

# Nasopharyngeal Microbiome Composition and its Clinical Correlates in Children Hospitalized with Severe Pneumonia in East Africa

Timothy O. Makori,<sup>1,2,3,6</sup> Elijah T. Gicheru,<sup>1,2</sup> Maureen W. Mburu,<sup>1,2</sup> Mercy S. Sada,<sup>1,2</sup> Omar Nyawa,<sup>1,2</sup> Martin Mutunga,<sup>1</sup> Clement Lewa,<sup>1</sup> Robinson Cheruiyot,<sup>1</sup> Sarah Kiguli,<sup>4</sup> Peter Olupot-Olupot,<sup>5,6</sup> Rita Muhindo,<sup>5</sup> Christabel Mogaka,<sup>6,7</sup> Thomas N. Williams,<sup>7,8</sup> Charles N. Agoti,<sup>1,2,6,7,8</sup> Kathryn Maitland,<sup>7,8</sup> and Charles J. Sande<sup>1,2,6,7,8</sup>

<sup>1</sup>Bioscience Department, KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya; <sup>2</sup>Epidemiology and Demography Department, KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya; <sup>3</sup>Laboratory of Medical Microbiology, Vaccine and Infectious Diseases Institute, University of Antwerp, Antwerp, Belgium; <sup>4</sup>Department of Paediatrics, School of Medicine, Makerere University and Mulago Hospital, Kampala, Uganda; <sup>5</sup>Mbale Regional Referral Hospital, Mbale Clinical Research Institute, Mbale, Uganda; <sup>6</sup>Centre for Tropical Medicine and Global Health, University of Oxford, Oxford, United Kingdom; <sup>7</sup>Centre for Geographic Medicine Research Coast, Kenya Medical Research Institute, Kilifi, Kenya; and <sup>8</sup>Department of Infectious Disease and Institute of Global Health and Innovation, Division of Medicine, Imperial College, London, United Kingdom

**Background.** Pneumonia remains the leading cause of infectious mortality in children under 5, with the highest burden in sub-Saharan Africa. Dysbiosis in nasopharyngeal (NP) microbiota may influence pneumonia susceptibility and progression, but little is known about its composition or clinical relevance in low- and middle-income countries. We characterized the NP microbiota of children hospitalized with severe pneumonia in East Africa and investigated associations with clinical outcomes.

**Methods.** We performed 16S rRNA partial gene sequencing of NP swabs collected at hospital admission from 876 children enrolled in the COAST trial across 5 sites in Kenya and Uganda. Clinical, demographic, and virological data were prospectively collected. Microbial profiles were analyzed using hierarchical clustering, nonmetric multidimensional scaling, and multivariable regression to assess associations with respiratory viral infections, sepsis, cyanosis, bacteremia, coma, HIV status, malnutrition, sickle cell disease, malaria, and mortality.

**Results.** The NP microbiome was structured in 6 distinct clusters, each dominated by different genera, including *Staphylococcus*, *Streptococcus*, *Haemophilus*, *Dolosigranulum*, *Corynebacterium*, and *Moraxella*. Multivariable models adjusting for study site and age showed a positive association between *Corynebacterium* and early mortality. Temporal analysis showed elevated *Corynebacterium* abundance in children who died within 48 hours of admission, then declined over longer 56 survival intervals, approaching levels observed in survivors. However, time-continuous models did not support this persistent association, suggesting a subgroup effect.

**Conclusions.** We provide one of the largest high-resolution surveys of the pediatric upper airway microbiome in Africa, identifying microbial patterns associated with viral infection, HIV status, early death, and bacteremia.

**Keywords.** nasopharyngeal microbiome; pediatric pneumonia; 16 seconds rRNA sequencing.

Pneumonia is the leading infectious cause of death among children under 5 years of age, accounting for approximately 900 000 deaths globally, the majority of which occur in low- and middle-income countries (LMICs) [1]. The mortality rate in

these settings is estimated to be over 60 times higher than in high-income countries [1, 2], reflecting disparities in risk factors, healthcare access, and disease management [3–5]. The etiology of childhood pneumonia is multifactorial [6, 7], with both viral and bacterial pathogens contributing to the burden of disease [7, 8]. Respiratory syncytial virus (RSV) and influenza viruses are among the most frequently identified viral causes [7, 9], while *Streptococcus pneumoniae* and *Haemophilus influenzae*, particularly nonvaccine types, are the most important bacterial causes of pneumonia [7, 10].

The upper respiratory tract serves as the first point of contact between the host and the pathogens that cause pneumonia [11]. It is a complex ecosystem comprising both commensal and potentially pathogenic bacteria and plays a central role in respiratory health and disease by acting as a reservoir from which the lower respiratory tract may be seeded with potentially pathogenic bacteria [11–14]. Studies in high-income settings have associated respiratory illness with increased abundance of

Received 23 September 2025; accepted 05 February 2026; published online 3 March 2026  
Correspondence: Charles J. Sande, PhD, Epidemiology and Demography Department, KEMRI-Wellcome Trust Research Programme, P.O. Box 230–80108, Hospital Road, Kilifi, Kenya (csande@kemri-wellcome.org).

The Journal of Infectious Diseases®

© The Author(s) 2026. Published by Oxford University Press on behalf of Infectious Diseases Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

<https://doi.org/10.1093/infdis/jia093>

pathobionts, such as *H. influenzae* and *S. pneumoniae* [11, 15–17], and reduced levels of protective taxa such as *Dolosigranulum pigrum*, *Corynebacterium spp.*, and *Moraxella lincolnii* [18–20]. However, these insights are derived almost exclusively from children in high-income countries, where demographic and social characteristics as well as environmental exposures differ substantially from children in LMICs, particularly sub-Saharan Africa [11].

In many African countries, additional risk factors such as severe malnutrition, HIV exposure, indoor air pollution, and household structure and size may exert profound effects on the respiratory microbiota and modify the risk of developing severe pneumonia [11, 21]. To date, very few studies have analyzed the association between the respiratory microbiome and these clinical and demographic factors within this setting. To address this gap, we conducted a comprehensive, multicenter, cross-sectional analysis of the nasopharyngeal (NP) microbiome in children hospitalized with severe pneumonia in East Africa. This study was nested within the Children’s Oxygen Administration Strategies Trial (COAST), a large randomized controlled trial evaluating oxygen delivery strategies in children with pneumonia in 6 hospitals in Kenya and Uganda [22]. Within this trial framework, we analyzed a subset of children whose upper airways were sampled by NP swabbing at admission and performed bacterial community profiling using partial 16S rRNA gene sequencing. Clinical and demographic data were collected prospectively, and multiplex PCR was performed for a standardized panel of respiratory viruses. In this descriptive analysis, we sought to characterize the composition and structure of NP microbiota in this population and explore its associations with key clinical features, including HIV status, bacteremia, malaria, sickle cell disease (SCD), sepsis, RSV (Influenza and hMPV), and mortality. We provide among the first high-resolution map of the pediatric respiratory microbiome in a large-scale LMIC pneumonia cohort.

## METHODS

### Study Design, Setting, and Participants

This cross-sectional respiratory microbiome profiling study was nested within the COAST (Clinical trial No: ISRCTN15622505), a multicenter randomized controlled trial that evaluated different oxygen administration strategies in children with severe pneumonia [22]. The COAST trial [22] was conducted in 6 hospitals in East Africa: 2 in Kenya (Kilifi County Hospital and Coast General Teaching and Referral Hospital) and 4 in Uganda (Mulago National Referral Hospital, Jinja Regional Referral Hospital, Mbale Regional Referral Hospital, and Soroti Regional Referral Hospital). A subset of children aged between 28 days and 12 years who were admitted with life-threatening pneumonia and had NP samples collected at the point of hospital admission were

selected for microbial profiling by partial 16S rRNA gene sequencing. Ethical approval was obtained from the Scientific and Ethics Review Unit of the Kenya Medical Research Institute and the Uganda National Council for Science and Technology. Written informed consent was obtained from parents and guardians in accordance with GCLP guidelines. NP samples were stored at  $-80^{\circ}\text{C}$  immediately after collection for later analysis. Alongside NP sampling, detailed clinical and demographic metadata were captured, including nutritional status, vital signs, malaria status, HIV status, and outcomes such as discharge from hospital or in-hospital death. Two radiologists independently read the chest radiographs (CXRs), and a third reader resolved discrepancies and classified for presence of consolidation. A combined approach was used to define sepsis, involving diagnosis by the admitting clinician and confirmed by laboratory data of abnormal white cell counts (total WBC  $>12\,000$  or  $<4000$  cells/ $\mu\text{L}$ ) [23]. SCD status was confirmed by molecular genotyping. In addition, multiplex PCR testing was performed for 15 respiratory pathogens, including RSV A and B, influenza A, B, and C, parainfluenza viruses 1–4, human metapneumovirus, rhinovirus, adenovirus, coronavirus OC43 and 229E, and *Mycoplasma pneumoniae*.

### DNA Extraction and 16S rRNA Gene Sequencing

Frozen NP swabs were thawed on ice and vortexed on a Vortex-Genie 2 mixer (Labgene Scientific). DNA was extracted using the QIAamp DNA Mini Kit (Qiagen), following mechanical lysis with glass beads on the TissueLyser II (Qiagen) and enzymatic digestion with proteinase K. DNA was eluted in AE buffer and quantified using a Qubit 4 Fluorometer (Invitrogen). Amplification of the V3–V4 hypervariable region of the bacterial 16S rRNA gene was performed using 341F and 785R primers in a 12.5  $\mu\text{L}$  PCR reaction. PCR products were visualized on a 2% agarose gel, purified with AMPure XP beads (Beckman Coulter), and indexed using the Nextera XT Index Kit v2 (Illumina). Indexed libraries were purified, quantified using the Qubit dsDNA HS Assay Kit, and their size distribution assessed on the Agilent 4150 TapeStation. Libraries were normalized, pooled, and denatured with NaOH, spiked with 5% PhiX control, and sequenced on the Illumina MiSeq platform (2  $\times$  300 bp, MiSeq v3 kit). Two negative controls (at extraction and PCR stages) were included for quality control.

### Data Availability

The 16S rRNA sequence data generated from MiSeq sequencing in this study is available at NCBI SRA under accession number PRJNA1285267.

### Data Analysis

All analyses were performed using R (v4.5.0). Raw paired-end reads were processed using the DADA2 pipeline (v1.22.0). Quality filtering and adapter trimming were performed using

filterAndTrim interface of the fastqPairedFilter function, followed by error model learning with learnErrors and denoising via dada function. Forward and reverse reads were merged using mergePairs, and chimeras were removed using removeBimeraDenovo functions. Taxonomic assignment of amplicon sequence variants (ASVs) was performed using the assignTaxonomy function and the SILVA reference database (v138). The resulting ASV table was aggregated to genus level, and samples were normalized to relative abundance. Microbiome composition was assessed using hierarchical clustering of Bray–Curtis dissimilarities computed from genus-level relative abundance profiles. Clustering was performed using Ward’s method and visualized using dendrograms and stacked bar plots of the top 50 genera. Samples were assigned to 6 compositional clusters, identified using the NbClust R package based on the community microbiome structure. To identify dominant taxa within each cluster, average genus abundance per cluster was computed, and genus names were rendered as bubble plots, scaled by mean relative abundance. Compositional data transformation was performed by applying centered log-ratio (CLR) on the genus features prior to Maaslin2 and temporal models analyses. This controls for the spurious correlations arising from relative abundance constraints. To explore overall microbial structure in relation to clinical and demographic metadata, 2-dimensional nonmetric multidimensional scaling (NMDS) was applied to Bray–Curtis dissimilarities. The envfit function from the vegan package was used to fit clinical and virological covariates as vectors into NMDS space with 999 permutations to assess significance. Significant genus vectors were overlaid on the NMDS ordination to examine alignment of different taxa and microbial clusters with ordination gradients. Associations between individual microbial taxa and clinical or virological variable were also tested using the Maaslin2 R package (v1.21.0). Relative abundances at the genus level were used as outcomes, and different clinical features (sepsis, cyanosis, death, coma, severe malaria, HIV status, SCD, asthma history, diarrhea, malnutrition, influenza A, and RSV) were passed as fixed effects in multivariable models, adjusting for age and study site as potential confounders. Significant associations were defined using an FDR-adjusted  $q$  value  $< 0.25$ . Results were visualized using coefficient plots (with standard errors) and annotated boxplots showing genus-level relative abundance stratified by metadata variable.

## RESULTS

### Study Population and Clinical Characteristics

We analyzed NP swab samples from 876 children aged 28 days to 12 years who were admitted with life-threatening pneumonia to 5 hospitals in Kenya and Uganda participating in the COAST trial. The sites included Kilifi, Mulago, Jinja, Mbale, and Soroti. Baseline characteristics varied by site, with

significant differences in age distribution, breastfeeding status, nutritional indicators, and mortality. Median age ranged from 7.5 months in Soroti to 14 months in Jinja. Overall mortality was highest in Kilifi and Mulago (23.1% each), and lowest in Soroti (1.7%). Severe malaria, sepsis, and coma were more frequently recorded in children from Mbale and Kilifi compared with other sites. Baseline clinical and demographic characteristics are summarized in [Table 1](#).

### Clinical Associations With Airway Microbiome Structure

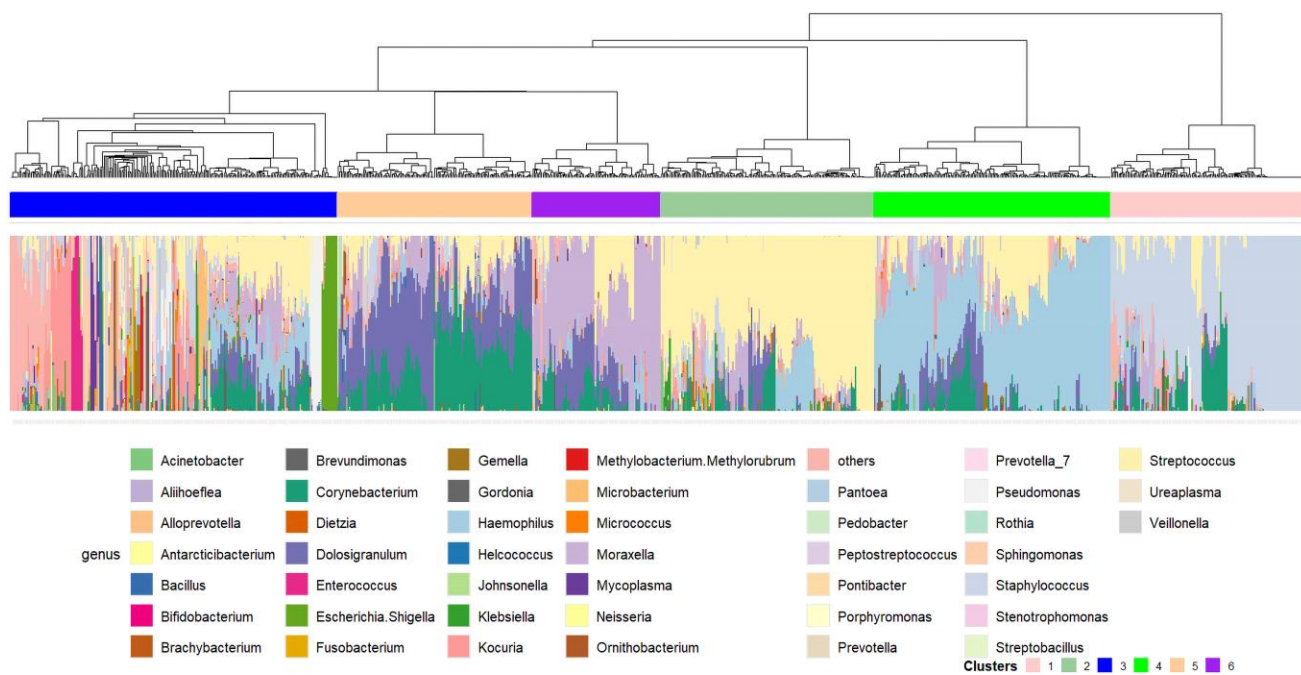
To identify associations between the overall structure of the airway microbiome and clinical factors, we first conducted hierarchical clustering of genus-level relative abundance profiles and identified 6 distinct microbiome clusters across the 876 samples. Clustering was based on Bray–Curtis dissimilarity and Ward’s linkage method, with visualization by dendrograms and stacked bar plots of the top 50 genera as shown in [Figure 1](#). Six distinct clusters were identified, each displaying a unique microbial composition pattern, with 1 or 2 dominant genera typically dominating specific clusters. Bubble plot visualizations ([Figure 2A](#)) showed that cluster 1 was dominated by *Staphylococcus*, cluster 2 by *Streptococcus*, while cluster 3 was heterogeneous and exhibited no clear patterns of dominance by a single taxon. Cluster 4 was dominated by *Haemophilus*, cluster 5 by *Dolosigranulum* and *Corynebacterium*, and cluster 6 by *Moraxella* ([Figure 2A](#)). To explore the relationships between this microbial structure and host factors, we applied NMDS to Bray–Curtis dissimilarities of genus-level profiles. Microbiota clusters 1, 2, and 5 showed clear spatial separation in ordination space, reflecting distinct compositional patterns described above ([Figure 2B](#)). Clinical and demographic covariates were then projected into the same ordination space. Clinical outcomes including influenza A infection and death were significantly associated with ordination gradients ( $P < .05$ ) as shown in [Figure 2B](#). Taxonomic vectors revealed directional alignments with clinical outcomes: death was directionally associated with cluster 5 (dominated by *Corynebacterium* and *Dolosigranulum*), while influenza A infection aligned with cluster 1 (dominated by *Staphylococcus*) as shown in [Figure 2B](#).

### Association Between Bacterial Genera and Clinical Outcomes

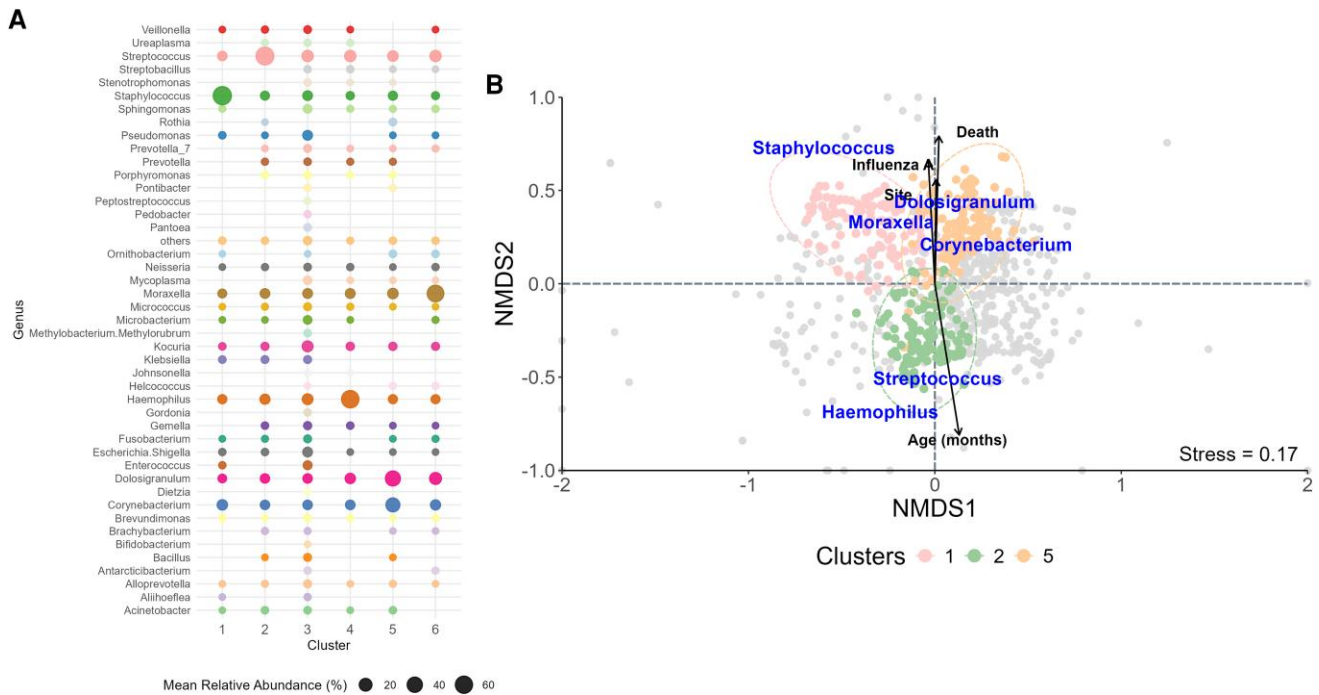
Microbiome data are compositional, with the taxa abundances expressed as relative proportions constrained to a fixed total, potentially creating false associations as changes in one taxon could reflect a shift in another rather than true biological differences. We performed centered log-ratio transformation to reduce this bias and improve interpretability, though results should still be considered within compositional data limitations. To identify associations between specific genera and clinical outcomes, we performed multivariable regression modeling using Maaslin2 after adjusting for age and study site [24].

**Table 1. Baseline Characteristics of the Study Participants**

	Jinja (N = 78)	Kilifi (N = 13)	Mbale (N = 348)	Mulago (N = 13)	Soroti (N = 424)	P Value
<b>Demographic characteristics</b>						
Age in months, median (IQR)	14 (7–38.2)	8 (3–15)	8 (4–22)	8 (3–15)	7.5 (3–18)	.002
Female, n (%)	36 (46.2)	4 (30.8)	150 (43.1)	2 (15.4)	176 (41.5)	.265
Breastfed, n (%)	78 (100%)	12 (92.3)	347 (99.7)	13 (100)	343 (80.9)	<.001
Siblings, median (IQR)	2 (1–3.8)	4 (2–5)	2 (1–3)	1 (0–2)	2 (1–4)	<.001
<b>Clinical characteristics</b>						
Temperature, median (IQR)	38 (37–38.5)	37 (37–38)	38 (37–38)	37 (36–37)	37 (37–38)	.006
Heart rate, median (IQR)	167.5 (143.8–182)	164(158–171)	164 (148–174)	142 (132–158)	157 (143–159)	<.001
Respiratory rate, median (IQR)	60 (48–72)	62 (58–64)	59.5 (50.8–66)	62 (58–66)	63 (54–73)	.005
Oxygen saturation, median (IQR)	86 (78.2–88)	89 (86–89)	87 (81.8–89)	83 (81–88)	87 (83–89)	.105
Weight, median (IQR)	9 (7–14)	5 (4–9)	8 (6–10)	8 (7–11)	8 (6–10)	.048
Comatose, n (%)	3 (3.8)	1 (7.7)	1 (0.3)	1 (7.7)	2 (0.5)	<.001
Diarrhea, n (%)	13 (16.7)	3 (23.1)	69 (19.8)	0	31 (7.3)	<.001
Severe malaria, n (%)	21 (26.9)	6 (46.2)	87 (25)	1 (7.7)	24 (5.7)	<.001
Sepsis, n (%)	16 (20.5)	2 (15.4)	46 (13.2)	0	11 (2.6)	<.001
Cyanosis, n (%)	2 (2.6)	0	11 (3.2)	0	2 (0.5)	.051
<b>Comorbidities</b>						
Asthma, n (%)	0	1 (7.7)	3 (0.9)	0	12 (2.8)	.079
HIV+, n (%)	0	0	2 (0.6)	0	2 (0.5)	.865
Sickle cell crisis, n (%)	11 (14.1)	1 (7.7)	15 (4.3)	0	18 (4.2)	.005
Malnutrition, muac < 11.5 cm, n (%)	11 (14.1)	5 (38.5)	34 (9.8)	0	27 (6.4)	<.001
Anemia, n (%)	45 (95.7)	4 (30)	313 (92.9)	12 (100)	386 (93.5)	.58
<b>Outcome</b>						
Died, n (%)	5 (6.4)	3 (23.1)	18 (5.2)	3 (23.1)	7 (1.7)	<.001



**Figure 1.** Hierarchical clustering of genus-level relative abundance profiles from 876 children revealed 6 distinct NP microbiome clusters based on Bray–Curtis dissimilarity and Ward’s linkage. Heatmap displays the top 50 genera; columns represent individual samples and rows represent bacterial genera.



**Figure 2.** Airway microbiota structure in children hospitalized with severe pneumonia. *A*, Bubble plots showing dominant genera in each cluster, scaled by mean relative abundance within cluster. Clusters were characterized by the following dominant genera: *Staphylococcus* (Cluster 1), *Streptococcus* (Cluster 2), *Haemophilus* (Cluster 4), *Corynebacterium* and *Dolosigranulum* (Cluster 5), and *Moraxella* (Cluster 6). Cluster 3 lacked a dominant genus and exhibited heterogeneous composition. *B*, Nonmetric multidimensional scaling (NMDS) of Bray–Curtis dissimilarities overlaid with microbial cluster assignments and genus vectors. Arrows represent clinical outcomes significantly associated with gradients in microbiota structure (envfit  $P < .05$ ), including death and influenza A infection. Death vector aligned directionally with Cluster 5, while influenza A infection aligned with Cluster 1. Age and study site were significantly associated with the ordination space, indicating their potential to confound associations with clinical outcomes such as mortality and viral infection.

Relative abundance of genera was modeled as the outcome, with fixed effects including sepsis, cyanosis, death, coma, severe malaria, HIV status, SCD, asthma history, diarrhea, malnutrition, RSV, and influenza A infection. Significant associations ( $P$  value  $< .05$ , FDR  $q < 0.25$ ) are shown in [Figure 3A](#). *Corynebacterium* showed a positive beta association in children who died, indicating an increased abundance of this genus in the airways of children with this outcome. *Streptococcus*, on the other hand, was negatively associated with influenza A ([Figure 3A](#)). Similarly, *Dolosigranulum* was negatively associated with Influenza A infections, reflecting lower bacterial abundance in infected children, while *Gemella* showed a positive beta association among HIV positive children, indicating higher abundance in these children relative to the rest of the study population. *Haemophilus* was negatively associated with culture-confirmed bacteremia. These associations are shown using boxplots stratified by clinical group in [Figure 3B](#).

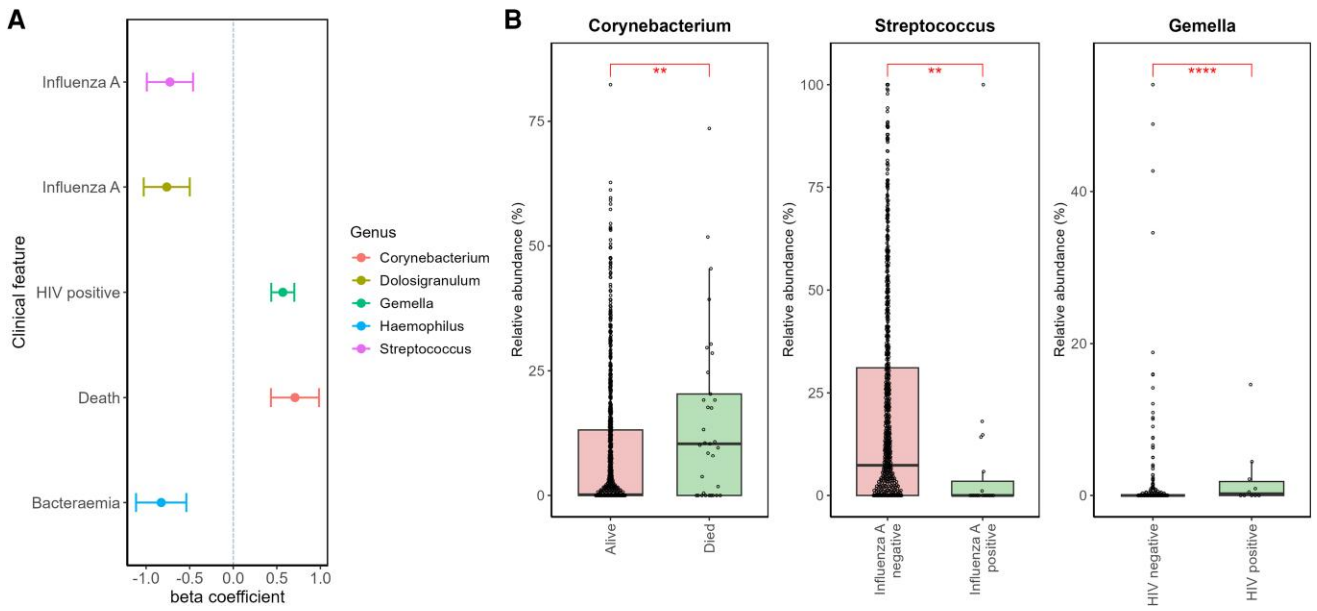
#### Alignment of Microbial Composition With Time to Death

To further explore the association between airway *Corynebacterium* abundance and death, we stratified children who died into 5 distinct time strata based on the duration of time from admission to death as follows: children who died

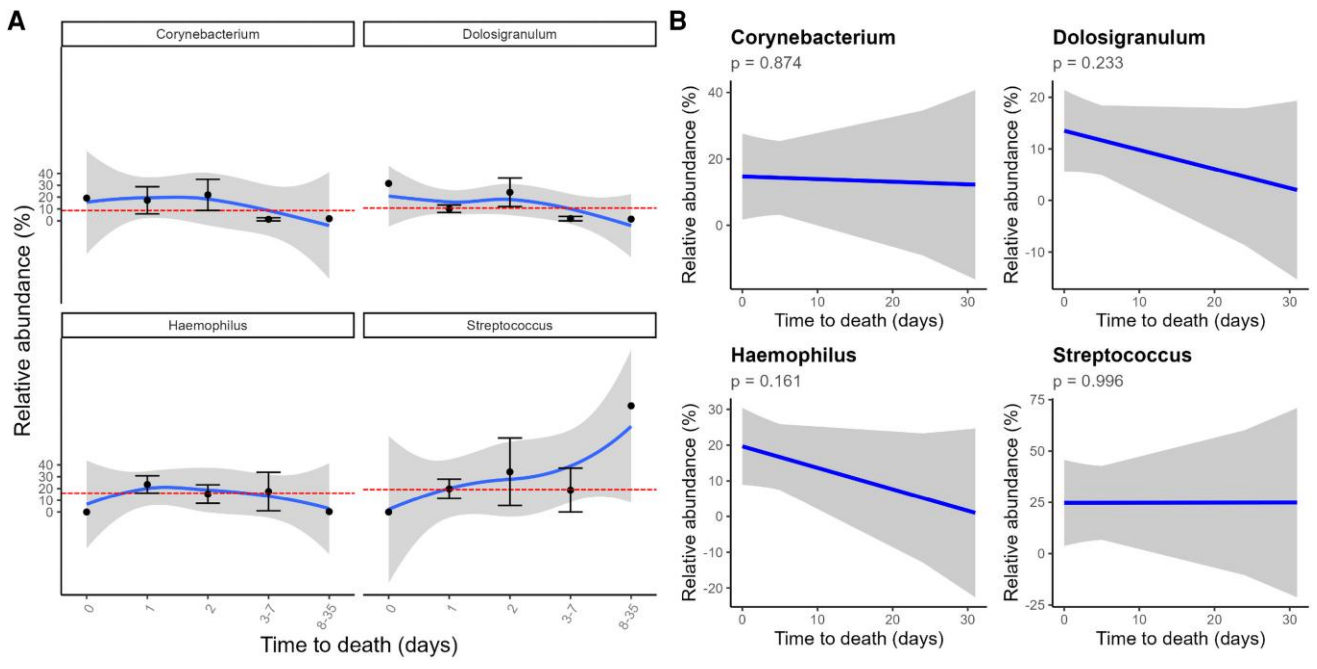
on the day of admission, children who died 1 or 2 days after admission, children who died between 3 and 7 days after admission, and children who died between 8 and 35 days after admission. We then used nonlinear regression modeling using loess fitting to visualize changes in microbial abundance relative to the time of death. Of all the genera analyzed, only *Corynebacterium* showed a clear temporal trend relative to the time of death: the highest abundance was observed among those who died on the day of admission, dropping progressively among children with longer survival times. Median *Corynebacteria* abundance in children who died  $\geq 3$  days after admission was similar to that of survivors ([Figure 4A](#)). However, this temporal trend was not statistically significant in the generalized additive mixed models treating time to death as a continuous variable and adjusting for age and study site ([Figure 4B](#)), suggesting that the observed elevated *corynebacterium* abundance in early deaths is driven by subgroup effects and not a continuous trend over time to death intervals.

#### DISCUSSION

This study presents one of the largest high-resolution surveys of the pediatric NP microbiome in sub-Saharan Africa, leveraging data from 876 children hospitalized with life-threatening



**Figure 3.** Associations between specific bacterial genera and clinical outcomes. *A*, Results from multivariable Maaslin2 models showing associations between genus-level relative abundance and clinical or virological covariates. Beta coefficients and standard errors are shown for selected genera with a false discovery rate (FDR)-adjusted  $q$  value  $< 0.25$ . *Corynebacterium* abundance was positively associated with death; *Gemella* was elevated in children living with HIV. *Streptococcus* was reduced in children infected with influenza A and *Haemophilus* was similarly reduced in children with culture-confirmed bacteremia. *B*, Annotated boxplots showing genus-level relative abundance stratified by clinical status for selected significant associations from panel (a).



**Figure 4.** Temporal patterns of genus-level abundance by time to death. *A*, LOESS smoothed trajectories of genus-level relative abundance in children who died, stratified by time from admission to death, compared with median abundance values in survivors (red dashed lines). *Corynebacterium* abundance was highest in early deaths (day 0–2), followed by a progressive decline among later deaths, converging toward survivor levels. *Dolosigranulum*, *Haemophilus*, and *Streptococcus* abundances showed no consistent temporal variation relative to time of death. Shaded ribbons denote 95% CIs for loess-smoothed estimates. *B*, Results from generalized additive mixed models (GAMMs) showing the relationship between time to death (days) from admission and genus-level relative abundance for *Corynebacterium*, *Dolosigranulum*, *Haemophilus*, and *Streptococcus*. Time to death was treated as a continuous variable and included age as a fixed effect and study site as a random effect. The blue lines denote the fitted GAMM curves and the gray shaded region show 95% confidence level. *Dolosigranulum* and *Haemophilus* show a declining trend with proximity to death, although not significant.

pneumonia in 5 hospital sites in Kenya and Uganda. Using partial 16S rRNA gene sequencing, clinical phenotyping, and virological testing, we identified 6 microbiome clusters, characterized by varying taxonomic composition, and explored their relationships with key clinical outcomes, including sepsis, cyanosis, death, coma, severe malaria, HIV status, SCD, asthma history, diarrhea, malnutrition, respiratory viral infections and death. This analysis showed significant associations between the airway microbiome structure and clinical outcomes, including death, HIV status, bacteremia and influenza A infection. Genus-level associations resolved using multivariable regression modeling, adjusting for study site and age showed outcomes such as death and HIV status were significantly associated with different microbial taxa, including *Corynebacterium* and *Gemella*, respectively.

An unexpected finding in our study was the association between *Corynebacterium* abundance and increased mortality, particularly in children who died early (< 3 days) during hospitalization. LOESS-smoothed temporal analysis showed striking early elevation in *Corynebacterium* abundance in children who died within 48 hours, followed by a progressive decline to levels that were comparable to those of survivors among children who died later. However, this observation was not statistically significant in the generalized additive mixed model by treating time to death as a continuous variable. This suggests the cross-sectional signal is driven by a local effect that is specific to early death and not a sustained trend across all death time points. This contrasts with the presumed beneficial role of this genus in HIC cohorts and experimental studies [25–29] and a recent study of South African children where *Corynebacterium* was associated with reduced risk of lower respiratory infection [30]. This may reflect differences in species level composition, underlying host factors, and geographic exposures. One speculative explanation is that the species composition of *Corynebacterium* differs substantially in LMICs. For example, *Corynebacterium* species such as *Corynebacterium ulcerans*, a known zoonotic pathogen [31], is more prevalent in rural African settings [32, 33] due to cultural nuances that enhance human-animal interaction. *C. ulcerans* has been implicated in severe respiratory infections [34], including pneumonia [35], and produces diphtheria toxin in some cases [31, 36, 37]. It is typically acquired through close contact with livestock or the consumption of unpasteurized milk [37–39]—both common in rural African households where cohabitation with animals and use of raw milk as a nutritional supplement are widespread practices [40]. Recent case series from Japan and Europe have reported increasing incidence and mortality associated with *C. ulcerans* respiratory infections [35, 39], underscoring its potential clinical impact. Given that our cohort included children from settings where such exposures are common, it is plausible that elevated *Corynebacterium* abundance may reflect carriage or infection with pathogenic species such as *C. ulcerans*. Unfortunately,

partial 16S rRNA gene sequencing is not sufficiently sensitive for species-level resolution [41], and this hypothesis will therefore require confirmation using higher-resolution shotgun metagenomics approaches or other targeted molecular methods. We also stratified nonsurvivors by time to death and compared genus-specific abundance profiles among children with different survival times to those of survivors. The airway abundance of *Dolosigranulum*, *Haemophilus*, and *Streptococcus*, on the other hand, did not exhibit a similar temporal pattern relative to the time of death and remained comparable to corresponding median survivor abundances at all time points. This temporal pattern suggests that elevated *Corynebacterium* abundance at admission may be a marker of acute disease severity, or a dysbiotic microbial state associated with imminent deterioration. In contrast, the absence of clear temporal variation in *Dolosigranulum*, *Haemophilus*, and *Streptococcus* abundance suggests that their associations with mortality are either weak, nonspecific, or masked by clinical heterogeneity. Together, these findings raise the possibility that the abundance of certain *Corynebacterium* species may serve as a context-sensitive indicator of early, fulminant disease trajectories in this setting, although species-specific attributions will require confirmation using shotgun approaches as discussed above.

Several limitations of this study warrant consideration. First, this was a cross-sectional analysis at hospital admission, and therefore causality could not be inferred. Second, partial 16S rRNA sequencing does not resolve species or functional potential, limiting taxonomic specificity and precluding inference of virulence mechanisms. For instance, *Haemophilus* abundance was negatively associated with bacteremia. While counterintuitive given the established role of *H. influenzae* as an invasive respiratory pathogen, our 16S partial sequencing approach cannot distinguish pathogenic *H. influenzae* from commensal *Haemophilus* species. Importantly, all the NP samples in this study are archived in our biobank and available for future shotgun metagenomics for strain-level resolution and functional profiling. Third, our data are drawn from hospitalized children with severe disease and may not generalize to mild community-acquired pneumonia or healthy controls. Nonetheless, this study provides one of the most comprehensive descriptions of the NP microbiome in African children with pneumonia to date. By combining deep clinical phenotyping with community profiling across diverse settings, we reveal both shared and context-specific features of microbial ecology in pediatric pneumonia. These findings set the stage for future mechanistic studies and highlight the importance of geographic and ecological diversity in microbiome research.

## Notes

**Author contribution.** T. O. M. performed conceptualization and design, methodology, laboratory and bioinformatics data

analysis, visualization, drafting and review of the article. S. K., P. O.-O., R. M. involved in Field data collection and review of the article. C. M. contributed to data management (data curation) and review of the article. M. M., C. L., and R. C. contributed to laboratory analysis (viral diagnostics). E. T. G., M. W. M., M. S. S., O. N. performed laboratory analysis (partial 16S rRNA sequencing) and article review. T. N. W., C. N. A. contributed to funding acquisition, conceptualization and design, data acquisition, and review of the article. K. M. involved in funding acquisition, conceptualization and design, data acquisition, supervision, review of the article. C. N. S. involved in funding acquisition, conceptualization and design, data acquisition, supervision, drafting and review of article.

**Acknowledgments.** We thank all the study participants and staff from all the centers participating in the COAST trial. We thank the Trial Steering Committee for its support of the trial.

**Financial support.** This study is sponsored by Imperial College London, P46493 and supported by the Joint Global Health Trials scheme (Medical Research Council, UK and Wellcome Trust; MR/L004364/1 and 102231, respectively). The views expressed in this publication are those of the author(s) and not those of funders.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

**Data Availability.** The data supporting the findings of this study are available from Harvard Dataverse (<https://doi.org/10.7910/DVN/L2C5E6>).

**Potential conflicts of interest.** All authors: No reported conflicts.

## References

1. McAllister DA, Liu L, Shi T, et al. Global, regional, and national estimates of pneumonia morbidity and mortality in children younger than 5 years between 2000 and 2015: a systematic analysis. *Lancet Glob Health* **2019**; 7:e47–57.
2. Marangu D, Zar HJ. Childhood pneumonia in low-and-middle-income countries: an update. *Paediatr Respir Rev* **2019**; 32:3–9.
3. Ahmed S, Mvalo T, Akech S, et al. Protecting children in low-income and middle-income countries from COVID-19. *BMJ Glob Health* **2020**; 5:50.
4. Simkovich SM, Underhill LJ, Kirby MA, et al. Resources and geographic access to care for severe pediatric pneumonia in four resource-limited settings. *Am J Respir Crit Care Med* **2022**; 205:183–97.
5. Nemet M, Gmehlin C, Vukoja M, et al. Ventilator-associated pneumonia in low- and middle-income vs. high-income countries: the role of ventilator bundle, ventilation practices, and healthcare staffing. *Chest* **2025**; 167:1628–38.
6. Yun KW. Community-acquired pneumonia in children: updated perspectives on its etiology, diagnosis, and treatment. *Clin Exp Pediatr* **2024**; 67:80–9.
7. Torres A, Cilloniz C, Niederman MS, et al. Pneumonia. *Nat Rev Dis Primers* **2021**; 7:25.
8. Yadav KK, Awasthi S. Childhood pneumonia: what's unchanged, and what's new? *Indian J Pediatr* **2023**; 90:693–9.
9. de Benedictis FM, Kerem E, Chang AB, Colin AA, Zar HJ, Bush A. Complicated pneumonia in children. *Lancet* **2020**; 396:786–98.
10. Cilloniz C, Martin-Loeches I, Garcia-Vidal C, Jose AS, Torres A. Microbial etiology of pneumonia: epidemiology, diagnosis and resistance patterns. *Int J Mol Sci* **2016**; 17: 2120.
11. Man WH, de Steenhuijsen Piter WAA, Bogaert D. The microbiota of the respiratory tract: gatekeeper to respiratory health. *Nat Rev Microbiol* **2017**; 15:259–70.
12. Kumpitsch C, Koskinen K, Schöpf V, Moissl-Eichinger C. The microbiome of the upper respiratory tract in health and disease. *BMC Biol* **2019**; 17:87.
13. Santacroce L, Charitos IA, Ballini A, et al. The human respiratory system and its microbiome at a glimpse. *Biology (Basel)* **2020**; 9:318.
14. Claassen-Weitz S, Lim KYL, Mullally C, Zar HJ, Nicol MP. The association between bacteria colonizing the upper respiratory tract and lower respiratory tract infection in young children: a systematic review and meta-analysis. *Clin Microbiol Infect* **2021**; 27:1262–70.
15. Seppanen EJ, Bayliss J, Clark SL, et al. Haemophilus influenzae remains the predominant otitis media pathogen in Australian children undergoing ventilation tube insertion in the PCV13 era. *J Infect* **2025**; 90:106478.
16. Bogaert D, De Groot R, Hermans PWM. Streptococcus pneumoniae colonisation: the key to pneumococcal disease. *Lancet Infect Dis* **2004**; 4:144–54.
17. Laufer AS, Metlay JP, Gent JF, Fennie KP, Kong Y, Pettigrew MM. Microbial communities of the upper respiratory tract and otitis media in children. *mBio* **2011**; 2:245–55.
18. de Steenhuijsen Piter WAA, Binkowska J, Bogaert D. Early life microbiota and respiratory tract infections. *Cell Host Microbe* **2020**; 28:223–32.
19. Henares D, Brotans P, de Sevilla MF, et al. Differential nasopharyngeal microbiota composition in children according to respiratory health status. *Microb Genom* **2021**; 7: 000661.
20. Lapidot R, Faits T, Ismail A, et al. Nasopharyngeal dysbiosis precedes the development of lower respiratory tract infections in young infants, a longitudinal infant cohort study. *Gates Open Res* **2024**; 6:48.

21. Dubowski K, Kaali S, Jack D, et al. Infant nasopharyngeal microbiota subphenotypes and early childhood lung function: evidence from a rural Ghanaian pregnancy cohort. *Int J Environ Res Public Health* **2021**; 18:7276.
22. Maitland K, Kiguli S, Opoka RO, et al. Children's oxygen administration strategies trial (COAST): a randomised controlled trial of high flow versus oxygen versus control in African children with severe pneumonia. *Wellcome Open Res* **2018**; 2:100.
23. Singer M, Deutschman CS, Seymour C, et al. The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* **2016**; 315:801–10.
24. Mallick H, Rahnavard A, McIver LJ, et al. Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput Biol* **2021**; 17:e1009442.
25. Bosch AATM, Levin E, van Houten MA, et al. Development of upper respiratory tract microbiota in infancy is affected by mode of delivery. *EBioMedicine* **2016**; 9:336–45.
26. Biesbroek G, Tsvitshivadze E, Sanders EAM, et al. Early respiratory microbiota composition determines bacterial succession patterns and respiratory health in children. *Am J Respir Crit Care Med* **2014**; 190:1283–92.
27. Kanmani P, Clua P, Vizoso-Pinto MG, et al. Respiratory commensal bacteria *Corynebacterium pseudodiphtheriticum* improves resistance of infant mice to respiratory syncytial virus and *Streptococcus pneumoniae* superinfection. *Front Microbiol* **2017**; 8(AUG): 273272.
28. Horn KJ, Jaber Vivar AC, Arenas V, Andani S, Janoff EN, Clark SE. *Corynebacterium* Species inhibit *Streptococcus pneumoniae* colonization and infection of the mouse airway. *Front Microbiol* **2022**; 12:804935.
29. Bomar L, Brugger SD, Yost BH, Davies SS, Lemon KP. *Corynebacterium accolens* releases antipneumococcal free fatty acids from human nostril and skin surface triacylglycerols. *mBio* **2016**; 7:e01725–15.
30. Claassen-Weitz S, Xia Y, Hannan L, et al. Nasopharyngeal microbiota in South African infants with lower respiratory tract infection: a nested case-control study of the Drakenstein Child Health Study. *Clin Infect Dis* **2025**; 81:e668–79.
31. Crestani C, Passet V, Rethoret-Pasty M, et al. Microevolution and genomic epidemiology of the diphtheria-causing zoonotic pathogen *Corynebacterium ulcerans*. *Nat Commun* **2025**; 16:1–11.
32. Addo PB, Dennis SM. *Corynebacteria* associated with diseases of cattle, sheep and goats in Northern Nigeria. *British Vet J* **1977**; 133:334–9.
33. World Health Organization. Disease outbreak news: diphtheria in Nigeria. World Health Organization. <https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON485>. Accessed 18 June 2025.
34. Otshudiema JO, Acosta AM, Cassidy PK, Hadler SC, Hariri S, Tiwari TSP. Respiratory illness caused by *Corynebacterium diphtheriae* and *C. ulcerans*, and use of diphtheria antitoxin in the United States, 1996-2018. *Clin Infect Dis* **2021**; 73:e2799–806.
35. Yasuda I, Matsuyama H, Ishifuji T, et al. Severe pneumonia caused by toxigenic *Corynebacterium ulcerans* infection, Japan. *Emerg Infect Dis* **2018**; 24:588–91.
36. Sekizuka T, Yamamoto A, Komiya T, et al. *Corynebacterium ulcerans* 0102 carries the gene encoding diphtheria toxin on a prophage different from the *C. diphtheriae* NCTC 13129 prophage. *BMC Microbiol* **2012**; 12:1–7.
37. Hillan A, Gibbs T, Weaire-Buchanan G, et al. Zoonotic transmission of diphtheria toxin-producing *Corynebacterium ulcerans*. *Zoonoses Public Health* **2024**; 71:157–69.
38. Hoefler A, Herrera-León S, Domínguez L, et al. Zoonotic transmission of diphtheria from domestic animal reservoir, Spain. *Emerg Infect Dis* **2022**; 28:1257.
39. Zendri F, Isgren CM, Sinovich M, et al. Case report: toxigenic *Corynebacterium ulcerans* diphtheria-like infection in a horse in the United Kingdom. *Front Vet Sci* **2021**; 8: 650238.
40. Mpatshwenumugabo JP, Mukasafari MA, Ndahetuye JB, Wredle E, Båge R. A systematic literature review of milk consumption and associated bacterial zoonoses in East Africa. *J Appl Microbiol* **2023**; 134:1–13.
41. Johnson JS, Spakowicz DJ, Hong BY, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun* **2019**; 10:1–11.