nature portfolio

Peer Review File

Endotracheal virome and metatranscriptome of preterm neonates at birth in relation with progression to bronchopulmonary dysplasia

Corresponding Author: Dr Gregory Destras

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Interesting and especially novel work. Authors should be commended for the idea and the deep level Omic analysis. This should be considered a FIH study preliminary to others. I do have some relatively minor comments that need to be addressed before publication might be granted:

1. BPD has been diagnosed with an old (2001) definition which is considered obsolete by INC and following consensus. Authors should reclassify cases using the Jensen definition, in order to increase reproducibility

2. Tracheal aspirates are well known to be poor samples unable to represent the alveolar biology (see Dargaville P Am J Resp Crit Care Med 1999 amongst the others). Conversely the injection of a minimal amount of normal saline, under controlled conditions allows to better retrieve epithelial lining fluid (see again Dargaville and others), and this has been recognised almost 25 years ago by the first European Respiratory Society recommendations of BAL in neonates and infants (de Blic. Eur Resp J 2000). The injection of normal saline (so calle dNB-BAL or mini-BAL) may be viewed as something invasive and this is totally wrong since many classical studies from the 80s have demonstrated its safety from different point of views. However, the injection does introduce variability and needs standardisation in order to have comparable data (there're several examples of this standardisation e.g. De Luca D. Intensive Care Med 2008). Now obviously it is not possible to improve the quality of specimens retrospectively. However authors should
- acknowledge the limitation due to he use of tracheal aspirates
- comment on a dual sample strategy (like Dargaville 's); i.e. if performed serially, tracheal aspirate and NB-BAL can give info on teh airways and alveolar district contributing to better localise the virome you are analysing.

3. "The more you know he more you ask". Being a pilot study this nice work cannot answer all teh questions. However some news to be highlighted and commented for future works:
- what is the effect of GA on virome? Too small sample size to tell and this can be important at the limit of viability
- was there an effect of maternal virome? this would need to be studied by associating maternal vaginal swab analyses
- what was the effect of maturation (i.e. postmentrual age)? This is crucial as it means studying the change of virome profile overtime and cannot be done with a single sampling study
- is there any link between virome and lung aeration or gas exchange metrics, or markers of alveolarization (SPD, IGF1, VEGF) or inflammation (IL6-8,m sPLA2 etc)? That would associate a given virome to a different BPD phenotype

Reviewer #2

(Remarks to the Author)
This is an interesting prospective study of endotracheal aspirates of virome, bacterial microbiome and transcriptome 31 very preterm neonates (<30 wks GA) collected within 24 hours of birth. The investigators report a low-abundance individualized virome, including mainly unclassified Caudoviricetes bacteriophages (71% of neonates). Multi-Omics Factor Analysis identified four microbial endotypes; one protective” endotype with low microbial load was associated with reduced BPD risk (29% vs.79%, p = 0.03), whereas high-risk clusters exhibited increased Pseudomonas load, elevated bacterial transcriptional activity, and greater viral diversity.

Strengths of this study: Virome study- very few virome studies in preterm neonates. Important study associating early life airway virome and microbiome to BPD progression. MOFA analysis using virome, bacteriome and transcriptome for endotypes
Limitations: Just one time point and how the virome changes longitudinally will be very important. Small sample size (n=31). Simultaneous evaluation of the gut virome/microbiome might throw more light on early life virome and subsequent morbidities in preterm infants.

Reviewer #3

(Remarks to the Author)
Setting aside the fact that this paper should not have been sent out for review with a typo in the title, I was intrigued by what the authors set out to accomplish in this manuscript.
Over the authors have made extensive strides to perform a cutting-edge mult-omic analysis of tracheal aspirates from premature infants and overall they have succeeded admirably. However, a fatal limitation still remains.
Primarily Limitation:
The principal limitation of this manuscript is the exceptionally low biomass of the supposed viral sequences. Given the current understanding of the field that low-biomass of previously reported “pregnancy” or “near-birth” microbiomes are actually representative of contamination and sequencing error, the burden of proof that these ultra-low biomass viral sequencing are representative of actual viral DNA falls on the authors. As the manuscript stands, they have failed to meet this critical barrier. Considerable additional positive and negative controls and methodological demonstration of the appropriate Ct thresholds representative of actual sequencing are required. As it stands the evidence presented in the manuscript suggests that what is actually being characterized is a mixture of contamination and sequencing error. Without better controls the use of post-hoc computational decontamination is insufficient to alleviate these significant methodological issues. While some of this appears to be included in the viral metagenomic sequencing section of the methods, showing this data in the supplement, and showing the computational decontamination data, is a requisite.
Major Corrections:
1. The authors acknowledge the absence of "sampling device flush controls" (lines 382-383) a s a limitation. Given the low-biomass nature of the samples and the potential for contamination during collection, this is a significant methodological omission. Please discuss i n the limitations section how the other decontamination steps (e.g., NTCs, PBS controls) might mitigate the impact of this specific missing control, or if a more robust justification for ruling out sampling device contamination can be provided.
2. Similarly while limitations in RNA detection sensitivity or the presence of non-viable bacteria" (Lines 453-454) are plausible, this explanation is insufficient and further evidence and expanded discussion is indicated.
3. The Kruskal-Wallis result in the summary at line 26-27 does not match what is reported in Table 2. Please clarify to limit potential confusion.
4. Line 515, includes a placeholder instead of the SRA number. Due to this the data could not be access and independently verified. Full availability of raw data is the standard of the field and must be corrected.
5. The manuscript states that "endotracheal virome profiles were found to be unable to distinguish between infants according to birth mode" (line 277), yet immediately after, it states that "bacteriophage abundance and viral richness at birth were significantly higher in vaginally delivered infants" (lines 278-279), referencing Figure 3E. Please clarify this apparent contradiction. If the overall virome profiles (e.g., as visualized by NMDS in Figure 3D) do not show clear separation by birth mode, but specific univariate analysis do (as I understand these graphs), then this should be more clearly and explicitly stated.
6. The authors utilize MOFA, when the significantly improved MOFA2 has been available for several years. Please update the analysis to MOFA2: https://github.com/bioFAM/MOFA2.
7. In Table 2, Kruskal-Wallis was used to determine significance. Best practices is to report adjusted and unadjusted P-values. Also, the methods used to perform this adjustment should be clearer.
8. The y-axis ranges for "Virome (VDR)" and "Bacteriophages (VDR)" in Figure 3A, and for "VDR" in Figure 4C, are truncated at 1.00. However, the text mentions values outside of these range, and correcting these axis would improve the data presentation.
9. While R packages and versions are provided for the bioinformatics analyses (e.g., MOFA, UMAP, clustering), providing more specific details on the bioinformatic analysis, or better yet, access to the analysis code on GitHub or other relatively permanent access website would greatly facilitate reproducibility of the analysis.
Minor Corrections:
1. Figures 5C, 6B and 6C are not called out in the text.
2. Figure 2B and 2C appear to be swapped in the figure versus the legend.
3. Use of the term “protective” is confusing, we would suggest less causative oriented language, such as “lower risk-associated” since this analysis can in no way imply causality.
4. The interpretation of the asterisks in Table 2, indicating "one vs all-
cluster comparison (Mann-Whitney test, pvalue‹0.05)", is ambiguous. The asterisks appear on specific median (QR) values within the table, implying that these specific clusters are significantly different from all other clusters combined. This is an unusual way to present post-hoc test results. Please clarify this interpretation in the table legend or footnote, and ensure the statistical methodology for these comparisons is fully described in the methods.
5. The "protective" endotype (C1) is characterized by a "low-abundant microbiome" (line 339), but the text later states that "higher loads of Staphylococcus sp. were observed in the protective 'cluster'" (line 344). While Table 2 shows a higher median Staphylococcus load in C1 compared to other clusters, the overall bacterial DNA load in C1 is lower than in C4. Please clarify this phrasing to avoid confusion about whether "higher loads" refers to absolute abundance or relative abundance within a low-biomass context, especially given the overall low abundance characteristic of C1.
6. Given the studies cross-sectional design, while acknowledged, the need for further longitudinal sampling to actually

demonstrate these findings should be further emphasized.
7. Reference 20 is related to MOFA not Papillomaviruses (line 421). Please correct.
8. There is another type at line 490.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
authors successfully addressed my comments and the paper is now ready for publication

Reviewer #3

(Remarks to the Author)
The authors have made a strong effort to address my concerns. Without sequencing standards proving they are actually able to sequence such ultra-low-abundance DNA with such extraordinarily high Ct values, that the results reported are representative of actual sequencing and not random error or contamination remains difficult to determine.

We sincerely thank the reviewers for their careful evaluation of our manuscript and their constructive comments. We provide below a detailed, point-by-point response to each of their remarks.

Reviewer #1 (Remarks to the Author):

Interesting and especially novel work. Authors should be commended for the idea and the deep level Omic analysis. This should be considered a FIH study preliminary to others. I do have some relatively minor comments that need to be addressed before publication might be granted:

1. BPD has been diagnosed with an old (2001) definition which is considered obsolete by INC and following consensus. Authors should reclassify cases using the Jensen definition, in order to increase reproducibility

We thank the reviewer for this comment. We agree that the definition proposed by Jobe and Bancalari is now considered outdated for grading BPD severity, and that more recent definitions have been introduced. Accordingly, we have revised our analyses using the Jensen classification to define BPD progression and severity.

However, we retained the concept of oxygen requirement at 36 weeks' PMA, as originally described by Jobe and Bancalari, as a clinically relevant outcome measure. Thus, BPD severity and progression were assessed according to the Jensen criteria, while oxygen requirement was evaluated following the framework established by Jobe and Bancalari.

Overall, the virome, bacteriome and endotypes did not differ according to BPD progression based on the Jensen criteria. However, endotypes remained associated with oxygen requirement at 36 weeks' PMA, even after more stringent decontamination of vOTUs.

2. Tracheal aspirates are well known to be poor samples unable to represent the alveolar biology (see Dargaville P Am J Resp Crit Care Med 1999 amongst the others). Conversely the injection of a minimal amount of normal saline, under controlled conditions allows to better retrieve epithelial lining fluid (see again Dargaville and others), and this has been recognised almost 25 years ago by the first European Respiratory Society recommendations of BAL in neonates and infants (de Blic. Eur Resp J 2000). The injection of normal saline (so calle dNB-BAL or mini-BAL) may be viewed as something invasive and this is totally wrong since many classical studies from the 80s have demonstrated its safety from different point of views. However, the injection does introduce variability and needs standardisation in order to have comparable data (there're several examples of this standardisation e.g. De Luca D. Intensive Care Med 2008). Now obviously it is not possible to improve the quality of specimens retrospectively. However authors should
- acknowledge the limitation due to he use of tracheal aspirates
- comment on a dual sample strategy (like Dargaville 's); i.e. if performed serially, tracheal aspirate and NB-BAL can give info on teh airways and alveolar district contributing to better localise the virome you are analysing.

We agree with the reviewer about the fact that endotracheal aspirates do not reflect the alveolar environment. In this specific population of preterm neonates, we took advantage of the administration of surfactant so that no additional sampling was done. The collection of the first endotracheal fluid in the first day was performed after injection of 4 droplets of normal saline. We

added the more detailed procedure of sampling in the material and methods section and added this limitation in the discussion section:

“Importantly, tracheal aspirates provide only a partial representation of the lower respiratory tract and may not fully reflect alveolar biology. Within the limits of feasibility in preterm neonates, future studies combining tracheal aspirates with standardized mini-BAL sampling could help better distinguish airway and alveolar viromes during the progression of BPD. However, because the abundance of the microbiome decreased along the respiratory tract, the alveolar virome is likely to be even less abundant than the observed endotracheal virome.”

3. "The more you know he more you ask". Being a pilot study this nice work cannot answer all teh questions. However some news to be highlighted and commented for future works:
- what is the effect of GA on virome? Too small sample size to tell and this can be important at the limit of viability
- was there an effect of maternal virome? this would need to be studied by associating maternal vaginal swab analyses
- what was the effect of maturation (i.e. postmentrual age)? This is crucial as it means studying the change of virome profile overtime and cannot be done with a single sampling study
- is there any link between virome and lung aeration or gas exchange metrics, or markers of alveolarization (SPD, IGF1, VEGF) or inflammation (IL6-8,m sPLA2 etc)? That would associate a given virome to a different BPD phenotype

We agree with the comments of the reviewer and have now added in the discussion section these sentences:

“Firstly, this is a pilot study, we did not compute the statistical power required for these exploratory analyses and the limited number of neonates may have reduced our ability to detect significant associations. Notably, the effects of gestational age on virome composition may likely influence viral colonization and dynamics. Similarly, the relationships between the lung function, alveolarization markers and inflammatory cytokines were not investigated.

[…]

The most important next step will consist of implementing longitudinal studies to explore whether these viral communities are transient or persistent, to pinpoint the timing of respiratory bacteriophages being induced from “pioneer” bacteria during the initial first weeks of life and the manner in which these bacteriophages, along with those inherited free, influence the progression to chronic lung diseases. Moreover, incorporating maternal virome investigation in these studies will be necessary to link how maternally transmitted and environmentally acquired viruses interact with host maturation, immunity, and bronchopulmonary outcomes.”

Reviewer #2 (Remarks to the Author):

This is an interesting prospective study of endotracheal aspirates of virome, bacterial microbiome and transcriptome 31 very preterm neonates (<30 wks GA) collected within 24 hours of birth. The investigators report a low-abundance individualized virome, including mainly unclassified Caudoviricetes bacteriophages (71% of neonates). Multi-Omics Factor Analysis identified four microbial endotypes; one protective" endotype with low microbial load was associated with reduced BPD risk (29% vs.79%, p = 0.03), whereas high-risk clusters exhibited increased Pseudomonas load, elevated bacterial transcriptional activity, and greater viral diversity.
Strengths of this study: Virome study- very few virome studies in preterm neonates. Important study associating early life airway virome and microbiome to BPD progression. MOFA analysis using virome, bacteriome and transcriptome for endotypes
Limitations: Just one time point and how the virome changes longitudinally will be very important. Small sample size (n=31). Simultaneous evaluation of the gut virome/microbiome might throw more light on early life virome and subsequent morbidities in preterm infants.

We thank the reviewer for his review of our work. We acknowledge the listed limitations of this work and have now expanded the discussion to better highlight these points:

" The study presents several limitations. Firstly, this is a pilot study, we did not compute the statistical power required for these exploratory analyses and the limited number of neonates may have reduced our ability to detect significant associations. Notably, the effects of gestational age on virome composition may likely influence viral colonization and dynamics. Similarly, the relationships between the lung function, alveolarization markers and inflammatory cytokines were not investigated. Although neonatal endotypes were not associated with the progression of BPD according to Jensen criteria, they were associated with oxygen requirement at 36 weeks PMA (Bancalari and Jobe criteria). Future studies with larger and independent cohorts are necessary to externally validate and confirm the generalizability of the existence of endotracheal microbiome endotypes at birth in preterm neonates with different trajectories of BPD progression and oxygen requirement, and to potentially guide new healthcare strategies. Secondly, we were unable to extrapolate our findings to the lower respiratory tract of healthy term infants, as obtaining such samples is very difficult in this population. Thirdly, tracheal aspirates provide only a partial representation of the lower respiratory tract and may not fully reflect alveolar biology. Within the limits of feasibility in preterm neonates, future studies combining tracheal aspirates with standardized mini-BAL sampling could help better distinguish airway and alveolar viromes during the progression of BPD. However, because the abundance of the microbiome decreased along the respiratory tract, the alveolar virome is likely to be even less abundant than the observed endotracheal virome. Finally, the cross-sectional design of the study may have hindered the identification of factors such as genetic, oxygen supplementation, breastfeeding, and respiratory bacterial or viral colonization, which may influence BPD progression[49]. The most important next step will consist of implementing longitudinal studies to explore whether these viral communities are transient or persistent, to pinpoint the timing of respiratory bacteriophages being induced from "pioneer" bacteria during the initial first weeks of life and the manner in which these bacteriophages, along with those inherited free, influence the progression to chronic lung diseases. Moreover, incorporating maternal virome investigation in these studies will be

necessary to link how maternally transmitted and environmentally acquired viruses interact with host maturation, immunity, and bronchopulmonary outcomes. “

Reviewer #3 (Remarks to the Author):

Setting aside the fact that this paper should not have been sent out for review with a typo in the title, I was intrigued by what the authors set out to accomplish in this manuscript.
Over the authors have made extensive strides to perform a cutting-edge mult-omic analysis of tracheal aspirates from premature infants and overall they have succeeded admirably. However, a fatal limitation still remains.
Primarily Limitation:
The principal limitation of this manuscript is the exceptionally low biomass of the supposed viral sequences. Given the current understanding of the field that low-biomass of previously reported "pregnancy" or "near-birth" microbiomes are actually representative of contamination and sequencing error, the burden of proof that these ultra-low biomass viral sequencing are representative of actual viral DNA falls on the authors. As the manuscript stands, they have failed to meet this critical barrier. Considerable additional positive and negative controls and methodological demonstration of the appropriate Ct thresholds representative of actual sequencing are required. As it stands the evidence presented in the manuscript suggests that what is actually being characterized is a mixture of contamination and sequencing error. Without better controls the use of post-hoc computational decontamination is insufficient to alleviate these significant methodological issues. While some of this appears to be included in the viral metagenomic sequencing section of the methods, showing this data in the supplement, and showing the computational decontamination data, is a requisite.

Major Corrections:
1. The authors acknowledge the absence of "sampling device flush controls" (lines 382-383) a s a limitation. Given the low-biomass nature of the samples and the potential for contamination during collection, this is a significant methodological omission. Please discuss i n the limitations section how the other decontamination steps (e.g., NTCs, PBS controls) might mitigate the impact of this specific missing control, or if a more robust justification for ruling out sampling device contamination can be provided.

We thank the reviewer for this thoughtful comment emphasizing the importance of controlling for sampling device contamination in low-biomass metagenomic studies.

We fully agree that the absence of sampling device flush controls collected concurrently with study samples represented a limitation of this work. We also acknowledge that, given recent debates on early-life microbiota acquisition and the risk of interpreting contamination or *in silico* false positives as genuine signals, it is crucial to demonstrate how our decontamination steps enabled the identification of specific and reliable signals in biological samples compared with controls.

1. Stringent decontamination procedures
To address this concern, we applied the most stringent decontamination steps to identify sample-specific signals. These steps included:

a. **Removal of vOTUs enriched in negative controls:** All viral operational taxonomic units (vOTUs) detected more frequently in no-template controls (NTCs) than in samples were removed, using the *decontam* R package, as previously described in the original submission.

b. **Subtraction of maximum control signal:** For additional stringency, we subtracted the maximum signal detected among all 35 controls (5 NTCs and 31 PBS-paired blanks) from all corresponding sample, regardless of pairing. This step replaced the previous pairwise decontamination approach and accounts for low-level signals that may arise sporadically across controls.

This process resulted in the exclusion of 20 vOTUs (out of the initial 89 from the first submission), corresponding to those sporadically found in PBS controls but not consistently in paired samples. To illustrate this process, we added new supplementary heatmaps (Figures S1 and S2) showing the abundance of the retained vOTUs across all NTCs. This analysis revealed that only 3 of the final 69 vOTUs were detectable in NTCs, and always at much lower abundance, while all others were specific to samples and undetected in any of the 35 controls.

c. **Post-hoc sequencing of sampling device flushes**
To further assess potential contamination from the sampling devices, we conducted post-hoc sequencing of six sterile flushes obtained from the same type of tracheal aspiration devices used in the study (though from a different manufacturing lot). These flushes were processed using identical extraction and library preparation protocols. We evaluated the homology of vOTUs retained at the precedent step with those identified in sampling device flushes (Table S1). None of the retained vOTU, except those corresponding to the MS2 IC shared a homology above 85% with the vOTUs of the sampling devices, confirming the likely absence of shared viral sequences and mitigating the risk of interpreting vOTUs originating from devices.

d. **Reanalysis of virome data:** Following this refined decontamination, we reanalyzed all data initially presented in the manuscript. Despite the removal of 20 vOTUs, the main findings remained unchanged:

- Higher abundance of viruses, particularly bacteriophages, in neonates delivered vaginally
- Only 2 of 23 neonates were finally negative for viral detection in endotracheal aspirates at birth
- The MOFA2 analysis again identified a cluster associated with a significantly lower risk of oxygen requirement
- The model retained strong generalizability, with 85% correct cluster prediction
- Host-response exploration identified one cluster with a high inflammatory profile, clusters with higher risk of oxygen requirement at 36 weeks' PMA associated with different metabolic activity, while a higher epithelial morphogenesis was found for the low-risk cluster.

All relevant figures, tables, and statistical results were updated accordingly throughout the manuscript.

This more stringent reanalysis and the additional sequencing of device flush controls mitigate the risk to misinterpret viral signals in these low-biomass samples. Nevertheless, we agree that the interpretation should remain cautious. Although we cannot fully exclude that a small fraction of viral sequences represents background noise or dry-lab artifacts, the strong and biologically plausible association between viral abundance and mode of birth

supports that contamination from sampling devices is unlikely to explain the observed viral signatures in tracheal aspirates.

We have expanded the discussion in the *Limitations* section:

"A key strength of our study was the meticulous removal of putative contaminant signals coming from the wet-lab process to ensure the identification of true viral communities, and face to the challenge of microbial investigations in ultra low-biomass samples[1]. Importantly, we focused only on viral contigs found specifically in endotracheal samples, using more NTCs than samples to remove potential contaminants and to mitigate the risk for contaminants introduced during the sample collection by sequencing device flush controls. Because the detected viral contigs were of very low abundance, we cannot formally exclude that they represent random contaminants or sequencing errors. Nonetheless, the inter-individual variability of detected viral contigs , and the differences observed regarding mode of delivery reinforce the hypothesis that these viruses might be indeed present and not background noise. Given the absence of shared viral contigs between the two twin pairs, it is also possible that viruses are acquired through transient stochastic exposure to the environment. Consistent with this, only ~15% of the infant gut virome is maternally derived, with substantial inter-individual variability even among co-twins, supporting a predominantly stochastic early-life assembly of the virome[17]. These signals might reflect real early-life exposures through vaginal delivery or environment, particularly in preterm infants, and whose immature barriers and heightened environmental contact increase susceptibility to such influences. "

2. Similarly while limitations in RNA detection sensitivity or the presence of non-viable bacteria" (Lines 453-454) are plausible, this explanation is insufficient and further evidence and expanded discussion is indicated.

We thank the reviewer for this suggestion. We have expanded the potential factors explaining "the absence of highly transcriptionally active bacteria in other endotypes":

1. RNA detection sensitivity:
The tracheal aspirate samples were of extremely low biomass and originated from premature infants in critical care, which limited both the available material and RNA yield. Despite optimized extraction protocols and immediate storage at –80 °C, partial RNA degradation during sampling and handling may have reduced sensitivity for bacterial transcript detection. Furthermore, as shown in our previous work (Destras et al., Frontiers 2026), exploring bacterial RNA through metatranscriptomics requires a particularly high sequencing depth, making the detection of low RNA levels challenging. Nonetheless, in both our previous and current studies, most predominant bacteria identified by 16S sequencing were also detected in metatranscriptomic data.

2. Presence of non-viable bacteria:
A portion of the bacterial DNA signal may reflect non-viable cells or extracellular DNA rather than active infection. This is biologically plausible, particularly in antibiotic-exposed or polymicrobial samples, and further studies assessing bacterial viability and RNA yield are warranted.

However, because of RNA detection sensitivity, the presence or absence of detectable RNA should not be interpreted strictly as bacterial viability, but rather as an indication of relative transcriptional activity among taxa.

We have added this paragraph in the discussion section:

“Although samples were promptly stored at –80 °C and processed with optimized extraction protocols, partial RNA degradation may have reduced metatranscriptomic sensitivity. As shown in our previous work (Destras *et al.*, *Frontiers 2026*), detecting bacterial RNA in low-biomass samples requires high sequencing depth, and undetected taxa may simply reflect low transcriptional activity or technical limits. Moreover, part of the bacterial DNA signal might originate from non-viable bacteria, which need to be demonstrated in future in vitro studies. Thus, DNA detection in tracheal aspirates should be viewed as evidence of microbial presence or exposure, whereas RNA detection reflects relative transcriptional activity rather than absolute viability.”

3. The Kruskal-Wallis result in the summary at line 26-27 does not match what is reported in Table 2. Please clarify to limit potential confusion.

We thank the reviewer for identifying this inconsistency. Upon review, we realized that the summary in the text (lines 26–27) contained an error and did not accurately reflect the values presented in Table 2. We have corrected the text to match the Kruskal-Wallis results reported in Table 2 and the new Table S3 (computation of statistical tests one cluster vs all others), and clarified the description of the statistical tests. These changes ensure consistency between the text and tables and reduce potential confusion for readers.

4. Line 515, includes a placeholder instead of the SRA number. Due to this the data could not be access and independently verified. Full availability of raw data is the standard of the field and must be corrected.

We thank the reviewer for highlighting this oversight. The placeholder in line 515 was unintentional. The raw sequencing data are fully deposited and publicly available in the NCBI Sequence Read Archive (SRA) under accession number **[PRJNA1373503]**. We have updated the manuscript to include the correct accession, ensuring that all data can be accessed and independently verified in accordance with field standards.

5. The manuscript states that "endotracheal virome profiles were found to be unable to distinguish between infants according to birth mode" (line 277), yet immediately after, it states that "bacteriophage abundance and viral richness at birth were significantly higher in vaginally delivered infants" (lines 278-279), referencing Figure 3E. Please clarify this apparent contradiction. If the overall virome profiles (e.g., as visualized by NMDS in Figure 3D) do not show clear separation by birth mode, but specific univariate analysis do (as I understand these graphs), then this should be more clearly and explicitly stated.

We thank the reviewer for highlighting this apparent contradiction. We agree that the distinction between overall virome composition and specific viral features was not clearly stated. The NMDS analysis (Figure 3D) indicates that the overall viral community composition does not segregate by birth mode, consistent with the lack of global clustering. In contrast, univariate analyses of individual features, such as bacteriophage absolute abundance and viral richness at birth (Figure 3e), reveal

significant differences between vaginally and cesarean-delivered infants. We have revised the manuscript in order to clarify the apparent contradiction:

"Endotracheal overall viral community composition was found to be unable to distinguish between infants according to progression of BPD or the birth mode **(Fig. 3b-3c)**. However, regarding specific viral features, we identified that viral normalized biomass and richness at birth were significantly higher in vaginally delivered infants ($p<0.01$ and $p=0.01$, respectively**, Fig 3D)**. This observation was primarily driven by bacteriophages as their normalized biomass (VDR) and relative abundances were significantly higher in vaginally delivered neonates ($p<0.05$, **Fig 3d, Supplementary Fig S3**)."

6. The authors utilize MOFA, when the significantly improved MOFA2 has been available for several years. Please update the analysis to MOFA2: https://github.com/bioFAM/MOFA2.

We thank the reviewer for this comment. We would like to clarify that MOFA2 was already used for the integrative multi-omics analysis presented in the manuscript, rather than the original MOFA. The Methods section and figure legends have been updated to explicitly indicate MOFA2 to avoid any confusion.

7. In Table 2, Kruskal-Wallis was used to determine significance. Best practices is to report adjusted and unadjusted P-values. Also, the methods used to perform this adjustment should be clearer.

We thank the reviewer for this suggestion. In the revised manuscript, Table 2 now includes both unadjusted and Benjamini-Hochberg-adjusted p-values for all Kruskal-Wallis tests. We have also clarified in the Methods section that the Benjamini-Hochberg procedure was applied to control for multiple testing. These updates ensure transparent reporting and alignment with best practices.

8. The y-axis ranges for "Virome (VDR)" and "Bacteriophages (VDR)" in Figure 3A, and for "VDR" in Figure 4C, are truncated at 1.00. However, the text mentions values outside of these range, and correcting these axis would improve the data presentation.

We thank the reviewer for pointing this out. In the revised figures, the y-axis ranges for "Virome (VDR)" and "Bacteriophages (VDR)" in Figure 3A and for "VDR" in Figure 4C have been adjusted to fully display all observed values. This correction ensures that the figures accurately reflect the data and improves clarity of data presentation. Values for virome are intended to be below 1 (expressed as VDR).

9. While R packages and versions are provided for the bioinformatics analyses (e.g., MOFA, UMAP, clustering), providing more specific details on the bioinformatic analysis, or better yet, access to the analysis code on GitHub or other relatively permanent access website would greatly facilitate reproducibility of the analysis.

We thank the reviewer for this important suggestion. To enhance reproducibility, we have made the full analysis scripts and workflows publicly available on GitHub at https://github.com/genepii/Multi_omics_analyses_Verdy, providing a permanent reference for reproducing the results.

Minor Corrections:
1. Figures 5C, 6B and 6C are not called out in the text.

We thank the reviewer for noting this oversight. The text has been updated, especially the text in legends.

2. Figure 2B and 2C appear to be swapped in the figure versus the legend.

We have corrected Figure 2 so that panels B and C match the figure legend and text.

3. Use of the term "protective" is confusing, we would suggest less causative oriented language, such as "lower risk-associated" since this analysis can in no way imply causality.

We agree that "protective" implies causality. The term has been replaced with "lower risk-associated" throughout the manuscript to better reflect associations rather than causation.

4. The interpretation of the asterisks in Table 2, indicating "one vs all-cluster comparison (Mann-Whitney test, pvalue‹0.05)", is ambiguous. The asterisks appear on specific median (QR) values within the table, implying that these specific clusters are significantly different from all other clusters combined. This is an unusual way to present post-hoc test results. Please clarify this interpretation in the table legend or footnote, and ensure the statistical methodology for these comparisons is fully described in the methods.

We have removed the asterisks in a Table 2 that indicated significant differences between the specified clusters. We performed additional analyses, summarized in a new Table S3, to explore features specifically associated with the different clusters. These analyses include Mann–Whitney or Fisher's exact tests comparing one cluster against all other clusters combined, with corresponding *P* values and adjusted *P* values calculated using the Benjamini–Hochberg method. Moreover, we added effect size measures for the different features (Cliff's delta for continuous variables and odds ratios for categorical variables), as well as summary statistics for the features (medians and direction). We specified in the Material and Methods section:

"Medians were compared using non-parametric Kruskal-Wallis or Mann-Whitney U tests. Proportions were compared using Fisher or Chi-square tests. To identify features associated with specific endotype clusters, Mann-Whitney U tests or Fisher's exact tests were used to compare one cluster against all others. Effect sizes were calculated using Cliff's delta for continuous variables and odds-ratio for categorical variables, and corresponding z-scores were derived. Clinical and microbial features associated with clusters were retained when P values were < 0.05 and Benjamini–Hochberg-adjusted P values were < 0.15. "

5. The "protective" endotype (C1) is characterized by a "low-abundant microbiome" (line 339), but the text later states that "higher loads of Staphylococcus sp. were observed in the protective 'cluster'" (line 344). While Table 2 shows a higher median Staphylococcus load in C1 compared to other clusters, the overall bacterial DNA load in C1 is lower than in C4. Please clarify this phrasing to avoid confusion about whether "higher loads" refers to absolute abundance or relative abundance within a low-biomass context, especially given the overall low abundance characteristic of C1.

We apologize if our wording was unclear. Higher loads refered to higher absolute abundance expressed as 16S rRNA copies/mL. We have revised the text according to the new results. Now there is no difference according the load of DNA bacteria. Nonetheless, we can still identify a higher absolute load of *Staphylococcus sp.* in the low-risk cluster.

"The "low-risk" endotype for oxygen requirement C2 was characterized by a higher absolute abundance of *Staphylococcus sp*., the absence of transcriptionally active bacteria and the presence of virulent phages dominating the virome."

6. Given the studies cross-sectional design, while acknowledged, the need for further longitudinal sampling to actually demonstrate these findings should be further emphasized.

We fully agree with the reviewer. We have emphasized in the Discussion that longitudinal sampling would be necessary to confirm these findings and assess temporal dynamics, reinforcing the study's limitations:

"The most important next step will consist on implementing longitudinal studies to explore whether these viral communities are transient or persistent, to pinpoint the timing of respiratory bacteriophages being induced from "pioneer" bacteria during the initial first weeks of life and the manner in which these bacteriophages, along with those inherited free, influence the progression to chronic lung diseases. Moreover, incorporating maternal virome investigation in these studies will be necessary to link how maternally transmitted and environmentally acquired viruses interact with host maturation, immunity, and bronchopulmonary outcomes."

7. Reference 20 is related to MOFA not Papillomaviruses (line 421). Please correct.

Reference 20 (now 25) has been corrected to reflect the appropriate citation.

8. There is another type at line 490.

The typographical error at line 490 has been corrected.