# Early Origins of Childhood Asthma:
# Role of Perinatal Microbial Exposure and Innate Immunity

BY

BENJAMIN TURTURICE
B.S., Loyola University Chicago, 2012

THESIS

Submitted as partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Microbiology and Immunology in the
Graduate College of the University of Illinois at Chicago, 2021

Chicago, Illinois

Defense Committee:
Patricia W. Finn, Co-advisor
David L. Perkins, Co-advisor
Nancy Frietag, Chair
Alan McLachlan
Deepak Shukla
Viswanathan Natarajan, Pharmacology

This dissertation is dedicated to those who have given their time to mentor me as a future clinician and scientist…

Nikolai Sopko
Marc Penn
Stefan Kanzok
Kim Williamson
Asha Rani
Ravi Ranjan
Christian Ascoli
Patricia Finn
David Perkins

And to my parents who have given me unconditional support.

Jim Turturice
Diane Lockard

## Acknowledgments

During the course of my graduate education I have had the pleasure of working with and being mentored by multitude of idividuals who I cannot thank enough. Each has helped me develop as future physician scientist in their own specific way. I could not have completed this dissertation without them and I am forever grateful.

I would like to acknowledge the tremendous support I have received from both my advisors, Patricia Finn and David Perkins. They have served role models on what it means to practice in medicine, science, and leadership. I would also like to thank the support and guidance I have received from my thesis committee: Nancy Frietag, Alan McLachlan, Deepak Shukla, and Viswanathan Natarajan. They have helped me navigate through the difficult path of doctoral research. I have had the joy working with great lab mates past and present. I wish thank Asha Rani, Ravi Ranjan, and Christian Ascoli for giving me guidance, assistance, and support in my more naïve years of graduate education. Last, I want to express my sincere gratitude to the leadership and administration – Larry Tobacman, Nancy Frietag, Roberta Bernstein, and Julie Mann -- of the Medical Scientist Training Program for giving me the opportunity and honor of doctoral research.

## Contribution of Authors

The authors who contributed to this dissertation are listed with their respective contributions under each work.

1. Ozturk AB, **Turturice BA**, Perkins DL, Finn PW. The Potential for Emerging Microbiome-Mediated Therapeutics in Asthma. *Curr Allergy Asthma Rep.* 2017 Aug 10;17(9):62. doi: 10.1007/s11882-017-0730-1.

Conception and design: O.A.B., D.L.P., and P.W.F.; manuscript writing: O.A.B., B.A.T., D.L.P., and P.W.F.; manuscript editing: O.A.B., B.A.T., D.L.P., and P.W.F.

2. **Turturice BA**, Ranjan R, Nguyen B, Hughes LM, Andropolis KE, Gold DR, Litonjua AA, Oken E, Perkins DL, Finn PW. Perinatal bacterial exposure contributes to IL-13 aeroallergen response. *Am J Respir Cell Mol Biol* 2017 Oct;57(4):419-427. doi: 10.1165/rcmb.2017-0027OC.

Conception and design: B.A.T., D.L.P., and P.W.F.; performed experiments: B.A.T., R.R., B.N., and K.E.A.; analysis and interpretation: B.A.T., L.M.H., D.L.P., and P.W.F.; reagents and materials: D.L.P. and P.W.F.; manuscript writing: B.A.T., D.L.P., and P.W.F.; manuscript editing: D.R.G.,A.A.L., and E.O.

3. **Turturice BA**, Gold DR, Litonjua AA, Oken E, Rifas-Shiman S, Perkins DL, Finn PW. Lower perinatal exposure to Proteobacteria is an independent predictor of early childhood wheezing. *J Allergy Clin Immunol.* 2019 Jan;143(1):419-421.e5. doi: 10.1016/j.jaci.2018.06.051. Epub 2018 Sep 8.

Conception and design: B.A.T., D.L.P., and P.W.F.; performed experiments: B.A.T.; analysis and interpretation: B.A.T., S.R.S., D.L.P., and P.W.F.; reagents and materials: D.L.P. and P.W.F.; manuscript writing: B.A.T., D.L.P., and P.W.F.; manuscript editing: D.R.G., A.A.L., and E.O.

4. **Turturice BA**, Theorell J, Koenig MD, Tussing-Humphreys L, Gold DR, Litonjua AA, Oken E, Rifas-Shiman SL, Perkins DL, Finn PW. Perinatal granulopoiesis and risk of pediatric asthma. Elife. 2021 Feb 10;10:e63745. doi: 10.7554/eLife.63745. PMID: 33565964; PMCID: PMC7889076.

Conception and design: B.A.T., D.L.P., and P.W.F.; performed experiments: B.A.T., J.T.; analysis and interpretation: B.A.T., S.R.S., D.L.P., and P.W.F.; reagents and materials: M.D.K., L.T.H., D.L.P., and P.W.F.; manuscript writing: B.A.T., D.L.P., and P.W.F.; manuscript editing: D.R.G., A.A.L., and E.O.

Work from Chapter 5 contains unpublished work that was contributed to by:
**Turturice BA**, Gold DR, Litonjua AA, Oken E, Rifas-Shiman S, Perkins DL, Finn PW

Conception and design: B.A.T., D.L.P., and P.W.F.; performed experiments: B.A.T.; analysis and interpretation: B.A.T., S.R.S., D.L.P., and P.W.F.; reagents and materials: D.L.P. and P.W.F.; manuscript writing: B.A.T., D.L.P., and P.W.F.; manuscript editing: D.R.G., A.A.L., and E.O.

## Table of Contents



## Table of Contents (continued)

**Chapter** **Page**

## Table of Contents (continued)

**Chapter** **Page**

## LIST OF TABLES

**Table** **Page**

## LIST OF FIGURES

**Figure** **Page**

## LIST OF FIGURES (continued)

**Figure** **Page**

## LIST OF ABREVIATIONS

| | |
|---|---|
| α | Alpha |
| aHR | Adjusted Hazard Ratio |
| AHR | Airway Hyper-reactivity |
| ANOSIM | Analysis of Similarities |
| ANOVA | Analysis of Variance |
| ANV | Amplicon Nucleotide Variants |
| aOR | Adjusted Odds Ratio |
| aβ | Adjusted Beta Coefficient |
| β | Beta |
| BAL | Broncho-alveolar Lavage |
| Bla G2 | Blattella germanica 2 |
| BMI | Body Mass Index |
| bp | base pairs |
| CBMC | Cord Blood Mononuclear Cell |
| CD | Cluster Differentiation |
| cfbDNA | Circulating Free Bacterial DNA |
| CI [95%] | 95% Confidence Interval |
| CM | Central Memory |
| CyTOF | Mass Cytometry |
| Der F1 | Dermatophagoides farine 1 |
| EM | Effector Memory |
| FBS | Fetal Bovine Serum |

| | |
|---|---|
| FEV1 | Forced Expiratory Volume 1 second |
| FVC | Forced Vital Capacity |
| FWER | Family Wise Error Rate |
| GBH | Group Bejamini-Hochberg |
| GI | Gastrointestinal |
| GO | Gene Ontology |
| GSDMA | Gasdermin A |
| GSDMB | Gasdermin B |
| HDM | House Dust Mite |
| HR | Hazard Ratio |
| HRV-C | Human Rhinovirus C |
| ICS | Inhaled Corticosteroid |
| IgE | Immunoglobulin Epsilon |
| lFNγ | Interferon Gamma |
| IL | Interleukin |
| LPS | Lipopolysaccharide |
| mAPI | modified Asthma Predictive Index |
| MAPK | Mitogen Activated Protein Kinase |
| mRNAseq | Messenger RNA Sequencing |
| NR | Non-Responders |
| OR | Odds Ratio |
| OTU | Operational Taxonomic Unit |
| OVA | Ovalbumin |

| | |
|---|---|
| PBMC | Peripheral Blood Mononuclear Cell |
| PC | Principal Component |
| PCR | Polymerase Chain Reaction |
| PGLYRP-1 | Peptidogylcan Recognition Protein 1 |
| PHA | Phytohemagglutinin |
| PRR | Pattern Recognition Receptors |
| qPCR | Quantitative PCR |
| R | Responders |
| RIN | RNA Integrity Number |
| rRNA | Ribosomal RNA Subunit |
| RSV | Respiratory Syncytial Virus |
| SCFA | Short Chain Fatty Acid |
| SD | Standard Deviation |
| SE | Standard Error |
| SI | Stimulation Index (log2 fold change over media) |
| SNP | Single Nucleotide Polymorphism |
| $T_h$ | T-cell helper |
| TLR | Toll-like Receptor |
| UIC | University of Illinois at Chicago |
| UIH | University of Illinois Hospital |
| UniFrac | Unique Fraction Metric |
| V1-V3 | Variable Regions 1-3 16s rRNA gene |
| V4 | Variable Region 4 16s rRNA gene |

## SUMMARY

Asthma—defined by reversible expiratory airflow obstruction-- is the most common chronic disease of childhood in the United States. While many triggers (e.g. bacteria, viruses, allergens, and pollution) have been identified as causative agents of asthma, there is no single etiology that drives reversible obstruction children. The pathogenesis of childhood asthma likely stems from an individual x environment interaction that is highly complex. The focus of the studies of the Finn-Perkins lab described here was to identify microbial exposures and immunological differences at birth that are predictive of asthma development, asthmatic response to triggers, and severity of asthmatic symptoms.

Although there is no known single etiologic agent, children with asthma have demonstrable skewing of T-cell responses toward $T_{h2}$ or “allergic” processes. Using an unbiased approach, we assessed relationships between perinatal microbial exposure and $T_{h2}$ responses to common air-borne allergens that are known asthma triggers. We identified a specific phyla of Gram negative organisms, Proteobacteria, with negative associations with $T_{h2}$ responses. Further, increased abundance of Proteobacteria was significantly associated with reduced risk of wheezing childhood. Together, this is suggestive of a protective role of early exposure to Proteobacteria.

While $T_{h2}$ responses are a hallmark of childhood asthma, it is not clear what predisposes to this immunologic skewing later in childhood. There are many risk factors for childhood asthma that have been identified, and many of these are discernable at birth (e.g. sex, mode of delivery, gestational age…). We demonstrated that the pre/perinatal risk factors have additive effects and in using an unbiased approach, we were able to identify a transcriptional signature associated with pre/perinatal risk stratification.

Additionally, we determined that umbilical cord serum PGLYRP-1 concentrations were predictive of this transcriptional signature, current asthma at 7 years old, and FEV1/FVC at 12 years old. The concentration of PGLYRP-1 was associated with perinatal exposure to Proteobacteria and several Proteobacteria genera. Collectively, this suggests that early exposure to Proteobacteria enhances innate immunity and, through enhanced innate immunity, may protect against future development of asthma. Additionally, these findings demonstrate that accurate prediction of asthma risk may be capable at birth using a serological marker, PGLYRP-1.

# CHAPTER 1

# BACKGROUND

* This chapter's text, tables, and figures contain portions of work that is reprinted from:

Ozturk AB, **Turturice BA**, Perkins DL, Finn PW. The Potential for Emerging Microbiome-Mediated Therapeutics in Asthma. *Curr Allergy Asthma Rep.* 2017 Aug 10;17(9):62. doi: 10.1007/s11882-017-0730-1.

And

**Turturice BA**, Theorell J, Koenig MD, Tussing-Humphreys L, Gold DR, Litonjua AA, Oken E, Rifas-Shiman SL, Perkins DL, Finn PW. Perinatal granulopoiesis and risk of pediatric asthma. Elife. 2021 Feb 10;10:e63745. doi: 10.7554/eLife.63745. PMID: 33565964; PMCID: PMC7889076.

## A. Childhood Asthma

### A.i. Clinical Features and Epidemiology

Asthma is one of the most common chronic diseases childhood, specifically in developed countries, and is defined by reversible expiratory airflow limitations. This airflow limitation is driven by chronic inflammation of the bronchi and produces symptoms such as wheezing and coughing (1). Asthma exacerbations, times of severe airflow limitation, are medical emergencies and can be life threatening (2). This obstructive physiology can be reversed by the use of medications targeting smooth muscle relaxation, such as albuterol (β-adrenergic agonist) and ipratropium (muscarinic cholinergic antagonist), and usually leads to abatement of symptoms. Long term management of asthma is maintained with inhaled corticosteroids to limit chronic inflammation (3). This reversibility differentiates asthma from other obstructive lung diseases of childhood like bronchopulmonary dysplasia and cystic fibrosis.

Reversible expiratory airflow limitation in children is largely assessed by auscultation of the lungs during times when symptoms are present and absent (4). This determination remains somewhat subjective as there are many different etiologies of wheezing in children. This is different from expiratory assessments of asthmatics in adults where airflow can be measured accurately by spirometry and specific diagnostic tests (e.g. bronchodilator reversibility and bronchial provocation) (4). This subjectivity and inability to perform validated testing leads to challenges in diagnosing children with asthma where many children are likely misdiagnosed or their symptoms are disregarded as being caused by another pathology, such as viral illness (5). The difficulty in asthma diagnosis of children is reflected in the epidemiology of the disease.

Childhood asthma is highly prevalent worldwide, and in the U.S alone there are approximately 9 million children who have been diagnosed with asthma. There are less children, approximately 6 million that report currently having asthma, which may be due to resolution or misdiagnosis (6). The prevalence of childhood asthma increased drastically, approximately 10 fold, from the 1960s into the 2000s in many developed countries around the world, although recently this trend has appeared plateau and may even be declining slightly (7). This rapid increase in prevalence is far too rapid to be explained solely by genetic factors, and it is believed that this increase is due to changing environments and exposures that substantially impact risk.

A.ii. Clinical Determinants of Risk

Asthma is a disease that does not have known origin of pathogenesis. Yet, there are many known factors that predispose an individual to higher risk of developing asthma

in childhood (8). These risk factors can be roughly grouped as genetic, pre/perinatal, and postnatal. Genetic risk factors that have been identified are largely restricted to single nucleotide polymorphisms (SNPs), have a dependence on ancestral background (e.g. Northern European vs. African descent), and remain under-explored in relationship to environmental interactions (9). Pre/perinatal risk factors are discernable differences that can be determined before or around the time of birth and are mostly *in utero* exposures (e.g. maternal smoking) and demographic differences at birth (e.g. mode of delivery, birth weight, etc). Postnatal risk factors occur throughout childhood—though tend to cluster around early infancy—and include environmental exposures (e.g. pollution, smoke, and allergen exposure) or life style differences (e.g. diet and sedentary lifestyles).

Genetic risk has been extensively studied in children who have an ancestral background of Northern European descent (9). While there have been over 100 SNPs associated with childhood risk of asthma, the effect of these individual SNPs is relatively small. Through meta-analysis only several loci have been identified having significant (*p*-value $\leq 5e10^{-7}$) associations with childhood onset of asthma: *IL18R1, IL33, IL13, GSDMA, GSDMB* (10). None of these SNPs were associated with later onset of asthma. Reflecting the small effect size seen in most SNPs, the highest confidence association (*GSDMB*, *p*-value = $6.4e10^{-23}$) only reduced the odds of asthma 25% (10). *GSDMA* and *GSDMB* are located on the 17q21 locus, which has been the most extensively studied with regards to asthma risk in children (11). This locus displays lower linkage disequilibrium, along with smaller effect sizes that are more inconsistent when assessing phenotype associations in populations of African descent (12-14). Additionally, this locus has been shown to interact with environmental exposures such tobacco smoke, viral infections, and animal

dander (15-18). While asthma has been predicted to have moderate heritability, the interaction between one's genetic ancestral background, environment, and specific SNPs have made it hard to assess an individual's risk for childhood asthma based on genetics.

Many of the demographic factors that have been associated with childhood asthma risk can be determined at the time of birth. Several of these are related to differences in birth characteristics: birth weight, gestational age at birth, and mode of delivery. In several meta-analyses assessing the risk factors stated above, the risk of childhood asthma associated with these factors approximately ranged from a 10-40% increase (19-22). Notably, these risk factors tend be correlated with each other; for example, individuals born pre-term are more likely have lower birth weights. This likely confounds individual assessment of these risk factors. Other pre/perinatal risk factors—such as a child's sex and race-- are identified from birth but likely have effects that extend throughout childhood. In the United Stated, approximately 9.5% of males under the age of 18 are asthmatic compared to 7.3% of females being asthmatics (6). Additionally, 12.8% those who are Black or African American are asthmatics compared to the population average of 8.4% (6). The effect of these pre/perinatal risk factors is likely mediated through a complex interaction between biology and socio-economics.

Postnatal risk factors are most closely associated with specific exposures that occur in the first several years of life. These include protective exposures such as breastfeeding and exposure to farm animals, whereas harmful exposures include viral infections, exposure to pollution/ tobacco smoke, and increased exposure to allergens (23-27). The current literature suggests that postnatal exposures likely influence immune tolerance and inflammatory response to foreign antigens (28). It is likely that postnatal

exposures interact with an individual's genetic and pre/perinatal susceptibility as a risk determinant for asthma.

B. Current Risk Prediction Strategies

Prediction of asthma remains an important and difficult issue to address. Early identification of individuals who are at elevated risk for asthma may allow for appropriate interventions that could alleviate the morbidity associated with asthma, whereas misdiagnosis could expose children to harmful effects of pharmacologic interventions, such as stunted growth due corticosteroids (29). The most well studied prediction model is the modified Asthma Predictive Index (mAPI). This test is meant for children under the age three with four or more wheezing episodes per year. Positive tests must include one major criterion (parent with asthma, child with eczema, or child with sensitivity to inhaled allergens) or two minor criterion (wheezing unrelated to infections, eosinophils ≥ 4% on complete blood count, or patient with a food allergy). The test is reported to have high specificity (>90%) but low sensitivity (<25%) (30). Thus the use of this test likely prevents the misdiagnosis of non-asthmatics but would lead to a large group of asthmatic children remaining undiagnosed. In a systematic review of current risk prediction models, twelve additional high quality prediction models were identified (31). Similar issues were noted in each of these risk prediction models. A large tradeoff between sensitivity and specificity was seen across these models. Additionally, pre-test probability of asthma appears to heavily influence many of these models. The use of clinical tests such as blood draws, skin prick tests, or exhaled nitric oxide were required for 8 of 12 models as well as for the mAPI (31). All of these clinical tests assess inflammation or immunity that is generated

through $T_{h2}$ mediated processes which only represents one phenotype of asthma. These factors can potentially limit the clinical utility of these test in different populations, especially those with different socio-economic status or in different geographic locations.

C. Impaired Immunity in Asthmatics

C.i. Inflammation of Asthmatics Airways of Children

Inflammation and immunity are critical to the pathophysiology of asthma. The symptoms of asthma are due to an underlying chronic inflammatory state in the bronchi. In childhood, the most common source of this inflammation is $T_{h2}$ $CD4^+$ T-cells that secrete cytokines IL-4, IL-5, and IL-13 (32). The effect of these cytokines is pleiotropic in nature, as they enhance features of asthma such as mucus hyper-secretion, goblet cell hyperplasia, increased bronchial smooth muscle tone, eosinophil survival, and class switching of immunoglobulins to the epsilon heavy chain (32). Additionally, a recent murine study has pointed to $IL\text{-}13^+$ $CD4^+$ T-cells as being critical to development of airway hyper-reactivity (AHR) in young mice whereas these cells were dispensable to induction AHR in adult mice (33). This mimics what is seen in humans where early onset of asthma is more associate with $T_{h2}$ inflammation and adult onset is associated with non-$T_{h2}$ inflammation (32). Importantly, many new biologic therapeutic options target these specific cytokines, their receptors, or their end products. These include monoclonal antibodies that target IL-5 signaling (mepolizumab and benralizumab), IgE (omalizumab), and IL-4/13 signaling (dupilumab and lebrikizumab) (34). Notably, these biologics are only effective in individuals experiencing signs of $T_{h2}$ inflammation: increased peripheral eosinophils, increased serum IgE, or increased serum periostin (a protein only expressed

in epithelium exposed to IL-13) (34). All of these therapies are reserved for individuals who fail first line treatment with inhaled corticosteroids (ICS) and currently, only omalizumab is approved for individuals under the age of 12 (35). Together, development of asthma during childhood leads to predominately $T_{h2}$ mediated airway inflammation which offers a therapeutic target for those with advanced and difficult to control disease.

C.ii. Development of Asthmatic Inflammation

Asthmatic symptoms are generally identified around age three to five, but the inflammation associated with diseases is hypothesized to begin years earlier. A previous study from the Finn-Perkins lab demonstrated that some individuals are predisposed at birth to lymphoproliferation and $T_{h2}$ cytokine responses in response to allergen stimulation. This predisposition was associated with risk factors for asthma such as maternal allergy, *in utero* smoke exposure, and African American race (36, 37). This suggests that there may be immunologic differences at birth that alter risk for development of asthma later in childhood.

Hematopoiesis, the generation of cellular components of blood, begins late the first trimester (38). The maternal and fetal immune system during pregnancy is highly biased towards $T_{h2}$ and immune-tolerant responses to foreign antigens (39-41). Additionally, throughout the majority of gestation, hematopoiesis mainly generates lymphoid and erythroid lineages. Towards term gestation there is a shift towards generating cells in the myeloid lineages (e.g. neutrophils) (42). This phenomenon is evident clinically, as pre-term (≤36 weeks gestation) infants have higher lymphocyte to neutrophil ratios compared to those born near term (≥36 weeks gestation) (43). The myeloid compartment produces several cytokines and co-stimulatory molecules that shift $CD4^+$ T-cells towards $T_{h1}$

response, including IL-18 whose receptor carries SNPs associated with asthma (10, 44). The lack of this switch towards the generation of $T_{h1}$ responses is associated with susceptibility to infections, such as lower respiratory illnesses (45). Moreover, differential methylation of promoter region of *RUNX1*, a critical regulator of myelopoeisis, in cord blood mononuclear cells (CBMCs) is associated with asthma in children (46).

Protective myeloid expansion has been reported later in childhood. A seminal study examining two genetically similar populations identified shifts towards increased neutrophil expansion and eosinophil contraction in the population with reduced rates of asthma (26). Concordantly, several meta-analyses of methylation in peripheral blood from children who develop asthma identified differential methylation in many eosinophil related genes (46, 47). Interestingly, several studies suggest that changes in myelopoeisis are caused by environmental exposures, mainly microbes (26, 48).

D. <u>The Role of the Microbiome in Childhood Asthma</u>

Microbial exposure has long been hypothesized to play an important role in many inflammatory diseases but the impact of the absence of microbial exposure in modern environments was first hypothesized by David Strachan and coined the “hygiene hypothesis”. The idea of the “hygiene hypothesis” has now taken on many different forms that generally resemble “the lack of microbe(s) x drives some immunological changes associated with disease y” (49). The diseases that this hypothesis is generally applied to show similar epidemiologic trends to asthma and include other allergic disease (e.g. atopic dermatitis, food allergy), autoimmune diseases (e.g. type I diabetes, multiple sclerosis), and certain cancers (e.g. acute myelogenous leukemia) (50). With regards to

childhood asthma, the evidence of microbes playing a critical role in disease development has become increasingly compelling.

Overcrowding, unhygienic conditions, and larger family size are postulated to be associated with a lower prevalence of allergy and asthma. Postulates have been raised that exposure to microorganisms and/or their components during early life may provide protection against asthma and allergic disease (51, 52). A low prevalence of allergies and asthma in the children of farmers suggest that environmental exposure to a diverse array of microbiota may provide a protective role against allergies and asthma (26, 52, 53). In contrast, exposure to pathogenic species do not prevent, but rather are highly associated with exacerbations in asthmatics, such as acute upper respiratory infections with respiratory viruses (e.g., RSV and HRV-C) and *Chlamydia pneumonia* (54, 55). The stimulation of the immune system by exposure to different microbes via pattern recognition receptors may mediate skewing toward specific T-cell phenotypes associated with health and/or disease (26, 51). Additionally, the composition of colonizing microbes is highly influenced by behaviors that have associations with asthma development, such as cesarean section, breast feeding, and diet (56).

An individual's microbiota varies across body sites in an anatomic specific manner (57). Prior studies of human microbiota have focused on feces as means to assess the gut microbiome. Notably, the composition of the gut microbiota also varies in a time-dependent manner, such that microbiome changes drastically over the first few years of life before stabilizing and becoming adult-like in composition. The initial composition of the gut microbiome is influenced by mode of birth (e.g., cesarean section vs. vaginal delivery), whereas early life exposure to antibiotics and breastfeeding alter the selection

of which microbes can colonize (56). More than 1000 species-level phylotypes can be found in the gut microbiota and only 400 bacterial species are culturable outside the gastrointestinal (GI) environment. In the GI tract, the majority of the bacterial species are anaerobes and most common genera are Bacteroides, Bifidobacterium, Lactobacillus, Eubacterium, Fusobacterium, and Clostridium (58). Breastfeeding is highly associated with larger abundances of the bacteria belonging to the genera Bifidobacterium and Lactobacillus, considered probiotic bacteria. Low abundance of Lactobacillus has also been implicated in the development of allergies. Populations that have low abundance of Bifidobacterium at early ages have a higher risk of developing asthma (59). This may explain the preventative effect of breastfeeding against the development of allergies and asthma in some instances (60). Additionally, these bacteria may allow for short-chain fatty acid producing bacteria to grow as well, which can stimulate T-regulatory cell development, known to suppress immune responses (61, 62).

In addition to the microbiota found in the gut, some of the first studies that provided strong evidence were those examining environmental exposures to microbes. The seminal studies, Prevention of Allergy Risk factors for Sensitization In children related to Farming and Anthroposophic Lifestyle (PARSIFAL) and A Multidisciplinary Study to Identify the Genetic and Environmental Causes of Asthma in the European Community (GABRIELA), identified that increasing diversity of microbes in mattress dust are associated with reduced risk of asthma and allergy. Notably there is geographic variation in this diversity, such that those living on farms are exposed to much higher levels of microbial diversity, which coincides with the prevalence of asthma and allergy being much lower in rural settings (63). These findings have now been replicated and expanded upon

with specific phylogenetic diversity (Gammaproteobacteria), bacterial components (endotoxin), and species derived from animals (*Acinetobacter iwoffi*) being recognized as protective environmental exposures (53, 64-66).

The mechanism behind this protection appears to be directly related to endotoxin exposure. Previous work by the Finn-Perkins lab demonstrated increased endotoxin exposure is associated with reduced $T_{h2}$ and proliferative responses of peripheral blood mononuclear cells (PBMCs) in children (51). In murine models of asthma, the Lipid A component of endotoxin could be used to decrease markers of $T_{h2}$ inflammation, including bronchoaveolar lavage (BAL) eosinophils, serum total IgE, and BAL IL-13 (67). This has now been replicated using dust from houses of Amish (low prevalence of asthma) and Hutterites (high prevalence of asthma) where the levels of endotoxin significantly differ. Dust from Amish houses when co-administered with ovalbumin in murine models of asthma reduced AHR and was dependent on the presence of the adaptor proteins for TLR4 signaling, MYD88 and TRIF (26).

Along with effect of environmental microbiota, endogenous microflora produce endotoxin. Notably, not all endotoxin is equivalent in structure and can vary in poly-saccharide structure (O-antigen), phosphorylation of sugars, or number/ structure of fatty acids (68). In population studies of three geographically similar locations that vary in autoimmune and atopy prevalence believed to be due to different levels of westernization, differences were found in the predominant source of endotoxin in children's fecal microbiomes. In high risk areas, the predominant source of endotoxin was *Bacteroides* species. *Bacteroides* species were shown to have a different endotoxin structure (lacking

phosphate group and altered fatty acid side chains) that did not induce robust cytokine production, endotoxin tolerance, or protect form autoimmune disease (69).

Interestingly, $CD4^+$ T-cells do express TLR4, the pattern recognition receptor for endotoxin, but endotoxin exposure does not alter $T_{h1}$ or $T_{h2}$ cytokine secretion in $CD4^+$ T-cells although it may alter other cytokines such IL-17 and adhesion related molecules (70, 71). This suggest an intermediary between endotoxin exposure and changes in $T_{h2}$ driven inflammation. In comparing peripheral blood granulocyte populations, Amish children had higher percentages of neutrophils whereas Hutterites had higher percentages of eosinophils, suggesting that differences in myelopoiesis may be associated with differential risk of asthma (26). Myeloid cells secrete IL-12 and IL-18, two critical cytokines necessary for the induction of $T_{h1}$ responses (72, 73). IL-12 and IL-18 act synergistically to induce IFN-γ with IL-12 responsible for the induction of IFN-γ mRNA and IL-18 responsible stabilization of transcripts and proliferation through MAPK and NFκB signaling (74). Together this suggest a strong role for microbial regulation on human immunity that can suppress or accentuate risk for asthma in childhood.

## E. Hypothesis and Aims of Study

The current literature suggests that microbial exposure likely influences risk of asthma through altered immunity. Most of the literature has focused on microbial exposures in the early postnatal period, yet it is known that some individuals have differential cytokine and proliferative responses to allergens at birth. We hypothesize that differences in microbial exposure and immunity at birth are predictive of asthma in childhood. The aims of this study are to investigate pre/peri-natal microbial exposures

associated with cytokine responses known allergens, identify novel differences in immunity at birth that are associated with risk, and determine which microbial exposures are associated with differences in immunity at birth.

# CHAPTER 2

# MATERIALS AND METHODS

* This chapter's text, tables, and figures contain portions of work that is reprinted from:

**Turturice BA**, Ranjan R, Nguyen B, Hughes LM, Andropolis KE, Gold DR, Litonjua AA, Oken E, Perkins DL, Finn PW. Perinatal bacterial exposure contributes to IL-13 aeroallergen response. *Am J Respir Cell Mol Biol* 2017 Oct;57(4):419-427. doi: 10.1165/rcmb.2017-0027OC.

And

**Turturice BA**, Gold DR, Litonjua AA, Oken E, Rifas-Shiman S, Perkins DL, Finn PW. Lower perinatal exposure to Proteobacteria is an independent predictor of early childhood wheezing. *J Allergy Clin Immunol.* 2019 Jan;143(1):419-421.e5. doi: 10.1016/j.jaci.2018.06.051. Epub 2018 Sep 8.

And

**Turturice BA**, Theorell J, Koenig MD, Tussing-Humphreys L, Gold DR, Litonjua AA, Oken E, Rifas-Shiman SL, Perkins DL, Finn PW. Perinatal granulopoiesis and risk of pediatric asthma. Elife. 2021 Feb 10;10:e63745. doi: 10.7554/eLife.63745. PMID: 33565964; PMCID: PMC7889076.

## A. Study Populations and Subject Recruitment

### A.i. Project VIVA

The current study was approved by the University of Illinois at Chicago IRB (#2016–0326). Volunteers were recruited from the eastern Massachusetts general population. The exclusion criteria were multiple gestation, inability to answer questions in English, gestational age ≥ 22 weeks at recruitment, and plans to move away before delivery. The cohort profile was previously described by Oken and colleagues (75).

### A.ii. University of Illinois Hospital Cohort

The current study was approved by the University of Illinois at Chicago IRB (#2015–0353). Volunteers were recruited from the University of Illinois Hospital and Health Services System (UIHHSS).

Inclusion criteria is as follows: obtaining prenatal care from UIC; planning to deliver at UIHHSS; 17-45 years of age; parity between 0-3; spontaneous/natural conception; singleton pregnancy; 26-33 weeks gestation; and willing to: a) fast at least 1 ½ hours at the baseline visit, b) ingest a naturally occurring stable iron isotope, c) refrain from iron and prenatal supplements for 1 day (24 hours) before the baseline visit, d) attend two research visits at UIC, e) undergo 3 venipunctures, f) allow biopsy of the placenta after delivery, and g) allow cord or placental blood collection. Able to read and write English, have access to a phone to inform the research team of admittance to UIC Labor and Delivery, live in Chicago or surrounding suburbs, be obese (pre-pregnancy BMI ≥ 30 kg/m2 - 50 kg/m2) or lean (pre-pregnancy BMI: 18.5 – 24.9 kg/m2) ; gestational diabetes is acceptable, but not necessary for inclusion. Lastly, over the course of the pregnancy, women must meet the minimum weight gain recommendations from the Institute of Medicine based on their pre-pregnancy BMI (normal weight: 25 lbs; obese weight: 11 lbs), and starting at 23/24 weeks of gestation normal weight women must be gaining at least 0.8 lbs weekly and obese weight women must be gaining at least 0.4 lbs weekly.

Exclusion criteria is as follows: Live birth or another pregnancy (including ectopic and molar pregnancies) in the previous 12 months; autoimmune disorder; current or previous premature rupture of membranes or chorioamnionitis; previous premature birth; current bacterial or viral infection; receiving steroid or anti-inflammatory treatment; previous bariatric surgery; malabsorptive condition; current hyperemesis; hematologic

disorder (i.e., sickle cell disease, sickle cell trait or hemochromatosis); or preeclampsia since these factors may influence the biological markers.

B. Specimen Collection, Cellular Isolation, and Storage

Umbilical cord blood was obtained by venipuncture shortly after time of delivery. Blood (approximately 5 mL per tube) was drawn into Red Top Serum Plus and Green Top Sodium Heparin 95 USP Units Blood Collection Tubes (BD Vacutainer). Red Tops were allowed to stand upright at room temperature for 30 min prior centrifugation at 1500x g for 10 minutes at room temperature. Supernatants (serum) were collected, aliquoted, and stored at -80C until further processing. Heparinized blood obtained in Green Tops was diluted 1:1 in 1x Phosphate Buffered Saline, pH 7.4 (PBS) and overlaid onto Ficoll-Paque™ Plus (GE Healthcare) density gradients. Density gradients were centrifuged at 400x g for 30 minutes at room temperature without brake. Upper phase (diluted plasma) was drawn off, aliquoted, and stored at -80C. Buffy coats (CBMCs) were drawn off, washed twice with 10 mL of 1x PBS, aliquoted and stored in 500 mL of RNAlater (Qiagen) or 1 mL of 10% DMSO in Fetal Bovine Serum (FBS) at $5x10^6$ per mL. Viability and number of cells isolated was determined by diluting cellular suspensions 1:1 with Trypan Blue Solution 0.4% (w/v) in PBS (Corning) and counting live/dead cells >7 µM using a TC20™ Automated Cell Counter (Biorad). The average viability of isolated cells was >90%. Time from delivery to storage was recorded for every sample. Maternal Feces were collected during the third trimester visit using fecal isolation kits. Feces was aliquoted and stored -80C.

C. DNA Isolation

C.i. Circulating DNA

Experiments related to Chapter 3: DNA was extracted from 1 mL of cord blood serum using QIAmp DNA Blood Midi Kit (Qiagen). All extractions were carried out under sterile conditions. Each extraction was recorded for kit lot number and time of extraction to test for any effect on our analysis. No effect was observed for lot number or date of extraction. To control for any possibility of bacterial DNA contamination introduced in our procedure, we used 1 mL of molecular-grade water through our extraction protocol as a negative control. Our negative control showed both a minimum 5-cycle time difference assessed by quantitative PCR from any sample and did not yield a 16S rRNA amplicon library. DNA concentration and base-pair size were quantified from all samples using Qubit dsDNA High Sensitivity Dye (Invitrogen) and on 2100 Bioanalyzer using High Sensitivity DNA kit (Agilent).

Experiments related to Chapter 5: To quantitate the abundance of 16s rRNA fragments in serum, we used a quantitative approach previously described (76). Prior to extraction of DNA synthetic spike-in 16s rRNA amplicons were added at $2.5x10^4$, $2.5x10^3$, $2.5x10^2$, $2.5x10^1$ copies per mL for EC5501, EC5001, CA5501, and BV5501 amplicons respectively. DNA was extracted from 0.4 mL of cord blood serum or molecular grade water (Negative Control) using MagMAX™ Cell-Free DNA Isolation Kit (Applied Biosystems). All extractions were carried out under sterile conditions. Each extraction was recorded for lot number on the kit and time of extraction and tested for any effect in our analysis. DNA concentration and base-pair size were quantified from all samples

using Qubit dsDNA High Sensitivity Dye (Invitrogen) and on 2100 Bioanalyzer using High Sensitivity DNA kit (Agilent).

C.ii. Fecal Microbial DNA

Approximately 100 mg of stool was transferred to an Eppendorf safe lock tube and processed with a PowerSoil DNA isolation kit (MO BIO Laboratories) using the method as described previously (77). The quality and quantity of the DNA was accessed using a spectrophotometer (NanoPhotometer Pearl, Denville Scientific, Inc), agarose gel electrophoresis, and fluorometer (Qubit® dsDNA High Sensitivity and dsDNA Broad Range assay, Life Technologies Corporation).

D. Metagenomic DNA Libraries

D.i. 16s rRNA Amplicon Libraries

For experiments in Chapter 3: V1-V3 regions of the 16s rRNA gene were amplified using the NEXTflex 16s V1-V3 Amplicon-Seq Kit (BIOO Scientific) with an initial 10-cycle step of 16s rRNA V1-V3 amplification followed by a 20-cycle step for amplification and barcoding. Libraries were pooled at equal molar concentrations and size selected for amplicons between 550 and 700 base pairs (bp) using Pippin Prep 1.5% Agarose Dye-Free Pippin Gel Cassette (Sage Science) to remove any nonspecific amplicons. The quality and quantity of all the DNA libraries were analyzed with an Agilent DNA 1000 Kit on the 2100 Bioanalyzer Instrument and Qubit.

For experiments in Chapter 5: V4 region of the 16s rRNA gene was amplified using a previously described two-step PCR protocol and previously designed primers 515F-806R and their barcoded counterparts (78, 79). Libraries were pooled at equal molar

concentrations and size selected for amplicons between 350 and 450 base pairs (bp) using Pippin Prep 2.0% Agarose Dye-Free Pippin Gel Cassette (Sage Science) to remove any nonspecific amplicons. The quality and quantity of all the DNA libraries were analyzed with an Agilent DNA 1000 Kit on the 2100 Bioanalyzer Instrument and Qubit.

D.ii. Shotgun Metagenomic Libraries

500 ng metagenomic DNA was mechanically sheared to average 450 bp fragments using a Covaris S220 instrument (Covaris). Fragmented metagenomic DNA was end-repaired and 3′-adenylated, ligated with Illumina adapters, and PCR enriched with Illumina sequencing indexes (barcodes) using the NEBNext Ultra DNA library prep kit for Illumina (New England BioLabs Inc). The quality and quantity of all the DNA libraries were analyzed with a DNA 1000 Kit on the 2100 Bioanalyzer Instrument (Agilent) and dsDNA Broad Range Kit and Qubit (Invitrogen).

E. RNA Isolation and mRNA Libraries

Total RNA was extracted from CBMCs using RNeasy kits (QIAGEN) following manufacturers protocol except for switching 70% ethanol for 100% ethanol. The quality and quantity of all the DNA libraries were analyzed with a RNA 6000 Nano Kit on the 2100 Bioanalyzer Instrument (Agilent) and ssRNA High Sensitivity Kit and Qubit (Invitrogen). RNA quality for samples was >8.5 RIN and showed two equally prominent peaks corresponding to the 18s and 28s rRNA. 500 ng of RNA was used for mRNA enrichment, fragmentation, conversion into cDNA, and construction into barcoded libraries using the TruSeq Stranded mRNA Library Prep Kit (Illumina).

F. High throughput Sequencing

F.i. 16s rRNA Amplicon Sequencing

For the experiments in Chapter 3: the pooled libraries were further diluted to 15 pM and 5% PhiX was added to increase the diversity, and sequenced for a paired-end 301 read length. The DNA libraries were sequenced on Illumina's MiSeq platform using MiSeq Reagent v3-600 cycle kit, following manufacturer's protocol.

For the experiments in Chapter 5: the pooled libraries were further diluted to 5 pM and 10% PhiX was added to increase the diversity, and sequenced for a paired-end 251 read length. The DNA libraries were sequenced on Illumina's MiSeq platform using MiSeq Reagent v2-500 cycle kit. Previously described read 1, indexing, and sequencing primers were diluted into wells 12, 13, and 14 of cartridge (80).

F.ii. Shotgun Metagenomic Sequencing

DNA libraries were pooled in equimolar concentration and were sequenced following manufacturer’s protocol by multiplexing on Illumina MiSeq using the v3-600 kit for 301 paired-end read length and included an additional 12 cycles for the index.

F.iii. mRNA Library Sequencing

The pooled libraries were further diluted and sequenced for a paired-end 151 read length. The DNA libraries were sequenced on Illumina's HiSeq X Ten platform using HiSeq Reagent v2.5 kit, following manufacturer's protocol.

G. Metagenomic Annotation

G.i. Operational Taxonomic Annotation (OTU)

For OTU annotation, we used the *QIIME* pipeline (81). Sequences were aligned to Greengenes Core reference alignment (82-84). Chimeric sequences were removed by UCHIME, and OTUs were picked by USEARCH (85, 86). Taxonomy was assigned using the Greengenes database, and a phylogenetic tree was constructed using FastTree 2.1.3 for the determination of UniFrac distances (87, 88).

G.ii. Amplicon Nucleotide Variant (ANV)

For ANV quality trimming and filtering of reads, chimera removal, and annotation we used the *dada2* pipeline (89). ANV sequences were annotated against the curated SILVA v132 data base that was modified to include the synthetic spike-in sequences (90).

G.iii. Contiguous Read Assembly and Annotation of Shotgun Metagenomics

The produced metagenomic sequencing reads were processed with our custom data analysis pipeline “WEVOTE” that is hosted on the UIC supercomputer “Extreme”. First, the sequences were quality controlled by filtering out all low-quality reads (<25 on Phred quality score), short reads (<100 bp), or any human reads. High-quality microbial reads were then assembled into longer contigs using MetaVelvet (91). For each sample, the microbial taxonomic profile was constructed using WEVOTE (92). Since WEVOTE is an ensemble classifier, we used Kraken, Clark, and BLASTN as base classifiers for WEVOTE (93-95).

H. mRNA Annotation

The sequences were quality controlled by filtering out all low-quality reads (<25 on Phred quality score), short reads (<50 bp). Transcripts were annotated using *salmon*

*v0.12.0* and *Ensembl Homo sapiens* Genome Assembly GRCh38.12 (96). Transcript counts were aggregated in gene level counts using the *tximport* package in R (97).

I. Multiplex Cytokine Assay

Serum was centrifuged at 10,000 × g for 10 minutes to clear any cellular debris. The serum was then assessed for concentrations of IFN-γ, -4, -5, and -13 using Bio-Plex Pro Human assays (Bio-Rad). Concentrations were determined using the Bio-Plex 200 system.

J. Enzyme-linked Immunosorbent Assays

PGLYRP-1 and sIL6Rα were assessed using Human PGLYRP1/PGRP-S DuoSet Elisa and Human IL-6 R alpha DuoSet Elisa (R&D Systems). Serum was diluted with 1% BSA in PBS, pH 7.2-7.4, 0.2 micron filtered at 1:100 for PGLYRP-1 and 1:300 for sIL6Rα. ELISAs were performed according to manufacturer's protocol. Optical Densities were assessed at 450 and 540 using a Spectra Max M5 (Molecular Devices).

K. Statistical Analysis

K.i. Chapter 3 Statistical Analysis

All statistical analysis was performed in R. The Mann–Whitney U test was used in all two-group comparisons unless otherwise stated. Permutational multivariate ANOVA was used to assess group differences in β diversity. Linear regression was used to correlate IL-13 and IFN-γ concentrations after stimulation with Der f1 or Bla g2 with bacterial family abundances, and P is the probability that the slope was nonzero. $P < 0.05$

was considered significant in all situations. Linear and logistic regression was used to correlate IL-13 and IFN-γ concentrations after stimulation with media, phytohemagglutinin, Der f1, or Bla g2 with the log ratio of Moraxellaceae to Proteobacteria.

Mann-Whitney U test was used in all 2-group comparisons unless otherwise stated. For testing associations between demographic data and cfbDNA composition, samples were normalized using median sum-scaling, pairwise sample distances were calculated using the weighted UNIFRAC method, ordinated using principal-component analysis, and tested for associations using the *vegan* envfit function (87, 98, 99). Maternal atopy was assessed by interview response to “Has a health professional ever told you that you have asthma, hay fever, or eczema?” Phylum-level abundances were calculated by combining normalized reads of each OTU assigned to each phylum. Proteobacteria abundance was dichotomized into low and high on the basis of normalized counts below or above the median, respectively. To determine associations between Proteobacteria abundance and early wheezing (response to annual questionnaire “Has your child had wheezing within the last 12 months?”) we assessed the time to first reported wheeze in univariate Cox-proportional hazards model and Cox-proportional hazards model adjusting for maternal atopy. To account for missing questionnaire responses, missing data were imputed using the following method: subjects not reporting at any time point were not assessed, missing data for wheezing in annual questionnaires years 1 through 7 were imputed 500 times with a larger subset of subjects from Project Viva (n = 293) using Amelia, models were estimated for each of the imputed data sets, and multiple

imputation model parameters were inferred using the method described by Rubin (100, 101).

K.ii. Chapter 4 Statistical Analysis

All statistical analysis was performed in R. Categorical outcomes were modeled using logistic regression. Continuous outcomes were modeled using linear regression. For regression models, PGLYRP-1 was log10 transformed and standardized. Subset analysis was performed by splitting the full data set by categorical variables and modeling outcomes as function of PGLYRP-1 in each subset.

To identify datasets studies used in the meta-analysis NCBI's Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo) was searched using the search "(cord blood) AND "Homo sapiens"[porgn:_txid9606]" and was limited to study types that included expression profiling by array. This search yielded 352 studies that were further examined for cell types assessed and metadata reported. 17 studies were included on the basis of reporting metadata regarding one pre/perinatal risk factor and expression data from either whole cord blood or cord blood mononuclear cells. Expression data and metadata was extracted using *geoquery* (102). If expression data was non-normalized, it was quantile normalized and log2 transformed. Six studies were excluded due to no variability in metadata (i.e. only males) or data quality, leaving 605 unique cord blood gene expression samples. To assess associations between gene expression and pre/perinatal risk factors, univariate, inverse variance weighted, random effects models were constructed for genes using the *GeneMeta* package (103). Fetal sex, maternal pre-pregnancy BMI, gestational age, and birthweight were assessed as pre/perinatal risk factors.

The z-statistic for each gene from each univariate test was oriented such that low risk was negative and high risk was positive. z-statistics were averaged and if a gene was not assessed in a univariate test the z-statistic was assumed to be 0 for that univariate test. The averaged z-statistic was used as pre-ranked list for gene set enrichment analysis for GO biologic processess using *GSEA* (104). Gene sets with p-values <0.05 and FWER<0.1 were considered significant and were used in leading edge analysis. Leading edge networks for genes and gene sets were constructed using Kullback-Leibler divergence. Network metrics closeness centrality and neighborhood connectivity were determined using *igraph* (105).

For analysis of mRNAseq expression, gene expression was modeled as function of the number of pre/perinatal risk factors using *DESeq2 (98)*. Four models were constructed to determine gene associated with risk. Replication between meta-analysis and mRNAseq was assessed by comparing z-statistics through spearman's correlation and non-parametric ranking method. For ranking replicating genes we multiplied the two z-statistics together. Genes with non-parametric rank greater than two were said to be replicating. GO term enrichment of high and low risk gene sets was performed using DAVID (106). P-value < 0.05 and Benjamini-Hochberg adjusted p-value <0.1 was considered statistically significant for GO term enrichment.

Bacterial taxa counts were normalized using the RLog transformation, DESeq2 (98). Taxa were tested as independent predictors for PGLYRP-1 concentration and eigenvalue of gene signature using generalized linear models. P-value <0.05 and Group Benjamini-Hochberg < 0.1 were considered statistically significant where groups were defined by a taxa's Genus (107).

K.iii. Chapter 5 Statistical Analysis

All statistical analysis was performed in R. To quantitate bacterial taxa, samples were first filtered on the detection of the spikes EC5001, EC5501, and CA5501. Next, spike counts were used to build model predicting known spike-in copy numbers. A mixed-effects model was used, where $Y$ is the known spike-in log10 copy number that is modeled by the fixed effects $\beta_{obs}$ and $\beta_{obs:sample}$--which are overall effect estimate of log10 observed count and log10 observed count interaction with each individual sample--and $Z_{sample}$, a random sample specific intercept. Samples with an absolute $\beta_{obs:sample}$ greater than 0.20 were removed from further analysis as these samples likely have under gone non-linear amplification. Removal of batch effect was performed using *limma* (108). To test associations between composition and demographic covariates, Euclidean distances were calculated between samples and then sub-grouped by demographic data to compare with-in and between group distances. Kruskal-Wallis and post-hoc Dunn's Test were used to test for statistical significance, $p < 0.05$ was considered significant. To test for associations with the concentration of umbilical cord blood serum PGLYRP-1, PGLYRP-1 was log10 transformed. Copy number and Phyla were test using linear models as predictors of PGLYRP-1 concentration, $p < 0.05$ was considered significant. Due to non-linear distributions of taxa at lower phylogenetic levels, spearman's correlation was used to determine relationships with PGLYRP-1, Stimulation Index, IL-13, and IFNγ. Spearman's rho and complete linkage hierarchical clustering was used to cluster taxa in their relation to immune response to stimulation and serum PGLYRP-1.

# CHAPTER 3

## INDENTIFICATION OF BACTERIAL TAXA ASSOCIATED WITH PERINATAL T-CELL $T_{H2}$ RESPONSES AND CHILDHOOD WHEEZING

* This chapter's text, tables, and figures contain portions of work that is reprinted from:

**Turturice BA**, Ranjan R, Nguyen B, Hughes LM, Andropolis KE, Gold DR, Litonjua AA, Oken E, Perkins DL, Finn PW. Perinatal bacterial exposure contributes to IL-13 aeroallergen response. *Am J Respir Cell Mol Biol* 2017 Oct;57(4):419-427. doi: 10.1165/rcmb.2017-0027OC.

And

**Turturice BA**, Gold DR, Litonjua AA, Oken E, Rifas-Shiman S, Perkins DL, Finn PW. Lower perinatal exposure to Proteobacteria is an independent predictor of early childhood wheezing. *J Allergy Clin Immunol.* 2019 Jan;143(1):419-421.e5. doi: 10.1016/j.jaci.2018.06.051. Epub 2018 Sep 8.

### A. Rationale

Atopic asthma is characterized by obstruction of air flow through the bronchial tree in response to specific aeroallergens, generally beginning early in life. Sensitization to common aeroallergens, such as those from the house dust mite (HDM) and cockroach, has been recognized as an important predictor of decreased pulmonary function and increased airway hyperresponsiveness (109, 110). Specifically, children sensitized to HDM allergens have an increased risk for respiratory infections, chronic wheeze, and the development of asthma and atopy that persist into adulthood (111-113). This sensitization is associated with higher levels of HDM allergen–specific IgE (113). An important relationship exists between altered immune responses to aeroallergens and the development of asthma and atopy in childhood.

Interestingly, many of these disordered immune responses, such as increased IgE levels, lymphoproliferation, and altered cytokine production in response to aeroallergen stimulation, are detected as early as the perinatal time period (36, 37, 45, 114, 115). Along

with altered immune responses, differential environmental exposure to microbes and colonization of mucosal sites by microbes early in life have been linked to the development of asthma (63, 116). This has led to postulations that early-life interactions between hosts and microbes are essential for development of the normal immune system, and perturbations of these interactions can lead to disordered immune responses such as those seen in asthmatics (117).

Recent examinations of mucosal sites other than the airways, such as the gut, have shown that the differential growth of specific bacteria is related to asthma development (61). Additionally, environmental factors (e.g., breastfeeding) can influence bacterial colonization of mucosal surfaces early in life (56). Airway colonization by certain bacterial pathogens, such as Streptococcus, Moraxella, and Haemophilus, in early life can enhance susceptibility to both upper and lower respiratory infections from bacteria and virus (118). Along with altered immune responses to aeroallergens, individuals who are sensitized to specific allergens (e.g., HDM allergens) also show disordered humoral recognition of bacterial antigens (119). Thus, there is a relationship between bacterial colonization and host immune recognition, and differential recognition may be involved in allergen sensitization and asthma development.

In addition to the traditional sites of bacterial colonization, such as the gut and the upper airways, there have been reports of bacteria and bacterial nucleic acids at solid tissue sites and in the blood stream. With the use of traditional techniques for bacterial detection and identification (e.g., Gram's staining and culture-independent techniques), several studies have detected both bacteria and bacterial DNA within placental tissue, and correlated the detection of nucleic acids derived from different bacterial species with

disease states, including urinary tract infections, preeclampsia, and preterm birth (120-123). In murine models, specific bacteria are capable of being enriched within the placental tissue (124). Along with potential fetal exposure to bacteria, many bacterial metabolites and cellular products appear capable of traversing the placental barrier (125, 126). Thus, exposure to microbial antigens and products may occur before birth and may influence the development of the immune system.

Along with localization within solid tissues, bacterial DNA has been reported to circulate in the blood stream of healthy individuals (127). The detection of bacterial DNA in blood does not appear to be limited only to adults, and can also be isolated from umbilical cord blood (128). The amount and composition of bacterial DNA in blood have been linked to several disease states in both adults and newborns (129, 130); however, the relation of the bacterial DNA in circulation to development of the immune system and asthma outcomes remains largely unexplored. The hypothesis for this study was that specific patterns of bacterial exposure are associated with umbilical cord blood mononuclear cell (CBMC) lymphoproliferation and cytokine production in response to allergen stimulation, along with specific features of asthma.

B. Bacterial 16S rRNA fragments circulate in perinatal umbilical cord blood and are derived from multiple taxonomic families

Consistent with previous reports, we isolated 10–100 ng of fragmented DNA from human serum. We were able to quantitate the 16S rRNA fragments (base pairs 926–1062) from all samples using universal bacterial 16S rRNA primers by qPCR assay, and all samples amplified with a minimum five-cycle difference compared with negative

controls (131, 132). Along with prior reports of bacterial DNA in adult peripheral blood, our results indicate that human umbilical cord blood contains bacterial DNA (127).

To assess the bacterial source of these fragments, we amplified the V1–V3 region (~500–700 bp) of the bacterial 16S rRNA from the isolated DNA. We successfully amplified DNA libraries of 16S rRNA amplicons in 27 out of 38 samples (71%). We speculate that we were not able to generate libraries in all of our samples due to the larger amplicon size compared with our qPCR reactions (~600 versus 140 bp) and because we started with DNA that contained fragments smaller than 600 bp. From each library we generated an average of ~560,000 paired-end reads per sample. These reads were assigned to 846 unique OTUs. The overwhelming majority (>99%) of reads assigned to OTUs were derived from four core phyla that were detected in a minimum of one read in ≥ 60% of all samples: Actinobacteria, Bacteroidetes, Firmicutes, and Proteobacteria (**Figure 1A**). Notably, we observed that a large portion of the bacterial families were sparsely distributed (minimum of one read in < 30% of all samples) and that only 13 families were identified as the core bacterial families in our samples (**Figure 1B**). Taken together, these results indicate that bacterial DNA is present in the circulation during the perinatal time period and derived from multiple bacterial sources.

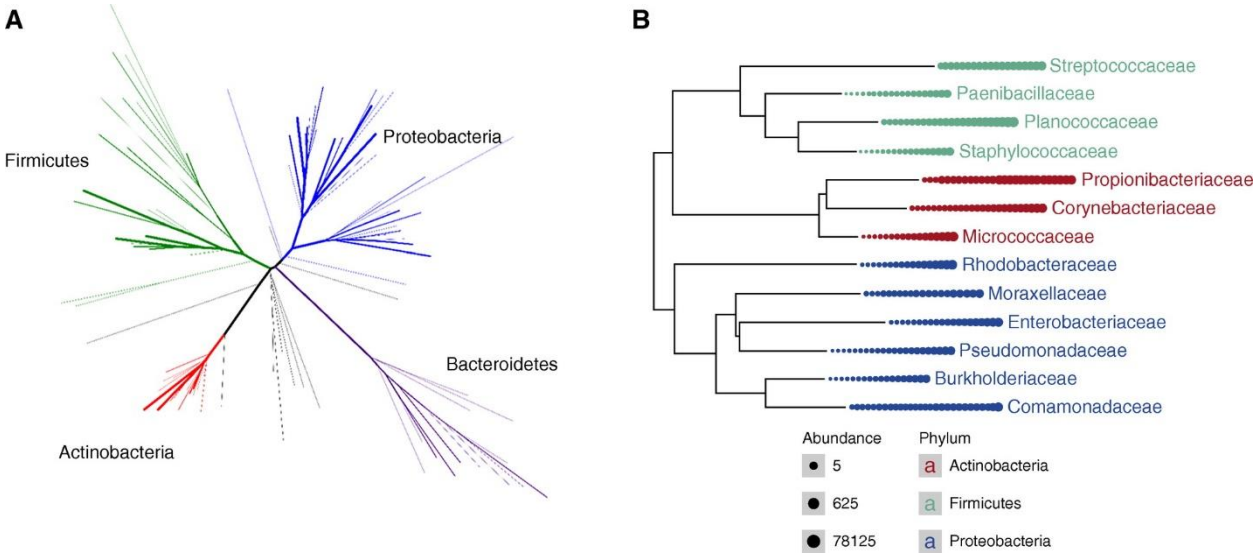


**Figure 1. Phylogenetics of bacterial DNA in umbilical cord blood. A)** Unrooted phylogenetic tree displaying all operational taxonomic units (OTUs). The line weighting is proportional to the percentage of samples in which one read was mapped to an OTU. Four core phylum (Firmicutes, Proteobacteria, Actinobacteria, and Bacteroidetes) are displayed by color (green, blue, red, and purple, respectively). **B)** Dendrogram of core bacterial families, colored by phylum, with size proportional to the normalized counts in each sample.

C. The IL-13 cytokine response to common aeroallergens Is related to the taxonomic structure and diversity of bacterial DNA

We previously found that peripheral and CBMCs exposed to aeroallergens (cockroach [Bla g2] and HDM [Der f1]) produce differential lymphoproliferative and $T_{h1}/T_{h2}$ cytokine responses, and that these responses are related to risk factors for the development of asthma (36, 37). Here, we assessed whether these allergen-specific responses correspond to unique signatures of bacterial DNA in the circulation. First, to categorize the samples, each sample with a known lymphoproliferative response was clustered by stimulation index to Bla g2 and Der f1. Similarly, to categorize the cytokine response, samples were grouped by the type of cytokine produced in response to Bla g2 and Der f1, and whether they produced detectable IL-13 or IFN-γ.

To identify differences in taxonomic structure, sample-wise distances were calculated using weighted unique fraction metric (UniFrac) distances, and group medoids were compared. Only the presence (responder [R]) or absence (nonresponder [NR]) of IL-13 production after allergen stimulation identified groups with significantly different medoids (**Figure 2A**). There were no significant differences between groups of low- and high-stimulation index or groups of IFN-γ responders and nonresponders, although there was a trend ($P < 0.1$) of different medoids categorizing IFN-γ. Thus, our data indicate that there is a relationship between allergen-specific IL-13 responses and the composition of bacterial DNA. Based on this observation, we focused our remaining analysis on the specific relationship between the IL-13 response and bacterial DNA.

We next examined the α diversity of each sample by calculating Fisher's α parameter for each sample. This parameter was selected because the majority of samples displayed logarithmic distributions on rank-abundance curves, which fits Fisher's model. When we

compared IL-13 nonresponders with responders, we observed a significant increase in Fisher's α parameter in responders (**Figure 2B**). By assessing sample-wise UniFrac distances, we showed that changes in family abundance (weighted UniFrac) were present when IL-13 nonresponders were compared with responders. Additionally, we found that IL-13 responders had an increased evenness of OTU abundances, and that nonresponders had significantly increased representation by fewer OTUs. These results indicate that $T_{h2}$ responses to common aeroallergens are related to the source of bacterial DNA, and not simply to the presence of microbial DNA.

To identify taxonomic differences in the composition of bacterial DNA between nonresponders and responders, we compared each group's normalized counts of the four core phylum. Only the bacterial abundance of Proteobacteria was increased in nonresponders (**Figures 2C, D**). The additional core bacterial phylum, Actinobacteria, Bacteroidetes, and Firmicutes, showed no significant differences between groups (**Figure 2D**).

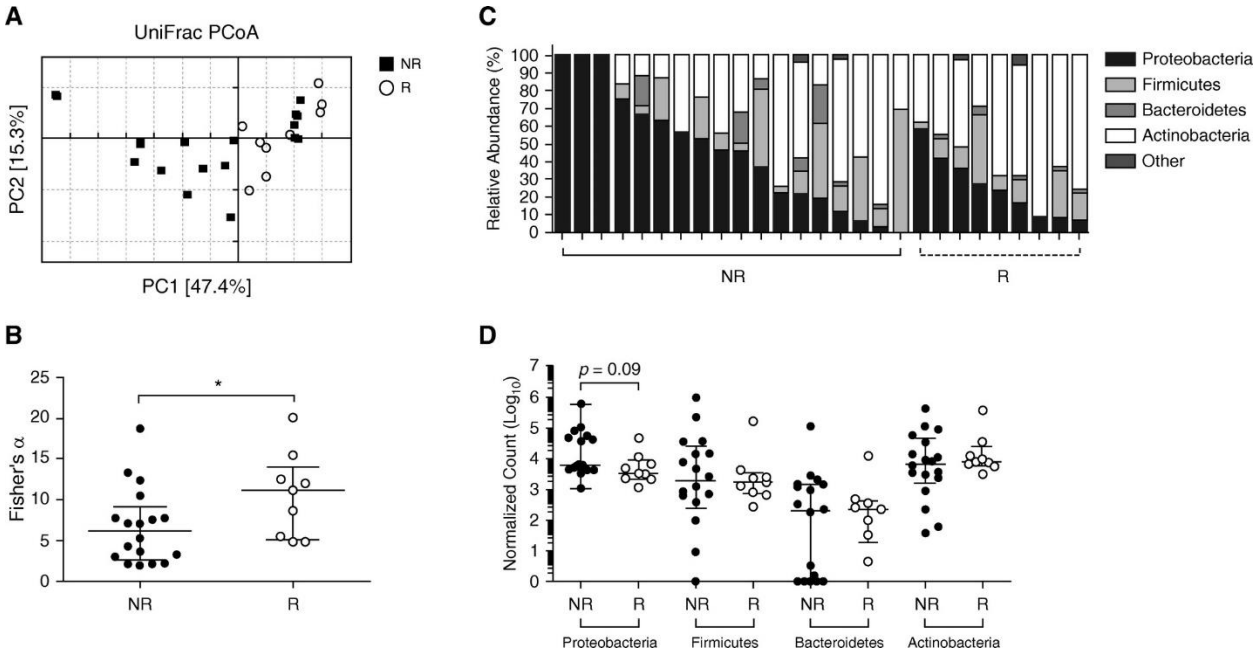


**Figure 2. Bacterial DNA diversity and composition differ in the IL-13 response to aeroallergens. A)** Principal component (PC) analysis plot (first and second components) of weighted UniFrac distances between IL-13 nonresponders (NR, solid square) and responders (R, open circle) as defined by detectable IL-13 after Bla g2 or Der f1 exposure. Significance ($P < 0.05$) was determined by permutational multivariate ANOVA. **B)** α-Diversity measurements between IL-13 nonresponders and responders, displayed by Fisher's α of individual samples. **C)** Relative abundances of core phyla and sparse phyla (other) between IL-13 nonresponders (NR) and responders (R). **D)** Normalized counts of Proteobacteria between IL-13 nonresponders (NR) and responders (R). Groups are shown as median with interquartile range. Counts were log10 transformed with a pseudocount of one added for visualization of samples with zero counts. *$P < 0.05$, Mann–Whitney U test. PCoA, principal coordinates analysis.

D. The common aeroallergen IL-13 response Is related to specific bacterial family abundances

Interestingly, when we examined the abundances of core bacterial families, we found three families that had significantly increased abundances in responders: Micrococcaceae, Moraxellaceae, and Rhodobacteraceae, the latter two of which are of the phylum Proteobacteria (**Figure 3A**). Although there was an overall increase in the abundance of Proteobacteria-derived DNA in samples with no response to aeroallergens, samples with $T_{h2}$-skewed responses to aeroallergens had increased DNA derived from specific families of Proteobacteria, notably the human pathogens in the family Moraxellaceae. Additionally, there was a trend ($P = 0.08$) toward increased abundance of sequences derived from Corynebacteriaceae in responders (**Figure 3A**). All other core bacterial families displayed no change in abundance between groups.

E. Circulating bacterial DNA in the perinatal time period correlates with IL-13 production in response to common aeroallergens

To assess whether differentially abundant families of bacterial DNA in the perinatal circulation are predictive of the quantity of IL-13, we correlated core bacterial families that had increased abundances in responders with the amount of IL-13 that was produced after Der f1 or Bla g2 stimulation. For two of the increased bacterial families in the responders, Corynebacteriaceae and Moraxellaceae, the DNA abundance significantly correlated with increased IL-13 production after Der f1 stimulation (**Figure 3B**). Interestingly, both of these bacterial families (Corynebacteriaceae and Moraxellaceae) are known to be colonizers and pathogens of the skin, oropharynx, and upper and lower respiratory tracts (118). There were no significant correlations between any of the four bacterial families tested with Bla g2–induced IL-13 production.

There was also a trend (P = 0.08) toward increasing IL-13 production after Der f1 stimulation in Rhodobacteraceae, and no significant correlation with Micrococcaceae (**Figure 3B**). Additionally, there were no significant correlations with any of these bacterial families and the production of IFN-γ. This is consistent with previous reports that increasing IL-13 production in CBMCs does not alter IFN-γ production after allergen stimulation (36).

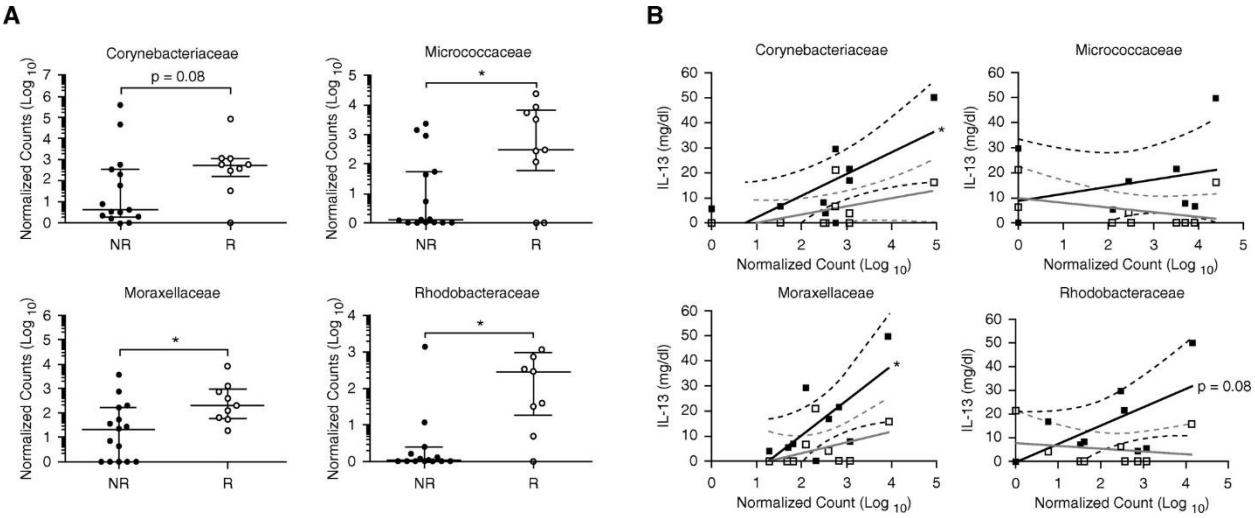


**Figure 3. Specific bacterial family abundances are increased with the IL-13 response to aeroallergens. A)** Normalized counts of bacterial families, with significant differences between IL-13 nonresponders and responders. Groups are shown as median with interquartile range. Counts were log10 transformed with a pseudocount of one added for visualization of samples with zero counts. *P < 0.05, Mann–Whitney U test. **B)** Scatter plots of normalized counts of bacterial families with corresponding cord blood mononuclear cell IL-13 production after allergen stimulation. Only samples with an IL-13 response to Der f1 (solid square) or Bla g2 (open square) are displayed. Linear regression of IL-13 responders is shown as a solid line, and 95% confidence intervals are shown as dashed lines for Der f1 (black) or Bla g2 (gray) response. *P < 0.05.

F. The Moraxellaceae to Proteobacteria ratio Is predictive of the IL-13 response to aeroallergens

To combine our findings into a predictive model of response, we compared the Moraxellaceae/Proteobacteria ratio with the quantity and presence of an IL-13 response. Moraxellaceae is a family within Proteobacteria, a diverse phylum of gram-negative bacteria, and this ratio represents the percentage of Proteobacteria that is derived from the family Moraxellaceae. The Moraxellaceae/Proteobacteria ratio was chosen to reconcile the fact that Proteobacteria was increased in IL-13 nonresponders, yet Moraxellaceae was increased in abundance in IL-13 responders and its abundance correlated positively with IL-13 production.

We found that the Moraxellaceae/Proteobacteria ratio was positively correlated with IL-13 production after Der f1 or Bla g2 exposure. Additionally, an increasing Moraxellaceae/Proteobacteria ratio was significantly associated with a greater probability of an IL-13 response after Der f1 exposure, and there was a trend ($P < 0.1$) toward greater probability after Bla g2 exposure. These associations were specific to aeroallergen exposure and the IL-13 response. We observed no significant associations when we assessed for IFN-γ response or IL-13 production after exposure to media or phytohemagglutinin (**Table 1**). We observed that there were two genera that composed the reads annotated in Moraxellaceae: Acinetobacter and Enhydrobacter. In IL-13 nonresponders and responders, 7 of 12 and 9 of 9 samples with detectable Moraxellaceae, respectively, had > 50% of the Moraxellaceae reads assigned to Acinetobacter ($P < 0.05$, Fisher's exact test).

| | *β* | CI [95%] | $R^2$ | *p* |
|---|---|---|---|---|
| **†IFN-γ** | | | | |
| Media | - 3.483 | - 9.155 to 2.189 | 0.060 | 0.217 |
| PHA | 1.5e3 | - 1038 to 4162 | 0.058 | 0.228 |
| Bla g2 | 46.91 | - 41.61 to 135.4 | 0.045 | 0.285 |
| Der f1 | 298.8 | - 199.7 to 797.3 | 0.057 | 0.228 |
| **†IL-13** | | | | |
| Media | 6.509 | - 2.797 to 15.82 | 0.077 | 0.162 |
| PHA | - 102.1 | - 1350 to 1146 | 0.001 | 0.868 |
| Bla g2 | 7.667 | 1.085 to 14.25 | 0.187 | 0.024 |
| Der f1 | 19.96 | 5.595 to 34.33 | 0.247 | 0.008 |

| | *β* | *SE* | *p* |
|---|---|---|---|
| **††IFN-γ** | | | |
| Media | 0.320 | 1.619 | 0.843 |
| PHA | 3.240 | 3.038 | 0.286 |
| Bla g2 | 0.821 | 1.894 | 0.665 |
| Der f1 | 0.116 | 1.898 | 0.951 |
| **††IL-13** | | | |
| Media | 2.362 | 1.892 | 0.212 |
| PHA | < 0.001 | > 1.0e5 | 1.000 |
| Bla g2 | 13.57 | 8.093 | 0.093 |
| Der f1 | 5.550 | 2.503 | 0.026 |

**Table 1. Association of Moraxellaceae to Proteobacteria ratio with cytokine production determined by univariate linear regression and logistic regression.** † Concentrations of cytokines treated as continuous variables and modeled using linear regression. †† Concentrations of cytokines treated as dichotomous variables and modeled using logistic regression.

G. Serum DNA microbial diversity Is associated with serum IL-13 cytokine concentrations

To assess the effect of the serum DNA microbial composition on in vivo $T_{h1}$ and $T_{h2}$ cytokine concentrations, we correlated the Moraxellaceae/Proteobacteria ratio with IL-4, -5, -13, and IFN-γ. We did not observe any significant associations between the cytokines and the Moraxellaceae/Proteobacteria ratio. To determine whether there is an association between diversity and cytokine concentrations, we correlated IL-4, -5, -13, and IFN-γ with Fisher's α parameter. Similar to our previous findings with allergen-stimulated CBMCs, IL-13 concentrations were significantly associated with Fisher's α parameter. There was also a trend (P = 0.08) toward an association between IL-4 and Fisher's α parameter. There were no significant associations between IL-5 or IFN-γ and Fisher's α parameter.

H. Lower perinatal exposure to Proteobacteria is an independent predictor of early childhood wheezing

We postulated that perinatal risk factors for asthma would be associated with early exposure to bacteria or bacterial products. To do so, we assessed circulating-free bacterial DNA (cfbDNA) in umbilical cord blood from mother-child pairs from Project Viva, a longitudinal study from birth through childhood (75). To assess cfbDNA we utilized 16s rRNA sequencing of the V1-3 region from circulating DNA isolated from umbilical cord blood serum. Of the 27 individuals in the cohort only four subjects reported doctors diagnosis of asthma, therefor we chose to assess association between cfbDNA composition, common risk factors for asthma, and asthma associated physiologies. We assessed associations between clinical features and cfbDNA composition. Of the demographcics tested, only maternal atopy was associated (P = 0.07, $R^2$ = 0.17) with the composition of cfbDNA (**Fig. 4A**). The four top abundant phyla were compared between

children of atopic and non-atopic mothers. Of these four phyla, only circulating Proteobacteria DNA was significantly different. Specifically, children of non-atopic mothers demonstrated increased Proteobacteria DNA as compared to children of atopic mothers ($P < 0.05$) (**Fig. 4B**).

Two lines of evidence suggest that abundance of Proteobacterial cfbDNA is related to risk of asthma; lower abundance is associated with offspring of atopic mothers (shown in this study) and with offspring who have increased likelihood of $T_{h2}$ responses to common aeroallergens. Children who had at least one response to annual questionnaire were assessed for Proteobacteria abundance. Proteobacteria was dichotomized by greater or less than the median and its association with wheezing within the first seven years after birth using Cox proportional hazards models (**Table 2**). Lower exposure to Proteobacteria was associated with an increased risk of wheezing in a univariate model (HR (95%CI) = 3.60 (1.70-7.64), $P < 0.001$). To assess if this association was independent of maternal atopic status, we adjusted for maternal atopy. Lower exposure to Proteobacteria was associated with an increased risk of wheezing in the multivariate model (aHR (95%CI) = 2.80 (1.46-5.36), $P < 0.01$). Children were evaluated by pulmonary function tests at mid-childhood follow up (median age 8.0 years, n = 9), and exposure to lower Proteobacteria trended towards an association with increased bronchodilator responses at mid-childhood ($P = 0.06$).

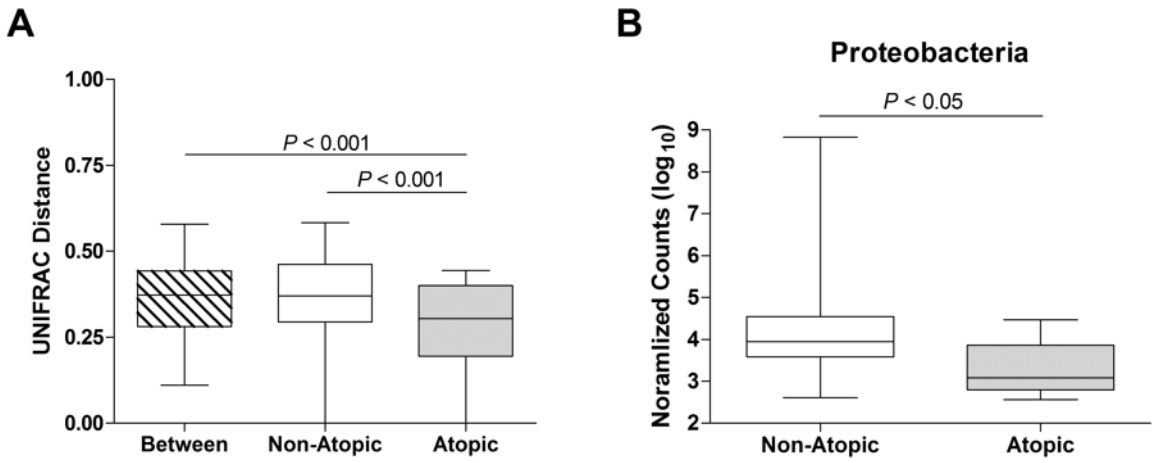


**Figure 4. Maternal atopy is associated with composition of cfbDNA and lower Proteobacteria cfbDNA. A)** unique fraction method (UNIFRAC) pairwise distances are shown for distances between children with atopic mothers and those without (between), and the within-group distances for children of atopic mothers (atopic, n = 9) and those of nonatopic mothers (nonatopic, n = 17). P < .05 was considered statistically significant, Kruskal-Wallis pre hoc test or Dunn's post hoc test. **B)** Abundance of Proteobacteria displayed as normalized counts for samples from nonatopic and atopic mothers. P < .05 was considered statistically significant, Mann-Whitney U test.

| | *β* | *SE* | *P* | 95%CI | *R*† |
|---|---|---|---|---|---|
| ***Imputed**** | | | | | |
| **Univariate** | 1.283 | 0.383 | 0.001 | (0.533, 2.032) | 0.433 |
| **Adjusted‡** | 1.030 | 0.331 | 0.002 | (0.381, 1.679) | 0.511 |
| ***No Missing Data*** ** | | | | | |
| **Univariate** | 1.466 | 0.956 | 0.125 | (-0.407, 3.339) | |
| **Adjusted‡** | 1.382 | 1.058 | 0.191 | (-0.692, 3.496) | |
| ***Imputed Wheeze***¶ | | | | | |
| **Univariate** | 0.910 | 0.496 | 0.066 | (-0.061, 1.881) | |
| **Adjusted‡** | 0.793 | 0.523 | 0.129 | (-0.230, 1.815) | |
| ***Imputed Non-Wheeze***# | | | | | |
| **Univariate** | 1.619 | 0.625 | 0.010 | (0.395, 2.843) | |
| **Adjusted‡** | 1.309 | 0.662 | 0.048 | (0.011, 2.607) | |

**Table 2. Associated risk for low Proteobacteria cfbDNA exposure and time to first parental reported wheezing event**. † R is estimated relative increase in variance due to nonresponse. ‡ Adjusted for maternal atopy. * Only subjects with complete questionnaire responses from year 1 to 7. N = 6 and N = 5 for Proteobacteria > and Proteobacteria < the median, respectively. ¶ All missing data was reported as wheezing event. # All missing data reported as non-wheezing event.

# CHAPTER 4

## DETERMINING PERINATAL IMMUNE MEDIATORS OF CHILDHOOD ASTHMA RISK

* This chapter's text, tables, and figures contain portions of work that is reprinted from:

**Turturice BA**, Theorell J, Koenig MD, Tussing-Humphreys L, Gold DR, Litonjua AA, Oken E, Rifas-Shiman SL, Perkins DL, Finn PW. Perinatal granulopoiesis and risk of pediatric asthma. Elife. 2021 Feb 10;10:e63745. doi: 10.7554/eLife.63745. PMID: 33565964; PMCID: PMC7889076.

### A. Rationale

While $T_{h2}$ driven inflammation has been recognized as important to pathogenesis of asthma in childhood, asthma is highly heterogeneous disease with regards to the physiology, inflammation, and severity (1). It has been recognized that not all individuals have signs of $T_{h2}$ inflammation (32). These individuals usually display later onset asthma, yet this does not preclude children with asthma from showing signs of non-$T_{h2}$ inflammation. Importantly, even with $T_{h2}$ type asthmatics there are further delineations between those with more eosinophilic asthma and those mediated by the effects of IL-4/13 on the bronchial epithelium (32).

With regards to childhood asthma, there is limited understanding in which pathways lead to asthma and not just $T_{h2}$ or allergic inflammation. Many of the strongest genetic associations with risk have been shown in genes that exist outside the $T_{h2}$ pathway, such *IL18R1*, *GSDMA*, *GSDMB*, and *SMAD3* (10). Additionally, a large proportion of asthma risk factors can be determined at or around the time of birth. Meta-analysis of CBMC methylation has indicated that differences exist in methylation of non-$T_{h2}$ associated genes (46). This indicates that there are likely immunologic differences at

birth between individuals who will or will not develop asthma in childhood, and that these differences largely exist outside $T_{h2}$ inflammation.

We hypothesize that assessing transcriptomic changes in CBMCs will allow for the identification of pathways, cell types, and genes that are associated with childhood asthma. A naïve approach would be to assess transcription differences between a child that does or does not develop asthma. This assumption ignores that there are many postnatal risk factors for asthma that may influence development of disease. The approach taken here focuses on identifying transcription differences that are associated with pre/perinatal risk profiles. This will allow for the elucidation of transcriptional differences that are associated with susceptibility to asthma.

B. Pre/perinatal risk factors for childhood asthma are additive in effect

To elucidate whether the simplistic view that greater number of pre/perinatal risk factors corresponds to greater risk for childhood asthma, questionnaire responses ascertaining whether a physician provided a diagnosis of asthma within the first 7 years of life were assessed in relationship to the number of risk factors in mother-child dyads from Project Viva, a longitudinal study from birth (75). Risk factors assessed were as follows: cesarean section, pre-term birth, small for gestational age, maternal smoking during pregnancy, race (black or Hispanic), maternal atopy, and maternal pre-pregnancy obesity. As expected, with increasing number of risk factors there was an increasing proportion of individuals who were diagnosed with asthma during childhood (**Figure 5A**). The odds of having asthma was linearly associated with the number of risk factors,

indicating that risk factors have an additive effect and potentially share a common mediator of risk (**Figure 5B**).

A common mediator of risk was hypothesized to be differences in the immune system and was assessed as changes in peripheral leukocyte populations at birth. Specifically, cord blood mononuclear cells (CBMCs) were assessed, as these have been thoroughly studied with regards to cytokine production and DNA methylation in relationship to later development of asthma (36, 37, 45, 46, 114). To increase power and generalizability, a meta-analysis was performed of existing CBMC transcriptomic profiles and used as a comparator with a cohort collected at University of Illinois Hospital (referred here after as UIH cohort) (**Figure 5C**). This approach was hypothesized to reveal transcriptional changes associated with low or high risk of asthma.

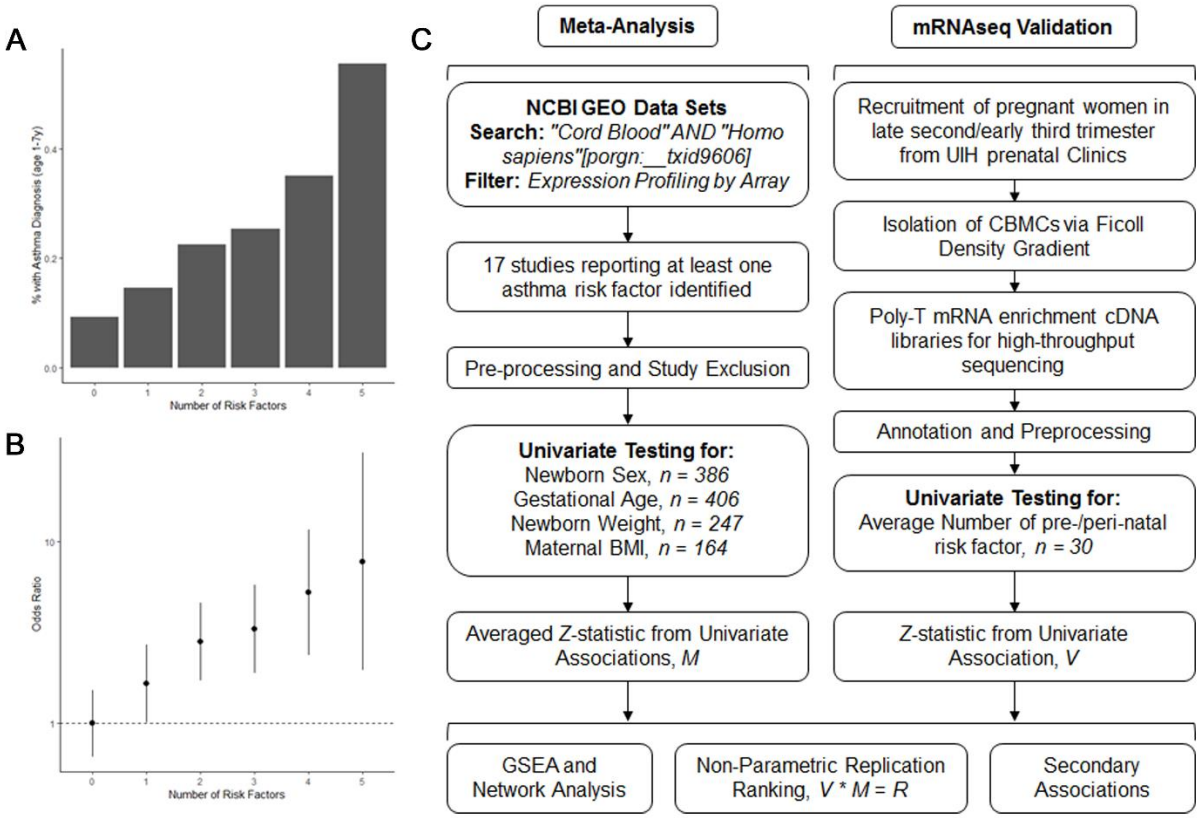


**Figure 5. Pre/Perinatal risk factors for childhood asthma produce additive effect and approach for mediator discovery. A)** Bar plot displaying the percentage of individuals reporting a physician's diagnosis of asthma within years 1-7 after birth (childhood asthma) in relation to number of pre/perinatal risk factors. **B)** Odds Ratio for childhood asthma corresponding to number of pre/perinatal risk factors. Pre/perinatal risk factors assessed were as follows: preterm, small for gestational age, race (Black or Hispanic), maternal pre-pregnancy obesity, maternal atopy, maternal smoking during pregnancy, male, and Cesarean-section mode of delivery. **C)** Overview of approach used to discover mediators of risk in cord blood mononuclear cells.

C. Meta-analysis of CBMCs identifies transcriptional changes in nuclear and innate immunity genes are associated with asthma risk profile

NCBI's Gene Expression Omnibus database was queried and 17 data sets that contained relevant metadata, including information about pre/perinatal risk factors, were identified (**Table 3**). Univariate random effects models were used to assess transcriptional changes with regards to fetal sex, gestational age, newborn weight, and maternal pre-pregnancy BMI. As expected several genes were identified to have large transcriptional changes when comparing sex and gestational age. With regards to sex-associated transcriptional changes, although there was no significant enrichment of genes located on sex chromosome, several genes located on X (*KDM5C*, *SMC1A*, *TXLNG*, and *KDM6A*) and Y (*KDM5D* and *EIF1AY*) sex chromosomes had the largest effect sizes along with the most significant differences (**Figure 6A**). In addition to expected sex-associated transcriptional changes, *HBE1,* a hemoglobin subunit associated with fetal erythropoiesis, was significantly associated with pre-term gestational ages (**Figure 6B**). Differential expression was observed in 717, 361, 90, and 89 genes when comparing fetal sex, gestational age, birthweight, and maternal pre-pregnancy BMI respectively (**Figure 7A**).

| Studies | GPL | Platform | N | Univariate | Location | Title |
|---|---|---|---|---|---|---|
| **GSE27272** | GPL6883 | Illumina HumanRef-8 v3.0 expression beadchip | 64 | Maternal PP_BMI, Gestational Age, Birth weight | Czech Republic | Comprehensive Study of Tobacco Smoke-Related Transcriptome Alterations in Maternal and Fetal Cells |
| **GSE30032** | GPL6883 | Illumina HumanRef-8 v3.0 expression beadchip | 47 | Maternal PP_BMI, Gestational Age, Birth weight | Czech Republic | Deregulation of Gene Expression induced by Environmental Tobacco Smoke Exposure in Pregnancy |
| **GSE83393** | GPL17077 | Agilent-039494 SurePrint G3 Human GE v2 8x60K Microarray 039381 (Probe Name version) | 146 | Sex | Belgium | Newborn Sex-specific Transcriptome Signatures and Gestational Exposure to Fine Particles: Findings from the ENVIRONAGE Birth Cohort |
| **GSE60403** | GPL570 | [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array | 16 | Sex, Maternal PP_BMI | USA | The obese fetal transcriptome |
| **GSE73685** | GPL6244 | [HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version] | 23 | Gestational Age | USA | Unique inflammatory transcriptome profiles at the maternal fetal interface and onset of human preterm and term birth |
| **GSE31836** | GPL6480 | Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version) | 111 | Excluded | Netherlands | Global gene expression analysis in cord blood reveals gender-specific differences in response to carcinogenic exposure in utero |
| **GSE37100** | GPL14550 | Agilent-028004 SurePrint G3 Human GE 8x60K Microarray (Probe Name Version) | 38 | Sex, Gestational Age, Birth weight | Singapore | Transcriptome changes affecting hedgehog and cytokine signalling in the umbilical cord in late pregnancy: implications for disease risk |
| **GSE53473** | GPL13667 | [HG-U219] Affymetrix Human Genome U219 Array | 128 | Sex, Gestational Age | Finland | Standard of hygiene and immune adaptation in newborn infants |
| **GSE21342** | GPL6947 | Illumina HumanHT-12 V3.0 expression beadchip | 37 | Maternal PP_BMI | USA | Maternal influences on the transmission of leukocyte gene expression profiles in population samples (mother and child) |
| **GSE36828** | GPL6947 | Illumina HumanHT-12 V3.0 expression beadchip | 48 | Gestational Age, Birth weight | USA | Genome-wide analysis of gene expression levels in placenta and cord blood samples from newborns babies |
| **GSE48354** | GPL16686 | [HuGene-2_0-st] Affymetrix Human Gene 2.0 ST Array [transcript (gene) version] | 38 | Sex, Gestational Age, Birth weight | Mexico | Prenatal arsenic exposure and the epigenome: altered gene expression profiles in newborn cord blood |
| **GSE53669** | GPL96 | [HG-U133A] Affymetrix Human Genome U133A Array | 16 | Excluded | USA | Fetal transcripts in maternal blood |
| **GSE29807** | GPL570 | [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array | 12 | Excluded | India | Gene expression study reveals compromised Pattern Recognition Receptors and Interferon Signaling in fullterm Low birth Weight newborns |
| **GSE51546** | GPL10558 | Illumina HumanHT-12 V4.0 expression beadchip | 12 | Excluded | Finland | Maternal-diabetes induced gene expression changes in the umbilical cord |
| **GSE32420** | GPL570 | [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array | 29 | Excluded | Slovakia | Expression data from 0Month (New Born) Slovak Children with high PCBs exposure |
| **GSE7967** | GPL570 | [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array | 65 | Excluded | Thailand | Activation of inflammation and nfkb signaling in infants born to arsenic exposed mothers |
| **GSE25504** | GPL570 | [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array | 20 | Sex | United Kingdom | Whole blood mRNA expression profiling of host molecular networks in neonatal sepsis |

**Table 3. Gene Expression Omnibus Data Sets Identified for Use in Meta-Analysis.**

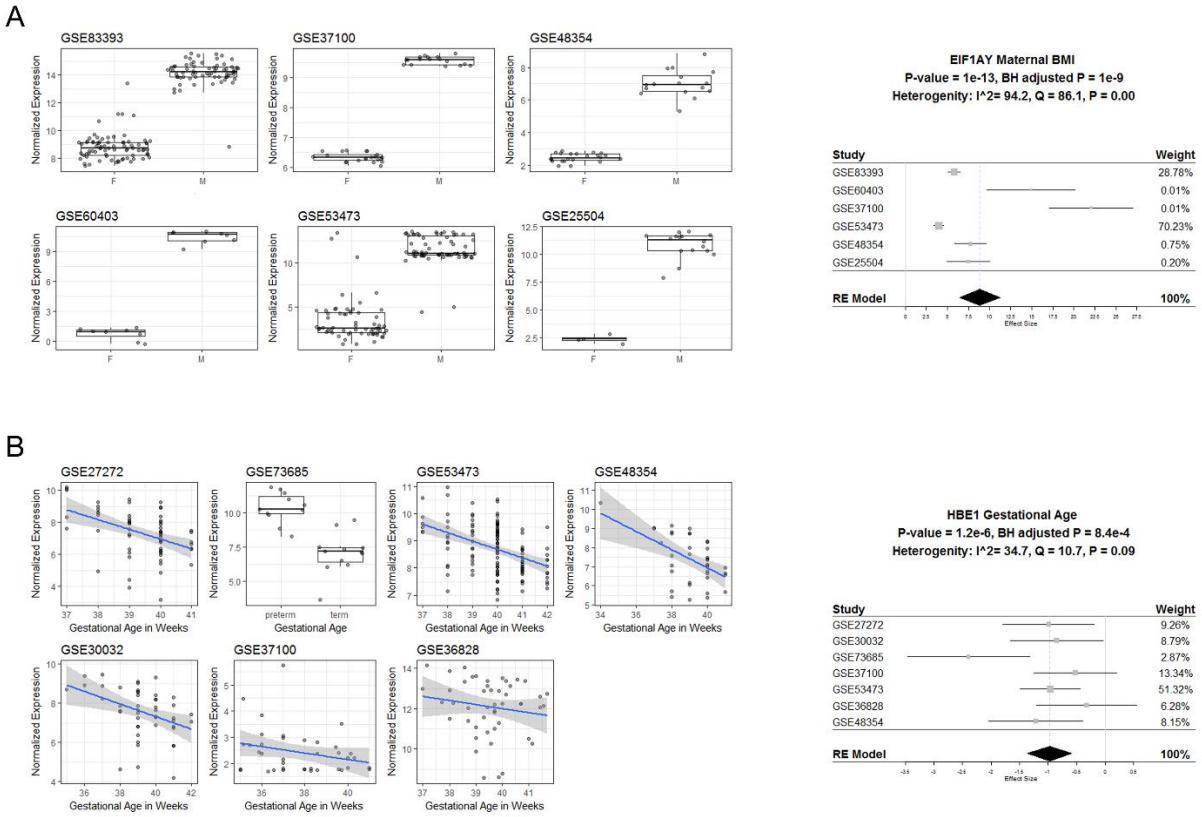


**Figure 6. Representative genes that have known associations with asthma risk factors.** Gene expression and forest plot of effect sizes from individual studies demonstrating associations between **A)** *EIF1AY* – Fetal Sex and **B)** *HBE1* – Gestational Age.

To identify biological processes enriched in genes with transcriptional changes associated with higher or lower risk of asthma, the *z*-statistics from the meta-analysis were averaged such that negative *z*-statistics were associated with lower risk (female, older gestational ages, higher birthweights, and lower maternal pre-pregnancy BMI) and positive z-statistics were associated with increased risk. This averaged *z*-statistic was used as a pre-ranked list for gene set enrichment analysis, GSEA (**Figure 7B**) (104). GO terms were assessed for enrichment, 41 and 31 GO terms were significantly enriched (FWER adjusted p-value < 0.1) with regards to low and high risk profiles, respectively. To identify pathways and genes that were most associated with the biological signal and least redundant, mutual information networks for genes and GO terms was constructed from leading edge genes and significantly enriched GO terms. Genes with the highest closeness centrality and absolute z-statistics should correlate with GO terms that have high neighborhood connectivity, thus allowing for identification of biological processes associated with increased or decreased risk. Specifically, increased expression of ribosomal proteins (*RPS14, RPS27A* and *RPS15A*) and RNA interacting proteins (*DDX1*, *TRMT112*, *MPHOSPH10*, *TSR2*, and *TARBP2*) were associated with higher asthma risk (**Figure 7C**). This matched the ranked importance of GO terms associated with higher asthma risk, as GO terms involving translation and RNA modifying proteins were ranked the highest (**Figure 7D**). Genes associated with the lower risk were related to innate immune processes such TLR signaling and IL1 secretion (**Figure 7C-D**). Specifically *PYCARD, IL1B,* and *TLR2* are key components of these pathways. As confirmation that these gene sets show associations with risk profiles and are not dependent on a single risk factor, z-statistics of leading edge genes from each data set were compared in the

meta-analysis. Low and high risk genes displayed similar behavior to the averaged *z*-statistic, with the exception of the low risk genes when comparing z-statistics in the model assessing maternal pre-pregnancy BMI (**Figure 8A**). These findings suggest that lower risk individuals have increased expression of genes involved in innate immune signaling and defense, whereas high risk individuals have increased expression of genes involved in translation.

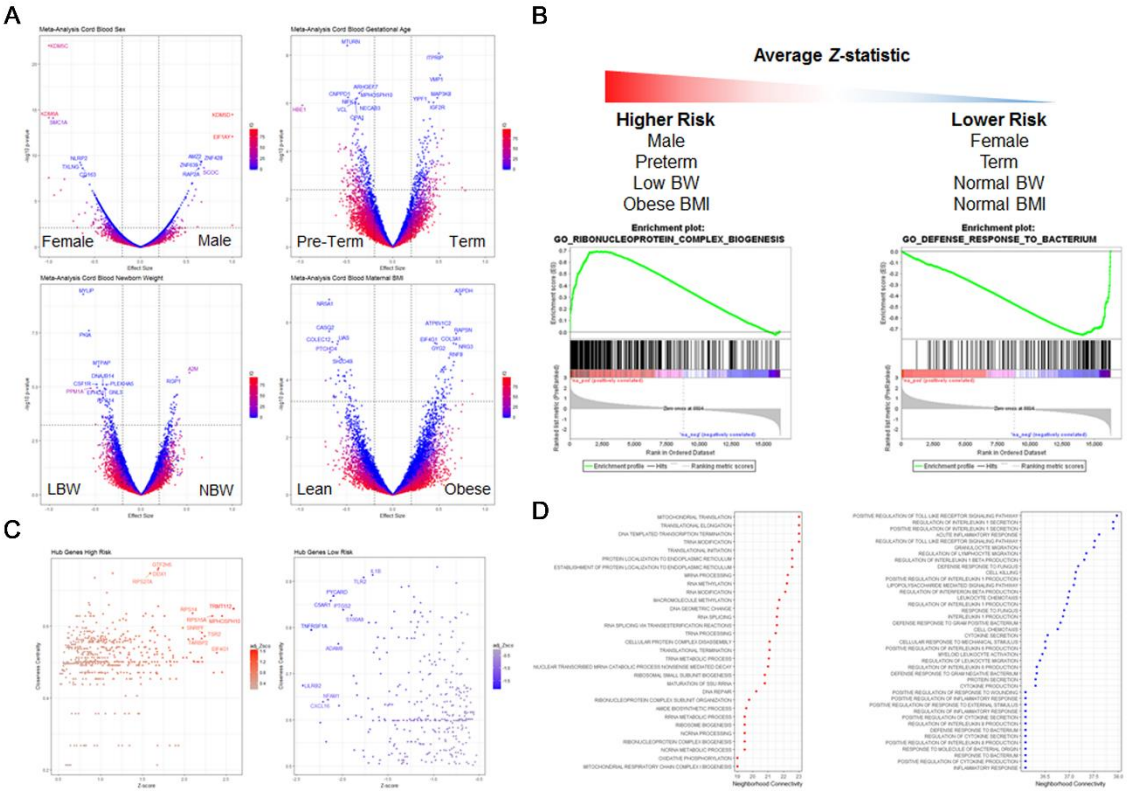


**Figure 7. Meta-Analysis identifies gene expression signatures related to childhood asthma risk profile. A)** Volcano plots displaying relationship between effect size and $-\log_{10}$(p-value). Points are colored by $I^2$ values (0-100%, blue-red). Each plot displays a single univariate test for one risk factor. **B)** Gene specific Z-statistics for each univariate were averaged to create gene association profile with high or low risk for childhood asthma. Association profiles were assessed for enrichment of GO terms using GSEA. **C)** Dot plot displaying to association of leading edge genes from significantly enriched GO terms average Z-statistics and closeness centrality. **D)** Significantly enriched GO terms ranked by their neighborhood centrality.

D. Validation of meta-Analysis identifies neutrophil development signature associated with asthma risk

To confirm the theoretical findings of gene expression changes associated asthma risk stratification, z-statistics were compared from the meta-analysis with a cohort of individuals in which common pre/perinatal risk factors were determined. Four different regression models were compared to the meta-analysis to identify genes that replicated between the meta-analysis and mRNAseq of the UIH cohort. Model 1 used maternal obesity, male, low birth weight (< 3000 g), pre-term birth, maternal allergy/asthma, caesarian section, and maternal smoking as risk factors. Model 2 used maternal obesity, male, low birth weight, and pre-term birth as risk factors. Model 3 used the same demographics as Model 1 minus maternal smoking as risk factors. Model 4 used the same demographics as Model 1 plus ethnicity/race (Black or Hispanic) as risk factors. We used a non-parametric statistic for ranking replication, the product of the z-statistic meta-analysis and z-statistic mRNAseq (**Figure 8B**). Genes with a replication score greater than two were identified as significantly replicating. Of the four models, z-statistics in model 2 had the strongest correlations between the UIC cohort and meta-analysis (**Figure 8C**). This is consistent with model 2 assessing the same risk factors as the meta-analysis, whereas the other models included other risk factors, such as maternal smoking and atopy.

Across all four models tested there was a set of genes that replicated highly (replication score > 3.5) with low risk (*CEACAM8, CAMP, RNASE2, MS4A3, SRGN, PGLYRP1, PTX3, CHI3L1, SLPI,* and *PADI4*) and with high risk (*MTURN, FANK1, VPS28, ELOVL6, CNPPD1*, and *HBE1*). There were several genes identified in the network analysis that had high replication (*TARBP2, LILRB2, TNFRSF1A, CXCL16,*

*PYCARD,* and *C5AR1*), thus validating the network analysis as an assessment tool. For further analysis model 4 was used as a comparator to the meta-analysis (**Figure 9A**). To test if the replicating genes were associated with the number risk factors, significantly replicating genes were decomposed into an eigenvector, where negative eigenvalues represent increased expression of low risk genes and positive eigenvalues represent increase expression of high risk genes (**Figure 9B, C**). It was observed the eigenvalue of the replicating genes was significantly associated ($p$-value < 0.01) with number of pre/perinatal risk factors in the UIH cohort.

To evaluate which gene sets were enriched in significantly replicating genes, GO term enrichment was performed. Top 10 significantly enriched (Bejamini-Hochberg adjusted $p$-value < 0.1) GO terms for high and low risk replicating genes are displayed (**Figure 9D**). Genes with increased expression in low risk individuals were highly enriched for GO terms involved in early neutrophil development. The data used for assessment of transcriptional changes is derived from pooled cell populations, therefore transcriptional differences are likely driven by differences in cell populations.

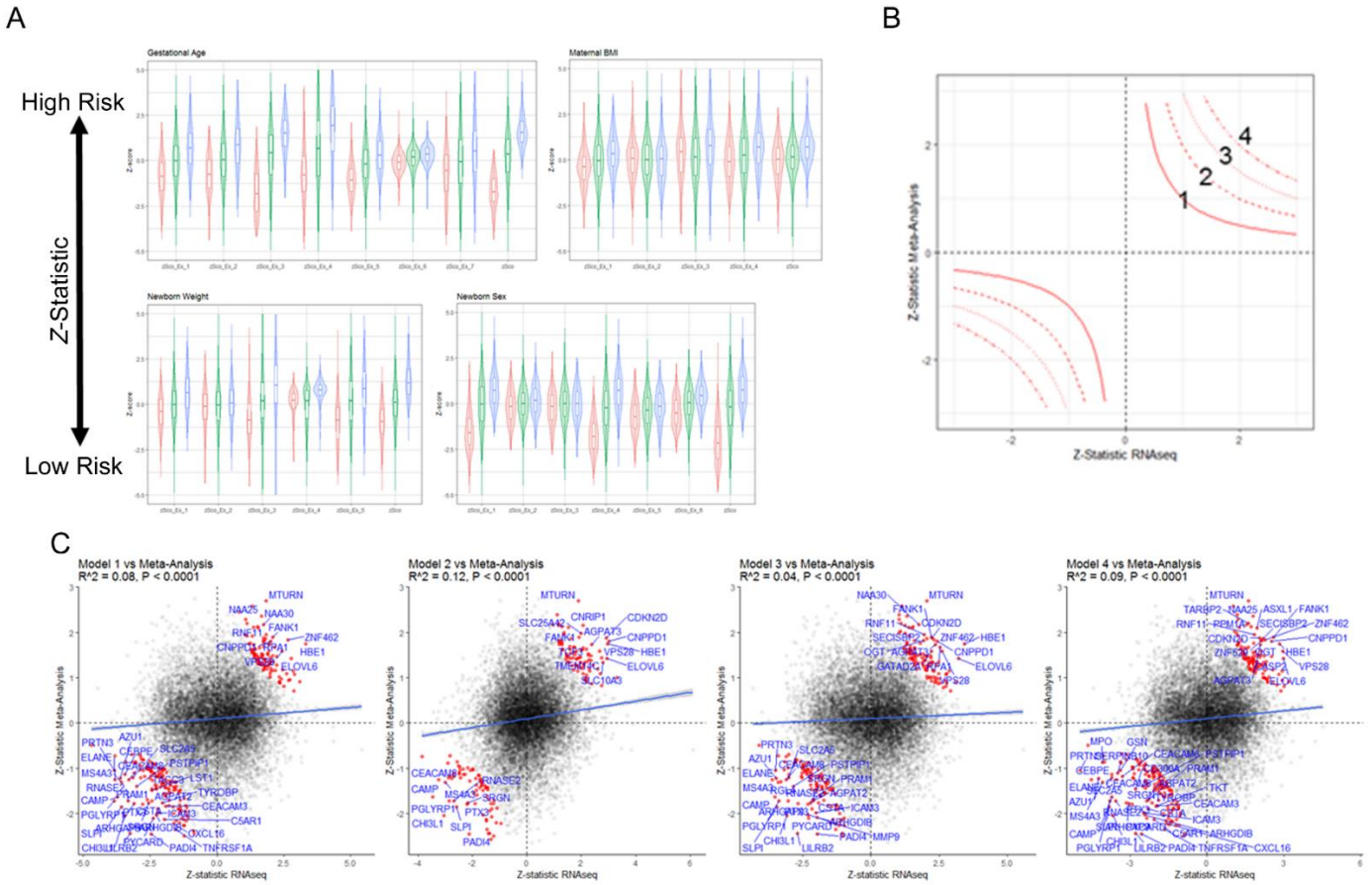


**Figure 8. Replication between meta-analysis and mRNAseq. A)** Replication of high (blue) and low (red) risk gene sets across all data sets used in the meta-analysis. Z-statistics are derived from univariate tests within each study or the meta-analysis. **B)** Curves for replication score cut-offs. **C)** Four mRNAseq models' Z-statistics compared to the meta-analysis Z-statistic. Dot-plot displaying replication between Meta-analysis and mRNAseq Z-statistics. Red dots and labeled dots indicate those with non-parametric replication score greater than two and 3.5 respectively.

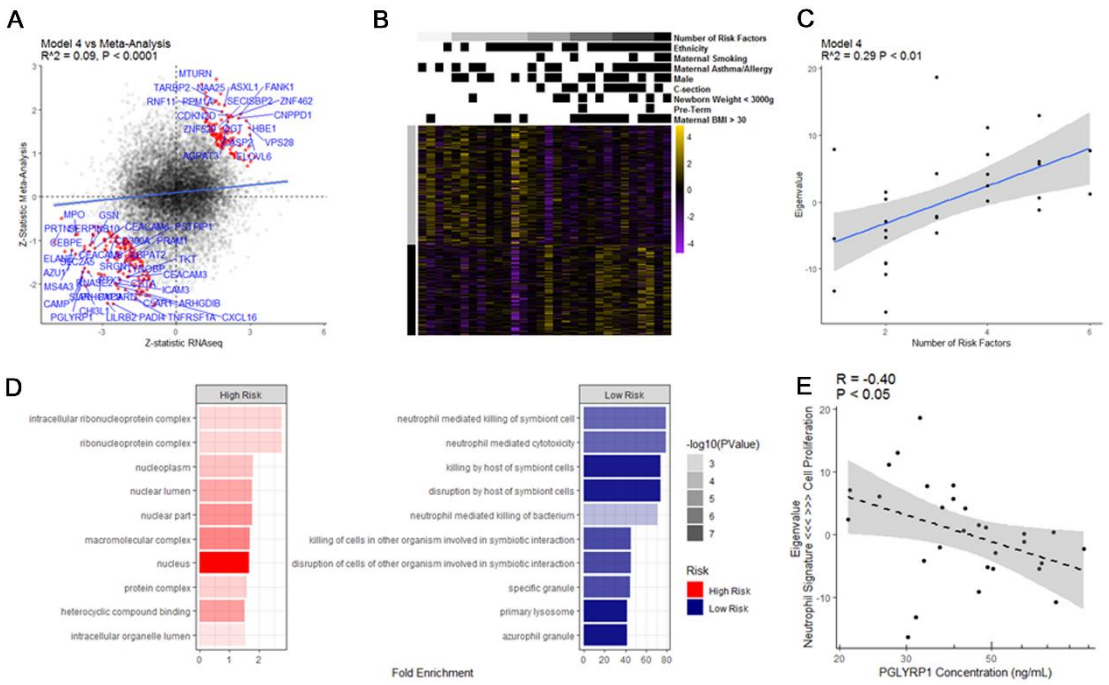


**Figure 9. Replication identifies gene signature associated with early neutrophil development. A)** Dot-plot displaying replication between Meta-analysis and mRNAseq Z-statistics. Red dots and labeled dots indicate those with non-parametric replication score greater than two and 3.5 respectively. **B)** Heatmap showing gene expression, displayed a z-scores, per sample. Samples a ranked from lowest to highest number of risk factors and genes grouped by association with low or high risk (grey or black). **C)** Correlation between number of risk factors and eigenvalue of gene signature. **D)** Association between umbilical cord blood serum PGLYRP-1 abundance and eigenvalue of gene signature**. E)** GO term enrichment associated with low and high risk genes.

To determine if these changes are due to differences in neutrophil abundance, CyTOF data published by Olin et al. underwent secondary analysis for neutrophil abundance in relationship to number of pre/perinatal risk factors (133). Neutrophil abundance significantly decreased with increasing number of risk factors (**Figure 10A**). To confirm and expand on these findings, reported umbilical cord plasma protein abundance data was correlated with number of risk factors. Genes that replicated with low risk had, on average, negative correlations with number of risk factors (**Figure 10B**). Several of the proteins (IL-18, PGLYRP-1, IL-6Rα, MMP-9, and OSM) had strong correlations with both neutrophil abundance and number of risk factors (**Figure 10B, C**). Umbilical cord serum PGLYRP-1 was significantly associated with eigenvalue of asthma risk signature genes in the UIH cohort (**Figure 10E**). Thus, umbilical cord serum PGLYRP-1 may represent novel biomarker for asthma risk and neutrophil development.

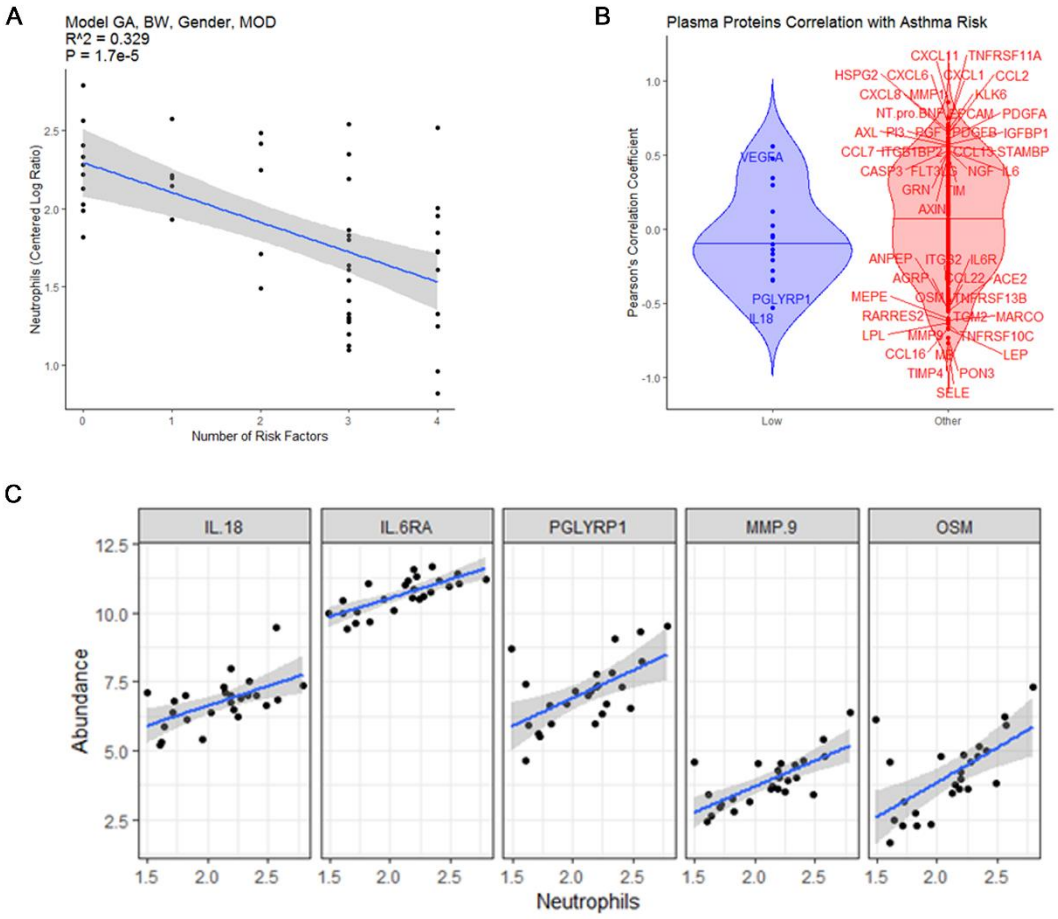


**Figure 10. Cellular and plasma proteomic differences associated with asthma risk factors.** All data are a re-analysis of publicly available data from Olin et al. (133). **A)** Percentages of neutrophils in cord blood were transformed using centered log-ratios and correlated with the number of risk factors. **B)** Pearson's correlation coefficients, *R*, for plasma protein concentration and number of risk factors displayed for low risk associated proteins (blue) and all other proteins (red). Proteins with $|R| > 0.5$ are displayed with text. **C)** Plasma proteins with $R > 0.5$ for correlations with cord blood neutrophil abundance and $R < -0.5$ with risk factors for asthma.

E. Umbilical cord serum PGLYRP-1 concentration is associated with odds of childhood asthma and pulmonary function differences

PGLYRP-1 concentration in umbilical cord serum was evaluated in Project VIVA to determine its relationship to longitudinal outcomes. The demographics of the subset of individuals from Project VIVA in which umbilical cord serum was available demonstrated a similar relationship between number of risk factors and proportion of individuals with childhood asthma (**Table 4**). The distribution for the concentration of PGLYRP-1 was right skewed, therefore it was log-transformed and scaled for the use in all regressions. There were no differences in the distributions between plates or cohorts. Asthmatic individuals within Project VIVA cohort were more often male (71 vs. 50%), pre-term (12 vs 2%), exposed to smoking during pregnancy (10 vs. 5%), identified as Black or Hispanic (31 vs. 16%) and offspring of atopic (40 vs. 32%) and pre-pregnancy obese (17 vs. 13%) mothers.

To evaluate if PGLYRP-1 was associated with specific risk factors, univariate regressions were performed (**Figure 11A**). Males, preterm births and individuals born by Cesarean section were associated with lower concentrations of PGLYRP-1. PGLYRP-1 was modeled as a predictor for physician diagnosed asthma and current asthma at mid-childhood (median age ~7.7 years old) and early-teen (median age ~12.8 years old) (**Figure 11B, C**). The abundance of PGLYRP-1 was significantly associated with physician diagnosed asthma and current asthma at mid-childhood (aOR [95% CI]: 0.429 [0.265-0.693], *p*-value < 0.001 and 0.444 [0.267-0.737], *p*-value < 0.01, per 1SD increase). PGLYRP1 was not associated with physician diagnosed asthma or current asthma at early-teen (aOR [95% CI]: 0.720 [0.481-1.078], *p*-value = 0.110 and 0.628 [0.294-1.339], *p*-value = 0.220, per 1SD increase). To identify moderating co-variates,

subset analysis was performed (**Figure 11D**). Small for gestational age and children identifying as other race subset of individuals displayed significantly different associations between PGLYRP-1 and mid-childhood asthma. Maternal smoking during pregnancy and maternal race identifying as other trended to have different associations between PGLYRP-1 and mid-childhood asthma. This suggest that PGLYRP-1's effect or predictive role in childhood asthma may be limited in populations with increased maternal smoking, high rates of low birthweights, and races other than White, Black, or Hispanic.

PGLYRP-1 was measured as a predictor of pulmonary function, measured by FEV1/FVC and bronchodilator response, at mid-childhood and early-teen follow-up time points (**Figure 11E, F**). There was trend towards an association between PGLYRP-1 concentration and FEV1/FVC at mid-childhood (aβ [SE]: 0.015 (0.007), *p*-value =0.055, per 1SD increase). Similar to FEV1/FVC associations at mid-childhood, there was significant association between PGLYRP-1 and FEV1/FVC at the early-teen follow up (aβ [SE]: 0.012 (0.005), *p*-value < 0.05, per 1SD increase). However, there was no association with bronchodilator response at either time-point. These findings indicate that the concentration PGLYRP-1 in umbilical cord serum is associated with probability of asthma at mid-childhood and expiratory flow throughout childhood and early teenage years. Umbilical cord serum sIL-6Rα, which was highly correlated with neutrophil abundance, was significantly associated with PGLYRP-1 (aβ [SE]: 0.191 (0.055), *p*-value < 0.001, per 1SD increase). Unlike PGLYRP-1, sIL-6Rα was not significantly associated with asthma outcomes or pulmonary function tests at any time-point. Taken together with the transcriptomic results, this suggest that the abundance of neutrophil vesicular proteins in umbilical cord serum are the most directly associated with risk for asthma.

| | No Response | | No Asthma | | Asthma | |
|---|---|---|---|---|---|---|
| ***N*** | 128 | | 188 | | 42 | |
| | | | | | | |
| **Mother** | | | | | | |
| *Pre-pregnancy BMI category, kg/m$^2$, %* | | | | | | |
| *. Underweight (<18.5)* | 7 | 0.05 | 7 | 0.04 | 1 | 0.02 |
| *. Normal (18.5-<25.0)* | 73 | 0.57 | 114 | 0.61 | 27 | 0.64 |
| *. Overweight (25.0-<30.0)* | 29 | 0.23 | 42 | 0.22 | 7 | 0.17 |
| *. Obese (>=30.0)* | 18 | 0.14 | 24 | 0.13 | 7 | 0.17 |
| *Pre-pregnancy BMI, kg/m$^2$* | 24.76 | 5.23 | 24.41 | 4.99 | 25.02 | 5.20 |
| *Race/ethnicity, %* | | | | | | |
| *. White* | 93 | 0.73 | 136 | 0.72 | 26 | 0.62 |
| *. Black* | 20 | 0.16 | 20 | 0.11 | 7 | 0.17 |
| *. Hispanic* | 6 | 0.05 | 7 | 0.04 | 6 | 0.14 |
| *. Other* | 7 | 0.05 | 25 | 0.13 | 3 | 0.07 |
| *>= college graduate, %* | | | | | | |
| *. No* | 53 | 0.41 | 59 | 0.31 | 16 | 0.38 |
| *. Yes* | 73 | 0.57 | 129 | 0.69 | 26 | 0.62 |
| *History of atopy, %* | | | | | | |
| *. No* | 78 | 0.61 | 128 | 0.68 | 25 | 0.60 |
| *. Yes* | 48 | 0.38 | 60 | 0.32 | 17 | 0.40 |
| *Any antibiotics during pregnancy, %* | | | | | | |
| *. No* | 81 | 0.63 | 137 | 0.73 | 32 | 0.76 |
| *. Yes* | 47 | 0.37 | 51 | 0.27 | 10 | 0.24 |
| *Pregnancy smoking status, %* | | | | | | |
| *. Never* | 82 | 0.64 | 141 | 0.75 | 30 | 0.71 |
| *. Former* | 21 | 0.16 | 35 | 0.19 | 8 | 0.19 |
| *. During pregnancy* | 25 | 0.20 | 9 | 0.05 | 4 | 0.10 |
| *Maternal smoking during pregnancy, 6m or 1y, %* | | | | | | |
| *. No* | 102 | 0.80 | 177 | 0.94 | 38 | 0.90 |
| *. Yes* | 26 | 0.20 | 11 | 0.06 | 4 | 0.10 |
| *C-section, %* | | | | | | |
| *. No* | 104 | 0.81 | 151 | 0.80 | 36 | 0.86 |
| *. Yes* | 24 | 0.19 | 37 | 0.20 | 6 | 0.14 |
| | | | | | | |
| **Child** | | | | | | |
| *Gestational age, weeks, mean (SD)* | 39.59 | 1.29 | 39.83 | 1.52 | 39.18 | 2.30 |
| *Gestational age category, %* | | | | | | |
| *. <37 w* | 4 | 0.03 | 4 | 0.02 | 5 | 0.12 |
| *. 37-<42 w* | 122 | 0.95 | 179 | 0.95 | 37 | 0.88 |
| *. >=42 w* | 2 | 0.02 | 5 | 0.03 | 0 | 0.00 |
| *BW for GA z-score, mean (SD)* | 0.24 | 0.90 | 0.30 | 0.95 | 0.35 | 0.95 |
| *Sex* | | | | | | |
| *. Male* | 67 | 0.52 | 94 | 0.50 | 30 | 0.71 |
| *. Female* | 61 | 0.48 | 94 | 0.50 | 12 | 0.29 |
| *Child race/ethnicity, %* | | | | | | |
| *. White* | 93 | 0.73 | 128 | 0.68 | 23 | 0.55 |
| *. Black* | 20 | 0.16 | 24 | 0.13 | 8 | 0.19 |
| *. Hispanic* | 6 | 0.05 | 6 | 0.03 | 5 | 0.12 |
| *. Other* | 7 | 0.05 | 30 | 0.16 | 6 | 0.14 |

**Table 4. Demographics of Project VIVA cohort dichotomized for mid-childhood asthma.**

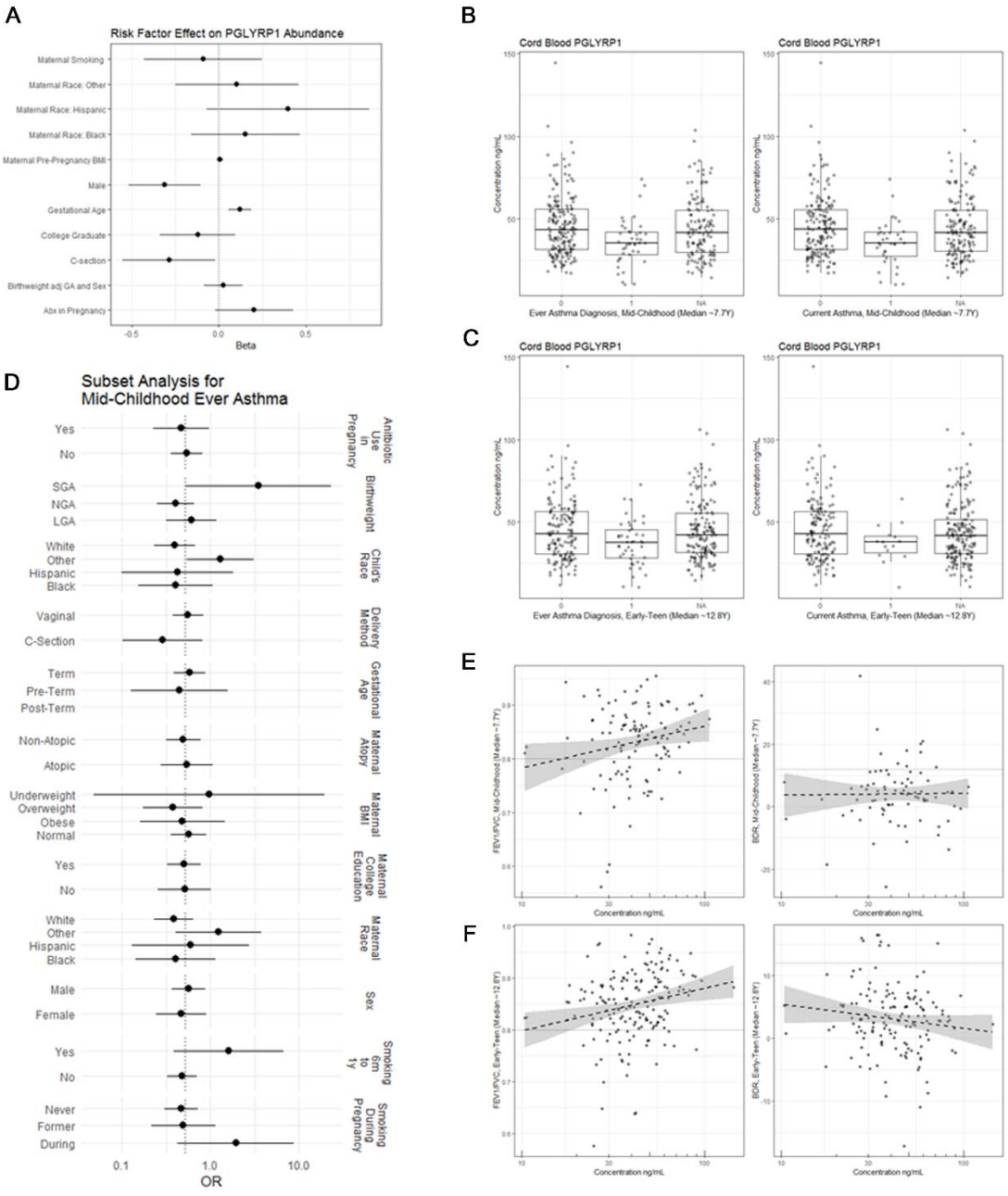


**Figure 11. Increased umbilical cord blood serum PGLYRP-1 is associated with reduced odds of childhood asthma and increased FEV1/FVC. A)** Regression effect estimates between pre/perinatal risk factors and PGLRYP-1 abundance z-scores. PGLYRP-1 concentration in umbilical cord blood serum in relationship to **B)** childhood asthma and **C)** early-teen asthma. **D)** Subset analysis for all covariates used in regression models. Odds ratios displayed for univariate association between PGLYRP-1 z-score and childhood asthma. PGLYRP-1 concentration in umbilical cord blood serum in relationship to FEV1/FVC and bronchodilator response at **E)** mid-childhood and **F)** early-teen.

F. Pre/perinatal risk and Umbilical cord serum PGLYRP-1 concentration are associated with reduced CBMC proliferative responses to common aeroallergens

CBMC proliferative response have previously been associated with cytokine production in response to allergen stimulation, other risk predictors of asthma (e.g. poly-unsaturated fatty acids), and risk of persistent wheezing (114). As expected, increasing number of pre/perinatal factors was significantly correlated with proliferation in response to Der f1 ($\rho$ = 0.161, $p$-value < 0.05) and Bla g2 ($\rho$ = 0.161, $p$-value < 0.05), but not to phytohemagglutinin (PHA) or ovalbumin (OVA) stimulation (**Figure 12A**). PGLYRP-1 concentrations were significantly associated with CBMC proliferation after Der f1 ($\rho$ = -0.149, $p$-value < 0.05) stimulation (**Figure 12B**).

To expand on the finding that PGLYRP-1 concentrations were associated with CBMC proliferative responses to a protein antigen, a cellular correlation network derived from reanalysis of a previously published CyTOF data set from umbilical cord and peripheral blood leukocyte populations at birth, 1 week, and 12 weeks after birth was evaluated (133). There were strong negative correlations between myeloid populations (e.g. neutrophils, CD14+ monocytes, plasmacytoid dendritic cells) and memory T-cell populations (**Figure 13A**). Cord blood neutrophil abundances were significantly associated with CD8+ effector memory abundance dynamics after birth, and individuals with high abundances of neutrophils showed significantly reduced CD8+ effector memory T-cell abundances at birth, 1 week, and 12 weeks after birth (**Figure 13B**). This suggests myeloid populations present at birth are capable of suppressing adaptive immunity several months after birth.

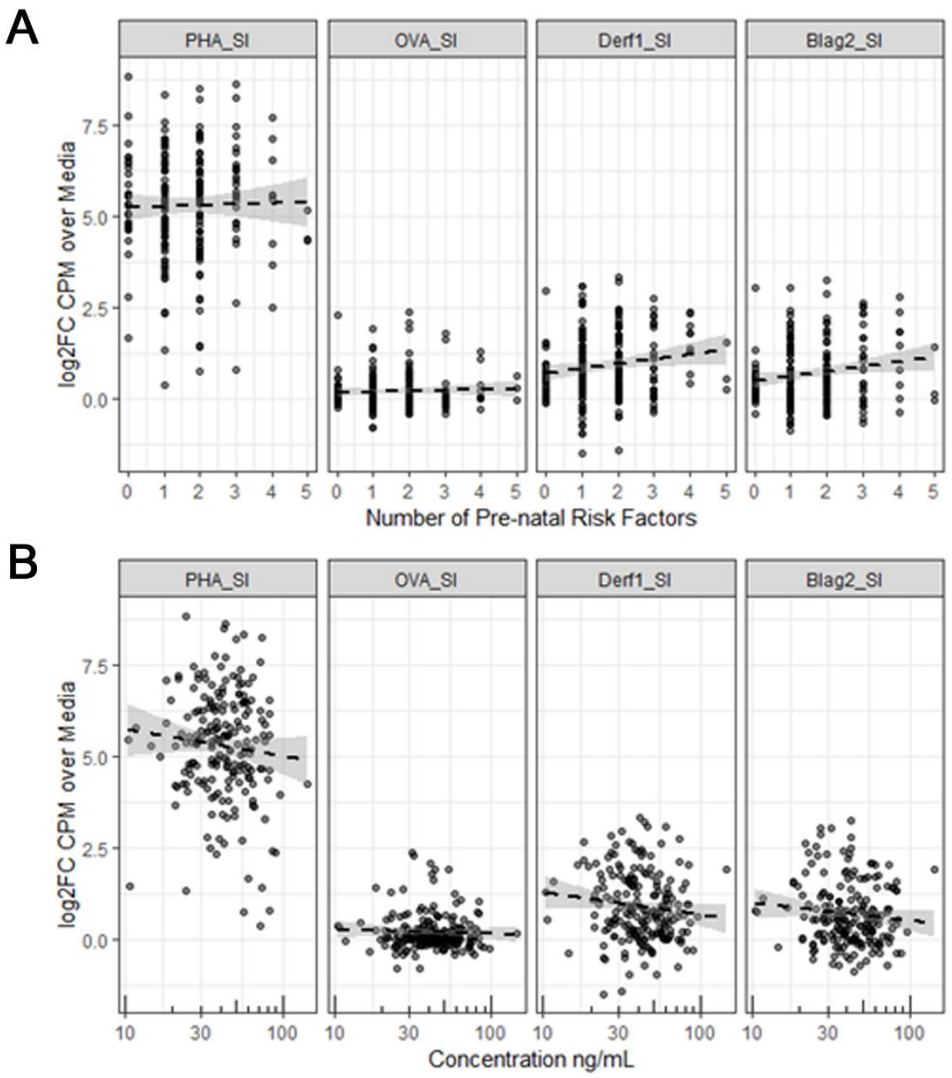


**Figure 12. Increasing number of pre/perinatal risk and lower PGLYRP-1 are associated with increased CBMC proliferation in response to common aeroallergens. A)** Association between number of pre/perinatal risk factors and cellular proliferation, as measured by $H^3$-thymidine incorporation, in response to stimulation with phytohemagglutinin, ovalbumin, Der f1, or Bla g2. **B)** Association between umbilical cord blood serum PGLYRP-1 concentration and cellular proliferation, as measured by $H^3$-thymidine incorporation, in response to stimulation with phytohemagglutinin, ovalbumin, Der f1, or Bla g2.

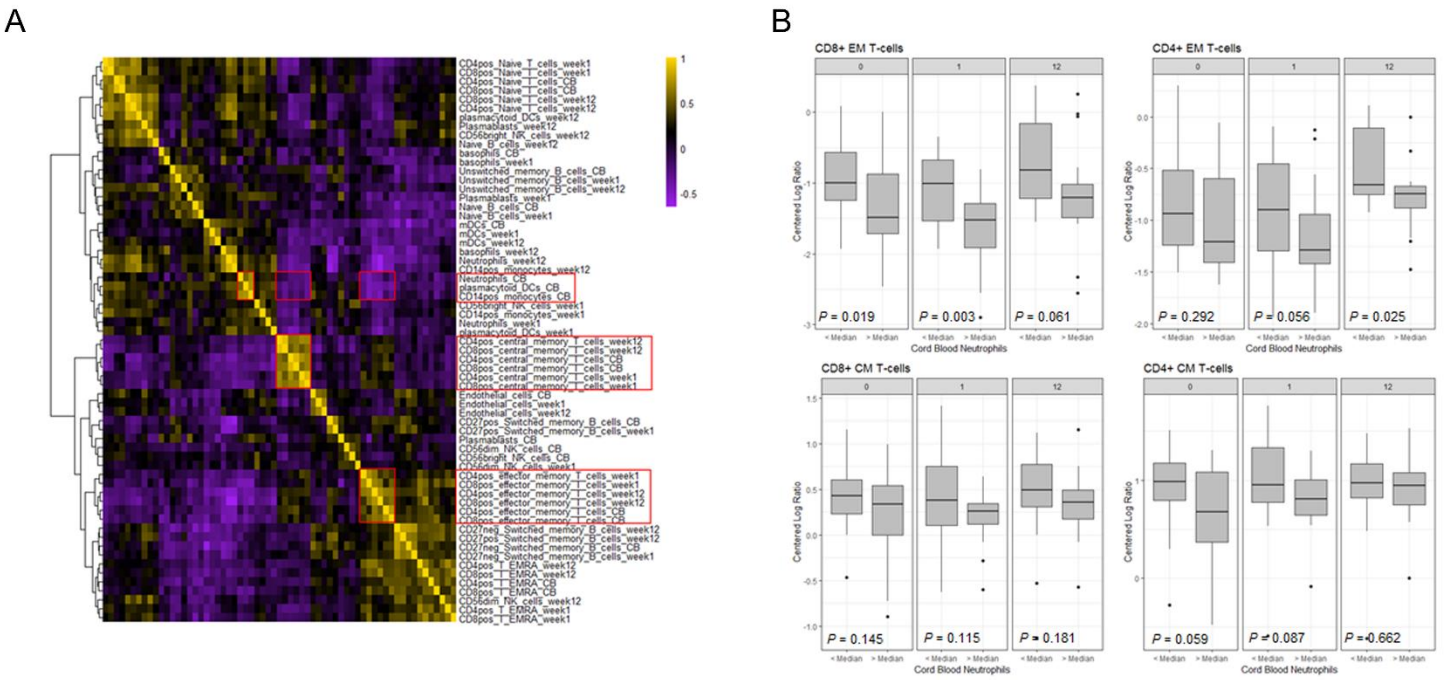


**Figure 13. Cellular dynamics of peripheral leukocytes during early life.** All data are a re-analysis of publicly available data from Olin et al. (133). **A)** Heatmap of Peason's correlation coefficients between whole blood cell populations at birth, 1, and 12 weeks. Red boxes are highlighting cell population's interactions of interest. **B)** Samples were dichotomized by the median abundance of cord blood neutrophils. Effector memory (EM) and central memory (CM) T-cell populations were assessed at birth, 1, and 12 weeks in greater or less than median abundance of cord blood neutrophil groups. $P < 0.05$ was considered statistically significant, Mann-Whitney U test.

G. Bacterial taxa in the maternal third trimester fecal microbiome are associated with umbilical cord serum PGLYRP-1 and neutrophil development signature in CBMCs

Difference in maternal fecal microbial compositions and diet, a major driver of gut microbiome composition, have been associated with wheezing illness in children (126). Fecal material in a subset of mothers in the UIH cohort was collected in the third trimester and assessed for microbial composition. There were no significant associations between α- or β-diversity and umbilical cord serum PGLYRP-1 or CBMC neutrophil development signature. Abundance of individual taxa were tested for associations with umbilical cord serum PGLYRP-1 and CBMC neutrophil development signature. Two genera of bacteria, *Clostridium* and *Ruminococcus*, had significant associations (GBH adjusted *p*-value < 0.05) with the concentration of PGLYRP-1 (**Figure 14A**). Both genera had species that had positive or negative associations with the concentration of PGLYRP-1. Highlighting this finding are *Rumininoccus albus* and *sp. SR1/5*, which were both found to be highly abundant and prevalent, yet had opposite associations with the concentration of PGLYRP-1. In contrast, only one genera, *Bacteroides*, had significant associations (GBH adjusted *p*-value < 0.05) with the neutrophil development signature (eigenvalue) identified in CBMCs (**Figure 14B**). Both highly abundant species, *Bacteroides vulgatus*, and low abundant species, *Bacteroides uniformis*, showed significant associations with increasing eigenvalue. This indicates that the abundance of *Bacteroides* is associated with reduced representation of mature neutrophil transcripts in CBMCs.

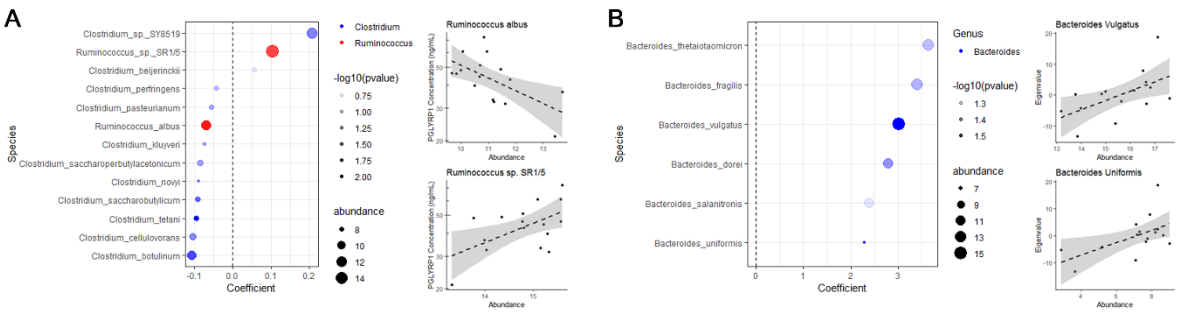


**Figure 14. Abundance of maternal third-trimester fecal Clostrdium, Ruminococcus, and Bacteroides species have significant associations with PGLYRP-1 abundance and neutrophil development signature.** Species with significant, GBH adjusted p-value < 0.05, association between bacterial abundance (rlog transformed counts) and **A)** PGLYRP-1 concentration or **B)** eigenvalue of asthma risk gene signature. Representative associations shown for PGLYRP-1 and eigenvalue.

# CHAPTER 5

# ASSESSING THE RELATIONSHIP BETWEEN PERINATAL MICROBIAL EXPOSURE AND IMMUNITY

## A. Rationale

Recent literature suggests that exposure to microbial products may occur throughout in utero development (48, 64, 125, 126). Previously, we have shown that there is bacterial DNA that circulates in umbilical cord blood. The composition of this circulating bacterial DNA is associated with $T_{h2}$ responses to allergens and early childhood wheezing. This data suggests that cfbDNA may have utility as biomarker for asthma in childhood.

Compositional approaches such as current methods for 16s rRNA sequencing have several limitations including that these approaches are not quantitative. Additionally, there have been serious concerns about contaminating sequences and batch effects in sequencing data that are erroneously associated with outcomes (134-136). There have been several methods developed to quantitate 16s sequences, including those that directly perform 16s rRNA qPCR, spiking-in bacterial sequences expected not to occur in the microbiome assessed, and spiking-in artificial bacterial sequences (76, 137). Unlike novel approaches to quantitative methods, batch correction has remained an underexplored area of sequencing methodologies. One study has identified a methodology for case-control comparisons but it does appear applicable for other analysis such as continuous variable or for multiple outcomes (138).

To address these issues, a quantitative spike-in approach using artificial bacterial sequences was adapted to assess cfbDNA (76). In addition to the goal of quantitation,

background contamination can be addressed by assessing negative controls with spike-in fragments. These important modifications to the previously used cfbDNA extraction and 16s rRNA sequencing methods allow for a better estimation of true microbial exposures at birth. Using the newly adapted method, we hypothesize that the quantitative abundance of microbial exposures will associate with specific clinical demographics, the concentration of PGLYRP-1, and risk of asthma in mid-childhood.

B. Synthetic spike-in amplicon approach allows for quantitation of femtogram quantities of microbial DNA

To test the value of spiking-in synthetic sequences for quantitation of microbial taxa, three synthetic spike-ins EC5001, EC55001, and CA5501 were mixed with the two different mock libraries of concentrations varying between (5 ng to 5 fg) at 1000, 100, and 10 copies. The mock libraries contained 20 identical species but the distribution of the composition between the two varies with one being evenly distributed (every species 5% relative abundance) whereas the other has a staggered distribution (species vary from 18-0.8%).

The total amount of spike-in was observed to a negative log-linear relationship with the starting material (**Figure 15A**). This indicates that as the starting concentration of microbial DNA decreases there is increased representation of the spikes with composition of 16s rRNA amplicon reads. It was observed that the 10 copy spike was inconsistently detected (**Figure 15B**), indicating incomplete sampling of low abundant taxa/spikes, which do not have a consistent sampling pattern, is not related to starting quantity. Most importantly, compositional differences, assessed by unweighted and weighted UNIFRAC distances, were maintained between the two mock libraries types (**Figure 16**).

Unweighted UNIFRAC distances showed an overlap of the two mock libraries, indicating that the same taxa made up their compositions, whereas weighted UNIFRAC demonstrated two distinct compositions. In assessing UNIFRAC distances, both mock libraries were easily distinguishable from negative controls. Although, the lowest dilution (5 fg) was observed to more similar to the negative control that other dilutions.

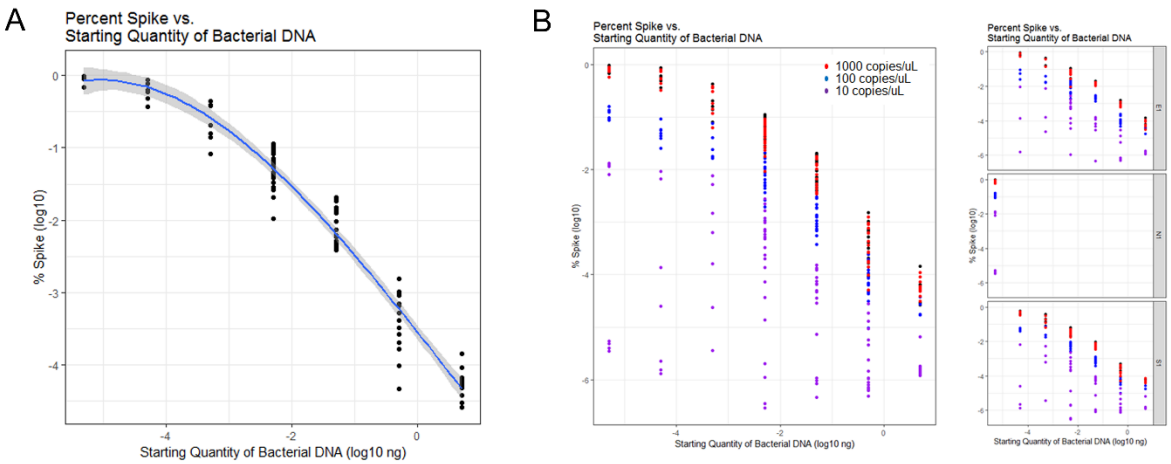


**Figure 15. Relationship between relative abundance of spike-in sequences and starting quantity of bacterial DNA. A)** Loess regression of log10 relative abundance of total spike detected in different mock library dilutions. Points less than $10^{-5}$ ng are negative controls and quantity is displayed as pseudocount. **B)** Points of individual spikes show colored by their initial copy number: 1000 (red), 100 (blue), or 10 (purple) red. All samples are shown (left) or split by mock library type: even (E), staggered (S), or negative (N) distributions.

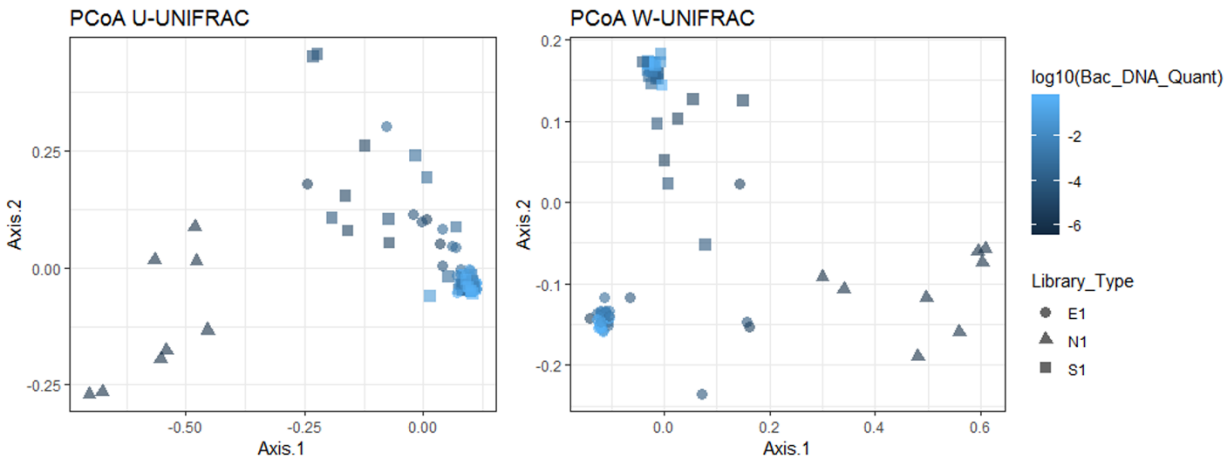


**Figure 16. Compositional effects of dilution and spike-ins in mock library sequencing.** β-diversity was assessed on diluted staggered and even mock libraries and compared to negative controls using principal component analysis (PCoA) with either unweighted or weighted UNIFRAC distances (U- or W-Unifrac). Samples are colored by their starting concentration of mock library and shaped by their library type: even (E), staggered (S), or negative (N).

C. Clinical Determinants of Microbial Exposures

406 samples were sequenced that included 342 cord blood samples and 64 negative controls. Of the 406 samples 249 samples had a least 3 of spikes detected and the spikes displayed log linear relationship. Notably, the majority of samples that were exclude from further analysis were 56 negative controls, largely due to detection of only one or two spikes suggesting that sampling lower concentration fragments in negative controls and serum samples is not equivalent. After filtering samples we assessed the prevalence and mean abundance of taxa detected in serum samples. The majority of taxa detected were sparse distributed with prevalence less than 10% in all of samples. 41 taxa were detected with a prevalence above 10% (**Figure 17**). These taxa represented 4 phyla, Actinobacteria, Bacteroidetes, Firmicutes, and Proteobacteria. When assessing the distribution of these 41 taxa among the 249 samples a concerning batch effect was noted (**Figure 18**), with several taxa displaying batch associated distributions: Stenotrophomonas, Shingobium, and Renibacterium. To correct for the batch effect, taxa were quantitated using mixed model regression estimated by the spike distribution and then batch corrected using *limma* (**Figure 19**). When assessing taxa distributions afterward, the batch effect was removed (**Figure 20**).

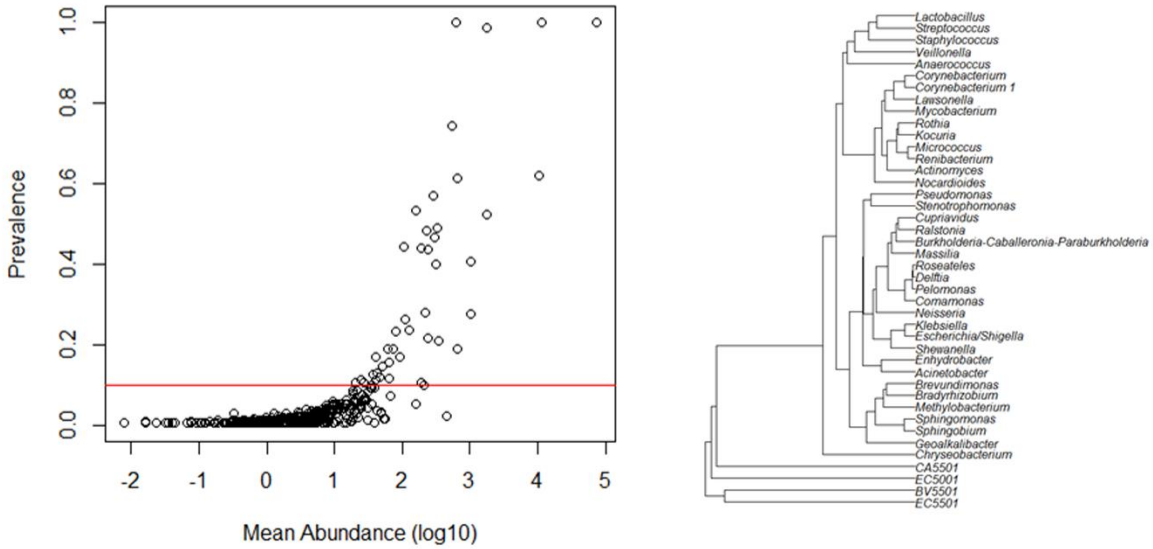


**Figure 17. Most prevalent cfbDNA sequences in umbilical cord blood serum.** Prevalence of cfbDNA sequences across all samples plotted as a function of mean abundance across all samples (left). Red Line indicates 10% prevalence threshold. Dendogram of cfbDNA sequences with great than 10% prevalence using UPGMA hierarchical clustering.

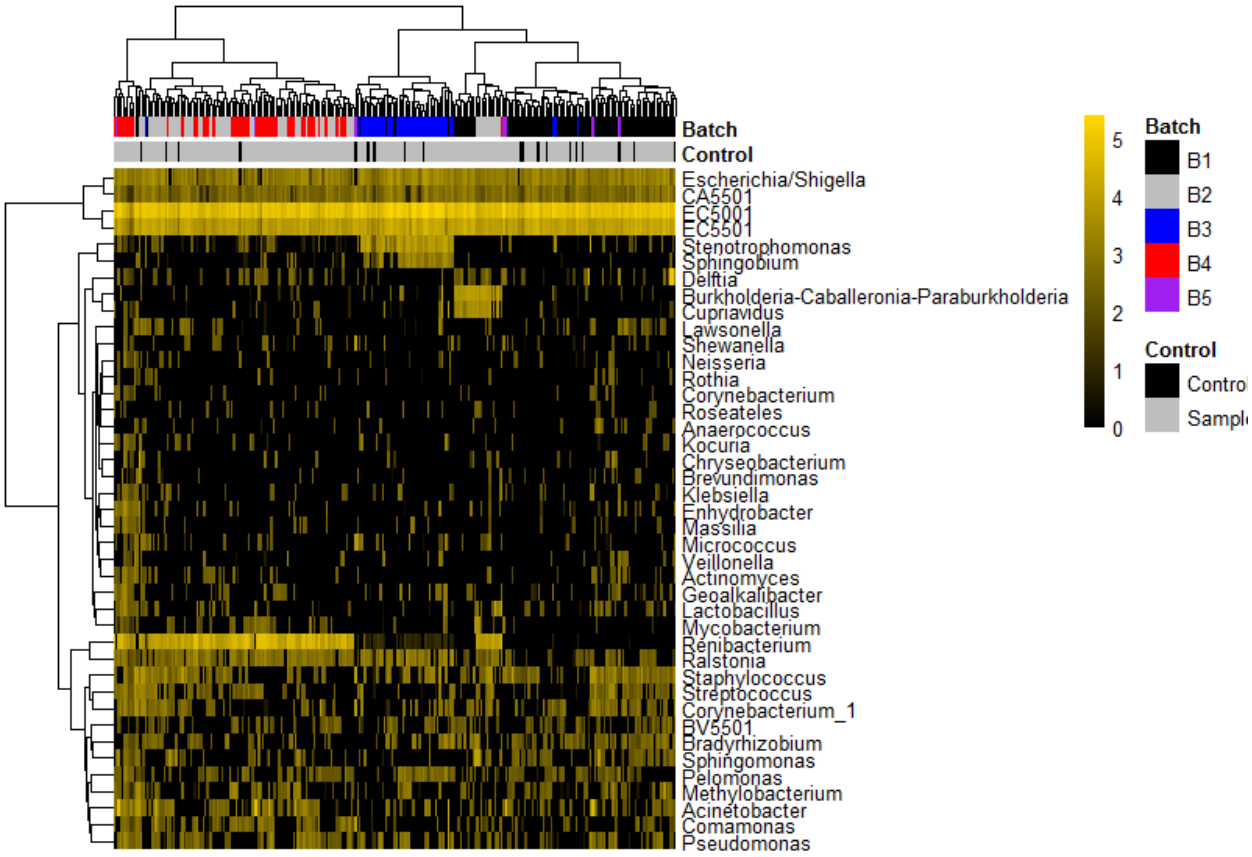


**Figure 18. Abundance of most prevalent genera raw read counts.** Heatmap displaying log10 raw read counts for bacterial genera and spike-ins with prevalence above 10%. Samples and genera are clustered by complete linkage clustering using Euclidean distances.

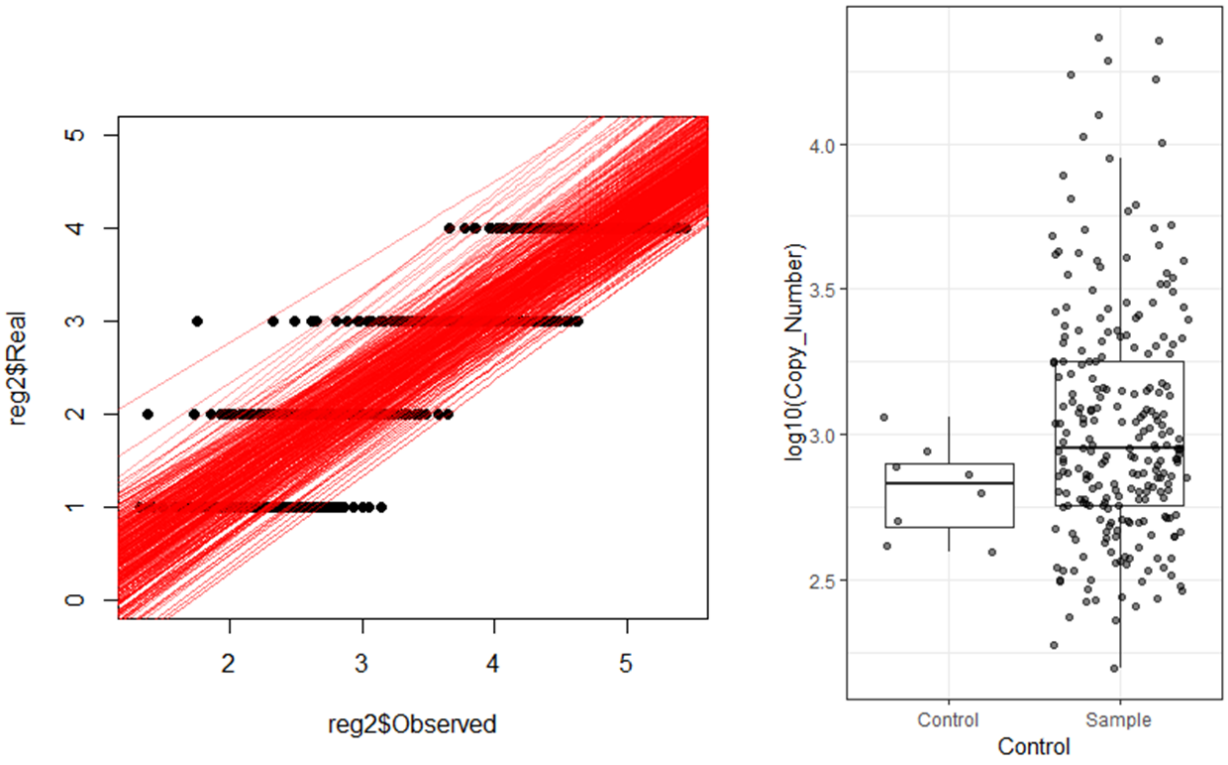


**Figure 19. Estimation of copy number of 16s rRNA in umbilical cord blood serum.** Mixed Effects model was fit on to spike-ins of known concentrations (left). Red Lines are estimated relationships of observed count and expected copy number for each individual sample. Axis are on log10 scale. Expected copy number for negative controls and umbilical cord blood serum (right).

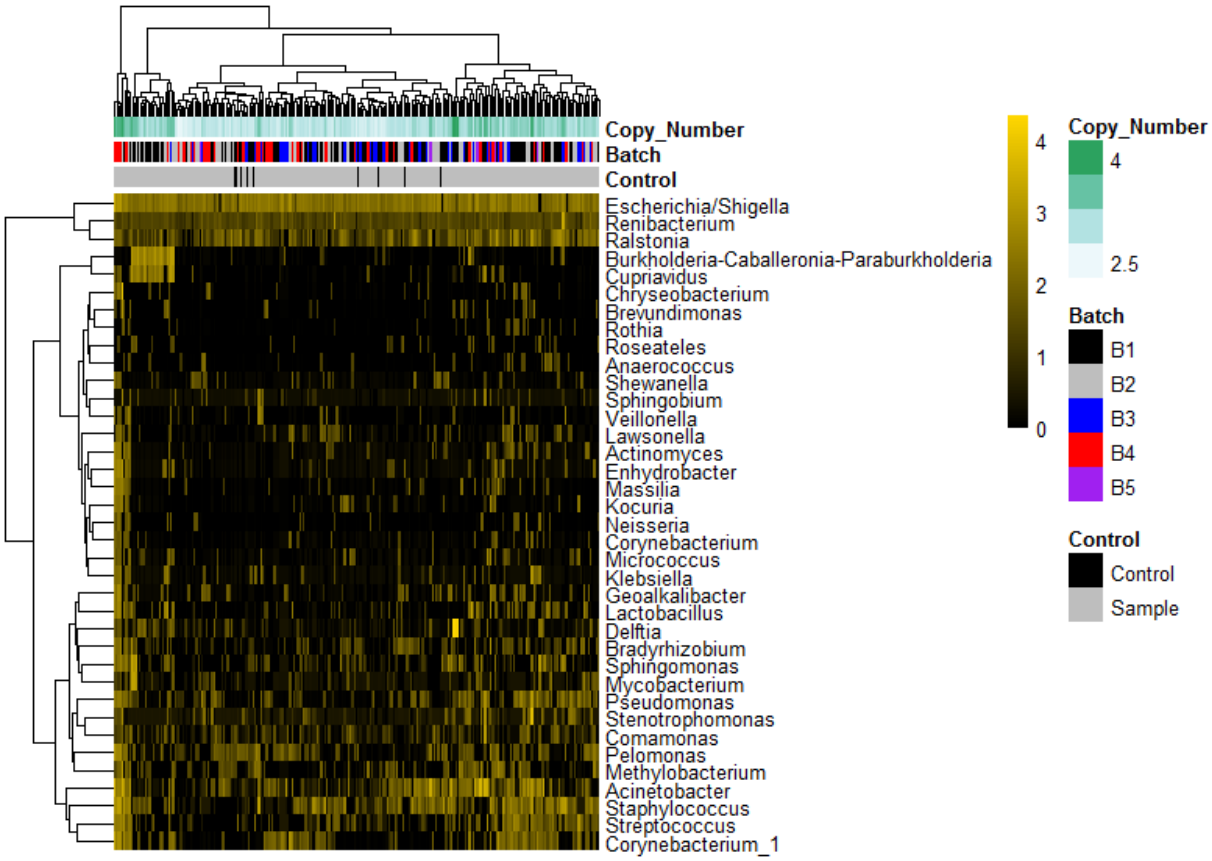


**Figure 20. Abundance of most prevalent genera after copy number estimation and batch correction.** Heatmap displaying log10 estimated and batch corrected copy number for bacterial genera with prevalence above 10%. Samples and genera are clustered by complete linkage clustering using Euclidean distances.

Analysis of Similarites (ANOSIM) was performed, assessing β-diversity associated with different asthma risk factors, and asthma risk factors showed increased dissimilarities within their risk categories. This indicates that there is not consistent microbiome composition that is associated with specific risk factor demographics. It was observed that vaginal delivery was associated with reduced within group β-diversity when compared to between group and caesarian section dissimilarities (**Figure 21**). This indicates that vaginal delivery is associated with specific microbial composition.

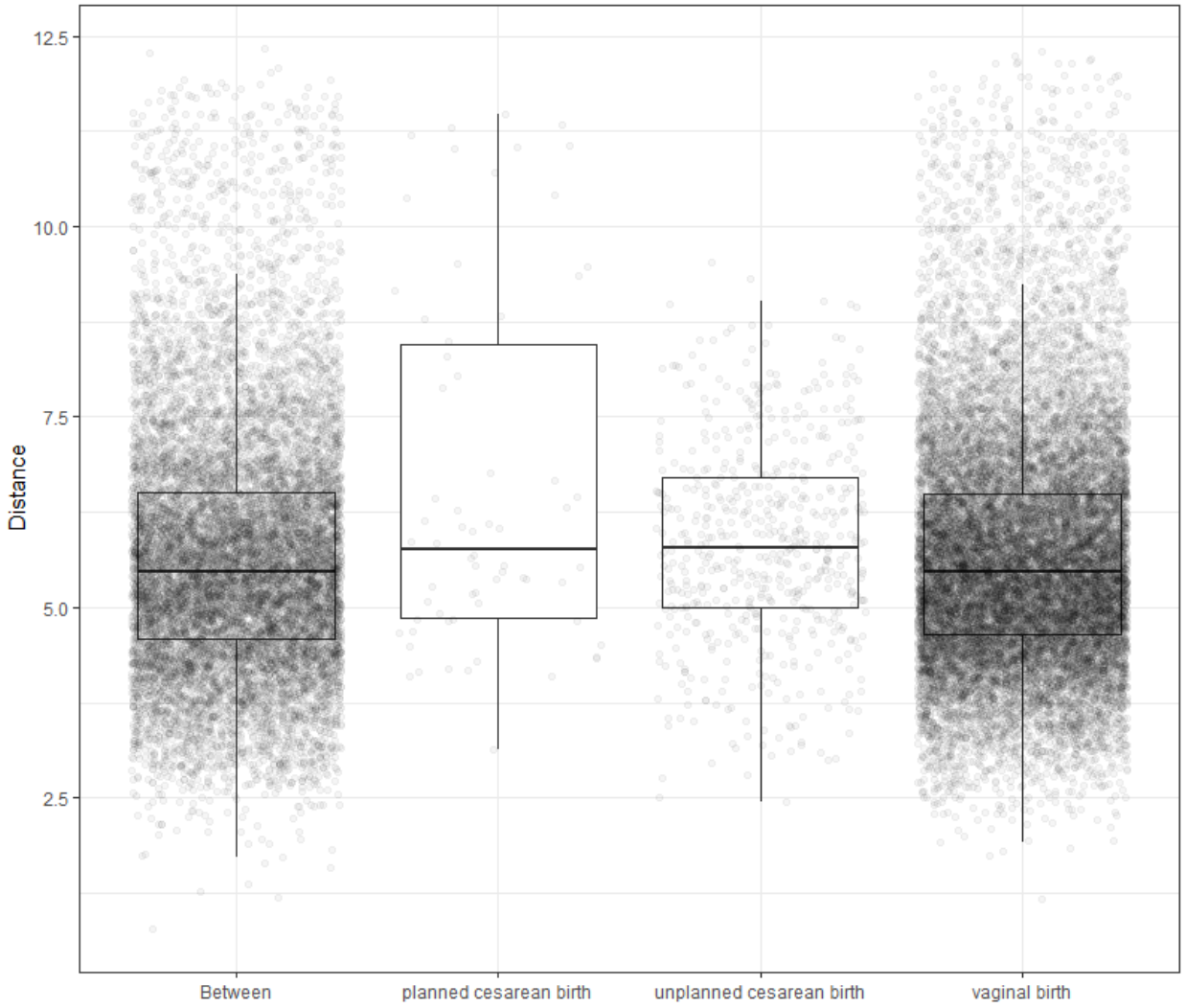


**Figure 21. Dissimilarity of microbial exposure in cesarean section delivered neonates.** Intra-sample distances for between all samples and those within cesarean section and vaginal deliveries.

D. Relationship between microbial exposures and umbilical cord serum PGLYRP-1

PGLYRP-1 was shown to be associated with several clinical risk factors for asthma and a major predictor of asthma at mid-childhood and pulmonary function at mid-childhood and early teenage years. PGLYRP-1 is believed to be produced predominately in neutrophils and stimulation with microbial products enhances its release. We hypothesized that the abundance of cfbDNA detected in umbilical cord blood is associated with the concentration of PGLYRP-1.

The concentration of PGLYRP-1 was assessed in relationship to copy number of 16s rRNA amplicons. Copy Number was significantly associated with the concentration of serum PGLYRP-1 (β [95% CI] = 0.12 [0.06-0.17], p <0.001) (**Figure 22**). This association was driven by the abundance of Proteobacteria which was also significantly associated with serum concentration of PGLYRP-1 (β [95% CI] = 0.13 [0.08-0.18], p <0.001) (**Figure 23**). There were no significant associations between PGLYRP-1 and Actinobacteria, Firmicutes, or Bacteroidetes. Further, genera were correlated with the concentration PGLYRP-1, 6 of 9 genera with significant positive correlations ($p < 0.05$, spearman's correlation) were Proteobacteria: *Roseateles, Pseudomonas, Ralstonia, Delftia, Acinetobacter*, and *Cupriavidus*. Cluster analysis was performed genera using their spearman's correlation coefficients with PHA, Der f1, and Bla g2 CBMC stimulation proliferative and cytokine response, as well as serum concentration of PGLYRP-1 (**Figure 24**). A cluster of taxa consisting of *Roseateles, Pseudomonas, Cupriavidus*, and several other gram negative taxa was identified as having moderate positive correlations serum PGLYRP-1 and IFN-γ concentrations after Der f1 and Bla g2 stimulation. Additionally, this

cluster had negative correlations with lymphproliferation and IL-13 concentrations after PHA, Der f1, and Bla g2 stimulation. This data is suggestive that microbial exposures at birth are influencing $T_{h1}$ and $T_{h2}$ skewing after foreign antigen exposure.

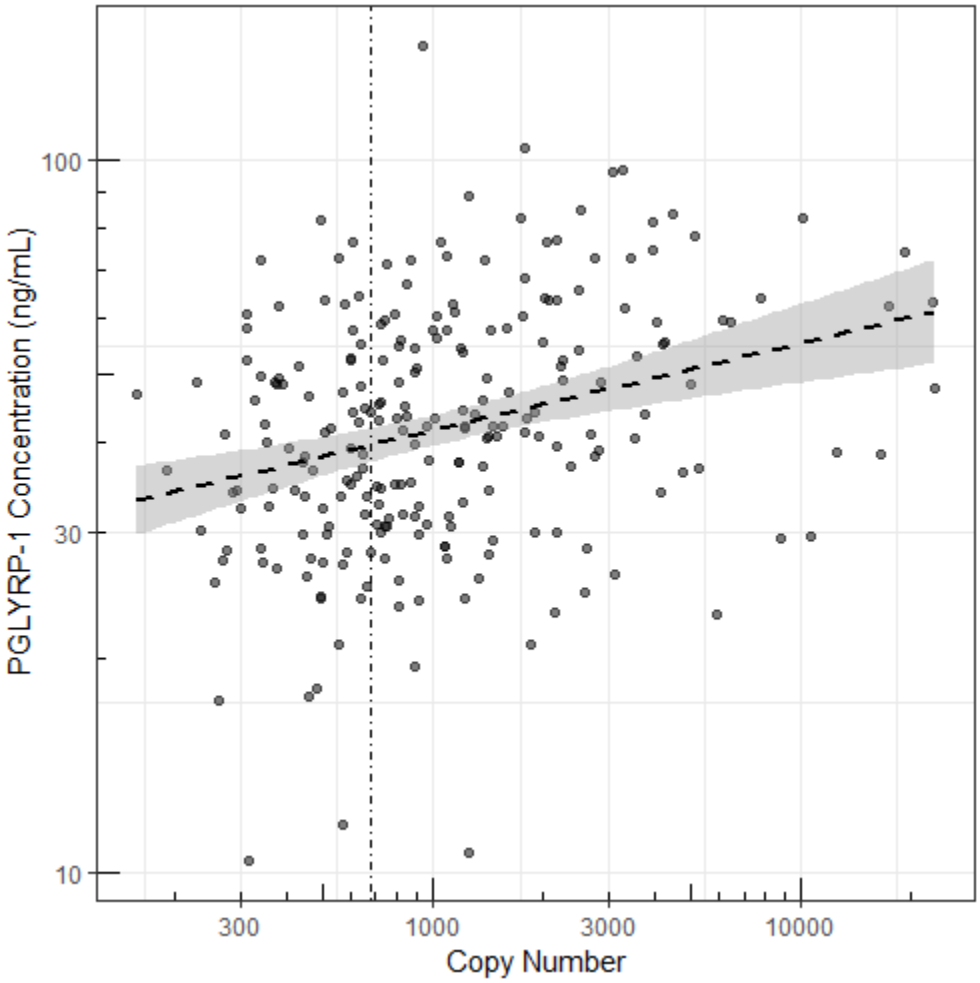


**Figure 22. Association between umbilical cord blood serum PGLYRP-1 concentration and estimated 16s rRNA copy number.** Scatterplot of estimated 16s rRNA copy number as predictor of PGLYRP-1 concentration in umbilical cord serum using linear regression. Axis are on log10 scales. Shaded grey is 95% confidence interval for estimate.

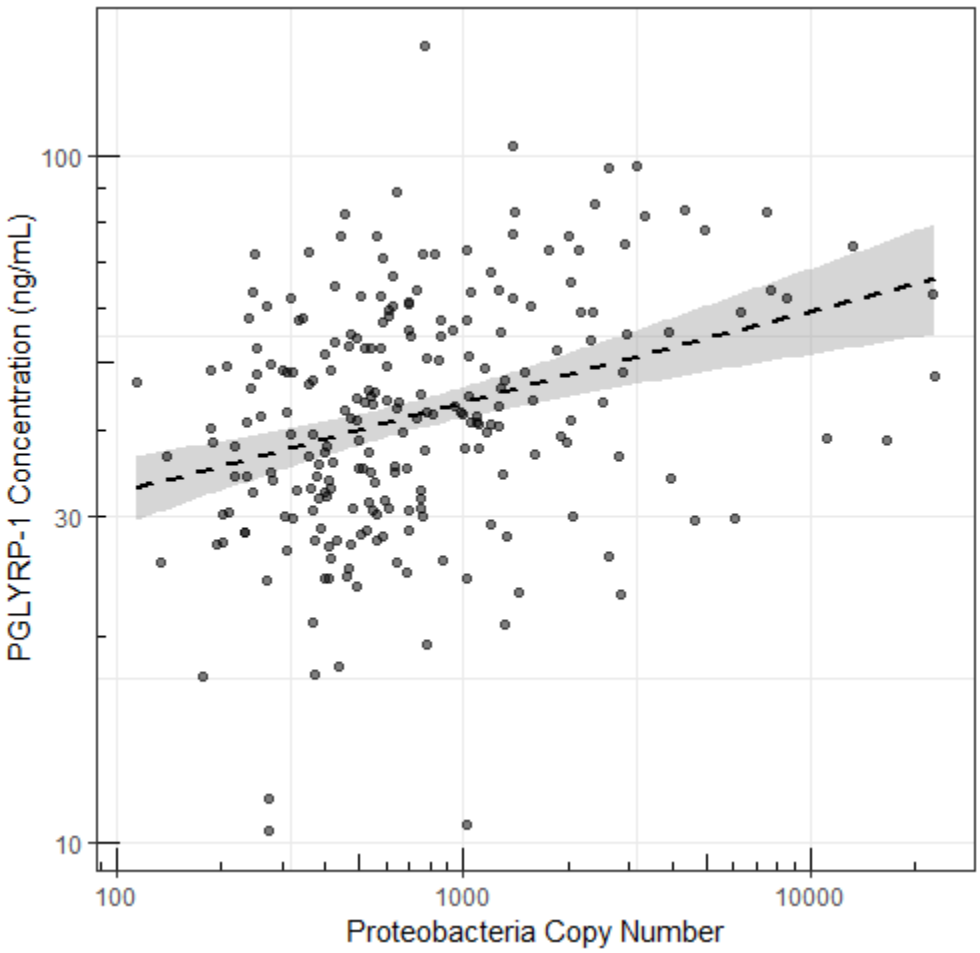


**Figure 23. Association between umbilical cord blood serum PGLYRP-1 concentration and sstimated Proteobacteria copy number.** Scatterplot of estimated Proteobacteria copy number as predictor of PGLYRP-1 concentration in umbilical cord serum using linear regression. Axis are on log10 scales. Shaded grey is 95% confidence interval for estimate.

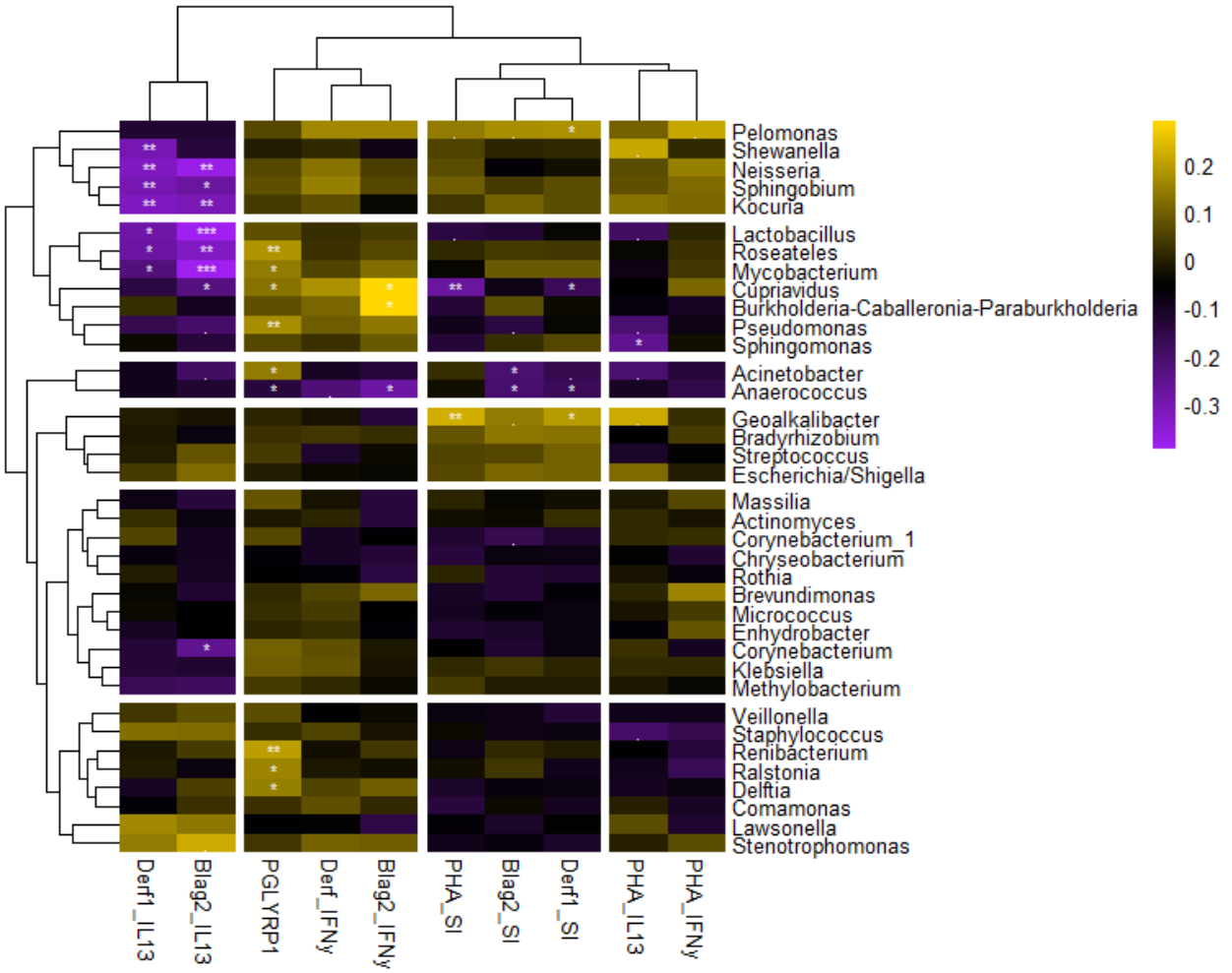


**Figure 24. Bacterial Taxa Associations with PGLYRP-1, CBMC Cytokine Production and Proliferation Post-Stimulation.** Heatmap displaying Spearman's correlation coefficients between bacterial Genera copy number and umbilical cord serum PGLYRP1 concentration, stimulation index (SI), IFNγ, or IL-13 post-stimulation with phytohaemmagglutinen (PHA), Der F1, or Bla G2. Cells are marked based on their p-values: . P < 0.1, * P < 0.05, ** P < 0.01, *** P < 0.001, Spearman's correlation. P-values are not exact due to ties.

# CHAPTER 6

# DISCUSSION

* This chapter's text, tables, and figures contain portions of work that is reprinted from:

**Turturice BA**, Ranjan R, Nguyen B, Hughes LM, Andropolis KE, Gold DR, Litonjua AA, Oken E, Perkins DL, Finn PW. Perinatal bacterial exposure contributes to IL-13 aeroallergen response. *Am J Respir Cell Mol Biol* 2017 Oct;57(4):419-427. doi: 10.1165/rcmb.2017-0027OC.

And

**Turturice BA**, Gold DR, Litonjua AA, Oken E, Rifas-Shiman S, Perkins DL, Finn PW. Lower perinatal exposure to Proteobacteria is an independent predictor of early childhood wheezing. *J Allergy Clin Immunol.* 2019 Jan;143(1):419-421.e5. doi: 10.1016/j.jaci.2018.06.051. Epub 2018 Sep 8.

And

Ozturk AB, **Turturice BA**, Perkins DL, Finn PW. The Potential for Emerging Microbiome-Mediated Therapeutics in Asthma. *Curr Allergy Asthma Rep.* 2017 Aug 10;17(9):62. doi: 10.1007/s11882-017-0730-1.

And

**Turturice BA**, Theorell J, Koenig MD, Tussing-Humphreys L, Gold DR, Litonjua AA, Oken E, Rifas-Shiman SL, Perkins DL, Finn PW. Perinatal granulopoiesis and risk of pediatric asthma. Elife. 2021 Feb 10;10:e63745. doi: 10.7554/eLife.63745. PMID: 33565964; PMCID: PMC7889076.

## A. Summary of Results

Recently there has been an increasing understanding of the impact of microbial exposures and myeloid cells on the development of many western diseases (e.g. asthma, inflammatory bowel disease and multiple sclerosis). Many studies have assessed differences that predate symptom development, particularly during the early childhood. There does appear to be differences in childhood, generally increased environmental exposure gram negative microbes as well as endogenous exposure to bacteria with more immunogenic LPS. These exposures are hypothesized to produce increased innate

imunnity through the production of antimicrobial peptides, increased neutrophils, and $T_{h1}$ responses.

Interestingly, very few studies have assessed microbial exposures before or around the time birth. In this study, microbial exposures and immunological differences associated with asthma risk and asthma risk factors were detected. Increased exposure to compositions high in Proteobacteria were associated with reduced probability of $T_{h2}$ responses to common aeroallergens in CBMCs. Using high-throughput transcriptomic techniques, increased development of early myeloid cells and specifically the concentration of PGLYRP-1 in umbilical cord blood serum was associated with reduced risk of asthma. Additionally, increased quantitative differences in Proteobacteria genera were associated with increased concentration of PGLYRP-1 and $T_{h1}$ responses to common aeroallergens. These findings indicate that the differences in microbial exposures and innate immunity that critical for asthma development can be identified at birth which is prior to what the majority of the current literature suggests.

B. Exposure to Unique Microbial Composition and Risk of $T_{h2}$ responses to allergens and childhood wheezing

We have demonstrated that a diverse array of bacterial DNA is present within the perinatal circulation. The bacterial DNA is derived from four main phyla: Actinobacteria, Bacteroidetes, Firmicutes, and Proteobacteria. Diversity of the taxonomy from which these sequences were derived was related to whether or not CBMCs produced IL-13 in response to Der f1 or Bla g2. Additionally, the ratio of Acinetobacter to Proteobacteria DNA is predictive of IL-13 responses to common aeroallergens. Along with in vitro IL-13

production in response to cytokines, IL-13 derived from serum significantly correlated with the diversity of bacterial DNA in serum.

IL-13, an IgE-inducing cytokine, is regulated by environmental factors such as microbial exposure at later time periods in childhood (139). Even as early as the perinatal time period, exposure to microorganisms and their products may be important for generating $T_{h2}$ responses to foreign antigens, which is important for class switching to IgE. Allergy to HDM aeroallergens is found in up to 85% of individuals with asthma (111). In addition to the high prevalence of IgE due to HDM allergens in asthmatics, children with higher levels of IgE due to HDM allergens are more likely to be susceptible to respiratory viruses and the development of wheezing (113). Asthmatics have a differential humoral response to bacterial antigens, specifically those that are induced by $T_{h2}$ cytokines (e.g., IgE), as well as IgG1 and IgG4 (119, 140, 141). We speculate that detection of bacterial DNA may represent exposure to specific bacteria.

Increased HDM-induced IL-13 production was strongly correlated with the abundance of Acinetobacter relative to Proteobacteria within a sample. Interestingly, in a study by Fyhrquist and colleagues involving nonatopic individuals (age 13–20 yr), exposure to Acinetobacter was associated with increased production of IL-10 and $T_{h1}$ cytokines (142). In the same study, it was observed that in those with high IgE levels to inhaled allergens, exposure and colonization with Acinetobacter was not associated with IL-10 and immunoregulatory gene expression. In atopic individuals, Acinetobacter may be inversely correlated with IL-10 production and expression (65, 142). Our model of the IL-13 response to aeroallergens predicts that an increasing representation of Acinetobacter within Proteobacteria enhances the likelihood and quantity of IL-13

production in response to Der f1 or Bla g2. Thus, both the quantity of Acinetobacter and the diversity of Proteobacteria are important for determining atopic-like responses to common aeroallergens. Our observation is similar to a previous finding that in the skin microbiome, Acinetobacter inversely correlates with IL-10 production, and there is reduced diversity within Gammaproteobacteria in atopic individuals (65).

There were several limitations to our study that may be addressed in future work. Future studies could include analyses of additional sites of interest (e.g., maternal serum and placental tissue) that could provide insight into origins of the bacterial DNA that we detected. Given our study's sample size (n = 27), we chose to focus our analyses on the most abundant bacterial families and phyla. Future work with a larger sample size may allow for further analysis of rarer bacterial families. Although multiple cytokines may be of interest in determining T cell phenotypes related to asthma, due to the limited sample volume, we focused on clinically relevant $T_{h2}$ (IL-4, -5, and -13) and $T_{h1}$ (IFN-γ) responses. Although we did not directly detect bacteria, the sequencing of DNA for indirect detection has been used in other situations, such as detection of fetal and tumor DNA, and bacterial DNA in serum may represent exposure to those bacteria at other sites on the body (143-145). Additionally, the detection of specific bacterial DNA within the circulation may be associated with specific disease states (130).

We demonstrate that the composition and diversity of bacterial DNA in perinatal cord blood serum correlates with the production of a $T_{h2}$ cytokine during allergen challenge to CBMCs and in serum. Notably, we observed stronger associations between bacterial DNA and IL-13 with an in vitro allergen challenge than in serum. This may be due to the many potential influences of T cells in vivo. The possible implications for health

outcomes remain to be determined. For example, we have shown a decrease in diversity and increase in Proteobacteria in nonresponders. There are differences in diversity and the abundance of Proteobacteria in bronchoalveolar lavages between asthmatics and healthy individuals (146-148). Our study adds to the growing picture that there is a relationship between how individuals respond to allergens and how they are colonized and respond to their own microbiota. Specific immune responses to allergens may be predictive of a potential colonization of bacteria later in life and susceptibility to asthma development. The importance of early-life exposure and interaction with bacteria as early predictors of asthma and atopy later in life merits further investigation.

C. Decreased Perinatal Innate Immunity and Susceptibility to Childhood Asthma

Recent findings have suggested that immunologic differences are apparent early in life—prior to the onset of symptoms-- between asthmatics and non-asthmatics (26, 46, 47, 119, 140). Our data suggests that the number of pre/perinatal risk factors is linearly related to one's odds of childhood asthma. Here we show that transcriptional changes associated with pre/perinatal risk are apparent in CBMCs and that PGLYRP-1 umbilical cord serological concentrations are predictive of childhood asthma. Umbilical cord serum PGLYRP-1, a neutrophil granule protein, concentration was strongly associated with this signature and significantly associated with odds of asthma at mid-childhood. Additionally, PGLYRP-1 was associated with FEV1/FVC in the early-teenage years and trended to an association at mid-childhood. These findings together suggest that impaired neutrophil development in early life results in greater risk for asthma in childhood.

We have identified differences involved in myeloid development as being related to asthma risk. The gene signature we identified was likely due to differences in neutrophil/ neutrophil precursor populations at birth. Notably, the key transcription factor in granule development (CEBPε), PADI4, (a gene shown to influence the transcription of CEBPε), and many neutrophil granule proteins were significantly associated with lower number of pre/perinatal risk factors (149-152). Additionally, during neutrophil development, there is a reduction in translational activity and nucleolar size (153). Genes involved in the high risk signature were enriched for GO terms involved in translation. Recent findings in meta-analysis of CBMC methylation studies identified a differentially methylated region upstream of RUNX1, a regulator of CEBPε expression during myelopoeis, as being associated with asthma development (46, 154). This suggest that the differentiation state of neutrophils is associated with early-life risk of asthma.

In replicating our meta-analysis, we used a smaller sample size from individuals living in an urban environment. There is potential that not all risk factors or their additive effects may have been adequately captured. Though the sample size for validation was small, PGLYRP1 would still have been identified as differentially expressed in regards to asthma risk status (p-value < 0.001, BH adj p-value < 0.1) in the UIH cohort. Notably, PGLYRP1 was associated with fetal sex (after false discovery rate correction), gestational age, and birthweight in the meta-analysis. These findings were confirmed by assessment of serum PGLYRP-1 which was found to associate with fetal sex and gestational age. Additionally, it was observed in three independent cohorts that mode of delivery also influences serum concentrations of PGLYRP-1. This result was not completely surprising as umbilical cord blood neutrophil abundance has been previously associated with fetal

sex, mode of delivery, and gestational age (43). It has been previously reported that suppressive granulocytic populations, LOX-1+/CD15+, vary with regards to abundance and function in relation to infant birthweight (155, 156). We observed LOX1 (OLR1) and LTF replicated between the meta-analysis and the mRNAseq data. Conversely, other genes associated with suppressive function in granulocytes, S100A9 and PTGS2 did not replicate and were only associated with asthma risk in the meta-analysis (157).

Importantly, through subset analysis, we identified individuals where serum PGLYRP-1 concentrations may not be predictive of asthma: small for gestational age, infants exposed to smoking early in utero, or children who did not identify as White, Black, or Hispanic. It has been shown that *in utero* exposure can change abundance and function of CBMCs (158, 159). Small for gestational age infants are more likely to experience neutropenia and infectious events (e.g. sepsis or necrotizing enterocolitis) (160, 161). Thus, functional differences in neutrophils may be explain the lack association between PGLYRP-1 and childhood asthma in these incidences.

It is not clear whether PGLYRP-1 has a direct casual role in asthma pathogenesis or if it is merely a biomarker for risk. PGLYRP-1 has antimicrobial function, although the concentration we observed in cord blood is far below those used for *in vitro* studies (162-164). PGLYRP-1 works synergistically in vitro with other antimicrobials (e.g. lysozyme), so its potential benefit as antimicrobial cannot be ruled out (163-165). Interestingly, mammalian PGLYRP-1 ortholog's appear divergent from ancestral PGLYRPs which contain enzymatic activity which is believed to be the mechanism of PGLYRPs antimicrobial mechanism (166). PGLYRP-1 appears to contain some functional activity as a cytokine as TREM-1 ligand even in the absence of peptidoglycan (167). Several

mouse studies investigating PGLYRP-1 null mice have reported conflicting results with regards to disease models. In asthma models null mice appear to be protected against inflammation whereas atopic dermatitis and inflammatory bowel disease models null mice seem to be more susceptible to disease (168-171). One potential confounder to the PGLYRP null models is there appears to be compensatory changes in expression of the other PGLYRPs when one of the other four is knocked out (168).

One limitation that could not be assessed in our study is the effect of environmental exposures, which are known to influence the risk of IL-13 response to allergens, wheezing in childhood, and asthma (26, 52, 53, 63, 65, 139, 142). The individuals from Project VIVA and UIH were recruited from Urban Centers. From the meta-analysis, it was not clear the exact recruitment locations for every study. Interestingly, we identified compositional differences in maternal microbiome that were associated with PGLYRP-1 concentration and the neutrophil development signature. We observed that increases in the Bacteroides genera were associated with gene signature representing the immature neutrophil state. As previously shown, microbiomes with increased Bacteroides are more susceptible to allergy and mice experience immunomodulation in response to changes in Bacteroides abundance (69, 172-174). Additionally, Clostridial species were associated with serum concentrations of PGLYRP-1. Thus, there may be different factors regulating neutrophil development and secretion of PGLYRP-1 that are environmentally controlled. It has been observed that children growing up in environments with higher levels of endotoxin have increased percentages of neutrophils and reduced prevelance of childhood asthma (26). Thus, future work will need to assess environmental impacts, including the differences in

microbiome, on umbilical cord serum concentrations of PGLYRP-1 and early neutrophil development.

In conclusion, we have identified a neutrophil development gene signature that is associated with pre/perinatal asthma risk. A soluble granule protein, PGLYRP-1, was strongly associated with odds of asthma in childhood and pulmonary function in childhood and early-teenage years. These differences were identified at birth. This indicates that in utero/ perinatal granulopoiesis plays a major role in the development of childhood asthma.

D. Microbial exposure and the induction of innate immunity and $T_{h1}$ responses

The findings in recent years suggest that microbial exposures early in life as important regulator of immunity (61, 64, 126). The most extreme of these findings is the number of the immune-deficiencies observed in germ free animals (175). The reductionist approaches of adding bacteria back into germ free mice have allowed for assessing individual bacterial species on the immune system (176). Yet, the microbial systems that colonize the surface of mucosal sites and the environments we are exposed to a complex mixtures of microbes (177).

To assess individual components of complex mixtures, a spike-in technique was adapted and utilized to determine microbial exposures at birth (76). The advantage of this technique is that by assessing individual components as log-ratios of known quantities allows for removal compositional effects and estimation of copy number of bacterial fragments that an individual was exposed to at birth. When assessing compositional exposures, there we no observed clustering of compositions. It was observed that for most clinical risk factors for asthma there was increased with-in group diversity, indicating

that there might be dysbiotic state associated with risk but it is not represented by a specific compositional pattern. When assessing similarities, vaginally delivered neonates were exposed to a more similar microbial pattern than those who were delivered via cesarean section.

When determining the relationship of specific microbial genera with immune responses to allergens and concentration of PGLYRP-1, Proteobacteria and several genera of Proteobacteria were positively correlated with increased $T_{h1}$ responses to allergens and PGLYRP-1 concentration. The neonatal immune system is polarized to produce predominately Th2 responses and the switch towards $T_{h1}$ responses is believed to be driven by exposure to microbes (39-41, 48, 64, 125). Increased exposure to Proteobacteria was associated with $T_{h1}$ responses in the Project VIVA cohort. Proteobacteria contain the most immunogenic LPS that is usually hexa- or penta-acylated and phosphorylated on both sugar groups which is thought to enhance its interactions with CD14 and MD2 increasing its signaling potential (68, 69). This is opposed to Bacteroidetes, also Gram negative, which have been reported to carry LPS with less than 5 acyl groups, are mono-phosphorylated, stimulate far less cytokine production and do not induce LPS tolerance (69). In addition, LPS stimulation of neutrophils and neutrophil precursors leads to release of granule contents. Proteobacteria was associated increased concentrations of PGLYRP-1 in umbilical cord serum. This may be due to increase granulopoeis or increased secretion of granule contents.

Proteobacteria may enhance the induction of Th1 and innate immunity. Increased production of Th1 cytokines in CBMCs is associated with reduced risk of lower respiratory illness (45). PGLYRP-1 concentration was associated with reduced risk of asthma.

Increased Proteobacteria compositionally was shown to be associated reduced risk of early wheezing. Thus increased Proteobacteria exposure at birth may enhance protection against respiratory illnesses which may enhance susceptibility to asthma.

E. Limitations and additional considerations

This thesis work identified novel relations between known risk factors for asthma, the neonatal immune system, and microbial exposures at birth that predictive of asthma in childhood. Although these findings represent a substantial increase in our understanding of the risk for childhood asthma, there are several limitations to these findings. First, the study design for all sampling was cross-sectional in nature with longitudinal follow-up in Project VIVA and no follow-up in other studies. Biological systems are dynamic and sampling a single time-point might not reflect the underlying dynamics of the system. To account for the sampling of single time point (birth), stimulation CBMCs with mitogens or allergens allows for assessment of these cells under different conditions. By assessing the microbial exposures, individuals that were studied were being assessed under different immunostimulatory conditions. Future work will need to address the long term effects immune outcomes in relation increased Proteobacteria exposure and PGLYRP-1 concentration at birth. There are some hints from Olin et al. where neonates were assessed at birth, 1, and 12 weeks. Importantly, increased cord myeloid populations were associated with reduced CD4+ and CD8+ EM T-cell abundance at birth, 1, and 12 weeks.

Due to the nature of working solely with human subjects and primary cells, there was limited potential to use reductionist techniques such as knockouts and over-

expression of genes. The reason for choosing to work with humans was that the study's focus was on the pathogenesis of disease whereas almost all murine models artificially induce asthma by the intratracheal installation of foreign antigens (26, 33, 64, 126, 170, 178-180). These murine models have been useful for studying the pathophysiology of asthma but have not been as fruitful about the origins of the disease. Additionally with concerns of microbial effects in mind, mice have been shown through fecal transplantation studies capable of harboring a fraction of the microbes found in the human intestine (181). Future work with longitudinal sampling of the microbiome and peripheral immune populations should allow for identification of causal relationships.

High throughput sampling of cfbDNA does not directly infer live microbial exposure. Circulating nucleic acids are sensitive to detecting even small biomasses of DNA such as lysis of tumor cells or tissue rejection but they do not impart spatial information on the origins of these nucleic acids (129, 131, 145). Thus, it is hard to infer more than that individual was exposed to a bacteria. Future studies may focus on more rigorous sampling of multiple mucosal sites where infants are exposed to microbes such as the maternal vaginal canal or neonatal meconium. Unfortunately there are limited options when trying to assess live bacterial exposure in high-throughput manner for multitude of reasons such not understanding culture conditions to host immune suppression of growth.

Last, it is important going forward to consider necessary sample size to determine effective power. Several issues plague power in human studies, including loss to follow-up, low prevalence of cases, and heterogeneous populations. In the Project VIVA cohort we estimated power curves for current asthma at mid-childhood and early-teen follow ups (**Figure 25**). No significant associations were determined at early-teen years and this is

believed to be due to low power at that time point. At early-teen follow up there are only 15 current asthmatics and 152 non-asthmatics (9% prevalence) whereas at mid-childhood there are 35 asthmatics and 171 non-asthmatics (17% prevalence). This lower prevalence at early-teen but not mid-childhood is noticeably different from the full cohort (19% and 15%, mid-childhood and early-teen prevalence). This likely due to high loss follow-up in the asthmatic group as 24 of the 35 (68%) where loss to follow up, whereas only 46 of 171 (27%) of non-asthmatics were lost to follow-up. This lead to a reduction in power at the early-teen follow-up time point. To address this issue future studies could include larger sample sizes to account for drop out or focus efforts on high risk individuals such as those with high number of risk factors or low serum concentrations of PGLYRP-1.

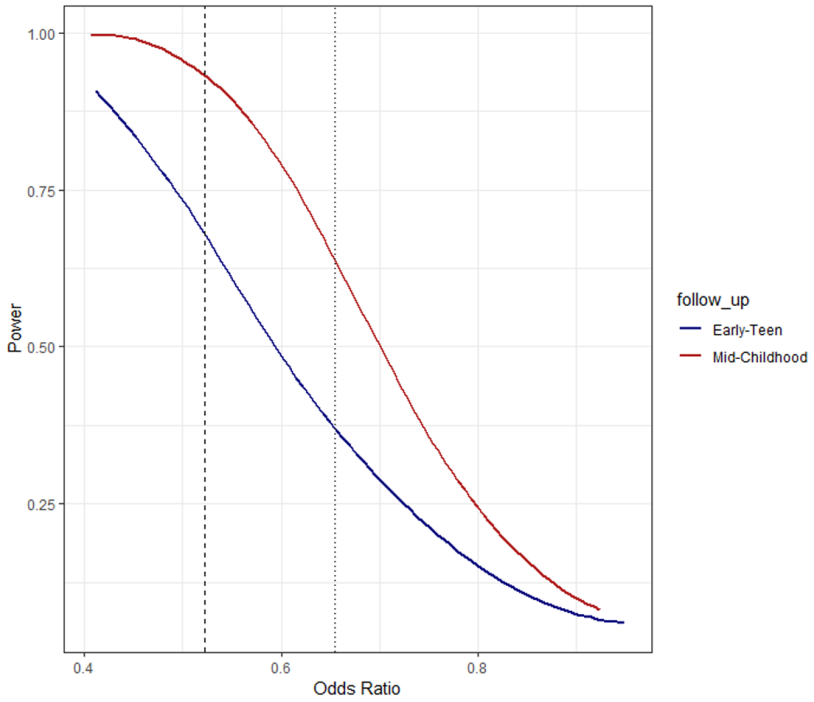


**Figure 25. Power calculation for Project VIVA current asthma sample sizes.** Loess Splines of average power – effect size relationships generated from 5000 random samples generated from a grid search of effect sizes 0.1 to 0.9 using logistic regression with sample sizes equivalent to those at mid-childhood (red) and early-teen (red) follow-ups. Dashed line is the effect size for mid-childhood and dotted line is the effect size for earl-teen follow-up.

F. Potential for therapeutic intervention in high risk children

Given data that supports the concept that the microbiome plays a fundamental role in asthma, manipulation of the microbiota with specific microbial products or dietary components may be a therapeutic strategy to enhance specific bacterial communities in the gut and/or lung (182). A change in the gut microbiota induced by such therapies may maintain a microbial balance that is a characteristic of healthy airways. Approaches vary from dietary administration of single bacterial strains (probiotic, prebiotic) to fecal transplantation. Further studies are needed to evaluate the impact of microbiota therapies on the airway microbiota in asthma.

Observational studies report an association between a high fat/low-fiber westernized diet and the prevalence of allergic disease, including asthma (183). On the other hand, the Mediterranean diet has been suggested as a healthy dietary pattern that may reduce the risk of asthma in both adults and children (184). The beneficial effect of breastfeeding in the development of fetal airways and the immune system is well-known (185). A recent systematic review provided evidence of negative associations between asthma/wheeze and dietary intake of vitamins C, E, and D, fruits, and adherence to a Mediterranean diet (184). In adults, increased intake of daily portions of fruits and vegetables was found to be associated with a reduction in markers of airway inflammation (184). The increased number of studies regarding microbiota and diet interactions in asthma development focuses on diet, a potential key factor in microbiota-based therapies. These studies are particularly carried out in pediatric populations and suggest that diet modulation might alter gut microbiota toward the treatment of asthma. Diet strongly influences the composition of the microbiome. Dietary fibers pass through the upper

intestine and are fermented by large-bowel anaerobic microbiota to produce certain short-chain fatty acids (SCFAs). SCFAs have some immune-modulating effects on the gut mucosa (186). Therefore, a diet-altered microbiome may lead to asthma and food allergies by the loss of SCFAs producing bacterial taxa in the gut. Murine studies support the therapeutic potential of a high-fiber diet which increases the number of bacteria that produce high levels of SCFAs (187). A reduced proportion of SCFAs has been documented in the fecal samples of allergic children, suggesting that this therapy may also be applicable for the treatment of atopic diseases and asthma (188). In a study by Wu et al., fecal communities were clustered into enterotypes distinguished by the levels of certain specific taxa. The enterotypes were strongly associated with protein and animal fat-containing diets. The microbiome composition exhibits detectably within 24 h of initiating a high-fat/low-fiber or low-fat/high-fiber diet (189). Prebiotics are non-digestible food ingredients that stimulate the growth and/or activity of one or a limited number of beneficial bacteria in the gut (190). There is limited data regarding the use of prebiotics in asthma. A recent Cochrane review evaluated the effect of prebiotic use in infants for the prevention of allergy (191). A meta-analysis of two studies found no significant difference in infant asthma, although significant heterogeneity was found between the studies. A meta-analysis of four studies found a significant reduction in eczema. Individual studies have reported a significant reduction in asthma and eczema with supplementation of prebiotics. Further studies are needed to clarify whether there is a protective role of prebiotics in asthma.

TLRs are the PRR members expressed in airway cells. Activation of some TLRs increases allergic sensitizations, while activation of other TLRs increases tolerance

toward aeroallergens (192). TLRs play roles in both innate immunity and adaptive immunity. Two TLRs, TLR2 and TLR4, have been implicated in allergic responses (192). TLR2 signaling may coordinate the Th1/Th2 balance (192). Murine asthma studies have shown that TLR2 agonists can act as either allergenic or anti-allergic agents based on the treatment schedule and materials (67, 178). TLR4 is the signaling receptor for bacterial lipopolysaccharide (LPS) and is a potent receptor for maturation of dendritic cells (67). TLR4 agonists can decrease and increase pulmonary allergic inflammation, in a similar manner to TLR2 agonists (67). Targeting of TLR4 to modify asthma and allergy is based on the use of TLR4 agonists as adjuvants in allergy vaccines. TLR4 agonists as adjuvants in allergy vaccines induce tolerance and can alter the course of respiratory allergy in either the upper or lower respiratory tract (193, 194). However, their effects were not tested alone in the previous studies. Intranasal administration of TLR4 agonists lead to a reduction in nasal symptoms in allergic rhinitis patients (195). Many synthetic TLR4 agonists have shown anti-asthmatic effects in murine models through the suppression of Th2 cytokines, airway hyperreactivity, airway inflammation, and airway eosinophilia (178, 180). Further clinical studies in humans are needed to define the therapeutic usefulness of TLR agonists in asthma.

## G. Conclusion

Early life events are critical to determining our susceptibly to wide variety of diseases. It is understood that early in infancy exposure to specific microbes can provide environmental ques to the immune system. These ques are necessary for the development of a functional immune system. Here we show that these microbial

exposures are relevant at birth and modulate the expression of a novel biomarker for childhood asthma. These findings suggest that prediction of high risk individuals for childhood asthma with serological biomarkers can be performed at birth which is earlier than most other currently used biomarkers.

## References


1. Grzelewska-Rzymowska I, Górski P. Asthma according to the GINA report 2014. Pediatria i Medycyna Rodzinna. 2015 Mar 31,;11(1):10-29.

2. Mazurek JM, Syamlal G. Prevalence of Asthma, Asthma Attacks, and Emergency Department Visits for Asthma Among Working Adults - National Health Interview Survey, 2011-2016. MMWR Morb Mortal Wkly Rep. 2018 Apr 06,;67(13):377-86.

3. Thomas A, Lemanske RF, Jackson DJ. Approaches to stepping up and stepping down care in asthmatic patients. J Allergy Clin Immunol. 2011 Nov;128(5):926.

4. Asthma, National Asthma Education and Prevention Program, Third Expert Panel on the Diagnosis and Management of. Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma. National Heart, Lung, and Blood Institute (US); 2007.

5. Tse, Sze Man, MDCM, MPH|Gold, Diane R., MD, MPH|Sordillo, Joanne E., ScD|Hoffman, Elaine B., PhD|Gillman, Matthew W., MD, SM|Rifas-Shiman, Sheryl L., MPH|Fuhlbrigge, Anne L., MD, MS|Tantisira, Kelan G., MD, MPH|Weiss, Scott T., MD, MS|Litonjua, Augusto A., MD, MPH. Diagnostic accuracy of the bronchodilator response in children. Journal of Allergy and Clinical Immunology, The. 2013;132(3):559.e5.

6. Most Recent Asthma Data | CDC [Internet].; 2019 [updated -02-04T04:48:03Z/; cited Feb 21, 2019]. Available from: https://www.cdc.gov/asthma/most_recent_data.htm.

7. Akinbami LJ, Simon AE, Rossen LM. Changing Trends in Asthma Prevalence Among Children. Pediatrics. 2016 Jan;137(1).

8. Bisgaard H, Bønnelykke K. Long-term studies of the natural history of asthma in childhood. J Allergy Clin Immunol. 2010 Aug;126(2):199.

9. Drake KA, Galanter JM, Burchard EG. Race, ethnicity and social class and the complex etiologies of asthma. Pharmacogenomics. 2008 Apr;9(4):453-62.

10. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. N Engl J Med. 2010 Sep 23,;363(13):1211-21.

11. Stein MM, Thompson EE, Schoettler N, Helling BA, Magnaye KM, Stanhope C, et al. A decade of research on the 17q12-21 asthma locus: Piecing together the puzzle. J Allergy Clin Immunol. 2018 Sep;142(3):764.e3.

12. Almoguera B, Vazquez L, Mentch F, Connolly J, Pacheco JA, Sundaresan AS, et al. Identification of Four Novel Loci in Asthma in European American and African American Populations. Am J Respir Crit Care Med. 2017 Feb 15,;195(4):456-63.

13. Galanter J, Choudhry S, Eng C, Nazario S, Rodríguez-Santana JR, Casal J, et al. ORMDL3 gene is associated with asthma in three ethnically diverse populations. Am J Respir Crit Care Med. 2008 Jun 01,;177(11):1194-200.

14. Sleiman PMA, Annaiah K, Imielinski M, Bradfield JP, Kim CE, Frackelton EC, et al. ORMDL3 variants associated with asthma susceptibility in North Americans of European ancestry. J Allergy Clin Immunol. 2008 Dec;122(6):1225-7.

15. Bouzigon E, Corda E, Aschard H, Dizier M, Boland A, Bousquet J, et al. Effect of 17q21 variants and smoking exposure in early-onset asthma. N Engl J Med. 2008 Nov 06,;359(19):1985-94.

16. Bräuner EV, Loft S, Raaschou-Nielsen O, Vogel U, Andersen PS, Sørensen M. Effects of a 17q21 chromosome gene variant, tobacco smoke and furred pets on infant wheeze. Genes Immun. 2012 Jan;13(1):94-7.

17. Calışkan M, Bochkov YA, Kreiner-Møller E, Bønnelykke K, Stein MM, Du G, et al. Rhinovirus wheezing illness and genetic risk of childhood-onset asthma. N Engl J Med. 2013 Apr 11,;368(15):1398-407.

18. Loss GJ, Depner M, Hose AJ, Genuneit J, Karvonen AM, Hyvärinen A, et al. The Early Development of Wheeze. Environmental Determinants and Genetic Susceptibility at 17q21. Am J Respir Crit Care Med. 2016 Apr 15,;193(8):889-97.

19. Jaakkola JJK, Ahmed P, Ieromnimon A, Goepfert P, Laiou E, Quansah R, et al. Preterm delivery and asthma: a systematic review and meta-analysis. J Allergy Clin Immunol. 2006 Oct;118(4):823-30.

20. Mu M, Ye S, Bai M, Liu G, Tong Y, Wang S, et al. Birth weight and subsequent risk of asthma: a systematic review and meta-analysis. Heart Lung Circ. 2014 Jun;23(6):511-9.

21. Thavagnanam S, Fleming J, Bromley A, Shields MD, Cardwell CR. A meta-analysis of the association between Caesarean section and childhood asthma. Clin Exp Allergy. 2008 Apr;38(4):629-33.

22. Xu X, Li Y, Sheng Y, Liu J, Tang L, Chen Z. Effect of low birth weight on childhood asthma: a meta-analysis. BMC Pediatr. 2014 Oct 23,;14:275.

23. Ciaccio CE, Gentile D. Effects of tobacco smoke exposure in childhood on atopic diseases. Curr Allergy Asthma Rep. 2013 Dec;13(6):687-92.

24. Guarnieri M, Balmes JR. Outdoor air pollution and asthma. Lancet. 2014 May 03,;383(9928):1581-92.

25. Oliver BGG, Robinson P, Peters M, Black J. Viral infections and asthma: an inflammatory interface? Eur Respir J. 2014 Dec;44(6):1666-81.

26. Stein MM, Hrusch CL, Gozdz J, Igartua C, Pivniouk V, Murray SE, et al. Innate Immunity and Asthma Risk in Amish and Hutterite Farm Children. N Engl J Med. 2016 Aug 04,;375(5):411-21.

27. Oddy WH. Breastfeeding, Childhood Asthma, and Allergic Disease. Ann Nutr Metab. 2017;70 Suppl 2:26-36.

28. Ober C, Sperling AI, von Mutius E, Vercelli D. Immune development and environment: lessons from Amish and Hutterite children. Curr Opin Immunol. 2017 Oct;48:51-60.

29. Sears MR. Predicting asthma outcomes. J Allergy Clin Immunol. 2015 Oct;136(4):836; quiz 837.

30. Amin P, Levin L, Epstein T, Ryan P, LeMasters G, Khurana Hershey G, et al. Optimum predictors of childhood asthma: persistent wheeze or the Asthma Predictive Index? J Allergy Clin Immunol Pract. 2014 Nov-Dec;2(6):709-15.

31. Smit HA, Pinart M, Antó JM, Keil T, Bousquet J, Carlsen KH, et al. Childhood asthma prediction models: a systematic review. Lancet Respir Med. 2015 Dec;3(12):973-84.

32. Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. Nat Med. 2012 May 04,;18(5):716-25.

33. Saglani S, Gregory LG, Manghera AK, Branchett WJ, Uwadiae F, Entwistle LJ, et al. Inception of early-life allergen-induced airway hyperresponsiveness is reliant on IL-13+CD4+ T cells. Sci Immunol. 2018 Sep 07,;3(27).

34. Pepper AN, Renz H, Casale TB, Garn H. Biologic Therapy and Novel Molecular Targets of Severe Asthma. J Allergy Clin Immunol Pract. 2017 Jul - Aug;5(4):909-16.

35. Chipps BE, Lanier B, Milgrom H, Deschildre A, Hedlin G, Szefler SJ, et al. Omalizumab in children with uncontrolled allergic asthma: Review of clinical trial and real-world experience. J Allergy Clin Immunol. 2017 May;139(5):1431-44.

36. Schaub B, Tantisira KG, Gibbons FK, He H, Litonjua AA, Gillman MW, et al. Fetal cord blood: aspects of heightened immune responses. J Clin Immunol. 2005 Jul;25(4):329-37.

37. Willwerth BM, Schaub B, Tantisira KG, Gold DR, Palmer LJ, Litonjua AA, et al. Prenatal, perinatal, and heritable influences on cord blood immune responses. Ann Allergy Asthma Immunol. 2006 Mar;96(3):445-53.

38. Pahal GS, Jauniaux E, Kinnon C, Thrasher AJ, Rodeck CH. Normal development of human fetal hematopoiesis between eight and seventeen weeks' gestation. Am J Obstet Gynecol. 2000 Oct;183(4):1029-34.

39. Krishnan L, Guilbert LJ, Russell AS, Wegmann TG, Mosmann TR, Belosevic M. Pregnancy impairs resistance of C57BL/6 mice to Leishmania major infection and causes decreased antigen-specific IFN-gamma response and increased production of T helper 2 cytokines. J Immunol. 1996 Jan 15,;156(2):644-52.

40. Lin H, Mosmann TR, Guilbert L, Tuntipopipat S, Wegmann TG. Synthesis of T helper 2-type cytokines at the maternal-fetal interface. J Immunol. 1993 Nov 01,;151(9):4562-73.

41. Wegmann TG, Lin H, Guilbert L, Mosmann TR. Bidirectional cytokine interactions in the maternal-fetal relationship: is successful pregnancy a TH2 phenomenon? Immunol Today. 1993 Jul;14(7):353-6.

42. Forestier F, Daffos F, Catherine N, Renard M, Andreux JP. Developmental hematopoiesis in normal human fetal blood. Blood. 1991 Jun 01,;77(11):2360-3.

43. Glasser L, Sutton N, Schmeling M, Machan JT. A comprehensive study of umbilical cord blood cell developmental changes and reference ranges by gestation, gender and mode of delivery. J Perinatol. 2015 Jul;35(7):469-75.

44. Dinarello CA. IL-18: A TH1-inducing, proinflammatory cytokine and new member of the IL-1 family. J Allergy Clin Immunol. 1999 Jan;103(1 Pt 1):11-24.

45. Ly NP, Rifas-Shiman SL, Litonjua AA, Tzianabos AO, Schaub B, Ruiz-Pérez B, et al. Cord blood cytokines and acute lower respiratory illnesses in the first year of life. Pediatrics. 2007 Jan;119(1):171.

46. Reese SE, Xu C, den Dekker HT, Lee MK, Sikdar S, Ruiz-Arenas C, et al. Epigenome-wide meta-analysis of DNA methylation and childhood asthma. J Allergy Clin Immunol. 2018 Dec 21,.

47. Xu C, Söderhäll C, Bustamante M, Baïz N, Gruzieva O, Gehring U, et al. DNA methylation in childhood asthma: an epigenome-wide meta-analysis. Lancet Respir Med. 2018 05;6(5):379-88.

48. Deshmukh HS, Liu Y, Menkiti OR, Mei J, Dai N, O'Leary CE, et al. The microbiota regulates neutrophil homeostasis and host resistance to Escherichia coli K1 sepsis in neonatal mice. Nat Med. 2014 May;20(5):524-30.

49. Strachan DP. Hay fever, hygiene, and household size. BMJ. 1989 Nov 18,;299(6710):1259-60.

50. Bach J. The effect of infections on susceptibility to autoimmune and allergic diseases. N Engl J Med. 2002 Sep 19,;347(12):911-20.

51. Abraham JH, Finn PW, Milton DK, Ryan LM, Perkins DL, Gold DR. Infant home endotoxin is associated with reduced allergen-stimulated lymphocyte proliferation and IL-13 production in childhood. J Allergy Clin Immunol. 2005 Aug;116(2):431-7.

52. Riedler J, Braun-Fahrländer C, Eder W, Schreuer M, Waser M, Maisch S, et al. Exposure to farming in early life and development of asthma and allergy: a cross-sectional survey. Lancet. 2001 Oct 06,;358(9288):1129-33.

53. Schuijs MJ, Willart MA, Vergote K, Gras D, Deswarte K, Ege MJ, et al. Farm dust and endotoxin protect against allergy through A20 induction in lung epithelial cells. Science. 2015 Sep 04,;349(6252):1106-10.

54. Bizzintino J, Lee W-, Laing IA, Vang F, Pappas T, Zhang G, et al. Association between human rhinovirus C and severity of acute asthma in children. Eur Respir J. 2011 May;37(5):1037-42.

55. Freymuth F, Vabret A, Brouard J, Toutain F, Verdon R, Petitjean J, et al. Detection of viral, Chlamydia pneumoniae and Mycoplasma pneumoniae infections in exacerbations of asthma in children. J Clin Virol. 1999 Aug;13(3):131-9.

56. Bäckhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, et al. Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. Cell Host Microbe. 2015 May 13,;17(5):690-703.

57. Structure, function and diversity of the healthy human microbiome. Nature. 2012 Jun 13,;486(7402):207-14.

58. Rajilić-Stojanović M, Smidt H, de Vos WM. Diversity of the human gastrointestinal tract microbiota revisited. Environ Microbiol. 2007 Sep;9(9):2125-36.

59. Fujimura KE, Sitarik AR, Havstad S, Lin DL, Levan S, Fadrosh D, et al. Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. Nat Med. 2016 10;22(10):1187-91.

60. Gdalevich M, Mimouni D, Mimouni M. Breast-feeding and the risk of bronchial asthma in childhood: a systematic review with meta-analysis of prospective studies. J Pediatr. 2001 Aug;139(2):261-6.

61. Arrieta M, Stiemsma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutsch S, et al. Early infancy microbial and metabolic alterations affect risk of childhood asthma. Sci Transl Med. 2015 Sep 30,;7(307):307ra152.

62. Berni Canani R, Sangwan N, Stefka AT, Nocerino R, Paparo L, Aitoro R, et al. Lactobacillus rhamnosus GG-supplemented formula expands butyrate-producing bacterial strains in food allergic infants. ISME J. 2016 Mar;10(3):742-50.

63. Ege MJ, Mayer M, Normand A, Genuneit J, Cookson, William O. C. M., Braun-Fahrländer C, et al. Exposure to environmental microorganisms and childhood asthma. N Engl J Med. 2011 Feb 24,;364(8):701-9.

64. Conrad ML, Ferstl R, Teich R, Brand S, Blümer N, Yildirim AO, et al. Maternal TLR signaling is required for prenatal asthma protection by the nonpathogenic microbe Acinetobacter lwoffii F78. J Exp Med. 2009 Dec 21,;206(13):2869-77.

65. Hanski I, von Hertzen L, Fyhrquist N, Koskinen K, Torppa K, Laatikainen T, et al. Environmental biodiversity, human microbiota, and allergy are interrelated. Proc Natl Acad Sci U S A. 2012 May 22,;109(21):8334-9.

66. von Mutius E. The microbial environment and its influence on asthma prevention in early life. J Allergy Clin Immunol. 2016 Mar;137(3):680-9.

67. Velasco G, Campo M, Manrique OJ, Bellou A, He H, Arestides RSS, et al. Toll-like receptor 4 or 2 agonists decrease allergic inflammation. Am J Respir Cell Mol Biol. 2005 Mar;32(3):218-24.

68. Steimle A, Autenrieth IB, Frick J. Structure and function: Lipid A modifications in commensals and pathogens. Int J Med Microbiol. 2016 Aug;306(5):290-301.

69. Vatanen T, Kostic AD, d'Hennezel E, Siljander H, Franzosa EA, Yassour M, et al. Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans. Cell. 2016 May 05,;165(4):842-53.

70. Reynolds JM, Martinez GJ, Chung Y, Dong C. Toll-like receptor 4 signaling in T cells promotes autoimmune inflammation. Proc Natl Acad Sci U S A. 2012 Aug 07,;109(32):13064-9.

71. Zanin-Zhorov A, Tal-Lapidot G, Cahalon L, Cohen-Sfady M, Pevsner-Fischer M, Lider O, et al. Cutting edge: T cells respond to lipopolysaccharide innately via TLR4 signaling. J Immunol. 2007 Jul 01,;179(1):41-4.

72. Kolb-Mäurer A, Kämmerer U, Mäurer M, Gentschev I, Bröcker EB, Rieckmann P, et al. Production of IL-12 and IL-18 in human dendritic cells upon infection by Listeria monocytogenes. FEMS Immunol Med Microbiol. 2003 Apr 01,;35(3):255-62.

73. Stober D, Schirmbeck R, Reimann J. IL-12/IL-18-dependent IFN-gamma release by murine dendritic cells. J Immunol. 2001 Jul 15,;167(2):957-65.

74. Mavropoulos A, Sully G, Cope AP, Clark AR. Stabilization of IFN-gamma mRNA by MAPK p38 in IL-12- and IL-18-stimulated human NK cells. Blood. 2005 Jan 01,;105(1):282-8.

75. Oken E, Baccarelli AA, Gold DR, Kleinman KP, Litonjua AA, De Meo D, et al. Cohort profile: project viva. Int J Epidemiol. 2015 Feb;44(1):37-48.

76. Tourlousse DM, Yoshiike S, Ohashi A, Matsukura S, Noda N, Sekiguchi Y. Synthetic spike-in standards for high-throughput 16S rRNA gene amplicon sequencing. Nucleic Acids Res. 2017 02 28,;45(4):e23.

77. Ranjan R, Rani A, Metwally A, McGee HS, Perkins DL. Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. Biochem Biophys Res Commun. 2016 Jan 22,;469(4):967-77.

78. Gohl DM, Vangay P, Garbe J, MacLean A, Hauge A, Becker A, et al. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. Nat Biotechnol. 2016 09;34(9):942-9.

79. Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G, Parada A, et al. Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. mSystems. 2016 Jan-Feb;1(1).

80. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol. 2013 Sep;79(17):5112-20.

81. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010 May;7(5):335-6.

82. Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R. PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics. 2010 Jan 15,;26(2):266-7.

83. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol. 2006 Jul;72(7):5069-72.

84. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. 2012 Mar;6(3):610-8.

85. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010 Oct 01,;26(19):2460-1.

86. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics. 2011 Aug 15,;27(16):2194-200.

87. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol. 2005 Dec;71(12):8228-35.

88. Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS ONE. 2010 Mar 10,;5(3):e9490.

89. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016 07;13(7):581-3.

90. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013 Jan;41(Database issue):590.

91. Namiki T, Hachiya T, Tanaka H, Sakakibara Y. MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. Nucleic Acids Res. 2012 Nov 01,;40(20):e155.

92. Metwally AA, Dai Y, Finn PW, Perkins DL. WEVOTE: Weighted Voting Taxonomic Identification Method of Microbial Sequences. PLoS ONE. 2016;11(9):e0163527.

93. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990 Oct 05,;215(3):403-10.

94. Ounit R, Wanamaker S, Close TJ, Lonardi S. CLARK: fast and accurate classification of metagenomic and genomic sequences using discriminative k-mers. BMC Genomics. 2015 Mar 25,;16:236.

95. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 2014 Mar 03,;15(3):R46.

96. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017 Apr;14(4):417-9.

97. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res. 2015;4:1521.

98. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

99. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. vegan: Community Ecology Package. 2019 -02-04.

100. Multiple imputation for nonresponse in surveys. Computers, Environment and Urban Systems. 1990 Jan;14(1):75.

101. Honaker J, King G, Blackwell M. Amelia II: A Program for Missing Data. Journal of Statistical Software. 2011 /12/12;45(1):1-47.

102. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. 2007 Jul 15,;23(14):1846-7.

103. GeneMeta: MetaAnalysis for High Throughput Experiments [Internet].; 2018 [cited Feb 21, 2019]. Available from: http://bioconductor.org/packages/GeneMeta/.

104. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005 Oct 25,;102(43):15545-50.

105. details, See AUTHORS file igraph author. igraph: Network Analysis and Visualization. 2019 -02-13.

106. Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol. 2007;8(9):R183.

107. Hu JX, Zhao H, Zhou HH. False Discovery Rate Control With Groups. J Am Stat Assoc. 2010 Sep 01,;105(491):1215-27.

108. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20,;43(7):e47.

109. Tepas EC, Litonjua AA, Celedón JC, Sredl D, Gold DR. Sensitization to aeroallergens and airway hyperresponsiveness at 7 years of age. Chest. 2006 Jun;129(6):1500-8.

110. Ulrik CS, Backer V. Markers of impaired growth of pulmonary function in children and adolescents. Am J Respir Crit Care Med. 1999 Jul;160(1):40-4.

111. Dust mite allergens and asthma--a worldwide problem. J Allergy Clin Immunol. 1989 Feb;83(2 Pt 1):416-27.

112. Peat JK, Tovey E, Mellis CM, Leeder SR, Woolcock AJ. Importance of house dust mite and Alternaria allergens in childhood asthma: an epidemiological study in two climatic regions of Australia. Clin Exp Allergy. 1993 Oct;23(10):812-20.

113. Soto-Quiros M, Avila L, Platts-Mills TAE, Hunt JF, Erdman DD, Carper H, et al. High titers of IgE antibody to dust mite allergen and risk for wheezing among asthmatic children infected with rhinovirus. J Allergy Clin Immunol. 2012 Jun;129(6):1505.e5.

114. Gold DR, Willwerth BM, Tantisira KG, Finn PW, Schaub B, Perkins DL, et al. Associations of cord blood fatty acids with lymphocyte proliferation, IL-13, and IFN-gamma. J Allergy Clin Immunol. 2006 Apr;117(4):931-8.

115. Scirica CV, Gold DR, Ryan L, Abulkerim H, Celedón JC, Platts-Mills TAE, et al. Predictors of cord blood IgE levels in children at risk for asthma and atopy. J Allergy Clin Immunol. 2007 Jan;119(1):81-8.

116. Bisgaard H, Hermansen MN, Buchvald F, Loland L, Halkjaer LB, Bønnelykke K, et al. Childhood asthma after bacterial colonization of the airway in neonates. N Engl J Med. 2007 Oct 11,;357(15):1487-95.

117. Wopereis H, Oozeer R, Knipping K, Belzer C, Knol J. The first thousand days - intestinal microbiology of early life: establishing a symbiosis. Pediatr Allergy Immunol. 2014 Aug;25(5):428-38.

118. Teo SM, Mok D, Pham K, Kusel M, Serralha M, Troy N, et al. The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development. Cell Host Microbe. 2015 May 13,;17(5):704-15.

119. Hales BJ, Chai LY, Elliot CE, Pearce LJ, Zhang G, Heinrich TK, et al. Antibacterial antibody responses associated with the development of asthma in house dust mite-sensitised and non-sensitised children. Thorax. 2012 Apr;67(4):321-7.

120. Aagaard K, Ma J, Antony KM, Ganu R, Petrosino J, Versalovic J. The placenta harbors a unique microbiome. Sci Transl Med. 2014 May 21,;6(237):237ra65.

121. Amarasekara R, Jayasekara RW, Senanayake H, Dissanayake VHW. Microbiome of the placenta in pre-eclampsia supports the role of bacteria in the multifactorial cause of pre-eclampsia. J Obstet Gynaecol Res. 2015 May;41(5):662-9.

122. Antony KM, Ma J, Mitchell KB, Racusin DA, Versalovic J, Aagaard K. The preterm placental microbiome varies in association with excess maternal gestational weight gain. Obstet Gynecol. 2015;212(5):653. e16.

123. Doyle RM, Alber DG, Jones HE, Harris K, Fitzgerald F, Peebles D, et al. Term and preterm labour are associated with distinct microbial community structures in placental membranes which are independent of mode of delivery. Placenta. 2014 Dec;35(12):1099-101.

124. Fardini Y, Chung P, Dumm R, Joshi N, Han YW. Transmission of diverse oral bacteria to murine placenta: evidence for the oral microbiome as a potential source of intrauterine infection. Infect Immun. 2010 Apr;78(4):1789-96.

125. Gomez de Agüero M, Ganal-Vonarburg SC, Fuhrer T, Rupp S, Uchimura Y, Li H, et al. The maternal microbiota drives early postnatal innate immune development. Science. 2016 Mar 18,;351(6279):1296-302.

126. Thorburn AN, McKenzie CI, Shen S, Stanley D, Macia L, Mason LJ, et al. Evidence that asthma is a developmental origin disease influenced by maternal diet and bacterial metabolites. Nat Commun. 2015 Jun 23,;6:7320.

127. Nikkari S, McLaughlin IJ, Bi W, Dodge DE, Relman DA. Does blood of healthy subjects contain bacterial ribosomal DNA? J Clin Microbiol. 2001 May;39(5):1956-9.

128. Jiménez E, Fernández L, Marín ML, Martín R, Odriozola JM, Nueno-Palop C, et al. Isolation of commensal bacteria from umbilical cord blood of healthy neonates born by cesarean section. Curr Microbiol. 2005 Oct;51(4):270-4.

129. Bossola M, Sanguinetti M, Scribano D, Zuppi C, Giungi S, Luciani G, et al. Circulating bacterial-derived DNA fragments and markers of inflammation in chronic hemodialysis patients. Clin J Am Soc Nephrol. 2009 Feb;4(2):379-85.

130. Mancini N, Poloniato A, Ghidoli N, Carletti S, Fomasi M, Barera G, et al. Potential role of the detection of enterobacterial DNA in blood for the management of neonatal necrotizing enterocolitis. J Med Microbiol. 2012 Oct;61(Pt 10):1465-72.

131. Rykova EY, Morozkin ES, Ponomaryova AA, Loseva EM, Zaporozhchenko IA, Cherdyntseva NV, et al. Cell-free and cell-bound circulating nucleic acid complexes: mechanisms of generation, concentration and content. Expert Opin Biol Ther. 2012 Jun;12 Suppl 1:141.

132. Yang Y, Chen M, Yang B, Huang X, Zhang X, He L, et al. Use of 16S rRNA Gene-Targeted Group-Specific Primers for Real-Time PCR Analysis of Predominant Bacteria in Mouse Feces. Appl Environ Microbiol. 2015 Oct;81(19):6749-56.

133. Olin A, Henckel E, Chen Y, Lakshmikanth T, Pou C, Mikes J, et al. Stereotypic Immune System Development in Newborn Children. Cell. 2018 Aug 23,;174(5):1292.e14.

134. Adams RI, Bateman AC, Bik HM, Meadow JF. Microbiota of the indoor environment: a meta-analysis. Microbiome. 2015;3(1):49.

135. Jervis-Bardy J, Leong LE, Marri S, Smith RJ, Choo JM, Smith-Vaughan HC, et al. Deriving accurate microbiota profiles from human samples with low bacterial content through post-sequencing processing of Illumina MiSeq data. Microbiome. 2015;3(1):19.

136. Sinha R, Abnet CC, White O, Knight R, Huttenhower C. The microbiome quality control project: baseline study design and future directions. Genome Biol. 2015;16(1):276.

137. Tkacz A, Hortala M, Poole PS. Absolute quantitation of microbiota abundance in environmental samples. Microbiome. 2018 06 19,;6(1):110.

138. Gibbons SM, Duvallet C, Alm EJ. Correcting for batch effects in case-control microbiome studies. PLoS Comput Biol. 2018 04;14(4):e1006102.

139. Michel S, Busato F, Genuneit J, Pekkanen J, Dalphin J-, Riedler J, et al. Farm exposure and time trends in early childhood may influence DNA methylation in genes related to asthma and allergy. Allergy. 2013 Mar;68(3):355-64.

140. Hales BJ, Pearce LJ, Kusel MMH, Holt PG, Sly PD, Thomas WR. Differences in the antibody response to a mucosal bacterial antigen between allergic and non-allergic subjects. Thorax. 2008 Mar;63(3):221-7.

141. Hollams EM, Hales BJ, Bachert C, Huvenne W, Parsons F, de Klerk NH, et al. Th2-associated immunity to bacteria in teenagers and susceptibility to asthma. Eur Respir J. 2010 Sep;36(3):509-16.

142. Fyhrquist N, Ruokolainen L, Suomalainen A, Lehtimäki S, Veckman V, Vendelin J, et al. Acinetobacter species in the skin microbiota protect against allergic sensitization and inflammation. J Allergy Clin Immunol. 2014 Dec;134(6):1309.e11.

143. Chen XQ, Stroun M, Magnenat JL, Nicod LP, Kurt AM, Lyautey J, et al. Microsatellite alterations in plasma DNA of small cell lung cancer patients. Nat Med. 1996 Sep;2(9):1033-5.

144. Lau TK, Lo KW, Chan LY, Leung TY, Lo YM. Cell-free fetal deoxyribonucleic acid in maternal circulation as a marker of fetal-maternal hemorrhage in patients undergoing external cephalic version near term. Am J Obstet Gynecol. 2000 Sep;183(3):712-6.

145. Lo YM, Corbetta N, Chamberlain PF, Rai V, Sargent IL, Redman CW, et al. Presence of fetal DNA in maternal plasma and serum. Lancet. 1997 Aug 16,;350(9076):485-7.

146. Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, et al. Disordered microbial communities in asthmatic airways. PLoS ONE. 2010 Jan 05,;5(1):e8578.

147. Huang YJ, Nelson CE, Brodie EL, Desantis TZ, Baek MS, Liu J, et al. Airway microbiota and bronchial hyperresponsiveness in patients with suboptimally controlled asthma. J Allergy Clin Immunol. 2011 Feb;127(2):3.

148. Yu W, Yuan X, Xu X, Ding R, Pang L, Liu Y, et al. Reduced airway microbiota diversity is associated with elevated allergic respiratory inflammation. Ann Allergy Asthma Immunol. 2015 Jul;115(1):63-8.

149. Borregaard N, Sørensen OE, Theilgaard-Mönch K. Neutrophil granules: a library of innate immunity proteins. Trends Immunol. 2007 Aug;28(8):340-5.

150. Dziarski R, Platt KA, Gelius E, Steiner H, Gupta D. Defect in neutrophil killing and increased susceptibility to infection with nonpathogenic gram-positive bacteria in peptidoglycan recognition protein-S (PGRP-S)-deficient mice. Blood. 2003 Jul 15,;102(2):689-97.

151. Evrard M, Kwok IWH, Chong SZ, Teng KWW, Becht E, Chen J, et al. Developmental Analysis of Bone Marrow Neutrophils Reveals Populations Specialized in Expansion, Trafficking, and Effector Functions. Immunity. 2018 02 20,;48(2):379.e8.

152. Kolodziej S, Kuvardina ON, Oellerich T, Herglotz J, Backert I, Kohrs N, et al. PADI4 acts as a coactivator of Tal1 by counteracting repressive histone arginine methylation. Nat Commun. 2014 May 29,;5:3995.

153. Zhu Y, Gong K, Denholtz M, Chandra V, Kamps MP, Alber F, et al. Comprehensive characterization of neutrophil genome topology. Genes Dev. 2017 01 15,;31(2):141-53.

154. Ng KP, Hu Z, Ebrahem Q, Negrotto S, Lausen J, Saunthararajah Y. Runx1 deficiency permits granulocyte lineage commitment but impairs subsequent maturation. Oncogenesis. 2013 Nov 04,;2:e78.

155. He Y, Li X, Perego M, Nefedova Y, Kossenkov AV, Jensen EA, et al. Transitory presence of myeloid-derived suppressor cells in neonates is critical for control of inflammation. Nat Med. 2018 Feb;24(2):224-31.

156. Condamine T, Dominguez GA, Youn J, Kossenkov AV, Mony S, Alicea-Torres K, et al. Lectin-type oxidized LDL receptor-1 distinguishes population of human polymorphonuclear myeloid-derived suppressor cells in cancer patients. Sci Immunol. 2016 Aug;1(2).

157. Heinemann AS, Pirr S, Fehlhaber B, Mellinger L, Burgmann J, Busse M, et al. In neonates S100A8/S100A9 alarmins prevent the expansion of a specific inflammatory monocyte population promoting septic shock. FASEB J. 2017 03;31(3):1153-64.

158. Mercelina-Roumans PE, Breukers RB, Ubachs JM, van Wersch JW. Hematological variables in cord blood of neonates of smoking and nonsmoking mothers. J Clin Epidemiol. 1996 Apr;49(4):449-54.

159. Rua, Ednildes de Almeida Olympio, Porto ML, Ramos JPL, Nogueira BV, Meyrelles SS, Vasquez EC, et al. Effects of tobacco smoking during pregnancy on oxidative stress in the umbilical cord and mononuclear blood cells of neonates. J Biomed Sci. 2014 Dec 30,;21:105.

160. Christensen RD, Yoder BA, Baer VL, Snow GL, Butler A. Early-Onset Neutropenia in Small-for-Gestational-Age Infants. Pediatrics. 2015 Nov;136(5):1259.

161. Warner BB, Tarr PI. Necrotizing enterocolitis and preterm infant gut bacteria. Semin Fetal Neonatal Med. 2016 12;21(6):394-9.

162. Kashyap DR, Wang M, Liu L, Boons G, Gupta D, Dziarski R. Peptidoglycan recognition proteins kill bacteria by activating protein-sensing two-component systems. Nat Med. 2011 Jun;17(6):676-83.

163. Liu C, Gelius E, Liu G, Steiner H, Dziarski R. Mammalian peptidoglycan recognition protein binds peptidoglycan with high affinity, is expressed in neutrophils, and inhibits bacterial growth. J Biol Chem. 2000 Aug 11,;275(32):24490-9.

164. Wang M, Liu L, Wang S, Li X, Lu X, Gupta D, et al. Human peptidoglycan recognition proteins require zinc to kill both gram-positive and gram-negative bacteria and are synergistic with antibacterial peptides. J Immunol. 2007 Mar 01,;178(5):3116-25.

165. Liu C, Xu Z, Gupta D, Dziarski R. Peptidoglycan recognition proteins: a novel family of four human innate immunity pattern recognition molecules. J Biol Chem. 2001 Sep 14,;276(37):34686-94.

166. Dziarski R, Gupta D. The peptidoglycan recognition proteins (PGRPs). Genome Biol. 2006;7(8):232.

167. Read CB, Kuijper JL, Hjorth SA, Heipel MD, Tang X, Fleetwood AJ, et al. Cutting Edge: identification of neutrophil PGLYRP1 as a ligand for TREM-1. J Immunol. 2015 Feb 15,;194(4):1417-21.

168. Park SY, Gupta D, Kim CH, Dziarski R. Differential effects of peptidoglycan recognition proteins on experimental atopic and contact dermatitis mediated by Treg and Th17 cells. PLoS ONE. 2011;6(9):e24961.

169. Park SY, Gupta D, Hurwich R, Kim CH, Dziarski R. Peptidoglycan recognition protein Pglyrp2 protects mice from psoriasis-like skin inflammation by promoting regulatory T cells and limiting Th17 responses. J Immunol. 2011 Dec 01,;187(11):5813-23.

170. Park SY, Jing X, Gupta D, Dziarski R. Peptidoglycan recognition protein 1 enhances experimental asthma by promoting Th2 and Th17 and limiting regulatory T

cell and plasmacytoid dendritic cell responses. J Immunol. 2013 Apr 01,;190(7):3480-92.

171. Saha S, Jing X, Park SY, Wang S, Li X, Gupta D, et al. Peptidoglycan recognition proteins protect mice from experimental colitis by promoting normal gut flora and preventing induction of interferon-gamma. Cell Host Microbe. 2010 Aug 19,;8(2):147-62.

172. Lei YM, Chen L, Wang Y, Stefka AT, Molinero LL, Theriault B, et al. The composition of the microbiota modulates allograft rejection. J Clin Invest. 2016 07 01,;126(7):2736-44.

173. McIntosh CM, Chen L, Shaiber A, Eren AM, Alegre M. Gut microbes contribute to variation in solid organ transplant outcomes in mice. Microbiome. 2018 05 25,;6(1):96.

174. Wu Q, Turturice B, Wagner S, Huang Y, Gupta PK, Schott C, et al. Gut Microbiota Can Impact Chronic Murine Lung Allograft Rejection. Am J Respir Cell Mol Biol. 2019 Jan;60(1):131-4.

175. Round JL, Mazmanian SK. The gut microbiota shapes intestinal immune responses during health and disease. Nat Rev Immunol. 2009 May;9(5):313-23.

176. Geva-Zatorsky N, Sefik E, Kua L, Pasman L, Tan TG, Ortiz-Lopez A, et al. Mining the Human Gut Microbiota for Immunomodulatory Organisms. Cell. 2017 02 23,;168(5):943.e11.

177. Layeghifard M, Hwang DM, Guttman DS. Disentangling Interactions in the Microbiome: A Network Perspective. Trends Microbiol. 2017 03;25(3):217-28.

178. Redecke V, Häcker H, Datta SK, Fermin A, Pitha PM, Broide DH, et al. Cutting edge: activation of Toll-like receptor 2 induces a Th2 immune response and promotes experimental asthma. J Immunol. 2004 Mar 01,;172(5):2739-43.

179. Yao X, Gao M, Dai C, Meyer KS, Chen J, Keeran KJ, et al. Peptidoglycan recognition protein 1 promotes house dust mite-induced airway inflammation in mice. Am J Respir Cell Mol Biol. 2013 Dec;49(6):902-11.

180. Bortolatto J, Borducchi E, Rodriguez D, Keller AC, Faquim-Mauro E, Bortoluci KR, et al. Toll-like receptor 4 agonists adsorbed to aluminium hydroxide adjuvant attenuate ovalbumin-specific allergic airway disease: role of MyD88 adaptor molecule and interleukin-12/interferon-gamma axis. Clin Exp Allergy. 2008 Oct;38(10):1668-79.

181. Arrieta M, Walter J, Finlay BB. Human Microbiota-Associated Mice: A Model with Challenges. Cell Host Microbe. 2016 May 11,;19(5):575-8.

182. Huang YJ. Asthma microbiome studies and the potential for new therapeutic strategies. Curr Allergy Asthma Rep. 2013 Oct;13(5):453-61.

183. Ellwood P, Asher MI, García-Marcos L, Williams H, Keil U, Robertson C, et al. Do fast foods cause asthma, rhinoconjunctivitis and eczema? Global findings from the International Study of Asthma and Allergies in Childhood (ISAAC) phase three. Thorax. 2013 Apr;68(4):351-60.

184. Garcia-Larsen V, Del Giacco SR, Moreira A, Bonini M, Charles D, Reeves T, et al. Asthma and dietary intake: an overview of systematic reviews. Allergy. 2016 Apr;71(4):433-42.

185. Elliott L, Henderson J, Northstone K, Chiu GY, Dunson D, London SJ. Prospective study of breast-feeding in relation to wheeze, atopy, and bronchial hyperresponsiveness in the Avon Longitudinal Study of Parents and Children (ALSPAC). J Allergy Clin Immunol. 2008 Jul;122(1):3.

186. Feehley T, Stefka AT, Cao S, Nagler CR. Microbial regulation of allergic responses to food. Semin Immunopathol. 2012 Sep;34(5):671-88.

187. Paturi G, Butts C, Monro J, Nones K, Martell S, Butler R, et al. Cecal and colonic responses in rats fed 5 or 30% corn oil diets containing either 7.5% broccoli dietary fiber or microcrystalline cellulose. J Agric Food Chem. 2010 May 26,;58(10):6510-5.

188. Böttcher MF, Nordin EK, Sandin A, Midtvedt T, Björkstén B. Microflora-associated characteristics in faeces from allergic and nonallergic infants. Clin Exp Allergy. 2000 Nov;30(11):1590-6.

189. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science. 2011 Oct 07,;334(6052):105-8.

190. Bengmark S, Martindale R. Prebiotics and synbiotics in clinical medicine. Nutr Clin Pract. 2005 Apr;20(2):244-61.

191. Osborn DA, Sinn JK. Prebiotics in infants for prevention of allergy. Cochrane Database of Systematic Reviews. 2013(3).

192. Suarez CJ, Parker NJ, Finn PW. Innate immune mechanism in allergic asthma. Curr Allergy Asthma Rep. 2008 Sep;8(5):451-9.

193. DuBuske LM, Frew AJ, Horak F, Keith PK, Corrigan CJ, Aberer W, et al. Ultrashort-specific immunotherapy successfully treats seasonal allergic rhinoconjunctivitis to grass pollen. Allergy Asthma Proc. 2011 May-Jun;32(3):239-47.

194. Rosewich M, Schulze J, Eickmeier O, Telles T, Rose MA, Schubert R, et al. Tolerance induction after specific immunotherapy with pollen allergoids adjuvanted by monophosphoryl lipid A in children. Clin Exp Immunol. 2010 Jun;160(3):403-10.

195. Casale TB, Kessler J, Romero FA. Safety of the intranasal toll-like receptor 4 agonist CRX-675 in allergic rhinitis. Ann Allergy Asthma Immunol. 2006 Oct;97(4):454-6.

# APPENDIX

RESEARCH ARTICLE

# Perinatal granulopoiesis and risk of pediatric asthma

**Benjamin A Turturice[1,2]*, Juliana Theorell[2], Mary Dawn Koenig[3], Lisa Tussing-Humphreys[4], Diane R Gold[5,6], Augusto A Litonjua[7], Emily Oken[8], Sheryl L Rifas-Shiman[8], David L Perkins[9,10]*, Patricia W Finn[1,2,10]***

[1]Department of Microbiology and Immunology, University of Illinois, Chicago, United States; [2]Department of Medicine, Division of Pulmonary, Critical Care, Sleep, and Allergy, University of Illinois, Chicago, United States; [3]Department of Women, Children and Family Health Science, College of Nursing, University of Illinois, Chicago, United States; [4]Department of Medicine and Cancer Center, University of Illinois, Chicago, United States; [5]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, United States; [6]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, United States; [7]Division of Pulmonary Medicine, Department of Pediatrics, University of Rochester, Rochester, United States; [8]Division of Chronic Disease Research Across the Life Course, Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, United States; [9]Department of Medicine, Division of Nephrology, University of Illinois, Chicago, United States; [10]Department of Bioengineering, University of Illinois, Chicago, United States

**Abstract** There are perinatal characteristics, such as gestational age, reproducibly associated with the risk for pediatric asthma. Identification of biologic processes influenced by these characteristics could facilitate risk stratification or new therapeutic targets. We hypothesized that transcriptional changes associated with multiple epidemiologic risk factors would be mediators of pediatric asthma risk. Using publicly available transcriptomic data from cord blood mononuclear cells, transcription of genes involved in myeloid differentiation was observed to be inversely associated with a pediatric asthma risk stratification based on multiple perinatal risk factors. This gene signature was validated in an independent prospective cohort and was specifically associated with genes localizing to neutrophil-specific granules. Further validation demonstrated that umbilical cord blood serum concentration of PGLYRP-1, a specific granule protein, was inversely associated with mid-childhood current asthma and early-teen $FEV_1$/FVCx100. Thus, neutrophil-specific granule abundance at birth predicts risk for pediatric asthma and pulmonary function in adolescence.

***For correspondence:**
bturtu2@uic.edu (BAT);
perkinsd@uic.edu (DLP);
pwfinn@uic.edu (PWF)

**Competing interests:** The authors declare that no competing interests exist.

**Funding:** See page 19

**Received:** 05 October 2020
**Accepted:** 05 February 2021
**Published:** 10 February 2021

**Reviewing editor:** Mone Zaidi, Icahn School of Medicine at Mount Sinai, United States



## Introduction

Several risk factors for pediatric asthma can be ascertained in the perinatal period. These risk factors include maternal characteristics (e.g., maternal atopy, maternal body mass index [BMI], race/ethnicity), demographics (e.g., newborn sex), and birth characteristics (e.g., birthweight, gestational age at birth, mode of delivery) (***Bisgaard and Bønnelykke, 2010***). Meta-analyses have provided strong evidence for associations between the variables stated above and risk for pediatric asthma (***Jaakkola et al., 2006***; ***Mu et al., 2014***; ***Thavagnanam et al., 2008***; ***Xu et al., 2014***). Many of these risk factors co-occur (e.g., low birthweight and preterm birth), and it has yet to be discerned whether their imparted risk is mediated through similar biologic processes.

## Lower perinatal exposure to Proteobacteria is an independent predictor of early childhood wheezing

**Author:**
Benjamin A. Turturice,Diane R. Gold,Augusto A. Litonjua,Emily Oken,Sheryl Rifas-Shiman,David L. Perkins,Patricia W. Finn

**Publication:** Journal of Allergy and Clinical Immunology

**Publisher:** Elsevier

**Date:** January 2019

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Feb 23, 2021

---

This Agreement between Dr. Benjamin Turturice ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

| | |
|---|---|
| License Number | 5015140232951 |
| License date | Feb 23, 2021 |
| Licensed Content Publisher | Springer Nature |
| Licensed Content Publication | Current Allergy and Asthma Reports |
| Licensed Content Title | The Potential for Emerging Microbiome-Mediated Therapeutics in Asthma |
| Licensed Content Author | Ayse Bilge Ozturk et al |
| Licensed Content Date | Aug 10, 2017 |
| Type of Use | Thesis/Dissertation |
| Requestor type | academic/university or research institute |
| Format | print and electronic |
| Portion | full article/chapter |
| Will you be translating? | no |

| | |
|---|---|
| Circulation/distribution | 1 - 29 |
| Author of this Springer Nature content | yes |
| Title | Early Origins of Childhood Asthma: Role of Perinatal Microbial Exposure and Innate Immunity |
| Institution name | University of Illinois |
| Expected presentation date | Mar 2021 |
| Requestor Location | Dr. Benjamin Turturice<br>2635 W Rice St<br>#1<br><br>CHICAGO, IL 60622<br>United States<br>Attn: Dr. Benjamin Turturice |
| Total | 0.00 USD |

Terms and Conditions

**Springer Nature Customer Service Centre GmbH**
**Terms and Conditions**

This agreement sets out the terms and conditions of the licence (the **Licence**) between you and **Springer Nature Customer Service Centre GmbH** (the **Licensor**). By clicking 'accept' and completing the transaction for the material (**Licensed Material**), you also confirm your acceptance of these terms and conditions.

**1. Grant of License**

**1. 1.** The Licensor grants you a personal, non-exclusive, non-transferable, world-wide licence to reproduce the Licensed Material for the purpose specified in your order only. Licences are granted for the specific use requested in the order and for no other use, subject to the conditions below.

**1. 2.** The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of the Licensed Material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

**1. 3.** If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

## 2. Scope of Licence

**2. 1.** You may only use the Licensed Content in the manner and to the extent permitted by these Ts&Cs and any applicable laws.

**2. 2.** A separate licence may be required for any additional use of the Licensed Material, e.g. where a licence has been purchased for print only use, separate permission must be obtained for electronic re-use. Similarly, a licence is only valid in the language selected and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence. Any content owned by third parties are expressly excluded from the licence.

**2. 3.** Similarly, rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to Journalpermissions@springernature.com/bookpermissions@springernature.com for these rights.

**2. 4.** Where permission has been granted **free of charge** for material in print, permission may also be granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

**2. 5.** An alternative scope of licence may apply to signatories of the STM Permissions Guidelines, as amended from time to time.

## 3. Duration of Licence

**3. 1.** A licence for is valid from the date of purchase ('Licence Date') at the end of the relevant period in the below table:

| Scope of Licence | Duration of Licence |
|---|---|
| Post on a website | 12 months |
| Presentations | 12 months |
| Books and journals | Lifetime of the edition in the language purchased |

## 4. Acknowledgement

**4. 1.** The Licensor's permission must be acknowledged next to the Licenced Material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.

## 5. Restrictions on use

**5. 1.** Use of the Licensed Material may be permitted for incidental promotional use and minor editing privileges e.g. minor adaptations of single figures, changes of format, colour and/or style where the adaptation is credited as set out in Appendix 1 below. Any other changes including but not limited to, cropping, adapting, omitting material that affect the meaning, intention or moral rights of the author are strictly prohibited.

**5. 2.** You must not use any Licensed Material as part of any design or trademark.

**5. 3.** Licensed Material may be used in Open Access Publications (OAP) before publication by Springer Nature, but any Licensed Material must be removed from OAP sites prior to final publication.

**6. Ownership of Rights**

**6. 1.** Licensed Material remains the property of either Licensor or the relevant third party and any rights not explicitly granted herein are expressly reserved.

**7. Warranty**

IN NO EVENT SHALL LICENSOR BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL OR INDIRECT DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND
WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

**8. Limitations**

**8. 1.** ***BOOKS ONLY:*** Where **'reuse in a dissertation/thesis'** has been selected the following terms apply: Print rights of the final author's accepted manuscript (for clarity, NOT the published version) for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/).

**8. 2.** For content reuse requests that qualify for permission under the STM Permissions Guidelines, which may be updated from time to time, the STM Permissions Guidelines supersede the terms and conditions contained in this licence.

**9. Termination and Cancellation**

**9. 1.** Licences will expire after the period shown in Clause 3 (above).

**9. 2.** Licensee reserves the right to terminate the Licence in the event that payment is not received in full or if there has been a breach of this agreement by you.

**Appendix 1 — Acknowledgements:**

**For Journal Content:**
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

For **Advance Online Publication papers:**
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)

**For Adaptations/Translations:**
Adapted/Translated by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

**Note: For any republication from the British Journal of Cancer, the following credit line style applies:**

Reprinted/adapted/translated by permission from [**the Licensor**]: on behalf of Cancer Research UK: : [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT]** (year of publication)

For **Advance Online Publication** papers:
Reprinted by permission from The [**the Licensor**]: on behalf of Cancer Research UK: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM])

**For Book content:**
Reprinted/adapted by permission from [**the Licensor**]: [**Book Publisher** (e.g. Palgrave Macmillan, Springer etc) [**Book Title**] by [**Book author**(s)] [**COPYRIGHT**] (year of publication)

**Other Conditions**:

Version 1.3

**Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.**

UIC

**Benjamin Turturice <bturtu2@uic.edu>**

# Permission for Matrerial Reuse in Thesis

2 messages

**Benjamin Turturice** <bturtu2@uic.edu> Mon, Mar 1, 2021 at 3:05 PM
To: permissions@thoracic.org

- Article Title: Perinatal Bacterial Exposure Contributes to IL-13 Aeroallergen Response
- Author(s): **Benjamin A. Turturice,** Ravi Ranjan, Brian Nguyen, Lauren M. Hughes, Kalista E. Andropolis, Diane R. Gold, Augusto A. Litonjua, Emily Oken, David L. Perkins, and Patricia W. Finn
- Page Range: 419–427
- Volume Number: 57(4)
- Journal Title: Am J Respir Cell Mol Biol
- I am requesting to use the material of the full article, including texts, figures, and tables in my thesis dissertation.
- I am the lead author on the publication

--
Benjamin A. Turturice
MD, PhD Candidate

University of Illinois at Chicago
College of Medicine
Medical Scientist Training Program

440.829.8328
bturtu2@uic.edu

**ATS Permission Requests** <permissions@thoracic.org> Tue, Mar 2, 2021 at 8:19 AM
To: Benjamin Turturice <bturtu2@uic.edu>

Dear Dr. Turturice,

Thank you for your request. Because you are one of the original authors of the material, and this request is for thesis use, permission is granted at no charge. Please complete the below and use it beneath the material. Thank you.

Reprinted with permission of the American Thoracic Society.



Cite: Author(s)/Year/Title/Journal title/Volume/Pages.

The American Journal of Respiratory Cell and Molecular Biology is an official journal of the American Thoracic Society.

Please note that this permission is for one time use. Thank you, and please let me know if you have any questions.

Best regards,

Megan

Megan Murphy

*Production Coordinator*

American Thoracic Society

25 Broadway, 4th Floor

New York, NY 10004

212-315-8643

mmurphy@thoracic.org

**Please donate to the ATS to make COVID-19 resources accessible and support respiratory health professionals worldwide. Wishing you, your community, and all health care professionals safety during this difficult time.**


[Quoted text hidden]

Cite: Benjamin A. Turturice, Ravi Ranjan, Brian Nguyen, Lauren M. Hughes, Kalista E.Andropolis, Diane R. Gold, Augusto A. Litonjua, Emily Oken, David L. Perkins, and Patricia W. Finn/2017/Perinatal Bacterial Exposure Contributes to IL-13 Aeroallergen Response/ American Journal of Respiratory Cell and Molecular Biology/Volume/419–427.
The American Journal of Respiratory Cell and Molecular Biology is an official journal of the American Thoracic Society