Urbanized microbiota in infants, immune constitution and later risk of atopic diseases

Lehtimäki, Jenni; Thorsen, Jonathan; Rasmussen, Morten Arendt; Hjelmso, Mathis ; Shah, Shiraz; Mortensen, Martin S.; Trivedi, Urvish; Vestergaard, Gisle Alberg; Bønelykke, Klaus; Chawes, Bo Lund

Total number of authors: 14

Published in:
Journal of Allergy and Clinical Immunology

Link to article, DOI:

Publication date:
2021

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
Urbanized microbiota in infants, immune constitution and later risk of atopic diseases

Jenni Lehtimäki, PhD, Jonathan Thorsen, MD, PhD, Morten Arendt Rasmussen, PhD, Mathis Hjelmsø, PhD, Shiraz Shah, PhD, Martin S. Mortensen, PhD, Urvish Trivedi, PhD, Gisle Vestergaard, PhD, Klaus Bønnelykke, MD, PhD, Bo Lund Chawes, MD, PhD, Susanne Brix, PhD, Søren J. Sørensen, PhD, Hans Bisgaard, MD, PhD, Jakob Stokholm, MD, PhD

PII: S0091-6749(20)32407-6
Reference: YMAI 14899
To appear in: Journal of Allergy and Clinical Immunology

Received Date: 17 September 2020
Revised Date: 11 November 2020
Accepted Date: 8 December 2020


This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier Inc. on behalf of the American Academy of Allergy, Asthma & Immunology.
URBANIZED MICROBIOTA IN INFANTS, IMMUNE CONSTITUTION AND LATER RISK OF ATOPIC DISEASES

Jenni Lehtimäki\textsuperscript{1,} PhD, Jonathan Thorsen\textsuperscript{1,2,} MD, PhD, Morten Arendt Rasmussen\textsuperscript{1,3,} PhD, Mathis Hjelmsø\textsuperscript{1,} PhD, Shiraz Shah\textsuperscript{1,} PhD, Martin S. Mortensen\textsuperscript{4,5,} PhD, Urvish Trivedi\textsuperscript{4,} PhD, Gisle Vestergaard\textsuperscript{4,6,} PhD, Klaus Bønnelykke\textsuperscript{1,} MD, PhD, Bo Lund Chawes\textsuperscript{1,} MD, PhD, Susanne Brix\textsuperscript{7,} PhD, Søren J. Sørensen\textsuperscript{4,} PhD, Hans Bisgaard*\textsuperscript{1,} MD, PhD, Jakob Stokholm\textsuperscript{1,} MD, PhD

\textsuperscript{1} COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Ledreborg Alle 34, 2820, Gentofte, Denmark

\textsuperscript{2} Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3b, 2100, Copenhagen, Denmark

\textsuperscript{3} Section of Chemometrics and Analytical Technologies, Department of Food Science, University of Copenhagen, Rolighedsvej 30, 1958, Frederiksberg C, Denmark

\textsuperscript{4} Section of Microbiology, Department of Biology, University of Copenhagen, Universitetsparken 15, 2100, Copenhagen, Denmark

\textsuperscript{5} Host-Microbiota Interactions Laboratory, Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, UK

\textsuperscript{6} Section for Bioinformatics, Department of Health Technology, Technical University of Denmark, 2800 Lyngby, Denmark

\textsuperscript{7} Department of Biotechnology and Biomedicine, Technical University of Denmark, Lyngby, Denmark.

*Correspondence: Professor Hans Bisgaard, MD, DMSc, address: Ledreborg Alle 34, 2820, Gentofte, Denmark, phone: +45 3867 7360, E-mail: bisgaard@copsac.com
Conflict of interest: All authors declare no potential, perceived, or real conflict of interest regarding the content of this manuscript. The funding agencies did not have any role in design and conduct of the study; collection, management, and interpretation of the data; or preparation, review, or approval of the manuscript. No pharmaceutical company was involved in the study.

Source of Funding: All funding received by COPSAC is listed on www.copsac.com. The Lundbeck Foundation (Grant no R16-A1694); The Ministry of Health (Grant no 903516); Danish Council for Strategic Research (Grant no 0603-00280B) and The Capital Region Research Foundation have provided core support to the COPSAC research center. Dr. Jonathan Thorsen is supported by the BRIDGE – Translational Excellence Programme (bridge.ku.dk) at the Faculty of Health and Medical Sciences, University of Copenhagen, funded by the Novo Nordisk Foundation. Grant agreement no. NNF18SA0034956.
ABSTRACT

Background Urbanization is linked with an increased burden of asthma and atopic traits. A putative mechanism is insufficient exposure to beneficial microbes early in life leading to immune dysregulation as previously shown for indoor microbial exposures.

Objective To investigate whether urbanization is associated with the microbiota composition in the infants’ body and early immune function, and whether these contribute to the later risk of asthma and atopic traits.

Methods We studied the prospective COPSAC2010 mother-child cohort of 700 children growing up in areas with different degrees of urbanization. During their first year of life, airway and gut microbiota as well as immune marker concentrations were defined. At six years of age, asthma and atopic traits were diagnosed by pediatricians.

Results In adjusted analyses, the risk of asthma and aeroallergen sensitization were increased in urban infants. The composition of especially airway, but also gut microbiota differed between urban and rural infants. The living environment related structure of the airway microbiota associated with immune mediator concentrations already at one month of age. An urbanized structure of airway and gut microbiota associated with an increased risk of asthma coherently during multiple time points, and also with the risks of eczema and sensitization.

Conclusion Our findings suggest that urbanization related changes in the infant microbiota may elevate the risk of asthma and atopic traits, probably via crosstalk with the developing immune system. The airways may facilitate this effect as they are open for colonization by environmental, airborne microbes and serve as immune interface.
Key messages

• Airway and gut microbiota as well as immune marker concentrations differ between rural and urban infants.

• Urbanized composition of microbiota in infancy increases the risk of asthma, eczema and allergic sensitization at six years of age.

• Urban environments may predispose to the development of a disease-promoting composition of the infant microbiota.

Capsule summary In a prospective cohort study, an urban living environment and urbanized infant microbiota associated with increased risks of asthma, eczema and aeroallergen sensitization at six years of age probably due to early immune dysregulation.

Keywords microbiome; residential environment; non-communicable diseases; urbanization; childhood

Abbreviations

aOR = adjusted odds ratio
ANOVA = Analysis of variance
CCL = (C-C motif) chemokine ligand
CRP = C-reactive protein
COPSAC = Copenhagen Prospective Studies on Asthma in Childhood
CXCL = (C-X-C motif) chemokine ligand
IL = Interleukin
INF = Interferon
TGF = Transforming growth factor
90   TNF = Tumor necrosis factor
91   OR = Odds ratio
92   PAM = Partitioning Around Medoids
93   PCA = Principal Component Analysis
94   PCoA = Principal Coordinates Analysis
95   PERMANOVA = Permutational Multivariate Analysis of variance
96   RMSE = Root Mean Squared Error
97   sIgE = specific Immunoglobulin E
98   SPT = skin prick test
99   sPLS = Sparse Partial Least Squares
INTRODUCTION

Urban children tend to have a higher risk of developing asthma and atopic traits such as allergic sensitization, rhinitis and eczema than rural children. Several factors differ between these populations including larger family size and increased exposure to animals in rural areas, which may contribute to the difference in disease risk. Also, the microbial exposures are different between rural and urban areas, which has been shown for the composition of outdoor, indoor and human microbiota. These microbial differences may mediate the protective effect of a rural environment, as demonstrated for farming-related indoor microbiota, which has been repeatedly associated with asthma protection, also in experimental settings.

Microbial exposures are especially important during infancy and early childhood, when the constitution of the immune system is primed by extensive immune-microbiota crosstalk. Indeed, accumulating evidence shows that infant microbiota is altered prior to the development of allergic sensitization, rhinitis, eczema and/or asthma, potentially influencing this early crosstalk. A recent report suggested that growing up in farms can support the maturation of infants’ gut microbiota and thus reduce their risk of asthma when compared to other rural children. However, it remains unclear how constantly increasing urbanization of our living environments influence on early immune-microbiota crosstalk and later risk of disease.

To address the hypothesis that dissimilar microbiota between urban and rural infants contribute to the development of their immune system and the later risk of disease, we utilized longitudinal data from the Copenhagen Prospective Studies on Asthma in Childhood 2010 (COPSAC) mother-child cohort. We defined the degree of urbanization in their living environment during infancy and studied its associations with asthma and atopic traits at age six years. We further investigated whether the early microbiota composition of airways and...
gut as well as the local airway and systemic immune function may convey associations between living environment and disease.

METHODS

Cohort

Children in the COPSAC2010 mother-child cohort (n=700) have been followed prospectively since birth with numerous scheduled visits at a dedicated research unit. Comprehensive descriptions of the cohort and the research approach are published\textsuperscript{26–28}. The COPSAC pediatricians were solely responsible for diagnosis and treatment of all respiratory, allergy and skin-related symptoms.

Definition of asthma and atopic traits

\textbf{Asthma} by age 6 years was diagnosed prospectively based on a previously detailed quantitative symptom algorithm\textsuperscript{26,29} (Online repository methods).

\textbf{Eczema} by age 6 years was diagnosed prospectively based on the criteria of Hanifin and Rajka\textsuperscript{30} requiring the presence of 3 of 4 major criteria and at least 3 of 23 minor criteria.

\textbf{Allergic rhinitis} at age 6 years was based on sensitization and clinical interviews of the parents on history of significant nasal congestion, sneezing and/or runny nose outside periods with common cold and in the relevant period of the sensitized aeroallergen.

\textbf{Allergic sensitization} was determined at age 6 years, defined as any positive skin prick test (SPT) $\geq 3$ mm (ALK-Abello, Horsholm, Denmark) or specific Immunoglobulin E (sIgE) in serum $\geq 0.35$ kUa/L against common inhalant and/or food allergens (ImmunoCAP; Thermo Fisher Scientific, Allerod, Denmark). "Not sensitized" were both SPT and sIgE negative (Online repository methods).
**Definition of urbanization gradient**

We utilized the land cover database CORINE (https://land.copernicus.eu/) from the year 2012 for an unbiased definition of the living environment. Addresses at birth (n=686) were translated to coordinates using Google API services. Land cover types around each coordinate point were extracted from the CORINE2012 raster map (100 m resolution) with a three-kilometer buffer (Table E1). The correlation matrix of proportions of five major land cover types within the buffer region were analyzed with Principal Component Analysis (PCA). The first axis explained 44% of the total variation and made clear separation between artificial and natural (i.e. agricultural and forested) land cover types (Fig E1a). The first axis was extracted for downstream analyses to represent a gradient of urbanization running from the most rural to the most urban environments in the study region. Land cover information was also dichotomized into rural and urban groups (n_rural=314, n_urban=372; Fig E1b) with Partitioning Around Medoids (PAM) clustering on Euclidean distances. Most children (n=577) did not move during their first year of life or moved later than 50 weeks after birth (n=5). From the 109 children who did move, 77% did not change between rural and urban groups (Online repository methods).

**Covariates**

Differences between rural and urban living environments were tested for common disease-associated factors including pet ownership, daycare attendance during the first year of life, length of breastfeeding period, passive smoking exposure, income of family, parental education, number of older siblings, home type, mode of delivery, parental diagnosis of asthma, eczema and rhinitis, and use of antibiotics during the first year of life (Fig E2).
**Microbiota samples**

Airway microbiota samples were collected at one week (n=549), one (n=647) and three months (n=657) of age\(^{34}\), gut microbiota samples were collected at one week (n=544), one month (n=615) and one year (n=625) of age\(^{22}\) (Online repository methods).

DNA was extracted from airway and gut samples using the PowerMag® Soil DNA Isolation Kit optimized for epMotion® (MO-BIO Laboratories, Inc., Carlsberg, CA, USA) using the epMotion® robotic platform model 5075 (Eppendorf, Hamburg, Germany). The manufacturer's protocol was mildly modified (described earlier\(^{34}\)). Extracted DNA was stored at -20 °C degrees prior to amplification of the 16S rRNA gene variable region 4 (V4) and sequencing with the Illumina MiSeq System (Illumina Inc., CA, USA) as previously described\(^{34}\).

The bioinformatics analysis pipeline is detailed in the Online repository methods.

**Immune mediators**

Airway immune mediator samples (upper airway mucosal lining fluid) were collected from one month old infants (n=620) with a synthetic absorptive matrix placed in both nostrils for 2 minutes\(^{35}\). Matrixes were stored immediately at -80 °C degrees. Blood for systemic immune mediator analysis was drawn in an EDTA tube from a cubital vein at the age of 6 months (n=568), centrifuged to separate plasma and cells, and plasma was immediately stored at -80°C until analysis\(^{36}\).

The levels of the *a priori* selected cyto- and chemokines\(^{37,38}\) were determined by using high-sensitivity immunoassays based on electrochemiluminescence in a multiplex setting. Samples were read in duplicates by using the Sector Imager 6000 (Meso Scale Discovery, Gaithersburg, MD).
Statistics

Odds ratios for atopic traits in response to urban and rural (baseline) classes as well as urbanization gradient were calculated with logistic regression models adjusted for lifestyle variables. Adjustment included other covariates associated with urbanization (Fig E2), while house type was excluded due to collinearity with the explanatory variable of interest (urban/rural groups).

The R-package phyloseq\(^{39}\) was used for handling the microbiota data. Observed richness and Shannon diversity were used as the alpha diversity metrics. The community compositions of microbiotas (beta diversity) were visualized with Principal Coordinates Analysis (PCoA) for Bray-Curtis\(^{40}\), Weighted UniFrac and Unweighted UniFrac\(^{41}\) distances and tested with Permutational Multivariate Analysis of variance (PERMANOVA)\(^{32}\) with 1000 permutations. The dissimilarities within rural and urban individuals were defined with the distance of each sample from group mean\(^{42}\).

Immune mediators were log transformed and centered log ratio scaled in order to remove biological variation due to differing amount of protein per volume of fluid from the dataset as well as to increase the comparability between immune mediators following methodology developed previously\(^{38}\). The normalized concentrations were tested between rural and urban groups in adjusted linear models.

Sparse Partial Least Squares (sPLS) models with relative abundances of taxa agglomerated to genus level, log-transformed, and cleaned to include those present at least in 5% of samples were used to predict the urbanization gradient\(^{38,43}\). The best model was selected based on the Root Mean Squared Error (RMSE) statistics, from repeated 10-fold cross-validation, in order to avoid overfitting. The predicted values from each model were extracted to represent a bacterial score for the urbanization gradient. All analyses were conducted in R version 3.5.2\(^{44}\).
Ethics statement

This study was approved by the local Ethics Committee (H-B-2008-093), and the Danish Data Protection Agency (2015-41-3696). Both parents gave verbal and written informed consent before enrolment.
RESULTS

Asthma and atopic traits in rural and urban infants

All subjects spent their first year of life in a westernized, densely populated, and geographically concise area in Zealand (7,031 km²), Denmark (Fig 1a). Few children lived on the farms. Several lifestyle variables such as pet ownership, number of older siblings, income in the family and education of the parents differed between rural and urban infants (Fig E2). Other screened variables, including mode of delivery, antibiotics during the first year of life, vitamin D or fish oil intervention groups, and atopic diseases in the parents, all previously associated with atopic diseases, did not differ between rural and urban infants.

The prevalence of asthma at six years of age or earlier was 22.3% (n=146), allergic rhinitis 6.8% (n=45), and eczema 31.7% (n=210). At six years of age, 31.7% (n=157) of children were sensitized, of which 23.7% (n=117) had aeroallergen and 14.7% (n=73) food sensitization. The prevalence of asthma, allergic rhinitis and aeroallergen sensitization at six years of age were higher in children spending their first year of life in an urban compared to a rural environment (Figs 1b and 1d). Allergic sensitization and eczema showed similar trends. Urban infants developed more specific sensitization to peanut, house dust mite, timothy, birch, and mugworth allergens (Table E2). After adjustment for lifestyle features, the odds ratios for asthma (aOR 2.31, 95% CI 1.47-3.68, P=0.0003) and aeroallergen sensitization (aOR 1.77, 95% CI 1.05-3.02, P=0.0348) were increased in urban infants (Figs 1c and 1e). The adjusted, predicted prevalence of asthma and aeroallergen sensitization showed a large increase along the urbanization gradient while only food sensitization remained completely unrelated (Fig E3).
Microbiota and immune mediator concentrations in rural and urban infants

The airway microbiota at all three time points showed significant differences between rural and urban infants regardless of the metric used to define the similarity of microbiota between subjects (Fig E4, Table E3). These differences remained significant after adjustment for each lifestyle variable as well as in a fully adjusted model (PERMANOVA, Table E4). A part of the observed rural-urban differences was explained by the larger family size and more common pet ownership in rural environments. Shannon diversity and richness in the airway microbiota were higher in urban infants at one month of age (ANOVA, P=0.004 and 0.005 respectively). Remarkably, urban infants had more homogenous airway microbiota than rural infants at all time points (i.e. urban children had more similar microbiota to each other than rural children; Fig E5). Among the fifteen most abundant genera in the airways, several showed dissimilar relative abundances between rural and urban infants (Fig E6). Veillonella showed a consistent association with an urban living environment through all time points. Haemophilus and Rothia associated with urban environment while Dolosigranulum and Moraxella associated with rural environment in more than one time point (Fig. E6).

The composition of the gut microbiota associated with the living environment only at one year of age regardless of distance metric (Fig E4, Table E3). Nevertheless, in fully adjusted model, this association attenuated (Table E5), mainly due to the influence of older siblings. The fifteen most abundant genera in the gut microbiota at different time points did not show differences between rural and urban environments after adjustment for multiple testing (Fig E7).

However, the Firmicutes/Bacteroidetes-ratio was lower in rural infants at one year of age (mean in rural: 3.4, urban: 4.1; ANOVA, P=0.0004). A closer look at this ratio revealed that the richness within Bacteroidetes was higher in rural than urban infants at the same time point (mean in rural: 13.4, urban: 11.5; ANOVA, P=0.0003).
Relative concentrations of both airway (1 month) and systemic (6 months) immune mediators differed between urban and rural infants in adjusted analyses for both sample types (Fig 2; Table E6). In airways (Fig 2a), Type 2-related cytokine and chemokine concentrations showed a large disparity between urban and rural infants. For example, urban infants had higher relative concentrations of CCL11, CCL13 and CCL17, while CCL22 and CCL26 were higher in rural infants. Furthermore, IL-12p70 (Type 1) was found in higher relative concentrations in rural infants as well as IL-10 (Regulatory), which has anti-inflammatory properties. A different set of immune mediators was measured in the systemic samples (Fig 2b). CXCL8 (Type 17) was measured in both airway and systemic samples and showed higher relative concentration in urban infants in both sample types. However, other immune mediators measured from both sample types, IL-1β and TNF-α, did not show similar responses.

**Urbanized infant microbiota**

We defined the bacterial community profiles most indicative of urbanization gradient separately for each time point and sampling site. We utilized cross-validated sPLS models to find a minimal (sparse) set of jointly contributing taxa that discriminated between urban and rural samples. The models predicted better during later time points in both airway and gut samples, indicating an accumulated influence of living environment with age (Fig 3). Higher relative abundances of the genera *Veillonella, Rothia, Gemella, Bergyella and Streptococcus* in airway microbial profiles were indicative for urban living environments, while *Moraxella, Paracoccus and Dolosigranulum* were characteristic of a rural environment (Fig 3a) at all time points corresponding to univariate analyses (Figs E6). Some airway bacteria such as *Bacillus* and *Listeria* associated with a rural environment at the first week of life but the direction shifted later to association with urban environments.
Moraxella was also indicative for rural environment and Veillonella for urban environment in the gut microbial profiles (Fig 3b), but only at a single time point each. Bifidobacterium in gut was indicative for rural environment, already in the first week of life. In the gut profiles, only Bacteroides switched direction from urban to rural environment at the last time point. Overall, living environment-related members of gut communities did not match much between time points primarily due to the low predictability at one week and one month as well as the strong shifts in community composition during the first year of life (as described previously²²).

Urbanized microbial profiles, risk of disease and concentrations of immune mediators

A bacterial score was extracted for each sample (how much did the composition of the sample resemble the sPLS trained bacterial signature), with higher values indicating more urbanized bacterial profiles in infants.

Having an urbanized bacterial profile in the samples was associated with higher risk of asthma for all but one week gut and three months airway bacterial scores (Airway 1 week: OR 1.25, 95% CI 1.01-1.55, P=0.0473; Airway 1 month: OR 1.22, 95% CI 1.00-1.48, P=0.0498; Gut 1 month: OR 1.29, 95% CI 1.05-1.59, P=0.0175; Gut 1 year: OR 1.24, 95% CI 1.02-1.53, P=0.0368; Fig 4). The risk of eczema increased only in response to an urbanized gut bacterial profile at one week of age (OR 1.24, 95% CI 1.04-1.48, P=0.0161). Similarly, the risk of sensitization was restricted to one time point and compartment. An urbanized bacterial profile in gut at one year of age was associated with the risk of any sensitization (OR 1.28, 95% CI 1.03-1.59, P=0.0281) and aeroallergen sensitization (OR 1.24, 95% CI 0.98-1.57, P=0.0812) though borderline significant (Fig 4).

The airway bacterial score at one month of age was correlated with immune mediator concentrations at the same time point (Fig 5; Table E6) and the direction of these associations
were similar to our comparison between urban and rural groups (**Fig 2**), but with important distinctions. The systemic CXCL8 concentration was higher with an urbanized bacterial profile (**Fig 5**) corresponding to results for living environment (**Fig 2**), but these associations were borderline after correction for multiple testing. In contrast to the living environment associations, airway IL-2 and IL-13 showed lower concentrations in infants with urbanized bacterial profile. Bacterial scores for the gut microbiota did not significantly associate with immune mediator profiles (**Fig E8; Table E6**).

**DISCUSSION**

*Primary Findings*

The higher risk of asthma and atopic traits in urban than in rural children is a well-established phenomenon. Our prospective, observational cohort study of 700 children supports the hypothesis that the early development of the human microbiota, and its influential crosstalk with the immune system, shapes the predisposition to these diseases in childhood. We found that an urban living environment in infancy was associated with a higher risk of asthma and aeroallergen sensitization in childhood. Urbanized microbiota both in the airways and the gut in infancy was positively associated with the risk of having asthma. The urbanized gut microbiota profile at one week of age also related with the risk of eczema and at one year of age with allergic sensitization. Furthermore, several immune markers showed significant associations with living environment related airway bacteria suggesting an influence of these bacteria in immune development.

*Strengths and Limitations*

A major strength of the study was the extensive and prospective clinical assessment of outcomes and environmental exposures. All children were uniformly diagnosed and repeatedly monitored for development of symptoms of asthma and atopic traits at the COPSAC clinic, in
which also microbial and immunological samples utilized in the study were collected. Our
definition of the living environment was based on land cover data providing unbiased
categorization of rural and urban individuals. However, our observational study is unable to
confirm causal relationships and depends on future studies with interventional or
experimental approaches to validate the mechanisms behind our findings. Further,
methodology we used for microbiota analysis can hide some details such as importance of
bacterial strains.

Even though our findings support the role of urbanized microbiota in disease development,
other explanatory variables can shape the risk. The examined diseases are believed to originate
from multiple triggers acting via a myriad of gene-environment interactions. These can
include exposures other than bacteria such as indoor fungi, which has been associated with
childhood asthma. Additionally, exposure to other molecules e.g. allergens may be crucial as
both allergens and bacterial exposure can be needed for normal immune development.

Further, our analysis lacks information on (air) pollution exposure that has previously been
associated with the development of asthma and is known to differ between rural and urban
Denmark. Therefore, potential differences in the exposure to non-bacterial items could
partly explain the identified differences between living in urban or rural environment. Finally,
we did not have information about mothers’ behavior, environment and health during
pregnancy, which can be important. Moreover, our cohort has limited ethnic diversity, which
reduces the generality of our findings.

**Interpretation**

Exposure to farming-associated indoor microbes, i.e. the “external” microbiota early in life is
associated with asthma protection. Here, we focused on the “internal” microbiota, i.e. the
infant’s own microbiota, in rural and urban children and showed corresponding findings,
especially for airway microbiota. Recently, living in farm has been shown to support the
development of gut microbiota in infants, which can decrease risk of asthma\(^{25}\). These associations are also suggested by *Biodiversity* and *Old Friends* hypotheses\(^{54,55}\). External and internal microbiotas are likely to have intimate relations. The airway microbiota is more open for colonization by air-borne microbes than the gut (because inhaled microbes continuously travel to the lungs) and thus reflects the external microbial exposures. Active mucosal layers in the airways and lungs can respond to microbial exposures and shape the function of immune system. However, as we did not sample from the gut mucosa, we cannot exclude its potential importance.

Bacteria associated with predisposing and protective effects vary between studies\(^{56}\). Members of urbanized bacterial communities probably act together, but *Veillonella* was an interesting genus associated with urbanization both in the airways and the gut. *Veillonella* in the gut at one year of age was previously associated with an increased risk of asthma in this cohort, whereas another study found a protective association\(^{22,23}\). In contrast to our findings, airway colonization with *Moraxella*, an opportunistic pathogen\(^{57}\), was previously associated with increased risk of asthma\(^{57–59}\), also in the previous COPSAC\(_{2000}\) cohort, in which culture-based methodology was applied\(^{59}\). We suspect that the opportunistic (pathogenic) nature of some urban and rural indicator bacteria, such as rural-associated *Moraxella, Acinetobacter*\(^{60,61}\) and *Mycobacterium*\(^{62}\), may explain their potential importance in atopic diseases as well as their cohort specific roles in diseases.

The hypothesized window of susceptibility, i.e. the period when extrinsic microbial stimuli can permanently influence immune responsiveness, is suggested to be during the first 100 days after birth\(^{23}\) or during the entire infancy including the prenatal period\(^{18}\). In this study, the window associated with asthma development was the first year of life while much narrower for eczema and allergic sensitization. Our study was not designed for defining such a window of susceptibility, however. Previously, the window associated with asthma development has
been restricted to shorter periods after birth (less than 100 days and less than 12 months)\textsuperscript{23,24} and longer period for eczema (until 12 month of age)\textsuperscript{21}, leaving it unclear when the exposure is most important.

We found large disparities in the immune mediator concentrations between urban and rural children as well as between urbanized and rural bacterial profiles. Therefore, we conclude that the living environment as well as associated bacteria may influence immune function during infancy. Further, our findings showed interesting similarities to our previous study discovering that asthma-associated bacteria at one-month of age were positively associated with CCL2 and CCL17 while negatively associated to TNF-\(\alpha\) and IL-1\(\beta\)\textsuperscript{38}. These same associations were found in response to urbanized airway bacteria at one-month of age. However, \textit{in vivo} cyto- and chemokine concentrations do not reflect easily delimited individual immune pathways. For example, both Type-2 and Type-17 related cytokines differed between rural and urban infants in a non-orchestrated manner. Some individual mediators showed expected patterns such as CXCL8, a neutrophil chemotactic Type-17 mediator that is enriched in moderate to severe asthma\textsuperscript{63}, which was increased in urban children as well as in response to urbanized airway microbiota. Deficiency in regulatory responses is suspected to be important for the trajectory towards hyperreactivity against allergens and development of manifest clinical allergy\textsuperscript{64}. Accordingly, urban infants had lower concentrations of the anti-inflammatory, regulatory type mediator IL-10. Having an urbanized airway microbiota associated with decreased IL-2 in the airways, which is interesting as suppressed IL-2 expression in infants has been associated with food allergy\textsuperscript{65}.

There are several potential explanations for some immune mediator concentrations differing from the expected. First, the influence of an individual mediator depends on other mediators. For example, a higher concentration of IL-13 (Type-2) in response to rural airway bacteria can indicate that when combined with IL-12p70, TNF-\(\alpha\), IL-1\(\beta\), IL-10 and IL-2, it does not provide
increased susceptibility to Type-2-based symptoms. Second, pro-inflammatory immune mediators can be beneficial during certain developmental stages such as higher IL-4, IL-5 and IL-13 in one-month old rural infants. Our results underscore that the current understanding about an adequately stimulated immune system during different developmental stages is not complete.

**Conclusions**

We found that an urbanized living environment, an urbanized microbiota in infants, *in vivo* immune mediators and the risk of childhood asthma and atopic traits were associated, in a geographically concise cohort of 700 unselected children. Our findings support an important role of the infant microbiota in the association between urbanization and disease, probably via crosstalk with the developing immune system.

**ACKNOWLEDGEMENTS**

We express our deepest gratitude to the children and families of the COPSAC 2010 cohort study for all their support and commitment. We acknowledge and appreciate the unique efforts of the COPSAC research team.

**Authors Contributions:** The guarantor of the study is HB, from conception and design, to conduction of the study and acquisition of data, data analysis, and interpretation of data. JL has written the first draft of the manuscript. All co-authors have provided important intellectual input and contributed considerably to the analyses and interpretation of the data. All authors guarantee that the accuracy and integrity of any part of the work have been appropriately investigated and resolved and all have approved the final version of the manuscript. The corresponding author had full access to the data and had final responsibility for the decision to
submit for publication. No honorarium, grant, or other form of payment was given to any of the authors to produce this manuscript.

**Governance:** We are aware of and comply with recognized codes of good research practice, including the Danish Code of Conduct for Research Integrity. We comply with national and international rules on the safety and rights of patients and healthy subjects, including Good Clinical Practice (GCP) as defined in the EU's Directive on Good Clinical Practice, the International Conference on Harmonisation's (ICH) good clinical practice guidelines and the Helsinki Declaration. Privacy is important to us which is why we follow national and international legislation on General Data Protection Regulation (GDPR), the Danish Act on Processing of Personal Data and the practice of the Danish Data Inspectorate.
REFERENCES


36. Chawes BL, Stokholm J, Bønnelykke K, Brix S, Bisgaard H. Neonates with reduced neonatal lung function have systemic low-grade inflammation. Journal of Allergy and Clinical Immunology. 2015;


Figure legends

**Figure 1. Prevalence of asthma and atopic diseases in rural and urban infants.** The proportions of main land cover types in a radius of 3km around the home at birth are condensed to a gradient from mostly rural to highly urbanized living environments by Principal Components Analysis (PCA). Figure a shows geographical distribution of children, colored by the urbanization gradient. Urban children are mainly clustered around the capital, Copenhagen, as expected. Coordinate information is slightly jittered on the map to preserve anonymity. Figures b and d show the crude prevalence of outcomes at six years of age in children living in urban and rural environment during infancy. Differences are tested with Fisher’s exact tests. Figures c and e show the adjusted odds ratios from logistic regression models for the outcomes in urban vs. rural (baseline) living environment. Urban infants have increased risk for asthma and aeroallergen sensitization. Analyses are adjusted for socio-economics, pet ownership, older siblings, passive smoking exposure and total length of breastfeeding.

**Figure 2. Relative concentration of airway and systemic immune mediators differ between urban and rural infants.** The airway immune profile at one month of age (n = 586, a) and systemic immune profile at six months of age (n = 665, b) in response to urban and rural classes defined by linear models adjusted for living environment associated lifestyle features. Positive estimates indicate higher relative concentration in urban infants while negative estimates indicate higher relative concentration in rural infants. Error bars represent the 95% CI. False positive rate corrected p-values are presented in Table E6.

**Figure 3. The composition of living environment-related microbial profiles for each time point.** sparse Partial Least Squares (sPLS) models were run for (sub-)genera present in at least 5% of samples. Selection of the best model for each time point was done by comparing Root Mean Square Error (RMSE) values between models with differing numbers of taxa and components. Each selected model for airway (a) and gut (b) microbiota had one component. Included taxa were ordered in the heatmap by their summed loadings across all time points. Taxa with positive loadings are more common in urban (gold) living environments, while taxa with negative loadings are more common in rural (green) living environments. Taxa colored white are not included in the present model.

**Figure 4. Airway and gut bacterial scores indicative of living environment and the risk of asthma and atopic traits at 6 years of age.** Bacterial scores were scaled to a standard deviation of one in order to ease interpretation. For airways (a-c), the risk of developing asthma increases with higher values of bacterial score, i.e. urban-like bacterial profiles at one week and one month of age. For gut (d-f), eczema associates with the bacterial score from one week (d), while the one-month and one-year time points are important for asthma development (e). The risk of sensitization increases with higher values of the gut bacterial score at one year of age (f). Error bars represent the 95% CI.

**Figure 5. Airway bacterial score versus the relative concentration of airway and systemic immune mediators.** The figure shows normalized and scaled relative concentrations of immune mediators in relation to airway bacterial score per standard deviation. Higher values indicate association with an urban-like bacterial community composition. Panel A shows the topical immune mediators at one month of age, and panel B the systemic immune mediators at six months of age in response to bacterial scores. Linear models were adjusted for living environment associated lifestyle features. Error bars represent 95% CI. False positive rate corrected p-values are presented in Table E6.
<table>
<thead>
<tr>
<th>Buffer</th>
<th>Rural</th>
<th></th>
<th>Urban</th>
<th></th>
<th>Silhouette width</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>3 km</td>
<td>314</td>
<td>45.8</td>
<td>372</td>
<td>54.2</td>
<td>0.61</td>
</tr>
<tr>
<td>1 km</td>
<td>204</td>
<td>29.7</td>
<td>482</td>
<td>70.3</td>
<td>0.70</td>
</tr>
<tr>
<td>500 m</td>
<td>134</td>
<td>19.5</td>
<td>552</td>
<td>80.5</td>
<td>0.78</td>
</tr>
<tr>
<td>Specific IgE</td>
<td>Rural (n = 238)</td>
<td>Urban (n = 289)</td>
<td>p-value (&gt;0.35)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>% &gt;0.35</td>
<td>n &gt;0.35</td>
<td>median</td>
<td>mean</td>
<td>range</td>
</tr>
<tr>
<td><strong>Food</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Egg</td>
<td>6.7</td>
<td>16</td>
<td>0</td>
<td>0.06</td>
<td>0;2.01</td>
</tr>
<tr>
<td>Milk</td>
<td>9.7</td>
<td>23</td>
<td>0</td>
<td>0.09</td>
<td>0;2.56</td>
</tr>
<tr>
<td>Wheat</td>
<td>5.0</td>
<td>12</td>
<td>0</td>
<td>0.05</td>
<td>0;1.70</td>
</tr>
<tr>
<td>Peanut</td>
<td>0.8</td>
<td>2</td>
<td>0</td>
<td>0.01</td>
<td>0;0.54</td>
</tr>
<tr>
<td>Seafood</td>
<td>0.8</td>
<td>2</td>
<td>0</td>
<td>0.04</td>
<td>0;7.49</td>
</tr>
<tr>
<td><strong>Aeroallergen</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>House dust mite</td>
<td>6.7</td>
<td>16</td>
<td>0</td>
<td>3.07</td>
<td>0;100</td>
</tr>
<tr>
<td>Cat</td>
<td>2.9</td>
<td>7</td>
<td>0</td>
<td>0.51</td>
<td>0;100</td>
</tr>
<tr>
<td>Dog</td>
<td>2.1</td>
<td>5</td>
<td>0</td>
<td>0.17</td>
<td>0;36.08</td>
</tr>
<tr>
<td>Horse</td>
<td>0.4</td>
<td>1</td>
<td>0</td>
<td>0.08</td>
<td>0;18.10</td>
</tr>
<tr>
<td>Timothy</td>
<td>6.7</td>
<td>16</td>
<td>0</td>
<td>0.12</td>
<td>0;6.08</td>
</tr>
<tr>
<td>Birch</td>
<td>3.8</td>
<td>9</td>
<td>0</td>
<td>0.9</td>
<td>0;100</td>
</tr>
<tr>
<td>Mugworth</td>
<td>1.7</td>
<td>4</td>
<td>0</td>
<td>0.02</td>
<td>0;0.81</td>
</tr>
<tr>
<td><strong>Molds</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Herbarum</td>
<td>0.4</td>
<td>1</td>
<td>0</td>
<td>0.02</td>
<td>0;5.28</td>
</tr>
<tr>
<td>A. Fumigatus</td>
<td>0.4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0;0.98</td>
</tr>
<tr>
<td>A. Tenuis</td>
<td>0.4</td>
<td>1</td>
<td>0</td>
<td>0.14</td>
<td>0;33.96</td>
</tr>
<tr>
<td>Body site</td>
<td>Age</td>
<td>n</td>
<td>Distance Metric</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>------</td>
<td>-----------------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Bray-Curtis R2</td>
<td>p</td>
<td>Weighted UniFrac R2</td>
</tr>
<tr>
<td>Gut</td>
<td>1 week</td>
<td>547</td>
<td>0.002</td>
<td>0.3</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>1 month</td>
<td>600</td>
<td>0.001</td>
<td>0.9</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>1 year</td>
<td>612</td>
<td>0.003</td>
<td>0.005</td>
<td>0.004</td>
</tr>
<tr>
<td>Airways</td>
<td>1 week</td>
<td>538</td>
<td>0.006</td>
<td>0.001</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>1 month</td>
<td>629</td>
<td>0.006</td>
<td>0.001</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>3 months</td>
<td>633</td>
<td>0.006</td>
<td>0.001</td>
<td>0.007</td>
</tr>
<tr>
<td>Living environment gradient +</td>
<td>1 week</td>
<td>1 month</td>
<td>3 months</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>R2</td>
<td>p</td>
<td>n</td>
<td>R2</td>
</tr>
<tr>
<td>Unadjusted model</td>
<td>538</td>
<td>0.005</td>
<td>0.001</td>
<td>629</td>
<td>0.006</td>
</tr>
<tr>
<td>Furry pet at home</td>
<td>530</td>
<td>0.004</td>
<td>0.001</td>
<td>622</td>
<td>0.003</td>
</tr>
<tr>
<td>Daycare attendance before age 1</td>
<td>538</td>
<td>0.006</td>
<td>0.001</td>
<td>629</td>
<td>0.006</td>
</tr>
<tr>
<td>Total breastfeeding duration</td>
<td>533</td>
<td>0.005</td>
<td>0.001</td>
<td>626</td>
<td>0.007</td>
</tr>
<tr>
<td>Passive smoking exposure</td>
<td>530</td>
<td>0.005</td>
<td>0.001</td>
<td>622</td>
<td>0.007</td>
</tr>
<tr>
<td>Socio-economics</td>
<td>516</td>
<td>0.005</td>
<td>0.001</td>
<td>604</td>
<td>0.007</td>
</tr>
<tr>
<td>Older siblings</td>
<td>531</td>
<td>0.005</td>
<td>0.001</td>
<td>622</td>
<td>0.005</td>
</tr>
<tr>
<td>Adjusted for all</td>
<td>514</td>
<td>0.004</td>
<td>0.002</td>
<td>602</td>
<td>0.005</td>
</tr>
<tr>
<td>Gut microbiota</td>
<td>1 week</td>
<td></td>
<td>1 month</td>
<td></td>
<td>1 year</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>--------</td>
<td>----------</td>
<td>---------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>R^2</td>
<td>p</td>
<td>n</td>
<td>R^2</td>
</tr>
<tr>
<td>Unadjusted model</td>
<td>547</td>
<td>0.002</td>
<td>0.3</td>
<td>600</td>
<td>0.002</td>
</tr>
<tr>
<td>Furry pet at home</td>
<td>540</td>
<td>0.002</td>
<td>0.2</td>
<td>591</td>
<td>0.002</td>
</tr>
<tr>
<td>Daycare attendance before age 1</td>
<td>547</td>
<td>0.002</td>
<td>0.4</td>
<td>600</td>
<td>0.001</td>
</tr>
<tr>
<td>Total breastfeeding duration</td>
<td>543</td>
<td>0.002</td>
<td>0.4</td>
<td>595</td>
<td>0.002</td>
</tr>
<tr>
<td>Passive smoking exposure</td>
<td>540</td>
<td>0.002</td>
<td>0.3</td>
<td>591</td>
<td>0.002</td>
</tr>
<tr>
<td>Socio-economics</td>
<td>527</td>
<td>0.002</td>
<td>0.3</td>
<td>574</td>
<td>0.002</td>
</tr>
<tr>
<td>Older siblings</td>
<td>541</td>
<td>0.002</td>
<td>0.3</td>
<td>592</td>
<td>0.002</td>
</tr>
<tr>
<td>Adjusted for all</td>
<td>525</td>
<td>0.002</td>
<td>0.2</td>
<td>572</td>
<td>0.001</td>
</tr>
<tr>
<td>Explanatory variable</td>
<td>Time</td>
<td>Immune marker</td>
<td>Immune response type</td>
<td>Estimate</td>
<td>Std.error</td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------------</td>
<td>---------------</td>
<td>----------------------</td>
<td>----------</td>
<td>-----------</td>
</tr>
<tr>
<td>Living environment (Rural/Urban)</td>
<td>One month</td>
<td>CCL2</td>
<td>Type 1</td>
<td>1.346</td>
<td>0.091</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCL4</td>
<td>Type 1</td>
<td>1.176</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CXCL10</td>
<td>Type 1</td>
<td>1.116</td>
<td>0.093</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IFNg</td>
<td>Type 1</td>
<td>1.027</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL12p70</td>
<td>Type 1</td>
<td>0.778</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TNFa</td>
<td>Type 1</td>
<td>1.104</td>
<td>0.093</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCL11</td>
<td>Type 1</td>
<td>1.344</td>
<td>0.091</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCL13</td>
<td>Type 2</td>
<td>1.360</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCL17</td>
<td>Type 2</td>
<td>1.301</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCL22</td>
<td>Type 2</td>
<td>0.724</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCL26</td>
<td>Type 2</td>
<td>0.792</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL13</td>
<td>Type 2</td>
<td>0.842</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL4</td>
<td>Type 2</td>
<td>0.903</td>
<td>0.093</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL5</td>
<td>Type 2</td>
<td>0.875</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CXCL18</td>
<td>Type 17</td>
<td>1.267</td>
<td>0.091</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL17</td>
<td>Type 17</td>
<td>1.097</td>
<td>0.093</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL1b</td>
<td>Type 17</td>
<td>0.942</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL10</td>
<td>Regulatory</td>
<td>0.749</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL2</td>
<td>Regulatory</td>
<td>0.975</td>
<td>0.093</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TGFb1</td>
<td>Regulatory</td>
<td>1.035</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TNFa</td>
<td>Type 1</td>
<td>0.867</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CXCL18</td>
<td>Type 17</td>
<td>1.450</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL1b</td>
<td>Type 17</td>
<td>0.704</td>
<td>0.089</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL6</td>
<td>Type 17</td>
<td>0.918</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CRP</td>
<td>Generic</td>
<td>1.129</td>
<td>0.09</td>
</tr>
<tr>
<td>Bacterial score (airways 1 month)</td>
<td>One month</td>
<td>CCL2</td>
<td>Type 1</td>
<td>1.107</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCL4</td>
<td>Type 1</td>
<td>0.96</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CXCL10</td>
<td>Type 1</td>
<td>1.043</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IFNg</td>
<td>Type 1</td>
<td>1.043</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL12p70</td>
<td>Type 1</td>
<td>0.909</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TNFa</td>
<td>Type 1</td>
<td>0.9</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCL11</td>
<td>Type 2</td>
<td>1.048</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCL13</td>
<td>Type 2</td>
<td>1.086</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCL17</td>
<td>Type 2</td>
<td>1.125</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCL22</td>
<td>Type 2</td>
<td>1.031</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCL26</td>
<td>Type 2</td>
<td>0.977</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL13</td>
<td>Type 2</td>
<td>0.877</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL4</td>
<td>Type 2</td>
<td>0.981</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL5</td>
<td>Type 2</td>
<td>1.066</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CXCL8</td>
<td>Type 17</td>
<td>0.919</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL17</td>
<td>Type 17</td>
<td>0.913</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL1b</td>
<td>Type 17</td>
<td>0.863</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL10</td>
<td>Regulatory</td>
<td>0.902</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL2</td>
<td>Regulatory</td>
<td>0.867</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TGFb1</td>
<td>Regulatory</td>
<td>1.007</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TNFa</td>
<td>Type 1</td>
<td>0.946</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CXCL8</td>
<td>Type 17</td>
<td>1.102</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL1b</td>
<td>Type 17</td>
<td>1.009</td>
<td>0.041</td>
</tr>
<tr>
<td>Bacterial score (gut 1 month)</td>
<td>One month</td>
<td>Six months</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacterial score (air 3 months)</td>
<td>Journal Pre-proof</td>
<td>Journal Pre-proof</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacterial score (gut 1 month)</td>
<td>Journal Pre-proof</td>
<td>Journal Pre-proof</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cytokine</th>
<th>Type</th>
<th>Value</th>
<th>SEM</th>
<th>t-value</th>
<th>p-value</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL6</td>
<td>Type 17</td>
<td>1.002</td>
<td>0.043</td>
<td>0.922</td>
<td>1.089</td>
<td>0.956</td>
</tr>
<tr>
<td>CRP</td>
<td>Generic</td>
<td>0.977</td>
<td>0.042</td>
<td>0.9</td>
<td>1.060</td>
<td>0.573</td>
</tr>
<tr>
<td>CCL2</td>
<td>Type 1</td>
<td>1.038</td>
<td>0.044</td>
<td>0.954</td>
<td>1.129</td>
<td>0.391</td>
</tr>
<tr>
<td>CCL4</td>
<td>Type 1</td>
<td>1.001</td>
<td>0.042</td>
<td>0.922</td>
<td>1.087</td>
<td>0.982</td>
</tr>
<tr>
<td>CXCL10</td>
<td>Type 1</td>
<td>1.025</td>
<td>0.043</td>
<td>0.942</td>
<td>1.115</td>
<td>0.57</td>
</tr>
<tr>
<td>IFNγ</td>
<td>Type 1</td>
<td>0.987</td>
<td>0.045</td>
<td>0.904</td>
<td>1.078</td>
<td>0.772</td>
</tr>
<tr>
<td>IL12p70</td>
<td>Type 1</td>
<td>0.978</td>
<td>0.044</td>
<td>0.897</td>
<td>1.066</td>
<td>0.617</td>
</tr>
<tr>
<td>TNFα</td>
<td>Type 1</td>
<td>0.968</td>
<td>0.043</td>
<td>0.888</td>
<td>1.054</td>
<td>0.418</td>
</tr>
<tr>
<td>CCL11</td>
<td>Type 2</td>
<td>1.043</td>
<td>0.043</td>
<td>0.959</td>
<td>1.134</td>
<td>0.328</td>
</tr>
<tr>
<td>CCL13</td>
<td>Type 2</td>
<td>1.054</td>
<td>0.041</td>
<td>0.972</td>
<td>1.142</td>
<td>0.204</td>
</tr>
<tr>
<td>CCL17</td>
<td>Type 2</td>
<td>1.034</td>
<td>0.041</td>
<td>0.953</td>
<td>1.121</td>
<td>0.419</td>
</tr>
<tr>
<td>CCL22</td>
<td>Type 2</td>
<td>0.975</td>
<td>0.041</td>
<td>0.899</td>
<td>1.057</td>
<td>0.539</td>
</tr>
<tr>
<td>CCL26</td>
<td>Type 2</td>
<td>1.001</td>
<td>0.043</td>
<td>0.92</td>
<td>1.089</td>
<td>0.979</td>
</tr>
<tr>
<td>IL13</td>
<td>Type 2</td>
<td>0.913</td>
<td>0.044</td>
<td>0.837</td>
<td>0.995</td>
<td>0.04</td>
</tr>
<tr>
<td>IL4</td>
<td>Type 2</td>
<td>0.963</td>
<td>0.044</td>
<td>0.884</td>
<td>1.049</td>
<td>0.386</td>
</tr>
<tr>
<td>IL5</td>
<td>Type 2</td>
<td>1.024</td>
<td>0.043</td>
<td>0.942</td>
<td>1.114</td>
<td>0.579</td>
</tr>
<tr>
<td>CXCL8</td>
<td>Type 17</td>
<td>1.023</td>
<td>0.043</td>
<td>0.941</td>
<td>1.112</td>
<td>0.589</td>
</tr>
<tr>
<td>IL17</td>
<td>Type 17</td>
<td>1.067</td>
<td>0.043</td>
<td>0.981</td>
<td>1.162</td>
<td>0.132</td>
</tr>
<tr>
<td>IL1b</td>
<td>Type 17</td>
<td>0.992</td>
<td>0.043</td>
<td>0.912</td>
<td>1.079</td>
<td>0.85</td>
</tr>
<tr>
<td>IL10</td>
<td>Regulatory</td>
<td>1.036</td>
<td>0.044</td>
<td>0.951</td>
<td>1.128</td>
<td>0.417</td>
</tr>
<tr>
<td>IL2</td>
<td>Regulatory</td>
<td>0.974</td>
<td>0.044</td>
<td>0.894</td>
<td>1.062</td>
<td>0.553</td>
</tr>
<tr>
<td>TGFβ1</td>
<td>Regulatory</td>
<td>1.021</td>
<td>0.043</td>
<td>0.938</td>
<td>1.112</td>
<td>0.627</td>
</tr>
<tr>
<td>TNFα</td>
<td>Type 1</td>
<td>1.002</td>
<td>0.043</td>
<td>0.922</td>
<td>1.089</td>
<td>0.969</td>
</tr>
<tr>
<td>CXCL8</td>
<td>Type 17</td>
<td>1.046</td>
<td>0.041</td>
<td>0.966</td>
<td>1.132</td>
<td>0.272</td>
</tr>
<tr>
<td>IL1b</td>
<td>Type 17</td>
<td>0.966</td>
<td>0.041</td>
<td>0.892</td>
<td>1.047</td>
<td>0.405</td>
</tr>
<tr>
<td>IL6</td>
<td>Type 17</td>
<td>1.010</td>
<td>0.043</td>
<td>0.928</td>
<td>1.100</td>
<td>0.812</td>
</tr>
<tr>
<td>CRP</td>
<td>Generic</td>
<td>1.060</td>
<td>0.042</td>
<td>0.976</td>
<td>1.152</td>
<td>0.165</td>
</tr>
</tbody>
</table>

*Note: SEM = Standard Error of the Mean, t-value = t-statistic, p-value = probability value, Effect Size = Cohen's d*
Relative concentration

Airway, 1 month

Type 1

Type 2

Type 17

Regulatory

Systemic, 6 months

Type 1

Type 17

Generic

Journal Pre-proof
Airway microbiota

Veillonella
Rothia
Gemella
Bergeyella
Streptococcus
Atopobium
Bifidobacterium
Clostridium (ss 1)
Lactobacillus
Campylobacter
Fusobacterium
Haemophilus
Prevotella
Scardovia
Enterococcus
Escherichia Shigella
Bacteroides
Actinomyces
Ureaplasma
Alloprevotella
Prevotella (7)
Lachnoanaerobaculum
Rhodanobacter
Family Burkholderiaceae
Deinococcus
Pseudomonas
Order Saccharimonadales
Anaerococcus
Novosphingobium
Prevotella (6)
Order Chloroplast
Porphyromonas
Rhodobacter
Megasphaera
Order Myxococcales
Aquabacterium
Psychrobacter
Helicobacter
Faecalibacterium
Bacillus
Leptotrichia
Methylbacterium
Tumebacillus
Prevotella (9)
Johnsonella
Streptobacillus
Succinivibrionaceae (UCG 001)
Granulicatella
Actinobacillus
Sediminibacterium
Acinetobacter
Acidibacter
Mycobacterium
Pelomonas
Listeria
Family Enterobacteriaceae
Staphylococcus
Corynebacterium (1)
Sphingomonas
Micrococcus
Dolosigranulum
Paracoccus
Moraxella

Gut microbiota

Senagalmassilia
Akkermansia
Campylobacter
Phascolarctobacterium
Veillonella
Lachnoclostridium
Proteus
Escherichia Shigella
Fusobacterium
Bacteroides
Atopobium
Family Enterobacteriaceae
Tyzzerella (4)
Eggerthella
Sutterella
Ruminiclostridium (5)
Order Rhodospirillales
Ruminococcaceae UCG002
Gemella
Ruminococcus torques group
Alistipes
Streptococcus
Moraxella
Barnesia
Negativococcus
Dorea
Tyzzerella
Dolosigranulum
Prevotella (9)
Collinsella
Bifidobacterium
Family Lachnospiraceae

Rural

Urban

sPLS loadings

-0.3 0 0.3 0.6 0.9

R² = 0.038  R² = 0.069  R² = 0.062
P = 2.35e-09  P = 3.21e-14  P = 1.26e-10
Airway bacterial scores

Asthma
Allergic rhinitis
Eczema
Any sensitization
Aeroallergen sensitization
Food sensitization

Gut bacterial scores

Asthma
Allergic rhinitis
Eczema
Any sensitization
Aeroallergen sensitization
Food sensitization

1 week

P = 0.047

1 month

P = 0.049

3 months

P = 0.10

1 year

P = 0.037

Odds ratio (log scale)
Airway immune profile (1 month) + airway bacterial score (1 month)

Systemic immune profile (6 months) + Airway bacterial score (3 months)