Human Sputum Microbiome Composition and Sputum Inflammatory Cell Profiles Are Altered with Controlled Wood Smoke Exposure as a Model for Wildfire Smoke

Catalina Cobos-Uribe, MS - https://orcid.org/0000-0002-6671-0780, Radhika Dhingra, PhD - http://orcid.org/0000-0003-0202-1860, Martha A. Almond, Neil E. Alexis, MHSc, PhD - https://orcid.org/0000-0002-9417-8269, David B. Peden, MD, MS - https://orcid.org/0000-0003-4526-4627, Jeffrey Roach, PhD https://orcid.org/0000-0001-9817-5877, Meghan E. Rebuli, PhD - https://orcid.org/0000-0003-1918-2257

Online Data Supplement

Methods

Study participants

The population analyzed in this study (N=54, Table 1) constitutes a subset of participants from the SmokeScreen clinical study (NCT02767973). Detailed inclusion and exclusion criteria were previously described (1). In brief, exclusion criteria for this study were smoking, antibiotic use within four weeks of wood smoke exposure, pregnancy, symptomatic allergic rhinitis or active allergies, upper respiratory infection within four weeks of the challenge, presence of contraindicated or interfering medical conditions, and medication use that may impact response to wood smoke exposure (e.g., steroids, beta-agonists, and immunosuppressive drugs). Participants with mild asthma on controller therapy were asked to withhold this therapy for two weeks before wood smoke exposure. Only samples with sufficient induced sputum volume for downstream microbiome analysis were included in this study. This study was approved by the University of North Carolina Institutional Review Board (UNC IRB Numbers: 05-2528 and 15-1775).

Controlled exposure to wood smoke

Participants were exposed to 500 $\mu g/m^3$ of wood smoke for two hours under controlled conditions at the U.S. EPA Human Studies Facility on the UNC-Chapel Hill campus from June 2016 to November 2022. Wood smoke was obtained by smoldering dried, untreated red oak logs. Generated smoke was injected into a human exposure chamber maintained at 22 °C and 40% humidity, as described by (2). To maximize wood smoke inhalation, participants were instructed to alternate between 15 minutes of exercise and 15 minutes of rest throughout the two-hour exposure period.

Induced sputum collection and processing

Induced sputum was collected and processed following the protocol outlined (3). Of note, participants rinsed their mouths and cleared their throats through gargling water prior to sputum induction to reduce potential oral contamination. Baseline samples were collected at a screening visit at least 24 hours prior to the wood smoke exposure visit, followed by collections at six- and 24 hours post-exposure (hpe). Following collection, samples were promptly transported on ice to the laboratory and processed. Briefly, plug selection was applied to the raw sputum samples. Then, samples were washed with DPBS Dulbecco's Phosphate-Buffered Saline (DPBS, Gibco, ThermoFisher Scientific, Cat# 14190144) and centrifuged. The supernatant was recovered and stored (-80 °C). Dithiothreitol (DTT, Sputolysin, EMD Millipore, CAS 578517) was used to treat the remaining cell mixture (15 minutes), followed by filtration to remove squamous cells, centrifugation, and recovery of DTT-supernatants and a cell pellet for hemocytometry evaluation and cytospin generation. DTT-treated supernatants were collected, aliquoted, and stored at -80°C. Sputum supernatants remained frozen until analysis. Samples containing fewer than 40% squamous cells were deemed to meet sputum quality control criteria. As only samples meeting quality

control criteria were stored in the biorepository utilized to generate the data in this study, all samples included in this study had already passed this quality control step by the biorepository, and no exclusions for sputum quality were necessary.

Markers of respiratory inflammatory response to wood smoke exposure

Total cell counts were determined from the sputum cell fraction using a Neubauer hematocytometer and Trypan Blue staining. Differential leukocyte analysis was performed on stained (Hema 3 stain) cytospin slides, and counts were expressed as a percentage of total non-squamous nucleated cells and as cells per sputum mg. The sputum cytokine profile (IL-1β, IL-6, IL-8, and TNFα) was determined from the DPBS treated sputum supernatant using a multiplex ELISA (MesoScale Discovery, Rockville, MD).

DNA extraction and 16S rRNA gene amplification and sequencing

Total DNA was extracted from the DTT-treated sputum supernatants. Mock sputum samples (n=3) and DNA extraction controls (n=7) were used as negative controls to identify possible contamination during sample processing. Mock samples were generated by subjecting sterile saline solutions (DPBS, Gibco, ThermoFisher Scientific, Cat# 14190144) to the sputum processing as described above and DNA extraction processing, followed by library preparation. Extraction controls consisted of sterile distilled ultrapure water aliquots, which were subjected to the same DNA extraction and library preparation protocol as the experimental samples. DNA extraction was completed using the DNeasy PowerSoil Pro kit (Qiagen, Germany) with a few modifications to the manufacturer's instructions. To process as much of the sample as possible, we used two PowerBead Pro tubes per sample and added 500-800µL of sample and 800µL of CD1 solution to each bead tube. Then, after vortexing and centrifuging, we transferred the entire supernatant to a clean microtube and added 400µL of CD2 solution, double the amount specified in the kit's instructions, to compensate for the larger supernatant volume. After centrifuging, we processed the entire supernatant solution (up to 700 µL at a time) as indicated in the instructions using the same spin column until all the solution was transferred. After this point, we followed the instructions provided by the manufacturer until the final elution step, where the extracted DNA was eluted to a final volume of 30 µL. DNA was quantified with the Qubit™ Flex Fluorometer and the Qubit™ dsDNA Broad Range Quantitation kit (ThermoFisher Scientific, USA) using 5 µl aliquots. The extracted DNA was stored at -80 °C until sequencing. Extracted DNA was sent to SeqCenter (SeqCenter LLC, PA, USA) for amplification and high-throughput sequencing. Samples were prepared using the Quick-16S™ NGS Library Prep Kit (Zymo Research, USA) with phased primers targeting the V3-V4 region of the 16S rRNA gene (Forward sequences: CCTACGGGDGGCWGCAG and CCTAYGGGGYGCWGCAG; Reverse sequence: GACTACHVGGGTATCTAATCC). Samples were sequenced on a P1 600cyc NextSeq 2000 Flowcell, producing 2x301 bp paired-end reads.

Bioinformatic analysis

Raw sequencing data were converted to the FASTQ format and demultiplexed with bcl-convert (4). These sequences were imported into QIIME2 (5), where primer sequences were removed using the Cutadapt plugin (6). Sequences were denoised using the dada2 plugin (7). Denoised sequences were assigned taxonomic identifiers using two databases: the commonly used SILVA database (8) and the extended Human Oral Microbiome Database (eHOMD) (9), which focuses on bacteria in the human aerodigestive tract; thus, enhancing taxonomic classification resolution. Hereafter, eHOMD is included in the main body of this paper, and SILVA is in the supplemental material. Finally, the metadata, phylogenetic tree, and feature and taxa tables were imported into RStudio using the *qiime2R* R package (version 0.99.6) (10) and used to create a phyloseq object (*phyloseq* package version 1.38.0). The *decontam* package (version 1.14.0) (11) was used to identify and remove contaminant sequences in our data, using mock sputum samples and extraction controls as negative controls (n = 10), using the prevalence method with a 0.1 threshold. Seventeen contaminant sequences were identified and removed from the phyloseq object. Contaminant sequences and their relative abundance and prevalence in sputum samples are reported in Supplementary Tables E1 and 2). This and all subsequent analyses and visualizations were carried out using R version 4.1.0 (12) in RStudio (13).

Data analysis

Host respiratory inflammatory response: Inflammatory cell and cytokine production data were analyzed using a repeated measures one-way ANOVA followed by pairwise paired *t*-tests to assess the statistical significance between the exposure times (aggregate analysis). For stratified analysis by sex, a repeated measures two-way ANOVA was used, followed by pairwise *t*-tests for comparison between treatment groups and between time points. Based on Q-Q plot assessments of normality (Supplemental Figure 1), percentage macrophage and neutrophil data were approximately normal and thus run via parametric tests. Subjects with missing inflammatory cell or cytokine data were excluded from the analysis, resulting in a final dataset of 45 subjects for inflammatory cell analysis (30 females, 15 males) and 35 subjects for cytokine analysis (20 females, 15 males). Absolute cell counts and cytokine data were log-transformed to more closely align with normal distribution.

Microbiome diversity: Alpha diversity metrics were determined using two different R packages. Shannon index was estimated with the *estimate_richness* function in *phyloseq* (14) and Faith's phylogenetic diversity was calculated with the *estimate_pd* function in *btools* (15). For statistical analysis, we used the same approach as with the host immune response with data log transformation to better align the associated Q-Q plots (Supplemental Figure 1). Beta diversity (Bray-Curtis dissimilarity) was calculated using the *adonis2* function from the *vegan* package (version 2.6-4) (16); results were visualized using a principal coordinate

analysis (PCoA). Beta diversity was evaluated between exposure groups with and without metadata variables (e.g. sex). To assess the impact of exposure and other factors on beta diversity, we employed a Permutation Multivariate Analysis of Variance (PERMANOVA). PERMANOVA was chosen for beta diversity and differs from the tests used for alpha diversity and inflammatory responses above and to account for the non-independence of samples within subjects, as the Bray-Curtis dissimilarity index is a distance matrix to control for within-subject variability, thus these data points lack the independence required by the assumptions of a one-way ANOVA. Given the repeated measures design of our study, we incorporated the 'strata' argument in our *adonis2* analysis to account for the non-independence of samples within subjects, thereby controlling for within-subject variability and preventing pseudoreplication.

Differential abundance analysis and host-microbiome associations: Relative abundance bar plots were created using the *microshades* package (17). To facilitate the visualization of relative abundance, we created plots at the phylum, genus, and species taxonomic levels. Differential abundance analysis was performed with a negative binomial mixed model (NBMM) controlling for age, sex, BMI, and asthma status using the *nlme* and *NBZIMM* packages and a minimum prevalence of 20% (18). Other demographic factors were assessed as sensitivity analyses in the model, but did not substantially alter the results (Supplementary Tables E6X, 8, 10, and 12). This model is recommended for longitudinal microbiome data because it accounts for the over-dispersion and sparsity present in microbiome count data while also handling the repeated measures from the same subject over time (i.e., random effect = subject), improving the reliability of statistical inferences about microbial abundance changes and the effects of covariates (18,19). Finally, using the same NBMM model, we explored potential host-microbiome associations by including immune cell data as model covariates. To investigate potential host-microbiome associations between sputum macrophages and microbiome taxa, we applied the same NBMM model, including macrophage per sputum mg data as a model covariates. Additionally, we performed a stratified analysis to examine host-microbiome associations further (54). This analysis compared cytokine and inflammatory cell data between groups defined by baseline microbiome observed richness. Richness groups were classified as high or low based on the study median (median ASV = 126). The low-richness group included samples with ASV <126, while the high-richness group included samples with ASV >126.

All *p*-values were corrected for multiple comparisons using the Benjamin-Hochberg false discovery rate (FDR) correction. Reported p-values throughout the manuscript are adjusted unless otherwise specified. For repeated measures one-way ANOVA, FDR correction was applied separately to each analysis set. For inflammatory cell analyses, FDR correction accounted for four comparisons (one per variable across three timepoints); for alpha diversity, two comparisons (one per diversity metric across three timepoints). For

the differential abundance analysis using a NBMM, FDR correction was applied to 170 comparisons at the genus level and 450 at the species level, corresponding to the number of taxa analyzed.

Supplementary Figure Legends

Supplementary Figure E1. Q-Q plots for macrophage and neutrophil percentages and absolute values, inflammatory cytokines (IL-1 beta, IL-6, IL-8, and TNF alpha), and alpha diversity measures (Shannon and Faith PD) at each timepoint (Pre-Exposure, 6 hours post-exposure, and 24 hours post-exposure). The plots assess the normality of the data distribution across the different variables and timepoints.

Supplementary Figure E2. Total cells per sputum mg observed following wood smoke exposure. Total cell counts were performed Pre-exposure, at six hours post-exposure, and at 24 hours post-exposure. Data is presented as individual data points connected by lines to represent paired analysis for each subject. The blue dashed line represents the mean. Analyzed with repeated measures one-way ANOVA.

Supplementary Figure E3. Sputum inflammatory cell host response following wood smoke exposure analyzed by sex. Induced sputum cell differentials were stained and counted for neutrophils and macrophages prior to exposure (Pre-exposure), at six hours post-exposure (6h post-exposure), and at 24 hours post-exposure (24h post-exposure). A) Relative percentage of neutrophils, B) absolute neutrophils per mg of sputum, C) relative percentage of macrophages, and D) absolute macrophages per mg of sputum are shown in box and whisker plots. Analyzed with repeated measures two-way ANOVA followed by pairwise paired t-tests. Purple (left) = females; Blue (right) = males. P-values shown (* $P \leq 0.05$, ** $P \leq 0.01$, and **** $P \leq 0.0001$) correspond to the paired t-test results. Purple lines below significance * are representative of significant differences between females across exposure groups. Black lines below significance * are representative of sex differences within exposure groups.

Supplementary Figure E4. Sputum cytokine response following wood smoke exposure. Cytokine profiles were measured from DPBS-treated sputum supernatants prior to exposure (Pre-exposure), at six hours post-exposure (6h post-exposure), and at 24 hours post-exposure (24h post-exposure). A) Interleukin-1β (IL-1β), B) Interleukin-6 (IL-6), C) Interleukin-8 (IL-8), and D) Tumor Necrosis Factor-α (TNF-α). Analyzed with repeated measures one-way ANOVA followed by pairwise paired t-tests.

Supplementary Figure E5. Sputum cytokine response following wood smoke exposure analyzed by sex. Cytokine profiles were measured from DPBS-treated sputum supernatants prior to exposure (Pre-exposure), at six hours post-exposure (6h post-exposure), and at 24 hours post-exposure (24h post-exposure). A) Interleukin-1β (IL-1β), B) Interleukin-6 (IL-6), C) Interleukin-8 (IL-8), and D) Tumor Necrosis Factor-α (TNF-α). Analyzed with repeated measures one-way ANOVA followed by pairwise paired t-tests. Purple (left) = females; Blue (right) = males. P-values shown (* $P \leq 0.05$) correspond to the paired t-test results. Black lines below significance * are representative of sex differences within exposure groups.

Supplementary Figure E6. Sputum microbiome alpha diversity with exposure to wood smoke analyzed by sex. Induced sputum supernatants were analyzed via 16S rRNA sequencing. Samples prior to exposure (Pre-exposure), at six hours post-exposure (6h post-exposure), and 24 hours post-exposure (24h post-exposure) were evaluated for differences in alpha diversity. Alpha diversity was analyzed by Shannon's (A) and Faith's (B) indices. Purple (left) = females; Blue (right) = males.

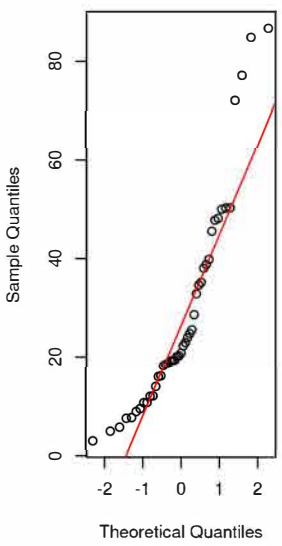
Supplementary Figure E7. Principal Coordinates Analysis (PCoA) of beta diversity. Lines between data points connect the three data points for each subject to illustrate individual changes across time points. Panels B, C, and D display only two data points at a time to improve visualization.

Supplementary Figure E8. Sputum microbiome composition. Induced sputum supernatants from study participants exposed to wood smoke were collected pre-exposure, 6 h post-exposure, and 24 h post-exposure. Samples were analyzed via 16S rRNA sequencing and annotated using SILVA. (A) Microbiome composition by relative abundance at the genus level for individual participants. (B-D) Aggregate median relative abundances at the (B) phylum, (C) genus, and (D) species level. Error bars represent the interquartile range.

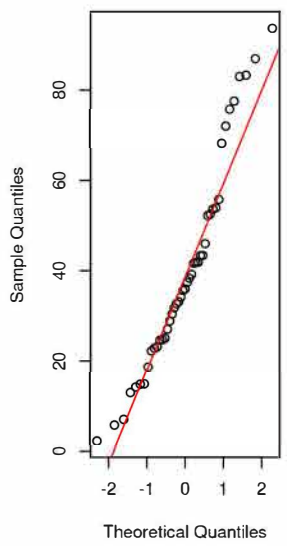
Supplementary Figure E9. Sputum microbiome differential abundance analysis for effects of wood smoke exposure at the (A) genus and (B) species level using SILVA and a negative binomial mixed model (NBMM) controlling for age, sex, BMI, and asthma status. Bacteria that were significantly affected by exposure are shown on the left of each plot, the direction and magnitude of effect shown in the middle of each plot, and the associated p-value on the right of each plot.

Supplementary Figure E10. Sputum cytokine response following wood smoke exposure stratified by baseline microbiome richness. Cytokine profiles were measured from DPBS-treated sputum supernatants prior to exposure (Pre-exposure), at six hours post-exposure (6h post-exposure), and at 24 hours post-exposure (24h post-exposure). A) Interleukin-1β (IL-1β), B) Interleukin-6 (IL-6), C) Interleukin-8 (IL-8), and D) Tumor Necrosis Factor-α (TNF-α). Analyzed with repeated measures one-way ANOVA followed by pairwise paired t-tests.

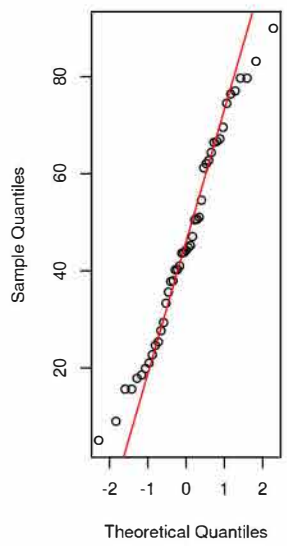
Neutrophil%: Pre-Exposure



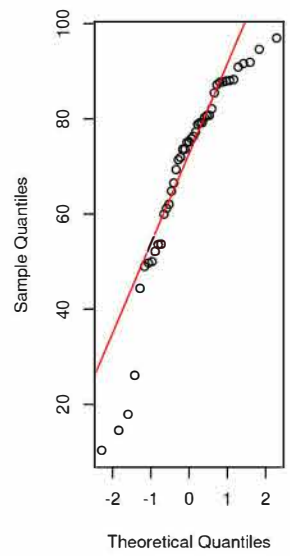
6h Post-Exposure



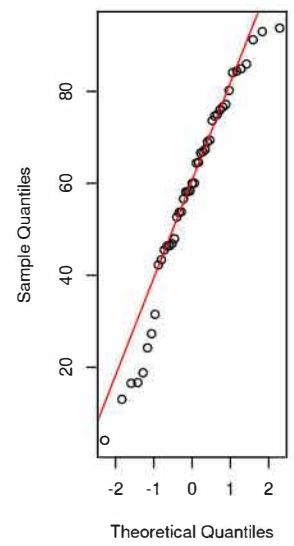
24h Post-Exposure



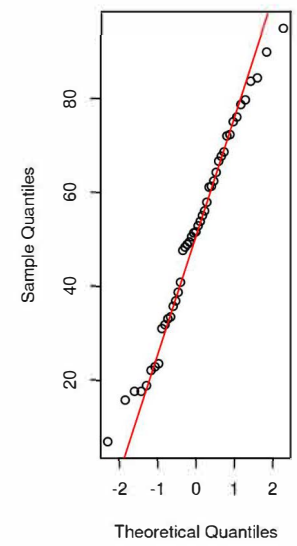
Macrophage%: Pre-Exposure



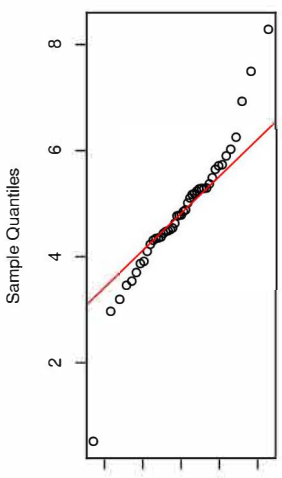
6h Post-Exposure



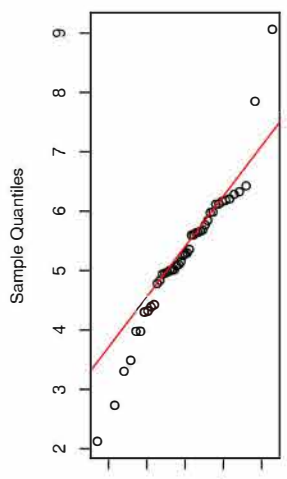
24h Post-Exposure



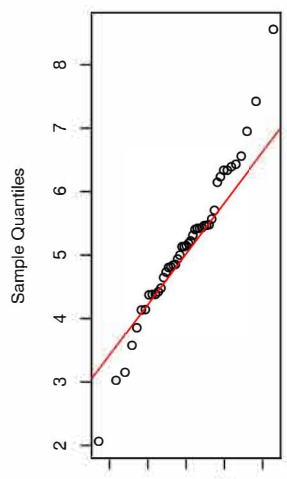
Neutrophil/mg: Pre-Exposure



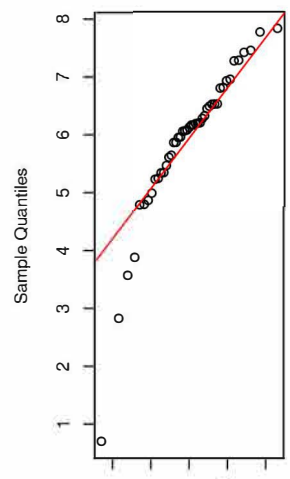
6h Post-Exposure



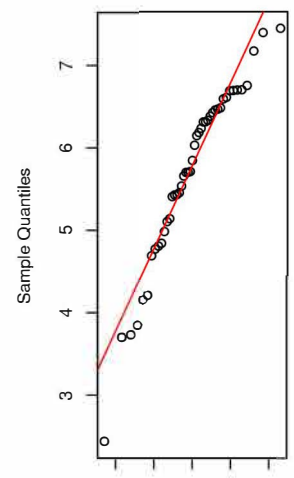
24h Post-Exposure



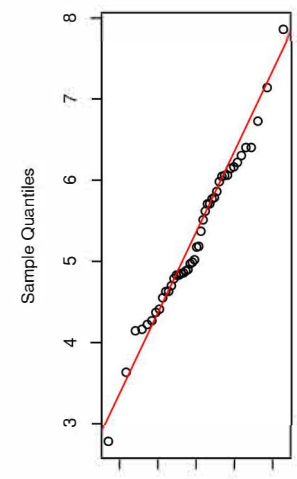
Macrophage/mg: Pre-Exposure



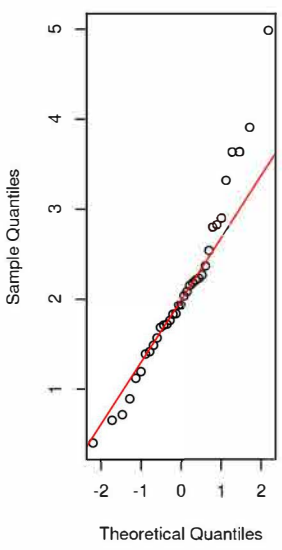
6h Post-Exposure



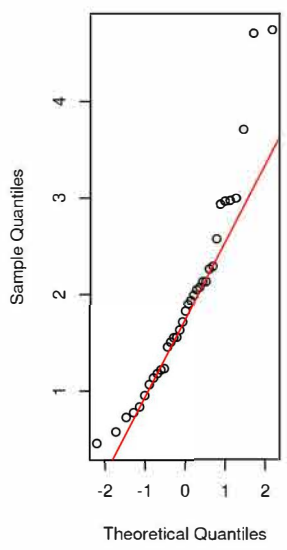
24h Post-Exposure



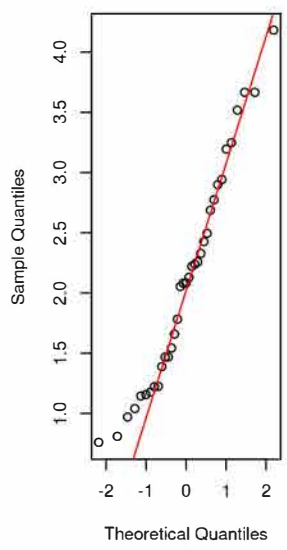
IL-1 beta: Pre-Exposure



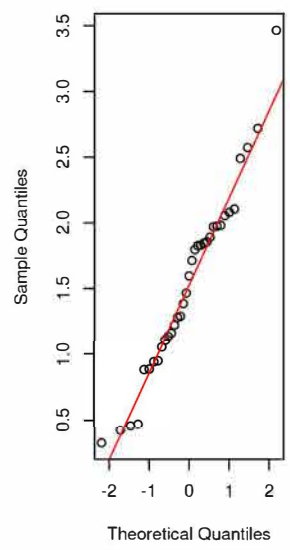
6h Post-Exposure



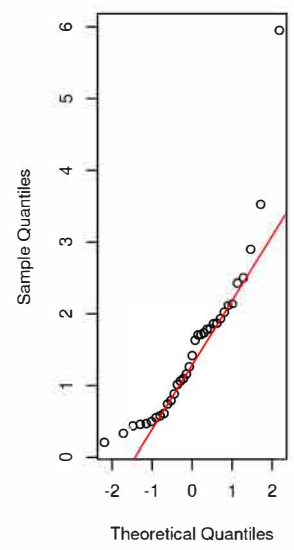
24h Post-Exposure



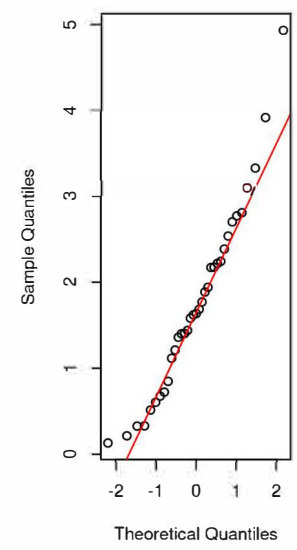
IL-6: Pre-Exposure



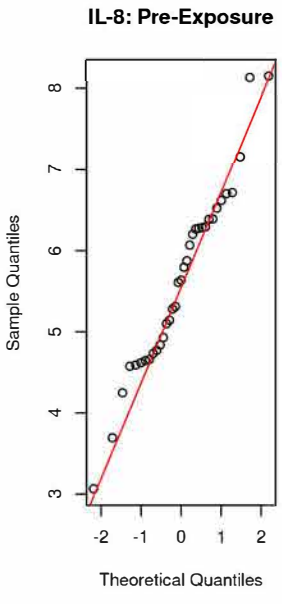
6h Post-Exposure



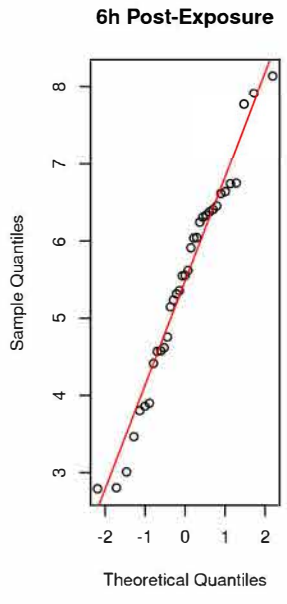
24h Post-Exposure



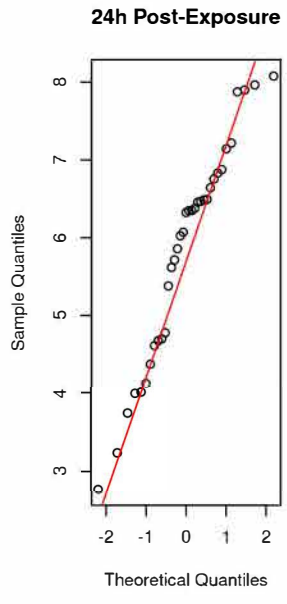
IL-8: Pre-Exposure



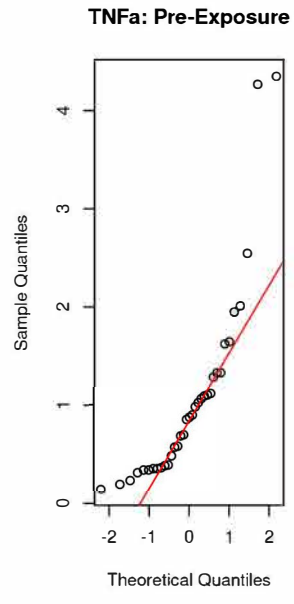
6h Post-Exposure



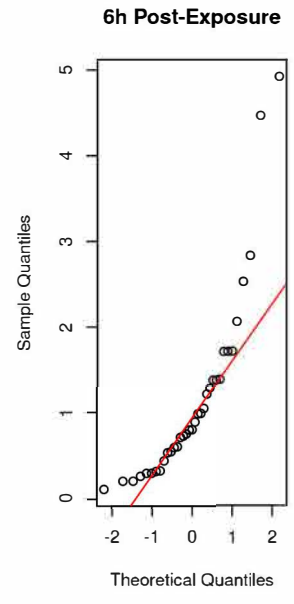
24h Post-Exposure



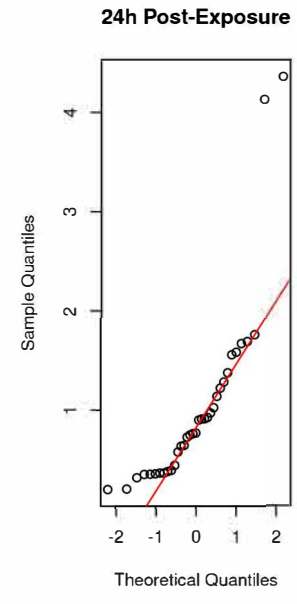
TNFa: Pre-Exposure



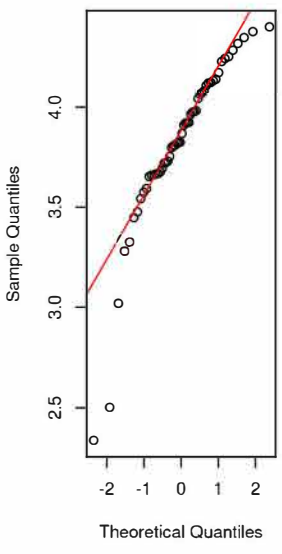
6h Post-Exposure



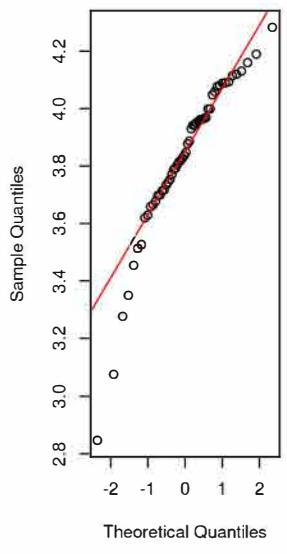
24h Post-Exposure



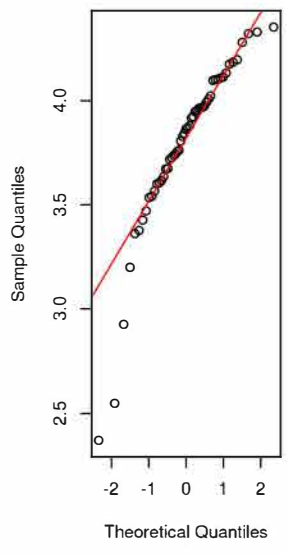
Shannon: Pre-Exposure



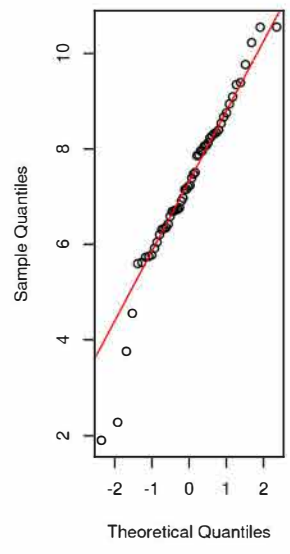
6h Post-Exposure



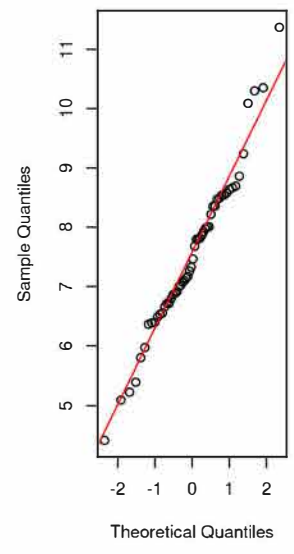
24h Post-Exposure



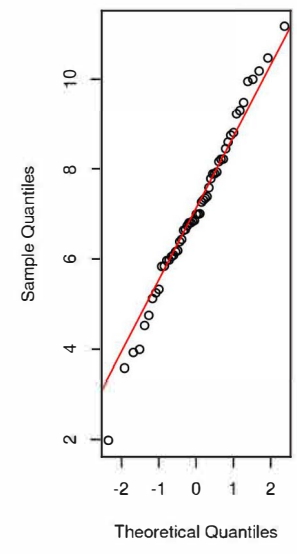
Faith PD: Pre-Exposure

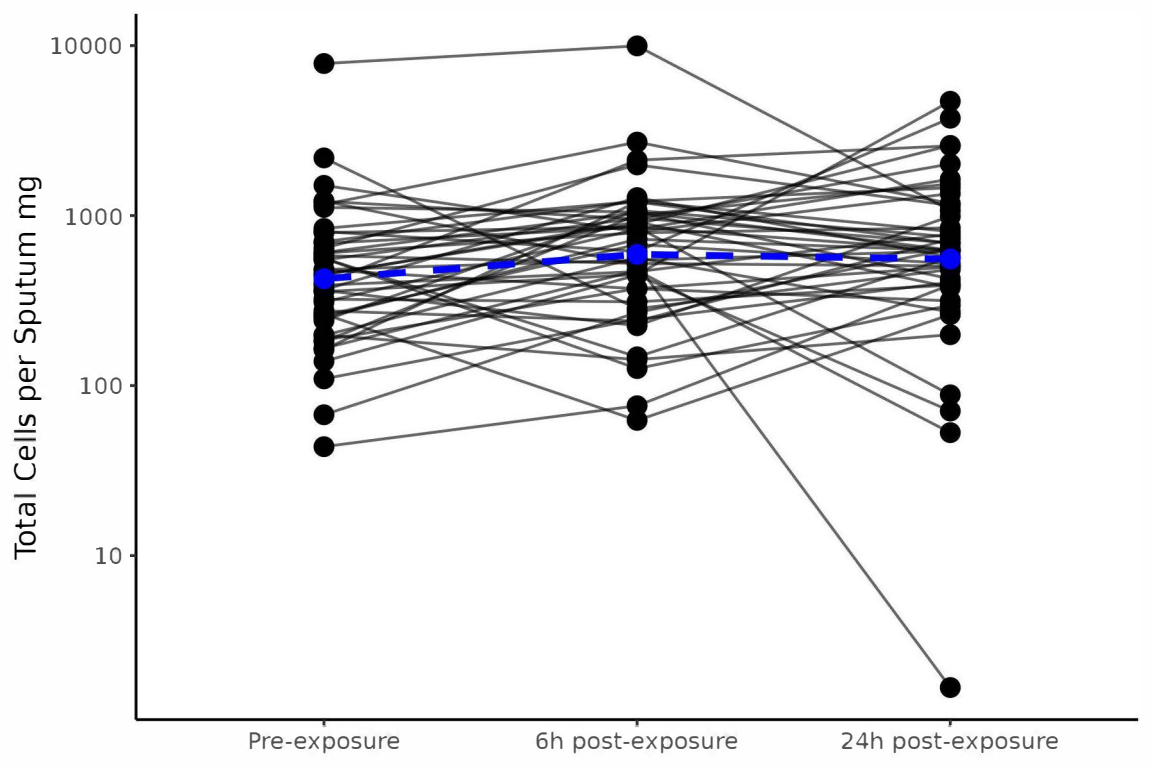


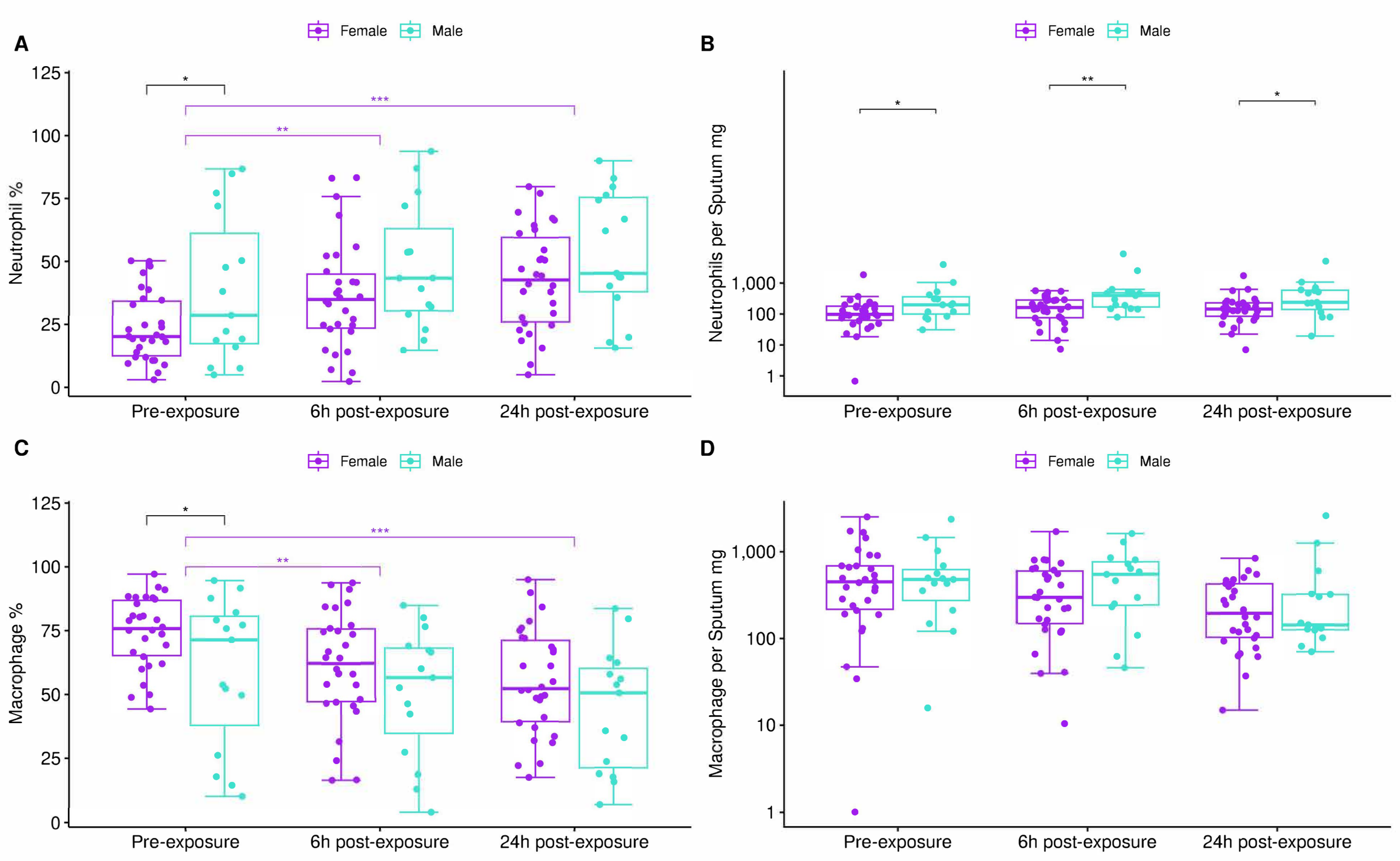
6h Post-Exposure

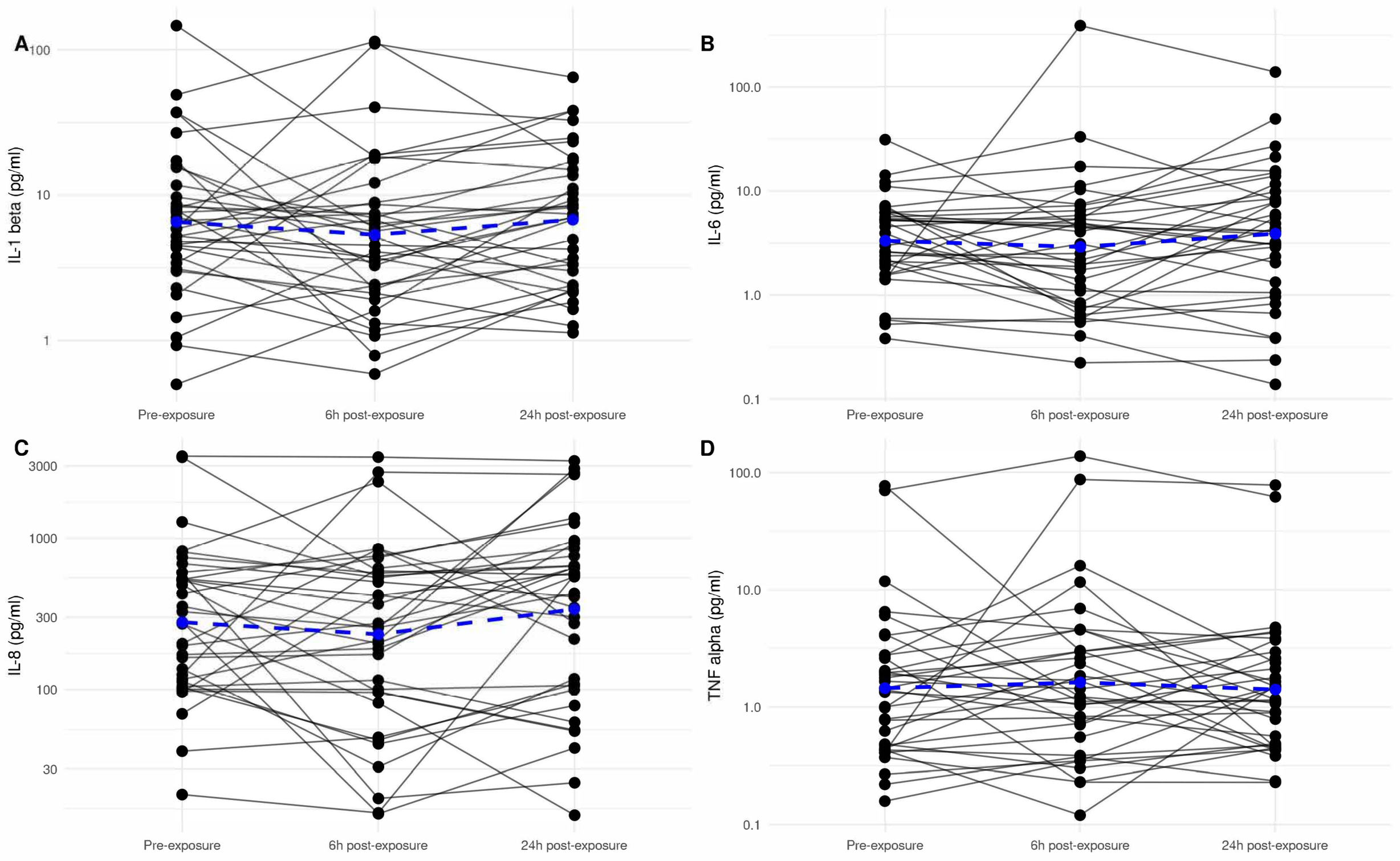


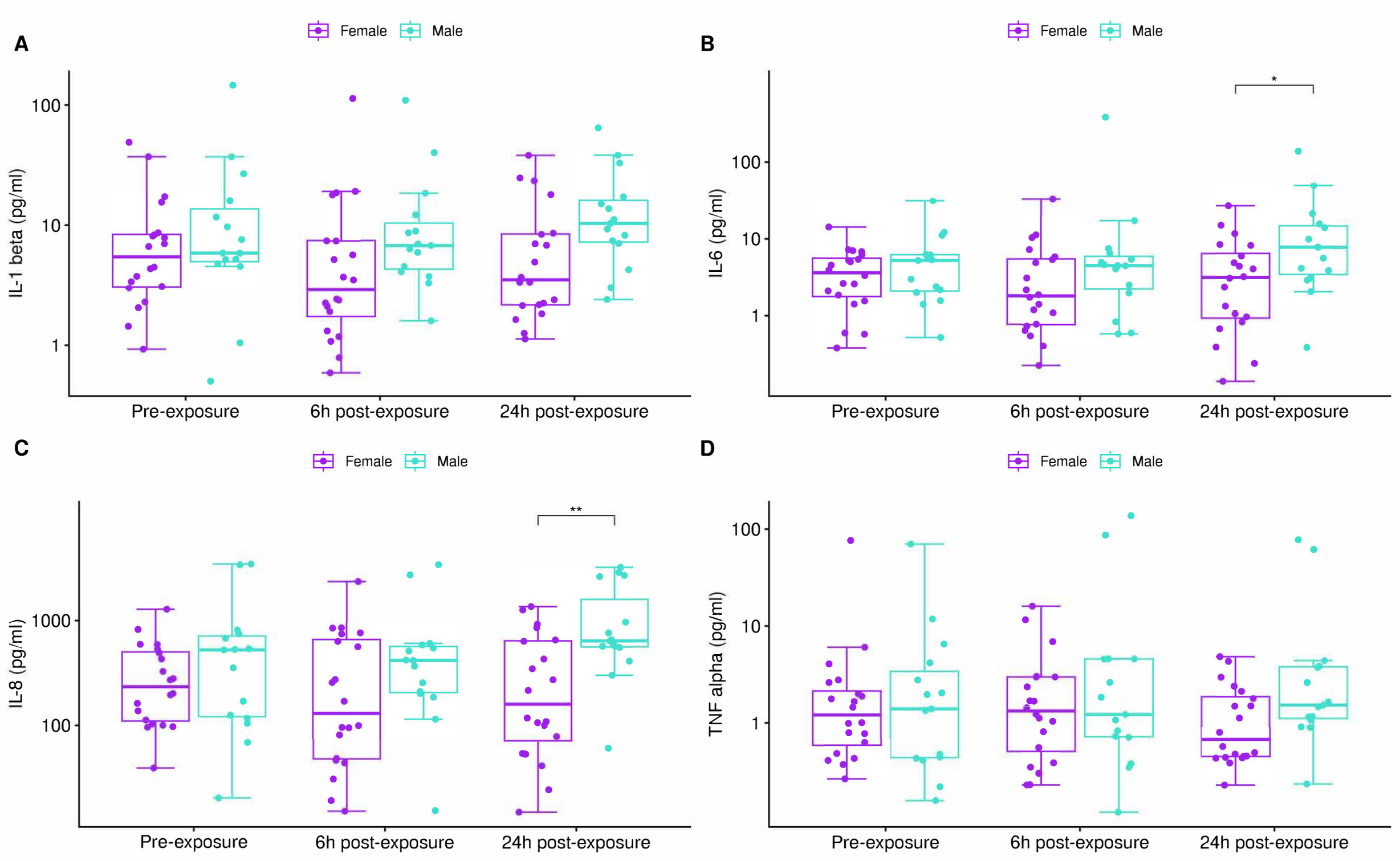
24h Post-Exposure

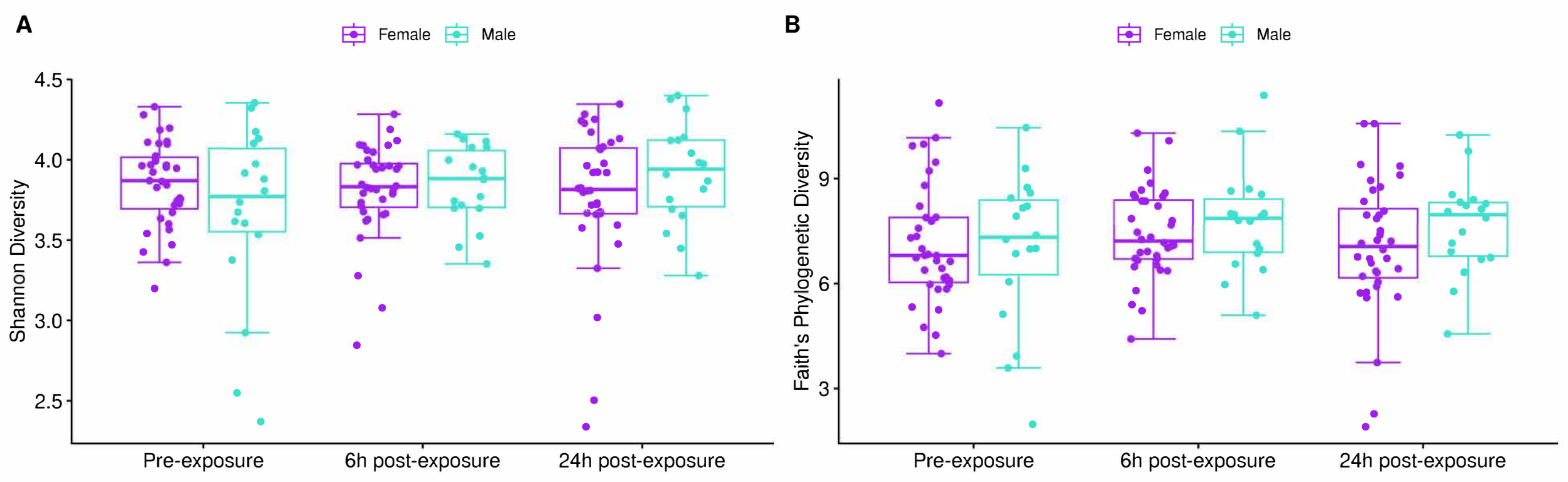


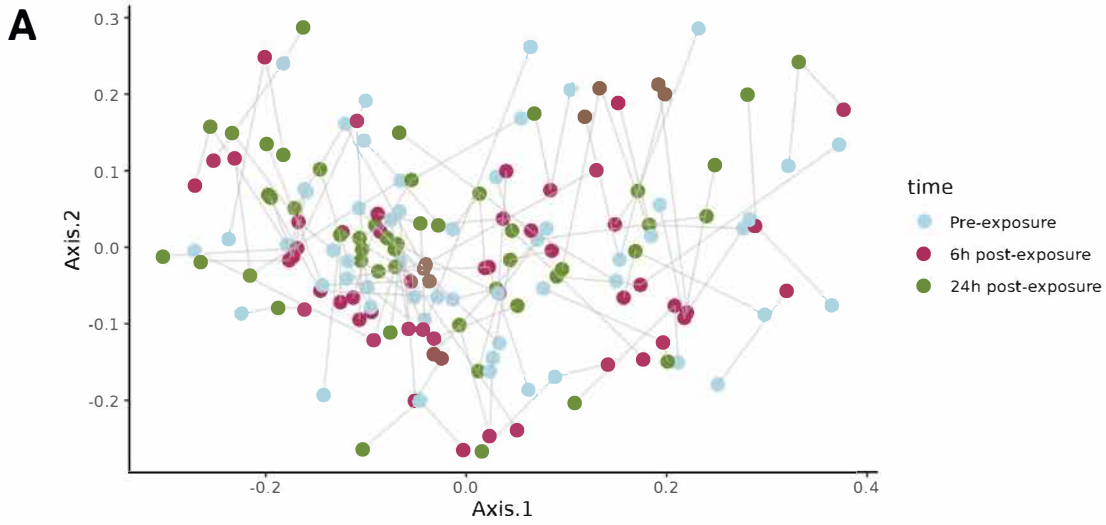

Total Cells per Sputum mg
10000
1000
100
10
Pre-exposure
6h post-exposure
24h post-exposure

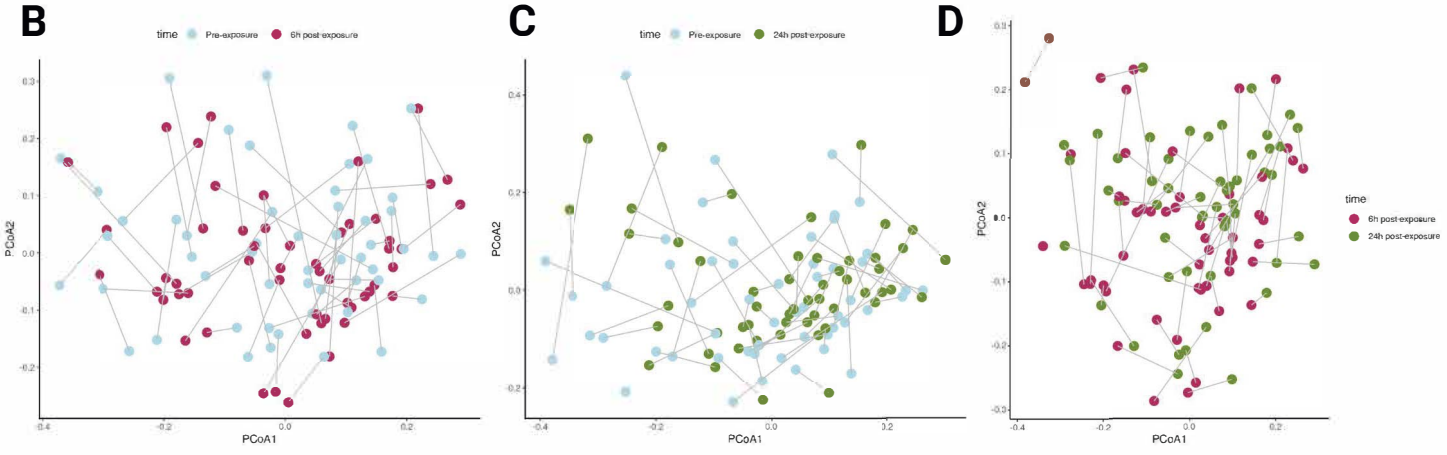

A
Female
Male
Neutrophil %
125
100
75
50
25
0
*
**
***
Pre-exposure
6h post-exposure
24h post-exposure
B
Female
Male
Neutrophils per Sputum mg
1,000
100
10
1
*
**
*
Pre-exposure
6h post-exposure
24h post-exposure
C
Female
Male
Macrophage %
125
100
75
50
25
0
*
**
***
Pre-exposure
6h post-exposure
24h post-exposure
D
Female
Male
Macrophage per Sputum mg
1,000
100
10
1
Pre-exposure
6h post-exposure
24h post-exposure

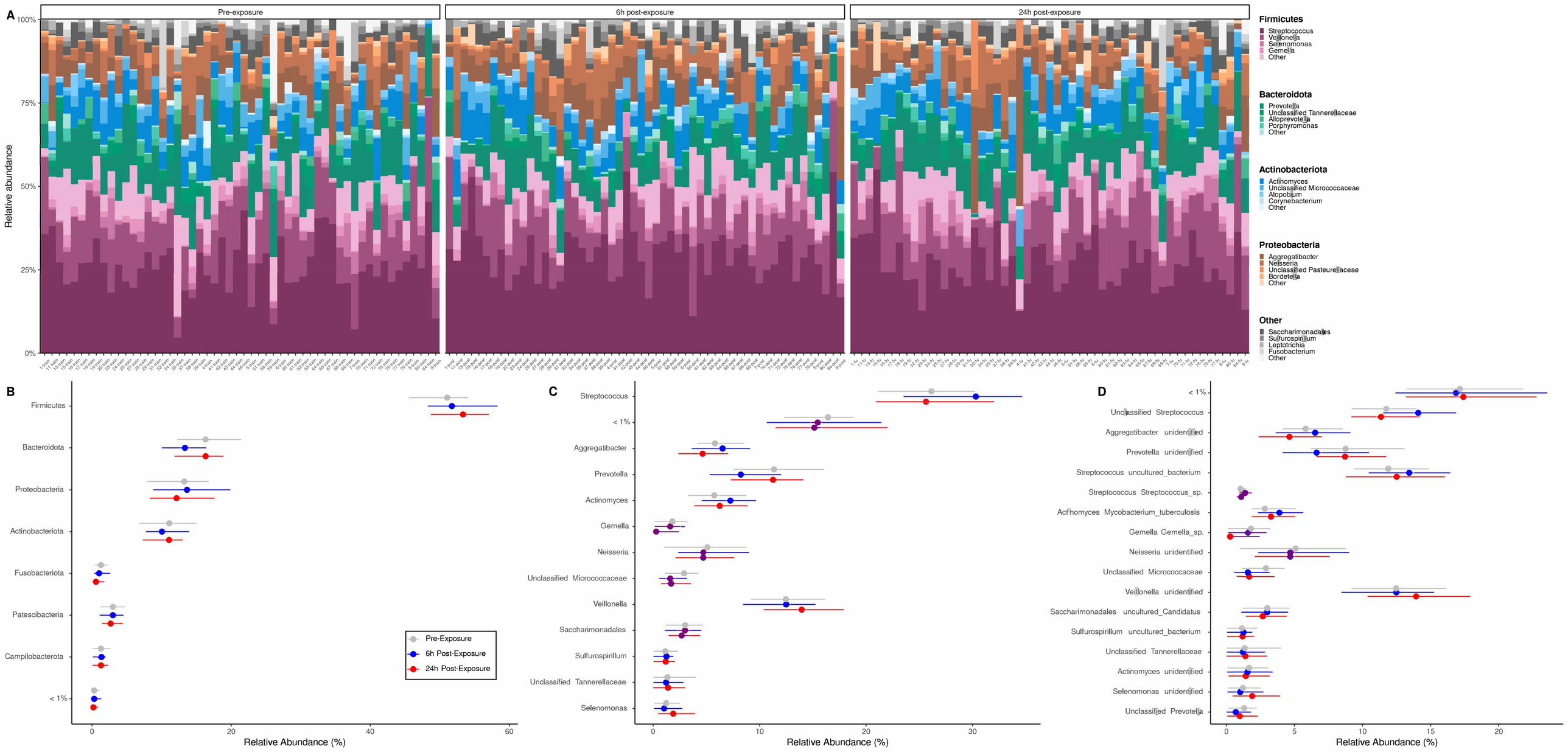

A
100
10
1
IL-1 beta (pg/ml)
Pre-exposure
6h post-exposure
24h post-exposure
B
100.0
10.0
1.0
0.1
IL-6 (pg/ml)
Pre-exposure
6h post-exposure
24h post-exposure
C
3000
1000
300
100
30
IL-8 (pg/ml)
Pre-exposure
6h post-exposure
24h post-exposure
D
100.0
10.0
1.0
0.1
TNF alpha (pg/ml)
Pre-exposure
6h post-exposure
24h post-exposure

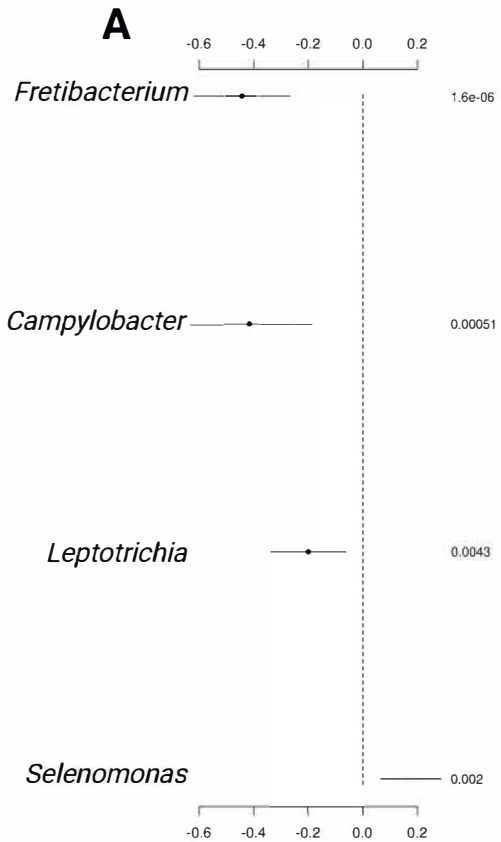

A
Female
Male
IL-1 beta (pg/ml)
100
10
1
Pre-exposure
6h post-exposure
24h post-exposure
B
Female
Male
*
IL-6 (pg/ml)
100
10
1
Pre-exposure
6h post-exposure
24h post-exposure
C
Female
Male
**
IL-8 (pg/ml)
1000
100
Pre-exposure
6h post-exposure
24h post-exposure
D
Female
Male
TNF alpha (pg/ml)
100
10
1
Pre-exposure
6h post-exposure
24h post-exposure

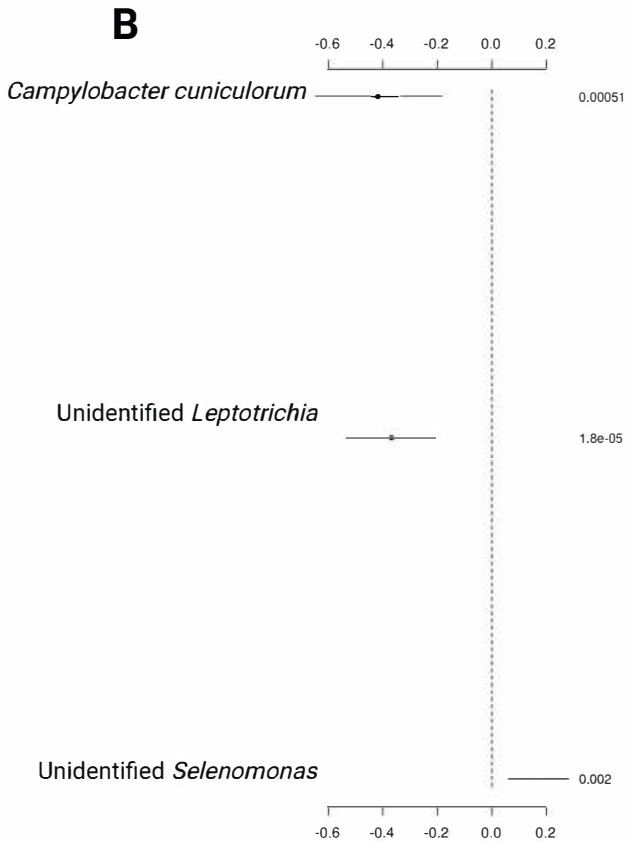

A
Female
Male
Shannon Diversity
4.5
4.0
3.5
3.0
2.5
Pre-exposure
6h post-exposure
24h post-exposure
B
Female
Male
Faith's Phylogenetic Diversity
9
6
3
Pre-exposure
6h post-exposure
24h post-exposure

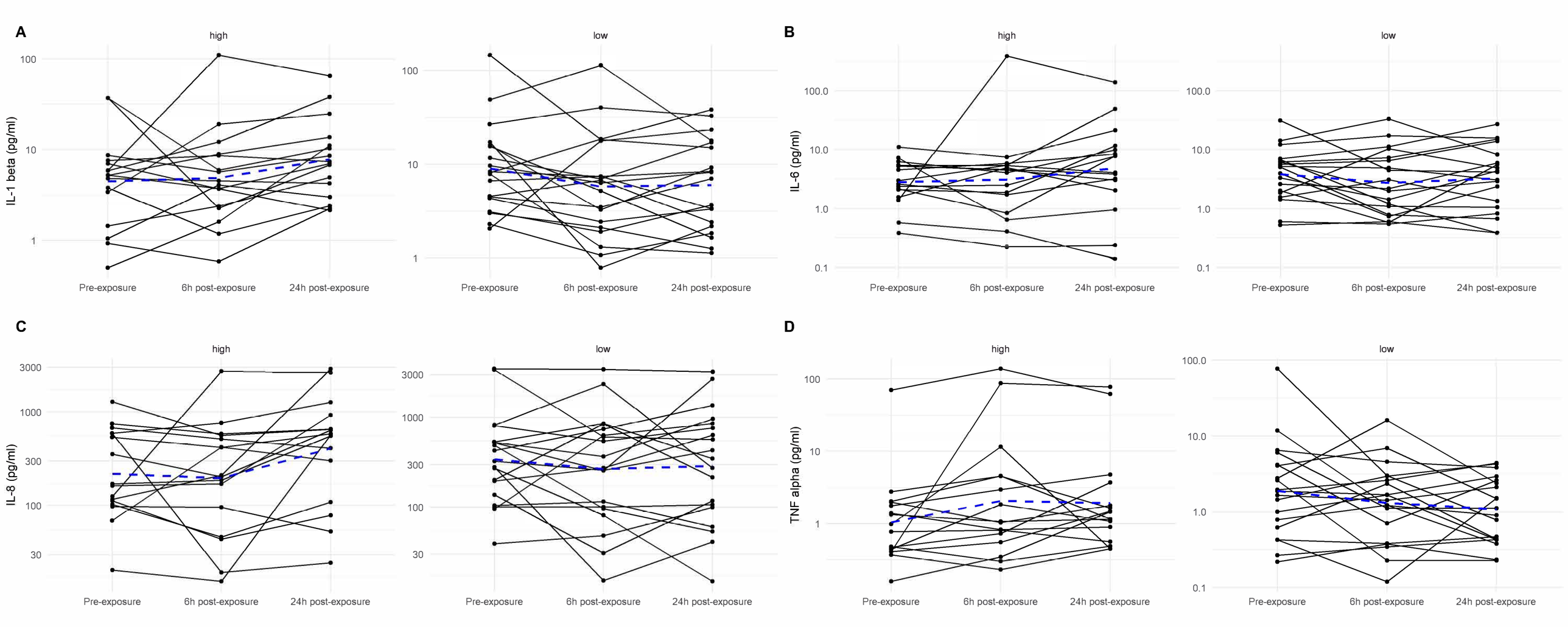

A
0.3
0.2
0.1
0.0
-0.1
-0.2
Axis.2
-0.2
0.0
0.2
0.4
Axis.1
time
Pre-exposure
6h post-exposure
24h post-exposure


B
time
Pre-exposure
6h post-exposure
PCoA2
PCoA1
C
time
Pre-exposure
24h post-exposure
PCoA2
PCoA1
D
PCoA2
PCoA1
time
6h post-exposure
24h post-exposure


A
Pre-exposure
6h post-exposure
24h post-exposure
Relative abundance
100%
75%
50%
25%
0%
Firmicutes
Streptococcus
Veillonella
Selenomonas
Gemella
Other
Bacteroidota
Prevotella
Unclassified Tannerellaceae
Alloprevotella
Porphyromonas
Other
Actinobacteriota
Actinomyces
Unclassified Micrococcaceae
Atopobium
Corynebacterium
Other
Proteobacteria
Aggregatibacter
Neisseria
Unclassified Pasteurellaceae
Bordetella
Other
Other
Saccharimonadales
Sulfurospirillum
Leptotrichia
Fusobacterium
Other
B
Firmicutes
Bacteroidota
Proteobacteria
Actinobacteriota
Fusobacteriota
Patescibacteria
Campilobacterota
< 1%
Pre-Exposure
6h Post-Exposure
24h Post-Exposure
0
20
40
60
Relative Abundance (%)
C
Streptococcus
< 1%
Aggregatibacter
Prevotella
Actinomyces
Gemella
Neisseria
Unclassified Micrococcaceae
Veillonella
Saccharimonadales
Sulfurospirillum
Unclassified Tannerellaceae
Selenomonas
0
10
20
30
Relative Abundance (%)
D
< 1%
Unclassified Streptococcus
Aggregatibacter unidentified
Prevotella unidentified
Streptococcus uncultured_bacterium
Streptococcus Streptococcus_sp.
Actinomyces Mycobacterium_tuberculosis
Gemella Gemella_sp.
Neisseria unidentified
Unclassified Micrococcaceae
Veillonella unidentified
Saccharimonadales uncultured_Candidatus
Sulfurospirillum uncultured_bacterium
Unclassified Tannerellaceae
Actinomyces unidentified
Selenomonas unidentified
Unclassified Prevotella
0
5
10
15
20
Relative Abundance (%)


A
-0.6
-0.4
-0.2
0.0
0.2
Fretibacterium
1.6e-06
Campylobacter
0.00051
Leptotrichia
0.0043
Selenomonas
0.002
-0.6
-0.4
-0.2
0.0
0.2


B
-0.6
-0.4
-0.2
0.0
0.2
Campylobacter cuniculorum
0.00051
Unidentified Leptotrichia
1.8e-05
Unidentified Selenomonas
0.002
-0.6
-0.4
-0.2
0.0
0.2


A
high
low
IL-1 beta (pg/ml)
100
10
1
Pre-exposure
6h post-exposure
24h post-exposure
B
high
low
IL-6 (pg/ml)
100.0
10.0
1.0
0.1
Pre-exposure
6h post-exposure
24h post-exposure
C
high
low
IL-8 (pg/ml)
3000
1000
300
100
30
Pre-exposure
6h post-exposure
24h post-exposure
D
high
low
TNF alpha (pg/ml)
100
10
1
100.0
10.0
1.0
0.1
Pre-exposure
6h post-exposure
24h post-exposure

References

1. Alexis NE, Zhou LY, Burbank AJ, Almond M, Hernandez ML, Mills KH, et al. Development of a screening protocol to identify persons who are responsive to wood smoke particle-induced airway inflammation with pilot assessment of GSTM1 genotype and asthma status as response modifiers. Inhal Toxicol. 2022;34(11–12):329–39.

2. Ghio AJ, Soukup JM, Case M, Dailey LA, Richards J, Berntsen J, et al. Exposure to wood smoke particles produces inflammation in healthy volunteers. Occup Environ Med. 2012 Mar 1;69(3):170–5.

3. Alexis N, Soukup J, Ghio A, Becker S. Sputum Phagocytes from Healthy Individuals Are Functional and Activated: A Flow Cytometric Comparison with Cells in Bronchoalveolar Lavage and Peripheral Blood. Clin Immunol. 2000 Oct 1;97(1):21–32.

4. Illumina. BCL Convert: a proprietary Illumina software for the conversion of BCL files to basecalls. [Internet]. 2021 [cited 2023 Nov 21]. Available from: https://support-docs.illumina.com/SW/BCL_Convert/Content/SW/FrontPages/BCL_Convert.htm

5. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019 Aug;37(8):852–7.

6. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011 May 2;17(1):10–2.

7. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High resolution sample inference from Illumina amplicon data. Nat Methods. 2016 Jul;13(7):581–3.

8. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013 Jan 1;41(D1):D590–6.

9. Escapa IF, Chen T, Huang Y, Gajare P, Dewhirst FE, Lemon KP. New Insights into Human Nostril Microbiome from the Expanded Human Oral Microbiome Database (eHOMD): a Resource for the Microbiome of the Human Aerodigestive Tract. mSystems. 2018 Dec 4;3(6):e00187-18.

10. Bisanz JE. qiime2R: Importing QIIME2 artifacts and associated data into R sessions [Internet]. 2018 [cited 2024 Jun 18]. Available from: https://github.com/jbisanz/qiime2R

11. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome. 2018 Dec 17;6(1):226.

12. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2021. Available from: https://www.R-project.org/

13. Posit team. RStudio: Integrated Development Environment for R [Internet]. Boston, MA: Posit Software, PBC; 2023. Available from: http://www.posit.co/

14. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE. 2013 abr;8(4):e61217.

15. Battaglia, Thomas. GitHub. 2015 [cited 2024 Jun 18]. btools: A suite of R function for all types of microbial diversity analyses. Available from: https://github.com/twbattaglia/btools/blob/master/README.md

16. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al. vegan: Community Ecology Package [Internet]. 2022. Available from: https://CRAN.R-project.org/package=vegan

17. Dahl EM, Neer E, Bowie KR, Leung ET, Karstens L. microshades: An R Package for Improving Color Accessibility and Organization of Microbiome Data. Microbiol Resour Announc. 2022;11(11):e00795-22.

18. Zhang X, Yi N. NBZIMM: negative binomial and zero-inflated mixed models, with application to microbiome/metagenomics data analysis. BMC Bioinformatics. 2020 Oct 30;21(1):488.

19. Zhang X, Mallick H, Tang Z, Zhang L, Cui X, Benson AK, et al. Negative binomial mixed models for analyzing microbiome count data. BMC Bioinformatics. 2017 Jan 3;18(1):4.

Supplementary Table E1. Contaminant Sequences Identified Using the Prevalence Method of the *decontam* R package

	Kingdom	Phylum	Class	Order	Family	Genus	Species
6dbcc01ebc2afd8182d70663db558527	Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	cepacia
8c346b3fac6d07017ead998035914b88	Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	cepacia
5f8b99d96b9fb65e12c944b3f8bd38b9	Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	cepacia
ffc36e27c82042664a16bcd4d380b286	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	Escherichia	coli
d55b64807e7ffc581671f5ce1f0695c0	Bacteria	Proteobacteria	Alphaproteobacteria	Hyphomicrobiales	Brucellaceae	Brucella	anthropi
dd8c8d60bfca3e1df01a4bbff8d3326e	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	glacialis
89a04ae48354976f71d111d52a8609e8	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium	humi
283be7225e0ecdd411525948112368a2	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium	humi
9db2817f5c42be6a7bcbca662959982d	Bacteria	Proteobacteria	Alphaproteobacteria	Hyphomicrobiales	Bradyrhizobiaceae	*NA*	*NA*
7f94ba24ae3488c6a0fcbe0587b78c59	Bacteria	Proteobacteria	Alphaproteobacteria	Hyphomicrobiales	Bradyrhizobiaceae	*NA*	*NA*
455bb9658aeda2552edddb8e918b1472	Bacteria	Proteobacteria	Alphaproteobacteria	Hyphomicrobiales	Bradyrhizobiaceae	*NA*	*NA*
7507ee1be078be2cf7444825cdc58b25	Bacteria	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Brevundimonas	diminuta
2cd30e9eda91687ba73dcae76f4f1880	Bacteria	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Brevundimonas	diminuta
b02a8d33d018119dedb2db15db887bfd	Bacteria	Actinobacteria	Actinomycetia	Propionibacteriales	Propionibacteriaceae	Cutibacterium	acnes
3c5f98e32eb24d418090f2d7f9f17015	Bacteria	Firmicutes	Clostridia	Eubacteriales	Peptostreptococcaceae	Peptostreptococcaceae[G5]	HMT493
3fe8608797c2760295ca61479f630cb6	Bacteria	Firmicutes	Clostridia	Eubacteriales	Peptostreptococcaceae	Mogibacterium	*NA*
5497318e515a8c328a68f95975d9c7d4	Bacteria	Firmicutes	Bacilli	Bacillales	Staphylococcaceae	Staphylococcus	capitis

Supplementary Table E2. Relative Abundance and Prevalence of Contaminant Sequences in Negative Controls and True Sputum Samples

				Mean Relative Abundance (%)		Prevalence	
	Family	Genus	Species	Negative controls	Sputum samples	Negative controls	Sputum samples
283be7225e0ecdd411525948112368a2	Sphingomonadaceae	Novosphingobium	humi	2.41105585	0	0.2	0
2cd30e9eda91687ba73dcae76f4f1880	Caulobacteraceae	Brevundimonas	diminuta	2.57331265	0	0.2	0
3c5f98e32eb24d418090f2d7f9f17015	Peptostreptococcaceae	Peptostreptococcaceae[G-5]	bacterium HMT493	0.94598605	0.002226525	0.1	0.00621118
3fe8608797c2760295ca61479f630cb6	Peptostreptococcaceae	Mogibacterium	NA	0.06083894	0.000701009	0.1	0.01242236
455bb9658aeda2552edddb8e918b1472	Bradyrhizobiaceae	NA	NA	0.21822894	0	0.2	0
5497318e515a8c328a68f95975d9c7d4	Staphylococcaceae	Staphylococcus	capitis	0.52114865	0.001443256	0.1	0.01242236
5f8b99d96b9fb65e12c944b3f8bd38b9	Burkholderiaceae	Burkholderia	cepacia	0.40582216	0	0.2	0
6dbcc01ebc2afd8182d70663db558527	Burkholderiaceae	Burkholderia	cepacia	5.67507778	0	0.3	0
7507ee1be078be2cf7444825cdc58b25	Caulobacteraceae	Brevundimonas	diminuta	0.06648661	0	0.2	0
7f94ba24ae3488c6a0fcbe0587b78c59	Bradyrhizobiaceae	NA	NA	0.31810694	0	0.4	0
89a04ae48354976f71d111d52a8609e8	Sphingomonadaceae	Novosphingobium	humi	0.23926992	0	0.2	0
8c346b3fac6d07017ead998035914b88	Burkholderiaceae	Burkholderia	cepacia	0.27966056	0	0.2	0
9db2817f5c42be6a7bcbca662959982d	Bradyrhizobiaceae	NA	NA	3.44972578	0	0.4	0
b02a8d33d018119dedb2db15db887bfd	Propionibacteriaceae	Cutibacterium	acnes	3.66583601	0	0.4	0
d55b64807e7ffc581671f5ce1f0695c0	Brucellaceae	Brucella	anthropi	1.32335596	0	0.2	0
dd8c8d60bfca3e1df01a4bbff8d3326e	Sphingomonadaceae	Sphingomonas	glacialis	5.9653503	0	0.4	0
ffc36e27c82042664a16bcd4d380b286	Enterobacteriaceae	Escherichia	coli	1.90998961	0	0.3	0

Supplementary Table E3. Beta Diversity PERMANOVA Results

	Df	SumOfSqs	R2	F	Pr(>F)
Exposure (hpe)	2	0.388	0.01024	0.817	0.001
Residual	158	37.480	0.98976		
Total	160	37.868	1		

Supplementary Table E4. Beta Diversity PERMANOVA Results by Exposure and Sex

	Df	SumOfSqs	R2	F	Pr(>F)
Exposure (hpe)	2	0.388	0.01024	0.8195	0.002
Sex	1	0.589	0.01554	2.4890	0.005
Exposure: Sex	2	0.234	0.00618	0.4951	0.584
Residual	155	36.658	0.96803		
Total	160	37.868	1		

Supplementary Table E5. Negative Binominal Mixed Model Results for Genus-Level Microbiome Differential Abundance Analysis Using eHOMD

	Estimate	Std.Error	p-value	p-adj	Exp.Estimate	Percent Change	Increase/ Decrease
Fretibacterium	-0.442817	0.0872207	1.65 x 10-06	6.76 x 10-05	0.642224723	35.78	↓
Selenomonas	0.1875918	0.0543699	8.05 x 10-04	1.65 x 10-02	1.206340986	20.63	↑

Supplementary Table E6. Negative Binominal Mixed Model Results for Genus-Level Microbiome Differential Abundance Analysis Using eHOMD and Incorporating Race as a Covariate

	Estimate	Std.Error	p-value	p-adj	Exp.Estimate	Percent Change	Increase/ Decrease
Fretibacterium	-0.4419975	0.0882453	2.20 x 10-06	9.01 x 10-05	0.642751242	35.72	↓
Selenomonas	0.1870556	0.0550483	9.57 x 10-04	1.96 x 10-02	1.20569432	20.57	↑

Supplementary Table E7. Negative Binominal Mixed Model Results for Species-Level Microbiome Differential Abundance Analysis Using eHOMD

	Estimate	Std.Error	p-value	p-adj	Exp.Estimate	Percent Change	Increase/ Decrease
Unclassified *Fretibacterium*	-0.5081605	0.1078981	7.56 x 10-06	0.0003212	0.601601207	39.84	↓
Unclassified *Parvimonas*	-0.3697333	0.0935324	1.40 x 10-04	0.0039531	0.690918574	30.91	↓
Porphyromonas catoniae	-0.3601067	0.1080873	1.19 x 10-03	0.0168479	0.697601888	30.24	↓
***Haemophilus influenzae*-Age**	-0.3426197	0.0872579	2.69 x 10-04	0.0229021	0.709908138	29.01	↓
Neisseria cinerea	-0.2761827	0.079379	7.31 x 10-04	0.0124233	0.758674308	24.13	↓
***Alloprevotella* sp. HMT 308-Age**	0.1275712	0.0362543	9.46 x 10-04	0.0402247	1.136065753	13.61	↑
***Selenomonas* sp. HMT 478**	0.1827091	0.0621528	4.03 x 10-03	0.0489787	1.200465142	20.05	↑
Unclassified *Selenomonas*	0.2200454	0.0587238	2.91 x 10-04	0.006194	1.246133304	24.61	↑
Haemophilus influenzae	0.3013354	0.0636323	6.80 x 10-06	0.0003212	1.351662613	35.17	↑

Supplementary Table E8. Negative Binominal Mixed Model Results for Species-Level Microbiome Differential Abundance Analysis Using eHOMD and Incorporating Race as a Covariate

	Estimate	Std.Error	p-value	p-adj	Exp.Estimate	Percent Change	Increase/ Decrease
Unclassified *Fretibacterium*	-0.507625	0.108974	9.30 x 10-06	7.81 x 10-04	0.60192327	39.81	↓
***Haemophilus influenzae*-Age**	-0.421349	0.07417	8.69 x 10-07	7.30 x 10-05	0.656161389	34.38	↓
Unclassified *Parvimonas*	-0.371045	0.094455	1.53 x 10-04	6.41 x 10-03	0.690013097	31.00	↓
Porphyromonas catoniae	-0.36099	0.108453	1.20 x 10-03	2.02 x 10-02	0.696985829	30.30	↓
Neisseria cinerea	-0.278123	0.080444	7.86 x 10-04	1.65 x 10-02	0.757203831	24.28	↓
Unclassified *Selenomonas*	0.220623	0.058271	2.54 x 10-04	7.11 x 10-03	1.24685278	24.69	↑
***Haemophilus influenzae*-BMI**	0.31028	0.076839	2.02 x 10-04	1.70 x 10-02	1.363806517	36.38	↑

Supplementary Table E9. Negative Binominal Mixed Model Results for Genus-Level Host-Microbiome Association Analysis Using eHOMD

	Variable	Estimate	Std.Error	p-value	p-adj	Exp Estimate	Percent Change	Increase/ Decrease
Fretibacterium	Exposure (hpe)	-0.55138	0.088516	1.31 x 10-08	5.37 x 10-07	0.576152	42.38	↓
Fretibacterium	Macrophages/mg	-0.0016	0.000421	2.60 x 10-04	5.34 x 10-03	0.998402	0.16	↓
Catonella	Macrophages/mg	-0.00097	0.00025	2.02 x 10-04	5.34 x 10-03	0.999033	0.10	↓
Unclassified *Neisseriaceae*	Macrophages/mg	-0.00085	0.000283	3.52 x 10-03	3.61 x 10-02	0.999152	0.08	↓
***Saccharibacteria* (TM7) [G-6]**	Macrophages/mg	0.00069	0.000223	2.59 x 10-03	3.54 x 10-02	1.000691	0.07	↑

Supplementary Table E10. Negative Binominal Mixed Model Results for Genus-Level Host-Microbiome Association Analysis Using eHOMD and Incorporating Race as a Covariate

	Variable	Estimate	Std.Error	p-value	p-adj	Exp Estimate	Percent Change	Increase/ Decrease
Fretibacterium	Exposure (hpe)	-0.55325	0.090658	2.32 x 10-08	9.26 x 10-07	0.575076	42.49	↓
Fretibacterium	Macrophages/mg	-0.00164	0.00043	2.48 x 10-04	4.95 x 10-03	0.998363	0.16	↓
Catonella	Macrophages/mg	-0.00097	0.000251	2.15 x 10-04	4.95 x 10-03	0.999033	0.10	↓
Unclassified *Neisseriaceae*	Macrophages/mg	-0.00086	0.000286	3.36 x 10-03	4.48 x 10-02	0.99914	0.09	↓

Supplementary Table E11. Negative Binominal Mixed Model Results for Species-Level Host-Microbiome Association Analysis using eHOMD

	Variable	Estimate	Std.Error	p-value	p-adj	Exp.Estimate	Percent Change	Increase/ Decrease
Unclassified *Fretibacterium*	Exposure (hpe)	-0.66859	0.111931	4.12 x 10-08	3.59 x 10-06	0.512431	48.76	↓
Campylobacter rectus	Exposure (hpe)	-0.49926	0.142113	6.83 x 10-04	1.51 x 10-02	0.606982	39.30	↓
Prevotella intermedia	Exposure (hpe)	-0.45469	0.133728	9.90 x 10-04	1.51 x 10-02	0.634647	36.54	↓
Haemophilus influenzae	Age	-0.36084	0.08969	2.28 x 10-04	1.23 x 10-02	0.697092	30.29	↓
***Haemophilus sp.* HMT 908**	Exposure (hpe)	-0.33559	0.099066	1.03 x 10-03	1.51 x 10-02	0.714913	28.51	↓
Neisseria cinerea	Exposure (hpe)	-0.28815	0.085118	1.04 x 10-03	1.51 x 10-02	0.749649	25.04	↓
Unclassified *Fretibacterium*	Macrophages/mg	-0.00189	0.000526	5.36 x 10-04	9.32 x 10-03	0.998114	0.19	↓
Prevotella intermedia	Macrophages/mg	-0.00153	0.000493	2.58 x 10-03	2.81 x 10-02	0.998474	0.15	↓
***Treponema* sp. HMT 237**	Macrophages/mg	-0.00137	0.000399	8.79 x 10-04	1.27 x 10-02	0.998628	0.14	↓
Catonella morbi	Macrophages/mg	-0.00107	0.000261	8.89 x 10-05	1.93 x 10-03	0.998932	0.11	↓
Unclassified *Neisseriaceae*	Macrophages/mg	-0.00085	0.000283	3.52 x 10-03	3.06 x 10-02	0.999152	0.08	↓
***Selenomonas* sp. HMT 478**	Macrophages/mg	-0.00082	0.000272	3.20 x 10-03	3.06 x 10-02	0.999176	0.08	↓
***Saccharibacteria* (TM7) [G6] HMT 870**	Macrophages/mg	0.00069	0.000223	2.59 x 10-03	2.81 x 10-02	1.000691	0.07	↑
***Alloprevotella* sp. HMT 473**	Macrophages/mg	0.001229	0.0003	8.77 x 10-05	1.93 x 10-03	1.00123	0.12	↑
Unclassified *Parvimonas*	Macrophages/mg	0.001553	0.000355	3.07 x 10-05	1.34 x 10-03	1.001554	0.16	↑
***Porphyromonas* sp. HMT 278**	Macrophages/mg	0.001696	0.000353	5.91 x 10-06	5.14 x 10-04	1.001697	0.17	↑
***Alloprevotella* sp. HMT 308**	Age	0.143162	0.04193	1.41 x 10-03	4.08 x 10-02	1.153917	15.39	↑
Stomatobaculum longum	Age	0.161942	0.040968	2.83 x 10-04	1.23 x 10-02	1.175792	17.58	↑
Unclassified *Selenomonas*	Exposure (hpe)	0.236566	0.072645	1.57 x 10-03	1.95 x 10-02	1.266891	26.69	↑
Haemophilus influenzae	Exposure (hpe)	0.401574	0.078013	1.44 x 10-06	6.28 x 10-05	1.494174	49.42	↑

Supplementary Table E12. Negative Binominal Mixed Model Results for Species-Level Host-Microbiome Association Analysis using eHOMD and Incorporating Race as a Covariate

	Variable	Estimate	Std.Error	p-value	p-adj	Exp.Estimate	Percent Change	Increase/ Decrease
Alloprevotella sp. HMT 308	Sex (Male)	-3.02855	0.439122	2.61 x 10-08	2.06 x 10-06	0.048386	95.16	↓
Streptococcus parasanguinis clade 411	Sex (Male)	-0.80719	0.238385	1.60 x 10-03	4.46 x 10-02	0.44611	55.39	↓
Unclassified Fretibacterium	Exposure (hpe)	-0.67436	0.114705	6.26 x 10-08	4.94 x 10-06	0.509481	49.05	↓
Campylobacter rectus	Exposure (hpe)	-0.50725	0.143344	6.27 x 10-04	2.16 x 10-02	0.602151	39.78	↓
Prevotella intermedia	Exposure (hpe)	-0.45061	0.134675	1.18 x 10-03	2.16 x 10-02	0.637241	36.28	↓
Haemophilus sp. HMT 908	Exposure (hpe)	-0.33869	0.101631	1.23 x 10-03	2.16 x 10-02	0.712701	28.73	↓
Neisseria cinerea	Exposure (hpe)	-0.28912	0.087623	1.37 x 10-03	2.16 x 10-02	0.748926	25.11	↓
Unclassified Fretibacterium	Macrophages/mg	-0.00196	0.000538	4.53 x 10-04	8.95 x 10-03	0.998045	0.20	↓
Prevotella intermedia	Macrophages/mg	-0.00155	0.000501	2.63 x 10-03	3.32 x 10-02	0.998452	0.15	↓
Treponema sp. HMT 237	Macrophages/mg	-0.00139	0.000416	1.16 x 10-03	1.83 x 10-02	0.998609	0.14	↓
Catonella morbi	Macrophages/mg	-0.00098	0.000266	3.70 x 10-04	8.95 x 10-03	0.99902	0.10	↓
Unclassified Neisseriaceae	Macrophages/mg	-0.00086	0.000286	3.36 x 10-03	3.32 x 10-02	0.99914	0.09	↓
Selenomonas sp. HMT 478	Macrophages/mg	-0.00083	0.000274	3.18 x 10-03	3.32 x 10-02	0.999169	0.08	↓
Alloprevotella sp. HMT 473	Macrophages/mg	0.001429	0.000299	6.37 x 10-06	2.52 x 10-04	1.00143	0.14	↑
Porphyromonas sp. HMT 278	Macrophages/mg	0.001812	0.000362	2.52 x 10-06	1.99 x 10-04	1.001814	0.18	↑
Stomatobaculum longum	Age	0.178938	0.040569	7.57 x 10-05	2.99 x 10-03	1.195946	19.59	↑
Alloprevotella sp. HMT 308	Age	0.190333	0.040621	3.21 x 10-05	2.53 x 10-03	1.209652	20.97	↑
Unclassified Selenomonas	Exposure (hpe)	0.231395	0.072505	1.93 x 10-03	2.53 x 10-02	1.260357	26.04	↑
Alloprevotella sp. HMT 473	Sex (Male)	1.822075	0.541339	1.69 x 10-03	4.46 x 10-02	6.184676	518.47	↑

Supplementary Table E13. Negative Binominal Mixed Model Results for Genus-Level Microbiome Differential Abundance Analysis Using SILVA

	Estimate	Std.Error	p-value	p-adj	Exp.Estimate	Percent Change	Increase/ Decrease
Fretibacterium	-0.442817	0.0872207	1.65 x 10-06	6.60 x 10-05	0.642224723	35.78	↓
Campylobacter	-0.4147313	0.115686	5.11 x 10-04	1.02 x 10-02	0.660517738	33.95	↓
Leptotrichia	-0.197746	0.0677126	4.27 x 10-03	4.27 x 10-02	0.820578254	17.94	↓
Selenomonas	0.1751485	0.0553511	2.03 x 10-03	2.71 x 10-02	1.19142313	19.14	↑

Supplementary Table E14. Negative Binominal Mixed Model Results for Species-Level Microbiome Differential Abundance Analysis Using SILVA

	Estimate	Std.Error	p-value	p-adj	Exp.Estimate	Percent Change	Increase/ Decrease
Campylobacter cuniculorum	-0.4147313	0.11568601	5.11 x 10-04	0.012014716	0.660517738	33.95	↓
Unidentified *Leptotrichia*	-0.3683424	0.08195423	1.79 x 10-05	0.000839771	0.691880241	30.81	↓
Unidentified *Selenomonas*	0.1751485	0.05535112	2.03 x 10-03	0.031787511	1.19142313	19.14	↑

Supplementary Table E15. Negative Binominal Mixed Model Results for Genus-Level Host-Microbiome Association Analysis Using SILVA

	Variable	Estimate	Std.Error	p-value	p-adj	Exp.Estimate	Percent Change	Increase/ Decrease
Fretibacterium	Exposure(hpe)	-0.55138	0.088516	1.31 x 1008	5.24 x 1007	0.576152	42.38	↓
Campylobacter	Exposure(hpe)	-0.52979	0.141547	3.13 x 1004	6.26 x 1003	0.588728	41.13	↓
Fretibacterium	Macrophages per mg	-0.0016	0.000421	2.60 x 1004	5.21 x 1003	0.998402	0.16	↓
Catonella	Macrophages per mg	-0.00097	0.00025	2.02 x 1004	5.21 x 1003	0.999033	0.10	↓
Unclassified *Burkholderiales*	Exposure(hpe)	0.294879	0.093209	2.10 x 1003	2.80 x 1002	1.342964	34.30	↑

Supplementary Table E16. Negative Binominal Mixed Model Results Species-Level Host-Microbiome Association Analysis Using SILVA

	Variable	Estimate	Std.Error	p-value	p-adj	Exp.Estimate	Percent Change	Increase/ Decrease
Campylobacter cuniculorum	Exposure (hpe)	-0.52979	0.141547	0.000313	0.014718	0.588728	41.13	↓
***Corynebacterium* unidentified**	Exposure (hpe)	-0.29927	0.096419	0.002524	0.029657	0.741362	25.86	↓
***Leptotrichia* unidentified**	Exposure (hpe)	-0.29654	0.094433	0.002257	0.029657	0.743388	25.66	↓
***Catonella* unidentified**	Macrophages per mg	-0.00097	0.00025	0.000202	0.009478	0.999033	0.10	↓
Unclassified *Burkholderiales*	Exposure (hpe)	0.294879	0.093209	0.002099	0.029657	1.342964	34.30	↑