

Informed development of a multi-species biofilm in chronic obstructive pulmonary disease

BRYN SHORT,¹  CHRISTOPHER DELANEY,¹ WILLIAM JOHNSTON,² GARY J. LITHERLAND,^{3,4} JOHN C. LOCKHART,^{3,4} CRAIG WILLIAMS,⁵ WILLIAM G. MACKAY^{3,4} and GORDON RAMAGE^{1,2,*}

¹School of Medicine, Dentistry and Nursing, College of Medical, Veterinary and Life Sciences (MVLS), University of Glasgow; ²Safeguarding Health through Infection Prevention (SHIP) Research Group, Research Centre for Health, Glasgow Caledonian University, Glasgow; ³Institute of Biomedical and Environmental Health Research, School of Health and Life Sciences, University of the West of Scotland, Paisley; ⁴Hamilton International Technology Park, Glasgow; and ⁵Microbiology Department, Lancaster Royal Infirmary, University of Lancaster, Lancaster, UK

Short B, Delaney C, Johnston W, Litherland GJ, Lockhart JC, Williams C, Mackay WG, Ramage G. Informed development of a multi-species biofilm in chronic obstructive pulmonary disease. APMIS. 2024.

Recent evidence indicates that microbial biofilm aggregates inhabit the lungs of COPD patients and actively contribute towards chronic colonization and repeat infections. However, there are no contextually relevant complex biofilm models for COPD research. In this study, a meta-analysis of the lung microbiome in COPD was used to inform development of an optimized biofilm model composed of genera highly associated with COPD. Bioinformatic analysis showed that although diversity matrices of COPD microbiomes were similar to healthy controls, and internal compositions made it possible to accurately differentiate between these cohorts (AUC = 0.939). Genera that best defined these patients included *Haemophilus*, *Moraxella* and *Streptococcus*. Many studies fail to account for fungi; therefore, *Candida albicans* was included in the creation of an interkingdom biofilm model. These organisms formed a biofilm capable of tolerating high concentrations of antimicrobial therapies with no significant reductions in viability. However, combined therapies of antibiotics and an antifungal resulted in significant reductions in viable cells throughout the biofilm ($p < 0.05$). This biofilm model is representative of the COPD lung microbiome and results from *in vitro* antimicrobial challenge experiments indicate that targeting both bacteria and fungi in these interkingdom communities will be required for more positive clinical outcomes.

Key words: COPD; biofilm; microbiome.

Gordon Ramage, Safeguarding Health through Infection Prevention (SHIP) Research Group, Research Centre for Health, School of Health and Life Sciences, Glasgow Caledonian University, 70 Cowcaddens Road, Glasgow G4 0BA, UK. e-mail: gordon.ramage@gcu.ac.uk

Chronic obstructive pulmonary disease (COPD) is an extremely debilitating disease characterized by airway inflammation, which leads to airflow obstruction and alveolar destruction [1, 2], making it one of the leading causes of death worldwide [3]. Hallmark symptoms of COPD include cough, dyspnoea and phlegm production [4, 5]. Individuals suffering from COPD also experience acute exacerbations of COPD (AECOPD) which are defined by significant worsening of symptoms, often resulting in hospitalization [2, 5, 6]. These periods of exacerbation are

traditionally thought to be caused by individual pathogens such as *Haemophilus influenzae* and *Streptococcus pneumoniae* [7].

Several studies have been published that have sought to define the microbiome of COPD patients. The findings of these studies are in agreement and report compositional changes in the microbiome of COPD patients in comparison with healthy controls [8–11]. Differences in microbiome have also been observed when comparing stable and exacerbated COPD [6, 12, 13]. Several studies have documented dysbiosis and have reported a decrease in organisms such as *Lactobacillus*, *Prevotella* and *Porphyromonas*

Received 4 December 2023. Accepted 2 February 2024

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

in COPD patients which is also accompanied with an increase in the abundance of genera that are associated with respiratory infections such as *Haemophilus*, *Moraxella*, *Staphylococcus* and *Pseudomonas* [8, 12, 14, 15]. Collectively, these findings suggest that the airway microbiome plays a major role in the pathogenesis of COPD. This is likely multifactorial by influencing susceptibility to infection as well as airway inflammation and immune system regulation [14, 16]. Generally, the microbiome of healthy lungs resembles that of the oral cavity, and this remains consistent in COPD [9, 17].

With an increasing number of studies investigating the COPD microbiome, we have a greater understanding of not only what organisms are present in the lungs but also the states in which they exist. For example, evidence supporting the presence of biofilms is increasing, with studies reporting on long-term or chronic bacterial colonization along with the identification of *H. influenzae* strains isolated from COPD patients exhibiting biofilm-forming capabilities [18]; further, antibodies to peroxiredoxin–glutaredoxin (pdgX), which is required for biofilm formation in *H. influenzae*, have been identified in COPD patients [19–21]. Despite several lines of evidence indicating that the lung microbiome in COPD patients exists in a biofilm state, few studies have incorporated COPD-associated organisms to a multi-species biofilm. Those that have are relatively simplistic and describe dual-species bacterial biofilms [22–24]. As such, there currently exist no multi-species biofilm models for the *in vitro* study of COPD that accurately represent the fungal and bacteria contents of the lungs. Therefore, the main aim of this study was to address this gap in the current literature by developing a multi-species biofilm model in the context of COPD by using a meta-analysis of 16S-amplified microbiome data to identify the most relevant organisms for this model. It is the hope that this novel biofilm model can provide insight into how complex microbial communities respond to antimicrobial challenge in COPD.

MATERIALS AND METHODS

Data set collection and interrogation

Studies relating to the microbiome of healthy and COPD-affected lungs were collected by using keyword searches on PubMed (Table 1). Searches were performed on 21 March 2021. Studies published before 2010 were excluded and remaining studies were exported to the reference manager EndNote (version X8). A cut-off of 2010 was implemented to coincide with the publication of the first NGS studies investigating the lung microbiome in COPD. All search results were then exported to EndNote where titles and abstracts could be screened for relevance. Studies were excluded if they were not related to the lungs in health or in COPD in humans or if they were not NGS studies.

Shortlisted studies were interrogated in a manner similar to Butcher *et al.* [25] where relevant information was recorded. This included experimental details which may influence the microbiome or downstream analyses such as sampling method, number of cases, number of controls, Global Initiative for Obstructive Lung Disease (GOLD) status, smoking status, DNA extraction method, NGS platform, primer sequences and region of 16S rRNA targeted. Data accession and DOI numbers were also recorded where possible for data retrieval (Table S1). Again, shortlisted articles were then subjected to a scoring system to determine whether the study could be used in final analyses as done previously [25].

Data retrieval and processing

Where possible, data were downloaded directly from the ENA database using SRA accession numbers as identifiers before quality control protocols were carried out using a standardized pipeline as described in detail by Butcher and colleagues [25]. Briefly, the standardized pipeline was performed in Qiime2 [26] as follows; primers and barcodes were removed where possible, poor-quality reads were trimmed to a minimum length of 100 bp, paired-end reads were merged and finally, reads were assigned to OTUs. A cutoff of 100 bp was selected to account for predicted lower quality scores that may accompany data from earlier studies and for those for which quality scores for raw data were not available. OTU clustering was performed at 97% confidence using the Silva reference database using the closest referenced method.

Individual runs, which are individual samples from each of the studies retrieved, were discarded if the total read depth was below 500 reads to account for the low biomass of microorganisms traditionally found in the lungs. Features were then filtered if they appeared less than 10 times in less than four studies. OTUs were normalized by centre log₂ ratio (CLR) transformation of OTUs for further analysis using the `make.CLR` function within the

Table 1. Keywords and terms used for literature searches

Condition	Search used
Chronic obstructive pulmonary disease	((microb* OR bacteri* OR archae* OR fung* OR mycob*) AND (structure OR composition OR diversity OR community) AND (sequencing OR metabarcoding OR amplicon OR metagenom* OR 16S OR “ITS”)) AND ((lung) AND (COPD) OR (chronic obstructive pulmonary disease) NOT (review))
Healthy lung	((microb* OR bacteri* OR archae* OR fung* OR mycob*) AND (structure OR composition OR diversity OR community) AND (sequencing OR metabarcoding OR amplicon OR metagenom* OR 16S OR “ITS”)) AND ((lung) AND (health*) NOT (review))

microbeR package (<https://github.com/jbisanz/MicrobeR>; v0.3.2). Filtered reads were utilized for alpha diversity measures.

Microbiome diversity and composition

Analyses were performed on all studies collectively and on a per-study basis. The phyloseq package (v1.30.0) was used to calculate the Shannon, Simpson and Chao1 diversity indices using the estimate_richness function [27]. The base aov function was used to perform a two-way analysis of variance (ANOVA) to test for statistical significance between presence of COPD and alpha diversity matrix. ANOVA results were viewed using the TukeyHSD function.

Beta diversity distance matrices were arranged as principal coordinate analysis (PCoA) plots using the ape (analyses of phylogenetics and evolution) package in RStudio [28]. PCoA plots were then visualized using ggplot2. Permutational multivariate analysis of variance was performed on beta diversity matrices using the adonis2 function in the vegan package (<https://github.com/vegandevs/vegan>).

Average OTU abundance was calculated in RStudio by calculating abundance as a percentage of total OTUs. The nine most abundant phyla and genera were shortlisted, and the remaining phyla and genera were listed as 'other'.

Random forest classifiers and linear discriminant analysis

Random forest classifiers were built using the 80:20 random assignment ratio method whereby the entire data set was randomly split and 80% was used to train the model which was then tested against the remaining 20%. CLR-normalized OTUs were used as predictor variables for healthy/disease cases and models were validated using 10-fold cross-validation as done so previously [25]. Random forest models were built using the randomforest package and receiver operator characteristic (ROC) curves were built based on random forest models using pROC and ggROC packages [29, 30].

LDA was performed in the microbial package (<https://github.com/guokai8/microbial>) in RStudio using the ldamarker function. LDA scores were plotted using the plotLDA function in the same package using a cut-off score of 4 to identify highly important genera.

Microbial growth and standardization

All bacteria used in this study were informed by LDA analysis described above and their growth conditions have been summarized in Table 2. All isolates were stored long-term in glycerol at -80°C and revived on solid media for 24–72 h before propagation in liquid media overnight. Following growth in liquid media, all organisms were washed by centrifugation and suspension in phosphate-buffered saline (PBS). Bacteria were standardized by optical density to give a concentration of 1×10^8 cells/mL, and yeast cells were then standardized by counting on a haemocytometer. Following counting and optical density measurements, all organisms were adjusted to assay specific concentrations in selected media.

Table 2. Microbe growth conditions and primers

Organism	Solid media	Liquid media	Growth conditions	Primer sequences
<i>Candida albicans</i> SC5314	Sabouraud Dextrose Agar	Yeast Peptone Dextrose broth	30 °C, Aerobic	F- GAGCGTGGTTTCTCCCTCAAACCCGCTGG R- GGTTGGACGTTACCCGCCGCAAGCAATGTT
<i>Sireptococcus pneumoniae</i> DSMZ 14377	Colombia Blood agar +5% Horse Blood	Tryptic Soy broth	37 °C, 5% CO ₂	F- GATACATAGCCGACCTGAG R- CCATTGGCCGAAGATTCC
<i>Moraxella catarrhalis</i> DSMZ 11994	Colombia Blood agar +5% Horse Blood	Brain Heart Infusion broth	37 °C, 5% CO ₂	F- ACACGCAACTCTTGACGAAG R- CTGAGCCTGTCAATTGAGGAA
<i>Haemophilus influenzae</i> DSMZ 11969	Chocolate agar +5% Horse Blood	Brain Heart Infusion broth + NAD + Hemin	37 °C, 5% CO ₂	F- CTTACGCTTCTATCTCGGTGATTAATAA R- TGTTTCGCCATAACTTCATCTTAGC
<i>Rothia dentocariosa</i> DSMZ 43762	Colombia Blood agar +5% Horse Blood	Brain Heart Infusion broth	37 °C, Aerobic	F- GGGTTGTAACCTCTGTTAGCATC R- CGTACCCACTGCAAAACCCAG
<i>Veillonella dispar</i> NCTC 11831	Fastidious Anaerobe agar +5% Horse Blood	Schaedlers broth	37 °C, Anaerobic	F- CCGTATGGGATGGAAACTGC R- CCTTCGCCACTGGTGTCTTC

Biofilm characterization

All microbial cells were adjusted to 1×10^6 cells/mL in biofilm growth media. Media were supplemented with menadione (Sigma Aldrich, Gillingham, Dorset, UK) and hemin (Sigma Aldrich, Gillingham, Dorset, UK) at 10 and 1 mg/mL, respectively. Biofilms were grown in microtitre plates at 37°C with 5% CO₂ for 24 h. Biofilm formation was then assessed by staining with 0.05% (w/v) crystal violet as previously [31].

To continue biofilm characterization, multi-species biofilms were grown for 24 h on Thermanox™ coverslips (Thermo Fisher, Renfrew, UK) and treated with relevant antimicrobials (erythromycin, amoxicillin and voriconazole) for a further 24 h. Biofilm susceptibilities to antimicrobials were assessed by treating 24 h biofilms with two-fold serial dilutions of each antimicrobial alone and in combination as previously [32]. Efficacy of antimicrobial challenges was assessed by initially washing treated biofilms with PBS and determining biofilm viability using the Alamar Blue assay (Thermo Fisher). Methods can be found in detail provided by Brown *et al.* [33].

Biofilm composition was assessed following 24 h antimicrobial challenge, alone and in combination, at 128 µg/mL as done so previously [32]. Following biofilm treatments, cells were harvested via sonication in an ultrasonic water bath at 35 kHz. Sonicated cells were then treated with propidium monoazide (PMA; Cambridge Bioscience, Cambridge, UK) to prevent the detection of DNA from non-viable cells in future analyses as described elsewhere [34, 35]. DNA was extracted using the MasterPure™ Yeast kit as per the manufacturer's instructions (CamBio, Reading, UK). DNA was amplified using species-specific primers (sequences provided in Table 2) and viable colony forming equivalent (CFE) numbers were calculated using a standard curve of serially diluted DNA as described previously [36]. Thermal profile used was as follows: 2 min at 50 °C, 2 min at 95 °C, followed by 40 cycles of 10 s at 95 °C and 30 s at 60 °C.

Statistical analysis

All statistical analyses and graph production relevant to the COPD microbiome were performed using the appropriate functions in R (version 4.1.2). GraphPad Prism (version 9, La Jolla, California, USA) was used for biofilm data analysis and visualization, and *t*-tests were used to compare means of biofilm formation and viability. In all instances, statistical significance was attained when $p < 0.05$.

RESULTS

Overview of eligible studies

Initially, search terms described in Table 1 were used to search for relevant studies. Terms were used to search for publications specific to COPD or healthy lungs which returned 188 COPD and 498 healthy lung studies which were then assessed for relevance. Studies that were relevant to the aims of the study were those which characterized the lung microbiome of COPD patients or healthy controls

via NGS. A total of 620 studies were excluded due to not being relevant to study aims, leaving 66 studies (50 COPD and 16 healthy lung studies; Table S1). Of the 66 relevant studies, 52 either did not provide access to raw sequencing data or sufficient patient metadata (Fig. 1). This left 14 studies that were included in final analyses which are summarized in Table 3. After compiling all COPD and healthy lung microbiome studies that were to be included in the final analyses, relevant study information was noted.

Relevant information of the remaining studies was collected, covering 3372 samples. This showed that study design varied greatly from study to study, including sequencing platform, DNA extraction method and 16S target region (Fig. 2). The V4 region was the most targeted region of the 16S hypervariable region (5 out of 14 studies), and the Illumina MiSeq was used in 64% of studies (Table 3).

Bacterial diversity of the COPD microbiome

Sequencing data were downloaded from the European Nucleotide Archive (ENA) database, and, after processing, operational taxonomic units (OTUs) were clustered at 97% similarity using the SILVA database [37]. Firstly, the beta diversity, which quantifies the diversity of the microbiome across all patients was viewed. Data points showed variation based on author and 16S sequencing region. Therefore, the observed variance is likely a result of study-specific factors (Fig. 3 and Table S2). As done so by previous studies, analyses were performed at the genus level to alleviate these effects [38, 39]. Separation of data points based on disease status was observed (PERMANOVA $R^2 = 0.025$ and $p = 0.001$; Fig. 3A); however, the disease status of patients explained little variance throughout the data. We found that samples produced from the same study clustered more closely (PERMANOVA $R^2 = 0.235$ and $p = 0.001$; Fig. 3B). Additionally, as previously reported in other studies, the amplified 16S region has been reported to have a significant impact on data distribution and we report the same (PERMANOVA $R^2 = 0.025$ and $p = 0.001$; Fig. S1). However, explained variance was less than that of study of origin.

Next, the alpha diversity of the microbiome, the measurement of intra-patient microbiome diversity and a subject of debate in recent years, was assessed (Fig. 4). From this analysis, it is clear that change in alpha diversity is not consistently altered in COPD. On average, 289 OTUs were observed in COPD compared with 281 in healthy controls ($p > 0.9$). Additional measures of alpha diversity (Shannon, Simpson and Chao1 indices) were also

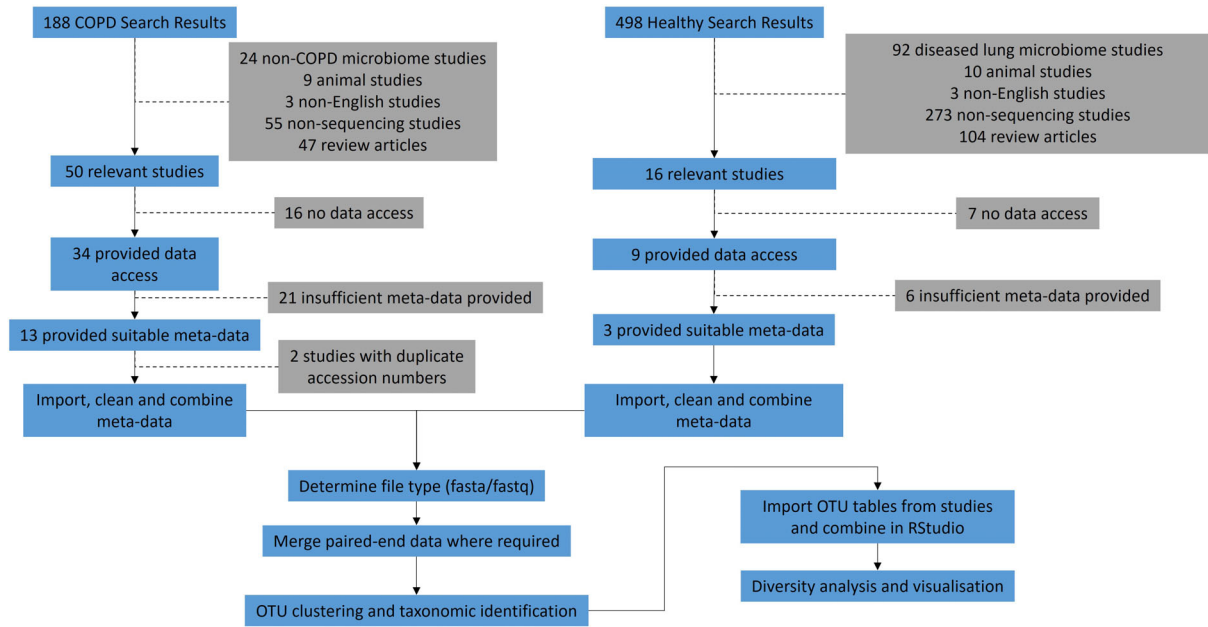


Fig. 1. Study acquisition pipeline. Searches returned a total of 686 results. Studies were screened for relevance, and all relevant studies were included in the final analysis where access to raw data was provided and sufficient patient metadata was provided. Final analyses were performed on a total of 14 studies.

Table 3. Studies in final analysis

Author	Controls	N (controls)	Cases	N (cases)	Region	Platform	Data score
Bassis, 2015	NA	NA	Healthy	28	V3-V5	Roche 454	3
Cameron, 2016	Healthy Smokers	10	COPD	8	ND	HiSeq	4
Dicker, 2018	NA	NA	COPD	141	V3-V4	MiSeq	3
Dickson, 2017	NA	NA	Healthy	8	V4	MiSeq	3
Einarsson, 2016	Healthy	19	COPD	18	V1-V2	MiSeq	3
Filho, 2019	NA	NA	COPD	112	V4	MiSeq	3
Jubenville, 2018	COPD	9	AECOPD	9	V6-V8	Roche 454	3
Mayhew, 2018	NA	NA	COPD	101	V4	MiSeq	3
McTaggart, 2019	NA	NA	Healthy	89	ITS1	MiSeq	3
Rylance, 2016	NA	NA	Healthy	44	V1-V3	Roche 454	3
Sun, 2020	AECOPD	15	AECOPD Recovery	15	V4-V5	Roche 454	3
Sze, 2016	NA	NA	COPD	21	V4	MiSeq	3
Wang, 2018	COPD	281	AECOPD	281	V4	MiSeq	4
Wang, 2019	Healthy	16	COPD	43	V4	MiSeq	4

calculated but revealed no significant differences. Shannon ($p = 0.1$) and Simpson ($p = 0.06$) indices showed slight but not statistically significant decreases in species evenness and richness between health and disease. Again, no significant differences in the Chao1 index between groups was observed ($p = 0.08$).

Microbial composition of the COPD lungs

The relative abundances of the most prevalent genera were then quantified (Fig. S2). *Veillonella* was identified as the most abundant genera across all

samples followed by *Streptococcus*. There was also a noticeable increase in the abundance of *Haemophilus* in COPD samples when compared to healthy controls. Considerable amounts of 'other' genera comprised a large percentage of the total composition which is representative of the increased (but not significant) alpha diversity in healthy cohorts.

To investigate whether the lung microbiome could be used as an alternative method of diagnosing COPD, random forest analysis was performed (Fig. S3). With a set of accurate training data, where the model was able to distinguish between a COPD and control microbiome 100% of the time

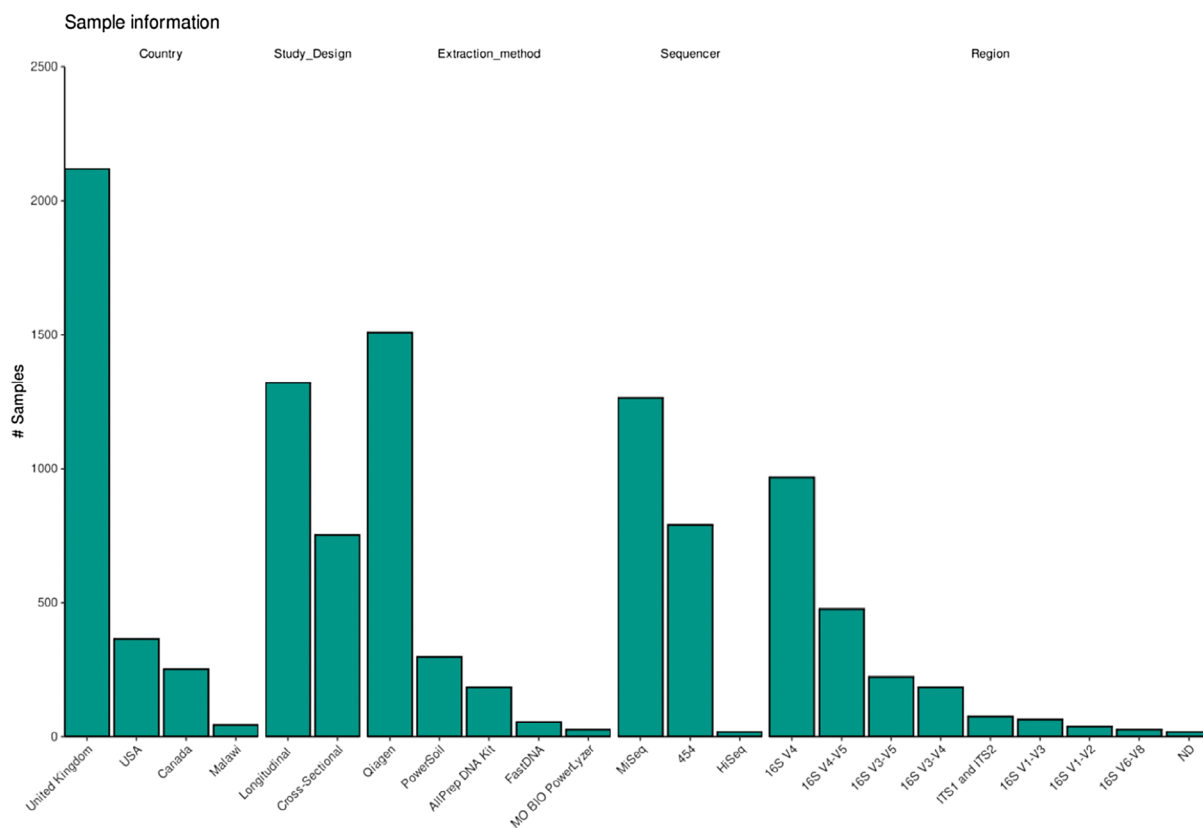


Fig. 2. Study characteristics. Following data processing, individual study metadata files were merged to allow for trend visualization. Samples from the 14 studies included in the final analysis were broken down by country of origin, study type, extraction method used, sequencing platform and 16S target region.

(area under the curve [AUC] = 1.0; Fig. S3), we then applied this model to the remaining data sets. We demonstrate that using OTUs at the genus level was the best predictor of a COPD microbiome (AUC = 0.939), followed by KO and OTUs spanning the species level (AUC = 0.902 and 0.895, respectively).

With the findings, we performed linear discriminant analysis (LDA) to identify key genera that help to best distinguish healthy from COPD microbiomes. Using cut-offs of 4 and -4 to identify highly associated genera, we observed that genera such as *Streptococcus*, *Haemophilus*, *Moraxella* and *Veillonella* were associated with COPD whereas *Porphyromonas*, *Lactobacillus*, *Staphylococcus* and *Prevotella* were linked to characterizing healthy microbiomes (Fig. 5).

Development and characterization of a novel COPD biofilm model

Using the genera that were identified as being highly associated with the microbiome in COPD, we developed a multi-species, interkingdom biofilm

model containing the five top bacterial genera, including *C. albicans* which was recently identified as the most dominant fungal component of the lung microbiome [40]. As expected, *C. albicans* produced robust biofilms, whereas the selected bacteria were not capable of widespread visible biofilm formation, with the exception of *V. dispar* (Fig. 6A). However, when grown in a multi-species biofilm alongside *C. albicans*, we observed a significant increase in biofilm biomass ($p < 0.05$). Next, we sought to determine the antimicrobial susceptibility profile of this newly described multi-species biofilm using antimicrobials that are common in the management of infections in COPD such as erythromycin, amoxicillin and voriconazole. Metabolic testing is widely used throughout the literature as a means to assess the potency of biofilm treatment strategies [41]. We therefore used this approach to determine both the sessile minimum inhibitory concentration 50 and 80 (SMIC₅₀ and SMIC₈₀, respectively), which are the drug concentrations required to inhibit 50% and 80% of biofilm metabolic viability. After exposure to antimicrobials either alone or in combination it was found that no treatment

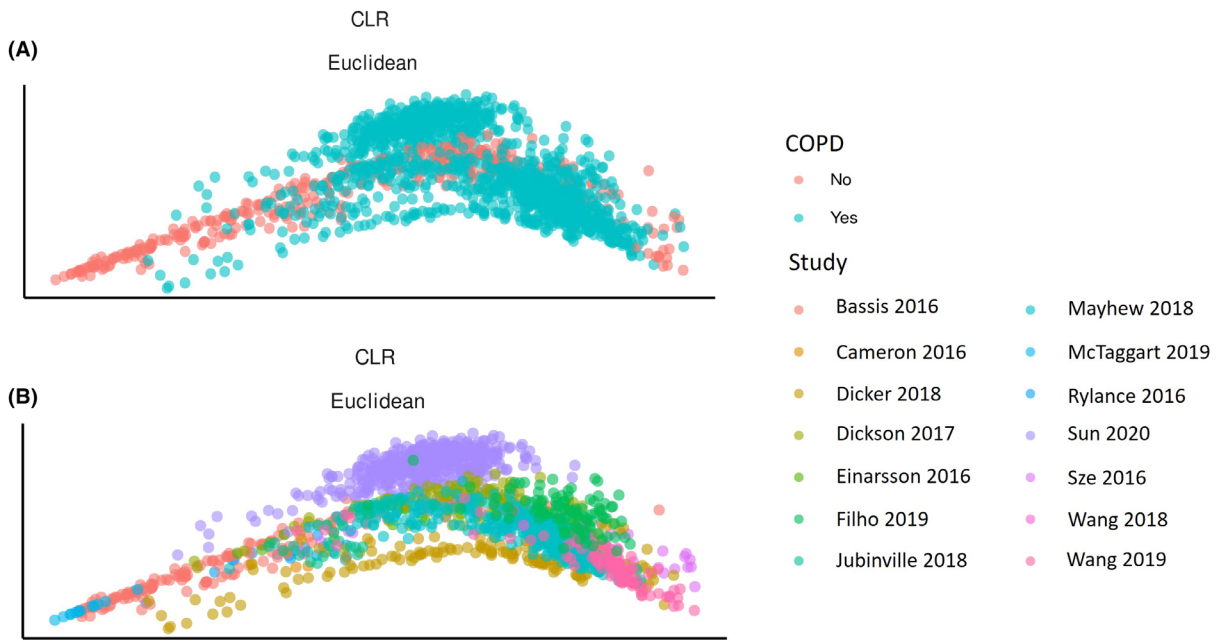


Fig. 3. Principal coordinate analysis of microbiome samples by disease status and study. Data points are coloured by study and provides evidence for study specific factors influencing the outcome of lung microbiome studies.

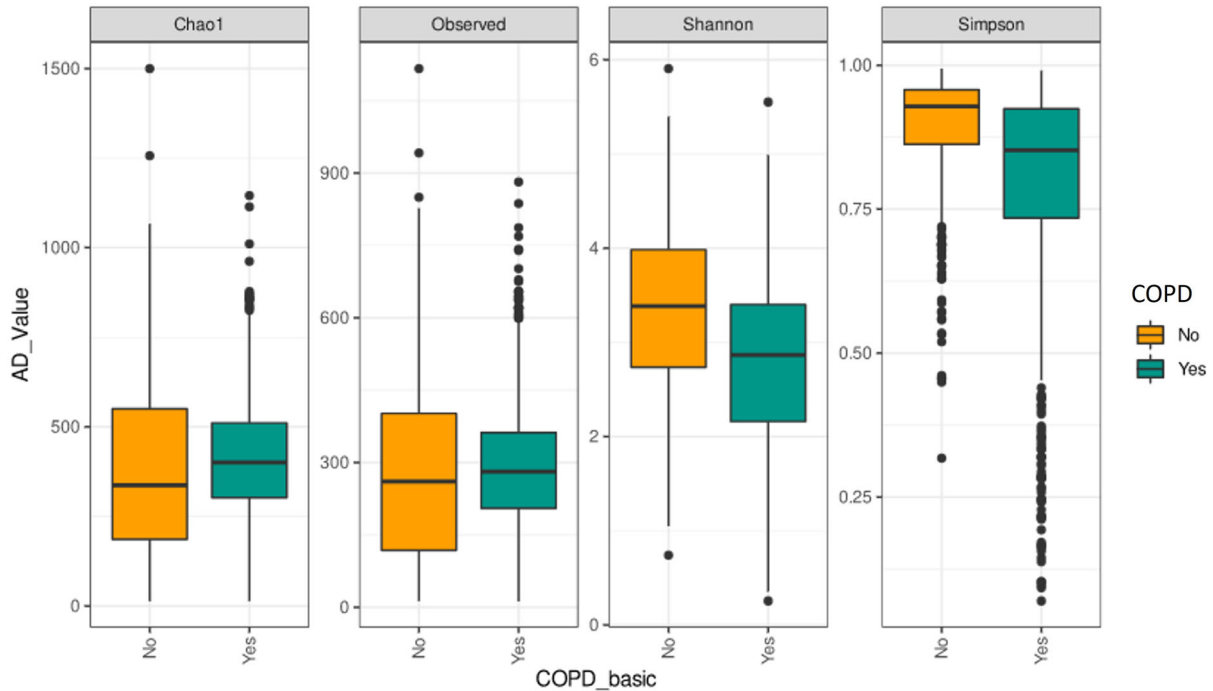


Fig. 4. Alpha diversity of the healthy and COPD lung microbiomes. The number of observed OTUs as well as Chao1, Shannon and Simpson diversity indices were calculated, demonstrating a decrease in microbiome richness and evenness in COPD. No statistical significance was observed when comparing healthy and diseased samples following one-way ANOVA and Tukey’s *post hoc* test. Upper and lower horizontal lines represent the upper and lower quartiles, respectively. Horizontal lines in the middle of the boxes represent the mean, whiskers (vertical lines outside each box) show data distribution beyond the upper and lower quartiles and individual dots represent outlier datapoints.

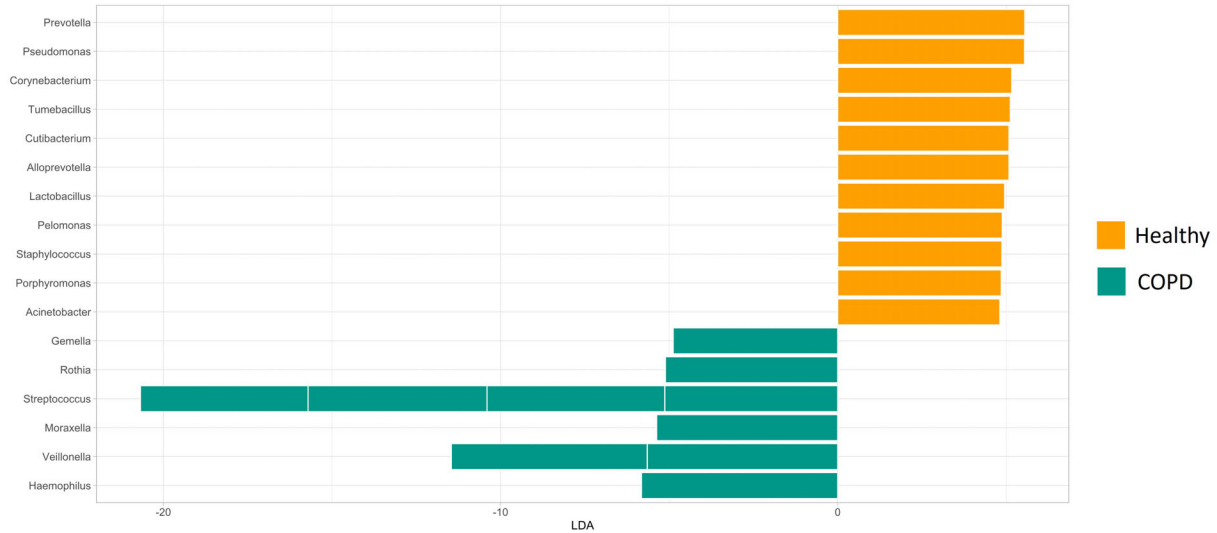


Fig. 5. Linear discriminant analysis displaying the genera that best classify health and disease. Cut-off points of -4 and 4 were set to identify the genera that were most enriched in health and disease. COPD enriched bacteria (blue) are designated negative values and health-enriched bacteria (yellow) are given positive scores.

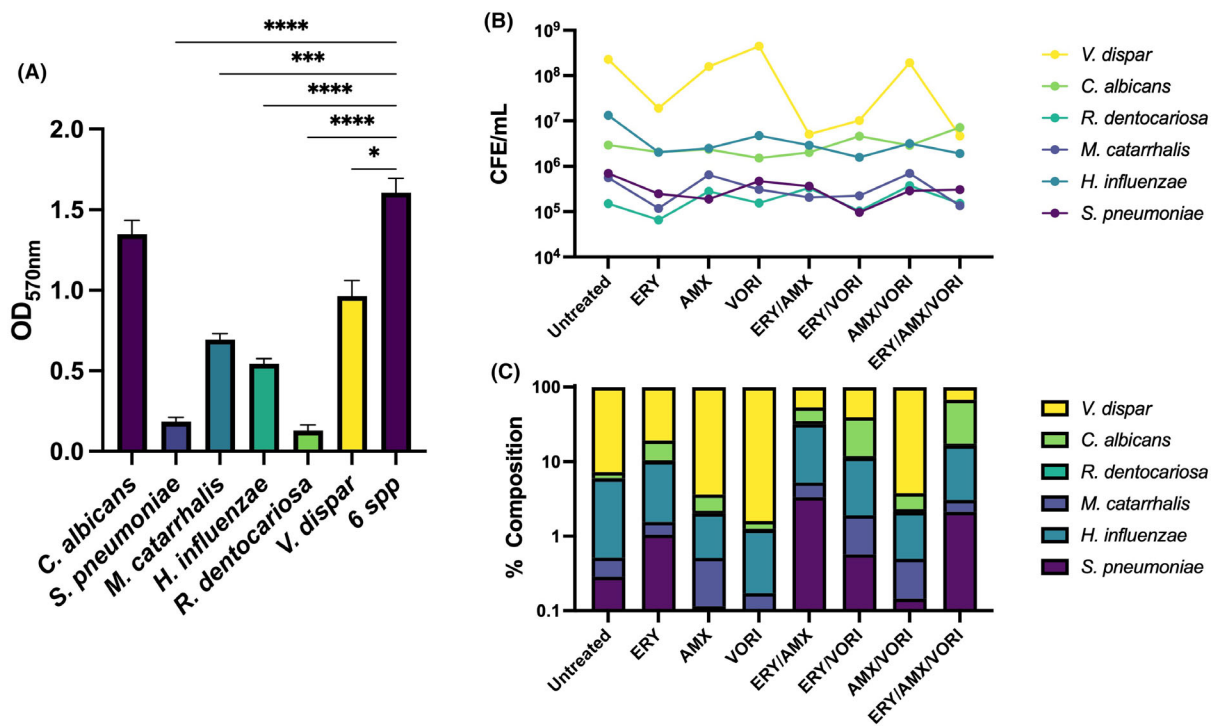


Fig. 6. Informed development of a multi-species COPD biofilm model. Bacteria and fungi were grown in combination for 24 h to form, mature biofilms representative of the COPD microbiome. Following growth, 0.05% crystal violet was used to assess overall biomass (A). The antimicrobials erythromycin (ERY), amoxicillin (AMX) and voriconazole (VORI) were used alone and in combination to treat mature biofilms for a further 24 h. Following antimicrobial challenge, biomass was removed via sonication and cells were treated with 10 mM PMA so that following DNA extraction, only DNA from live cells would be detected. True CFE values for each organism (B) and percentage composition is shown in the bar graph (C) with and without treatments. Viable CFEs are presented as the means of all experiments. Each experiment was performed in triplicate, all on separate occasions (*; $p < 0.05$, **; $p < 0.01$, ***; $p < 0.001$, ****; $p < 0.0001$).

method was sufficient to reduce multi-species biofilm viability below 50% and all SMIC for each treatment was found to be $>128 \mu\text{g/mL}$. However, combined treatment with all three antimicrobials was most effective and reduced biofilm viability by 33% (Fig. S4).

When assessing the composition of the described biofilms, we found that the composition of a mature biofilm commensurate with published literature, where *Veillonella* was the most dominant genera both in COPD patient microbiomes and in untreated biofilms (Fig. 6B,C). However, following treatments, the internal composition of the biofilms underwent shifts, specifically in the abundances of *V. dispar*, which significantly decreased from 2.2×10^8 CFE/mL to 4.6×10^6 CFE/mL when treated with a combination of erythromycin, amoxicillin and voriconazole ($p < 0.05$). Treatment with combinations of all antimicrobials also resulted in an increase in the percentage abundance of *C. albicans* from 1 to 50% (Fig. 6C). Although it was not statistically significant, this corresponded to a 2.4-fold increase in true-viable CFEs (2.9×10^7 to 7.1×10^7 CFE/mL). Additionally, treatments with erythromycin and voriconazole showed the highest efficacy against *H. influenzae*, reducing viable cells from 1.3×10^7 to 1.5×10^6 CFE/mL, equating to an 89% reduction ($p < 0.05$). We also found that only through using a combined treatment approach with two or more antimicrobials were we able to observe a significant reduction in total biofilm cells whereby combinations of erythromycin-amoxicillin, erythromycin-voriconazole and erythromycin-amoxicillin-voriconazole resulted in reductions of 96, 93 and 94%, respectively ($p < 0.05$; Fig. S5).

DISCUSSION

With interest surrounding the microbiological factors in COPD growing, we are now in a position where we must develop relevant, *in vitro* models to study these aspects of disease biology. Here we utilize a bioinformatics pipeline that has been previously described in the context of identifying key microorganisms in dental caries, to interrogate similar COPD microbiome data sets. We identify organisms which are characteristic of the COPD microbiome signature, making them prime candidates for integration into a multi-species biofilm model that can be utilized to study the microbiological nature of COPD *in vitro*.

Firstly, we report on a large-scale meta-analysis on publicly available data sets regarding the lung microbiome of COPD patients and healthy individuals. Initial findings from literature searches

identified large variances in study design, likely contributing to the large-scale data heterogeneity observed in Fig. 2. However, at present, there are no methods to effectively account for study-specific batch effects in NGS studies. To best mitigate these effects, analyses were largely confined to the genus level as done in previous studies [38, 39, 42]. Although this alleviated some study-specific effects, some biases would still remain such as the use of different DNA extraction methods can select for different organisms due to different cell lysis methods [43] and the targeting of different 16S regions can influence the outcomes of sequencing runs [44].

Through community composition and diversity analyses, no significant differences in alpha diversity matrices were observed between the lung microbiomes of COPD patients and healthy individuals. This has been a subject of dispute throughout the literature in recent years with some studies reporting decreases in microbial diversity in COPD whereas others suggest the opposite [9, 45]. While we could speculate that additional studies may accentuate the differences observed in alpha diversity, it could also be revealed that it is not the differences in diversity that influence COPD presence and progression but a compositional and functional change instead. This could be reinforced by our findings that random forest modelling could still accurately differentiate between healthy and COPD microbiomes and that particular genera such as *Haemophilus* and *Moraxella* were responsible for this finding. Similar findings were reported by a previous study that found these same genera were part of a group of consistently, differentially abundant genera which could be used as a means of aiding the diagnosis of COPD in future [39].

Our findings regarding beta diversity distributions in Fig. 3 are in agreement with previous meta-analyses on COPD as well as others with focus on other health conditions, which have found that a combination of study-specific factors such as DNA extraction method, study location and sequencing platform, contribute most to data grouping [25, 39, 42, 46].

Using the genera that were found to best discern the microbiome of COPD patients from that of healthy controls, we designed a complex, multi-species biofilm model in which *C. albicans* was introduced as the most abundant fungal species in COPD lungs [40]. Studies have shown that this a critical, acting in the strictest sense as a keystone microorganism within biofilms, structurally acting as a scaffold to support bacterial adhesion [47]. Within our biofilms, *C. albicans* forms a dense hyphal network closely associated with bacteria, a

phenomenon commonly seen in dual- and multi-species biofilms of fungi and bacteria [48–50]. This close, interkingdom, association has previously been shown to increase antimicrobial tolerance to both antibiotics and antifungals [48, 51]. In the data presented above, we showed that high concentrations of antimicrobials traditionally used in the management of COPD were not sufficient to achieve complete biofilm killing. Studies on other interkingdom biofilms have produced similar results whereby a combined therapy of antibiotics and antifungals results in a shift in internal microbe levels. Indeed, we have presented similar results elsewhere [32].

The model described above provides a composition similar to that of the lung microbiome in COPD, as we identified in Fig. S5. This is characterized by a large abundance of *Veillonella*. However, despite antimicrobial challenges, the biofilm remained viable. Similar to findings of previous studies, multi-species biofilms remain largely unaffected by antimicrobial therapy challenges and, for example, when treating with an antibiotic, the abundance of *C. albicans* increases [32]. Such responses to antimicrobial challenge could have severe implications on patient outcome as commonly used antimicrobials in COPD management are not sufficient for complete removal of biofilm biomass, which could leave patients susceptible to future infections. This fluctuating composition of the biofilm can be attributed to therapies preferentially targeting particular organisms, creating the opportunity for others to prosper. With this being said, the lack of responsiveness from the biofilm to antimicrobials that we observed is not entirely unexpected. Interkingdom biofilms containing *C. albicans* and bacteria are often shown to be highly tolerant to treatments [48, 51]. This is due to the increased production of extracellular matrix components, preventing access of antimicrobials to biofilm cells [52, 53].

Our observations (Fig. 6) are reinforced by that of Townsend *et al.* [32], who showed that a combined antimicrobial approach is necessary to induce a significant effect on these biofilm consortia by affecting internal biofilm composition. Although these authors did not find combination therapy to influence overall total cell numbers in the biofilm, we determined that combining certain antibiotics (ERY/AMX, ERY/VORI and ERY/AMX/VORI) was sufficient to significantly reduce viable biofilm CFEs. This may be attributed to the more complex nature of this biofilm.

This study does possess limitations in that a small panel of active agents was tested which was limited to antimicrobials. Previous studies have shown that alternative therapies such as DNase in

combination with antimicrobials increases potency against dual-species biofilms of *C. albicans* and *Staphylococcus aureus* [48]. Therefore, employing additional therapies could yield promising results in the treatment of multi-species biofilms such as the one described herein. Moreover, *S. pneumoniae* produces the autolysin *lytA* which has previously been reported to degrade streptococcal biofilms which could explain why such little biomass was detected in the *S. pneumoniae* only biofilms [54].

Herein, we have described the first report of a biofilm model that is representative of the lung microbiome of COPD patients that is both robust and easily reproducible. This model can be utilized in future as a basis to begin studying the complex network of microbial interactions in the COPD lungs, and more importantly as a means to investigate the efficacy of novel therapies before use in a more developed infection model. Fundamentally, this supports the 3Rs initiative which aims to identify alternative research methods to reduce the use of animal models wherever possible and provides an evidence-based biofilm model platform which can be used by academics and industry as a clinically relevant testbed.

FUNDING INFORMATION

Bryn Short was supported by a matched Ph.D. studentship provided by the University of the West of Scotland to support the Borders and Regions Airways Training Hub project (BREATH; INTVA/045) which is funded by the European Union (EU), under the INTERREG VA Programme, managed by the Special EU Programmes Body. The funders had no role in study design, data collection and interpretation or the decision to submit this work for publication.

DATA AVAILABILITY STATEMENT

Sequencing data can be accessed via SRA accession numbers which can be found in the original study. All other *in vitro* biofilm data is available from the authors upon request.

REFERENCES

1. Mannino DM, Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. *Lancet*. 2007;370(9589):765–73.
2. Rabe KF, Watz H. Chronic obstructive pulmonary disease. *Lancet*. 2017;389(10082):1931–40.
3. WHO. The Top 10 Causes of Death 2018. Available from: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>

4. Burkhardt R, Pankow W. The diagnosis of chronic obstructive pulmonary disease. *Dtsch Arztebl Int.* 2014;111(49):834–46.
5. Han MK, Agusti A, Calverley PM, Celli BR, Criner G, Curtis JL, et al. Chronic obstructive pulmonary disease phenotypes. *Am J Respir Crit Care Med.* 2010;182(5):598–604.
6. Jubinville E, Veillette M, Milot J, Maltais F, Comeau AM, Levesque RC, et al. Exacerbation induces a microbiota shift in sputa of COPD patients. *PLoS One.* 2018;13(3):e0194355.
7. Sethi S, Evans N, Grant BJB, Murphy TF. New strains of bacteria and exacerbations of chronic obstructive pulmonary disease. *N Engl J Med.* 2002;347(7):465–71.
8. Cameron SJS, Lewis KE, Huws SA, Lin W, Hegarty MJ, Lewis PD, et al. Metagenomic sequencing of the chronic obstructive pulmonary disease upper bronchial tract microbiome reveals functional changes associated with disease severity. *PLoS One.* 2016;11(2):e0149095.
9. Erb-Downward JR, Thompson DL, Han MK, Freeman CM, McCloskey L, Schmidt LA, et al. Analysis of the lung microbiome in the “healthy” smoker and in COPD. *PLoS One.* 2011;6(2):e16384.
10. Wang Z, Maschera B, Lea S, Kolsum U, Michalovich D, Van Horn S, et al. Airway host-microbiome interactions in chronic obstructive pulmonary disease. *Respir Res.* 2019;20(1):113.
11. 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.
12. Wang Z, Bafadhel M, Haldar K, Spivak A, Mayhew D, Miller BE, et al. Lung microbiome dynamics in COPD exacerbations. *Eur Respir J.* 2016;47(4):1082–92.
13. Williams NP, Ostridge K, Devaster J-M, Kim V, Coombs NA, Bourne S, et al. Impact of radiologically stratified exacerbations: insights into pneumonia aetiology in COPD. *Respir Res.* 2018;19(1):143.
14. Wang L, Hao K, Yang T, Wang C. Role of the lung microbiome in the pathogenesis of chronic obstructive pulmonary disease. *Chin Med J (Engl).* 2017;130(17):2107–11.
15. Garcia-Nuñez M, Marti S, Puig C, Perez-Brocá V, Millares L, Liñares F, et al. Influence of biofilm on the bronchial microbiome in COPD patients colonized or infected by *Pseudomonas aeruginosa*. *Eur Respir J.* 2015;46(Suppl 59):PA5028.
16. Dickson RP, Martinez FJ, Huffnagle GB. The role of the microbiome in exacerbations of chronic lung diseases. *Lancet.* 2014;384(9944):691–702.
17. Bassis CM, Erb-Downward JR, Dickson RP, Freeman CM, Schmidt TM, Young VB, 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.
18. Short B, Carson S, Devlin A-C, Reihill JA, Crilly A, MacKay W, et al. Non-typeable *Haemophilus influenzae* chronic colonization in chronic obstructive pulmonary disease (COPD). *Crit Rev Microbiol.* 2021;47:192–205.
19. Ahearn CP, Gallo MC, Murphy TF. Insights on persistent airway infection by non-typeable *Haemophilus influenzae* in chronic obstructive pulmonary disease. *Pathogens Dis.* 2017;75(4):ftx042.
20. Murphy TF, Brauer AL, Schiffmacher AT, Sethi S. Persistent colonization by *Haemophilus influenzae* in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med.* 2004;170(3):266–72.
21. Murphy TF, Kirkham C. Biofilm formation by non-typeable *Haemophilus influenzae*: strain variability, outer membrane antigen expression and role of pili. *BMC Microbiol.* 2002;2:7.
22. Bair KL, Campagnari AA. *Moraxella catarrhalis* promotes stable polymicrobial biofilms with the major otopathogens. *Front Microbiol.* 2020;10:3006.
23. Domenech M, Garcia E. N-acetyl-L-cysteine and cysteamine as new strategies against mixed biofilms of nonencapsulated *Streptococcus pneumoniae* and nontypeable *Haemophilus influenzae*. *Antimicrob Agents Chemother.* 2017;61(2):1–10. <https://doi.org/10.1128/aac.01992-16>
24. Sempere J, Llamasí M, Román F, Lago D, González-Camacho F, Pérez-García C, et al. Clearance of mixed biofilms of *Streptococcus pneumoniae* and methicillin-susceptible/resistant *Staphylococcus aureus* by antioxidants N-acetyl-L-cysteine and cysteamine. *Sci Rep.* 2022;12(1):6668.
25. Butcher MC, Short B, Veena CLR, Bradshaw D, Pratten JR, McLean W, et al. Meta-analysis of caries microbiome studies can improve upon disease prediction outcomes. *APMIS.* 2022;130(12):763–77.
26. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol.* 2019;37(8):852–7.
27. McMurdie PJ, Holmes S. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One.* 2013;8(4):e61217.
28. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics.* 2018;35(3):526–8.
29. Breiman L. Random forests. *Machine Learning.* 2001;45(1):5–32.
30. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* 2011;12(1):77.
31. Sherry L, Rajendran R, Lappin DF, Borghi E, Perdoni F, Falleni M, et al. Biofilms formed by *Candida albicans* bloodstream isolates display phenotypic and transcriptional heterogeneity that are associated with resistance and pathogenicity. *BMC Microbiol.* 2014;14:182.
32. Townsend EM, Sherry L, Kean R, Hansom D, Mackay WG, Williams C, et al. Implications of antimicrobial combinations in complex wound biofilms containing fungi. *Antimicrob Agents Chemother.* 2017;61(9):e00672-17.
33. Brown JL, Short B, Ware A, Sherry L, Kean R, Ramage G. Cell viability assays for *Candida auris*. *Methods Mol Biol.* 2022;2517:129–53.
34. Sherry L, Lappin G, O'Donnell LE, Millhouse E, Millington OR, Bradshaw DJ, et al. Viable compositional analysis of an eleven species oral polymicrobial biofilm. *Front Microbiol.* 2016;7:912.

35. Brown JL, Townsend E, Short RD, Williams C, Woodall C, Nile CJ, et al. Assessing the inflammatory response to in vitro polymicrobial wound biofilms in a skin epidermis model. *NPJ Biofilms Microbiomes*. 2022;8(1):19.
36. O'Donnell LE, Smith K, Williams C, Nile CJ, Lappin DF, Bradshaw D, et al. Dentures are a reservoir for respiratory pathogens. *J Prosthodont*. 2016;25(2):99–104.
37. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2012;41(D1):D590–6.
38. Duvallat C, Gibbons SM, Gurry T, Irizarry RA, Alm EJ. Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat Commun*. 2017;8(1):1784.
39. Wang Z, Yang Y, Yan Z, Liu H, Chen B, Liang Z, et al. Multi-omic meta-analysis identifies functional signatures of airway microbiome in chronic obstructive pulmonary disease. *ISME J*. 2020;14(11):2748–65.
40. Tiew PY, Dicker AJ, Keir HR, Poh ME, Pang SL, Mac Aogáin M, et al. A high-risk airway mycobiome is associated with frequent exacerbation and mortality in COPD. *Eur Respir J*. 2020;57:2002050.
41. Wilson C, Lukowicz R, Merchant S, Valquier-Flynn H, Caballero J, Sandoval J, et al. Quantitative and qualitative assessment methods for biofilm growth: a mini-review. *Res rev. J Eng Technol*. 2017;6(4):1–42.
42. Bisanz JE, Upadhyay V, Turnbaugh JA, Ly K, Turnbaugh PJ. Meta-analysis reveals reproducible gut microbiome alterations in response to a high-fat diet. *Cell Host Microbe*. 2019;26(2):265–272.e4.
43. Brooks JP, Edwards DJ, Harwich MD, Rivera MC, Fettweis JM, Serrano MG, et al. The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiol*. 2015;15(1):66.
44. Cai L, Ye L, Tong AH, Lok S, Zhang T. Biased diversity metrics revealed by bacterial 16S pyrotags derived from different primer sets. *PLoS One*. 2013;8(1):e53649.
45. Sze MA, Dimitriu PA, Hayashi S, Elliott WM, McDonough JE, Gosselink JV, et al. The lung tissue microbiome in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2012;185(10):1073–80.
46. Delaney C, Veena CLR, Butcher MC, McLean W, Shaban SMA, Nile CJ, et al. Limitations of using 16S rRNA microbiome sequencing to predict oral squamous cell carcinoma. *APMIS*. 2023;131:262–76.
47. Young T, Alshanta OA, Kean R, Bradshaw D, Pratten J, Williams C, et al. *Candida albicans* as an essential “Keystone”. Component within polymicrobial oral biofilm models? *Microorganisms*. 2020;9(1):1–23.
48. Kean R, Rajendran R, Haggarty J, Townsend EM, Short B, Burgess KE, et al. *Candida albicans* mycofilms support *Staphylococcus aureus* colonization and enhances miconazole resistance in dual-species interactions. *Front Microbiol*. 2017;8:258.
49. Short B, Delaney C, McKloud E, Brown JL, Kean R, Litherland GJ, et al. Investigating the transcriptome of *Candida albicans* in a dual-species *Staphylococcus aureus* biofilm model. *Front Cell Infect Microbiol*. 2021;11:791523.
50. Bamford CV, d'Mello A, Nobbs AH, Dutton LC, Vickerman MM, Jenkinson HF. *Streptococcus gordonii* modulates *Candida albicans* biofilm formation through intergeneric communication. *Infect Immun*. 2009;77(9):3696–704.
51. Harriott MM, Noverr MC. *Candida albicans* and *Staphylococcus aureus* form polymicrobial biofilms: effects on antimicrobial resistance. *Antimicrob Agents Chemother*. 2009;53(9):3914–22.
52. Souza JGS, Bertolini M, Thompson A, Mansfield JM, Grassmann AA, Maas K, et al. Role of glucosyltransferase R in biofilm interactions between *Streptococcus oralis* and *Candida albicans*. *ISME J*. 2020;14(5):1207–22.
53. Kong EF, Tsui C, Kucharíková S, Andes D, Van Dijk P, Jabra-Rizk MA. Commensal protection of *Staphylococcus aureus* against antimicrobials by *Candida albicans* biofilm matrix. *mBio*. 2016;7(5):e01365-16.
54. Domenech M, García E, Moscoso M. In vitro destruction of *Streptococcus pneumoniae* biofilms with bacterial and phage peptidoglycan hydrolases. *Antimicrob Agents Chemother*. 2011;55(9):4144–8.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Shortlisted studies.

Table S2. PERMANOVA results of beta diversity matrices.

Fig. S1. Principle co-ordinate analysis of microbiome samples by 16S sequencing region.

Fig. S2. Relative abundances of the most abundant phyla and genera in COPD and healthy patients.

Fig. S3. Receiver operator characteristic curves in the classification of lung microbiomes.

Fig. S4. Multi-species biofilms are highly tolerant to antimicrobial therapies.

Fig. S5. Combined treatments required to reduce biofilm biomass.