Sputum Microbiota in Coal Workers Diagnosed With Pneumoconiosis as Revealed by 16S Metagenomic Sequencing

Vladimir Druzhinin  
Kemerovo State University, Kemerovo, Russian Federation

Liudmila Matskova (✉ liudmila.matskova@ki.se)  
Kemerovo State University, Kemerovo, Russian Federation  
https://orcid.org/0000-0002-3174-1560

Pavel Demenkov  
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russian Federation

Elizaveta Baranova  
Kemerovo State University, Kemerovo, Russian Federation

Valentin Volobaev  
Kemerovo State University, Kemerovo, Russian Federation

Varvara Minina  
Kemerovo State University, Kemerovo, Russian Federation

Alexey Larionov  
Kemerovo State University, Kemerovo, Russian Federation

Snezhana Paradnikova  
Kemerovo State University, Kemerovo, Russian Federation

Research

Keywords: coal worker’s pneumoconiosis, sputum microbiome, lung fibrosis, next generation sequencing, 16S rRNA genes

DOI: https://doi.org/10.21203/rs.3.rs-56277/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Objectives: The microbiome of sputum from former and active coal miners diagnosed with coal worker’s pneumoconiosis (CWP) as compared to healthy controls

Methods: Next Generation Sequencing of bacterial 16S rRNA genes obtained from the sputum of CWP subjects.

Results: Differences were detected between the sputum microbiomes from the healthy and CWP subjects. We noted a significant decrease in Bacteroidetes and an increase in the level of Proteobacteria.

Conclusions: The microbiomes found in sputum from CWP subjects are enriched in bacterial species previously reported to induce pro-inflammatory responses. The profile of the microbiomes correlated mainly to the occupational activity and not to the age of the coal miners.

Introduction

Coal worker’s pneumoconiosis (CWP) is an occupationally induced progressive fibrotic lung disease, caused by the deposition of coal mine dust in the lung parenchyma and by the reaction of tissues to its presence. This public health problem typically occurs in the coal mining industry, including opencast mining, around the world [1–2]. In many countries coal is still used as an important source of energy, and coal mining remains a major industry. This irreversible but preventable disease currently affects millions across the world [3]. Numerous studies have shown that inhalation of coal dust containing crystalline silica (silicon dioxide), usually in the form of quartz or silica, is the primary cause of silicosis, leading to progressive pulmonary fibrosis, which is the main clinical and pathological feature of CWP [4]. Recent studies have shown that the basic level of chromosome damage is increased in blood lymphocytes of CWP patients [5–6]. CWP is also associated with an increased risk of malignant neoplasms [7].

The main mechanisms of CWP include silica-induced macrophage cytotoxicity, activation of leukocytes to produce active oxygen radicals, and damage to alveolar epithelial cells stimulating fibroblast proliferation. Deregulation of DNA methylation is also pointed out as a possible mechanism of CWP pathogenesis [8]. However, the exact mechanisms of progressive pulmonary fibrosis in CWP remains to be elucidated. In particular, the possible effects of the respiratory tract microbiota on the etiology and pathogenesis of CWP needs further investigation.

Numerous recent studies using metagenomic sequencing have shown that the respiratory microbiota plays an important role in maintaining lung health and can differ significantly in various diseases associated with the lungs [9–10]. Changes in the taxonomic composition of respiratory microbiota were evaluated in patients with various pulmonary disorders: COPD [11], asthma [12], community-acquired pneumonia [13], cystic fibrosis [14], lung cancer [15], idiopathic pulmonary fibrosis [16].
By analogy with the above diseases, it can be assumed that professional exposure to coal dust changes the composition of the respiratory microbiota. These changes may, in turn, be associated with progressive pulmonary fibrosis, which are regarded as the main clinical and pathological feature of CWP. Recently, several reports have highlighted the role of the microbiota in fibrosis affecting several human organs, i.e.: intestine, cardiac tissue, liver, skin and breast tissue. Thus, data on the changes of the microbiome composition in the respiratory tract during CWP pathogenesis may be of importance to clarify the role of microbiota in lung fibrosis.

To test this hypothesis, we first performed an analysis of the taxonomic composition of the sputum microbiome of coal miners suffering from CWP and from healthy subjects using 16S ribosomal RNA sequencing. Our results showed that the taxonomic profile of the respiratory microbiome in patients with CWP is different from that in healthy subjects. This may be useful for the early diagnosis of CWP and for the development of a method for suppressing pulmonary fibrosis caused by prolonged exposure to coal dust.

**Methods**

**Cohort information**

The composition of the bacterial microbiome in sputum samples was studied in 21 patients with CWP diagnosis (men only, average age 59.13 ± 8.27 years) who were admitted to the Department of Occupational Disease Pathology, Kemerovo Regional Clinical Hospital (Kemerovo, Russian Federation). The diagnosis of CWP (code J60 according to ICD-10) was made on the basis of chest x-ray and spirometry. All patients worked as underground coal miners. Of these, 5 (23.8%) were actively working in coal mines at the time of the survey, 16 participants (76.2%) had ceased working due to the onset of the disease. Mining work experience in CWP patients varied from 18 to 37 years (average value 26.9 ± 5.3 years). As a control group, we examined 21 healthy men – donors at a blood transfusion station, who were residents of Kemerovo (average age 53.3 ± 5.36 years). Among CWP patients there was one active smoker, among the controls – 65.2%. The summarised information on CWP patients and controls is shown in Table 1. An individual questionnaire was filled out for each survey participant, containing information about the place and date of birth, profession, exposure to occupational hazards, health status, diet features, medications, X-ray records and harmful habits (smoking and alcohol use).

**Ethics statement**

All procedures followed the ethical standards of the Helsinki Declaration (1964, amended 2008) of the World Medical Association. All participants (CWP patients and controls) were informed about the aim, methodology and possible risks of the study; informed consent was signed by each donor. The design of this study was approved by the Ethics Committee of the Kemerovo State University.

**Sample Collection, Processing and Storage**
To analyze the composition of the microbiome of the respiratory tract, sputum samples from CWP patients and controls were obtained prior to all diagnostic or therapeutic procedures. Sputum samples were collected non-invasively through participant-induced coughing (i.e., without induction) and represented the oropharyngeal secretion. The resulting samples were immediately placed in sterile plastic vials and frozen (-20 °C). Frozen samples were transported to the laboratory and stored at -80 °C.

DNA extraction, 16S rRNA amplification and 16S rRNA sequencing

Sample DNA was extracted using FastDNA Spin Kit For Soil (MP Biomedicals) based on the manufacturer's recommendation. Forty two 16S rRNA gene amplicon libraries were prepared by PCR amplification of the 467 bp fragment within the hypervariable (V3-V4) region of the bacterial 16S rRNA genes from 50 ng of each of the extracted and purified sputum DNAs. The initial PCR was performed with broad-spectrum 16S rRNA primers.

Forward primer: 5′-TCGTCGCGACGGTATGATGTAAGACAGCCTACGGGNGGCWGCAG-3′

Reverse primer: 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC-3′

using BioMaster Hi-Fi LR 2 × ReadyMix DNA polymerase (BiolabMix company, Novosibirsk, Russia). Cycle conditions were 94 °C (3 min 30 s), followed by 25 cycles of 94 °C (30 s), 55 °C (30 s), 68 °C (40 s), and a final extension at 68 °C (5 min). Libraries were purified using Agencourt AMPure XP beads (Beckman Coulter, Bray, USA). Dual indices and Illumina sequencing adapters from the Illumina Nextera XT index kits v2 B and C (Illumina, San Diego, USA) were added to the target amplicons in a second PCR step using BioMaster Hi-Fi LR 2 × ReadyMix DNA polymerase (BiolabMix company, Novosibirsk, Russia). Cycle conditions were 94 °C (3 min 30 s), then 8 cycles of 94 °C (30 s), 55 °C (30 s), 68 °C (40 s), and a final extension of 68 °C (5 min). Libraries were again purified using XP beads. Preparation of 16S rRNA libraries were done according to the Illumina 16S metagenomic sequencing library protocol. Sample PCR products were pooled in equimolar ratio, purified using XP Beads, and quantified using a fluorometer (Quantus Fluorometer dsDNA (Promega, Madison, WI, USA). Molarity of the libraries was brought to 4 nM, the libraries were denatured, and diluted to a final concentration of 8 pM with a 10% PhiX spike buffer for sequencing on the Illumina MiSeq [17].

Taxonomy quantification using 16S rRNA gene sequences and statistical methods

The resulting data was processed using the program QIIME2 [18]. A quality check was carried out and a sequence library was generated.

The sequences were combined into operational taxonomic units (OTUs) based on a 99% nucleotide similarity threshold using the Greengenes reference sequence library (versions 13 – 8) and SILVA (version 132), followed by removal of singletons (OTUs containing only one sequence).
The total diversity of prokaryotic sputum communities (alpha diversity) as estimated by the number of allocated OTUs (analogue of species richness) and Shannon indices ($H = \Sigma p_i \ln p_i$, $p_i$ – part of $i$-th species in community) according to the UniFrac method [19].

When calculating sample diversity indices, 351 sequences were normalized (the minimum number of received sequences per sample). The variation in the structure of the bacterial community of different samples (beta diversity) was also analyzed using UniFrac [19] – a method common in microbial ecology that estimates the difference between communities based on the phylogenetic relationships of the presented taxa.

We used a version of the unweighted UniFrac method that takes into account only the presence of taxa, but not their share in the community. The significance of differences between groups of samples was evaluated by the PERMANOVA method (Adonis).

In addition, to assess the significance of differences in the relative percentage of individual bacterial taxa in the sputum samples the Mann-Whitney U test was used. To estimate the difference in the frequencies of occurrence, the Fisher exact test was used. Calculations were performed using the software package STATISTICA.10, Statsoft, USA.

**Results**

In our sequencing approach (16S rRNA V3–V4) for CWP and controls using sputum samples, we were able to identify a total of 8 phyla with relative frequencies above 0.1%. The prevailing phyla in our dataset were Firmicutes, Bacteroidetes, Actinobacteria and Proteobacteria (Fig. 1), as expected from previous studies [20–21]. A comparison of the frequencies of the main bacteria phyla in sputum revealed a significant decrease in the representatives of Bacteroidetes in the CWP compared to the controls ($22.13 \pm 9.8$ vs $28.7 \pm 8.42\%$, respectively; $p = 0.022$). For the remaining 7 phyla of bacteria, there were no differences between patients and controls (Fig. 2).

Regarding alpha diversity, neither the number of allocated OTUs nor the Shannon indices showed any significant differences between CWP and controls. Overall, bacterial communities in the two groups in the study were fairly diverse as indicated by the Shannon index at the genus level (5.842 in CWP vs 5.791 in controls). This suggests that the changes in the sputum microbiome as result of progressive pulmonary fibrosis were not large-scale community shifts.

Differences in the structure of bacterial communities in sputum samples of CWP and controls are shown in Fig. 3. The PERMANOVA (Adonis) test using the difference matrix constructed by the unweighted UniFrac method showed a significant difference in the prokaryotic communities in the sputum of healthy subjects as compared to miners with pneumoconiosis (pseudo-$F = 4.42$; $p = 0.001$). The taxonomic structure of the microbiomes at the genera level are shown in Fig. 4. Sequencing statistics for 44 genera (with no less than 0.1% relative frequency) are summarised in Table 2, alongside with the corresponding U-rank Mann-Whitney $p$ values. Among genera, *Streptococcus*, *Prevotella* (f. *Prevotellaceae*), *Veillonella* and *Anaerosinus* were the most common in the two pools.
In the sputums of CWP subjects, compared to controls, there was a significant increase in abundance (by percentage) of the following genera: *Streptococcus* (*21.32 ± 8.9 vs 15.42 ± 9.2; p = 0.006); *Gemella* (*3.12 ± 2.19 vs 2.01 ± 2.37; p = 0.04); *Bacillus* (*3.11 ± 2.19 vs 1.99 ± 2.39; p = 0.036) and *Pasteurellaceae* (*3.27 ± 2.35 vs 1.91 ± 1.43; p = 0.041). At the same time, the genus *Prevotella* (*f. Prevotellaceae*) was significantly less represented in the microbiomes of CWP subjects compared to the controls (*14.77 ± 5.93 vs 19.36 ± 5.66; p = 0.029). Three more bacterial genera were also significantly less represented in the microbiome of CWP subjects compared to the controls: *Selenomonas* (*2.94 ± 2.21 vs 6.17 ± 3.69; p = 0.004); *Megasphaera* (*2.22 ± 1.62 vs 3.74 ± 2.26; p = 0.021) and *Dialister* (*0.13 ± 0.36 vs 0.43 ± 0.65; p = 0.038).

Sequencing statistics for 31 species (which met with a relative frequency of not less than 0.1%) are summarised in Table 3, alongside with the corresponding U-rank Mann-Whitney p values. Differences between patients and controls were also found in the contents of the three most common bacterial species: *Streptococcus agalactiae*, *Selenomonas bovis* and *Megasphaera micronuciformis*. The average percentage of *Streptococcus agalactiae* in CWP sputum samples was significantly higher than in controls (*22.35 ± 10.07 vs 16.55 ± 9.42; p = 0.011). In contrast, two other microbes were more abundant in the sputum of healthy subjects compared to miners with pneumoconiosis. Specifically, for *Selenomonas bovis* − *5.96 ± 5.28 in controls and 2.6 ± 2.1 in CWP subjects (p = 0.028) and for *Megasphaera micronuciformis* − *3.77 ± 2.3 in controls and 2.19 ± 1.67 in CWP subjects (p = 0.017).

Given the fact that age can affect the composition of the microbiome, we used the Spearman coefficient to assess the influence of age on the representation of those bacterial taxa for which there were differences between CWP subjects and controls. In the total sample collection (CWP subjects and controls), a slight increase in the content of the genus *Streptococcus* was observed with age, however, the significance of this increase was found only at the confidence limit (*r = 0.2916; p = 0.05). The correlation between age and the content of other genera in the sputum samples also turned out to be insignificant. *Prevotella* (*f. Prevotellaceae*) (*r =-0.1935; p = 0.2) *Selenomonas* (*r =-0.265; p = 0.08) *Megasphaera* (*r =-0.279; p = 0.06).

We could not evaluate smoking status as a factor affecting the composition of the bacterial flora in CWP subjects as there was only one smoker in this group, but we studied it separately in the controls (Fig. 5). A comparison of genera in the sputums of smokers and non-smokers, all in the control group, revealed a significant decrease in the presence of the following genera in smokers: *Neisseria* (*0.63 ± 1.39 vs 2.77 ± 2.13; p = 0.003); *Bulleidea* (*0.2 ± 0.3 vs 0.58 ± 0.44; p = 0.03) and *Peptostreptococcus* (*0.08 ± 0.3 vs 0.82 ± 1.01; p = 0.03).

There were large differences in disease duration among miners suffering from pneumoconiosis, from 2 to 24 years from the date of diagnosis. To assess the possible correlations of disease duration to the content of different bacteria in sputum, the Spearman coefficient was used. No significant correlations were found between the duration of the disease and the content of any bacterial genus or species. In addition to the duration of pneumoconiosis, patients were further divided into two subgroups: working miners (*n = 5*) and former miners (*n = 16*), who stopped working due to the onset of the disease. A
comparison of the composition of the microbiomes (Table 4) revealed an increase of: *Lachnoanaerobaculum orale* (0.99 ± 0.12 vs 0.6 ± 0.7; p = 0.043); *Prevotella tannarae* (2.05 ± 1.94 vs 0.7 ± 1.28; p = 0.026) and *Uncultured eubacterium E1-K12* (2.0 ± 1.98 vs 0.51 ± 1.25; p = 0.005) in the sputums of still active miners as compared to the former miners.

**Discussion**

Differences in bacterial populations in the healthy and diseased human airways have already been recognized as a possible contributing factor in the pathogenesis of diseases of the respiratory tract, however, the bacterial population in the airways of subjects diagnosed with coal worker’s pneumoconiosis (CWP) has not yet been investigated.

The lung microbiome during health overlaps mainly with the microbiome of the oral cavity, not with those from nose and gastric tract, which provides evidence for microaspiration as a common way of formation of the lung microbiome in healthy individuals, but it is less species-rich and shows signs of specific elimination of some bacterial species from the upper respiratory tract. The lung microbiome shows great interindividual variability [9, 22].

The «healthy» lung microbiome may be perturbed by pulmonary diseases and some environmental factors, for example, by the diverse microbial exposure during development of childhood asthma [23].

In the course of our investigation, we have determined certain regularities in the composition of microbial communities in the sputum of coal miners suffering from occupational lung fibrosis.

The data obtained by us do not allow to determine the primary cause of disease: changes in the microbiota under the influence of environmental factors or the pulmonary pathology. At the same time, we suggest that the risk of developing CWP is connected to the composition of the microbial community in the sputum since it is known that certain bacterial taxa are capable of producing pro- or anti-inflammatory lipopolysaccharides. Pro-inflammatory lipopolysaccharides bind the CD14 / TLR4 / MD2 receptor complex in many cell types, but especially in monocytes, dendritic cells, macrophages and B cells, which contributes to the secretion of pro-inflammatory cytokines, nitric oxide and eicosanoids [24]. Changes in the balance of lipopolysaccharide production in the direction of an inflammatory response is likely to influence the immunologic reaction significantly, leading to a pathologic development.

When evaluating the relative abundance of bacterial types in the sputums of CWP patients, we noted a significant decrease in Bacteroidetes, primarily the genus *Prevotella*. Previously, a decrease in the relative level of Bacteroidetes and especially representatives of the genus *Prevotella* was reported in patients with bronchial asthma [25] and COPD [26]. It is known that bacteria from the *Bacteroidales* order produce anti-inflammatory forms of lipopolysaccharides thereby providing immune escape for the entire microbiota community [27]. In particular, the proportion of anti-inflammatory lipopolysaccharides produced by representatives of Bacteroidetes can reach 79% of all anti-inflammatory lipopolysaccharides present in healthy subjects [28].
An increased level of pro-inflammatory lipopolysaccharides in blood plasma was previously observed in many inflammatory diseases, including COPD [29]. We noticed an increase in the level of *Pasteurellaceae* (Proteobacteria) in patients with CWP and an increase in the levels of the potentially pathogenic genus *Haemophilus* also belonging to this family. Proteobacteria, being the main producers of pro-inflammatory lipopolysaccharides, have been repeatedly associated with various inflammatory and allergic diseases [30]. *Pasteurellaceae* spp. are commensals of the mucous membranes but can act as opportunistic pathogens following a decrease in immunity due to various factors. They have the ability to synthesize pore-forming toxins (RTX), pro-inflammatory lipopolysaccharides and immunogenic lipoproteins [31].

In assessing beta diversity, a significant increase in the genus *Streptococcus* and, in particular, of the *Streptococcus agalactiae* was observed in individuals with CWP diagnosis.

*Streptococcus agalactiae* (also known as GBS) is an important opportunistic bacterium that can cause pneumonia, sepsis and meningitis in newborns and in patients with weakened immunity. GBS bacteria effectively attach to pulmonary epithelial cells and are capable of invasion. This is initiated by attachment to extracellular matrix molecules such as agglutinin, fibronectin, fibrinogen and laminin, which facilitates their attachment to host cell surface proteins, such as integrins. Thus, the invasive potential of GBS is influenced by changes in the surface proteome of the host cells, which can be caused by various lung pathologies [32]. The molecular mechanisms of human cytopathology caused by GBS bacteria is under intensive investigation currently [33–34].

In the sputum of men with CWP diagnosis we found, in addition to *Streptococcus*, a small, but statistically significant increase in the abundance of representatives of the genera *Gemella* and *Bacillus* as compared to the control group. *Gemella*, a genus of Gram-variable motionless asporogenic bacteria, is a commensal of the oral cavity in the healthy population but it can act as a causative agent of lung abscess [35]. *Gemella* can also cause endocarditis [36], can play a role in exacerbating pneumonia and act as a biomarker in patients with cystic fibrosis [37].

In the microbiome of miners with CWP diagnosis, we noted a significant decrease in the abundance of representatives of the genera *Selenomonas*, *Megasphaera*, and *Dialister* compared to the controls. Interestingly, while the bulk of emerging data about these bacterial species concern their association with human pathologies, one study demonstrates that *Dialister* is locally reduced in tumor biopsies from lung cancer patients. The *Dialister* population was larger in material taken from an unaffected area from the same patients. Moreover, the level of *Dialister* was higher in the control group than in any of the samples from cancer patients [38].

We have evaluated the impact of factors like smoking and age, on the differential representation of the bacterial taxa in the sputum of coal miners with CWP diagnosis. Our correlation analysis showed no significant age-dependent differences in the representation of the *Streptococcus*, *Prevotella* (f. *Prevotellaceae*), *Selenomonas* and *Megasphaera* taxa. Thus, we infer that the variation in sputum microbiomes that we observe is not age related, but that exposure to occupational hazards, such as coal dust may be the causative factor for the variation in sputum microbiomes. The impact of smoking was
evaluated only in the control group. We confirmed a decrease of *Neisseria* in the sputum of smokers as previously reported [39]. Additionally, we revealed a significant decrease of *Bulleidea* and *Peptostreptococcus* species in the sputums of the smokers.

In sputum samples from active coal miners we detect significantly more of three bacterial taxa: *Lachnoanaerobaculum orale*, *Prevotella Tannarae* and *Uncultured eubacterium E1-K12*, as compared to former miners, all with a CWP diagnosis. This may provide evidence for specific damages to the respiratory tract of the coal miners not coupled to age, smoking or exposure to hazardous occupational factors. Recently, an increased level of *Lachnospiraceae* and *Lachnoclostridium* was reported in patients with silicosis, albeit in the gut microbiome [40].

We describe the differential representation of certain bacterial species in the sputum of coal miners with CWP diagnosis. The causative contribution of this to CWP pathogenesis requires further investigation.

**Conclusion**

This pilot study for the first time presents the results of an analysis of sputum microbiomes in a small group of current and former coal miners suffering from CWP, living in Kuzbass, a coal-mining region in Russia.

The method of massive parallel sequencing of 16S rRNA genes has been used for the first time to obtain a taxonomic characteristic of the sputum samples from the coal miners with CWP diagnosis.

The differential representation of bacterial taxa revealed in this study of sputum samples from coal miners with CWP diagnosis will be further confirmed in a larger group of samples. In addition, GOC and KEGG analysis will be employed to evaluate the functional relevance of these variations. The sputum microbiome may serve as an important source for non-invasive biomarkers of CWP.

**Declarations**

**Ethics approval and consent to participate.** All procedures followed the ethical standards of the Helsinki Declaration (1964, amended 2008) of the World Medical Association. All participants (CWP patients and controls) were informed about the aim, methodology and possible risks of the study; informed consent was signed by each donor. The design of this study was approved by the Ethics Committee of the Kemerovo State University.

**Consent for publication.** Consent for publication was obtained from each person in the study. All authors approved the final version of the manuscript.

**Availability of supporting data.** The datasets obtained and analysed during the current study available from the corresponding author on reasonable request.

**Competing interests.** Authors declare no conflict of interest.
Funding. The authors declare that they have no competing interests.

Authors' contributions. D.V.G. and M.L.V. conceived the study; D.V.G. and M.L.V. wrote the manuscript; B.E.D., V.V.P., M.V.I. performed laboratorial work; D.P.S. and P.S.A. carried out bioinformatics and statistical analyses; D.V.G. provided samples and managed clinical data. L.A.V. critically revised the manuscript.

Acknowledgements. We would like to thank all patients and controls for donating their samples and for collaborating in this study. This work was supported by Russian Science Foundation Grant No. 18-14-00022 and by the Russian Academic Excellence Project at the Immanuel Kant Baltic Federal University (5-100).

References

1. Han S, Chen H, Harvey MA, Stemn E, Cliff D. Focusing on Coal Workers’ Lung Diseases: A Comparative Analysis of China, Australia, and the United States. Int J Environ Res Public Health. 2018;15(11):E2565.

2. Blackley DJ, Halldin CN, Laney AS. Continued Increase in Prevalence of Coal Workers’ Pneumoconiosis in the United States, 1970–2017. Am J Public Health. 2018;108(9):1220–2.

3. Leonard R, Zulfikar R, Stansbury R. Coal mining and lung disease in the 21st century. Curr Opin Pulm Med. 2020;26(2):135–41.

4. Beer C, Kolstad HA, Sondergaard K, et al. A systematic review of occupational exposure to coal dust and the risk of interstitial lung diseases. Eur Clin Respir J. 2017;4(1):1264711.

5. Volobaev VP, Sinitsky MY, Larionov AV, et al. Modifying influence of occupational inflammatory diseases on the level of chromosome aberrations in coal miners. Mutagenesis. 2016;31(2):225–9.

6. Druzhinin VG, Apalko SV, Baranova ED, et al. Micronuclei in blood lumphocytes of existing and former coal miners: evaluation of the effect of anthracosilicosis. Ecological genetics. 2019;17(4):55–62. [Article in Russian].

7. Tomaskova H, Jirak Z, Splichalova A, Urban P. Cancer incidence in Czech black coal miners in association with coal workers’ pneumoconiosis. Int J Occup Med Environ Health. 2012;25(2):137–44.

8. Zhang N, Liu K, Wang K, et al. Dust induces lung fibrosis through dysregulated DNA methylation. Environ Toxicol. 2019;34(6):728–41.

9. Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. The Microbiome and the Respiratory Tract. Annu Rev Physiol. 2016;78:481–504.

10. Caverly LJ, Huang YJ, Sze MA. Past, Present, and Future Research on the Lung Microbiome in Inflammatory Airway Disease. Chest. 2019;156(2):376–82.

11. Qi YJ, Sun XJ, Wang Z, et al. Richness of sputum microbiome in acute exacerbations of eosinophilic chronic obstructive pulmonary disease. Chin Med J (Engl). 2020;133(5):542–51.
12. Kozik AJ, Huang YJ. The microbiome in asthma: Role in pathogenesis, phenotype, and response to treatment. Ann Allergy Asthma Immunol. 2019;122(3):270–5.
13. Wootton DG, Cox MJ, Gloor GB, et al. A Haemophilus sp. dominates the microbiota of sputum from UK adults with non-severe community acquired pneumonia and chronic lung disease. Sci Rep. 2019;9(1):2388.
14. Acosta N, Heirali A, Somayaji R, et al. Sputum microbiota is predictive of long-term clinical outcomes in young adults with cystic fibrosis. Thorax. 2018;73(11):1016–25.
15. Maddi A, Sabharwal A, Violante T, et al. The microbiome and lung cancer. J Thorac Dis. 2019;11(1):280–91.
16. Hewitt RJ, Molyneaux PL. The respiratory microbiome in idiopathic pulmonary fibrosis. Ann Transl Med. 2017;5(12):250.
17. Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7:335–6.
18. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37:852–7.
19. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol. 2005;71:8228–35.
20. Dickson RP, Huffnagle GB. The Lung Microbiome: New Principles for Respiratory Bacteriology in Health and Disease. PLoS Pathog. 2015;11(7):e1004923.
21. Cameron SJS, Lewis KE, Huws SA, et al. A pilot study using metagenomic sequencing of the sputum microbiome suggests potential bacterial biomarkers for lung cancer. PLoS One. 2017;25:e0177062.
22. Bassis CM, Erb-Downward JR, Dickson RP, et al. Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. mBio. 2015;6(2):e00037.
23. Cui L, Morris A, Huang L, et al. The microbiome and the lung. Ann Am Thorac Soc. 2014;11(Suppl 4(Suppl 4):227–32.
24. Skirecki T, Cavaillon JM. Inner sensors of endotoxin - implications for sepsis research and therapy. FEMS Microbiol Rev. 2019;43(3):239–56.
25. Fedosenko SV, Ogorodova LM, Karnaushkina MA, Kulikov ES, Deev IA, Kirillova NA. The Airways Microbial Community Composition in Healthy Individuals and Bronchial Asthma Patients. Vestn Ross Akad Med Nauk. 2014; (3–4): 71 – 6. [Article in Russian].
26. Sze MA, Dimitriu PA, Suzuki M, et al. The host response to the lung microbiome in chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2015;192(4):438–45.
27. Lin TL, Shu CC, Chen YM, et al. Like Cures Like: Pharmacological Activity of Anti-Inflammatory Lipopolysaccharides From Gut Microbiome. Front Pharmacol. 2020;11:554.
28. Alexandre A, Maria KD, Alexandre PD, et al. Knowledge for Health and Consumer Safety. The Human Gut Microbiota: Overview and analysis of the current scientific knowledge and possible impact on
healthcare and well-being. 361. Luxembourg: Publications Office of the European Union; 2018. p. k2179.

29. Gupta V, Banyard A, Mullan A, Sriskantharajah S, Southworth T, Singh D. Characterization of the inflammatory response to inhaled lipopolysaccharide in mild to moderate chronic obstructive pulmonary disease. Br J Clin Pharmacol. 2015;79(5):767–76.

30. Salguero MV, Al-Obaide MAI, Singh R, Siepmann T, Vasylyeva TL. Dysbiosis of Gram-negative gut microbiota and the associated serum lipopolysaccharide exacerbates inflammation in type 2 diabetic patients with chronic kidney disease. Exp Ther Med. 2019;18(5):3461–9.

31. Dousse F, Thomann A, Brodard I, et al. Routine phenotypic identification of bacterial species of the family Pasteurellaceae isolated from animals [published correction appears in J Vet Diagn Invest. 2009; Mar; 21(2): 288. J Vet Diagn Invest. 2008; 20(6): 716–724.

32. Sharma P, Lata H, Arya DK, et al. Role of Pilus Proteins in Adherence and Invasion of Streptococcus agalactiae to the Lung and Cervical Epithelial Cells. J Biol Chem. 2012;288(6):4023–34.

33. Costa AF, Moraes JA, Oliveira JS, et al. Reactive oxygen species involved in apoptosis induction of human respiratory epithelial (A549) cells by Streptococcus agalactiae. Microbiology. 2016;162:94–9.

34. Wu X, Deng G, Li M, et al. Wnt/β-catenin signaling reduces Bacillus Calmette-Guerin-induced macrophage necrosis through a ROS -mediated PARP/AIF-dependent pathway. BMC Immunol. 2015; 16 (1).

35. Takayanagi N, Kagiyama N, Ishiguro T, Tokunaga D, Sugita Y. Etiology and Outcome of Community-Acquired Lung Abscess. Respiration. 2010;80(2):98–105.

36. Shinha T. Endocarditis Due to Gemella Morbillorum. Intern Med. 2017;56(13):1751.

37. Carmody LA, Zhao J, Schloss PD, et al. Changes in cystic fibrosis airway microbiota at pulmonary exacerbation. Ann Am Thorac Soc. 2010;10:179–87.

38. Liu HX, Qu JM. Characterization of lower airway microbiome and its correlation with lung cancer. Alterations in microbiome and virome in lung disease. Am J Respir Crit Care Med. 2017;195:A2940.

39. Morris A, Beck JM, Schloss PD. et al. Comparison of the respiratory microbiome in healthy nonsmokers and smokers. Am J Respir Crit Care Med. 2013;187(10):1067-75.

40. Zhou Y, Chen L, Sun G, Li Y, Huang R. Alterations in the gut microbiota of patients with silica-induced pulmonary fibrosis. J Occup Med Toxicol. 2019; 14.

Tables

Table 1. Characteristics of the study cohorts
| Cohort                        | N   | Age (years) | Smoking status, % |
|------------------------------|-----|-------------|-------------------|
|                              |     | Mean ± SD   | Min-max           | Smokers | Non-smokers |
| Coal worker's pneumoconiosis:|     |             |                   |         |             |
| Current miners               |     |             |                   |         |             |
| Former miners                | 5   | 53.2 ± 8.11 | 42-61             | 0       | 100         |
| Total                        | 16  | 60.78 ± 7.74| 48-77             | 6.3     | 93.7        |
| Controls                     | 21  | 59.3 ± 8.62 | 42-77             | 5.3     | 94.7        |
|                              | 21  | 53.3 ± 5.36 | 46-63             | 65.2    | 34.8        |

Table 2. Composition of the “core” microbiome at the genera level in sputum from subjects in the control and CWP groups (by average percentage).
| Genus                                      | Controls | Controls | CWP       | CWP       | P Value |
|--------------------------------------------|----------|----------|-----------|-----------|---------|
|                                            | Mean ± SD| Count    | mean ± SD | Count    |         |
| **Streptococcus**                         | 15.42±9.2| 0/21     | 21.32±8.9 | 0/21     | **0.006**|
| **Prevotella (f. Prevotellaceae)**         | 19.36±5.66| 0/21    | 14.77±5.93| 0/21     | **0.029**|
| **Veillonella**                           | 12.71±7.67| 0/21     | 9.74±5.61 | 0/21     | > 0.05 |
| **Anaerosinus**                           | 12.37±6.75| 1/20     | 8.99±5.42 | 0/21     | > 0.05 |
| **Selenomonas**                           | 6.17±3.69 | 1/20     | 2.94±2.21 | 3/18     | **0.004**|
| **Porphyromonas**                         | 4.21±3.29 | 2/19     | 2.38±1.69 | 2/19     | > 0.05 |
| **Actinomyces**                           | 3.75±2.13 | 2/19     | 4.09±2.86 | 1/20     | > 0.05 |
| **Megasphaera**                           | 3.74±2.26 | 3/18     | 2.22±1.62 | 3/18     | **0.021**|
| **Alloprevotella**                        | 3.19±2.36 | 2/19     | 2.68±3.47 | 6/15     | > 0.05 |
| **Streptobacillus**                       | 3.17±2.98 | 4/17     | 2.89±3.22 | 5/16     | > 0.05 |
| **Leptotrichia**                          | 2.63±2.64 | 4/17     | 3.09±3.14 | 4/17     | > 0.05 |
| **Granulicatella**                        | 2.28±1.68 | 3/18     | 1.85±1.34 | 2/19     | > 0.05 |
| **Gemella**                               | 2.01±2.37 | 5/16     | 3.12±2.19 | 1/20     | **0.04** |
| **Rothia**                                | 1.99±1.95 | 6/15     | 3.1±3.09  | 4/17     | > 0.05 |
| **Bacillus**                              | 1.99±2.39 | 6/15     | 3.11±2.19 | 1/20     | **0.036**|
| **Atopobium**                             | 1.98±1.42 | 0/21     | 1.78±1.42 | 2/19     | > 0.05 |
| **Pasteurellaceae**                       | 1.91±1.43 | 2/19     | 3.27±2.35 | 3/18     | **0.041**|
| **Fusobacterium**                         | 1.91±1.82 | 8/13     | 1.65±1.42 | 4/17     | > 0.05 |
| **Macellibacteroides**                    | 1.69±1.62 | 7/14     | 1.02±1.28 | 11/10    | > 0.05 |
| **Neisseria**                             | 1.45±1.97 | 10/11    | 2.45±5.46 | 8/13     | > 0.05 |
| **Bacteroides**                           | 1.43±1.57 | 7/14     | 1.29±2.18 | 10/11    | > 0.05 |
| **Prevotella (f. Paraprevotellaceae)**     | 1.43±1.8  | 8/13     | 1.17±1.58 | 8/13     | > 0.05 |
| **Vestibaculum**                          | 1.17±1.82 | 11/10    | 0.51±0.78 | 13/8     | > 0.05 |
| **Stomatobaculum**                        | 0.94±0.99 | 6/15     | 0.6±1.02  | 9/12     | > 0.05 |
| **Campylobacter**                         | 0.72±0.9  | 6/15     | 0.39±0.43 | 7/14     | > 0.05 |
| **Lachnoanaerobaculum**                   | 0.62±0.54 | 6/15     | 0.67±0.66 | 5/16     | > 0.05 |
| **Clostridium (f. Lachnospiraceae)**       | 0.6±0.72  | 10/11    | 0.59±0.89 | 10/11    | > 0.05 |
| Species                     | Percentage Control | p-value Control  | Percentage CWP | p-value CWP  |
|----------------------------|--------------------|------------------|----------------|--------------|
| *Treponema*                | 0.59±0.93          | 7/14             | 0.79±1.13      | 7/14         | 0.05          |
| *Solobacterium*           | 0.45±0.42          | 5/16             | 0.38±0.43      | 9/12         | 0.05          |
| *Oribacterium*            | 0.44±0.62          | 12/9             | 0.74±0.99      | 8/13         | 0.05          |
| *Dialister*               | 0.43±0.65          | 11/10            | 0.13±0.36      | 16/5         | 0.038         |
| *Bulleidea*               | 0.43±0.43          | 7/14             | 0.37±0.47      | 9/12         | 0.05          |
| *Capnocytophaga*          | 0.41±0.68          | 9/12             | 0.67±1.2       | 12/9         | 0.05          |
| *Peptostreptococcus*      | 0.36±0.74          | 16/5             | 0.57±0.76      | 10/11        | 0.05          |
| *Mycoplasma*              | 0.34±0.95          | 14/7             | 0.34±0.47      | 10/11        | 0.05          |
| *Clostridium (f. Clostridiaceae)* | 0.32±0.87      | 17/4             | 0.17±0.4       | 16/5         | 0.05          |
| *Moriella*                | 0.28±0.59          | 17/4             | 0.27±0.65      | 18/3         | 0.05          |
| *Haemophilus*             | 0.23±0.53          | 17/4             | 2.2±9.81       | 19/2         | 0.05          |
| *Actinobacillus*          | 0.16±0.73          | 20/1             | 0.78±1.87      | 16/5         | 0.05          |
| *Bifidobacterium*         | 0.15±0.52          | 2/19             | 0.25±0.5       | 14/7         | 0.05          |
| *Lactobacillus*           | 0.14±0.33          | 17/4             | 0.17±0.79      | 20/1         | 0.05          |
| *Bordetella*              | 0.13±0.39          | 17/4             | 0.11±0.32      | 18/3         | 0.05          |
| *Filifactor*              | 0.12±0.22          | 15/6             | 0.27±0.43      | 12/9         | 0.05          |
| *Bergeyella*              | 0.1±0.21           | 11/10            | 0.15±0.23      | 4/17         | 0.05          |

Table 3. Composition of the “core” microbiome at the species level in sputum from subjects in the control and CWP groups (by average percentage).
| Species                        | Controls (21) | CWP (21) | P Value |
|-------------------------------|--------------|----------|---------|
|                               | Mean ± SD    | Count    | mean ± SD | Count |         |
| **Streptococcus agalactiae**  | 16.55±9.42   | 0/21     | 22.35±10.07 | 0/21  | 0.011   |
| **Anaerosinus glycerini**     | 13.11±6.99   | 1/20     | 8.93±4.87  | 0/21  | > 0.05  |
| **Selenomonas bovis**         | 5.96±5.28    | 1/20     | 2.6±2.1    | 3/18  | 0.028   |
| **Megasphaera micronuciformis** | 3.77±2.3    | 3/18     | 2.19±1.67  | 5/16  | 0.017   |
| **Prevotella histicola**      | 2.72±2.46    | 7/14     | 2.88±2.51  | 5/16  | > 0.05  |
| **Actinomyces hyovaginalis**  | 2.36±1.62    | 3/18     | 2.63±2.27  | 5/16  | > 0.05  |
| **Granulicatella balaenopterae** | 2.23±1.7  | 3/18     | 1.83±1.32  | 2/19  | > 0.05  |
| **Atopobium rimae**           | 2.09±1.46    | 0/21     | 1.95±1.37  | 2/19  | > 0.05  |
| **Prevotella pallens**        | 2.07±2.12    | 7/14     | 0.84±0.76  | 7/14  | > 0.05  |
| **Rothia terrae**             | 1.93±1.89    | 7/14     | 3.03±3.04  | 4/17  | > 0.05  |
| **Macellibacteroides fermentans** | 1.82±1.69 | 7/14     | 1.03±1.29  | 11/10 | > 0.05  |
| **Bacteroides nordii**        | 1.6±1.5      | 5/16     | 1.57±1.91  | 9/12  | > 0.05  |
| **Prevotella tannarae**       | 0.89±1.32    | 10/11    | 1.02±1.53  | 10/11 | > 0.05  |
| **Lachnoanaerobaculum orale** | 0.69±0.51    | 5/16     | 0.69±0.64  | 3/18  | > 0.05  |
| **Prevotella intermedia**     | 0.69±1.27    | 13/8     | 0.6±0.76   | 10/11 | > 0.05  |
| **Prevotella nigrescens**     | 0.64±1.49    | 14/7     | 0.24±0.43  | 15/6  | > 0.05  |
| **Prevotella nanceiensis**    | 0.52±0.93    | 13/8     | 0.56±0.64  | 9/12  | > 0.05  |
| **Bulleidia moorei**          | 0.44±0.43    | 7/14     | 0.37±0.46  | 9/12  | > 0.05  |
| **Clostridium bolteae**       | 0.33±0.61    | 15/6     | 0.46±0.66  | 12/9  | > 0.05  |
| **Porphyromonas endodontalis**| 1.19±1.67    | 9/12     | 0.58±0.99  | 11/10 | > 0.05  |
| **Vestibaculum illigatum**    | 1.19±1.84    | 10/11    | 0.52±0.77  | 12/9  | > 0.05  |
| **Clostridium acidurici**     | 0.33±0.87    | 17/4     | 0.16±0.4   | 17/4  | > 0.05  |
| **Mycoplasma zalophi**        | 0.3±0.94     | 15/6     | 0.34±0.47  | 11/10 | > 0.05  |
| **Moryella indoligenes**      | 0.28±0.59    | 17/4     | 0.28±0.66  | 16/5  | > 0.05  |
| **Peptostreptococcus anaerobius** | 0.25±0.67 | 17/4     | 0.35±0.59  | 14/7  | > 0.05  |
| **Treponema amulovorum**      | 0.23±0.47    | 11/10    | 0.12±0.28  | 17/4  | > 0.05  |
| **Oribacterium sinus**        | 0.2±0.6      | 18/3     | 0.13±0.3   | 17/4  | > 0.05  |
| Species                    | Current (5)  | Former (16) | P   |
|---------------------------|--------------|-------------|-----|
| *Lachnoanaerobaculum orale* | 0.99±0.12    | 0.6±0.7     | 0.043 |
| *Prevotella Tannarae*     | 2.05±1.94    | 0.7±1.28    | 0.026 |
| *Uncultured eubacterium E1-K12* | 2.0±1.98     | 0.51±1.25   | 0.005 |

Table 4. Species showing distinctive differences in the sputum microbiomes between currently active and former coal miners, all with CWP diagnosis.

Figures

![Bar graph showing differences in species abundances between non-smokers and smokers](image)

**Figure 1**

Taxonomic structure of the sputum microbiomes from CWP subjects and controls at the phyla level.
Figure 2

Occurrence frequencies of the main bacterial phyla in the sputum of CWP subjects compared to controls.
Figure 3

Two-dimensional diagram constructed by the method of principal components demonstrating the phylogenetic similarity of prokaryotic sputum communities in CWP subjects and controls.
Figure 4

Taxonomic structure of sputum microbiomes from CWP subjects and controls at the genera level.
Figure 5

Significant differences in the content of bacterial genera between smoking and non-smoking controls.