Associations of air pollution and greenness with the nasal microbiota of healthy infants: A longitudinal study

Amanda Gisler a,b,1, Insa Korten a,b,1, Kees de Hoogh c,d, Danielle Vienneau c,d, Urs Frey a,b, Fabienne Decru c,e, Olga Gorlanova a,b, Andras Soti a,b, Markus Hilty e, Philipp Latzin a,b, Jakob Usemann a,b,*, on behalf of the BILD study group

a University Children’s Hospital Basel, University of Basel, Basel, Switzerland
b Division of Paediatric Respiratory Medicine and Allergology, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland
c Institute for Infectious Diseases, University of Bern, Bern, Switzerland
d University of Basel, Basel, Switzerland
e Institute for Tropical and Public Health Institute, Basel, Switzerland

ABSTRACT

Keywords:
Infancy
Respiratory disease
Microbiota
NDVI
Particulate matter
Pneumococcal conjugate vaccine; rRNA, ribosomal RNA; PCR, Polymerase chain reaction; NDVI, Normalized difference vegetation index; PM2.5, Particulate matter ≤2.5 μm; NO2, Nitrogen dioxide; BILD, Bern-Basel Infant Lung Development.

Background: Air pollution and greenness are associated with short- and long-term respiratory health in children but the underlying mechanisms are only scarcely investigated. The nasal microbiota during the first year of life has been shown to be associated with respiratory tract infections and asthma development. Thus, an interplay between greenness, air pollution and the early nasal microbiota may contribute to short- and long-term respiratory health. We aimed to examine associations between fine particulate matter (PM2.5), nitrogen dioxide (NO2) and greenness with the nasal microbiota of healthy infants during the first year of life in a European context with low-to-moderate air pollution levels.

Methods: Microbiota characterization was performed using 16 S rRNA pyrosequencing of 846 nasal swabs collected fortnightly from 47 healthy infants of the prospective Basel-Bern Infant Lung Development (BILD) cohort. We investigated the association of satellite-based greenness and an 8-day-average exposure to air pollution (PM2.5, NO2) with the nasal microbiota during the first year of life. Exposures were individually estimated with novel spatial-temporal models incorporating satellite data. Generalized additive mixed models adjusted for known confounders and considering the autoregressive correlation structure of the data were used for analysis.

Results: Mean (SD) PM2.5 level was 17.1 (3.8 μg/m³) and mean (SD) NO2 level was 19.7 (7.9 μg/m³). Increased PM2.5 and increased NO2 were associated with reduced within-subject Ruzicka dissimilarity (PM2.5: per 1 μg/m³ -0.004, 95% CI -0.008, −0.001; NO2: per 1 μg/m³ -0.004, 95% CI -0.007, −0.001). Whole microbial community comparison with nonmetric multidimensional scaling revealed distinct microbiota profiles for different PM2.5 exposure levels. Increased NO2 was additionally associated with reduced abundance of Corynebacteriaceae (per 1 μg/m³: −0.027, 95% CI -0.053, −0.001). No associations were found between greenness and the nasal microbiota.

Conclusion: Air pollution was associated with Ruzicka dissimilarity and relative abundance of Corynebacteriaceae. This suggests that even low-to-moderate exposure to air pollution may impact the nasal microbiota during the first year of life. Our results will be useful for future studies assessing the clinical relevance of air-pollution-induced alterations of the nasal microbiota with subsequent respiratory disease development.

Abbreviations: GAMM, Generalized additive mixed model; SDI, Shannon diversity index; OT, Oligotype; NMDS, Non-metric multidimensional scaling; PCV, Pneumococcal conjugate vaccine; rRNA, ribosomal RNA; PCR, Polymerase chain reaction; NDVI, Normalized difference vegetation index; PM2.5, Particulate matter ≤2.5 μm; NO2, Nitrogen dioxide; BILD, Bern-Basel Infant Lung Development.

* Corresponding author. University Children’s Hospital Basel (UKBB) Spitalstrasse 33, 4056, Basel, Switzerland.
E-mail address: jakob.usemann@akbb.ch (J. Usemann).

1 These authors contributed equally.

https://doi.org/10.1016/j.envres.2021.111633
Received 30 March 2021; Received in revised form 29 June 2021; Accepted 30 June 2021
Available online 10 July 2021
0013-9361/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
1. Introduction

Respiratory infections and asthma constitute major health issues in infants and children (Kyu et al., 2016; Lozano et al., 1999; Zar and Ferkol, 2014). Several studies in these age groups have shown that air pollution and residential greenness are critical risk factors for acute and chronic respiratory diseases (Bowatte et al., 2015; Fuertes et al., 2020; Lambert et al., 2017, 2019; Sbihi et al., 2017; Stern et al., 2013; To et al., 2020; Usemann et al., 2019; Zhang et al., 2019). The relevance of these environmental risk factors is amplified as, by nature, they are omnipresent and able to affect all individuals (Kim et al., 2014; Mihireach et al., 2016). In particular, infants belong to the most vulnerable group due to immaturity of their immune and respiratory system (Bowatte et al., 2015).

The underlying mechanisms for the association of air pollution and greenness with respiratory health are only scarcely investigated. However, the nasal microbiota of the first year of life has been identified as an important player in the progression of respiratory infections (Bosch et al., 2017; Toivonen et al., 2019) and asthma development (Bisgaard et al., 2007; To et al., 2019). Certain microbiota profiles have been linked to disease progression; others have been associated with improved respiratory health. Previous studies have found that the development of the nasal microbiota in the first year of life is driven by age of infant, mode of delivery, feeding type, childcare attendance, viruses and presence of siblings (Bosch et al., 2016, 2017; Hasegawa et al., 2017; Mika et al., 2015). In addition, first studies revealed an association of air pollution with the nasal microbiota profile in adults (Xue et al., 2020). However, evidence on the association of air pollution and greenness with the early nasal microbiota is missing. Since this time period is crucial for later respiratory health (Teo et al., 2015) and exposure to air pollution in particular is omnipresent, a better understanding of the association between air pollution and greenness and the infant’s nasal microbiota is relevant to potentially improve respiratory health in the short- and long-term.

The aim of the present study was to examine the association of short-term exposure to particulate matter with an aerodynamic diameter \( \leq 2.5 \ \mu m (PM_{2.5}) \), nitrogen dioxide (NO\(_2\)) and residential greenness (normalized difference vegetation index (NDVII)) with the nasal microbiota of healthy infants during the first year of life. We focused on microbiota measures previously identified as clinically relevant (primary outcomes: bacterial diversity, bacterial density; secondary outcomes: relative abundance of bacterial families) (Acosta et al., 2018; Bisgaard et al., 2007; Hoggard et al., 2017; To et al., 2015; Valenzi et al., 2020). We took advantage of an available longitudinal dataset, including fortnightly obtained nasal swabs (Mika et al., 2015).

2. Materials and methods

2.1. Study design

A total of 778 healthy infants were enrolled in the prospective Basel-Bern Infant Lung Development (BILD) cohort study (Fuchs et al., 2012) between April 1999 and July 2013 (Fig. S1). Pregnant women were recruited at maternity hospitals and obstetric practices. Due to relevant differences in lung parameters, infants with ethnicity other than Caucasian, preterm delivery (<37 weeks’ gestation), major birth defects, disease or later diagnosis of airway malformation or specific chronic respiratory disease were excluded (Fuchs et al., 2012). Longitudinal microbiome data for the first year of life were collected and analysed in previous studies (Korten et al., 2016, 2019; Mika et al., 2015, 2017) for a subgroup of 48 infants of the BILD cohort born between February 2010 and July 2013. For the present study, this dataset was used and combined with the following exposure data calculated for the residential address of each infant and date of swab: \( PM_{2.5} \), \( NO_2 \) and NDVI. The study was approved by the Ethics Committee of the Canton of Bern and written consent was obtained at enrolment.

2.2. Microbiome data

Nasal swabs were collected fortnightly by the parents starting from the fifth week of life until completion of the first year of life. Study nurses instructed parents how to obtain the nasal swabs. Swabs were sent back to the study centre in transport tubes and medium (nasal swabs from Copan, Italy; UTMS® tubes from Verridial E. Muller, Switzerland) on average within 3 days and stored in a –80 °C freezer until further processing. Nasal swabs obtained during respiratory tract infection or antibiotic therapy were excluded (Fig. S1).

Methods for PCR amplification and amplicon sequencing are described in detail elsewhere (Hilty et al., 2012; Mika et al., 2015). Briefly, V3 to V5 regions of the bacterial 16S rRNA genes were amplified using the primer pair 341 F/907 R. After purification with the Wizard SV PCR cleanup system (Promega, Madison, WI), PCR mixtures were eluted in 40 µL of double-distilled water. For a reliable threshold, 1 pg/µL bacterial DNA corresponding to a PCR product concentration of 1 ng/µL is recommended (Biesbroek et al., 2012). PCR products with a concentration below that threshold were excluded from the study. Furthermore, samples with less than 70 sequence reads and samples displaying >5% sequence reads identical to two negative control samples were excluded (Fig. S1). Taxa that were recently identified as contaminants (Salter et al., 2014) were not detected.

From each PCR product, 40 ng/µL was pooled and every multiplex identifier was used once. The resulting 8 pools were sequenced with the 454-sequencing platform. Reads were submitted to the National Center for Biotechnology Information Sequence Read Archive. Analysis of the sequencing products was performed with PyroTagger, which comprises the definition of operational taxonomic units (OTUs) based on 97% sequencing identity, estimation of chimeras and taxonomic assignments (Kunin and Hugenholtz, 2010).

Concentration of the 16S rRNA PCR product (PCR) in ng/µL was used to estimate bacterial density. Alpha-diversity was assessed using the Shannon Diversity Index (SDI) and species richness (Xia et al., 2018). SDI was calculated in R (version 3.0.2) using the “vegan” package. The “Vegdist” function of the “vegan” package was used to assess the Manhattan-type Ruzicka dissimilarity (abundance-based) between 2 samples. Within-subject dissimilarity was calculated by comparison of two consecutive samples from one infant. As this dissimilarity depends on the length of the time interval (Δt) between the two consecutive samples (Mika et al., 2015), we excluded dissimilarity values derived from samples with <10 days or >21 days distance. In addition, we calculated Ruzicka dissimilarities between all samples of the study population. These data were used as input matrix for the whole microbial community comparison by nonmetric multidimensional scaling (nMDS).

Relative abundances of bacterial families were defined through taxonomic assignment of PyroTagger. Mean and 95% CI were calculated for all bacterial families in each sample. The five most abundant bacterial families (Pasteurellaceae, Corynebacteriaceae, Moraxellaceae, Staphylococcaceae, Streptococcaceae) were analysed separately and remaining families were grouped to “others”.

2.3. Air pollution exposure

Daily average concentrations of \( PM_{2.5} \) and \( NO_2 \) for April 2010 to December 2013 were estimated using a 4-stage spatiotemporal modeling framework, as described previously (de Hoogh et al., 2018, 2019). Since air pollution data for later years were not available, we excluded nasal swabs obtained after December 2013 (Fig. S1). Briefly, data from the ozone monitoring instrument (OMI), Copernicus Atmosphere Monitoring Service, land use and meteorological variables were included to estimate daily \( NO_2 \) concentrations during the sampling period. A multistage framework with mixed-effect and random forest models were used to downscale satellite measurements and incorporate local sources. These models explain 73% of the overall variation in \( NO_2 \).
exposure measured at monitoring sites in Switzerland (de Hoogh et al., 2019). Daily ambient PM$_{2.5}$ concentration at a spatial resolution of 100 × 100 m was estimated with geostatistical hybrid models taking advantage of the multivariate implementation of atmospheric correction (MAIAC) spectral aerosol optical depth (AOD) data combined with other spatiotemporal predictor variables (de Hoogh et al., 2018). These models explain 89% (total), 95% (spatial) and 88% (temporal) of the local variation in PM$_{2.5}$ concentrations.

Previous studies have shown that short-term air pollution exposure is associated with changes in the nasal microbiota occurring within one week (Mariani et al., 2018; Qin et al., 2019). Therefore, we focused on a 1-week exposure window and calculated mean PM$_{2.5}$ and NO$_2$ levels for the week prior to each nasal swab. To capture potential delayed effects, we additionally generated time lags by shifting back the exposure window 1–10 days (e.g., lag 1: exposure window starts 1 day prior to nasal swab, and so forth). In addition, we calculated yearly subject means for PM$_{2.5}$ and NO$_2$ to facilitate comparison with annual exposure limits suggested by the World Health Organization (World Health Organization (WHO) 2006).

### 2.4. Greenness exposure

Greenness surrounding the child’s home was determined using the Normalized Difference Vegetation Index (NDVI) obtained from the MODIS VI MOD13Q1 product (data generated every 16 days at 250 × 250 m resolution (Didan et al., 2015). The NDVI is a common metric used to define residential greenness in epidemiological studies (Lambert et al., 2017). This index approximates greenness based on the visible (red) and near-infrared light reflected by the land surface. NDVI is measured on a scale from −1 to 1 (NDVI ≤ 0.1: water, barren areas, sand or snow; 0.2–0.3: shrub, grassland; > 0.3 indicate increasing greenness) (Weier and Herring, 2000). The post-processed dataset was adjusted for cloud coverage. On the basis that greenness changes slowly, to further reduce a potential bias from cloud coverage, we averaged the last two NDVI estimates prior to each nasal swab.

### 2.5. Additional risk factors

Analyses were adjusted for known covariates based on previous findings (Mika et al. 2015, 2017). Information about presence of siblings (binary), childcare at date of swab (binary), any breastfeeding at date of swab (binary), parental smoking (binary) and delivery mode (vaginal delivery vs. caesarean section) was derived from hospital records, questionnaires and weekly phone interviews performed by trained study nurses. Pneumococcal conjugate vaccination (PVC) era was defined by questionnaires and weekly phone interviews performed by trained study nurses. Pneumococcal conjugate vaccination (PVC) era was defined by

### 2.6. Exposure assignment

Environmental exposures (PM$_{2.5}$, NO$_2$, NDVI) were linked to each child’s home address. Addresses were geocoded using the building registry of the Swiss Federal Statistical Office (Neuchâtel). Address changes within the study period were considered for the exposure assignment.

### 2.7. Statistical analysis

Time-series data including repeated nasal swabs and environmental exposures were used. The associations of PM$_{2.5}$, NO$_2$ and NDVI with bacterial diversity (SDI, species richness, within-subject Ruzička dissimilarity) and bacterial density (PCR concentration) were assessed using a generalized additive mixed model (GAMM) with Gaussian distribution and autoregressive structure (lag 1) to account for temporal correlation within subjects. Log-transformation was performed for species richness and PCR concentration to approach normal distribution. Coefficient estimates were back transformed and shown as percentage change ((exp(β)−1)×100). The associations of PM$_{2.5}$, NO$_2$ and NDVI with the relative abundance of bacterial families were assessed using a GAMM with autoregressive structure (lag 1) and negative binomial distribution to account for the zero-inflated outcome data. Simple models were adjusted for date of swab. The following covariates were used in the adjusted models: age at swab, sex, siblings, parental smoking, childcare, breastfeeding, delivery mode, vaccination era and date of swab. Age and date of swab were introduced into the model as smoothing terms in case a non-linear relationship with the outcome was identified. PM$_{2.5}$ and NO$_2$ were analysed in separate models to avoid multicollinearity. NDVI was studied in the same model as PM$_{2.5}$, respectively NO$_2$, as correlation between NDVI and the air pollutants was moderate (Pearson correlation ≤ −0.6). For the main analysis, we focused on the short-term effect of PM$_{2.5}$ and NO$_2$ exposure without time lag (lag 0: 1-week exposure window starting at date of swab). We additionally performed a sensitivity analysis for time-lagged PM$_{2.5}$ and NO$_2$ exposures to assess delayed effects of air pollution on SDI, species richness, bacterial density and relative abundance of bacterial families. Therefore, exposure windows were shifted backwards by 1–10 days prior to the date of swab.

NMDS was applied for whole microbial community comparison, using Ruzička dissimilarities as input matrix and considering repeated measures. Phylogenetic variation among all samples with respect to air pollution and greenness levels was assessed using “metaMDS” of the “vegan” package in R. For the mNDS the population was divided into terciles by individual levels of PM$_{2.5}$, NO$_2$ and NDVI (1st tercile: “low-exposed”, 2nd tercile: “mid-exposed”, 3rd tercile: “high-exposed”). The “adonis” function of the “vegan” package was used as PERMANOVA analysis for the clustering (p < 0.05 indicates significant clustering due to exposure group) and “envfit” was used to fit vectors which indicate the most rapid change (arrow direction).

Statistical analyses were performed with STATA (version 16.1, STATA Corp, Texas) and R (version 3.0.2. and 4.0.3).

### 3. Results

#### 3.1. Study population

The final data set comprised 846 nasal swabs from the first year of life of 47 healthy infants. Nasal swabs were obtained between April 2010 and December 2013. The mean (SD) of included swabs per infant was 18 (3.35). Details on the study population are outlined in Table 1.

#### 3.2. Environmental exposures

We found seasonal fluctuations for air pollution levels and greenness (Fig. 1). PM$_{2.5}$ and NO$_2$ concentrations decreased during warmer months (4-year mean April–September: PM$_{2.5}$ 14.4 μg/m$, NO_2 13.4 μg/m$) and increased during colder months (4-year mean October–March: PM$_{2.5}$ 19.2 μg/m$, NO$_2 25.8 μg/m$). In contrast, NDVI values increased during warmer months (4-year mean April–September: 0.6) and decreased during colder months (4-year mean October–March: 0.5).

For each infant we calculated the yearly subject mean of the individual air pollution and greenness levels. Mean (SD) PM$_{2.5}$ level was 17.1 (3.8 μg/m$)^2$ and mean (SD) NO$_2$ level was 19.7 (7.9 μg/m$)^2$. Across all subjects, the mean (SD) greenness was 0.5 (0.1) during the first year of life.

#### 3.3. Nasal microbiota composition

Within the 846 nasal swabs, SDI ranged from 0.0 to 3.5 with a mean (SD) of 1.2 (0.8), species richness ranged from 2.0 to 120.0 with a mean (SD) of 31.9 (20.6) and PCR concentration ranged from 1.0 to 154.6 ng/
3.4. Association of air pollution and greenness with the nasal microbiota

We found associations of air pollution with Ruzicka dissimilarity and the relative abundance of Corynebacteriaceae in healthy infants during the first year of life. Results from simple and adjusted models are shown in Tables 2 and 3.

In the adjusted model an increase in PM$_{2.5}$ and NO$_2$ was associated with reduced within-subject Ruzicka dissimilarity (PM$_{2.5}$: per 1 μg/m$^3$ -0.004, 95% CI -0.008, −0.001; NO$_2$: per 1 μg/m$^3$ -0.004, 95% CI -0.007, −0.001) (Table 2, Fig. 2). In the whole community comparison among all samples using nMDS, we observed significant differences by PM$_{2.5}$ exposure for the “mid-exposed” group (PM$_{2.5}$ range: 12.8–18.5 μg/m$^3$; “adonis” function: p = 0.001) and “high-exposed” group (PM$_{2.5}$ range: 18.6–45.8 μg/m$^3$; “adonis” function: p = 0.032). No significant differences could be identified for the “low-exposed”, “mid-exposed” and “high-exposed” NO$_2$ exposure groups (Fig. S3).

Our adjusted models allowed us to estimate the association of Corynebacteriaceae, Moraxellaceae and Streptococcaceae with air pollution and greenness. However, the models did not converge for Pasteurellaceae and Staphylococcaceae, indicating inadequate fit for these data. We found that time-lagged NO$_2$ exposure was associated with decreased abundance of Corynebacteriaceae. Associations were significant for lags of 3–10 days (NO$_2$ lag 3: per 1 μg/m$^3$ -0.027, 95% CI -0.053, −0.001) (Table 3, Fig. 2). No association could be identified for PM$_{2.5}$ and the abundance of Corynebacteriaceae (Table 3, Fig. 2). Results for lag 3 are shown in Table 3 and Fig. 2. Results for all time lags (0-10 days) are shown in Figure S6 and S7.

No associations could be found for PM$_{2.5}$ and NO$_2$ with SDI, bacterial density (PCR concentration), species richness and the relative abundance of Moraxellaceae and Streptococcaceae (Tables 2 and 3, Fig. 2). Also, the sensitivity analysis for time-lagged PM$_{2.5}$ and NO$_2$ exposure did not reveal any associations with these measures (Figs. S4–S7). Furthermore, no associations between greenness and any of the examined microbiota measures were identified in models mutually adjusted for air pollution (Tables 2 and 3, Fig. 2, Fig. S3).

4. Discussion

In this study we examined the association of air pollution (PM$_{2.5}$ and NO$_2$) and greenness (NDVI) with the nasal microbiota of healthy infants using longitudinal data. This data structure enabled us to study the associations of air pollution and greenness with the nasal microbiota during the first year of life in detail.

We found that increasing PM$_{2.5}$ and NO$_2$ levels were associated with a decreased within-subject Ruzicka dissimilarity comparing two consecutive swabs of one subject, indicating increased bacterial stability. There is evidence that decreased within-subject dissimilarity is associated with a reduced susceptibility to respiratory tract infections, but there is no evidence on the association between air pollution and within-subject Ruzicka dissimilarity (Biesbroek et al., 2014; Bosch et al., 2017). As air pollution has been positively associated with respiratory symptoms and respiratory tract infections in infants (Horne et al., 2018;
### Table 2
Association of air pollution (lag 0) and NDVI with α- and β-diversity of the nasal microbiota.

| SDI | PCR Concentration | Species Richness | Within-Subject Bručka Dissimilarity |
|-----|-------------------|------------------|-----------------------------------|
|     | Coef               | 95% CI           | p-Value                           | Coef               | 95% CI           | p-Value |
|     |                   |                  |                                   |                   |                  |        |
| Simple model |                   |                  |                                   |                   |                  |        |
| PM$_{2.5}$ μg/m$^3$ | -0.005            | -0.014, 0.003    | 0.229                             | 0.148             | -1.102, 1.415    | 0.817   | -0.099             | -0.841, 0.649    | 0.794   | -0.004             | -0.007, 0.000    | 0.043   |
| NO$_2$ μg/m$^3$ | -0.001            | -0.008, 0.006    | 0.788                             | -0.014            | -1.174, 0.888    | 0.778   | -0.320             | -0.883, 0.246    | 0.267   | -0.003             | -0.005, 0.000    | 0.051   |
| NDVI | 0.009              | -0.051, 0.069    | 0.771                             | -2.83          | -11.546, 6.743   | 0.549   | 0.149              | -4.727, 5.274    | 0.953   | 0.007              | -0.016, 0.020    | 0.545   |
| Adjusted Model |                   |                  |                                   |                   |                  |        |
| PM$_{2.5}$ μg/m$^3$ | -0.007            | -0.015, 0.002    | 0.143                             | 0.034            | -1.243, 1.327    | 0.959   | -0.224             | -0.998, 0.555    | 0.571   | -0.004             | -0.008, -0.001   | 0.018   |
| NO$_2$ μg/m$^3$ | -0.003            | -0.011, 0.005    | 0.443                             | -0.049           | -1.597, 0.712    | 0.446   | -0.472             | -1.141, 0.202    | 0.170   | -0.004             | -0.007, -0.001   | 0.017   |
| NDVI | 0.005              | -0.064, 0.051    | 0.828                             | -4.621           | -12.570, 4.050   | 0.286   | -0.517             | -5.323, 4.532    | 0.837   | -0.002             | -0.026, 0.022    | 0.847   |
| PM$_{2.5}$ adjusted | -0.007           | -0.073, 0.060    | 0.844                             | -0.679           | -15.683, 3.024   | 0.168   | -2.361             | -7.766, 3.361    | 0.411   | -0.012             | -0.040, 0.015    | 0.389   |
| NO$_2$ adjusted | -0.007           | -0.103, 0.103    | 0.326                             | -0.103           | -0.310, 0.103    | 0.797   | 0.074              | -0.027, 0.176    | 0.151   |

Simple models were adjusted for date of swab (seasonal component).

Adjusted models were adjusted for date of swab, siblings, gender, childcare at swab, age at swab, feeding type at swab, delivery mode, vaccination era and parental smoking.

a Average values of day 0–7 prior to swab (lag 0: exposure window of 1 week starts at day of swab, coefficients for all lags are shown in Figures S4 and S5).

b Average values of last two NDVI estimates prior to swab.

β $\Delta$ represents the percent increase per 1 μg/m$^3$ increase in PM$_{2.5}$ and NO$_2$, and per 0.1 increase in NDVI: PCR concentration and species richness: (exp($β$)-1)*100.

### Table 3
Association of air pollution (lag 3) and NDVI with relative abundance of bacterial families.

| Family              | Corynebacteriaceae | Moraxellaceae | Streptococcaceae |
|---------------------|--------------------|---------------|------------------|
|                     | Coef               | 95% CI        | p-Value          | Coef               | 95% CI        | p-Value          | Coef               | 95% CI        | p-Value          |
| Simple Model |                   |                |                  |                   |                |                  |                   |                |                  |
| PM$_{2.5}$ μg/m$^3$ | -0.001            | -0.033, 0.031 | 0.965            | -0.003            | -0.014, 0.007  | 0.561            | -0.001            | -0.013, 0.012  | 0.918            |
| NO$_2$ μg/m$^3$ | -0.003            | -0.027, 0.021 | 0.800            | -0.003            | -0.013, 0.006  | 0.497            | -0.003            | -0.013, 0.006  | 0.497            |
| NDVI | -0.039            | -0.239, 0.161 | 0.699            | -0.003            | -0.092, 0.086  | 0.947            | 0.042             | -0.038, 0.121  | 0.302            |
| Adjusted Model |                   |                |                  |                   |                |                  |                   |                |                  |
| PM$_{2.5}$ μg/m$^3$ | -0.005            | -0.034, 0.024 | 0.749            | -0.003            | -0.014, 0.009  | 0.673            | 0.003             | -0.010, 0.016  | 0.621            |
| NO$_2$ μg/m$^3$ | -0.027            | -0.053, -0.001 | 0.041           | -0.002            | -0.015, 0.011  | 0.762            | 0.003             | -0.009, 0.015  | 0.624            |
| NDVI | -0.003            | -0.015, 0.017 | 0.839            | -0.017            | -0.120, 0.087  | 0.754            | 0.066             | -0.020, 0.153  | 0.134            |
| PM$_{2.5}$ adjusted | 0.018              | -0.156, 0.192 | 0.839            | -0.017            | -0.134, 0.103  | 0.797            | 0.074             | -0.027, 0.176  | 0.151            |
| NO$_2$ adjusted | -0.103            | -0.310, 0.103 | 0.326            | -0.015            | -0.134, 0.103  | 0.797            | 0.074             | -0.027, 0.176  | 0.151            |

Simple models were adjusted for date of swab (seasonal component).

Adjusted models were adjusted for date of swab, siblings, gender, childcare at swab, age at swab, feeding type at swab, delivery mode, vaccination era and parental smoking.

a Average values of day 0–7 prior to swab (lag 3: exposure window of 1 week starts 3 days prior to swab, coefficients for all lags are shown in Figures S6 and S7).

b Average values of last two NDVI estimates prior to swab.

c Coef represent the difference in the logs of relative abundance per 1 μg/m$^3$ increase in PM$_{2.5}$ and NO$_2$, and per 0.1 increase in NDVI, while holding the other variables in the model constant.
We expect lower bacterial stability for increased exposure levels. The negative association between air pollution and within-subject Růžička dissimilarity identified in the present study is therefore counterintuitive. We speculate that the exposure to a constantly dominating external factor, such as higher levels of air pollution, may trigger a typical bacterial profile. The clinical relevance of reduced within-subject dissimilarity for respiratory health remains unclear and has to be investigated in longitudinal studies assessing the relationship between environmental factors and the nasal microbiome in parallel to the association with respiratory disease development.

In the whole microbial community comparison among all samples we observed significant differences in the phylogenetic variation by PM$_{2.5}$ exposure for groups exposed to medium and high levels of PM$_{2.5}$, suggesting distinct microbiota profiles. To our knowledge, this is the first report of such differences in the nasal microbiome.

**Fig. 2.** Associations of air pollution and greenness with the nasal microbiota during the first year of life. a) Effect of PM$_{2.5}$, NO$_2$ and NDVI on within-subject Růžička dissimilarity shown as $\beta$ coefficients with 95% CI derived from the adjusted models (Table 2) (per 1 $\mu$g/m$^3$ change in PM$_{2.5}$ and NO$_2$, per 0.1 change in NDVI). b) Assessment of phylogenetic variation among all samples using nMDS showed significant difference by PM$_{2.5}$ exposure for the “mid-exposed” group (12.8–18.5 $\mu$g/m$^3$, “adonis” function: $p = 0.001$) and “high-exposed” group (18.6–45.8 $\mu$g/m$^3$, “adonis” function: $p = 0.032$). Each dot represents a sample. Vectors indicate the most rapid change (arrow direction). c) Associations of 3-days-lagged air pollution exposure and NDVI with the abundance of bacterial families (Str: Streptoccocaceae, Mor: Moraxellaceae, Cor: Corynebacteriaceae) shown as $\beta$ coefficients with 95% CI derived from the adjusted models (Table 3).
first study investigating the association of PM$_{2.5}$ and NO$_2$ with Ružička dissimilarity of the nasal microbiota in infants. However, our results are supported by other studies which found that the exposure to PM$_{2.5}$ was related with a distinct composition of the pharyngeal microbiome in adults (Qin et al., 2019) and a distinct composition of the lung microbiome in mice (Li et al., 2020).

Furthermore, we found that a 3-to-10-day time-lagged exposure to NO$_2$ was negatively associated with the relative abundance of Corynebacteriaceae. Our finding is supported by Mariani et al., (2018), who identified a negative association between time-lagged PM$_{2.5}$ exposure and the abundance of Corynebacteriaceae (% change per 1 µg/m$^3$: −2.52). However, comparability is limited, as they assessed single-day exposures to PM$_{2.5}$, while we studied the effect of averaged NO$_2$ (mean of 8 days). Higher abundance of Corynebacteriaceae has been previously related to respiratory health in infants and children (Biesbroek et al., 2014; McCauley et al., 2019; Neumann et al., 2018; Rawls and Ellis, 2019; Teo et al., 2015), which suggests that the association between increasing air pollution exposure and reduced abundance of this commensal bacterial family might be of clinical relevance.

The association of PM$_{2.5}$ and NO$_2$ with Ružička dissimilarity and bacterial composition identified herein supports our hypothesis that there may be an interplay between the nasal microbiota and air pollution with potential consequences for short- and long-term respiratory health. To put the air pollution levels in our study in context, and compare with the WHO annual guideline limits, the mean annual exposure for all infants was 17 µg/m$^3$ PM$_{2.5}$ (guideline 10 µg/m$^3$) and 20 µg/m$^3$ NO$_2$ (guideline 40 µg/m$^3$). Our results indicate that low-to-moderate air pollution levels may still impact the nasal microbiota profile and thus affect respiratory health (World Health Organization (WHO) 2006).

Our study has considerable strengths. Most relevant is the high-quality exposure data that has been estimated for each nasal swab included in the analysis. We have used novel modeling techniques which allowed us to capture fine-scale spatial and temporal variation of PM$_{2.5}$ and NO$_2$. The daily air pollution data has enabled us to study time-lagged air pollution exposures. Further, we used fortnightly data as a basis for the NDVI exposure. Another major strength of our study is the dense longitudinal sampling of nasal swabs. Samples were collected every two weeks in 47 infants, resulting in 846 samples available for the analysis. Previous studies investigating the association of air pollutant with the nasal microbiota, included one (Mariani et al., 2018) to two (Qin et al., 2019) swabs per subject, which limits the conclusion on effects. Our longitudinal dataset with several repeated measures per subject increases the power to detect potential associations between the dynamic nasal microbiota and highly variable environmental exposures.

Certain limitations should be addressed. First, variation in our exposure data was small. More variation in exposure data and a larger sample size might be necessary to detect additional associations of air pollution and greenness with the nasal microbiota that are rather weak compared to the associations with other previously described factors such as age or presence of siblings (Bosch et al. 2016, 2017; Hasegawa et al., 2017; Mika et al., 2015). In addition, more distinct microbiota profiles could be obtained using novel metagenome sequencing techniques which would enable more insights on the association between environmental exposures and the microbiota. Furthermore, associations with air pollution and greenness could be examined only for Corynebacteriaceae, Moraxellaceae and Streptococcaceae. The GAMM did not control for Pseudomonadaceae and Staphylococcaceae, most likely due to zero inflation and low abundance of these bacterial families (Wood 2021). This indicates that the GAMM model, even when accounting for negative binomial distribution, does not provide the ideal fit for zero-inflated microbiome data. It is well-known that the analysis of this type of data is challenging, and is even more complex for longitudinal data (Zhang et al., 2020). Nonetheless, the GAMM is a good fit for investigating our primary outcomes (bacterial diversity, bacterial density) under consideration of the longitudinal data structure and was therefore viewed as a reasonable compromise.

5. Conclusion

This study shows that exposure to PM$_{2.5}$ and NO$_2$ is related to the early nasal microbiota. The association of air pollution with Ružička dissimilarity and the abundance of Corynebacteriaceae suggests that even low-to-moderate exposure levels may impact the nasal microbiota during the first year of life, which has been associated with respiratory infections and asthma development. Our results can be useful for future studies assessing the clinical relevance of air-pollution-induced alterations of the nasal microbiota and subsequent disease development. However, further longitudinal studies with more distinct exposure levels are needed to extend upon our findings.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: PL reports personal fees from OM Pharma, Polyphor, Santhera, Vertex and Vifor, outside the submitted work. JU reports personal fees from Vertex outside the submitted work. MH reports personal fees and grants from Pfizer, outside the submitted work. The remaining authors declare no conflicts of interest.

Acknowledgments

The authors thank the study participants; the BILD study group and the BILD cohort team (University of Basel Children’s Hospital, Switzerland; Division of Respiratory Medicine, Department of Pediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland); M. Graf, S. Lüscher and L. Beul for data collection; M. Mika for PCR amplification of 16S rRNA genes and 454 amplicon sequencing of the samples and Fiona Beck for critical reading of the manuscript.

BILD study group: Kees de Hoogh, Basel; Fabienne Decrue, Basel; Urs Frey, Basel; Oliver Fuchs, Bern; Amanda Gisler, Basel; Olga Gorlanova, Basel; Insa Korten, Bern; Johanna Kurz, Bern; Claudia Kühni, Bern; Philipp Latzin, Bern; Loretta Müller, Bern; Uri Nahum, Basel; Marc-Alexander Oestreich, Bern; Martin Röösli, Basel; Yasmin Salem, Bern; András Soti, Bern; Jakob Usemann, Basel; Danielle Vienneau, Basel; Corin Willers, Bern; Sophie Yammine, Bern.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envres.2021.111633.

Funding

The samples used in this study are part of the Basel-Bern Infant Lung Development (BILD) cohort study, which is funded by the Swiss National Science Foundation (grant 320030_163311). AG is funded by the “Forschungsfond exzellenter Nachwuchsforshender” issued to JU.

Author contributions

PL, UF, IK and JU designed the study. IK, JU, FD, OG, AS assisted in collection of the clinical and metadata. MH was responsible for sample amplicon sequencing and bioinformatics analyses. AG performed the data analysis and wrote the main manuscript with input from the co-authors. JU, PL, IK, OG and UP contributed to the statistical interpretation of results. DV and KdH provided the environmental exposure data and helped in interpretation of the data. UF is the principal investigator of the BILD cohort. All authors read and approved the final manuscript.
Air quality guidelines global update 2005: particulate matter, ozone, nitrogen dioxide and sulfur dioxide Available from: https://apps.who.int/iris/handle/10665/107823 (Accessed).

Xia, Y., Sun, J., Chen, D.-G., 2018. Community Diversity Measures and Calculations. Statistical Analysis of Microbiome Data with R. Springer Singapore, Singapore.

Xue, Y., Chu, J., Li, Y., Kong, X., 2020. The influence of air pollution on respiratory microbiome: a link to respiratory disease. Toxicol. Lett. 334, 14-20. https://doi.org/10.1016/j.toxlet.2020.09.007.

Zar, H.J., Ferkol, T.W., 2014. The global burden of respiratory disease-impact on child health. Pediatr. Pulmonol. 49 (5), 430-434. https://doi.org/10.1002/ppul.23030.

Zhang, B., Li, Y., Chen, Q., Jiang, Y., Chu, C., Ding, Y., et al., 2019. The relationship between air quality and respiratory pathogens among children in Suzhou City. Ital. J. Pediatr. 45 (1), 123. https://doi.org/10.1186/s11861-019-0702-2.

Zhang, X., Guo, B., Yi, N., 2020. Zero-Inflated Gaussian mixed models for analyzing longitudinal microbiome data. PLoS one 15 (11), e0242073. https://doi.org/10.1371/journal.pone.0242073.