The Influence of Cigarette Smoke on Human and Murine Lung Microbiota

Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Draginja Kovačević

Borstel, 2020

Erster Gutachter: Prof. Dr. Holger Heine
Zweiter Gutachter: Prof. Dr. Thomas Roeder

Tag der mündlichen Prüfung: 11.02.2021

Table of content

I Abstract

Microbes are predominant inhabitants in the world and are in direct contact with our various body surfaces including skin, gut, genitals and respiratory tract. Host and microbes are interconnected and they both affect and depend on each other. However, this close connection can be disturbed by diverse external factors. Imbalance in microbial composition has already been associated with chronic inflammatory diseases, mostly of gut and skin, but recently these associations have been made regarding chronic lung diseases (CLD) as well. Hence, an increased interest in lung microbiome research has emerged in the last decade. It is known that smoking is one of the major risk factors for the development of CLDs, however its impact on microbial composition is still poorly understood. Therefore, the present work aimed to delineate effects of cigarette smoke (CS) exposure on the respiratory microbiota and its capacity to recover after smoking cessation.

To examine the role of CS on the microbiota of the upper (URT) and the lower respiratory tract (LRT), bronchoscopy was performed in healthy smokers, ex-smokers and never-smokers. Additionally, to get insights into early effects of smoking on the immune system and the microbial compositions of the lung and gut, mice were exposed to mainstream CS for four different treatment periods (7, 21, 37 or 56 days).

Results from the human study confirmed strong overlap of the oropharyngeal and the lung microbiota, but this dynamic relationship was disturbed by smoking. CS had an impact on the relative abundances of certain potentially pathogenic bacterial taxa in different areas of the respiratory system. Certain taxa were detected exclusively in the lungs and correlated positively with all smoking related parameters. In the mouse experiment, the most important changes upon CS exposure were loss of body weight**,** an increase of neutrophils and lymphocytes at later timepoints, and KC/GRO and VEGF inflammatory mediators in the bronchoalveolar lavage fluid. Also, an overexpression of Cyp1A1 and MMP12 genes and overproduction of MMP12 protein was detected in the lung tissue. Bacterial communities were dominated in all groups by Firmicutes and Bacteroidetes in the gut and by Firmicutes, Actinobacteria, Proteobacteria and Bacteroidetes in the lungs. However, only the gut microbiota changed significantly upon CS exposure, with lower relative abundance of *Prevotella* and higher relative abundance of *Clostridium* and *Mucispirillum*.

In conclusion, CS showed different effects on the microbial composition in the URT and the LRT, leading to a higher relative abundance of potential pathogens. Further, the mouse data indicated that longer exposure to mainstream CS induces a loss of body weight and an increases number of inflammatory cells with modulations of bacterial communities in the gut and lungs. However, the effect was more prominent in the gut than in the lungs. Of note, the obtained data regarding microbial composition are descriptive and thus limited to draw more detailed conclusions about functional meaning of observed changes.

II Zusammenfassung

Mikroben sind in unserer Welt die vorherrschenden Bewohner und stehen in direktem Kontakt mit verschiedenen Körperoberflächen wie Haut, Darm, Uro-Genitaltrakt und den Atemwegen. Dabei sind Wirt und Mikroben miteinander verbunden, beeinflussen sich gegenseitig und sind voneinander abhängig. Diese enge Verbindung kann jedoch durch verschiedene externe Faktoren gestört werden. Ein Ungleichgewicht in der mikrobiellen Zusammensetzung wird vor allem mit chronischen Entzündungskrankheiten, wie Darm- und Hautkrankheiten, in Verbindung gebracht. In letzter Zeit wurden diese Zusammenhänge jedoch auch bei chronischen Lungenkrankheiten (CLK) festgestellt. Daher ist in den letzten zehn Jahren ein zunehmendes Interesse an der Lungenmikrobenforschung entstanden. Es ist bekannt, dass Rauchen einer der Hauptrisikofaktoren für die Entstehung von CLK ist. Der Über den Einfluss auf die mikrobielle Zusammensetzung ist jedoch noch wenig bekannt.

Daher zielte die vorliegende Arbeit darauf ab, die Auswirkungen der Rauchexposition auf die respiratorische Mikrobiota und deren Regeneration zu untersuchen.

Um den Einfluss des Zigarettenrauchs (ZR) auf die Mikrobiota der oberen (OAW) und unteren Atemwege (UAW) zu untersuchen, wurde eine Bronchoskopie bei gesunden Rauchern, Ex-Rauchern und Nicht-Rauchern durchgeführt. Um tiefere Einblicke in die frühen Auswirkungen des Rauchens auf das Immunsystem und die Lungen- und Darm-Mikrobiota zu erhalten, wurden zusätzlich im Tierversuch Mäuse in vier verschiedenen Zeitspannen beraucht.

Die Untersuchungen der Probanden bestätigten eine starke Überlappung der Mikrobiota der Lunge, mit der des Mund-Rachenraums Diese dynamische Beziehung wurde durch das Rauchen gestört. ZR hatte einen Einfluss auf die relative Häufigkeit verschiedener potenziell pathogener Bakteriengruppen in verschiedenen Bereichen die Atemwege. Bestimmte Taxa wurden ausschließlich in der Lunge nachgewiesen und korrelierten positiv mit allen Rauch-Parametern der Probanden.

Im Tierversuch waren die wichtigsten Veränderungen nach ZR-Exposition der Verlust des Körpergewichts, der Anstieg der Entzündungsmediatoren KC/GRO und VEGF in der Lungenspülflüssigkeit und eine Überexpression von Cyp1A1- und MMP12-Genen im Lungengewebe. Nach längerem Berauchen (37 und 56 Tage) konnte zusätzlich ein Anstieg der Neutrophilen und Lymphozyten beobachtet werden (Lungenspülung).

Die Mikrobiota der berauchten und nicht-berauchten Tiere wurde durch Firmicutes und Bacteroidetes (Darm), bzw. von Firmicutes, Actinobacteria, Proteobakteria und Bacteroidetes (Lunge) dominiert. Eine signifikante Änderung nach ZR-Exposition war allerding nur in dem Darm zu finden. Hier reduzierte sich die relative Häufigkeit von *Prevotella* und es erhöhte sich die relative Häufigkeit von *Clostridium* und *Mucispirillum*.

Zusammenfassend lässt sich sagen, dass sowohl in der OAW als auch in der UAW deutliche bakterielle Reaktionen auf das Rauchen von Zigaretten beobachtet werden konnten. Dies führte zu einer höheren Häufigkeit von möglicherweise opportunistischen Pathogenen. Darüber hinaus wiesen die Daten des Tierversuchs darauf hin, dass eine längere Exposition gegenüber dem ZR einen Verlust des Körpergewichts und eine erhöhte Anzahl von Entzündungszellen mit Modulationen der Darm- und Lungen-Mikrobiota induziert. Die Wirkung war jedoch im Darm stärker ausgeprägt als in der Lunge. Erwähnenswert ist, dass die erhaltenen Daten bezüglich der mikrobiellen Zusammensetzung deskriptiv sind und daher nur beschränkt Rückschlüsse auf die funktionelle Bedeutung der beobachteten Veränderungen zulassen.

III List of Abbreviations

AhR	Aryl hydrocarbon receptor
ANOVA	Analysis of variance
APC	Allophycocyanin
BAL	Bronchoalveolar lavage
BUV	Brilliant ultraviolet
BV	Brilliant violet
CAST	Computer assisted stereology toolbox
CD	Cluster of differentiation
cDNA	Complementary deoxyribonucleic acid
COPD	Chronic obstructive pulmonary disease
CPD	Cigarettes per day
CS	Cigarette smoke
Ct	Cycle threshold
Cyp1A1	Cytochrome P450 family 1 subfamily A member 1
DAPI	4',6-diamidino-2-phenylindole
db-RDA	Distance-based redundancy analysis
DNA	Deoxyribonucleic acid
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
EPO	Erythropoietin
ExS	Ex-smoker
FCS	Fetal calf serum
FCS-A	Forward scatter area
FCS-H	Forward scatter height
FITC	Fluorescein isothiocyanate
g	Gram
GM-CSF	Granulocyte-macrophage colony-stimulating factor
h	Hours
HPRT	Hypoxanthine phosphoribosyltransferase
IFN	Interferon
IL	Interleukin
IP-10	Interferon gamma-induced protein 10

KC/GRO	Keratinocyte chemoattractant/human growth-regulated oncogene
KCl	Potassium chloride
KH2PO4	Potassium dihydrogen orthophosphate
KHCO3	Potassium bicarbonate
LRT	Lower respiratory part
M	Molar concentration
MDS	Multidimensional scaling
mg	Milligram
min	Minute
ml	Milliliter
mm	Millimeter
MMP12	Matrix metalloproteinase 12
Na2HPO4	Sodium dihydrogen phosphate monohydrate
NaCl	Sodium chloride
NaN_3	Sodium azide
NEC	Negative extraction control
NGS	Next generation sequencing
NH4Cl	Ammonium chloride
nM	Nanomolar
nm	Nanometer
nMDS	Nonmetric multidimensional scaling
NS	Never-smoker
NSW	Nasal swabs
NTC	Non-template controls
OCLN	Occludin
PAS	Periodic Acid-Schiff
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PERMANOVA	Permutational multivariate analysis of variance
PHS	Oropharyngeal swabs
pM	Picomolar
qRT-PCR	Quantitative real-time polymerase chain reaction
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid

RT	Reverse transcription
S	Smoker
SB	Super bright
SD	Standard deviation
SSC-A	Side scatter area
TBE	Tris-borate-EDTA
TBP	TATA-Box Binding Protein
TJP1	Tight junction protein 1
TMB	3,3',5,5'-Tetramethylbenzidine
TNF	Tumor necrosis factor
TPM	Total particulate matter
URT	Upper respiratory tract
UniFrac	Unique fraction metric
VEGF	Vascular endothelial growth factor
zOTUs	Zero-radius operational taxonomic units

1 Introduction

1.1 Morphology of the Respiratory Tract

The respiratory system can be divided into an upper and a lower part. The upper respiratory tract (URT) is located outside the thorax and composed of the nasal cavity, sinuses, pharynx and larynx. It conducts, warms, humidifies and partially filters passaging air. The lower respiratory part (LRT) contains the trachea and the bronchial tree formed by two bronchi that enter lungs and divide further into series of smaller bronchioles. The last division occurs at the level of the smallest alveolar sacs where gas exchange occurs[1]. The human respiratory tree consists of 20-23 generations, while that of the mouse has on average 13-17, with some morphological dissimilarities among different strains[2,3]. Beside the differences in the airways' generations and diameter size, human and mouse bronchial trees also show distinct branching patterns (dichotomic vs. monopodial)[2,4].

The main functional unit of the respiratory system are the lungs, paired pyramid-shaped organs located in the thorax. Each lung is composed of smaller units called lobes separated by fissures. In humans, the right lung has three (superior, middle and inferior) and the left lung two (superior and inferior) lobes[1]. In contrast, mouse lungs have four lobes on the right (superior, middle, inferior and post-caval) and one big lobe on the left side[5].

1.1.1 The Lung Epithelium

The airway epithelium acts as a frontline defense against inhaled exogenous insults. However, not only does it physically impede potential pathogens, it also actively contributes to the immune response[6,7]. The conductive airways are lined by a pseudostratified epithelium comprised of secretory, ciliated and basal cells, the proportion of which varies along the proximal-distal axis. Foreign particles normally get trapped in mucus secreted by goblet cells and are subsequently removed from the airways by the movement of ciliated cells. Basal cells are located at the bottom of the epithelium and serve as stem cells for secretory and ciliated cells[8].

Differences in the epithelium between humans and mice are reflected primarily in terms of thickness and presence of particular cell types. The intrapulmonary bronchioles of mice have very thin epithelium with majority of non-ciliated cells and very few goblet cells. In contrast, the human epithelium is much thicker and has considerably more goblet cells[9].

With successive divisions of the bronchi, the epithelium decreases in height and pseudostratification, both in humans and mice. Thus, the cell arrangement of the respiratory surface in the alveoli is quite distinct from the rest of the airways. The thin wall of alveolar compartment is highly vascularized and consists of two types of alveolar epithelial cells: major type I, which are essential for gas-exchange, and minor type II pneumocytes which produce surfactant[10,11].

1.2 Microbiota of the Respiratory Tract

Commensal microbes live in close contact with the host, and their ability to be immune-tolerated depicts an impressive characteristic of microbial survival and resilience[12]. Abundant studies have revealed the relevance of microbiota in maintaining host health[13–16]. To date, more research has been done in the field of the gut microbiome and its microbial composition is much better characterized than that of the respiratory tract. However, the development of new sophisticated approaches for the extraction of the low mass specimen has allowed us to get more detailed insight into the lung communities and their relationships with the host. Despite common embryonic origin of these two organs, and their similar organization of columnar epithelium, they differ fundamentally in terms of microbial composition and dynamics. In contrast to the unidirectional transport of microbes in the gut, there is a bidirectional flow into and out of the airways via breathing. Furthermore, while the entire gastrointestinal tract has a uniform temperature of 37 °C, inhaled air follows a gradient from the ambient temperature to the body temperature in the alveoli. Finally, oxygen levels and pH values differ considerably between these two organs. The lung microbiota is more impermanent and more dynamic compared to the gut. As a result of different local conditions, microbial compositions of these two organs are quite distinct[17–19].

Respiratory microbiota cover airway epithelium, forming a barrier between outer environment and the tissues. Microbes from the oral and nasal cavities presumably act as reservoirs for further colonization of the lower airways where the host defense system regulates immigration and elimination[20]. Besides air inhalation, microaspiration is considered to be another source of microbial immigration[18,21–23]. It has been confirmed that this event normally occurs in healthy subjects[24,25], yet it is more frequent in those with lung disease[26–28]. Despite the overlap of general bacterial communities, some of the healthy individuals showed distinct bacterial species present in the lungs which highly unlikely originated from the URT[29–32]. Additionally, the lung microbiota

seem relevant for baseline organ physiology and immune activity in spite of its lesser presence[16,22,33].

The importance of certain microbes for our health has been a focus of intensive research. Defined microbial strains have been frequently used as probiotics in animal models of respiratory infections. It has been well known that genera *Lactobacillus* and *Bifidobacterium* have protective role in certain lung inflammatory diseases[33–40]. Likewise, the application of *Acinetobacter lwoffi* to mouse skin induced Th1-associated anti-inflammatory cytokines, reducing allergic airway inflammation in mice[41]. All these findings indicate a valuable role of microbes in the development of the immune response.

1.2.1 Composition of the Respiratory Microbiota

In a healthy host, a balance exists among members of the microbiota, and the dominant commensal bacteria belong to the phyla Bacteroidetes and Firmicutes, followed by Proteobacteria and Actinobacteria[20] (Figure 1). The nasal cavity and nasopharynx considerably overlap in the bacterial communities with *Moraxella spp*., *Staphylococcus spp*., *Streptococcus* spp. and *Corynebacterium spp*. as their common inhabitants[42]. However, nasopharyngeal microbiota are more diverse and also contain *Dolosigranulum spp.* and *Haemophilus spp*.[43]. Going deeper into the oropharynx, bacterial diversity increases, and it is mostly characterized by *Neisseria* spp., *Rothia* spp., *Veillonella* spp., *Prevotella* spp. and *Leptotrichia* spp.[44,45]. In the human LRT, the most frequently identified genera are *Prevotella*, *Veillonella*, and *Streptococcus* with a lower contribution of *Fusobacterium* and *Haemophilus*[46,47].

The four main phyla present in the human lungs were also found to be dominant in mice, but with predominance of Firmicutes and Proteobacteria. The diversity of murine lung microbiota increases with age and it is very dynamic in the first eight weeks of life, mostly at the genus level[48,49]. After this turbulent period, the microbial composition is mainly considered as stabile[49,50].

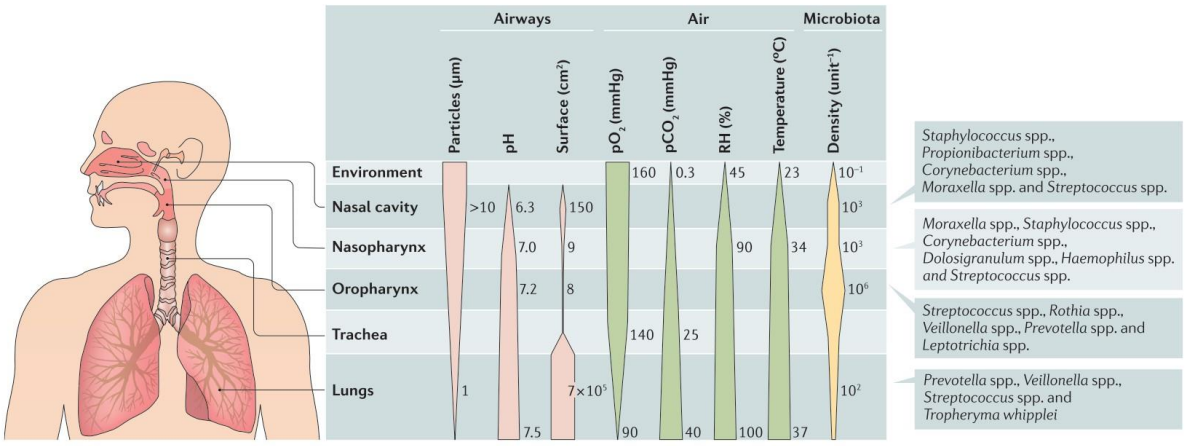


Figure 1. Physiological and microbial gradients along the respiratory tract.
Taken from: Man, W., de Steenhuijsen Piters, W. & Bogaert, D. The microbiota of the respiratory tract: gatekeeper to respiratory health. Nat Rev Microbiol 15, 259–270 (2017)[51]

Despite the commonly present dominant phyla, the lung microbiota display greater variation between individuals than between different parts of the respiratory tract. Therefore, it is difficult to agree on definition of core microbiota in healthy individuals.

1.2.2 Alterations of the Lung Microbiota

In healthy lungs, there is no selective pressure and all microbes have an equal potential to compete within the local environment, with lower importance of growth conditions[52]. In contrast, microbes of diseased lungs are under active selection by altered local growth conditions which afterwards change their composition[31,46]. Thus, the ability to adapt to a new status gives certain bacteria an advantage that exceeds the influence of immigration and elimination. Alterations in the diversity of the lung microbiota have been implicated in chronic respiratory diseases such as chronic obstructive pulmonary disease (COPD) and asthma[53–56] (Figure 2). In one of the first lung microbiome studies, Hilty et al.[46] demonstrated that oropharyngeal and lung samples of healthy controls had distinct microbiome from those found in the samples of asthmatic and COPD patients.

Several studies on COPD patients showed predominance of Proteobacteria (mostly *Haemophilus* or *Proteobacteria spp.*) or Firmicutes (*Veillonella*) and decrease of Bacteroidetes (*Prevotella* sp.)[46,56–59]. Still, the relationship the between microbiota and the disease status is not straightforward. Even though higher abundance of certain microorganisms in the lungs may contribute to airway inflammation, their role in clinically stable patients remains unclear. More severe COPD patients had lower bacterial diversity in the lungs[57,60], mostly with a dominance of

Pseudomonas spp.[61], but those with mild symptoms were more similar to the healthy controls[57]. Interestingly, patients with the advanced disease had distinct microbiota at different micro-anatomic sites of the same lung[57]. In another study, both mild and severe COPD patients showed less diverse microbiota which notably differed from healthy groups of smokers and non-smokers[62]. While Mayhew at al.[58] demonstrated that a decrease of microbial stability and diversity in sputum samples was associated with a higher frequency of exacerbations in COPD, Pragman et al.[54] associated the disease progression with higher diversity found in BAL fluid. All this evidence clearly demonstrates the heterogeneity of COPD and highlights the importance of bacteria in exacerbations. Furthermore, the treatment during exacerbations influenced microbial composition differently when based on antibiotics, reducing the abundance of Proteobacteria, and corticosteroids, which favored an over-representation of certain taxa[63,64].

Asthmatic patients were mostly characterized with dominance of Proteobacteria and Firmicutes phyla, mainly due to an increase of their pathogenic species *Haemophilus spp.* and *Staphylococcus spp.*. Conversely, Bacteroidetes were less abundant, particularly because of fewer *Prevotella spp.*[46,65,66]. As asthma is also a heterogeneous disease, studies conducted until now could not show uniform changes of the lung microbiota. Thus, some of them reported an increase of the diversity in asthmatics[67], whereas others showed a decrease with a predominance of certain taxa[16,46,66,68]. Moreover, the lung microbiome differed between the patients with neutrophilic and eosinophilic asthma[66,69].

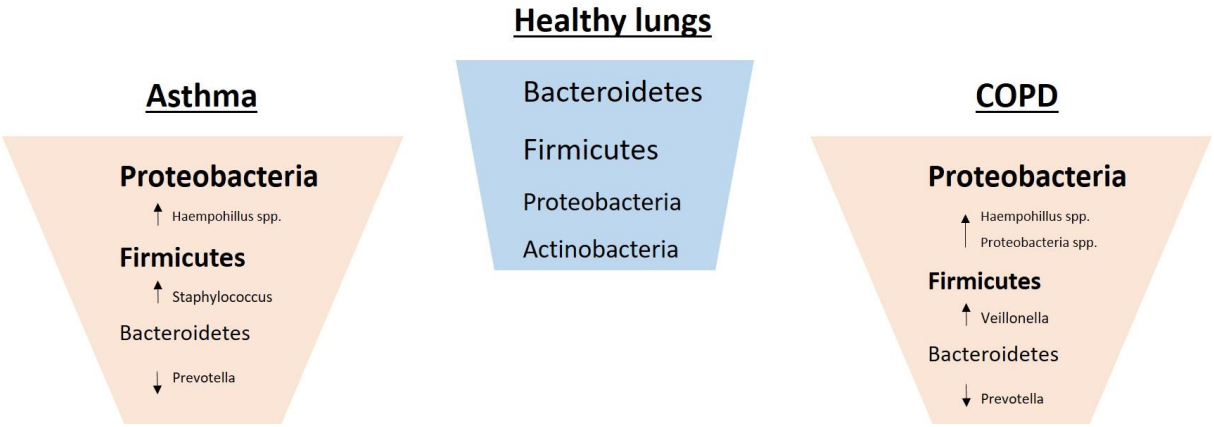


Figure 2. Lung microbiota composition

1.3 Host-bacteria Interactions at the Epithelium

As described above, polarized airway epithelium is in constant communication with the external environment and represents the first barrier to inhaled microbes and potentially harmful

particles. Emerging data have shown that aberrant regulation of the interaction between microbes and the host in different organs and at different body surfaces impacts pathophysiology of chronic inflammatory disorders[70–73]. Therefore, the cross talk between the epithelium and the commensal microbes is clearly essential in the maintenance of tissue homeostasis[12,74,75]. Important components of the lung and gut epithelium are intercellular tight junctions and extracellular matrix (ECM) which preserve the tissue integrity and prevent entrance of the pathogens. As these components represent binding sites for many bacteria, their aberrations can disrupt binding affinities and completely change bacterial composition in the tissue. In gut microbiome research, several studies have reported that commensal bacteria and their metabolites enhance tight junction integrity, thus protecting the intestinal epithelial lining[76–79]. Additionally, the mucus layer protects the epithelium from injuries and provides beneficial nutrients to maintain microbial coexistence and diversity[80–82]. However, airway obstruction can lead to an excess production of mucus which is more difficult for ciliated cells to remove. Its longer residence on the epithelium may favor the selection of certain bacteria, causing microbial imbalance[83]. Therefore, it is conceivable that changes in the ratio of commensals and pathogens modulate their interactions with the host epithelium which could then trigger dysregulated immune responses.

1.4 Impact of Cigarette Smoke on the Lungs and its Microbiota

The respiratory tract can be easily affected by various particles carried by air. Among them, one of the most common and widely inhaled is cigarette smoke (CS), which is well known to cause respiratory diseases[84]. It has been known that inflammation plays an important role in the development of these diseases, yet the exact mechanisms remain unclear. Numerous studies have reported that CS activates and expands infiltration of inflammatory cells in the lungs, resulting in increased levels of proinflammatory cytokines and chemokines[85–87]. Additionally, smoking and air pollution exposure reduce mucociliary clearance of bacteria[88] and promote the creation of pathogen-rich biofilms at the epithelium[89]. By altering the tissue conditions in the URT and LRT, CS can disturb bacterial homeostasis and affect its composition. Therefore, an increased interest in the link between lung disease and its microbiota has emerged in recent years. In healthy individuals, cigarette smoking has so far been linked either to microbial changes of the oral[90,91] and nasal[92,93] microbiomes or to that of the LRT[57,59,62]. One of the common hypotheses is that CS favors an unstable microbial community which is prone to inflammatory

host response. However, there is a lack of information whether the lung microbiome retrieves after smoking cessation, and if does, how much time is necessary for the recovery. Characterization of early microbial shifts could open new possibilities for secondary prevention or therapeutic modulation for respiratory diseases.

2 Aim of the Study

Historically, microbial research has been mostly focused on the impact and mechanisms of certain pathogens on host's health, leaving commensal microbes neglected in some way. Today it is clear that commensals, which reside in various body tissues, exert a crucial role in homeostasis and diseases of the host.

Furthermore, most lung microbiome studies have been based on samples from the patients with COPD, asthma or cystic fibrosis while insights into the microbiome of healthy individuals remain scarce.

Additionally, the current understanding of how CS impacts airway microbiota is based on the studies restricted to one or two compartments of the tract. Thus, comprehensive analyses of the whole respiratory tract are needed to elucidate its microbial dynamics upon smoking. Also, there is a lack of knowledge on the capacity of the respiratory microbiota to recover from smoking-induced insults.

Therefore, the specific objectives of this study were:

1. To characterize the microbial composition in the nasal and oropharyngeal cavities and the lungs in well-characterized smoking and never-smoking individuals without any evidence of lung disease.
2. To assess whether microbial changes observed in smokers revert to a "normal" state after smoking cessation.

In addition, as it is hardly possible to obtain information about early effects of CS exposure, the human study was mimicked in the mouse smoking-model,

3. To test how smoking affects body weight of mice.
4. To test whether smoke exposure disrupts lung epithelium and causes lung injury.
5. To assess inflammatory response upon smoke exposure in a murine model.
6. To check how smoking affects microbial composition of the murine lungs and gut.

3 Material and Methods

3.1 Material

3.1.1 Chemicals and Reagents

Substance	Manufacturer	Country
Agar-Agar, Bioscience, granulated	Carl Roth	Karlsruhe (GER)
Ammonium chloride	Merck	Darmstadt (GER)
Aqua B. Braun	B. Braun	Melsungen (GER)
Bench Top 1kb DNAladder	Promega	Madison, WI (USA)
Collagenase from *Clostridium histolyticum*	Sigma-Aldrich	St. Louis, MO (USA)
DAPI	Invitrogen™ Thermo Fisher	Carlsbad, CA (USA)
Ethanol ≥ 99.8 %	Carl Roth	Karlsruhe (GER)
Glycogen, RNA grade	ThermoScientific	Waltham, MA (USA)
MACS buffer	Sigma-Aldrich	St. Louis, MO (USA)
Fetal calf serum, FCS	PAN-Biotech	Aidenbach (GER)
Ketamine (Ketamidor 100 mg/ml)	WDT	Garbsen (GER)
Lysozyme	Roche Diagnostics	Mannheim (GER)
QIAzol lysis reagent	Qiagen	Hilden (GER)
Paraformaldehyde	Sigma-Aldrich	St. Louis, MO (USA)
Potassium bicarbonate	Merck	Darmstadt (GER)
Potassium dihydrogen orthophosphate	Carl Roth	Karlsruhe (GER)
Potassium chloride	Carl Roth	Karlsruhe (GER)
Rotiphorese 10xTBE buffer	Carl Roth	Karlsruhe (GER)
Roti®-Histofix, formaldehyde 4 %	Carl Roth	Karlsruhe (GER)
SYBR Green II	Lonza	Basel (CH)
Sodium acetate solution	ThermoScientific	Waltham, MA (USA)
Sodium chloride	Sigma-Aldrich	St. Louis, MO (USA)
Sodium dihydrogen phosphate monohydrate	Merck	Darmstadt (GER)
Trypan blue	Sigma-Aldrich	St. Louis, MO (USA)
UltraPureTM agarose	Invitrogen™ Thermo Fisher	Carlsbad, CA (USA)
Xylazine (Xylavet 20 ng/ml)	CP-Pharma	Burgdorf (GER)

Table 1. Chemical and reagents

3.1.2 Buffer and Solutions

Names	Substance	Volume
FACS wash buffer	PBS	225 ml
	FCS	25 ml
	NaN_3	2.5 g
Gey'sche lysis buffer 1 l, pH 7.2	$KHCO_3$	0.5006 g
	NH_4Cl	4.14547 g
	EDTA	18.612 mg
	Aqua dest.	500 ml

MSD wash buffer	PBS	47.5 ml
	Tween-20	2.5 ml
Paraformaldehyde 4 %	PBS	960 ml
	Paraformaldehyde	40 ml
Phosphate-buffered saline - PBS	KCl	2 g
	KH_2PO_4	2.4 g
	NaCl	80 g
	Na_2HPO_4	14.4 g
RNA loading buffer	Formamide, deionized	18.0 µl
(per sample)	5 x TBE Buffer	3.0 µl
	SYBR Green II (1:100)	1.5 ml

Table 2. Buffers and solutions

3.1.3 Commercial Kits

Kit	Manufacturer	Country
Agencourt AMPure XP-kit	Beckman Coulter	Brea (USA)
DNF-473 Standard sensitivity NGS fragment analysis kit	Advanced Analytical Technologies	Ankeny, IA (USA)
MiSeq reagent kit v3	Illumina	San Diego, CA (USA)
Mouse MMP12 SimpleStep ELISA kit	Abcam	Cambridge (UK)
Mouse/rat cotinine ELISA kit	Calbiotech	El Cajon, CA (USA)
MSD U-Plex platform kit	Meso Scale Diagnostics	Rockville, MD (USA)
NEBNext high fidelity 2x PCR master mix	New England Biolabs	Ipswich (UK)
Nextera XT index kit v2 (Set A & C)	Illumina	San Diego, CA (USA)
Periodic Acid Schiff's staining kit	Carl Roth	Karlsruhe (GER)
PureLink genomic DNA mini kit	ThermoFisher Scientific	Waltham, MA (USA)
PureLink microbiome DNA purification kit	ThermoFisher Scientific	Waltham, MA (USA)
QIAamp PowerFecal DNA Kit	Qiagen	Venlo (NL)
QuantiNova reverse transcription kit	Qiagen	Hilden (GER)
RAF Diff-Quik kit	RAL Diagnostics	Martillac (FR)
RNeasy mini kit	Qiagen	Hilden (GER)

Table 3. Commercial kits

3.1.4 Commercial Control Genomes

Name	Manufacturer	Country
HM-782D	BEI Resources	Manassas, VA (USA)
HM-783D	BEI Resources	Manassas, VA (USA)
PhiX control v3	Illumina	San Diego, CA (USA)

Table 4. Commercial control genomes

3.1.5 Oligonucleotides

Gene name	Forward (5' - 3')	Reverse (5' – 3')
AhR	TTCTATGCTTCCTCCACTATCCA	GGCTTCGTCCACTCCTTGT
Cyp1A1	CGTTACCTGCCTAACTCTTC	ATGCTCAATGAGGCTGTCTG
HPRT	CAGGCCAGACTTTGTTGGAT	ACGTGATTCAAATCCCTGAAGT
MMP12	CTGGGCAACTGGACAACTCA	AGATGCTGTACATCGGGCAC

OCLN	CCTCCACCCCCATCTGACTA	GCTTGCCATTCACTTTGCCA
TBP	AATTGTACCGCAGCTTCAAAAT	ATGATGACTGCAGCAAATCG
TJP1	ACGAGATGCTGGGACTGACC	AACCGCATTTGGCGTTACAT

Table 5. Primer sequences for qRT-PCR reactions

Name	Target region	Sequence (5' – 3')
347F	16s rRNA V3 – V4 hypervariable regions	GGAGGCAGCAGTRRGGAAT
803R		CTACCRGGGTATCTAATCC

Table 6. Primer sequences for 16s rRNA amplicon PCR

3.1.6 Antibodies

Marker	Fluorophore	Clone	Manufacturer	Country
B220	BV510	RA3-6B2	BioLegend	San Diego, CA (USA)
CD19	SB600	eBio1D3	Miltenyi Biotec	Cologne, Germany
CD3	BUV737	17A2	BioLegend	San Diego, CA (USA)
CD4	APC-H7	GK1.5	BioLegend	San Diego, CA (USA)
CD45	APC	30-F11	BioLegend	San Diego, CA (USA)
CD8b	FITC	YTS156.7.7	BioLegend	San Diego, CA (USA)
Fc Block (CD16/CD32)	-	2.4G2	BD Pharmingen	San Diego, CA (USA)
Zombie UV Dye	DAPI	-	BioLegend	San Diego, CA (USA)

Table 7. Antibodies

3.1.7 Analytes

Cytokine/Chemokine	Manufacturer	Country
EPO	Meso Scale Diagnostics	Rockville, MD (USA)
GM-SCF	Meso Scale Diagnostics	Rockville, MD (USA)
IFN-γ	Meso Scale Diagnostics	Rockville, MD (USA)
IL-1β	Meso Scale Diagnostics	Rockville, MD (USA)
IL-33	Meso Scale Diagnostics	Rockville, MD (USA)
IL-6	Meso Scale Diagnostics	Rockville, MD (USA)
IP-10	Meso Scale Diagnostics	Rockville, MD (USA)
KC/GRO	Meso Scale Diagnostics	Rockville, MD (USA)
TNF-α	Meso Scale Diagnostics	Rockville, MD (USA)
VEGF	Meso Scale Diagnostics	Rockville, MD (USA)

Table 8. Analytes for MSD Assay

3.1.8 Miscellaneous Consumables

Consumables	Manufacturer	Country
15 ml conical centrifuge tubes, 120 x 17 mm	Sarstedt, Inc.	Nümbrecht (GER)
20G needles	BD	Heidelberg (GER)
50 ml conical centrifuge tubes, 114 x 28 mm	Sarstedt, Inc.	Nümbrecht (GER)
70 µm cell strainer	Corning	Wiesbaden, GER
96 Well cell culture plate	Corning	Wiesbaden, GER
Cryotubes	Sarstedt, Inc.	Nümbrecht (GER)
Cytofuge filters	ThermoScientific	Runcorn (UK)
DNA free water PCR grade	Molzym	Bremen (GER)

Microscope slides	Langenbrinck	Emmendingen (GER)
Microscope SuperFrost®Plus slides	Langenbrinck	Emmendingen (GER)
PCR plates	Roche Diagnostics	Mannheim (GER)
Petri dish 92x16mm	Sarstedt, Inc.	Nümbrecht (GER)
Reaction tubes (0.2 ml, 0.5 ml, 1.5 ml)	Eppendorf	Hamburg (GER)
Research cigarettes 3R4F	University of Kentucky	Lexington, KY (USA)
Sealing foil for PCR plate	Roche Diagnostics	Mannheim (GER)
Surgical disposable scalpels	B. Braun	Melsungen (GER)
Syringes (1 ml, 2ml, 5 ml)	BD	Heidelberg (GER)
Tracheal cannulas (No. 730029)	Harvard Apparatus	Cambridge (UK)

Table 9. Miscellaneous consumables

3.1.9 Equipment and Devices

Equipment/device	Name	Manufacturer	Country
Absorbance plate reader	Tecan Sunrise multiplate reader	Tecan Trading AG	Männedorf (GER)
Blood cell differential blood counter	Counter AC-8	Glaswarenfabrik Karl Hecht GmbH & Co KG	Sondheim vor der Rhön (GER)
Bronchoscopes	- -	Fujifilm Pentax	Tokyo (JPN) Tokyo (JPN)
Capillary electrophoresis instrument	Fragment analyzer	Advanced Analytical Technologies	Ankeny, IA (USA)
Centrifuges	Eppendorf centrifuge 5415 R Multifuge X1R (plate centrifuge) Cytospin 3 (cytospin centrifuge)	Eppendorf AG Heraeus Holding GmbH SHANDON	Hamburg (GER) Hanau (GER) Cambridge (UK)
Electrophoresis power supply	Powerpac basic	Biorad	Hercules, CA (USA)
Embedding machine	HistoStar	ThermoFisher Scientific	Waltham, MA (USA)
Flow cytometer	BD Accuri C6 FlowCytometer	BD Biosciences	Franklin Lakes, NJ (USA)
Fluorospectrometer	NanoDrop1000	Thermo Fisher Scientific	Mariette, OH (USA)
Freezers/Fridges	Privileg öko Energiesparer (-20°C) VWR60086V (-80°C) Privileg de lux (4°C)	Quelle GmbH Thermo Fisher Scientific Quelle GmbH	Fürth (GER) Mariette, OH (USA) Fürth (GER)
Fume hood	Köttermann Systemlabor Typ 2-453-DAND	Köttermann Systemlabor	Uelze-Hänigsen (GER)
Gel chamber	PEQLAB, comb 1.5 mm 24 well	PEQLAB Biotechnologie GmbH	Erlangen (GER)

Gel imaging system	ChemiDoc touch imaging system doc	BIORAD	Hercules, CA (USA)
Haemocytometer	Neubauer	Marienfeld Superior	Lauda-Königshofen (GER)
Homogenizer	Precellys24	Bertin Technologies	Montigny le Bretoneux (FR)
Incubator	C02-AUTO-ZERO	Heraeus Holding GmbH	Hanau (GER)
Laminar flow hoods	BDK SCANLAF MARS	Weiss Pharmatechnik LaboGene (Scandinavian by Design)	Oldenburg (GER) Allerød (DK)
Liquid nitrogen tank	Liquid nitrogen	Westfalen AG	Münser (GER)
Microscopes	BX40 Zeiss	Olympus Zeiss	Hamburg (GER) Oberkochen (GER)
Microtome	Microtome HM340E (with waterslide and cooling block)	Thermo Fisher Scientific	Waltham, MA (USA)
Microwave	MW 7873	Severin	Sundern (GER)
Mouse cage system	IVC Greenline SEALSAFE® PLUS	Tecniplast	Buguggiate (IT)
Multichannel pipettes	Eppendorf multichannel pipettes	Eppendorf	Hamburg (GER)
Multi-array reader	MESO QuickPlex SQ120	Meso Scale Diagnostics	Rockville, MD (USA)
Sequencer	MiSeq	Illumina Inc	San Diego (CA, USA)
Smoking robot	SCIREQ InExpose system	SCIREQ	Montreal (CA)
Swabs	Mast swab	Mast Group Ltd	Bootle (UK)
Thermocyclers	Biometra TRIO thermal cycler Light cycler 480 II	Biometra biomedizinische Analytik GmbH Hoffmann-La Roche	Göttingen (GER) Basel (CH)
Tissue dehydrator	Autotechnicon	Thermo Fisher Scientific	Mariette, OH (USA)
Tissue lyser	Tissue lyser II	Qiagen	Venlo (NL)
Tissue processor	Microm STP120	Thermo Fisher Scientific	Waltham, MA (USA)
Tool Sterilizer	Hot bead sterilizer	Fine Science Tools	Heidelberg (GER)
TPM detector	MicroDust Pro	Casella	Scarborough (UK)
Warming plate	ThermoLux warming plate	Witte+Sutor GmbH	Murrhardt (GER)
Water bath	Haake W13	Thermo Fisher Scientific	Waltham, MA (USA)

Table 10. Equipment and devices

3.1.10 Software

Name	Version	Source
Adobe Illustrator	CS6	Adobe (San Jose, CA, USA)
BD FACSDiva	v8.0.1	BD Bioscience-US (San Jose, USA)
DISCOVERY WORKBENCH	v4.0	Meso Scale Diagnostics (Rockville, MD, USA)
Ensembl	-	https://www.ensembl.org/index.html
FCS Express	v6.06.0021	De Novo Software (Pasadena, CA, USA)
FlexiWare (inExpose)	v6.1	Scireq (Montreal, CA)
GraphPad Prism	v6	GraphPad Software, Inc. (San Diego, CA, USA)
Image Lab	v2.0.1 build 18	Bio-Rad Laboratories (Hercules, CA, USA)
LightCycler® 480	v1.5.1	Roche (Mannheim, GER)
Magellan	-	Tecan (Männedorf, GER)
newCAST	-	Visiopharm (Hoersholm, DK)
NIS-Elements D	v3.1	Nikon (Tokyo, JPN)
Primer-BLAST	-	https://www.ncbi.nlm.nih.gov/tools/primer-blast/
Standard Nucleotide-BLAST		https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch
R	v3.4.3 phyloseq Rhea pipeline	R Fondation[94] McMurdie et al.[95] Lagkouvardos et al.[96]
RDP database	rdp_16s_v16 (v.16)	Ribosomal Database Project[97]
Reverse complement	-	http://reverse-complement.com/
Usearch	v.10.2.240	Edgar et al.[98]
Vsearch	v.2.80	Rognes et al.[99]

Table 11. Software

3.1.11 Mouse Strain

Strain	Vendor	Country
Balb/c	Charles River	Sulzfeld (GER)

Table 12. Mouse stain

3.2 Methods

3.2.1 Human Study

3.2.1.1 Cohort Recruitment

Subjects from two cohorts were included in this study and both were recruited in the Center for Clinical Studies at the Research Center Borstel, Germany. Cohort 1 ("Borstel Healthy Cohort") was recruited from February 2016 to January 2017, while recruitment of Cohort 2 ("LuMEn cohort") started in March 2017 and finished in February 2020. Sample collection and processing were performed at the Research Center Borstel, while microbiome complete analyses were done

at the Helmholtz Center Munich in the **Research Unit for Comparative Microbiome Analysis.** Caucasian adult individuals (≥ 18 years) were enrolled in the study based on defined eligibility criteria: if they had no evidence of (1) a respiratory infection or systemic glucocorticoid therapy in the month, (2) antibiotic therapy in the last two months, (3) diabetes mellitus, (4) pregnancy or lactation, (5) active tuberculosis or history thereof, (6) immunosuppression, (7) known pulmonary disease (except for COPD stage GOLD I/II). Participants were excluded if they had a medical contraindication for bronchoscopy as judged by the investigator (e.g. allergy to sedative medication).

Participants were assigned either to smokers, ex-smokers or never-smokers, according to their smoking history as defined by the Centers for Disease Control and Prevention[100]. Smokers were further classified according to smoking intensity and duration. Long-term smokers were defined by ≥ 10 cigarettes per day, ≥ 10 packyears. Short-term smokers differed from long-term smokers by a shorter smoking history: ≥ 10 daily cigarettes, but < 10 packyears. Mild smokers consumed less than 10 daily cigarettes and had < 5 packyears. The study was approved by the independent Ethics Committee of the University of Luebeck (cohort 1, ref.15-194; cohort 2, ref. 16-145). Oral and written informed consent from all study subjects was obtained according to ICH/GCP standards. Cohort 1 is registered at the German Register for Clinical Trials (DRKS00016932). Cohort 2 is registered at clinicaltrials.gov (NCT03562442).

3.2.1.2 Sample Collection and Processing

In total, samples from 58 participants were collected in the study (Figure 3). SOP-guided biosampling and processing included deep bilateral nasal swabs (NSW; n=46), bilateral oropharyngeal swabs (PHS; n=54) and bronchoalveolar lavage (BAL) samples (n=52). PHS and NSW (Mast swab, Mast Group Ltd., UK) were frozen in cryotubes on dry ice immediately after sampling. Unused sterile swabs were included for microbiome analyses as controls per lot. Participants underwent flexible bronchoscopy according to current German guidelines[101]. The bronchoscope was wedged into a sub-segmental bronchus of the middle lobe. BAL was performed in 15 x 20 ml fractions up to a total volume of 300 ml using sterile saline solution (0.9%). Sterile saline solution was added as control for the microbiome analyses. Due to possible contamination, the recovery of the first BAL fraction was discarded. The remaining fractions were pooled and frozen at -80°C until they were transported to the Helmholtz Center Munich. Nicotine and its metabolites cotinine, 3-hydroxy-cotinine and anabasine were determined in unprocessed urine samples from each participant by liquid chromatography tandem mass spectrometry.

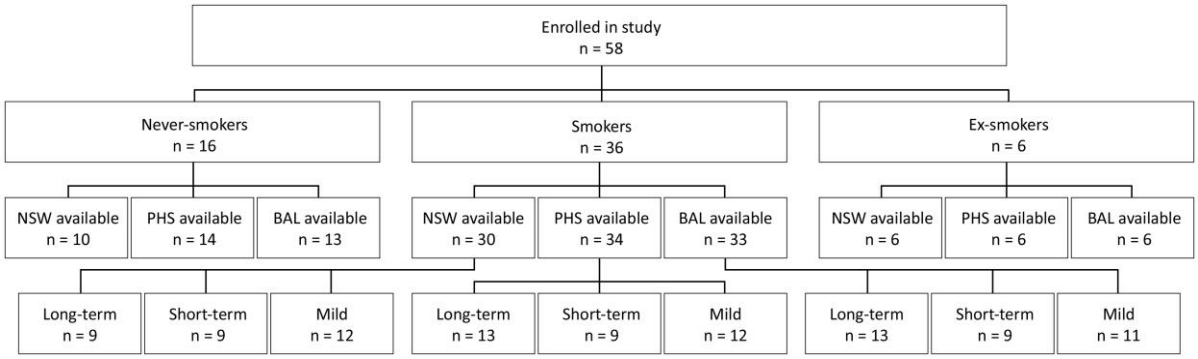


Figure 3. Graphic summary of human sample distribution.

3.2.1.3 Microbiome Analysis

3.2.1.3.1 DNA Extraction

Due to the low bacterial content of the samples and therefore high risk of contamination, all steps for deoxyribonucleic acid (DNA) extraction, polymerase chain reaction (PCR), and amplicon library preparation were performed using biosafety cabinets, as it is recommended for low biomass samples[102]. 5ml BAL was unfrozen and centrifuged at 24000 x g for 10 min. Thereupon, cell pellets were dissolved in 180 µl lysozyme solution (10 mg/ml) and incubated for 45 min at 37°C. DNA extraction from the pellets was performed using the PureLink Genomic DNA Mini kit (ThermoFisher Scientific, USA) according to the manufacturer's protocol for Gram-positive bacteria. As the supernatants from the initial centrifugation step could contain free DNA released during the freezing and thawing of the BAL, the extraction was performed via standard ethanol precipitation of nucleic acids. The extracts collected from the supernatants and the pellets were then combined. DNA from the NSW and PHS samples was extracted using the PureLink Microbiome DNA Purification kit (ThermoFisher Scientific, USA) following the manufacturer's instructions for buccal swabs. Additionally, to identify contaminants deriving from the kits and the working environment, DNA extraction control samples were included. For NSW and PHS, 2 PureLink Microbiome DNA Purification kit controls (per lot) were used as well as 2 unused sterile NSW and PHS. Within the extraction procedure from the BAL samples, four controls were included: two for the ethanol precipitation of the supernatant (one per each extraction procedure) and two controls of the PureLink Genomic DNA Mini kit (one per each extraction procedure). In addition, to avoid confounding by potential contamination of the saline solution and bronchoscope flushing prior to use, DNA was extracted from 5 ml of sterile saline solution and 5ml of bronchoscope flushings of three samples.

3.2.1.3.2 16s rRNA Gene Library Construction

16S ribosomal ribonucleic acid (rRNA) gene analysis is a common amplicon sequencing method used for identification and classification of bacteria present within a sample. This gene has extremely slow rate of evolution and its hypervariable region demonstrates diversity among different bacterial species[103,104].

3.2.1.3.2.1 Amplicon PCR

As presence of human nuclear and mitochondrial DNA can strongly affect the accuracy of amplicon sequencing, primers were chosen based on their exclusive specificity for bacteria and broad taxonomic coverage. Therefore, primer pair 347F-803R (Table 6), which amplifies the hypervariable regions V3-4 of the 16S rRNA gene, was used[105]. Illumina overhang adapter sequence was allocated to each sample at 5' primer end. The library preparation as well as amplicon sequencing were performed using a MiSeq sequencer (Illumina Inc., USA) according to the slightly modified manufacturer's guidelines[106]. All samples were run in triplicates and pooled afterwards. Each amplicon PCR reaction contained 12.5 µl NEBNextHighFidelity 2xPCRMasterMix (New England Biolabs, Ipswich, USA), 0.5 µl primers and between 10-80 ng of template DNA. To assess the presence of contaminants and to accurately quantify bacterial abundances, exogenous spike-in bacteria were added to the samples. Therefore 1/100 dilutions of the mock community HM-782D (BEI Resources, NIAID, NIH, USA) were used for one negative extraction control (NEC) of each sample type, one non-template PCR control (NTC) and one sample (BAL from never-smoker) representing a positive control. Operating conditions are described in the Table 13.

Step	Cycle	Time [sec]	Temperature [°C]
Initial denaturation	1	30	98
Denaturation		30	98
Annealing	28	45	58
Elongation		30	72
Final elongation	1	300	72

Table 13. Amplicon PCR Protocol for human samples.

Success of each PCR was verified by agarose gel electrophoresis and pooled triplicates were purified using Agencourt AMPure XP paramagnetic beads (Beckman Coulter, USA). Products were checked for dimers and were quantified using the DNF-473 Standard Sensitivity NGS Fragment Analysis kit on the Fragment Analyzer (Advanced Analytical, USA).

3.2.1.3.2.2 Index PCR

Sample-specific indices, Nextera XT Index kit v2 Set A and Set C (Illumina, Inc., USA), were added to compatible overhang adapters to enable identification after multiplex sequencing. The index

PCR reactions contained 2.5 µl of forward and reverse indexing primers, 12.5 µl 2xPCR Master Mix and 10 ng the NEBNext High-Fidelity 2xPCR Master Mix (New England Biolabs, USA). Operating conditions are described in the Table 14.

Step	Cycle	Time [sec]	Temperature [°C]
Initial denaturation	1	30	98
Denaturation		30	98
Annealing	8	30	55
Elongation		30	72
Final elongation	1	300	72

Table 14. Index PCR program for human samples.

Verification of PCR success, purification and quantification of PCR products were performed as it was described for the amplicon PCR. The products were diluted to 4 nM/l and pooled prior to the sequencing.

3.2.1.3.3 Sequencing

Sequencing was performed on a MiSeq System sequencer (Illumina, Inc., USA) using the MiSeq® Reagent kit v3 (600 cycles) for paired-end sequencing. For the run, 8pM of DNA spiked with 20 % PhiX were loaded and the sequencing was performed according to the MiSeq System User Guide (Illumina, Inc., USA). The obtained reads are available under the BioProject accession PRJNA501806 of the NCBI. Processing of the raw reads was done according to the FASPA[107] protocol. First, raw reads were demultiplexed and merged into a single file using Usearch v.10.2.240 software[98]. Next, primer removal and quality filtering were done via Vsearch software where all reads with more than one error were removed. In the dereplication step, redundant replicates were reduced to unique sequences, which were sorted by decreasing abundance. All following steps were performed using Usearch. Sequencing errors such as chimeras were identified and excluded via denoising with Unoise3[108] command and reads were clustered into zero-radius operational taxonomic units (zOTUs). Further, zOTUs below 350 bp were removed. The resulting zOTU table was reanalyzed using the UNCROSS algorithm[109] which removes sequencing errors that could lead to a wrong barcoding assignment through cross-talk[110]. In the next step, taxonomy down to the genus level was predicted using RDP reference database v16[111] using the SINTAX algorithm[112] with a confidence cut-off at 0.5. Further taxonomy classification to the species level was performed for selected zOTUs using blastn[113] tool. As short reads of the 16S rRNA gene are not reliable in allowing phylogenetic resolution to the species level, the obtained sequence variants were assigned to species if the following criteria were met: (1) 100% percent

sequence match to an isolate/cultivate of which the full 16S rRNA gene sequence was available; (2) unambiguous assignment – the zOTU sequence matches only to one particular bacterial species. Further, a taxonomic tree in the Newick format was generated with the cluster–agg command. Finally, non-prokaryotic zOTUs (eukaryotes, chloroplast, and mitochondria) as well as singletons were removed from the zOTU table. Total number of zOTUs was 7057192 in 250 total samples (28229 zOTUs per sample).

As described above, negative controls were included at all stages of laboratory processing as it was recommended for respiratory microbiota studies[114]. DNA extraction control samples had a much lower read number and the taxa which were most abundant in the samples were mainly not detected in the associated controls. Also, PCR controls had a none or almost no reads. The usage of a denoising algorithm further allowed distinction of contaminating zOTUs from sample zOTUs of the same lineage via oligotyping[115]. zOTUs which were detected in all NECs and the mock community NECs were removed as it was recommended by Salter *et al*[114].

To further estimate the potential impact of contamination, the controls spiked with mock community were compared. Spiking of negative PCR and extraction controls showed an expected number and taxonomic assignment of zOTUs of the expected mock strains. In particular, NTC had 48981 of 49343 of the reads assigned to the mock community strains, the NEC had 94100/95692 zOTUs assigned to mock strains. The sample spiked with and without the mock community showed that the used amount of mock DNA was comparable to that of the sample (19551/43698). Data processing also successfully discerned between mock strains and real sample strains showing an overlap of the sample specific zOTUs in the spiked and not spiked sample.

3.2.2 Mouse Model

3.2.2.1 Animals

The study was performed under the guidelines for the use and care of laboratory animals and was approved by the Government of the District of Schleswig-Holstein (V244-4782/2017 (34-3/17)). Female BALB/c mice were purchased from Charles River Laboratories (Suzfeld, Germany) at the age of six weeks and were housed in individually ventilated cages (Sealsafe PLUS, Green Line, Techniplast, Italy) with unlimited access to food and water. The mice were obtained and housed in our specific pathogen-free facility one week prior the experiments.

3.2.2.2 CS Exposure

To investigate changes in composition of fully developed lung microbiota upon CS exposure, the mice were whole-body exposed to mainstream CS (3R4F research cigarettes, Univ. of Kentucky, Lexington USA) over 60 min five days a week using smoking robot (SCIREQ InExpose System, Canada). Control group was exposed to room air. To adapt the mice to the smoke, for the first three days CS was generated using 6 cigarettes. Afterwards, 24 cigarettes were used to generate CS. The exposure was monitored by light scattering (MicroDust Pro, Casella, UK) of total particulate matter (TPM) concentrations in the chamber and cotinine levels in blood serum. Additionally, the body weight of the mice was controlled on a weekly basis. At days 7, 21, 37 and 56, mice were sacrificed 2 h after exposure (Figure 4).

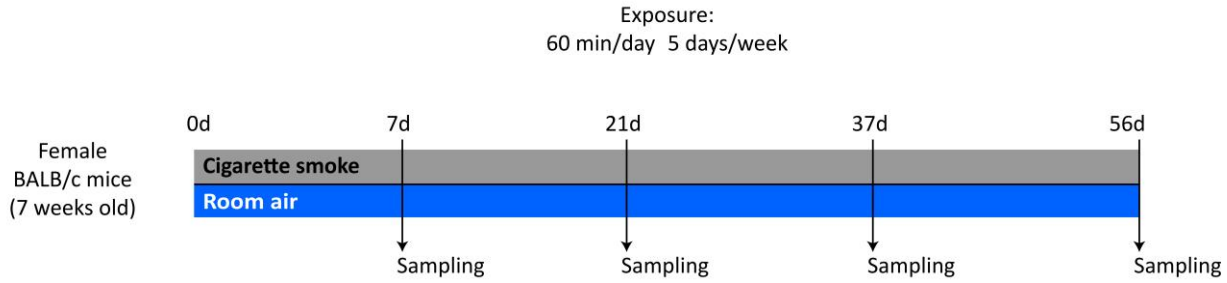


Figure 4. Experimental setup of CS exposure

3.2.2.3 Organ Harvesting and Preparation

In order to avoid cross-contamination during organ harvesting, the surgical tools were cleaned after each use. Tools were first rinsed in sterile water (Braun, Germany) and then sterilized in a hot bead sterilizer (Fine Science Tools, Cat No. 18000-45). Further, they were left to cool down as hot tools could kill microorganisms which would give false results.

To minimize discomfort to the animal, narcosis was administered intraperitoneally (10 µl/g of body weight; 13.8 mg/ml ketamine + 0.7 mg/ml xylazine in phosphate-buffered saline) before sacrificing the mice by cutting the abdominal vein. Mice were fixed on styrofoam panel and dampened with 70 % ethanol. Using scissors, the skin was cut from abdomen to neck and torn with forceps to expose thoracic cage. The dissected organs, including lungs, caeca, ears and tongues, were snap frozen in liquid nitrogen and later stored at -80 °C. Thereafter, lung and caecum samples were sent to the Helmholtz Centre Munich on dry ice to remain frozen and were immediately stored at -80 °C upon arrival.

3.2.2.4 Bronchoalveolar Lavage

The neck muscles were gently removed to expose the trachea. Then, a cannula was inserted and firmly fixed with a thread. A 1 ml syringe was attached and 800 µl of PBS were injected into the lungs. The fluid recovered from BAL was collected into 1.5 ml tubes and held on ice. After centrifugation at 1800 rpm at 4 °C for 10 min, the supernatant was collected and stored at - 80 °C, while the pellet was resuspended in 200 µl of PBS. Further, the cell suspension was mixed with trypan blue in a ratio of 1:1 to quantify the cell count and to measure the total number of cells using microscope with a haemocytometer. 100 µl of BAL fluid were deposited onto cytospin slides and centrifuged using a cytofuge (Cytospin 3, Shandon, UK) at 900 rpm at room temperature for 8 min. After drying overnight at room temperature, the cytospins were stained with Diff-Quik cell dye (Medions Diagnostics, France). 500 cells were counted by microscopy (BX40, Olympus, Germany) at a magnification of 40x and cells were differentiated into macrophages, neutrophils, lymphocytes and eosinophils based on morphological and staining characteristics.

3.2.2.5 Lymphocyte Analysis

At the timepoints 21, 37 and 56, the left lung lobe was used for flow cytometry analysis of different lymphocyte subtypes. After cutting, the lobe was minced with a scalpel and processed in digestion buffer (1 mg/ml collagenase (Sigma-Aldrich, USA) in 1xPBS) at 37 °C (water bath) for 45 min. After digestion, the samples were homogenized using a 10 ml syringe with a 20G needle (Braun, Germany) and filtered through a 70 µm cell strainer (Corning, USA). Further, erythrocytes in single cell suspensions were lysed by Gey'sche lysis buffer (10 mM KHCO3 (Merck, Germany), 155 mM NH_4Cl (Merck, Germany), 100 µM EDTA (Sigma-Aldrich, USA). All single cell suspensions were counted via a haemocytometer and adjusted to $2x10^7$ cells/ml. Single cell suspensions of lung samples were incubated with Fc Block (CD16/CD32) to prevent non-specific binding of antibodies. Afterwards, cells were stained with fluorochrome-conjugated antibodies against surface markers and intracellular markers (Table 7). After staining, cells were fixed with 4% paraformaldehyde before measurements. Flow cytometric analysis was conducted using a LSRII (BD Bioscience, USA), equipped with 405 nm, 488 nm and 633 nm lasers. FCS Express software version 6 (De Novo Software, USA)) was used for the analysis and gating strategies for B and T lymphocytes are presented within the Figure 5**.**

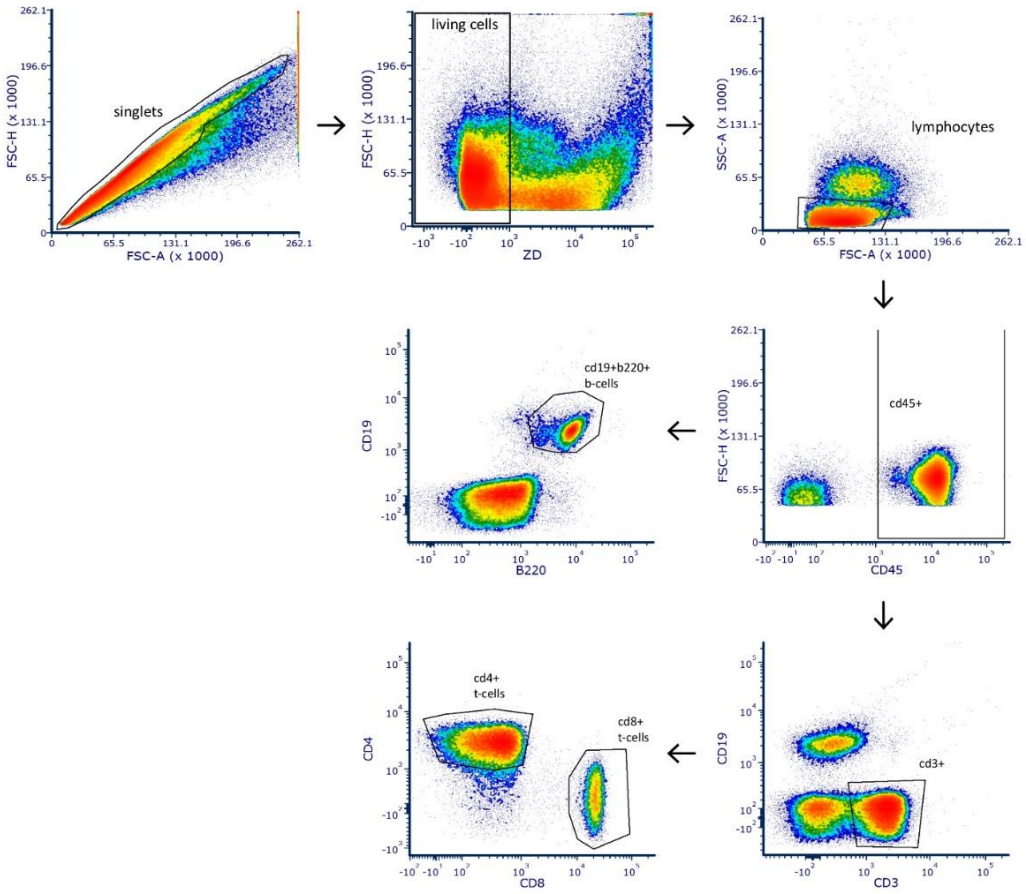


Figure 5. Gating strategy for flow cytometry analysis of B and T lymphocytes.
Forward scatter height (FSC-H), forward scatter area (FSC-A), side scatter area (SSC-A), haematopoietic cells (CD45+), B lymphocytes (CD19+B220+), T lymphocytes (CD3+), helper T lymphocytes (CD4+), cytotoxic T lymphocytes(CD8+).

3.2.2.6 Cotinine Levels Measurement

Rat/mouse cotinine enzyme-linked immunosorbent essay (ELISA) kit (Calbiotech, USA) was used for detection of nicotine's metabolite cotinine in the murine blood sera. All steps were performed at room temperature. The samples and serially diluted standards were tested in duplicates and firstly incubated with the enzyme conjugate in the dark for 1 h. After washing the plate 3 times with washing buffer and drying it via vigorous slapping on absorbent paper, substrate reagent was added. After 30 min of incubation in the dark, stop solution was added and the absorbance was read on multiplate reader (Tecan Sunrise, Tecan, Germany) at 450 nm. Results were calculated by interpolating data from the standard curve generated by serial diluted standard proteins with known concentrations.

3.2.2.7 Gene Expression

3.2.2.7.1 Tissue Processing for RNA Isolation

Approximately 5 ml PBS were instilled in the right ventricle to drain blood from the lungs. The whole tissue samples were put into a fresh 2 ml reaction tubes containing 700 µl Qiazol® (Qiagen, Germany) and stored at -80 °C before performing RNA extraction.

3.2.2.7.2 RNA Isolation

Using steel beads and the TissueLyser II (Qiagen, Germany), the lung tissue samples were homogenized for 2 min at room temperature with a frequency of 30 Hz. The samples were further treated according to the manufacturer's instructions using RNeasy Mini kit (Qiagen, Germany). At the final step, the ribonucleic acid (RNA) was eluted with 30 µl nuclease-free water and the concentration was measured with the NanoDrop1000 (Thermo Fisher Scientific, USA) at 260 nm and impurities at 280 nm. The isolated RNA samples were stored at - 80 °C. If necessary, purity of RNA was increased by precipitation.

3.2.2.7.3 RNA Precipitation

Mix of 3M sodium acetate (ThermoScientific, USA), glycogen (ThermoScientific, USA) and 100 % ethanol were added to samples and stored overnight at -80 °C. Thereafter, samples were centrifuged at 10000 rpm for 15 min and supernatant was discarded. After washing with 70 % ethanol, pellets were air-dried and dissolved in nuclease-free water.

3.2.2.7.4 RNA Quality Control

Quality of RNA was checked on 1.2 % agarose gel electrophoresis. Therefore, 1 µg RNA was mixed with loading buffer (containing formamide and SYBR Green II in TBE buffer) and denatured at 65 °C for 15 min. Afterwards, the samples were separated in agarose gel in TBE buffer at 60 V for 1.5 h. Finally, the gel was analyzed on ChemiDoc Touch imaging system (Biorad, USA) for the 28S and 18S rRNA bands identified by the marker (BenchTop 1kb DNA Ladder, Promega, USA). Only the samples with 2:1 ratio of 28S:18S bands were considered as intact and were used for further analysis.

3.2.2.7.5 RNA Reverse Transcription

DNA was synthetized from the RNA template using QuantiNova Reverse Transcription kit (Qiagen, Germany). According to the manufacturer’s protocol, 500 ng RNA was adjusted to 12 µl in nuclease-free water. Prior the transcription, samples were incubated in thermocycler (TProfesional TRIO, Biometra, Germany) with gDNA wipeout buffer at 45 °C for 2 min to ensure that residual genomic RNA was removed. Then, together with the master mix and reverse

transcriptase, template RNA samples were incubated under conditions shown in the Table 15. Complementary DNA (cDNA) was diluted 10 times for the further use in PCR reactions.

Step	Time [min]	Temperature [°C]
Annealing	3:00	25
Reverse-transcription	10:00	45
Inactivation of reaction	5:00	85

Table 15. Thermocycler protocol for reverse transcription

3.2.2.7.6 qRT-PCR

Quantitative real-time PCR (qRT-PCR) was performed on LightCyler 480 (La Roche, Switzerland) via mixing 7.4 µl SYBRGreen I master mix (Roche Life Science, Germany) and primer mix with 2.4 µl cDNA samples following the protocol shown in the Table 16. All samples were run in triplicates, and NTCs were included for each target gene. To compensate variations occurring during the PCR run, the expression of all target genes was normalized to endogenous reference genes, such as TATA-Box binding protein (TBP) and Hypoxanthine phosphoribosyltransferase (HPRT). The $2^{-\Delta\Delta Ct}$ values were used to indicate the fold change in gene expression relative to the air control samples. Target genes tested in this experiment encode following proteins: Cytochrome P450, family 1, subfamily A, polypeptide 1 (Cyp1A1), Aryl hydrocarbon receptor (AhR), Tight junction protein 1 (TJP1), Occludin (OCLN) and Matrix metalloproteinase 12 (MMP12). Primer sequences used in this experiment are described in the Table 5.

Step	Cycle	Time [min]	Temperature [°C]
Initial denaturation	1	10:00	95
Amplification	45	00:10	95
		00:15	60
		00:10	72
		00:01	78
Melting curve	1	00:05	95
		01:00	60
	0.11 °C/sec	continuous	99
Cooling	1	00:10	40

Table 16. LightCycler protocol for qRT-PCR

3.2.2.8 Analysis of MMP12 Protein Levels

BAL supernatants were used for enzyme-linked immunosorbent assay (ELISA) to quantify levels of MMP12 protein. Mouse MMP12 SimpleStep ELISA kit (Abcam, United Kingdom) was used according to the manufacturer's instructions. First, all samples and standards were added into precoated ready-to-use 96-wel plate. Next, antibody cocktail containing both capture and

detector antibody was added and incubated on a plate shaker at 400 rpm at room temperature for 1 hour. After washing the plate 3 times with washing buffer PT, TMB solution was added for color development and incubated for 10 minutes under previously described conditions. In the final step, reaction was interrupted by adding stop solution to the wells and optical density was read at 450 nm using multiplate reader (Tecan Sunrise, Tecan, Germany).

3.2.2.9 Analysis of Inflammatory Mediators

Cytokine and chemokine levels in BAL supernatants were measured via Meso Scale Diagnostics U-Plex Platform (MSD, USA). This method is based on a sandwich ELISA and allows simultaneous quantification of multiple analytes in each well. All steps were performed at room temperature following the manufacturer's protocol. Prior the experiment, coating of U-PLEX plate was necessary. First, 200 µl of each biotinylated capture antibody were coupled with 300 µl of unique linker. After 30 min of incubation, the linking process was stopped by the addition of 200 µl of the stop solution and further incubation of 30 min. 600 µl of each U-PLEX-coupled antibody solution were combined and 50 µl of the pool was added to each well. The sealed plate was shaken for 1 h and afterwards washed tree times with 150 µl PBS-Tween. Measured cytokines and chemokines (Table 8) were selected to reflect proinflammatory mediators implicated in airway diseases. All values less than the detectable range were assigned to zero.

3.2.2.10 Lung Histology

The lungs from the mice sacrificed at day 56 were instilled via the tracheas under pressure in 4 % formaldehyde for 20 min and afterwards fixed in 4 % formaldehyde (Roti HistoFix, Carl Roth, Germany) overnight at room temperature. The samples were then embedded in 2 % Agar-Agar and sliced using tissue slicer. After dehydration in the spin tissue processor (ThermoFisher Scientific, USA) by immersion in increasingly concentrated ethanol, the slices were embedded in paraffin using an embedding station (HistoStar, ThermoFisher Scientific). Paraffin-embedded tissue samples were cut on Microtom HM340E (ThermoFisher Scientific, USA) into 2 µm sections and dried on SuperFrost®Plus slides (Langenbrinck, Germany) at 37 °C in an incubator overnight prior to the staining with Periodic Acid-Schiff (PAS) dye following manufacturer's kit (Carl Roth, Germany) instructions[116]. Periodic acid oxidizes polysaccharides into aldehyde structures, giving the tissue a blue color, which is thereon dyed red after Schiff's reagents binds to these structures. Structural changes of lung sections were analyzed under a microscope (BX40, Olympus, Germany). Additionally, morphometric analysis of the alveolar space was performed using

Computer Assisted Stereology Toolbox (newCAST, Visiopharm, Denmark) and measuring random sampling test points on microscopic images.

3.2.2.11 Microbiome Analysis

Upon harvesting, all organs for microbiome analysis were first snap frozen in liquid nitrogen and afterwards stored at -80 °C before sending them to the Helmholtz Center in Munich. DNA extraction, 16S rRNA gene library preparation, sequencing using MiSeq Illumina system and further bioinformatic analysis were performed by Dr. Stefan Pfeiffer at the **Research Unit for Comparative Microbiome Analysis**.

3.2.2.11.1 DNA Extraction

The library preparation for sequencing and the sequencing run itself were performed separately for lung and caecum samples. The frozen caeca were weighed and homogenized in 400 µl double-distilled water using a Precellys24 homogenizer (Bertin Technologies, France) at 5.5 m/sec for 30 seconds. Following the manufacturer's recommendations, DNA was extracted out of 100 µl of the homogenized tissue via QIAamp PowerFecal DNA kit (Qiagen, Netherlands). All caecum samples were extracted with a NEC included. Frozen lung samples were likewise weighted and homogenized. 180µl lysozyme solution (20 mg/ml) (Roche Diagnostics, Germany) was added to each 200 µl tissue homogenate and the mixture was incubated at 37 °C for 30 minutes under slow agitation. Further steps were conducted according to the manufacturer's protocol for Gram-positive bacteria of the PureLink™ Genomic DNA Mini kit (ThermoFisher Scientific, USA). All lung samples were extracted with one NEC included.

3.2.2.11.1.1 Amplicon PCR

All steps and material for the amplicon PCR were the same as described for the human samples except operating conditions described in the Tables 17 and 18.

Step	Cycle	Time [sec]	Temperature [°C]
Initial denaturation	1	30	98
Amplification	33	10	98
		30	58
		30	72
Final elongation	1	300	72

Table 17. Amplicon PCR protocol for murine lung samples

Step	Cycle	Time [sec]	Temperature [°C]
Initial denaturation	1	30	98
Denaturation		10	98
Annealing	20	30	58
Elongation		30	72
Final elongation	1	300	72

Table 18. Amplicon PCR protocol for murine caecum samples

In addition to the samples, NEC and PCR NTCs were tested. Further, one NEC and NTC control that was spiked with 10 ng of the commercial bacterial mock community HM-783D (BEI Resources, NIAID, NIH, USA) were included with each lung samples. This step was performed to estimate the contamination risks and chimeric biases that could occur due to the high number of amplification cycles. Additionally, success of each of the PCRs was verified with agarose gel electrophoresis and pooled triplicates were purified with Agencourt AMPure XP paramagnetic beads (Beckman Coulter, Brea, USA). Furthermore, purified PCR products were checked for dimers and were quantified using the DNF-473 Standard Sensitivity NGS Fragment Analysis kit on the Fragment Analyzer (Advanced Analytical, USA).

3.2.2.11.1.2 Index PCR

As well as in the previous step, the index PCR protocol was same as described for human samples with modifications of operating conditions (Table 19).

Step	Cycle	Time [sec]	Temperature [°C]
Initial denaturation	1	30	98
Denaturation		10	98
Annealing	8	30	55
Elongation		30	72
Final elongation	1	300	72

Table 19. Index PCR program for mouse samples

Verification of PCR, purification and quantification of PCR products was conducted as described for the amplicon PCR. PCR products were diluted to 4 nmol/L and pooled. For sequencing, 8 pM of DNA spiked with 20 % PhiX were loaded and the paired-end sequencing was performed on a MiSeq® System (Illumina, Inc., CA, USA) using the MiSeq® Reagent kit v3 (600 cycle) for paired-end sequencing according to the manufacturer's guidelines.

3.2.2.11.2 Sequencing

Amplicons were processed through the same FASPA protocol described for human samples.

3.2.2.12 Statistical Analysis

3.2.2.12.1 Body Weight, Gene Expression and Inflammatory Responses

Data were tested for normal distribution with D'Agostino-Person omnibus normality test. Normally distributed data were analyzed by unpaired student's t test (t test) or 1way analysis of variance (ANOVA) with Tukey post-hoc test for more than two groups. For data being compared to two factors, 2way ANOVA with Tukey's or Sidak's post-hoc test was used. Results are expressed as mean ± standard deviation (SD). Data were analyzed using Graph Pad Prism version 6.0 (GraphPad Software, USA) and statistical significance was accepted with p-values <0.05. In all gene expression analysis, fold change from the air group was set to 1.

3.2.2.12.2 Bioinformatic Data Analysis

The output files generated by sequence processing were imported into R[94] studio to carry out the statistical analysis and data visualization of microbial diversity and composition. First, the data was transformed into a phyloseq[95] object, a format suitable for the *Rhea* pipeline[96]. As some steps of the library preparation differed for the human samples, murine lungs and caecum, data analysis of each sample type was executed individually using modified Rhea pipeline. In the first step, normalization of read counts was done to control the sufficiency of sequencing depth and to remove the most undersampled specimen. Further, alpha diversity was calculated based on observed species richness, evenness as well as Simpson and Shannon effective diversities. Existing differences between categories were depicted with box plots. Generalized UniFrac distances were used to calculate beta diversity and visualization in two-dimensional space was performed by metric or nonmetric multi-dimensional scaling (MDS; nMDS) plot. To test if the separation of sample groups is significant, differences of categorical metadata between all samples was calculated using a permutational multivariate analysis of variance (PERMANOVA with 999 permutations). On the taxa and zOTU levels, pairwise categorical differences were calculated using the Wilcoxon-Rank-Sum-Test with FDH correction. For Pearson correlation (FDH corrected) analysis, the quantitative relative abundance data of zOTUs and taxa was centre log-transformed to remove compositional constraints.

4 Results

4.1 Human Study in Smokers, Never-smokers and Ex-smokers

The data obtained in this study were processed and analyzed **by Dr. Stefan Pfeiffer at the Helmholtz Center in Munich.**

	Smokers				Never-smokers	Ex-smokers
	Total	**Long-term**	**Short-term**	**Mild**	**Total**	**Total**
Number, n	36	15	9	12	16	6
Male, n (%)	23 (63.9)	9 (60)	5 (55.6)	9 (75)	6 (37.5)	5 (83.3)
Age in years, median (IQR)	27.5 (23-34)	33 (28-49.5)	21 (19-28)	26 (22.8-26)	26 (24-39)	59 (49.5-63.5)
Body mass index [kg/m^2], median (IQR)	24.2 (21.5-26.5)	26.7 (24.7-29.5)	21.5 (21.2-23.7)	24.7 (22.5-25.4)	24.3 (20.6-26.2)	24.9 (23.9-26.3)
Current cigarettes per day, median (IQR)	14.5 (8-19)	16.3 (13.5-20)	14.5 (9.5-17.9)	4 (2-7)	-	-
Maximum cigarettes per day, median (IQR)	20 (16-30)	30 (22-37.5)	20 (17-30)	10 (8-15)	-	30 (7-30)
Packyears, median (IQR)	7 (5-17)	18.5 (14.3-31.5)	5 (5-6.3)	1 (0.8-4)	-	4.5 (3-25)
Years of smoking, median (IQR)	10 (5-18)	17 (13-21.5)	5 (5-8)	8 (0.8-11.5)	-	14.5 (4.5-24)
Years since smoking cessation, median (IQR)	-	-	-	-	-	19 (18-23)

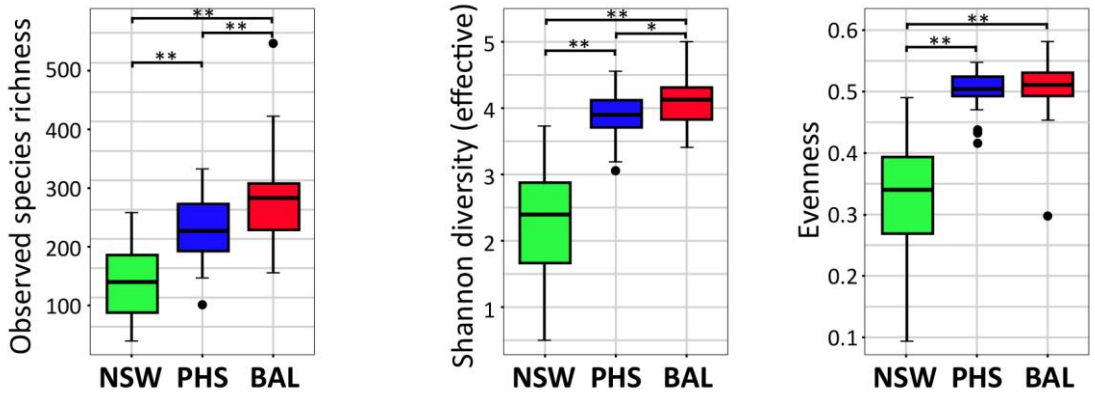
Data are presented as n, n (%) and median (interquartile range).

Table 20. Participants' characteristics stratified by smoking history.

4.1.1 Microbiota of Human Respiratory Tract

After quality filtering, denoising and removal of short sequences, sufficient sequencing depth and removal of the most under-sampled specimen were obtained via rarefaction. All reads were normalized to 10000/sample. To uncover bacterial compositions of the different sampling spots, α-diversity metrics were measured. Shannon diversity and observed species richness of PHS were significantly higher than for NSW and significantly lower than for BAL. Evenness was comparably low in NSW and it showed strong inter-individual differences (Figure 6a). Furthermore, β-diversity analysis was used to measure dissimilarities between the sampling spots. nMDS plot, based on generalized UniFrac distances, hence showed a significant influence of the body site on the microbial community composition (Figure 6b). Although the PHS and BAL samples clustered close to each other in comparison to the NSW samples, pairwise comparison showed that communities at these two sites differed significantly with respect to their bacterial community composition (p=0.006). Calculation of constrained distance-based redundancy analysis (db-RDA) with 999 permutations estimated that the sampling spot explains 18.9 % of the total variance.

A



B

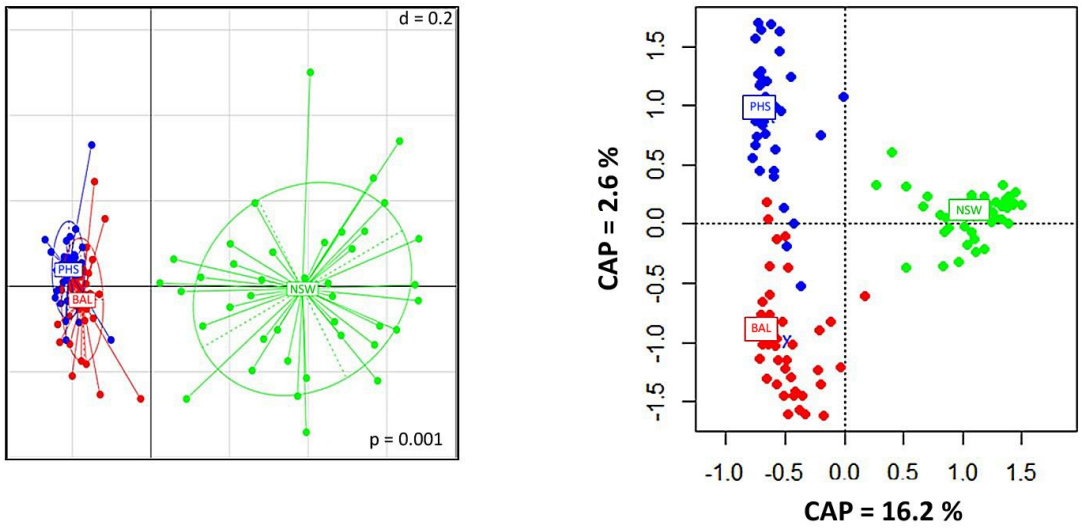


Figure 6. Diversity and compositional differences of microbial communities among different body sites.
A) Box plot of α-diversity metrics showing observed species richness, Shannon diversity and Evenness of NSW, PHS and BAL samples. Only samples that were available from all three body sites were included for comparison (n=39). Pairwise differences are calculated based on the non-parametric Wilcoxon-Rank-Sum test and corrected for multiple testing with the Benjamin-Hochberg method to decrease the False Discovery Rate (FDR). *($p<0.05$) or **($p<0.005$). B) β-diversity metrics. Left: unconstrained nMDS plot; Right: constrained (right) db-RDA plot of generalized UniFrac distances of NSW (green), PHS (blue) and BAL (blue) samples. Constrained dispersion is maximized by sampling site. Constrained analysis of principal coordinates (CAP) axes displays the amount of the total variation that is explained by the sampling site. PERMANOVA (999 perturbations) based levels of significance are displayed on the lower-right hand edge of the plots.

Additionally, highly abundant genera in PHS and BAL samples were *Prevotella*, *Veillonella,* *Streptococcus* and *Actinomyces* (Figure 7). These four genera were also present in the NSW samples, but there the most dominant were *Corynebacterium* and *Staphylococcus* with 47.1% of the obtained bacterial reads (Supplementary tables 21 and 22).

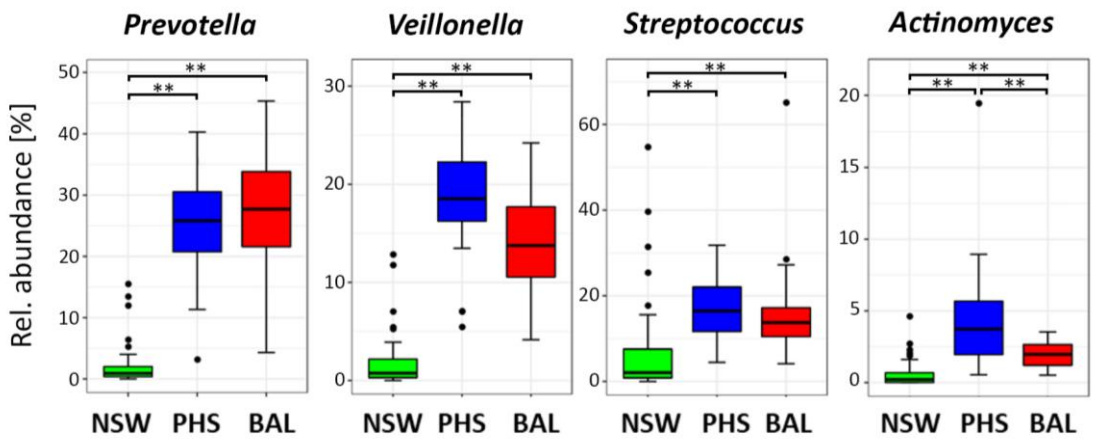


Figure 7. Comparison of the relative abundances of certain taxa in NSW, PHS and BAL samples.
Box plots median relative abundances of the 4 genera which were found in all samples of the three body sites. (n=39). Pairwise Wilcoxon Rank Sum test (*p < 0.05 or **p<0.005, FDR corrected).

4.1.2 Influence of CS on the Nasal Microbiota

When assessing smoking parameters, data from NSW samples showed that the relative abundance of *Corynebacterium* and *Dolosigranulum* correlated negatively with the maximum number of cigarettes smoked per day, smoking years and packyears. Also, nasal samples showed constant and exclusive presence of *Propionibacterium* sp., which correlated positively with smoking years. Further, the relative abundance of *Prevotella* correlated positively with the maximum number of cigarettes used per day.

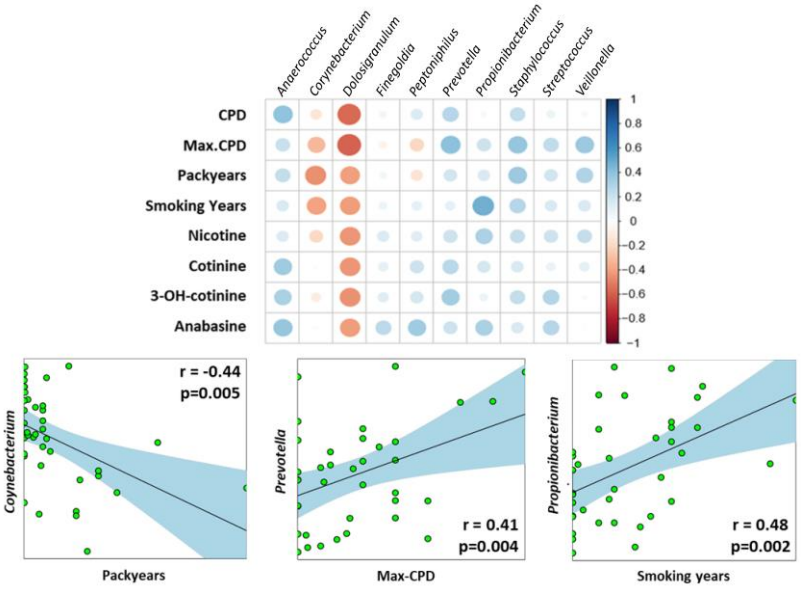


Figure 8. Correlation between certain taxa in NSW with smoking characteristics.
Upper panel: Correlation array of smoking factors with NSW (n=46) genera across smoking parameters and transformed sequencing data (centered log-ratio transformation). The size of the circles shows the significance of the correlation (bigger circle corresponds to a lower p-value). The color of the circle indicates the Pearson correlation between 1 (blue, positive correlation) and -1 (red, negative correlation). CPD = cigarettes per day, Max-CPD = maximum cigarettes per day; Lower panel: Pearson correlation of bacterial taxa with smoking related parameters. Each dot represents an individual sample. The line corresponds to the fitted linear model. Confidence intervals are shown by the light blue area. The correlation coefficient (upper row), the FDR-corrected p-value (middle row) and the number of observations (lower row) are shown in the corners of the plots.

Another taxa that was positively associated with these three smoking factors was *Staphylococcus*, which was mainly *S. epidermidis.* While this strain was present in all NSW samples, it was significantly more abundant in those obtained from long-term and short-term smokers compared with never- and mild smokers (Figure 9a). Additionally, many samples of never- and mild smokers had a high relative abundance of Carnobacteriaceae, in particular *Dolosigranulum pigrum*, which correlated negatively with all smoking related parameters (Figure 8a and Figure 9b).

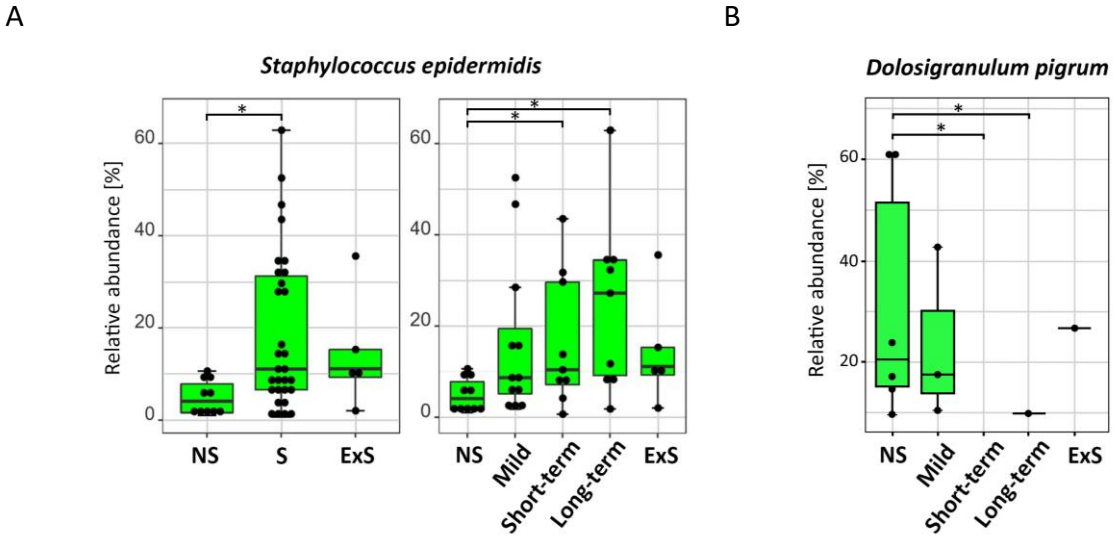


Figure 9. Relative abundances of certain taxa in NSW upon CS.
A) Left: Box plots of median relative abundances of S. epidermidis among different groups. never-smokers (NS) n=10, smokers (S) n=30 and ex-smokers (ExS): n=6; Right: Box plot includes groups based on smoking history. B) Box plot of median relative abundance of D. pigrum among smoking groups based on smoking status and on smoking history. NS: n=10, heavy long-term smokers (long-term: n=9), heavy smokers (short-term: n=9), mild/ occasional smokers (mild: n=12) and ExS: n=6. Significant differences in relative abundance were determined using pairwise Wilcoxon Rank Sum test (*$p < 0.05$, FDR corrected).

4.1.3 Influence of CS on the Oropharyngeal Microbiota

Results showed that Firmicutes phylum was significantly higher in oropharyngeal samples of smokers compared with never-smokers (Figure 10). Yet, its most abundant genera *Streptococcus* and *Veillonella* were not associated neither with smoking intensity nor duration (Figure 11). Furthermore, genus *Centipeda* correlated negatively with short-term smoking parameters, as did *Leptotrichia* from Fusobacteria phylum. Among Actinobacteria, the significantly higher abundance in smokers and ex-smokers compared with never-smokers was related to the higher abundance of *Atopobium* (Figure 12a). This genus correlated positively with cotinine and numbers of daily cigarettes, but not with long-term smoking (Figure 11). In contrast to Firmicutes and Actinobacteria, the relative abundance of Proteobacteria was lower in the oropharynx from smokers compared with ex- and never-smokers (Figures 10 and 12b). On the genus level,

although *Neisseria* was detected only in few samples from smokers (31 % of smokers, 71 % of never-smokers and 90 % of ex-smokers), the mean relative abundance was lowest in ex-smokers (Figure 12b). This genus correlated negatively with the maximum number of cigarettes smoked per day (Figure 6). The abundance of Bacteroidetes did not differ in PHS samples of smokers and never-smokers. Also, its genus Prevotella did not correlate notably with any smoking parameters.

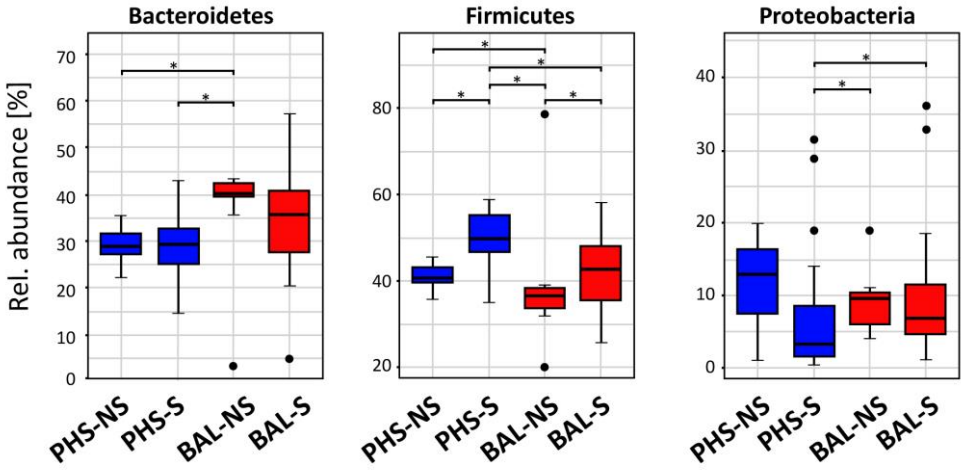


Figure 10. Comparison of relative abundance of the most common taxa in BAL and PHS samples from never-smokers and smokers.
BAL samples of never-smokers (BAL-NS), PHS samples of never-smokers (PHS-NS), BAL samples of smokers (BAL-S) and PHS samples of smokers (PHS-S). Pairwise Wilcoxon Rank Sum test (*p < 0.05, Benjamin-Hochberg corrected).

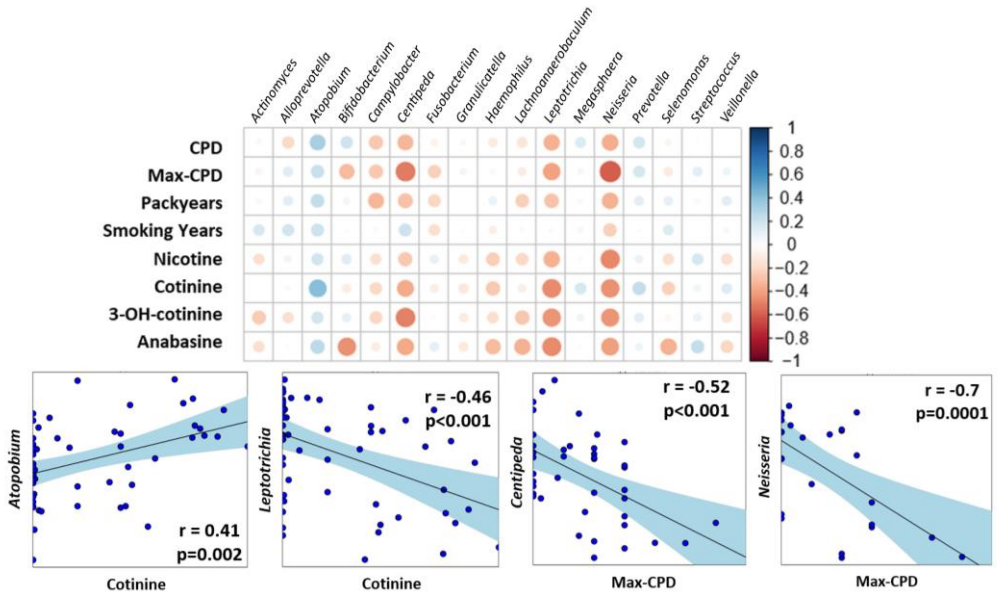
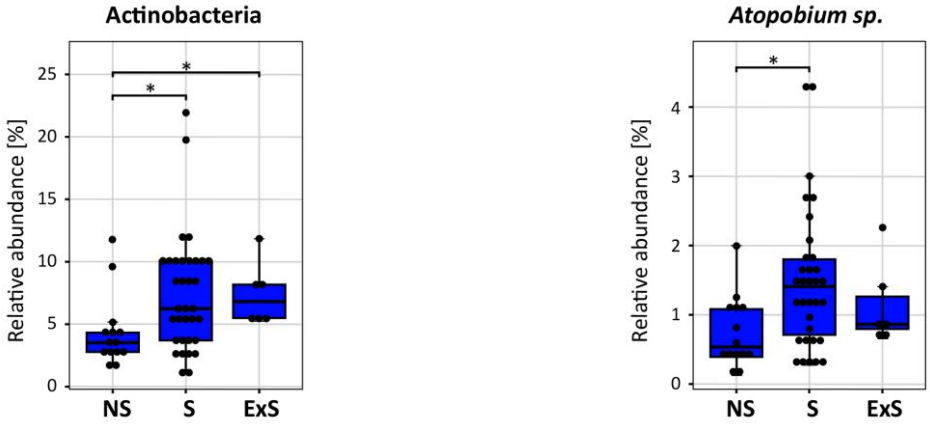


Figure 11. Correlation between certain taxa in PHS with smoking characteristics.
Upper panel: correlation array of smoking factors with PHS (n=54) genera across smoking parameters and transformed sequencing data (centered log-ratio transformation) to remove constrains introduced by the compositionality of the data. The size of the circles shows the significance of the correlation (bigger circle corresponds to a lower p-value). The color of the circle indicates the Pearson correlation between 1 (blue, positive correlation) and -1 (red, negative correlation). CPD = cigarettes per day, Max-CPD = maximum cigarettes per day; Lower panel: Pearson correlation of bacterial taxa with smoking related parameters. Each dot represents an individual sample. The line corresponds to the fitted linear model. Confidence intervals are shown by the light blue area. The correlation coefficient (upper row), the FDR-corrected p-value (middle row) and the number of observations (lower row) are shown in the corners of the plots.

A



B

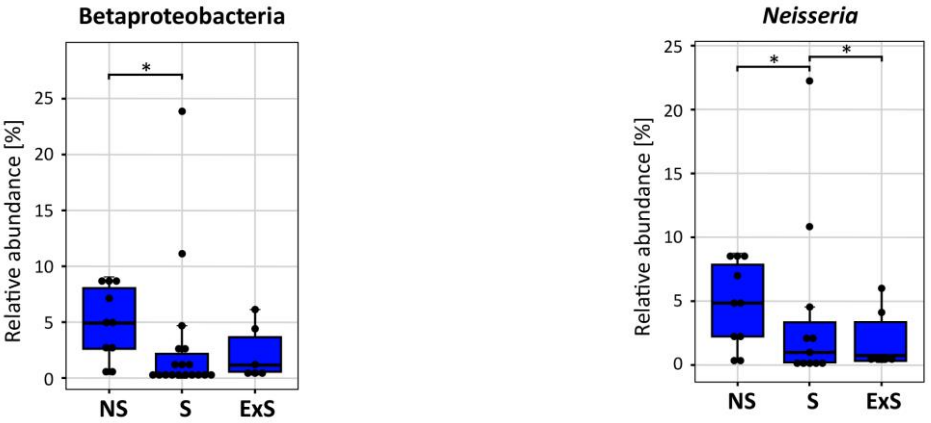


Figure 12. Relative abundances of certain taxa in PHS samples upon CS exposure.
A) Actinobacteria (Left: Phylum level; Right: Species level). B) Proteobacteria (Left: Class level; Right: Species level). Box plots of median relative abundances of ZOTUs related to different taxa among smoking groups. NS (n=14), S (n=34) and ExS (n=6), of paired samples. Significant differences in relative abundance were determined using pairwise Wilcoxon Rank-Sum test (*$p < 0.05$, FDR corrected).

4.1.4 Influence of CS on the Lung Microbiota

Like in PHS microbiota, relative abundance of Firmicutes phyla notably increased in BAL samples upon smoking (Figure 5). However, its genus *Veillonella* negatively correlated with packyears. Decrease in relative abundances of Proteobacteria was observed as well in BAL samples taken from smokers, however it was not significant (Figure 5). Additionally, relative abundance of its genus *Campylobacter* differed among BAL samples of smokers and never-smokers (Figure 13). Several β- and γ- proteobacterial species were specific for the microbial community composition of BAL samples (Supplementary table 23) and were further positively correlated with all smoking related parameters. Specifically, *Acinetobacter berezinae, A. johnsonii, Cupriavidus metallidurans*, *Serratia marcescens* and *Stenotrophomonas maltophilia* were strongly influenced both by long-term smoking and smoking intensity (Figure 14). Aside from Proteobacteria, the

relative abundance of Bacteroidetes also decreased in smokers, but again without a significance (Figure 10). In line with that, the genus *Prevotella* negatively correlated with packyears in BAL samples (Figure 14). In contrast, BAL-specific Bacteroidetes correlated positively with long-term smoking.

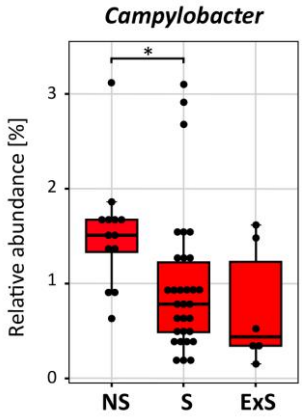


Figure 13.Relative abundance on Genus level in BAL samples.
Box plots of median relative abundances of ZOTUs related to different taxa among smoking groups. NS (n=13), S (n=31), ExS (n=6), of paired samples. Significant differences in relative abundance were determined using pairwise Wilcoxon Rank Sum test (*p < 0.05, FDR corrected).

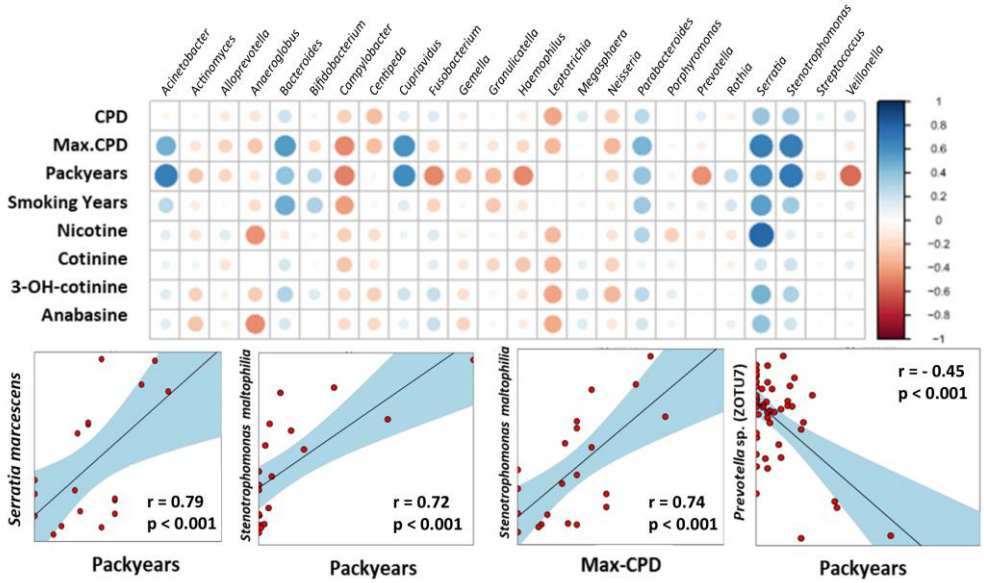


Figure 14. Correlation between certain taxa in BAL with smoking characteristics.
Upper panel: Correlation array of smoking factors with BAL (n=52) genera across smoking parameters and transformed sequencing data (centered log-ratio transformation) to remove constrains introduced by the compositionality of the data. The size of the circles shows the significance of the correlation (bigger circle corresponds to a lower p-value). The color of the circle indicates the Pearson correlation between 1 (blue, positive correlation) and -1 (red, negative correlation). CPD = cigarettes per day, Max-CPD = maximum cigarettes per day; Lower panel: Pearson correlation of bacterial taxa with smoking related parameters. Dots represent samples. The line corresponds to the fitted linear model. Confidence intervals of the model are shown by the light blue area. The correlation coefficient (upper row), the FDR-corrected p-value (middle row) and the number of observations (lower row) are shown in the corners of the plots.

4.2 Murine Smoke Model

Since human studies do not allow testing whole lung tissue, a murine smoke model was established in order to complement insight into changes of the microbiome in lungs.

4.2.1 Influence of CS Exposure on Body Weight of Mice

Smoke-exposed mice are known to have a lower weight gain which can reflect the intensity of the smoke exposure[117]. Therefore, mice were weighted once per week over 56 days. From day 14, a significant decline in the body weight was observed in the CS-exposed mice relative to the air control group. This divergence continued until the end of the experiment (Figure 15). While the mean weight gain of the air-exposed group was 3.41 g (18.6 %), the CS-exposed group gained on average 1.47 g (8.34 %) over 56 days (Figure 16). Aside from this effect, no other visual differences e.g. behavioral anomalies and fur shedding were detected.

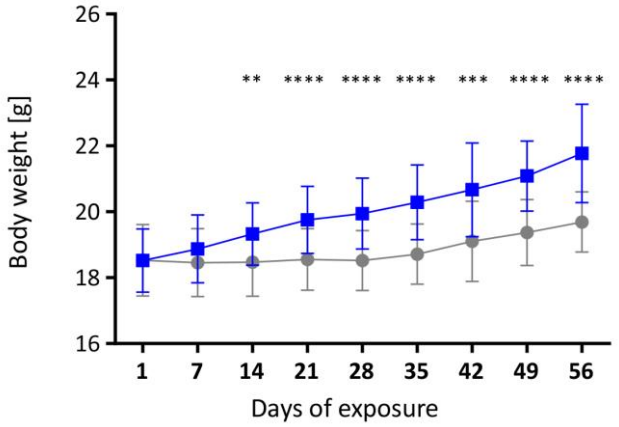


Figure 15. Total body weight of BALB/c mice.
Data are shown as mean ± SD (n (d1-7) = 78 animals/group, n (d7-21) = 60 animals/group, n (d21-35) = 40 animals/group, n (d35-56) = 22 animals/group). (■ air, ● CS). Two-way ANOVA with Sidak's multiple comparison test. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

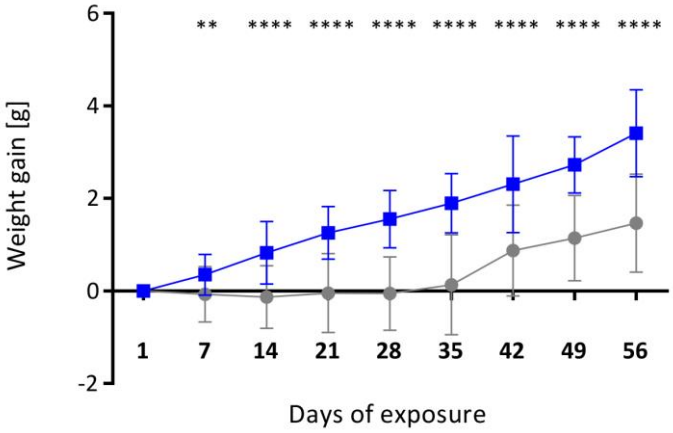


Figure 16. Body weight gain of BALB/c mice.
Data are shown as mean ± SD (n (d1-7) = 78 animals/group, n (d7-21) = 60 animals/group, n (d21-35) = 40 animals/group, n (d35-56) = 22 animals/group). (■ air, ● CS). Two-way ANOVA with Sidak's multiple comparison test. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

4.2.2 Cotinine Serum Levels in CS-exposed Mice

In order to monitor exposure to CS more objectively than by body weight gain, concentrations of nicotine's more stable metabolite cotinine were measured in murine blood sera. As a result, cotinine was highly present only in the samples of CS-mice. Of note, the protein levels were heterogeneous among different CS-exposed groups and differed notably between different time points (Figure 17).

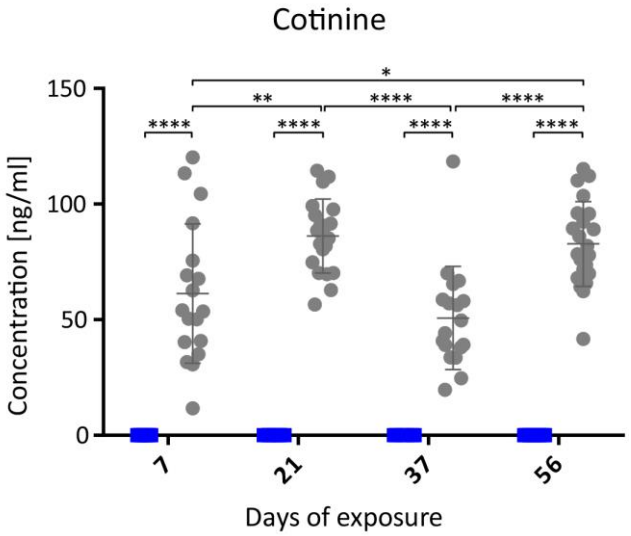


Figure 17. Cotinine concentration in blood serum.
Each data point represents an individual animal; data are shown as mean ± SD (n (air) = 9/group, n (smoke) = 18 /group). (■ air, ● CS). Two-way ANOVA with Tukey's multiple comparison test. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

4.2.3 Gene Expression in the Lungs of CS-exposed Mice

In humans, CS activates the expression of the cytochrome P450 gene Cyp1a1 in bronchial epithelial cells. This enzyme is regulated by AhR and metabolizes polycyclic aromatic hydrocarbons present in CS. Cyp1A1 is thus an important regulator in xenobiotic metabolism. To further objectify CS exposure, its gene expression in the lung homogenates was therefore analyzed by qPCR. While in the air-exposed groups this gene was not upregulated, a notably high expression was observed in the lungs of CS-exposed mice (Figure 18). The levels increased from day 7 until day 21 and afterwards gradually declined to day 56.

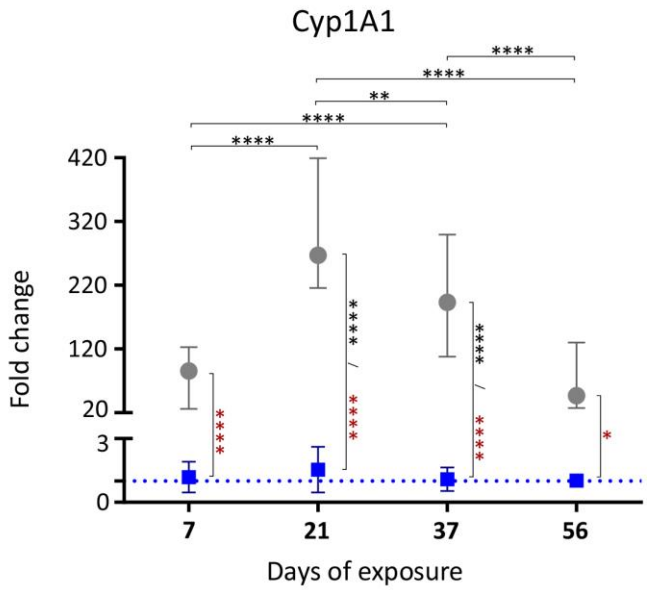


Figure 18. Fold change in Cyp1A1 gene expression in mouse lungs.
Data are shown as mean ± SD (n = 5-9 animals/group). (■ air, ● CS). Two-way ANOVA with Tukey's multiple comparison test (black stars). Unpaired t-test (red stars). Blue dotted line represents fold change of 1. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

Since Cyp1a1 was upregulated in CS exposed mice, expression of AhR gene was tested to clarify whether changes observed in Cyp1A1 levels originated from the protein involved in its activation. However, CS exposure did not cause any significant changes in the levels of this gene (Figure 19).

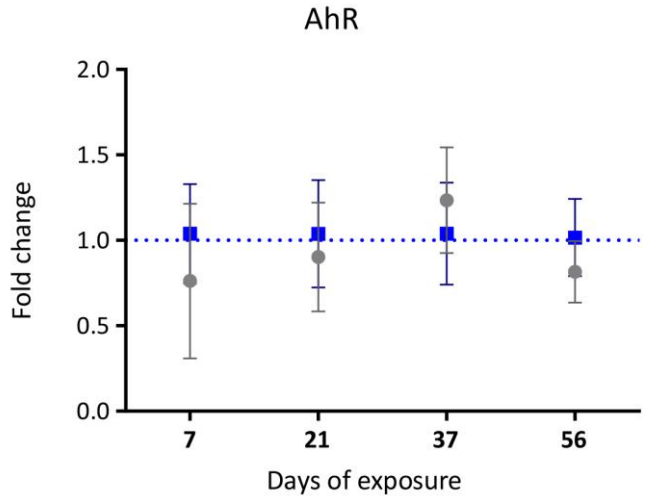


Figure 19. Fold change in AhR gene expression in mouse lungs.
Data are shown as mean ± SD (n = 5-9 animals/group). (■ air, ● CS). Two-way ANOVA with Tukey's multiple comparison test. Blue dotted line represents fold change of 1. No significant differences were observed.

To test whether CS exposure affected lung epithelial integrity, gene expression for tight junction molecules was analyzed. Despite lower expression of OCLN in the CS-exposed group than in the air control, this decline did not reach statistical significance. Similarly, the expression of TJP1 gene

was similar between both groups at all timepoints indicating that the epithelial barriers were still intact (Figure 20).

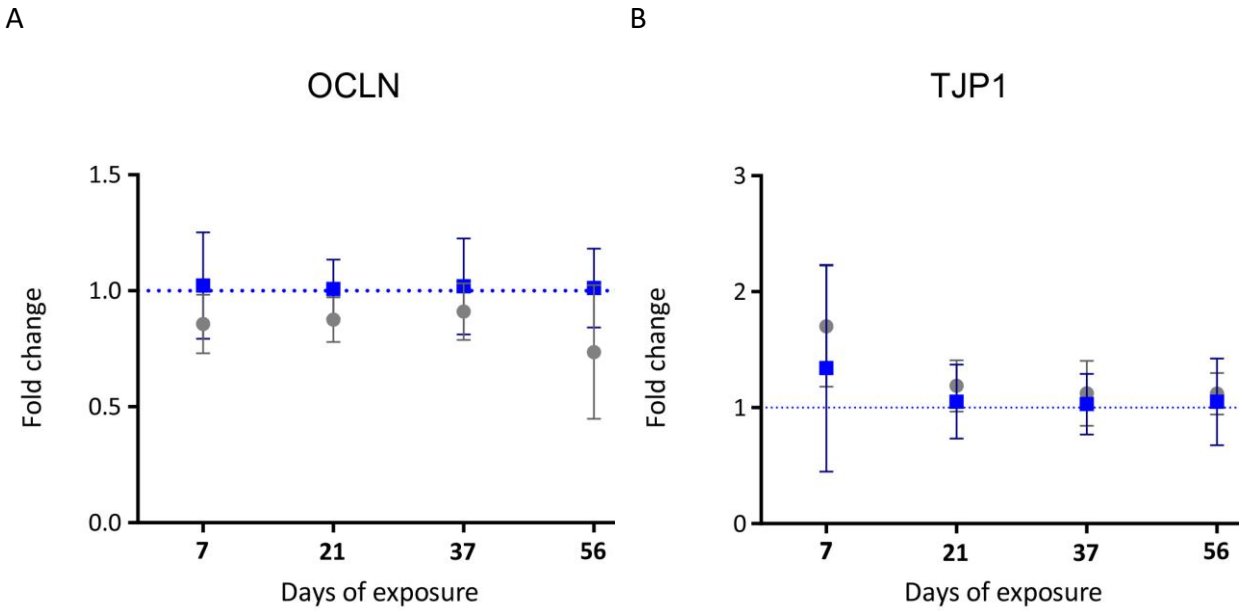


Figure 20. Fold change in the expression of tight junction genes in mouse lungs.
A) OCLN; B) TJP1. Data are shown as mean ± SD (n = 5-9 animals/group). (■ air, ● CS). Two-way ANOVA with Tukey's multiple comparison test. Blue dotted line represents fold change of 1. No significant differences were observed.

MMP12 is involved in the homeostasis of ECM. To assess if CS could affect the turnover of ECM, MMP12 gene expression in the lungs was quantified by qPCR. Starting with day 21, CS exposure caused a gradual increase of the gene expression in the lungs with remarkably higher values at the latest time point (Figure 21).

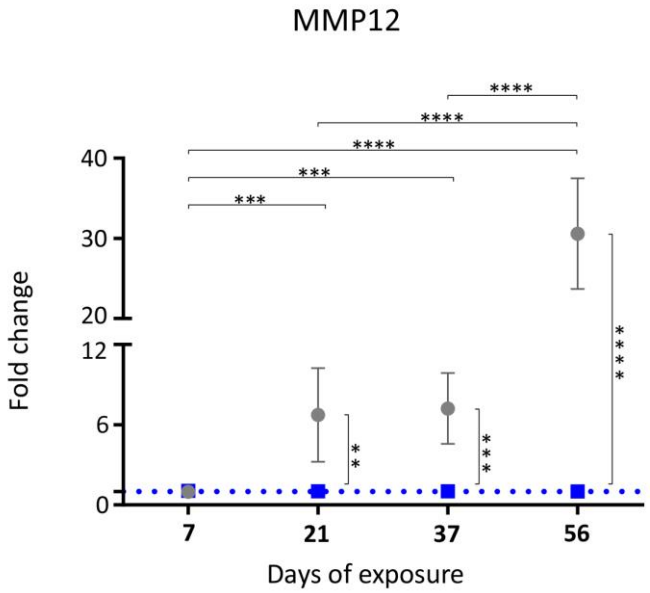


Figure 21. Fold change in MMP12 gene expression in mouse lungs.
Data are shown as mean ± SD (n = 5-9 animals/group). (■ air, ● CS). Two-way ANOVA with Tukey's multiple comparison test. Blue dotted line represents fold change of 1. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

4.2.4 Cell Count in BAL Fluid of CS-exposed Mice

Inflammatory cell infiltration in the lungs upon CS exposure was analyzed using standard BAL protocol. Total number of cells (Figure 22), macrophage and eosinophil number (Figures 23a and 23b) in BAL fluid did not differ notably between air- and CS-exposed groups through the complete experiment. Yet, the level of neutrophils and lymphocytes started to increase significantly in the samples of CS-exposed group after 37 or 56 days, respectively (Figure 23c and 23d).

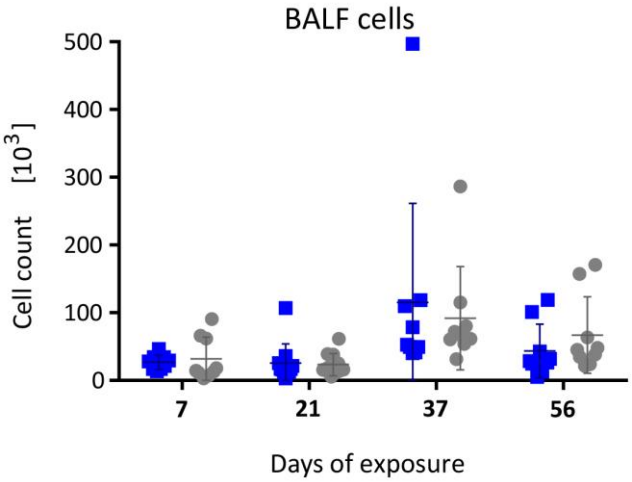
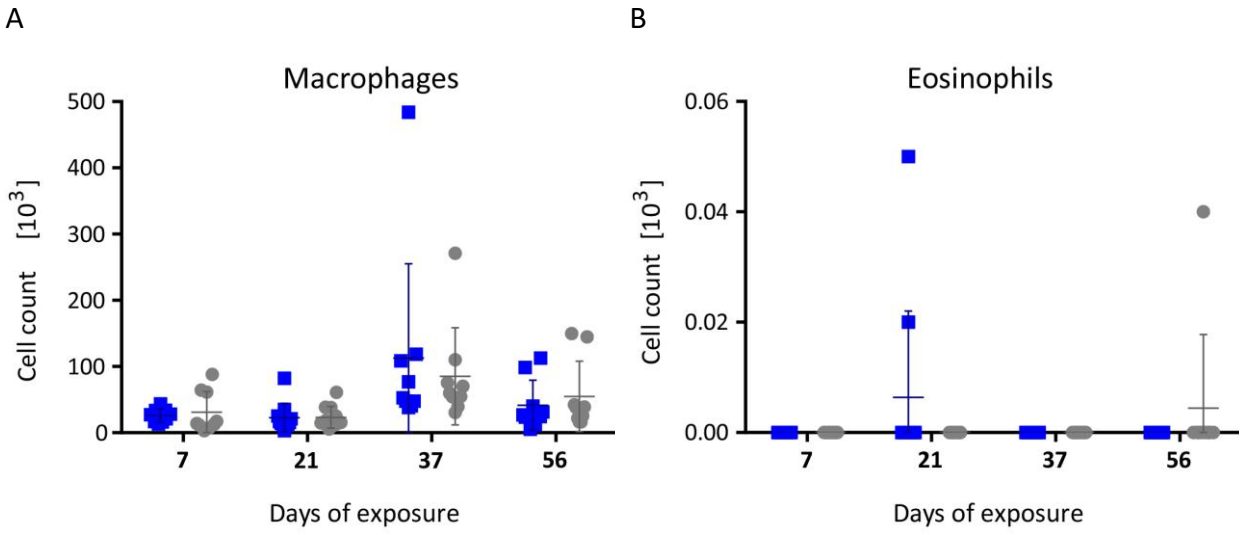


Figure 22. Absolute cell count in BAL fluid.
Each data point represents an individual animal; data are shown as mean ± SD (n = 9/group). (■ air, ● CS). Two-way ANOVA with Tukey's multiple comparison test. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.



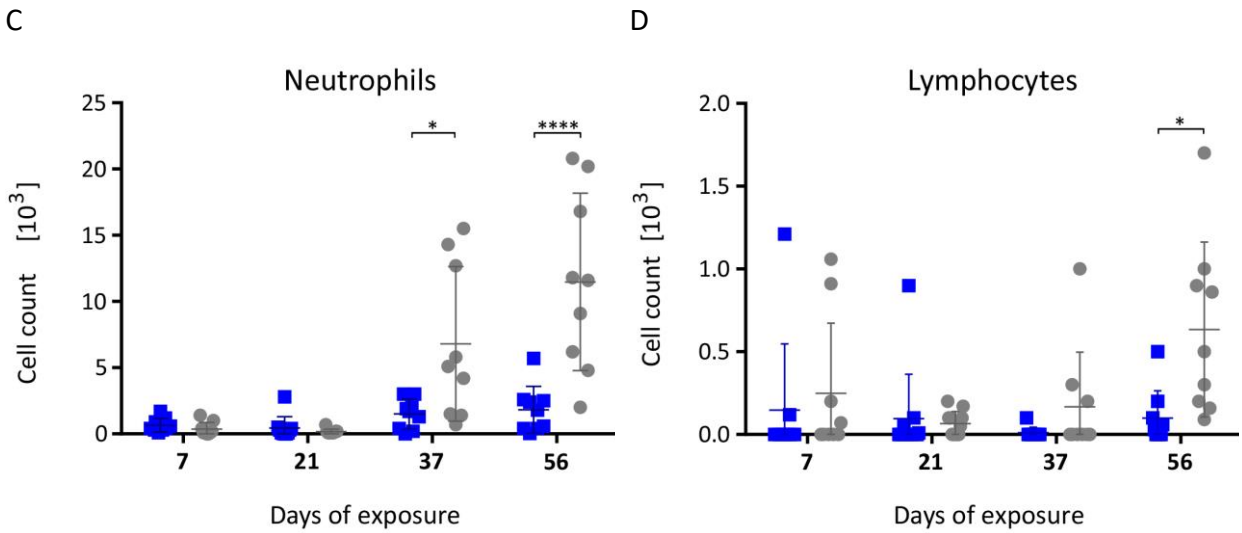


Figure 23. Number of inflammatory cells in BAL fluid.
A) Number of macrophages; B) Number of neutrophils; C) Number of lymphocytes; D) Number of eosinophils. Each data point represents an individual animal; data are shown as mean ± SD (n = 9/group). (■ air, ● CS). Two-way ANOVA with Tukey's multiple comparison test. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

4.2.5 Lymphocyte Levels in the Lungs of CS-exposed Mice

Further, as the level of lymphocytes in the lungs increased after longer CS exposure, their different subtypes were assessed at later timepoints. Within hematopoietic cells (Figure 24), the percentage of B-cells was slightly elevated upon smoke exposure, but the increase did not reach statistical significance. Among T-cell population (CD3+), the ratio of CD4+ to CD8+ was higher in the air groups in contrast to the samples from CS-exposed mice (Figure 25). However, statistical difference between the air and the smoke group was observed only at day 21 (Figure 16). While the ratio did not change among the smoking groups at all timepoints, over the time, a decline occurred in the air groups.

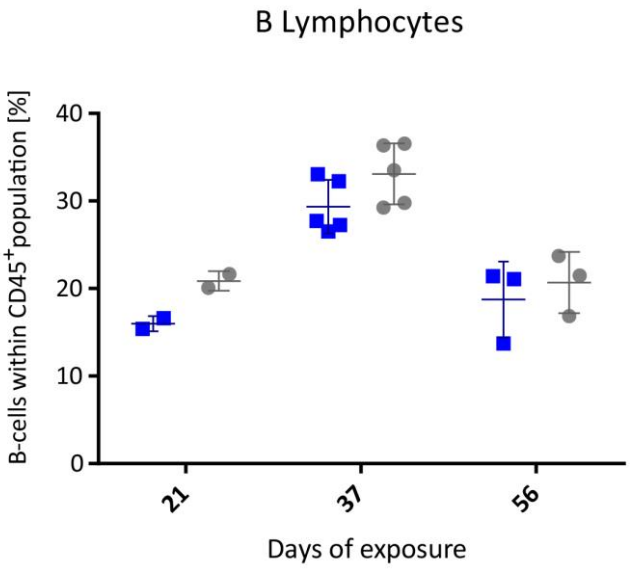


Figure 24. Percentage of B cells within haematopoietic cell population.
Each data point represents an individual animal; data are shown as mean ± SD (n = 2-5/group). (air, CS). Two-way ANOVA with Tukey's multiple comparison test. ***p < 0.001

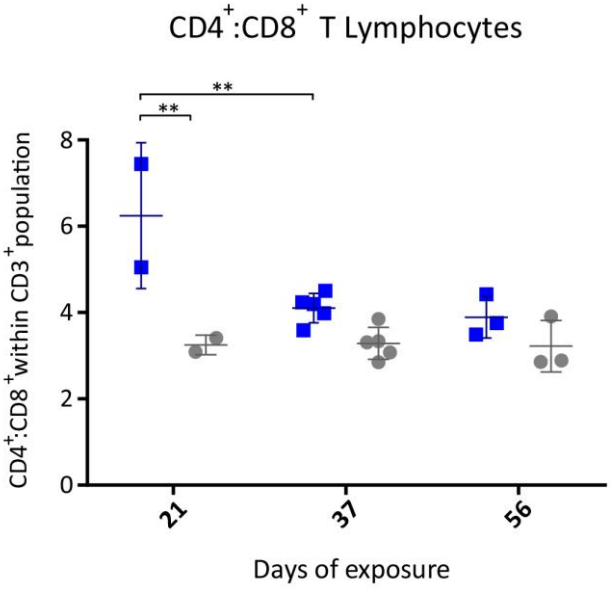


Figure 25. Ratio of CD4+ to CD8+ T cells in the lung samples.
Each data point represents an individual animal; data are shown as mean ± SD (n = 2-5/group). (air, CS). Two-way ANOVA with Tukey's multiple comparison test. **p < 0.01

4.2.6 MMP12 Protein Levels in BAL Fluid of CS-exposed Mice

To corroborate the increase of MMP12 at the gene expression level (Figure 21), protein concentrations of MMP12 was tested in BAL fluid. Results showed that the air control groups had constantly low levels of the protein. In contrast, the data from CS-exposed mice were in line with the previously described gene results. BAL fluid samples of these mice had notably higher MMP12 concentrations from day 21 which gradually increased with the amount of exposures (Figure 26).

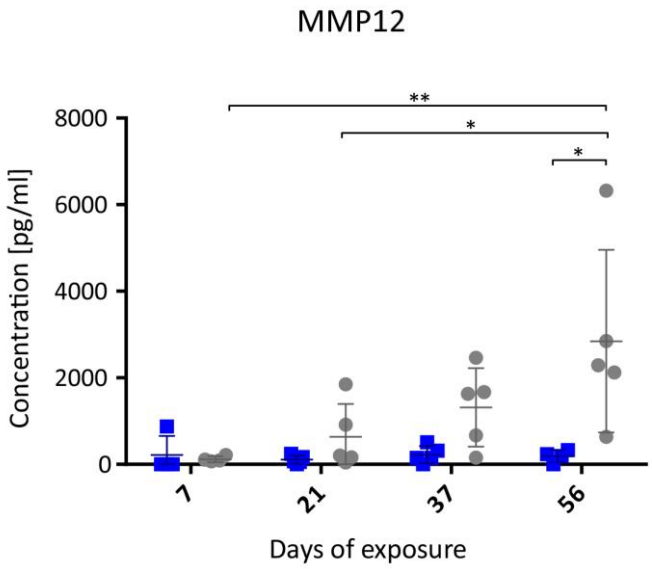


Figure 26.Concentration of MMP12 in BAL fluid.
Each data point represents an individual animal; data are shown as mean ± SD (n = 5/group). (air, CS). Two-way ANOVA with Tukey's multiple comparison test. *p < 0.05, **p < 0.01.

4.2.7 Inflammatory Mediators in BAL Fluid of CS-exposed Mice

Since CS had led to an influx of inflammatory cells (Figure 23), it was analyzed which proinflammatory mediators accompany the cellular inflammatory response. To this end, several cytokines and chemokines were quantified in BAL fluid. Keratinocyte chemoattractant/human growth-regulated oncogene (KC/GRO) chemokine significantly increased in the CS-exposed mice after 37 days (Figure 27a). Furthermore, the levels of vascular endothelial growth factor (VEGF) increased upon CS exposure (Figure 27b). However, other inflammatory mediators did not show clear pattern in response to the treatment (Supplementary figure 45). Interleukin (IL)-6 levels increased in CS-exposed mice at day 56, while in all previous timepoints the levels were lower compared with the control. Granulocyte-macrophage colony-stimulating factor (GM-CSF) levels significantly raised in the CS group at day 37. At all other timepoints, no differences among the groups were observed. Similarly, concentrations of interferon (IFN)-γ showed notable difference only after 37 days of CS exposure. Tumor necrosis factor (TNF)-α did not differ between the groups over the entire experiment. At earlier timepoints, mean levels of IL-1β were higher in the control samples, but later the increase occurred in CS-exposed groups. However, observed differences were not significant. Additionally, no obvious patterns or significances were observed in interferon gamma-induced protein (IP)-10, IL-33 and erythropoietin (EPO) levels.

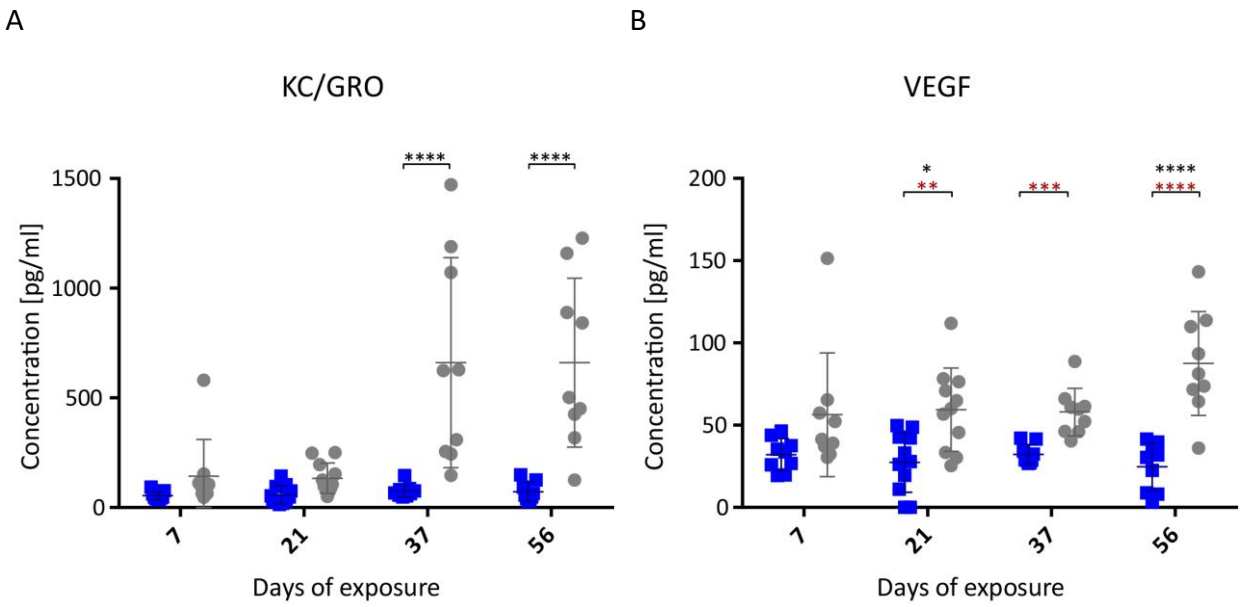


Figure 27. Concentration of inflammatory mediators in BAL fluid.
A) KC/GRO; B) VEGF. Each data point represents an individual animal; data are shown as mean ± SD (n = 9-11/group). (■ air, ● CS). Two-way ANOVA with Tukey's multiple comparison test (black stars). Unpaired t-test (red stars). *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

4.2.8 Lung Histology in CS-exposed Mice

On the one hand, the analysis of tight junction molecules did not indicate a disruption of the epithelial barrier, however the upregulation of MMP12 might be related to the remodeling of ECM. Hence the present thesis addressed the question whether CS induces morphological changes in the lungs. The analysis of lung sections stained with PAS did not reveal histological differences between the groups after 56 days of exposure. No noteworthy disruptions of alveoli or airspace enlargement in CS-exposed mice were observed (Figures 28 and 29). Quantification of the surface area also showed that CS- and air-exposed lungs had same frequency distribution of the alveolar size (Supplementary figure 46).

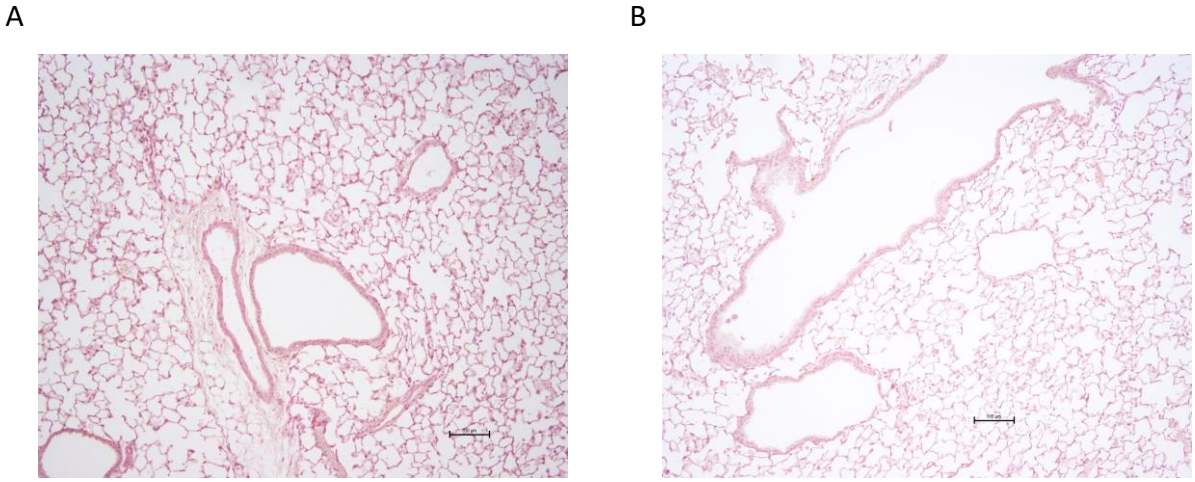


Figure 28. PAS histology staining of lung sections after 56 days of exposure (10-fold magnification).
A) Air-exposed mouse; B) CS-exposed mouse

A B

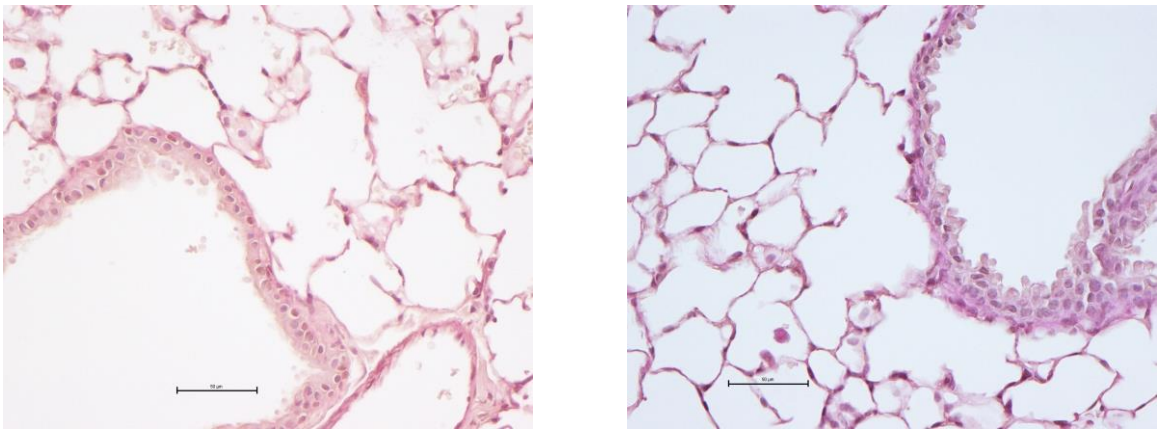


Figure 29. PAS histology staining of lung sections after 56 days of exposure (40-fold magnification).
A) Air-exposed mouse; B) CS-exposed mouse

4.2.9 Microbiome Analysis

4.2.9.1 Lungs

4.2.9.1.1 Normalization

Rarefaction analysis were used to control for the sufficiency of sequencing depth and remove the most under-sampled specimen. As most lung samples were reaching the maximum number of zOTUs at around 5000 reads, all reads were normalized to this value (Figure 30).

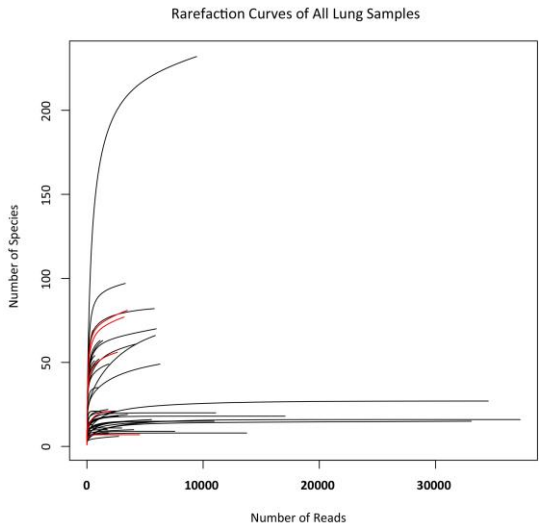
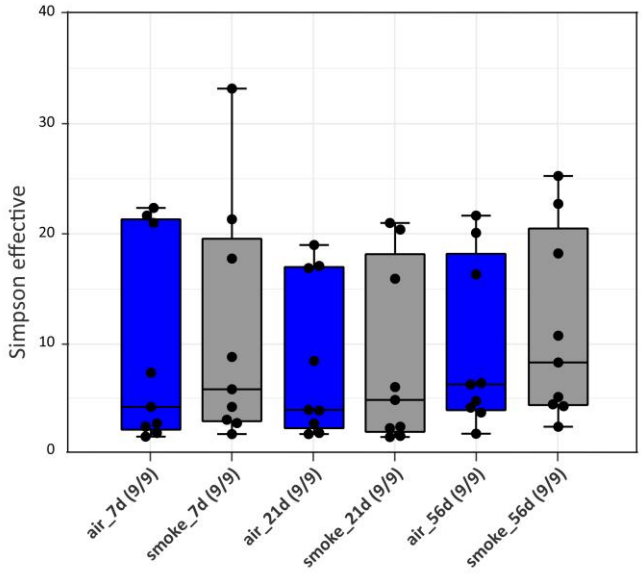


Figure 30. Rarefaction curves of the lung samples.
5 most undersampled cases were highlighted in red

4.2.9.1.2 Alpha Diversity

The diversity of zOTUs within each lung sample were calculated based on richness, Shannon and Simpson indices (Supplementary table 24) and data were visualized via box plot (Figure 31). However, alpha diversity did not differ significantly when all groups were compared together using Kruskal-Wallis rank test.



Kruskal–Wallis Rank Sum Test – all groups

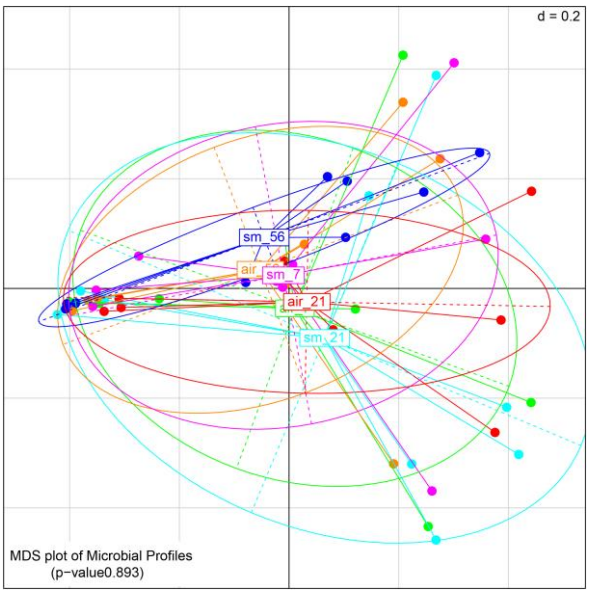
	P–value	Adj. p–value
Simpson.effective	0.7448	0.7448
Shannon.effective	0.6812	0.7448
Richness	0.6041	0.7448

Figure 31. Alpha diversity of lung samples.
Left: Box plot presentation of Simpson effective diversity index among lung samples (n = 9 animals/group); Right: Table of measured alpha diversity values. Kruskal-Wallis Rank Sum Test for group comparison (no significance was observed).

4.2.9.1.3 Beta Diversity

Generalized UniFrac distance values were calculated to evaluate similarities between microbial profiles of different groups. However, beta diversity among the lung samples' groups did not differ significantly, both regarding the exposure or the timepoints (Figure 32). Pairwise comparison of air- and CS-exposed mice from the same timepoints showed that these groups shared similar microbial profiles (Figure 33).

A



B

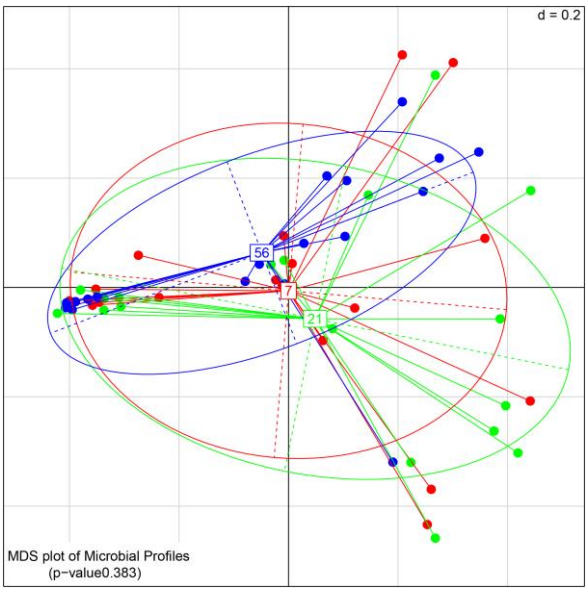


Figure 32. MDS plots of beta-diversity microbial profiles.
A) Comparison of all air- and CS-exposed groups at all timepoints; B) Comparison all lung samples grouped based on timepoints. n = 9 animals/group. Based on generalized UniFrac distances. PERMANOVA (999 perturbations) based levels of significance are displayed on the lower-left hand edge of the plots.

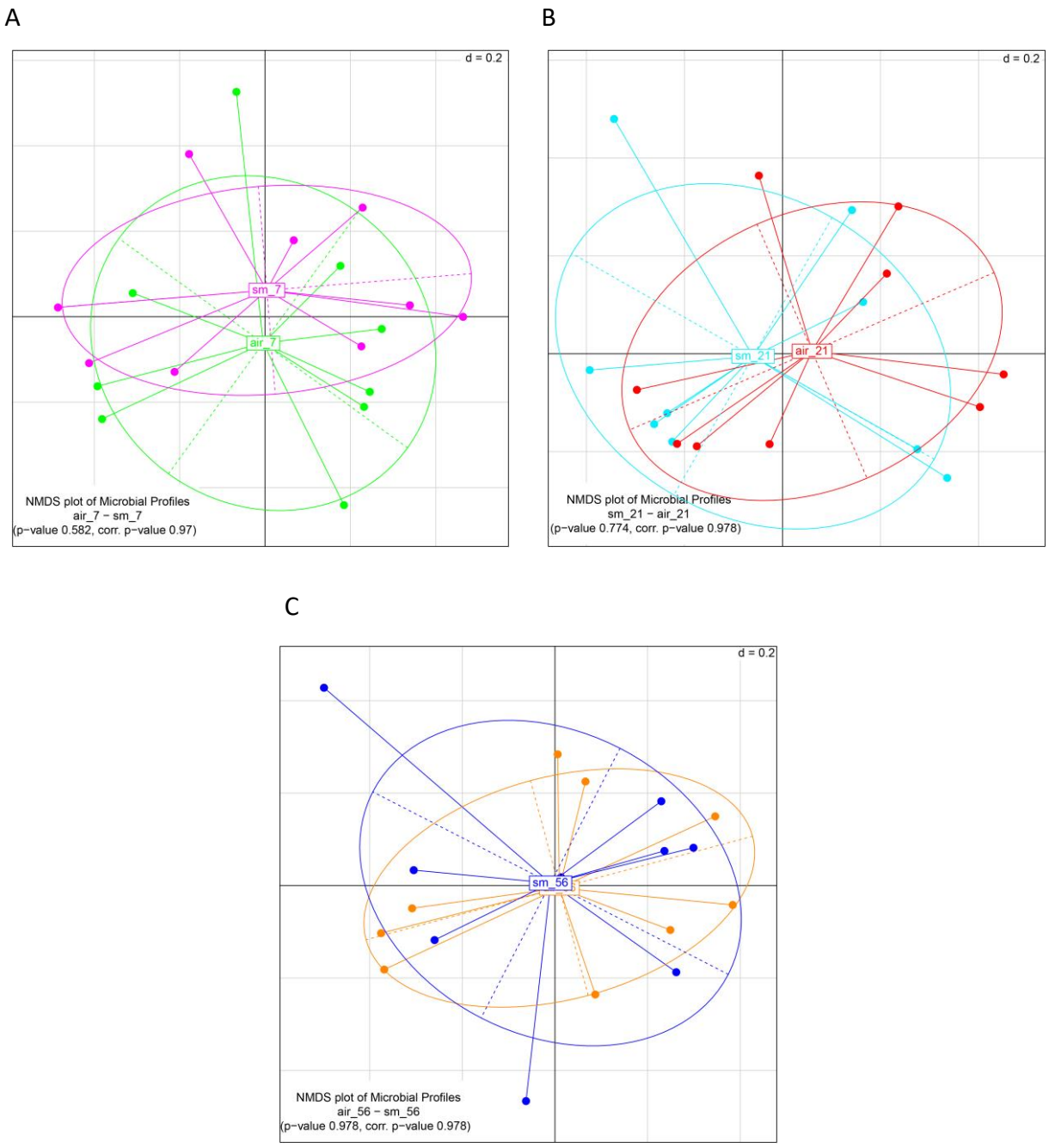


Figure 33. nMDS plots comparing groups from the same timepoint.
A) 7 days; B) 21 days; C) 56 days. n = 9 animals/group. Based on generalized UniFrac distances. PERMANOVA (999 perturbations) based levels of significance are displayed on the lower-left hand edge of the plots.

4.2.9.1.4 Taxonomy

Microbial profiles highly varied between lung samples within the same group, as illustrated at the phylum level (Supplementary figure 47). However, the most abundant phylum in all samples was Firmicutes followed by Proteobacteria, Bacteroidetes and Fusobacteria (Figure 34). The latter was not detected in the CS group on day 56. While the level of the Actinobacteria decreased in the air group over the time, the opposite was observed in the smoking groups.

Although the air group bar on day 21 shows higher abundance of the phylum Gemmatimonades, it was actually detected in only one sample.

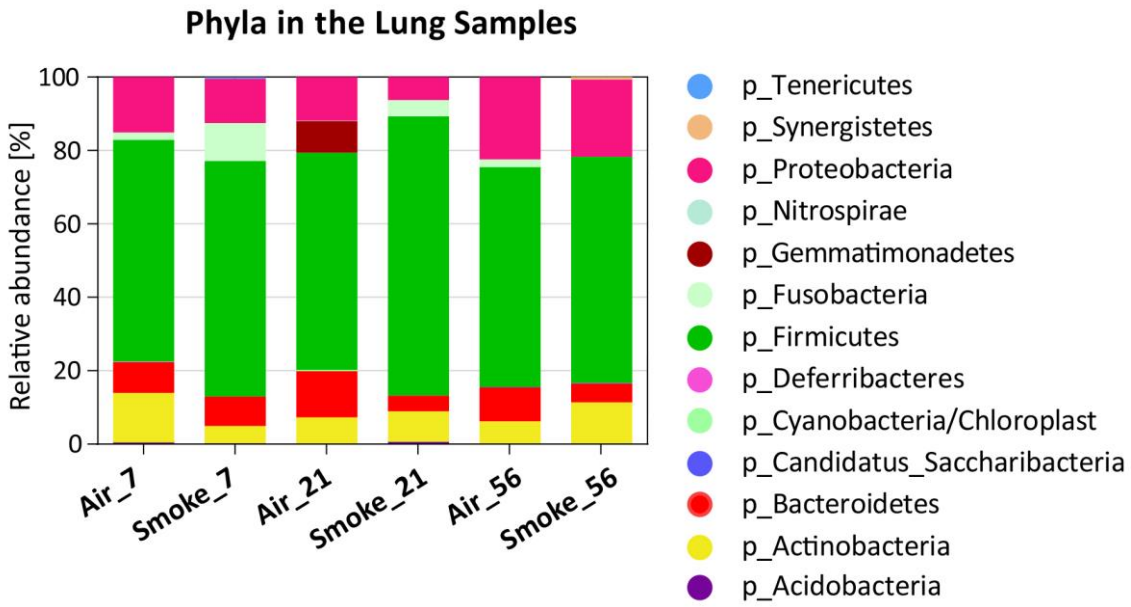


Figure 34. Taxonomy at phyla level of the lung samples.
Bar plots on the left present samples from air-exposed and on the left CS-exposed mice from three different time points.

4.2.9.1.5 Correlations

The correlation was measured between days of smoking, different alpha diversity values and the taxa identified in the lungs. In the presented plot, the color of the dot represents the direction of the correlation, with blue as positive, and red as negative correlation. Also, the size of the circles corresponds to the *p* values of the correlations (Figure 35). Bacteroidetes correlated positively with alpha diversity parameters throughout all taxonomic levels from phylum to the genus Barnesiella. With an addition of the genus Bacteroides, these taxa correlated negatively with smoking parameters. Among Firmicutes, family Ruminococcaceae with unclassified genus showed positive correlation with alpha diversity, while genera Lactobacillus and Streptococcus correlated negatively with these parameters. Firmicutes families Veillonellaceae and Clostridiales showed positive correlation with cotinine levels and smoking duration, together with Micrococcaceae family from Actinobacteria phylum. The opposite correlation had unclassified genera belonging to the families of Ruminococcaceae and Enterobacteriaceae.

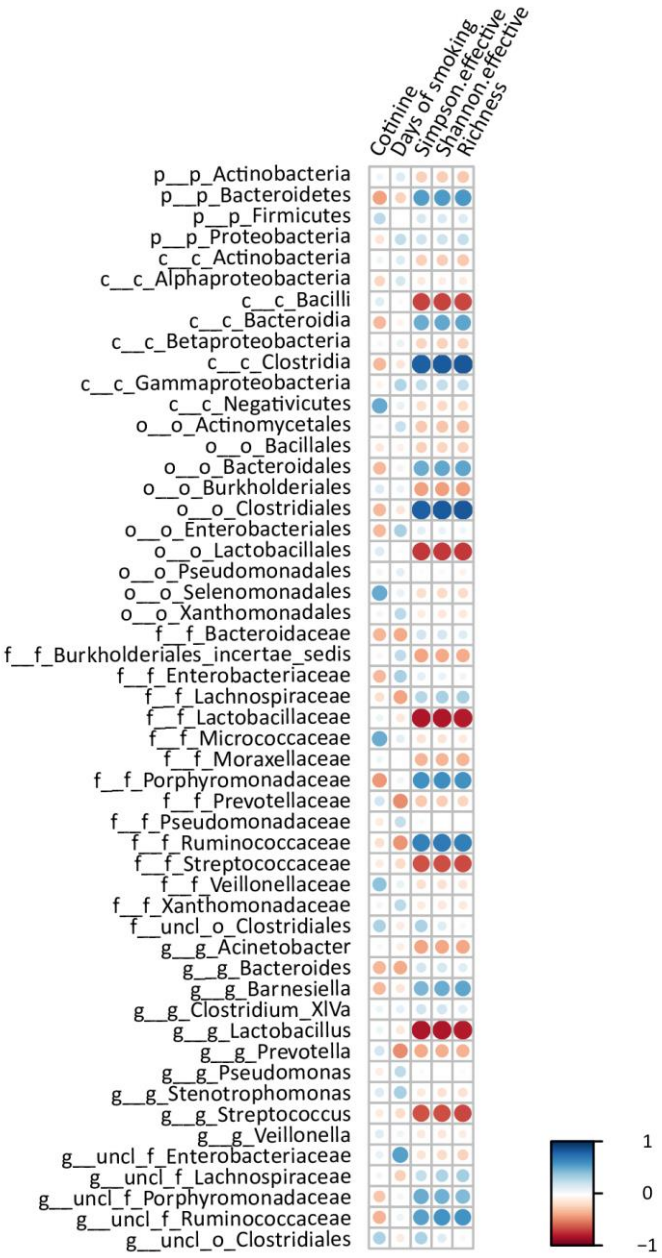


Figure 35. Correlation plot of certain taxa detected in lungs.
The size of the circles shows the significance of the correlation (bigger circle corresponds to a lower p-value). The color of the circle indicates the Pearson correlation between 1 (blue, positive correlation) and -1 (red, negative correlation) between taxa and alpha diversity values.

4.2.9.2 Caecum

4.2.9.2.1 Normalization

Caecum data were rarefied to remove the effect of discrepancy in number of sequences between samples. Therefore, mean number of reads was calculated based on the point where most of the samples reached maximum number of OUTs and all samples were normalized to this value (Figure 36).

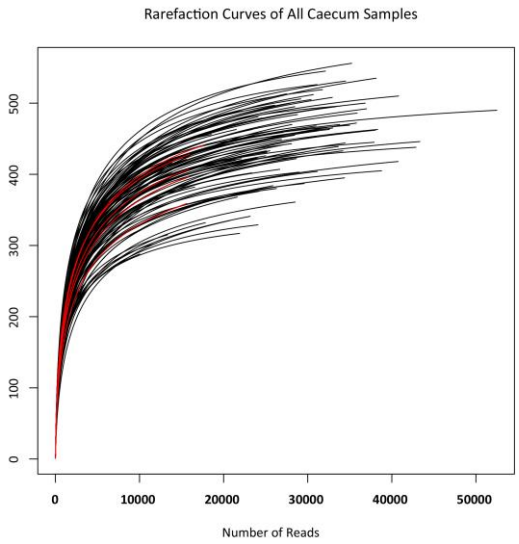
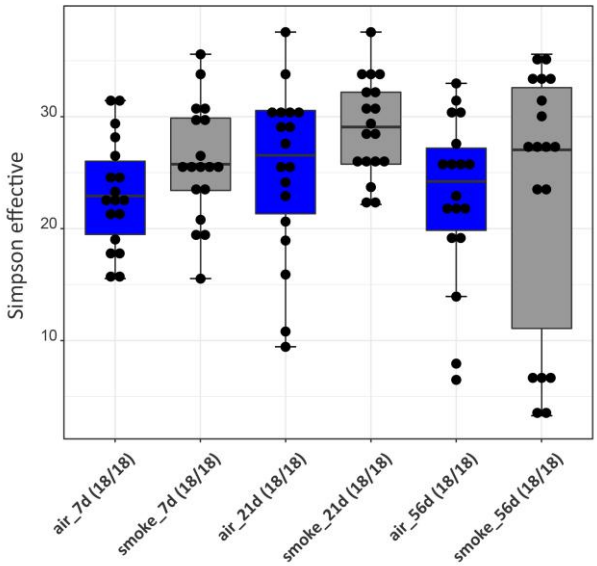


Figure 36. Rarefaction curves of the caecum samples.
5 most undersampled cases were highlighted in red.

4.2.9.2.2 Alpha Diversity

As seen on the box plot graph (Figure 37), alpha diversity (Supplementary table 25) was slightly higher in the caecum samples from CS-exposed mice compared with the air control group. Using Kruskal-Wallis statistical test, the differences in the diversity were quantified. As presented in the table, the mean Simpson effective index of all compared groups showed significant difference. However, the Shannon's diversity and richness of the caecum microbiome did not differ notably. When the groups were compared pairwise via Wilcoxon Rank Sum test, the alpha diversity differed significantly only between CS-exposed samples from day 21 and air groups from day 7 and 56.



Kruskal–Wallis Rank Sum Test – all groups

	P-value	Adj. p-value
Simpson.effective	0.0468	0.0849
Shannon.effective	0.0757	0.1163
Richness	0.1339	0.1919

Wilcoxon Rank Sum Test – pairwise

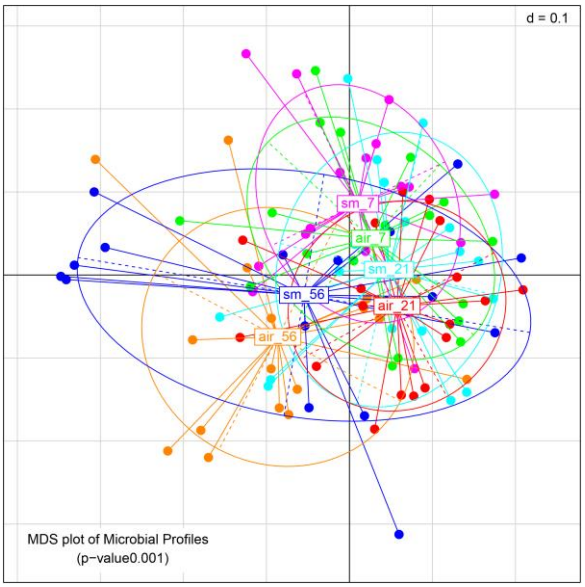
Groups	P-value	Adj. p-value
air 56_smoke 21	0.0058	0.0435
air 7_smoke 21	0.0012	0.0180

Figure 37. Alpha diversity of caecum samples.
Left: Box plot presentation of Simpson effective diversity index among caecum samples (n = 18 animals/group); Right: Table of measured alpha diversity values. Kruskal-Wallis Rank Sum Test for group comparison (no significance was observed).

4.2.9.2.3 Beta Diversity

Differences in the identified microbes among the groups were presented in the MDS plots. Microbial profiles differed significantly when comparison was performed between the groups based on the treatment and timepoints (Figure 38a). Statistical significance was reached when the groups were compared based on timepoints (Figure 38b). However, pairwise comparison of the air and smoke groups at each timepoint did not show significant differences (Figure 39).

A



B

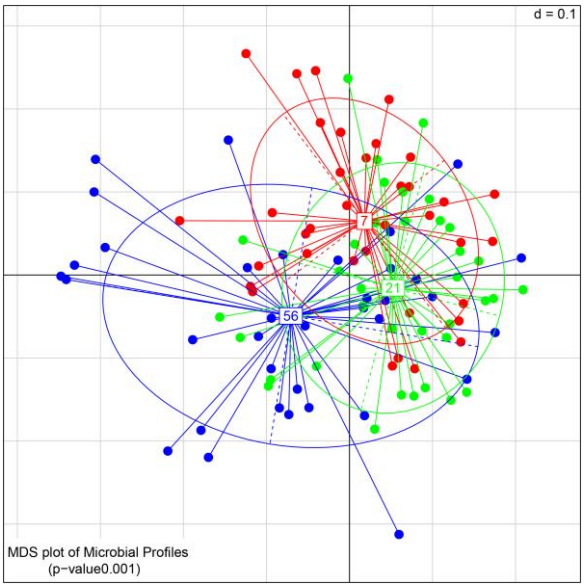
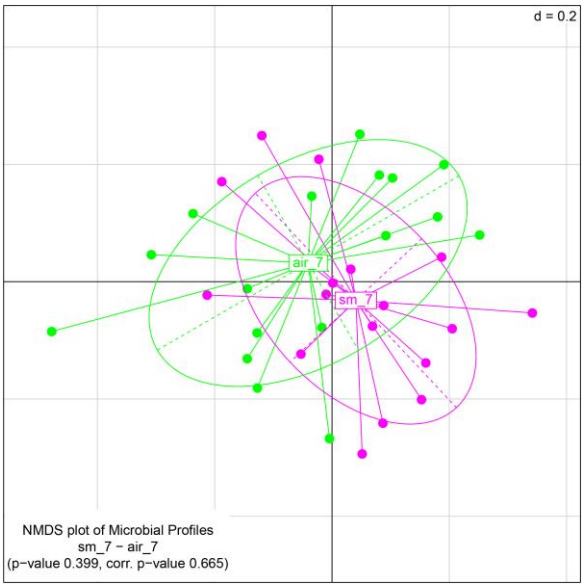
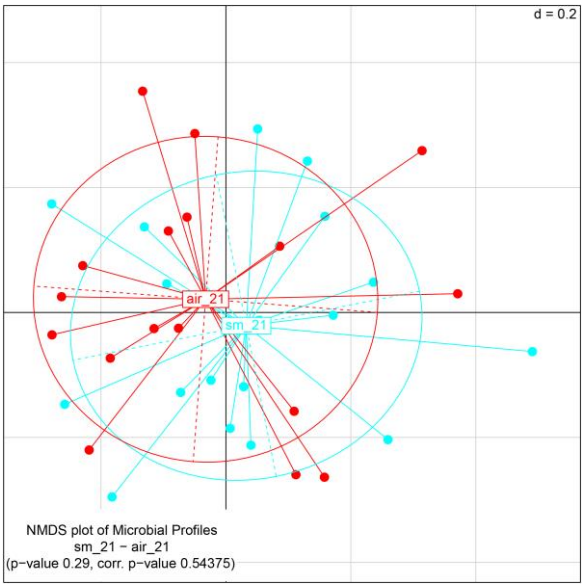


Figure 38. MDS plots of beta-diversity microbial profiles of caecum samples.
A) Comparison of all air- and CS-exposed groups at all timepoints; B) Comparison all caecum samples grouped based on timepoints (n = 18 animals/group). Based on generalized UniFrac distances. PERMANOVA (999 perturbations) based levels of significance are displayed on the lower-left hand edge of the plots.

A



B



C

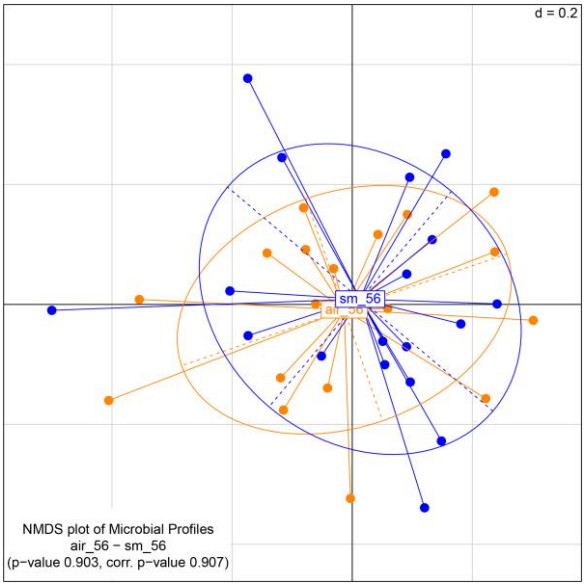


Figure 39. nMDS plots comparing groups from the same timepoint.
A) 7 days; B) 21 days; C) 56 days (n = 18 animals/group). Based on generalized UniFrac distances. PERMANOVA (999 perturbations) based levels of significance are displayed on the lower-left hand edge of the plots.

4.2.9.2.4 Taxonomy

The dominance of Firmicutes and Bacteroidetes was observed among the caecum samples. Proteobacteria, Deferribacteres, Candidus Saccharibacteria and Actinobacteria were also detected, but to a much lower extent (Figure 40). Unlike the lung samples, the caecum samples within the same groups did not show high variations in the microbial abundance (Supplementary figure 48). However, the abundance was also similar between different groups.

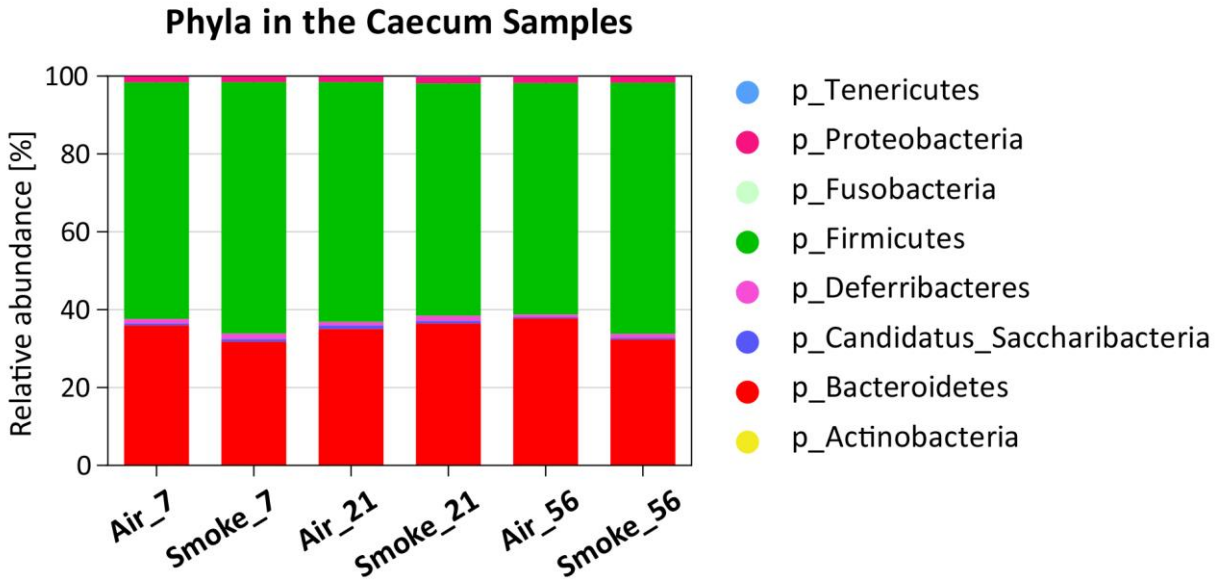
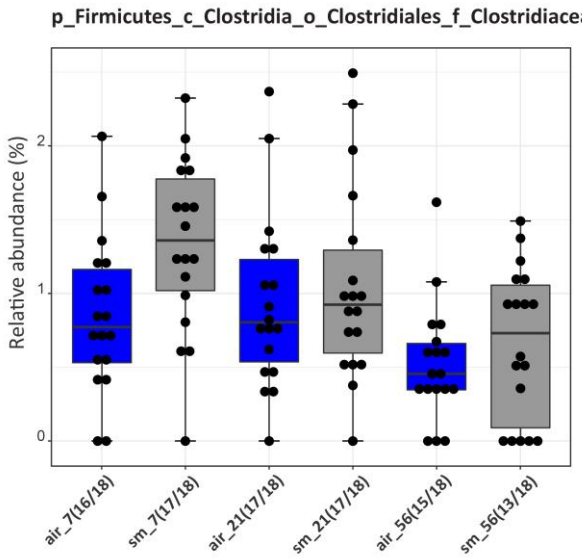


Figure 40. Taxonomy at phyla level of the caecum samples.
Bar plots present grouped samples.

4.2.9.2.5 Group Comparison of Certain Genera

The relative abundance of *Clostridium XIVb* was increased in the gut of the CS-exposed mice at all timepoints. The strain had the highest mean of the relative abundance after 7 days of smoking, which significantly differed from the air group (Figure 41).



Kruskal–Wallis Ranks Sum Test – all groups

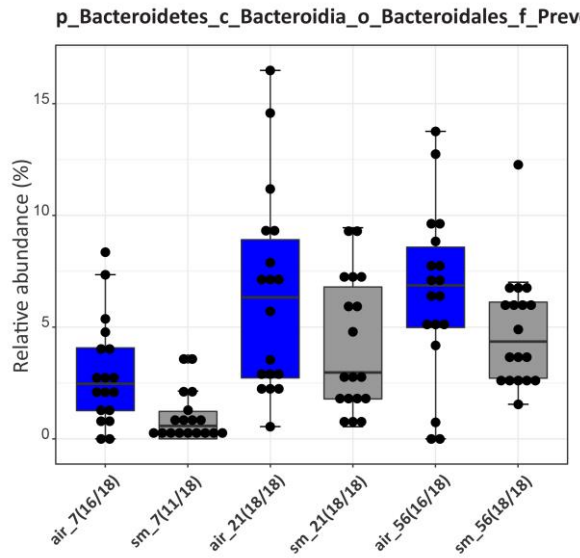
Groups	P–value	Adj. p–value
g_Clostridium XIVb	9e-04	0.0097

Wilcoxon Rank Sum Test – pairwise

Groups	P–value	Adj. p–value
air 7_air 56	0.0226	0.0678
air 21_air 56	0.0277	0.0692
smoke 7_smoke 56	0.0011	0.0082
air7_smoke 7	0.0155	0.0581

Figure 41. Relative abundance of Clostridium XIVb in the caecum samples.
Left: Box plot presentation of relative abundance in the caecum samples (n = 18 animals/group). Right: Table of measured statistical significance among the groups. Kruskal-Wallis Rank Sum Test for the comparison of all groups. Wilcoxon Rank Sum Test for pairwise comparison.

Conversely, *Prevotella* decreased in the CS-groups, which was statistically different compared with the air group at day 7 (Figure 42).



Kruskal–Wallis Ranks Sum Test – all groups

Groups	P–value	Adj. p–value
g_Prevotella	0	0

Fisher's Exact Test – pairwise

Groups	P–value	Adj. p–value
smoke 7_smoke 21	0.0076	0.038
smoke 7_smoke 56	0.0076	0.038

Wilcoxon Rank Sum Test – pairwise

Groups	P–value	Adj. p–value
air 7_air 21	0.0099	0.0212
air 7_air 56	0.0046	0.0115
smoke 7_smoke 21	0.0002	0.0008
smoke 7_smoke 56	0.0000	0.0000
air7_smoke 7	0.0032	0.0096

Figure 42. Relative abundance of Prevotella in the caecum samples.
Left: Box plot presentation of relative abundance in the caecum samples (n = 18 animals/group). Right: Table of measured statistical significance among the groups. Kruskal-Wallis Rank Sum Test for the comparison of all groups. Wilcoxon Rank Sum Test for pairwise comparison.

Significant difference between air group at day 7 to day 56 was observed at *Mucispirillum* genus, which was also observed at each hierarchical level from the genus to the phylum Deferribacteres (Figure 43).

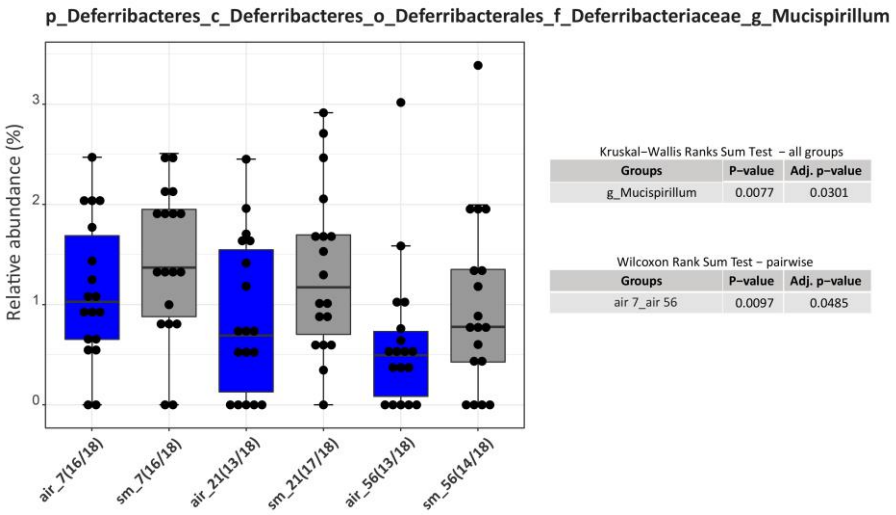


Figure 43. Relative abundance of Mucispirillum in the caecum samples.
Left: Box plot presentation of relative abundance in the caecum samples (n = 18 animals/group). Right: Table of measured statistical significance among the groups. Kruskal-Wallis Rank Sum Test for the comparison of all groups. Wilcoxon Rank Sum Test for pairwise comparison.

4.2.9.2.6 Correlation

Same as for the lung samples, correlation test was performed to test the association between taxa detected in the caecum and alpha diversity and smoking parameters (Figure 44). Most of the taxa which correlated positively with alpha diversity belonged to Firmicutes, including genera Anaerotruncus, Oscillibacter, and unclassified genera from families Ruminococcaceae, Lachnospiraceae and Clostridiales. Clostridium XIVa and Mucispirillum showed negative correlation with alpha diversity. Regarding cotinine levels, genera Acetatifactor, Intestinimonas and Pseudoflavonifractor showed positive correlation, while phylum Bacteroidetes, class Bacteroidia and order Bacteroidales showed negative correlation. Genera Anaerotruncus and Clostridium XIVb correlated negatively with days of smoking, in contrast to Prevotella which showed positive correlation.

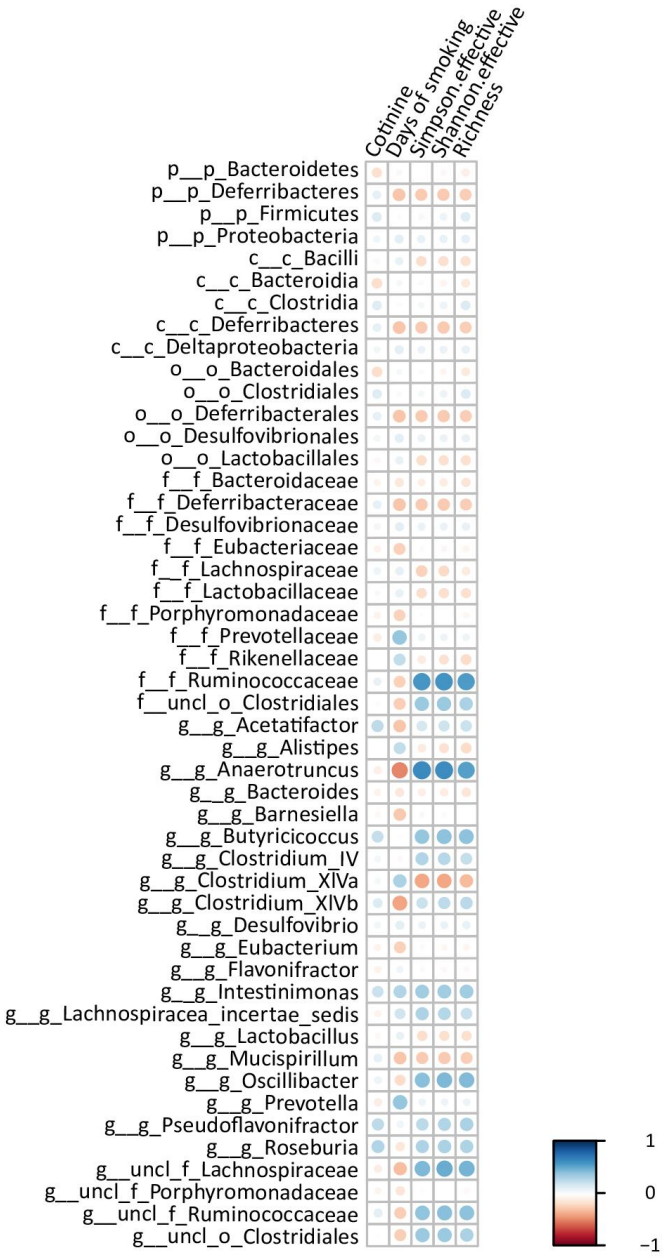


Figure 44. Correlation plot of certain taxa detected in caecum.
The size of the circles shows the significance of the correlation (bigger circle corresponds to a lower p-value). The color of the circle indicates the Pearson correlation between 1 (blue, positive correlation) and -1 (red, negative correlation) between taxa and alpha diversity values.

5 Discussion

5.1 Human Study

Smoke from a burning cigarette consists of thousands of chemicals, of which at least 250 have been identified as harmful[118]. Despite knowing the dangerous effects of smoking, a certain population will probably never be willing or able to quit. As CS is one of the major causes of chronic lung diseases, which have been associated with altered microbial composition, this study aimed to expand the understanding of the relationship between CS and lung microbiota.

5.1.1 Microbiota of Human Respiratory Tract at Different Sampling Sites

Different sites of the respiratory tract contain notably different microbiota, as demonstrated by alpha and beta diversity analyses. All three sampling spots showed distinct microbial composition with the lowest diversity in the nasal cavities. Bacteroidetes (*Prevotella*), Firmicutes (*Veillonella* and *Streptococcus*) and Actinobacteria (*Actinomyces*) were the most abundant phyla found in PHS and BAL samples. NSW samples had Actinobacteria (*Corynebacterium*) and Firmicutes (*Staphylococcus*) as the most dominant ones, but aforementioned were also detected. These findings are in agreement with the study of Stearns *et al.*[44] where bacterial composition of the URT was investigated. It has been assumed that the majority of the lung microflora directly originate from oral and nasal microbiota derived via microaspiration[31,119]. This study confirmed the strong overlap of the oropharyngeal and lung microbiota, but smoking disturbed this dynamic relationship.

5.1.2 Changes of the Nasal Microbiota upon Smoke Exposure

Among Firmicutes, *Dolosigranulum pigrum* was more abundant in never-smokers and mild smokers, and correlated negatively with smoking years, the maximum number of cigarettes smoked per day and packyears. On the other hand, *S. epidermidis* correlated positively with these three factors and was more abundant among long- and short-term smokers. It has been shown that *S. epidermidis* can serve as an important reservoir for antibiotic resistance genes[120], and is a commensal bacteria of the skin and mucous membranes in many individuals[121,122]. To the knowledge of the author of the present thesis, this strain has not been associated with smoking so far. Nicotine can enhance *S. epidermidis* biofilm formation[123], which might explain its higher relative abundance in this study. Among Actinobacteria phylum, *Propionibacterium* was correlating positively with smoking years (increase with years) and *Corynebacterium* negatively

with smoking years, max cigarettes per day and packyears. *Corynebacterium* is also a common resident of normal nasal microflora[32] and it was earlier reported that its abundance decreased in the nasopharynx upon smoking[124,125]. Given that colonization with *Dolosigranulum* or *Corynebacterium* in the URT of infants has been associated with more stable microbial profile[126,127], the relative reduction of these strains could indicate a negative influence of smoking on the nasal microbiota. Among Bacteroidetes, *Prevotella* correlated positively with the maximum number of cigarettes smoked per day.

5.1.3 Changes of the Oropharyngeal Microbiota upon Smoke Exposure

The relative abundance of Firmicutes phylum was significantly higher in PHS of smokers compared with never-smokers. Actinobacteria were significantly more abundant in smokers and ex-smokers compared with never-smokers and it was related to the higher relative abundance of its genus *Atopobium*. It correlated positively with cotinine and numbers of daily cigarettes, but not with long-term smoking. While Firmicutes and Actinobacteria were more abundant among smokers, the opposite occurred for Proteobacteria. Their relative abundance was lower in the PHS samples from smokers compared with ex- and never-smokers. When compared on the genus level, although *Neisseria* was the least detected in the samples of smokers, its relative abundance was lowest in ex-smokers. This genus correlated negatively with maximum number of cigarettes used per day. These observations are in line with previous reports[124,128] confirming that CS shifts oropharyngeal microbiome. Smoking could favor depletion of oxygen[129], and consequently anaerobic and facultative anaerobic Actinobacteria and Firmicutes could outgrow aerobic Proteobacteria. Even though CS could promote anaerobic environment, facultative anaerobic Fusobacterial genus *Leptotrichia* and the strictly anaerobic Firmicutes strain *Centipeda periodontii* correlated negatively with smoking parameters. This finding is still in the agreement with the comprehensive study of Wu *et al.*[129] where it was suggested that the depletion of these microbes could occur or directly by CS toxicants or indirectly through interaction with other microbes.

5.1.4 Changes of the Lung Microbiota upon Smoke Exposure

The relative abundance of Bacteroidetes decreased in BAL samples taken from smokers, but it was not significant. Hilty et al.[46] suggested that the reduction of this phyla in asthmatic and COPD patients may be relevant to disease. The genus *Prevotella* negatively correlated to packyears in BAL samples, whereas BAL specific Bacteroidetes correlated positively with long-term smoking.

Like in the oropharyngeal microbiota, the relative abundance of Firmicutes phyla notably increased upon smoking. However, its genus *Veillonella* negatively correlated with packyears. A decrease in relative abundances of Proteobacteria occurred in smokers, however without statistical significance. Its genus Campylobacter correlated negatively with all smoking parameters and the relative abundance differed among the samples of smokers and never-smokers. Several β- and γ- proteobacterial species were detected exclusively in the BAL samples and were in positive correlation with all smoking related parameters. Specifically, *Acinetobacter berezinae, A. johnsonii, Cupriavidus metallidurans*, *Serratia marcescens* and *Stenotrophomonas maltophilia* were strongly influenced both by long-term smoking and smoking intensity. Disruption of the lung epithelium by CS[130] can increase chances for biofilm-forming bacteria to outgrow the surface. As β- and γ-proteobacteria are known to be potential pathogens and biofilm producers[131–135], their presence in the lungs might indicate a colonization of injured airways. Concordantly, the present study showed that the abundance of these taxa correlated strongly with smoking intensity and duration. The parallel decrease of *Prevotella* sp., which has been suggested to play a role in shaping of immune response[136], could further indicate a dysbiosis of the lung microbiome which is driven by the severity and duration of smoking exposure.

5.2 Study in Mice

Based on the last World Health Organization report, cigarette smoking kills more than eight million people annually[137]. Hence, smoking poses a vast threat to public health worldwide, affecting both active and passive smokers. For ethical reasons, it is not possible to investigate CS effects in first-time smokers and therefore animal models are an indispensable research tool. This mouse study aimed to broaden knowledge about initial smoke-induced changes in the body and in the lung and gut microbiota. To avoid potential gender influence[138,139] on microbial diversity, only female BALB/c mice were used in the experiment.

5.2.1 Body Weight Change upon Smoke Exposure

In the present study, a significant decline in body weight of the CS-exposed mice was observed starting from the second week and continued until the end of the experiment. Body mass of the smoking group barely changed before day 35 when the mice started gaining weight. This clear differences are in line with previous findings showing the impact of smoking on body weight of mice[140,141]. Studies from Chen *et al.*[142,143] demonstrated that even short-term whole-body CS exposure of BALB/c mice led to reduced food intake and fat mass, thus promoting weight loss. In

another study with the same mouse strain, a sharp reduction in weight gain occurred after initial CS exposures and recovered within a month after removal from the smoke, reaching similar values of the control group[144]. In contrast to the effect of whole-body smoke exposures, nose-only exposures usually had minimal impact on body weight[145,146].

5.2.2 Cotinine Levels in Blood of CS-exposed Mice

To measure the amount of smoke inhaled by the mice, the levels of cotinine were measured in serum obtained 2 hours after last exposure. This primary metabolite of nicotine is used as an indicator of CS exposure due to its longer half-life compared with that of nicotine[147,148]. In this study, air-exposed mice did not have cotinine in serum, which confirmed that CS was not present in the control chember. In contrast, serum samples of CS-exposed groups always displayed high concentrations of cotinine. Interestingly, the levels were significantly higher on the 21st and 56th day of smoking compared with the 7th and 37th day. As the half-life of cotinine is approximately 16 hours[149], which is shorter than the time between two CS exposures, no cumulative effect of smoking on cotinine levels was expected.

5.2.3 Expression of Xenobiotic Related Genes upon Smoke Exposure

Cyp1A1 belongs to superfamily of cytochrome P450 enzymes and plays a role in the metabolism of xenobiotics such as those found in CS. By analysing differences in gene expression between the lungs of air- and CS-exposed mice, this study showed that Cyp1A1 was upregulated only in the smoking group. This, together with cotinine levels, confirmed that CS was not present in the chamber of the air control group. Nevertheless, the gene expression did not continuously increase throughout the experiment but declined after 21 days. To discover the reason for this decrease, it was tested whether some changes occur at the expression of AhR gene. This gene codes for the protein which binds tobacco xenobiotics and regulates the expression of the Cyp1A1 gene[150]. However, this study did not show differences in AhR gene expression between the control and the CS-exposed mice. This still does not exclude possible posttranscriptional and translational changes. Thus, it would be necessary to test the protein levels to clarify at which step of xenobiotic metabolism the need for higher production of Cyp1A1 was reduced.

5.2.4 Inflammatory Cells and Mediators in CS-exposed Mice

BAL technique has allowed sampling and examination of free lung cells and characterization of immune processes of the LRT[151]. In general, animal and human studies have shown consistency

regarding the effect of CS on the cells recovered by BAL[152–154]. Results from a large cohort[155] of healthy current, ex- and never-smokers showed a striking effect of smoking on the number of inflammatory cells in BAL fluid, which correlated positively with smoking history. The total cell number was almost four times higher in current smokers compared with the other two groups, mainly due to an increase of alveolar macrophages, but also neutrophils and lymphocytes. Furthermore, BAL fluid parameters did not differ between ex- and never-smokers, indicating a normalization after smoking cessation. However, long-term cessation effects have not been fully investigated so far. In the work of Vlahos *et al.*[156], when BALB/c mice were exposed to CS generated by 3, 6 or 9 cigarettes per day for 4 consecutive days, the total number of inflammatory cells and neutrophils in BAL fluid increased in a dose-dependent manner. In another study, both acute (2-weeks) and sub-chronic (8-weeks) whole-body smoke exposure of BALB/c mice led to a significant increase of neutrophils[157]. Botelho *et al.*[158] showed that both BABL/c and C57BL/6 mice manifested an influx of neutrophils and an increase of inflammatory cytokines in BAL fluid after only 4 days of smoking. Neutrophilia was more pronounced in BALB/c mice and it persisted after prolonged exposure, while in C57BL/6 it attenuated. Consistent with the aforementioned data, the present study showed that mice exposed to CS had a significant increase in neutrophils and lymphocytes. However, these responses occurred only after exposure over a longer period of time. Moreover, the levels of total inflammatory cells and macrophages were not significantly changed. These effects could be explained by the kinetic of changes. Morris *et al.*[159] demonstrated that in BALB/c mice, the peak of neutrophils in BAL fluid occurred 24 hours after CS exposure and was afterwards followed by an influx of lymphocytes and macrophages. In the present study, BAL was always performed 2 hours after last exposure which could be argued as short time for a detectable inflammatory response. The mice seemed better able to compensate for the effect of smoking early in the experiment as the response was not observed shortly after the last exposure. At later timepoints, the inflammatory response commenced earlier and could be clearly seen after only few hours.

Airways of smokers with COPD have been characterized with increased infiltration T lymphocytes, particularly CD8+[160,161]. Several studies have looked for the impact of active smoking on the ratio of CD4+ to CD8+ in peripheral blood, however with inconsistent results, showing or increase[162], decrease[163,164] or no significant effect[165]. In the present study, the CD4+/CD8+ ratio was higher in the lungs of the air-exposed compared with the CS-exposed mice.

This difference is mostly due to slightly higher CD4+ and lower CD8+ in the air group. While in the CS group the ratio was at constant level at days 21, 37 and 56, in the control group it was higher at the beginning and declined later.

To determine the cause of the influx of inflammatory cells in BAL fluid, the levels of pro-inflammatory cytokines and chemokines were also explored. Their secretion has important effects on multiple components of the immune system, and it shapes inflammatory responses. Emerging data has shown that CS induces secretion of pro-inflammatory mediators including TNF-α, INF-γ, IL-1β, IL-6, IL-33, IP-10, GM-SCF, VEGF and KC/GRO[85,166–168]. TNF- α is multifunctional cytokine with important roles in pathogenesis of inflammation[169–171], lymphoid organogenesis[172], tumor cell death[173,174] and defense against intracellular pathogens[175–177]. It is activated by various cell types including activated macrophages and lymphocytes[178,179]. IL-6 is central mediator of the immune system and a potent inducer of inflammation. It has a key role in the acute phase of infection or tissue damage[180,181] and regulates lymphocyte effector responses to pathogens[182–184]. It is produced by various haematopoietic and non-haematopoietic cells[185]. Another mediator involved in acute phase response is IL-1β which often acts together with IL-6. Its activity has been shown to be crucial for the development of inflammatory diseases[186,187]. IL-33 belongs to the IL-1 superfamily and not only controls pathogen removal but also promotes tissue repair[188]. GM-CSF has been associated with the progression of inflammation by inducing macrophage and granulocyte differentiation and survival[189,190]. INF-γ is essential in host defense against microbial infections. It stimulates macrophages for enhanced antimicrobial response and dendritic cells for stronger antigen presentation[191]. IP-10 is a chemokine secreted by several cell types in response to IFN-γ and exerts pro-inflammatory and anti-angiogenic properties[192]. Its inhibition decreased the expression of inflammatory mediators and improved lung function in COPD-mouse model[193]. VEGF has been suggested as important factor for the development of CLD, but without clear roles due to contradictory reports about its protective and harmful mechanisms of action[194]. Studies have demonstrated that CS induced production of KC/GRO[140,159,195] (also known as CXCL1). This mouse chemokine is related to human IL-8 protein and is responsible for neutrophil recruitment[196]. The results from this thesis agree with previous findings. The significant increase in KC/GRO after 37 days of smoking correlated with the neutrophil influx, once again confirming that this chemokine attracts them to a site of inflammation. Maybe it could be surmised that neutrophils, recruited via CS-induced release of

KC/GRO, contributed to the increase of VEGF, coordinating thus lymphangiogenesis which is an important response to inflammatory insults[197,198]. Unlike previous findings, CS exposure did not alter the general proinflammatory milieu, as indicated by measurement of TNF-α, IL-6, IFN-γ and IL-1β. The mean values of IL-1β were lower in the CS-groups after 7 and 21 days of smoking, with the opposite result at the later timepoints. However, this change did not reach statistical significance probably due to high distribution in the groups.

5.2.5 Expression of Tight Junction Genes and Histology of the Lungs Upon CS Exposure

CS can affect the airway epithelial barrier by disrupting tight junction proteins and modifying the expression of their genes[199,200]. To elucidate whether this occurred in the lungs of CS-exposed mice, the expression of OCLN and TJP1 genes was analyzed in the present study. While the OCLN gene was downregulated upon CS exposure, the difference was not significant when compared with the air group. Regarding the TJP1 gene, the expression did not differ between the groups at each timepoint. Further, smoke exposure did not cause lung injury nor enlargement of airspace. Vlahos et al[156] agreed that 8 weeks of CS exposure can cause mild lung inflammation, but not acute changes. The results regarding tight junction genes' expression and histology are comparable, confirming that this smoking model did not cause notable disruptions of airspace.

5.2.6 MMP12 Gene Expression in the Lungs and Protein Regulation

MMP12 belongs to the family of matrix metalloproteinases which are involved in ECM remodeling[201] and are involved in the migration of macrophages and neutrophils into the lungs[202,203]. Hence, dysregulation of MMP12 levels and activity may result in tissue damage. It has been speculated that MMP12 could contribute to regulation of inflammation by controlling macrophage activation[204]. One mouse model, which mimicked some COPD features, demonstrated that instillation of a recombinant human form of MMP-12 in the airways induced early accumulation of neutrophils followed by a delayed recruitment of macrophages[203]. As this protease has also been associated with smoke-induced pathological changes in the lungs[205–207], the impact of CS exposure on the MMP12 at the gene and protein levels were tested in this thesis. The data regarding the gene expression were consistent with other studies which demonstrated that CS exposure increased the expression of the MMP12 gene in the lungs[156,166]. For example, Morris et al.[159] showed that the MMP12 gene was highly expressed in the CS-exposed mice while genes encoding for other MMPs did not differ. Regarding the protein levels, the same trend was observed, indicating an upregulation of MMP12 in the BAL fluid of CS-exposed mice. To the

knowledge of the author, there were no similar studies in the BAL fluid of mice, but the findings are in line with studies of sputum and BAL samples from COPD smokers[208–210]. The increase in MMP12 protein levels correlated with neutrophil recruitment and KC/GRO release in BAL fluid.

The general picture emerging from the analysis is that CS exposure over longer period of time induced the release of the MMP12 protein in the airways, followed by neutrophil and lymphocyte recruitment, mostly due to the influx of pro-inflammatory cytokines. However, the triggered response did not cause alveolar enlargement or disruption.

5.2.7 Murine Lung and Gut Microbiota upon CS Exposure

As expected, microbes in the lung samples showed lower Simpson diversity compared with those of the caecum. The taxonomic results of the lung microbiota were in general agreement with other mouse studies showing the dominance of Firmicutes, Proteobacteria, Bacteroidetes, Actinobacteria and Fusobacteria[48,50]. However, alpha diversities among all samples, as well as beta diversities between the groups, did not differ significantly. In general, most of the taxa identified in the lungs showed opposite correlations with alpha diversity and smoking parameters. Thus, genus Barnesiella correlated positively with alpha diversity, but negatively with cotinine levels and smoking duration. The same correlation was observed for its family Porphyromonadaceae, order Bacteroidales, class Bacteroidia and phylum Bacteroidetes. Another genus belonging the same phylum, Bacteroides, showed negative correlation with smoking parameters, which is in contrast with the result observed in the human BAL samples. Unclassified genus from Clostridiales family showed positive correlation with cotinine levels. Of note, difficulties to define core lung microbiota stemmed from large inter-individual variations in the composition.

In the caecum samples, CS-exposed groups had slightly higher alpha diversity than the control groups. However, only the smoking group from day 21 statistically differed from the air groups from days 7 and 56. Microbial composition differed notably among the samples grouped based on the treatment and timepoints together, but did not differ between smoking and air groups from the same timepoints. Taxonomic results showed the dominance of Firmicutes and Bacteroidetes in all caecum samples which is line with other murine gut microbiota studies[211,212]. To much lower extent, Proteobacteria, Deferribacteres, Candidus Saccharibacteria and Actinobacteria were present as well. Inter-individual variations of microbial composition were

not high as those detected in the lungs. The identified taxa did not show clear trend in the correlation with cotinine and the days of smoking. The genus Clostridium XIVb was more abundant in the CS-mice and in positive correlation with cotinine levels. Its relative abundance decreased with time in the controls and CS-groups and the strain correlated negatively with days of smoking. Another strain which correlated positively with cotinine and negatively with smoking days was Mucispirillum and it was always slightly more abundant in the CS-exposed groups. Also, its relative abundance decreased over time in both treatment groups which was significant between the air groups from day 7 and day 56. Prevotella decreased in the smoking groups and showed slightly negative correlation with cotinine levels, yet it correlated positively with days of smoking. The relative abundance increased over time in both treatment groups.

5.2.8 Outlook

This thesis described bacterial community compositions of three sampling spots in the respiratory tract of healthy individuals. The strong overlap of the oropharyngeal and lung microbiota was confirmed, and it was shown that this dynamic relationship was disturbed by smoking. Further, smoking histories were correlated with the microbial composition in the URT and LRT. CS had an impact on the relative abundances of distinct potentially pathogenic bacterial groups in different areas of the respiratory system. Certain taxa were detected exclusively in the lungs and correlated positively with all smoking parameters. Surely, these results should be further explored in a larger cohort to understand the resilience/vulnerability of certain bacterial genera to smoking. Mouse study in this thesis could not show significant impact of CS exposure on the lung microbial community, while inflammatory response was clearly triggered. This can be improved by usage of another mouse strain e.g. C57BL/6J which have shown higher dynamic of microbial community. Another limitation in this thesis is that obtained microbiome data, together with other published studies, are largely descriptive. Therefore, it is not possible to assign functions to those findings and to conclude to what extent changes in bacterial community structure would be related to clinical outcomes. More studies e.g. in gnotobiotic mice are necessary to understand specific functions of members of the microbial network.

The long-term overarching goal would be to understand the detailed effects of CS on the microbial composition and what consequences these effects have on lung health. That knowledge could be an essential prerequisite for eventual therapeutic agents in a treatment of CLDs.

6 Bibliography

1. Ionescu CM. The human respiratory system : an analysis of the interplay between anatomy, structure, breathing and fractal dynamics. Springer; 2013.

2. Thiesse J, Namati E, Sieren JC, et al. Lung structure phenotype variation in inbred mouse strains revealed through in vivo micro-CT imaging. *J Appl Physiol* 2010;109(6):1960–1968.

3. Chao C-M, Agha E El, Tiozzo C, Minoo P, Bellusci S. A Breath of Fresh Air on the Mesenchyme: Impact of Impaired Mesenchymal Development on the Pathogenesis of Bronchopulmonary Dysplasia. *Front Med* 2015;2(APR):1.

4. Valerius K-P. Size-dependent morphology of the conductive bronchial tree in four species of myomorph rodents. *J Morphol* 1996;230(3):291–297.

5. Sato S, Bartolák-Suki E, Parameswaran H, Hamakawa H, Suki B. Scale dependence of structure-function relationship in the emphysematous mouse lung. *Front Physiol* 2015;6:146.

6. Iwasaki A, Foxman EF, Molony RD. Early local immune defences in the respiratory tract. *Nat Rev Immunol* 2017;17(1):7–20.

7. Jacquet A. Interactions of airway epithelium with protease allergens in the allergic response. *Clin Exp Allergy* 2011;41(3):305–311.

8. Rock JR, Onaitis MW, Rawlins EL, et al. Basal cells as stem cells of the mouse trachea and human airway epithelium. *Proc Natl Acad Sci* 2009;106(31):12771–12775.

9. Rackley CR, Stripp BR. Building and maintaining the epithelium of the lung. *J Clin Invest* 2012;122(8):2724–2730.

10. Fehrenbach H. Alveolar epithelial type II cell: Defender of the alveolus revisited. Respir. Res. 2001;2(1):33–46.

11. Whitsett JA, Kalin T V., Xu Y, Kalinichenko V V. Building and Regenerating the Lung Cell by Cell. *Physiol Rev* 2019;99(1):513–554.

12. Ayres JS. Cooperative Microbial Tolerance Behaviors in Host-Microbiota Mutualism. *Cell*

2016;165(6):1323–1331.

13. Sonnenburg ED, Smits SA, Tikhonov M, Higginbottom SK, Wingreen NS, Sonnenburg JL. Diet-induced extinctions in the gut microbiota compound over generations. *Nature* 2016;529(7585):212–215.

14. Schwarzer M, Makki K, Storelli G, et al. Lactobacillus plantarum strain maintains growth of infant mice during chronic undernutrition. *Science (80-)* 2016;351(6275):854–857.

15. Subramanian S, Huq S, Yatsunenko T, et al. Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* 2014;510(7505):417–421.

16. Russell SL, Gold MJ, Hartmann M, et al. Early life antibiotic-driven changes in microbiota enhance susceptibility to allergic asthma. *EMBO Rep* 2012;13(5):440–447.

17. Dickson RP, Huffnagle GB. The Lung Microbiome: New Principles for Respiratory Bacteriology in Health and Disease. *PLoS Pathog* 2015;11(7).

18. Huffnagle GB, Dickson RP, Lukacs NW. The respiratory tract microbiome and lung inflammation: A two-way street. *Mucosal Immunol* 2017;10(2):299–306.

19. Enaud R, Prevel R, Ciarlo E, et al. The Gut-Lung Axis in Health and Respiratory Diseases: A Place for Inter-Organ and Inter-Kingdom Crosstalks. *Front Cell Infect Microbiol* 2020;10(9).

20. Dickson RP, Erb-Downward JR, Freeman CM, et al. Spatial variation in the healthy human lung microbiome and the adapted island model of lung biogeography. *Ann Am Thorac Soc* 2015;12(6):821–830.

21. Charlson ES, Bittinger K, Haas AR, et al. Topographical continuity of bacterial populations in the healthy human respiratory tract. *Am J Respir Crit Care Med* 2011;184(8):957–963.

22. Segal LN, Clemente JC, Tsay JCJ, et al. Enrichment of the lung microbiome with oral taxa is associated with lung inflammation of a Th17 phenotype. *Nat Microbiol* 2016;1(5):16031.

23. Pragman AA, Lyu T, Baller JA, et al. The lung tissue microbiota of mild and moderate chronic obstructive pulmonary disease. *Microbiome* 2018;6(1):7.

24. Gleeson K, Eggli DF, Maxwell SL. Quantitative aspiration during sleep in normal subjects. *Chest* 1997;111(5):1266–1272.

25. Huxley EJ, Viroslav J, Gray WR, Pierce AK. Pharyngeal aspiration in normal adults and patients with depressed consciousness. *Am J Med* 1978;64(4):564–568.

26. Field SK, Underwood M, Brant R, Cowie RL. Prevalence of gastroesophageal reflux symptoms in asthma. *Chest* 1996;109(2):316–322.

27. Koh WJ, Lee JH, Kwon YS, et al. Prevalence of gastroesophageal reflux disease in patients with nontuberculous mycobacterial lung disease. *Chest* 2007;131(6):1825–1830.

28. Maqbool A, Pauwels A. Cystic Fibrosis and gastroesophageal reflux disease. *J Cyst Fibros* 2017;16:S2–S13.

29. 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.

30. 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.

31. Venkataraman A, Bassis CM, Beck JM, et al. Application of a neutral community model to assess structuring of the human lung microbiome. *MBio* 2015;6(1):e02284-14.

32. Tang Rasmussen T, Kirkeby LP, Poulsen K, Reinholdt J, Kilian M. Resident aerobic microbiota of the adult human nasal cavityNote. *APMIS* 2000;108(10):663–675.

33. Yadava K, Pattaroni C, Sichelstiel AK, et al. Microbiota promotes chronic pulmonary inflammation by enhancing IL-17A and autoantibodies. *Am J Respir Crit Care Med* 2016;193(9):975–987.

34. Waki N, Yajima N, Suganuma H, et al. Oral administration of Lactobacillus brevis KB290 to mice alleviates clinical symptoms following influenza virus infection. *Lett Appl Microbiol* 2014;58(1):87–93.

35. Kawase M, He F, Kubota A, Yoda K, Miyazawa K, Hiramatsu M. Heat-killed Lactobacillus gasseri TMC0356 protects mice against influenza virus infection by stimulating gut and

respiratory immune responses. *FEMS Immunol Med Microbiol* 2012;64(2):280–288.

36. Jung YJ, Lee YT, Ngo V Le, et al. Heat-killed Lactobacillus casei confers broad protection against influenza A virus primary infection and develops heterosubtypic immunity against future secondary infection. *Sci Rep* 2017;7(1):1–12.

37. Khailova L, Petrie B, Baird CH, Dominguez Rieg JA, Wischmeyer PE. Lactobacillus rhamnosus GG and Bifidobacterium longum attenuate lung injury and inflammatory response in experimental sepsis. *PLoS One* 2014;9(5).

38. Sagar S, Morgan ME, Chen S, et al. Bifidobacterium breve and Lactobacillus rhamnosus treatment is as effective as budesonide at reducing inflammation in a murine model for chronic asthma. *Respir Res* 2014;15(1):46.

39. Kawahara T, Takahashi T, Oishi K, et al. Consecutive oral administration of Bifidobacterium longum MM-2 improves the defense system against influenza virus infection by enhancing natural killer cell activity in a murine model. *Microbiol Immunol* 2015;59(1):1–12.

40. MacSharry J, O'Mahony C, Shalaby KH, et al. Immunomodulatory effects of feeding with Bifidobacterium longum on allergen-induced lung inflammation in the mouse. *Pulm Pharmacol Ther* 2012;25(4):325–334.

41. Fyhrquist N, Ruokolainen L, Suomalainen A, et al. Acinetobacter species in the skin microbiota protect against allergic sensitization and inflammation. *J Allergy Clin Immunol* 2014;134(6):1301-1309.e11.

42. Mika M, Mack I, Korten I, et al. Dynamics of the nasal microbiota in infancy: A prospective cohort study. *J Allergy Clin Immunol* 2015;135(4):905-912.e11.

43. SM T, D M, K P, et al. The Infant Nasopharyngeal Microbiome Impacts Severity of Lower Respiratory Infection and Risk of Asthma Development. *Cell Host Microbe* 2015;17(5).

44. Stearns JC, Davidson CJ, McKeon S, et al. Culture and molecular-based profiles show shifts in bacterial communities of the upper respiratory tract that occur with age. *ISME J* 2015;9(5):1246–1259.

45. Steenhuijsen Piters WAA de, Huijskens EGW, Wyllie AL, et al. Dysbiosis of upper respiratory tract microbiota in elderly pneumonia patients. *ISME J* 2016;10(1):97–108.

46. Hilty M, Burke C, Pedro H, et al. Disordered Microbial Communities in Asthmatic Airways. *PLoS One* 2010;5(1):e8578.

47. Segal LN, Alekseyenko A V., Clemente JC, et al. Enrichment of lung microbiome with supraglottic taxa is associated with increased pulmonary inflammation. *Microbiome* 2013;1(1):19.

48. Singh N, Vats A, Sharma A, Arora A, Kumar A. The development of lower respiratory tract microbiome in mice. *Microbiome* 2017;5(1):61.

49. Kostric M, Milger K, Krauss-Etschmann S, et al. Development of a Stable Lung Microbiome in Healthy Neonatal Mice. *Microb Ecol* 2018;75(2):529–542.

50. Klingenberg Barfod K, Roggenbuck M, Hansen LH, et al. The murine lung microbiome in relation to the intestinal and vaginal bacterial communities. *BMC Microbiol* 2013;13(303):303.

51. Man WH, Steenhuijsen Piters WAA de, Bogaert D. The microbiota of the respiratory tract: gatekeeper to respiratory health. *Nat Rev Microbiol* 2017;15.

52. Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. The Microbiome and the Respiratory Tract. *Annu Rev Physiol* 2016;78(1):481–504.

53. Faust K, Sathirapongsasuti JF, Izard J, et al. Microbial Co-occurrence Relationships in the Human Microbiome. *PLoS Comput Biol* 2012;8(7):e1002606.

54. Pragman AA, Kim HB, Reilly CS, Wendt C, Isaacson RE. The Lung Microbiome in Moderate and Severe Chronic Obstructive Pulmonary Disease. *PLoS One* 2012;7(10):e47305.

55. Turturice BA, McGee HS, Oliver B, et al. Atopic asthmatic immune phenotypes associated with airway microbiota and airway obstruction. *PLoS One* 2017;12(10):e0184566.

56. Huang YJ, Kim E, Cox MJ, et al. A Persistent and Diverse Airway Microbiota Present during Chronic Obstructive Pulmonary Disease Exacerbations. *Omi A J Integr Biol* 2010;14.

57. Erb-Downward JR, Thompson DL, Han MK, et al. Analysis of the lung microbiome in the "healthy" smoker and in COPD. *PLoS One* 2011;6(2).

58. Mayhew D, Devos N, Lambert C, et al. Longitudinal profiling of the lung microbiome in the AERIS study demonstrates repeatability of bacterial and eosinophilic COPD exacerbations. *Thorax* 2018;0:1–9.

59. Sze MA, Dimitriu PA, Hayashi S, et al. The lung tissue microbiome in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2012;185(10):1073–1080.

60. Diao W, Shen N, Du Y, et al. Symptom-related sputum microbiota in stable chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis* 2018;13:2289–2299.

61. Millares L, Ferrari R, Gallego M, et al. Bronchial microbiome of severe COPD patients colonised by Pseudomonas aeruginosa. *Eur J Clin Microbiol Infect Dis* 2014;33(7):1101–1111.

62. Einarsson GG, Comer DM, Mcilreavey L, et al. Community dynamics and the lower airway microbiota in stable chronic obstructive pulmonary disease, smokers and healthy non-smokers. *Thorax* 2016;71:795–803.

63. Huang YJ, Sethi S, Murphy T, Nariya S, Boushey HA, Lynch S V. Airway microbiome dynamics in exacerbations of chronic obstructive pulmonary disease. *PLoS One* 2010;5:e8578.

64. Wang Z, Bafadhel M, Haldar K, et al. Lung microbiome dynamics in COPD exacerbations. *Eur Respir J* 2016;47(4):1082–92.

65. Zhang Q, Cox M, Liang Z, et al. Airway microbiota in severe asthma and relationship to asthma severity and phenotypes. *PLoS One* 2016;11(4).

66. Simpson JL, Daly J, Baines KJ, et al. Airway dysbiosis: Haemophilus influenza and Tropheryma in poorly controlled asthma. *Eur Respir J* 2016;47(3):792–800.

67. Huang YJ, Nelson CE, Brodie EL, et al. Airway microbiota and bronchial hyperresponsiveness in patients with suboptimally controlled asthma. *J Allergy Clin Immunol* 2011;127(2).

68. Goleva E, Jackson LP, Harris JK, et al. The effects of airway microbiome on corticosteroid responsiveness in asthma. *Am J Respir Crit Care Med* 2013;188(10):1193–1201.

69. Taylor SL, Leong LEX, Choo JM, et al. Inflammatory phenotypes in patients with severe asthma are associated with distinct airway microbiology. *J Allergy Clin Immunol* 2018;141(1):94-103.e15.

70. Amaral FA, Sachs D, Costa V V, et al. Commensal microbiota is fundamental for the development of inflammatory pain. *Proc Natl Acad Sci* 2008;105(6):2193–2197.

71. Molyneaux PL, Willis-Owen SAG, Cox MJ, et al. Host-microbial interactions in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2017;

72. Boris S, Suárez JE, Vázquez F, Barbés C. Adherence of Human Vaginal Lactobacilli to Vaginal Epithelial Cells and Interaction with Uropathogens. *Infect Immun* 1998;66(5):1985–1989.

73. Herbst T, Sichelstiel A, Schär C, et al. Dysregulation of Allergic Airway Inflammation in the Absence of Microbial Colonization. *Am J Respir Crit Care Med* 2011;184(2):198–205.

74. Artis D. Epithelial-cell recognition of commensal bacteria and maintenance of immune homeostasis in the gut. *Nat Rev Immunol* 2008;8(6):411–420.

75. Martin C, Burgel PR, Lepage P, et al. Host–microbe interactions in distal airways: Relevance to chronic airway diseases. Eur. Respir. Rev. 2015;78–91.

76. Karczewski J, Troost FJ, Konings I, et al. Regulation of human epithelial tight junction proteins by Lactobacillus plantarum in vivo and protective effects on the epithelial barrier. *Am J Physiol - Gastrointest Liver Physiol* 2010;298(6).

77. Resta-Lenert S, Barrett KE. Probiotics and commensals reverse TNF-α- and IFN-γ-induced dysfunction in human intestinal epithelial cells. *Gastroenterology* 2006;130(3):731–746.

78. Zyrek AA, Cichon C, Helms S, Enders C, Sonnenborn U, Schmidt MA. Molecular mechanisms underlying the probiotic effects of Escherichia coli Nissle 1917 involve ZO-2 and PKC? redistribution resulting in tight junction and epithelial barrier repair. *Cell Microbiol* 2007;9(3):804–816.

79. Ewaschuk JB, Diaz H, Meddings L, et al. Secreted bioactive factors from Bifidobacterium infantis enhance epithelial cell barrier function. *Am J Physiol - Gastrointest Liver Physiol* 2008;295(5).

80. Donaldson GP, Lee SM, Mazmanian SK. Gut biogeography of the bacterial microbiota. *Nat Rev Microbiol* 2016;14(1):20–32.

81. Kirjavainen P V, Ouwehand AC, Isolauri E, Salminen SJ. The ability of probiotic bacteria to bind to human intestinal mucus. *FEMS Microbiol Lett* 1998;167(2):185–189.

82. Werlang C, Cárcarmo-Oyarce G, Ribbeck K. Engineering mucus to study and influence the microbiome. *Nat Rev Mater* 2019;4(2):134–145.

83. Flynn JM, Niccum D, Dunitz JM, Hunter RC. Evidence and Role for Bacterial Mucin Degradation in Cystic Fibrosis Airway Disease. *PLOS Pathog* 2016;12(8):e1005846.

84. ERS White Book. Chapter 9: Tobacco smoking.

85. Churg A, Dai J, Tai H, Xie C, Wright JL. Tumor Necrosis Factor-α Is Central to Acute Cigarette Smoke–induced Inflammation and Connective Tissue Breakdown. *Am J Respir Crit Care Med* 2002;166(6):849–854.

86. Cozen W, Diaz-Sanchez D, Gauderman WJ, et al. Th1 and Th2 cytokines and IgE levels in identical twins with varying levels of cigarette consumption. *J Clin Immunol* 2004;24(6):617–622.

87. Torres de Heens GL, Velden U van der, Loos BG. Cigarette smoking enhances T cell activation and a Th2 immune response; an aspect of the pathophysiology in periodontal disease. *Cytokine* 2009;47(3):157–161.

88. Leopold PL, O'Mahony MJ, Lian XJ, Tilley AE, Harvey B-G, Crystal RG. Smoking Is Associated with Shortened Airway Cilia. *PLoS One* 2009;4(12):e8157.

89. Shah SA, Ganesan SM, Varadharaj S, Dabdoub SM, Walters JD, Kumar PS. The making of a miscreant: tobacco smoke and the creation of pathogen-rich biofilms. *npj Biofilms Microbiomes* 2017;3(1):26.

90. Wu J, Peters BA, Dominianni C, et al. Cigarette smoking and the oral microbiome in a

large study of American adults. *ISME J* 2016;10:2435–2446.

91. Stewart CJ, Auchtung TA, Ajami NJ, et al. Effects of tobacco smoke and electronic cigarette vapor exposure on the oral and gut microbiota in humans: A pilot study. *PeerJ* 2018;6:e4693.

92. Yu G, Phillips S, Gail MH, et al. The effect of cigarette smoking on the oral and nasal microbiota. *Microbiome* 2017;5(1):3.

93. Charlson ES, Chen J, Custers-Allen R, et al. Disordered Microbial Communities in the Upper Respiratory Tract of Cigarette Smokers. *PLoS One* 2010;5(12):e15216.

94. R Development Core Team, Vienna, A. R: A language and environment for statistical computing. R Foundation for Statistical Computing.
Available from: https://www.r-project.org/

95. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* 2013;8(4):e61217.

96. Lagkouvardos I, Fischer S, Kumar N, Clavel T. Rhea: A transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons. *PeerJ* 2017;2017(1):e2836.

97. Cole JR, Wang Q, Fish JA, et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 2014;42(D1):D633–D642.

98. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;26(19):2460–2461.

99. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: A versatile open source tool for metagenomics. *PeerJ*.
Available from: https://github.com/torognes/vsearch

100. NHIS - Adult Tobacco Use - Glossary.
Available from: https://www.cdc.gov/nchs/nhis/tobacco/tobacco_glossary.htm

101. Häussinger K, Ballin A, Becker H, et al. Empfehlungen zur Sicherung der Qualität in der Bronchoskopie. *Pneumologie* 2004;58(05):344–356.

102. Marsh RL, Nelson MT, Pope CE, et al. How low can we go? The implications of low bacterial load in respiratory microbiota studies. *Pneumonia* 2018;10(1):7.

103. Pei AY, Oberdorf WE, Nossa CW, et al. Diversity of 16S rRNA genes within individual prokaryotic genomes. *Appl Environ Microbiol* 2010;76(12):3886–3897.

104. Johnson JS, Spakowicz DJ, Hong BY, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun* 2019;10(1):1–11.

105. Nossa CW, Oberdorf WE, Yang L, et al. Design of 16S rRNA gene primers for 454 pyrosequencing of the human foregut microbiome. *World J Gastroenterol* 2010;16(33):4135–44.

106. Illumina. 16S Metagenomic Sequencing Library Preparation. Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System.

107. FASPA.github.io/README.md at master · StefanPfeiffer80/FASPA.github.io · GitHub Available from: https://github.com/StefanPfeiffer80/FASPA.github.io/blob/master/README.md

108. Edgar RC. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing.

109. Edgar R. UNCROSS: Filtering of high-frequency cross-talk in 16S amplicon reads. *bioRxiv* 2016;088666.

110. Wright ES, Vetsigian KH. Quality filtering of Illumina index reads mitigates sample cross-talk. *BMC Genomics* 2016;17(1):876.

111. Maidak BL, Cole JR, Lilburn TG, et al. The RDP-II (Ribosomal Database Project). 2001.

112. Edgar RC. Accuracy of taxonomy prediction for 16S rRNA and fungal ITS sequences. *PeerJ* 2018;6(4):e4652.

113. Nucleotide BLAST: Search nucleotide databases using a nucleotide query. Available from: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch

114. Salter SJ, Cox MJ, Turek EM, et al. Reagent and laboratory contamination can critically

impact sequence-based microbiome analyses. *BMC Biol* 2014;12(1):87.

115. Eren AM, Maignien L, Sul WJ, et al. Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol Evol* 2013;4(12):1111–1119.

116. Mcmanus JFA. Histological and histochemical uses of periodic acid. *Biotech Histochem* 1948;23(3):99–108.

117. Wolff M, Danov O, Bartel S, et al. Cigarette Smoke Affects Dendritic Cell Populations, Epithelial Barrier Function and the Immune Response to Viral Infection with H1N1. *Front Med* 2020;7:605.

118. Adams JD, O'mara-adams KJ, Hoffmann D. Toxic and carcinogenic agents in undiluted mainstream smoke and sidestream smoke of different types of cigarettes. *Carcinogenesis* 1987;8(5):729–731.

119. Dickson RP, Erb-Downward JR, Freeman CM, et al. Bacterial Topography of the Healthy Human Lower Respiratory Tract. *MBio* 2017;8(1):e02287-16.

120. Ziebuhr W, Hennig S, Eckart M, Kränzler H, Batzilla C, Kozitskaya S. Nosocomial infections by Staphylococcus epidermidis: how a commensal bacterium turns into a pathogen. *Int J Antimicrob Agents* 2006;28(SUPPL. 1):14–20.

121. Kim HJ, Jo A, Jeon YJ, et al. Nasal commensal Staphylococcus epidermidis enhances interferon-λ-dependent immunity against influenza virus. *Microbiome* 2019;7(1):80.

122. Kloos WE, Musselwhite MS. Distribution and Persistence of Staphylococcus and Micrococcus Species and Other Aerobic Bacteria on Human Skin1. *Appl Microbiol* 1975;30(3):381–395.

123. Wu Y, Ma Y, Xu T, et al. Nicotine Enhances Staphylococcus epidermidis Biofilm Formation by Altering the Bacterial Autolysis, Extracellular DNA Releasing, and Polysaccharide Intercellular Adhesin Production. *Front Microbiol* 2018;9:2575.

124. Charlson ES, Chen J, Custers-Allen R, et al. Disordered Microbial Communities in the Upper Respiratory Tract of Cigarette Smokers. *PLoS One* 2010;5(12):e15216.

125. Ramakrishnan VR, Feaze LM, Gitomer SA, Ir D, Robertson CE, Frank DN. The microbiome

of the middle meatus in healthy adults. *PLoS One* 2013;8(12):85507.

126. Bosch AATM, Levin E, Houten MA van, et al. Development of Upper Respiratory Tract Microbiota in Infancy is Affected by Mode of Delivery. *EBioMedicine* 2016;9:336–345.

127. Biesbroek G, Tsivtsivadze E, Sanders EAM, et al. Early respiratory microbiota composition determines bacterial succession patterns and respiratory health in children. *Am J Respir Crit Care Med* 2014;190(11):1283–1292.

128. Yu G, Phillips S, Gail MH, et al. The effect of cigarette smoking on the oral and nasal microbiota. *Microbiome* 2017;5(1):3.

129. Wu J, Peters BA, Dominianni C, et al. Cigarette smoking and the oral microbiome in a large study of American adults. *ISME J* 2016;10(10):2435–2446.

130. Heijink IH, Brandenburg SM, Postma DS, Oosterhout AJM Van. Cigarette smoke impairs airway epithelial barrier function and cell–cell contact recovery.

131. Berdah L, Taytard J, Leyronnas S, Clement A, Boelle PY, Corvol H. Stenotrophomonas maltophilia: A marker of lung disease severity. *Pediatr Pulmonol* 2018;53(4):426–430.

132. K C, S V, A G. Stenotrophomonas maltophilia in Lower Respiratory Tract Infections. *J Clin DIAGNOSTIC Res* 2014;8(12).

133. Pauly JL, Waight JD, Paszkiewicz GM. Tobacco flakes on cigarette filters grow bacteria: A potential health risk to the smoker? *Tob Control* 2008;17(SUPPL. 1):i49–i52.

134. Zhou Y, Lin F, Cui Z, et al. Correlation between either Cupriavidus or Porphyromonas and primary pulmonary tuberculosis found by analysing the microbiota in patients' bronchoalveolar lavage fluid. *PLoS One* 2015;10(5).

135. Jain AL, Harding CM, Assani K, et al. Characteristics of invasive Acinetobacter species isolates recovered in a pediatric academic center. *BMC Infect Dis* 2016;16(1).

136. Bernasconi E, Pattaroni C, Koutsokera A, et al. Airway Microbiota Determines Innate Cell Inflammatory or Tissue Remodeling Profiles in Lung Transplantation. *Am J Respir Crit Care Med* 2016;194(10):1252–1263.

137. Tobacco.
Available from: https://www.who.int/news-room/fact-sheets/detail/tobacco

138. Ma Z (Sam), Li W. How and Why Men and Women Differ in Their Microbiomes: Medical Ecology and Network Analyses of the Microgenderome. *Adv Sci* 2019;6(23):1902054.

139. Haro C, Rangel-Zúñiga OA, Alcalá-Díaz JF, et al. Intestinal Microbiota Is Influenced by Gender and Body Mass Index. *PLoS One* 2016;11(5):e0154090.

140. Hepworth ML, Passey SL, Seow HJ, Vlahos R. Losartan does not inhibit cigarette smoke-induced lung inflammation in mice. *Sci Rep* 2019;9(1):1–11.

141. Zhang R, Chen L, Cao L, et al. Effects of smoking on the lower respiratory tract microbiome in mice. *Respir Res* 2018;19(1).

142. Chen H, Vlahos R, Bozinovski S, Jones J, Anderson GP, Morris MJ. Effect of short-term cigarette smoke exposure on body weight, appetite and brain neuropeptide Y in mice. *Neuropsychopharmacology* 2005;30(4):713–719.

143. Chen H, Hansen MJ, Jones JE, et al. Cigarette smoke exposure reprograms the hypothalamic neuropeptide Y axis to promote weight loss. *Am J Respir Crit Care Med* 2006;173(11):1248–1254.

144. Hanspeter W, Imelda Espiritu, Stephanie T. Dance MSM. A mouse lung tumor model of tobacco smoke carcinogenesis. *Toxicol Sci* 2002;68(2):322–30.

145. Lee KM, Hoeng J, Harbo S, et al. Biological changes in C57BL/6 mice following 3 weeks of inhalation exposure to cigarette smoke or e-vapor aerosols. *Inhal Toxicol* 2018;30(13–14):553–567.

146. Eijl S Van, Oorschot R Van, Olivier B, Nijkamp F, Bloksma N. Stress and hypothermia in mice in a nose-only cigarette smoke exposure system. *Inhal Toxicol* 2006;18(11):911–918.

147. Seccareccia F, Zuccaro P, Pacifici R, et al. Serum cotinine as a marker of environmental tobacco smoke exposure in epidemiological studies: The experience of the MATISS project. *Eur J Epidemiol* 2003;18(6):487–492.

148. Moreno-Gonzalez I, Estrada LD, Sanchez-Mejias E, Soto C. Smoking exacerbates amyloid pathology in a mouse model of Alzheimer's disease. *Nat Commun* 2013;4(1):1–10.

149. Shiverick KT. Cigarette Smoking and Reproductive and Developmental Toxicity. In: Reproductive and Developmental Toxicology. Elsevier Inc.; 2011. p. 319–331.

150. Ye W, Chen R, Chen X, et al. AhR regulates the expression of human cytochrome P450 1A1 (CYP1A1) by recruiting Sp1. *FEBS J* 2019;286(21):4215–4231.

151. Meyer KC. Bronchoalveolar lavage as a diagnostic tool. *Semin Respir Crit Care Med* 2007;28(5):546–560.

152. Hunninghake GW, Gadek JE, Kawanami O, Ferrans VJ, Crystal RG. Inflammatory and immune processes in the human lung in health and disease: Evaluation by bronchoalveolar lavage. *Am J Pathol* 1979;97(1):149–205.

153. Hunninghake GW, Crystal RG. Cigarette smoking and lung destruction. Accumulation of neutrophils in the lungs of cigarette smokers. *Am Rev Respir Dis* 1983;128(5):833–838.

154. Gairola CG. Free lung cell response of mice and rats to mainstream cigarette smoke exposure. *Toxicol Appl Pharmacol* 1986;84(3):567–575.

155. Karimi R, Tornling G, Grunewald J, Eklund A, Sköld CM. Cell recovery in bronchoalveolar lavage fluid in smokers is dependent on cumulative smoking history. PLoS One. 2012;7(3).

156. Vlahos R, Bozinovski S, Jones JE, et al. Differential protease, innate immunity, and NF-κB induction profiles during lung inflammation induced by subchronic cigarette smoke exposure in mice. *Am J Physiol - Lung Cell Mol Physiol* 2006;290(5):931–945.

157. Talbot M, Hamel-Auger M, Beaulieu MJ, et al. Impact of immunization against OxLDL on the pulmonary response to cigarette smoke exposure in mice. *Respir Res* 2018;19(1):131.

158. Botelho FM, Gaschler GJ, Kianpour S, et al. Innate Immune Processes Are Sufficient for Driving Cigarette Smoke–Induced Inflammation in Mice. *Am J Respir Cell Mol Biol* 2010;42(4):394–403.

159. Morris A, Kinnear G, Wan WYH, Wyss D, Bahra P, Stevenson CS. Comparison of cigarette

smoke-induced acute inflammation in multiple strains of mice and the effect of a matrix metalloproteinase inhibitor on these responses. *J Pharmacol Exp Ther* 2008;327(3):851–862.

160. Saetta M, Stefano A Di, Turato G, et al. CD8+ T-lymphocytes in peripheral airways of smokers with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998;157(3 PART I):822–826.

161. O'Shaughnessy TC, Ansari TW, Barnes NC, Jeffery PK. Inflammation in bronchial biopsies of subjects with chronic bronchitis: inverse relationship of CD8+ T lymphocytes with FEV1. *Am J Respir Crit Care Med* 1997;155(3):852–857.

162. Tollerud DJ, Clark JW, Brown LM, et al. The Effects of Cigarette Smoking on T Cell Subsets: A Population-based Survey of Healthy Caucasians. *Am Rev Respir Dis* 1989;139(6):1446–1451.

163. Ginns LC, Goldenheim PD, Miller LG, et al. T-lymphocyte subsets in smoking and lung cancer. Analysis by monoclonal antibodies and flow cytometry. *Am Rev Respir Dis* 1982;126(2):265–269.

164. Miller LG, Goldstein G, Murphy M, Ginns LC. Reversible alterations in immunoregulatory T cells in smoking. Analysis by monoclonal antibodies and flow cytometry. *Chest* 1982;82(5):526–529.

165. Hughes DA, Haslam PL, Townsend PJ, Turner-Warwick M. Numerical and functional alterations in circulatory lymphocytes in cigarette smokers. *Clin Exp Immunol* 1985;61(2):459–66.

166. John G, Kohse K, Orasche J, et al. The composition of cigarette smoke determines inflammatory cell recruitment to the lung in COPD mouse models. *Clin Sci* 2014;126(3):207–221.

167. Kuschner WG, D'Alessandro A, Wong H, Blanc PD. Dose-dependent cigarette smoking-related inflammatory responses in healthy adults. *Eur Respir J* 1996;9(10):1989–1994.

168. Lee JH, Hailey KL, Vitorino SA, Jennings PA, Bigby TD, Breen EC. Cigarette smoke triggers IL-33-associated inflammation in a model of late-stage chronic obstructive pulmonary

disease. *Am J Respir Cell Mol Biol* 2019;61(5):567–574.

169. Kontoyiannis D, Pasparakis M, Pizarro TT, Cominelli F, Kollias G. Impaired on/off regulation of TNF biosynthesis in mice lacking TNF AU- rich elements: Implications for joint and gut-associated immunopathologies. *Immunity* 1999;10(3):387–398.

170. Kojouharoff G, Hans W, Obermeler F, et al. Neutralization of tumour necrosis factor (TNF) but not of IL-1 reduces inflammation in chronic dextran sulphate sodium-induced colitis in mice. *Clin Exp Immunol* 1997;107(2):353–358.

171. Zhou J, Kawai T, Yu Q. Pathogenic role of endogenous TNF-α in the development of Sjögren's-like sialadenitis and secretory dysfunction in non-obese diabetic mice. *Lab Investig* 2017;97(4):458–467.

172. Furtado GC, Pacer ME, Bongers G, et al. TNFa-dependent development of lymphoid tissue in the absence of RORgt lymphoid tissue inducer cells. *Mucosal Immunol* 2014;7(3):602–614.

173. Petersen SL, Wang L, Yalcin-Chin A, et al. Autocrine TNFα Signaling Renders Human Cancer Cells Susceptible to Smac-Mimetic-Induced Apoptosis. *Cancer Cell* 2007;12(5):445–456.

174. Carswell EA, Old LJ, Kassel RL, Green S, Fiore N, Williamson B. An endotoxin induced serum factor that cuases necrosis of tumors. *Proc Natl Acad Sci U S A* 1975;72(9):3666–3670.

175. Gonçalves NS, Ghaem-Maghami M, Monteleone G, et al. Critical role for tumor necrosis factor alpha in controlling the number of lumenal pathogenic bacteria and immunopathology in infectious colitis. *Infect Immun* 2001;69(11):6651–6659.

176. Kindler V, Sappino AP, Grau GE, Piguet PF, Vassalli P. The inducing role of tumor necrosis factor in the development of bactericidal granulomas during BCG infection. *Cell* 1989;56(5):731–740.

177. Flynn JAL, Goldstein MM, Chan J, et al. Tumor necrosis factor-α is required in the protective immune response against mycobacterium tuberculosis in mice. *Immunity* 1995;2(6):561–572.

178. Parameswaran N, Patial S. Tumor necrosis factor-a signaling in macrophages. Crit. Rev. Eukaryot. Gene Expr. 2010;20(2):87–103.

179. Aggarwal BB. Signalling pathways of the TNF superfamily: A double-edged sword. Nat. Rev. Immunol. 2003;3(9):745–756.

180. Velazquez-Salinas L, Pauszek SJ, Stenfeldt C, et al. Increased virulence of an epidemic strain of vesicular stomatitis virus is associated with interference of the innate response in pigs. *Front Microbiol* 2018;9(AUG):1891.

181. Kopf M, Baumann H, Freer G, et al. Impaired immune and acute-phase responses in interleukin-6-deficient mice. *Nature* 1994;368(6469):339–342.

182. Naseem S, Manzoor S, Javed A, Abbas S. Interleukin-6 Rescues Lymphocyte from Apoptosis and Exhaustion Induced by Chronic Hepatitis C Virus Infection. *Viral Immunol* 2018;31(9):624–631.

183. Yang Y, Ochando J, Yopp A, Bromberg JS, Ding Y. IL-6 Plays a Unique Role in Initiating c-Maf Expression during Early Stage of CD4 T Cell Activation. *J Immunol* 2005;174(5):2720–2729.

184. Rosser EC, Oleinika K, Tonon S, et al. Regulatory B cells are induced by gut microbiota–driven interleukin-1β and interleukin-6 production. *Nat Med* 2014;20(11):1334–1339.

185. Kishimoto T. INTERLEUKIN-6: From Basic Science to Medicine—40 Years in Immunology. *Annu Rev Immunol* 2005;23(1):1–21.

186. Matsuki T, Nakae S, Sudo K, Horai R, Iwakura Y. Abnormal T cell activation caused by the imbalance of the IL-1/IL-1R antagonist system is responsible for the development of experimental autoimmune encephalomyelitis | International Immunology | Oxford Academic. *Int Immunol* 2006;18(2):399–407.

187. Pascual V, Allantaz F, Arce E, Punaro M, Banchereau J. Role of interleukin-1 (IL-1) in the pathogenesis of systemic onset juvenile idiopathic arthritis and clinical response to IL-1 blockade. *J Exp Med* 2005;201(9):1479–1486.

188. Liew FY, Girard J-P, Turnquist HR. Interleukin-33 in health and disease. *Nat Rev Immunol*

2016;16(11):676–689.

189. Shibata Y, Berclaz P-Y, Chroneos ZC, Yoshida M, Whitsett JA, Trapnell BC. GM-CSF Regulates Alveolar Macrophage Differentiation and Innate Immunity in the Lung through PU.1. *Immunity* 2001;15(4):557–567.

190. Martinez FO, Gordon S, Locati M, Mantovani A. Transcriptional Profiling of the Human Monocyte-to-Macrophage Differentiation and Polarization: New Molecules and Patterns of Gene Expression. *J Immunol* 2006;177(10):7303–7311.

191. Cheung DS, Ehlenbach SJ, Kitchens RT, et al. Cutting Edge: CD49d + Neutrophils Induce FcεRI Expression on Lung Dendritic Cells in a Mouse Model of Postviral Asthma. *J Immunol* 2010;185(9):4983–4987.

192. Gao N, Liu X, Wu J, et al. CXCL10 suppression of hem- and lymph-angiogenesis in inflamed corneas through MMP13. *Angiogenesis* 2017;20(4):505–518.

193. Jing H, Liu L, Zhou J, Yao H. Inhibition of C-X-C motif chemokine 10 (CXCL10) protects mice from cigarette smoke-induced chronic obstructive pulmonary disease. *Med Sci Monit* 2018;24:5748–5753.

194. Voelkel NF, Vandivier RW, Tuder RM. Vascular endothelial growth factor in the lung. *Am J Physiol Cell Mol Physiol* 2006;290(2):L209–L221.

195. Thatcher TH, McHugh NA, Egan RW, et al. Role of CXCR2 in cigarette smoke-induced lung inflammation. *Am J Physiol - Lung Cell Mol Physiol* 2005;289(2 33-2).

196. Sawant K V., Poluri KM, Dutta AK, et al. Chemokine CXCL1 mediated neutrophil recruitment: Role of glycosaminoglycan interactions. *Sci Rep* 2016;6.

197. Tan KW, Chong SZ, Wong FHS, et al. Neutrophils contribute to inflammatory lymphangiogenesis by increasing VEGF-A bioavailability and secreting VEGF-D. *Blood* 2013;122(22):3666–3677.

198. Gong Y, Koh DR. Neutrophils promote inflammatory angiogenesis via release of preformed VEGF in an in vivo corneal model. *Cell Tissue Res* 2010;339(2):437–448.

199. Schamberger AC, Mise N, Jia J, et al. Cigarette Smoke–Induced Disruption of Bronchial

Epithelial Tight Junctions Is Prevented by Transforming Growth Factor-β. *Am J Respir Cell Mol Biol* 2014;50(6):1040–1052.

200. Fransen F, Zagato E, Mazzini E, et al. BALB/c and C57BL/6 Mice Differ in Polyreactive IgA Abundance, which Impacts the Generation of Antigen-Specific IgA and Microbiota Diversity. *Immunity* 2015;43(3):527–540.

201. Madala SK, Pesce JT, Ramalingam TR, et al. Matrix Metalloproteinase 12-Deficiency Augments Extracellular Matrix Degrading Metalloproteinases and Attenuates IL-13–Dependent Fibrosis. *J Immunol* 2010;184(7):3955–3963.

202. Dean RA, Cox JH, Bellac CL, Doucet A, Starr AE, Overall CM. Macrophage-specific metalloelastase (MMP-12) truncates and inactivates ELR + CXC chemokines and generates CCL2, -7, -8, and -13 antagonists: Potential role of the macrophage in terminating polymorphonuclear leukocyte influx. *Blood* 2008;112(8):3455–3464.

203. Nénan S, Planquois J-M, Berna P, et al. Analysis of the inflammatory response induced by rhMMP-12 catalytic domain instilled in mouse airways. *Int Immunopharmacol* 2005;5(3):511–524.

204. Dufour A, Bellac CL, Eckhard U, et al. C-terminal truncation of IFN-γ inhibits proinflammatory macrophage responses and is deficient in autoimmune disease. *Nat Commun* 2018;9(1):1–18.

205. Churg A, Wang RD, Tai H, et al. Macrophage Metalloelastase Mediates Acute CigaretteSmoke–induced Inflammation via Tumor NecrosisFactor-α Release. *Am J Respir Crit Care Med V* 2002;167:1083–1089.

206. Houghton AMG, Quintero PA, Perkins DL, et al. Elastin fragments drive disease progression in a murine model of emphysema. *J Clin Invest* 2006;116(3):753–759.

207. Hautamaki RD, Kobayashi DK, Senior RM, Shapiro SD. Requirement for Macrophage Elastase for Cigarette Smoke-Induced Emphysema in Mice. *Science (80-)* 1997;277(5334):2002–2004.

208. Ilumets H, Rytilä P, Demedts I, et al. Matrix metalloproteinases -8, -9 and -12 in smokers and patients with Stage 0 COPD. *Int J COPD* 2007;2(3):369–379.

209. Babusyte A, Stravinskaite K, Jeroch J, Lötvall J, Sakalauskas R, Sitkauskiene B. Patterns of airway inflammation and MMP-12 expression in smokers and ex-smokers with COPD. *Respir Res* 2007;8(1):81.

210. Molet S, Belleguic C, Lena H, et al. Increase in macrophage elastase (MMP-12) in lungs from patients with chronic obstructive pulmonary disease. *Inflamm Res* 2005;54(1):31–36.

211. Lagkouvardos I, Pukall R, Abt B, et al. The Mouse Intestinal Bacterial Collection (miBC) provides host-specific insight into cultured diversity and functional potential of the gut microbiota. *Nat Microbiol* 2016;1(10):16131.

212. Tanca A, Manghina V, Fraumene C, et al. Metaproteogenomics Reveals Taxonomic and Functional Changes between Cecal and Fecal Microbiota in Mouse. *Front Microbiol* 2017;8(MAR):391.

7 Supplementary material

7.1 Supplementary results

	Core genus abundance/presence			Group comparison core genera[3]		
GENUS	**Site**	**<x>[1]**	**p/a[2]**	**Site 1**	**Site 2**	**FDH corrected p-value**
Actinomyces	BAL	2.00	0	BAL	NSW	< 0.001
Actinomyces	NSW	0.66	0	BAL	PSH	< 0.001
Actinomyces	**PHS**	**4.26**	**0**	**NSW**	**PSH**	**< 0.001**
Alloprevotella	**BAL**	**1.77**	**0**	**BAL**	**NSW**	**0.027**
Alloprevotella	NSW	0.04	14	BAL	PSH	0.1305
Alloprevotella	PHS	0.81	5	NSW	PSH	0.1305
Atopobium	BAL	0.92	0	BAL	NSW	< 0.001
Atopobium	**NSW**	**0.19**	**8**	**BAL**	**PSH**	
Atopobium	PHS	1.28	0	NSW	PSH	< 0.001
Campylobacter	**BAL**	**1.07**	**0**	**BAL**	**NSW**	**< 0.001**
Campylobacter	NSW	0.18	8	BAL	PSH	
Campylobacter	PHS	0.91	0	NSW	PSH	< 0.001
Corynebacterium	BAL	0.04	8	BAL	NSW	**< 0.001**
Corynebacterium	**NSW**	**23.93**	**0**	**BAL**	**PSH**	**1**
Corynebacterium	PHS	0.02	8	NSW	PSH	0,0769
Fusobacterium	**BAL**	**4.60**	**0**	**BAL**	**NSW**	**< 0.001**
Fusobacterium	NSW	0.27	5	BAL	PSH	0.2798
Fusobacterium	PHS	3.84	1	NSW	PSH	< 0.001
Gemella	**BAL**	**1.34**	**0**	**BAL**	**NSW**	**< 0.001**
Gemella	NSW	0.38	2	BAL	PSH	0,011
Gemella	PHS	0.65	0	NSW	PSH	< 0.001
Granulicatella	BAL	0.98	0	BAL	NSW	0.1249
Granulicatella	NSW	0.39	3	BAL	PSH	0.0806
Granulicatella	PHS	1.34	0	**NSW**	**PSH**	**0.0072**
Haemophilus	BAL	1.90	1	BAL	NSW	0.0928
Haemophilus	NSW	1.05	2	BAL	PSH	0.336
Haemophilus	PHS	2.85	0	NSW	PSH	**0.0474**
Lachnoanaerobaculum	BAL	0.31	2	BAL	NSW	0.6352
Lachnoanaerobaculum	NSW	0.15	8	BAL	PSH	< 0.001
Lachnoanaerobaculum	PHS	1.45	0	**NSW**	**PSH**	**0.0126**
Leptotrichia	BAL	2.76	1	BAL	NSW	< 0.001
Leptotrichia	NSW	0.20	3	BAL	PSH	0.0868
Leptotrichia	PHS	5.55	0	**NSW**	**PSH**	**< 0.001**
Megasphaera	**BAL**	**2.84**	**0**	**BAL**	**NSW**	**< 0.001**
Megasphaera	NSW	0.23	4	BAL	PSH	0.4379
Megasphaera	PHS	2.34	0	NSW	PSH	< 0.001
Porphyromonas	**BAL**	**1.69**	**1**	**BAL**	**NSW**	**< 0.001**
Porphyromonas	NSW	0.28	7	BAL	PSH	0.0122
Porphyromonas	PHS	1.05	1	NSW	PSH	0.03
Prevotella	**BAL**	**27.70**	**0**	**BAL**	**NSW**	**< 0.001**
Prevotella	NSW	2.39	0	BAL	PSH	0.1073
Prevotella	PHS	25.07	0	NSW	PSH	< 0.001
Staphylococcus	BAL	0.02	24	BAL	NSW	< 0.001
Staphylococcus	**NSW**	**24.16**	**0**	**BAL**	**PSH**	**NA**
Staphylococcus	PHS	0.00	35	NSW	PSH	NA
Streptococcus	BAL	15.30	0	BAL	NSW	< 0.001
Streptococcus	NSW	7.51	0	BAL	PSH	0.0546
Streptococcus	**PHS**	**17.04**	**0**	**NSW**	**PSH**	**< 0.001**
Veillonella	BAL	13.89	0	BAL	NSW	< 0.001
Veillonella	NSW	2.11	0	BAL	PSH	< 0.001
Veillonella	**PHS**	**18.76**	**0**	**NSW**	**PSH**	**< 0.001**

<x>: mean genus abundance. [2] p/a: number of samples where the genus was not detected.

Table 21. Comparison of core genera between body sites.

Nasopharynx			Oropharynx			Bronchoalevolar lavage		
Genus	**NSW <x>[1]**	**NSW p/a[2]**	**Genus**	**PHS <x>[1]**	**PHS p/a[2]**	**Genus**	**BAL <x>[1]**	**BAL p/a[2]**
Staphylococcus	24.16	0	*Prevotella*	25.07	0	*Prevotella*	27.70	0
Corynebacterium	23.93	0	*Veillonella*	18.76	0	*Streptococcus*	15.30	0
Dolosigranulum	9.32	20	*Streptococcus*	17.04	0	*Veillonella*	13.89	0
Streptococcus	7.51	0	*Leptotrichia*	5.55	0	*Fusobacterium*	4.60	0
Peptoniphilus	5.16	6	*Actinomyces*	4.26	0	*Megasphaera*	2.84	0
Moraxella	4.73	23	*Fusobacterium*	3.84	1	*Leptotrichia*	2.76	1
Anaerococcus	3.33	2	*Haemophilus*	2.85	0	*Actinomyces*	2.00	0
g_Actinomycetales	2.74	9	*Megasphaera*	2.34	0	*Haemophilus*	1.90	1
Finegoldia	2.51	9	*g__uncl_f_Pasteurellaceae*	1.98	22	g_Prevotellaceae	1.86	3
Prevotella	2.39	0	*Lachnoanaerobaculum*	1.45	0	*Alloprevotella*	1.77	0
Veillonella	2.11	0	*Neisseria*	1.36	12	*Neisseria*	1.71	8
g_ Neisseriaceae	1.41	16	*Granulicatella*	1.34	0	*Porphyromonas*	1.69	1
Propionibacterium	1.36	0	*Atopobium*	1.28	0	*Acinetobacter*	1.59	1
Haemophilus	1.05	2	*Porphyromonas*	1.05	1	*Gemella*	1.34	0
Rothia	0.85	3	*Centipeda*	0.96	1	*g__uncl_f_Pasteurellaceae*	1.11	13
Actinomyces	0.66	0	*Rothia*	0.94	0	*Centipeda*	1.11	2
Granulicatella	0.39	3	*Stomatobaculum*	0.94	1	*Campylobacter*	1.07	0
Gemella	0.38	2	*Campylobacter*	0.91	0	*Rothia*	1.07	0
Porphyromonas	0.28	7	*g__uncl_f_Prevotellaceae*	0.86	4	*Granulicatella*	0.98	0
Fusobacterium	0.27	5	*Alloprevotella*	0.81	5	*Atopobium*	0.92	0
Lactobacillus	0.26	12	*Lactobacillus*	0.66	26	*g__uncl_g_Saccharibacteria_genera_incertae_sedis*	0.89	0
Micrococcus	0.26	4	*Gemella*	0.65	0	*Parvimonas*	0.87	2
Megasphaera	0.23	4	*Oribacterium*	0.61	2	*Selenomonas*	0.64	2
Neisseria	0.21	16	*Selenomonas*	0.53	2	*Oribacterium*	0.61	1
Leptotrichia	0.20	3	*Peptostreptococcus*	0.45	7	*Parabacteroides*	0.43	8
Atopobium	0.19	8	*g__uncl_g_Saccharibacteria_genera_incertae_sedis*	0.40	2	*Dialister*	0.39	3
Pseudomonas	0.18	10	*g__uncl_f_Lachnospiraceae*	0.38	19	*Capnocytophaga*	0.35	3
Campylobacter	0.18	8	*Bifidobacterium*	0.33	26	*g__uncl_g_SR1_genera_incertae_sedis*	0.32	21
Gordonia	0.16	27	*Capnocytophaga*	0.26	6	*Lachnoanaerobaculum*	0.31	2
Acinetobacter	0.15	6	*Parvimonas*	0.25	6	*Lactobacillus*	0.30	21
Arthrobacter	0.15	18	*Eubacterium*	0.20	5	*Stenotrophomonas*	0.29	2
Lachnoanaerobaculum	0.15	8	*g__uncl_p_Bacteroidetes*	0.16	20	*Eubacterium*	0.26	8
Dialister	0.12	14	*g__uncl_g_SR1_genera_incertae_sedis*	0.15	27	*Cupriavidus*	0.26	2
Stomatobaculum	0.11	12	*Dialister*	0.12	4	*Catonella*	0.26	4
Sphingomonas	0.10	12	*Morococcus*	0.12	10	*g__uncl_o_Clostridiales*	0.25	1

<x>: mean genus abundance. [2] p/a: number of samples where the genus was not detected.

Table 22. Most abundant genera at different body sites.

Genus	Phylum/class	Mean relative abundance			Not detected[1]		
		BAL	NSW	PHS	BAL (%)	NSW (%)	PHS (%)
Alloprevotella	Bacteroidetes	1.79	0.04	0.82	0	37	13
Acinetobacter	γ-proteobacteria	1.62	0.14	-	3	16	100
Stenotrophomonas	β-proteobacteria	0.30	0.01	-	5	66	100
Cupriavidus	β-proteobacteria	0.27	-	-	5	74	100
Solobacterium	Firmicutes	0.03	-	0.01	11	68	26
Serratia	γ-proteobacteria	0.14	-	0.00	11	71	100
Unclassified Enterobacteriaceae	γ-proteobacteria	0.15	0.03	-	13	55	100
Tannerella	Bacteroidetes	0.12	-	0.09	13	68	26
Treponema	Spirochaetae	0.18	-	0.06	18	76	32
Fretibacterium	Bacteroidetes	0.16	-	0.03	21	68	47
Unclassified Comamonadaceae	β-proteobacteria	0.11	-	-	21	68	92
Parabacteroides	Bacteroidetes	0.45	-	-	21	74	100

[1]Not detected = Percentage of samples in which the genus is absent. For comparison, data is based on samples that were available for NSW, PHS and BAL

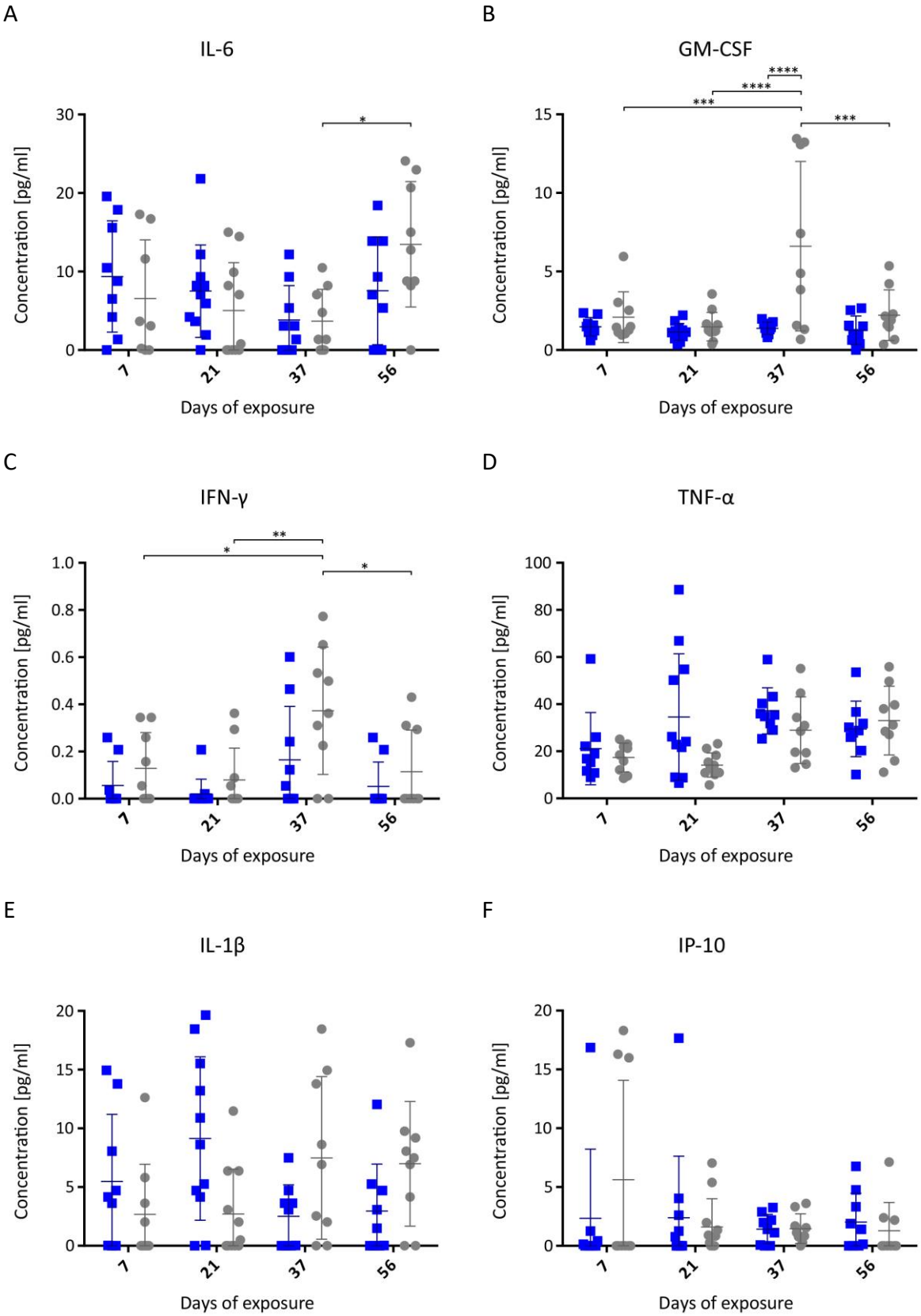
Table 23. Genera prevalently or exclusively found in BAL samples.

Exposure_days	Alpha diversity			Exposure_days	Alpha diversity			Exposure_days	alpha diversity		
	Richness	Shannon.effective	Simpson.effective		Richness	Shannon.effective	Simpson.effective		Richness	Shannon.effective	Simpson.effective
air_7	10	6.06	4.1	air_21	8	5.06	3.76	air_56	10	7.59	6.23
	5	3.16	2.28		4	2.11	1.62		6	2.41	1.65
	7	2.01	1.39		5	4.26	3.81		7	5.49	4.63
	12	8.94	7.19		12	9.54	8.24		8	5.57	4.02
	7	3.3	2.58		6	2.36	1.7		10	5.11	3.55
	4	2.14	1.72		3	2.76	2.57		9	6.91	6.12
	29	24.81	21.46		24	19.98	16.9		29	22.37	16.1
	32	26.56	20.8		27	22.78	18.78		31	25.34	19.88
	29	25.23	22.16		25	20.72	16.69		28	24.44	21.46
smoke_7	13	8.04	5.7	smoke_21	3	1.63	1.36	smoke_56	19	13.93	10.56
	5	3.49	2.93		12	7.69	5.88		7	5.29	4.33
	6	3.62	2.62		7	2.12	1.45		14	10.17	8.1
	13	10.07	8.61		4	2.7	2.16		9	5.34	4.16
	8	5.16	4.07		3	2.53	2.27		10	6.4	4.99
	4	2.03	1.6		8	5.65	4.73		6	3.05	2.3
	32	26.34	21.12		25	19.43	15.72		26	21.42	18.03
	40	36.64	32.98		26	22.71	20.18		33	27.78	22.53
	33	25.56	17.56		25	22.79	20.8		32	28.35	25.04

Table 24. Alpha diversity of the murine lung samples.

Exposure_days	Alpha diversity			Exposure_days	Alpha diversity			Exposure_days	Alpha diversity		
	Richness	Shannon.effective	Simpson.effective		Richness	Shannon.effective	Simpson.effective		Richness	Shannon.effective	Simpson.effective
air_7	34	28.2	21.81	air_21	40	35.68	30.73	air_56	31	26.08	21.46
	30	25.23	20.8		33	29.27	25.51		33	26.44	19.31
	29	22.4	17.56		43	40.5	37.56		33	27.97	22.92
	34	26.42	18.03		41	36.59	30.73		29	16.92	7.95
	34	30.1	26.51		39	36.4	33.8		36	31.07	26
	31	26.65	22.92		40	35.44	30.04		36	32.02	27.59
	37	32.89	28.17		32	26.04	20.48		35	30.56	26
	33	28.01	22.16		32	23.75	15.91		41	36.59	30.73
	33	26.03	19.04		29	18.98	10.82		42	37.58	31.44
	37	33.27	29.39		38	33.11	27.59		35	30.36	25.51
	41	36.64	31.44		38	34.39	30.73		30	23.92	19.04
	39	35.44	31.44		39	34.69	29.39		23	12.17	6.5
	31	23.01	15.91		34	28.36	22.92		35	29.06	22.16
	34	29.27	24.58		32	27.78	24.14		41	37.38	32.98
	33	27.97	22.92		37	29.6	20.8		39	34.74	30.04
	31	22.63	15.54		28	17.87	9.45		35	30.26	25.51
	36	30.36	24.58		34	29.6	25.51		30	21.92	13.94
	34	28.65	23.31		38	33.49	28.77		31	26.08	21.46
smoke_7	29	23.96	19.59	sm_21	38	33.44	28.17	sm_56	23	12.41	6.53
	39	33.49	26.51		32	26.96	22.16		26	14.35	7.01
	34	29.6	25.51		34	28.64	23.72		23	9.17	3.81
	38	32.6	26		43	40.5	37.56		32	27.47	23.31
	39	35.09	30.73		38	33.6	28.77		25	13.09	6.35
	35	30.67	26		35	30.71	26.51		19	7.29	3.29
	40	35.44	30.04		39	35.09	30.73		38	33.05	27.59
	38	34.39	30.73		36	30.87	25.51		42	38.01	32.98
	34	29.6	25.51		37	33.27	29.39		39	36.4	33.8
	29	21.89	15.54		37	31.75	25.51		38	32.89	27.04
	33	27.93	23.31		41	37.63	33.8		39	34.62	30.04
	31	25.65	20.8		41	37.76	33.8		38	34.74	31.44
	39	36.4	33.8		34	30.2	26.51		42	38.39	33.8
	42	39.04	35.58		40	36.4	32.19		40	34.39	27.04
	36	28.72	19.31		39	35.68	32.19		36	30.1	23.72
	34	28.94	23.72		40	37.01	33.8		41	35.09	27.04
	38	33.82	29.39		39	35.09	30.73		43	39.69	35.58
	37	31.43	25.04		30	26.04	22.53		41	38.01	34.67

Table 25. Alpha diversity of the murine caecum samples.


A
IL-6
Concentration [pg/ml]
Days of exposure
B
GM-CSF
Concentration [pg/ml]
Days of exposure
C
IFN-γ
Concentration [pg/ml]
Days of exposure
D
TNF-α
Concentration [pg/ml]
Days of exposure
E
IL-1β
Concentration [pg/ml]
Days of exposure
F
IP-10
Concentration [pg/ml]
Days of exposure

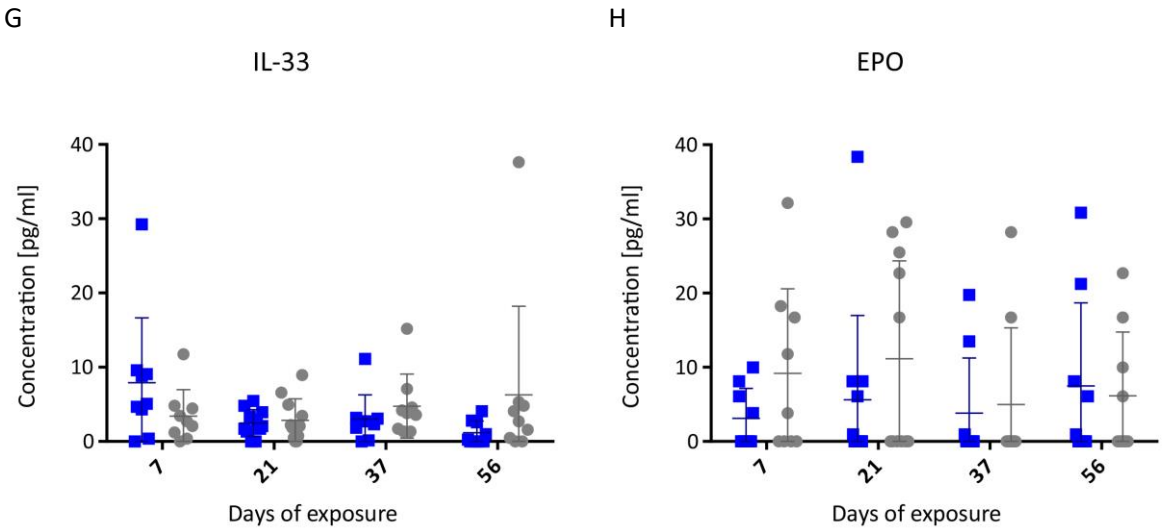


Figure 45. Concentration of inflammatory mediators in BAL fluid.
A) IL-6; B) GM-CSF; C) IFN-γ; D) TNF-α; E) IL-1β; F) IP-10; G) IL-33; H) EPO. Each data point represents an individual animal; data are shown as mean ± SD (n = 9-11/group). (■ air, ● CS). Two-way ANOVA with Tukey's multiple comparison test (black stars). Unpaired t-test (red stars). *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

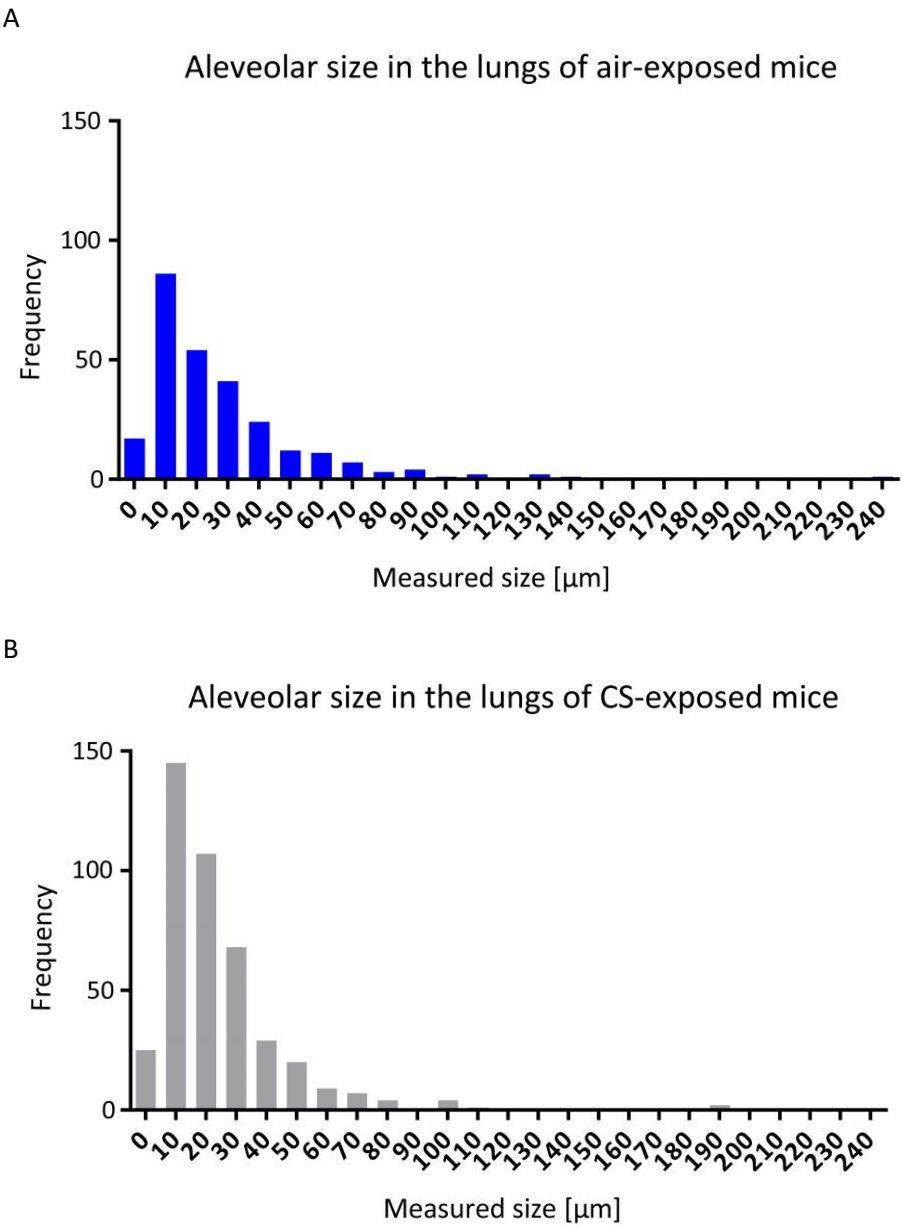


Figure 46. Distribution of the alveolar size of murine lungs.
A) Air-exposed lungs. n = 2; B) CS-exposed lungs. n = 4

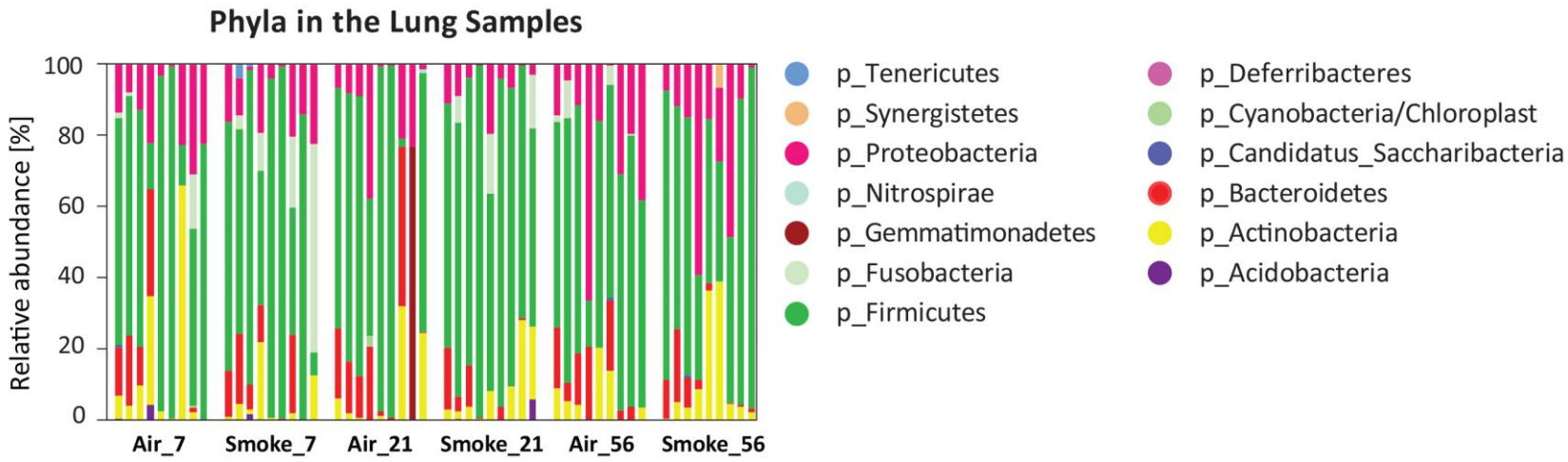


Figure 47. Taxonomic binning at the phylum level. Each bar plot represents one lung sample

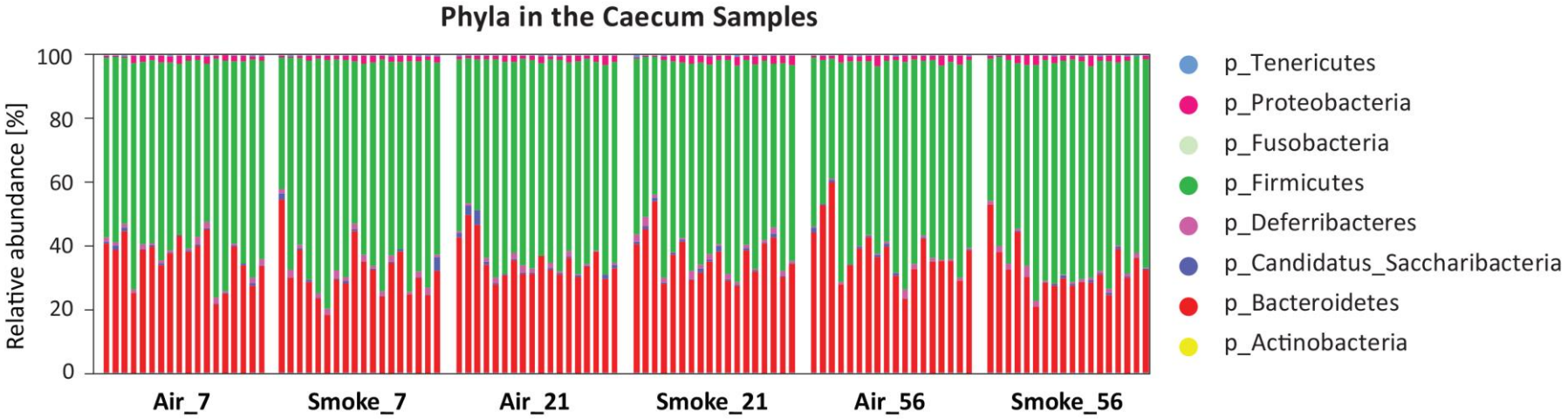


Figure 48. Taxonomic binning at the phylum level. Each bar plot represents one caecum sample.

7.2 Table of Figures

7.3 List of Tables

7.4 Acknowledgments

First of all, I want to thank Prof. Dr. Susanne Krauss-Etschmann for giving me the opportunity to work on this project and providing appreciated support and supervision throughout my PhD. Working with her encouraged me to follow my interest and expended my horizons.

Furthermore, my sincere thanks goes to Prof. Dr. Holger Heine for supervising me and sharing his knowledge during discussions we had. I am also thankful to Dr. Sabine Bartel for her guidance and help.

A very special gratitude goes to my former and present teammates from *Early Life Origins of Chronic Lung Diseases* and *DZL Experimental Microbiome Research* groups: Arne Krüger, Barbara Hammer, Beate Höschler, Dr. Eistine Boateng, Gabi Huß, Gregor Jatzlauk, Dr. Hanna Angstmann, Huan Ma, Janin Braun, Dr. Joni Valeska Lund, Dr. Karin Uliczka, Martin Wolff, Masha-Melissa Spauszus, Dr. Natalia El-Merhie and Dr. Sebastian Reuter. This work would have been impossible without their assistance and ideas. I am grateful for this awesome time we have spent together and still counting on many more to come.

My sincere appreciation goes to Prof. Dr. Michael Schloter, Dr. Stefan Pfeiffer, and their team from the *Research Unit for Comparative Microbiome Analysis* at the Helmholtz Center Munich for performing sequencing steps and thorough analysis of the microbiome data.

Further I would like to thank other LuMEn consortium members: Prof. Dr. Ulrich Schaible from the *Cellular Microbiology*, Andrea Glaewe, Dr. Christian Herzmann, Johanna Döhling, Lenka Krabbe and Steffi Fox from the *Study Center*, and Birgit Kullmann, Dr. Karoline Gaede and Romina Pritzkow from the *Biobank* at the Research Center Borstel, and other collaboration partners from IPHT in Jena and DMSZ in Braunschweig.

An important thanks belongs to Philipp Hagemann for helping me with the flow cytometry analysis, and Franziska Beyersdorf for performing CAST measurement. In addition, I would like to acknowledge other research groups at the Research Center Borstel for providing their lab space and equipment: *Innate Immunity* (Heine), *Experimental Pneumology* (Fehrenbach), *Asthma Exacerbation and Regulations* (Wegmann), *Molecular and Experimental Mycobacteriology* (Niemann), *Infection Immunology* (Hölscher), as well as the animal caretakers.

Also, my gratitude goes to other colleagues from the Research Center Borstel, as well as Gudrun Lehwark-Yvetot who is keeping our Ausländerbande active.

I thank all colleagues and friends from the Institute of Molecular Genetics and Genetic Engineering in Belgrade, but especially Dr. Aleksandra Divac-Rankov. They helped me while I was making my baby steps in research and showed that science can still be fun despite potential troubles.

A big thank you goes to Tracy Newton for improving my writing skills, as well as to Uroš Janković and David Carter for proofreading. With their contribution, this thesis looks much better.

There would not be enough words and pages to express my deepest love and gratitude to my parents, brother, sister-in-law, and to all of my friends at home and those scattered across the world. A big thank you for all your unconditional support.

8 Statutory declaration

I hereby declare that the content and design of the thesis named “The Influence of Cigarette Smoke on Human and Murine Lung Microbiota” is all my own work except clearly indicated and acknowledged collaborative contributions. No support apart from the supervisor's guidance and sources other than those listed were used.

I also declare that this thesis has not been submitted, neither partially nor wholly, as part of a doctoral examination procedure to another examining body. Part of results from the thesis have been submitted for publication.

This work was carried out in compliance with the rules of good scientific practice of the German Research Foundation.

Further, I confirm that no academic degree has ever been withdrawn.

Borstel, 14.10.2020

Place, Date

Signature