

# *Sphingomonas* and *Phenylobacterium* As Major Microbiota in Thymic Epithelial Tumors

**Rumi Higuchi**

Yamanashi Kenritsu Chuo Byoin

**Taichiro Goto** (✉ [taichiro@1997.jukuin.keio.ac.jp](mailto:taichiro@1997.jukuin.keio.ac.jp))

Yamanashi Kenritsu Chuo Byoin <https://orcid.org/0000-0002-2560-3639>

**Yosuke Hirotsu**

Yamanashi Kenritsu Chuo Byoin

**Sotaro Otake**

Yamanashi Kenritsu Chuo Byoin

**Toshio Oyama**

Yamanashi Kenritsu Chuo Byoin

**Kenji Amemiya**

Yamanashi Kenritsu Chuo Byoin

**Hiroshi Ohyama**

Yamanashi Kenritsu Chuo Byoin

**Hitoshi Mochizuki**

Yamanashi Kenritsu Chuo Byoin

**Masao Omata**

Yamanashi Kenritsu Chuo Byoin

---

## Research

**Keywords:** thymoma, microbiome, 16S RNA sequencing, genera, driver mutation

**Posted Date:** July 1st, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-666392/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

Microbiota has been reported to be closely associated with carcinogenesis and cancer progression. However, its involvement in the pathology of thymoma remains unknown. In this study, we aimed to identify thymoma-specific microbiota using resected thymoma samples.

## Methods

Nineteen thymoma tissue samples were analyzed through polymerase-chain-reaction amplification and 16S rRNA gene sequencing. The subjects were grouped according to histology, driver mutation status in the *GTF2I* gene, PD-L1 status, and smoking habits. To identify the taxa composition of each sample, operational taxonomic units (OTUs) were classified on the effective tags with 97% identity. The Shannon index of the 97% identity OTUs was calculated to evaluate alpha diversity. The linear discriminant analysis effect size (LEfSe) method was used to compare the relative abundances of all bacterial taxa.

## Results

We identified 107 OTUs in the tumor tissues, which were classified into 26 genera. *Sphingomonas* and *Phenylobacterium* were identified as abundant genera in almost all samples. No significant difference was determined in the alpha diversity within these groups; however, type A thymoma exhibited higher bacterial diversity than type B thymoma. Through LEfSe analysis, we identified the following differentially abundant taxa: Bacilli, Firmicutes, and Lactobacillales in type A thymoma; Proteobacteria in type B thymoma; Gammaproteobacteria in tumors harboring the *GTF2I* mutation; and Alphaproteobacteria in tumors without the *GTF2I* mutation.

## Conclusions

*Sphingomonas* and *Phenylobacterium* were identified as dominant genera in thymic epithelial tumors. These genera appear to comprise thymoma-specific microbiota involved in tumor progression; thus, they could serve as targets for the prevention of thymoma.

## Introduction

Early microbiome research focused primarily on gastrointestinal diseases, such as pseudomembranous enterocolitis, inflammatory bowel disease, and irritable colitis [1]. Recently, human intestinal microbiota has been reported to be involved in carcinogenesis and cancer progression, and this phenomenon has been attracting attention [2, 3]. In addition, microbiota has been identified in tissues of the pancreas, lung, and breast through advanced sequencing technology [4–8].

Thymoma is a relatively rare mediastinal tumor with malignant potential that is difficult to treat [9, 10]. According to the histological classification by the World Health Organization, thymomas can be categorized into types A, AB, B1, B2, and B3, depending on the tumor cell morphology and proportion of coexisting lymphocytes [11]. Type A thymomas are the least aggressive with the best prognosis; the extent of aggressiveness increases and the prognosis worsens according to the following order: type A, AB, B1, B2, and B3 [12, 13]. Thymoma has been reported to commonly occur in people aged 40 to 60 years [14]. The development of thymoma is not associated with smoking habits or sex; its causes are unknown [15]. However, thymoma coexists in approximately 20% of patients with myasthenia gravis [16, 17]. Owing to the absence of effective treatment other than surgical resection, there is an urgent need to elucidate the pathology and establish preventive measures and new treatment strategies for thymoma [18–21].

Although recent reports indicate the association between the microbiome and the development of colorectal, oral, pancreatic, lung, and other cancers [22–27], there is no report on the involvement of the microbiome in thymoma. Unlike oral, gastrointestinal, and respiratory cancers, which have been previously reported, thymoma is anatomically located in the anterior mediastinum, and it does not communicate with the outer environment. Because the tumor environment of thymoma has been theoretically assumed to be sterile, microbiome research has not been conducted in the past. Consequently, little progress has been made in research on the involvement of the microbiota in the pathology of thymoma.

In this study, we performed polymerase chain reaction (PCR) to amplify the 16S ribosomal RNA (rRNA) region in the bacterial genome in resected thymoma samples. Subsequently, we performed 16S rRNA sequence and metagenomic analyses using next-generation sequencing to investigate the composition and diversity of the microbiota and to identify thymoma-specific microbiota. On the basis of the results from these analyses, we presented a predictive model of pathogenesis and evaluated its potential for prevention and control of the development of thymoma.

## Materials And Methods

### Patients and sample preparation

In this study, we enrolled 19 patients in an unbiased manner, who underwent surgical resection for thymoma at our hospital between January 2014 and August 2020. Since antibiotics would affect the microbiome, patients who had used oral or systematic antibiotics in the past 3 months were not included in this study. We obtained written informed consent for genetic research from all patients in accordance with the protocols approved by the Institutional Review Board at Yamanashi Central Hospital. The specimens were categorized histologically according to the classification guidelines by the World Health Organization [28, 29] and staged according to the Masaoka staging system [19, 30, 31]. Sections of formalin-fixed and paraffin-embedded (FFPE) tissues were stained with hematoxylin–eosin and microdissected using the ArcturusXT laser-capture microdissection system (Thermo Fisher Scientific,

Waltham, MA, USA), as previously reported [32-37]. For type AB thymomas, the type A and B portions were microdissected and examined separately. We analyzed 22 samples obtained from all 19 patients, including three patients with type AB thymoma. The GeneRead DNA FFPE Kit (Qiagen, Hilden, Germany) was used according to the manufacturer's instructions, and the DNA quality was evaluated using primers against ribonuclease P, as previously reported [38]. In the same manner, tumor DNA was extracted from FFPE samples obtained from patients with pancreatic cancer in our hospital (n=30) and used as a control.

## **16S rRNA amplification and targeted sequencing**

The 16S rDNA V4 region was amplified using PCR and sequenced as described previously with minor modifications [7]. FFPE DNA was amplified using the Platinum PCR SuperMix High Fidelity (Thermo Fisher Scientific) with the forward primer 5'-GTGYCAGCMGCCGCGGTAA-3'(16S\_rRNA\_V4\_515F) and reverse primer 5'-GGACTACNVGGGTWTCTAAT-3' (16S\_rRNA\_V4\_806R). The PCR products were confirmed using agarose gel electrophoresis and purified using Agencourt AMPure XP reagents (Beckman Coulter, Brea, CA, USA). End repair and barcode adaptors were ligated with the Ion Plus Fragment Library Kit (Thermo Fisher Scientific) in accordance with the manufacturer's instructions to construct libraries. The library concentration was determined using an Ion Library Quantitation Kit (Thermo Fisher Scientific), and the same number of libraries was pooled for one sequence. Emulsion PCR and chip loading was performed on the Ion Chef with the Ion PGM Hi-Q View Chef Kit; sequencing was performed on the Ion PGM Sequencer (Thermo Fisher Scientific). Sequence data were transferred to the IonReporter local server using the IonReporterUploader plugin. Data was analyzed using Metagenomics Research Application using a custom primer set. Analytical parameters were set as default.

## **Data analysis**

The original raw tags were obtained through splicing the reads using FLASH (v1.2.7) and subsequently filtered to acquire clean tags using QIIME (Version 1.9.1). To identify the taxa composition of each sample, operational taxonomic units (OTUs) were classified on the effective tags with 97% identity using the Usearch (Uparse v7.0.1001) software. The representative sequence of each OTU was annotated using the RDP classifier against the SILVA (SSU123) 16S rRNA database using a confidence threshold of 80%, obtaining taxonomic classification at the phylum, class, order, family, genus and species levels. Multiple sequence alignment was performed using MUSCLE3.6 (Version 3.8.31) to further explore the phylogenetic relationships among different OTUs. The Shannon index was performed using QIIME to determine alpha diversity. Linear discriminant analysis (LDA) effect size (LEfSe) analyses were performed using the online LEfSe tool (<http://huttenhower.sph.harvard.edu/lefse/>). The LDA (linear discriminant analysis) threshold score was set at 2.

## **Targeted deep sequencing of GTF2I mutation**

In this study, the presence of point mutations in the *GTF2I* gene was investigated in thymomas using targeted sequencing coupled with molecular barcoding, as we previously reported [39].

## Immunohistochemistry for PD-L1

Specimens from the 19 patients were fixed using 10% buffered formalin. The formalin-fixed paraffin-embedded tissues were cut into 5  $\mu\text{m}$  sections, deparaffinized, rehydrated, and stained in an automated system (Ventana Benchmark ULTRA system; Roche, Tucson, AZ, USA) using commercially available detection kits and antibodies against PD-L1 (28–8, ab205921; Abcam, Cambridge, MA, USA). PD-L1 was primarily localized to the cell membrane of tumor cells, and its expression was determined quantitatively by two pathologists on the basis of the proportion of PD-L1-positive tumor cells. Cells were considered PD-L1-positive based on  $\geq 1\%$  PD-L1 expression.

## Statistics

Continuous variables were presented as the mean  $\pm$  standard deviation (SD), and compared using unpaired Student's *t* tests. One-way analysis of variance and the Tukey–Kramer multiple comparison test were used to detect significant differences between groups. *P*-values less than 0.05 in the two-tailed analyses were considered to denote statistical significance.

## Results

### Patient characteristics

We analyzed 22 samples from 19 patients with thymomas who had undergone surgery at Yamanashi Central Hospital between January 2014 and August 2020. Table 1 presents the clinicopathologic characteristics of the patients, such as the age, sex, histology, tumor size, stage, smoking status, and diagnosis of myasthenia gravis. Among the 19 patients, 11 and 8 were males and females, respectively, and 12 and 7 were smokers and non-smokers, respectively. According to histological examination, five, three, five, four, and two patients exhibited type A, AB, B1, B2, and B3 tumors, respectively (Table 1). There were no cases of micronodular thymoma. The 19 patients recruited in this study were divided according to the Masaoka stages: stage I ( $n=7$ ), II ( $n=9$ ), III ( $n=2$ ), and IV ( $n=1$ ). The maximum tumor diameter ranged from 20 mm to 95 mm (mean tumor diameter,  $43.6 \pm 22.8$  mm). The age of the patients ranged between 42 and 81 years ( $68.2 \pm 12.9$  years). One patient with type B2 thymoma exhibited comorbidity with myasthenia gravis (Case 20; Figure 1 and 2).

### OTU analyses

A total of 136 OTUs were identified in the 22 samples. The dominant ( $>1\%$  average relative abundance) classifiable OTUs belonged to four families, namely, Sphingomonadaceae (abundance:  $62.0 \pm 12.2\%$ ), Caulobacteraceae (abundance:  $23.9 \pm 7.6\%$ ), Bradyrhizobiaceae (abundance:  $6.7\% \pm 5.3\%$ ), and Phyllobacteriaceae (abundance:  $2.0\% \pm 1.4\%$ ) (Figure 1). We identified 107 genera ( $>1\%$  average relative abundance); the predominant genera are presented in Figure 2. The top two genera with high abundance and composition were *Sphingomonas* (abundance:  $66.9 \pm 10.8\%$ ) and *Phenylobacterium* (abundance:  $26.0 \pm 8.7\%$ ). *Sphingomonas* was detected in all samples, and *Phenylobacterium* was detected in all

samples except in case 18. Both bacterial genera were significantly more abundant than the others (Supplementary Figure 1).

### **Differences in microbiota between thymomas and pancreatic cancers**

To identify thymoma-specific microbiota, we compared the microbiota between thymoma and pancreatic cancer. In the thymoma samples, compared with the pancreatic cancer samples, *Phenylobacterium*, *Phyllobacterium*, and *Sphingomonas* were significantly more abundant (Figure 3). Because *Phenylobacterium*, *Phyllobacterium*, and *Sphingomonas* were detected only in 4, 3, and 8 of the 30 pancreatic cancer samples, respectively, the composition of these genera in the 22 thymoma samples (detected in 21, 18, and 22 samples, respectively) was significantly higher.

### **Analysis of microbial diversity within groups**

The Shannon index was calculated to evaluate the bacterial diversity within different groups. No significant differences were observed in terms of histology, presence or absence of *GTF2I* mutation, PD-L1 expression, and smoking habits (Figure 4). However, type A samples exhibited a tendency toward increased microbiome diversity, compared with type B samples ( $p=0.059$ , Figure 4A).

### **Analysis of differentially abundant taxa**

To further identify specific species in every group, we used the LEfSe method to identify the differentially abundant taxa at each level. First, in type A and B histological groups, we identified four differential bacterial taxa including two phyla, Firmicutes and Proteobacteria; one class, Bacilli; and one order, Lactobacillales (Figure 5A). The differential features were Firmicutes, Bacilli, and Lactobacillales in type A thymomas and Proteobacteria in type B thymomas (Figure 5B). Alphaproteobacteria was dominant in thymomas without *GTF2I* mutation, while Gammaproteobacteria was dominant in thymomas harboring *GTF2I* mutation (Figure 5C, D). No differential bacterial composition and abundance were observed in association with the stage, PD-L1 expression, or smoking habits.

## **Discussion**

In this study, sequencing of microbiota in resected thymoma samples identified two genera, *Sphingomonas* and *Phenylobacterium*, in almost all thymoma samples; the bacterial composition and abundance of these genera were markedly high. We separately analyzed Type AB thymoma for type A and type B components and detected *Sphingomonas* and *Phenylobacterium* in both components. Although the oral microbiome is likely to affect and contaminate the lung microbiome, thymoma is anatomically unlikely to be affected by the oral microbiome [40, 41]. The composition and abundance of these two genera were significantly higher in the microbiota of thymoma tissues than in the microbiota of pancreatic cancer tissues. Our results suggest that these two genera are thymoma-specific microbiota. In addition, the pancreatic cancer and thymoma tissue samples were analyzed in the same process at the genome analysis center of our hospital during the same period. This analysis suggested that the

presence of the two genera was not a result of contamination during the analysis process. In contrast, *Sphingomonas* and *Phenylobacterium* have not been detected in lung cancer tissues according to recent reviews on the microbiota in patients with lung cancer [4, 42–44]. In this study, because these two genera were detected in almost all thymoma samples, it was suggested that *Sphingomonas* and *Phenylobacterium* may represent differential microbiome function in thymoma development.

*Sphingomonas* is a bacterial genus that was separated from *Pseudomonas* approximately 30 years ago. Members of the former are Gram-negative bacteria; however, they do not contain lipopolysaccharides specific to Gram-negative bacteria [45]. Instead, these bacteria contain glycosphingolipids, which are found in eukaryotic cells [45]. They are common microorganisms inhabiting various environments, such as water environments (e.g., freshwater and seawater), soil, and plant root systems. The wide ecological distribution of these bacteria is attributed to their ability to use diverse organic compounds and their strong vitality, allowing them to survive in nutrient-poor environments [46]. Although several bacteria in the genus *Sphingomonas* were isolated in relatively clean environments, certain bacterial species were isolated in contaminated environments containing toxic organic compounds, such as polychlorinated biphenyl, creosote, and pentachlorophenol [47]. Subsequent studies revealed that these bacteria take up certain organic contaminants and use them as energy sources [48]. On the basis of these findings, progress has been made in elucidating the mechanism through which *Sphingomonas* metabolizes organic contaminants. Furthermore, several attempts have been made worldwide for applying this mechanism in environmental cleanup (bioremediation). Meanwhile, with respect to the microbiome, *Sphingomonas* has been reported to be enriched as blood microbiota in the serum of healthy patients and patients with breast cancer who exhibit a favorable prognosis [49, 50].

The genus *Phenylobacterium* comprises a single species called *P. immobile*, which is remarkable for its extremely limited nutritional spectrum. All strains isolated and described hitherto grow optimally only on artificial compounds such as chloridazon, antipyrin, and pyramidon [51]. Species within the genus *Phenylobacterium*, are capable of degrading xenobiotic compounds with a phenyl moiety such as chloridazon, antipyrine, pyramidon, or their analogues [52]. Additionally, these bacteria can degrade polycyclic aromatic hydrocarbons [53]. *Phenylobacterium* has now been used in the bioremediation of a petroleum-contaminated soil to degrade polycyclic aromatic hydrocarbons and their analogues [54]. Unlike *Sphingomonas*, there has been no report of the detection of *Phenylobacterium* as blood microbiota. Future studies are expected to elucidate how *Sphingomonas* and *Phenylobacterium*, which are two genera of environmentally indigenous bacteria used for bioremediation, coexist in thymoma and how they are involved in the carcinogenic mechanism of thymoma.

Several indigenous microorganisms exist in the epithelium of the whole human body (e.g., the mouth, ear, nasal cavity, respiratory organs, digestive tract, skin, and reproductive organs), form microbiota, play various roles in the body, and form a symbiotic relationship with humans [1, 2]. In recent years, it has been considered that disturbance in the microbiota composition (dysbiosis) may alter the risk of disease development, and there is a growing number of reports on the association between intestinal microbiota and several diseases, such as allergy, cancer, multiple sclerosis, Parkinson's disease, depression,

inflammatory bowel disease, and rheumatism [42]. Furthermore, sterilization and specific-pathogen-free breeding have been reported to alleviate or cure these diseases in pathological mouse models [55]. Improvement of the microbiota may additionally prevent the development of diseases in humans [56]. If one or several species of bacteria cause a disease, they can be potential therapeutic targets. For example, eradication of *Helicobacter pylori* is the standard of care for the prevention of gastric cancer in infected patients at present [57]. This study clearly identified the microbiota involved in the development of thymoma. Considering the future clinical application of this result, the development of thymoma can be prevented through controlling the bacterial genera *Sphingomonas* and *Phenylobacterium*. Patients with myasthenia gravis are at a high risk of developing thymoma [16, 17], and prevention of thymoma is important for their long-term survival. In this study, case 20 involved a patient with thymoma complicated by myasthenia gravis (Figs. 1 and 2); this patient was positive for *Sphingomonas* and *Phenylobacterium*, which were abundant. The development of probiotic models for antibiotics, vaccines, and other therapies targeting these genera identified in this study may be important for the prevention of thymoma.

Bacterial diversity tended to be higher in type A thymoma (least aggressive type) than in type B (more aggressive type). A study comparing the microbiota between tumor and normal peritumoral tissues in lung cancer demonstrated that the bacterial diversity was significantly higher in the normal peritumoral tissues [6]. According to these data, cancer aggressiveness and alpha diversity are negatively correlated. Because the cancer microenvironment is more perturbed, dysbiosis might be enhanced; consequently, bacterial diversity might decrease. In addition, because the lymphocyte counts in the tissues are higher in type B thymoma than in type A thymoma, the immunity against these bacteria may fundamentally differ between these types.

Using LEfSe analysis, we identified variation in specific species between type A and B tumors and between tumors with and without *GTF2I* mutation, indicating the differential microbiome function in the development of each type of tumor [58]. We determined that Firmicutes, Bacilli, and Lactobacillales were common Gram-positive bacteria in type A thymoma and Proteobacteria were common Gram-negative bacteria in type B thymoma. When the *P*-value based on the Kruskal–Wallis test was increased from 0.05 to 0.1 (Supplementary Fig. 2), 15 of 20 bacteria that were significantly detected in the microbiota of type A thymoma were Gram-positive bacteria, and all four bacteria significantly detected in the microbiota of type B thymoma were Gram-negative bacteria. Although these findings indicated a correlation between the histological types (types A and B) and Gram staining results for the microbiota, the biological significance of this correlation is unknown. Because Gram-negative bacteria are generally more pathogenic than Gram-positive bacteria, the former may be involved in carcinogenesis in type B thymoma, which is a more aggressive phenotype. Additionally, it is unclear from our observational study whether the identified bacterial differences are causally related to lung carcinogenesis or merely reflective of the disease process in thymoma. In the future, detailed studies with a larger sample size may be needed.

We previously reported that *GTF2I* mutation is a driver mutation in thymoma [39]. In the present study, specific species were identified between tumors with and without *GTF2I* mutation. While

Alphaproteobacteria were detected in significantly more cases without the driver mutation in the *GTF2I* gene, Alphaproteobacteria are reported to induce natural genetic transformation, a process in which genetic material passes from one bacterium to another, as reported in *Neisseria gonorrhoeae* and *Haemophilus influenzae* [59]. Natural genetic transformation is a process involving DNA transfer from one bacterial cell to another through the integration of the donor sequence into the recipient genome. Such an alternative pathway may be involved in the oncological development of thymoma without driver mutation. Additionally, the mechanisms through which the microbiota contributes to carcinogenesis need to be examined in detail using a large sample size in the future.

This study is associated with some limitations. First, the patient cohort was relatively small owing to the rarity of the tumor. Second, patient survival could not be analyzed as no patients have shown recurrence in the cohort. Third, no blood samples were analyzed for microbiota containing the two genera, *Sphingomonas* and *Phenylobacterium*. Analysis of blood samples might have elucidated the reasons for the presence of the microbiota in the sterile anterior mediastinal environment. In addition, the higher abundance of *Sphingomonas* and *Phenylobacterium* may be related to the impaired immunity of the tumor microenvironment, which may cause proliferation of these bacteria in the blood. Thus, they may be clinically applicable as serum biomarkers for thymoma. In this context, a larger series of studies need to be performed for evaluating the microbiome landscape of thymomas more comprehensively and elucidate associations with clinical parameters through a more exhaustive multivariate analysis. Nevertheless, since the major aim of this preliminary analysis was identification of the thymoma-specific microbiota that should be prioritized for clinical development, the modestly sized samples provide useful insights.

## Conclusions

This is the first study that examined the microbiota in thymomas and revealed two genera specific to thymomas: *Sphingomonas* and *Phenylobacterium*. The development of probiotic models for controlling these two genera may enable the prevention of thymoma. Particularly, antibiotic and vaccine therapies for *Sphingomonas* and *Phenylobacterium* are expected to be clinically applicable for preventing thymoma in patients with myasthenia gravis who are likely to develop thymoma.

## Abbreviations

FFPE, Formalin-fixed paraffin-embedded; LEfSe, linear discriminant analysis effect size; OTU, operational taxonomic units

## Declarations

## Acknowledgments

The authors appreciate Yoshihiro Miyashita, Yumiko Kakizaki, and Toshiharu Tsutsui for helpful scientific discussions.

### **Author contributions**

**Rumi Higuchi:** Conceptualization; Methodology; Data curation; Formal analysis; Investigation; Writing - original draft. **Taichiro Goto:** Conceptualization; Methodology; Data curation; Formal analysis; Investigation; Writing - review & editing. **Yosuke Hirotsu:** Conceptualization; Methodology; Data curation; Investigation; Resources; Writing - original draft. **Sotaro Otake:** Data curation; Investigation. **Toshio Oyama:** Data curation. **Kenji Amemiya:** Data curation; Investigation. **Hiroshi Oyama:** Data curation; Investigation. **Hitoshi Mochizuki:** Data curation; Methodology; Formal analysis. **Masao Omata:** Conceptualization; Methodology; Formal analysis; Resources; Supervision.

### **Funding**

This study was supported by a Grant-in-Aid for Genome Research Project from Yamanashi Prefecture (to Y.H. and M.O.).

### **Availability of data and materials**

The data that support the findings of this study are available from the authors upon reasonable request.

### **Ethics Approval and Consent to participate**

This study was approved by the Institutional Review Board of Yamanashi Central Hospital.

### **Consent for publication**

Not applicable.

### **Competing interests**

The authors declare they have no competing interests.

### **Availability of data and materials**

All data used in this study are present in the main text and supplementary materials. Additional data and materials can also be requested from first author or corresponding author.

## **References**

1. Power SE, O'Toole PW, Stanton C, Ross RP, Fitzgerald GF. Intestinal microbiota, diet and health. *Br J Nutr.* 2014;111(3):387–402.

2. Kovaleva OV, Romashin D, Zborovskaya IB, Davydov MM, Shogenov MS, Gratchev A. Human Lung Microbiome on the Way to Cancer. *J Immunol Res*. 2019;2019:1394191.
3. Mao Q, Jiang F, Yin R, Wang J, Xia W, Dong G, Ma W, Yang Y, Xu L, Hu J. Interplay between the lung microbiome and lung cancer. *Cancer Lett*. 2018;415:40–8.
4. Goto T. **Airway Microbiota as a Modulator of Lung Cancer**. *Int J Mol Sci* 2020, 21(9).
5. Laborda-Illanes A, Sanchez-Alcoholado L, Dominguez-Recio ME, Jimenez-Rodriguez B, Lavado R, Comino-Méndez I, Alba E. Queipo-Ortuño MI: **Breast and Gut Microbiota Action Mechanisms in Breast Cancer Pathogenesis and Treatment**. *Cancers (Basel)* 2020, 12(9).
6. Peters BA, Hayes RB, Goparaju C, Reid C, Pass HI, Ahn J. The Microbiome in Lung Cancer Tissue and Recurrence-Free Survival. *Cancer Epidemiol Biomarkers Prev*. 2019;28(4):731–40.
7. Riquelme E, Zhang Y, Zhang L, Montiel M, Zoltan M, Dong W, Quesada P, Sahin I, Chandra V, San Lucas A, et al. Tumor Microbiome Diversity and Composition Influence Pancreatic Cancer Outcomes. *Cell*. 2019;178(4):795–806 e712.
8. Wei MY, Shi S, Liang C, Meng QC, Hua J, Zhang YY, Liu J, Zhang B, Xu J, Yu XJ. The microbiota and microbiome in pancreatic cancer: more influential than expected. *Mol Cancer*. 2019;18(1):97.
9. Engels EA. Epidemiology of thymoma and associated malignancies. *J Thorac Oncol*. 2010;5(10 Suppl 4):260–5.
10. Venuta F, Anile M, Diso D, Vitolo D, Rendina EA, De Giacomo T, Francioni F, Coloni GF. Thymoma and thymic carcinoma. *Eur J Cardiothorac Surg*. 2010;37(1):13–25.
11. Marx A, Chan JK, Coindre JM, Detterbeck F, Girard N, Harris NL, Jaffe ES, Kurrer MO, Marom EM, Moreira AL, et al. The 2015 World Health Organization Classification of Tumors of the Thymus: Continuity and Changes. *J Thorac Oncol*. 2015;10(10):1383–95.
12. Marx A, Strobel P, Badve SS, Chalabreysse L, Chan JK, Chen G, de Leval L, Detterbeck F, Girard N, Huang J, et al. ITMIG consensus statement on the use of the WHO histological classification of thymoma and thymic carcinoma: refined definitions, histological criteria, and reporting. *J Thorac Oncol*. 2014;9(5):596–611.
13. Moon JW, Lee KS, Shin MH, Kim S, Woo SY, Lee G, Han J, Shim YM, Choi YS. Thymic epithelial tumors: prognostic determinants among clinical, histopathologic, and computed tomography findings. *Ann Thorac Surg*. 2015;99(2):462–70.
14. Thomas CR, Wright CD, Loehrer PJ. Thymoma: state of the art. *J Clin Oncol*. 1999;17(7):2280–9.
15. Conforti F, Pala L, Giaccone G, De Pas T. Thymic epithelial tumors: From biology to treatment. *Cancer Treat Rev*. 2020;86:102014.
16. Kumar R. Myasthenia gravis and thymic neoplasms: A brief review. *World J Clin Cases*. 2015;3(12):980–3.
17. Romi F. Thymoma in myasthenia gravis: from diagnosis to treatment. *Autoimmune Dis*. 2011;2011:474512.

18. Girard N, Lal R, Wakelee H, Riely GJ, Loehrer PJ. Chemotherapy definitions and policies for thymic malignancies. *J Thorac Oncol.* 2011;6(7 Suppl 3):1749–55.
19. Litvak AM, Woo K, Hayes S, Huang J, Rimner A, Sima CS, Moreira AL, Tsukazan M, Riely GJ. Clinical characteristics and outcomes for patients with thymic carcinoma: evaluation of Masaoka staging. *J Thorac Oncol.* 2014;9(12):1810–5.
20. Schmitt J, Loehrer PJ. Sr.: The role of chemotherapy in advanced thymoma. *J Thorac Oncol.* 2010;5(10 Suppl 4):357–60.
21. Zhao Y, Shi J, Fan L, Hu D, Yang J, Zhao H. Surgical treatment of thymoma: an 11-year experience with 761 patients. *Eur J Cardiothorac Surg.* 2016;49(4):1144–9.
22. Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, Barnes R, Watson P, Allen-Vercoe E, Moore RA, et al. *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res.* 2012;22(2):299–306.
23. Gethings-Behncke C, Coleman HG, Jordao HWT, Longley DB, Crawford N, Murray LJ, Kunzmann AT. *Fusobacterium nucleatum* in the Colorectum and Its Association with Cancer Risk and Survival: A Systematic Review and Meta-analysis. *Cancer Epidemiol Biomarkers Prev.* 2020;29(3):539–48.
24. Gur C, Ibrahim Y, Isaacson B, Yamin R, Abed J, Gamliel M, Enk J, Bar-On Y, Stanietsky-Kaynan N, Copenhagen-Glazer S, et al. Binding of the Fap2 protein of *Fusobacterium nucleatum* to human inhibitory receptor TIGIT protects tumors from immune cell attack. *Immunity.* 2015;42(2):344–55.
25. Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, Clancy TE, Chung DC, Lochhead P, Hold GL, et al. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe.* 2013;14(2):207–15.
26. Kostic AD, Gevers D, Pedamallu CS, Michaud M, Duke F, Earl AM, Ojesina AI, Jung J, Bass AJ, Taberero J, et al. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res.* 2012;22(2):292–8.
27. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. *Fusobacterium nucleatum* promotes colorectal carcinogenesis by modulating E-cadherin/beta-catenin signaling via its FadA adhesin. *Cell Host Microbe.* 2013;14(2):195–206.
28. Gibbs AR, Thunnissen FB. Histological typing of lung and pleural tumours: third edition. *J Clin Pathol.* 2001;54(7):498–9.
29. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, Chirieac LR, Dacic S, Duhig E, Flieder DB, et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. *J Thorac Oncol.* 2015;10(9):1243–60.
30. Chansky K, Detterbeck FC, Nicholson AG, Rusch VW, Vallieres E, Groome P, Kennedy C, Krasnik M, Peake M, Shemanski L, et al. The IASLC Lung Cancer Staging Project: External Validation of the Revision of the TNM Stage Groupings in the Eighth Edition of the TNM Classification of Lung Cancer. *J Thorac Oncol.* 2017;12(7):1109–21.

31. Ruffini E, Fang W, Guerrera F, Huang J, Okumura M, Kim DK, Girard N, Bille A, Boubia S, Cangir AK, et al. The International Association for the Study of Lung Cancer Thymic Tumors Staging Project: The Impact of the Eighth Edition of the Union for International Cancer Control and American Joint Committee on Cancer TNM Stage Classification of Thymic Tumors. *J Thorac Oncol*. 2020;15(3):436–47.
32. Amemiya K, Hirotsu Y, Goto T, Nakagomi H, Mochizuki H, Oyama T, Omata M. Touch imprint cytology with massively parallel sequencing (TIC-seq): a simple and rapid method to snapshot genetic alterations in tumors. *Cancer Med*. 2016;5(12):3426–36.
33. Goto T, Hirotsu Y, Amemiya K, Nakagomi T, Shikata D, Yokoyama Y, Okimoto K, Oyama T, Mochizuki H, Omata M. Distribution of circulating tumor DNA in lung cancer: analysis of the primary lung and bone marrow along with the pulmonary venous and peripheral blood. *Oncotarget*. 2017;8(35):59268–81.
34. Higuchi R, Nakagomi T, Goto T, Hirotsu Y, Shikata D, Yokoyama Y, Otake S, Amemiya K, Oyama T, Mochizuki H, et al: **Identification of Clonality through Genomic Profile Analysis in Multiple Lung Cancers**. *J Clin Med* 2020, 9(2).
35. Nakagomi T, Goto T, Hirotsu Y, Shikata D, Yokoyama Y, Higuchi R, Otake S, Amemiya K, Oyama T, Mochizuki H, et al: **Genomic Characteristics of Invasive Mucinous Adenocarcinomas of the Lung and Potential Therapeutic Targets of B7-H3**. *Cancers (Basel)* 2018, 10(12).
36. Nakagomi T, Hirotsu Y, Goto T, Shikata D, Yokoyama Y, Higuchi R, Otake S, Amemiya K, Oyama T, Mochizuki H, et al: **Clinical Implications of Noncoding Indels in the Surfactant-Encoding Genes in Lung Cancer**. *Cancers (Basel)* 2019, 11(4).
37. Oyama T, Goto T, Amemiya K, Hirotsu Y, Omata M. Squamous Cell Carcinoma of the Lung With Micropapillary Pattern. *J Thorac Oncol*. 2020;15(9):1541–4.
38. Goto T, Hirotsu Y, Oyama T, Amemiya K, Omata M. Analysis of tumor-derived DNA in plasma and bone marrow fluid in lung cancer patients. *Med Oncol*. 2016;33(3):29.
39. Higuchi R, Goto T, Hirotsu Y, Yokoyama Y, Nakagomi T, Otake S, Amemiya K, Oyama T, Mochizuki H, Omata M. **Primary Driver Mutations in GTF2I Specific to the Development of Thymomas**. *Cancers (Basel)* 2020, 12(8).
40. Hujoel PP, Drangsholt M, Spiekerman C, Weiss NS. An exploration of the periodontitis-cancer association. *Ann Epidemiol*. 2003;13(5):312–6.
41. Yang J, Mu X, Wang Y, Zhu D, Zhang J, Liang C, Chen B, Wang J, Zhao C, Zuo Z, et al. Dysbiosis of the Salivary Microbiome Is Associated With Non-smoking Female Lung Cancer and Correlated With Immunocytochemistry Markers. *Front Oncol*. 2018;8:520.
42. Maddi A, Sabharwal A, Violante T, Manuballa S, Genco R, Patnaik S, Yendamuri S. The microbiome and lung cancer. *J Thorac Dis*. 2019;11(1):280–91.
43. Ramirez-Labrada AG, Isla D, Artal A, Arias M, Rezusta A, Pardo J, Galvez EM. The Influence of Lung Microbiota on Lung Carcinogenesis, Immunity, and Immunotherapy. *Trends Cancer*. 2020;6(2):86–97.

44. Xu N, Wang L, Li C, Ding C, Li C, Fan W, Cheng C, Gu B. Microbiota dysbiosis in lung cancer: evidence of association and potential mechanisms. *Transl Lung Cancer Res.* 2020;9(4):1554–68.
45. Stolz A. Molecular characteristics of xenobiotic-degrading sphingomonads. *Appl Microbiol Biotechnol.* 2009;81(5):793–811.
46. Aso Y, Miyamoto Y, Harada KM, Momma K, Kawai S, Hashimoto W, Mikami B, Murata K. Engineered membrane superchannel improves bioremediation potential of dioxin-degrading bacteria. *Nat Biotechnol.* 2006;24(2):188–9.
47. Miller TR, Delcher AL, Salzberg SL, Saunders E, Detter JC, Halden RU. Genome sequence of the dioxin-mineralizing bacterium *Sphingomonas wittichii* RW1. *J Bacteriol.* 2010;192(22):6101–2.
48. Asaf S, Numan M, Khan AL, Al-Harrasi A. *Sphingomonas*: from diversity and genomics to functional role in environmental remediation and plant growth. *Crit Rev Biotechnol.* 2020;40(2):138–52.
49. Dong Z, Chen B, Pan H, Wang D, Liu M, Yang Y, Zou M, Yang J, Xiao K, Zhao R, et al. Detection of Microbial 16S rRNA Gene in the Serum of Patients With Gastric Cancer. *Front Oncol.* 2019;9:608.
50. Huang YF, Chen YJ, Fan TC, Chang NC, Chen YJ, Midha MK, Chen TH, Yang HH, Wang YT, Yu AL, et al. Analysis of microbial sequences in plasma cell-free DNA for early-onset breast cancer patients and healthy females. *BMC Med Genomics.* 2018;11(Suppl 1):16.
51. Zuo J, Hu L, Shen W, Zeng J, Li L, Song L, Gan N. **The involvement of  $\alpha$ -proteobacteria *Phenylobacterium* in maintaining the dominance of toxic *Microcystis* blooms in Lake Taihu, China.** *Environ Microbiol* 2020.
52. Li X, Yu Y, Choi L, Song Y, Wu M, Wang G, Li M. *Phenylobacterium soli* sp. nov., isolated from arsenic and cadmium contaminated farmland soil. *Int J Syst Evol Microbiol.* 2019;69(5):1398–403.
53. Yang S, Wen X, Zhao L, Shi Y, Jin H. Crude oil treatment leads to shift of bacterial communities in soils from the deep active layer and upper permafrost along the China-Russia Crude Oil Pipeline route. *PLoS One.* 2014;9(5):e96552.
54. Huang Y, Pan H, Wang Q, Ge Y, Liu W, Christie P. Enrichment of the soil microbial community in the bioremediation of a petroleum-contaminated soil amended with rice straw or sawdust. *Chemosphere.* 2019;224:265–71.
55. Aron-Wisnewsky J, Warmbrunn MV, Nieuwdorp M, Clément K. **Metabolism and Metabolic Disorders and the Microbiome: The intestinal microbiota associated with obesity, lipid metabolism and metabolic health: pathophysiology and therapeutic strategies.** *Gastroenterology* 2020.
56. Sommariva M, Le Noci V, Bianchi F, Camelliti S, Balsari A, Tagliabue E, Sfondrini L. The lung microbiota: role in maintaining pulmonary immune homeostasis and its implications in cancer development and therapy. *Cell Mol Life Sci.* 2020;77(14):2739–49.
57. Liou JM, Malfertheiner P, Lee YC, Sheu BS, Sugano K, Cheng HC, Yeoh KG, Hsu PI, Goh KL, Mahachai V, et al. Screening and eradication of *Helicobacter pylori* for gastric cancer prevention: the Taipei global consensus. *Gut.* 2020;69(12):2093–112.
58. Yu G, Gail MH, Consonni D, Carugno M, Humphrys M, Pesatori AC, Caporaso NE, Goedert JJ, Ravel J, Landi MT. Characterizing human lung tissue microbiota and its relationship to epidemiological and

clinical features. *Genome Biol.* 2016;17(1):163.

59. Demaneche S, Kay E, Gourbiere F, Simonet P. Natural transformation of *Pseudomonas fluorescens* and *Agrobacterium tumefaciens* in soil. *Appl Environ Microbiol.* 2001;67(6):2617–21.

## Tables

Table 1  
Patient Characteristics

Parameter	Number of patients	Overall percentage
Total number	19	
Age, median (range)	68 (42–81)	
Sex		
Male	11	57.9%
Female	8	42.1%
Histology		
Type A	5	26.3%
Type AB	3	15.8%
Type B1	5	26.3%
Type B2	4	21.1%
Type B3	2	10.5%
Tumor size (cm)		
≤ 3	7	36.8%
3 < size ≤ 5	8	42.1%
5 <	4	21.1%
Masaoka Stage		
I	7	36.8%
II	9	47.4%
III	2	10.5%
IV	1	5.3%
Smoking Status (Pack year)		
0	7	36.8%
0 < PY ≤ 30	8	42.1%
> 30	4	21.1%
Myasthenia gravis		
+	1	5.3%

Parameter	Number of patients	Overall percentage
-	18	94.7%

## Figures

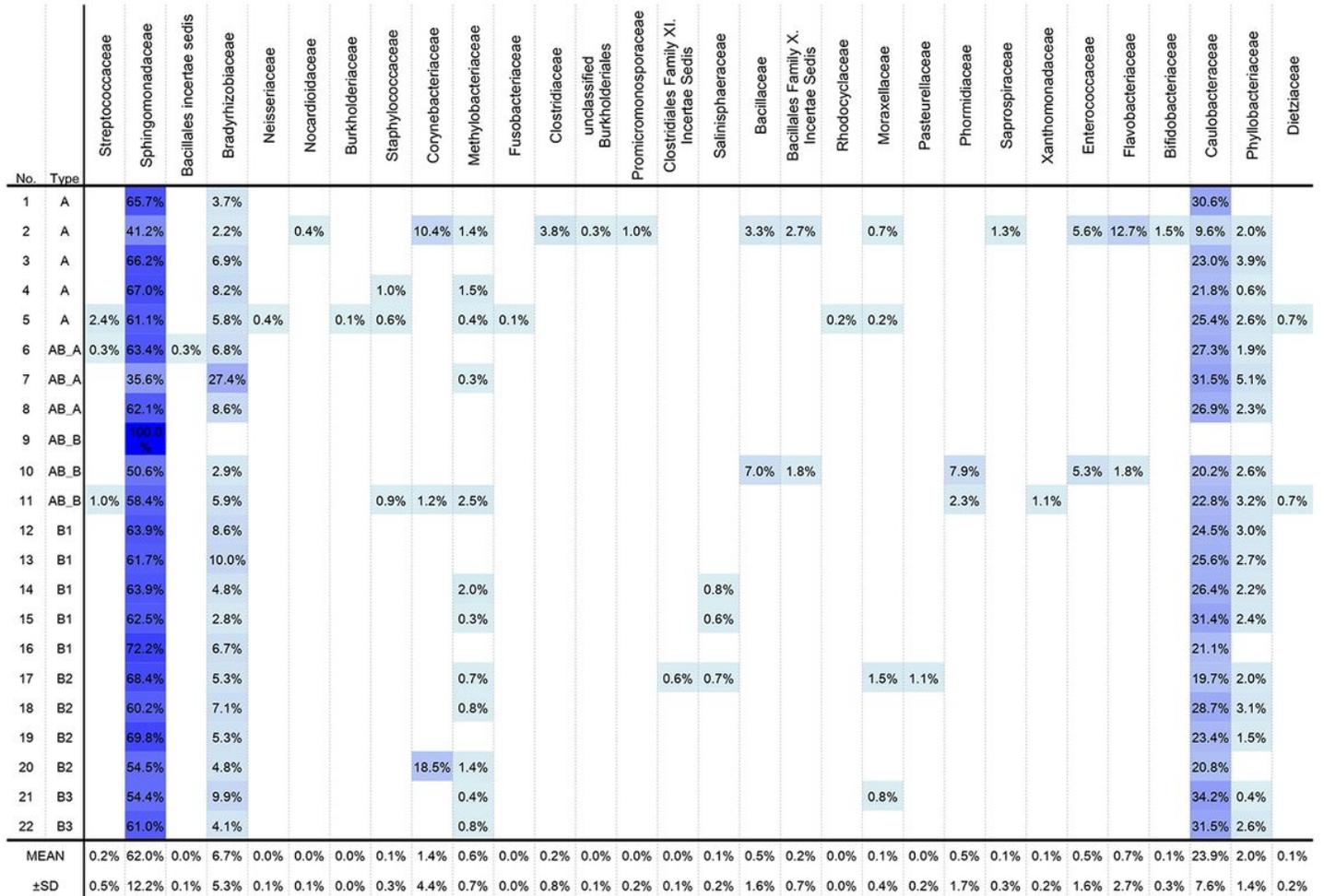


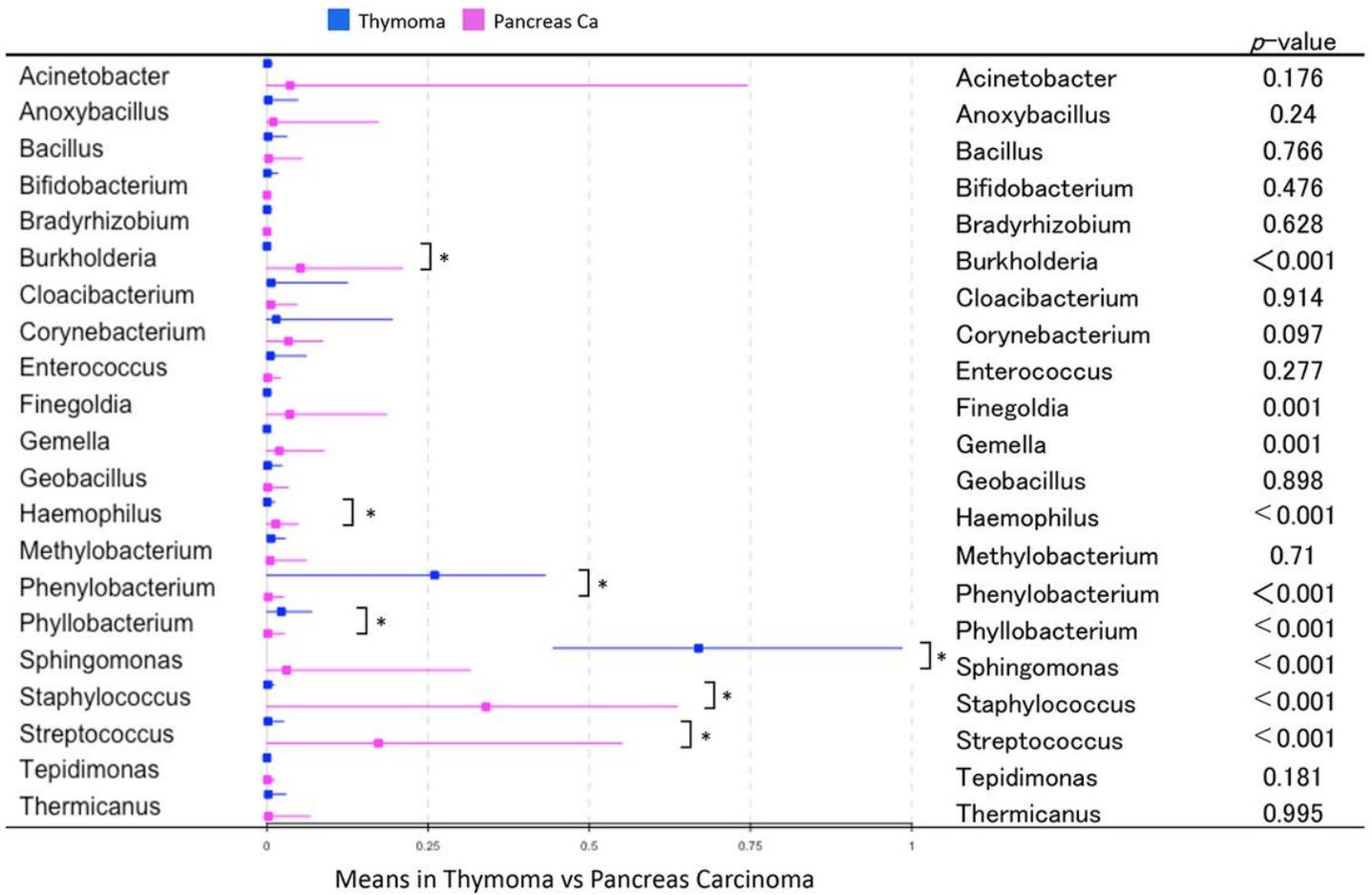
Figure 1

Composition and abundance of dominant families in all samples. The abundance values of the detected families are presented using a heatmap.

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22			
Type	A	A	A	A	A	AB_A	AB_A	AB_A	AB_B	AB_B	AB_B	B1	B1	B1	B1	B1	B2	B2	B2	B2	B3	B3			
Stage	II	I	III	I	I	I	II	I	I	I	I	II	II	II	II	II	IV	II	I	I	I	I			
Pack Year	0	2	0	1	2	1	1	0	1	1	0	1	0	2	0	0	1	0	1	1	2	1			
GTF2I	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	-	-	+	-	-	-	-	+		
PD-L1	-	+	-	-	+	-	-	-	-	+	+	+	-	+	-	+	+	+	+	+	+	+	+	MEAN	±SD
Acinetobacter	0.8%						0.2%				0.8%												0.1%	0.2%	
Anoxybacillus		4.8%																					0.2%	1.0%	
Bradyrhizobium													0.7%										0.0%	0.1%	
Phenylobacterium	10.4%	22.6%	24.4%	24.7%	21.3%	24.7%	26.6%	30.9%	28.4%	32.2%	37.5%	23.7%	43.2%	28.9%	27.9%	26.8%	31.8%			22.6%	32.4%	29.5%	21.9%	26.0%	8.7%
Tepidimonas	0.3%																						0.0%	0.1%	
Phyllobacterium	2.2%	2.9%	3.6%	1.6%	2.2%	4.2%	2.9%	3.4%	3.0%	2.5%	0.4%	0.6%	6.9%	2.1%	2.4%	3.3%					2.7%	2.6%	2.2%	1.7%	
Sphingomonas	44.4%	56.7%	64.9%	73.8%	74.0%	71.1%	66.3%	64.8%	68.6%	65.0%	60.8%	73.0%	48.8%	68.4%	67.6%	69.9%	68.2%	100.0%	77.4%	64.0%	68.0%	57.2%	66.9%	10.8%	
Staphylococcus			1.0%				0.6%					1.1%											0.1%	0.3%	
Bifidobacterium	1.6%																						0.1%	0.3%	
Bacillus	1.2%	3.1%																					0.2%	0.7%	
Haemophilus					1.2%																		0.1%	0.3%	
Finegoldia					0.6%																		0.0%	0.1%	
Gemella														0.3%									0.0%	0.1%	
Chryseobacterium	1.2%																						0.1%	0.3%	
Corynebacterium	11.2%		1.3%																			19.4%	1.5%	4.7%	
Streptococcus			1.1%				2.6%							0.3%									0.2%	0.6%	
Thermicanus	2.9%	2.0%																					0.2%	0.7%	
Isoptericola	1.1%																						0.0%	0.2%	
Geobacillus	2.3%																						0.1%	0.5%	
Burkholderia							0.1%																0.0%	0.0%	
Enterococcus	6.1%	5.9%																					0.5%	1.8%	
Pseudoxanthomonas			0.9%																				0.0%	0.2%	
Methyloversatilis							0.2%																0.0%	0.0%	
Methylobacterium	1.5%		2.8%		0.7%		0.5%	0.9%		0.3%	0.4%	1.6%	0.4%		2.1%						0.9%	1.4%	0.6%	0.8%	
Nocardioides	0.4%																						0.0%	0.1%	
Cloacibacterium	12.5%	2.0%																					0.7%	2.7%	

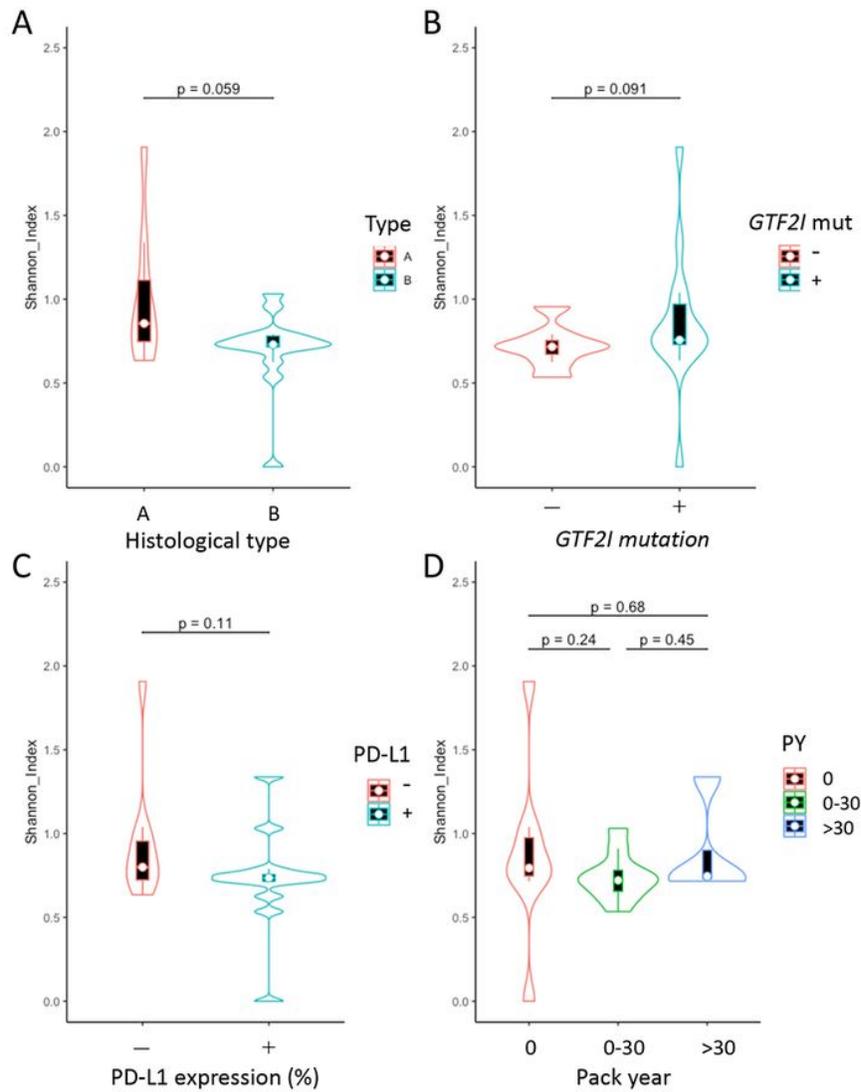
**Figure 2**

Composition and abundance of dominant genera in all samples. In total, 26 genera were identified. The heatmap visualizes the abundance of detected genera. PY (pack-year) 0 represents non-smokers; PY1, smokers with >0 to ≤30 pack-year history; and PY2, smokers with >30 pack-year history.



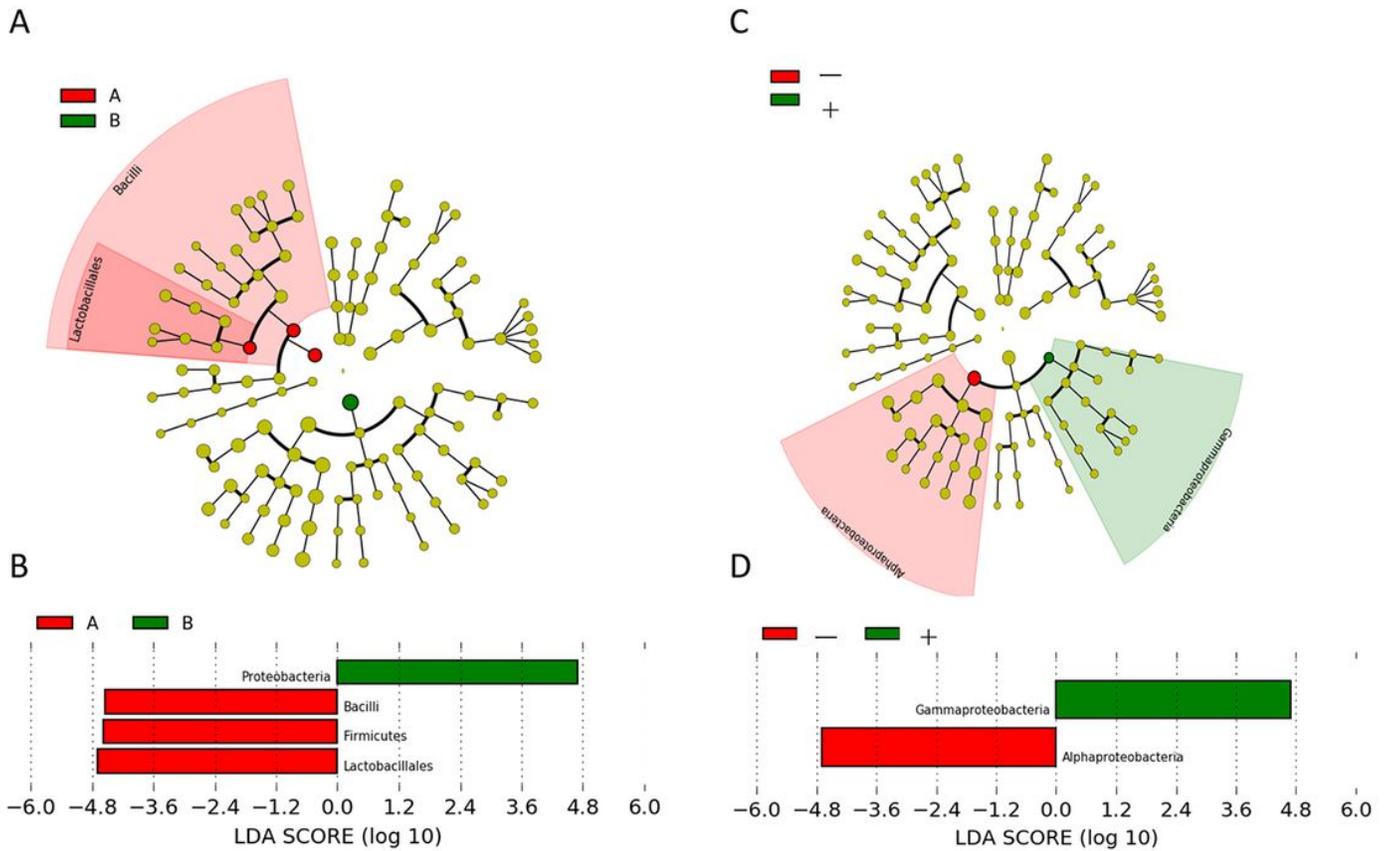
**Figure 3**

Microbiome differences between thymoma and pancreatic cancer samples. \*p<0.05



**Figure 4**

Taxonomic alpha diversity of thymoma microbiomes within samples in different groups. (A) Comparison of the Shannon index between Type A and B histology groups. (B) Comparison of the Shannon index between tumors exhibiting the presence and those exhibiting an absence of GTF2I driver mutation. (C) Comparison of the Shannon index between tumors exhibiting the presence and those exhibiting an absence of PD-L1 expression on tumor cells. (D) Comparison of the Shannon index among non-smokers, light smokers, and heavy smokers. No significant difference was found among these groups.



**Figure 5**

Differential taxa in the histology and driver mutation groups. (A) Cladogram of differential taxa between Type A and B histology. Dominant taxa are indicated in red for the Type A group and in green for the Type B group. (B) Kruskal–Wallis test results on the relative abundance between Type A and Type B histology. Type A is presented in the red column and Type B in the green column. \* $p < 0.05$ . (C) The result of LefSe analysis between tumors exhibiting the presence and those exhibiting the absence of GTF2I driver mutation. (D) Kruskal–Wallis test results for relative abundance between tumors exhibiting the presence and those exhibiting the absence of GTF2I driver mutation. \* $p < 0.05$ .

## Supplementary Files

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

- [SuppleFigure1.pdf](#)
- [SuppleFigure2.pdf](#)

