Cancer Science

WILEY

Cancer Science

REPORT OPEN ACCESS

Association of Intratumoral Bacterial Abundance With Lung Cancer Prognosis in Chiba University Hospital Cohort

Takahiro Ochi[1,2] | Ryoji Fujiki[2] | Masaki Fukuyo[2] | Bahityar Rahmutulla[2] | Takuya Nakagawa[2,3] | Masayuki Ota[4] | Jun-ichiro Ikeda[4] | Yukiko Matsui[1] | Ichiro Yoshino[1] | Hidemi Suzuki[1] | Atsushi Kaneda[2,3]

[1]Department of General Thoracic Surgery, Graduate School of Medicine, Chiba University, Chiba, Japan | [2]Department of Molecular Oncology, Graduate School of Medicine, Chiba University, Chiba, Japan | [3]Health and Disease Omics Center, Chiba University, Chiba, Japan | [4]Department of Diagnostic Pathology, Graduate School of Medicine, Chiba University, Chiba, Japan

Correspondence: Atsushi Kaneda (kaneda@chiba-u.jp)

Received: 12 December 2024 | **Revised:** 1 April 2025 | **Accepted:** 4 April 2025

Funding: This work was supported by grant from the Japan Agency for Medical Research and Development (22zf0127008s0301) (A.K.), the IAAR Research Support Program (A.K.) and the Futuristic Medical Fund (T.O.) from Chiba University, and Grants from the Chiba Foundation for Health Promotion & Disease Prevention (H.S).

Keywords: intratumoral bacteria | lung cancer | poor prognostic factor

ABSTRACT

The relationship between cancer prognosis and intratumoral microbiome has recently gained attention. Regarding lung cancer, most studies have focused on bacteria outside tumors, such as sputum or lavage fluid, with few examining intratumoral bacteria and their impact on prognosis. In this study, we extracted DNA from lung tumor samples of 507 patients undergoing surgery at Chiba University Hospital and quantified intratumoral bacterial abundance using bacteria-specific PCR primers. Bacteria were detected in 77.1% of cases, and bacterial abundance was significantly higher in lung adenocarcinoma than in squamous cell carcinoma. Patients were categorized into three groups (High, Low, and Very-Low) based on bacterial abundance, and associations with clinicopathological factors were analyzed. In lung squamous cell carcinoma, higher bacterial abundance was significantly associated with worse recurrent-free survival and overall survival and was found to be a poor prognostic factor independent of pathological tumor stage. In conclusion, intratumoral bacterial abundance was found in the majority of lung cancer tissues, with variations based on pathology. This abundance may serve as a useful marker for stratifying lung squamous cell carcinoma with distinct prognoses.

1 | Introduction

Bacteria are known to contribute to cancer development as external factors: *Helicobacter pylori* induces chronic inflammation in gastric mucosa, leading to atrophic gastritis and intestinal metaplasia, which increases the risk of gastric cancer development [1]. Specific bacteria within tumor tissues have also been implicated in carcinogenesis and metastasis in colorectal and esophagus cancers [2, 3]. Microbiome has been detected from lung cancer tumors [4], but studies have mainly focused on bacteria outside the tumor, such as sputum and bronchoalveolar lavage fluid [5, 6], with limited investigation of intratumoral bacteria [7, 8]. Quantitative and qualitative evaluation methods vary across reports, and sample sizes are often limited. The association between intratumoral bacterial abundance and lung cancer prognosis is not fully validated.

Abbreviations: bgDNA, bacterial genomic DNA; bPrimer, Primer for bgDNA; COPD, chronic obstructive pulmonary disease; Cq, quantification cycle; CuhT, Chiba University Hospital Thoracic surgery; hgDNA, human genomic DNA; hPrimer, Primer for hgDNA; IP, interstitial pneumonia; LUAD, Lung adenocarcinoma; LUSC, Lung squamous cell carcinoma; OS, overall survival; pStage, pathological stage; R^2, coefficient of determination; RFS, recurrence-free survival.

This study aimed to establish an accurate quantitative method to analyze a large number of lung cancer cases and evaluate its impact on prognosis. The cohort included samples from Chiba, a major industrial area in Japan with higher air pollution levels [9].

2 | Materials and Methods

Detailed information can be found in Doc. S1.

2.1 | DNA Preparation of Tumor Samples

Cancer tissue samples (3 mm) were collected during thoracic surgery at Chiba University Hospital under clean conditions. DNA was extracted using the NucleoSpin VET kit (MACHEREY-NAGEL GmbH & Co. KG, Duren, Germany).

2.2 | Quantification of Genomic DNA

DNA concentration was measured using QuantiFluor ONE dsDNA System (Promega Corporation, Madison, WI) (Table S1). Quantitative PCR (qPCR) primers specific for bacteria genomic DNA (bgDNA) and human genomic DNA (hgDNA) were designed (bPrimer and hPrimer, respectively) (Tables S2, S3).

2.3 | FISH

For cases with a high amount of bgDNA detected, localization within the tumor tissue was examined using FISH performed on FFPE samples.

2.4 | 16S rDNA Amplicon Analysis

Taxonomic classification was performed based on qPCR amplicons generated using the designed primer sequences for bgDNA.

2.5 | Statistical Analysis

Student's *t*-test, Fisher's exact test, Mann–Whitney *U* test, Kruskal–Wallis test, Cox proportional hazards analysis, and log-rank test were performed.

3 | Results and Discussion

3.1 | Establishment of CuhT Cohort

Two most common lung cancer subtypes were analyzed: lung adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC). Between January 2016 and December 2023, 537

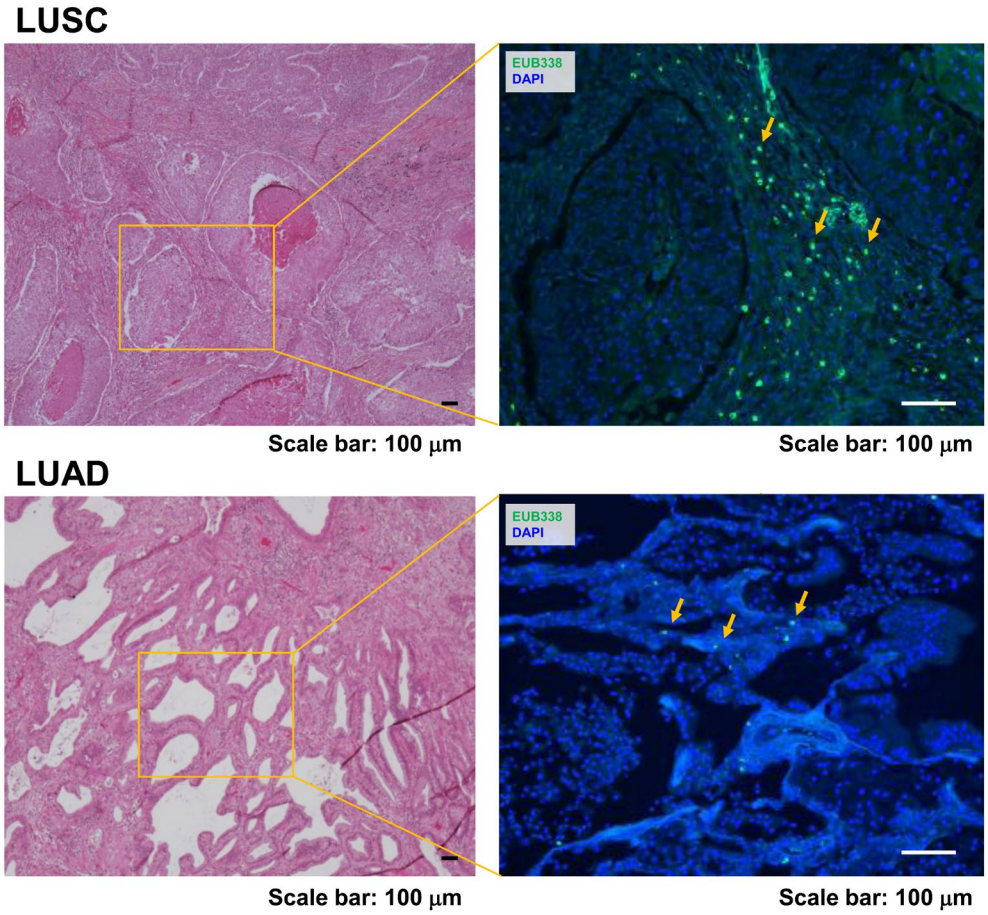

FIGURE 1 | Detection of bacteria in lung cancer tissue using FISH. H&E staining and FISH results against 16S rRNA with EUB338 in representative LUSC (CuhT cohort #241) and LUAD (CuhT cohort #138) cases. As highlighted by arrows, bacteria were confirmed in cancer tissue. CuhT, Chiba University Hospital Thoracic surgery; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma.

patients underwent surgery and consented to participate. After excluding cases of surgery for recurrence and those with preoperative chemotherapy, 507 cases remained (Chiba University Hospital Thoracic surgery cohort: CuhT cohort). The mean age of the CuhT cohort was 70.5 ± 8.6 years; 346 patients were male (68.2%) and 379 (74.8%) had a smoking history. Pathological results revealed LUAD in 369 patients (72.8%) and LUSC in 138 patients (27.2%). Patient characteristics in the CuhT cohort appeared comparable to those reported by the National Cancer Center [10]. The median yield of extracted DNA was 30.56 μg (Range: 4.02–53.34) (Table S1).

3.2 | Lung Tumors Contain Abundant Bacteria

We designed bPrimers targeting conserved regions of 16S rDNA from universal primers [11], and hPrimers specific to hgDNA at the *LMNA* locus. We used gDNA derived from *E. coli* as a positive control for bgDNA and gDNA derived from gastric epithelial cells as a positive control for hgDNA. The primer evaluation included: (1) specific amplification of target bgDNA or hgDNA, (2) PCR product size (50–200 bp), and (3) linearity of the qPCR standard curve using a 10-fold serial dilution series of the standardized template.

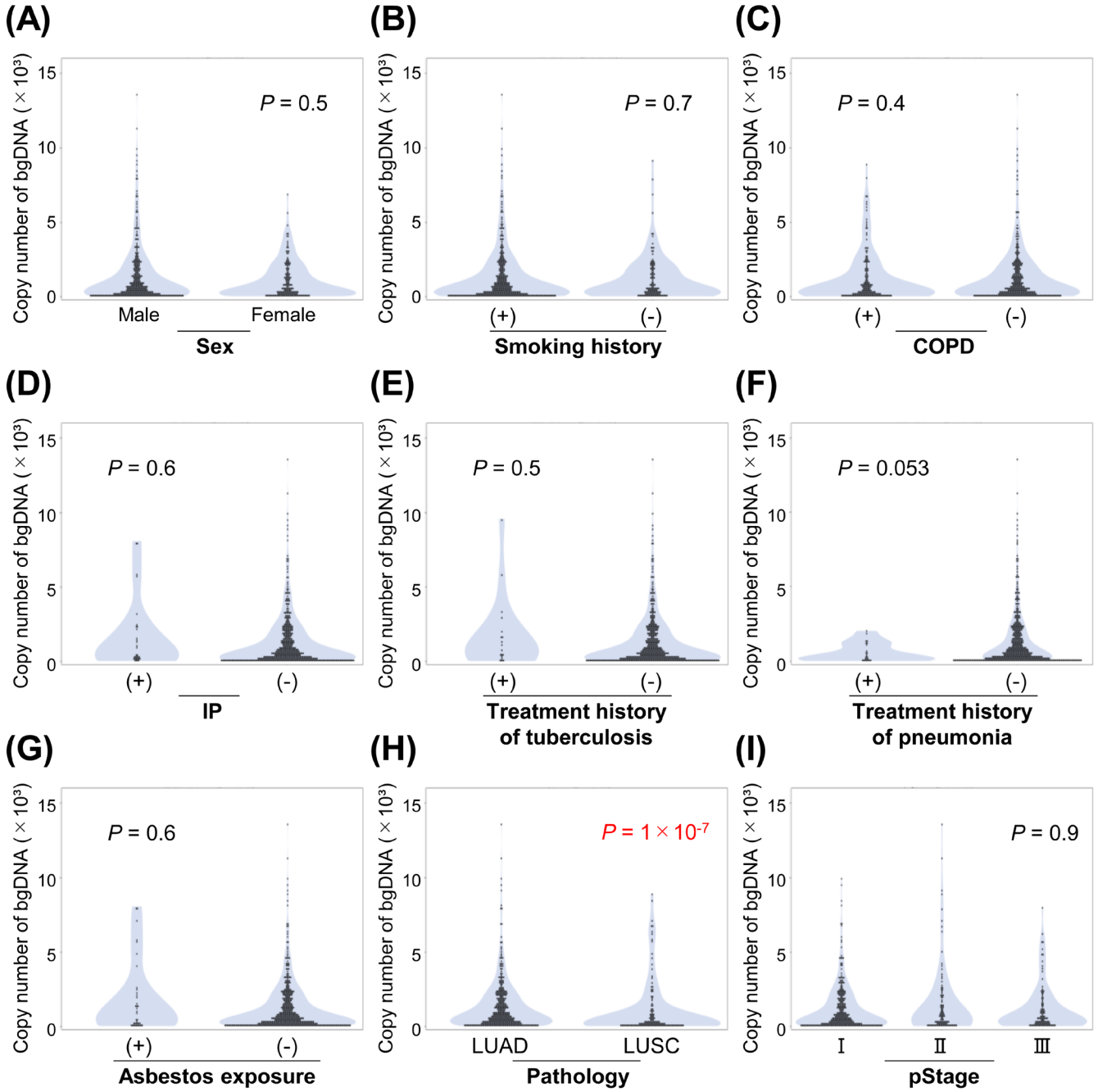

FIGURE 2 | Association between intratumoral bacterial abundance and clinicopathological characteristics. Comparison of intratumoral bacterial abundance by sex (A), smoking history (B), presence/absence of COPD (C), presence/absence of IP (D), treatment history of tuberculosis (E), treatment history of pneumonia (F), asbestos exposure (G), pathology (H), and pStages (I). COPD, chronic obstructive pulmonary disease; IP, interstitial pneumonia; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; pStage, pathological stage.

TABLE 1 | Comparison of bacterial abundance and patient characteristics in the patients with LUAD ($n = 369$) or LUSC ($n = 138$).

Characteristic (number)	Bacterial abundance (median [range])	*p*
LUAD		
Sex, male/female (223/146)	7.3 [0.5–135.6]/5.4 [0.5–68.5]	0.076
Smoking history, yes/no (244/125)	7.5 [0.5–135.6]/5.3 [0.5–91.2]	0.169
COPD, yes/no (104/265)	5.5 [0.5–79.6]/7.6 [0.5–135.6]	0.105
IP, yes/no (17/352)	4.1 [0.5–79.6]/6.8 [0.5–135.6]	0.992
Treatment history of tuberculosis, yes/no (15/354)	4.3 [0.5–94.9]/6.8 [0.5–135.6]	0.670
Treatment history of pneumonia, yes/no (16/353)	6.0 [0.5–20.3]/6.9 [0.5–135.6]	0.587
Asbestos exposure history, yes/no (23/346)	13.2 [0.5–79.6]/6.5 [0.5–135.6]	0.208
pStage, I/II/III (223/67/79)	7.3 [0.5–99.2]/6.9 [0.5–135.6]/5.6 [0.5–79.6]	0.878
LUSC		
Sex, male/female (123/15)	2.1 [0.5–88.7]/0.5 [0.5–33.5]	0.083
Smoking history, yes/no (135/3)	1.8 [0.5–88.7]/0.5 [0.5–3.9]	0.260
COPD, yes/no (71/67)	2.8 [0.5–88.7]/1.0 [0.5–84.3]	0.010
IP, yes/no (14/124)	3.3 [0.5–57.9]/1.7 [0.5–88.7]	0.156
Treatment history of tuberculosis, yes/no (4/134)	13.6 [3.1–57.9]/1.7 [0.5–88.7]	0.045
Treatment history of pneumonia, yes/no (8/130)	0.5 [0.5–2.0]/2.1 [0.5–88.7]	0.013
Asbestos exposure history, yes/no (15/123)	0.7 [0.5–70.8]/1.8 [0.5–88.7]	0.816
pStage, I/II/III (76/36/26)	1.4 [0.5–84.3]/2.6 [0.5–88.7]/2.3 [0.5–62.0]	0.279

Abbreviations: COPD, chronic obstructive pulmonary disease; IP, interstitial pneumonia; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; pStage, pathological stage.

bPrimers 12, 13, 14, and hPrimer 2 specifically and exclusively amplified 16S rDNA-derived bgDNA and *LMNA*-derived hgDNA (Figure S1). With bPrimer 14, qPCR revealed a standard curve slope of 4.05, a coefficient of determination (R^2) of 0.99, within 10^2–10^6 copies. With hPrimer 2, qPCR showed a standard curve slope of 3.32 (R^2: 0.99), within 10^1–10^4 copies. We performed qPCR using mixture samples containing a constant concentration (10^4 copies/2 μL of PCR template) of hgDNA and a 10-fold serial dilution of spike-in bgDNA to evaluate the accuracy of bgDNA quantification in DNA samples extracted from lung tumors. qPCR with bPrimer 14 maintained the quantifiability of bgDNA with linearity (R^2:0.99) within 10^2–10^6 copies. Using hPrimer 2, 10^4 copies of hgDNA in the template were consistently quantified, regardless of bgDNA levels (Figure S1).

In the CuhT cohort, hgDNA was detected in all samples, with a median copy number of 12933.33 (range: 925.33–39320.00) in 2 μL of template. As for bgDNA, qPCR showed that 391 samples (77.1%) had quantities above the lower limit of quantification range, while the remaining 116 samples (22.9%) were defined as "Very-Low-group," (45 below and 71 undetected) (Figure S2). For bgDNA, the quantity in 2 μL of template was 534 0.47 copies (range: 29.17–13560.67). Using negative controls, where empty sample tubes subjected to the same procedure as lung specimen collection were washed with PBS and underwent DNA extraction, qPCR results showed that all eight samples were below the measurement range, corresponding to the Very-Low-group. The median of bacteria copies/40 ng of DNA was 486.90 (range: 15.41–173799.19) (Figure S3). The higher values comparable to the previous report [4] may result from the rigorous DNA extraction method using steel and corundum beads and bacteria-specific primer.

FISH analysis on LUSC and LUAD samples revealed the presence of bacteria in the cancer tissues (Figure 1). Interestingly, bacteria were more abundant in the stroma than in the tumor cells in the LUSC. It was indicated that the results of bacterial quantification were not attributable to bacteria present on the tissue surface or the alveolar epithelium.

3.3 | Bacteria Within Lung Tumors Affect Patient Characteristics and Prognosis

Comparison of bacterial abundance and patient characteristics showed significantly higher bacterial abundance in LUAD than in LUSC ($p = 1 \times 10^{-7}$), with no significant differences in other clinicopathological characteristics (Figure 2). Patient characteristics were therefore examined separately within LUAD and LUSC cohorts (Table 1). No significant differences were observed in any variables within LUAD. In LUSC, however, patients with chronic obstructive pulmonary disease (COPD) exhibited significantly higher bacterial abundance, which may be explained by chronic airway inflammation that compromises epithelial barrier integrity, facilitating bacterial infiltration into

TABLE 2 | Univariate and multivariate analyses of prognosis in the patients with LUAD ($n = 235$) or LUSC ($n = 76$) who underwent surgery between 2016 and 2019.

	Univariate analysis		Multivariate analysis	
Characteristic (number)	**HR (95% CI)**	***p***	**HR (95% CI)**	***p***
LUAD, OS				
Bacterial abundance group				
High (97) vs. Low (99)	0.660 (0.359–1.211)	0.180	0.718 (0.390–1.320)	0.286
High (97) vs. Very-low (39)	0.778 (0.347–1.747)	0.544	0.897 (0.398–2.025)	0.794
Low (99) vs. Very-low (39)	1.180 (0.555–2.508)	0.668	1.250 (0.587–2.663)	0.562
pStage				
I (150) vs. II (38)	0.532 (0.253–1.118)	0.096	0.543 (0.257–1.146)	0.109
I (150) vs. III (47)	0.306 (0.168–0.559)	1.1×10^{-4}*	0.313 (0.171–0.572)	1.6×10^{-4}*
II (38) vs. III (47)	0.576 (0.269–1.231)	0.155	0.576 (0.269–1.230)	0.154
LUAD, RFS				
Bacterial abundance group				
High (97) vs. Low (99)	0.981 (0.589–1.636)	0.942	1.126 (0.674–1.881)	0.651
High (97) vs. Very-low (39)	0.575 (0.319–1.035)	0.065	0.703 (0.389–1.271)	0.244
Low (99) vs. Very-low (39)	0.585 (0.326–1.050)	0.073	0.624 (0.348–1.122)	0.115
pStage				
I (150) vs. II (38)	0.264 (0.146–0.479)	1.1×10^{-5}*	0.267 (0.147–0.486)	1.5×10^{-5}*
I (150) vs. III (47)	0.148 (0.088–0.249)	6.5×10^{-13}*	0.150 (0.089–0.253)	1.1×10^{-12}*
II (38) vs. III (47)	0.559 (0.317–0.987)	0.045*	0.560 (0.317–0.989)	0.046*
LUSC, OS				
Bacterial abundance group				
High (15) vs. Low (24)	2.812 (0.888–8.902)	0.079	4.020 (1.211–13.349)	0.023*
High (15) vs. Very-low (37)	2.958 (1.059–8.264)	0.039*	3.703 (1.277–10.739)	0.016*
Low (24) vs. Very-low (37)	1.052 (0.343–3.223)	0.929	0.921 (0.297–2.856)	0.887
pStage				
I (42) vs. II (21)	1.224 (0.377–3.977)	0.736	1.673 (0.498–5.628)	0.406
I (42) vs. III (13)	0.353 (0.131–0.949)	0.039*	0.328 (0.120–0.895)	0.030*
II (21) vs. III (13)	0.288 (0.084–0.986)	0.047*	0.196 (0.054–0.706)	0.013*
LUSC, RFS				
Bacterial abundance group				
High (15) vs. Low (24)	1.904 (0.613–5.918)	0.265	2.209 (0.690–7.069)	0.182
High (15) vs. Very-low (37)	3.394 (1.085–10.616)	0.036*	3.540 (1.099–11.402)	0.034*
Low (24) vs. Very-low (37)	1.782 (0.573–5.541)	0.318	1.603 (0.514–4.997)	0.416
pStage				
I (42) vs. II (21)	0.624 (0.190–2.046)	0.437	0.809 (0.239–2.738)	0.733
I (42) vs. III (13)	0.206 (0.069–0.614)	0.005*	0.214 (0.072–0.640)	0.006*
II (21) vs. III (13)	0.330 (0.105–1.042)	0.059	0.265 (0.081–0.862)	0.027*

Note: Univariate and multivariate analyses were performed using the Cox proportional hazard model. Asterisks indicate statistically significant differences.
Abbreviations: HR, hazard ratio; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; OS, overall survival; pStage, pathological stage; RFS, recurrent-free survival.

tumor tissues. Patients with tuberculosis treatment history showed significantly higher bacterial abundance, possibly reflecting residual structural lung damage following tuberculosis, which could favor bacterial colonization. In contrast, the lower bacterial abundance observed in patients with pneumonia history may reflect the effects of broad-spectrum antibiotic treatment and enhanced bacterial clearance mechanisms.

The 391 quantified samples were divided into high-volume (High-group) and low-volume (Low-group) groups, based on the median value of bgDNA quantification. Comparison of patient characteristics between High- and Low-groups, as well as Very-Low-group, revealed that LUAD was significantly more prevalent in the High-group (Table S4). LUAD primarily arises in the peripheral airways, while LUSC preferentially develops in the central airways. Compared to the peripheral airways, the central airways exhibit faster airflow, denser cilia, and higher secretion levels. These disparities in airway clearance may underlie observed differences in intratumoral bacterial abundance between LUAD and LUSC.

For prognosis, cases from 2016 to 2019 with > 5 years of follow-up were pathologically classified. In LUAD, bacterial abundance did not significantly correlate with either overall survival (OS) or recurrent-free survival (RFS) (Table 2). In LUSC, however, univariate and multivariate analyses using the Cox proportional hazard model showed that pathological stage (pStage) and bacterial abundance were both significantly associated with prognosis in OS and RFS. While the prognostic relevance of pStage is well-established, bacterial abundance was identified as a prognostic factor independent from pStage. The Kaplan–Meier curve also revealed that the High-group showed significantly poorer OS ($p = 0.03$) and RFS ($p = 0.02$) than the Very-Low-group in LUSC (Figure S4). The impact of intratumoral bacteria on cancer prognosis reportedly involved mechanisms, for example, chronic inflammation induced by cytokine production, DNA damage caused by increased reactive oxygen species, and deterioration of the tumor metabolic environment due to the imbalance of lactic acid bacteria and anaerobic bacteria [12]. Increased bacterial abundance may reflect the amplification of these effects in LUSC.

Since bPrimer14 used for quantitative evaluation includes the hypervariable V5 region of 16S rDNA, we performed overhang sequence PCR and index PCR on the qPCR amplicon, followed by 16S rDNA amplicon analysis on LUAD-High, LUAD-Low, LUSC-High, and LUSC-Low samples (Figure S5). Alpha diversity analysis showed that microbial diversity was significantly higher in LUSC-High than in both LUAD-High ($p = 0.01$) and LUSC-Low ($p = 0.04$) (Figure S6). Beta diversity analysis revealed significant differences in microbial community composition between LUAD-High and LUAD-Low ($p = 0.004$), LUAD-High and LUSC-High ($p = 0.013$), and LUSC-High and LUSC-Low ($p = 0.044$), indicating distinct phylogenetic structures among groups (Table S5). These findings were notable, as High-group of LUSC also demonstrated significantly worse prognosis in OS and RFS, independent of pStage. The limitations of this method include the limited taxonomic resolution due to targeting the hypervariable region V5 of the 16S rDNA, which prevents reliable identification at the family or genus level, and the use of a high number of PCR cycles, which may compromise the accurate reflection of the true bacterial population. Further analyses are necessary to clarify the influence of bacterial abundance and microbial diversity on clinical outcomes in LUSC.

Taken together, this study highlighted specific trends in bacterial abundance itself across pathological types and prognosis of lung cancer using quantitative evaluations. Meanwhile, qualitative evaluations, for example, 16S rDNA amplicon analysis in detail, are also important. Notably, specific oral microbiome influences the pathogenesis of esophageal and colorectal cancer [2, 3]. We believe that the tumor DNA samples from this cohort could support future studies investigating the association between specific bacteria or microbiome factors and the initiation, progression, and prognosis of lung cancer.

Author Contributions

Takahiro Ochi: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, resources, visualization, writing – original draft. **Ryoji Fujiki:** data curation, formal analysis, investigation, writing – original draft. **Masaki Fukuyo:** data curation, formal analysis. **Bahityar Rahmutulla:** data curation, formal analysis. **Takuya Nakagawa:** investigation, writing – review and editing. **Masayuki Ota:** investigation, resources. **Jun-ichiro Ikeda:** investigation, resources. **Yukiko Matsui:** resources. **Ichiro Yoshino:** conceptualization, resources, supervision. **Hidemi Suzuki:** conceptualization, funding acquisition, resources, supervision. **Atsushi Kaneda:** conceptualization, funding acquisition, supervision, writing – review and editing.

Acknowledgments

We thank Haruka Maruyama, Tsubasa Matsusaka, and Kyoko Hirano for their technical assistance.

Ethics Statement

This research was approved by the Institutional Review Board of Chiba University (#1203, #HS202301-01).

Consent

Lung cancer samples were obtained with written informed consent.

Conflicts of Interest

A.K. is Associate Editor of Cancer Science. All other authors declare no conflicts of interest.

References

1. T. Ushijima, S. J. Clark, and P. Tan, "Mapping Genomic and Epigenomic Evolution in Cancer Ecosystems," *Science* 373, no. 6562 (2021): 1474–1479.

2. A. D. Kostic, D. Gevers, C. S. Pedamallu, et al., "Genomic Analysis Identifies Association of Fusobacterium With Colorectal Carcinoma," *Genome Research* 22, no. 2 (2012): 292–298.

3. K. Yamamura, Y. Baba, S. Nakagawa, et al., "Human Microbiome *Fusobacterium nucleatum* in Esophageal Cancer Tissue Is Associated With Prognosis," *Clinical Cancer Research* 22, no. 22 (2016): 5574–5581.

4. D. Nejman, I. Livyatan, G. Fuks, et al., "The Human Tumor Microbiome Is Composed of Tumor Type-Specific Intracellular Bacteria," *Science* 368, no. 6494 (2020): 973–980, https://doi.org/10.1126/science.aay9189.

5. J. J. Tsay, B. G. Wu, M. H. Badri, et al., "Airway Microbiota Is Associated With Upregulation of the PI3K Pathway in Lung Cancer," *American Journal of Respiratory and Critical Care Medicine* 198, no. 9 (2018): 1188–1198, https://doi.org/10.1164/rccm.201710-2118OC.

6. J. J. Tsay, B. G. Wu, I. Sulaiman, et al., "Lower Airway Dysbiosis Affects Lung Cancer Progression," *Cancer Discovery* 11, no. 2 (2021): 293–307, https://doi.org/10.1158/2159-8290.cd-20-0263.

7. G. Yu, M. H. Gail, D. Consonni, et al., "Characterizing Human Lung Tissue Microbiota and Its Relationship to Epidemiological and Clinical Features," *Genome Biology* 17, no. 1 (2016): 163.

8. K. L. Greathouse, J. R. White, A. J. Vargas, et al., "Interaction Between the Microbiome and TP53 in Human Lung Cancer," *Genome Biology* 19, no. 1 (2018): 123.

9. Ministry of the Environment, Japan, Annual Report on Environmental Statistics 2017, accessed November 26, 2024, https://www.env.go.jp/en/statistics/contents/2017/E2017_all.pdf.

10. National Cancer Center Japan, Cancer Information Service, Cancer Statistics: Lung Cancer (2023), https://ganjoho.jp/reg_stat/statistics/stat/summary.html.

11. D. A. Soergel, N. Dey, R. Knight, and S. E. Brenner, "Selection of Primers for Optimal Taxonomic Classification of Environmental 16S rRNA Gene Sequences," *ISME Journal* 6, no. 7 (2012): 1440–1444.

12. Y. Cao, H. Xia, X. Tan, et al., "Intratumoural Microbiota: A New Frontier in Cancer Development and Therapy," *Signal Transduction and Targeted Therapy* 9, no. 1 (2024): 15.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.