INTEGRATIVE MICROBIOMICS REVEALS A DISRUPTED INTERACTOME IN BRONCHIECTASIS EXACERBATIONS

Micheál Mac Aogáin[#1], PhD, Jayanth Kumar Narayana[#1], BS-MS, Pei Yee Tiew[1,2], MD, Nur A'tikah Binte Mohamed Ali, BSc[1], Valerie Yong Fei Lee, BSc[1], Tavleen Kaur Jaggi, MD[1], Albert Lim Yick Hou, MD[3], Holly R. Keir, BSc[4], Alison J. Dicker, PhD[4], Thng Kai Xian, BSc[1], Akina Tsang[1], Fransiskus Xaverius Ivan, PhD[1], Mau Ern Poh, MD[5], Martina Oriano, BSc[6], Stefano Aliberti, MD[6], Francesco Blasi, MD[6], Teck Boon Low, MD[7], Ong Thun How, MD[2], Brian Oliver, PhD[8,9], Giam Yan Hui, BSc[4], Augustine Tee, MD[7], Mariko Siyue Koh, MD[2], John Arputhan Abisheganaden, MD[3], Krasimira Tsaneva-Atanasova, PhD[10], James D. Chalmers, PhD[4] and Sanjay H. Chotirmall, PhD[1]*.

#These authors contributed equally

* Corresponding author: Sanjay H. Chotirmall, Lee Kong Chian School of Medicine, Nanyang Technological University, 11 Mandalay Road, Singapore 308232.

Email: schotirmall@ntu.edu.sg

[1]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore.
[2]Department of Respiratory and Critical Care Medicine, Singapore General Hospital, Singapore.
[3]Department of Respiratory and Critical Care Medicine, Tan Tock Seng Hospital, Singapore.
[4]University of Dundee, Ninewells Hospital and Medical School, Dundee, Scotland.
[5]Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia
[6]Respiratory Unit and Cystic Fibrosis Adult Center, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy; Department of Pathophysiology and Transplantation, University of Milan, Sassari, Italy.
[7]Department of Respiratory and Critical Care Medicine, Changi General Hospital, Singapore.
[8]Woolcock Institute of Medical Research, University of Sydney, Sydney, NSW Australia
[9]School of Life Sciences, University of Technology Sydney, Sydney, NSW Australia.
[10]Department of Mathematics, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK.

ABSTRACT

We report the first 'multi-biome' analysis integrating bacterial, viral, and fungal communities in bronchiectasis (651 microbiomes in 217 patients) employing weighted Similarity Network Fusion (wSNF): codified as an online webtool (https://integrative-microbiomics.ntu.edu.sg/) to identify patients at highest risk of exacerbation. Frequent exacerbators exhibit less complex microbial co-occurrence networks, reduced diversity and a higher degree of antagonistic interactions. Interactome dynamics, derived longitudinally (153 microbiomes in 17 patients), illustrate significantly increased and antagonistic interactions during exacerbations which resolve following treatment within an otherwise stable multi-biome. Assessment of the *Pseudomonas*-interactome reveals that its network rather than *Pseudomonas* abundance is key in determining bronchiectasis exacerbation risk, and, that incorporating microbial interactions improves models that predict clinical outcome. Validating interactomes and demonstrating their clinical applicability was achieved through functional metagenomics and confirmatory experimentation. A whole-genome shotgun metagenomic validation in an independent cohort of 166 patients, including bacteriophage assessment, served to validate the multi-biome interactions detected by targeted amplicon sequencing and confirmed microbial interactions linked to exacerbation risk. 'Integrative microbiomics' captures microbial interactions determining exacerbation risk which cannot be appreciated by studying a single microbial group. Antibiotic strategies likely target interaction networks rather than individual microbes providing a fresh approach to understanding respiratory infection.

Key words: Bronchiectasis; Exacerbation; Microbiome; Interactome; Microbiomics

Word count: 4,176 (Abstract 197)

References: 30

ABBREVIATIONS

BMI: Body mass index

BSI: Bronchiectasis severity index

CAMEB: Cohort of Asian and Matched European Bronchiectasis

CF: Cystic Fibrosis

COPD: Chronic Obstructive Pulmonary Disease

CRS: Chronic rhinosinusitis

DD: Dundee

HRCT: high resolution computed tomography

IQR: Interquartile range

ITS: Internal Transcribed Spacer

KL: Kullback-Leibler

LEfSe: Linear Discriminant Analysis Effect Size

NCBI: National Center for Biotechnology Information

PBS: Phosphate-buffered Saline

RDP: Ribosomal Database Project

RT-qPCR: Real-time quantitative Polymerase Chain Reaction

SG-KL: Singapore-Kuala Lumpur

SNF: Similarity Network Fusion

SRA: Sequence Read Archives

INTRODUCTION

The global burden of bronchiectasis is increasing, and, there remains a lack of proven treatment options due to disease heterogeneity, however, this is being addressed through endo-phenotyping efforts and the identification of treatable traits [1-4]. Recurrent infection and inflammation result in progressive irreversible airway dilatation characterised by an altered airway microbiome [2,5,6]. Bacterial, viral and fungal communities in bronchiectasis have been investigated and associated with clinical outcomes including exacerbations [7-9]. While specific bacteria, viruses and fungi are implicated in bronchiectasis exacerbations, prior bacterial microbiome studies illustrate minimal actual change during exacerbations based on simplistic single kingdom analyses focused on dominant bacterial taxa or dissimilarity metrics demonstrating our incomplete understanding of the microbiome's role [10]. Exacerbation occurrence and frequency in bronchiectasis and other respiratory diseases remain a major cause of morbidity and key driver of mortality, however, the precise microbial relationships underpinning exacerbations remain complex with most accepting simplistic single-kingdom models of bacterial overgrowth causing infection which is then suppressed by antibiotics [5,10,11]. Disease heterogeneity in bronchiectasis has hindered clinical trials, and, patient stratification based on microbiomes may provide focused and precision-based therapy given the variability in clinical, immunological and inflammatory phenotypes, aetiologies and therapeutic responses [1,7,12]. Critically, however, most microbiome studies to date have considered bacteria, viruses, and fungi as separate entities, but, the true microbiome, consisting of all microorganisms and their genes within a single body space includes bacteria, viruses, and fungi. Prior studies are therefore incomplete, and a greater understanding of disease likely gained through holistic and integrated 'multi-biome' analysis, which more accurately represents the *in-vivo* state. Here, we perform the first integrated 'multi-biome' analysis of the bronchiectasis airway combining bacterial, viral, and fungal community profiles from individual patients including longitudinal

assessment during exacerbations. We demonstrate that 'integrative microbiomics' provides a novel framework for understanding exacerbations with potential application across respiratory diseases.

RESULTS

Multi-biome data integration by weighted Similarity Network Fusion (wSNF) identifies exacerbators: To evaluate the bronchiectasis microbiome, we assessed respiratory specimens in 217 patients capturing bacterial, fungal and viral microbiomes in each patient (3 profiles per patient, total; 651 microbiomes). These patients were recruited as part of the CAMEB study; a cross-sectional Cohort of Asian and Matched European Bronchiectasis [8]. Patients had a median age of 68 (range: 60-74 years old) with equal gender distribution. Most had idiopathic or post-infection (non-mycobacterial) bronchiectasis and classified as moderate to severe disease (median Bronchiectasis Severity Index (BSI) = 9; range 6-13). For inclusion patients had confirmed radiological bronchiectasis by high resolution computed tomography (HRCT) and were recruited during outpatient attendance when clinically stable (further details in the online methods; full patient demographics and clinical details are described in supplementary table S1). Having previously characterized the fungal mycobiome in the CAMEB cohort, we first generated bacterial and viral microbiome profiles for all patients to assess a more holistic microbiome in each individual (Supplementary figures S1 and S2, Supplementary table S4) [8]. We subsequently integrated the derived bacterial, viral, and fungal community profiles in a novel analysis by implementing a weighted SNF approach as each -biome differentially influences the overall multi-biome based on its individual taxonomic composition and richness (Figure 1a-b and Supplementary figure S3). Weighting each -biome relative to taxonomic richness was achieved according to the total number of observed taxa present in a particular -biome, with filtering based on a prevalence of at least 5% across the entire patient cohort; i.e.

bacteriome (62 genera) > mycobiome (52 genera) > virome (4 viral species) observed across 217 patients. Assessment of weighting based on these stable observed inter-kingdom taxa (n=118) resulted in assigned weightings of 53% (52/118) for bacteria, 44% for fungi (52/118) and 3% for viruses (4/118) in our network fusion consistent with the breadth of information content underlying each -biome network (Figure 1b). After spectral clustering of the resultant similarity matrix, we identify two patient clusters (Figure 1c). The mean misclassification ratio over 100 iterations was found to be 12.43%, indicating a cluster robustness of 87.6%. The weighted SNF method employed has been codified and is openly available as an online webtool (https://integrative-microbiomics.ntu.edu.sg/; see online methods). Microbes within each cluster reveal a range of discriminant bacterial, fungal, and viral taxa highlighting potential interaction between them to define the observed clinical state (Figure 1d). Patients from the larger cluster (cluster 1, n = 134) exhibit greater microbial diversity (Figure 1e) and had better clinical outcomes than those in the smaller cluster (cluster 2, n = 83) in terms of exacerbations and symptoms (Figure 1f-g). Additional geographic and clinical differences between the clusters included a higher proportion of European patients in the frequently exacerbating cluster (82% versus 16%, $p<0.001$) with patients in this cluster also exhibiting a higher BMI (20-31 versus 18-24, $p<0.001$), higher prevalence of chronic rhinosinusitis (CRS) (36% vs 18%, $p<0.001$) greater inhaled corticosteroid (ICS) use (50% versus 26%, $p<0.001$), and likelihood of a smoking history (41% versus 22%, $p<0.005$). The presence of CRS was increased in Cluster 2 but not associated with a greater presence of taxa known to be over-represented in sinus disease (Supplementary figure S4) suggesting that our sampling methodologies were robust to such influence[13]. The influence of therapy (ICS and antibiotic use) on microbiome composition is illustrated in Supplementary figure S5. Patients in cluster 2 carried a relative risk of 2.36 (95% CI 1.6-3.5; $p<0.0001$) for exhibiting the 'frequent exacerbator' phenotype, defined as experiencing >=3 exacerbations annually [14]. Importantly,

while lung function and disease severity between the clusters is comparable (Figure 1h-i); evaluating the multi-biome provides a clinically meaningful patient stratification by accurately identifying 'high-frequency' exacerbators in comparison to analysis of the bacterial microbiome alone (Supplementary table S5), a clear demonstration of the clinical utility of our integrated analysis. In contrast to existing bronchiectasis paradigms, the 'high frequency' exacerbation cluster had lower prevalence (31% vs 66%, p>0.001) and relative abundance (1.7% vs 10.8%, p>0.001) of *Pseudomonas*.

Co-occurrence analyses reveals an antagonistic interactome in high-frequency exacerbators: To characterise microbial interactions (the interactome) within each cluster, a weighted co-occurrence approach using an ensemble of similarity measures and regression techniques was employed to generate microbial association networks (Figure 2a-b). Leveraging methodology of Faust *et al.* [15] mitigated against compositionality of relative abundance data and provided a framework (based on graph theory) where microbes (described as nodes) may be assessed in terms of their interconnection with (predicted) interacting partners (edges) which can be positively or negatively correlated (see supplementary material). Therefore, a positive interaction between microbes is defined by the consensus ensemble correlative score whereby a positive value represents the co-occurrence of microbes and a negative value co-exclusion. While the low frequency exacerbation cluster had a higher total number of microbes and microbial interactions, the high frequency exacerbation cluster exhibits lower diversity and a greater proportion of negative interactions between constituent microbes (Figure 2c). An altered interactome is therefore evident in the high frequency exacerbation cluster suggestive of opposing microbial interactions which potentially drives this observed clinical state (Figure 2d-e).

Busy, critical and influential microbes within the interactome: Adopting network-based approaches permits an assessment of alternate metrics to characterise microbiomes for potential clinical applicability [16]. We next evaluated network metrics including node degree, stress- and betweenness-centrality (of the nodes) to describe microbes within a network that we refer to as 'busy' (microbes with an increased number of direct interactions with other microbes), 'critical' (key microbes to maintain network integrity) and 'influential' (microbes influencing other microbes within a network including indirectly). Using this approach, we identified key taxa of clinical relevance and potential targets for antimicrobial intervention in our clusters (Figure 3a-b). With these metrics, a different view of cluster-specific interactomes is appreciated with *Rothia*, *Streptococcus, Candida, Actinomyces* and *Haemophilus* the highest ranked taxa in the low frequency exacerbation cluster demonstrating characteristics of being busy, critical and influential within their network. Of these taxa, only *Haemophilus* is similarly ranked in the high frequency exacerbation cluster alongside *Cryptococcus*, *Leptotrichia*, *Poryphyromonas, Prevotella* and *Veillonella* (Figure 3b, Supplementary table S6). Critically, while some of the top taxa identified in each respective cluster commonly share the busy, influential and critical network characteristics (e.g. *Streptococcus*, *Haemophilus*, *Candida* and *Cryptococcus*), all importantly exhibit markedly different interaction networks when assessed independently within their respective clusters where they also exhibit variable prevalence (Supplementary figures S6 and S7). This suggests that a microbe's interactome rather than the microbe itself dictates clinical status such as the presence of a high risk for exacerbations. Therefore, classification based on the interactome provides superior resolution to that provided by a single microbe (e.g. *Pseudomonas* spp.) to identify patient populations at highest risk of adverse clinical outcome. Conversely, when interaction networks were assessed in a supervised manner, by splitting patients into groups based on their known exacerbation frequency (<3 vs >2) as opposed to weighted-SNF, the observed network configurations from the unsupervised

analysis were comparable but with the notable appearance of *Pseudomonas* spp. among the top taxa of frequent exacerbators (Supplementary figure S8).

As *Pseudomonas* spp. are strongly associated to exacerbations in bronchiectasis, we next specifically assessed *Pseudomonas*-interaction networks [17]. Interestingly, *Pseudomonas* spp. were identified in both unsupervised exacerbation clusters, suggesting that its presence alone does not adequately account for the different exacerbation frequencies between clusters (Figure 3a-b). *Pseudomonas* spp. exhibits distinct interactomes based on a patient's exacerbation frequency, and, while an overall trend toward more negative interactions is observed in the high frequency exacerbation cluster (Figure 2c-e), the *Pseudomonas*-interaction network within this cluster instead exhibits a greater number of positive interactions (Figure 3c). Notably however, the *Pseudomonas*-interaction network in the low frequency exacerbation cluster exhibits greater negative interactions (n=18 vs n=5, Figure 3d). Several important differences in *Pseudomonas*-interaction networks between clusters were evident: *Pseudomonas* exerts a greater negative influence on *Haemophilus* and lesser negative influence on *Streptococcus* in the high frequency exacerbation cluster. Critically, and of interest, antithesis relationships were observed in respect to *Aspergillus, Prevotella*, *Veillonella, Neisseria* and human parainfluenza virus 3 between clusters where positive interactions predominate in the high frequency and negative interactions in the low frequency exacerbation clusters respectively (Figure 3d). Therefore, *Pseudomonas* spp. presence alone does not adequately explain the published links between this microbe and bronchiectasis exacerbations.

Differential network analysis during exacerbation reveals 'core' and 'ancillary' interactomes: As interactomes differ based on exacerbation frequency, we next evaluated the interactome prospectively across the course of exacerbations in an independent bronchiectasis

cohort recruited from two hospitals in the East of Scotland. Patients had a median age of 72 (range: 68-74 year) and were predominantly female (65%). Most had idiopathic bronchiectasis (65%) and classified as moderate to severe disease (median BSI 10; range 6-14) (further details in the online methods; patient demographics and clinical details are described in supplementary table S2). Here, we assessed bacterial, fungal, and viral microbiome profiles generated for these 17 patients across three timepoints (total; 153 microbiomes). We assessed the interactome at baseline (pre-exacerbation), during exacerbation and post-exacerbation (following antibiotic therapy). A comparison of detected microbes revealed broad comparability of longitudinal multi-biome signatures across the assessed timepoints with no significant differences observed in microbial composition, α- and β-diversity suggesting overall stability of the microbiome across exacerbation and recovery (Figure 4a-c, Supplementary figure S9). In contrast however, co-occurrence analysis revealed major changes in interactomes with an increase in the number and strength of negative interactions during exacerbations when compared to baseline (pre-exacerbation) or following treatment (post-exacerbation) (Figure 4d-f). A detailed comparison of changes from baseline to exacerbation and from exacerbation to post-exacerbation clearly illustrates a dynamic shift to a new 'post-exacerbation' network (Figure 4g). Fewer interactions are observed during and post-exacerbation compared to baseline, likely explained by broad spectrum antibiotic usage which eliminates potentially interacting microorganisms and, greatest overlap is observed between the exacerbation and post-exacerbation state (Figure 5a). Importantly, a 'core' interactome of 64 'conserved' microbial interactions exist with the strongest interactions noted between *Prevotella* , *Leptotrichia* and *Veillonella* (Figure 5a). To further characterise the key microbial interactions related specifically to the exacerbation state, differential network analysis was implemented which illustrates 'core' and 'ancillary' networks (Figure 5b). The 'core' network remains unaltered by exacerbation or therapy and includes some principal microbiota including bacteria such as *Streptococcus*, *Prevotella*,

Veillonella, *Neisseria, Leptotrichia,* and *Rothia*. Some conserved fungal and viral based interactions involve *Cryptococcus* and Rhinovirus, respectively. The interactions most susceptible to variability with exacerbation and treatment ('ancillary' network) includes more established respiratory pathogens such as the bacteria *Pseudomonas, Haemophilus, Stenotrophomonas, Moraxella* and *Staphylococcus* but also *Saccharomyces, Candida* (fungi), Influenza virus B and Metapneumovirus (viruses) (Figure 5b).

Next, we assessed the clinical utility of our derived network-based interactomes by predicting the influence of antibiotic exposure on its contained microbiome. Patient therapy was driven by culture-based microbiologic work up that compared well with the 16S rRNA analysis of the bacteriome (Supplementary figure S10), leading to several patients in the longitudinal study (n=12) receiving β-lactam antibiotics for treatment of their initial exacerbation (Supplementary table S8). We used the baseline (pre-β-lactam exposure) interactome network (Figure 5c) to predict network re-configuration post β-lactam treatment by artificially reducing the abundance of β-lactam-sensitive microbes by 75% (Figure 5d, see Online methods for further details). We then compared our 'simulated' network to that actually observed among our β-lactam-treated patients following therapy (Figure 5e). Our network-based prediction had reliable comparability to the network actually observed in β-lactam-treated patients with respect to several microbial nodes. Notably, the rank order difference in key microbial taxa post antibiotic treatment was correctly predicted for 10 out of 13 taxa in our simulated model further validating our derived interactomes and their potential for clinical relevance and translatability (Figure 5c-e, Supplementary table S7). As a prognostic indicator, we also found that interactions rather than individual microbial abundance served as a better predictor of time to next exacerbation among patients in the longitudinal arm of the study (Supplementary figure S11).

Functional and microbiological validation of low and high exacerbation frequency clusters: To assess function in identified clusters, we performed initial metagenomic sequencing on a subset of n=20 patients from each cluster (total, n = 40; Supplementary table S9). Linear discriminant analysis (LDA) identified a significant number of genes enriched in the high exacerbation frequency cluster highlighting potential genetic components related to exacerbation phenotypes and observed differences in the corresponding microbial interactome (Figure 6a). Functional mapping of these genes identified several microbial virulence related pathways enriched in the high exacerbation frequency cluster including functional categories related to quorum sensing, biofilm formation and antibiotic resistance (Figure 6b). To further assess and validate specific interactions within our derived interactomes, we selected the interaction between *P. aeruginosa* and *A. fumigatus* for further interrogation. These two organisms exhibit net opposing interactions from our originally derived clusters (Figure 3c-d): co-exclusion in the low and co-occurrence in the high exacerbation frequency clusters respectively (Figure 6c). Comparisons of *P. aeruginosa* clinical isolates derived from patients belonging to the low and high exacerbation frequency clusters respectively reveal these contrasting interactions (Figure 6d-e). Consistent with the observation from our derived interactomes (Figure 3c-d and 6c), the low exacerbation frequency cluster isolate (LEF) exhibited negative interactions with *A. fumigatus* whereas no such inhibitory effect was observed with the high exacerbation frequency cluster isolate (HEF) (Figure 6d-e). Further, though their growth was comparable, the HEF and LEF isolates exhibited clear variability in surface motility (Figure 6d – compare control plates) suggesting significant phenotypic differences between strains. Taken together, these *in-vitro* observations are consistent with our *in vivo*-derived interactomes further validating their accuracy and clinical relevance.

Whole genome shotgun metagenomics independently identifies and validates clinical exacerbation risk and associated microbial interactions in bronchiectasis: To further validate the association between microbial interactions and exacerbation, we performed whole-genome shotgun sequencing on an independently recruited bronchiectasis cohort from four separate jurisdictions (n=166; Singapore, Malaysia, Scotland and Italy) (Figure 7, Supplementary table S10). Functional analysis led to the independent identification of two patient clusters, again distinguished by the overrepresentation of microbial virulence functions including chemotaxis, two-component systems, secretion systems and siderophore production pathways co-incident with higher exacerbation frequency (Figure 7a-c). Patient clusters exhibited significant differences in lung function (FEV1 % predicted) (p=0.035) while disease severity (BSI) and symptoms (MMRC scores) were comparable (Figure 7d, Supplementary figure S12).

Next, to independently re-evaluate the multi-biome analysis performed previously using targeted amplicon sequencing data and validate the interactome (Figures 1-3), we first performed a greater in-depth analysis of the virome (including bacteriophages) generating a rich viral profile based on metagenomic data (Supplementary Results, Supplementary figure S13). Microbiome integration was then achieved by wSNF applying weights as described previously and assigned as follows: bacteriome (992 genera) > virome (703 viral contigs) > Mycobiome (16 genera) (Figure 7e-h). Implementation of integrative microbiomics again resulted in successful stratification of bronchiectasis patients into two clusters separated by exacerbation risk, but in this case with even greater precision than the initial functional metagenomic analysis (supplementary table S11). In addition to a greater frequency of exacerbations (akin to our earlier HEF cluster), patients in the higher risk cluster by metagenomics ('SC2') demonstrate significantly reduced lung function (as FEV1 % predicted;

p=0.0179) while symptoms (MMRC score) and disease severity (BSI) were comparable between clusters (SC1 versus SC2). Although an increased prevalence of CRS was observed in our initial clustering (derived from targeted amplicon sequencing analysis), this was not replicated by metagenomic analysis (26.9% vs 36.2%, p = 0.5565). Similarly, metagenomic validation did not detect differences in ICS use (43.3% vs 40.6%, p = 0.850) or antibiotic therapy (52.6 vs 37.8% p = 0.08207) which were both comparable between clusters. The key difference in exacerbation frequency between our identified clusters however was consistent between both sequencing approaches, especially regarding their interactomes, which in turn represents a strong validation of its clinical relevance (Figure 7i-k).

Co-occurrence analysis of the high-risk cluster from the metagenomics sequencing approach reveals important interaction networks and keystone taxa (which now include bacteriophages) which themselves exhibit a marked shift in phage profile between clusters (Supplementary figure S13a-e). There are also clear differences in overall network configuration between the low (SC1) and high (SC2) exacerbation clusters, which now can also provide additional information on an increased abundance of antimicrobial resistance determinants in SC2 (Figure 7i-j, Supplementary figure S13f).

The derivation of network configurations from two independent bronchiectasis cohorts and using two sequencing approaches further facilitated a direct comparison between interactomes generated by targeted and metagenomic sequencing, respectively, in relation to the high exacerbation risk clusters (i.e. HEF versus SC2). Direct comparison between microbes detected by both approaches revealed that 89.9% of interactions (i.e. 267 interactions between 18 microbes) were common between HEF and SC2 clusters which strongly validates associations

of these networks with clinical exacerbation risk confirming the overall importance and relevance of interactome analysis.

DISCUSSION

Here, we present to the best of the authors knowledge, the first 'multi-biome' analysis using 'integrative microbiomics' combining bacterial, viral, and fungal communities in individual patients. By developing a novel weighted-SNF methodology, accessible through an open access online platform, we identify frequent exacerbators with high precision and classify microbes within an 'interactome' as 'busy', 'influential' and/or 'critical'. Frequent exacerbators exhibit antagonistic interactomes, and longitudinal assessment over an exacerbation reveals disrupted interactomes, undetectable by assessing microbial identity alone, and which can predict subsequent exacerbation. The 'interactome' during exacerbations demonstrates 'core' and 'ancillary' networks, amenable to therapy, and, by use of simulation followed by confirmatory validation, we demonstrate its clinical relevance for modelling microbiome re-configuration in response to antibiotic exposure. Functional validation of interactomes was achieved by metagenomics which illustrates high activity in microbial virulence pathways indicative of chronic lung adaptation in frequent exacerbators. Microbiological evidence supports the interactome approach by demonstrating differential interaction between *P. aeruginosa* and *A. fumigatus* in frequent compared to infrequent exacerbators, an inter-kingdom pairing recognized and validated in subsequent metagenomic analysis. Taken together, our findings reveal a novel aspect of the microbiome with potential implications for the use of antibiotics in clinical practice.

Infection is central to bronchiectasis pathogenesis and based upon conceptual frameworks such as the 'vicious cycle and vicious vortex' [2,3]. Therefore, targeting bacteria with antibiotics reduces bacterial load, accompanying inflammation and therefore exacerbation risk, which, in turn alleviates symptoms and improves clinical outcomes, however, the role of co-existing, even commensal or 'pathobiont' microbes, is not considered in this model [18,19]. Further, this

model fails to explain how or why patients improve despite receiving antibiotics not necessarily targeting their dominant pathogen, exemplified by macrolide use in *Pseudomonas* infection. Interactomes therefore alter how we think about antibiotic use and treat exacerbations in bronchiectasis offering reasons for unexplained clinical observations. For instance, it is well recognised that patients receiving amoxicillin or a macrolide in the presence of *Pseudomonas* improve, the latter thought to be due to the drugs anti-inflammatory properties or the presence of co-infection. Our interactome analysis during exacerbations, however, suggests that the prescribed antibiotic may change the interactome, conferring the observed clinical benefits. Therefore, modulating *Pseudomonas*-related interaction, within a susceptible interactome, rather than directly targeting the culprit organism itself, may restrict its pathogenic potential and provide good clinical outcome as observed with amoxicillin and macrolides. Additionally, the longstanding clinical phenomenon of benefit gained by administration of antibiotics to which a target organism may be resistant, established in cystic fibrosis, is potentially also explained by antibiotic-related effects on the interactome, where microorganisms susceptible to the prescribed drug influence interactions with the target pathogen thereby modulating its pathogenic ability indirectly. Additionally, the lack of detectable change in respiratory microbiomes during bronchiectasis exacerbations, and even following antibiotics suggests that microbial abundance alone provides an incomplete view of airway microbial ecology [10,20]. Our analysis also offers the first metagenomic survey of bacteriophages in bronchiectasis, uncovering a striking change to bacteriophage profiles between our clinical exacerbation clusters. A higher burden of antimicrobial resistance genes related to these bacteriophages is demonstrated despite an absence of significant difference in antimicrobial therapy received between the clusters. Relationships between resident airway microbes and increased bacterial load during exacerbations, including mechanisms driving evolution from stability to

exacerbation are lacking, and an improved understanding of interactomes (including bacteriophages) provides key insight reflective of the *in-vivo* state.

The value of data integration using SNF for multidimensional datasets (such as multi-omics) in airways disease such as COPD has been demonstrated, however, these methods have not been previously applied to microbiome integration [21]. Conventional SNF is not optimized for biological systems such as multi-kingdom microbiomes where dynamism and potential dominance of one kingdom over the others needs to be considered. Employing a weighted SNF approach based on richness, we demonstrate improved patient stratification in bronchiectasis by identifying high frequency exacerbators with accuracy exceeding that of using a single microbial group. The methods described have been made accessible to the research community through our online webtool (https://integrative-microbiomics.ntu.edu.sg/) which has been road tested with several publicly available datasets suggesting broad applicability beyond respiratory disease.

Traditionally, exacerbations are considered to occur when an increased bacterial load or acquisition of a new virus ensues, however, analysis of a single microbial group by bacterial abundance or viral PCR has been shown as inadequate to discriminate between the stable and exacerbation states in bronchiectasis [9,10]. Interactome analysis goes deeper by identifying changing inter-kingdom interactions during an exacerbation. Despite identifying clinically relevant patient clusters tied to exacerbation frequency by integrative microbiomics; even the low frequency exacerbator group was enriched for key bronchiectasis pathogens associated to exacerbations including *Pseudomonas* and *Aspergillus,* suggesting that presence alone (including abundance) of a particular organism does not sufficiently explain the microbial dynamics occurring during exacerbations [8,22]. To better understand this, we employed network

427 analysis which provides insight into microbiome architecture, and, identifies keystone
428 microbial taxa based on relationships within an overall community rather than on their
429 occurrence or abundance alone [15,16]. This captures, with greater accuracy, microbiome
430 complexity and provides novel insights into events conferring clinical change such as
431 exacerbations and/or therapeutic response [16]. Our study is the first to employ such an approach
432 to the airway microbiome, and, demonstrates that bronchiectasis patients at highest risk of
433 exacerbations have an 'interactome' dominated by antagonistic interaction between microbial
434 kingdoms, explaining their lower α-diversity, where microbes compete rather than co-operate
435 with one another. Assessing the 'interactome' as a network of 'busy', 'critical' and 'influential'
436 microbes within an airway ecosystem highlights the relevance of established bronchiectasis
437 pathogens such as *Haemophilus,* however, particularly in the high-frequency exacerbation
438 cluster, relationships with other bacteria such as anaerobes (*Prevotella* and *Veillonella*) or other
439 kingdoms such as fungi (*Cryptococcus)* are novel and previously unrecognised in
440 bronchiectasis. The uncovered relationship to anaerobes is particularly interesting as anaerobes
441 are detected at high frequencies in the cystic fibrosis (CF) airway with conflicting results in
442 attempts to link them with disease outcomes or exacerbations [23,24]. Significantly, however, key
443 pathogenic taxa such as *Pseudomonas*, with established links to bronchiectasis exacerbations
444 demonstrate contrasting 'interactomes' between the low- and high-frequency exacerbators
445 confirming the importance of appreciating this phenomenon to best apply it to precision
446 microbiology.
447
448 To further validate our findings, we temporally assessed the interactome in a prospective
449 longitudinal bronchiectasis cohort experiencing exacerbations. This first confirmed findings of
450 prior microbiome studies in bronchiectasis indicating stability of the microbiome across
451 exacerbation and then following treatment with little change in microbial composition, α- or

β-diversity [5,10]. Significantly however, we detected a changed 'interactome', not assessed by prior works, which did not employ integrative approaches. A clear shift toward antagonistic microbial relationships during exacerbation was evident, comparable to that observed in our high-frequency exacerbation cluster, a finding unexplained by linear increases in pathogen dominance as reflected by comparable diversity indices. These findings, further validated by the 'core' and 'ancillary' interactomes during exacerbation, underlines the advantages conferred by network analysis, which reveals relationships undetectable by microbial abundance or identity assessment alone. Our novel approach highlights the importance of inter-kingdom 'interactomes' that varies during exacerbations offering deeper insight into potential triggers of microbial virulence. Importantly, the 'interactome' provides significant, new, and previously unrecognised targets for antimicrobial therapy that may be considered as alternative or in combination to established regimens to increase efficacy. We importantly demonstrate and validate that simulated microbial networks can be re-configured in response to antibiotic therapy highlighting the clinical potential and applicability of the interactome approach as a model to predict therapy-induced microbial dynamics. What remains unknown and unaddressed by this work are the respective benefits of targeting 'busy', 'critical' or 'influential' microbes within an 'interactome' and should be the subject of future studies.

Through application of metagenomics, we demonstrate important functional differences in gene profiles across our identified patient clusters. Discriminating genes characterising frequent exacerbators include genes implicated in biofilm formation (*glpP, glgC*), quorum-sensing and others with established roles in recalcitrant chronic infections [25,26]. Against this backdrop, and to further demonstrate plausibility of interactions within an interactome based on underlying exacerbation frequency, we selected *P. aeruginosa* isolates from patients belonging to the high and low exacerbation frequency clusters respectively, and, assessed their

477 growth ability in co-culture with a laboratory strain of *A. fumigatus*. Interaction between *P.*
478 *aeruginosa* and *A. fumigatus* has been shown to be clinically relevant in several prior works [27-]
479 [29]. We successfully replicated *in-vitro* the expected *in-vivo* co-exclusion and co-occurrence
480 interactions based on patient cluster membership highlighting a strong consistency between
481 our derived interactomes and observed functional and microbiological outcomes. Our
482 microbiological validation was however restricted to the well-established and easily cultivated
483 microbes *Pseudomonas* and *Aspergillus*; long-studied exemplars of inter-kingdom
484 communication in chronic respiratory disease. Further work is required to assess less
485 established interactions emerging from our network analysis, aided by continued
486 improvements in culture metagenomics [30].
487
488 Our study represents, to the best of our knowledge, the first description of an 'integrative
489 microbiomics' approach to the 'multi-biome' in chronic airways disease however does have
490 limitations. First, the patients were recruited from the established CAMEB cohort which by
491 design is cross-sectional, hence we use largely static data to predict dynamic interaction [8,12].
492 This is partially overcome by inclusion of a longitudinal arm to our analysis to better assess
493 temporal dynamics in association to exacerbation and antibiotic treatment. Next, although 16S
494 methodologies are well established, there are inherent limitations, including under-
495 representation of mycobacteria, an important group of organisms in bronchiectasis [31].
496 Additionally, fungal ITS sequencing approaches are challenged by under-developed reference
497 databases [32]. Our initial virome analysis, while broad, comprehensive, and informed by
498 established literature, targets a known virus panel and therefore is subject to bias. This resulted
499 in a much lower weighting of the virome in our SNF approach due to lower observed taxonomic
500 richness, constrained by the employed methodologies. We at least partially attempted to
501 overcome this through use of a metagenomics validation approach that assessed

bacteriophages. Future work and alternative approaches assessing viromes such as RNAseq may yield different results, and be more comprehensive allowing greater weighting of the viral contribution to the overall integrated microbiome, an important area of future exploration given the relatively poorly defined role of viruses in bronchiectasis. In addition, only young (healthy) controls were evaluated in our comparison of viral loads with bronchiectasis, and, additional older controls, comparable to age groups afflicted by bronchiectasis may have been of value. Further, while networks were weighted based on species richness, their true influence on the microbiome is not necessarily captured by richness alone, but rather a function of functional genes, competition, substrate utilization and energy flux through the ecosystem, traits that can't be comprehensively assessed by sequencing alone. While metagenomics potentially represents a less biased alternative approach while we have performed as validation, it itself underestimates fungal presence given the significantly higher airway bacterial burden hence obscuring the influence that fungi have on the interactome. We further acknowledge that sputum is an imperfect matrix, and, make no inference about lower airway ecology, noting only the clinical associations between sputum as a surrogate, readily obtainable, non-invasive upper airway sample. Finally, while observational data suggests potential causal association, other factors may drive observed effects. Observed interactions may represent epiphenomena of a selectively operating immune system, for example, and our work did not include any assessment of host responses: another avenue for future work.

Disrupting microbial networks through alteration of the 'interactome' is a novel consideration for chronic respiratory disease complicated by infection. The airway microbiome (and its accompanying interactome) is likely a critical predictor of antibiotic treatment response and provides a theoretical basis for understanding several phenomena associated with antibiotics that remain unexplained clinically including antimicrobial responses in apparently resistant

organisms. Manipulating microbiomes by means other than antibiotics are being explored and the effect of probiotics on the interactome should be considered. Holistic analytical approaches reflective of the *in-vivo* state, and, that which go beyond microbial identity alone must consider the complexity of inter-kingdom interactions demonstrated by 'integrative microbiomics' which may improve patient stratification, clinical trial design and therapeutic outcomes in bronchiectasis and other respiratory diseases.

Acknowledgements: This research is supported by the Singapore Ministry of Health's National Medical Research Council under its Transition Award (NMRC/TA/0048/2016) (S.H.C); Clinician-Scientist Individual Research Grant (MOH-000141) (S.H.C); the NTU Integrated Medical, Biological and Environmental Life Sciences (NIMBELS) [NIM/03/2018] (S.H.C); the British Lung Foundation through the GSK/British Lung Foundation Chair of Respiratory Research (J.D.C) and the Scottish Government Chief Scientist Office (SCAF/17/03).

Author contributions: MMA and JKN: performance and design of experiments, data analysis and interpretation, statistical analysis, writing the manuscript. PYT, AYHL, MEP, TBL, OTH, AT, MSK, JAA: patient recruitment and procurement clinical data and specimens for Asian patients. PYT, NABMA, TKX, AT, GYH, VFY, TKJ: performance of experiments. FXI: bioinformatic support. HRK, AJD, JDC: patient recruitment and procurement clinical data and specimens for European patients and the longitudinal cohort (Scotland). MO, FB, SA: patient recruitment and procurement clinical data and specimens for European patients (Italy). BO, KTA, JDC, SHC: intellectual contributions. KTA: data analysis. SHC: conception and design of overall study and experiments, data analysis and interpretation, statistical analysis, writing manuscript, procurement of funding.

Data availability statement.

All sequence data described in this study has been uploaded to the National Centre for Biotechnology Information (NCBI) Sequence read archives (SRA) under project accession PRJNA590225. Other associated data including bacterial, fungal and viral profiles for all patients, as well as patients clinical attributes are available via the study's code repository **https://github.com/translational-respiratory-lab/The_Interactome**.

Declaration of interests: JDC has received research grants from Glaxosmithkline, Boehringer-Ingelheim, Astrazeneca, Gilead Sciences, Grifols and Insmed and has received personal fees from GSK, BI, AZ, Chiesi, Grifols, Napp, Novartis, Insmed and Zambon all outside the submitted work. SA has received personal fees from Bayer Healthcare, Grifols, Astra Zeneca, Zambon and research grants and personal fees from Chiesi, Insmed, GlaxoSmithKline, Menarini, ZetaCube Srl, Fisher & Paykel all outside the submitted work. FB has received research grants and personal fees from Astrazeneca, Bayer, Chiesi, GlaxoSmithKline, Guidotti, Grifols, Insmed, Menarini, Novartis, Pfizer, Zambon all outside the submitted work. All other authors have no conflicts of interest to declare.

REFERENCES

1 Chalmers, J. D. & Chotirmall, S. H. Bronchiectasis: new therapies and new perspectives. *The Lancet. Respiratory medicine* **6**, 715-726, doi:10.1016/s2213-2600(18)30053-5 (2018).

2 Flume, P. A., Chalmers, J. D. & Olivier, K. N. Advances in bronchiectasis: endotyping, genetics, microbiome, and disease heterogeneity. *Lancet (London, England)* **392**, 880-890, doi:10.1016/s0140-6736(18)31767-7 (2018).

3 Chalmers, J. D., Chang, A. B., Chotirmall, S. H., Dhar, R. & McShane, P. J. Bronchiectasis. *Nature reviews. Disease primers* **4**, 45, doi:10.1038/s41572-018-0042-3 (2018).

4 Chotirmall, S. H. & McShane, P. J. Time to acknowledge, address, and take action against bronchiectasis. *Lancet Glob Health* **7**, e1162-e1163, doi:10.1016/s2214-109x(19)30311-0 (2019).

5 Woo, T. E. *et al.* A longitudinal characterization of the Non-Cystic Fibrosis Bronchiectasis airway microbiome. *Sci Rep* **9**, 6871, doi:10.1038/s41598-019-42862-y (2019).

6 Rogers, G. B., Bruce, K. D., Martin, M. L., Burr, L. D. & Serisier, D. J. The effect of long-term macrolide treatment on respiratory microbiota composition in non-cystic fibrosis bronchiectasis: an analysis from the randomised, double-blind, placebo-controlled BLESS trial. *The Lancet. Respiratory medicine* **2**, 988-996, doi:10.1016/s2213-2600(14)70213-9 (2014).

7 Rogers, G. B. *et al.* A novel microbiota stratification system predicts future exacerbations in bronchiectasis. *Annals of the American Thoracic Society* **11**, 496-503, doi:10.1513/AnnalsATS.201310-335OC (2014).

8 Mac Aogain, M. *et al.* Immunological corollary of the pulmonary mycobiome in bronchiectasis: the CAMEB study. *The European respiratory journal* **52**, doi:10.1183/13993003.00766-2018 (2018).

9 Mitchell, A. B. *et al.* Viruses in bronchiectasis: a pilot study to explore the presence of community acquired respiratory viruses in stable patients and during acute exacerbations. *BMC pulmonary medicine* **18**, 84, doi:10.1186/s12890-018-0636-2 (2018).

10 Cox, M. J. *et al.* Longitudinal assessment of sputum microbiome by sequencing of the 16S rRNA gene in non-cystic fibrosis bronchiectasis patients. *PloS one* **12**, e0170622, doi:10.1371/journal.pone.0170622 (2017).

11 Crichton, M. L., Aliberti, S. & Chalmers, J. D. A systematic review of pharmacotherapeutic clinical trial end-points for bronchiectasis in adults. *Eur Respir Rev* **28**, doi:10.1183/16000617.0108-2018 (2019).

12 Mac Aogain, M. *et al.* Distinct "Immunoallertypes" of Disease and High Frequencies of Sensitization in Non-Cystic Fibrosis Bronchiectasis. *American journal of respiratory and critical care medicine* **199**, 842-853, doi:10.1164/rccm.201807-1355OC (2019).

13 Kumpitsch, C., Koskinen, K., Schöpf, V. & Moissl-Eichinger, C. The microbiome of the upper respiratory tract in health and disease. *BMC Biol* **17**, 87, doi:10.1186/s12915-019-0703-z (2019).

14 Chalmers, J. D. *et al.* Characterization of the "Frequent Exacerbator Phenotype" in Bronchiectasis. *American journal of respiratory and critical care medicine* **197**, 1410-1420, doi:10.1164/rccm.201711-2202OC (2018).

15 Faust, K. *et al.* Microbial co-occurrence relationships in the human microbiome. *PLoS Comput Biol* **8**, e1002606, doi:10.1371/journal.pcbi.1002606 (2012).
16 Layeghifard, M., Hwang, D. M. & Guttman, D. S. Disentangling Interactions in the Microbiome: A Network Perspective. *Trends Microbiol* **25**, 217-228, doi:10.1016/j.tim.2016.11.008 (2017).
17 Finch, S., McDonnell, M. J., Abo-Leyah, H., Aliberti, S. & Chalmers, J. D. A Comprehensive Analysis of the Impact of Pseudomonas aeruginosa Colonization on Prognosis in Adult Bronchiectasis. *Annals of the American Thoracic Society* **12**, 1602-1611, doi:10.1513/AnnalsATS.201506-333OC (2015).
18 Laska, I. F., Crichton, M. L., Shoemark, A. & Chalmers, J. D. The efficacy and safety of inhaled antibiotics for the treatment of bronchiectasis in adults: a systematic review and meta-analysis. *Lancet Respir Med* **7**, 855-869, doi:10.1016/S2213-2600(19)30185-7 (2019).
19 Sibila, O. *et al.* Airway Bacterial Load and Inhaled Antibiotic Response in Bronchiectasis. *American journal of respiratory and critical care medicine* **200**, 33-41, doi:10.1164/rccm.201809-1651OC (2019).
20 Budden, K. F. *et al.* Functional effects of the microbiota in chronic respiratory disease. *The Lancet. Respiratory medicine*, doi:10.1016/s2213-2600(18)30510-1 (2019).
21 Li, C. X., Wheelock, C. E., Skold, C. M. & Wheelock, A. M. Integration of multi-omics datasets enables molecular classification of COPD. *Eur Respir J* **51**, doi:10.1183/13993003.01930-2017 (2018).
22 Aliberti, S. *et al.* Clinical phenotypes in adult patients with bronchiectasis. *The European respiratory journal* **47**, 1113-1122, doi:10.1183/13993003.01899-2015 (2016).
23 Tunney, M. M. *et al.* Detection of anaerobic bacteria in high numbers in sputum from patients with cystic fibrosis. *Am J Respir Crit Care Med* **177**, 995-1001, doi:10.1164/rccm.200708-1151OC (2008).
24 Muhlebach, M. S. *et al.* Anaerobic bacteria cultured from cystic fibrosis airways correlate to milder disease: a multisite study. *Eur Respir J* **52**, doi:10.1183/13993003.00242-2018 (2018).
25 Jackson, D. W. *et al.* Biofilm formation and dispersal under the influence of the global regulator CsrA of Escherichia coli. *J Bacteriol* **184**, 290-301, doi:10.1128/jb.184.1.290-301.2002 (2002).
26 Singh, P. K. *et al.* Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature* **407**, 762-764, doi:10.1038/35037627 (2000).
27 Sass, G. *et al.* Intermicrobial interaction: Aspergillus fumigatus siderophores protect against competition by Pseudomonas aeruginosa. *PLoS One* **14**, e0216085, doi:10.1371/journal.pone.0216085 (2019).
28 Mowat, E. *et al.* Pseudomonas aeruginosa and their small diffusible extracellular molecules inhibit Aspergillus fumigatus biofilm formation. *FEMS Microbiol Lett* **313**, 96-102, doi:10.1111/j.1574-6968.2010.02130.x (2010).
29 Briard, B., Heddergott, C. & Latgé, J. P. Volatile Compounds Emitted by Pseudomonas aeruginosa Stimulate Growth of the Fungal Pathogen Aspergillus fumigatus. *mBio* **7**, e00219, doi:10.1128/mBio.00219-16 (2016).
30 Whelan, F. J. *et al.* Culture-enriched metagenomic sequencing enables in-depth profiling of the cystic fibrosis lung microbiota. *Nat Microbiol* **5**, 379-390, doi:10.1038/s41564-019-0643-y (2020).

31 Sulaiman, I. *et al.* Evaluation of the airway microbiome in nontuberculous mycobacteria disease. *The European respiratory journal* **52**, doi:10.1183/13993003.00810-2018 (2018).
32 Ali, N., Mac Aogain, M., Morales, R. F., Tiew, P. Y. & Chotirmall, S. H. Optimisation and Benchmarking of Targeted Amplicon Sequencing for Mycobiome Analysis of Respiratory Specimens. *International journal of molecular sciences* **20**, doi:10.3390/ijms20204991 (2019).

Figure 1. Integration of multi-biome data through weighted-SNF identifies a cluster of bronchiectasis patients who frequently exacerbate. Overview of the multi-biome 'integrative microbiomics' strategy for analysis of 'bacteriome', 'mycobiome' and 'virome' datasets employing weighted Similarity Network Fusion (SNF) in bronchiectasis. (a) A heatmap illustrating the relative abundance of the top 20 identified taxa within the bacterial (blue) and fungal (green) communities respectively, and, the 17 viruses examined (red). Collectively, these kingdoms form the 'multi-biome' in the airway of patients with stable bronchiectasis (n=217). Relative taxonomic abundance is expressed in the heatmap according to depth of colour (0-100%). (b) A schematic overview of conventional (unweighted) and our derived weighted Similarity Network Fusion (SNF) approach to assess the airway multi-biome. Weightage is assigned to each -biome dataset based on its taxonomic richness. The weight for each biome is calculated by the number of observed taxa present with a prevalence of at least 5% across the patient cohort. Weighted SNF reflects the *in-vivo* state and overcomes weaknesses of conventional SNF methodologies. (c) A heatmap illustrating pairwise patient weighted-SNF similarity scores (range; 0 - 0.5, blue - red) assessed by spectral clustering. Two distinct patient clusters are illustrated by blue and purple bars above the heatmap respectively. Clusters are colour-coded as (1) blue and (2) purple. (d) Linear discriminant effect size (LEfSe) analysis of the observed clusters illustrating taxa that discriminate between the multi-biome profiles of each group. A bar plot details each of the identified discriminant taxa ranked by their effect size. Discriminant taxa with a log-transformed effect size of >3 are presented (n = 31). Alpha diversity and clinical features of the identified multi-biome clusters are illustrated by box and whisker plots showing (e) alpha diversity (Shannon diversity index), (f) number of exacerbations in the preceding year, (g) breathlessness score (MMRC), (h) lung function (as FEV1 % predicted) and (i) bronchiectasis severity index (BSI) for patients in each cluster. Prefixes indicate whether identified taxa are bacterial (B), fungal (F) or viral (V) and

699 significance levels for observed between-group differences are indicated as follows: ns: non-
700 significant; *p<0.05; **p<0.01; ***p<0.001.
701
702 **Figure 2.** Co-occurrence analysis reveals a less complex but distinct 'multi-biome' network
703 among high frequency bronchiectasis exacerbators characterised by a greater proportion of
704 negative microbial interactions. (a-b) Co-occurrence network maps of low (blue) and high
705 (purple) exacerbation frequency clusters illustrating all identified inter-kingdom interactions
706 between bacteria, fungi and viruses of the multi-biome. Interactions between microbes (or
707 nodes) are represented by connecting lines (edges) where the number of interactions for each
708 microbe (or node) is reflected by node size linked to the scale bar provided. Selected bacterial
709 (light blue), fungal (green) and viral (red) taxa of clinical relevance are indicated by node
710 coloration. (c) Summary table illustrating the network characteristics of the low and high
711 exacerbation frequency networks illustrating the total number of detectable microbes (nodes)
712 within each network and the total number of interactions (edges) separating out the number of
713 negative interactions (negative edges) observed as a proportion of the overall detected
714 interactions. (d-e) Visualization of positive and negative interactions between the most
715 abundant taxa in each respective cluster is illustrated and coloured according to the figure
716 legend. Interactions between microbes are classified as negative if the sign of the edge weights
717 between them is negative (negative correlation) and *vice versa.*
718
719

Figure 3. 'Busy', 'Critical' and 'Influential' microbes (nodes) are characterized by distinct interaction networks among frequently exacerbating bronchiectasis patients. Network visualization of key microbial taxa in (a) low and (b) high exacerbation frequency bronchiectasis clusters. Coloured circles represent microbes and grey lines their associated interactions within the network. Taxa present at >1% relative abundance in at least 5% of the patient cohort are included. Circle size (degree) reflects the number of direct interactions for a given microbe (termed 'busy'). Circle outline thickness represents the calculated stress centrality for each microbe termed 'critical' while circle colour depth reflects betweenness centrality or the 'influence' of the microbe within the network. Two bacterial (*Streptococcus* and *Haemophilus*) and two fungal (*Candida* and *Cryptococcus*) genera all demonstrate high calculated network metrics (in both clusters) and are therefore considered 'busy', 'critical' and 'influential' within each network and are indicated by red borders. (c) Co-occurrence analysis illustrates distinct *Pseudomonas*-interaction networks associated with increased exacerbation frequency in bronchiectasis. An overview of microbial interaction networks in the low- (left side) and high- (right side) exacerbation frequency clusters. Microbes not interacting directly with *Pseudomonas* (i.e. not part of the *Pseudomonas*-interaction network) are coloured according to their respective clinical cluster membership as blue (low exacerbation frequency cluster) and purple (high exacerbation frequency cluster). Lines connecting microbes directly interacting with *Pseudomonas* are coloured to reflect positive or negative interaction as indicated in the colour legend. Colour depth reflects strength of the interaction (edge weight). (d) Co-occurrence analysis highlighting the *Pseudomonas*-interaction network in the low- and high- exacerbation frequency clusters (left and right sided respectively). Interactions are positioned according to their positive (green; top) or negative (red; bottom) association with *Pseudomonas* spp and where indicated in bold are reflective of an antithesis relationship.

Figure 4. Longitudinal analysis of the integrated multi-biome during bronchiectasis exacerbations. (a) Bacterial, fungal and viral community status were assessed longitudinally in n=17 bronchiectasis patients at baseline (pre-exacerbation) ('B'), during an established pulmonary exacerbation ('E') and then post-exacerbation ('P') following completion of antibiotic therapy. Pie charts illustrate aggregate microbial composition of the bacterial, fungal and viral community profiles across each timepoint with the most abundant taxa indicated by colour legend. (b) Boxplots illustrating comparable α-diversity across baseline ('B'), exacerbation ('E') and post-exacerbation ('P') specimens. Dotted lines indicate the longitudinal pattern of each individual patient (n=17). (c) Non-metric Multi-Dimensional Scaling (NMDS) plot illustrating comparable multi-biome β-diversity across baseline ('B'), exacerbation ('E') and post-exacerbation ('P') specimens. Samples are grouped according to their respective longitudinal time-point and timepoints indicated by coloured planes (d-f) Visualization of the interactome's positive and negative interactions between the most abundant taxa at (d) baseline (pre-exacerbation), (e) during exacerbation and (f) post-exacerbation. Interactions between microbes are classified as negative if the sign of the edge weights between them is negative and *vice versa*, as indicated in the colour legend. To study the stability of interactions longitudinally across the three timepoints, the relative change in strength of an interaction (defined as "maximal (interaction strength) – minimal (interaction strength)") across timepoints was assessed. (g) Relative interaction change is plotted comparing the changes occurring between baseline and exacerbation ('B vs E') and exacerbation vs post exacerbation ('E v P'). Pairwise matrices indicate the comparative change in interaction observed between individual bacteria, fungi or viruses. Magnitude of change is indicated by the presented colour scale.

Figure 5. 'Integrative microbiomics' of the multi-biome reveals a core and ancillary microbial network in bronchiectasis exacerbations. (a) A Venn diagram summarizing the observed interactions of the multi-biome across the longitudinal sampling timepoints (baseline: blue, exacerbation: red, post exacerbation: green) and their intersections. (b) Network analysis illustrating that a 'core' microbial network is present across the three longitudinal timepoints assessed (pre-exacerbation, during exacerbation and post-exacerbation). This occurs in parallel to the 'ancillary' microbial network implicated in bronchiectasis exacerbation. The presented network summary captures the common interactions to that in the baseline (pre-exacerbation) reference network and outlines condition specific networks (during exacerbation and post-exacerbation). Levels of conservation for each specific interaction within the network is colour-coded with blue indicating highly conserved interactions and purple highly variable interactions across the course of a bronchiectasis exacerbation. (c) Baseline network analysis of bronchiectasis patients who subsequently received β-lactam therapy for treatment of an exacerbation (n=12). (d) a simulated network based on 75% reduction in the abundance of β-lactam-susceptible organisms and calculation of the re-configured network. (e) observed network reconfiguration in patients following β-lactam therapy. Circle size, outline thickness and colour respectively represent node importance based on network metrics; degree, stress centrality and betweenness centrality (*c.f.* figure 3).

Figure 6. Functional characterisation of the high-exacerbation frequency cluster and its associated interactome. (a) Functional mapping of metagenomic data from n=20 patients from the low-exacerbation (LEF, blue) and high exacerbation frequency (HEF, purple) clusters was performed identifying 113 discriminant genes in the HEF cluster compared to 16 genes in the LEF cluster by Linear discriminant effect size (LEfSe) analysis. A bar plot illustrates effect size observed for discriminant genes in each group. Discriminant genes with a log-transformed effect size of >2 are presented (n = 129) (b) KEGG Pathway mapping of identified genes indicating enriched functional pathways in the HEF and LEF clusters. (c) Node and edge plots extracted from the LEF and HEF network cluster analysis (figure 3) highlighting opposing interactions between *P. aeruginosa* and *A. fumigatus* related to exacerbation frequency. Edges are coloured green or red reflecting a positive (co-occurrence) or negative (co-exclusion) interaction, respectively. Circle size, outline thickness and colour respectively represent node importance based on network metrics; degree, stress centrality, and betweenness centrality (*c.f.* figure 3). (d) Demonstration of strain-dependant inter-kingdom interaction between *P. aeruginosa* and *A. fumigatus*. Comparison of direct interactions between *P. aeruginosa* laboratory strain ('PAO1'; grey) and isolates obtained from patients from the LEF and HEF clusters respectively ('LEF'; blue, 'HEF'; purple) with *A. fumigatus* (Af293) by disk inhibition assays. Colony zone diameter is indicated by a red circle for *P. aeruginosa* strains grown in the presence (+) or absence of (-) Af293 at 24h and 48h timepoints, respectively. (e) Analysis of *P. aeruginosa* zone diameters observed following co-culture with Af293 following 24h and 48h incubation. Bars are coloured according to the respective *P. aeruginosa* strain as described above. Open bars indicate zone diameters observed in the absence of *A. fumigatus* and filled bars indicate zone diameters observed on co-culture. Error bars represent the standard deviation of triplicate determinations. ns: non-significant; **p<0.01; ***p<0.001.

Figure 7. Metagenomics reveals shifts in gene function and validates microbial interactions associated with clinical exacerbation in bronchiectasis. (a) Heatmap illustrating pairwise patient similarity scores (range; 0 - 0.5, cyan - turquoise) assessed by spectral clustering of functional gene assignments. Two distinct patient clusters are indicated (FC1 and FC2). (b) Linear discriminant effect size (LEfSe) analysis of the observed gene functional pathways that discriminate between each group (FC1 and FC2). A bar plot details each of the identified discriminant taxa ranked by their effect size. Discriminant pathways with a log-transformed effect size of >3 are presented (n = 42). Clinical features of the identified multi-biome clusters are illustrated by box and whisker plots showing (c) number of exacerbations in the preceding year and (d) bronchiectasis severity index (BSI) for patients in each cluster; ns: non-significant; ***$p<0.001$. The aggregate relative abundance of (e) bacteria, (f) viruses and (g) fungi determined by taxonomic assignment of metagenomic reads across an independently recruited cohort of bronchiectasis patients (n=166). (h) Heatmap illustrating pairwise similarity scores and resultant patient stratification based on spectral clustering according to each individual - biome view from metagenomics analysis (bacteria: blue, virus: red and fungi: green) and the weighted-SNF integrated microbiome (purple). (i) Co-occurrence analysis of taxa identified in wSNF clusters SC1 (low exacerbation) and (j) SC2 (high exacerbation). Circles represent microbial nodes and grey lines associated interactions within the network. Circle size (degree) reflects the number of direct interactions for a given microbe (termed 'busy'). Circle outline thickness represents the calculated stress centrality for each microbe termed 'critical' while circle colour depth reflects betweenness centrality or the 'influence' of the microbe within the network. Bacteriome, virome and mycobiome nodes are respectively indicated by blue. red, green borders. (k) Network analysis illustrating conserved microbial interactions (bacteria and fungi only) present across targeted and metagenomic interactome analysis. Common

838 interactions to both networks are colour-coded according to the strength of the conserved
839 interactions.