Am J Cancer Res 2015;5(10):3111-3122
www.ajcr.us /ISSN:2156-6976/ajcr0009363

*Original Article*

# Discovery and validation of potential bacterial biomarkers for lung cancer

Xinmin Yan[1*], Mingxia Yang[1*], Juan Liu[1], Ruichen Gao[1], Jihong Hu[1], Jiong Li[2], Lijun Zhang[3], Yujia Shi[1], Hongrong Guo[1], Jinluo Cheng[1], Miriam Razi[2], Shen Pang[3], Xiaowei Yu[1], Shen Hu[2]

[1]*Changzhou Second People's Hospital, Nanjing Medical University, Changzhou, 213003, China;* [2]*School of Dentistry and Jonsson Comprehensive Cancer Center, University of California, Los Angeles, CA 90095, USA;* [3]*David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA.* [*]*Equal contributors.*

Received April 20, 2015; Accepted June 25, 2015; Epub September 15, 2015; Published October 1, 2015

**Abstract:** Microbes are residents in a number of body sites, including the oral and nasal cavities, which are connected to the lung via the pharynx. The associations between oral diseases and increased risk of lung cancer have been reported in previous prospective studies. In this study, we measured variations of salivary microbiota and evaluated their potential association with lung cancer, including squamous cell carcinoma (SCC) and adenocarcinoma (AC). A three-phase study was performed: First, we investigated the salivary microbiota from 20 lung cancer patients (10 SCC and 10 AC) and control subjects (n=10) using a deep sequencing analysis. Salivary *Capnocytophaga*, *Selenomonas*, *Veillonella* and *Neisseria* were found to be significantly altered in patients with SCC and AC when compared to that in control subjects. Second, we confirmed the significant changes of *Capnocytophaga*, *Veillonella* and *Neisseria* in the same lung cancer patients using quantitative PCR (qPCR). Finally, these bacterial species were further validated on new patient/control cohorts (n=56) with qPCR. The combination of two bacterial biomarkers, *Capnocytophaga* and *Veillonella*, yielded a receiver operating characteristic (ROC) value of 0.86 with an 84.6% sensitivity and 86.7% specificity in distinguishing patients with SCC from control subjects and a ROC value of 0.80 with a 78.6% sensitivity and 80.0% specificity in distinguishing patients with AC from control subjects. In conclusion, we have for the first time demonstrated the association of saliva microbiota with lung cancer. Particularly, the combination of the 16S sequencing discovery with qPCR validation studies revealed that the levels of *Capnocytophaga* and *Veillonella* were significantly higher in the saliva from lung cancer patients, which may serve as potential biomarkers for the disease detection/classification.

**Keywords:** Lung cancer, 16S rDNA sequencing, *Capnocytophaga, Veillonella,* saliva microbiota

## Introduction

Lung cancer is considered a terminal illness, with a five-year survival rate of about 11% [1]. It is the most common cause of cancer-related deaths in North America and worldwide [2, 3]. Tobacco smoking represents one of the main risk factors of lung cancer, since tobacco contains carcinogens that may induce cell transformation [4]. However, studies have also demonstrated that lung cancer involves immune responses [5], viral infections [6], and other factors that are not related to tobacco [6, 7].

Recent studies have reported the identification of potential biomarkers in saliva, including oral bacteria, for cancer detection [8-13]. Particularly, salivary microbiota has been associated with oral squamous cell carcinoma and pancreatic ductal adenocarcinoma [12, 13], suggesting that salivary microbiota may serve as an informative source for discovering non-invasive biomarkers of cancer diseases. Earlier studies have also shown serological evidence of an association between *Chlamydia pneumoniae* infection and lung cancer, and *Chlamydia pneumoniae* infection may serve as a biomarker for increased risk of lung cancer [14, 15]. Most salivary bacteria are derived from the oral cavity, although some may also arise from the esophagus and the upper respiratory tract. Survival and growth of bacteria in the oral cavity are dependent on the oral environment, which can be affected by the composition of the

Table 1. Enrolled lung cancer (n=61) and control (n=25) subjects for the present study

| | Group | Subjects under study | | | Tumor stage | | |
|---|---|---|---|---|---|---|---|
| | | n | Male | Age | T4 | T3 | T2 |
| Discovery | SCC | 10 | 9 | 67.4±8.99 | 6 | 3 | 1 |
| | AC | 10 | 8 | 61.3±15.13 | 7 | 3 | 0 |
| | Controls | 10 | 8 | 59.20±9.84 | | N/A | |
| Validation | SCC | 13 | 12 | 69.8±6.03 | 12 | 1 | 0 |
| | AC | 28 | 24 | 63.8±11.78 | 24 | 2 | 2 |
| | Controls | 15 | 10 | 57.5±9.01 | | N/A | |

Table 2. Primers for qPCR of 16S rDNA

| Name | Taxonomy | Sequences |
|---|---|---|
| *Bacteroidales* | Class | F WTYATTGGGTTTAAAGGG |
| | | R GGTAAGGTTCCTCGCGTA |
| *Neisseria* | Genus | F CTGTTGGGCARCWTGAYTGC |
| | | R GATCGGTTTTRTGAGATTGG |
| *Capnocytophaga* | Genus | F TGGWCAATGGTCGGAAGACTG |
| | | R CCGCTACACTACACATTCCA |
| *Selenomonas* | Genus | F ACRCGTAGRCAACCTGCCG |
| | | R CGATCCGAAGACCTTCTTCAC |
| *Veillonella* | Genus | F CGGGTGAGTAACGCGTAATCA |
| | | R CCAACTAGCTGATGGGACGC |
| Common | For all bacteria | F ATTAGATACCCTGGTAGTCC |
| | | R CCCCGTCAATTCATTTGAGT |

sputum, and habits such as cigarette smoking. These factors are related to lung cancer occurrence and progression.

Based on these considerations, as well as previously reported involvement of bacteria in lung cancer tumorigenesis [14-16], we investigated the global variations of salivary microbiota in lung cancer patients. We quantified bacterial flora composition of saliva samples obtained from lung cancer patients and controls by sequencing V3 and V6 of 16S rDNA, using the Illumina HiSeq 2000 sequencer. This approach can identify more than 500 prevalent human bacterial species, with up to 100,000 bacterial sequences per sample. This high sensitivity and comprehensive analyses provide a new approach to investigating the relationship between lung cancer and bacterial composition.

Non-small cell lung cancer (NSCLC) accounts for more than 80% of lung cancer cases, most of which are either squamous cell carcinoma (SCC, ~30-35%) or adenocarcinoma (AC, ~50%). Therefore, we mainly enrolled patients with SCC or AC for the present study. To discover candidate bacterial biomarkers for SCC or AC, we performed the sequencing analysis of 10 samples from SCC patients, 10 samples from patients with AC and 10 samples from control subjects (n=30), and identified bacteria at the genus level that showed significant differences between both lung cancer groups and the control group. The results were subsequently confirmed using qPCR to quantify the relative abundance of bacteria. Finally, we further validated these potential bacterial biomarkers on new patient/control cohorts including a total of 41 lung cancer patients and 15 contro

## Materials and methods

### Sample collection

The study was approved by the Medical Ethics Committee at the Changzhou Second People's Hospital (CSPH), Nanjing Medical University, and all saliva samples were collected according to the approved protocol. Written informed consent was obtained from all participants. Since smoking may affect the growth of bacteria in the oral cavity, only subjects with a smoking history of over 10 years were selected. Saliva samples were obtained after cancer detection and before treatment. None of the subjects, including the patients and the controls, manifested other diseases related to salivary bacteria, such as diabetes, immune dysfunction, herpes viral infections, or oral mucosal ulcers. Patient information is summarized in **Table 1**.

### Sequencing analysis of bacterial flora of saliva samples

DNA content in saliva was isolated by SDS lysis, following phenol extraction to remove proteins in saliva. The DNA content of samples recovered by ethanol precipitation was amplified using PCR, as previously described [17]. The amplified samples were sequenced using the Illumina HiSeq 2000, and V3 and V6 of 16S rDNA were analyzed. To ensure high quality

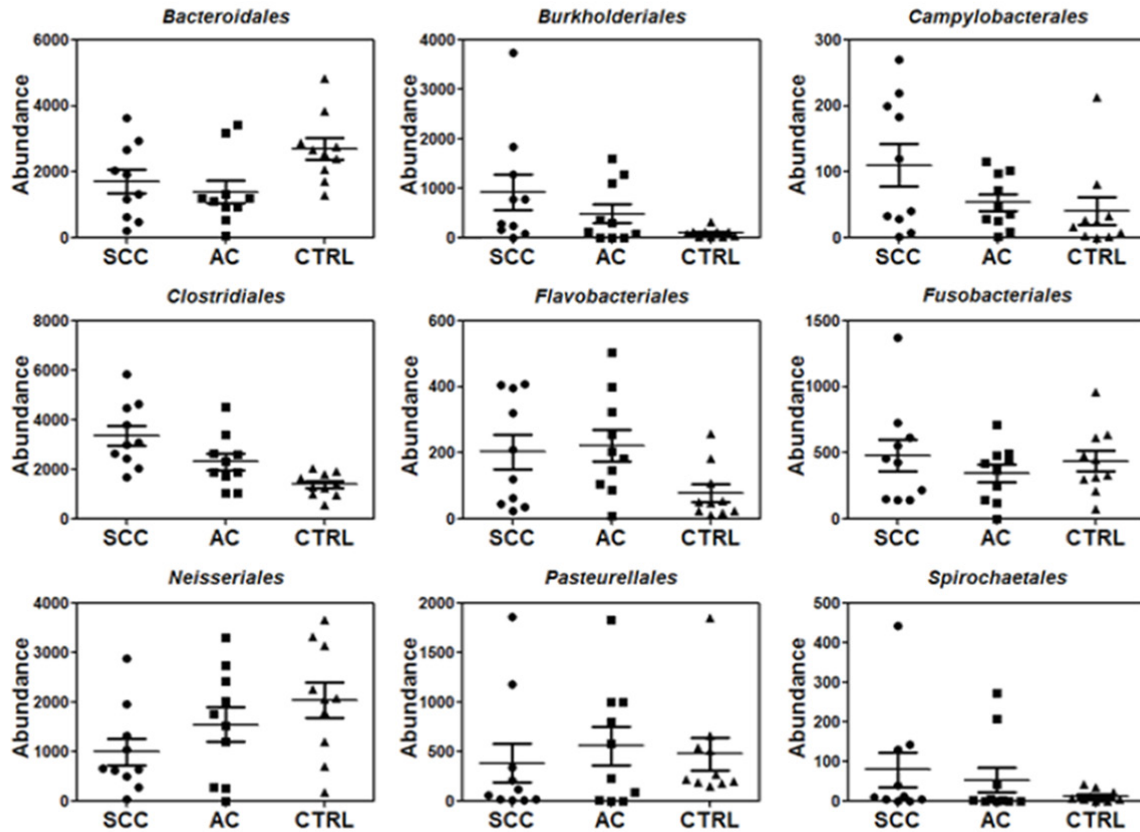


**Figure 1.** Relative abundance for a partial list of salivary bacteria at the order level among SCC, AC and control groups (n=10 each group).

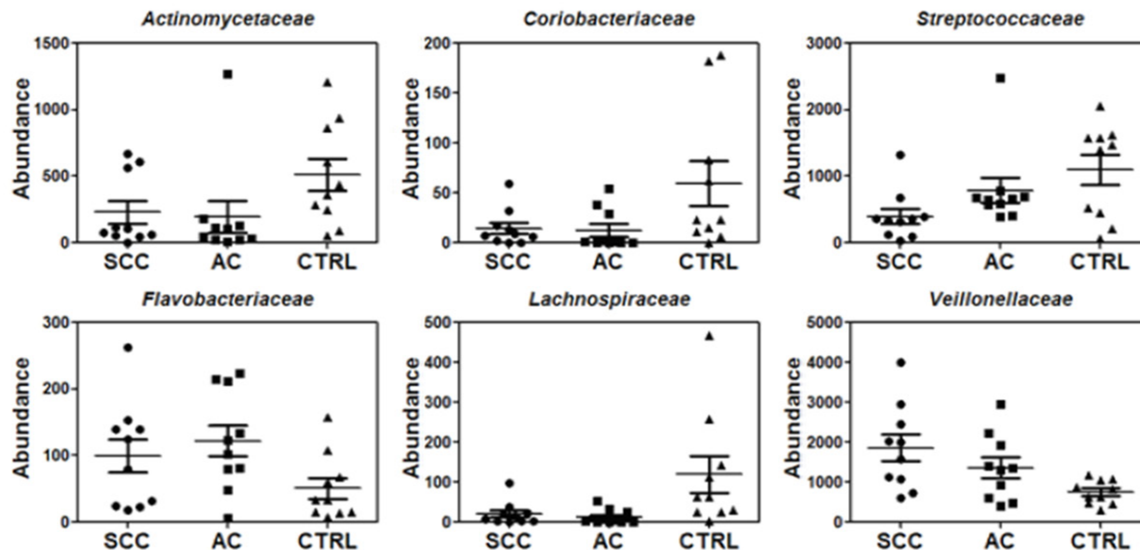


**Figure 2.** Relative abundance for a partial list of saliva bacteria at the family level among SCC, AC and control groups (n=10 each group).

readings, we adopted stringent conditions to process the sample sequences. In brief, the following processing steps were performed: 1) all readings were assigned to corresponding samples by allowing one mismatch to the sample barcode and two mismatches to the adjacent

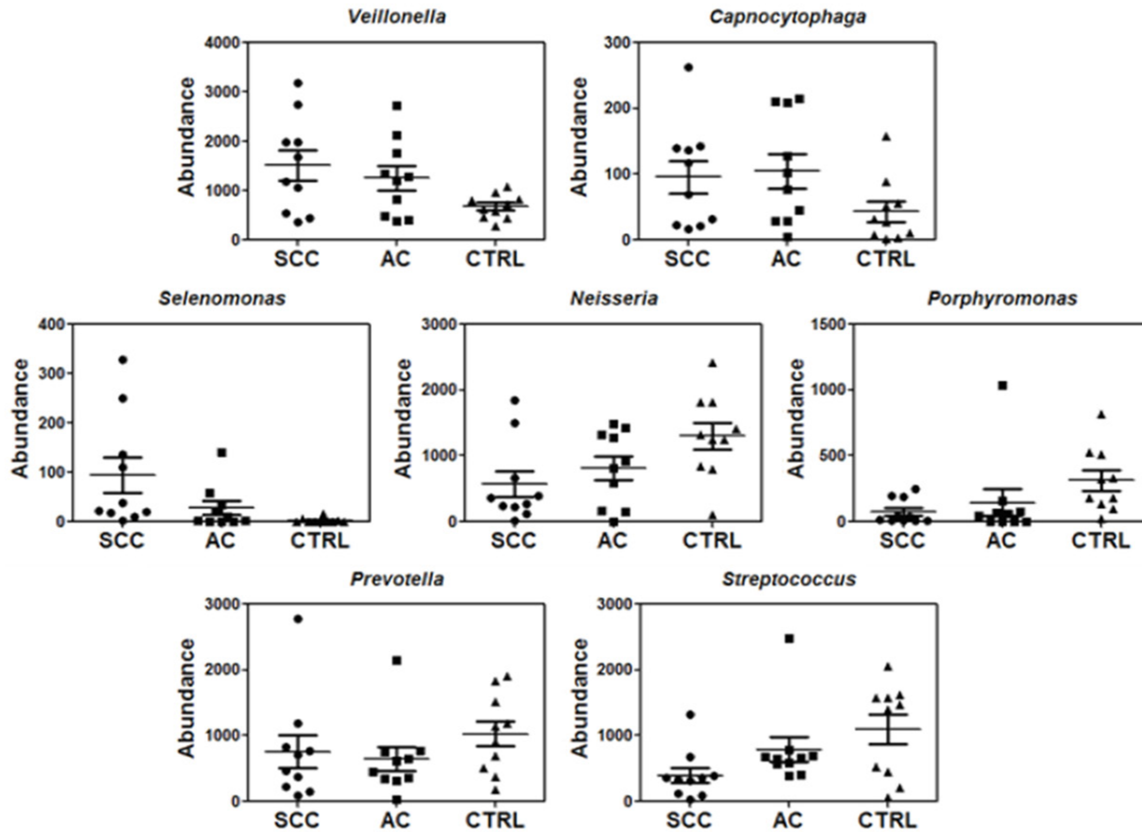


**Figure 3.** Relative abundance for a partial list of saliva bacteria at the genus level among SCC, AC and control groups (n=10 each group).

PCR primer; 2) the readings were subsequently de-noised by the PyroNoise algorithm [18]; 3) readings containing ambiguous nucleotides or a homopolymer longer than 8 base pairs (bp) were removed, as were sequences shorter than 200 bp or longer than 1000 bp; 4) the readings were aligned using a nearest alignment space termination (NAST)-based sequence aligner to a custom reference based on the SILVA alignment [19], and sequences that did not align to the anticipated region of the reference alignment were discarded; 5) chimeric sequences identified by the UCHIME algorithm were removed [20]; and 6) readings were classified using a Bayesian classifier with the Ribosomal Database Project (RDP). Sequences of mitochondria or unknowns (those readings that could not be classified at the kingdom level) were removed. Finally, all of the effective readings were clustered into operational taxonomic units (OTUs) based on 97% sequence similarity, using the MOTHUR program [21]. The taxonomy profiling of samples at different taxonomic levels (phylum, class, order, family, genus) were created using QIIME [22].

*Quantitative PCR*

The candidate bacterial biomarkers identified from the sequencing analysis were further confirmed by qPCR on the same set of saliva samples used in the discovery phase (n=30). Specific PCR primers to detect 16S rDNA were designed and shown in **Table 2**. These primers were chosen from previously published literature [23] or by searching the RDP database. Verified potential biomarkers were further validated on a new patient (n=41) and control (n=15) cohort using qPCR.

*Statistical analyses*

Data analysis was performed with the GraphPad Prism and MedCalc software. Receiver operating characteristic (ROC) analysis was used to estimate the performance of potential biomarkers and evaluate if combining biomarkers

Table 3. Relative abundance of a partial list of salivary bacteria at the order level between SCC and control groups or between AC and control groups based on sequencing analysis (n=10 per group)

| Taxonomy/Order | Fold change (SCC vs CTRL) | *P* value (SCC vs CTRL) | Fold change (AC vs CTRL) | *P* value (AC vs CTRL) |
|---|---|---|---|---|
| *Bacteroidales* | 0.63 | 0.03 | 0.52 | 0.006 |
| *Flavobacteriales* | 2.58 | 0.02 | 2.81 | 0.009 |
| *Fusobacteriales* | 1.1 | 0.39 | 0.79 | 0.19 |
| *Burkholderiales* | 9.01 | 0.02 | 4.76 | 0.03 |
| *Neisseriales* | 0.49 | 0.02 | 0.76 | 0.17 |
| *Campylobacterales* | 2.71 | 0.04 | 1.32 | 0.3 |
| *Pasteurellales* | 0.81 | 0.36 | 1.16 | 0.38 |
| *Spirochaetales* | 5.76 | 0.08 | 3.88 | 0.12 |
| *Clostridiales* | 2.41 | 0.0002 | 1.66 | 0.01 |

Table 4. Relative abundance of a partial list of salivary bacteria at the family level between SCC and control groups or between AC and control groups based on sequencing analysis (n=10 per group)

| Taxonomy/Family | Fold change (SCC vs CTRL) | *P* value (SCC vs CTRL) | Fold change (AC vs CTRL) | *P* value (AC vs CTRL) |
|---|---|---|---|---|
| *Actinomycetaceae* | 0.45 | 0.08 | 0.38 | 0.04 |
| *Coriobacteriaceae* | 0.25 | 0.07 | 0.21 | 0.03 |
| *Flavobacteriaceae* | 1.96 | 0.11 | 2.4 | 0.01 |
| *Streptococcaceae* | 0.36 | 0.01 | 0.72 | 0.16 |
| *Lachnospiraceae* | 0.18 | 0.05 | 0.11 | 0.02 |
| *Veillonellaceae* | 2.46 | 0.006 | 1.8 | 0.02 |

Table 5. Relative abundance of a partial list of salivary bacteria at the genus level between SCC and control groups or between AC and control groups based on sequencing analysis (n=10 per group)

| Taxonomy/Genus | Fold change (SCC vs CTRL) | *P* value (SCC vs CTRL) | Fold change (AC vs CTRL) | *P* value (AC vs CTRL) |
|---|---|---|---|---|
| *Porphyromonas* | 0.24 | 0.005 | 0.46 | 0.1 |
| *Prevotella* | 0.74 | 0.2 | 0.63 | 0.08 |
| *Capnocytophaga* | 2.22 | 0.04 | 2.42 | 0.03 |
| *Neisseria* | 0.44 | 0.009 | 0.63 | 0.04 |
| *Streptococcus* | 0.36 | 0.007 | 0.72 | 0.16 |
| *Selenomonas* | 34.7 | 0.01 | 10.15 | 0.04 |
| *Veillonella* | 2.24 | 0.008 | 1.85 | 0.02 |

improves sensitivity and specificity of disease detection.

## Results

*Differential salivary microflora profiles between lung cancer and control subjects*

Our deep sequencing analysis indicated that the abundance of salivary bacteria is different between lung cancer and control groups. This can be seen at all taxonomic levels (**Figures 1-3**, Supplemental Figures 1-4). At the order-level, the most significantly altered ones included *Flavobacteriales, Burkholderiales, Campylobacterales, Spirochaetales* (more abundant in lung cancer) *and Bacteroidales* (less abundant in lung cancer) (**Table 3**). At the family-level, *Veillonellaceae* is signifi cantly over-expressed whereas *Lachnospiraceae* is under-expressed in lung cancer patients (**Table 4**). More importantly, at the genus level, *Capnocytophaga*, *Selenomonas*, and *Veillonella* were found to be more abundant in both SCC and AC patients whereas *Neisseria* was less abundant in both SCC and AC patients than the controls (**Table 5**). On the other hand, the levels of *Steptococcus,* and *Porphyromonas* were only significantly lower in patients with SCC and *Prevotella* was not significantly changed in either cancer groups. All the sequencing results at different taxonomy levels (Phylums, Class, Order, Family and Genus) are shown in Supplemental Figures 1-4.

*qPCR verification of the candidate bacterial biomarkers*

To verify the finding from sequencing analysis, we used qPCR to quantify the levels of the identified bacteria genera in the same cancer and control subjects used for sequencing analysis. A common primer pair that can amplify most bacterial 16S rDNA was used for qPCR and subsequently the qPCR results based on the common primer pair were used to normalize the levels of individual bacterial rDNA in each control and cancer samples (**Table 2**). qPCR analysis showed significantly ($p<0.05$) higher levels of *Capnocytophaga* and

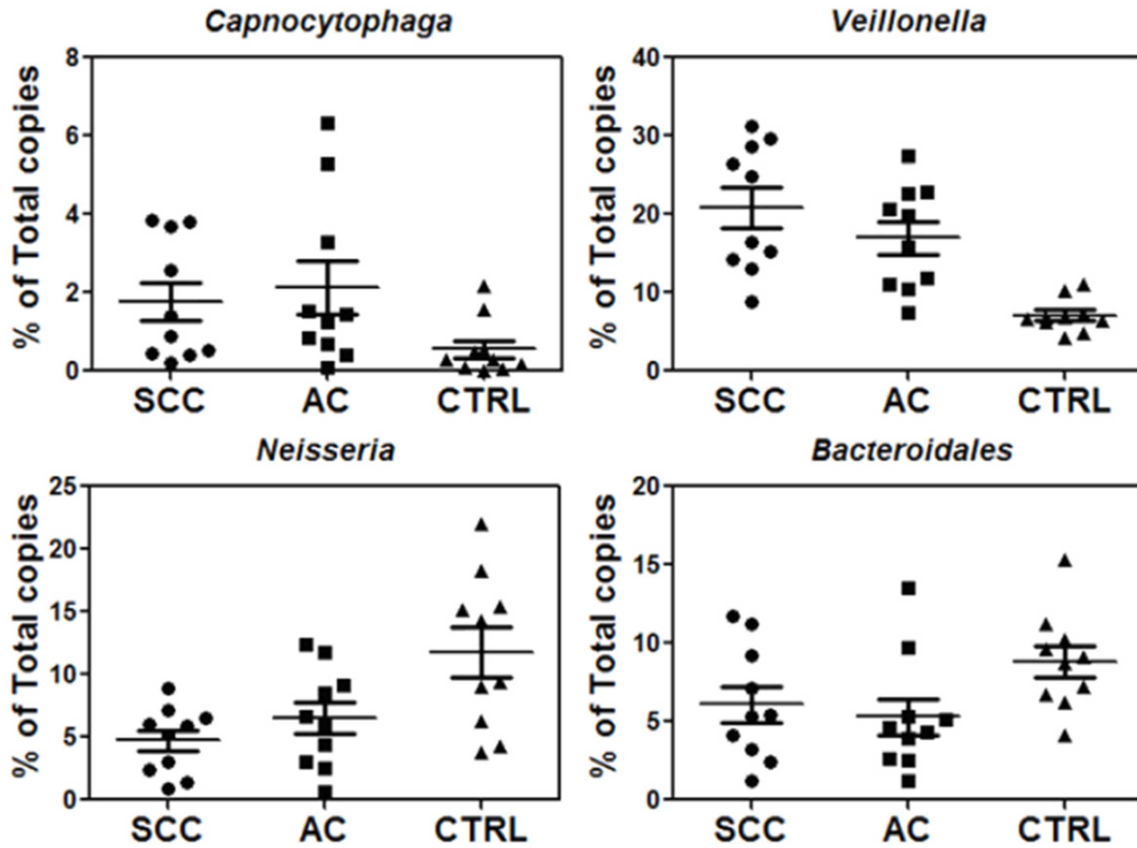


**Figure 4.** Verification of three potential biomarkers, *Capnocytophaga, Veillonella,* and *Neisseria*, as well as *Bacteroidales* by qPCR of 16S rDNA (n=10 each group).

**Table 6.** qPCR verification of three potential salivary bacterial biomarkers, Capnocytophaga, Veillonella, and Neisseria, as well as bacteroidales between lung cancer and control groups (n=10 per group)

| | Fold change (SCC vs CTRL) | *P* value (SCC vs CTRL) | Fold change (AC vs CTRL) | *P* value (AC vs CTRL) |
|---|---|---|---|---|
| *Neisseria* | 0.4 | 0.004 | 0.55 | 0.04 |
| *Capnocytophaga* | 3.18 | 0.04 | 3.78 | 0.04 |
| *Veillonella* | 2.94 | 0.00006 | 2.39 | 0.0003 |
| *Bacteroidales* | 0.69 | 0.09 | 0.6 | 0.03 |

*Veillonella* but lower levels of *Neisseria* in cancer samples than the control samples (**Figure 4**, **Table 6**). This confirms the sequencing analysis results and supports the feasibility of using qPCR to quantify the levels of selected bacterial clusters. Due to its very low level, *Selenomonas* was not detectable by qPCR in some of saliva samples, so it was not included in subsequent qPCR validation studies on independent patient populations. We also verified the alterations of *Bacteroidales* in SCC and AC patients with qPCR, and the results were similar to those obtained from the deep sequencing analysis (**Figure 4**, **Table 6**).

*Validation of potential bacterial biomarkers in new patient cohorts*

To validate the three potential bacterial markers (*Neisseria, Capnocytophaga* and *Veillonella*), we used qPCR to quantify their levels in a new patient/control cohort, including 41 cancers (13 SCC, 28 AC) and 15 controls. The qPCR results indicated that the levels of *Capnocytophaga* and *Veillonella* were significantly higher in both SCC and AC groups, consistent with our qPCR verification analysis of previous 30 samples (**Figure 5**). However, the validation result of *Neisseria* was not as promising as those found in the qPCR verification study. Although the levels of *Neisseria* in the AC

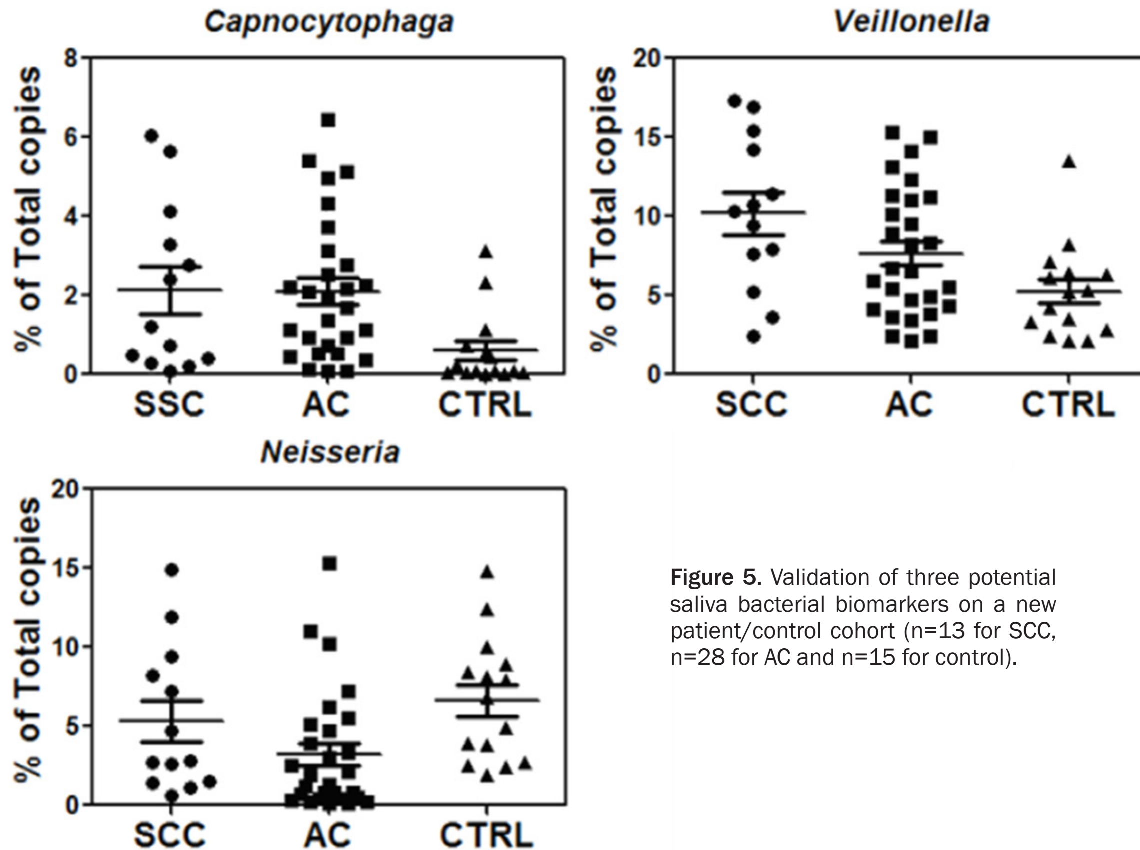


**Figure 5.** Validation of three potential saliva bacterial biomarkers on a new patient/control cohort (n=13 for SCC, n=28 for AC and n=15 for control).

patients were significantly lower (p=0.009) than those in the controls, the comparison of SCC with control groups did not show significance difference.

As shown in **Figure 6**, the ROC analysis was performed to evaluate pre-clinical utility of these potential biomarkers and to assess if combining biomarkers may improve the sensitivity and specificity. The overall performance of the three potential biomarkers in detecting SCC and AC is summarized in **Table 7**. The ROC values of *Veillonella* were determined to be 0.81 for SCC and 0.68 for AC, and the ROC values of *Capnocytophaga* were 0.79 for SCC and 0.81 for AC. ROC value or sensitivity can be improved by combining *Veillonella* with *Capnocytophaga*. However, the ROC values significantly decreased when adding the third potential biomarker, *Neisseria*, for both SCC and AC.

## Discussion

Our intensive sequencing analysis demonstrated that the salivary microbiota profiles are significant altered in the patients with lung cancer when compared to those from control subjects. Detailed analysis of the deep sequencing results showed that the abundance of salivary bacteria varied between lung cancer and control groups at multiple taxonomic levels. Particularly, at the genus-level, *Capnocytophaga*, *Selenomonas*, and *Veillonella* were found to be more abundant in both SCC and AC patients whereas *Neisseria* was less abundant in both SCC and AC patients than the controls (**Table 5**). In addition, the abundance of *Steptococcus,* and *Porphyromonas* were both significantly lower in patients with SCC. Based on these results, we performed qPCR verification of *Capnocytophaga*, *Selenomonas*, *Veillonella*, and *Neisseria* in the same group of SCC and AC patients used for deep sequencing analysis and further validated the significant changes of *Capnocytophaga*, *Veillonella* and *Neisseria* in an independent patient/control cohort.

Based on ROC analysis, *Veillonella* and *Capnocytophaga* are more valuable biomarkers than *Neisseria* because they exhibit higher sensitivity/specificity and their levels are significantly higher in both cancer types than the control

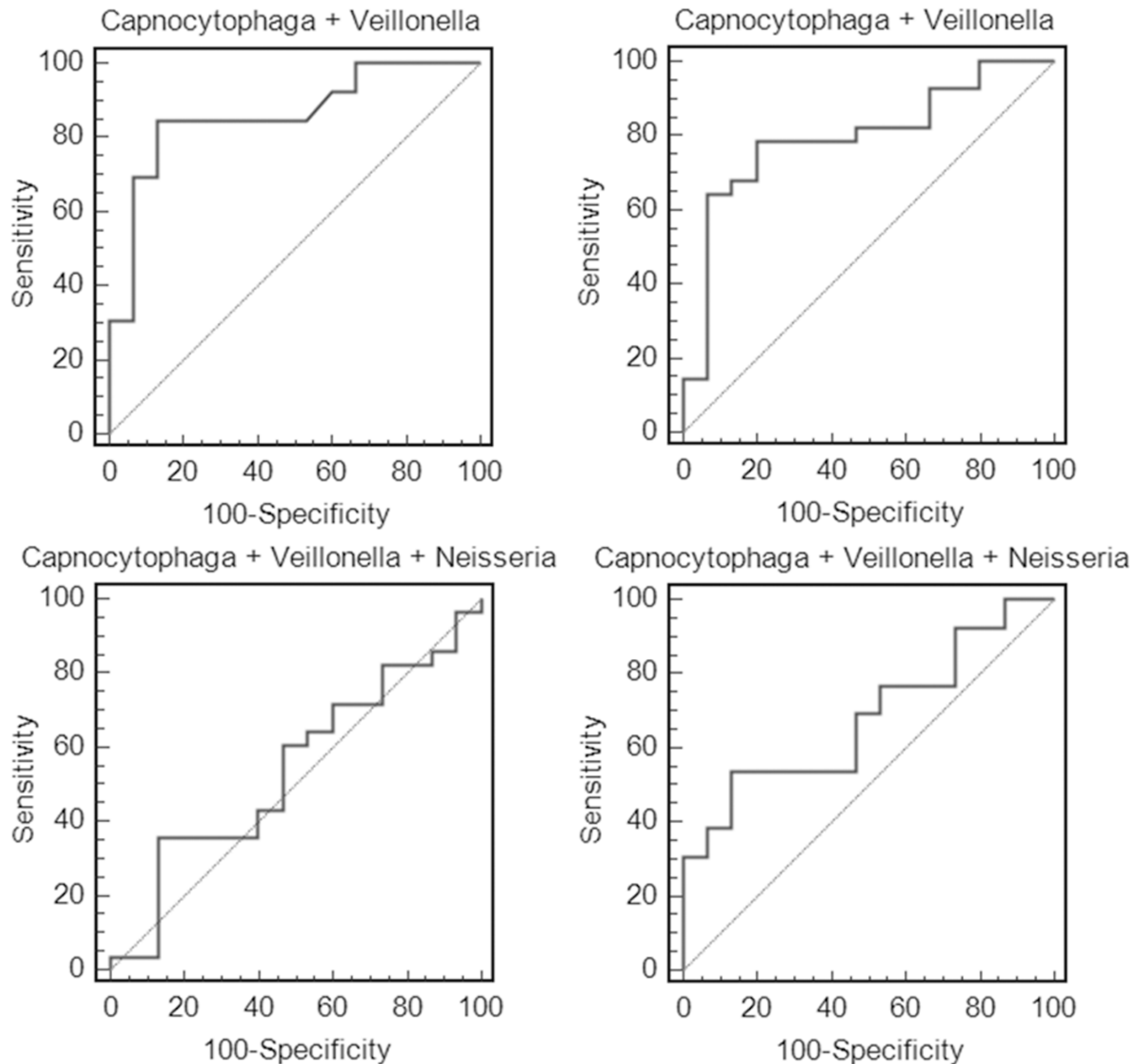


**Figure 6.** Receiver operating characteristic (ROC) analysis of the three potential biomarkers (*Capnocytophaga, Veillonella,* and *Neisseria*) based on the validation data. The combination of *Capnocytophaga with Veillonella (*but not *Neisseria)* improved the ROC value or sensitivity.

subjects (instead of being lower in cancer groups). In fact, if we combine *Veillonella* with *Capnocytophaga*, the ROC value reached 0.86 for SCC with a sensitivity of 84.6% and a specificity of 86.7%. Although the ROC value (0.80) became slightly lower than *Capnocytophaga* itself when combining these two biomarkers for AC, the sensitivity was improved (78.6%) and specificity reached 80%.

The intrinsic relationship between salivary microbiota and lung cancer is currently under investigation. Indeed, previous studies have provided critical insights on how bacteria may be related to abnormal growth of mammalian cells. Bacterium-produced toxins can disturb the cell cycle, resulting in altered cell growth [15, 24, 25]. This is caused by alterations to the genes that control normal cell division and apoptosis [26, 27]. It is well known that gut microbiota is involved in the pathogenesis of inflammatory bowel disease due to the disruption of immune tolerance [28]. Similarly, salivary microbiota may also affect lung cells by inducing long-term immune response.

In our study, although more research is needed to establish the relationship of *Neisseria*, *Steptococcus,* and *Porphyromonas* to lung cancer, our findings clearly show that there is sig-

**Table 7.** Performance of salivary bacterial markers for SCC or AC

| | *p* value (SCC vs CTRL) | *p* value (AC vs CTRL) | ROC value (SCC vs CTRL) | Cut-off (SCC vs CTRL) | Sensitivity (SCC vs CTRL) | Specificity (SCC vs CTRL) | ROC value (AC vs CTRL) | Cut-off (AC vs CTRL) | Sensitivity (AC vs CTRL) | Specificity (AC vs CTRL) |
|---|---|---|---|---|---|---|---|---|---|---|
| *Neisseria* | 0.260 | 0.009 | 0.62 | ≤1.5% | 30.80% | 100% | 0.78 | ≤2.1% | 57.10% | 93.30% |
| *Capnocytophaga* | 0.020 | 0.004 | 0.79 | ≥0.1% | 92.30% | 53.30% | 0.81 | ≥0.7% | 71.40% | 80.00% |
| *Veillonella* | 0.003 | 0.030 | 0.81 | ≥7.1% | 76.90% | 86.70% | 0.68 | ≥8.3% | 42.90% | 93.30% |
| *Capnocytophaga + Veillonella* | - | - | 0.86 | - | 84.60% | 86.70% | 0.80 | - | 78.60% | 80.00% |
| *Capnocytophaga + Veillonella + Neisseria* | - | - | 0.68 | - | 53.80% | 86.70% | 0.54 | - | 35.70% | 86.70% |

nificant association of salivary *Capnocytophaga* and *Veillonella* with lung cancer (both SCC and AC). In fact, there have been increasing interests in studying the link between oral bacteria and respiratory infection [29-33]. An association between oral conditions such as periodontal disease and several respiratory conditions has been reported previously [33]. Oral periodontopathic bacteria can be aspirated into the lung to cause aspiration pneumonia. The teeth may also serve as a reservoir for respiratory pathogen colonization and subsequent nosocomial pneumonia. Once established in the mouth, these pathogens may cause lung infection with the following possible mechanisms: 1. aspiration of oral pathogens into the lung to cause infection; 2. periodontal disease-associated enzymes in saliva may modify mucosal surfaces to promote adhesion and colonization by respiratory pathogens, which are then aspirated into the lung; 3. periodontal disease-associated enzymes may destroy salivary pellicles on pathogenic bacteria to hinder their clearance from the mucosal surface; and 4. cytokines originating from periodontal tissues may alter respiratory epithelium to promote infection by respiratory pathogens [33]. In patients with Cystic fibrosis (CF), there is a high prevalence of oropharyngeal anaerobic bacteria in sputum. Recently, Rivas Caldas *et al* demonstrated that the same *Pseudomonas aeruginosa* clonal types were present in both saliva and sputum samples of CF patients. This suggest that the oral cavity is a possible reservoir for lung infection [32]. *Haemophilus pittmaniae* (*H. pittmaniae*), a member of the human saliva microbiota, was identified by mass spectrometry and 16S rRNA gene sequencing in the sputum specimen collected from a patient with a massive fibrotic form of siderosis. *H. pittmaniae* was found to be responsible for the worsening of the patient's chronic respiratory failure, and the patient's condition rapidly improved when he was treated with oral amoxicillin, an antibiotic therapy guided by the in vitro susceptibility pattern of the *H. pittmaniae* isolate. These results suggest that *H. pittmaniae* is a possible new pathogen responsible for respiratory tract infection in patients with chronic lung diseases [30]. In addition, post-operative pulmonary infection often appears to result from aspiration of pathogens colonizing the oral cavity. In order to investigate if impaired periodontal status and pathogenic oral bacteria significantly contribute to development of aspiration pneumonia following neurosurgical operations, Bagyi *et al* compared a matched cohort of 18 patients without postoperative lung complications to a cohort of 5 patients who developed pneumonia within 48 hours after brain surgery. They found that the number and severity of coexisting periodontal diseases were significantly greater in patients with postoperative pneumonia in comparison to the control group. Therefore, dental examination of the salivary microbiota may be needed in order to identify patients at high risk of developing postoperative respiratory infections [29].

Although *Capnocytophaga* and *Veillonella* are significantly associated with both SCC and AC, *Veillonella* appears to be a better biomarker for SCC whereas *Capnocytophaga* serves as a better biomarker for AC (**Table 7**). *Veillonella* was previously isolated from the lower airways of lung cancer patients, suggesting that *Veillonella* species may be related to lung cancer [34]. Normally found in the oropharyngeal tract, *Capnocytophaga* species have been reported to be involved in lung cancer, as they have been shown to be involved in the formation of lung abscesses [35] and lower respiratory tract infections [36]. These findings, including ours, seem to suggest that either these bacteria induce long-term immune response/infection to the organ or cancer growth environment favors the growth of these bacteria in the airway or oropharyngeal tracts. As discussed earlier, *Neisseria* was found at significantly lower levels in both SCC and AC patients than control subjects. It should be noted that similar results were also reported in pancreatic cancer saliva samples [12], suggesting that *Neisseria* may be related to inhibited growth of cancer cells.

As a summary, our studies have provided insights on the potential role of oral cavity as a reservoir of bacterial pathogens in lung cancer, and the discovery of potential bacterial biomarkers may lead to a non-invasive method for helping detect/classify the disease. Nevertheless, the exact role of these salivary bacteria in lung cancer occurrence and progression is largely unknown and certainly warrants further investigation.

#### Acknowledgements

We thank Wendy Aft, Elliot Razi, and Ethan Razi for editing. This study was supported by the Changzhou International Science and Techno-

logy Collaboration Program (#CZ20110021). The funders have no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Address correspondence to:** Dr. Xiaowei Yu, Changzhou Second People's Hospital, Nanjing Medical University, Changzhou, 213003, China. E-mail: yxwlyy8180102@sina.com; Dr. Shen Hu, School of Dentistry and Jonsson Comprehensive Cancer Center, University of California, Los Angeles. E-mail: shenhu@ucla.edu

## References

[1] Verdecchia A, Francisci S, Brenner H, Gatta G, Micheli A, Mangone L, Kunkler I and Group E-W. Recent cancer survival in Europe: a 2000-02 period analysis of EUROCARE-4 data. Lancet Oncol 2007; 8: 784-796.
[2] Jemal A, Bray F, Center MM, Ferlay J, Ward E and Forman D. Global cancer statistics. CA Cancer J Clin 2011; 61: 69-90.
[3] Siegel R, Ward E, Brawley O and Jemal A. Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. CA Cancer J Clin 2011; 61: 212-236.
[4] Hecht SS. Lung carcinogenesis by tobacco smoke. Int J Cancer 2012; 131: 2724-2732.
[5] Pine SR, Mechanic LE, Enewold L, Chaturvedi AK, Katki HA, Zheng YL, Bowman ED, Engels EA, Caporaso NE and Harris CC. Increased levels of circulating interleukin 6, interleukin 8, C-reactive protein, and risk of lung cancer. J Natl Cancer Inst 2011; 103: 1112-1122.
[6] Samet JM, Avila-Tang E, Boffetta P, Hannan LM, Olivo-Marston S, Thun MJ and Rudin CM. Lung cancer in never smokers: clinical epidemiology and environmental risk factors. Clin Cancer Res 2009; 15: 5626-5645.
[7] Yang P. Lung cancer in never smokers. Semin Respir Crit Care Med 2011; 32: 10-21.
[8] Gorugantula LM, Rees T, Plemons J, Chen HS and Cheng YS. Salivary basic fibroblast growth factor in patients with oral squamous cell carcinoma or oral lichen planus. Oral Surg Oral Med Oral Pathol Oral Radiol 2012; 114: 215-222.
[9] Elashoff D, Zhou H, Reiss J, Wang J, Xiao H, Henson B, Hu S, Arellano M, Sinha U, Le A, Messadi D, Wang M, Nabili V, Lingen M, Morris D, Randolph T, Feng Z, Akin D, Kastratovic DA, Chia D, Abemayor E and Wong DT. Prevalidation of salivary biomarkers for oral cancer detection. Cancer Epidemiol Biomarkers Prev 2012; 21: 664-672.
[10] Turner DO, Williams-Cocks SJ, Bullen R, Catmull J, Falk J, Martin D, Mauer J, Barber AE, Wang RC, Gerstenberger SL and Kingsley K. High-risk human papillomavirus (HPV) screening and detection in healthy patient saliva samples: a pilot study. BMC Oral Health 2011; 11: 28.
[11] Zhang L, Xiao H, Zhou H, Santiago S, Lee JM, Garon EB, Yang J, Brinkmann O, Yan X, Akin D, Chia D, Elashoff D, Park NH and Wong DT. Development of transcriptomic biomarker signature in human saliva to detect lung cancer. Cell Mol Life Sci 2012; 69: 3341-3350.
[12] Farrell JJ, Zhang L, Zhou H, Chia D, Elashoff D, Akin D, Paster BJ, Joshipura K and Wong DT. Variations of oral microbiota are associated with pancreatic diseases including pancreatic cancer. Gut 2012; 61: 582-588.
[13] Byakodi R, Krishnappa R, Keluskar V, Bagewadi A and Shetti A. The microbial flora associated with oral carcinomas. Quintessence Int 2011; 42: e118-123.
[14] Laurila AL, Anttila T, Laara E, Bloigu A, Virtamo J, Albanes D, Leinonen M and Saikku P. Serological evidence of an association between Chlamydia pneumoniae infection and lung cancer. Int J Cancer 1997; 74: 31-34.
[15] Littman AJ, White E, Jackson LA, Thornquist MD, Gaydos CA, Goodman GE and Vaughan TL. Chlamydia pneumoniae infection and risk of lung cancer. Cancer Epidemiol Biomarkers Prev 2004; 13: 1624-1630.
[16] Moghaddam SJ, Ochoa CE, Sethi S and Dickey BF. Nontypeable Haemophilus influenzae in chronic obstructive pulmonary disease and lung cancer. Int J Chron Obstruct Pulmon Dis 2011; 6: 113-123.
[17] Paster BJ, Boches SK, Galvin JL, Ericson RE, Lau CN, Levanos VA, Sahasrabudhe A and Dewhirst FE. Bacterial diversity in human subgingival plaque. J Bacteriol 2001; 183: 3770-3783.
[18] Quince C, Lanzen A, Curtis TP, Davenport RJ, Hall N, Head IM, Read LF and Sloan WT. Accurate determination of microbial diversity from 454 pyrosequencing data. Nat Methods 2009; 6: 639-641.
[19] Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J and Glockner FO. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 2007; 35: 7188-7196.
[20] Edgar RC, Haas BJ, Clemente JC, Quince C and Knight R. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 2011; 27: 2194-2200.
[21] Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ and Weber CF. Introducing mothur: open-source, platform-

independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 2009; 75: 7537-7541.
[22] Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J and Knight R. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010; 7: 335-336.
[23] Muhling M, Woolven-Allen J, Murrell JC and Joint I. Improved group-specific PCR primers for denaturing gradient gel electrophoresis analysis of the genetic diversity of complex microbial communities. ISME J 2008; 2: 379-392.
[24] Koyi H, Branden E, Gnarpe J, Gnarpe H and Steen B. An association between chronic infection with Chlamydia pneumoniae and lung cancer. A prospective 2-year study. APMIS 2001; 109: 572-580.
[25] Kocazeybek B. Chronic Chlamydophila pneumoniae infection in lung cancer, a risk factor: a case-control study. J Med Microbiol 2003; 52: 721-726.
[26] Lara-Tejero M and Galan JE. A bacterial toxin that controls cell cycle progression as a deoxyribonuclease I-like protein. Science 2000; 290: 354-357.
[27] Nougayrede JP, Taieb F, De Rycke J and Oswald E. Cyclomodulins: bacterial effectors that modulate the eukaryotic cell cycle. Trends Microbiol 2005; 13: 103-110.
[28] Fava F and Danese S. Intestinal microbiota in inflammatory bowel disease: friend of foe? World J Gastroenterol 2011; 17: 557-566.
[29] Bagyi K, Haczku A, Marton I, Szabo J, Gaspar A, Andrasi M, Varga I, Toth J and Klekner A. Role of pathogenic oral flora in postoperative pneumonia following brain surgery. BMC Infectious Diseases 2009; 9: 104.
[30] Boucher M, Bedotto M, Couderc C, Gomez C, Reynaud-Gaubert M and Drancourt M. Haemophilus pittmaniae respiratory infection in a patient with siderosis: a case report. J Med Case Rep 2012; 6: 120.
[31] Foweraker JE, Cooke NJ and Hawkey PM. Ecology of Haemophilus influenzae and Haemophilus parainfluenzae in sputum and saliva and effects of antibiotics on their distribution in patients with lower respiratory tract infections. Antimicrob Agents Chemother 1993; 37: 804-809.
[32] Rivas Caldas R, Le Gall F, Revert K, Rault G, Virmaux M, Gouriou S, Héry-Arnaud G, Barbier G and Boisramé S. Pseudomonas aeruginosa and periodontal pathogens in the oral cavity and lungs of cystic fibrosis patients: a case control study. J Clin Microbiol 2015; 53: 1898-907.
[33] Scannapieco FA. Role of Oral Bacteria in Respiratory Infection. J Periodontol 1999; 70: 793-802.
[34] Rybojad P, Los R, Sawicki M, Tabarkiewicz J and Malm A. Anaerobic bacteria colonizing the lower airways in lung cancer patients. Folia Histochem Cytobiol 2011; 49: 263-266.
[35] Thirumala R, Rappo U, Babady NE, Kamboj M and Chawla M. Capnocytophaga lung abscess in a patient with metastatic neuroendocrine tumor. J Clin Microbiol 2012; 50: 204-207.
[36] Forman MA, Johnson LR, Jang S and Foley JE. Lower respiratory tract infection due to Capnocytophaga cynodegmi in a cat with pulmonary carcinoma. J Feline Med Surg 2005; 7: 227-231.