Metagenomics Reveals a Core Macrolide Resistome Related to Microbiota in Chronic Respiratory Disease

Micheál Mac Aogáin PhD[1]*, Kenny J. X. Lau BSc[2]*, Zhao Cai PhD[2], Jayanth Kumar Narayana BS-MS[1], Rikky W. Purbojati MSc[2], Daniela I. Drautz-Moses PhD[2], Nicolas E. Gaultier BSc[2], Tavleen K. Jaggi MD[1], Pei Yee Tiew MD[1,3], Thun How Ong MD[3], Mariko Siyue Koh MD[3], Albert Lim Yick Hou MD[4], John A. Abisheganaden MD[4], Krasimira Tsaneva-Atanasova PhD[5], Stephan C. Schuster PhD[2] and Sanjay H. Chotirmall MD PhD[1+].

*These authors contributed equally

[1]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore.

[2]Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore.

[3]Department of Respiratory and Critical Care Medicine, Singapore General Hospital, Singapore

[4]Department of Respiratory and Critical Care Medicine, Tan Tock Seng Hospital, Singapore.

[5]Department of Mathematics, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, UK.

Corresponding author: Sanjay H. Chotirmall, Lee Kong Chian School of Medicine, Nanyang Technological University, 11 Mandalay road, Singapore 308232. Email: schotirmall@ntu.edu.sg

+ Associate Editor, AJRCCM (participation complies with American Thoracic Society requirements for recusal from review and decisions for authored works).

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Data availability: All sequence data from this study has been uploaded to the National Center for Biotechnology Information (NCBI) Sequence read archives (SRA) under project accession PRJNA595703.

Author contributions: MMA and KJXL: performance and design of experiments, data analysis and interpretation, statistical analysis, writing the manuscript. CZ and JKN: performance of antimicrobial resistance gene analysis and co-occurrence network inference. TPY, OTH, AYHL: patient recruitment and procurement of clinical data and specimens. RWP, DID-M: metagenomic whole genome shotgun sequencing and analytics. NEG and TJ: curation of clinical and environmental samples and data. JAA, MSK: intellectual contributions, patient recruitment and procurement of clinical data and specimens. KTA: oversight of mathematical methods and statistics. SCS and SHC: conception and design of overall study and experiments, data analysis and interpretation, statistical analysis, writing manuscript, procurement of funding.

Support: This research is supported by the Singapore Ministry of Health's National Medical Research Council under its Clinician-Scientist Individual Research Grant (MOH-000141) (S.H.C) and Research Training Fellowship (NMRC/Fellowship/0049/2017) (P.Y.T) ; the Singapore Ministry of Education under its Singapore Ministry of Education Academic Research Fund Tier 1 (2016-T1-001-050) (S.H.C) and Tier 3 (2013-T3-013) (S.C.S and S.H.C) and the NTU Integrated Medical, Biological and Environmental Life Sciences (NIMBELS), Nanyang Technological University, Singapore [NIM/03/2018] (S.C.S. and S.H.C)

Running head: A core airway macrolide resistome relates to microbiota

Descriptor number: 10.3 Chronic Bronchial Suppurative Diseases

Word count main body: 4,287 (Abstract 233)

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

AT A GLANCE COMMENTARY

Scientific knowledge on the subject: Long-term and prophylactic antibiotics including macrolides are being increasingly used in the management of frequently exacerbating patients with chronic respiratory disease including severe asthma, COPD and bronchiectasis. The airway resistome, while recognised, is poorly characterised and its relationship to the host microbiome unknown.

What this study adds to the field: We, for the first time, in the largest and deepest metagenomics assessment of the airway evaluate airway resistomes across chronic respiratory disease states, and, relate them to host microbiomes. We identify a 'core' airway resistome, harboured by the host microbiome, and, dominated by macrolide resistance genes but with high prevalence of β-lactam, fluoroquinolone and tetracycline resistance. This 'core' resistome is independent of health status or antibiotic exposure and shares significant overlap with resistomes detected on paired patient inhaler devices where the latter represents a proxy for the host microbiome. Metagenomic analysis of the airway reveals a core macrolide resistome with implications for potential macrolide antibiotic resistance in the management of respiratory disease.

Online data supplement statement: This article has an online data supplement, which is accessible from this issue's table of contents online at www.atsjournals.org

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

ABSTRACT

Rationale: Long-term antibiotic use for managing chronic respiratory disease is increasing however the role of the airway resistome and its relationship to host microbiomes remains unknown

Objective: To evaluate airway resistomes, and, relate them to host and environmental microbiomes using ultra-deep metagenomic shotgun sequencing **Methods:** Airway specimens from n=85 individuals with and without chronic respiratory disease (severe asthma, COPD and bronchiectasis) were subjected to metagenomic sequencing to an average depth exceeding twenty million reads. Respiratory and device-associated microbiomes were evaluated based on taxonomical classification and functional annotation including the Comprehensive Antibiotic Resistance Database (CARD) to determine airway resistomes. Co-occurrence networks of gene-microbe association were constructed to determine potential microbial sources of the airway resistome. Paired patient-inhaler metagenomes were compared (n=31) to assess for the presence of airway-environment overlap in microbiomes and/or resistomes.

Results: Airway metagenomes exhibit taxonomic and metabolic diversity and distinct antimicrobial resistance patterns. A 'core' airway resistome dominated by macrolide but with high prevalence of β-lactam, fluoroquinolone and tetracycline resistance genes exist, and, is independent of disease status or antibiotic exposure. *Streptococcus* and *Actinomyces* are key potential microbial reservoirs of macrolide resistance including the *ermX, ermF* and msrD genes. Significant patient-inhaler overlap in airway microbiomes and their resistomes is identified where the latter may be a proxy for airway microbiome assessment in chronic respiratory disease.

Conclusion: Metagenomic analysis of the airway reveals a core macrolide resistome harboured by the host microbiome.

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Key words: Respiratory Disease; Metagenomics; Antimicrobial resistance; Macrolides; Resistome

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

INTRODUCTION

The development, progression and associated phenotypes of chronic airways disease has been associated with perturbation of the airway microbiome. Such airway dysbiosis is often characterised by predominance of pathogenic bacteria such as *Haemophilus, Streptococcus* and/or *Pseudomonas* which co-exist with a gamut of other taxa forming the airway 'pathobiome' (1, 2). Patients with airways disease and frequent exacerbations are therefore prescribed long-term antibiotic regimes aiming to reduce bacterial burden, inflammation and improve clinical symptoms with variable results (3-5). The selective pressure resulting from such an approach however promotes antimicrobial resistance, a key global concern and serious threat to public health (5).

The impact of antimicrobial exposure on lung microbiome architecture, and, mechanisms promoting antimicrobial resistance remain an intense area of clinical and research interest. In addition, the broader implications of antibiotic exposure on the resident airway resistome, beyond that related to the target bacterial pathogen alone, has lacked study. This is important as interspecies interactions are plentiful in the airway, and, the emergence of resistance in non-pathogenic organisms are key factors potentially influencing therapeutic outcomes (6). Moreover, the environment remains a vast, mobilizable reservoir of resistance determinants with great potential to seed the airway resistome but remains inadequately characterised in the setting of chronic respiratory disease states (7).

As next generation sequencing becomes cheaper, the era of clinical metagenomics represents an emerging and robust molecular tool allowing characterisation of airway microbiomes in tandem to assessment of their functional properties for use in diagnosis, treatment and/or patient risk stratification (8, 9). Work performed in Cystic Fibrosis (CF) reveals the cumulative effect of antibiotic exposure on the airway microbiome and recent data in chronic obstructive

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

pulmonary disease (COPD) confirms the airway as an important reservoir for antimicrobial resistance genes, however, no work has been performed using metagenomics and no dedicated studies including severe asthma or bronchiectasis (10, 11). Unlike targeted amplicon sequencing approaches (e.g. 16S rRNA) metagenomic shotgun sequencing remains less susceptible to polymerase chain reaction (PCR) amplification bias, is not influenced by copy number variation and, critically, provides scope to probe the functional aspects of the microbiome including its resistome (9, 12, 13).

Here, we report the largest application of deep metagenomic shotgun sequencing to airway samples across a range of chronic respiratory disease states (severe asthma, COPD and bronchiectasis) and include non-diseased (healthy) individuals to characterise resident airway resistomes. We provide taxonomic and functional insight to airway-based resident antibiotic resistance in health and disease, and, further assess patient inhaler devices as a potential environmental reservoir of resistance. Some of the results of these studies have been previously reported in the form of an abstract (14).

METHODS

Study Population(s)

Patients with respiratory disease were prospectively recruited during routine attendance at respiratory outpatient clinics at two tertiary hospital sites in Singapore (Table 1): Singapore General Hospital (SGH) and Tan Tock Seng Hospital (TTSH). Non-diseased (healthy) individuals were recruited through an established voluntary exercise program at Nanyang Technological University (NTU), Singapore. Severe asthma patients were graded as being at least step four of the Global Initiative for Asthma (GINA) guideline treatment ladder and met current criteria for severe asthma. COPD was defined according to the global initiative for chronic obstructive lung disease (GOLD) criteria and bronchiectasis (Non-CF) was defined by

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

radiological confirmation of bronchiectasis by either dynamic Computed tomography (CT) or High-resolution computed tomography (HRCT) thorax in accordance with British Thoracic Society guidelines along with the absence of any other major concurrent chronic respiratory disease state (15-18). Non-diseased (healthy) individuals had no active or past history of any respiratory or other medical disease and normal spirometry measured in accordance with ERS/ATS criteria. Non-diseased individuals were free from any exposure to inhaled medications or antibiotic use in the preceding 12-month period. Demographics and associated clinical data were collated including age, sex, ethnicity, body mass index (BMI), lung function (FEV_1 % predicted), smoking status and antibiotic use in the preceding six-month period. In addition, the number of exacerbations in the year preceding study recruitment was recorded for all patients with respiratory disease (severe asthma, COPD and bronchiectasis) and these patient groups had their respective disease severity/control scores recorded respectively by the Asthma Control Test (ACT), Global Initiative for Obstructive Lung Disease (GOLD) score and Bronchiectasis Severity Index (BSI) (15, 19, 20).

Whole-genome shotgun sequencing of clinical and environmental samples

Representative, spontaneously expectorated sputum samples were obtained from all participants through directed coughing using the Huff cough manoeuvre. An independent and prospective cohort of patients with severe asthma, COPD or bronchiectasis (as defined above) using regular inhaled bronchodilator and/or corticosteroid therapy were recruited and the inhaler used most frequently sampled (Supplementary Table E1). Patients had documented self-reported inhaler adherence over the preceding six-month period and demonstrated objective evidence of stable and/or improved pulmonary function on spirometry. Paired airway specimen (sputum) and an inhaler swab was obtained and subjected to metagenomic analysis. DNA was extracted from airway and environmental samples according to previously described clinical and environmental sampling methods (21, 22). DNA was used in the preparation of

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

metagenomic shotgun sequencing libraries as described (23). Resultant sequence data was processed and quality trimmed before subjecting it to secondary analysis to derive taxonomic and functional genomic profiles including analysis with specific reference to the Comprehensive Antimicrobial Resistance Database (CARD) to assess the metagenomic resistome in all patient and environmental specimens (24).

Full details on DNA extraction, metagenomic sequencing, functional and taxonomic assignment of metagenomic sequence reads, data analysis, visualization, statistical analysis and details on ethical approvals are provided in the supplementary material.

RESULTS

The airway metagenome exhibits functional metabolic dysbiosis and antibiotic resistance: Functional classification of microbial gene content based on read assignment to functional categories illustrates variability in patients with COPD and bronchiectasis in contrast to relatively comparable profiles in healthy individuals and severe asthma patients (Figure 1A). A dysbiotic shift in the abundance of carbohydrate and amino acid related pathways toward lipid-associated pathways is evident in COPD and bronchiectasis with the latter exhibiting greatest change (Figure 1A). Classification of metagenomic data with reference to the KEGG functional database further reveals alteration in genes associated with antibiotic-associated pathways including degradation, detoxification and anti-microbial resistance (25). While drug-related and xenobiotic metabolism is increased in COPD and bronchiectasis, genes involved in streptomycin, butirosin, neomycin and β-lactamase biosynthesis are abundant in severe asthma and healthy individuals (Figure 1A). To probe specifically for the composition of antibiotic resistance genes within the airway metagenome, reads were classified with reference to the CARD database – a dataset of curated antibiotic-resistance genes (24). Derived antibiotic resistance gene profiles exhibit variability across the healthy and diseased states with

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

contrasting abundances of aminoglycoside, bicyclomycin, diaminopyrimidine, multi-drug, peptide, phenicol, sulphonamides, sulfones and triclosan resistance genes (Figure 1B). Critically, recent antibiotic exposure (selective pressure) did not translate to a detectable higher relative abundance of anti-microbial resistance (AMR) and importantly, AMR profiles of healthy individuals and patients unexposed to antibiotics in the preceding six month period exhibit a significant presence of AMR suggesting the presence of a core airway resistome (Figure 1C, Table 2).

The core airway resistome: Variable assemblages of antibiotic resistance genes are identified across individual disease cohorts including healthy individuals. Patients with COPD and bronchiectasis have the greatest repertoire of antibiotic resistance determinants, although this was potentially biased by the greater number of patients in these groups (Figure 2A). Patients with COPD harboured the highest diversity of resistance genes (n=92; 85% shared with other cohorts) distinguished by the presence of specific β-lactam, fluoroquinolone, macrolide, multi-drug, phenicol and tetracycline resistance determinants. While healthy individuals and severe asthmatics had a slightly lower number of cumulative AMR sequences, analysis revealed a strikingly consistent subset of 18 AMR genes common to all cohorts (including healthy) representing a 'core resistome' (Figure 2B, Table 2). Importantly, this 'core resistome' was predominated by AMR genes from the β-lactam, fluoroquinolone, macrolide and tetracycline classes (Figure 1B, 2B and 2C, Table 2) and detectable in every individual recruited into our study irrespective of health or disease status and did not differ by type of chronic respiratory disease. In aggregate, genes encompassing a 'core macrolide resistome' were most abundant and included *msrD (mel), ermB, ermX and ermF* accompanied by genes encoding tetracycline (t*etW, tetA(46), tetB (46)*), β-lactam (*cfxA2*), and fluoroquinolone (*pmrA*) resistance (Figure 2C, Table 2).

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

The microbial ecology of sputum samples across respiratory disease states: Deep sequencing of sputum revealed a microbiome profile dominated by bacteria, with fungi and viruses accounting for <0.01% and <0.25% average relative abundance respectively (Supplementary Figure E1). Dysbiosis of the respiratory microbiome was evident across chronic respiratory disease in line with established literature (Figure 3A) (26). Microbiome α-diversity between groups varied with healthy individuals exhibiting highest Shannon diversity compared to disease (Supplementary Figure E3A). Simpson diversity was comparable across cohorts while patients with severe asthma and bronchiectasis exhibited reduced Chao1 (Supplementary Figure E3B and E3C). Among the respiratory disease cohorts, microbiome profiles were distinguished by Actinobacteria (*Rothia* spp.) and proteobacteria (*Pseudomonas* and *Haemophilus* spp.) which exhibited increased abundance in diseased cohorts with corresponding reductions in *Prevotella* spp., *Treponema* spp. *Fusobacteria* spp. and diverse firmicutes compared to healthy individuals based on linear discriminant analysis (Supplementary Figure E3D and E3E). Among diseased patients, the increased abundance of *Rothia mucilaginosa* and *Pseudomonas aeruginosa* were among the most striking species-level differences associated with disease status in our analysis (Figure 3A). Fungi were detected sporadically and at low abundance and microbiome β-diversity reveals heterogeneity to be highest in bronchiectasis, followed by COPD (Figure 3B). Conversely, healthy and severe asthma have more evenly distributed and diverse bacterial species in their airways further supported by analysis of average centroid distances, which is greatest in COPD and bronchiectasis (Figure 3C).

Gene-microbe co-occurrence in the core resistome: We next looked at correlation between microbial taxa and the 'core' resistome by constructing a co-occurrence network of microbes and their respective antibiotic resistance genes (Figure 4). This revealed significant relationships (including correlation) between resistance genes (Figure 4A) and airway

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

microbiota (Figure 4B) allowing inference of gene-microbe associations within an integrated holistic network. We focused on the 'core' macrolide resistome because it formed a key component of the 'core' resistome (Figure 2B and Figure 2C) and is an antibiotic class gaining widespread use across chronic respiratory disease states (17, 27, 28). The macrolide resistance gene *ermX*, represents a highly connected node associated with several microbial taxa (Figure 4C). These predominantly consist of upper airway commensals containing only few overtly pathogenic species (Figure 4C). Network inference further identifies a lesser number of microbial associations with *ermF* (Figure 4D) and *msrD* (Figure 4E) respectively. *Streptococci* and *Actinomyces* were associated with *ermX* and *ermB* while the strongest association detected is between *ermF* and the gut microbe *Bacteroidetes thetaiotaomicron*. Associations between *msrD* and the gut pathogen *Clostridioidies difficile* and the largely unstudied *Morococcus cerebrosus* were also detected (Figure 4D-E).

Metagenomics identifies inhaler devices as potential sinks for antibiotic resistance : To assess for potential sites of resistance host-environment transfer, patient inhaler devices with respectively paired airway specimens were obtained in an independent prospectively recruited cohort of patients with chronic respiratory disease and subjected to metagenomics sequencing and analysis (Supplementary table E1). We identify significant overlap between the airway metagenome and that detectable on patient-paired inhaler devices with a significant number of microbial taxa found on both specimens (Figure 5A). This suggests that an inhaler swab, when subjected to metagenomics could represent a surrogate measure of a patient's airway microbiome. The number of microbes observed on inhaler devices (n=207) exceeds that detected from the airways (n=116) suggestive of other potential environmental influences however, 80 overlapping microbial taxa were observable in paired airway-inhaler specimens with 36 and 127 microbial taxa found in airway and inhaler devices respectively, suggesting increased microbial diversity on inhaler devices including microbes of potential environmental

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

origin (Figure 5B). Among the 80 co-occurring microbial taxa, most (63 species and 78.8% of all co-occurring taxa) were detected at the individual level in specimens obtained from the same patient, illustrated for the most abundant taxa in Figure 5C. Similar analyses focused on antibiotic resistance genes reveals comparable metagenomic profiles from paired airway-inhaler device specimens but with lower resistance gene abundance on inhaler devices (Figure 5D). Here, multi-drug, macrolide and tetracycline antibiotic resistance determinants were most frequently observed, a finding consistent with our previously described 'core' resistome (Figure 5D). A slightly higher number of resistance determinants (n=98) were found associated to inhaler devices as compared to airway specimen (n=89), a finding consistent with the greater diversity of microbes seen on inhalers, while 53 resistance genes were overlapping between inhalers and sputum (Figure 5E). Among overlapping resistance genes between the airway-inhaler device, 86.8% (46 resistance genes) co-occurred in paired airway-inhaler device specimens illustrated for the most abundant resistance genes in Figure 5F. A subset of identified resistance genes was associated with inhaler devices without detectable levels in the sputum of any patient or non-diseased control. These genes may reflect potential environmental sources of resistance gene diversity, 12 of which were independently observed on at least two inhaler devices (Supplementary Table E2).

Co-occurrence analysis of gene-microbe associations in sputum and inhaler devices: We next sought to further assess microbial correlates of resistance among the subset of microbes (n = 63) and resistance genes (n = 46) with confirmed co-occurrence detected across the paired airway-inhaler devices (Figures 5C and 5F). Consistent with resistance gene abundance, the most highly correlated gene-microbe pairs included genes conferring resistance to macrolides (n=4), tetracyclines (n=4), β-lactams (n=1) and fluoroquinolones (n=1). Several species exhibit significant correlation with co-occurring resistance genes and predominantly comprise of Firmicutes (*Carnobacterium, Granulicatella, Prevotella* and *Streptococcus*) and

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Actinobacteria (*Actinomyces, Corynebacterium* and *Rothia*) (Figure 6A). Correlation analysis reveals significant association for all four macrolide genes (*msrD, ermF, ermB* and *ermX*) as well as fluoroquinolone (*pmrA)*, tetracycline (*tetA_46* and *tetB_46*) and β-lactam (*cfxA2*) resistance determinants. The most highly correlated species associated with *msrD* was *Carnobacterium maltaromanticum*; a lactic acid bacterium frequently found in food products including fish, meat, and some dairy. *Prevotella intermedia* was most closely correlated with *ermF* while *Actinomyces* ICM47 was associated with *ermF* and *ermX* gene abundance. *Streptococci* associated with the presence of fluoroquinolone and tetracycline resistance determinants while *Prevotella pallens* was identifiable as the most likely microbial source of *cfxA2* β-lactamase based on maximal correlation coefficients and significance compared to other taxa. The abundance of these putative microbial resistance vehicles and their associated resistance genes are illustrated in Figure 6B.

DISCUSSION

We describe the largest airway clinical metagenomics analyses performed to date linking the antimicrobial resistome to host microbiomes in chronic respiratory disease. Our work highlights the versatility and usefulness of metagenomics in assessing functional aspects of the human microbiome, including antibiotic resistance. The airway metagenome exhibits functional metabolic dysbiosis including increased antibiotic resistance, predominant in COPD and bronchiectasis. Variability exists in antibiotic-associated functional pathways in healthy and diseased airways, mirrored by a presence and high abundance of resistance genes. Critically, we uncover, even in a healthy state, the presence of a 'core' resistome; dominated by genes from common antibiotic classes including macrolides, β-lactams and fluroquinolones unrelated to antibiotic exposure. By assessing host microbiomes with increased robustness provided by metagenomics, we link the presence of specific bacterial taxa to the 'core' resistome and specifically genes conferring macrolide resistance. Analysis of paired patient-

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

inhaler metagenomes illustrates significant overlap suggesting the latter as a surrogate for the host microbiome. Patient-inhaler overlap confirms our identified resistance gene-microbe associations as well as identifies resistance determinants unique to inhaler devices aligning with the concept of the wider environment as source of resistance determinants linked to microbial, and, therefore resistance transfer between environment and host.

A key observation from this work is the high abundance of resistance genes within the Macrolide-Lincosamide-Streptogramin (MLS) axis which dominates the airway resistome. This is important given the widespread and increasing use of macrolides across a variety of clinical respiratory diseases (4, 29-31). Prior work illustrates that macrolide resistance is amongst the most prevalent in the wider environment, representing a rich source of resistance determinants with strong potential for horizontal transfer (32). Our detection of a 'core' airway resistome dominated by macrolide resistance is therefore of particular concern considering the selective pressure exerted by macrolides toward resistance in both the airway and wider environment (30, 33, 34). Recent work focused on environmental resistomes illustrates their distinct core and discriminatory elements (35). We, in similar fashion, detected core and discriminatory elements within the airway metagenome, where discriminant resistance genes had higher occurrence in diseased states (e.g multidrug resistance in COPD and bronchiectasis) which associated with their microbiomes. A core airway resistome is also evident comprising of genes from the macrolide, fluoroquinolone, β-lactam and tetracycline antibiotic classes. This core resistome demonstrates stability across all respiratory diseases including healthy individuals, and, therefore explains the detected relative microbiome stability during exacerbations despite antibiotic pressure (30, 36, 37). Our description of the 'core' airway resistome also provides a novel perspective on reasons why pathogens, expected to exhibit *in-vivo* susceptibility to a particular antibiotic, are not necessarily eradicated by appropriate antimicrobial therapy (38).

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Predominance of macrolide resistance within the core resistome is of relevance in respiratory disease. Macrolide use is advocated in severe asthma, COPD and more recently bronchiectasis, particularly for patients demonstrating recurrent and persistent exacerbations (27, 28, 39). The 'core' macrolide resistome includes *msrD (mel), ermB, ermX* and *ermF* genes which incorporates efflux (*msrD*) and rRNA methylases (*erm*), mechanisms previously established in the human gut and sewage effluent (32, 34, 40). Of all macrolide resistance genes identified, *ermX* exhibits greatest abundance in respiratory disease, and, presents as a highly connected node in our co-occurrence analysis, its presence relating to several resident airway bacteria. Interestingly, *ermF* and *mstD* exhibit association with low abundance gut microbiota including *Clostridiodes difficile* and *Bacteroides thetaiotaomicron* which are reported to harbour these genes in association with mobile genetic elements (41, 42). This suggests seeding of the airway resistome may occur, through aspiration of gut microbes, in the case of *ermF* and *msrD*, contrasting *ermX* which strongly associates more directly with the respiratory microbiota. Silent aspiration and/or gastroesophageal reflux is proposed to occur in chronic respiratory disease including asthma, COPD and bronchiectasis and therefore should be considered as a contributor to the airway resistome (43-46).

To better understand direct relationships and potential transfer between host and environmental resistomes, we next assessed metagenomes in paired patient-inhaler specimens in individual patients. We demonstrate that an inhaler swab, when subjected to metagenomics, may act as a surrogate of the airway microbiome and/or resistome: a feature of relevance in dry non-productive patients commonly seen in severe asthma but also COPD and bronchiectasis (47). Co-occurrence of microbe-resistance gene combinations was evident, strongly indicative of environmental contributors and microbial sources of the airway resistome. By generating comparable microbiome and resistome profiles between airway and inhalers, we confirm the potential for such therapeutic devices to act as resistance reservoirs, allowing trafficking of

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

pathogenic microbes and resistance determinants between the environment and airway. Recent work illustrates the COPD airway to be an important reservoir for antibiotic resistance genes, linking their abundance to bacterial colonization (11). Interestingly, existing data further proposes that macrolide resistance genes, including *ermX* may be aerosolised and associate with COPD patient filter masks in the hospital setting (48). Our work further builds on such concepts demonstrating the potential of clinical metagenomics in inhaler devices to uncover microbe-resistance gene associations through gene-species co-occurrence in environmental and airway specimens. We identify a subset of microbes and genes co-occurring between the airway and inhaler surface when integrated into a co-occurrence network leading to the identification of bacteria highly correlated to genes conferring macrolide, tetracycline, β-lactam and fluoroquinolone resistance. Interestingly, this reveals taxa previously associated with antimicrobial resistance including *Actinomyces*, *Streptococcus* and *Prevotella* species, implicating them as potential key vehicles for resistance (36, 49). *Prevotella* is known to exhibit reduced susceptibility to β-lactams in the CF lung, findings consistent with our observed association between *P. pallens* and the β-lactamase-encoding *cfxA2* gene, while the presence of *tet* genes in *Streptococcus* is also previously described (50). Association between *Carnobacterium maltaromaticum* and *msrD* is not previously described although the clinical relevance of *Carnobacterium* species remains to be fully established. Our detected relationship between *Actinomyces* species and resistance genes appears in contradiction to its described macrolide sensitivity however it should be noted that molecular (sequence) data related to *Actinomyces* spp. ICM47 has yet to be taxonomically confirmed, and, therefore potentially represents an exception to the general trend in this genera favouring macrolide susceptibility (30).

Our work is novel and represents the largest clinical metagenomics study performed on airway specimens using robust state-of-the-art methodologies effectively applied to low biomass

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

samples such as outdoor air (23). Using this approach, we uncover a core airway resistome dominated by macrolide resistance linked to the host microbiome. We further illustrate that inhaler devices may act as a surrogate of the host airway microbiome and remain a key resistance reservoir. Despite our study's strengths and novelty, we acknowledge its limitations. First, despite being the largest clinical metagenomics study to date, we include only n=85 individuals which are further nested into healthy and diseased groups, within a cross-sectional study design requiring validation in larger longitudinal studies given the myriad of confounders that could possibly influence microbiome profiles. For instance, while diseased patients were relatively well matched for age, the healthy cohort were significantly younger. Even with a limited sample size, we could identify microbiome perturbation characteristic of respiratory disease (e.g. COPD and bronchiectasis). Heterogeneity and relationships between the resistome, microbiome and specific disease phenotypes (e.g. very frequent exacerbators) within the diseased cohorts, most prominent for the bronchiectasis group, could not be fully resolved by this work because of the limited sample size. Recently, Taylor and colleagues have demonstrated the effect of macrolide exposure on the resistome in a longitudinal study of patients with severe asthma. Following azithromycin exposure, the abundance of several macrolide genes (including those observed in our analysis) *ermB, ermF, mef* and *mel (msrD)* were detected, supporting the functional importance of these genes to the resistome in response to antibiotic exposure (30). Interestingly, shifts in the resistome in response to therapy were accompanied by clinical resistance in *H. influenzea* isolates suggesting metagenomic resistome profiling may reflect clinically observed resistance that warrants further analysis in terms of diagnostic implications in the context of the findings from our study. Next, while metagenomic shotgun sequencing is a powerful tool, it remains relatively expensive with slower turnaround times compared to other sequencing approaches. This poses challenges for real-time diagnostic or therapeutic use and translation into everyday clinical practice (51). Though challenging due

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

to the nature of sample variability in particular conditions such as asthma or (dry) bronchiectasis compared to non-diseased controls, as a matrix, sputum is advantageous in terms of accessibility and scope for broad application in large studies across a range of clinical settings. This clearly contrasts with more invasive BAL or tissue biopsy sampling where sample acquisition and control subject recruitment are major limiting factors. Our study recruited several non-diseased participants from which airway specimens were readily obtained by applying the 'huff cough' technique in a protocol applied previously to non-diseased control samples in assessment of their airway microbiomes (21). Though clearly important, the degree to which sputum reflects the true microbial ecosystem of the lower airway has been the subject of debate, and, its accessibility likely comes at a cost of reduced resolution of lower airway taxa which may be particularly relevant in diseased states (52). Consistency of sputum sampling in terms of the relative proportions of upper and lower airway biomass in a given sample, variability in sampling during acquisition and also changes with respect to disease severity are all likely to influence clinical association and remain important areas for future exploration. In our assessment of patient inhalers, we included rigorous swab-sampling contamination controls but lacked the additional experimental control of a swab taken from an unused inhaler device – an important consideration given the sensitivity of metagenomics to detect minute levels of background contamination. The high abundance of human DNA in sputum is an additional hurdle with implications for adequate and unbiased detection of resistance genes, as sequencing depth will influence detection. However, recent work in the area of host DNA removal may address this issue, allowing for greater microbial read depth and scalability of sputum metagenomics in this field (53). Furthermore, metagenomic data processing requires specialised personnel with bioinformatic skills and a facility with the capability to perform high-performance computing, both significant barriers to clinical implementation. The use of short read sequencing also largely precludes a definitive

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

assessment of mobile genetic elements associated with resistance genes and long-read metagenomic workflows currently being developed may offer better insights going forward. Finally, our study detected change to genes controlling lipid metabolism across a range of chronic respiratory disease states. Alterations in lipid metabolism are identified in COPD and bronchiectasis, and, potentially contribute to their pathogenesis through initiation and resolution of inflammation. Emerging data suggests that microbial dysbiosis associates with such metabolic change and remains an important avenue for future study particularly in regard to effect on host immune function (26, 54, 55).

Despite economic, analytical and resource related challenges to the application of clinical metagenomics into respiratory practice, its ability to concurrently capture individual microbial taxonomy, function and resistance in an unbiased robust manner using a single specimen makes it an attractive tool for the delivery of precision respiratory medicine. This is important in the current era of patient endo-phenotyping including disease overlap. Our identification of a core airway resistome, dominated by macrolide resistance, is an important cautionary warning worthy of clinical consideration. Despite advances in the use of antimicrobials to improve clinical outcomes across a range of chronic respiratory diseases, we must be cognisant of their potential long-term resistance implications and weigh this against the perceived short-term clinical benefit in individual respiratory patients.

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

REFERENCES

1. Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. The Microbiome and the Respiratory Tract. *Annual review of physiology* 2016; 78: 481-504.

2. Pitlik SD, Koren O. How holobionts get sick-toward a unifying scheme of disease. *Microbiome* 2017; 5: 64.

3. Sibila O, Laserna E, Shoemark A, Keir HR, Finch S, Rodrigo-Troyano A, Perea L, Lonergan M, Goeminne PC, Chalmers JD. Airway Bacterial Load and Inhaled Antibiotic Response in Bronchiectasis. *American journal of respiratory and critical care medicine* 2019; 200: 33-41.

4. Herath SC, Normansell R, Maisey S, Poole P. Prophylactic antibiotic therapy for chronic obstructive pulmonary disease (COPD). *The Cochrane database of systematic reviews* 2018; 10: Cd009764.

5. Miravitlles M, Anzueto A. Antibiotics for Acute and Chronic Respiratory Infection in Patients with Chronic Obstructive Pulmonary Disease. *American Journal of Respiratory and Critical Care Medicine* 2013; 188: 1052-1057.

6. Vandeplassche E, Tavernier S, Coenye T, Crabbé A. Influence of the lung microbiome on antibiotic susceptibility of cystic fibrosis pathogens. *European Respiratory Review* 2019; 28: 190041.

7. Surette MD, Wright GD. Lessons from the Environmental Antibiotic Resistome. *Annu Rev Microbiol* 2017; 71: 309-329.

8. Forbes JD, Knox NC, Peterson CL, Reimer AR. Highlighting Clinical Metagenomics for Enhanced Diagnostic Decision-making: A Step Towards Wider Implementation. *Comput Struct Biotechnol J* 2018; 16: 108-120.

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

9. Chiu CY, Miller SA. Clinical metagenomics. *Nat Rev Genet* 2019; 20: 341-355.

10. Feigelman R, Kahlert CR, Baty F, Rassouli F, Kleiner RL, Kohler P, Brutsche MH, von Mering C. Sputum DNA sequencing in cystic fibrosis: non-invasive access to the lung microbiome and to pathogen details. *Microbiome* 2017; 5: 20.

11. Ramsheh MY, Haldar K, Bafadhel M, George L, Free RC, John C, Reeve NF, Ziegler-Heitbrock L, Gut I, Singh D, Mistry V, Tobin MD, Oggioni MR, Brightling C, Barer MR. Resistome analyses of sputum from COPD and healthy subjects reveals bacterial load-related prevalence of target genes. *Thorax* 2019.

12. Pinto AJ, Raskin L. PCR Biases Distort Bacterial and Archaeal Community Structure in Pyrosequencing Datasets. *PLOS ONE* 2012; 7: e43093.

13. Hong S, Bunge J, Leslin C, Jeon S, Epstein SS. Polymerase chain reaction primers miss half of rRNA microbial diversity. *Isme j* 2009; 3: 1365-1373.

14. Mac Aogáin M, Zhao C, Purbojati RW, Drautz-Moses DI, Lim AYH, Low TB, Ong TH, Koh MS, Abisheganaden JA, Liang Y, Schuster SC, Chotirmall SH. The airway 'resistome' in chronic respiratory disease: A metagenomics approach. *Eur Respir J* 2019.

15. GOLD. Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2018. ; 2018 (http://goldcopd.org/).

16. Pasteur MC, Bilton D, Hill AT, British Thoracic Society Bronchiectasis non CFGG. British Thoracic Society guideline for non-CF bronchiectasis. *Thorax* 2010; 65 Suppl 1: i1-58.

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

17. Chung KF, Wenzel SE, Brozek JL, Bush A, Castro M, Sterk PJ, Adcock IM, Bateman ED, Bel EH, Bleecker ER, Boulet LP, Brightling C, Chanez P, Dahlen SE, Djukanovic R, Frey U, Gaga M, Gibson P, Hamid Q, Jajour NN, Mauad T, Sorkness RL, Teague WG. International ERS/ATS guidelines on definition, evaluation and treatment of severe asthma. *The European respiratory journal* 2014; 43: 343-373.

18. Global Initiative for Asthma. Global Management and Prevention 2019. 2019 17/10/2019]. Available from: www.ginasthma.org.

19. Chalmers JD, Goeminne P, Aliberti S, McDonnell MJ, Lonni S, Davidson J, Poppelwell L, Salih W, Pesci A, Dupont LJ, Fardon TC, De Soyza A, Hill AT. The bronchiectasis severity index. An international derivation and validation study. *American journal of respiratory and critical care medicine* 2014; 189: 576-585.

20. Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, Marcus P, Murray JJ, Pendergraft TB. Development of the asthma control test: a survey for assessing asthma control. *J Allergy Clin Immunol* 2004; 113: 59-65.

21. Mac Aogáin M, Chandrasekaran R, Lim AYH, Low TB, Tan GL, Hassan T, Ong TH, Hui Qi Ng A, Bertrand D, Koh JY, Pang SL, Lee ZY, Gwee XW, Martinus C, Sio YY, Matta SA, Chew FT, Keir HR, Connolly JE, Abisheganaden JA, Koh MS, Nagarajan N, Chalmers JD, Chotirmall SH. Immunological corollary of the pulmonary mycobiome in bronchiectasis: the CAMEB study. *The European respiratory journal* 2018; 52.

22. Luhung I, Wu Y, Ng CK, Miller D, Cao B, Chang VW-C. Protocol Improvements for Low Concentration DNA-Based Bioaerosol Sampling and Analysis. *PLOS ONE* 2015; 10: e0141158.

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

23. Gusareva ES, Acerbi E, Lau KJX, Luhung I, Premkrishnan BNV, Kolundzija S, Purbojati RW, Wong A, Houghton JNI, Miller D, Gaultier NE, Heinle CE, Clare ME, Vettath VK, Kee C, Lim SBY, Chenard C, Phung WJ, Kushwaha KK, Nee AP, Putra A, Panicker D, Yanqing K, Hwee YZ, Lohar SR, Kuwata M, Kim HL, Yang L, Uchida A, Drautz-Moses DI, Junqueira ACM, Schuster SC. Microbial communities in the tropical air ecosystem follow a precise diel cycle. *Proceedings of the National Academy of Sciences of the United States of America* 2019.

24. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, Lago BA, Dave BM, Pereira S, Sharma AN, Doshi S, Courtot M, Lo R, Williams LE, Frye JG, Elsayegh T, Sardar D, Westman EL, Pawlowski AC, Johnson TA, Brinkman FS, Wright GD, McArthur AG. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res* 2017; 45: D566-d573.

25. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000; 28: 27-30.

26. Budden KF, Shukla SD, Rehman SF, Bowerman KL, Keely S, Hugenholtz P, Armstrong-James DPH, Adcock IM, Chotirmall SH, Chung KF, Hansbro PM. Functional effects of the microbiota in chronic respiratory disease. *The Lancet Respiratory medicine* 2019; 7: 907-920.

27. Wedzicha JA, Ritchie AI, Martinez FJ. Can Macrolide Antibiotics Prevent Hospital Readmissions? *Am J Respir Crit Care Med* 2019; 200: 796-798.

28. Chalmers JD, Boersma W, Lonergan M, Jayaram L, Crichton ML, Karalus N, Taylor SL, Martin ML, Burr LD, Wong C, Altenburg J. Long-term macrolide antibiotics for the

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

treatment of bronchiectasis in adults: an individual participant data meta-analysis. *Lancet Respir Med* 2019; 7: 845-854.

29. Laska IF, Crichton ML, Shoemark A, Chalmers JD. The efficacy and safety of inhaled antibiotics for the treatment of bronchiectasis in adults: a systematic review and meta-analysis. *The Lancet Respiratory medicine* 2019.

30. Taylor SL, Leong LEX, Mobegi FM, Choo JM, Wesselingh S, Yang IA, Upham JW, Reynolds PN, Hodge S, James AL, Jenkins C, Peters MJ, Baraket M, Marks GB, Gibson PG, Rogers GB, Simpson JL. Long-Term Azithromycin Reduces Haemophilus influenzae and Increases Antibiotic Resistance in Severe Asthma. *American journal of respiratory and critical care medicine* 2019; 200: 309-317.

31. Welte T. Azithromycin: The Holy Grail to Prevent Exacerbations in Chronic Respiratory Disease? *Am J Respir Crit Care Med* 2019; 200: 269-270.

32. Hendriksen RS, Munk P, Njage P, van Bunnik B, McNally L, Lukjancenko O, Roder T, Nieuwenhuijse D, Pedersen SK, Kjeldgaard J, Kaas RS, Clausen P, Vogt JK, Leekitcharoenphon P, van de Schans MGM, Zuidema T, de Roda Husman AM, Rasmussen S, Petersen B, Global Sewage Surveillance project c, Amid C, Cochrane G, Sicheritz-Ponten T, Schmitt H, Alvarez JRM, Aidara-Kane A, Pamp SJ, Lund O, Hald T, Woolhouse M, Koopmans MP, Vigre H, Petersen TN, Aarestrup FM. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat Commun* 2019; 10: 1124.

33. Milakovic M, Vestergaard G, Gonzalez-Plaza JJ, Petric I, Simatovic A, Senta I, Kublik S, Schloter M, Smalla K, Udikovic-Kolic N. Pollution from azithromycin-manufacturing promotes macrolide-resistance gene propagation and induces spatial and seasonal

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

bacterial community shifts in receiving river sediments. *Environ Int* 2019; 123: 501-511.

34. Roberts MC. Environmental macrolide-lincosamide-streptogramin and tetracycline resistant bacteria. *Front Microbiol* 2011; 2: 40.

35. Gupta S, Arango-Argoty G, Zhang L, Pruden A, Vikesland P. Identification of discriminatory antibiotic resistance genes among environmental resistomes using extremely randomized tree algorithm. *Microbiome* 2019; 7: 123.

36. Choo JM, Abell GCJ, Thomson R, Morgan L, Waterer G, Gordon DL, Taylor SL, Leong LEX, Wesselingh SL, Burr LD, Rogers GB. Impact of Long-Term Erythromycin Therapy on the Oropharyngeal Microbiome and Resistance Gene Reservoir in Non-Cystic Fibrosis Bronchiectasis. *mSphere* 2018; 3.

37. Cox MJ, Turek EM, Hennessy C, Mirza GK, James PL, Coleman M, Jones A, Wilson R, Bilton D, Cookson WO, Moffatt MF, Loebinger MR. Longitudinal assessment of sputum microbiome by sequencing of the 16S rRNA gene in non-cystic fibrosis bronchiectasis patients. *PloS one* 2017; 12: e0170622.

38. Chalmers JD. Macrolide resistance in Pseudomonas aeruginosa: implications for practice. *Eur Respir J* 2017; 49.

39. Holguin F, Cardet JC, Chung KF, Diver S, Ferreira DS, Fitzpatrick A, Gaga M, Kellermeyer L, Khurana S, Knight S, McDonald VM, Morgan RL, Ortega VE, Rigau D, Subbarao P, Tonia T, Adcock IM, Bleecker ER, Brightling C, Boulet LP, Cabana M, Castro M, Chanez P, Custovic A, Djukanovic R, Frey U, Frankemolle B, Gibson P, Hamerlijnck D, Jarjour N, Konno S, Shen H, Vitary C, Bush A. Management of Severe Asthma: a European Respiratory Society/American Thoracic Society Guideline. *Eur Respir J* 2019.

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

40. Ohashi Y, Fujisawa T. Detection of antibiotic resistance genes in the feces of young adult Japanese. *Biosci Microbiota Food Health* 2017; 36: 151-154.

41. Lofmark S, Jernberg C, Jansson JK, Edlund C. Clindamycin-induced enrichment and long-term persistence of resistant Bacteroides spp. and resistance genes. *J Antimicrob Chemother* 2006; 58: 1160-1167.

42. Isidro J, Menezes J, Serrano M, Borges V, Paixao P, Mimoso M, Martins F, Toscano C, Santos A, Henriques AO, Oleastro M. Genomic Study of a Clostridium difficile Multidrug Resistant Outbreak-Related Clone Reveals Novel Determinants of Resistance. *Front Microbiol* 2018; 9: 2994.

43. Raghu G, Meyer KC. Silent gastro-oesophageal reflux and microaspiration in IPF: mounting evidence for anti-reflux therapy? *Eur Respir J* 2012; 39: 242-245.

44. Lee AL, Goldstein RS. Gastroesophageal reflux disease in COPD: links and risks. *Int J Chron Obstruct Pulmon Dis* 2015; 10: 1935-1949.

45. Mandal P, Morice AH, Chalmers JD, Hill AT. Symptoms of airway reflux predict exacerbations and quality of life in bronchiectasis. *Respir Med* 2013; 107: 1008-1013.

46. Leggett JJ, Johnston BT, Mills M, Gamble J, Heaney LG. Prevalence of gastroesophageal reflux in difficult asthma: relationship to asthma outcome. *Chest* 2005; 127: 1227-1231.

47. Aliberti S, Lonni S, Dore S, McDonnell MJ, Goeminne PC, Dimakou K, Fardon TC, Rutherford R, Pesci A, Restrepo MI, Sotgiu G, Chalmers JD. Clinical phenotypes in adult patients with bronchiectasis. *The European respiratory journal* 2016; 47: 1113-1122.

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

48. Kennedy M, Ramsheh MY, Williams CML, Auty J, Haldar K, Abdulwhhab M, Brightling CE, Barer MR. Face mask sampling reveals antimicrobial resistance genes in exhaled aerosols from patients with chronic obstructive pulmonary disease and healthy volunteers. *BMJ Open Respir Res* 2018; 5: e000321.

49. Baron S, Diene S, Rolain J-M. Human microbiomes and antibiotic resistance. *Human Microbiome Journal* 2018.

50. Montanari MP, Cochetti I, Mingoia M, Varaldo PE. Phenotypic and molecular characterization of tetracycline- and erythromycin-resistant strains of Streptococcus pneumoniae. *Antimicrob Agents Chemother* 2003; 47: 2236-2241.

51. Greninger AL. The challenge of diagnostic metagenomics. *Expert review of molecular diagnostics* 2018; 18: 605-615.

52. Carney SM, Clemente JC, Cox MJ, Dickson RP, Huang YJ, Kitsios GD, Kloepfer KM, Leung JM, LeVan TD, Molyneaux PL, Moore BB, O'Dwyer DN, Segal LN, Garantziotis S. Methods in Lung Microbiome Research. *American journal of respiratory cell and molecular biology* 2019.

53. Nelson MT, Pope CE, Marsh RL, Wolter DJ, Weiss EJ, Hager KR, Vo AT, Brittnacher MJ, Radey MC, Hayden HS, Eng A, Miller SI, Borenstein E, Hoffman LR. Human and Extracellular DNA Depletion for Metagenomic Analysis of Complex Clinical Infection Samples Yields Optimized Viable Microbiome Profiles. *Cell reports* 2019; 26: 2227-2240.e2225.

54. Chen H, Li Z, Dong L, Wu Y, Shen H, Chen Z. Lipid metabolism in chronic obstructive pulmonary disease. *International journal of chronic obstructive pulmonary disease* 2019; 14: 1009-1018.

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

55. Olveira G, Olveira C, Dorado A, Garcia-Fuentes E, Rubio E, Tinahones F, Soriguer F, Murri M. Cellular and plasma oxidative stress biomarkers are raised in adults with bronchiectasis. *Clinical nutrition (Edinburgh, Scotland)* 2013; 32: 112-117.

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

FIGURE LEGENDS

Figure 1. Airway shotgun metagenomics reveals functional metabolic dysbiosis and an increased antibiotic resistance gene abundance across chronic respiratory disease states. (A) Heatmap illustrating the relative abundance of functionally classified sequence reads assigned to functional categories (KEGG) in each microbiome profile. Values are expressed as Z-scores (calculated based on the deviation from the mean abundance in each group and scaled to the standard deviation). Higher abundance (indicated in red) associated with specific functional pathways including lipid metabolism, xenobiotic biodegradation and antibiotic-associated biosynthetic pathways. These are highest in patients with COPD and bronchiectasis. (B) Heatmap illustrating specific antibiotic-resistance gene abundance (by class) based on read alignment to the CARD antibiotic resistance database. β-lactam, fluoroquinolone, macrolide and tetracycline resistance genes are detectable in diseased and non-diseased subjects. Patients with COPD and bronchiectasis have the highest load of antibiotic resistance determinants. (C) Patient antibiotic usage and respective class (in the preceding six months preceding airway sampling) is indicated by black dots (•). COPD: Chronic Obstructive Pulmonary Disease.

Figure 2. A core airway resistome exists across respiratory disease states including antibiotic-naïve and non-diseased (healthy) individuals. (A) Venn diagram illustrating the number of individual antibiotic resistance genes among the study cohorts and their intersections; ND: Non-Diseased, SA: Severe Asthma, COPD: Chronic Obstructive Pulmonary Disease and BE: Bronchiectasis. (B) An *Upset* plot, corresponding to the presented Venn diagram (2A) illustrating the antibiotic resistance gene composition across individual cohorts and their intersections. Stacked bar charts reflect the detected antibiotic resistance genes coloured according to antibiotic class. Individual groups and their intersections are indicated for each cohort separately (ND; SA; COPD and BE) followed by their respective intersection by a matrix (located below stacked bars). Set size, i.e. the number of resistance genes detected per

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

group is indicated by horizontal bars (ND<SA<COPD<BE). Black dots (•) indicate sets and connecting lines indicate relevant intersections related to each stacked bar chart. An 18-gene 'core' resistome was identified (across all four cohorts) and largely comprises of genes conferring macrolide, tetracycline, β-lactam and aminoglycoside resistance, while the 32 genes shared by COPD and bronchiectasis patients are predominantly multi-drug and triclosan resistance classes. (C) Heatmap illustrating specific antibiotic resistance genes by class and individual cohort. Specific antibiotic resistance genes grouped by coloured class (x-axis) are plotted against individual cohorts (ND; SA; COPD and BE) (y-axis). Genes are presented in order of detected abundance with *msrD (mel)*, *ErmB*, *ErmF*, and *ErmX* macrolide resistance genes most frequently observed across all four cohorts followed by genes encoding tetracycline, β-lactam and fluoroquinolone resistance.

Figure 3. Metagenomic microbiome taxonomic composition exhibits disease-associated signatures with greatest heterogeneity in COPD and bronchiectasis. (A) Bubble chart illustrating microbial abundance of discriminant taxa in non-diseased *vs* diseased cohorts based on species-level classification. Bubble size corresponds to read count and phylum membership is colour-coded. *Rothia mucilaginosa* was consistently increased in diseased versus non-diseased subjects (Dunn's test, $p = 0.01$) while *Pseudomonas aeruginosa* was also increased, most notably among bronchiectasis patients (Dunne test's, $p = 0.02$) (B) A non-metric multi-dimensional scaling plot (NMDS) illustrating β-diversity between the study groups including ND: Non-Diseased (dark blue), SA: Severe Asthma (light blue), COPD: Chronic Obstructive Pulmonary Disease (purple) and BE: Bronchiectasis (red), each highlighted by a coloured ellipse (C) The average distances to centroid from the NMDS plot was measured using the PERMDISP test of homogeneity and plotted for each respective study group illustrating the heterogeneity of their respective microbiome profiles (error bars reflect standard deviation).

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Difference in average distance to centroid was formally confirmed by ANOVA and Tukey post-hoc analysis; ***p <0.001.

Figure 4. Network inference through co-occurrence analysis reveals gene-microbe associations of the 'core' macrolide resistome. (A) Antibiotic resistance genes within the co-occurrence network are colour-coded with respect to antibiotic class while microbes are coloured black. Grey lines denote interactions between nodes (representing both microbes and resistance genes) with line thickness reflecting their observed interaction strength respectively. Interactions between resistance genes are highlighted by red lines. (B) Microbes within the co-occurrence network are colour-coded with respect to their species while antibiotic resistance genes are coloured black. Grey lines denote interactions between nodes (microbes or resistance genes) with thickness reflecting interaction strength. Interactions between species are highlighted by red lines. (C-E) Identified nodes of the macrolide resistome are highlighted indicating the specific microbes (by species) that associate with (C) *ermX*, (D) *ermF* and (E) *msrD.* Line thickness reflects the observed interaction strength between microbial nodes and the central resistance gene while arrow heads depict directionality of the co-occurrence prediction.

Figure 5. Metagenomics assessment of inhaler devices as potential antibiotic resistance reservoirs. Metagenomics shotgun analyses was performed on paired (airway-inhaler) specimens obtained through sputum collection and swabbing of mouthpieces of patient inhaler devices (n=total 31 pairs consisting of n=16 (severe asthma); n=11 (COPD) and n=4 (bronchiectasis). (A) Microbiome profiles of paired airway-inhaler devices exhibits a comparable overall pattern illustrated by stacked bar plots of a species-level relative abundance. (B) Venn diagram illustrating the observed metagenomics-derived microbial taxa present in the airway (green set, 'A', n = 116) and inhaler device (grey set, 'I', n = 207) and the co-occurrence of microbial species that are detectable in both groups (intersect, n = 80). Thirty-

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

six and 127 species were therefore unique to the airway sputum and inhaler metagenomics profiles respectively. (C) Horizontal bar plot indicating microbial species confirmed to co-occur in paired specimens, i.e. species found in both the airway and inhaler device of the same patient (n = 63 species). (D) Resistance gene profiles for paired airway-inhaler devices demonstrate comparability with a higher abundance of resistance genes (measured in RPKM - Reads Per Kilobase Million) detected in airway specimens. (E) Venn diagram illustrating the observed diversity of resistance genes detected in airway specimens (green set, 'A', n = 89) and inhaler devices (grey set, 'I', n = 98) by metagenomics. Co-occurrence of a significant number of microbial species were detected (intersect, n = 53 species). Thirty-six and 45 resistance genes were unique to the airway sputum and inhaler metagenomics profiles respectively. (F) Horizontal bar plot indicating resistance genes confirmed to co-occur in paired specimens, i.e. genes found in both the airway and inhaler device of the same patient (n = 46 genes). Genes co-occurrences observed in $n \geq 2$ subjects are plotted.

Figure 6. Metagenomic derivation of microbe-gene associations highlighting a potential source of resistance implicated in airway-inhaler device cross-over. (A) Correlation plot of microbes and resistance genes co-identified in metagenomic profiles from paired patient airway and inhaler device specimens. The presence of a circle indicates significant association ($p<0.05$) while circle size and colour intensity reflect observed Pearson's correlation for all pairwise comparisons indicating the strong positive correlations detected between microbes and resistance genes. The antibiotic resistance genes are colour coded according to their respective antibiotic class (B) Bubble chart illustrating the co-occurrence of the most highly correlated microbe (empty circle) and resistance gene (filled circle) combinations illustrated by disease i.e. Severe Asthma (SA), COPD and bronchiectasis (BE). Bubble size represents the number of classified reads while colour indicates the antibiotic class. Black bars along x-

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

axis indicate each individual paired airway and inhaler specimen respectively (from left to right).

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Table 1. Patient Demographics

Demographic	Non-diseased (ND)	Diseased (D)			*p-value*	
		Severe Asthma	COPD	Bronchiectasis	ND vs D	D v D
N	13	11	15	15		
Age	34±8	70±17	70±9	64±15	**<0.001**	0.472
Gender					0.528	**0.012**
Male	62% (8)	73% (8)	100% (15)	53% (8)		
Female	38% (5)	27% (3)	0% (0)	47% (7)		
BMI	23±3	27.9±8	21.7±8	18.4±6	0.237	**0.009**
FEV_1 predicted	109±14	72±14	45±13	65±20	**<0.001**	**0.005**
Disease severity †						
Mild	-	0% (0)	7% (1)	6% (1)	-	0.797
Moderate	-	27% (3)	40% (6)	27% (4)		
Severe	-	73% (8)	53% (8)	67% (10)		
Exacerbations in year preceding study recruitment	0±0	0±1.5	3±3	2±1	**0.032**	**0.013**
Smoking status					**<0.001**	**<0.001**
Never smoker	100% (13)	91% (10)	0% (0)	67% (10)		
Ex-smoker	0% (0)	0% (0)	47% (7)	27% (4)		
Current smoker	0% (0)	9% (1)	53% (8)	7% (1)		
Antibiotic use in 6 months preceding recruitment					0.050	**<0.001**
Yes	0% (0)	0% (0)	67% (10)	20% (3)		
No	100% (13)	100% (11)	33% (5)	80% (12)		
Inhaled corticosteroid use						
Yes	-	100% (11)	53% (8)	53% (8)	-	0.064
No	-	0% (0)	47% (7)	47% (7)		
Inhaled bronchodilator use						
Yes	-	100% (11)	100% (15)	53% (8)	-	**0.009**
No	-	0% (0)	0% (0)	47% (7)		

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

† Defined according to disease-specific criteria. Severe asthma: ACT score; ≥ 20 = “mild”, 19-15 = “moderate”, <15 = “severe”. COPD: GOLD stage; 1 = “mild”, 2 = “moderate”, ≥ 3 = “severe”. Bronchiectasis: Bronchiectasis Severity Index (BSI); 0-4 = “mild”, 5-8 = “moderate”, ≥ 9 = “severe”.

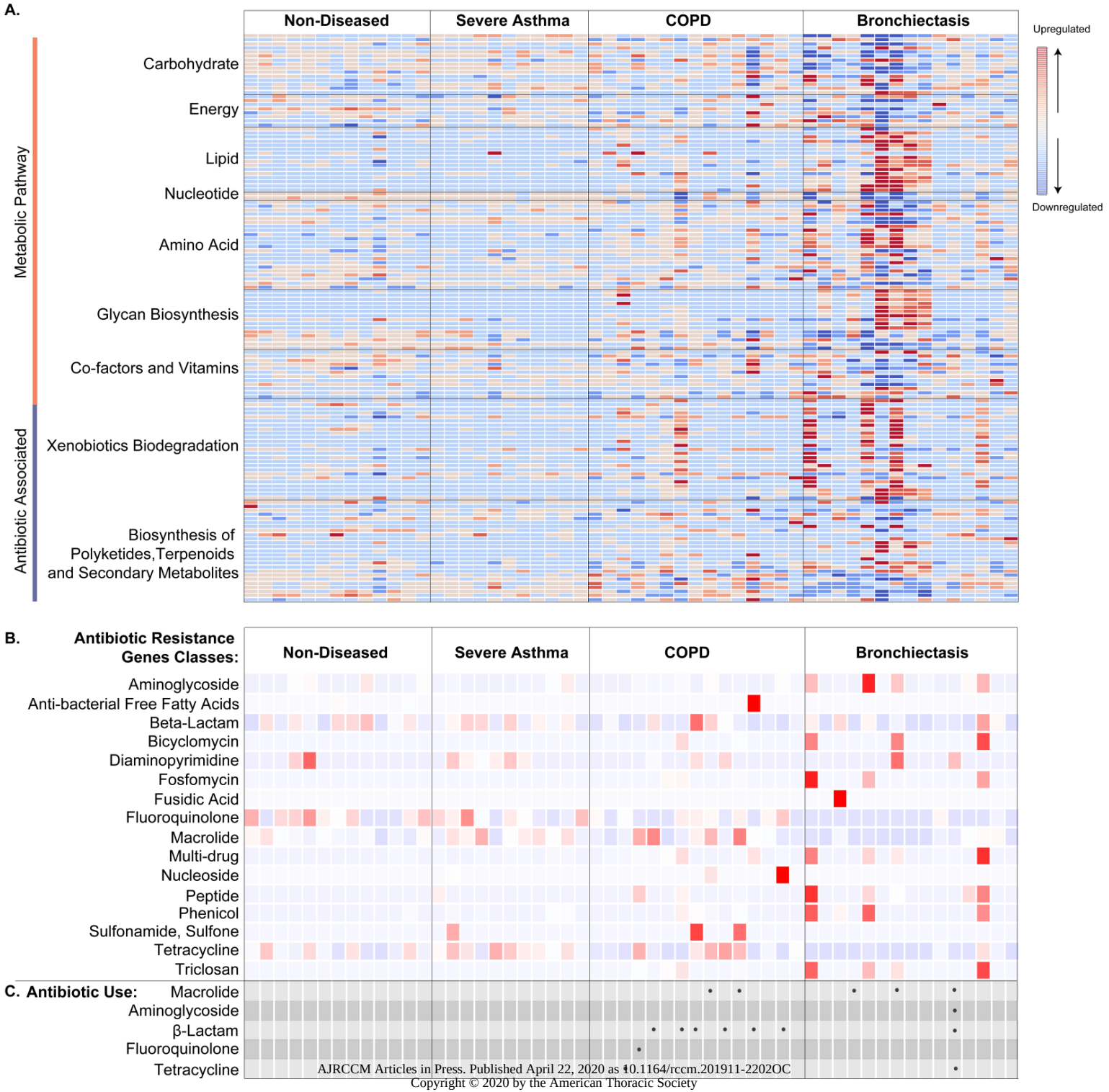
AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Table 2. Genes of the core 'resistome' (n =18) identified across study cohorts.

Gene name	Drug class	Resistance mechanism
aac(3)-VIIa	aminoglycoside	antibiotic inactivation
aph(3)-IIIa	aminoglycoside	antibiotic inactivation
oxa-255	cephalosporin, penam (β-lactam)	antibiotic inactivation
cfxA2	cephamycin (β-lactam)	antibiotic inactivation
dfrA1	diaminopyrimidine	antibiotic target replacement
pmrA	fluoroquinolone	antibiotic efflux
mel (msrD)	macrolide, lincosamide, streptogramin, tetracycline, phenicol, oxazolidinone, pleuromutilin	antibiotic target protection
ermB	macrolide, lincosamide, streptogramin	antibiotic target alteration
ermF	macrolide, lincosamide, streptogramin	antibiotic target alteration
ermX	macrolide, lincosamide, streptogramin	antibiotic target alteration
lnuC	lincosamide	antibiotic inactivation
efrB	rifamycin, macrolide, fluoroquinolone	antibiotic efflux
catS	phenicol	antibiotic inactivation
tetA(46)	tetracycline	antibiotic efflux
tetW	tetracycline	antibiotic target protection
tetB(46)	tetracycline	antibiotic efflux
tet(D)	tetracycline	antibiotic efflux
tetO	tetracycline	antibiotic target protection

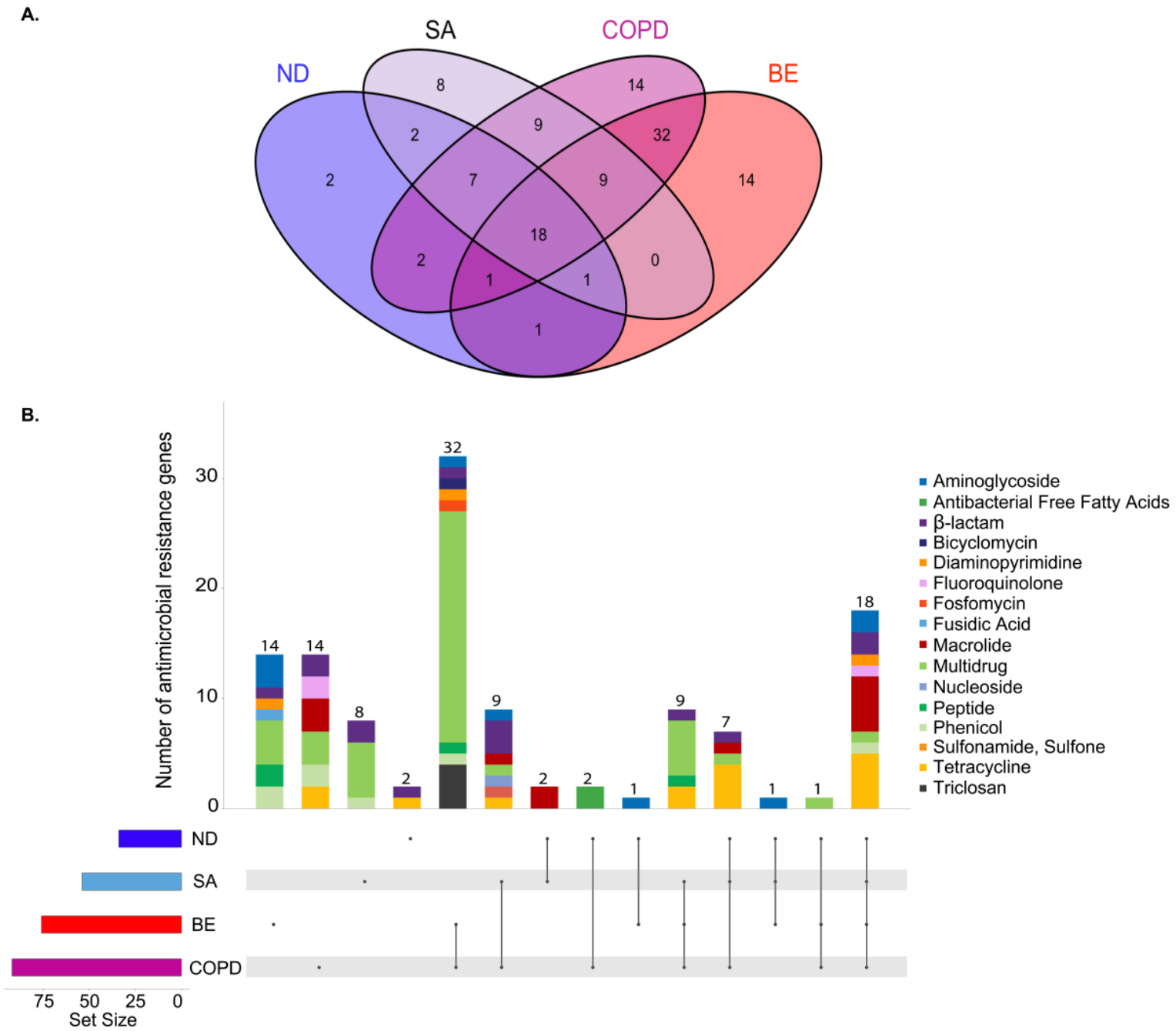
AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Figure 1



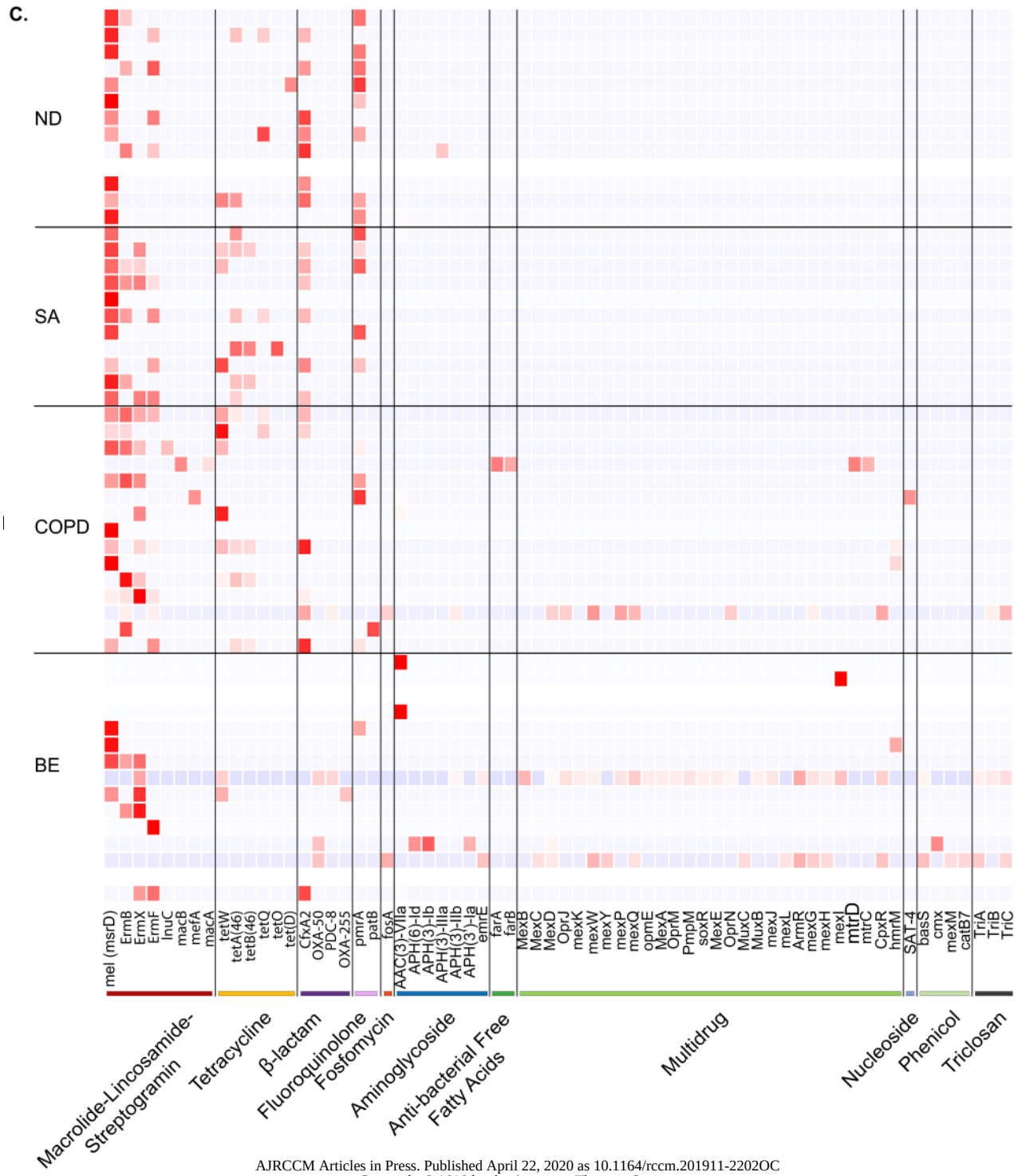
AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Figure 2 (a-b)



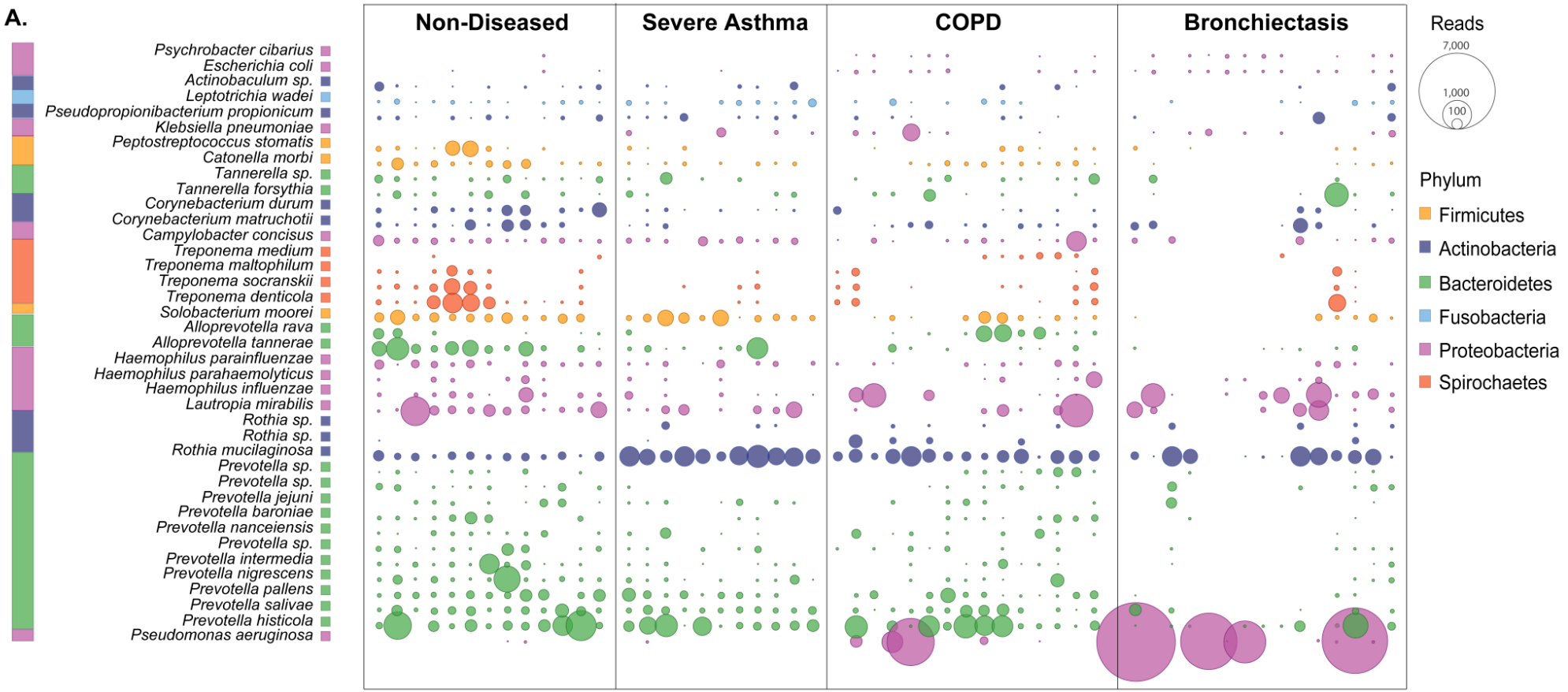
AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Figure 2 (c)



AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

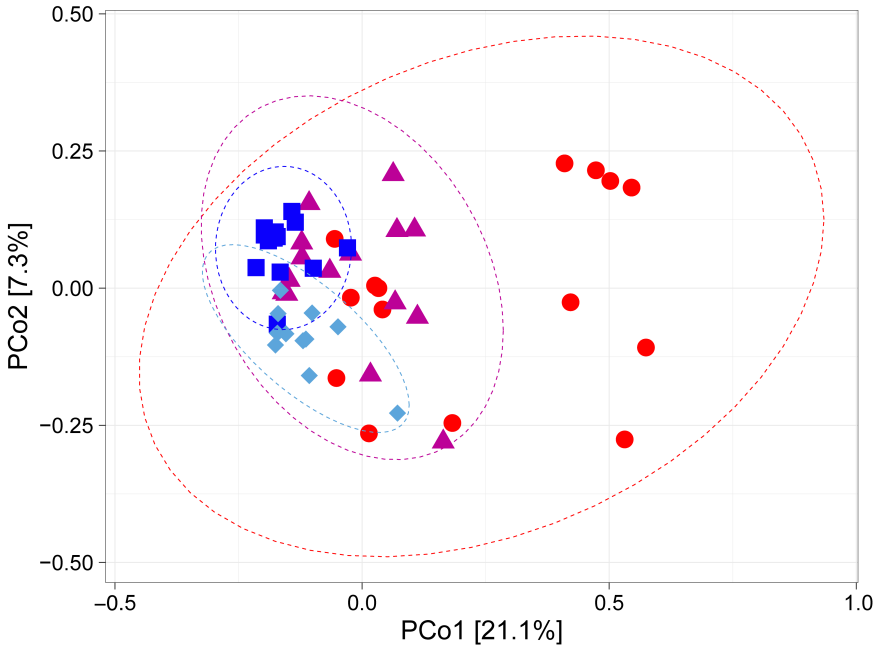
Figure 3 (a)



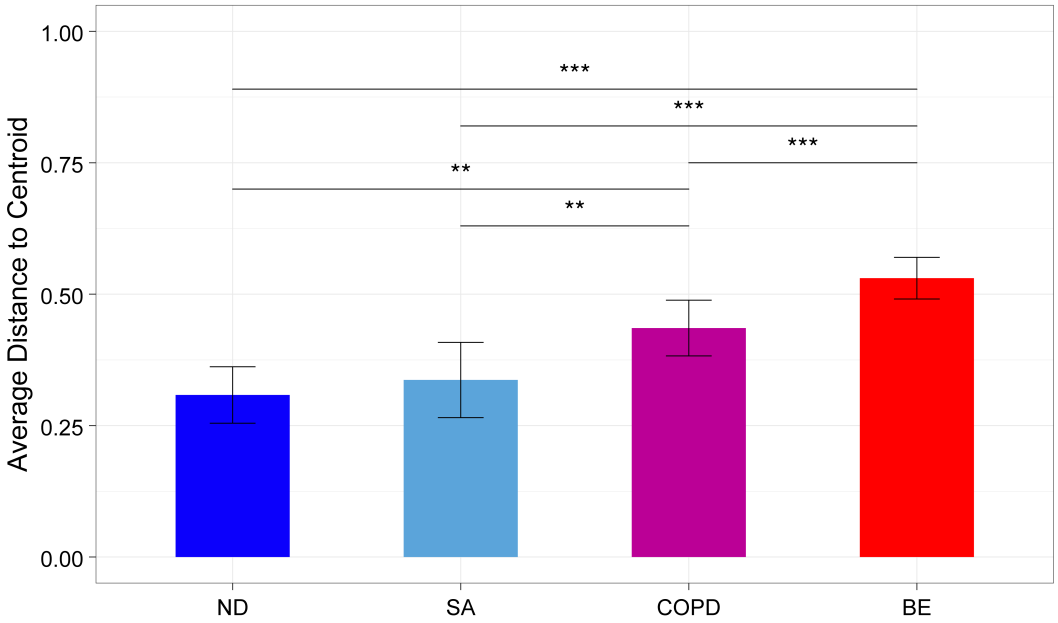
AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Figure 3 (b-c)

B.

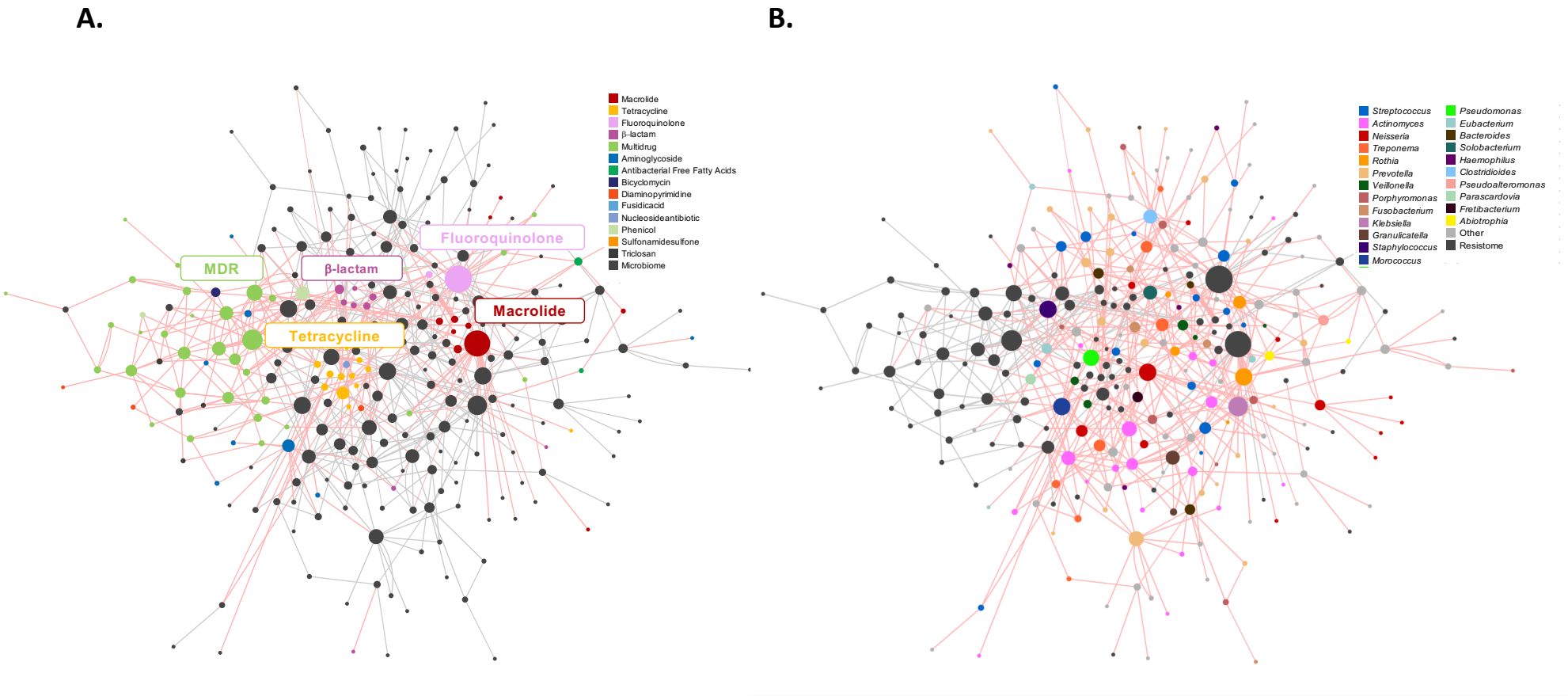


C.



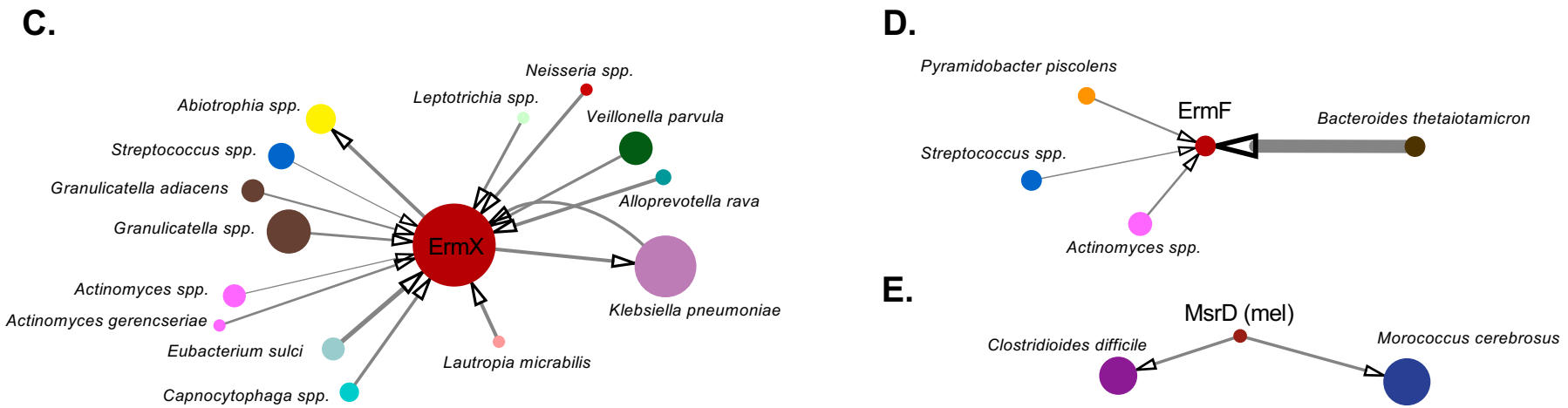
AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Figure 4 (a-b)



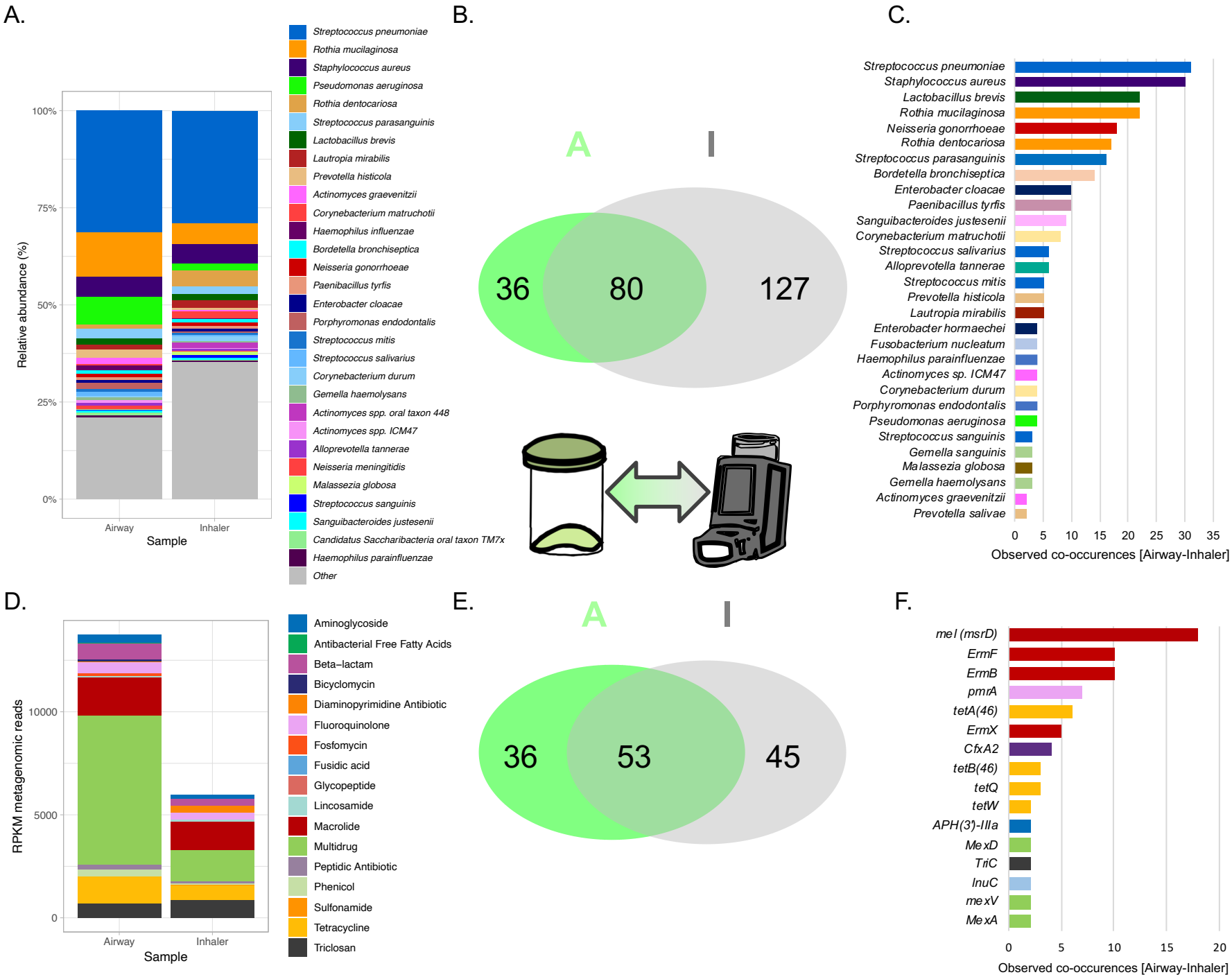
AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Figure 4 (d-e)



AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

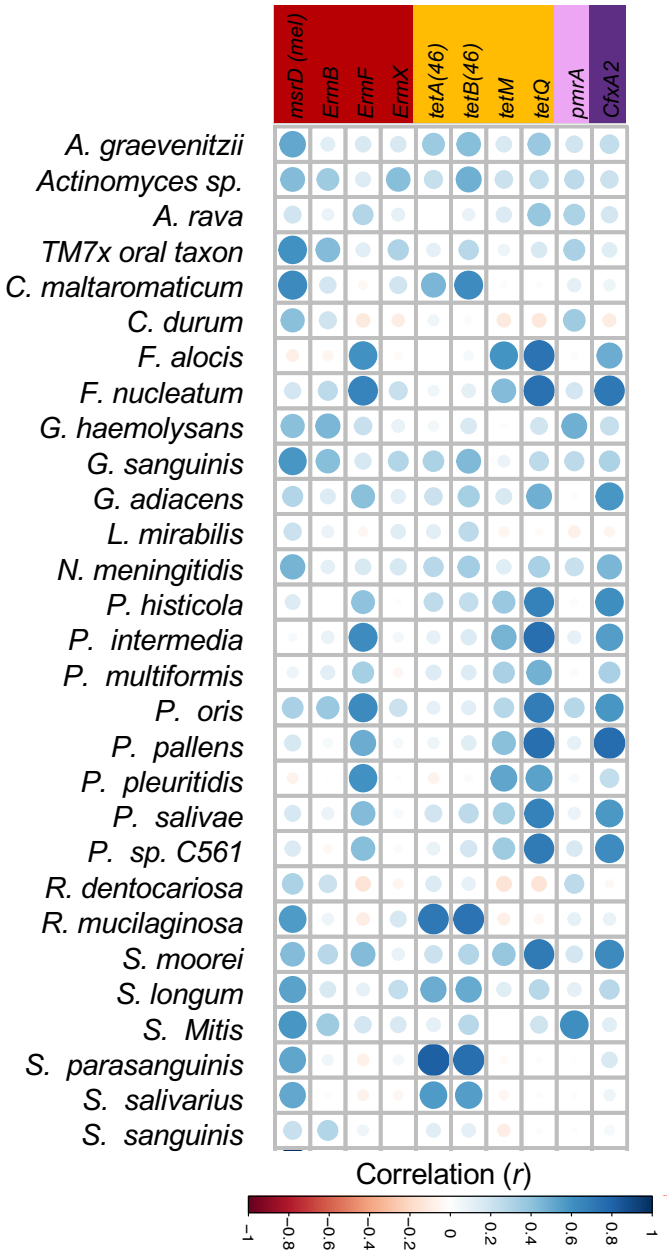
Figure 5



AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

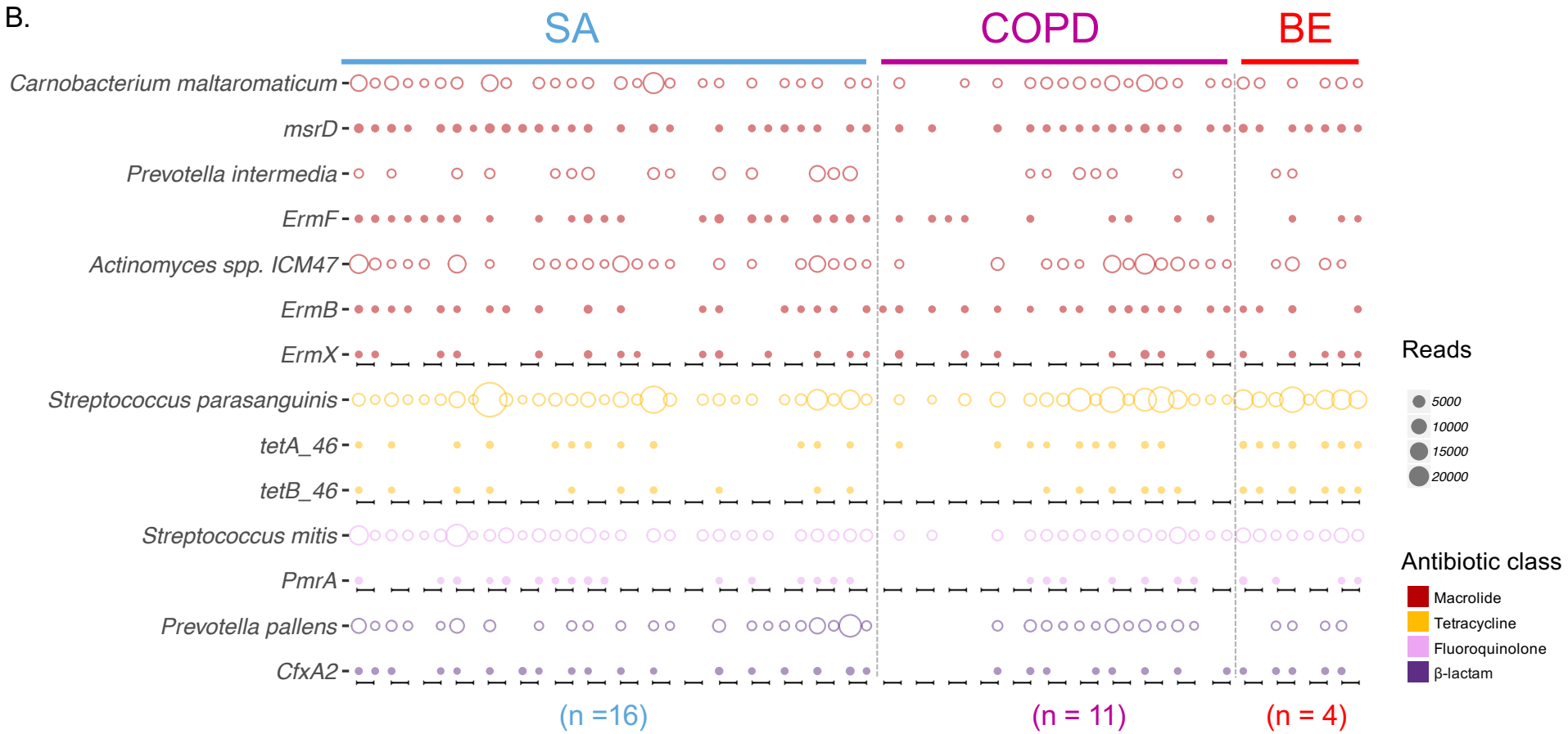
Figure 6a

A.



AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Figure 6b



AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Online Data Supplement

Metagenomics Reveals a Core Macrolide Resistome Related to Microbiota in Chronic Respiratory Disease

Micheál Mac Aogáin, Kenny J. X. Lau, Zhao Cai, Jayanth Kumar Narayana,

Rikky W. Purbojati, Daniela I. Drautz-Moses, Nicolas E. Gaultier, Tavleen K. Jaggi,

Pei Yee Tiew, Thun How Ong, Mariko Siyue Koh, Albert Lim Yick Hou,

John A. Abisheganaden, Krasimira Tsaneva-Atanasova,

Stephan C. Schuster and Sanjay H. Chotirmall.

*These authors contributed equally

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

SUPPLEMENTARY MATERIALS AND METHODS

Ethical approval: This study was approved by the institutional review boards of all participating sites as follows and all included participants provided written informed consent: CIRB 2016/2628 (severe asthma), CIRB 2016/2715, CIRB 2016/2549 (COPD), CIRB 2016/2073 (bronchiectasis) (all mutually recognised by DSRB) and NTU IRB-2016-01-031 and IRB-2017-12-010 (healthy individuals and diseased patients).

DNA extraction and sequencing

Spontaneously expectorated and representative sputum was obtained from all study participants according to established, standardised and published collection and assessment protocols (1, 2). DNA was extracted from representative sputum using the Roche High-pure PCR Template Preparation Kit (Roche) as previously described (3-5). To assess the environmental metagenome present on the mouthpieces of patient inhalers, 4N6Floq swabs (Copan, USA) pre-moistened in phosphate buffered saline (PBS) with 0.1% Triton-X100 (Sigma-Aldrich, USA) were used and the inhaler used most frequently sampled. After sampling, the swabs were snapped into a DNeasy PowerWater kit (Qiagen, Germany) bead tube. DNA was extracted and processed according to the manufacturer's protocol with the addition of proteinase K (Sigma-Aldrich, USA) and sonication at 65ºC (6). Sterile swabs and reagents were processed simultaneously as extraction controls to assess the levels of background contamination and subjected to metagenomic analysis which confirmed lower-level background contamination in sputum (11428-59428 reads, n =4) and swab (1175-4487 reads, n =6) extraction controls (Supplementary Figure E2). Extracted DNA was quantitated using the Qubit dsDNA High Sensitivity (HS) Assay Kit (Invitrogen, USA) and sequenced on a HiSeq 2500 platform (Illumina, USA) according to library preparation and DNA sequencing methods described by Gusareva, *et al.* (7). All sequence data from this study has been uploaded

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

to the National Center for Biotechnology Information (NCBI) Sequence read archives (SRA) under project accession PRJNA595703.

Functional and taxonomic assignment of metagenomic sequence reads

Raw sequencing reads with a minimum Phred score 20 and >30 bp in length were selected and adapter-trimmed using Cutadapt (version 1.14) (8). Trimmed reads were then mapped against the GRCh38 human reference genome with bowtie2 (version 2.3.2) (9). Unmapped non-host reads were separated for further analysis and aligned against the NCBI non-redundant protein database (7 August 2017) with Diamond (version 0.9.9) (10). Based on these alignments, microbial taxonomical classification was generated using the Lowest Common Ancestor (LCA) algorithm implemented in MetaGenome Analyzer (MEGAN, v6.8.18) with a minimum score of 100 and a support of ≥25 (11). Functional annotation according to the Kyoto Encyclopedia of Genes and Genomes (KEGG) was performed in MEGAN based on the '*acc2kegg-Dec2017'* database visualizing metabolism-associated functions. Resistance genes present among non-host reads were identified using ShortBRED (12). Unique protein markers were created by applying the 'shortbred_identify.py' script with a 95% cluster identity using antimicrobial resistance (AMR) protein sequences from Comprehensive Antibiotic Resistance Database (CARD) and Uniref90 database as a reference (13, 14). The abundance of unique protein markers was determined using the shortbred_quantify.py script based on average read length of 200 bp. Drug class assignments were manually validated against the CARD database. Normalized gene abundances were expressed in reads per kilobase million mapped reads (RPKM) relative to each AMR gene.

Data analysis, visualisation and statistical analysis

All continuous data was tested for normality by the Kolmogorov-Smirnoff test. Categorical data was assessed by Chi-squared or Fisher's exact test as appropriate. For non-normal data,

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Mann-Whitney U-testing was used for group comparisons. For comparison of three or more groups of non-normal measures, the Kruskal-Wallis test was employed with Dunn's *post hoc* test and Benjamini-Hochberg correction for multiple comparisons. Differences were considered significant at $p < 0.05$.

Read functional assignments were visualised using heatmaps with abundance coloured according to Z-scores in MEGAN while the RPKM values of AMR genes were plotted using the '*heatmap.2'* function from the R package '*gplots'* (15). A list of AMR genes was extracted and plotted in a Venn diagram using both the '*venndiagram'* and '*upsetR'* packages to examine the similarities in AMR profiles across healthy and diseased groups (16, 17). Principal coordinate analysis (PCoA) plots of taxonomic data were generated using '*vegan'* and '*phyloseq'* packages to observe clustering patterns among the four studied groups (18, 19). The average distance to centroid of each group was determined using PERMDISP tests of homogeneity of dispersions by the *'betadisper'* function in '*vegan'*. The significance of distances to centroid differences obtained between groups was assessed by analysis of variance (ANOVA) with Tukey's *post hoc* analysis. Linear discriminant analysis between cohorts was performed with lefse (version 1.0) with default parameters (20). Assigned reads were visualised in MEGAN as bubble charts representing normalised read counts. Co-occurrence networks were generated by combining normalised AMR and taxonomic data and assessed several network metrics using generalized boosted linear models (GBLMs) as described by Faust *et al.* (21). Pearson's correlation for co-occurring taxa and resistance genes were visualised in R using the *'rcorr'* function from the *'Hmisc'* package, with consideration of correlation strength and significance ($p > 0.05$) (22).

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

REFERENCES

1. Murray PR, Washington JA. Microscopic and baceriologic analysis of expectorated sputum. *Mayo Clin Proc* 1975; 50: 339-344.

2. Van Scoy RE. Bacterial sputum cultures. A clinician's viewpoint. *Mayo Clin Proc* 1977; 52: 39-41.

3. Chotirmall SH, O'Donoghue E, Bennett K, Gunaratnam C, O'Neill SJ, McElvaney NG. Sputum Candida albicans presages FEV(1) decline and hospital-treated exacerbations in cystic fibrosis. *Chest* 2010; 138: 1186-1195.

4. Mac Aogain M, Chandrasekaran R, Lim AYH, Low TB, Tan GL, Hassan T, Ong TH, Hui Qi Ng A, Bertrand D, Koh JY, Pang SL, Lee ZY, Gwee XW, Martinus C, Sio YY, Matta SA, Chew FT, Keir HR, Connolly JE, Abisheganaden JA, Koh MS, Nagarajan N, Chalmers JD, Chotirmall SH. Immunological corollary of the pulmonary mycobiome in bronchiectasis: the CAMEB study. *The European respiratory journal* 2018; 52.

5. Ali N, Mac Aogain M, Morales RF, Tiew PY, Chotirmall SH. Optimisation and Benchmarking of Targeted Amplicon Sequencing for Mycobiome Analysis of Respiratory Specimens. *Int J Mol Sci* 2019; 20.

6. Luhung I, Wu Y, Ng CK, Miller D, Cao B, Chang VW-C. Protocol Improvements for Low Concentration DNA-Based Bioaerosol Sampling and Analysis. *PLOS ONE* 2015; 10: e0141158.

7. Gusareva ES, Acerbi E, Lau KJX, Luhung I, Premkrishnan BNV, Kolundzija S, Purbojati RW, Wong A, Houghton JNI, Miller D, Gaultier NE, Heinle CE, Clare ME, Vettath

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

VK, Kee C, Lim SBY, Chenard C, Phung WJ, Kushwaha KK, Nee AP, Putra A, Panicker D, Yanqing K, Hwee YZ, Lohar SR, Kuwata M, Kim HL, Yang L, Uchida A, Drautz-Moses DI, Junqueira ACM, Schuster SC. Microbial communities in the tropical air ecosystem follow a precise diel cycle. *Proceedings of the National Academy of Sciences of the United States of America* 2019.

8. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* 2011; 17: 3.

9. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 2012; 9: 357-359.

10. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nature Methods* 2014; 12: 59.

11. Huson DH, Auch AF, Qi J, Schuster SC. MEGAN analysis of metagenomic data. *Genome research* 2007; 17: 377-386.

12. Kaminski J, Gibson MK, Franzosa EA, Segata N, Dantas G, Huttenhower C. High-Specificity Targeted Functional Profiling in Microbial Communities with ShortBRED. *PLOS Computational Biology* 2015; 11: e1004557.

13. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, Lago BA, Dave BM, Pereira S, Sharma AN, Doshi S, Courtot M, Lo R, Williams LE, Frye JG, Elsayegh T, Sardar D, Westman EL, Pawlowski AC, Johnson TA, Brinkman FS, Wright GD, McArthur AG. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res* 2017; 45: D566-d573.

14. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 2019; 47: D506-d515.

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

15. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M, Magnusson A, Moeller S. gplots: Various R programming tools for plotting data. *R package version* 2009; 2: 1.

16. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* 2011; 12: 35.

17. Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* 2017; 33: 2938-2940.

18. Oksanen J, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Simpson G, Solymos P, Stevens M, Wagner H. vegan: Community Ecology Package. R package version 2.4-6. 2018. *There Is No Corresponding Record for This Reference* 2018.

19. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013; 8: e61217.

20. Wickham H. ggplot2: elegant graphics for data analysis. Springer; 2016.

21. Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, Huttenhower C. Microbial co-occurrence relationships in the human microbiome. *PLoS Comput Biol* 2012; 8: e1002606.

22. Harrell Jr FE, Harrell Jr MFE. Package 'Hmisc'. *CRAN2018* 2019: 235-236.

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

SUPPLEMENTARY FIGURE LEGEND

Supplementary Figure E1. Bar plot of metagenomic read assignments across cohorts. The abundance of non-human reads assigned to bacterial (purple), fungal (red) and viral taxa (yellow) is illustrated. Dark and light grey coloration denotes non-human reads with non-deterministic or no hits in the reference database.

Supplementary Figure E2. Metagenomic sequence analysis of control samples to assess background DNA contamination profiles of DNA extraction blanks. Four negative 'blank' DNA extractions, using only PBS and following our sputum extraction protocol ("sputum") and six DNA extractions from sterile swabs following our swab extraction protocol ("Swab") were performed and subjected to metagenomic sequencing. Assigned taxonomy for the most abundant species is colour-indicated in the legend provided within the figure.

Supplementary Figure E3. Box and whisker plots illustrating the α-diversity of the respiratory microbiome in Non-diseased (ND), Severe asthma (SA), Chronic obstructive pulmonary disease (COPD) and bronchiectasis (BE) patients included in this study. Plots indicating calculated (A) Shannon, (B) Simpson and (C) Chao1diversity indices for each group are analysed by ANOVA with Tukey post-hoc analysis; *$p<0.05$, **$p<0.01$ (D) Identification of discriminant taxa based on linear discriminant analysis (LDA) effect size (LEfSe). The observed effect sizes of discriminant taxa identified among diseased patients ('D') are indicated by the upper right-facing bars while those discriminant for non-diseased subjects ('ND') are indicated by lower left-facing bars. Colour represents phylum membership. (E) Bubble chart summarizing the taxonomic profiles of all four study groups, classified to genus level. Bubble size reflects read abundance and colour indicates phylum-level association.

AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

SUPPLEMENTARY TABLES

Supplementary Table E1: Demographics for paired patient-inhaler metagenomes

Demographic	Disease		
	Severe Asthma	**COPD**	**Bronchiectasis**
N	16	11	4
Age	64±19	68±5	65+11
Gender			
Male	19% (3)	100% (11)	25% (1)
Female	81% (13)	0% (0)	75% (3)
BMI	24.7±9	22.3±3	19.4±5
FEV_1 % predicted	78±18	58±33	66±14
Disease severity †			
Mild	32% (5)	0% (0)	50% (2)
Moderate	32% (5)	55% (6)	0% (0)
Severe	36% (6)	45% (5)	50% (2)
Exacerbations in year preceding study recruitment	2±5	0±0.3	2±0.1
Smoking status			
Never smoker	75% (12)	0% (0)	75% (3)
Ex-smoker	19% (3)	55% (6)	25% (1)
Current smoker	6% (1)	45% (5)	0% (0)
Antibiotic use in 6 months preceding recruitment			
Yes	0% (0)	0% (0)	0% (0)
No	100% (16)	100% (11)	100% (4)
Inhaled corticosteroid use			
Yes	100% (16)	36% (4)	25% (1)
No	0% (0)	64% (7)	75% (3)
Inhaled Bronchodilator use			
Yes	100% (16)	100% (11)	75% (3)
No	0% (0)	0% (0)	25% (1)

† Defined according to disease-specific criteria. Severe asthma: ACT score; ≥ 20 = "mild", 19-15 = "moderate", <15 = "severe". COPD: GOLD stage; 1 = "mild", 2 = "moderate", ≥ 3 = "severe". Bronchiectasis: Bronchiectasis Severity Index (BSI); 0-4 = "mild", 5-8 = "moderate", ≥ 9 = "severe".

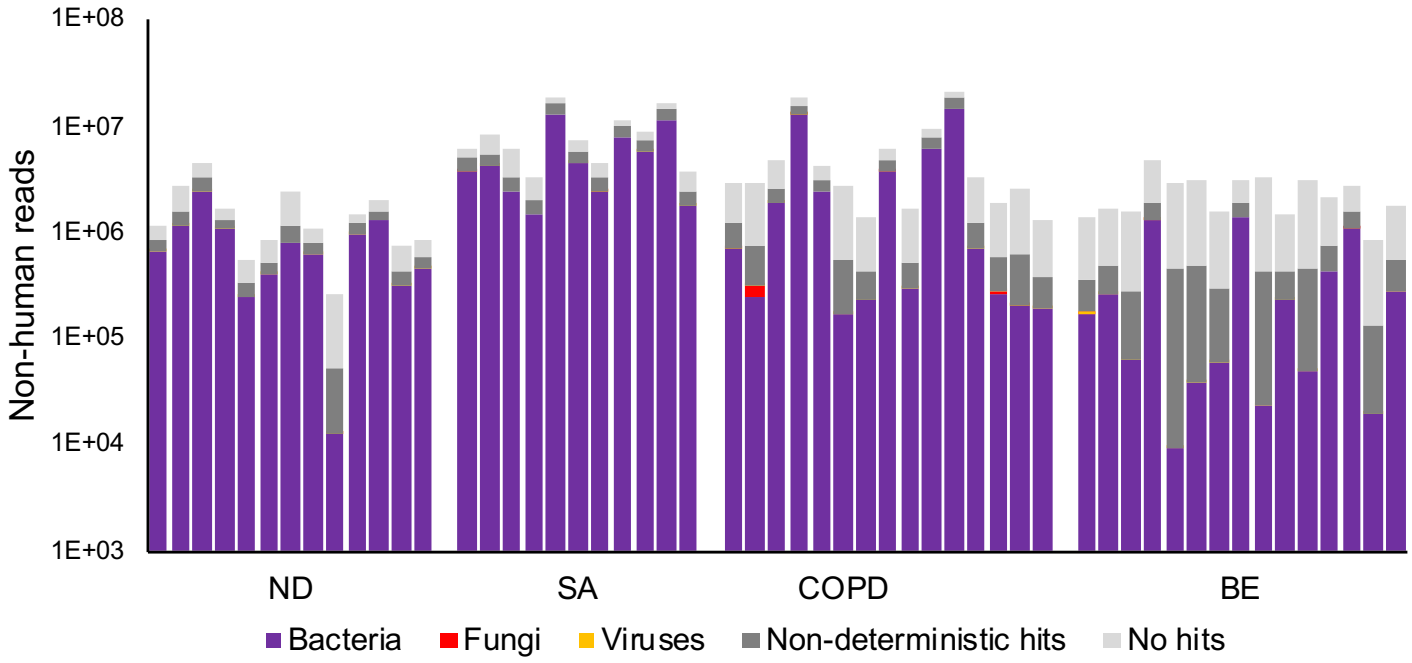
AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Supplementary Table E2: Putative environmental resistance genes found in association with inhaler devices but absent in sputum.

Gene name	Drug class	Resistance mechanisms
aac(3)-VIIa	aminoglycoside	antibiotic inactivation
aph(3'')-Ib	aminoglycoside	antibiotic inactivation
BUT-1	cephalosporin (β-lactam)	antibiotic inactivation
qacA	fluoroquinolone	antibiotic efflux
lfrA	fluoroquinolone	antibiotic efflux
lnuA	lincosamide	antibiotic inactivation
mgrA	antibiotic peptide, fluoroquinolone, tetracycline, penam, cephalosporin, acridine dye	antibiotic efflux
vgaALC	streptogramin, pleuromutilin, oxazolidinone, macrolide, tetracycline, lincosamide, phenicol	antibiotic target protection
msrA	streptogramin, pleuromutilin, oxazolidinone, macrolide, tetracycline, lincosamide, phenicol	antibiotic target protection
sdiA	tetracycline, penam, phenicol, rifamycin, cephalosporin, glycylcycline, triclosan, fluoroquinolone	antibiotic efflux
cmx	phenicol	antibiotic efflux
tet(38)	tetracycline	antibiotic efflux

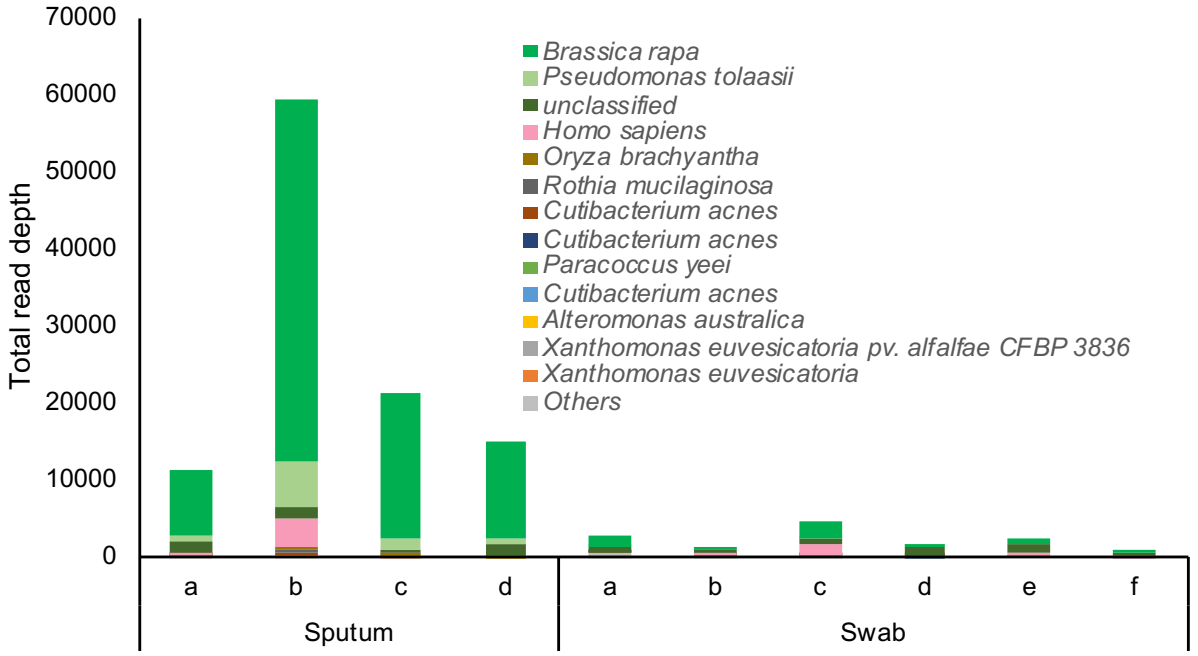
AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Supplementary Figure E1



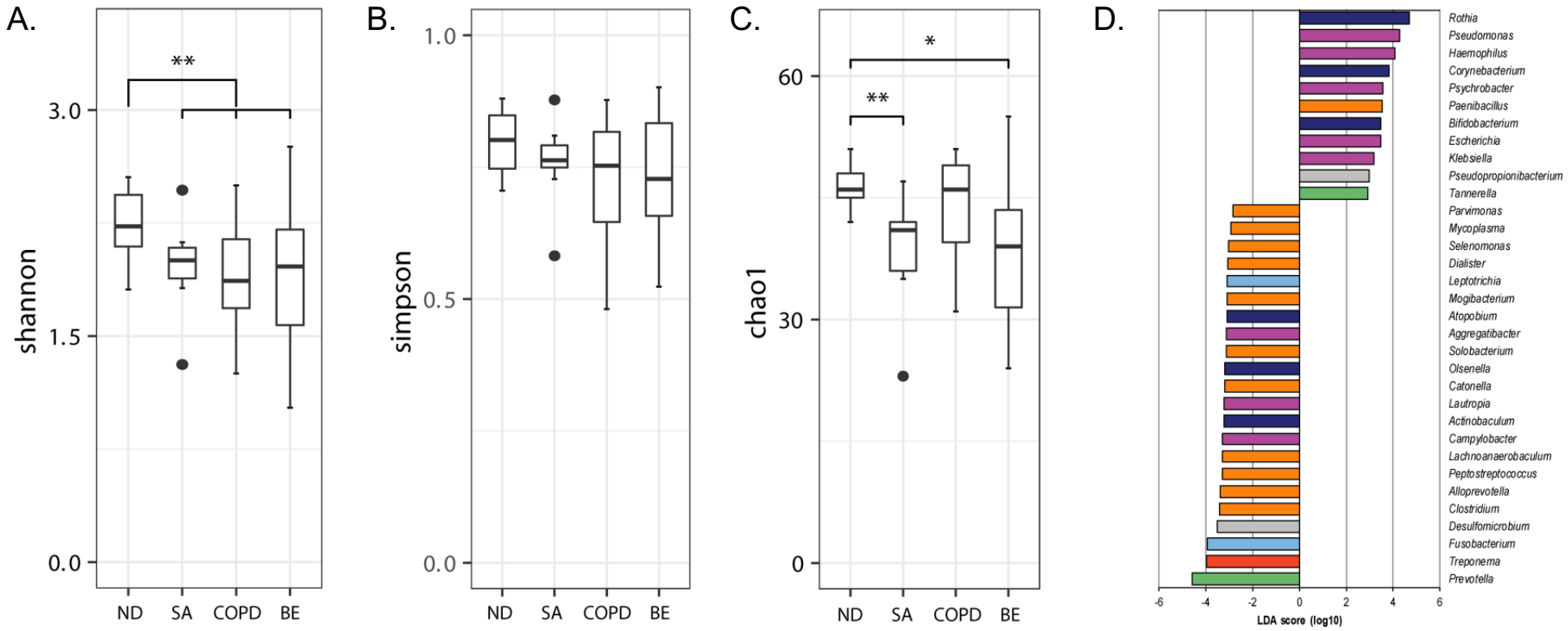
AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Supplementary Figure E2



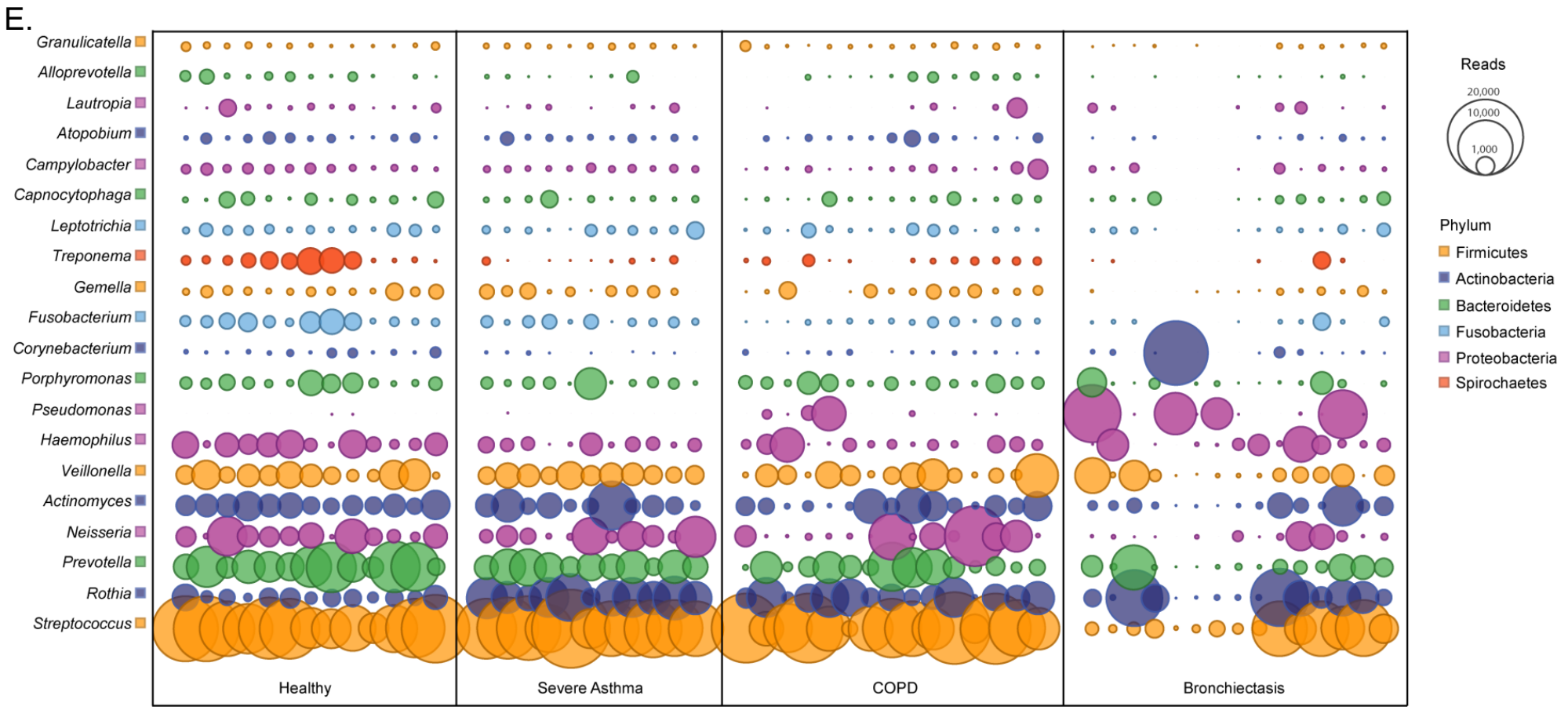
AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Supplementary Figure E3 (a-d)



AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC

Supplementary Figure E3 (e)



AJRCCM Articles in Press. Published April 22, 2020 as 10.1164/rccm.201911-2202OC