



Early View

Original article

Dissimilarity of the gut-lung axis and dysbiosis of the lower airways in ventilated preterm infants

David Gallacher, Emma Mitchell, Dagmar Alber, Richard Wach, Nigel Klein, Julian R. Marchesi, Sailesh Kotecha

Please cite this article as: Gallacher D, Mitchell E, Alber D, *et al.* Dissimilarity of the gut-lung axis and dysbiosis of the lower airways in ventilated preterm infants. *Eur Respir J* 2020; in press (<https://doi.org/10.1183/13993003.01909-2019>).

This manuscript has recently been accepted for publication in the *European Respiratory Journal*. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJ online.

Copyright ©ERS 2020

Dissimilarity of the gut-lung axis and dysbiosis of the lower airways in ventilated preterm infants

¹David Gallacher*, ¹Emma Mitchell*, ²Dagmar Alber, ³Richard Wach, ²Nigel Klein, ^{4,5}Julian R Marchesi, ¹Sailesh Kotecha.

¹Department of Child Health, School of Medicine, Cardiff University, Cardiff, UK.

²Institute of Child Health, University College London, London, UK.

³Neonatal Unit, North Bristol NHS Trust, Bristol, UK.

⁴School of Biosciences, Cardiff University, Cardiff, UK.

⁵Division of Integrative Systems Medicine and Digestive Disease, Imperial College London, London, UK

*Joint first authors

Address for Correspondence:

Professor Sailesh Kotecha

Department of Child Health

School of Medicine

Cardiff University

Heath Park

Cardiff CF14 4XN

Email: KotechaS@cardiff.ac.uk

Abstract

Background: Chronic lung disease of prematurity (CLD), also called bronchopulmonary dysplasia, is a major consequence of preterm birth but the role of the microbiome in its development remains unclear. We, therefore, assessed the progression of the bacterial community in ventilated preterm infants over time in the upper and lower airways, and assessed the gut-lung axis by comparing the upper and lower airways bacterial communities with the stool findings. Finally, we assessed if the bacterial communities were associated with lung inflammation to suggest dysbiosis.

Methods: We serially sampled multiple anatomical sites including the upper airway (nasopharyngeal aspirates, NPA), lower airways (tracheal aspirate fluid, TAF, and bronchoalveolar lavage fluid, BAL) and the gut (stool) of ventilated preterm-born infants. Bacterial DNA load was measured in all samples and sequenced using the V3-V4 region of the 16S rRNA gene

Results: From 1,102 (539 NPA, 276 TAF, 89 BAL, 198 stool) samples from 55 preterm infants, 352 (32%) amplified suitably for 16s RNA gene sequencing. Bacterial load was low at birth, quickly increased with time but was associated with predominant operational taxonomic units (OTUs) in all sample types. There was dissimilarity in bacterial communities between the upper and lower airways and the gut with a separate dysbiotic inflammatory process occurring in the lower airways of infants. Individual OTUs were associated with increased inflammatory markers.

Conclusions: Taken together, these findings suggest that targeted treatment of the predominant organisms, including those not routinely treated such as *Ureaplasma* spp., may decrease the development of CLD in preterm-born infants.

Introduction

Chronic lung disease of prematurity (CLD), also called bronchopulmonary dysplasia (BPD), is a major consequence of preterm birth being influenced by many early life [1]. We have previously shown that the presence of predominant bacteria is associated with the development of CLD most likely via the pulmonary inflammatory pathway [2], [3]. Furthermore, specific organisms such as *Ureaplasma* are also strongly implicated in the pathogenesis of CLD [4], [5] with some evidence suggesting that eradication of *Ureaplasma* may decrease rates of CLD [6]. In the term-born infant, early bacterial colonisation of the upper respiratory tract appears to occur within minutes of birth [7] or possibly antenatally but the natural history of colonisation of the upper and lower airways of the preterm lung is less clear [8]. A recent systematic review [9] investigating the development of the airway microbiome in preterm infants raised more questions about the bacterial colonisation of the preterm lung than have been answered with existing studies; no study has investigated the links between the upper airways and the gastrointestinal tract, i.e. the proposed gut-lung axis, with the lower airways of preterm infants. The included six studies [10]–[15] showed variable findings e.g. one reported progression from Proteobacteria to Firmicutes but another showed reverse progression. To provide a more comprehensive assessment of the bacterial community in preterm infants who are at risk of developing CLD, (a) we serially studied multiple anatomical sites, including obtaining tracheal aspirate fluid (TAF) and bronchoalveolar lavage (BAL) fluid from ventilated preterm infants to assess the tracheal and lower airways respectively; nasopharyngeal aspirates (NPA) and stool samples to assess the progression of the bacterial community in ventilated preterm infants over time in the upper and lower airways. (b) We assessed the microbial relationship between the stool and upper/lower airways, i.e. the gut-lung axis, for which there are no studies thus far in term or preterm-born infants. (c) We also assessed the relationship between the bacterial communities and clinical factors including mode of delivery, sex, clinical site of sampling and effect of antibiotics. (d) Finally, we measured proinflammatory cytokines interleukin (IL)-6 and -8 in all samples types and related the findings to the bacterial communities observed.

Methods

Sample Collection

Preterm infants of ≤ 32 weeks' gestation were recruited from two tertiary regional neonatal intensive care units in the United Kingdom. TAF and NPA were collected at the time of routine endotracheal tube (ETT) suctioning. BAL samples were collected as per our previous studies [2], [16] and according to the European Respiratory Society taskforce guidelines [17] at the time of routine ETT suctioning as described in the online supplement. TA, BAL and NPA samples were collected daily during the first week then twice weekly or until extubation, whichever occurred earlier. Stool samples were collected weekly until 28 days of life. Ethical approval (Ref 14/WA/0190) was obtained from the Wales Research Ethics Committee 2 and written informed consent was obtained from the parents.

DNA Extraction, Bacterial Load and Sequencing

DNA extraction, bacterial load and 16S sequencing are described in more detail in the online supplement. Total bacterial load was quantified by a qPCR assay based on published protocols [18] amplifying the V3-V4 region of the 16S rRNA gene. Purified PCR product were quantified and samples with adequate amplification pooled to create an amplicon library. Separate libraries were prepared for TAF, BAL, NPA and stool samples. Each library was sequenced on an Illumina MiSeq device [19].

Data Analysis

Mothur v1.39.5 [20] was used to process the sequencing data, excluding chimeric sequences and low quality reads and to group the DNA sequence data into OTUs at 3% dissimilarity. Samples with less than 1,000 reads were excluded. The reference database from the Ribosomal Database Project [21] was used for phylogenetic identification of each OTU from kingdom to genus level. Samples were rarefied to 1,000 reads per sample prior to further analysis. Statistical analysis was performed within

R v3.4.2 [22] using Phyloseq v1.20.0 [23] and Vegan [24]. STAMP v2.13 [25] was used to compare bacterial profiles between different groups. Datasets were merged at genus level for comparison of microbiome profiles from different anatomical sites. Samples with a single OTU representing >50% reads were considered to contain a predominant organism.

Results

1,102 samples were collected from 55 preterm infants born at ≤ 32 weeks' gestation (median \pm IQR: 26.0, 24.7 to 27.5 weeks) during the first 28 days of age including 539 NPA, 276 TA, 89 BAL and 198 stool samples. The demographics are shown in **Table 1** including comparison of data between the two recruitment centres (University Hospital of Wales, UHW; and North Bristol Trust, NBT).

Marginally more immature infants were recruited at NBT although the difference was not statistically significant (26.8 vs 25.9 weeks, $p=0.07$). More patent ductus arteriosus were identified at NBT since they screen for this condition. As expected with recruitment of only ventilated preterm infants, most infants developed CLD defined using the NIHD definitions [26] with 84% developing moderate to severe CLD.

352 (32%) of the 1,102 samples were successfully sequenced for 16S rRNA genes with the greatest success rate for stool (73%) samples and least for TAF (17%) samples (**Figure 1**). Bacterial load was low at days 1-3 with only 1.1% and 6.7% of TAF and NPA samples successfully sequenced for bacterial 16S rRNA genes with a success rate of 35.7% for stool samples suggesting a low bacterial load at birth especially in the upper airways. Although 19.4% of BAL samples were available for DNA sequencing, a lower threshold for bacterial DNA sequencing was used for these samples (concentration of DNA post-PCR amplification of 0.3ng/ μ L for TAF and NPA, 0.1ng/ μ L for BAL, and 0.5ng/ μ L for stool). Bacterial load peaked at day 4-7 for BAL and TAF samples and at days 8-14 for NPA and stool samples. Interestingly, a similar pattern of increase was observed for the TAF and BAL samples, and for the stool and NPA samples (**Figure 1**).

After low quality reads and chimeric sequences were removed, the mean reads were 109,457 (range 8,665 to 426,696) for TAF samples; 44,548 (27 to 177,875) for BAL samples; 60,295 (5,951 to 222,539) for NPA samples; and 59,329 (1 to 481,757) for stool samples with 0, 2, 0 and 39 samples removed respectively for reads of <1,000 per sample. The maximum reads for negative controls were 7,558, 588, 580 and 206 respectively with one TAF sample discarded due to co-clustering with a negative control sample.

The majority of phyla in all samples, as shown in **Figure 2** and summarised in **Online Figure 1**, were Proteobacteria and Firmicutes; Tenericutes were also present particularly in TAF and BAL samples and to a lesser extent in NPA samples; Bacteroidetes were noted in stool samples with lesser presence in NPA, TAF and BAL samples. The phyla TAF and BAL pattern, over the first month of life, remained static with a mean relative abundance ~40-50% of Proteobacteria; whilst Proteobacteria in NPA and stool samples increased from a mean relative abundance of 15% in NPA and 30% in stool in the first week of life to 46% and 55% respectively by the third/fourth weeks of age (**Figure 2**, summarised data shown in **Online Figure 2**). Most samples contained a single dominant phylum (>50% of reads from a single phylum). At genus level, *Staphylococcus*, as shown in **Figure 3**, was the dominant genus (>50% of reads from a single genus) in nearly 30% of all samples with *Klebsiella*, *Escherichia*, *Unclassified Enterobacteriaceae*, *Enterococcus* and *Serratia* predominating in 35% and the Tenericutes *Ureaplasma* and *Mycoplasma* for 5.2% all of samples. Whilst Proteobacteria increased and Firmicutes decreased with age especially in NPA (Proteobacteria $p < 0.001$, Firmicutes $p = 0.008$, Kruskal-Wallis test with Benjamini Hochberg FDR) and stool samples, there was no obvious patterns of change at genus level although dominance of the above-mentioned genera was evident. Both at genus and OTU level, a dominant genus/OTU was present in most sample with 31/34 (91%), 43/47 (92%), 122/126 (97%) and 130/145 (90%) of BAL, TAF, NPA and stool samples respectively having >50% of their reads for a single dominant OTU as shown in **Figure 3**.

Interesting, each infant had their own individual bacterial pattern which differed between individual infants and also between anatomical sites. (**Figure 4A and 4C**). Since there was no overall pattern over time, we collated the samples from each anatomical site for further analyses. The number of observed genera and Chao1 index were different between the sample types being greater in the BAL and TAF samples ($p < 0.001$ for both indices) (**Figure 4B**) but not for Shannon or Simpson's alpha diversity indices (data not shown) when calculated at genus level. However, there were no differences noted for observed OTUs or Chao1 when calculated at OTU level (**Figure 4B**). This data of similar indices for genera and OTU level for BAL and TAF samples suggest that individual OTUs predominate in these samples but the increase in indices between the genera and OTUs for NPA and stool samples suggest greater richness in these samples. Beta-diversity analysis for TAF and NPA samples did not show any significant pattern over time (Online Figure 4).

When bacterial communities were examined between the four sample types by PERMANOVA, all differed significantly from each other as shown in **Figure 4C**: BAL vs TAF; BAL vs NPA; BAL vs stool; TAF vs NPA; TAF vs stool; and NPA vs stool; all $p < 0.01$.

Delivery mode influences early colonisation across many anatomical sites in term new-born infants and in the limited data for preterm infants in stool samples [7], [27], [28]. Mode of delivery was associated with an increase in Gram-negative genera (*Acinetobacter* ($p = 0.04$) and marginally with *Pseudomonas* ($p = 0.06$)); *Mycoplasma* ($p = 0.04$) in TAF samples; and *Serratia* ($p = 0.01$) in NPA from infants delivered vaginally. For infants delivered by caesarean section, *Staphylococcus* were increased in NPA ($p = 0.02$), and marginally in TAF samples ($p = 0.056$) (**Online Figure 2A**). However, sex did not influence TAF or NPA bacterial communities (**Online Figure 2B**). A significant difference for bacterial community structure was noted between the recruitment centres ($p < 0.05$) (**Online Figure 2C**). This may be partially explained by the outbreak of *Acinetobacter* at one site resulting in greater

Acinetobacter observed at that recruitment centre.

At both recruiting centres the initial prophylactic antibiotic regimen of benzylpenicillin and gentamicin was altered according to subsequent microbiological assessments. 219/276 (79%) TAF and 335/539 (62%) NPA samples were taken while the infants were receiving antibiotic therapy (**Online Figure 3**). Samples taken when not on antibiotics were significantly more likely to be successfully sequenced for 16S rRNA genes than samples taken whilst on antibiotics (OR=2.91 (95% confidence interval 1.48, 5.73), $p=0.002$ for TAF; and 3.10 (2.05 to 4.67), $p<0.0001$ for NPA samples). Antibiotics did not affect alpha diversity in successfully sequenced TAF and NPA samples suggesting that they globally suppressed bacterial growth rather than suppressing specific species. Nevertheless, antibiotics were associated with increased proportion of Tenericutes ($p=0.03$), most likely due to *Mycoplasma* which were marginally increased ($p=0.059$) in TAF samples compared to the no antibiotic group.

Since a single OTU predominated (>50% of reads) in most of the successfully sequenced samples, potentially those samples with greater initial biomass, we measured IL-6 and IL-8 in all samples to evaluate if successful bacterial DNA detection in samples was associated with an inflammatory response. IL-6 and IL-8 were largely undetectable in the NPA and faecal fluid extracts but were noted in considerable quantities in the BAL and TAF samples if they could be sequenced as shown in **Figure 5A** in TAF and BAL samples: TAF successful sequencing vs unsuccessful sequencing samples: IL-6: 1,222 (534, 3,224) vs 539 (217, 1,377); and IL-8: 26,383 (3,460, 51,740) vs 9,031 (2,535, 20,511) pg/ml; and BAL: IL-6 3,629 (2,029, 8,538) vs 1,213 (485, 2,459); and IL-8 80,689 (12,303, 196,477) vs 9520 (3,060, 18,572) pg/ml, all $p<0.001$. We next tracked the dynamic changes in individual infants as shown in **Figure 5B**. Each individual infant had episodic increases in IL-6 and IL-8 which corresponded to increased bacterial load and to presence of predominant OTUs. From 20 episodic TAF IL-8 peaks of >50,000 pg/ml observed in 19 infants, 13 episodes (65%) were associated with

presence of bacterial DNA within 24 hours of the peak cytokine concentration. Both IL-6 and IL-8 showed peak concentration at 7 days of age as we have shown previously in infants developing CLD[2], [3], [16] (**Figure 5C**). Finally, the predominance of OTUs such as *Acinetobacter*, unclassified *Enterobacteriaceae* and Mollicutes including *Mycoplasma* and *Ureaplasma* were strongly associated with the greatest increases in IL-6 and IL-8 in BAL, being several-fold greater than in samples which could not be sequenced (**Figure 5D**).

Discussion

In this comprehensive longitudinal study, we have studied the upper and lower airways of preterm infants who were at risk of developing CLD and related the findings to serial stool samples to assess the gut-lung axis. We have made several important observations: firstly, low bacterial DNA load was noted in the first few days of life which increased during the first month of life. Secondly, all anatomical sites sampled including the upper and lower airways had their own distinct patterns of bacterial community, in most cases with the presence of a dominant OTU, and with dissimilarity between the anatomical sites. Thirdly, bacterial communities in the lower airways were associated with an active proinflammatory phase. Fourthly, specific OTUs were associated with marked increases in proinflammatory cytokines.

Overall, from 1,102 samples, 31.9% (352) were satisfactorily amplified for bacterial sequencing. Bacterial load was low in all samples during the first three days of life but peaked at days 4-7 of life in BAL and TAF samples, and at days 8-14 in stool and NPA samples. Bacterial acquisition is most likely to be from the mother with vaginally associated microbes including Tenericutes predominating in vaginally delivered infants, and *Staphylococcus* predominating when delivery was by caesarean section.

At phyla level, the NPA and stool samples changed similarly over time with an increase in

Proteobacteria which remained relatively static in the BAL/TAF samples. The Tenericutes, *Ureaplasma* and *Mycoplasma*, were mainly confined to the lower airway samples with only some presence in NPA samples but were almost absent from stool samples. In contrast, *Acinetobacter* was present in stool and in NPA samples but not in lower airway samples reflecting the outbreak in that unit. Almost all samples had a predominant OTU, with >90% of samples having a predominant OTU of >50% of reads with *Staphylococcus*, *Mycoplasma*, and *Ureaplasma* predominating especially in the lower airways. Interestingly, all infants had their own bacterial pattern with little overlap between the sample types suggesting an active infective episodic process. Significant differences were present in the community structure of the samples from each anatomical site as was shown by the PERMANOVA analysis. Alpha diversity indices estimating richness showed greater number of genera present in BAL and TAF than in NPA or stool samples, but the differences disappeared at OTU level suggesting greater richness at OTU level for NPA and stool samples. Since measures of evenness (Shannon diversity index) did not show a difference between the sample sites, the data suggest that the degree of dominance of a single OTU was no different between anatomical sites. Differences between the NPA and BAL/TAF samples may be partly explained by the presence of an uncuffed ET tube which partially separates the upper and lower airways. Obtaining lower airways samples from non-ventilated preterm or term new-born infants is not possible thus limited data exist for the microbial evolution of the lower airways [29].

As previously observed for term-born infants, mode of delivery was associated with increased proportion of *Staphylococcus* in NPA ($p=0.022$) and marginally in TAF ($p=0.056$) in those who were delivered by caesarean section and increased *Mycoplasma* ($p=0.039$) and *Acinetobacter* ($p=0.042$) in TAF samples and *Serratia* ($p=0.008$), *Ureaplasma* ($p=0.024$) and *Corynebacterium* ($p=0.028$) in NPA samples from vaginally delivered infants. Sex did not affect the bacterial colonisation but the centre at which the sample was obtained did influence the results. These results may be partly explained by the outbreak of *Acinetobacter* in one unit but the differences between the two centres remained in

sensitivity analyses when the infants colonised with *Acinetobacter* were removed thus suggesting that the local environment or local personnel appear to influence colonisation in addition to the maternal milieu. Antibiotics, as expected, decreased bacterial load with fewer samples being suitable for bacterial gene sequencing. The diversity and number of OTUs were, however, similar when the babies were on or off antibiotics. *Tenericutes* presence was increased in TAF samples ($p=0.03$) and a trend towards *Tenericutes* increased in NPA samples ($p=0.059$) in the antibiotic group.

Most interestingly, we noted several fold differences in IL-6 and IL-8 concentrations in BAL and TAF when bacterial sequencing was possible against when it was not. Our data suggest that an infective dysbiotic environment is occurring in the lungs of babies who are at risk of developing CLD. Low biomass samples may have lower cytokine concentrations due to dilution caused by variation in volume of samples returned. However, as per our previous observations, the collated data for IL-6 and IL-8 showed a peak concentration of both cytokines at approximately 7 days of age [2], [3], [16]. Increased cytokine peaks associated with predominant OTUs in each individual infant suggest an infective process. This was further supported by the marked increases in TAF IL-6 and IL-8 for individual OTUs confirming that infective processes were common in the lungs of babies who are at risk of developing CLD.

Taken together, these data suggest that both the bacterial community and pulmonary inflammation are potential targets with antimicrobials such as azithromycin which has both anti-infective and anti-inflammatory activities to decrease rates of CLD [6].

There are several limitations to this study, most notably the presence of an ET tube to support the infant's respiratory system. Most importantly, this precludes the acquisition of samples from "healthy" preterm- and term-born infants which is not ethically possible. The presence of an ET tube

may physically separate the lower and upper airways but, despite prolonged ventilation in many infants, we noted separate processes in the upper and lower airways and in the GI tract. The strength of the study is the large number of samples from multiple anatomical sites obtained longitudinally making this by far the largest study of preterm infants. Furthermore, we have related the findings between recruitment centres and with the inflammatory processes occurring in the lungs of infants at risk of developing CLD.

In summary, bacterial load was low at birth but quickly increases and microbes are most likely to be acquired from the mother. There was dissimilarity between the upper and lower airways and the gut in bacterial community structures with a separate dysbiotic inflammatory process occurring in the lower airways of infants at increased risk of developing CLD. Additionally, the pulmonary inflammatory processes were occurring despite the infants being on antibiotics suggesting that microbes not susceptible to routinely used antibiotics such as *Ureaplasma* and *Mycoplasma* may predominate and potentially provide a therapeutic target to decrease the development of CLD in preterm-born infants [6].

Acknowledgments

SK receives funds from the MRC (MR/M022552/1) and NIHR/HTA (16/111/106). For JRM, the Division of Integrative Systems Medicine and Digestive Disease at Imperial College London receives financial support from the National Institute of Health Research (NIHR) Imperial Biomedical Research Centre (BRC) based at Imperial College Healthcare NHS Trust and Imperial College London. This article is independent research funded by the NIHR BRC, and the views expressed in this publication are those of the authors and not necessarily those of the NHS, NIHR, or the Department of Health.

Author contributions

SK was responsible for initial design together with JRM and NK. DG and RW were responsible for sample collection overseen by SK. DA, DG and EM were responsible for the sample analyses overseen by NK and JRM. All were responsible for data collection and data analyses. The manuscript was written by all.

Data sharing – All sequences have been uploaded to EBI's ENA (<https://www.ebi.ac.uk/ena>).

References

- [1] M. Chakraborty, E. P. McGreal, and S. Kotecha, "Acute lung injury in preterm newborn infants: mechanisms and management.," *Paediatr. Respir. Rev.*, vol. 11, no. 3, pp. 162–70; quiz 170, Sep. 2010.
- [2] M. L. Beeton *et al.*, "Role of pulmonary infection in the development of chronic lung disease of prematurity.," *Eur. Respir. J.*, vol. 37, no. 6, pp. 1424–30, Jun. 2011.
- [3] P. L. Davies, O. B. Spiller, M. L. Beeton, N. C. Maxwell, E. Remold-O'Donnell, and S. Kotecha, "Relationship of proteinases and proteinase inhibitors with microbial presence in chronic lung disease of prematurity.," *Thorax*, vol. 65, no. 3, pp. 246–251, Mar. 2010.
- [4] J. Lowe *et al.*, "Association between pulmonary *Ureaplasma* colonization and bronchopulmonary dysplasia in preterm infants: updated systematic review and meta-analysis.," *Pediatr Infect Dis J*, vol. 33, no. 7, pp. 697–702, 2014.
- [5] S. Kotecha, R. Hodge, J. A. Schaber, R. Miralles, M. Silverman, and W. D. Grant, "Pulmonary *Ureaplasma urealyticum* Is Associated with the Development of Acute Lung Inflammation and Chronic Lung Disease in Preterm Infants.," *Pediatr. Res.*, vol. 55, no. 1, pp. 61–68, 2004.
- [6] V. Nair, P. Loganathan, and A. S. Soraisham, "Azithromycin and other macrolides for prevention of bronchopulmonary dysplasia: a systematic review and meta-analysis.," *Neonatology*, vol. 106, no. 4, pp. 337–47, 2014.
- [7] M. G. Dominguez-Bello *et al.*, "Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 26, pp. 11971–5, Jun. 2010.
- [8] D. J. Gallacher and S. Kotecha, "Respiratory Microbiome of New-Born Infants.," *Front. Pediatr.*, vol. 4, no. February, p. 10, Feb. 2016.
- [9] M. Pammi *et al.*, "Airway Microbiome and Development of Bronchopulmonary Dysplasia in Preterm Infants: A Systematic Review," *J. Pediatr.*, vol. 204, pp. 126-133.e2, 2018.
- [10] P. M. Mourani, J. K. Harris, M. K. Sontag, C. E. Robertson, and S. H. Abman, "Molecular

- identification of bacteria in tracheal aspirate fluid from mechanically ventilated preterm infants," *PLoS One*, vol. 6, no. 10, p. e25959, Jan. 2011.
- [11] C. V. Lal *et al.*, "The Airway Microbiome at Birth.," *Sci. Rep.*, vol. 6, p. 31023, Jan. 2016.
- [12] P. Lohmann *et al.*, "The airway microbiome of intubated premature infants: characteristics and changes that predict the development of bronchopulmonary dysplasia.," *Pediatr. Res.*, vol. 76, no. 3, pp. 294–301, Sep. 2014.
- [13] F. A. Stressmann *et al.*, "The use of culture-independent tools to characterize bacteria in endo-tracheal aspirates from pre-term infants at risk of bronchopulmonary dysplasia.," *J. Perinat. Med.*, vol. 38, no. 3, pp. 333–7, May 2010.
- [14] M. S. Payne *et al.*, "Molecular microbiological characterization of preterm neonates at risk of bronchopulmonary dysplasia," *Pediatr. Res.*, vol. 67, no. 4, pp. 412–418, 2010.
- [15] B. D. Wagner *et al.*, "Airway Microbial Community Turnover Differs by BPD Severity in Ventilated Preterm Infants," *PLoS One*, vol. 12, no. 1, p. e0170120, 2017.
- [16] S. Kotecha, L. Wilson, A. Wangoo, M. Silverman, and R. J. Shaw, "Increase in interleukin (IL)-1 beta and IL-6 in bronchoalveolar lavage fluid obtained from infants with chronic lung disease of prematurity.," *Pediatr. Res.*, vol. 40, no. 2, pp. 250–6, Aug. 1996.
- [17] J. de Blic *et al.*, "Bronchoalveolar lavage in children. ERS Task Force on bronchoalveolar lavage in children. European Respiratory Society.," *Eur. Respir. J.*, vol. 15, no. 1, pp. 217–31, Jan. 2000.
- [18] C. M. Liu *et al.*, "BactQuant: An enhanced broad-coverage bacterial quantitative real-time PCR assay," *BMC Microbiol.*, vol. 12, no. 1, p. 56, 2012.
- [19] J. J. Kozich, S. L. Westcott, N. T. Baxter, S. K. Highlander, and P. D. Schloss, "Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform.," *Appl. Environ. Microbiol.*, vol. 79, no. 17, pp. 5112–20, Sep. 2013.
- [20] P. D. Schloss *et al.*, "Introducing mothur: open-source, platform-independent, community-

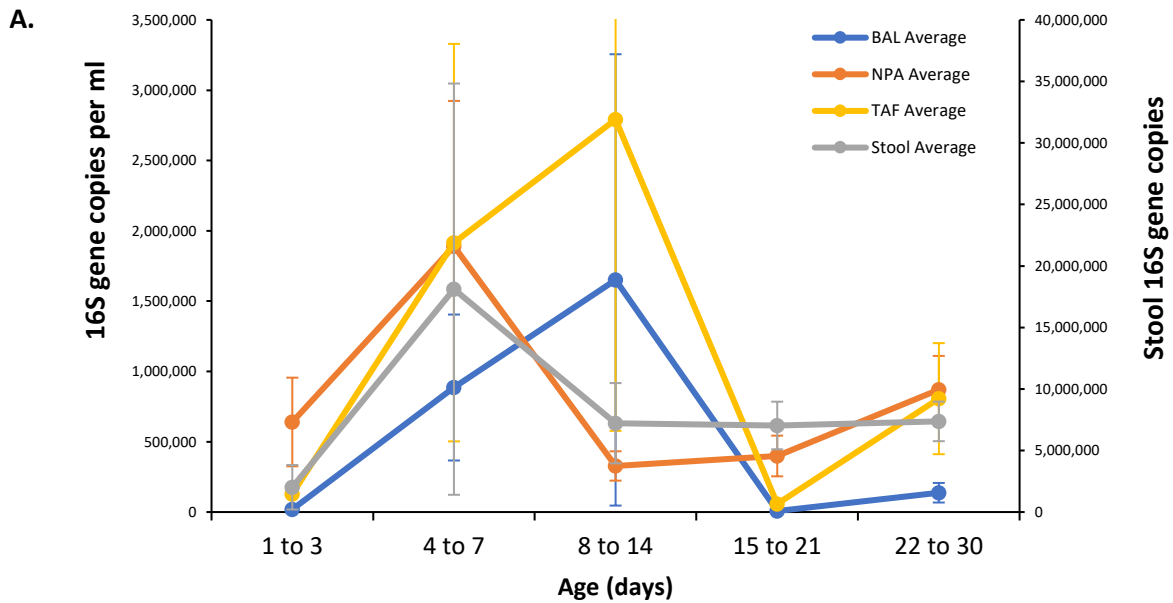
- supported software for describing and comparing microbial communities.," *Appl. Environ. Microbiol.*, vol. 75, no. 23, pp. 7537–7541, Dec. 2009.
- [21] J. R. Cole *et al.*, "Ribosomal Database Project: data and tools for high throughput rRNA analysis," *Nucleic Acids Res.*, vol. 42, no. D1, pp. D633–D642, Jan. 2014.
- [22] R Core Team, "R: Language and Environment for Statistical Computing." 2017.
- [23] P. J. McMurdie and S. Holmes, "Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data," *PLoS One*, vol. 8, no. 4, 2013.
- [24] J. Oksanen *et al.*, "CRAN - Package vegan," 2016. [Online]. Available: <https://cran.r-project.org/web/packages/vegan/index.html>. [Accessed: 17-May-2019].
- [25] D. H. Parks and R. G. Beiko, "Identifying biologically relevant differences between metagenomic communities," *Bioinformatics*, vol. 26, no. 6, pp. 715–721, Mar. 2010.
- [26] A. H. Jobe and E. Bancalari, "Bronchopulmonary dysplasia.," *Am. J. Respir. Crit. Care Med.*, vol. 163, no. 7, pp. 1723–9, Jun. 2001.
- [27] G. Biasucci, M. Rubini, S. Riboni, L. Morelli, E. Bessi, and C. Retetangos, "Mode of delivery affects the bacterial community in the newborn gut," *Early Hum. Dev.*, vol. 86, no. 1, pp. 13–15, 2010.
- [28] E. Rutayisire, K. Huang, Y. Liu, and F. Tao, "The mode of delivery affects the diversity and colonization pattern of the gut microbiota during the first year of infants' life: a systematic review," *BMC Gastroenterol.*, vol. 16, no. 1, p. 86, Dec. 2016.
- [29] R. P. Dickson *et al.*, "Bacterial Topography of the Healthy Human Lower Respiratory Tract," *MBio*, vol. 8, no. 1, pp. e02287-16, Mar. 2017.

Table 1 – Demographics of recruited preterm infants from two sites

	All infants	Infants recruited from UHW	Infants recruited from NBT	p-value (UHW vs NBT)
Number of infants	55	20 (36%)	35 (64%)	-
Number of samples				
Nasopharyngeal aspirates	539	145	394	-
Tracheal aspirate fluid	276	62	214	-
Bronchoalveolar lavage fluid	89	89	0	-
Stool	198	64	134	-
Total	1102	360	742	-
Sex (male)	36 (65%)	12 (60%)	24 (69%)	0.73
Gestation (weeks)	26.0 (24.7-27.5)	26.8 (25.3-29.4)	25.9 (24.7-26.6)	0.07
Birth weight (g)	764 (680-918)	835(695-1082)	746 (677-880)	0.16
Antenatal steroids	51/55 (93%)	18/20 (90%)	33/35 (94%)	0.18
Delivery mode (percent vaginal delivery)	29 (53%)	10 (50%)	19 (54%)	0.98
Multiple births	17 (31%)	6 (30%)	11 (31%)	1
Maternal antibiotic in labour	14 (25%)	4 (20%)	10 (29%)	0.70
Surfactant administration	55/55 (100%)	20/20 (100%)	35/35 (100%)	0.27
Chronic lung disease severity				
None	5 (9%)	4 (20%)	1 (3%)	0.02
Mild	4 (7%)	3 (15%)	1 (3%)	
Moderate	18 (33%)	3 (15%)	15 (43%)	
Severe/Died	28 (51%)	10 (50%)	18 (51%)	
Survival to discharge	47 (85%)	17 (85%)	30 (86%)	1
Ventilation days	17 (4-32.5)	25 (3.5-37.8)	17 (5-28.5)	0.55
Non-invasive ventilation days	47 (24.5-64.5)	38(19.8-53.5)	55(31.5-66)	0.14
Length of hospital stay (days)	93 (69.5-130)	104.5 (52.8-136.5)	93 (80-122)	0.95
Necrotising enterocolitis				
Grade 1	7 (13%)	4 (20%)	3 (9%)	1
Grade 2	1 (2%)	0 (0%)	1 (3%)	
Grade 3	6 (11%)	1 (5%)	5 (14%)	
Patent ductus arteriosus	40 (73%)	11 (55%)	29 (83%)	0.06
Initial breast milk	55 (100%)	20 (100%)	35 (100%)	1
Discharge breast milk	17 (31%)	6 (35%)	11 (37%)	1

Medians \pm interquartile ranges are shown unless otherwise stated. UHW – University Hospital Wales, NBT – North Bristol Trust. Necrotizing Enterocolitis Grades were categorized according to Bell’s staging criteria and CLD severity according to ref: 16.

Figure 1 - Total bacterial load in the airways and gut in preterm infants



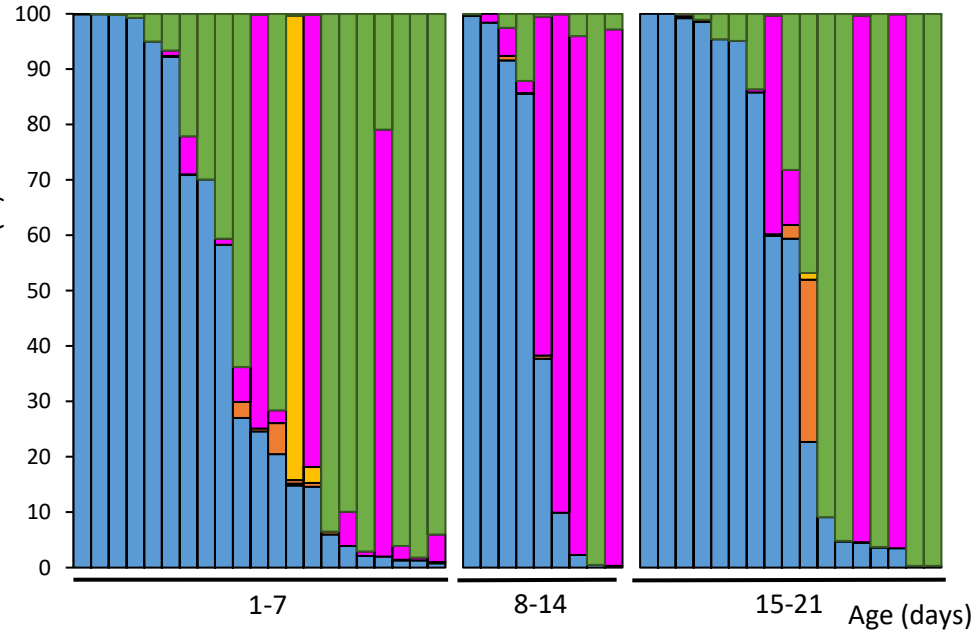
B.

	D1-3	D4-7	D8-14	D15-21	D22-28	Overall
NPA	9/133 (6.7%)	41/160 (25.6%)	20/82 (24.4%)	23/80 (28.8%)	33/84 (39.7%)	126/539 (23.4%)
TAF	1/94 (1.1%)	20/91 (21.9%)	9/35 (25.7%)	2/26 (7.7%)	15/30 (50%)	47/276 (16.9%)
BAL	6/31 (19.4%)	12/18 (66.7%)	5/15 (33.3%)	4/11 (36.4%)	7/14 (50%)	34/89 (38.2%)
Stool	5/14 (35.7%)	22/35 (62.8%)	35/50 (70.0%)	37/45 (82.2%)	46/54 (85.1%)	145/198 (73.2%)

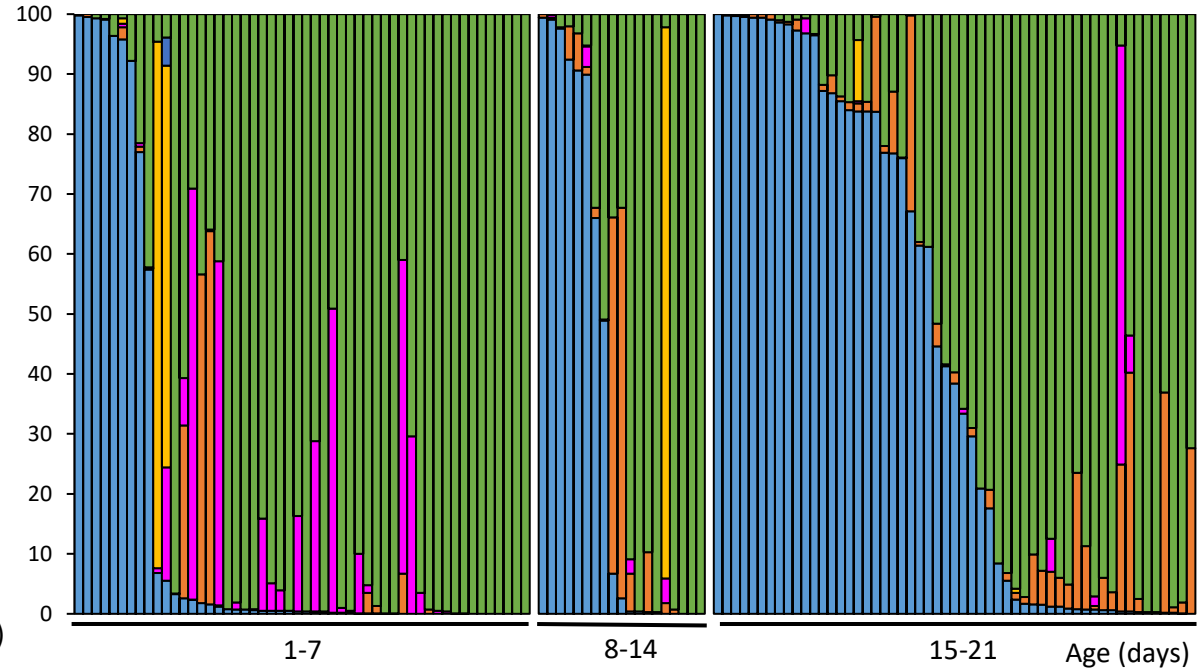
Figure 1. Total bacterial load in the gut and airways of preterm infants. (A) Change in total bacterial load over time shown for the four anatomical sites BAL (Bronchoalveolar Lavage), NPA (Nasopharyngeal Aspirate), TAF (Tracheal Aspirate Fluid) and stool samples. BAL, TAF and NPA samples were measured in copies of the 16S RNA gene per ml of supernatant. Stool samples were measured in copies of the 16S DNA gene per mg of stool. Mean \pm standard errors shown. (B) Table showing the number of samples available at each time point for each sample type, and the percent of the samples from which the 16 rRNA gene could be successfully amplified at each time point of each sample type.

Figure 2 – Microbiome analyses for each sample site shown by phyla

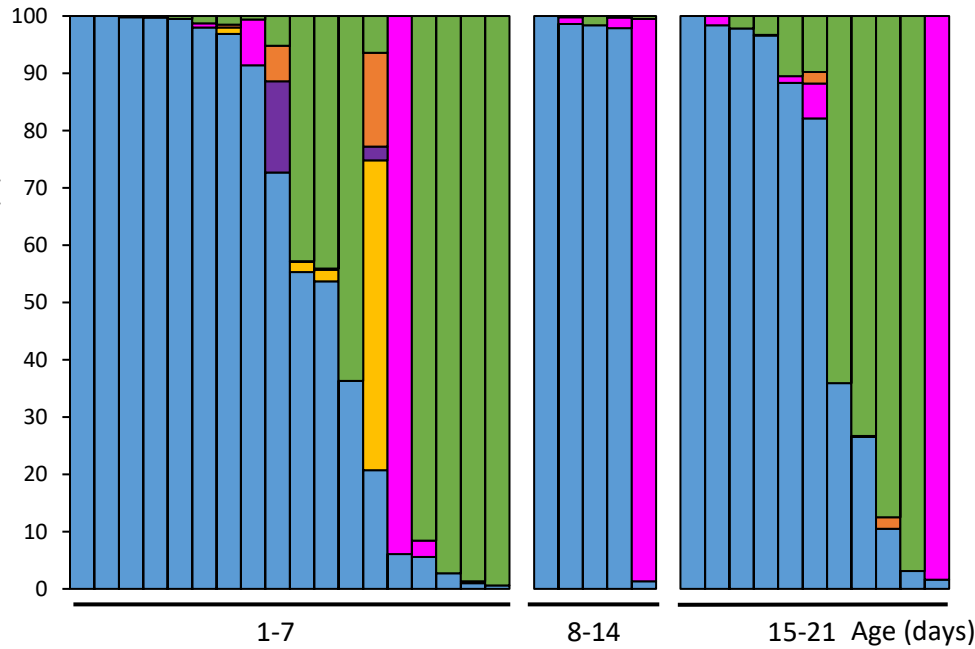
TAF



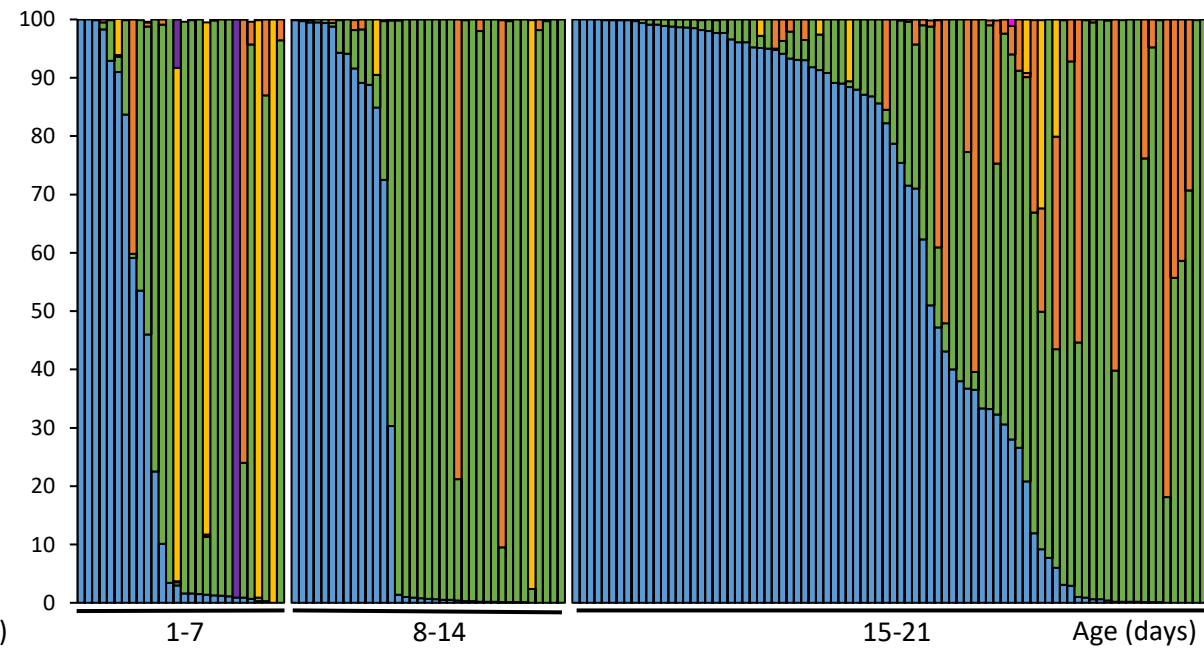
NPA



BAL



Stool



■ Proteobacteria
 ■ Actinobacteria
 ■ Tenericutes
 ■ Bacteroidetes
 ■ Fusobacteria
 ■ Planctomycetes
 ■ Firmicutes

Figure 2. Microbiome analyses for each sample site shown by phyla. The relative abundance of the predominant phyla of the preterm infant is shown for each time period. Each bar represents an individual sample with relative abundance shown after all OTUs from each phylum were combined and averaged. Abbreviations: BAL – Bronchoalveolar Lavage, NPA – Nasopharyngeal Aspirate, TAF – Tracheal Aspirate Fluid.

Figure 3 – Microbiome analyses for each sample site shown at genus level

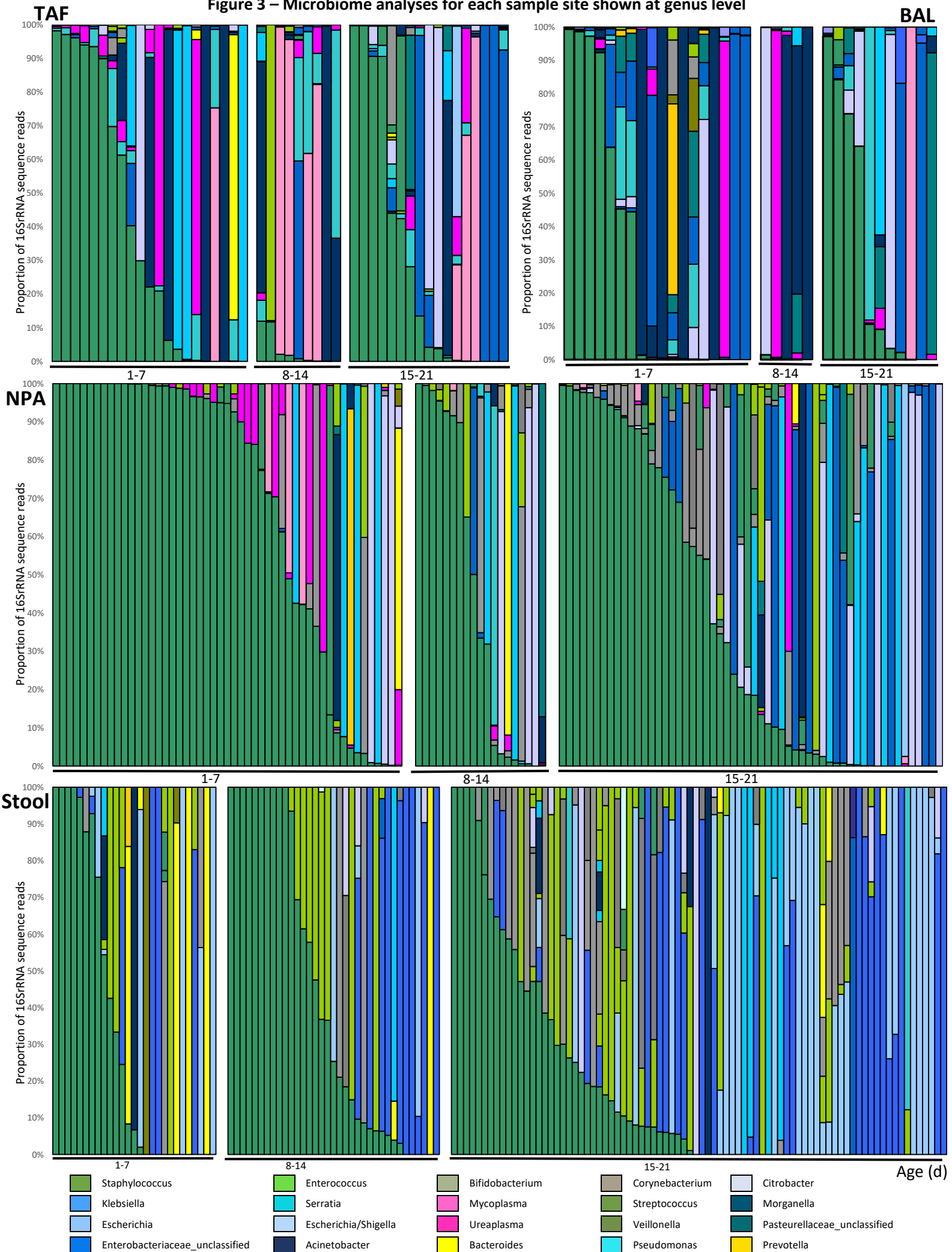


Figure 3. Microbiome analyses for each sample site shown at genus level.

The relative abundance of the predominant genera of the preterm infant is shown for each time period. Each bar represents an individual sample with relative abundance shown after all OTUs from each genus were combined and averaged. Abbreviations: BAL – Bronchoalveolar Lavage, NPA – Nasopharyngeal Aspirate, TAF – Tracheal Aspirate Fluid.

Figure 4 – Microbiome community structure between sample sites

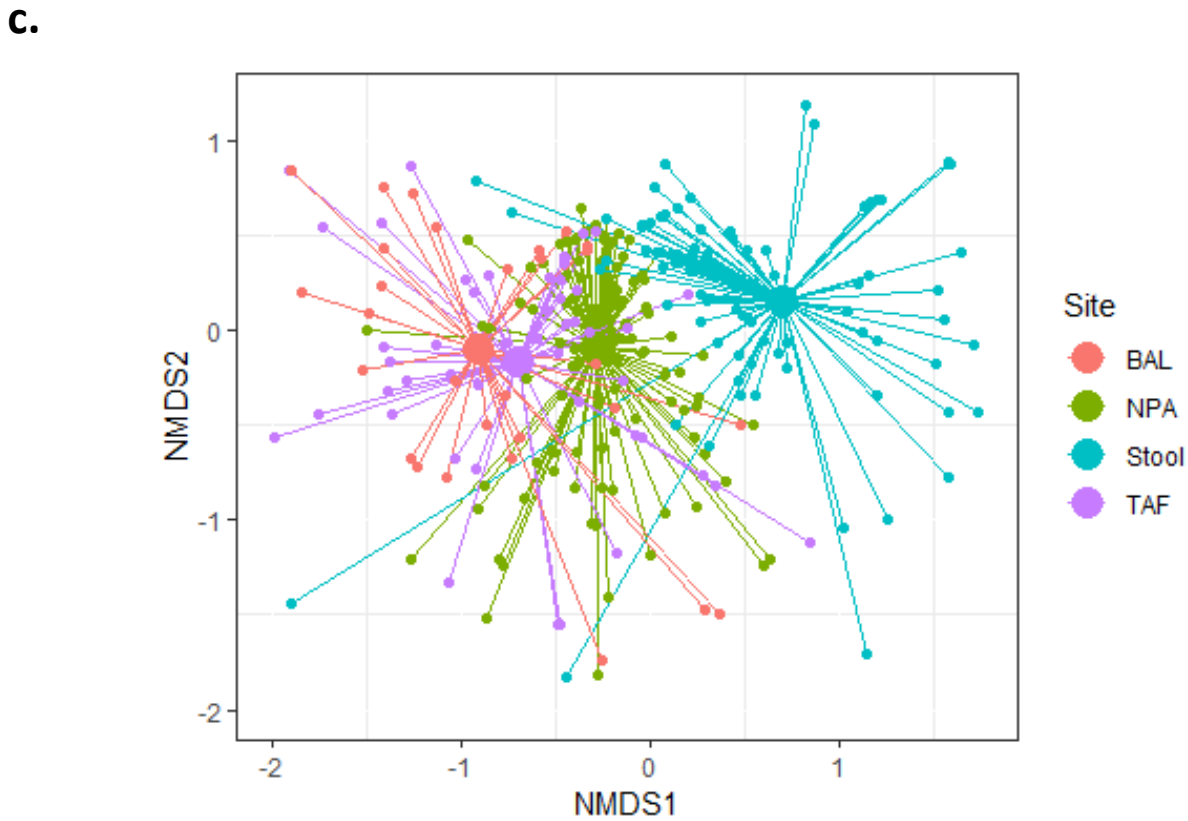
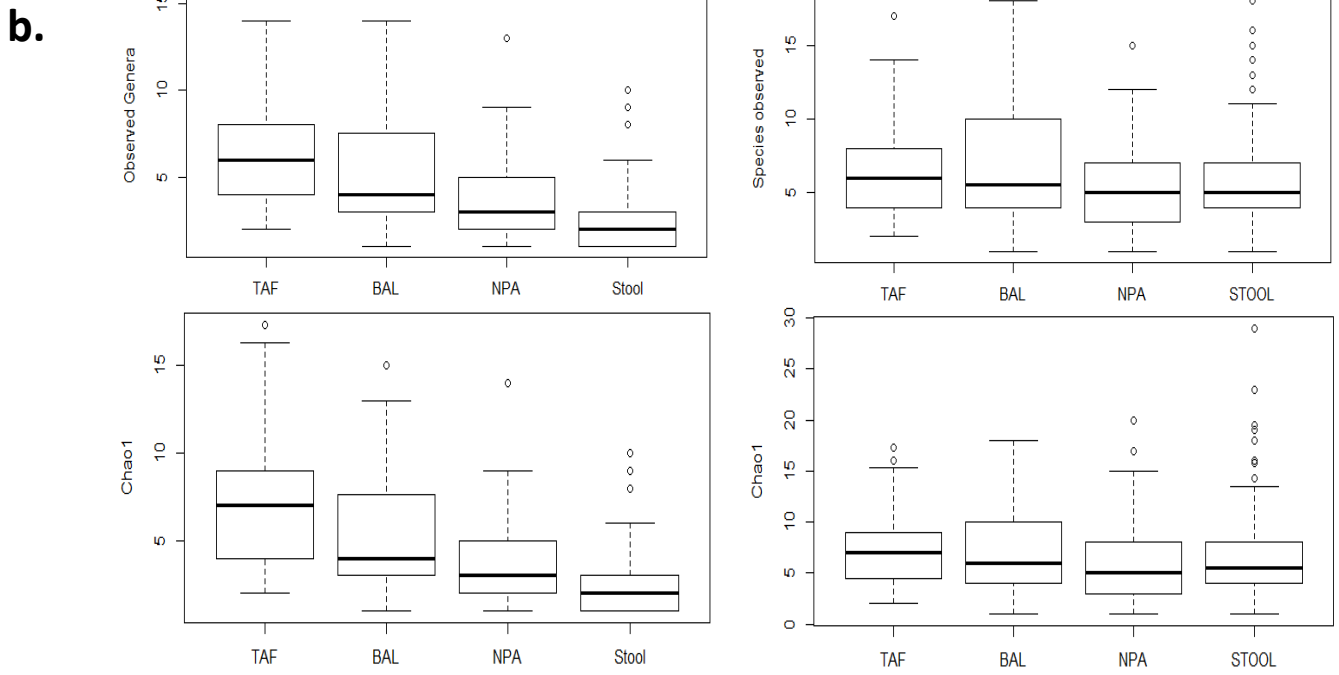
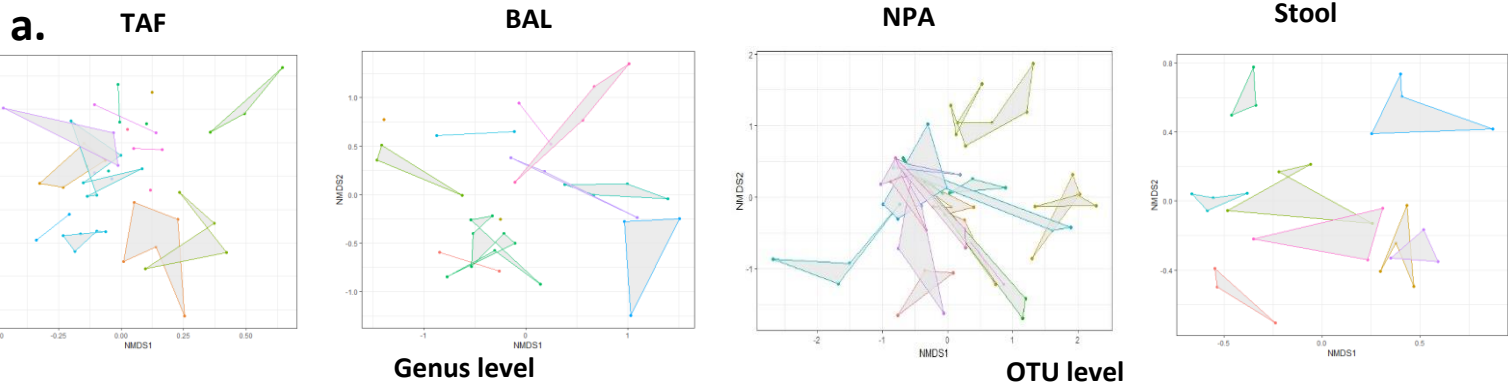
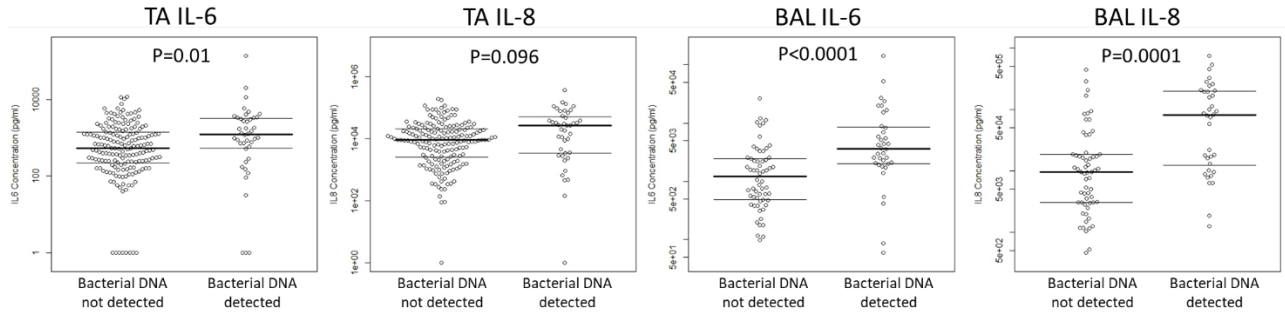


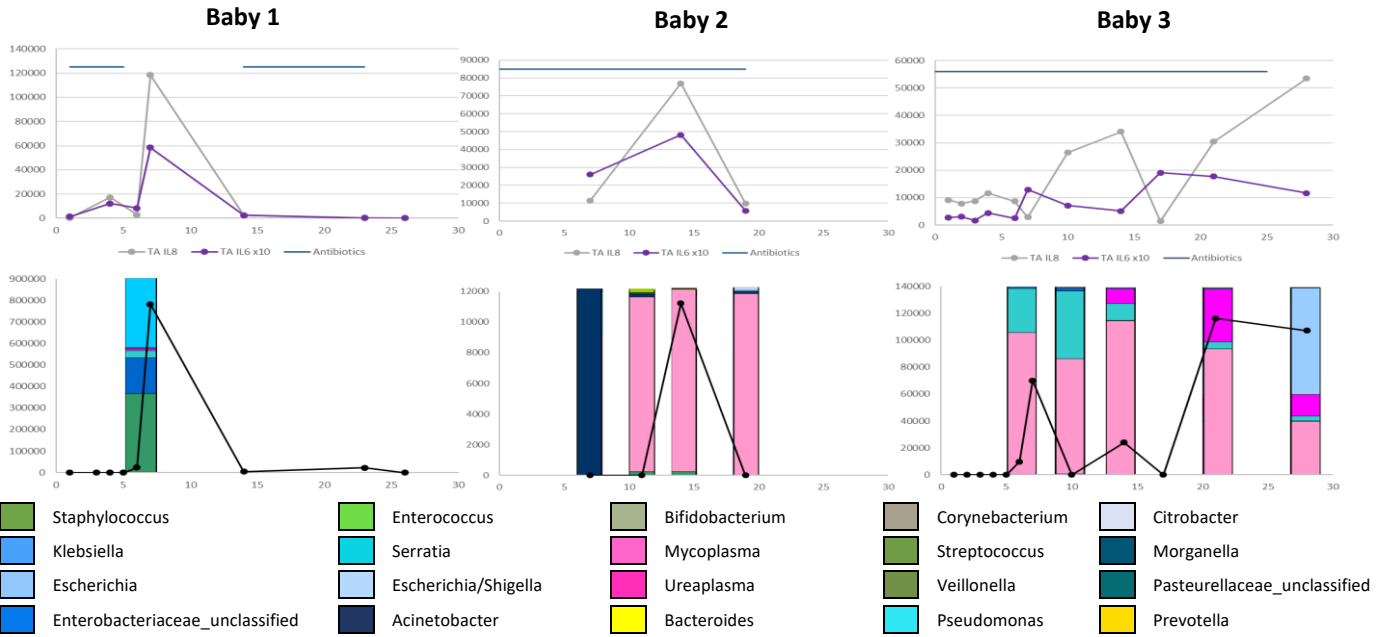
Figure 4. Microbiome community structure between sample sites. (A) NMDS plots for individual babies for each sample site shown. Each coloured polygon joins all samples from an individual infant. (B) Alpha diversity of observed and Chao1 shown at genus and OTU level. (C) NMDS plot showing the relationship between the bacterial community between the anatomical sites. PERMANOVA showed significant differences between the BAL vs TAF, BAL vs NPA, BAL vs Stool, TAF vs NPA, TAF vs Stool, NPA vs Stool using Bonferroni correction, all $p < 0.01$. Abbreviations: BAL – Bronchoalveolar Lavage, NPA – Nasopharyngeal Aspirate, TAF – Tracheal Aspirate Fluid.

Figure 5 - Association of the lower airway microbiome with markers of inflammatory

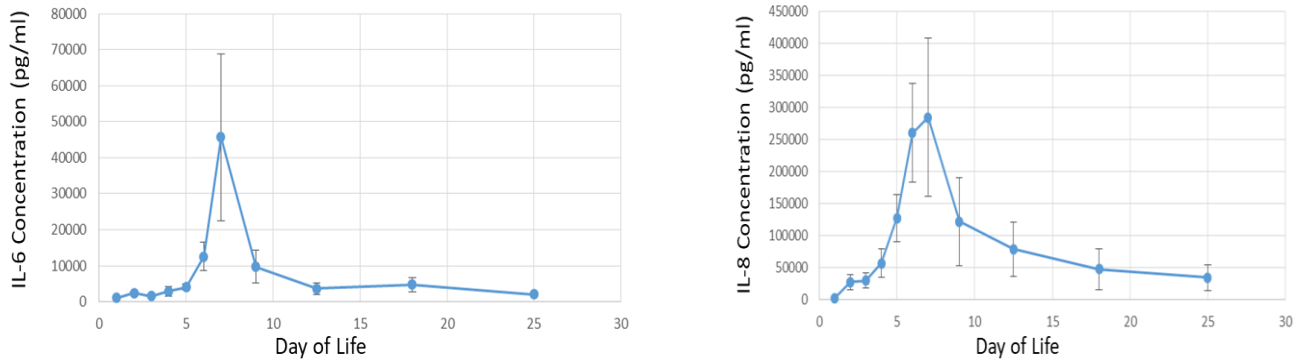
a.



b.



c.



d.

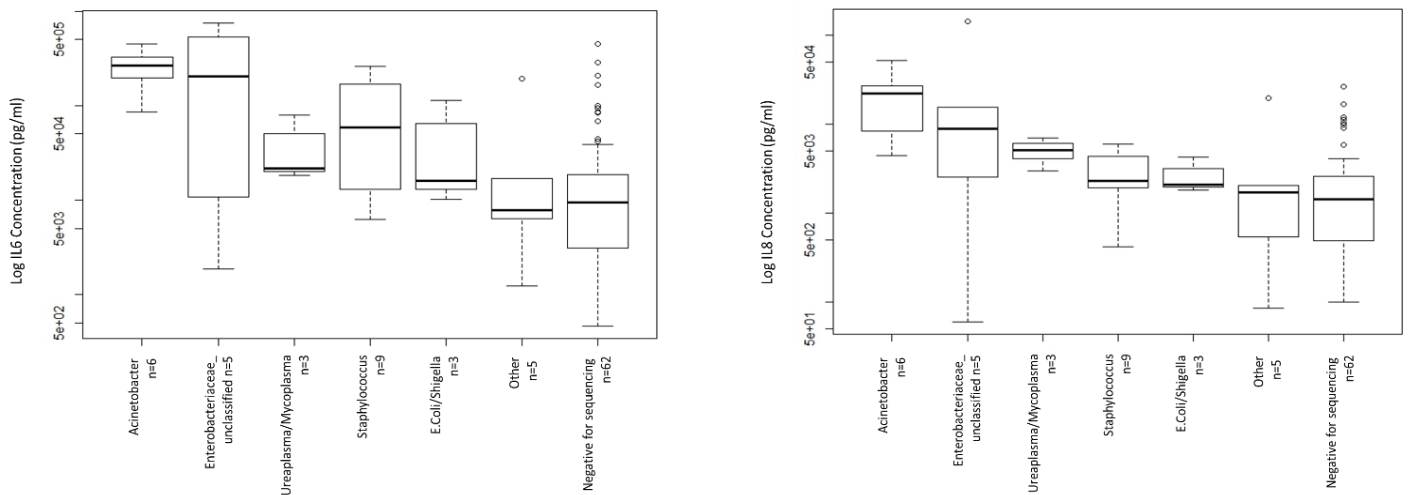


Figure 5. Association of the lower airway microbiome with markers of inflammatory. (A) Log_{10} scale for interleukin-6 (IL-6) and interleukin-8 (IL-8) in tracheal aspirate fluid (TAF) and bronchoalveolar lavage (BAL) fluid samples for samples where bacterial sequencing was positive or negative. (B) Shows the relationship between IL-6 (grey), IL-8 (purple), antibiotic treatment (blue) in upper graph and bacterial load (black) and bars showing genus constituents in three individual babies (Baby 1 – 3). Generally, parallel increases and decreases were observed for all markers despite antibiotic treatment. (C) Averaged IL-6 and IL-8 in TAF samples shown for each time point regardless of bacterial presence showing a peak of both cytokines at day 7 as previously shown^{2,3,19}. Mean \pm SEM shown. (D) Association is shown between the presence of IL-6 and IL-8 in BAL samples with specific OTUs showing increased IL-6 and IL-8 if an OTU was identified when compared to the negative for sequencing group.

Dissimilarity of the gut-lung axis and dysbiosis of the lower airways in ventilated preterm infants –

Online Supplement

¹Gallacher D*, ¹Mitchell E*, ²Alber D, ³Wach R, ²Klein N, ^{4,5}Marchesi JR, ¹Kotecha S.

¹Department of Child Health, School of Medicine, Cardiff University, Cardiff, UK.

²Institute of Child Health, University College London, London, UK.

³Neonatal Unit, North Bristol NHS Trust, Bristol, UK.

⁴School of Biosciences, Cardiff University, Cardiff, UK.

⁵Division of Integrative Systems Medicine and Digestive Disease, Imperial College London, London, UK

*Joint first authors

Address for Correspondence:

Professor Sailesh Kotecha

Department of Child Health

School of Medicine

Cardiff University

Heath Park

Cardiff CF14 4XN

Email: KotechaS@cardiff.ac.uk

Methods

Collection of Clinical Samples

BAL fluid was collected as previously described¹⁻³. Briefly, with the infant supine, with the head turned to the left, a Fr6 catheter was advanced through the ET tube until resistance was felt and 1ml/kg of sterile saline instilled. The fluid was then aspirated. After allowing the baby to recover, the procedure was repeated, and the samples pooled. The BAL samples were centrifuged at 10,000g for 10 mins within 60 minutes to separate the cells and supernatant before storage at -80°C.

TA samples were collected at the time of clinically indicated routine endotracheal tube suctioning and the collection catheter flushed with 2ml of sterile saline. Nasopharyngeal aspirates (NPA) samples were collected by suctioning mucous from the nasopharynx and flushing the collected sample with 2 ml of sterile saline. Collected TAF and NPA samples were initially frozen at -20°C then transferred without thawing to -80°C. TA, BAL and NPA samples were obtained daily for the first week of life then twice weekly until 28 days of age or until extubation, whichever occurred earlier. Stool samples were collected twice weekly until 28 days of age. Control samples were obtained flushing sterile saline through the same suction catheters used to collect samples.

DNA Extraction

TAF and NPA samples were thawed on ice and centrifuged at 10,000g for 10 mins to separate the cell pellets and supernatant. Cell pellets from all respiratory samples were resuspended in 0.5ml of DNA extraction buffer containing 5M guanidine thiocyanate. The samples were bead beaten with 0.1mm zirconia beads for 3 x 30 seconds at 6.5m/s using a FastPrep 24 device (MP Biomedicals, CA, USA) and placed on ice for 5 mins between each episode. Samples were then loaded into a Maxwell 16 automated DNA extraction device (Promega UK, Southampton, UK) using the tissue DNA extraction kit. DNA was eluted into 300ul of the supplied elution buffer. Stool samples were thawed on ice for 1 hour before being weighed before homogenising. The homogenate was then centrifuged, and DNA extracted from the faecal pellet using the Qiagen Stool Mini Kit (Qiagen,

Hildenberg, Germany) as per manufacturer's instructions, with the following optimizations. 2 ml of InhibitEX buffer was added to the faecal pellet and vortexed continuously for 1 minute, or until the sample was fully homogenous. 1.5 ml of this solution was then transferred into a 2 ml MPBiomedicals Lysing Matrix tube E, which contained 1.4 ceramic spheres, 0.1 mm silica spheres and one 4 mm glass bead. The tubes were then homogenised in the FastPrep 24 instrument for 60 seconds at 5.0 m/s. This process was repeated 3 times with 5-minute incubations on ice between each beating. The samples were then heated for 5 minutes at 70°C in a heat block. The heated samples were then vortexed for 15 seconds, before being centrifuged at 20,000 x g for 1 minute. 200 µl of the supernatant was then taken forward into the remainder of the protocol.

Quantification of Bacterial Load

Total bacterial load was quantified by a qPCR assay based on published protocols⁴ amplifying the V3-V4 region of the 16S rRNA bacterial gene. Primers sequences used were: forward 5'- CCTACGGDGGCWGCA-3' and reverse: 5'- GGACTACHVGGTMTCTAATC -3'. The TaqMan probe sequence was (6FAM) 5'-CAGCAGCCGCGGTA-3' (MGBNFQ). 2.5 µL of DNA was added to a total reaction of 10 µL. The thermal cycling protocol was heating for 3 min at 50°C for UNG treatment, then 10 min at 95°C for Taq activation followed by 40 cycles of 15s at 95°C for denaturation and 1 min at 60°C for annealing and extension. A standard curve was generated in each assay using serially diluted *E. Coli* genomic DNA. DNA used generating standard curves was extracted from pure growths of *E. Coli* organisms with DNA quantification using a Cubit analyser. Samples, standards and controls were run in triplicate and a set cT threshold maintained for each run of the experiment.

Sequencing

The V3-V4 region of the 16S rRNA gene was amplified from each DNA extract using the Taq Core PCR kit (Qiagen, Hilden, Germany) with forward and reverse primers at 0.5µM each. Barcoded primers were designed according to Kozich *et al.*⁵ to include Illumina adapter, an 8 nucleotide barcode

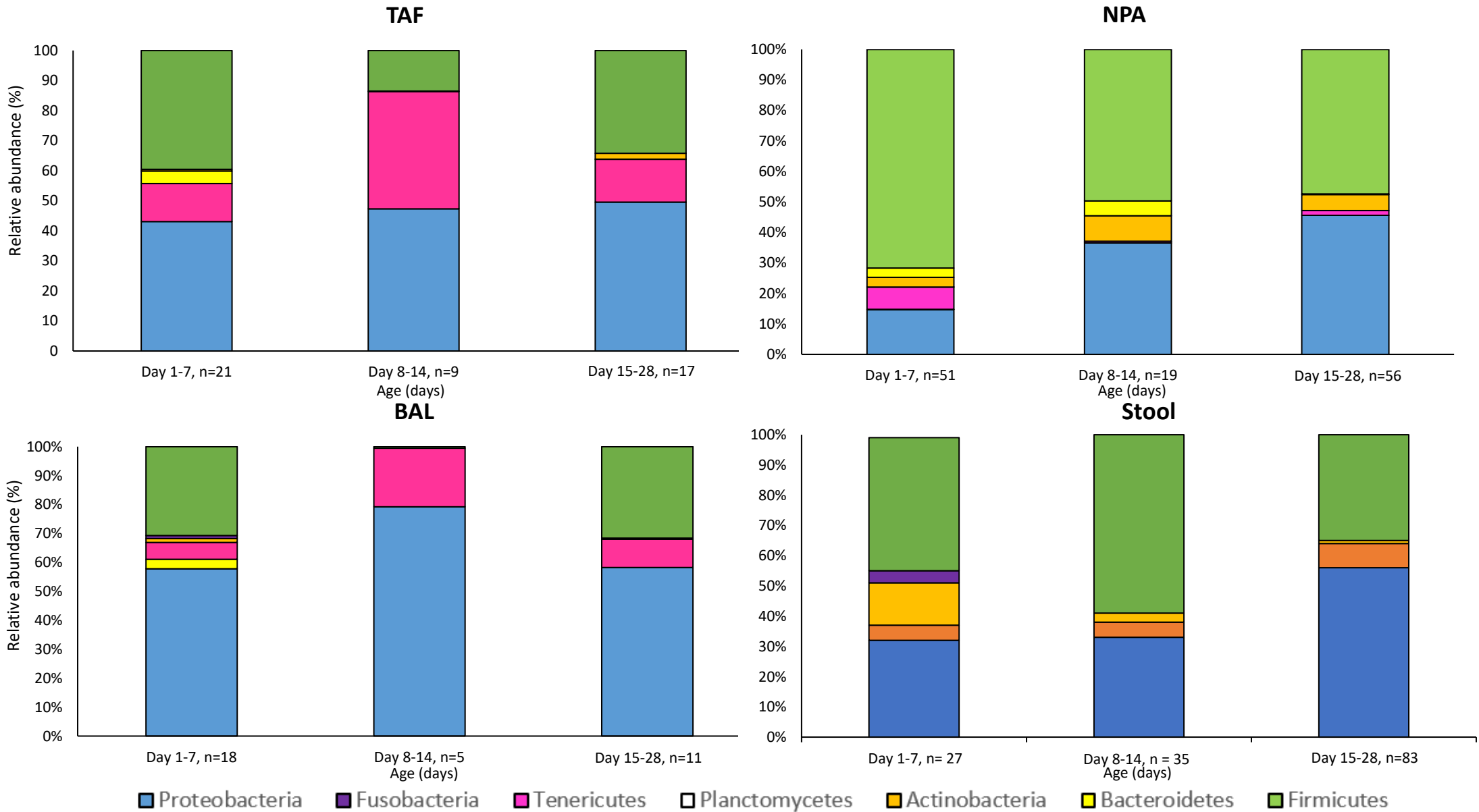
sequence, a 10 nucleotide pad sequence, a 2 nucleotide linker, and a gene-specific primer:: 341F-CCTACGGGNGGCWGCAG or 805R-GACTACHVGGGTATCTAATCC. (Sigma-Aldrich, Dorset, UK) and used to amplify each DNA sample using different barcode combination. The cycling conditions were as follows: initial denaturation at 95°C for 3min, 30 cycles of 95°C for 30 sec, 54°C for 30 sec, 72°C for 10 min and a final extension at 72°C for 10 min. Purification of the PCR products was achieved using AMPure beads (Beckman Coulter (UK), High Wycombe, UK).

Purified PCR product were quantified using the Qubit dsDNA High Sensitivity Assay Kit (ThermoFisher, Leicestershire, UK), and pooled to create a library with approximately equal concentrations of 16S rRNA amplicons from each sample. Separate libraries were prepared for TAF, BAL, NPA and stool samples. Pooled libraries were quality and quantity checked using the The High Sensitivity D1000 ScreenTape assay (Agilent Technologies, CA, USA) and a NEBNext library quantification kit (New England Biolabs). Each library spiked with 10% PhiX was sequenced on an Illumina MiSeq using the Reagent Kit V2 with 500 cycles (Illumina UK, Cambridge, UK) and custom primers as previously described⁵.

Data Analysis

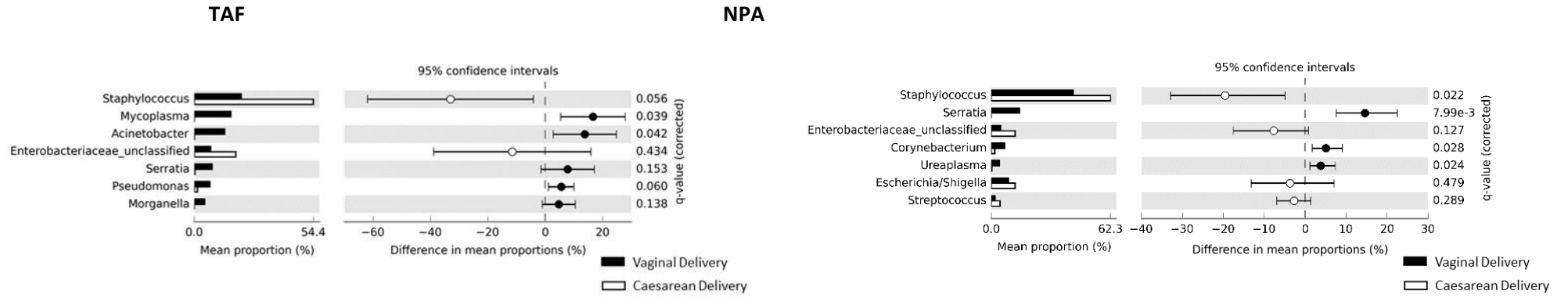
STAMP (Statistical Analysis of Metagenomic Profiles) software was used to produce error bar plots using White's test with Benjamini-Hochberg False Discovery Rate method used to correct for multiple comparisons.

Online Figure 1 – Averaged microbiome data for each sample site shown by phyla

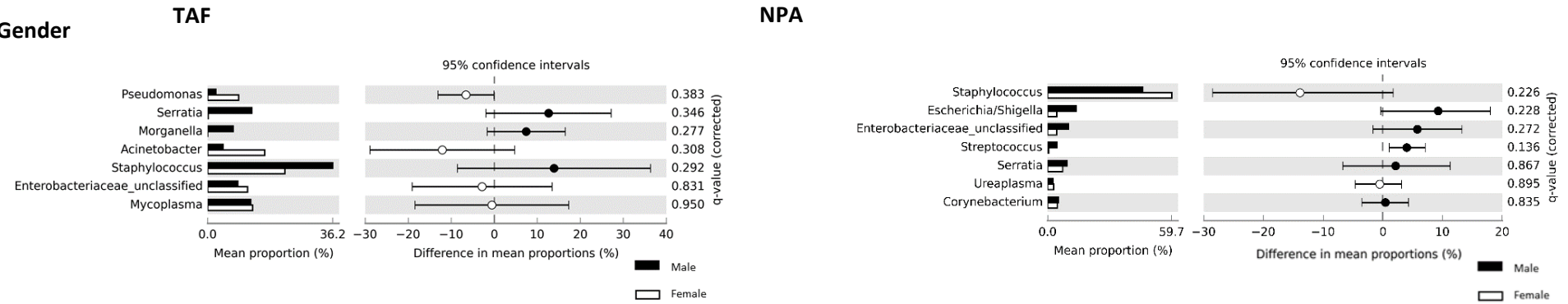


Online Figure 2 – The effect of mode of delivery, gender and centre of recruitment on the lung microbiome

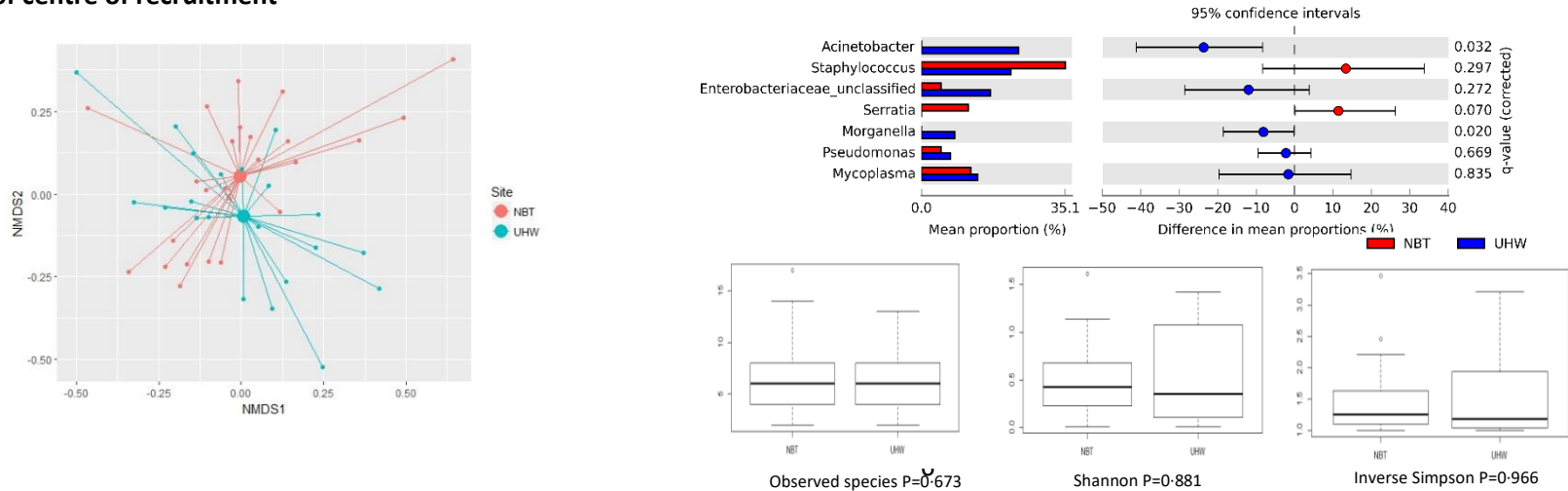
a. Mode of delivery



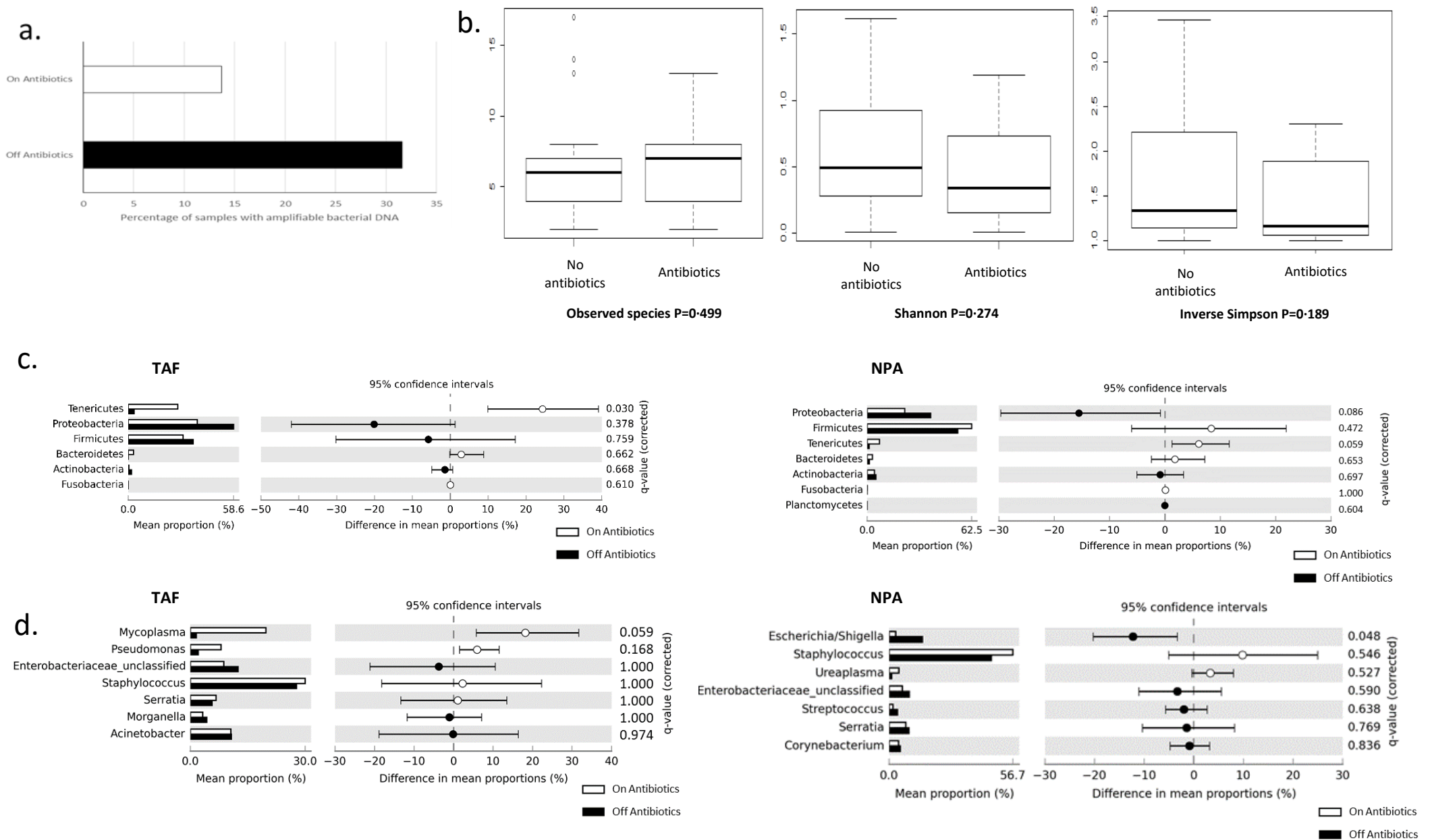
b. Gender



c. Effect of centre of recruitment

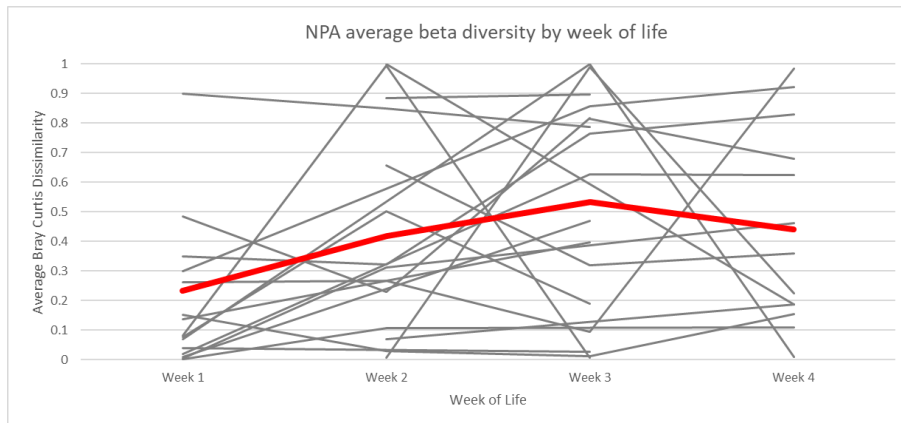


Online Figure 3 – The effect of antibiotics on tracheal aspirate fluid (TAF) and nasopharyngeal aspiration (NPA)

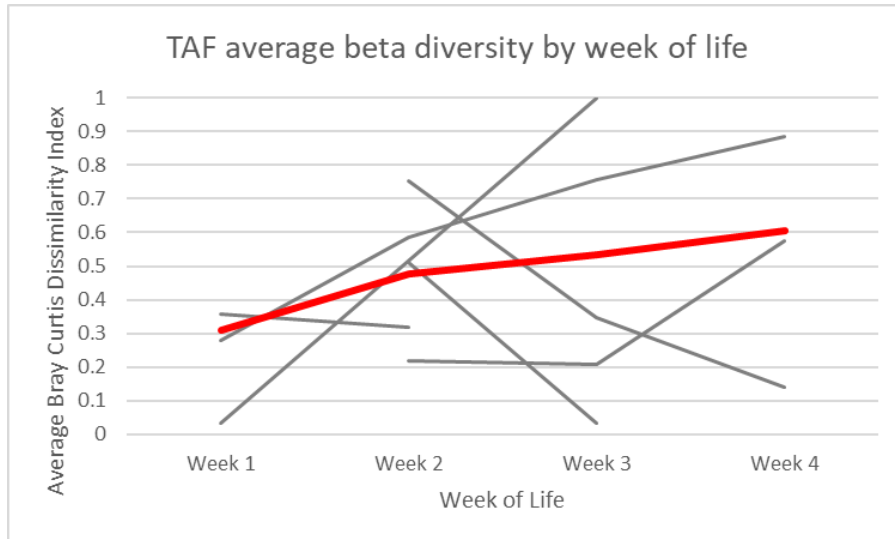


Online Figure 4: Beta-Diversity Changes over time

A - NPA



B - TAF



Legends

Online Figure 1. Averaged microbiome data for each sample site shown by phyla. The relative abundance of each phylum averaged for all samples at each time-point for each sample type is shown. Abbreviations: BAL – Bronchoalveolar Lavage, NPA – Nasopharyngeal Aspirate, TAF – Tracheal Aspirate Fluid.

Online Figure 2. The effect of mode of delivery, gender and centre of recruitment on the upper (NPA) and lower airway (TAF) microbiome. Shown are the effect of the (A) mode of delivery, (B) Gender and (C) centre from which sample was collected on bacterial OTUs in upper (NPA) and lower (TAF) airway samples. In addition, alpha and beta diversity shown for sample site (red NBT, North Bristol Trust, UHW, University Hospitals of Wales).

Online Figure 3. The effect of antibiotics on tracheal aspirate fluid (TAF) and nasopharyngeal aspirate (NPA) samples. (A) percentage of samples which had sufficient for bacterial 16s rRNA gene sequencing, whilst on or off antibiotics. (B) Alpha diversity shown for TAF samples when taken when the infant was or was not on antibiotics. (Alpha diversity for NPA similarly similar between the antibiotic and no antibiotic samples.) (C) shows the effect of antibiotics on the lung bacterial community at the phylum level and (D) at genus level, [for the most abundant 7 genera](#), for samples taken whilst on or off antibiotic treatment.

Online Figure 4: Beta-diversity changes over time

[Changes in Bray-Curtis Dissimilarity index in samples from individual infants by week of life. Bray-Curtis index calculated between each consecutively collected successfully sequenced sample from each infant. The median day between the samples used to allocate week of life. For weeks with more than 1 beta diversity value calculated the average value was taken. Grey lines represent the](#)

results for individual infants and the bold red line shows the average from each week. A= NPA(Kruskal Wallis $p = 0.098$) , B= TAF(Kruskal Wallis $p = 0.46$).

References

1. Kotecha, S., Wilson, L., Wangoo, A., Silverman, M. & Shaw, R. J. Increase in interleukin (IL)-1 beta and IL-6 in bronchoalveolar lavage fluid obtained from infants with chronic lung disease of prematurity. *Pediatr. Res.* **40**, 250–6 (1996).
2. Beeton, M. L. *et al.* Role of pulmonary infection in the development of chronic lung disease of prematurity. *Eur. Respir. J.* **37**, 1424–30 (2011).
3. de Blic, J. *et al.* Bronchoalveolar lavage in children. ERS Task Force on bronchoalveolar lavage in children. European Respiratory Society. *Eur. Respir. J.* **15**, 217–31 (2000).
4. Liu, C. M. *et al.* BactQuant: An enhanced broad-coverage bacterial quantitative real-time PCR assay. *BMC Microbiol.* **12**, 56 (2012).
5. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–20 (2013).