in an oxygen-free atmosphere (Bactron 300, Shel Lab) and then stored in 20% glycerol tryptic soy broth at −80°C. Bacteria strains were thawed and streaked on anaerobic PRAS-Brucella Blood agar plates (Anaerobe Systems). The plates were incubated at 37°C in an oxygen-free atmosphere (Bactron 300, Shel Lab) and then stored in 20% glycerol tryptic soy broth at −80°C. The colonies were collected from the plate and resuspended in 1 mL of sterile PBS. OD620 was measured to calculate the approximate concentration prior to use.

**Intratracheal Microbial and Control Challenge**

Mice were assigned to receive the microbial challenge with Veillonella parvula twice a week via intratracheal inoculation starting two weeks after the inoculation with lung cancer. First, mice were sedated with the use of isoflurane anesthesia. The mice were then suspended by their dorsal incisors upon an elastic cord; a blunt pair of forceps was used to ventrally pull the tongue forward to expose the larynx. Then, a pneumatic otoscope (Welch-Allyn; cat. #71000C) with a 2-mm ear specula was advanced until the vocal cords were visualized. Using a gel-loading tip, a 50-μL volume of the Veillonella parvula was deployed into the trachea of the mouse. These exposures occurred twice a week, spaced three to four days/week apart. Mice were monitored during this process; no mice died due to the inoculation procedure. A control procedure to inoculate mice with PBS was performed in the same manner.

**Immune Inhibition Experiment**

Two weeks after KP cell inoculation, mice were challenged intratracheally with Veillonella parvula similar to above. At this time, mice were randomized 1:1 to receive anti-IL17 (1 mg/mL; Bio X Cell), anti-IL17 isotype control (2 mg/mL; Bio X Cell). Antibody dose was diluted in 100 μL and given via intraperitoneal injection twice a week for a total of two weeks.

**Organization and Measurements on Mice**

Once lung tumor development was detected by IVIS (two weeks post inoculation), mice were randomized according to tumor burden to receive either PBS or dysbiosis with V. parvula while maintaining co-housing conditions. For the KP mice, those with median lumens of 8 × 10^3 to 7 × 10^4 photon-flux (photons/s/cm²/spectral) at two weeks were utilized for the experiments. Wild-type mice from the same strain and no KP exposure were used as control mice and were exposed to sterile PBS or V. parvula. Thus, in all experiments, mice were randomized to the following groups: (i) wild-type with PBS control (WT), (ii) wild-type with dysbiosis with V. parvula (Dys), (iii) KP lung cancer with PBS control (LC), and (iv) KP lung cancer with V. parvula (LC + Dys). Imaging the mice utilizing luciferin expression (luminescence) occurred two weeks after inoculation with KP lung cancer cells. The platform we used to image the mice was PerkinElmer IVIS Spectrum (PerkinElmer, cat. # 124262). Luciferin (1.5 mg; PerkinElmer, Xeno-Light D-Luciferin Potassium Salt; cat. # 122799) was given intraperitoneally. Mice received 50 μL of their respective inoculum with the Veillonella condition receiving 1.5 × 10^6 cfu/mL. The mice were organized into groups based upon their median lumens to establish experimental groups of mice with the same luminescence for a baseline. The imaging of the mice occurred twice every week on the day prior to inoculation. For the survival experiment, we utilized 60 mice that were followed for six weeks after initiation of microbial challenge or PBS control. Forty additional mice were divided in the same four conditional groups for immune phenotyping on lung homogenate, including lung transcriptomics, flow cytometry, and cytokine measurement. For this experiment, mice were sacrificed after two weeks post initiation of microbial or PBS exposure. For host RNA transcriptome, flash-frozen lung samples were defrosted and then homogenized utilizing a hand TissueRuptor II on the second lowest setting (Qiagen). Then samples were spun down on a tabletop centrifuge 14,000 rpm for two minutes, and the pellet was collected and sent for RNA processing. RNA was extracted from collected supernatant using the QiaGen miRNeasy Mini Kit (Qiagen; cat. #74135). Quality control was established with RNA integrity number cutoff >6. RNA-seq was performed using Hi-Seq (Illumina) at the NYU Genomic Technology Center. RNA-seq library prep was made using the Illumina TruSeq Stranded mRNA LT Kit (Illumina; cat. #RS-1222-2101) on a Beckman Biomek FX instrument, using 250 ng of total RNA as input, amplified by 12 cycles of PCR, and run on an Illumina 2500 (v4 chemistry), as single-read 50 bp. Sequences from the murine lung homogenate were aligned against the murine ensemble reference genome utilizing STAR, RRID:SCR_015899 (v2.5) aligner (80). Gene counting of each sample was performed using featureCounts, RRID:SCR_012919 (v1.5.3; refs. 81, 82). FACS was performed on single-cell suspension derived from lung homogenate. First, lung samples were minced and dissociated utilizing Liberase (Hoffmann-La Roche) for 35 minutes in a 37°C water bath and followed by mechanical disruption through a 70-μm filter. Liberase was used at a concentration of 0.5 mg/mL in DMEM supplemented with 10% FBS. For intracellular cytokine staining, the
cells were treated with a cell stimulation and protein transport inhibition cocktail containing PMA, Tonomycin, Brefeldin A, and Monensin (500x eBioscience Affymetrix) for four hours. The cells were surface stained, fixed in 2% paraformaldehyde, and permeabilized with 0.5% saponin. Cell staining with fluorochrome-conjugated antibodies was performed targeting CD3, CD4, CD8, CD69, PD1, and IL17 (Thermo-Fisher), and measurements were performed on a BD LSR II flow cytometer (BD Biosciences). Acquired data were analyzed using FlowJo, RRD-SCR_008520 version 10.3 (Tree Star Inc.). Cytokines and chemokines were measured using Luminex (Murine Cytokine Panel II, EMD Millipore). Lung homogenates were thawed and processed according to the recommended protocol using the Murine Cytokine/Chemokine Magnetic Bead Panel (# MCYTMA 70K-PXld52). All cytokines/chemokines concentrations were normalized by the gram of lung homogenate and included those with dynamic range: G-CSF, Eotaxin, IFNg, IL1a, IL1b, IL3, IL4, IL5, IL6, IL7, IL9, IL10, IL12p40, IL12p70, LIF, IL17, IP10, KC, MCP1, MIP1a, MIP1b, M-CSF, MIP2, MIG, RANTES, VEGF, and TNFa.

**Multiplex Immunostaining**

Five-micrometer sections of paraffin-embedded preserved lung tissue were stained with Akoya Biosciences Opal multiplex automation kit reagents unless stated otherwise. Automated staining was performed on Leica BondRX autostainer. The protocol was performed according to the manufacturers’ instructions with the antibodies specified in Supplementary Table S7. Brieﬂy, all slides underwent sequential epitope unmixing and analysis using InForm version 2.4.10 software from (Murine Cytokine Panel II, EMD Millipore). All cytokines/chemokines concentrations were normalized by the gram of lung homogenate and included those with dynamic range: G-CSF, Eotaxin, IFNg, IL1a, IL1b, IL3, IL4, IL5, IL6, IL7, IL9, IL10, IL12p40, IL12p70, LIF, IL17, IP10, KC, MCP1, MIP1a, MIP1b, M-CSF, MIP2, MIG, RANTES, VEGF, and TNFa.

**Image Acquisition and Analysis**

Semiautomated image acquisition was performed on a Vectra Polaris multispectral imaging system. After whole-slide scanning at 20x, the tissue was manually outlined to select fields for spectral unmixing and analysis using InForm version 2.4.10 software from Akoya Biosciences. Fields of view for analysis were separated as containing tumor only or areas of pulmonary parenchyma where tumor was not apparent. For each field of view, cells were segmented based on nuclear signal (DAPI). Cells were phenotyped after segmentation using InForm’s trainable algorithm based on gmnnet (83) package in R. Four algorithms were created to classify cells as Ly6g (neutrophils) or “other,” CD4+ or “other,” CD8+ or “other” and F4/80 or “other.” Phenotypes were reviewed for different samples during training iterations. Data were exported as text containing sample names, field of acquisition coordinates, individual cell information including coordinates and identified phenotype. Each image was analyzed with all four algorithms, so that every cell was classified four times. Concatenation of all phenotyping information was performed in R using the Phenoptr Reports package (Kent S. Johnson 2020). phenoptr: InForm Helper Functions. R package version 0.2.7. https://akoyaibo.github.io/phenoptr/ in RStudio software (RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc.; URL http://www.rstudio.com/). Statistical analysis (Mann-Whitney U test) was run for the following groups: lung cancer vs. lung cancer + dysbiosis (n = 4 and 8 mice, respectively, Fig. 3E), and lung cancer + dysbiosis vs. lung cancer + dysbiosis + anti-IL17 (n = 8 and 6 mice, respectively, Fig. 4D), taking each field as an independent value.

**Statistical and Multiomic Analyses**

In Supplementary Table S2, categorical variables were presented as frequencies and percentages and their distribution difference between groups with dead or alive OS status were assessed by the Fisher exact test. The Cox proportion hazards models (84) were used to evaluate each variable’s marginal association with the time to death. Hazard ratio and P value were reported.

The MiRKAT (85) was used to investigate whether the community-level microbial profile among lower airway samples was different between any paired samples from right main, involved, or noninvolved locations, and between stage I–IIIA and IIIB–IV while adjusting for smoking status within each location samples. The survival version of MiRKAT test: MiRKAT-S (86) was used to investigate whether the community-level microbial profile is associated with the OS while adjusting for smoking status, stage, and surgery within each location sample. The paired Bray–Curtis dissimilarity was used in all tests.

For the taxonomic level analysis, we used the linear mixed-effect model on the arc sin square root–transformed relative abundance at genus level for their associations with stage (I–IIIA/IIIB–IV; Supplementary Table S3). In the model, the subject was set as the random effect to take care of the correlation among three location samples from the same subjects. The stage was set as fixed effect while adjusting for smoking status. We used the two-stage linear mixed-effect model (87) on the arc sin square root–transformed relative abundance at genus level for their associations of the OS (Supplementary Table S4) while adjusting for smoking status, stage, and surgery. In the first stage, the linear mixed-effect model was used to take care of the correlation among three location samples from the same subjects. The random intercept estimates from the first stage were used in the Cox proportional hazards model in the second stage to investigate their association with the OS.

Because the distributions of microbiome data are nonnormal, and no distribution-specific tests are available, we used nonparametric tests of association. For association with discrete factors, we used either the Mann–Whitney test (in the case of two categories) or the Kruskal–Wallis ANOVA (in the case of >2 categories). For tests of association with continuous variables, we used the nonparametric Spearman correlation tests. FDR was used to control for multiple testing (88). To evaluate for taxonomic or transcriptomic differences between groups, we utilized DESeq2 (89).

Differential abundance of microbes related to lung cancer stage (IIIB–IV vs. I–IIIA) was calculated using Songbird as previously described (90). We then computed the microbe-transcript co-occurrence probability (probability of observing a transcriptomic pathway when a microbe is observed) using mmvec (91). A probability matrix of the top 10 transcriptome-related pathways for each microbe was generated and used to create a network based on the Fruchterman–Reingold force-directed algorithm using R package ggnnet v 0.1.0. (https://cran.r-project.org/web/packages/GGally/index.html). Microbe nodes were colored based on differential analysis of stage IIIB–IV versus I–IIIA NSCLC.

**Data Storage**

Sequencing data are available at Sequence Read Archive (92, 93) under accession number 16S Microbiome PRJNA592147, Human RNA-seq PRJNA600487, and Murine RNA-seq PRJNA600489. Codes utilized for the analyses presented in the current article are available at https://github.com/segalmicrobiomelab/lung_cancer_prognosis_microbiome.

**Authors’ Disclosures**

J. Carpenito reports grants from NIH during the conduct of the study. J. Pass reports grants from NCI EDRN U01 and AACR/Kr Grant during the conduct of the study. K.-K. Wong reports being a founder and equity holder of G1 Therapeutics. K.-K. Wong has sponsored research agreements with Pfizer, Janssen, Takeda, Targimmune, AstraZeneca, Novartis, Merck, Ono, and Array. L. Segal reports grants from NCI/NIH, FNIH, and AACR Grant during the conduct of the study. No disclosures were reported by the other authors.
Lung Microbiome and Lung Cancer Prognosis

Authors’ Contributions

J.-C.J. Tsay: Conceptualization, resources, data curation, formal analysis, funding acquisition, validation, investigation, methodology, writing—original draft, writing—review and editing. B.G. Wu: Investigation, methodology, writing—original draft. I. Sulaiman: Conceptualization, data curation, software, formal analysis, investigation, visualization, and methodology. K. Gershner: Data curation. R. Schluger: Data curation. Y. Li: Resources, data curation, and supervision. T.-A. Yie: Data curation. P. Meyn: Data curation and software. E. Olsen: Data curation, software, and formal analysis. L. Perez: Data curation. B. Franca: Data curation, software, and formal analysis. J. Carpenito: Data curation and investigation. T. Iizumi: Investigation. M. El-Ashmawy: Data curation, software, and formal analysis. M. Badri: Conceptualization, data curation, software, formal analysis, visualization, methodology, and writing—original draft. J.T. Morton: Conceptualization, data curation, software, investigation, visualization, methodology, and writing—original draft. N. Shen: Data curation, software, and formal analysis. L. He: Data curation, software, and formal analysis. G. Michelaud: Data curation. S. Rafiq: Data curation. J.L. Bessich: Data curation. R.L. Smith: Data curation. H. Sauthoff: Data curation. K. Felner: Data curation. R. Pillai: Investigation. A.-M. Zavitsanou: Investigation. S.B. Koralov: Methodology. V. Mezzano: Software, investigation, and visualization. C.A. Loomis: Conceptualization and resources. A.L. Moreira: Software, investigation, and visualization. W. Moore: Formal analysis. A. Tsirogi: Conceptualization, supervision, and methodology. A. Heguy: Conceptualization, supervision, and methodology. V.N. Rom: Conceptualization and supervision. D.H. Sterman: Conceptualization and supervision. H.I. Pass: Data curation, software, formal analysis, supervision, and validation. H. Li: Conceptualization, formal analysis, supervision, and validation. R. Bonneau: Conceptualization and supervision. K.-K. Wong: Conceptualization, resources, supervision, funding acquisition, methodology, writing—review and editing. T. Papagiannakopoulos: Conceptualization, supervision, funding acquisition, validation, investigation, methodology, writing—review and editing. L.N. Segal: Conceptualization, resources, data curation, software, formal analysis, supervision, funding acquisition, investigation, visualization, methodology, writing—original draft, project administration, writing—review and editing.

Acknowledgments

This study was supported by R37 CA244775 (L.N. Segal, NIH/NCI); PACT grant (L.N. Segal, FNHI); K23 AI02970 (L.N. Segal, NIH/NAID); EDNR SU01CA08137–13 (W.N. Rom); DoD W81XWH-16-1-0324 (J.-C. Tsay); the 2018 AACR-Johnson and Johnson Lung Cancer Innovation Science Grant Number 18-90-52-ZHAN (H.I. Pass/L.N. Segal); A Breath of Hope Foundation (J.-C. Tsay); Simmons Foundation (A. Bonneau); and CTIS Grant #UL1 TR000038 (L.N. Segal). The Genome Technology Center and ExpPath Core are partially supported by the Cancer Center Support Grant P30CA016087 at the Laura and Isaac Perlman Cancer Center (A. Heguy and A. Tsirogi); T32 CA193111 (B.G. Wu); UL1TR000145 (B.G. Wu); FAMRI Young Clinical Scientist Award (B.G. Wu); Stony Wold-Herbert Fund Grant Fund-in-Aid/Fellowship (B.G. Wu, I. Sulaiman, and K. Gershner); R01 HL125816 (L.N. Segal and S.B. Koralov, NIH/NHLBI); and R01 DK110014 (H. Li and L. He). We would like to thank the Genome Technology Center (GTC) for expert library preparation and sequencing, the Applied Bioinformatics Laboratories (ABL) for providing bioinformatics support and helping with the analysis and interpretation of the data, and Experimental Pathology Research Laboratory for histopathology services and imaging. GTC, ExpPath Core, and ABL are shared resources partially supported by the Cancer Center Support Grant P30CA016087 at the Laura and Isaac Perlman Cancer Center. This work has used computing resources at the NYU School of Medicine High Performance Computing Facility. Financial support for the PACT project is possible through funding support provided to the FNHI by AbbVie Inc., Amgen Inc., Boehringer-Ingelheim Pharma GmbH and Co. KG, Bristol-Myers Squibb, Celgene Corporation, Genentech Inc., Gilead, GlaxoSmithKline plc, Janssen Pharmaceutical Companies of Johnson and Johnson, Novartis Institutes for Biomedical Research, Pfizer Inc., Sanofi, and Shared Instrumentation Grant S10 OD021747 award for the Vectra Imaging system in Experimental Pathology.

Received March 8, 2020; revised September 15, 2020; accepted October 27, 2020; published first November 11, 2020.

REFERENCES


Lung Microbiome and Lung Cancer Prognosis


Lower Airway Dysbiosis Affects Lung Cancer Progression


Updated version
Access the most recent version of this article at:
doi:10.1158/2159-8290.CD-20-0263

Supplementary Material
Access the most recent supplemental material at:
http://cancerdiscovery.aacrjournals.org/content/suppl/2020/11/05/2159-8290.CD-20-0263.DC1

Cited articles
This article cites 91 articles, 21 of which you can access for free at:
http://cancerdiscovery.aacrjournals.org/content/11/2/293.full#ref-list-1

Citing articles
This article has been cited by 2 HighWire-hosted articles. Access the articles at:
http://cancerdiscovery.aacrjournals.org/content/11/2/293.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link
http://cancerdiscovery.aacrjournals.org/content/11/2/293.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.