medRχiv
THE PREPRINT SERVER FOR HEALTH SCIENCES

The fungal intestinal microbiota predict the development of bronchopulmonary dysplasia in very low birthweight newborns

Kent A. Willis[1,2], Mary Silverberg[1], Isaac Martin[1], Ahmed Abdelgawad[1], Kosuke Tanaka[1], Ibrahim Karabayir[3], Brian A. Halloran[1], Erin D. Myers[4], Jay P. Desai[5], Catrina T. White[5], Charitharth V. Lal[1], Namasivayam Ambalavanan[1], Brian M. Peters[6], Viral G. Jain[1], Oguz Akbilgic[3,7], Laura Tipton[8], Tamás Jilling[1], Stephania A. Cormier[9,10], Joseph F. Pierre[11], and Ajay J. Talati[5,12]

[1]Division of Neonatology, Department of Pediatrics, Heersink School of Medicine, The University of Birmingham, Birmingham, AL, USA.
[2]Microbiome Center, The University of Alabama at Birmingham, Birmingham, AL, USA.
[3]Division of Cardiology, Department of Medicine, Wake Forest University School of Medicine, Wake Forest University, Winston-Salem, NC, USA.
[4]College of Medicine, The University of Tennessee Health Science Center, Memphis, TN, USA.
[5]Division of Neonatology, Department of Pediatrics, College of Medicine, The University of Tennessee Health Science Center, Memphis, TN, USA.
[6]Department of Clinical Pharmacy and Translational Science, College of Pharmacy, The University of Tennessee Health Science Center, Memphis, TN, USA.
[7]Epidemiological Cardiology Research Center, Wake Forest University, Winston-Salem, NC, USA.
[8]Department of Data Science, Analytics, and Visualization, Chaminade University of Honolulu, Honolulu, HI, USA.
[9]Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, USA.
[10]Pennington Biomedical Research Center, Baton Rouge, LA, USA.
[11]Department of Nutritional Sciences, College of Agricultural and Life Sciences, The University of Wisconsin – Madison, Madison, WI, USA.
[12]Department of Obstetrics and Gynecology, College of Medicine, The University of Tennessee Health Science Center, Memphis, TN, USA.

Abstract

Rationale: Bronchopulmonary dysplasia (BPD) is the most common morbidity affecting very preterm infants. Gut microbial communities contribute to multiple lung diseases, and alterations of the gut microbiome may be a factor in BPD pathogenesis.

Objective: To determine if features of the multikingdom gut microbiome predict the development of BPD in very low birthweight newborns.

Methods: We performed a prospective, observational cohort study comparing the multikingdom fecal microbiota of 147 preterm infants with BPD or post-prematurity respiratory disease (PPRD) by sequencing bacterial 16S and fungal ITS2 ribosomal RNA gene. To address the potential causative relationship between gut dysbiosis and BPD, we used fecal microbiota transplant in an antibiotic-pseudohumanized mouse model. Comparisons were made using RNA sequencing, quantitative confocal microscopy, lung morphometry, and oscillometry.

Measurements and Main Results: We analyzed 100 fecal microbiome samples collected during the second week of life. Infants who later developed BPD showed an obvious fungal dysbiosis as compared to infants with PPRD (P = 0.0209). Instead of fungal communities dominated by *Candida* and *Saccharomyces*, the microbiota of infants who developed BPD were

characterized by the abundance of *Aureobasidium* and a greater diversity of rarer fungi in less interconnected community architectures. On successful colonization, the gut microbiota from infants with BPD augmented lung injury in the offspring of recipient animals. We identified alterations in the murine lung and intestinal microbiomes and transcriptional alterations associated with augmented lung injury.

Conclusions: The gut fungal microbiome of infants who will develop BPD is dysbiotic and may contribute to disease pathogenesis.

Trial Registration: NCT03229967.

Keywords: Multikingdom microbiome, gut microbiome; fungal microbiome; mycobiome; preterm infant; chronic lung injury

Introduction

Bronchopulmonary dysplasia (BPD) is the most common morbidity of prematurity (2, 3). Disproportionately affecting the most premature newborns, BPD exacts a devastating toll on the immature lung, impairing alveolarization and disrupting vascularization (4). Despite the considerable, life-long morbidity and staggeringly high mortality of BPD (2, 3), meaningful interventions to modulate lung injury in preterm newborns remain beyond reach (5). While environmental triggers, such as exposure to hyperoxia, have been described, modulators of the disease remain elusive.

The gut-lung axis is a recently described phenomenon whereby intestinal commensal microbes may project their influence on lung physiology and pathology(6, 7). Examples of this connection are described in models of neonatal pneumonia (8, 9) and asthma (10, 11). Intestinal microbes can influence the lung directly via the production of metabolites and indirectly by informing immune mechanisms that module lung injury and recovery (8, 10). While most previous work has focused on the bacterial microbiota, other microbiota components, such as fungi, may prove to be critical mediators of immunity (10, 12).

To determine if the multikingdom (bacterial and fungal) gut microbiota predicts the development of BPD, we performed a prospective, observational cohort study of very low birthweight (VLBW) preterm newborns. The principal outcome was moderate to severe BPD or death before BPD adjudication. To interrogate causality, we also performed fecal microbiota transplant (FMT) into antibiotic-pseudohumanized mice. We hypothesized that characteristics of the fungal microbiota would predict BPD in VLBW infants and that this effect would be transferable to *in vivo* models of BPD. Here, we show the gut-lung axis may influence the development of BPD and that dysbiotic gut microbiota can augment BPD-like illness via FMT.

Methods

Detailed methods are provided in the online supplement.

Study Subjects:

This study was a prospective observational cohort study of inborn, VLBW (< 1500g) preterm newborns without known congenital anomalies or immunodeficiencies (MiBPD study: NCT03229967). We defined BPD by the NICHD Workshop definition (13). For dichotomous analyses, 'BPD' represents the composite outcome of Grade II to III BPD or death before 36 weeks of gestation, with the remainder defined as 'PPRD' (post-prematurity respiratory disease) (14). At the time of this study, probiotics were not administered in this unit. This analysis used the first non-transitional stool collected during the second week of life. A second cohort provided the samples for FMT to create antibiotic-pseudohumanized mice.

Animal experimental protocol:

To create pseudo-humanized mice (15–17), we purchased conventional specific pathogen-free C57Bl6/J mice from the Jackson Laboratory and randomized them into cages of four female mice each. After acclimatization, we exposed the mice to a week of systemically absorbable antibiotics, followed by gastric gavage of polyethylene glycol, then another week of non-absorbable antibiotics. With the native microbiota thus disrupted, we engrafted human microbiota from infants with BPD or PPRD by FMT. We bred these mice to generate neonatal offspring we termed BPD-FMT or PPRD-FMT. Pooled litters of neonates with the same exposure

were then randomized to either normoxia (FiO_2 0.21) or hyperoxia (FiO_2 0.85) exposure from postnatal day (P) 3-14 using our standardized model of BPD (18, 19).

Transcriptomics:
RNA-seq was performed as described (19) from the left lower lobe.

Oscillometry:
We performed forced oscillometry on anesthetized 14-day-old mice using the flexiVent FX System (SCIREQ) (23).

Histology and immunohistochemistry:
We used 5µm hematoxylin and eosin-stained formalin-fixed paraffin-embedded gravity-inflated to 20 cm H_2O lung sections for morphometry (20). Morphometry was performed with standardized techniques (18, 21, 22) compatible with the ATS guidelines (23). We stained sections with von Willebrand Factor (vWF) and alpha-smooth muscle actin (aSMA) to quantify vascular density.

Bacterial and fungal metagenomic sequencing:
Human microbiome samples were collected in oxygen-reduced PBS with a sterile spatula immediately following evacuation into a clean diaper. For mouse samples, we collected the whole left lung and terminal ileum (20). Microbial DNA was extracted, amplified, and sequenced as described (21, 22). Sequence data were processed with *QIIME 2* (23), *DADA2* (24), and *SCRuB* (25) and then analyzed in R. Details regarding data processing and ecological analyses are available in the online supplement.

Machine learning models:
We developed a light gradient boosted machine (LightGBM) model using clinical and phylum level microbial data on a randomly selected 80%-subset of participants. We then validated the model using a test set of the remainder.

Statistical analysis:
Statistical analyses were performed in SPSS (IBM), Prism 9 (GraphPad), Python, or R. Animal experiments are representative of 2-3 pooled mouse litters per human donor with at least 3 human donors per experimental group and analyzed by two-way ANOVA with a Šidák correction.

Results

We performed a prospective, observational cohort study of 147 VLBW infants in neonatal intensive care. 100 infants produced their first non-meconium stool during the second week of life, from which we characterized the intestinal microbiota. No samples were excluded due to sequencing quality. Cohort characteristics are described in Table 1.

Dysbiosis of the gut microbiota in infants with BPD.
We first compared the intestinal microbiota of preterm infants who later developed BPD to infants who did not develop BPD (PPRD cohort). As shown in Figure 1, the fungal intestinal microbiota (mycobiome) of BPD infants differed in community diversity, composition, and interconnectivity from PPRD infants. We did not identify differences in the multikingdom microbiome.

For alpha diversity, the within-group variation exceeded the between-group differences for both the mycobiome and multikingdom microbiome. The alpha diversity of the mycobiome was greater in BPD as quantified by the Shannon and Chao1 indices (P = 0.0293 and 0.0008, Mann-Whitney, Figure 1B). The combined increase in both indices indicates an increase in microbial richness produces altered alpha diversity in fungal samples. The diversity of the multikingdom microbiome was similar (Figure 1C).

We used complementary supervised clustering methods to quantify beta diversity. As visualized by principal coordinates analysis, disease status results in divergent clustering (Figure 1D). The mycobiome is differently composed in BPD infants (P = 0.0209, permutational multivariate ANOVA (PERMANOVA)). We verified this was not a false positive due to unequal dispersion with permutational multivariate analysis of dispersion (PERMDISP, P = 0.3997). However, the multikingdom microbiome was not statistically different (P = 0.0580, PERMANOVA, P = 1, PERMDISP). We performed biplot analysis of principal components analysis graphs to identify which amplicon sequence variants (ASVs) produced these compositional differences. We determined that multiple ASVs belonging to the genera *Candida* and an ASV aligning to *Aureobasidium* made the most significant contributions to the structure of the mycobiome. A similar analysis for the multikingdom microbiome showed *Klebsiella* and *Enterococcus* made the

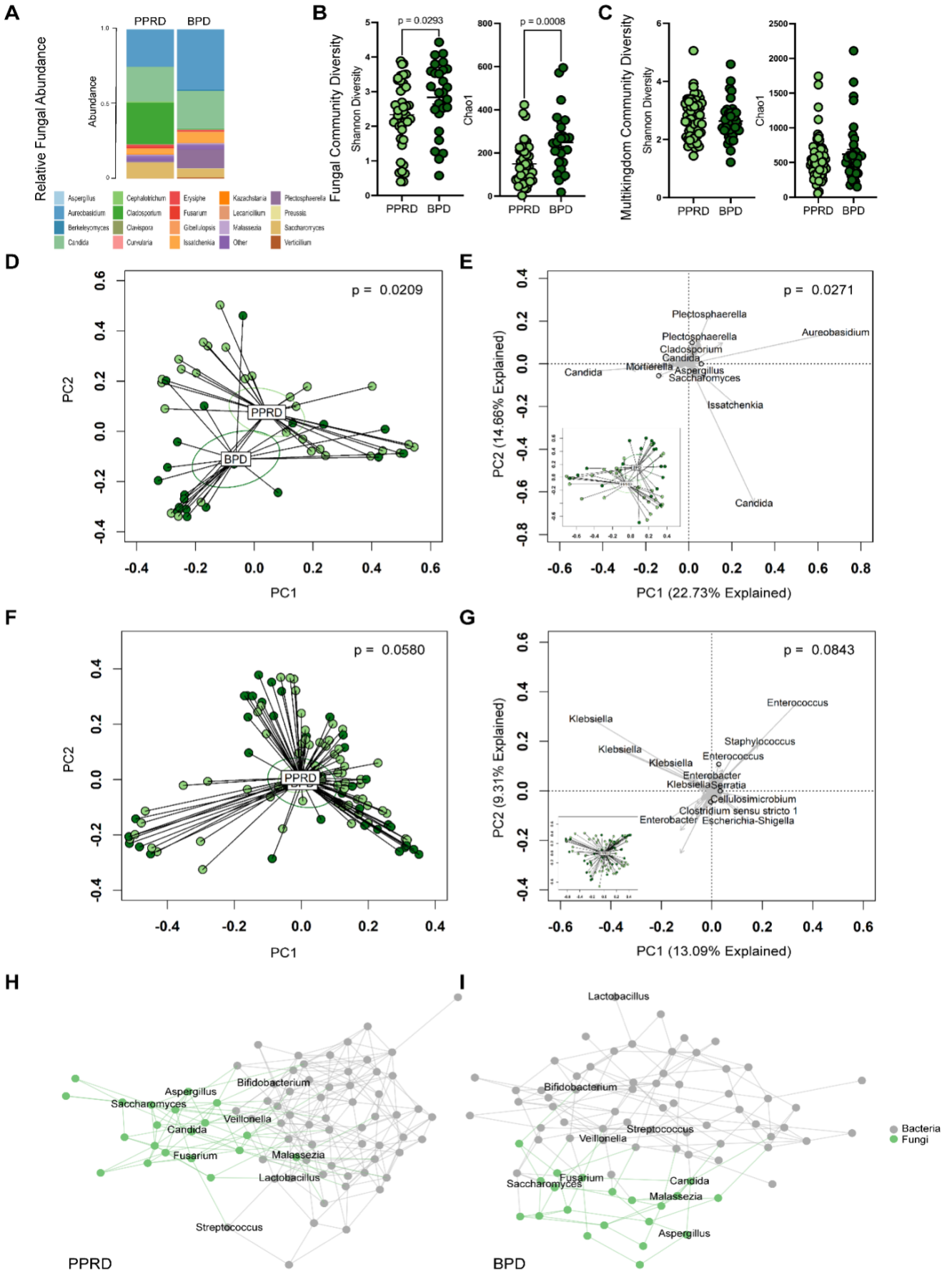
A
PPRD
BPD
Relative Fungal Abundance
Abundance
Aspergillus
Aureobasidium
Berkeleyomyces
Candida
Cephalotrichum
Cladosporium
Clavispora
Curvularia
Erysiphe
Fusarium
Gibellulopsis
Issatchenkia
Kazachstania
Lecanicillium
Malassezia
Other
Plectosphaerella
Preussia
Saccharomyces
Verticillium
B
Fungal Community Diversity
Shannon Diversity
p = 0.0293
Chao1
p = 0.0008
PPRD
BPD
C
Multikingdom Community Diversity
Shannon Diversity
Chao1
PPRD
BPD
D
p = 0.0209
PC1
PC2
PPRD
BPD
E
p = 0.0271
PC1 (22.73% Explained)
PC2 (14.66% Explained)
Plectosphaerella
Plectosphaerella
Aureobasidium
Cladosporium
Candida
Candida
Mortierella
Aspergillus
Saccharomyces
Issatchenkia
Candida
F
p = 0.0580
PC1
PC2
PPRD
G
p = 0.0843
PC1 (13.09% Explained)
PC2 (9.31% Explained)
Enterococcus
Klebsiella
Staphylococcus
Klebsiella
Enterococcus
Klebsiella
Enterobacter
Klebsiella
Serratia
Cellulosimicrobium
Clostridium sensu stricto 1
Enterobacter
Escherichia-Shigella
H
Aspergillus
Saccharomyces
Bifidobacterium
Veillonella
Candida
Fusarium
Malassezia
Lactobacillus
Streptococcus
PPRD
I
Lactobacillus
Bifidobacterium
Streptococcus
Veillonella
Fusarium
Saccharomyces
Candida
Malassezia
Aspergillus
BPD
Bacteria
Fungi

Figure 1. The fungal gut microbiota is altered in infants that will develop bronchopulmonary dysplasia (BPD). (*A*) Compared to infants that will develop post-prematurity respiratory disease (PPRD), BPD infants have different relative abundance of certain fungi. BPD infants have greater fungal (*B*) but not multikingdom (*C*) community diversity. Principal coordinates analysis revealed that BPD infants' mycobiome is composed differently (*D*), and biplots of principal components analysis show this is driven by members of the *Candida* and *Aureobasidium* genera (*E*). Differences are not observed in the multikingdom microbiome (*F* and *G*). SPIEC-EASI network analysis shows that BPD infants formed sparser networks than PPRD infants (*H* and *I*). Ellipses indicate the 95% confidence interval.

strongest contributions to the overall community structure (Figure 1G).

To quantify the interconnectivity of these microbial communities, we used a novel cross-domain extension of the SPIEC-EASI method (12) to build multikingdom interaction networks (Figure 1H-I). Compared to PPRD infants, BPD infants have fewer fungal nodes but the same number of multikingdom nodes (Figure E1D). Similarly, they have fewer fungal and multikingdom edges (Figure E1E). Together, this creates multikingdom interaction networks that were sparser for BPD infants, with fewer fungi-to-fungi and fungi-to-bacterial interactions than PPRD infants.

We used DESeq2 to identify differentially abundant genera. For the mycobiome, the genera *Candida* and *Saccharomyces* were more abundant in PPRD infants (Figure E1). In contrast, for the multikingdom microbiome, while nine ASVs were differentially abundant, we did not detect any significantly different genera by disease status. These analyses indicate the mycobiome of infants that will develop BPD was unique. It was characterized by a lower abundance of *Candida* and *Saccharomyces* and a higher number of rarer fungal taxa arranged in more sparse and less interconnected multikingdom networks.

Fungal dysbiosis predicts BPD development.

To test the predictive potential of the mycobiota, we first developed random forest machine-learning models using fungal or multikingdom microbiome and clinical variables as an exploratory analysis. The fungal model (Figure E3A-B), which had an area under the receiver operating curve of 0.81, identified several rare fungal genera, such as *Plectosphaerella*, *Gamszarea*, and *Verticillium,* that outperform several clinical criteria. No bacterial genera in the multikingdom model are more important than the major clinical criteria (Figure E3C-D).

To test prognostic potential at greater depth, we developed a LightGBM model using clinical variables known to predict the development of BPD from the Neonatal BPD Outcomes Predictor (26, 27) and higher-order taxa microbiota data (Figure 2). After five-fold cross-validation and Bayesian hyperparameter tuning, the model predicts the risk of developing BPD in the unseen validation set with an area under the curve of 0.81 ($CI^{95\%}$ 0.61-1.00), with a sensitivity of 80% and specificity of 89%, positive predictive value of 89%, and an F1-score of 0.89 (Figure 2A). For the composite outcome (severe BPD or death), a second model obtained an area under the curve of 0.89 ($CI^{95\%}$ 0.76-1.00), with a sensitivity of 75% and specificity of 86%, a positive predictive value of 90%, and F1-score of 0.82. Intriguingly, in variable importance analysis, several fungal phyla's relative abundance strongly contributed to the model (Figure E3E), including the minor phyla Mortierellomycota and Rozellomycota and the major phyla Ascomycota and Basidiomycota. The fungal taxa outperformed several well-described clinical characteristics and all bacterial phyla except Bacteroidata.

Sex-specific differences in the composition of the mycobiome.

Due to known sex-specific differences in BPD (28, 29), we investigated if sex-specific features exist in the mycobiome. To test this, we performed supervised clustering using the combination of disease status and sex assigned at birth (Figures 3 and E4-5). Shannon diversity of the mycobiome was lower in infants assigned female at birth (AFAB) than those assigned male at birth (AMAB, Figure 3B), which may indicate decreased evenness in these samples. Fungal

community composition differed for all groups (P = 0.0046, PERMANOVA; P = 0.6935, PERMDISP, Figure 3D). In principal coordinates analysis, the AFAB infants with BPD and PPRD clustered closer than infants AMAB

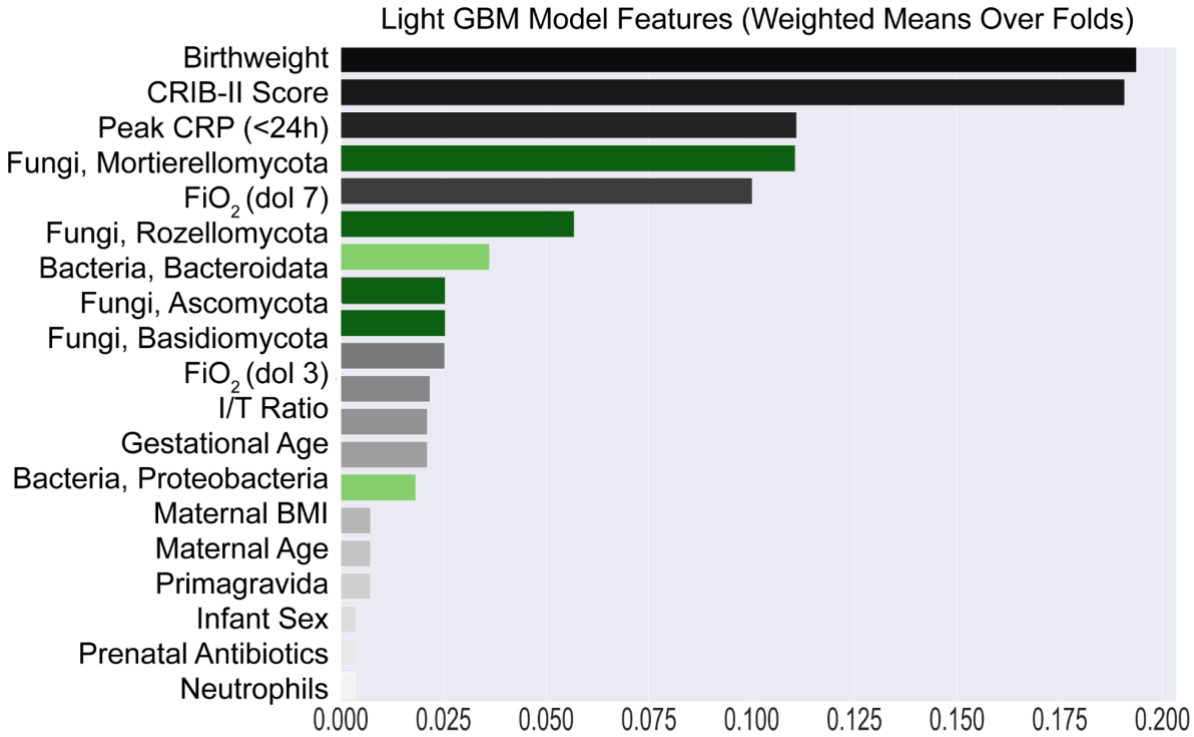

Figure 2. The mycobiota predicts the development of BPD. (*A*) A light gradient-boosted machine learning model (LightGBM) with Bayesian hyperparameter tuning developed from clinical data and phylum-level microbiome data indicates that key members of the mycobiome predict the development of BPD. CRIB-II, critical risk index for babies-II. CRP, c-reactive protein. FiO_2, fraction of inspired oxygen. I/T, immature to total lymphocyte ratio on admission complete blood count. BMI, body mass index. Neutrophils, neutrophil count on admission complete blood count.

(Figure 3D). DESeq2 identified differences in the abundance of multiple fungal genera. Among infants who later developed BPD, differences in the relative abundance of *Trichoderma* and *Saccharomyces* were characteristic of the mycobiome of infants AFAB, while in AMAB infants, differences in the abundance of *Trichoderma*, *Plectosphaerella*, *Issatchenkia*, and *Geotrichum* were the most important. We detected no differences in overall community diversity or composition of the multikingdom microbiome (Figure 3C, E). SPIEC-EASI network analysis identified greater interconnectivity of infants with PPRD regardless of sex assigned at birth (Figure 3F-I). Infants AMAB had fewer fungal edges than infants AFAB, but the number of edges was reduced only in infants AMAB that would develop BPD (Figure E5D). These differences produced sparser networks in infants AMAB, with the least interconnected networks in infants AMAB with BPD (Figure E5E). From these analyses, we conclude that sex-specific differences in the composition of the mycobiome are the strongest in infants AMAB with BPD.

Distinct clusters in fungal community composition are linked to clinical outcomes.

Having shown that fungal microbiota features are associated with BPD development, we sought to identify other clinical characteristics that may contribute to heterogeneity in the composition of the neonatal mycobiome. First, we used PERMANOVA to screen potential additional contributors to fungal community composition. Of the potential mycobiome covariates we examined, only total days of oxygen exposure remained significant after adjustment for false discovery (Q = 0.016, Table E2, Figure E6). When we tested potential multikingdom microbiome covariates, significant covariates included birthweight, prenatal and postnatal antibiotic exposure, culture-

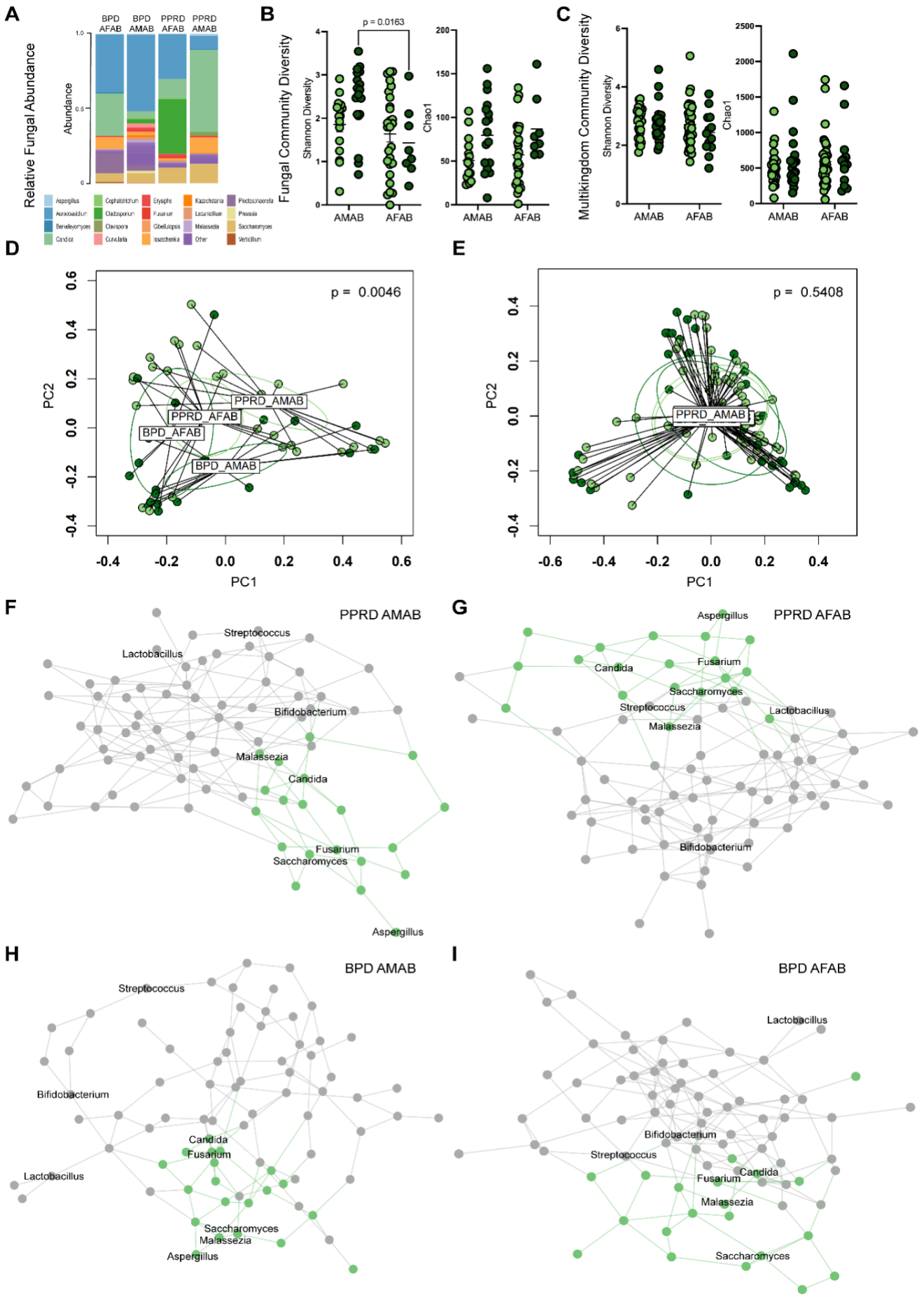
A
BPD AFAB
BPD AMAB
PPRD AFAB
PPRD AMAB
Relative Fungal Abundance
Abundance
B
Fungal Community Diversity
Shannon Diversity
p = 0.0163
Chao1
AMAB
AFAB
C
Multikingdom Community Diversity
Shannon Diversity
Chao1
AMAB
AFAB
D
p = 0.0046
PPRD_AMAB
PPRD_AFAB
BPD_AFAB
BPD_AMAB
PC1
PC2
E
p = 0.5408
PPRD_AMAB
PC1
PC2
F
PPRD AMAB
Streptococcus
Lactobacillus
Bifidobacterium
Malassezia
Candida
Fusarium
Saccharomyces
Aspergillus
G
PPRD AFAB
Aspergillus
Candida
Fusarium
Saccharomyces
Streptococcus
Malassezia
Lactobacillus
Bifidobacterium
H
BPD AMAB
Streptococcus
Bifidobacterium
Candida
Fusarium
Lactobacillus
Saccharomyces
Malassezia
Aspergillus
I
BPD AFAB
Lactobacillus
Bifidobacterium
Streptococcus
Candida
Fusarium
Malassezia
Saccharomyces

Figure 3. Sex-specific differences in infants that will develop BPD. Differences in the relative abundance of certain fungi in infants assigned male (AMAB) or female at birth (AFAB) with PPRD and BPD. (*B*) Sex-specific differences in fungal Shannon diversity in infants with BPD are not observed in the multikingdom microbiome (*C*). Principal coordinates analysis revealed that the mycobiome differed for each sex and disease state (*D*) but not the multikingdom microbiome (*E*). SPIEC-EASI network analysis showed that BPD infants AMAB formed the sparsest networks (*F-I*). Ellipses indicate the 95% confidence interval.

negative sepsis, and relative abundance of fungi (Table E3).

We next pursued unsupervised clustering to explore heterogeneity in fungal community features using Dirichlet-Multinomial Mixtures (DMM) modeling (30) (Figure 4). Laplace analysis for the whole mycobiome identifies 2 or 5 components as the optimal fit points (Figure E7A). Membership in Component 1 of the two-component DMM model (Figure 4A-C) was driven by the relative abundance of a *Candida* ASV and a *Saccharomyces* ASV. In contrast, membership in Component 2 was due to a different *Candida* ASV and an *Aureobasidium* ASV (Figure E7B). We then used canonical correspondence analysis (CCA) to map clinical characteristics that may be associated with the fungal composition profiles identified by the DMM components. Component 2 was closely associated with BPD and increased illness severity (CRIB-II). In contrast, Component 1 was associated with increasing gestational age and birthweight. Interestingly, high fungal evenness (arbitrarily defined as >50% of reads aligning to one fungal ASV) was also linked to Component 1 (Figure E7C).

In a five-component DMM model (Figure 4D-F), Component 1 and 2 membership was driven by the same ASVs as the two-component model (Figure E7D). Clustering into Component 3 was driven primarily by the abundance of an *Aureobasidium* ASV. Component 5 was related but dominated by *Aureobasidium* ASV to a greater degree. However, Component 4 was uniquely dominated by a *Saccharomyces* ASV. CCA indicates similar clinical associations with Components 1 and 2 as the two-component model (Figure E7E). BPD was very closely aligned with Component 3. Component 5 was aligned with BPD to a slightly lesser degree and with critical illness to a greater degree. Component 2 was only moderately aligned to BPD compared to Components 3 and 5. The divergence of Component 2 was likely explained by the more significant influence of mortality and critical illness with the composition of this cluster.

Finally, we developed three-component DMM models independently for infants with either PPRD (Figure 4G-I) or BPD (Figure 4J-L). These results, described in greater detail in the online supplement, confirmed that Components 3 and 5 from the whole mycobiome DMM model are subsets of the BPD-associated mycobiome that align most closely with the diagnosis of BPD. These unsupervised analyses confirm that differences in the relative abundance of the genera *Candida*, *Saccharomyces*, and *Aureobasidium* characterize the BPD-associated mycobiome.

Fecal microbiota transfer transmits a severe phenotype of BPD.

To test if alterations in the gut microbiome could augment lung injury in BPD, we transferred the gut microbiome from infants with either BPD or PPRD into mice. Following antibiotic preparation, recipient mice were randomly orally inoculated with anoxic fecal samples from donor infants. These mice were bred, and the resulting newborn mice received a second FMT from the same human donor after birth. We then exposed the neonatal mice to hyperoxia to induce BPD-like lung injury (Figure 5A).

Neonatal mice that received BPD-associated microbiomes (BPD-FMT) had more significant histological lung injury when exposed to hyperoxia (greater mean linear intercept and reduced radial alveolar count, $P < 0.0001$ and $P = 0.0005$, two-way ANOVA, respectively, Figure 5B-C) than mice that received PPRD-associated microbiomes (PPRD-FMT). Functionally, BPD-FMT mice had increased pulmonary resistance ($P = 0.0015$) and decreased compliance ($P = 0.0196$). We used immunohistochemistry and morphometry to measure neointimal changes

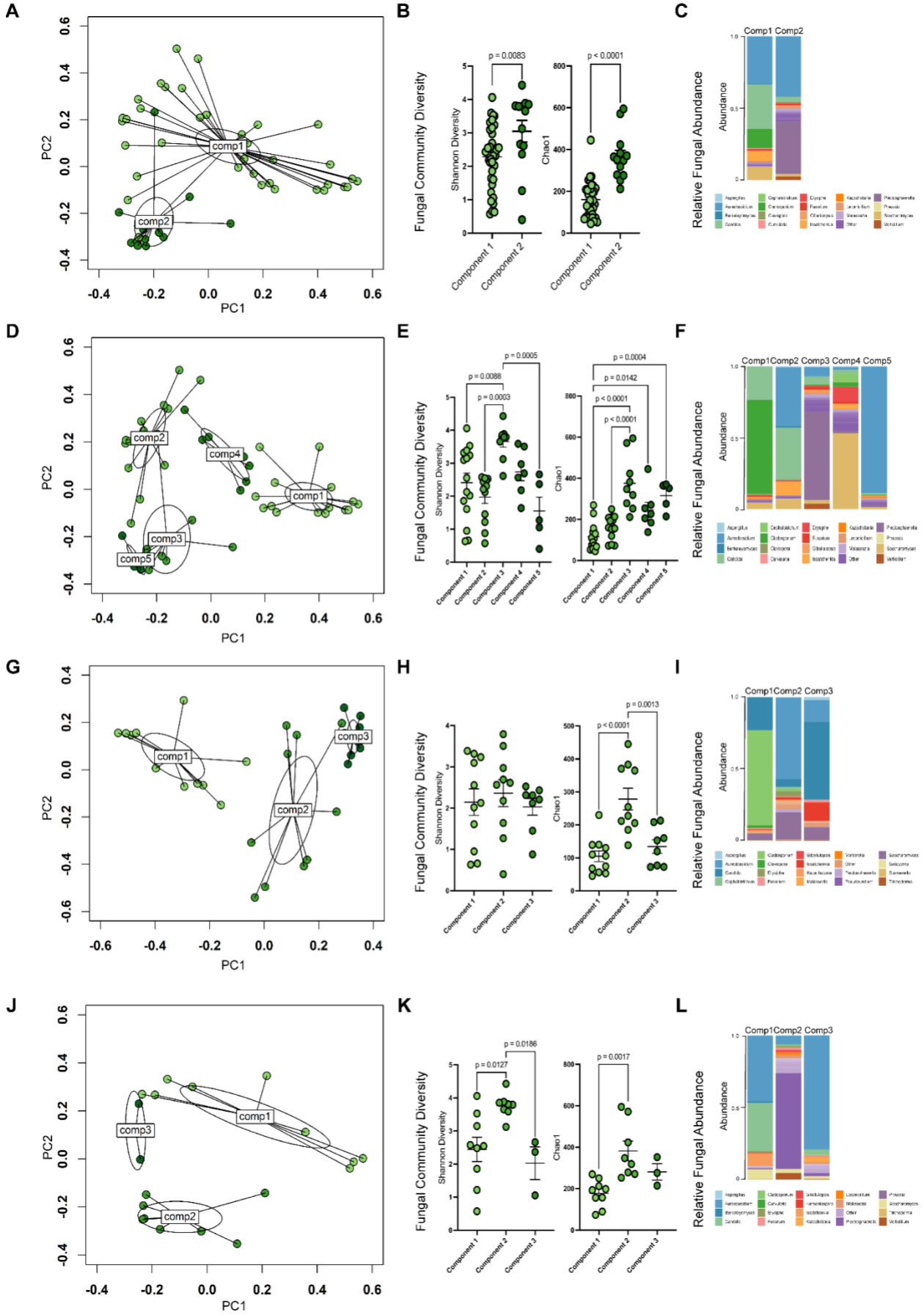

Figure 4. Unsupervised clustering analysis of heterogeneity in the mycobiome. (*A*) Two-cluster Dirichlet-Multinomial Mixture (DMM) model of the mycobiome. (*B*) Fungal community diversity analyses indicate higher microbial richness in Component 2

(Comp2). (*C*) Relative abundance in each DMM cluster. (*D-F*) Five-cluster DMM model with the highest microbial richness in Component 3. (*G-I*) A three-cluster DMM model in infants with post-prematurity respiratory disease with the highest richness in Component 2. (*J-L*) A three-cluster DMM model in infants with bronchopulmonary dysplasia with the highest microbial richness in Component 2. Comp1-5, Component 1-5. Ellipses indicate the 95% confidence interval.

consistent with pulmonary hypertension (Figure 5D) and identified BPD-FMT mice had reduced vascular density of 20-50 µm diameter vessels (P = 0.0065, Figure 5E).

We next used RNAseq to quantify transcriptomic changes in the lungs of these mice. Principal coordinates analysis confirmed unique global transcriptomic changes in the lungs of mice by both FMT and hyperoxia exposure (Figure 5F). We also visualized differences in transcriptome as a heatmap (Figure 5G). Volcano plots show a clear difference in the response of PPRD-FMT (Figure 5H) and BPD-FMT (Figure 5I) to hyperoxia exposure.

Fecal microbiota transfer alters the neonatal mouse microbiome.

We also characterized the mouse lung and gut microbiomes to identify how FMT and hyperoxia exposure influences their ecology. In the mouse gut and lung, 79% and 69% of the genera were shared with human donor microbiomes. Beta but not alpha diversity was significantly different for the gut mycobiome (P = 0.0062, Figure 6A) and multikingdom microbiome (P = 0.0167, Figure 6B). In the lungs, only the multikingdom microbiome was uniquely composed (P = 0.0103, Figure 6C-D). From these analyses, we conclude that a BPD-associated microbiome is transferable to neonatal mice and results in augmented lung injury from hyperoxia exposure.

Discussion

BPD is the most common morbidity and severe chronic respiratory disease in newborns (5). Considerable heterogeneity exists in its clinical phenotype, and effective means to predict and modulate disease remain unavailable because the underlying etiology remains unclear. Here, we sought new answers in the gut. We found that in preterm neonates who would eventually develop BPD the fungal intestinal microbiota was different several weeks before diagnosis from neonates who did not develop BPD (PPRD cohort). Furthermore, these differences were driven by fungal commensals. With FMT, we showed that gut microbiota transferred from the BPD cohort could augment lung injury as compared to transfer from PPRD. Our findings suggest the gut-lung axis may play a crucial role in the etiology of BPD.

Microeukaryotes are vital members of the human microbiota (31). However, unlike bacterial commensals, their non-pathological and non-parasitic functions remain poorly understood (31). So while the microbiome has recently emerged as a contributor to clinical heterogeneity in adult patients in intensive care (32–40), our knowledge of fungal commensals in newborn respiratory disease is limited (41). Here, in one of the most extensive studies of the multikingdom intestinal microbiome in preterm newborns, we show that the composition of intestinal fungal communities predicts the development of BPD. The presence of fungi alone and the specific abundance of certain higher-order fungal taxa were similarly predictive of the risk of developing BPD as well-established clinical criteria (26, 27, 42). Intriguingly, we also identified robust sex-specific differences between infants who would develop BPD that were not observed in infants who did not develop BPD. In light of recent sex-specific differences in the development of BPD detected by single-cell transcriptomics (28, 43–45), this is an interesting finding worth further investigation.

Overall, our analyses found that the composition of the intestinal mycobiota of infants that would develop PPRD were more uniform, whereas those that eventually developed BPD had more disparate mycobiota. This suggests that a particular pattern of mycobiota development may be necessary to impart resistance to the development of BPD, and failure to do so in various ways is associated with disease development. A more robust health-associated microbiome with a less specific disease-associated

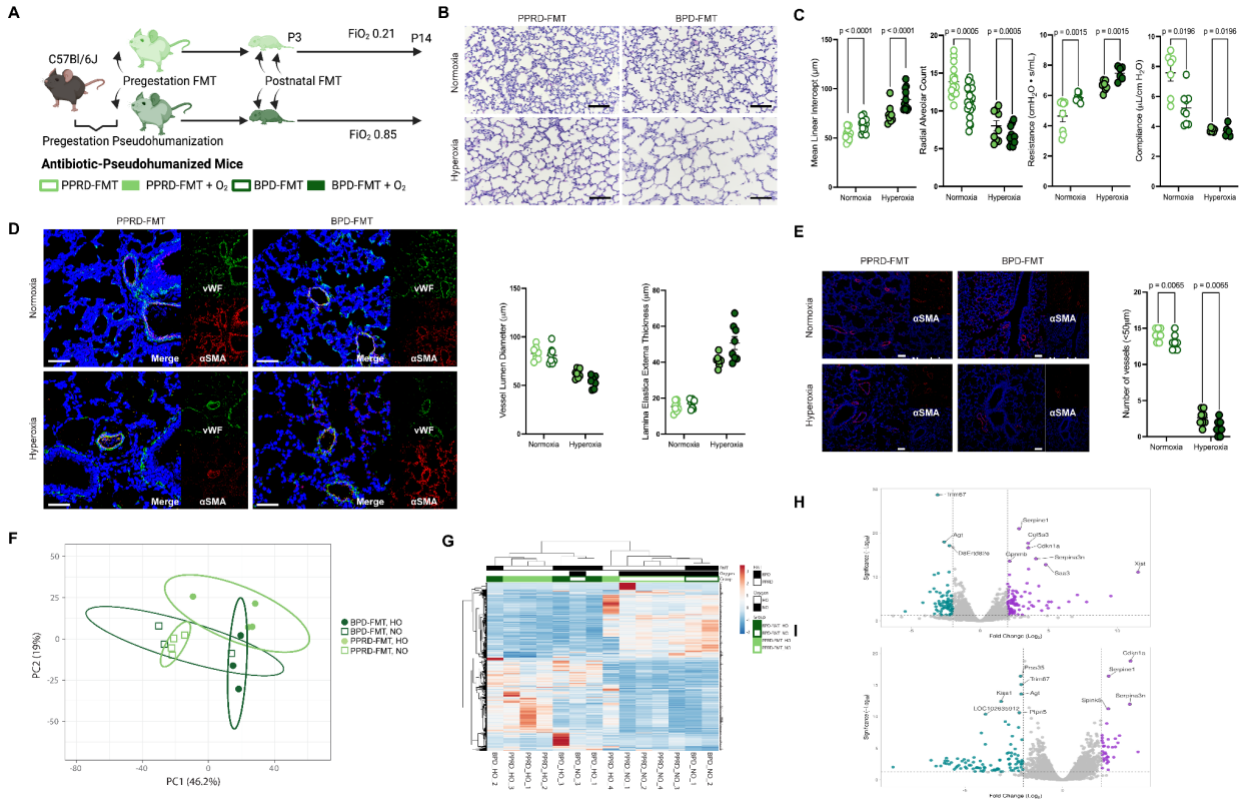

Figure 5. Fecal microbiota transfer from infants with BPD augments lung injury in mice. (*A*) Experimental schematic. (*B-C*) Fecal microbiota transfer from BPD-associated microbiomes (BPD-FMT) augments lung injury as assessed by histomorphological analysis and forced oscillometry. Scale bars, 100 µm. (*D-E*) Immunohistochemistry is consistent with vascular changes consistent with early pulmonary hypertension (*D*) and reduced vascular density of 20-50 µm vessels (*E*). Scale bars, 20 and 50 µm. (*F-I*) Hyperoxia exposure induces different transcriptomic profiles in BPD-FMT or PPRD-FMT mice by RNA-seq. (*F*) Principal components analysis shows unique clustering by both disease and FMT. (*G*) Heatmap of global transcriptomic changes. (*H*) Volcano plot of the hyperoxia-induced transcriptomic alterations in PPRD-FMT. (*I*) Volcano plot of hyperoxia-induced transcriptomic alterations in BPD-FMT. FiO_2, fraction of inspired oxygen. Schematic created using BioRender.com.

microbiome is a well-described ecological pattern termed the Anna Karenina phenomenon (46). This paradigm holds for other aspects of the pathogenesis of BPD (47), therefore, this finding fits within our broader understanding of the disease.

A previous study of the bacterial microbiota identified a divergence over time between the bacterial community composition in the gut and lungs of preterm infants that would eventually develop BPD (48), suggesting that a gut-lung axis may influence the development of BPD. While the mechanisms underlying the gut-lung axis require further exploration in BPD, work in other childhood diseases has suggested some potential mechanisms. Mouse models of childhood asthma suggest intestinal fungi are more immunostimulatory than bacteria, while intestinal bacteria are more responsible for producing metabolites detectable in the host's serum (10). Gut intestinal fungi, mainly the yeast *Candida albicans*, are known to provoke strong immune responses in the host (49). Indeed, monocolonization with *Candida* can stimulate postnatal immune development (50). The greater abundance of *Candida* dominated communities we observed in PPRD infants is intriguing in this regard.

Limitations of this study include that we transplanted whole fecal samples in the mouse FMT experiment. While we analyzed the microbiota transferred to the neonatal mice, it remains unclear which components of this complex mix of microbes, metabolites, and nutrients modulate lung injury. Future experiments

might use a multiomic approach to identify and validate potential candidates to address this limitation in more reductionist experiments. Similarly, it is currently not feasible to confine the transfer of microbes to only one organ system. Quantifying how spillover from FMT might affect other host organs, such as the lungs, remains challenging. A strength of this study was the rigorous exploration of the multikingdom microbiota using the most up-to-date bioinformatic techniques (41) and the demonstration that the microbiome effect was transferable by FMT.

Conclusions:

In this prospective, observational cohort study of VLBW preterm newborns, baseline compositional differences in the fungal intestinal microbiome predicted the eventual development of BPD. Furthermore, we showed that this dysbiotic microbiota could augment lung injury in a mouse model of BPD. These findings suggest that the gut microbiota can modulate lung injury and represents a therapeutic target in newborn lung disease.

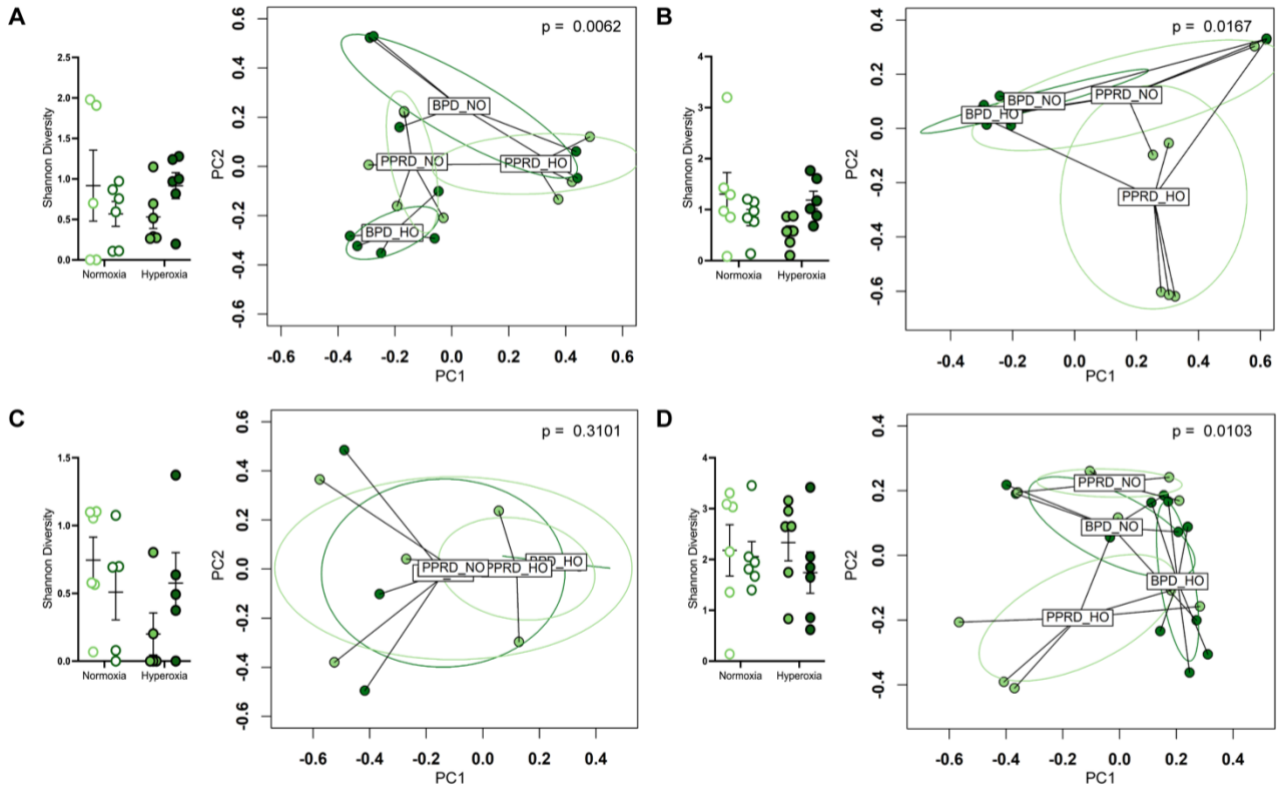

Figure 6. Fecal microbiota transfer alters the lung and gut microbiome of recipient mice. (*A-B*) The gut fungal (*A*) and multikingdom microbiomes (*B*) differed by fecal microbiota transfer. (*C-D*) The multikingdom lung microbiota is different but not the lung mycobiome. BPD_NO, BPD normoxia. BPD_HO, BPD hyperoxia. PPRD_NO, PPRD normoxia. PPRD_HO, BPD hyperoxia. Ellipses indicate the 95% confidence interval.

Table 1. Cohort Clinical Characteristics and Demographics

	PPRD	BPD	*P* Value
n	64	38	
Gestational age (weeks)	29.0 (27.6 - 30.8)	25.7 (25 - 27.8)	< 0.001
Birthweight (g)	1145 (840 - 1330)	775 (670 - 860)	< 0.001
Female sex assigned at birth	37 (57.8%)	14 (36.8%)	0.041
Antenatal Steroids (No)	8 (17%)	4 (13.3%)	0.33
Surfactant (None)	27 (43.5%)	1 (2.6%)	<0.001
Apgar score (1 min)	6 (3 - 7)	4 (2 - 5)	< 0.001
Apgar score (5 min)	7 (6 - 8)	6 (5 - 7)	< 0.001
Small for gestational age*	6 (9.4%)	8 (21.1%)	0.098
Ceasarean section	28 (59.6%)	20 (66.7%)	0.531
Prolonged rupture of membranes‡	7 (15%)	5 (16.7%)	0.834
Critical Risk Index for Babies II (CRIB II) score	7 (4 - 9)	12 (9 - 14)	< 0.001
Antenatal antibiotic exposure	22 (46.8%)	15 (50%)	0.785
Duration of antenatal antibiotic exposure (d)	1 (0 - 1)	0 (0 - 1)	0.542
Postnatal antibiotic exposure	55 (85.9%)	38 (100%)	0.015
Duration of postnatal antibiotic exposure (d)	2 (1 - 2)	2 (1 - 4)	0.016
Respiratory Support on Day 3			
Mechanical ventilation	20 (37%)	33 (89.2%)	< 0.001
Fraction of inspired oxygen	0.21 (0.21 - 0.21)	0.25 (0.21 - 0.34)	< 0.001
Predicted risk of severe BPD†	4.40% (1.68 - 10.14%)	10.70% (8.73 - 14.49%)	< 0.001
Predicted risk of death†	2.84% (1.36 - 6.30%)	10.90% (7.69 - 16.00%)	< 0.001
Respiratory Support on Day 7			
Mechanical ventilation	15 (28.3%)	30 (83.4%)	< 0.001
Fraction of inspired oxygen	0.21 (0.21 - 0.23)	0.29 (0.24 - 0.30)	< 0.001
Predicted risk of severe BPD†	4.35% (1.54 - 9.40%)	11.09% (8.22 - 21.50%)	< 0.001
Predicted risk of death†	3.35% (1.1 - 5.52%)	10.21% (7.01 - 13.2%)	< 0.001
Endotracheal intubation	41 (66.1%)	37 (100%)	< 0.001
Ventilation duration (d)	1 (0 - 4)	34 (22 - 50)	< 0.001
Non-invasive ventilation duration (d)	5 (1 - 17)	14 (1 - 38)	0.053
Oxygen exposure duration (d)	13 (4 - 41)	92 (30 - 119)	< 0.001
Supplemental oxygen on day 28	23 (37.1%)	33 (91.7%)	<0.001
Death	0 (0%)	6 (15.8%)	< 0.001

Data are presented as median (interquartile range) or n (%) with significance testing by Mann-Whitney U or Pearson chi-squared test, as appropriate.
BPD, Bronchopulmonary dysplasia; PPRD, Post-prematurity respiratory disease. *< 3rd percentile, ‡ > 18 h, †NICHD Neonatal Outcomes Estimator

Author Contributions: Conceptualization: KW, NA, BP, SC, JP, AT. Experimentation: KW, MS, AA, BH, KT, CW. Clinical sample and Data Collection: KW, MS, IM, EM, JD. Analysis: KW, IM, IK, OA, LT. Manuscript: KW, NA, TJ, SC, JP, AT. All authors have approved the final manuscript as submitted.

Funding: Research reported in this article was supported by the National Institutes of Health under award numbers K08 HL151907 (KW), L40 HL159581 (KW), KL2 TR3097 (VJ), K08 HL141652 (CVL), R01 HL156275 (NA), R01 HL129907 (NA), R21 DK129890K08 (JP), and R21 AI163503 (JP); the Dixon Fellowship (MS), the Microbiome Center of the University of Alabama at Birmingham (KW), and the Division of Neonatology at the University of Tennessee Health Science Center (AT).

Data Availability: Raw sequencing files are deposited at the National Center for Biotechnology Information Sequence Read Archive (BioProject: *Pending*). Processed data files are available at github.com/WillisLungLab/MiBPD

Disclosures: CL is the founder of ResBiotic Nutrition, Inc. and Alveolus Bio, Inc. KW and NA are advisors.

Inclusion and Diversity:
We support the inclusive, diverse, and equitable conduct of research. Multiple authors of this paper self-identify as underrepresented ethnic or gender minorities, or as a member of the LGBTQA+ community. While citing

references scientifically relevant to this work, we also actively worked to promote gender balance in our reference list.

Acknowledgments:
We thank the infants, their parents, and the study coordinators, without whom this study would have been impossible. Confocal microscopy was performed at the UAB High-Resolution Imaging Facility with the assistance of Robert Grabski.

References

1. Willis KA, Silverberg M, Abdelgawad A, Martin I, Tanaka K, Halloran B, *et al.* The Fungal Intestinal Microbiota Predict the Development of Bronchopulmonary Dysplasia in Very Low Birthweight Newborns. *B96 Best Pediatrics* 2023;A4223–A4223.doi:10.1164/ajrccm-conference.2023.207.1_meetingabstracts.a4223.
2. Lapcharoensap W, Lee HC, Nyberg A, Dukhovny D. Health Care and Societal Costs of Bronchopulmonary Dysplasia. *Neoreviews* 2018;19:e211–e223.
3. Lui K, Lee SK, Kusuda S, Adams M, Vento M, Reichman B, et al. Trends in Outcomes for Neonates Born Very Preterm and Very Low Birth Weight in 11 High-Income Countries. *The Journal of Pediatrics* 2019;215:32-40.e14.
4. Kinsella JP, Greenough A, Abman SH. Bronchopulmonary dysplasia. *The Lancet* 2006;367:1421–1431.
5. Thébaud B, Goss KN, Laughon M, Whitsett JA, Abman SH, Steinhorn RH, et al. Bronchopulmonary dysplasia. *Nat Rev Dis Primers* 2019;5:78.
6. Dang AT, Marsland BJ. Microbes, metabolites, and the gut–lung axis. *Mucosal Immunol* 2019;12:843–850.
7. Willis KA, Physiology-Lung JS of, 2020. Recent Advances in Understanding the Ecology of the Lung Microbiota and Deciphering the *Gut*-Lung Axis. *journals.physiology.org* doi:10.1152/ajplung.00360.2020?casa_token=vzqvpyfk__gaaaaa:xicgvv8haq3a6xe0hnzg7rpv7okk41wvnifpwf6mm17xb7y8ztncebzzxkzp6ouwsm-etae_wn0.
8. Stevens J, Steinmeyer S, Bonfield M, Peterson L, Wang T, Gray J, et al. The balance between protective and pathogenic immune responses to pneumonia in the neonatal lung is enforced by gut microbiota. *Sci Transl Med* 2022;14:.
9. Deshmukh HS, Liu Y, Menkiti OR, Mei J, Dai N, O'Leary CE, et al. The microbiota regulates neutrophil homeostasis and host resistance to Escherichia coli K1 sepsis in neonatal mice. *Nature Medicine* 2014;20:524–530.
10. Bernardes E van T, Pettersen VK, Gutierrez MW, Laforest-Lapointe I, Jendzjowsky NG, Cavin J-B, et al. Intestinal fungi are causally implicated in microbiome assembly and immune development in mice. *Nature communications* 2020;11:2577–16.
11. Arrieta M-C, Stiemsma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutsch S, et al. Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Science translational medicine* 2015;7:307ra152-307ra152.
12. Tipton L, Müller CL, Kurtz ZD, Huang L, Kleerup E, Morris A, et al. Fungi stabilize connectivity in the lung and skin microbial ecosystems. *Microbiome* 2018;6:12.
13. Higgins RD, Jobe AH, Koso-Thomas M, Bancalari E, Viscardi RM, Hartert TV, et al. Bronchopulmonary Dysplasia: Executive Summary of a Workshop. *J Pediatrics* 2018;197:300–308.
14. Cristea AI, Ren CL, Amin R, Eldredge LC, Levin JC, Majmudar PP, et al. Outpatient Respiratory Management of Infants, Children, and Adolescents with Post-Prematurity Respiratory Disease: An Official American Thoracic Society Clinical Practice Guideline. *Am J Resp Crit Care* 2021;204:e115–e133.
15. Staley C, Kaiser T, Beura LK, Hamilton MJ, Weingarden AR, Bobr A, et al. Stable engraftment of human microbiota into mice with a single oral gavage following antibiotic conditioning. *Microbiome* 2017;5:87.
16. Chen X, Li P, Liu M, Zheng H, He Y, Chen M-X, et al. Gut dysbiosis induces the development of pre-eclampsia through bacterial translocation. Gut 2020;gutjnl-2019-319101.doi:10.1136/gutjnl-2019-319101.

17. Wrzosek L, Ciocan D, Borentain P, Spatz M, Puchois V, Hugot C, et al. Transplantation of human microbiota into conventional mice durably reshapes the gut microbiota. *Scientific reports* 2018;8:6854.
18. Nardiello C, Mižíková I, Silva DM, Ruiz-Camp J, Mayer K, Vadász I, et al. Standardisation of oxygen exposure in the development of mouse models for bronchopulmonary dysplasia. *Disease models & mechanisms* 2017;10:185–196.
19. Willis KA, Siefker DT, Aziz MM, White CT, Mussarat N, Gomes CK, et al. Perinatal maternal antibiotic exposure augments lung injury in offspring in experimental bronchopulmonary dysplasia. *American Journal of Physiology - Lung Cellular and Molecular Physiology* 2019;3:21.
20. Baker JM, Hinkle KJ, McDonald RA, Brown CA, Falkowski NR, Huffnagle GB, et al. Whole lung tissue is the preferred sampling method for amplicon-based characterization of murine lung microbiota. Microbiome 2021;9:99.
21. Pierre JF, Phillips GJ, Chandra LC, Rendina DN, Thomas-Gosain NF, Lubach GR, et al. Lyticase Facilitates Mycobiome Resolution Without Disrupting Microbiome Fidelity in Primates. *J Surg Res* 2021;267:336–341.
22. Willis KA, Purvis JH, Myers ED, Aziz MM, Karabayir I, Gomes CK, et al. Fungi form interkingdom microbial communities in the primordial human gut that develop with gestational age. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2019;33:12825–12837.
23. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;37:852–857.
24. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods* 2016;13:581–583.
25. Austin GI, Park H, Meydan Y, Seeram D, Sezin T, Lou YC, et al. Contamination source modeling with SCRuB improves cancer phenotype prediction from microbiome data. *Nat Biotechnol* 2023;1–9.doi:10.1038/s41587-023-01696-w.
26. Greenberg RG, McDonald SA, Laughon MM, Tanaka D, Jensen E, Meurs KV, et al. Online clinical tool to estimate risk of bronchopulmonary dysplasia in extremely preterm infants. *Archives Dis Child - Fetal Neonatal Ed* 2022;107:638–643.
27. Laughon MM, Langer JC, Bose CL, Smith PB, Ambalavanan N, Kennedy KA, et al. Prediction of Bronchopulmonary Dysplasia by Postnatal Age in Extremely Premature Infants. *Am J Resp Crit Care Med* 2011;183:1715–1722.
28. Cantu A, Gutierrez MC, Dong X, Leek C, Sajti E, Lingappan K. Remarkable sex-specific differences at Single-Cell Resolution in Neonatal Hyperoxic Lung Injury. *Am J Physiol-lung C* 2022;doi:10.1152/ajplung.00269.2022.
29. Trembath A, Laughon MM. Predictors of Bronchopulmonary Dysplasia. *Clinics in perinatology* 2012;39:585–601.
30. Holmes I, Harris K, Quince C. Dirichlet Multinomial Mixtures: Generative Models for Microbial Metagenomics. *Plos One* 2012;7:e30126.
31. Laforest-Lapointe I, Arrieta M-C. Microbial Eukaryotes: a Missing Link in Gut Microbiome Studies. *Msystems* 2018;3:e00201-17.
32. Bongers KS, Chanderraj R, Woods RJ, McDonald RA, Adame MD, Falkowski NR, et al. The Gut Microbiome Modulates Body Temperature Both in Sepsis and Health. Am J Resp Crit Care 2022;doi:10.1164/rccm.202201-0161oc.
33. Moutsoglou DM, Tatah J, Prisco SZ, Prins KW, Staley C, Lopez S, et al. Pulmonary Arterial Hypertension Patients Have a Proinflammatory Gut Microbiome and Altered Circulating Microbial Metabolites. Am J Resp Crit Care 2022;doi:10.1164/rccm.202203-0490oc.
34. Tashiro H, Shore SA. The Gut Microbiome and Ozone-induced Airway Hyperresponsiveness. Mechanisms and Therapeutic Prospects. *Am J Resp Cell Mol* 2021;64:283–291.

35. Wang Z, Locantore N, Haldar K, Ramsheh MY, Beech AS, Ma W, et al. Inflammatory Endotype–associated Airway Microbiome in Chronic Obstructive Pulmonary Disease Clinical Stability and Exacerbations: A Multicohort Longitudinal Analysis. Am J Resp Crit Care 2021;203:1488–1502.
36. Dickson RP, Schultz MJ, Poll T van der, Schouten LR, Falkowski NR, Luth JE, et al. Lung Microbiota Predict Clinical Outcomes in Critically Ill Patients. Am J Resp Crit Care 2020;201:555–563.
37. O'Dwyer DN, Ashley SL, Gurczynski SJ, Xia M, Wilke C, Falkowski NR, et al. Lung Microbiota Contribute to Pulmonary Inflammation and Disease Progression in Pulmonary Fibrosis. Am J Resp Crit Care 2019;199:1127–1138.
38. Ashley SL, Sjoding MW, Popova AP, Cui TX, Hoostal MJ, Schmidt TM, et al. Lung and gut microbiota are altered by hyperoxia and contribute to oxygen-induced lung injury in mice. Science translational medicine 2020;12:eaau9959.
39. Dickson RP, Singer BH, Newstead MW, Falkowski NR, Erb-Downward JR, Standiford TJ, et al. Enrichment of the lung microbiome with gut bacteria in sepsis and the acute respiratory distress syndrome. *Nat Microbiol* 2016;1:16113.
40. Combs MP, Wheeler DS, Luth JE, Falkowski NR, Walker NM, Erb-Downward JR, et al. Lung microbiota predict chronic rejection in healthy lung transplant recipients: a prospective cohort study. *Lancet Respir Medicine* 2021;9:601–612.
41. Alcazar CG-M, Paes VM, Shao Y, Oesser C, Miltz A, Lawley TD, et al. The association between early-life gut microbiota and childhood respiratory diseases: a systematic review. *Lancet Microbe* 2022;3:e867–e880.
42. Jensen EA, Dysart K, Gantz MG, McDonald S, Bamat NA, Keszler M, et al. The Diagnosis of Bronchopulmonary Dysplasia in Very Preterm Infants. An Evidence-based Approach. *American Journal of Respiratory and Critical Care Medicine* 2019;200:751–759.
43. Xia S, Ellis LV, Winkley K, Menden H, Mabry SM, Venkatraman A, et al. Neonatal hyperoxia induces activated pulmonary cellular states and sex-dependent transcriptomic changes in a model of experimental bronchopulmonary dysplasia. Am J Physiol-lung C 2022;doi:10.1152/ajplung.00252.2022.
44. Lao JC, Bui CB, Pang MA, Cho SX, Rudloff I, Elgass K, et al. Type 2 immune polarization is associated with cardiopulmonary disease in preterm infants. Sci Transl Med 2022;14:eaaz8454.
45. Hurskainen M, Mižíková I, Cook DP, Andersson N, Cyr-Depauw C, Lesage F, et al. Single cell transcriptomic analysis of murine lung development on hyperoxia-induced damage. *Nat Commun* 2021;12:1565.
46. Zaneveld JR, McMinds R, Thurber RV. Stress and stability: applying the Anna Karenina principle to animal microbiomes. Nat Microbiol 2017;2:17121.
47. Lal CV, Ambalavanan N. Biomarkers, Early Diagnosis, and Clinical Predictors of Bronchopulmonary Dysplasia. Clinics in perinatology 2015;42:739–754.
48. Gallacher D, Mitchell E, Alber D, Wach R, Klein N, Marchesi JR, et al. Dissimilarity of the gut-lung axis and dysbiosis of the lower airways in ventilated preterm infants. *The European respiratory journal* 2020;1901909.doi:10.1183/13993003.01909-2019.
49. Underhill DM, Iliev ID. The mycobiota: interactions between commensal fungi and the host immune system. *Nat Rev Immunol* 2014;14:405–416.
50. Jiang TT, Shao T-Y, Ang WXG, Kinder JM, Turner LH, Pham G, et al. Commensal Fungi Recapitulate the Protective Benefits of Intestinal Bacteria. *Cell host & microbe* 2017;22:809-816.e4.
51. Doron I, Mesko M, Li XV, Kusakabe T, Leonardi I, Shaw DG, et al. Mycobiota-induced IgA antibodies regulate fungal commensalism in the gut and are dysregulated in Crohn's disease. Nat Microbiol 2021;6:1493–1504.
52. Ost KS, O'Meara TR, Stephens WZ, Chiaro T, Zhou H, Penman J, et al. Adaptive immunity induces mutualism between commensal eukaryotes. Nature 2021;596:114–118.

The fungal intestinal microbiota predict the development of bronchopulmonary dysplasia in very low birthweight newborns

Kent A. Willis, Mary Silverberg, Isaac Martin, Ahmed Abdelgawad, Kosuke Tanaka, Ibrahim Karabayir, Brian A. Halloran, Erin D. Myers, Jay P. Desai, Catrina T. White, Charitharth V. Lal, Namasivayam Ambalavanan, Brian M. Peters, Viral Jain, Oguz Akbilgic, Laura Tipton, Tamás Jilling, Stephania A. Cormier, Joseph F. Pierre, Ajay J. Talati

ONLINE DATA SUPPLEMENT

Key Resource Table

Reagent	Company - Catalog #	RRID
Teklad S-2335 Mouse Breeder Sterilizable Diet	Envigo - 7904.15	
Penicillin VK	Thermo Fisher Scientific (TFS) - AC455230050	
Cefoperazone Sodium	TFS - AAJ6518506	
Clindamycin Sodium	TFS - 1G 11-101-3689	
Polyethylene glycol 3350 (Miralax)	TFS - J61495.AK	
Ertapenem Sodium	Fisher - 50-997-777 5G	
Vancomycin HCL	Milipore Sigma - 1404-93-9	
Neomycin Sodium	TFS - 50-186-3466	
Fecal Microbiota Transplant	This study only	
Lamina Propria Isolation Kit, Mouse	Miltenyi Biotec - 130-097-410	
Hank's Balanced Salt Solution	TFS - 14170112 or 14025092	
DNase I	Sigma - D5025	
Collagenase, Type I	Invitrogen - 7100017	
Fetal Bovine Serum	Millipore Sigma - F2442-500ML	
C-Tube	Miltenyi Biotec - 130-096-334	
RBC lysis buffer	Biolegend - 420301	
IC Fixation Buffer	TFS - 00-8222-49	
Permeabilization Buffer	TFS - 00-8333-56	
VMR Cell Strainers, 40, 70 µm	Avantor - 76327-098	
V-bottom, 96 well plate	Greiner Bio-One - 651901	
Sabouraud Dextrose Broth 4% (pH 5.6)	TFS - R064416	
zymoBIOMICS microbial community standards	Zymo Research - D6300 and D6310	
zymoBIOMICS microbial community DNA standards	Zymo Research - D6305/6306 and D6311	

zymoBIOMICS spike-in control I	Zymo Research - D6320-10	
zymoBIOMICS spike-in control II	Zymo Research - D6321-10	
zymoBIOMICS gut microbiome standard	Zymo Research - D6331	
RNase-free DNase Set	Qiagen - 79254	
QIAshredder	Qiagen - 79656	
PrimeScript RT Master Mix	Takara Biosciences - RR036B	
Bioanalyzer High Sensitivity DNA Analysis Kit	Agilent Technologies - 5067-4626	
Tissue Protein Extraction Reagent	Pierce Biotechnology - 78510	
Roche Complete Protease Inhibitor	Milipore Sigma - 11836153001	
Bradford Protein Assay Kit I	Bio-Rad - 5000001	
Protein Assay Standard I	Bio-Rad - 5000005	
Immun-Star™ WesternC™ Chemiluminescence Kit	Bio-Rad - 170-5070	
Paraformaldehyde, 4% w/v aq. soln.	TFS - 047392.9M	
Antigen Retrieval Buffer	Abcam - ab93678	
ProLong Diamond Anti-Fade Mounting Media with DAPI	TFS - P36962	RRID:AB_2307445
ArC™ Amine Reactive Compensation Bead Kit	Invitrogen - A10346	
UltraComp eBeads	Invitrogen - 01-2222-41	
Antibodies		
Ghost Dye Violet 510	TonboBio - 13-0870-T100	
LIVE/DEAD Fixable Blue Dead Cell Stain Kit, for UV removal	Invitrogen - L23105	
Mouse BD Fc Block (CD16/32)	BD Biosciences - 553142	RRID:AB_394656
Ly-6-G Pacific Blue™	Biolegend - 127611	RRID:AB_1877212
MERTK (Mer) BV 711	Biolegend - 151515	RRID:AB_2876505
CD64 (FcγRI) FITC	Biolegend - 139315	RRID:AB_2566555
F4/80 BUV737	BD Biosciences - 749283	RRID:AB_2873658

Ly-6C Brilliant Violet 605™	Biolegend - 128035	RRID:AB_2562352
CD11c - BV650	Biolegend - 117339	RRID:AB_2562414
CD169 (Siglec-1) APC	Biolegend - 142417	RRID:AB_2565640
CX3CR1 Recombinant PE	Biolegend - 153705	RRID:AB_2734221
CD170 (Siglec-F) - BV421	Biolegend - 155509	RRID:AB_2810421
CD3ε (17A2) violetFluor™ 450	TonboBio - 75-0032	RRID:AB_2621923
CD3ε BUV 737	BD - 612771	RRID:AB_2870100
CD19 Pacific Blue™	Biolegend - 115526	RRID:AB_493341
CD11b Pacific Blue™	Biolegend - 101224	RRID:AB_755986
CD49b (pan-NK cells) Pacific Blue™	Biolegend - 108918	RRID:AB_2265144
F4/80 Pacific Blue™	Biolegend - 123123	RRID:AB_893487
FcεRIα Pacific Blue™	Biolegend - 134313	RRID:AB_10612933
CD90.2 (Thy1.2) FITC	Biolegend - 105306	RRID:AB_313177
CD127 (IL-7Rα) APC	Biolegend - 135012	RRID:AB_1937216
CD45 APC/Cyanine7	Biolegend - 103116	RRID:AB_312981
IL-33R (ST2) (RMST2-33), PE	TFS - 12-9333-82	RRID:AB_2572706
RORγt, PerCP-Cy™5.5	BD Biosciences - 562683	RRID:AB_2737720
GATA3, PE-Cy™7	BD Biosciences - 560405	RRID:AB_1645544
αSMA Monoclonal Anti-mouse, IgG	Abcam - ab7817	RRID:AB_262054
von Willebrand Factor (vWF), Polyclonal Rabbit Anti-mouse, IgG	TFS, Invitrogen - PA5-16634	RRID:AB_10982615
CD169 (Siglec1), Clone 3D6.112, Monoclonal Rabbit Anti-mouse, IgG	Novus Biologicals - NBP3-08989-0.1mg	RRID:AB_2923219
CD11c Monoclonal Armenian Hamster Anti-mouse, IgG	TFS, Invitrogen - 14-0114-82	RRID:AB_467115
F4/80 Purified anti-mouse Antibody	Biolegend - 123102	RRID:AB_893506
Alexa Fluor™ 488 Tyramide SuperBoost™ Kit, goat anti-rabbit IgG	TFS - B40922	
Alexa Fluor™ 647 Tyramide SuperBoost™ Kit, goat anti-mouse	TFS - B40916	

IgG		
Goat anti-Hamster IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 647	TFS, Invitrogen - A-21451	RRID:AB_2535868
Goat anti-Rat IgG - Alexa Fluor™ 555	Invitrogen - A-21434	RRID:AB_2535855
Chicken anti-Mouse IgG - AF™ 647	Invitrogen- A-21463	RRID:AB_2535869
Goat anti-Rabbit IgG - AF™ 555	Invitrogen- A-21428	RRID:AB_2535849
Goat anti-Hamster IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 647	TFS, Invitrogen - A-21451	RRID:AB_2535868
Goat anti-rabbit-HRP antibody	Cell Signaling - 7074	RRID:AB_2099233
Donkey anti-goat-HRP antibody	SouthernBiotech - 6402	RRID:AB_2796335
Primer Sequences		
Bacterial 16S, V4-V5, 450 bp	515F	GTGCCAGCGCCGCGGTAA
	907R	CCGTCAATTCCTTTGAGTTT
Fungal, ITS1, 200-400 bp	ITS5-1737F	GGAAGTAAAAGTCGTAACAAGG
	ITS2-2043R	GCTGCGTTCTTCATCGATGC
Fungal, ITS2, 380 bp	ITS3-2024F	GCATCGATGAAGAACGCAGC
	ITS4-2409R	TCCTCCGCTTATTGATATGCTCCTCCGCTTATTGATATGC
Organisms		
C57Bl/6J mouse	Jackson Laboratory - 000664	RRID:IMSR_JAX:000664
Equipment		
Luminex 200 System	Merck Millipore	RRID:SCR_018025
ProOx 360 controller	BioSpherix	RRID:SCR_021130
gentleMACS	Miltenyi Biotec	RRID:SCR_020271
FACS Symphony A5 SE	BD Biosciences	RRID:SCR_022674

A1R HD confocal microscope	Nikon	RRID:SCR_020317
flexiVent	SCIREQ	RRID:SCR_022673
Vevo 770 High-Resolution *in vivo* Micro-Imaging System	Fujifilm Visual Sonics	RRID:SCR_022921
2100 Bioanalyzer Instrument	Agilent	RRID:SCR_018043
QuantStudio 6	Applied Biosciences	RRID:SCR_020239
Rotor-Gene Q Series	Qiagen	RRID:SCR_018976
NovoSeq 6000 system	Illumina	RRID:SCR_016387
MiSeq platform	Illumina	RRID:SCR_020133
HiSeq 4000 platform	Illumina	RRID:SCR_016386
Software		
NICHD Neonatal Outcomes Estimator	https://neonatal.rti.org	
FACSDiva	BD Biosciences - v9.0	RRID:SCR_001456
FlowJo for Mac	BD Biosciences - v10.8.1	RRID:SCR_008520
Rotor-Gene Q Series Software	Qiagen	RRID:SCR_015740
FIJI	https://fiji.sc	RRID:SCR_002285
Prism 9 for Mac	GraphPad - v9.4.1	RRID:SCR_002798
SPSS Statistics	IBM - 28.0.1	RRID:SCR_002865
R	R project for statistical computing - 4.2.2	RRID:SCR_001905
RStudio	Posit - 2022.12.0	RRID:SCR_000432
phyloseq	1.42.0	RRID:SCR_013080
vegan	2.6-4	RRID:SCR_011950
DESeq2	v1.38.2	RRID:SCR_015687
mia	1.6.0	RRID:SCR_023619
SpiecEasi	1.1.2	RRID:SCR_022712
igraph	1.4.1	RRID:SCR_019225
caret	6.0-94	RRID:SCR_021138
D2 calculator	https://github.com/ jamescarson3/computeD2	

QIIME 2	2022.11	RRID:SCR_021258
DADA2	v1.26	RRID:SCR_023519
SCRuB	0.0.1	RRID:SCR_023518
SILVA	138.1	RRID:SCR_006423
UNITE	9.0	RRID:SCR_006518
NCBI MOLE-BLAST		RRID:SCR_004870
CLC Genomic Workbench	Qiagen - v22	RRID:SCR_011853
ClustVis	https://biit.cs.ut.ee/clustvis/	RRID:SCR_017133
Ingenuity Pathway Analysis	Qiagen (July 2022 edition)	RRID:SCR_008653
Datasets		
RNAseq, Mouse		
RNAseq, hBECs		
16S, human		
16S, mouse		
ITS2, human		
ITS2, mouse		

Resource Availability

Lead Contact: Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Kent Willis (kawillis@uab.edu).

Materials Availability: This study did not generate any unique reagents or transgenic animals.

Data Availability:

- Raw RNA-seq, 16S and ITS2 sequencing data are deposited at the NCBI Sequence Read Archive (Access Numbers: *Pending*).
- Processed data files and original code are available at github.com/WillisLungLab/MiBPD.
- Any additional data required to reanalyze the data presented in this paper are available from the lead contact upon request.

Experimental Model and Subject Details

Study Subjects:
University of Tennessee Health Science Center cohort: This study was approved by the Institutional Review Board of The University of Tennessee under protocol 17-05311-XP and conducted in accordance with the Declaration of Helsinki. We performed a prospective observational cohort study at

the regional NICU at Regional One Health (Memphis, TN, USA) from July 2017-Jan 2020, pre-registered with clinicaltrials.gov (NCT03229967). The Microbial Influence on the Development of Bronchopulmonary Dysplasia (MiBPD) study approached the mothers of all inborn very low birthweight infants (VLBW, < 1500g) without known major congenital anomalies or immunodeficiency for potential consent for enrollment during the first week of life unless they were enrolled in a concurrent trial of probiotics that enrolled < 10 infants during this period. Mothers provided written consent on behalf of themselves and assent for inclusion of their newborn prior to enrollment. Samples were collected weekly for the first month of life. Meconium samples collected within the first 72 h have been described(1). Analysis in this study was constrained to fecal samples collected during the second week of life. Demographic and clinical data were collected to describe the clinical course and to identify the severity of bronchopulmonary dysplasia (BPD) according to the NICHD workshop definition (2). The risk to develop BPD was predicted using the NICHD Neonatal BPD Outcomes Estimator (neonatal.rti.org, 2022 updated formula)(3). Human subjects data is reported in accordance with the extended STROBE guidelines (4).

University of Alabama at Birmingham cohort: This study was approved under protocol IRB-300003994, by the IRB at the University of Alabama at Birmingham (UAB, Birmingham, AL, USA). Samples from this cohort were collected in oxygen-reduced PBS immediately prior to FMT to the mice. The samples were used for the generation of the pseudohumanized mouse model and were collected from Aug 2021-June 2022.

Animals: Animal experiments were performed at UAB under IACUC protocol 22042. 8-week-old female C57Bl/6J mice (Stock Number 000664, RRID:IMSR_JAX:000664) were purchased from the Jackson Laboratory and allowed to acclimate in specific-pathogen free housing with 4 animals per cage, on a 12-hour light/dark cycle, at 20-24C, with *ad libitum* access to irradiated rodent chow (Teklad 7904, Envigo) in the vivarium at the UAB for two weeks. Animal studies were performed in accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health under protocol 22042, approved by the Institutional Animal Care and Use Committee at UAB. Animal experimentation is reported in accordance with the ARRIVE guidelines (5).

Generation of pseudo-humanized mice: Pseudo-humanized mice were generated by fecal microbiota transplant (FMT) from human infant donors with or without severe BPD, following antibiotic preconditioning, as previously described (6–8), generating mice we named PPRD-FMT and BPD-FMT, respectively. Briefly, 9-week-old dams were subjected to antibiotic administration for two weeks in their drinking water which they were allowed to access *ad libitum*. The first week consisted of systemically absorbable antibiotics: penicillin VK, cefoperazone sodium, and clindamycin sodium (all at a concentration of 1mg/mL). This was followed by a bowel cleansing regimen by oral-gastric gavage with 200 uL of polyethylene glycol 3350 at 425 g/L. This was repeated twice. Mice were placed on the second week of antibiotic therapy in the drinking water which consisted of the non-absorbable antibiotics ertapenem sodium, vancomycin sodium, and neomycin sodium (all at a concentration of 1mg/mL). Cage bottoms were changed daily, and after each gavage of polyethylene glycol, as mice are coprophagous and could reinoculated themselves with stool. After two weeks of antibiotic treatment, female mice were given stool from neonatal human donors with or without BPD. The stool was preserved in oxygen-reduced sterile PBS. We then performed FMT. Mice in each group were gavaged twice with 200 uL of sample from the same human donor. Stool samples were collected after antibiotic therapy, after FMT engraftment, or at the time of harvest. After engraftment, dams inoculated with the same donors FMT were bred simultaneously to generate neonatal mice for further experimentation. On day 2 after birth the pups were inoculated with the same donor FMT to create BPD-FMT and PPRD-FMT mice, respectively.

Each experiment was conducted with at least two litters arising from the same human donor, and a minimum of three human donors in each experimental group (6 litters=3 donors/group).

Neonatal hyperoxia exposure model: We utilized our established, standardized hyperoxia-exposure model of BPD (9, 10). We time-mated the mice to produce multiple simultaneous birth cohorts with the same perinatal exposure. On the day of birth, pups from two consecutively born litters with the same perinatal exposure were pooled and the pups were re-distributed evenly between the two dams. Litter sizes for all experiments were adjusted to 5–8 pups per treatment group to minimize nutritional effects on lung development. One pooled litter was randomly assigned to hyperoxia [fraction of inspired oxygen (FiO_2) 0.85]) and the other to air (FiO_2 0.21). Pups were exposed continuously for 11 days, starting on the third day after birth (P3-14). Oxygen concentrations were maintained using a ProOx monitor (Biospherix). Dams were rotated daily between the two pooled litters to limit the complications of hyperoxia. Growth was monitored daily for all pups throughout the experiment.

Method Details

Data collection: Clinical data were collected from the medical record that described clinical illness severity, infant characteristics, maternal and infant demographics, the extent of prematurity, respiratory parameters, and known microbiome modulators when the infant reached 36 weeks corrected gestational age or was discharged. BPD was adjudicated at 36 weeks' corrected gestational age using the NICHD workshop definition (2). We quantified clinical illness severity with the Critical Risk Index for Babies II (CRIB II)(11). The risk to develop BPD was predicted using the NICHD Neonatal BPD Outcomes Estimator (neonatal.rti.org, 2022 updated formula)(3).

RNA-seq: Bulk RNA-seq was performed on isolated tissue from the left middle lobe using established methods (12). Briefly, 25 mg of tissue was collected in RLT buffer (Qiagen) with 1% (v/v) 2-Mercaptoethanol (BME, Sigma Aldrich), homogenized using a QIAshredder (Qiagen), and isolated with a RNeasy Kit (Qiagen) with on column removal of DNase. DNA purity and concentration were validated with the Bioanalyzer High Sensitivity DNA Analysis chip (Agilent Technologies) at the Heflin Genomics Core Facility. Samples were sent to Novogene for RNA library preparation and transcriptome sequencing on the HiSeq 2500 platform (Illumina) using a 150-nucleotide paired-end read length (12).

FlexiVent pulmonary function testing: The P14 mice used for this analysis were not used for other analyses. Forced ossilometry was performed as described (13, 14). Briefly, we performed a tracheostomy on the sedated mice with the appropriate cannula and secured it with two 3-0 sutures. We performed the neonatal pulmonary function assessment program on the flexiVent system (SCIREQ). We calibrated the system with the tracheal cannula prior to each mouse.

Histology and immunohistochemistry: We gravity inflated the right lung of each mouse to 20 cm H_2O with 4% formalin (Thermo Fisher Scientific) and were placed into 100% ethanol within 16 h. After paraffin embedding, we cut 4-5 m slides at 3 random levels through the block that were stained with hematoxylin and eosin for morphometry. At each level, three slides were made from adjacent sections for immunohistochemistry. These slides were deparaffinized and rehydrated with graduated concentrations of ethanol. We performed antigen retrieval with heated 10 µM Sodium Citrate Buffer with 20% Tween. Slides were, fenestrated, blocked in 1% BSA and anti-CD16/32 (BD Biosciences) for 30 min, then incubated overnight with either anti-vWF and or anti-αSMA at 4C in incubation buffer (1% BSA, 1% normal donkey serum, 0.3% Triton X-100, and 0.01% sodium azide in PBS). The next day they were washed in PBS, blocked again, counterstained with the appropriate secondary antibodies, washed, and

finally counterstained with ProLong Diamond Anti-Fade Mounting Media with DAPI (Thermofisher Scientific). Images were captured using an A1R HD inverted confocal microscope (Nikon).

Bacterial and fungal microbiota analysis: First non-transitional stool sample, collected during the second week of life were used to assess the human gut microbiota. In mice, liquid nitrogen snap-frozen homogenized whole left lung or a 1 cm section of terminal ileum were used because of the advantages in quantifying the adherent mucosal microbiota (15). We used ZymoBIOMICS whole organism and DNA microbial community standards (Zymo Research) as well as sterile-filtered PBS to create pairs of positive and negative controls appropriate for each step of the sample collection and isolation procedure. Microbial DNA was extracted using the ZymoBIOMICS DNA Miniprep Kit (Zymo Research) with the addition of a lyticase (200 Units per 1 mL sample volume) predigestion with 10 minute bead bashing on a Fisher Vortex Genie 2 (Thermo Fisher Scientific) at RT, then 37C with gentle agitation for 20 min, as previously validated to facilitate extraction of fungal DNA (16). Extracted DNA was sequenced using the MiSeq platform (Illumina) at Novogene (human samples) or Argonne National Laboratory (mouse samples) using 16S V3-4 (bacterial) rRNA with primers 515F-806R and the internal transcribed spacer (ITS) 2 (fungal) rRNA gene primers. Primer design was chosen based on prior published optimization for mycobiome amplification in respiratory samples (17, 18). Quantitative RT-PCR of the ITS1, ITS2, and 16S genes was performed as described (19), using TaqMan probes (Thermo Fisher Scientific) validated with ZymoBIOMICS controls (Zymo Research) and additional human and mouse samples. QC was performed at the UAB Genomics Core.

Bioinformatics: Microbiomics - We used QIIME 2 (20) to import and trim ITS2 and 16S rRNA gene amplicon sequences. DADA2 (21) was used to denoise and cluster amplicon sequence variants (ASVs). ASV taxonomy was assigned with the RDP naive Bayesian Classifier using the SILVA 138.1 and UNITE 9.0 databases, for bacteria and fungi, respectively. All ASVs aligning to the division Chytridiomycota or unassigned at the phylum level were then removed because confirmatory analysis with BLAST revealed these sequences to be either host mitochondrial or plant material contaminates. Contamination source modeling was then performed using the R package *SCRuB* (22) at the default settings. Decontaminated ASV tables were then exported into R where they were normalized, transformed, and analyzed with the following packages: *vegan*, *phyloseq*, *DESeq2*, *SpiecEasi* (Sparse Inverse Covariance Estimation for Ecological Association Inference), *tidyverse, igraph, mia,* and *caret*. Alpha diversity was quantified with the Shannon and Chao1 Indices. We displayed beta diversity with both principal components analysis of Euclidean distances and principal coordinates analysis of Bray-Curtis dissimilarity matrices. Significance testing was performed using PERMANOVA and evaluated confounding by differential dispersion using PERMDISP. Feature selection was performed using DESeq2 and Random Forest using *ranger* and *caret*. Unsupervised clustering was performed using Dirichlet-Multinomial Mixture modeling using the packages *mia* and *DirichletMultinomial*. Feature selection was performed using the *fitted* function within *DirichletMultinomial*. Mapping of clinical data to DMM clusters was performed using Canonical Correspondence Analysis (CCA) using the *vegan* package.

RNAseq - Raw sequencing data was aligned, processed, normalized and differential gene expression was calculated using the CLC Genomic Workbench (Qiagen). Genes with adjusted p-value (q) < 0.05 and log(FoldChange) > 0.301 were considered differentially expressed. Heatmaps and principal components analyses were graphed using ClustVis. Differentially expressed genes were input into Ingenuity IPA (Qiagen) to identify significantly regulated networks and pathways.

Machine learning models: Gradient boosting algorithms utilize a group of shallow trees and assemble them into a single prediction score to improve their predictive properties (23). We developed a Light

Gradient Boosted Machine (LightGBM) (24) model to predict BPD using both clinical demographics at birth and the corresponding microbiome data. Clinical parameters were curated from those previously validated in the Neonatal BPD Outcome Predictor (3), and included gestational age at birth, maternal-self reported race, birth weight, temperature on admission to the NICU, and mode of ventilation. Since the sample size of the dataset is modest in comparison to the number of unique variables in the microbiota data, we only utilized the kingdom and phylum-level microbiota data to maintain generalizability and reduce the risk of over-fitting. To construct the model, we split the dataset into two sets: 80% into model building discovery set and 20% as a holdout validation set. We used 5-fold cross validation in the discovery set to tune the hyperparameters of the algorithm. We then assessed the accuracy of the training phase using the validation set. A Bayesian algorithm (25) was chosen for hyper-parameter tuning. To tune the model, we adjusted the number of trees, feature fraction, learning rate, and the maximum depth of the trees. In addition, the training phase was stopped if there was not a consecutive improvement in the validation set. Finally, the LightGBM model from the cross-validation was performed on the holdout set and we calculated the resulting AUC of the ensemble model.

We also implemented default variable importance analysis of LightGBM algorithm to evaluate the contribution of the variables into models, in which their contribution was weighted over folds according to the models' performance on the corresponding validation set. We also evaluated the directional effect of the variables on the BPD risk prediction score within a 95% confidence interval (inverse or direct relationship) by systematically increasing by one unit for categorical variables and one standard deviation for continuous variables.

Statistical analysis: Descriptive statistical analysis of the human data was performed in SPSS Statistics (IBM). Murine data was analyzed primarily in R and Prism 9 (GraphPad, Inc.).

Supplemental Information

Extended Data:

Dirichlet-Multinomial Models of PPRD and BPD infants:

Laplace analyses performed independently within the BPD or PPRD cohorts identified 3 clusters as the best fit for each of these mycobiome subsets (Figure E7F,I). In our three-component Dirichlet-Multinomial Mixture (DMM) model of PPRD infants (Figure 4G-I), membership in Component 1 is driven by *Candida* and *Saccharomyces* ASVs as above. Component 2 is driven by the abundance of *Saccharomyces* and *Aureobasidium* ASVs, while Component 3 is driven by a different *Candida* ASV and an *Aureobasidium* ASV (Figure E7G). Component 1 is associated with increasing gestational age and birthweight and high fungal evenness by CCA. In contrast, Component 2 is associated with lower gestational age and a greater degree of clinical illness (Figure E5H).

Finally, we developed a three-component DMM model of BPD infants (Figure 4I-K). Membership in Component 1 was driven by *Candida* and *Aureobasidium*, and to a lesser degree *Saccharomyces* ASVs. Components 2 and 3 were driven by the presence of an *Aureobasidium* ASV, like the BPD associated components we identified in the five-component DMM model of the whole mycobiome (Figure E7J). CCA identified that Component 1 is aligned with lower gestational age and birthweight and a greater degree of clinical illness, while Component 2 has the opposite drivers (Figure E7K). From these unsupervised analyses, we conclude that in addition to the associations with decreased abundance of the genera

Candida and *Saccharomyces* we identified in our supervised analyses, these analyses also identified a strong link with the abundance of *Aureobasidium*.

Multikingdom DMM models:
DMM modeling using the multikingdom microbiome data is shown in Figures E8 and E9. Of note, Laplace clustering for the whole microbiome indicated optimal clustering with 4 components (Figure E9A), while clustering for the PPRD and BPD infants was optimal with 3 and 2 clusters, respectively (Figures E9F, I). We then performed variable importance analysis and identified only bacterial ASVs with significant contribution to component formation (Figure 9). Clustering in the whole microbiome model is driven by *Enterococcus*, *Enterobacter*, *Cellulosimicrobium*, and multiple *Klebsiella* ASVs. While clustering in PPRD infants was driven by *Enterococcus, Enterobacter* and *Cellulosimicrobium*, while the BPD infant clusters were driven by *Enterococcus* and *Enterobacter*. CCA analysis did not show any clear associations between BPD or PPRD and the composition of the microbiome (Figure E9C, F, I). Together, the DMM clustering between the fungal and multikingdom microbiomes models is not indicative of any significant overlap.

Tables:

Table E1. Cohort characteristics and demographics

	PPRD	**BPD**	***p***
n	64	38	
Gestational age (weeks)	29.0 (27.6 - 30.8)	25.7 (25 - 27.8)	< 0.001
Birthweight (g)	1145 (840 - 1330)	775 (670 - 860)	< 0.001
Female sex assigned at birth	37 (57.8%)	14 (36.8%)	0.041
Antenatal Steroids (No)	8 (17%)	4 (13.3%)	0.33
Surfactant (None)	27 (43.5%)	1 (2.6%)	<0.001
Apgar score (1 min)	6 (3 - 7)	4 (2 - 5)	< 0.001
Apgar score (5 min)	7 (6 - 8)	6 (5 - 7)	< 0.001
Small for gestational age*	6 (9.4%)	8 (21.1%)	0.098
Caesarean section	28 (59.6%)	20 (66.7%)	0.531
Prolonged rupture of membranes‡	7 (15%)	5 (16.7%)	0.834
Critical Risk Index for Babies II (CRIB II) score	7 (4 - 9)	12 (9 - 14)	< 0.001
Antenatal antibiotic exposure	22 (46.8%)	15 (50%)	0.785

Duration of antenatal antibiotic exposure (d)	1 (0 - 1)	0 (0 - 1)	0.542
Postnatal antibiotic exposure	55 (85.9%)	38 (100%)	0.015
Duration of postnatal antibiotic exposure (d)	2 (1 - 2)	2 (1 - 4)	0.016
Respiratory Support on Day 3			
Mechanical ventilation	20 (37%)	33 (89.2%)	< 0.001
Fraction of inspired oxygen	0.21 (0.21 - 0.21)	0.25 (0.21 - 0.34)	< 0.001
Predicted risk of severe BPD†	4.40% (1.68 - 10.14%)	10.70% (8.73 - 14.49%)	< 0.001
Predicted risk of death†	2.84% (1.36 - 6.30%)	10.90% (7.69 - 16.00%)	< 0.001
Respiratory Support on Day 7			
Mechanical ventilation	15 (28.3%)	30 (83.4%)	< 0.001
Fraction of inspired oxygen	0.21 (0.21 - 0.23)	0.29 (0.24 - 0.30)	< 0.001
Predicted risk of severe BPD†	4.35% (1.54 - 9.40%)	11.09% (8.22 - 21.50%)	< 0.001
Predicted risk of death†	3.35% (1.1 - 5.52%)	10.21% (7.01 - 13.2%)	< 0.001
Endotracheal intubation	41 (66.1%)	37 (100%)	< 0.001
Ventilation duration (d)	1 (0 - 4)	34 (22 - 50)	< 0.001
Non-invasive ventilation duration (d)	5 (1 - 17)	14 (1 - 38)	0.053
Oxygen exposure duration (d)	13 (4 - 41)	92 (30 - 119)	< 0.001
Supplemental oxygen on day 28	23 (37.1%)	33 (91.7%)	<0.001
Death	0 (0%)	6 (15.8%)	< 0.001

Data are presented as median (interquartile range) or n (%) with significance testing by Mann-Whitney U or Pearson chi-squared test, as appropriate.

BPD, Bronchopulmonary dysplasia; PPRD, Post-prematurity respiratory disease. *< 3rd percentile, ‡ > 18 h, †NICHD Neonatal Outcomes Estimator

Table E2. PERMANOVA analysis of potential fungal covariates.

ITS Covariates	Sum of Squares	R^2	F	*p*	*q* (FDR)
Oxygen (days)	1.1213	0.07027	3.3253	0.00079992	0.0159984
Death	0.3704	0.02155	1.035	0.38806119	0.6874313
Sex	0.6877	0.04001	1.9587	0.03159684	0.1053228
Gestational Age	0.5335	0.03104	1.5055	0.11068893	0.2213779
Birthweight	0.7706	0.04483	2.2061	0.01649835	0.1053228
CRIB-II	0.6563	0.03818	1.8659	0.04059594	0.1159884
Intubation	0.5435	0.03208	1.5248	0.10318968	0.2213779
Vent (days)	0.6854	0.04295	1.9746	0.02909709	0.1053228
Maternal UTI	0.2663	0.02064	0.822	0.60543946	0.8072526
Preeclampsia	0.7999	0.062	2.5779	0.00949905	0.0949905
Cesarean Section	0.242	0.01876	0.7455	0.66913309	0.8176829
Multiples	0.1859	0.01441	0.5701	0.84761524	0.8922266
IUGR	0.3229	0.02503	1.0012	0.41245875	0.6874313
Feeding	0.9411	0.05556	0.8627	0.6950305	0.8176829
Prenatal Abx	3.2965	0.25552	1.3729	0.03159684	0.1053228
Postnatal Abx	3.3554	0.19521	1.2128	0.0849915	0.2124788
NEC	0.2633	0.01554	0.7263	0.75822418	0.8424713
Culture+ Sepsis	0.3378	0.01994	0.936	0.48175182	0.7411567
Culture- Sepsis	0.3211	0.01895	0.8888	0.54034597	0.7719228

Table E3. PERMANOVA analysis of potential fungal covariates.

Multikingdom Covariates	Df	Sum of Sqs	R^2	F	*p*	*q* (FDR)
Death	1	0.154	0.0045	0.4286	0.9962	0.9962
Sex	1	0.206	0.006	0.5761	0.9531	0.9962

Gestational age	1	0.606	0.0177	1.7119	0.0281	0.0923
Birthweight	1	0.773	0.0226	2.1976	0.004	0.0299
CRIB-II	1	0.656	0.0192	1.8582	0.0165	0.0632
Feeding	4	1.404	0.0429	0.9976	0.4667	0.6941
Oxygen (days)	1	0.5229	0.0167	1.4963	0.0735	0.1458
Vent (days	1	0.3389	0.0108	0.9642	0.4829	0.6941
Multiples	1	0.3002	0.0123	0.8999	0.5491	0.7017
IUGR	1	0.3327	0.0137	0.9987	0.4226	0.6941
Maternal UTI	1	0.2843	0.0117	0.8516	0.6217	0.715
Preeclampsia	1	0.437	0.018	1.3175	0.1715	0.3034
Cesarean section	1	0.1907	0.0078	0.5689	0.928	0.9962
Prenatal Abx	10	4.747	0.1995	1.5448	1E-04	0.0023
Postnatal Abx	13	5.535	0.1618	1.2324	0.0065	0.0299
NEC	1	0.309	0.0095	0.8783	0.5954	0.715
Intubation	1	0.521	0.0159	1.4889	0.0761	0.1458
Culture+ sepsis	1	0.58	0.0177	1.6594	0.0386	0.111
Culture- sepsis	1	0.738	0.0226	2.1242	0.0053	0.0299
Fungal abundance	1	0.951	0.0278	2.7151	0.0005	0.0057

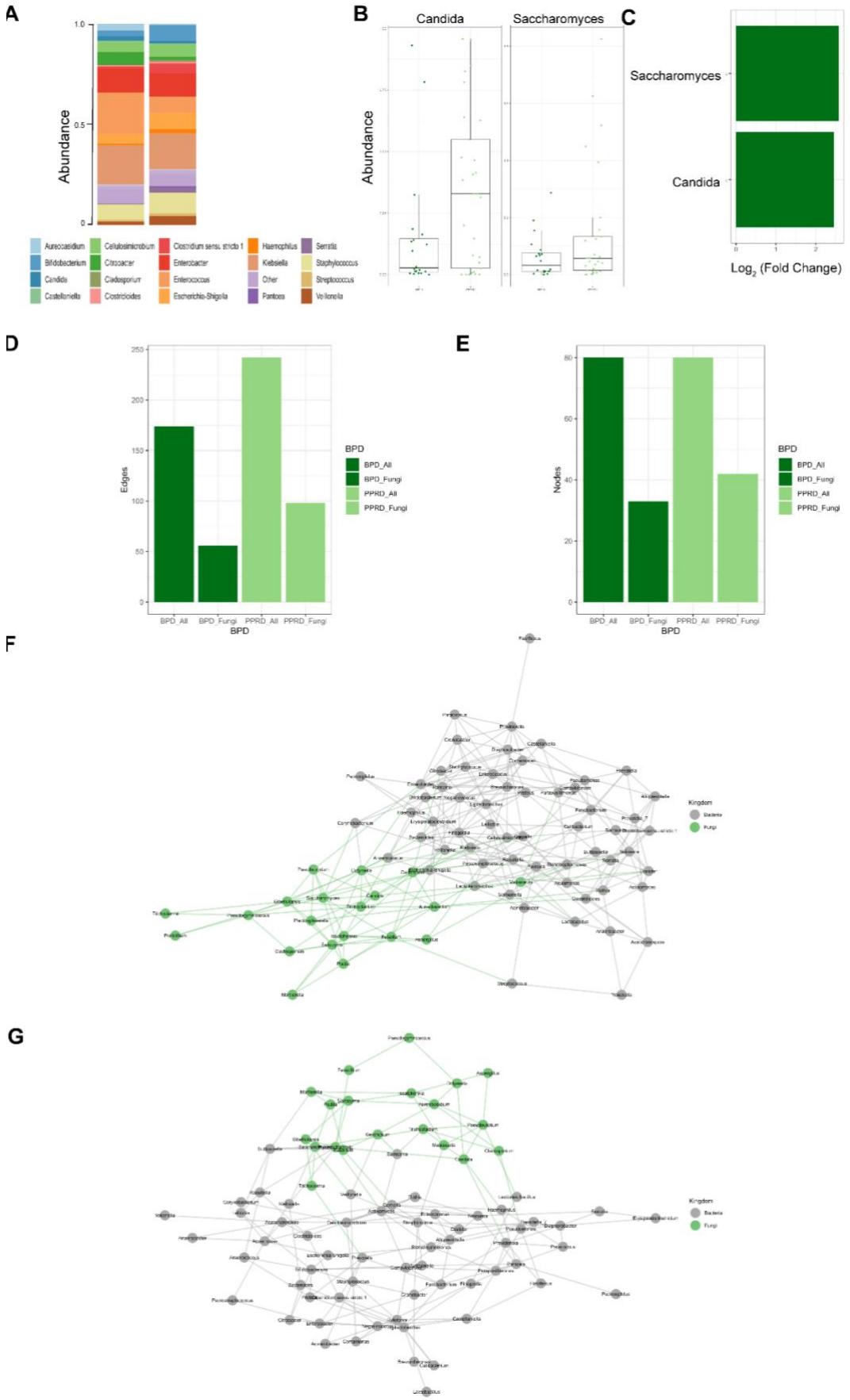

Figure E1. The BPD-associated mycobiome is unique. (A) Relative abundance of the top 20 most abundant microbes in the multikingdom microbiome. (B, C) relative abundance of key fungi by DESeq2.

(D) Number of edges in each SPIEC-EASI network. (E) Number of nodes in each SPIEC-EASI network. (F) Complete PPRD network with all genera labeled. (G) Complete BPD network with all genera labeled.

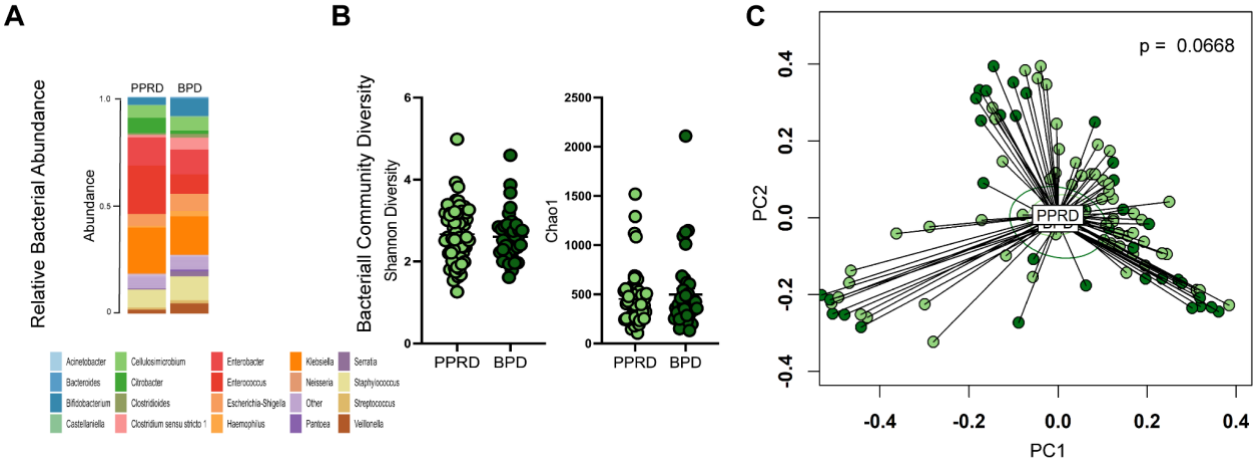

Figure E2. Overall composition of the bacterial microbiome. (A) Relative abundance of the top 20 bacterial genera. (B) Alpha diversity. (C) Beta diversity (PCoA of Bray-Curtis dissimilarity, p = 0.0668, PERMANOVA, p = 0.2412, PERMDISP). Ellipses indicate the 95% confidence interval.

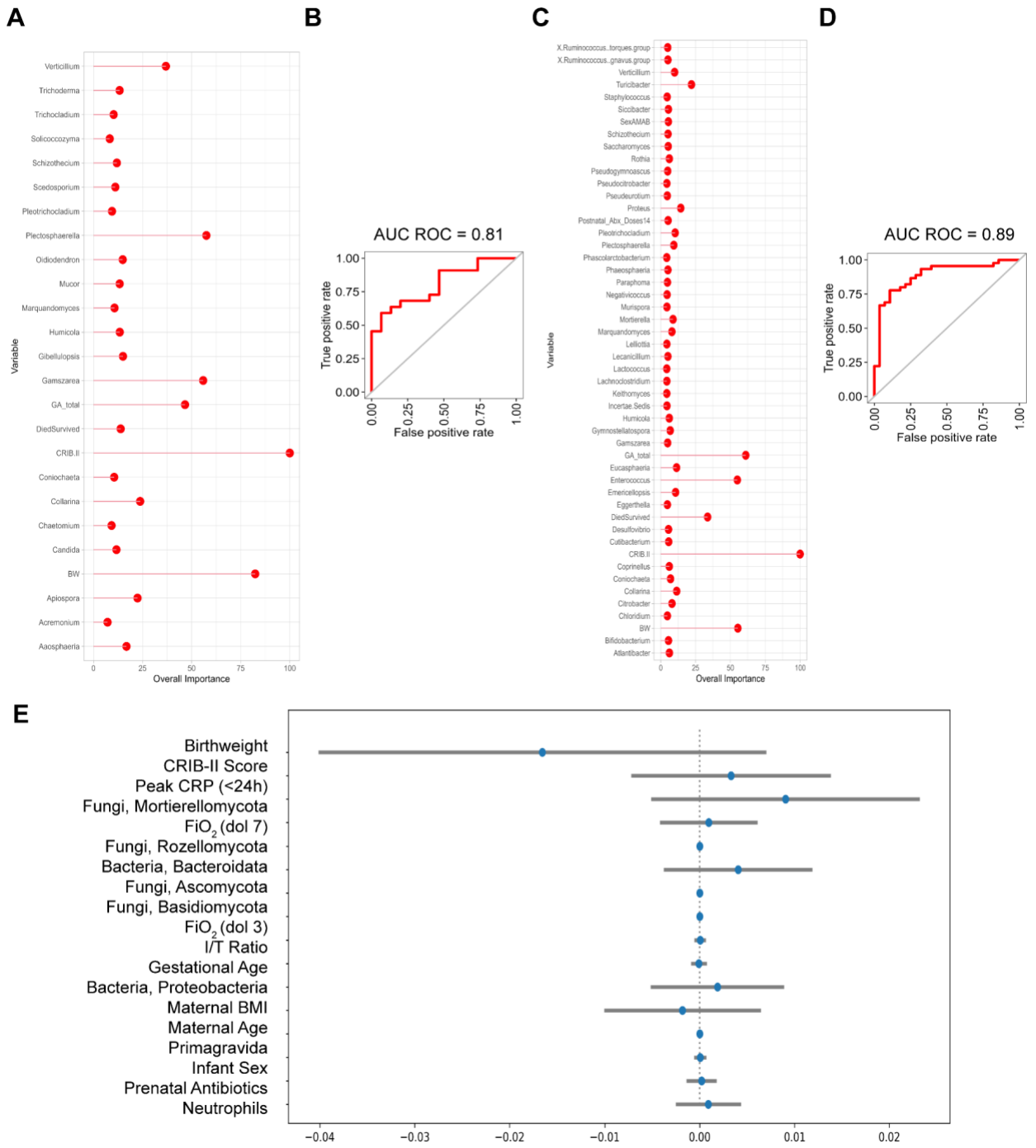

Figure E3. BPD status is predictable using machine-learning models of the fungal microbiota prior to the adjudication of the clinical diagnosis at 36 weeks corrected gestational age. (A-B) Exploratory random forest machine learning model using genera level mycobiome and clinical data (Should be considered exploratory due to high risk of overfitting). (C-D) Exploratory random forest machine learning model using genera level multikingdom microbiome and clinical data. (E) Directional variable importance analysis for the light gradient boosted machine learning model.

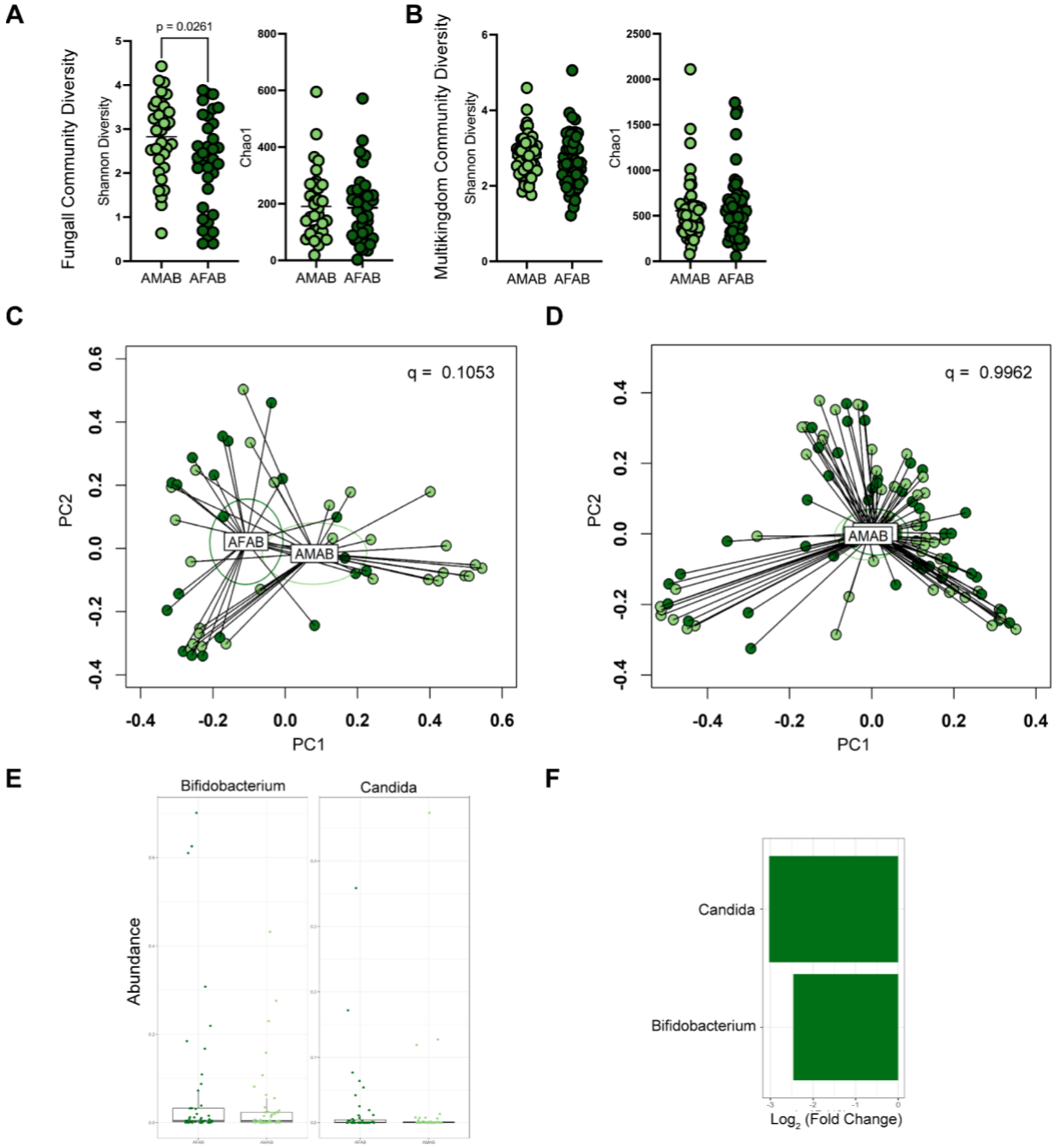

Figure E4. Mycobiome feature differences by sex assigned at birth. (A) Fungal alpha diversity. (B) Multikingdom alpha diversity. Shannon diversity was decreased in infants assigned female at birth (AFAB) as compared to infants assigned male at birth (AMAB, p = 0.0261, Mann-Whitney) (C) Principal coordinates analysis of Bray-Curtis dissimilarity of fungal beta diversity (PERMANOVA, q = 0.1053). (D) Principal coordinates analysis of Bray-Curtis dissimilarity of multikingdom beta diversity (PERMANOVA, q = 0.9962). (E) Relative abundance of *Bifidobacterium* and *Candida*. (F) DESeq2 selected differential features of the multikingdom microbiome.

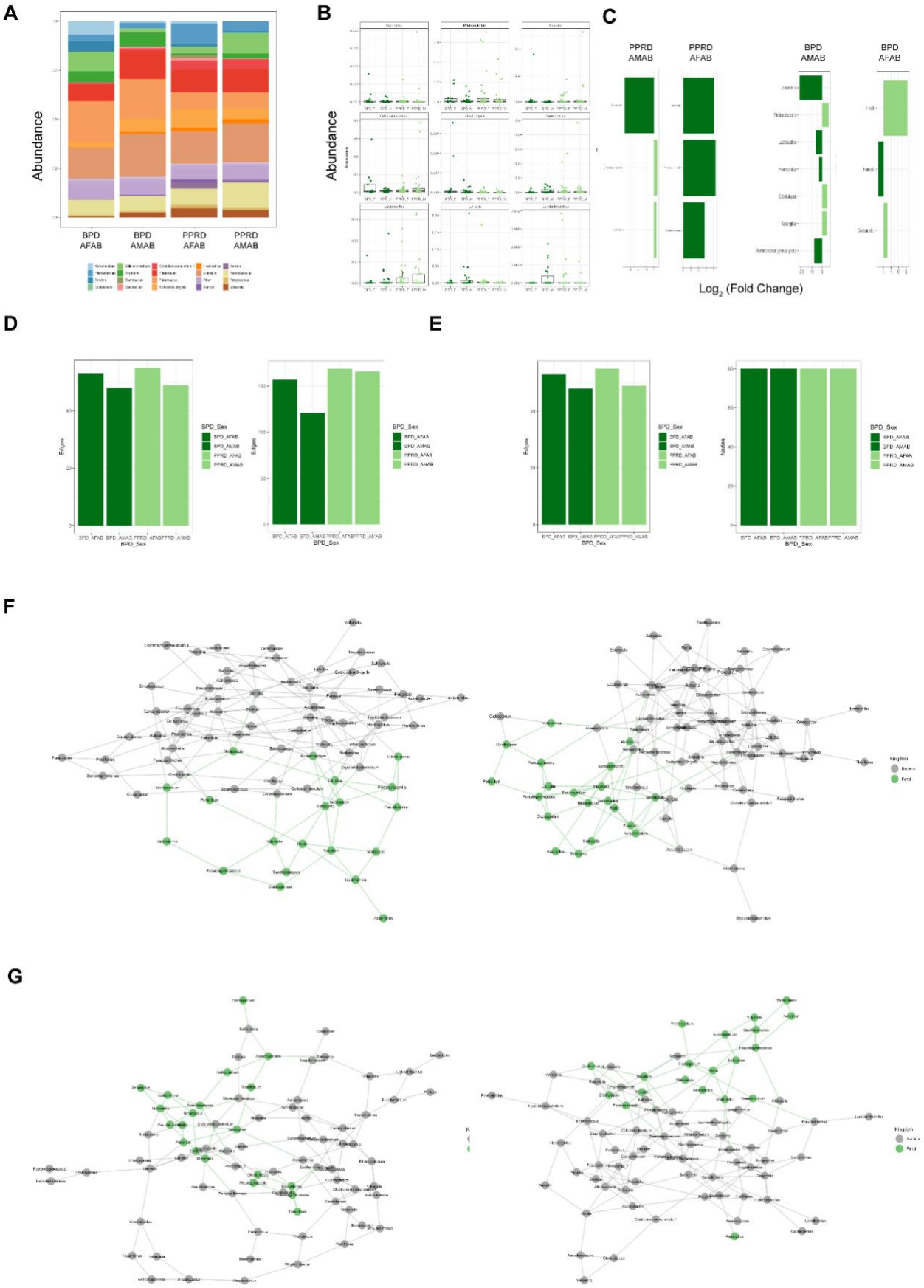

Figure E5. Microbiota alterations by disease status and sex assigned at birth. (A) Multikingdom relative abundance of the top 20 most abundant species. (B-C) DESeq2 selected features. (D) SPIEC-EASI

network edges. (E) SPIEC-EASI network nodes. (F) SPIEC-EASI network for PPRD infants assigned male at birth (AMAB) or assigned female at birth (AFAB). (G) SPIEC-EASI network for BPD infants.

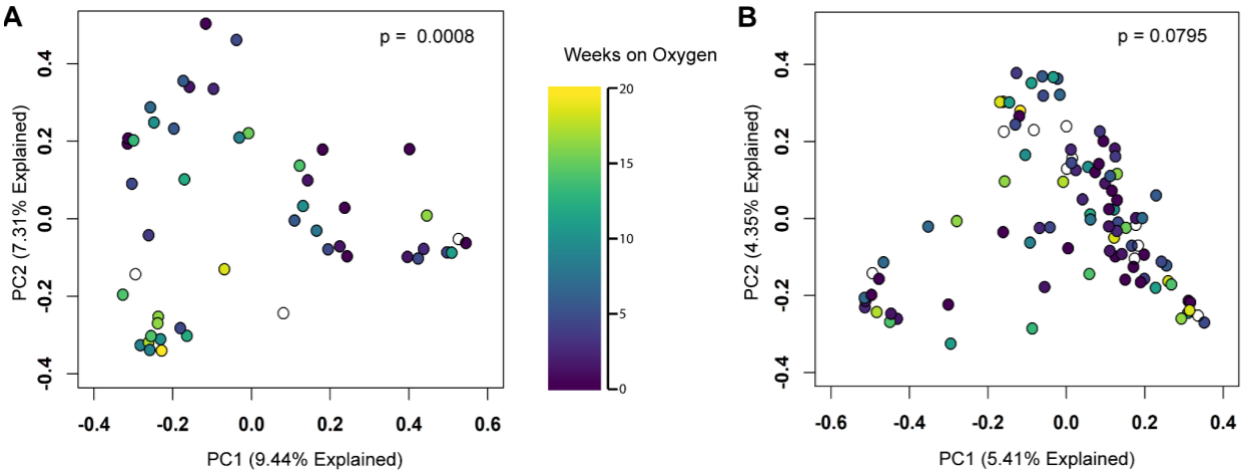

Figure E6. Oxygen concentration influences the composition of the intestinal microbiome. (A) PCoA of Bray-Curtis dissimilarity showing that the composition of the intestinal mycobiome is significantly associated with total weeks of exposure to oxygen (p = 0.0008, PERMANOVA). (B) PCoA of Bray-Curtis dissimilarity showing that the composition the intestinal microbiome is not associated with total weeks of exposure to oxygen (p = 0.0795, PERMANOVA).

.

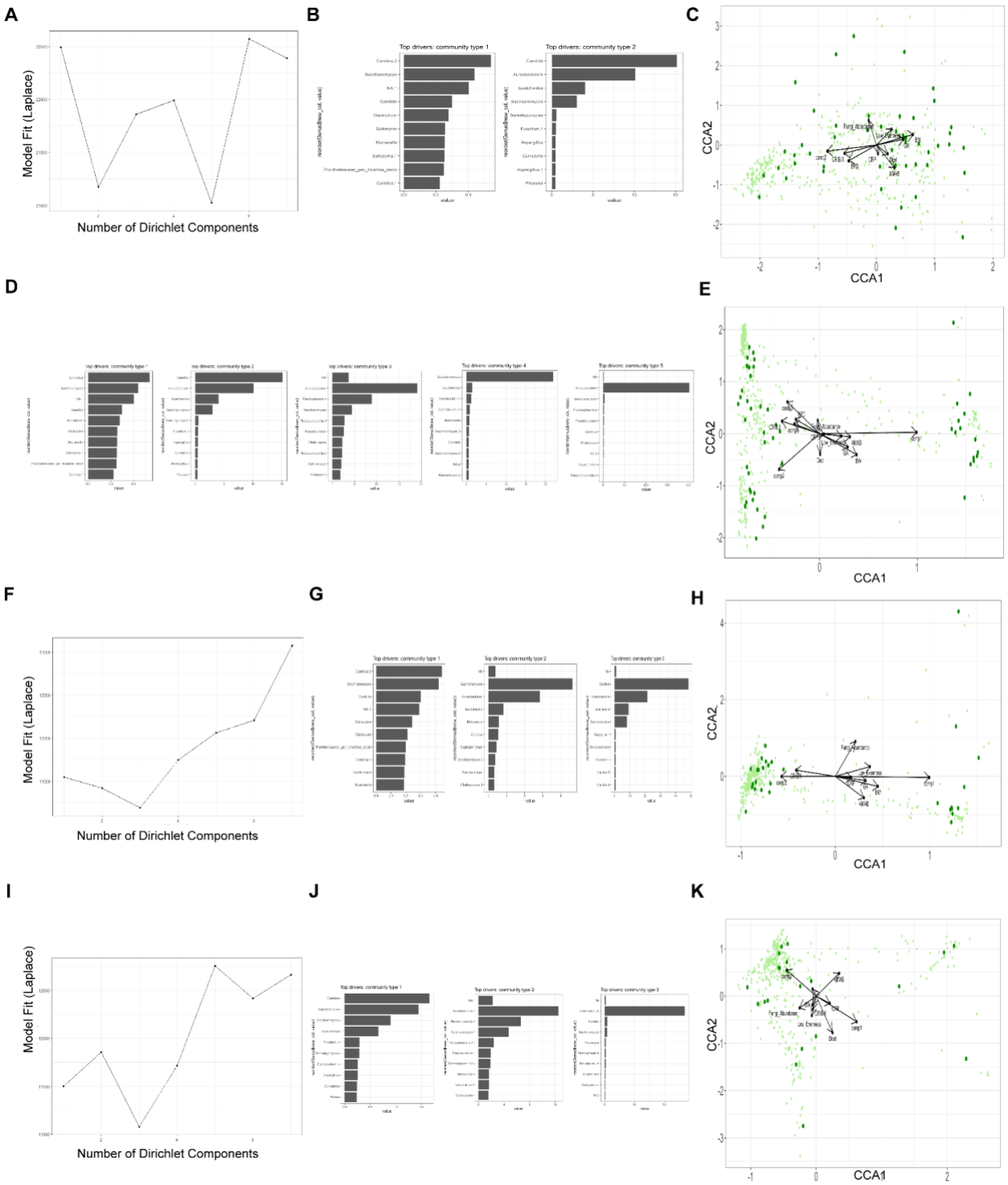

Figure E7. Dirichlet-Multinomial Mixture (DMM) modeling of the mycobiome. (A) Laplace analysis of the mycobiome indicates optimal model fit points with 2 or 5 components. (B) Variable importance analysis of the two-component DMM model. (C) Canonical Correspondence Analysis (CCA) for the two-component DMM model. (D) Variable importance analysis of the five-component DMM model. (E) CCA for the two-component DMM model. (F) Laplace analysis of mycobiome samples from PPRD infants showing an optimal fit point at three components. (G) Variable importance analysis of the PPRD DMM model. (H) CCA for the PPRD DMM model. (I) Laplace analysis of mycobiome samples from BPD infants showing an optimal fit point at three components. (J) Variable importance analysis of the BPD DMM model. (K) CCA for the BPD DMM model.

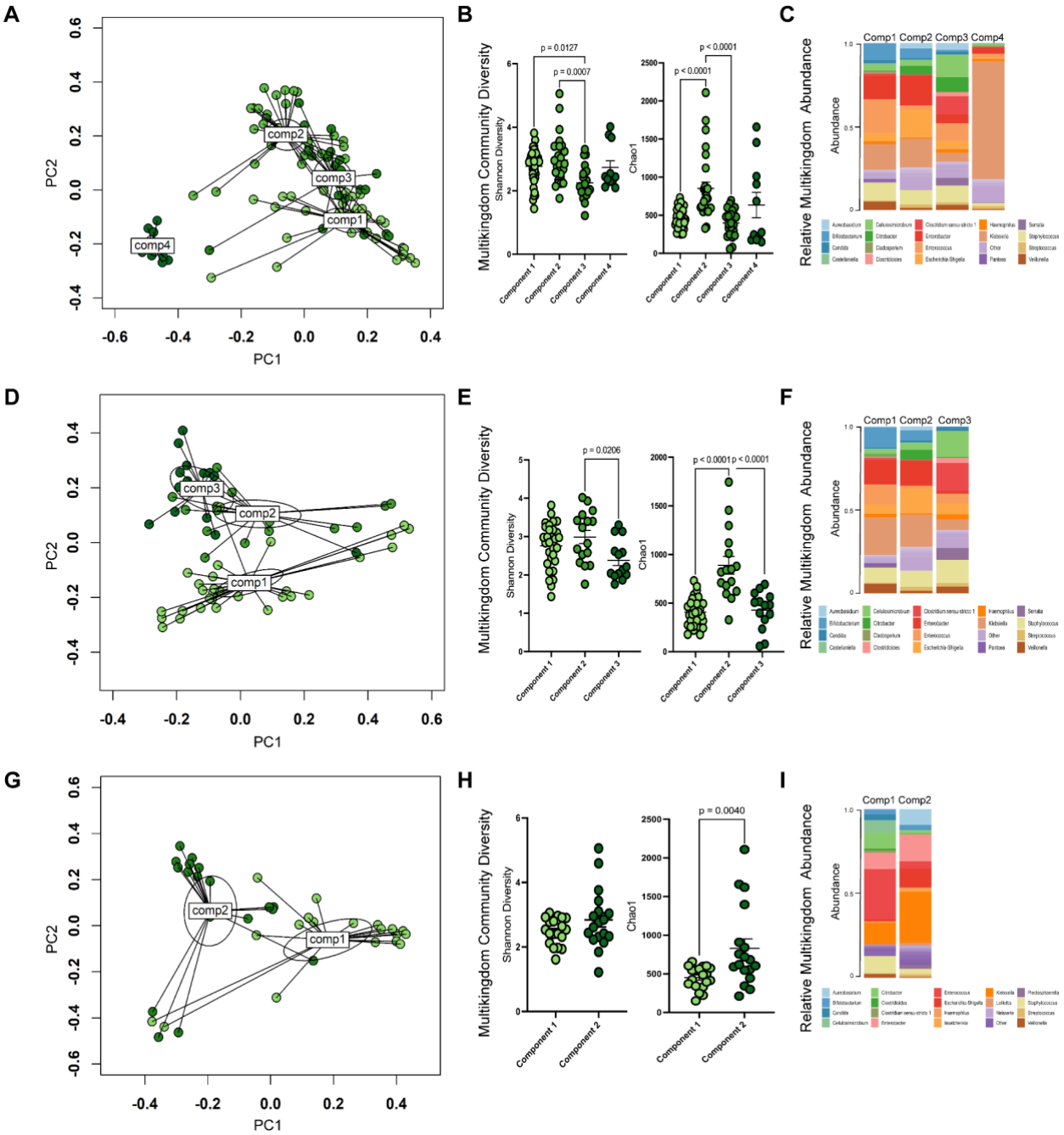

Figure E8. Dirichlet-Multinomial Mixture (DMM) modeling of the multikingdom microbiome. (A) Principal Coordinates Analysis (PCoA) of the multikingdom microbiome DMM model. (B) Alpha diversity of the multikingdom microbiome DMM model. (C) Relative abundance of the multikingdom microbiome in the DMM components. (D) PCoA of the DMM model of the multikingdom microbiome of infants with PPRD. (E) Alpha diversity of the multikingdom microbiome. (F) Relative abundance of the multikingdom microbiome in the PPRD DMM model. (G) PCoA of the DMM model of the multikingdom microbiome of infants with BPD. (H) Alpha diversity of the multikingdom microbiome. (I) Relative abundance of the multikingdom microbiome in the BPD DMM model. Ellipses indicated the 95% confidence interval.

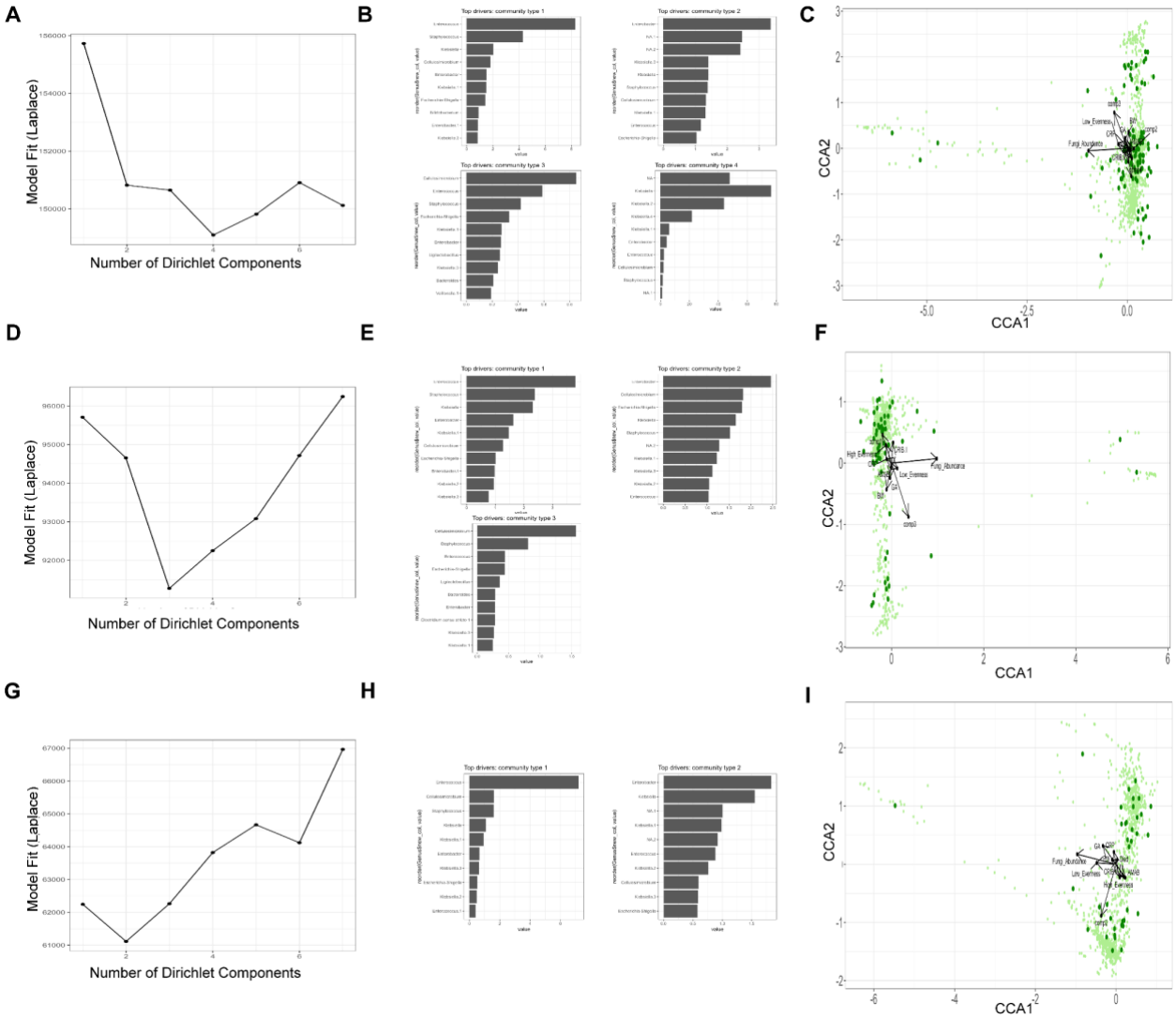

Figure E9. DMM modeling of the multikingdom microbiome. (A) Laplace analysis of the mycobiome indicates an optimal model fit point with four components. (B) Variable importance analysis of the multikingdom microbiome DMM model. (C) Canonical Correspondence Analysis (CCA) for the multikingdom microbiome DMM model. (D) Laplace analysis of multikingdom microbiome samples from PPRD infants showing an optimal fit point at three components. (E) Variable importance analysis of the PPRD DMM model. (F) CCA for the PPRD DMM model. (G) Laplace analysis of multikingdom microbiome samples from BPD infants showing an optimal fit point at two components. (H) Variable importance analysis of the BPD DMM model. (I) CCA for the BPD DMM model.

References:

1. Willis KA, Purvis JH, Myers ED, Aziz MM, Karabayir I, Gomes CK, *et al.* Fungi form interkingdom microbial communities in the primordial human gut that develop with gestational age. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2019;33:12825–12837.
2. Higgins RD, Jobe AH, Koso-Thomas M, Bancalari E, Viscardi RM, Hartert TV, et al. Bronchopulmonary Dysplasia: Executive Summary of a Workshop. *J Pediatrics* 2018;197:300–308.
3. Laughon MM, Langer JC, Bose CL, Smith PB, Ambalavanan N, Kennedy KA, et al. Prediction of Bronchopulmonary Dysplasia by Postnatal Age in Extremely Premature Infants. *Am J Resp Crit Care* 2011;183:1715–1722.
4. Elm E von, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. *Plos Med* 2007;4:e296.
5. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. 2010;
6. Staley C, Kaiser T, Beura LK, Hamilton MJ, Weingarden AR, Bobr A, et al. Stable engraftment of human microbiota into mice with a single oral gavage following antibiotic conditioning. *Microbiome* 2017;5:87.
7. Chen X, Li P, Liu M, Zheng H, He Y, Chen M-X, et al. *Gut* dysbiosis induces the development of pre-eclampsia through bacterial translocation. Gut 2020;gutjnl-2019-319101.doi:10.1136/gutjnl-2019-319101.
8. Wrzosek L, Ciocan D, Borentain P, Spatz M, Puchois V, Hugot C, et al. Transplantation of human microbiota into conventional mice durably reshapes the gut microbiota. *Scientific reports* 2018;8:6854.
9. Willis KA, Siefker DT, Aziz MM, White CT, Mussarat N, Gomes CK, et al. Perinatal maternal antibiotic exposure augments lung injury in offspring in experimental bronchopulmonary dysplasia. *American Journal of Physiology - Lung Cellular and Molecular Physiology* 2019;3:21.
10. Nardiello C, Mižíková I, Silva DM, Ruiz-Camp J, Mayer K, Vadász I, et al. Standardisation of oxygen exposure in the development of mouse models for bronchopulmonary dysplasia. *Disease models & mechanisms* 2017;10:185–196.
11. Ezz-Eldin ZM, Hamid TAA, Youssef MRL, Nabil HE-D. Clinical Risk Index for Babies (CRIB II) Scoring System in Prediction of Mortality in Premature Babies. *J Clin Diagnostic Res* 2015;9:SC08-11.
12. Wang W, Yan T, Guo W, Niu J, Zhao Z, Sun K, et al. Constitutive GLI1 expression in chondrosarcoma is regulated by major vault protein via mTOR/S6K1 signaling cascade. *Cell Death Differ* 2021;28:2221–2237.
13. Nicola T, Ambalavanan N, Zhang W, James ML, Rehan V, Halloran B, et al. Hypoxia-induced inhibition of lung development is attenuated by the peroxisome proliferator-activated receptor-γ agonist rosiglitazone. American Journal of Physiology - Lung Cellular and Molecular Physiology 2011;301:L125-34.
14. Nicola T, Hagood JS, James ML, MacEwen MW, Williams TA, Hewitt MM, et al. Loss of Thy-1 inhibits alveolar development in the newborn mouse lung. *Am J Physiol-lung C* 2009;296:L738–L750.
15. Baker JM, Hinkle KJ, McDonald RA, Brown CA, Falkowski NR, Huffnagle GB, et al. Whole lung tissue is the preferred sampling method for amplicon-based characterization of murine lung microbiota. Microbiome 2021;9:99.
16. Pierre JF, Phillips GJ, Chandra LC, Rendina DN, Thomas-Gosain NF, Lubach GR, et al. Lyticase Facilitates Mycobiome Resolution Without Disrupting Microbiome Fidelity in Primates. *J Surg Res* 2021;267:336–341.
17. Tiew PY, Dicker AJ, Keir HR, Poh ME, Pang SL, Aogáin MM, et al. A high-risk airway mycobiome is associated with frequent exacerbation and mortality in COPD. *Eur Respir J* 2021;57:2002050.
18. Ali NABM, Aogáin MM, Morales RF, Tiew PY, Chotirmall SH. Optimisation and Benchmarking of

Targeted Amplicon Sequencing for Mycobiome Analysis of Respiratory Specimens. *Int J Mol Sci* 2019;20:4991.
19. Jilling T, Ren C, Yee A, Aggarwal S, Halloran B, Ambalavanan N, et al. Exposure of neonatal mice to bromine impairs their alveolar development and lung function. American Journal of Physiology - Lung Cellular and Molecular Physiology 2018;314:L137–L143.
20. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;37:852–857.
21. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods* 2016;13:581–583.
22. Austin GI, Park H, Meydan Y, Seeram D, Sezin T, Lou YC, et al. Contamination source modeling with SCRuB improves cancer phenotype prediction from microbiome data. Nat Biotechnol 2023;1–9.doi:10.1038/s41587-023-01696-w.
23. Doupe P, Faghmous J, Basu S. Machine Learning for Health Services Researchers. *Value Health* 2019;22:808–815.
24. Karabayir I, Goldman SM, Pappu S, Akbilgic O. Gradient boosting for Parkinson's disease diagnosis from voice recordings. *Bmc Med Inform Decis* 2020;20:228.
25. Ansotegui C, Sellmann M, Shah T, Tierney K. Learning How to Optimize Black-Box Functions With Extreme Limits on the Number of Function Evaluations. *Arxiv* 2021;doi:10.48550/arxiv.2103.10321.