THE UNIVERSITY OF CHICAGO

GENETIC AND EPIGENETIC MECHANISMS UNDERLYING MATERNAL TRANSMISSION OF ASTHMA RISK

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES

AND THE PRITZKER SCHOOL OF MEDICINE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF HUMAN GENETICS

BY

KEVIN MATTHEW MAGNAYE

CHICAGO, ILLINOIS

DECEMBER 2020

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my advisor, Carole Ober, for her incredible mentorship. I could not have selected a better mentor to guide me academically, scientifically, and personally. She has never dismissed my ideas or doubted my capabilities and has supported me in all of my endeavors. Most importantly, I came into graduate school with a strong passion for science and I am leaving with an even stronger passion for science, which I attribute all to Carole.

I would also like to thank the Ober lab for providing a collaborative and scientifically engaging environment that has allowed me to succeed these past five years. Thank you to past members, Jessie Nicodemus-Johnson, Dara Torgerson, Catherine Igartua and Michelle Stein, for their previous scientific explorations, which have inspired the work presented in this dissertation. Thank you to William Wentworth-Sheilds, Rebecca Anderson, Catherine Stanhope, Katie Naughton, Raluca Nicolae, Christine Billstrand, Emma Thompson and Matt Dapas for their help in generating, processing, and maintaining high-quality sequencing and phenotype data. Thank you to my fellow graduate students, Sahar Mozaffari, Marcus Soliai, Selene Clay, Liana Hernandez, Charles Washington and Mitchell Conery, for their friendship and support.

I would like to acknowledge the research groups who have shared their data and scientific expertise and the study participants for whom these studies would not be possible. Thank you to my thesis committee (Anna Di Rienzo, Dan Nicolae, and John Novembre) for supporting my training. Thank you to the Department of Human Genetics and the University of Chicago for providing a wonderful academic environment. Finally, thank you to my partner, my family, my friends, and to all of the personal and scientific adventures and the people I've met along the way for uplifting me and helping me close this fulfilling chapter of my life.

ABSTRACT

Asthma in the mother is one of the most reproducible risk factors for asthma in her child, however the underlying mechanisms are unknown. In this thesis, I identified and characterized genetic and epigenetic mechanisms underlying maternal transmission of asthma risk. First, I examined ADAR-mediated microRNA (miRNA) editing in bronchial epithelial cells (BECs) from 142 asthma cases and controls. I identified A-to-I editing of a miRNA, miR-200b-3p, that was significantly associated with moderate and severe, but not mild, asthma. I also found that the expression of the primary A-to-I editing gene, *ADAR*, was significantly decreased in asthma cases with an asthmatic mother. These results suggested that maternal asthma may alter the gene regulatory landscape in BECs at the *ADAR* locus *in utero*, with long-lasting effects on ADAR-mediated editing of miR-200b-3p and leading to more severe asthma in adulthood. Second, I identified genome-wide DNA methylation patterns in BECs from asthmatic children with an asthmatic mother that differed compared to asthmatic children without an asthmatic mother and to controls. Maternal asthma-associated DNA methylation signatures reflected different endotypes of asthma severity. Overall, I suggested that the maternal asthma *in utero* environment alters epigenetically-mediated developmental pathways in the lower airways that lead to subtypes of severe asthma in adulthood. Lastly, I tested for maternal asthma-associated mitochondrial DNA (mtDNA) variants and haplogroups in nine ethnically diverse cohorts. Associations between mtDNA variants and haplogroups with asthma were primarily in maternal asthma dyads and not in paternal asthma dyads. These results indicated that some of the risk for asthma in children of asthmatic mothers may be due to variation inherited from the maternally-inherited mtDNA genome.

CHAPTER 1

INTRODUCTION

1.1 Asthma – a complex, heterogeneous disease of the airways

Asthma is the most common chronic disease in childhood, affecting 5.5 million children and 19.2 million adults in the United States, with annual health care costs exceeding $80 billion (*1*). Its clinical diagnosis is based on symptoms of wheeze, cough, and chest tightness, which reflect the cardinal features of bronchial hyperresponsiveness and lower airway inflammation and remodeling (*2*). Although most individuals with asthma share these features, there is extensive clinical heterogeneity and inter-individual variation in age of onset, co-morbidities, and response to therapies. For example, nearly 15% of individuals with asthma have severe asthma, defined by uncontrollable symptoms, persistent exacerbations and poor response to standard asthma treatments (*3*).

Moreover, risks for asthma vary by age and sex. Among children under the age of 17 years, asthma is more prevalent among boys compared to girls (9.9% vs. 6.9%, respectively) (*4*). Because asthma symptoms often resolve in boys (*5*) and more girls develop asthma during and after puberty (*6*), asthma is more common in women than in men (*7*). Differences in immune responses in males and females have been observed in the first few years of life, suggesting a developmental component to the sex-biased disease risk in asthma (*8*).

Asthma prevalence also varies across ethnic groups. For example, in 2007, the prevalence of childhood asthma in non-Hispanic Blacks and Puerto Ricans was 12.7% and 19.2%, respectively, compared to 8.0% and 6.4% in non-Hispanic Whites and Mexican Americans, respectively (*9*). In fact, African ancestry per se was associated with asthma risk in African Americans (*10*) and with asthma and total serum IgE (a marker of allergic sensitization) in

Caribbean, North American, and South American populations with variable levels of African ancestry (*11-13*). By leveraging the differences in risk for asthma, I can potentially gain greater insight into disparities in asthma risk, identify disease endotypes and understand shared and distinct biological pathways. In this thesis, I will explore differences in asthma risk by disease severity, atopy (allergic sensitization), and ethnicity.

1.2 Early life origins of asthma

Accumulating evidence suggests that the trajectory to asthma begins early in life. The fetus is exposed to a number of intrauterine challenges that must be overcome for sustained growth and development (*14*), and the first few years of life represents a critical window for susceptibility to environmental exposures associated with asthma risk or protection. While the fetal immune system must elicit an immune response to protect against harmful environmental stimuli, immunological tolerance, or the state of unresponsiveness of the immune system, must be established in the early life environment to prevent potentially damaging inflammatory responses to "self" (*15*). Below, I highlight different environmental exposures that support the idea of an early life origin of asthma.

Maternal behaviors, such as diet and smoking, have been shown to alter outcomes of allergic disease in childhood. In a longitudinal study of 460 Spanish children, a high Mediterranean diet during pregnancy was protective against wheeze and atopy at age 6.5 (*16*). A large component of the Mediterranean diet includes fish intake, containing a high content of n-3 polyunsaturated fatty acids (PUFAs). In a randomized, controlled trial of 98 pregnant women with allergic sensitization, those who received n-3 PUFA supplementation had infants who had reduced cytokine responses to allergen challenges (*17*). A more recent study demonstrated that pregnant women who received supplementation of n-3 PUFAs during the third trimester of pregnancy had offspring with reduced persistent wheeze and asthma (*18*). Together, these studies show that

maternal or infant dietary supplementation may protect against wheezing in infancy and asthma in childhood. It is also well established that maternal smoking in pregnancy and infant exposure to environmental tobacco smoke modifies lung development and immune responses (*19*) and are associated with the increased risk of wheezing illnesses and asthma (*20, 21*). Maternal smoking has additionally been associated with low birth weight and prematurity (*22*), which can both alter lung development (*23*) and contribute to childhood asthma risk (*24, 25*).

Among the many environmental factors that influence risk for asthma, early-life microbial exposures are some of the most significant. Human rhinovirus (RV) is the most common respiratory pathogen and the primary cause of the common cold (*26*). RV infections are responsible for over half of asthma exacerbations throughout life (*27*), contributing to the morbidity of chronic respiratory illnesses including asthma, as well as chronic obstructive pulmonary disease and cystic fibrosis (*28*). Multiple studies have demonstrated that infants who wheeze with RV infections are at greater risk for asthma (*29*). In a genetic study from the Childhood Origins of Asthma (COAST) birth cohort from Madison, Wisconsin, a SNP at the 17q12-21 asthma locus, the most consistently associated locus for childhood-onset asthma (*30*), was significantly associated with the number of RV-associated wheezing illnesses during the first three years of life. Moreover, RV-associated wheezing illnesses in the first 3 years of life modified genotype effects of this SNP on childhood onset asthma risk (*29*). This interaction between RV-associated wheeze and 17q12-21 genotype on the development of asthma was replicated in a birth cohort from Denmark (*29*). These studies highlight the importance of responses to respiratory viruses, host genotype, and their interactions on early life wheezing illness and the subsequent development of asthma in childhood.

A link between the early-life host microbiome and childhood asthma has been reproduced in various North American and European birth cohorts (*31-35*). In general, depletion of diversity

in the gut microbiota of infants and neonates is associated with increased risk of asthma. In a Canadian cohort of 319 infants, the relative abundances of *Lachnospira*, *Veillonella*, *Faecalibacterium*, and *Rothia* were significantly decreased in children who developed wheezing illness in the first year of life (*31*). High-risk infants had reduced levels of fecal acetate and enterohepatic metabolites, and reduced levels of anti-inflammatory PUFAs and enrichment of monohydroxy fatty acids, such as 12,13-diHOME (*33*), which has been associated with immunological intolerance and risk of allergic disease in childhood (*36*). In a Danish birth cohort of 700 infants, microbial diversity and the relative abundances of *Veillonella* and *Prevotella* in the airway microbiota at age one month was associated with asthma at age 6 years (*34*). Higher relative abundances of both genera modified an airway immune profile that was itself an independent predictor of asthma.

The connection between early life exposures and subsequent development of asthma fits with the "developmental origins of health and disease" hypothesis, also known as the Barker hypothesis (*37*). This hypothesis posits that prenatal and postnatal exposures induce permanent changes that promote disease development later in life (*38*). Epigenetic changes provide a substrate through which early life exposures can influence the trajectory to asthma. Indeed, early life exposures have been associated with epigenetic changes that may regulate the expression of nearby genes with effects on immunomodulatory pathways (*39-41*).

1.3 Maternal asthma as a risk factor

Heritability estimates of asthma are on average ~50% (*42*) and many asthma-associated SNPs have been identified in genome-wide association studies (GWAS) (*43-47*). However, asthma-associated GWAS variants have modest effects and only account for a small portion of the heritability (*43-47*). In fact, a recent GWAS of childhood onset and adult onset asthma in over 300,000 participants

in the UK Biobank resource (*48*) estimated the SNP-based heritability to be 33% and 9.8%, respectively (*43*). The genome-wide significant SNPs only account for one-third of the total heritability. This "missing heritability" suggests that there are remaining genetic contributions to asthma risk to be discovered.

One of the most consistent risk factors for childhood-onset asthma is a maternal history of asthma – a finding that has been replicated across many epidemiological studies and is significantly stronger than risk associated with paternal history of asthma (*49*). This finding was also suggested for other allergic and atopic phenotypes (*50*). Maternal asthma has also been associated with low birth weight (*51*) and prematurity in her newborns (*52*). Despite the recognized importance of maternal asthma as a risk factor for asthma in her children, relatively few genetic studies of asthma have considered maternal asthma in their analyses.

The few genetic studies to investigate maternal asthma have largely focused on one candidate asthma-susceptibility gene, *HLA-G*. This initial discovery came from a linkage analysis of asthma by Nicolae *et al. (53)*, which ultimately revealed an interaction between genotype at a SNP in the promoter region of this gene and asthma, but with opposite alleles associated with asthma in children from families with an asthmatic mother compared to children from families with a non-asthmatic mother. This interaction was replicated by Tan *et al.* (*54*) in the COAST birth cohort, where they further implicated another SNP, in LD with the promoter SNP, located in the 3'UTR. They showed that the 3'UTR SNP disrupted a target site for three related miRNAs. Finally, Nicodemus-Johnson *et al (55)*. showed that *HLA-G* genotype was associated with the concentration of HLA-G protein in bronchoalveolar lavage (BAL) fluid only in asthma cases with an asthmatic mother but not in asthma cases without an asthmatic mother. They further demonstrated that one of the three miRNAs, *miR-148b*, was increased in bronchial epithelial cells

from asthmatics with an asthmatic mother compared to asthmatics without an asthmatic mother. Notably, these associations were observed in adults, suggesting that maternal asthma effects on gene regulation persist into adulthood and that at least some of these effects are mediated through miRNAs.

Epigenetic studies of maternal asthma effects have highlighted DNA methylation (DNAm) differences between children of asthmatic mothers and children of non-asthmatic mothers. Gunawardhana *et al.* (*56*) reported altered DNA methylation profiles in peripheral blood cells from 12-month-old infants with and without a mother with asthma. The authors highlighted hypermethylation of the gene encoding CHFR, a ubiquitin ligase that is associated with persistent wheezing in children (*57, 58*). DeVries *et al.* (*59*) reported 589 differentially methylated regions (DMRs) between cord blood immune cells from 18 children with an asthmatic mother who developed asthma by age 9 compared to 18 children with an asthmatic mother who did not develop asthma by age 9. The most significant DMRs were near genes that clustered in immunoregulatory and proinflammatory pathways. The most connected gene in the network was *SMAD3*, which lies in a highly replicated asthma GWAS locus (*60*). The promoter of *SMAD3* was also hypermethylated in asthmatic children with an asthmatic mother compared to asthmatic children without an asthmatic mother. This finding with a *SMAD3* methylation site was replicated in two additional birth cohorts (*59*).

Collectively, these studies support the "developmental origins" of asthma theory, with maternal asthma-associated epigenetic regulation of gene expression by miRNAs and DNAm altering immune pathways in early development that lead to the development of asthma later in life.

1.4 Dissertation overview

The primary goal of this dissertation was to identify and characterize genetic and epigenetic contributors to risk of asthma in children of asthmatic mothers. To do this, I use various next-generation sequencing datasets in ethnically-diverse cohorts to investigate this theory, and to provide biological insights into the origins of asthma. In Chapter 2, I perform the first study of a novel epigenetic mechanism, ADAR-mediated microRNA editing, in bronchial epithelial cells and identify associations between editing and asthma severity and between *ADAR* expression and maternal asthma. In Chapter 3, I investigate DNA methylation profiles in the same bronchial epithelial cells from asthmatic cases with an asthmatic mother and identify epigenetic signatures in these subjects that reflect multiple potential endotypes of severe disease. In Chapter 4, I describe a study aimed at testing associations between mitochondrial variants and haplogroups with asthma in ethnically diverse mother-child dyads.

Overall, this research will provide information on asthma risk that is personalized with respect to the presence of disease in the mother. For example, risk from other asthma loci or due to environmental exposures may depend on mitochondrial variants inherited from the individual's mother or involve ADAR-mediated microRNA editing or DNA methylation that are perturbed in asthmatic children of asthmatic mothers. In addition, I provide unique insights into the genetic and epigenetic architecture of asthma as well as uncover mechanisms contributing to maternal transmission of asthma risk, potentially contributing to the development of precision medicine diagnostics and treatments.

CHAPTER 2

ADAR-mediated editing of miR-200b-3p in airway cells is associated with moderate-to-severe asthma in children of asthmatic mothers

2.1 Abstract

One of the most consistent risk factors for childhood-onset asthma is having a mother with asthma. A possible explanation may be altered gene regulation *in utero* that is long-lasting. Here, I performed the first genome-wide analysis of ADAR-mediated microRNA editing using microRNA-seq in primary bronchial epithelial cells from 142 adult asthma cases and non-asthma controls. Of 19 A-to-I edited sites detected in these microRNAs, 16 were in seed regions. Four of the 16 edited sites were observed in >10 individuals and were tested for differential editing (% A-to-I) between groups. One site at position 5 of miR-200b-3p was edited less frequently in asthma cases compared to controls ($P = 0.013$). A-to-I editing of this site was then compared between asthma severity groups (mild, moderate and severe) based on lung function and medication use. The moderate ($P = 0.029$) and severe ($P = 3.9x10^{-4}$), but not mild ($P = 0.38$), asthma cases had significantly less A-to-I editing of the 5th position of miR-200b-3p compared to controls. Bioinformatic prediction revealed 232 in silico target genes of the edited miR-200b-3p, which included the *SOCS1* gene that was itself more highly expressed in moderate ($P = 0.017$) and severe ($P = 5.4x10^{-3}$) asthma. Moreover, the primary editing enzyme, *ADAR*, was significantly downregulated in adults with a mother with asthma ($P = 0.024$), but not in asthma cases overall. Together, these results suggest that maternal asthma may alter the gene regulatory landscape in BECs *in utero*, with long-lasting effects on ADAR-mediated editing of the 5^{th} position of miR-200b-3p. I further suggest that this "exposure" results in reduced suppression of *SOCS1* in BECs of asthmatics, leading to more symptoms and more severe asthma.

2.2 Introduction

Asthma is a common chronic inflammatory disease of the airways, affecting as many as 300 million people worldwide (*1*). The clinical diagnosis of asthma is based on symptoms of wheezing, coughing, and chest tightness, which reflect the cardinal features of bronchial hyperresponsiveness and airway remodeling. As many as 15% of individuals with asthma have severe asthma, which is defined by uncontrollable symptoms, persistent exacerbations and poor response to standard therapies (*3*). Among the many epidemiologic risk factors for asthma, asthma in the mother is one of the most reproducible (*49*), although the mechanism(s) underlying this association are unknown.

MicroRNA (miRNA)-mediated gene silencing has been implicated in the pathogenesis of asthma. MiRNAs are approximately 22-nucleotide RNAs that bind to 3' untranslated regions of target genes and ultimately downregulate their expression. Using array-based studies of 876 miRNAs, Solberg and colleagues showed that the percentage of differentially expressed (DE) miRNAs far exceeded the percentage of DE mRNAs in bronchial epithelial cells (BECs) from 47 subjects (*61*). In a study of 190 selected miRNAs measured by quantitative PCR, Simpson *et al.* reported altered expression of miR-19a and its gene targets that promoted Type 2 (T2) cytokine production in airway T cells from 43 asthmatic subjects (*62*). Nicodemus-Johnson *et al.* also found that the abundance of miR-148b, a miRNA that binds to and downregulates the expression of an asthma-associated gene, *HLA-G (53)*, differed in BECs between adult asthma cases with and without a mother with asthma (*55*). These studies highlighted miRNA abundances as potentially important gene regulatory mechanisms in the lower airways of individuals with asthma.

MiRNA-mediated gene silencing is often altered by adenosine deaminase acting on RNA (ADAR)-mediated editing, a novel epigenetic mechanism that involves the deamination of adenosine to inosine (A-to-I). The ADAR family of enzymes contributes to one of the most

prevalent forms of post-transcriptional modifications of the primary, precursor and mature forms of miRNAs (*63*). Because most ADAR-mediated editing events in miRNAs occur in the seed regions (*64*), the critical determinants of target complementarity, A-to-I edited sites are expected to downregulate the expression of gene targets that differ from the gene targets of the non-edited form. *ADAR* expression can be induced by interferons (*65*), suggesting a particularly important role for ADAR-mediated editing in innate inflammatory responses. In fact, elevated *ADAR* expression and subsequent A-to-I editing activity have significant effects on the transcriptomes of cancer cells and may even correlate with patient survival (*66, 67*). Furthermore, widespread A-to-I editing by *ADAR* is present in immune tissues and cells, such as spleen, thymus and peripheral lymphocytes, in response to inflammation in endotoxin-treated mice (*68*). Despite these advances, the role of ADAR-mediated editing of miRNAs in asthma has not previously been explored.

Here, I present the first genome-wide analysis of ADAR-mediated editing of miRNAs using small RNA sequencing in primary BECs from 96 asthma cases and 46 non-asthma controls. I identified A-to-I editing of the 5th position of miR-200b-3p, which was less frequent in BECs from asthma cases, particularly those with moderate-to-severe asthma. Using RNA-seq in cells from the same subjects, I examined the expression of predicted targets of the unedited and edited versions of miR-200b-3p and characterized effects of maternal asthma on the expression of the ADAR genes. Overall, my study identified ADAR-mediated editing of miR-200b-3p in the lower airway as a novel contributor to asthma.

2.3 Results

2.3.1 Identifying A-to-I edited sites in bronchial epithelial cells

Our lab obtained small RNA sequences in freshly isolated BECs from 142 adults. Of these, 138 subjects (94 asthma cases; 44 non-asthma controls) had >2 million small RNA reads and were

included in downstream analyses. I identified 19 A-to-I edited sites in 17 miRNAs (**Table S2.1**); the number of edited sites is consistent with a study of glioblastoma that included healthy lung tissue (*64*). The mean A-to-I editing at the 19 sites ranged from 0.14% to 34.09% across individuals in my study with detectable levels of editing (see **Methods in 2.5.5**); 16 of the 19 A-to-I edited sites were in miRNA seed regions (positions 2 to 8). Nearly all of the A-to-I edited sites (18 of 19) were detected in fewer than 20% of individuals (range = 1 to 27). The 5th position of miR-200b-3p was the most frequently edited site in the sample: 97.10% of individuals had edited sites at this position.

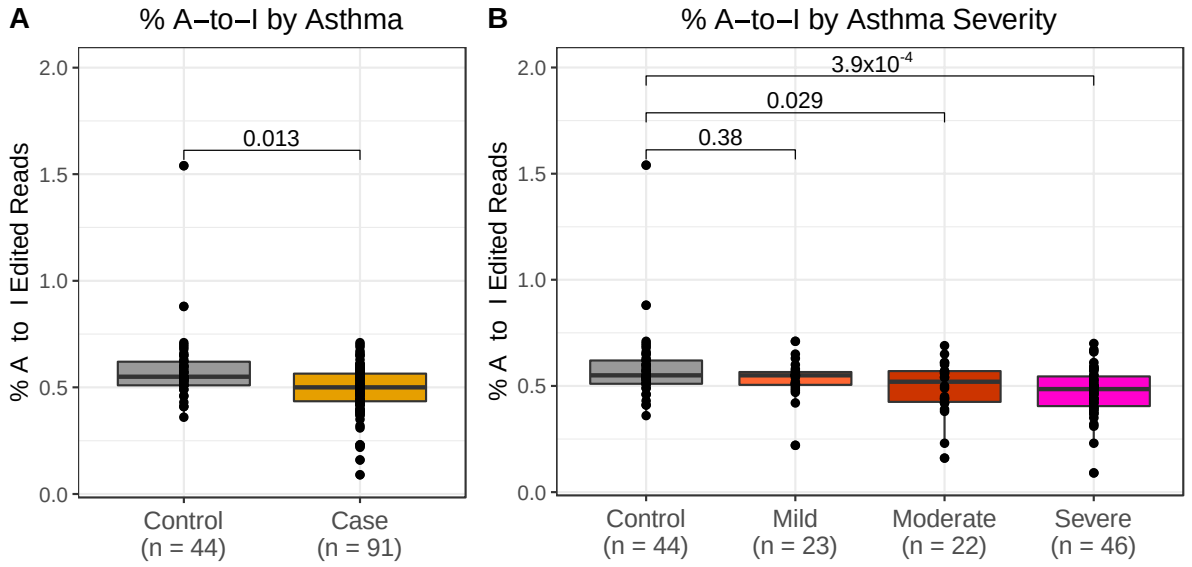
2.3.2 Differential A-to-I editing of miR-200b-3p between cases and controls

Table 2.1. Top four A-to-I edited sites. A-to-I edited sites detected in bronchial epithelial cells from at least 10 of 138 subjects. The edited adenosine within the mature miRNA sequence is in bold font. All 19 A-to-I edited sites are shown in **Table S2.1**.

miRNA	Position	Sequence (seed sequence)	Average Mapped Counts	Average % A-to-I Reads	# Subjects with Edited miRNA
miR-200b-3p	5	T<u>AAT</u>**A**<u>CTG</u>CCTGGTAATGATGA	75358	0.52	135
miR-191-5p	3	C<u>AA</u>**A**<u>CGGAA</u>TCCCAAAAGCAGCTG	58512	0.17	27
miR-200b-3p	3	T<u>AA</u>**T**<u>ACTG</u>CCTGGTAATGATGA	91380	0.15	22
miR-186-5p	3	C<u>AA</u>**A**<u>AGAAT</u>TCTCCTTTTGGGCT	24646	0.18	13

Four edited sites were observed in >10 subjects (**Table 2.1**). For 135 subjects that passed QC and had available genotypes for estimating ancestry principal components, differential editing (% A-to-I) was tested between 91 asthma cases and 44 non-asthma controls. The most edited site, the 5th position of miR-200b-3p, was significantly less frequently edited in cases compared to controls (P = 0.013 after correcting for four tests; age, sex, current smoking status, and the first three ancestry PCs were included as covariates) (**Figure 2.1A**). This significant decrease in A-to-I editing remained after excluding one outlier among the controls (P = 0.023 after correcting for four tests) (**Figure S2.1**). The other three sites were not differentially edited between asthma cases and controls (**Figure S2.2**).

Figure 2.1. A-to-I editing of position 5 of miR-200b-3p in asthma cases and controls. Associations of percentage (%) of A-to-I editing of position 5 of miR-200b-3p in bronchial epithelial cells with (**A**) asthma and (**B**) asthma severity groups based on STEP classification scores. One outlier with the highest % of edited reads in the control group (>3 SD) was included here (see **Figure S2.1** with the outlier excluded). The number of subjects is shown below each risk group. Adjusted p-values are shown above the horizontal bars. The p-values for (**A**) and (**B**) were corrected for four and three tests, respectively, using a Bonferroni correction.



I next assessed whether A-to-I editing of the 5th position of miR-200b-3p was associated with asthma severity. A-to-I editing was less frequent in the moderate (N = 22) and severe (N = 46) asthma groups compared to controls (N = 44) (P = 0.029 and P = 3.9x10^{-4}, respectively, after correcting for three tests) (**Figure 2.1B**, see also **Figure S2.1**). The frequency of editing did not differ between mild asthma cases (N = 23) and controls (P = 0.38 after correcting for three tests).

2.3.3 Downstream targets of unedited and edited miR-200b-3p

The A-to-I edited site at the 5th position is in the seed region of miR-200b-3p and is predicted to downregulate the expression of novel target genes. I used TargetScanHuman (*69*) to bioinformatically predict the target genes of the unedited and edited miR-200b-3p. Of the 798 *in silico* target genes of the unedited miR-200b-3p, 604 (75.7%) were detected as expressed in BECs; of the 320 *in silico* target genes for the edited miR-200b-3p, 232 (72.5%) were detected as

expressed in BECs (**Figure 2.2A**). Only 48 expressed genes overlapped between the two sets of target genes. Pathway analysis (TopFunn (*70*)) of the 604 expressed target genes of the unedited miR-200b-3p, which was more frequent in the cases, resulted in more than 39 significant pathways, most of which were involved in kinase signaling, proliferation and differentiation (**Table S2.2**). In contrast, the 232 expressed target genes of the edited miR-200b-3p, which was more frequent in the controls, were significantly enriched for four pathways (**Figure 2.2B**). Two of the pathways of the edited miR-200b-3p were the same as two pathways of the unedited miR-200b-3p: "C-MYB transcription factor network" from two databases and "FoxO family signaling" (**Figure 2.2C**). The two other pathways were specific to the edited miR-200b-3p: "IL-4 mediated signaling events" from two databases and "IFN-gamma pathway" (**Figure 2.2D**).

RNA-sequences were available for 124 subjects from the same cells in which I measured miRNAs. I first examined the average normalized expression of the expressed target genes in each of the four enriched pathways of the edited miR-200b-3p. The average expression of the eight target genes (*CEBPB*, *SPI1*, *ETS1*, *PIK3CA*, *CBL*, *SOCS5*, *SOCS1*, and *MYB*) of the edited miR-200b-3p in the pathway enriched for "IL-4 mediated signaling events" and the six target genes (*CEBPB, PIK3CA, RAP1A, RAP1B, CBL,* and *SOCS1*) of the edited miR-200b-3p in the "IFN-gamma pathway" were both significantly increased in asthma cases (N = 83) compared to controls (N = 41; P = 2.9×10^{-3} and 1.8×10^{-3}, respectively, after correcting for four tests) (**Figure 2.3A**). In contrast, the average expression was not different between asthma cases and controls for the nine target genes *(CEBPB*, *SPI1*, *ETS1*, *KRAS*, *SKI*, *CSF1R*, *ATP2B1*, *MYB, CDK6*) in the pathways enriched for "C-MYB transcription factor network" or for the six target genes (*ZFAND5, FOXO1, YWHAE, YWHAZ*, *SIRT1*, *CSNK1G2*) in the pathways enriched for "FoxO family signaling" (P = 1.00 after correcting for four tests) (**Figure 2.3B**). The overall increased average expression of the

target genes of both the IL-4 signaling and interferon gamma pathways in the asthma cases is consistent with the edited miRNA binding to and downregulating the expression of these genes, and with the observed higher frequency of editing of this miRNA in the controls.

Figure 2.2. Significant pathways for the target genes of the unedited and edited miR-200b-3p. (**A**) The 604 gene targets of the unedited miR-200b-3p and the 232 gene targets of the edited miR-200b-3p that were expressed in BECs (CPM > 1) were included as input for pathway analysis, as implemented in TopFunn. (**B**) The 604 gene targets of the unedited miR-200b-3p were enriched for at least 39 pathways (**Table S2.2**) and the 232 gene targets of the edited miR-200b-3p were enriched for four pathways. (**C**) Of the four significant pathways enriched for gene targets of the edited miR-200b-3p, two pathways that included "C-MYB transcription factor network" from two databases and "FoxO family signaling" were also enriched for the target genes of both the unedited and edited miR-200b-3p. (**D**) The two pathways that were only enriched for the target genes of the edited miR-200b-3p were "IL-4-mediated signaling events" from two databases and "IFN-gamma pathway".

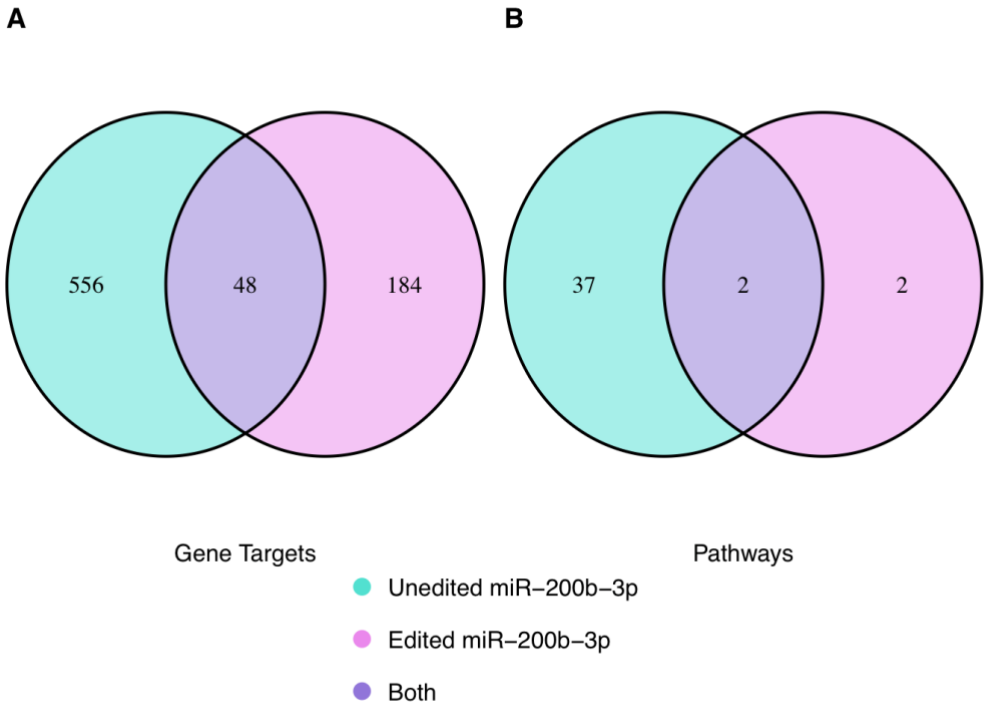


Figure 2.2. Significant pathways for the target genes of the unedited and edited miR-200b-3p (continued).

C Pathways shared by the target genes of the edited and unedited miR-200b-3p

ID	Name	Database	Adjusted P-value	Number of genes from input	Number of genes in annotation	Genes from input
M195	C-MYB transcription factor network	MSigDB C2 BIOCARTA (v7.1)	1.85E-03	9	84	*CEBPB, SPI1, ETS1, KRAS, SKI, CSF1R, ATP2B1, MYB, CDK6*
138073	C-MYB transcription factor network	BioSystems: Pathway Interaction Database	1.02E-02	8	78	*CEBPB, SPI1, ETS1, SKI, CSF1R, ATP2B1, MYB, CDK6*
138036	FoxO family signaling	BioSystems: Pathway Interaction Database	4.81E-02	6	48	*ZFAND5, FOXO1, YWHAE, YWHAZ, SIRT1, CSNK1G2*

D Pathways specific to the target genes of the edited miR-200b-3p

ID	Name	Database	Adjusted P-value	Number of genes from input	Number of genes in annotation	Genes from input
137933	IL-4-mediated signaling events	BioSystems: Pathway Interaction Database	1.99E-03	8	63	*CEBPB, SPI1, ETS1, PIK3CA, CBL, SOCS5, SOCS1, MYB*
M28	IL-4-mediated signaling events	MSigDB C2 BIOCARTA (v7.1)	2.25E-03	8	64	*CEBPB, SPI1, ETS1, PIK3CA, CBL, SOCS5, SOCS1, MYB*
M161	IFN-gamma pathway	MSigDB C2 BIOCARTA (v7.1)	1.64E-02	6	40	*CEBPB, PIK3CA, RAP1A, RAP1B, CBL, SOCS1*

Figure 2.3. Average normalized expression of the genes in each of the four pathways of the edited miR-200b-3p in asthma cases and controls. Only gene targets that were expressed (CPM > 1) in BECs in at least 25% of the 124 subjects (83 cases and 41 controls) with RNA-seq data were included. (**A**) The average normalized expression of the two pathways that were enriched only for target genes of the edited miR-200b-3p: "IL-4 mediated signaling events" pathway (left) and "IFN-gamma" pathway (right). (**B**) The average normalized expression of the two pathways that were enriched for target genes of both the unedited and edited forms of miR-200b-3p: "C-MYB transcription factor network" pathway (left) and "FoxO family signaling" pathway (right). Adjusted p-values (for four tests) are shown above the horizontal bars.

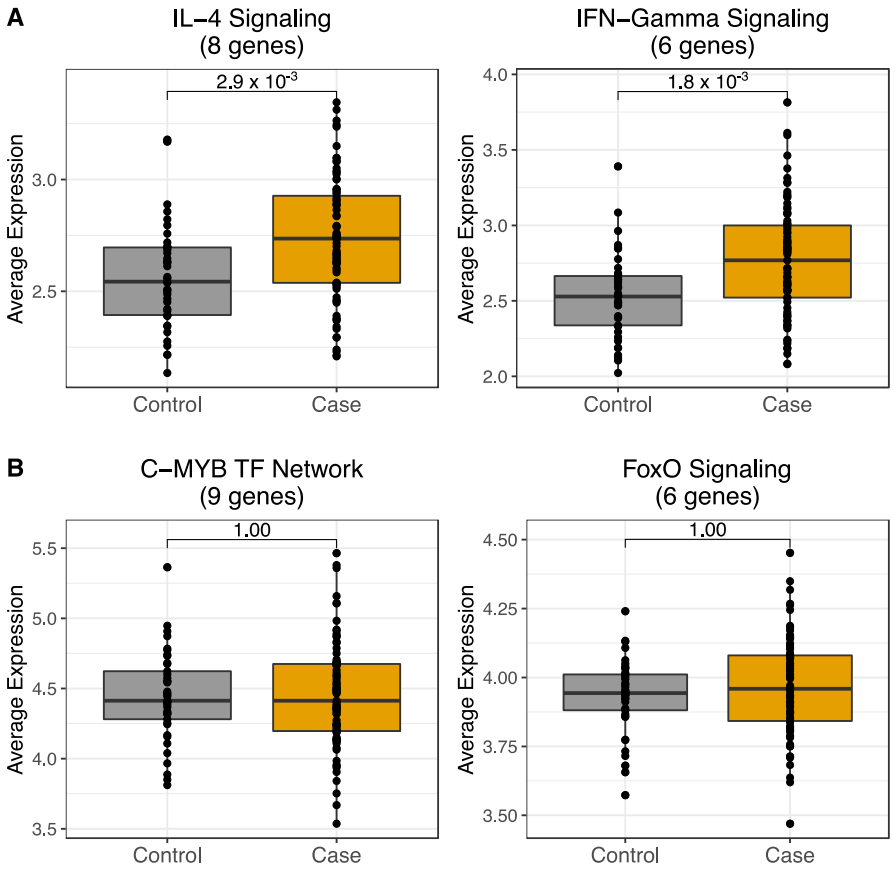
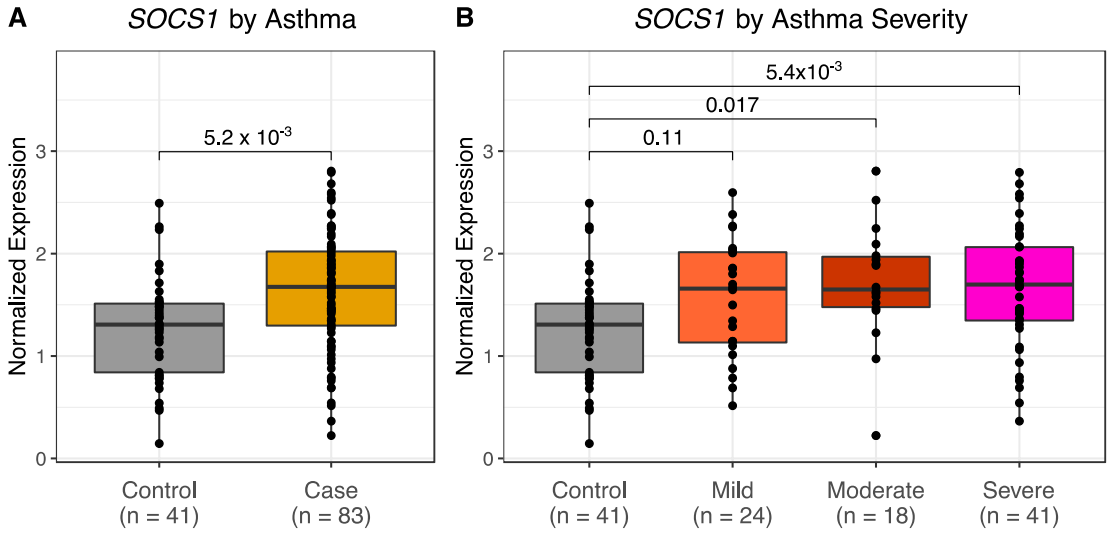


Figure 2.4. Expression of *SOCS1*, a target of the edited miR-200b-3p, in asthma cases and controls. Normalized gene expression of *SOCS1* in bronchial epithelial cells (**A**) between asthma cases and controls, and (**B**) asthma severity groups based on STEP classification. The numbers of subjects are shown below the risk groups. Adjusted p-values (for three tests) are shown above the horizontal bars.



Of the ten target genes in the IL-4 and interferon gamma pathways of the edited miR-200b-3p, *SOCS1*, a negative regulator of both Type 2 (T2) cytokine signaling (*71*, *72*) and interferon signaling induction (*73*, *74*), was the only individual gene that was significantly overexpressed in the asthma cases (N = 83) ($P = 5.2x10^{-3}$ after correcting for ten tests) (**Figure 2.4A**), and in moderate (N = 18) and severe (N = 41), but not in mild (N = 24) asthma cases compared to controls (N = 41) ($P = 0.017$, $5.4x10^{-3}$, and 0.11, respectively, after correcting for three tests) (**Figure 2.4B**). None of the other nine target genes in these pathways were differentially expressed between asthma cases and controls (**Figure S2.3**). Thus, the decreased editing of the 5th position of miR-200b-3p in BECs from moderate and severe asthma cases is associated with overexpression of a centrally important negative regulator of cytokine signaling, *SOCS1*.

2.3.4 MiR-200b-3p editing, SOCS1 expression, and measures of asthma

Because both A-to-I editing of the 5th position of miR-200b-3p and *SOCS1* expression were associated with asthma severity (**Figures 2.1 and 2.4**), I next examined more closely the potential effects of editing and *SOCS1* expression on clinical measures of asthma. I first assessed their relationships with asthma and asthma severity (**Table 2.2**). Unsurprisingly, compared to the controls, the asthma cases had significantly lower lung function, as measured by FEV1% predicted and the FEV1/FVC ratio, and higher FeNO, BAL eosinophils, blood eosinophils, and T2 gene signatures (*CLCA1*, *POSTN*, and *SERPINB2*) (*75*) (**Table 2.2A**). IgE was also higher in the cases but the differences were not significant after correction for seven tests. I next examined associations between these measures and STEP classification of asthma severity, based on corticosteroid use and lung function (*76*) (**Table 2.2B**). Two measures, in addition to the lung function measures, were associated with STEP classification. BAL eosinophils were highest in the moderate and severe asthma cases, whereas the T2 gene signature was highest in the moderate asthma cases but similar in cases with mild and severe disease. These profiles suggest that the asthma cases classified as having moderate disease have a T2-high endotype, whereas the cases classified as severe asthma have a T2-independent endotype.

Table 2.2. Clinical characteristics of subjects at the time of bronchoscopy. Af Am, African American; Eur Am, European American**.** *Continuous variables were evaluated with a Wilcoxon rank-sum test; categorical variables were evaluated with a Fisher's Exact Test. &All variables were assessed using an ordinal logistic regression. Five severe asthma cases had missing FeNO measurements. One control each was missing BAL and blood eosinophil measurements. Normalized and covariate-adjusted gene expression was available for 41 controls and 83 cases (24 mild, 18 moderate, and 41 severe) to assess the Type 2 gene signature (sum of *POSTN*, *CLCA1*, and *SERPINB2*). Significant correlations after correcting for seven tests (clinical measures) are shown in bold font.

	A. Asthma			**B. Asthma Severity**				
	Control (N = 44)	**All Cases** (N = 94)	**P-value***	**Control** (N = 44)	**Mild** (N = 23)	**Moderate** (N = 23)	**Severe** (N = 48)	**P-value&**
Covariates								
Age (mean yr ± SD)	37.25 ± 11.54	40.16 ± 12.77	0.21	37.25 ± 11.54	34.83 ± 12.61	37.96 ± 13.83	43.77 ± 11.37	9.48×10^{-3}
Gender (% female)	64	73	0.32	64	61	70	81	0.06
Ethnicity (Af Am/Eur Am/Other)	28/12/4	54/39/1	0.03	28/12/4	16/7/0	16/7/0	22/25/1	0.02
% smoker at bronchoscopy	16	4	0.04	16	9	0	4	0.03
Clinical Measures								
ICS use (%)	-	74	-	-	9	91	98	-
OCS use (%)	-	37	-	-	0	0	73	-
Mean FEV1% Predicted (± SD)	95.18 ± 11.31	73.67 ± 19.20	2.03×10^{-10}	95.18 ± 11.31	83.61 ± 16.60	79.96 ± 18.84	65.90 ± 17.47	**5.34×10^{-10}**
Mean FEV1/FVC (± SD)	0.81 ± 0.05	0.76 ± 0.30	1.88×10^{-7}	0.81 ± 0.05	0.80 ± 0.25	0.74 ± 0.08	0.75 ± 0.37	**2.05×10^{-6}**
Median total serum IgE (IU/mL) (lower, upper quartile)	56.50 (27.25, 169.00)	121.50 (37.00, 360.25)	0.02	56.50 (27.25, 169.00)	95.00 (28.50, 171.50)	116.00 (70.50, 465.50)	194.50 (24.75, 387.00)	0.02
Median FeNO (ppb) (lower, upper quartile)	14.00 (10.75, 17.75)	25.00 (14.00, 45.00)	9.75×10^{-5}	14.00 (10.75, 17.75)	25.00 (17.00, 54.00)	31.00 (19.00, 47.50)	20.00 (13.00, 34.00)	0.04
Median BAL eosinophilia (%) (lower, upper quartile)	0.002 (0, 0.005)	0.029 (0.013, 0.058)	3.92×10^{-14}	0.002 (0, 0.005)	0.019 (0.010, 0.048)	0.033 (0.017, 0.072)	0.031 (0.015, 0.060)	**1.21×10^{-5}**
Median blood eosinophilia (cells/µl) (lower, upper quartile)	0.100 (0.070, 0.170)	0.190 (0.080, 0.318)	1.43×10^{-3}	0.100 (0.070, 0.170)	0.140 (0.080, 0.250)	0.220 (0.135, 0.335)	0.200 (0.078, 0.303)	0.01
Median T2 gene (lower, upper quartile)	1.64 (0.15, 5.53)	7.38 (2.12, 11.68)	7.55×10^{-6}	1.64 (0.15, 5.53)	5.49 (0.80, 12.07)	10.03 (7.13, 11.39)	5.45 (1.89, 11.48)	**4.35×10^{-4}**

Table 2.3. MiR-200b-3p editing, *SOCS1* expression and measures of asthma. Significant correlations after correcting for five tests are shown in bold font. Number of subjects with measurements for each test of association is reported.

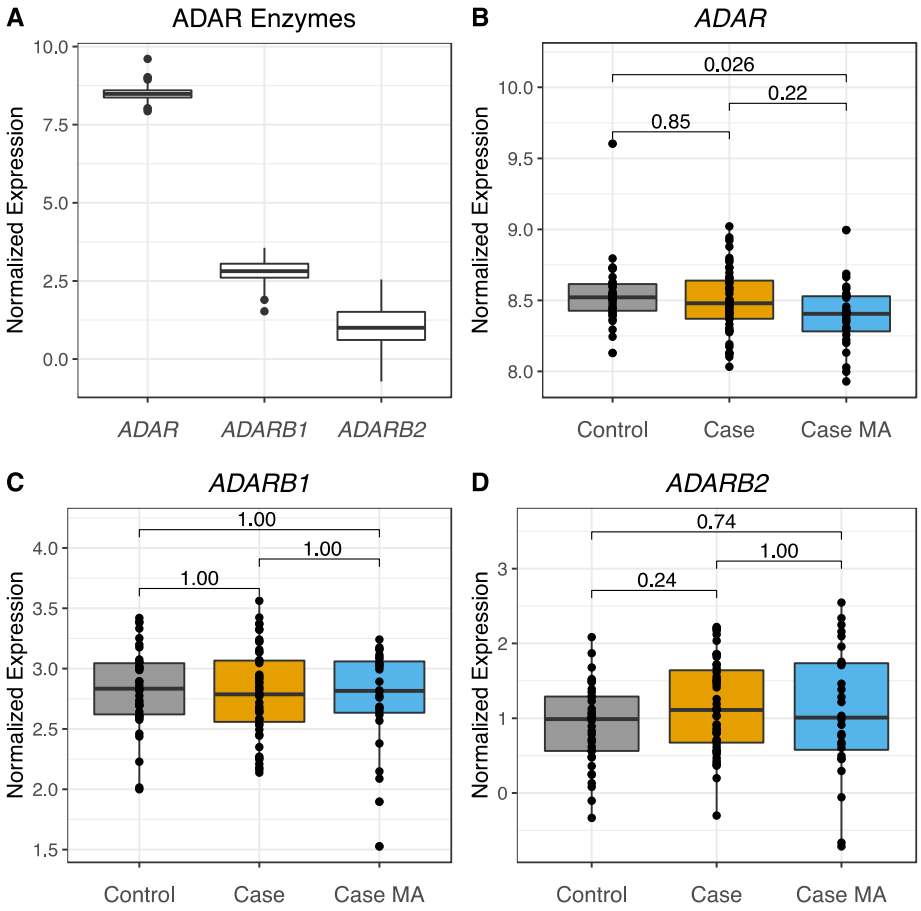
Variables	% A-to-I of miR-200b-3p				*SOCS1* expression			
	N	Rho	P-value	Adjusted P-value	N	Rho	P-value	Adjusted P-value
Total Serum IgE (IU/mL)	135	-0.07	0.41	1.00	124	0.21	0.017	0.09
FeNO (ppb)	130	-0.15	0.089	0.09	**119**	**0.24**	**$8.5x10^{-3}$**	**0.04**
BAL Eosinophilia (%)	**134**	**-0.24**	**$5.9x10^{-3}$**	**0.03**	**123**	**0.30**	**$6.2x10^{-3}$**	**0.03**
Blood Eosinophilia (cells/µl)	134	-0.05	0.60	1.00	123	0.08	0.37	1.00
T2 gene signature	**109**	**-0.25**	**$8.4x10^{-3}$**	**0.04**	**124**	**0.37**	**$2.8x10^{-5}$**	**$1.4x10^{-4}$**

I next tested for correlations between five clinical measures and A-to-I editing of the 5th position of miR-200b-3p and *SOCS1* expression (**Table 2.3**). Because I previously showed that both A-to-I editing of the 5th position of miR-200b-3p and *SOCS1* expression were correlated with asthma severity based on STEP classification, I excluded lung function measures from these analyses because lung function was used to define STEP classification. A-to-I editing of the 5th position of miR-200b-3p was negatively correlated with both BAL eosinophils and the T2 gene signature (*75*). Consistent with the results described above, *SOCS1* expression was positively correlated with BAL eosinophils and the T2 gene signature, as well as with FeNO, a marker of airway inflammation. Neither total serum IgE nor blood eosinophil count was associated with either editing or *SOCS1* expression. These data support a link between decreased ADAR-mediated editing of the 5th position of miR-200b-3p and increased expression of *SOCS1* in BECs from asthmatics with different endotypes of moderate-to-severe asthma.

2.3.5 Maternal asthma and ADAR abundance

The decreased editing of the 5th position of miR-200b-3p in BECs from asthma cases suggested a novel epigenetic mechanism of ADAR-mediated miRNA editing regulating asthma-promoting pathways. Because earlier studies suggested that maternal asthma may epigenetically modify asthma-promoting programs in her child, I next asked whether the expression of the *ADAR*

Figure 2.5. *ADAR* gene expression in bronchial epithelial cells (BECs). (**A**) Normalized expression of three ADAR genes; (**B**) *ADAR*, (**C**) *ADARB1* and (**D**) *ADARB2* were compared between non-asthma controls (n = 37), asthma cases without a mother with asthma (n = 51) and asthma cases with a mother with asthma (Case MA; n = 27). One outlier of *ADAR* and *ADARB1* expression (>3 SD) was removed from each of the Control and Case MA groups, respectively, to better visualize the distributions (see **Figure S2.4** for figures with the outliers excluded).



genes differed between asthma cases with and without an asthmatic mother. To test this hypothesis, I compared the expression of the three ADAR genes, *ADAR*, *ADARB1* and *ADARB2*, between three asthma risk groups (37 non-asthma controls without an asthmatic mother, 51 asthma cases without an asthmatic mother, and 27 asthma cases with an asthmatic mother). Four controls with a mother with asthma were excluded from these analyses. *ADAR* and *ADARB1* have editing ability; *ADARB2* does not have editing function and was included as a negative control. Of 13,770 genes detected as expressed in BECs, *ADAR* was in the top 2% of expressed genes overall and the most

highly expressed of the three ADAR genes (**Figure 2.5A**); only *ADAR* was significantly less abundant in asthma cases with an asthmatic mother compared to controls (P = 0.026 after correcting for three tests) (**Figure 2.5B)**. The significantly lower *ADAR* expression remained after including one outlier among the controls (P = 0.024 after correcting for three tests) (**Figure S2.4)**. *ADAR* did not significantly differ in asthma cases with an asthmatic mother compared to asthma cases without an asthmatic mother (P = 0.22 after correcting for three tests) or in asthma cases without an asthmatic mother compared to controls (P = 0.85 after correcting for three tests). *ADARB1* and *ADARB2* abundance did not differ between any of the groups (**Figure 2.5C-D**). These results indicate that the abundance of *ADAR*, an epigenetic mediator of miRNA editing, is less abundant in BECs from adult asthma cases with an asthmatic mother compared to controls.

2.4 Discussion

Although asthma symptoms in most individuals can be controlled with standard therapies, approximately 15% of asthmatics do not respond to these treatments and account for a disproportionate burden on health care costs and quality of life (*3*). Genetic studies of response to the most common therapies, corticosteroids and beta agonists, have revealed few clues to genetic mechanisms underlying inter-individual variation in response (*77, 78*), and a recent genome-wide association study of moderate-to-severe asthma generally mirrored the associations with asthma (*60, 79*). These results suggest that determinants of moderate-to-severe asthma may not be due to classical genetic mechanisms.

Using an unbiased, genome-wide approach, I identified for the first time endogenous post-transcriptional A-to-I editing of a miRNA in BECs as an epigenetic mechanism potentially leading to more severe clinical manifestations and poorer response to standard therapies in individuals with asthma. A-to-I editing of one site, the 5th position of miR-200b-3p, was significantly more

frequent in BECs from non-asthma control subjects compared to cases with moderate and severe asthma. This same A-to-I edited site has been reported in various cancer tumors, including glioblastoma (*80*) and lung adenocarcinoma (*81*), and is consistently increased in tumor samples compared to normal tissue and inversely correlated with patient survival (*67*). Functional characterization of miR-200b-3p in this context identified its role as a tumor suppressor that becomes an oncogenic miRNA after A-to-I editing of the 5th position. In my study, reduced A-to-I editing of the 5th position of miR-200b-3p was associated with increased risk of asthma. These combined results suggest that edited miR-200b-3p is critical to different disease processes, in which editing of the same site can be promoting (cancer) or protective (asthma) of disease. Editing at this position may therefore be highly regulated to optimize expression levels of key genes targeted by the edited form of miR-200b-3p.

Indeed, the predicted target genes of the edited miR-200b-3p were enriched for IL-4 signaling and interferon gamma pathways, both of which included the gene, *SOCS1*. Consistent with decreased A-to-I editing of the 5th position of miR-200b-3p, *SOCS1* was expressed at significantly higher levels in BECs from adult cases with moderate or severe asthma compared to controls. *SOCS1* is a centrally important transcription factor that has dual roles in both suppressing T2 cytokine (*71, 72*) and interferon (*73, 74*) signaling pathways. Although the evidence is conflicting, most studies have reported increased expression of *SOCS1* in BECs of subjects with severe asthma (*82-84*), similar to my study, whereas only one study reported decreased expression of *SOCS1* in cells from subjects with severe asthma (*85*).

Asthma severity, miR-200b-3p editing and *SOCS1* expression were also correlated with BAL eosinophils and the T2 gene signature. Whereas eosinophils were most abundant in BAL from moderate and severe asthmatics, the T2 gene signature was increased among asthma cases

with moderate disease but was expressed at similar levels in asthma cases with severe and mild disease. This suggests that the moderate asthma cases, who are by definition more responsive to corticosteroids than the severe asthmatics, have a profile consistent with a T2-high asthma endotype. In contrast, the severe asthma cases have features consistent with T2-independent disease. These results therefore indicate that editing of the 5th position of miR-200b-3p, and the resulting over expression of *SOCS1*, contributes to both T2-high and T2-independent endotypes of moderate-to-severe asthma (*86*). These findings highlight the dual roles of *SOCS1* in suppressing both T2 cytokine and interferon signaling. Taken together, I suggest that decreased A-to-I editing of the 5th position of miR-200b-3p in the asthma cases leads to increased *SOCS1* expression, which contributes to moderate and severe asthma, with both T2-dependent and T2-independent features resulting from either impaired T2 cytokine signaling or impaired interferon responses, respectively.

Finally, I show that *ADAR* itself was significantly less abundant in lower airway cells from asthma cases with an asthmatic mother compared to controls, suggesting less editing activity in BECs from these subjects. Although unexpected, this finding is consistent with an earlier study in a subset of these same individuals that reported maternal asthma-associated differences in the abundance of specific miRNAs (measured by qPCR) that target an asthma-associated gene, *HLA-G*, which also showed patterns of genetic associations that differed between children with or without mothers with asthma (*53-55*). My observation suggested that either *ADAR* expression itself is sensitive to the unique *in utero* environment of pregnancies in women with asthma, or that *ADAR* expression is a downstream target of other epigenetic programs that are sensitive to maternal asthma in pregnancy.

Despite the novelties of my study, there are limitations. First, direct measures of T2 cytokines or interferons were not available in these subjects so I could not test for associations between cytokine profiles and editing or *SOCS1* expression. Such measurements could provide information that would allow us to better link the moderate and severe asthma phenotypes to increased inflammation via IL-4 signaling or to asthma exacerbations mediated by impaired interferon signaling, respectively. Second, my sample sizes were relatively small, which may have prevented us from detecting significant differences in the mean expression of all target genes of the unedited or edited miR-200b-3p. It is possible that subtle changes in the expression of target genes of the edited miR-200b-3p other than *SOCS1* contribute to moderate or severe asthma phenotypes that could potentially be detected in larger sample sizes or in cases with different endotypes of asthma. Third, there were only four controls in my study who had a mother with asthma. Thus, I could not assess whether the maternal asthma effects that I observed in the cases would be similar or not in the controls, or disentangle the effects of maternal asthma and asthma in the child on my observations.

In summary, I provide the first evidence of associations between ADAR-mediated miRNA editing in bronchial epithelial cells with moderate-to-severe asthma, and implicate an epigenetic mechanism that may mediate the *in utero* effects of maternal asthma on immune modulatory pathways and childhood onset asthma in pregnancies of mothers with asthma. Presumably, prenatal effects persist into adulthood, reflecting long-lasting consequences of *in utero* exposures that are specific to maternal asthma. In this study, I demonstrate an important role for ADAR-mediated miRNA editing on asthma pathogenesis and identify new therapeutic targets for treatment of asthma that is difficult to control with standard therapies.

2.5 Methods

2.5.1 Data collection

Adults were recruited and clinically evaluated at the University of Chicago Medical Center. Whole blood was collected for genotyping and primary BECs were obtained via endobronchial brushings for small and total RNA studies, as previously described (*55*). A total of 161 subjects had either available genotypes for estimating ancestry principal components (n = 158), small RNA sequences (n = 142) or RNA sequences (n = 128). These studies were approved by the University of Chicago's Institutional Review Board.

2.5.2 Genotyping and quality control

SNPs were genotyped using the Illumina Omni2.5-8v1A, Omni1MDuo, or Human Core arrays. SNPs on each array were excluded with HWE < 0.0001 within each ethnic group (European American, African American), MAF < 0.05, SNP call rate < 0.95, and subject call rates < 0.95. Ancestry principal component analysis (PCA) was performed using 676 ancestral informative markers included on the arrays that overlap with the HapMap release 3.

2.5.3 Small RNA and total RNA extraction and sequencing

RNAs were extracted from bronchial epithelial cells and purified using the QIAGEN AllPrep DNA/RNA/miRNA Universal Kit (Hilden, Germany), following manufacturer's instructions. Quality and quantity assessment of small RNA and total RNA were measured at the University of Chicago Functional Genomics Core using an Agilent RNA 6000 Pico assay and the Agilent 2100 Bioanalyzer. Small RNA-seq libraries were prepared using Illumina small RNA-SEQ library kits. cDNA libraries were constructed using the Illumina TruSeq RNA Library Prep Kit v2. Small RNA and RNA sequencing were performed at the University of Chicago Genomics Core on either the Illumina HiSeq 2500 or 4000 platforms.

2.5.4 Quality control of small and total RNA-seq

For the small and total RNA-seq data, potential sample contamination and sample swaps were assessed using VerifyBamID (*87*). No cross-contamination was detected. For the RNA-seq data, two sample swaps between individuals were identified and corrected. Quality control checks were performed on the small RNA and total RNA-seq data using FastQC. Small RNA and total RNA sequences were then aligned and annotated to known miRNA and RNA sequences using miRge2.0 (*88*) and Spliced Transcripts Alignment to a Reference using STAR (*89*). miRNAs and genes with low count data (<1 CPM in at least 25% of the sample) and those on the X, Y, and mitochondrial chromosomes were removed. Samples containing > 2M mapped miRNA reads or > 8M mapped total RNA reads were kept, retaining 138 and 124 subjects for downstream analyses.

Raw miRNA and mRNA counts were normalized using quantile normalization and the trimmed mean of M-values (TMM) methods, respectively. Mean-variance trend of both datasets was adjusted using variance modeling in voom (*90*)). Technical sources of variation were identified using principal components analysis (PCA) for the global miRNA and total RNA expression profiles. For differential expression, surrogate variable analysis (*91*) was performed to predict latent covariates to include in the model. All covariates are described below.

2.5.5 Detection of A-to-I edited sites

The percentages of A-to-I editing were calculated for mapped miRNAs, using the A-to-I analysis module in miRge2.0 (*88*). Briefly, putative A-to-I edited signals were removed if: 1) they were located at miRNA SNPs with A/G differences; 2) they were in the 455 miRNAs located in repeat elements; 3) they were in a lowly expressed miRNA (RPM < 1); or 4) their edited sequences aligned to ≥1+ locations in the genome. For the retained reads that belonged to the canonical and edited miRNAs, all nucleotide positions in the canonical miRNA, except the terminal 5 bp were

screened for A to G changes based on a binomial test considering the expected sequencing error rate (0.1%), as described (*92*). The A-to-I editing frequency (% A-to-I) was defined as the proportion of the mapped reads containing the edited nucleotide (G) relative to the total mapped reads at the given location.

2.5.6 Analysis of A-to-I edited sites

A-to-I edited sites within miRNAs were identified using the A-to-I analysis module in miRge2.0 (*88*). Differential A-to-I editing was examined using a linear model, as implemented in limma (*93*). The four A-to-I edited sites that were present in at least 10 subjects were tested for differential editing between asthma cases and controls adjusted for age, sex, current smoking status and the first three ancestry PCs. Pairwise comparisons between control and asthma severity groups using STEP classification (*76*) were performed using the makeContrasts function in limma.

2.5.7 Target gene prediction and pathway analysis

Bioinformatic prediction of target genes was performed for the unedited and edited miR-200b-3p using TargetScanHuman5.2 and TargetScanHuman5.2 Custom, respectively (*69*). For the unedited miR-200b-3p, *in silico* target genes were identified using a search for the unedited "miR-200b-3p". For the edited miR-200b-3p, *in silico* target genes were identified using a search for the seed sequence at positions 2-8 including the edited base, "AAT**G**CTG". Pathway analysis was performed using the TopFunn function within the ToppGene Suite for gene list enrichment analysis (*70*). Pathway enrichment analyses used a hypergeometric distribution with Bonferroni correction (p-value < 0.05). Venn diagrams of the genes and pathways of the unedited and edited forms of miR-200b-3p were generated using the VennDiagram package in R. Because TopFunn only reports the top 50 pathway enrichments, the unique pathways among the top 50 pathway enrichments of the unedited miR-200b-3p were used for comparison.

2.5.8 Analyses of gene expression

Target genes for the unedited and edited miR-200b-3p were bioinformatically predicted using TargetScanHuman5.2 and TargetScanHuman5.2 Custom, respectively (*69*). Pathway analysis was performed for the expressed target genes (CPM > 1 in at least 25% of 124 subjects with processed total RNA expression data) of the unedited and edited miR-200b-3p using the TopFunn function within the ToppGene Suite for gene list enrichment analysis (*70*).

The average normalized expression of expressed gene targets of the four enriched pathways of the edited miR-200b-3p was tested for differential expression between 83 asthma cases and 41 controls. For pathways reported from multiple databases, the pathway enrichment with the greatest enrichment of target genes was selected for downstream analyses. Average expression was measured as the mean over the standard deviation of normalized gene expression. Differential expression analysis was performed using a linear model in limma, adjusted for age, sex, current smoking status, sequencing pool, the first three ancestry PCs and 15 surrogate variables. Average expression and individual gene expression of the ten *in silico* gene targets of the edited miR-200b-3p that were enriched for either IL-4 signaling or IFN-gamma pathways were also tested for differential expression between asthma cases and controls and for differences between STEP classification groups using the makeContrasts function in limma.

The three ADAR enzymes, *ADAR*, *ADARB1* and *ADARB2*, were separately tested for associations with maternal asthma by interrogating three asthma risk groups: controls, asthma cases without a mother with asthma, and asthma cases with a mother with asthma. Four controls with a mother with asthma and five subjects without maternal asthma information were removed. Pairwise comparisons of expression between asthma risk groups were performed using the makeContrasts function in limma.

2.5.9 Correlations with measures of asthma

Asthma-related measures were tested for associations in my sample with both asthma and asthma severity (STEP classification (*76*)). To test differences between asthma cases and controls, continuous variables were evaluated with a Wilcoxon rank-sum test and categorical variables were evaluated with a Fisher's Exact Test. To test associations between all variables and asthma severity, ordinal logistic regressions were used, as implemented in the MASS package in R.

Spearman's rank correlation coefficients were used to assess correlations of A-to-I editing at the 5th position of miR-200b-3p and *SOCS1* expression with five clinical measures of asthma. These variables included total serum IgE, fractional exhaled nitric oxide (FeNO), bronchoalveolar lavage (BAL) eosinophilia, blood eosinophilia, and an epithelial cell signature of T2 inflammation that is a marker for T2-high asthma (*86*). The latter was measured as the sum of the normalized and covariate-adjusted gene expression levels of *POSTN*, *CLCA1*, and *SERPINB2* transcripts (T2 gene signature) (*75*). Asthma severity (STEP classification) was based on steroid use and lung function measures, so I did not include ICS (inhaled corticosteroid), OCS (oral corticosteroid), forced expiratory volume at 1 minute (FEV1) % predicted, or FEV1/Forced Vital Capacity (FVC) in these analyses. Age, sex, current smoking status and the first three ancestry PCs were regressed out from the % of A-to-I editing at the 5th position of miR-200b-3p, and the residuals were used to test for correlations with each clinical measure. I also tested the normalized and covariate-adjusted expression of *SOCS1* with the clinical measures. Significant correlations were determined after Bonferroni-corrections of the p-values, considering the five tests.

2.6 Supplementary information

2.6.1 Supplementary tables

Table S2.1. Nineteen A-to-I edited sites in bronchial epithelial cells from 138 subjects. The edited adenosine within the mature miRNA sequence is in bold.

miRNA	Position	Sequence (seed sequence)	Average Mapped Counts	Average % A-to-I Reads	# Subjects with Edited miRNA
miR-200b-3p	5	T<u>AAT</u>**A**<u>CTG</u>CCTGGTAATGATGA	75358	0.52	135
miR-191-5p	3	C<u>AA</u>**A**<u>CGGAA</u>TCCCAAAAGCAGCTG	58512	0.17	27
miR-200b-3p	3	T<u>AA</u>**A**<u>TACTG</u>CCTGGTAATGATGA	91380	0.15	22
miR-186-5p	3	C<u>AA</u>**A**<u>AGAAT</u>TCTCCTTTTGGGCT	24646	0.18	13
miR-148a-3p	3	T<u>C</u>**A**<u>GTGCA</u>CTACAGAACTTTGT	154711	0.14	6
miR-411-5p	5	T<u>AGT</u>**A**<u>GAC</u>CGTATAGCGTACG	372	5.40	5
miR-497-5p	2	C**A**<u>GCAGCA</u>CACTGTGGTTTGT	355	5.21	5
miR-421	14	A<u>TCAACAG</u>ACATT**A**ATTGGGCGC	663	0.89	5
miR-21-3p	3	C<u>A</u>**A**<u>CACCA</u>GTCGATGGGCTGT	22650	0.21	4
miR-376c-3p	6	A<u>ACAT</u>**A**<u>GA</u>GGAAATTCCACGT	27	34.09	3
miR-381-3p	4	T<u>AT</u>**A**<u>CAAG</u>GGCAAGCTCTCTGT	254	3.09	3
miR-200a-3p	3	T<u>AA</u>**A**<u>CACTG</u>TCTGGTAACGATGT	24079	0.19	2
miR-339-3p	15	T<u>GAGCGCC</u>TCGACG**A**CAGAGCCG	822	9.62	2
miR-451a	3	A<u>A</u>**A**<u>CCGTT</u>ACCATTACTGAGTT	36137	0.16	2
miR-191-5p	8	C<u>AACGGA</u>**A**TCCCAAAAGCAGCTG	56489	0.15	1
miR-20a-5p	4	T<u>AA</u>**A**<u>GTGC</u>TTATAGTGCAGGTAG	81	3.7	1
miR-561-5p	13	A<u>TCAAGGA</u>TCTT**A**AACTTTGCC	225	2.22	1
miR-589-3p	6	T<u>CAGA</u>**A**<u>CA</u>AATGCCGGTTCCCAGA	14	21.43	1
miR-664a-5p	8	A<u>CTGGCT</u>**A**<u>GGGAAAATGATTGGAT</u>	153	3.92	1

Table S2.2. Top 50 pathway enrichments for the gene targets of unedited miR-200b-3p. Only the 604 gene targets of unedited miR-200b-3p that were expressed in BECs (CPM > 1) were included as input for pathway analysis, as implemented in TopFunn. Thirty-nine unique pathways are listed.

ID	Name	Database	Bonferroni Adjusted P-value	Number of genes from input	Number of genes in annotation
M16763	Neurotrophin signaling pathway	MSigDB C2 BIOCARTA (v7.1)	1.53E-11	27	126
M186	PDGFR-beta signaling pathway	MSigDB C2 BIOCARTA (v7.1)	2.30E-10	26	129
101143	Neurotrophin signaling pathway	BioSystems: KEGG	2.06E-09	24	119
M48	Signaling events mediated by Hepatocyte Growth Factor Receptor (c-Met)	MSigDB C2 BIOCARTA (v7.1)	1.89E-08	19	79
137930	Signaling events mediated by Hepatocyte Growth Factor Receptor (c-Met)	BioSystems: Pathway Interaction Database	3.30E-07	17	72
M295	Genes related to PIP3 signaling in cardiac myocytes	MSigDB C2 BIOCARTA (v7.1)	8.95E-06	15	67
1269443	Signalling by NGF	BioSystems: REACTOME	9.03E-06	43	483
M281	Signaling events mediated by focal adhesion kinase	MSigDB C2 BIOCARTA (v7.1)	1.32E-05	14	59
1457794	Signaling by MET	BioSystems: REACTOME	1.38E-05	15	69
M121	mTOR signaling pathway	MSigDB C2 BIOCARTA (v7.1)	1.38E-05	15	69
1269380	Signaling by EGFR	BioSystems: REACTOME	1.46E-05	36	367
M2404	Mechanism of Gene Regulation by Peroxisome Proliferators via PPARa(alpha)	MSigDB C2 BIOCARTA (v7.1)	2.35E-05	13	52
169348	Signaling events mediated by focal adhesion kinase	BioSystems: Pathway Interaction Database	3.86E-05	13	54
83107	Renal cell carcinoma	BioSystems: KEGG	5.05E-05	14	65

Table S2.2. Top 50 pathway enrichments for the gene targets of unedited miR-200b-3p (continued).

1269654	Transcriptional Regulation by TP53	BioSystems: REACTOME	7.93E-05	35	374
1269507	Signaling by Rho GTPases	BioSystems: REACTOME	9.33E-05	38	430
M13266	Renal cell carcinoma	MSigDB C2 BIOCARTA (v7.1)	1.38E-04	14	70
M164	ErbB1 downstream signaling	MSigDB C2 BIOCARTA (v7.1)	1.45E-04	17	105
M187	Neurotrophic factor-mediated Trk receptor signaling	MSigDB C2 BIOCARTA (v7.1)	1.85E-04	13	61
138001	mTOR signaling pathway	BioSystems: Pathway Interaction Database	1.85E-04	13	61
1270302	Developmental Biology	BioSystems: REACTOME	2.22E-04	69	1081
1269460	NGF signalling via TRKA from the plasma membrane	BioSystems: REACTOME	2.25E-04	35	390
1404799	Endocrine resistance	BioSystems: KEGG	2.34E-04	16	96
1270303	Axon guidance	BioSystems: REACTOME	4.81E-04	43	554
138057	ErbB1 downstream signaling	BioSystems: Pathway Interaction Database	5.57E-04	16	102
694606	Hepatitis B	BioSystems: KEGG	7.02E-04	19	144
1269479	Downstream signal transduction	BioSystems: REACTOME	7.36E-04	32	355
1269478	Signaling by PDGF	BioSystems: REACTOME	1.06E-03	33	379
PW:0000578	scatter factor/hepatocyte growth factor signaling	Pathway Ontology	1.22E-03	6	11
M195	C-MYB transcription factor network	MSigDB C2 BIOCARTA (v7.1)	1.47E-03	14	84
83067	Focal adhesion	BioSystems: KEGG	1.88E-03	22	199
1269650	Generic Transcription Pathway	BioSystems: REACTOME	2.06E-03	57	879

Table S2.2. Top 50 pathway enrichments for the gene targets of unedited miR-200b-3p (continued).

1268855	Diseases of signal transduction	BioSystems: REACTOME	2.21E-03	32	373
P00047	PDGF signaling pathway	PantherDB	2.40E-03	17	127
83105	Pathways in cancer	BioSystems: KEGG	2.70E-03	33	395
1269284	DAP12 signaling	BioSystems: REACTOME	2.86E-03	31	359
83065	Axon guidance	BioSystems: KEGG	3.54E-03	20	175
138073	C-MYB transcription factor network	BioSystems: Pathway Interaction Database	3.71E-03	13	78
1383076	Regulation of TP53 Activity	BioSystems: REACTOME	4.37E-03	19	162
137977	Neurotrophic factor-mediated Trk receptor signaling	BioSystems: Pathway Interaction Database	4.40E-03	11	56
M19118	Keratinocyte Differentiation	MSigDB C2 BIOCARTA (v7.1)	4.88E-03	10	46
M100	SHP2 signaling	MSigDB C2 BIOCARTA (v7.1)	5.29E-03	11	57
M13863	MAPKinase Signaling Pathway	MSigDB C2 BIOCARTA (v7.1)	5.76E-03	13	81
137989	FGF signaling pathway	BioSystems: Pathway Interaction Database	6.01E-03	10	47
852705	MicroRNAs in cancer	BioSystems: KEGG	6.06E-03	27	299
1269283	DAP12 interactions	BioSystems: REACTOME	6.75E-03	31	374
M7253	Focal adhesion	MSigDB C2 BIOCARTA (v7.1)	7.11E-03	21	199
138036	FoxO family signaling	BioSystems: Pathway Interaction Database	7.37E-03	10	48
M136	FoxO family signaling	MSigDB C2 BIOCARTA (v7.1)	8.99E-03	10	49
P00018	EGF receptor signaling pathway	PantherDB	9.17E-03	15	111

2.6.2 Supplementary figures

Figure S2.1. A-to-I editing of position 5 of miR-200b-3p in asthma cases and controls excluding one outlier. Associations of percentage (%) of A-to-I editing of position 5 of miR-200b-3p in bronchial epithelial cells with (**A**) asthma and (**B**) asthma severity groups based on STEP classification scores. One outlier (>3 SD) was included from the control group in **Figure 2.1**. The number of subjects is shown below risk group. Adjusted p-values are shown above the horizontal bars. The p-values for (**A**) and (**B**) were corrected for four and three tests, respectively, using a Bonferroni correction.

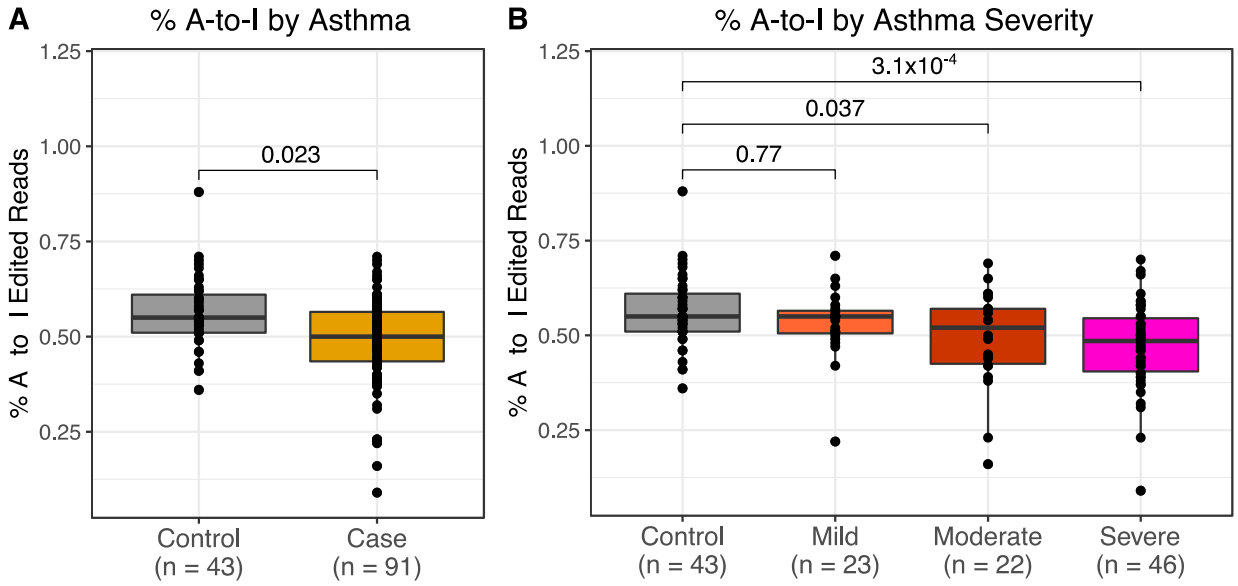


Figure S2.2. Differential editing analysis of the top four A-to-I edited sites. Comparisons of the percentage (%) of A-to-I editing at the 5th position of miR-200b-3p (top left), the 3rd position of miR-191-5p (top right), the 3rd position of miR-200b-3p (bottom left), and the 3rd position of miR-186-5p (bottom right) were made between 91 asthma cases and 44 non-asthma controls. Adjusted p-values are shown above the horizontal bars. P-values were corrected for four tests using a Bonferroni correction.

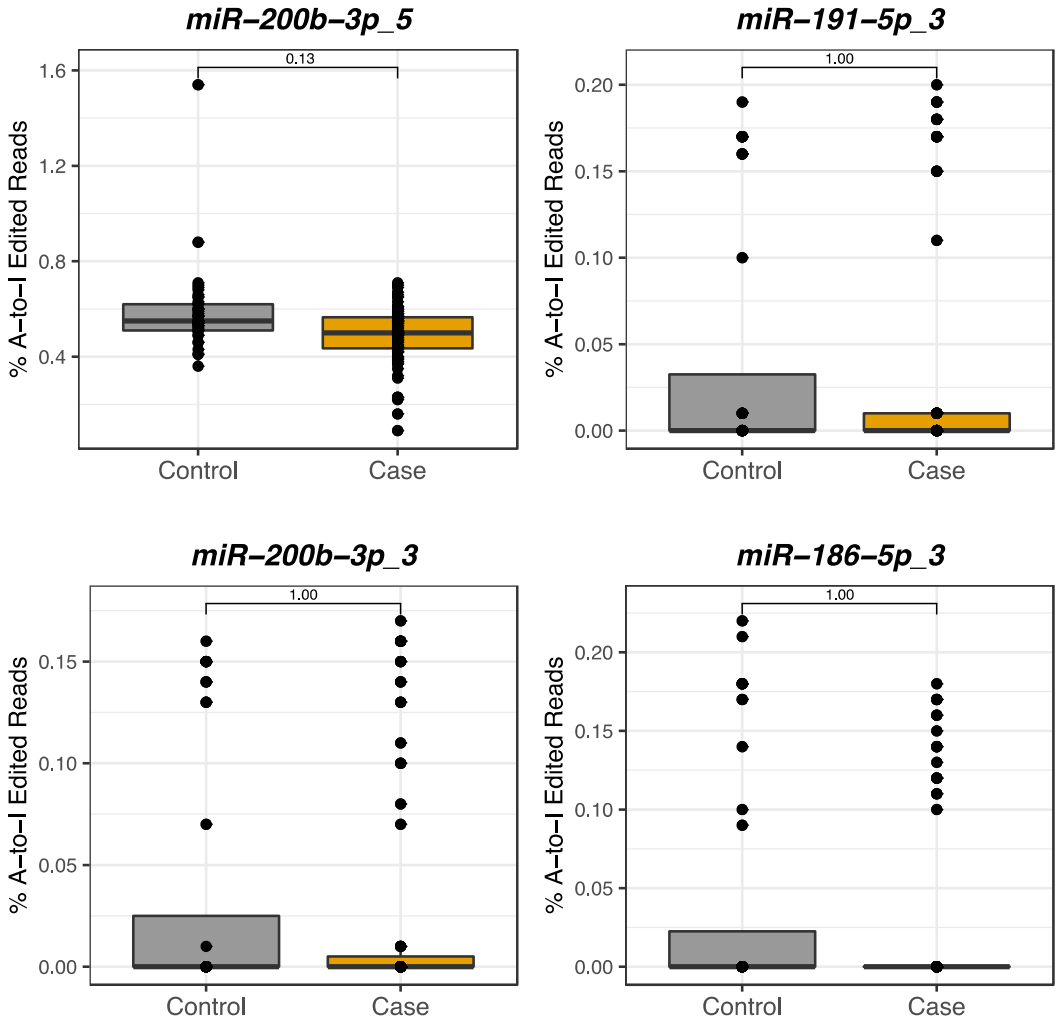


Figure S2.3. Normalized expression of ten gene targets for the two enriched pathways of the edited miR-200b-3p in 83 asthma cases and 41 controls. Gene targets that were expressed (CPM > 1) in BECs in at least 25% of the 124 subjects with RNA-seq were retained for analysis. Ten gene targets of the "IL-4 mediated signaling events" and "IFN-gamma" pathways are displayed. Adjusted p-values are shown above the horizontal bars. P-values were corrected for ten tests using a Bonferroni correction.

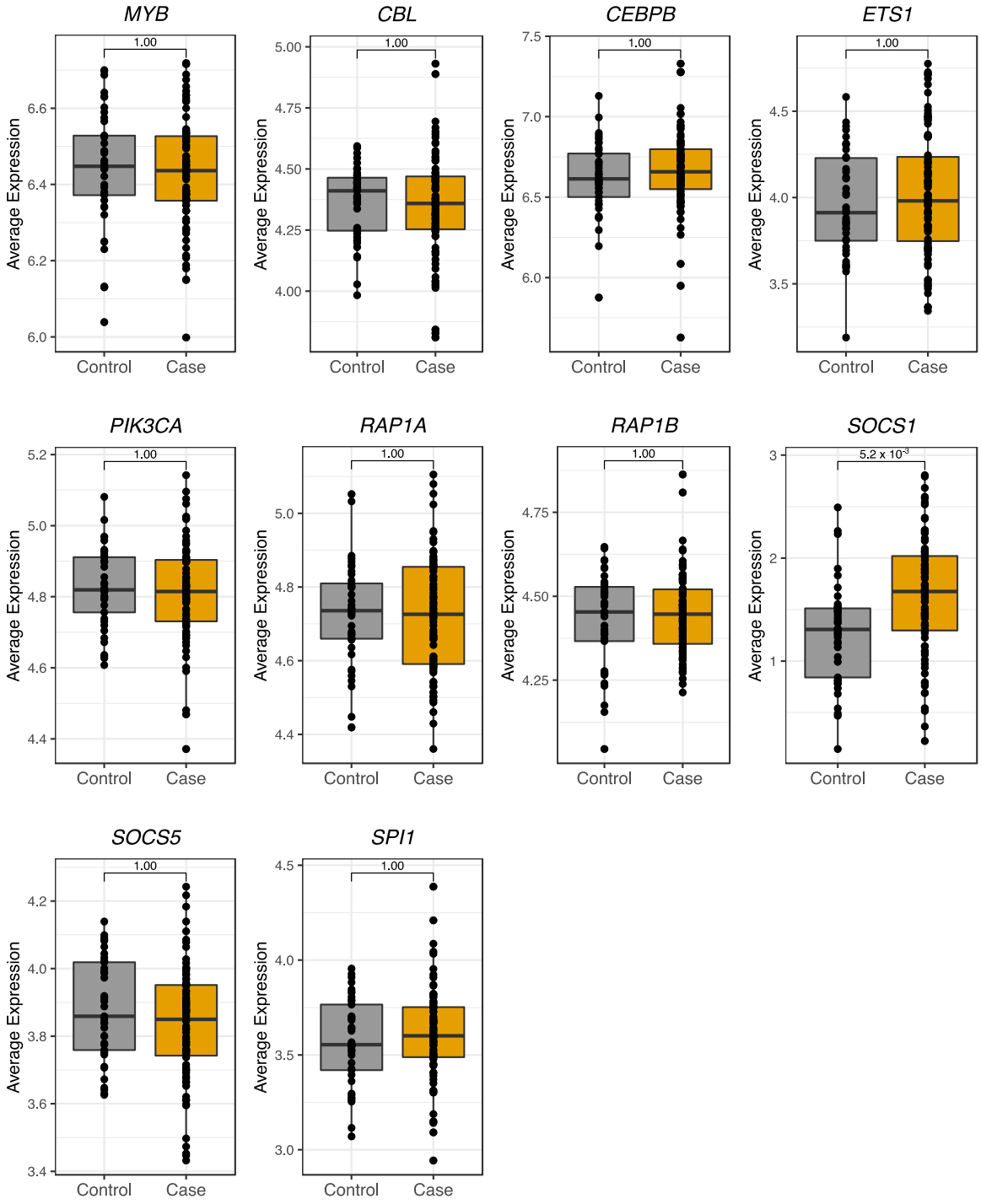
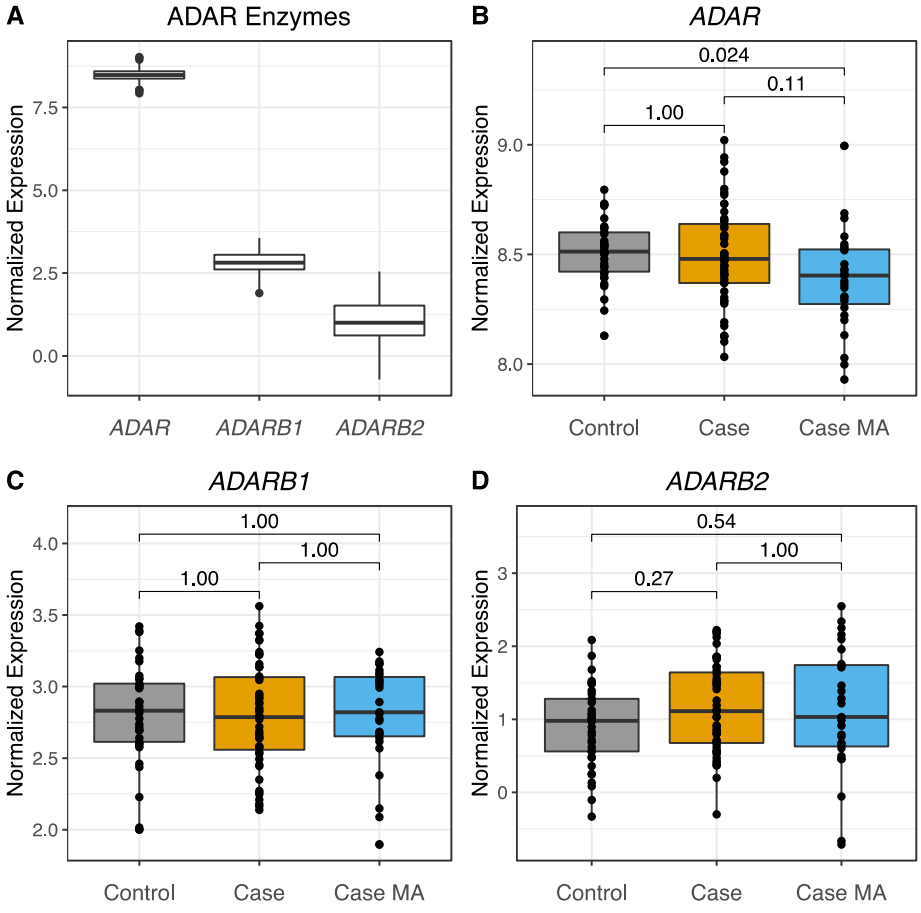


Figure S2.4. *ADAR* gene expression in bronchial epithelial cells (BECs) and in asthma cases and controls excluding two outliers. (**A**) Among 13,770 expressed genes in bronchial epithelial cells (CPM > 1%), median gene expression of *ADAR*, *ADARB1* and *ADARB2* (control) were at the 98th, 22nd, and 9th percentiles, respectively. Normalized expression of (**B**) *ADAR*, (**C**) *ADARB1*, and (**D**) *ADARB2* were compared between non-asthma controls (n = 38), asthma cases without a mother with asthma (n = 51) and asthma cases with a mother with asthma (Case MA; n = 28). One outlier of *ADAR* and *ADARB1* expression (>3 SD) were included from the Control and Case MA groups, respectively, in **Figure 2.5**.



CHAPTER 3

Maternal asthma-associated DNA methylation signatures in airway cells reflect subtypes of severe asthma

3.1 Abstract

Maternal asthma (MA) is one of the most consistent risk factors for childhood-onset asthma. A possible mechanism for this observation is through epigenetic modifications *in utero* that have lasting effects on gene regulation in children of asthmatic mothers. Here, I performed differential DNA methylation analyses of 398,186 CpGs in primary bronchial epithelial cells (BECs) between non-asthma controls (N=42) and all asthma cases (N=88), cases without MA (N=56), and cases with MA (N=32). Only 34 CpGs were differentiated methylated (DMCs) in cases without MA, whereas 1,428 were DMCs in the cases with MA, compared to controls. To examine the correlation structure of these 1,428 DMCs, referred to as MA-DMCs, I performed weighted gene co-expression network analysis, which grouped 935 (65.48%) of the MA-DMCs into six co-methylation modules. Correlations were assessed between the module eigenvectors and the 13,757 genes expressed in the same BECs and eight clinical measures. All of the MA-associated co-methylation modules were associated with severe asthma and lung function measures. Five modules were additionally associated with features of atopic asthma, including total serum IgE and eosinophilia. One module was also associated with exhaled nitric oxide and BMI, linking airway inflammation and obesity, respectively, to atopy and severity. One module was not associated with atopic asthma features, and its correlated genes were enriched for diverse immune-related processes and diseases. Overall, these data suggest that the MA *in utero* or early life environment alters diverse epigenetically-mediated developmental pathways that lead to subtypes of severe asthma in adulthood, including hard-to-treat, non-atopic asthma.

3.2 Introduction

Asthma is the most common chronic disease in childhood, affecting 5.5 million children and 19.2 million adults in the United States, with annual health care costs exceeding $80 billion (*94*). Asthma is a heterogeneous disease with significant inter-individual in degree of airway inflammation and response to therapies. Nearly 15% of individuals with asthma suffer from severe manifestations of the disease, including hard-to-control symptoms, persistent exacerbations, and poor response to standard asthma treatments (*3*). Among the many epidemiologic risk factors for asthma, having a mother with asthma is one of the most reproducible (*49*), although the mechanism(s) underlying this risk are unknown.

It has been suggested that the effects of maternal asthma (MA) on subsequent risk for asthma in her children result from exposure to a unique environment in pregnancies of asthmatic mothers that modifies the epigenetic landscape of the fetus and impacts downstream developmental programs and subsequent risk for asthma in the child (*95*). This idea has been supported by previous studies (*55, 56, 59, 96*). For example, in the Infant Immune Study, a longitudinal birth cohort study, DeVries et al. first showed 589 differentially methylated regions between cord blood immune cells from 18 children with MA who developed asthma by age 9 compared to 18 children with MA who did not have a diagnosis of asthma by age 9 (*59*). The genes near the differentially methylated regions clustered in immunoregulatory and proinflammatory pathways that were most connected by the transcription factor SMAD3, whose gene, *SMAD3*, resides within a highly replicated asthma genome-wide association study (GWAS) locus (*97*). They next showed that the promoter of *SMAD3* was hypermethylated in cord blood cells from asthmatic children with MA compared to asthmatic children without MA from IIS and two additional birth cohorts (*59*). These findings support the idea that MA modifies childhood asthma

risk through epigenetic mechanisms, which are detectable as early as birth. Yet, no previous studies have examined the effects of MA in airway cells or in adult children of asthmatic and non-asthmatic mothers, or included paired gene expression data to empirically characterize the transcriptional pathways associated with epigenetic programs in the child.

Here, I present the first epigenome-wide analysis of MA and DNA methylation (DNAm) in primary bronchial epithelial cells (BECs) from 96 adult asthma cases and 46 adult non-asthmatic controls. I identified 1,428 CpGs that were differentially methylated only in asthma cases with MA, but not in cases with no maternal asthma (NMA), compared to controls. These CpGs clustered into six co-methylation modules. The eigenvalues of all six modules were correlated with asthma severity in addition to module-specific sets of expressed genes and clinical measures of asthma. I highlight one module that was not correlated with features of atopic asthma but was correlated with genes enriched for diverse immune-related processes and diseases, including 25 genes that were differentially expressed between cases with and without MA. Overall, I describe MA-associated DNAm signatures that reflect endotypes of severe asthma, which may be epigenetically programmed *in utero* or in early life, with long-lasting effects on gene regulatory programs in the lower airways.

3.3 Results

3.3.1 Identification of maternal asthma-associated differentially methylated CpGs (MA-DMCs)

Our lab obtained array-based DNAm measures from freshly isolated BECs from 150 adults who were evaluated at the University of Chicago Asthma Clinics. I retained 130 subjects (42 non-asthma controls, 56 asthma cases without MA, and 32 asthma cases with MA) who passed DNAm QC and had MA information (**Table 3.1**). The cases with and without MA were similarly different

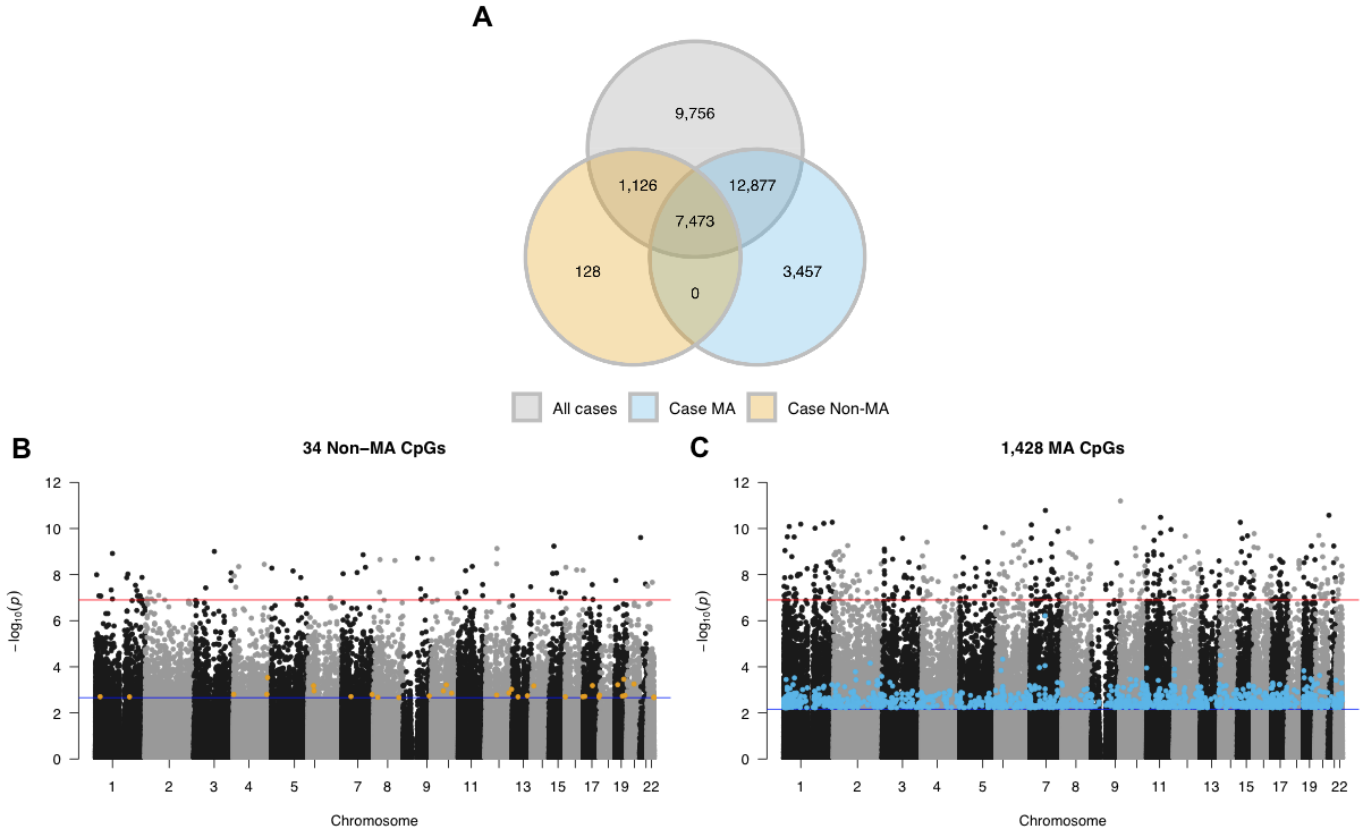
from the controls with respect to most clinical measures, except only the MA cases had significant higher serum IgE levels compared to controls. Overall, median IgE levels, a marker of allergic sensitization (atopy) and atopic asthma (*98*), were highest in the MA cases (156.50 IU/mL) compared to NMA cases (117.50 IU/mL) and non-asthma controls (56.50 IU/mL). This suggests that adult asthma cases with MA may be enriched for atopic asthma compared to those without MA.

Table 3.1. Characteristics of 130 subjects at the time of bronchoscopy by asthma risk group. Severity was defined by STEP classification of mild, moderate and severe asthma. MA, Maternal Asthma; Af Am, African American; Eur Am, European American; ICS, Inhaled Corticosteroid; OCS, Oral Corticosteroid; FVC, Forced Vital Capacity. Asthma cases with no maternal asthma (Case NMA) and cases with maternal asthma (Case MA) were each compared to controls (*p<0.05, **p<0.01, ***p<0.001). Continuous variables were tested with a Wilcoxon rank-sum test and categorical variables were tested using a Fisher's Exact Test. [A]52 asthma cases with no maternal asthma and 30 cases with maternal asthma had FeNO measurements. [B]41 controls and 55 asthma cases with no maternal asthma had BAL eosinophil measurements. [C]41 controls had blood eosinophil measurements.

	Control (N = 42)	**Case NMA** (N = 56)	**Case MA** (N = 32)
Age (mean yr ± SD)	37.45 ± 11.75	41.41 ± 12.24	37.56 ± 13.39
Gender (% female)	62	79	66
Ethnicity (%) (Af Am/Eur Am/ Other)	64/29/7	55/45/0**	66/34/0*
% smoker at bronchoscopy	17	4*	6
ICS use (%)	-	75	75
OCS use (%)	-	45	25
STEP Category (% mild/% moderate/% severe)	-	22/23/55	28/25/47
Mean FEV1% Predicted (± SD)	95.36 ± 11.63	75.45*** ± 19.73	70.41*** ± 18.15
Mean FEV1/FVC (± SD)	0.82 ± 0.05	0.73*** ± 0.10	0.83*** ± 0.48
Median total serum IgE (IU/mL) (lower, upper quartile)	56.50 (22.00, 169.00)	117.50 (22.00, 305.25)	156.50** (70.25, 626.75)
Median FeNO (ppb) (lower, upper quartile)[A]	14.00 (10.50, 17.75)	23.50** (13.00, 45.75)	27.00** (13.25, 54.75)
Median BAL eosinophilia (%) (lower, upper quartile)[B]	0.000 (0.000, 0.004)	0.029*** (0.015, 0.056)	0.036*** (0.015, 0.071)
Median blood eosinophilia (cells/µl) (lower, upper quartile)[C]	0.018 (0.010, 0.029)	0.023 (0.010, 0.048)	0.031* (0.016, 0.049)
Mean Body Mass Index (± SD)	28.52 ± 5.71	34.64*** ± 8.83	33.68* ± 10.48

I first identified MA-associated differentially methylated CpGs (DMCs) by comparing DNAm levels between non-asthma controls (N=42; none with MA) to all asthma cases (N=88), to MA cases (N=56), and to NMA cases (N=32). All analyses included age, sex, current smoking status, and the first three genetic ancestry PCs to account for population stratification and ancestry imbalances between groups. These three comparisons revealed 31,232, 8,727 and 23,807 DMCs, respectively (FDR<10%) (**Figure 3.1A**). Among the 9,756 DMCs that were significant (FDR<10%) only in the analysis including all cases, 66.85% (N=6,522) had larger effects and smaller p-values in the MA cases compared to NMA cases, despite the smaller sample size of MA cases (Paired signed-rank test; $P<2.2x10^{-16}$) (**Figure S3.1**). Even at a stringent Bonferroni threshold, there were significantly more DMCs in the analysis of MA cases (N=446) compared to the analysis of NMA cases (N=76) (Exact binomial test; $P<2.2x10^{-16}$). Among the 128 CpGs that were differentially methylated between controls and NMA cases at an FDR<10%, only 34 CpGs (26.6%; orange points in **Figure 3.1B**) did not differ between cases and controls in the other two analyses at an FDR>20%. In contrast, among the 3,457 CpGs that were differentially methylated between controls and MA cases at an FDR<10%, 1,428 CpGs (41.3%; blue points in **Figure 3.1C**) did not differ between cases and controls in the other two analyses at an FDR>20%. I refer to these 1,428 DMCs as "maternal asthma-associated differentially methylated CpGs" (MA-DMCs) and consider nearly all the remaining 33,389 DMCs to be shared across study groups. These data indicate that global DNAm levels in BECs differ between adult asthma cases with and without MA, consistent with the hypothesis that exposure to a MA environment *in utero* or in early life has long-lasting effects on the epigenetic landscapes in cells from exposed children.

Figure 3.1. Comparisons of differentially methylated CpGs (DMCs) in maternal asthma risk groups. (**A**) Genome-wide differential methylation analyses were performed for controls (n = 42) vs. all asthma cases (n = 88; All cases; silver), controls vs. asthma cases with no maternal asthma (n = 56; Case NMA; yellow) and controls vs. asthma cases with maternal asthma (n = 32; Case MA; blue). These comparisons revealed 31,232, 8,727 and 23,807 DMCs, respectively (FDR < 10%). Manhattan plots show results of 398,186 CpGs compared between non-asthma controls to (**B**) cases NMA and (**C**) cases MA. P-values (y axis) correspond to the differences in methylation between asthma cases and non-asthma controls. Red horizontal lines correspond to the Bonferroni-adjusted p-value threshold ($p = 1.3x10^{-7}$) and blue horizontal lines correspond to the q value threshold (FDR 10%). Orange points in (B) indicate the 34 CpGs that were differentially methylated in cases NMA at FDR < 10% but not in all cases or in cases MA at FDR > 20%. Blue points in (C) indicate the 1,428 CpGs that were differentially methylated in cases MA at FDR < 10% but not in all cases or in cases NMA at FDR > 20%.



3.3.2 Methylation signatures and gene expression pathways

Table 3.2. Maternal asthma-associated co-methylation module eigenvectors. Of the 1,428 maternal asthma-associated CpGs, 935 (65.48%) formed six co-methylation modules (32 to 344 CpGs in each module). See **Figure S3.2** for a cluster dendrogram of the six modules. (**A)** The number of genes correlated (|rho| > 0.20) with each module eigenvector, the number of module-specific genes (|rho| > 0.20 only in that module), and the most significant pathway of the module-specific genes identified in TopFunn. **Table S3.1** lists all the pathways enriched among the module-specific genes in each module. (**B)** Correlation coefficients and p-values are shown between clinical phenotypes and module eigenvectors. Significant associations after correcting for eight variables are shown in bold font. The numbers of subjects with measurements for each variable is shown. For all correlations with P > 0.05, ns is noted.

Variables	**WGCNA co-methylation modules (number of CpGs)**					
	Black (48)	Blue (304)	Green (140)	Magenta (32)	Red (76)	Turquoise (344)
A. Correlated Genes and Pathway Enrichments						
Number of correlated genes	1235	1035	827	1969	881	2180
Number of module-specific genes	92	374	51	703	100	762
Number of pathways	0	41	0	30	0	0
Most significant pathway	-	Interferon signaling	-	Selenocysteine synthesis	-	-
B. Clinical Phenotypes						
Asthma Severity (N = 142)	**3.56** **$5.1x10^{-4}$**	**-3.90** **$1.5x10^{-4}$**	**-3.34** **$1.1x10^{-3}$**	**3.47** **$7.1x10^{-4}$**	**2.97** **$3.5x10^{-3}$**	**-4.77** **$4.6x10^{-6}$**
FEV1% Predicted (N = 142)	-0.22 $9.2x10^{-3}$	**0.29** **$5.4x10^{-4}$**	0.20 0.017	**-0.24** **$4.1x10^{-3}$**	**-0.23** **$6.2x10^{-3}$**	**0.33** **$6.5x10^{-5}$**
FEV1/FVC (N = 142)	ns	0.19 0.025	ns	ns	ns	ns
Total Serum IgE (N = 141)	**0.23** **$5.2x10^{-3}$**	ns	-0.19 0.027	0.20 0.015	0.17 $8.0x10^{-3}$	**-0.28** **$7.4x10^{-4}$**
FeNO (N = 135)	ns	ns	ns	**0.24** **$4.7x10^{-3}$**	ns	ns
BAL Eosinophilia (N = 140)	0.18 0.031	ns	ns	0.22 $8.1x10^{-3}$	ns	-0.22 $8.7x10^{-3}$
Blood Eosinophilia (N = 141)	ns	ns	ns	ns	ns	ns
Body Mass Index (N = 142)	ns	ns	ns	**0.31** **$2.1x10^{-4}$**	ns	ns

To characterize the correlation structure of the 1,428 MA-DMCs and further examine their relationship to gene expression in BECs, I used a systems biology approach, as implemented in Weighted Gene Co-Expression Network Analysis (WGCNA) (*99*). WGCNA grouped 935

(65.48%) of the MA-DMCs into six co-methylation modules, with 32 to 344 CpGs in each module (**Table 3.2**; **Figure S3.2**). There were too few CpGs (N=34) that were differentially methylated only between controls and NMA cases to perform WGCNA with those CpGs.

Among the 13,757 genes that were detected as expressed in BECs, 4,025 genes were correlated (|rho|>0.20) with the eigenvector of one or more modules. I then defined a set of uniquely correlated genes for each module as those that had |rho|<0.20 with the eigenvectors of the other five modules. I refer to these 2,082 genes as module-specific genes, which were then subjected to pathway analyses (TopFunn (*70*)) (**Table 3.2A**). The genes in two of the modules (blue and magenta) were correlated with distinct pathways that have known roles in asthma pathobiology (**Table S3.1**). The blue module was correlated with 374 unique genes that were significantly enriched in 41 pathways after Bonferroni correction, with the most significant being "Interferon signaling" ($P_{adjusted}=1.52x10^{-10}$; 26/202 genes), that itself included *ADAR*, an interferon-inducible gene (*100*). "Interferon signaling" is an important pathway involved in viral responses, which is often impaired in individuals with asthma, particularly those with allergic asthma (*101*). The other enriched pathways of the blue module genes were also involved in immune responses and immune-mediated diseases, including asthma, inflammatory bowel disease, and Type I diabetes mellitus. The magenta module was correlated with 703 unique genes that were significantly enriched in 30 pathways after Bonferroni correction, with the most significant being "Selenocysteine synthesis" ($P_{adjusted}=8.99x10^{-15}$; 28/96 genes), involved in the production of antioxidant selenoproteins (*102*), which are potent anti-oxidants that augment oxidative stress and influence immune responses (*103, 104*). "Selenocysteine synthesis" and 28 other pathways consisted primarily of genes that encode ribosomal proteins, which have significant roles in protein synthesis, cell growth, apoptosis and immunosurveillance (*105*). All of the pathways enriched for

MA-DMC-correlated genes in the blue and magenta modules are shown in **Table S3.1**. The correlated genes that were specific to the black, green, red and turquoise modules were not enriched for any pathways. Overall, these results suggest that MA-associated DNAm patterns regulate diverse gene expression networks, potentially reflecting different mechanistic pathways (endotypes) of asthma.

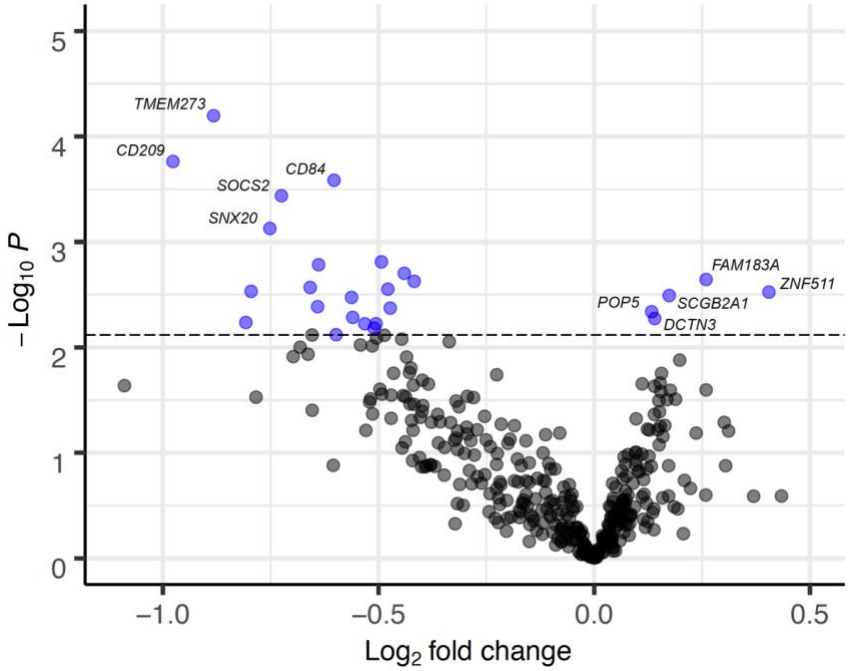
3.3.3 Methylation signatures and clinical measures of asthma severity

To test directly the hypothesis that the MA-DMCs in different modules underlie distinct clinical features and endotypes of asthma, I examined correlations between the eigenvectors of the six MA-DMC-associated modules and eight clinical measures. This revealed six distinct patterns of correlated phenotypes (**Table 3.2B**). Surprisingly, the eigenvectors of all six modules of MA-DMCs were significantly correlated with asthma severity, defined by STEP classification of mild, moderate and severe asthma (*76*). All six modules were correlated with forced expiratory volume at 1 second (FEV1) % predicted, a measure of lung function that is used in the STEP classification, but only correlations with the blue, magenta, red and turquoise modules were significant after correction for multiple testing. The blue module, which was enriched for interferon and other immune response genes, was nominally associated with the FEV1/FVC ratio but not with any other measures. In contrast, the magenta module, which was enriched for selenocysteine synthesis genes, was additionally significantly correlated with fractional exhaled nitric oxide (FeNO), a marker of airway inflammation, and BMI, and with total serum IgE and BAL eosinophilia at nominal levels of significance. Two of the modules (black and turquoise) were significantly correlated with total serum IgE, and nominally with BAL eosinophilia. Thus, whereas all modules of MA-associated DMCs are associated with asthma severity, only the blue module is not associated with features of allergic or atopic asthma. Taken together, these studies establish potential mechanistic links

between MA-associated DNAm patterns in BECs with clinical measures of asthma severity, reflecting both atopic and non-atopic endotypes of severe asthma.

3.3.4 A non-atopic maternal asthma-associated endotype

Figure 3.2. Volcano plot of the 25 differentially expressed (DE) genes of the blue module. Of 374 genes that were correlated only with the blue module eigenvector (|rho| > 0.20), 25 were DE (blue points) between asthma cases with maternal asthma (MA; N = 30) and no maternal asthma (NMA; N = 51). The horizontal line corresponds to the q value threshold (FDR < 10%). See **Figure S3.3** for boxplots of each gene and **Table S3.2** for complete summary statistics and citations of representative studies.



I next asked whether any of the 2,082 genes that were correlated with only one module of MA-DMCs (|rho|>0.20) were also differentially expressed (DE) between MA and NMA cases (MA-DE). I identified one MA-DE gene (of 92 genes) in the black module and 25 MA-DE genes (of 374 genes) in the blue module (FDR<10%) (**Table S3.2**). None of the module-specific genes correlated with the green, magenta, red and turquoise eigenvalues were DE between MA and NMA cases. The one MA-DE gene correlated with the black module was *SAMSN1* (SAM Domain, SH3 Domain and Nuclear Localization Signals 1) ($P_{adjusted}$=0.042), a negative regulator of B-cell

activation that was previously identified as a DE gene in nasal epithelial cells from atopic asthma cases and non-atopic healthy controls (*106*).

Of the 25 MA-DE genes correlated with the blue module, twenty were downregulated and five were upregulated in cases with MA compared to cases without MA (**Figure 3.2; Figure S3.3**). Remarkably, 23 of the 25 MA-DE genes have previously been implicated in asthma (see **Table S3.2** for citations of representative studies); nine of the downregulated genes were in 17 of the 40 enriched pathways (one to five pathways for each gene) (**Table S3.3**). Four downregulated MA-DE genes (*WIPF1, IKZF1*, *GNG2*, *DOCK2*) were also DE genes in a study comparing peripheral blood leukocytes of children with severe therapy-resistant asthma compared to children with controlled persistent asthma (*107*). Six other MA-DE genes (1 downregulated [*TMEM273*] and 5 upregulated [*FAM183A*, *ZNF511*, *SCGB2A1*, *POP5*, *DCTN3*]) were reported as IL-6-inducible genes that were also not stimulated with IL-4 and IL-13, two canonical Type 2 (T2) cytokines (*108*). These results indicate that the MA-DE genes of the blue module are associated with different endotypes of asthma, including severe therapy-resistant and non-atopic forms.

3.4 Discussion

It has been proposed that the prenatal environment is critical in establishing risk for asthma with onset in childhood (*96, 109-114*). Maternal asthma is itself one of the most replicated risk factors for the development of asthma in her child (*49*). In an earlier study, in a subset of the subjects included in this study, I reported differential expression of a microRNA that downregulated the expression of an asthma-associated gene, *HLA-G*, in bronchial epithelial cells from adults with MA but not in bronchial epithelial cells from adults without MA (*55*). I suggested then that prenatal or early life exposures in fetuses of asthmatic mothers have lifelong effects on the regulatory landscape of bronchial epithelial cells. In this study, I extended this mechanism to include a

widespread epigenetic mark, DNAm, and to a genome-wide survey. My study revealed global DNAm patterns that differ between BECs from adult asthmatics with and without MA. Co-methylated sets of MA-DMCs and their correlated genes were further associated with distinct sets of genes, biological pathways, and features of severe asthma.

I identified five epigenetic signatures of MA-DMCs that were associated with severe asthma, and some with measures of allergic sensitization. One signature (magenta module) was additionally associated with exhaled NO and BMI, linking airway inflammation and obesity to asthma severity. The DNAm signature of this module was also correlated with genes enriched in selenocysteine synthesis, a pathway that produces antioxidant selenoproteins (*102*). Dietary selenium has been suggested to ameliorate oxidative stress in the lower airways of individuals with atopic asthma (*104, 115*). This idea has been supported by the association of reduced selenium with the development and/or severity of asthma (*116, 117*), although this association has been poorly replicated (*118, 119*). Possible explanations for these conflicting studies are the inclusion of asthma cases with variable clinical manifestations of atopic disease and potential different effects of MA. Previous studies (*120-122*) have suggested that prenatal selenium supplementation may reduce risk of childhood asthma, but did not examine the effects of MA on supplementation effects. My study further suggests that these beneficial effects may be most effective in children of asthmatic mothers.

I also identified an epigenetic signature (blue module) of MA-DMCs that was associated with severe asthma but not with measures of allergic sensitization. This is consistent with the genes that were associated with the blue module eigenvector, which were enriched for diverse immune-related pathways that did not include the canonical T2 cytokines or signaling molecules (*75*). The presence of multiple interferon signaling molecules further supports the idea that impaired innate

immune responses to viral infections lead to damaged epithelium, asthma exacerbations and more severe asthma (*101*). The blue module eigenvector was also associated with 25 module-specific genes that were differentially expressed in bronchial epithelial cells from asthma cases with MA compared to asthma cases without MA. Among the 25 MA-DE genes of the blue module, 23 have been linked to asthma or allergies in previous studies. The expression for four of the downregulated MA-DE genes was also differentially expressed in peripheral blood leukocytes from cases with severe therapy-resistant compared to cases with controlled persistent asthma (*107*). Another study showed that six other MA-DE genes (5 upregulated; 3 downregulated in my study) formed part of a transcriptional signature that defined a subset of asthmatic patients with impaired lung epithelial IL-6 trans-signaling in the absence of T2 inflammation (*108*). All five blue module genes that were upregulated in cases with MA in my study were also upregulated after stimulation with IL-6, but not with the canonical T2 cytokines, IL-4 or IL-13 (*108*), suggesting an additional link between maternal asthma and IL-6 signaling. Combined with results of published studies, I interpret the results of my study to suggest that the CpGs in the blue module represents a DNAm signature in bronchial epithelial cells that perturbs immunomodulatory gene networks in adult children of asthmatic mothers. Moreover, this signature may underlie a severe, hard-to-treat non-atopic asthma endotype through T2-independent pathways that reflect both altered interferon signaling and IL-6 trans-signaling activation.

Finally, I provide evidence for the link between two maternal asthma-associated epigenetic mechanisms: ADAR-mediated editing and DNA methylation. In chapter 2, I demonstrated that the expression of the primary A-to-I editing enzyme, *ADAR*, was decreased only in asthma cases with an asthmatic mother compared to controls. Here, I found *ADAR* among the blue module-specific genes of the most enriched pathway, “Interferon signaling”. None of the CpGs in the *ADAR* gene

were part of the blue module, suggesting that interferon-inducible *ADAR* is a downstream target of other genes in this pathway. This finding indicates that maternal asthma-associated effects on ADAR-mediated editing in BECs may be mediated by a distinct DNA methylation profile.

The results of my study support my hypothesis that the epigenetic landscape in bronchial epithelial cells differs between adult asthmatics with and without MA, however, there are limitations. First, I did not have a sufficient number of controls with MA to assess whether the observed MA effects on DNAm patterns would be similar or not in controls, or to disentangle the effects of MA and asthma in the child. Second, I used a modest threshold ($|rho| < 0.20$) to identify module-specific correlated genes. Because the six MA-associated modules were formed from CpGs that were altered only in cases with MA, the module eigenvectors were highly correlated with each other and with many of the same genes. Despite this less stringent threshold, the enriched pathways of the module-specific genes and most of the MA-DE genes of the black and blue modules have been previously implicated in asthma, and subsets of the DE genes in this module are supported by previous published studies. Together, these results suggest that I captured relevant genes that were module specific, although they require validation in independent studies. Third, because this was a cross-sectional study with all assessments performed in adulthood, I was not able to determine the longitudinal development of MA-associated DNA methylation patterns and their effects on gene expression, prior to the inception of asthma. Thus, I could not directly show that the DNA methylation signatures were the cause or effect of severe asthma in the cases. However, the fact that these signatures differed between asthma cases with MA compared to controls, but not between asthma cases without MA compared to controls, suggests that they reflect early life, or *in utero*, exposures that also differ between these groups of asthmatics. Addressing these important questions in longitudinal studies will provide insights into MA effects on both

epigenetic and transcriptional programs and their associations with the clinical features of severe asthma.

In summary, my study of DNAm profiles in lower airway cells revealed distinct patterns in children of asthmatic mothers that may mediate the prenatal or early life effects of MA on immunomodulatory gene expression pathways and subsequent risk of severe asthma symptoms. My data further indicate that these effects persist into adulthood, reflecting long-lasting consequences of an *in utero* or early life environment that is specific to MA. Overall, I suggest that the MA environment alters epigenetically-mediated developmental pathways that lead to subtypes of endotypes of asthma in adulthood that include both atopic and non-atopic endotypes.

3.5 Methods

3.5.1 Study design

Adults were recruited and clinically evaluated before undergoing bronchoscopy at the University of Chicago Medical Center between March 2010 and March 2014. Endobronchial brushings for these studies were obtained during bronchoscopy, as previously described (*55, 123*). Subjects with asthma had a current doctor's diagnosis of asthma, no conflicting pulmonary diagnoses, and were using asthma medications. Controls had no current or previous diagnosis of asthma, normal spirometry, and no evidence of bronchial hyperresponsiveness by methacholine challenge tests. A total of 169 subjects had either available genotypes for estimating ancestry principal components (N=162), RNA sequences (N=128), or array-based DNAm measurements (N=150). Processing and quality control of the genotyping and RNA-sequencing were described in Chapter 2 (see **Methods in 2.5**). A flowchart of the sample selection for this study is shown in **Figure S3.4**. These studies were approved by the University of Chicago's Institutional Review Board. Written informed consent was obtained from all research participants.

3.5.2 Processing and quality control of DNA methylation arrays

DNA was isolated and purified from epithelial cells using the QIAzol lysis reagent or the AllPrep DNA/RNA/miRNA Universal Kit (QIAGEN). DNAm profiles were measured on the Infinium Human Methylation 450K Bead Chip (N=482,421) or the Illumina Infinium MethylationEPIC BeadChip (N=853,307) at the University of Chicago Functional Genomics Core. Only probes that were present on both arrays were retained using the combineArrays function in the minfi package (*124*) (N=398,186). Probes located on the sex chromosomes and those with detection p-values greater than 0.01 in more than 10% of samples were removed from the analysis. Cross-reactive probes and probes within two nucleotides of a SNP with a MAF greater than 0.05 were removed using the function rmSNPandCH() from the R package DMRcate (*125*), retaining 398,186 probes for preprocessing. A preprocessing control normalization function was applied to correct for raw probe values or background and single-sample Noob (ssNoob) was used to correct for technical differences between the Infinium type I and type II probes and across the two generations of Infinium methylation arrays (*126*). Finally, eight outlier samples (>2.5 absolute deviations from the median of methylation) were removed, retaining DNAm profiles for 142 subjects and 398,186 CpGs for downstream analyses. PCA was used to determine the effects of known confounding variables on global DNAm profiles. Chip, ancestry, age, and sex were significantly correlated with at least one PC. The effects of chip were regressed out using COMBAT (*126*).

3.5.3 Statistical analyses

Differential methylation analysis was performed using a linear model in limma (*93*), corrected for age, sex, current smoking status, and the first three ancestry PCs. Three differential methylation analyses were performed to identify DMCs: non-asthma controls without MA were compared to all asthma cases, to cases without MA, and to cases with MA. I required DMCs to be significant

in each group at FDR<10%. Eight subjects without MA information and four control subjects with MA were excluded from these analyses, retaining 130 subjects.

The CpGs that were differentially methylated only between controls and cases with MA (FDR<10%) but not in the other two groups (FDR>20%) were clustered into co-methylation modules using WGCNA (*99*). The soft thresholding power was 10; all other settings were kept at default values. Modules that had highly correlated eigenvectors were merged using the mergeCloseModules function (**Figure S3.2**). Spearman's rank correlation coefficient was used to assess correlations between the eigenvectors for each co-methylation module with all genes that were expressed in BECs (CPM>1 in at least 25% of 118 subjects with both DNAm and gene expression data) and eight clinical phenotypes. Only correlated genes (|rho|>0.20) that were not correlated with the eigenvectors of the other five modules (|rho|<0.20) were retained for pathway analysis using the TopFunn function within the ToppGene Suite (*70*). The correlated genes were required to have the same direction of effect in all subjects and in non-asthma controls. Significant pathways with Bonferroni-corrected p-values < 0.05 were retained. Age, sex, current smoking status and the first three ancestry PCs were regressed out from each of seven clinical phenotypes (excluding STEP severity) and the residuals were used to test for correlations with module eigenvectors. For STEP severity categories (control, STEP=0; mild, STEP=1,2; moderate, STEP=3,4; severe, STEP=5,6) (*76*), the eigenvectors were tested for association using an ordinal logistic regression correcting for the same covariates. Significant correlations were determined after Bonferroni-corrections of the p-values, considering eight tests.

For the correlated genes of each module, normalized and covariate-adjusted expression data was compared between asthma cases with (N=30) and without MA (N=51) using a linear model, as implemented in limma. Sequencing batch, age, sex, current smoking status, and the first

three ancestry principal components were regressed out from the gene expression data and the residuals were used to test for differences between cases with and without MA. Differentially expressed genes were detected with an FDR-correction (FDR<10%) and referred to as MA-DE genes.

3.6 Supplementary information

3.6.1 Supplementary tables

Table S3.1. Significant pathways for the genes that were correlated with the blue and magenta maternal asthma-associated co-methylation modules. The 374 and 703 genes that were correlated (|rho| > 0.20) only with the eigenvectors of the blue or magenta co-methylation modules, respectively, were included as input for pathway analysis (TopFunn (*70*)). Forty-one and thirty pathways were significantly enriched for the correlated genes of the blue and magenta modules, respectively, after a Bonferroni-adjusted p-value of 0.05. No pathways were enriched for the correlated genes of the black, green, red or turquoise modules.

Blue module (374 correlated genes)						
ID	Name	Source	P-value	Bonferroni-adjusted P-value	# of input genes / # of total genes in annotation	Genes from input
1269311	Interferon Signaling	BioSystems: REACTOME	8.21E-14	1.52E-10	26 / 202	*DDX58, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB2, HLA-DRA, B2M, HLA-DRB5, GBP1, IRF1, IRF4, IRF6, TRIM21, ADAR, TRIM22, STAT1, JAK2, UBA52, GBP4, KPNB1, ICAM1, OAS2, IFIT2, IFIT3, ISG15, MX1*
1269174	Translocation of ZAP-70 to Immunological synapse	BioSystems: REACTOME	1.46E-13	2.69E-10	11 / 22	*HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB2, HLA-DRA, HLA-DRB5, PTPN22, CD3D, CD3E, CD3G, CD247*
1269175	Generation of second messenger molecules	BioSystems: REACTOME	1.51E-13	2.78E-10	13 / 36	*HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB2, HLA-DRA, HLA-DRB5, GRAP2, LCP2, CD3D, CD3E, CD3G, CD247, FYB1*
1269310	Cytokine Signaling in Immune system	BioSystems: REACTOME	2.36E-13	4.35E-10	50 / 763	*CDKN1A, IL12RB1, IL12RB2, CRLF1, DDX58, TNFSF14, HLA-DPA1, HLA-DPB1, CXCL10, HLA-DQA1, HLA-DQB2, HLA-DRA, INPP5D, B2M, HLA-DRB5, IRAK2, GBP1, IRF1, IRF4, IRF6, TRIM21, ADAR, ITGB2, TRIM22, STAT1, STAT5A, SOCS2, CCR2, JAK2, UBA52, GBP4, CCR5, APBB1IP, KPNB1, ANXA1, DUSP2, ICAM1, OAS2, FASLG, IFIT2, IFIT3, CSF1, CSF1R, CSF3, ISG15, CD40, LTB, TNF, TNFRSF1B, MX1*
1269314	Interferon gamma signaling	BioSystems: REACTOME	6.38E-13	1.18E-09	18 / 94	*HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB2, HLA-DRA, B2M, HLA-DRB5, GBP1, IRF1, IRF4, IRF6, TRIM21, TRIM22, STAT1, JAK2, GBP4, ICAM1, OAS2*
1269173	Phosphorylation of CD3 and TCR zeta chains	BioSystems: REACTOME	8.72E-13	1.61E-09	11 / 25	*HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB2, HLA-DRA, HLA-DRB5, PTPN22, CD3D, CD3E, CD3G, CD247*
1269182	PD-1 signaling	BioSystems: REACTOME	1.48E-12	2.74E-09	11 / 26	*HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB2, HLA-DRA, HLA-DRB5, CD3D, CD3E, CD3G, CD247, PDCD1LG2*

Table S3.1. Significant pathways for the genes that were correlated with the blue and magenta maternal asthma-associated co-methylation modules (continued).

137922	IL12-mediated signaling events	BioSystems: Pathway Interaction Database	1.86E-11	3.43E-08	14 / 61	*IL12RB1, IL12RB2, HLA-DRA, B2M, STAT1, STAT5A, JAK2, CCR5, FASLG, CD3D, CD3E, CD3G, CD247, CD8B*
M54	IL12-mediated signaling events	MSigDB C2 BIOCARTA (v7.1)	2.98E-11	5.50E-08	14 / 63	*IL12RB1, IL12RB2, HLA-DRA, B2M, STAT1, STAT5A, JAK2, CCR5, FASLG, CD3D, CD3E, CD3G, CD247, CD8B*
M1462	CTL mediated immune response against target cells	MSigDB C2 BIOCARTA (v7.1)	3.68E-11	6.78E-08	8 / 13	*B2M, ITGB2, ICAM1, FASLG, CD3D, CD3E, CD3G, CD247*
1458885	Th1 and Th2 cell differentiation	BioSystems: KEGG	5.87E-11	1.08E-07	16 / 92	*IL12RB1, IL12RB2, HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, STAT1, STAT5A, JAK2, NOTCH2, CD3D, CD3E, CD3G, CD247*
M4319	IL12 and Stat4 Dependent Signaling Pathway in Th1 Development	MSigDB C2 BIOCARTA (v7.1)	8.66E-11	1.60E-07	9 / 20	*IL12RB1, IL12RB2, ETV5, JAK2, CCR5, CD3D, CD3E, CD3G, CD247*
83078	Hematopoietic cell lineage	BioSystems: KEGG	1.35E-10	2.48E-07	16 / 97	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, CD3D, CD3E, CD3G, CD5, CSF1, CSF1R, CD8B, CSF3, CD9, TNF*
M6231	NO2-dependent IL 12 Pathway in NK cells	MSigDB C2 BIOCARTA (v7.1)	3.48E-10	6.42E-07	8 / 16	*IL12RB1, IL12RB2, JAK2, CCR5, CD3D, CD3E, CD3G, CD247*
1469482	Th17 cell differentiation	BioSystems: KEGG	6.11E-10	1.13E-06	16 / 107	*IL12RB1, HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, IRF4, STAT1, STAT5A, JAK2, CD3D, CD3E, CD3G, CD247, IL21R*
213306	Measles	BioSystems: KEGG	2.40E-09	4.42E-06	17 / 134	*DDX58, TNFSF10, ADAR, CD209, STAT1, STAT5A, JAK2, HSPA1A, HSPA1B, OAS2, FASLG, SLAMF1, CD3D, CD3E, CD3G, MX1, CDK2*
217173	Influenza A	BioSystems: KEGG	3.18E-09	5.86E-06	19 / 173	*DDX58, HLA-DMB, TNFSF10, HLA-DPA1, HLA-DPB1, CXCL10, HLA-DQA1, HLA-DRA, HLA-DRB5, ADAR, STAT1, JAK2, HSPA1A, HSPA1B, ICAM1, OAS2, FASLG, TNF, MX1*
1269172	TCR signaling	BioSystems: REACTOME	5.59E-09	1.03E-05	16 / 124	*HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB2, HLA-DRA, INPP5D, HLA-DRB5, PTPN22, UBA52, GRAP2, LCP2, CD3D, CD3E, CD3G, CD247, FYB1*
83074	Antigen processing and presentation	BioSystems: KEGG	5.62E-09	1.04E-05	13 / 77	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, B2M, HLA-DRB5, NFYA, HSPA1A, HSPA1B, CD8B, CD74, TNF*
M16519	HIV Induced T Cell Apoptosis	MSigDB C2 BIOCARTA (v7.1)	6.06E-09	1.12E-05	6 / 9	*CCR5, FASLG, CD3D, CD3E, CD3G, CD247*

Table S3.1. Significant pathways for the genes that were correlated with the blue and magenta maternal asthma-associated co-methylation modules (continued).

172846	Staphylococcus aureus infection	BioSystems: KEGG	1.65E-08	3.04E-05	11 / 56	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, ITGB2, C2, SELPLG, KRT10, ICAM1*
1269177	Costimulation by the CD28 family	BioSystems: REACTOME	3.01E-08	5.54E-05	12 / 73	*HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB2, HLA-DRA, HLA-DRB5, GRAP2, CD3D, CD3E, CD3G, CD247, PDCD1LG2*
M18615	Allograft rejection	MSigDB C2 BIOCARTA (v7.1)	4.71E-08	8.69E-05	9 / 37	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, FASLG, CD40, TNF*
83123	Allograft rejection	BioSystems: KEGG	6.06E-08	1.12E-04	9 / 38	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, FASLG, CD40, TNF*
M6427	T Helper Cell Surface Molecules	MSigDB C2 BIOCARTA (v7.1)	6.33E-08	1.17E-04	6 / 12	*ITGB2, ICAM1, CD3D, CD3E, CD3G, CD247*
M13247	T Cytotoxic Cell Surface Molecules	MSigDB C2 BIOCARTA (v7.1)	6.33E-08	1.17E-04	6 / 12	*ITGB2, ICAM1, CD3D, CD3E, CD3G, CD247*
842771	Inflammatory bowel disease (IBD)	BioSystems: KEGG	8.39E-08	1.55E-04	11 / 65	*IL12RB1, IL12RB2, HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, STAT1, IL21R, TNF*
M13950	Asthma	MSigDB C2 BIOCARTA (v7.1)	1.23E-07	2.27E-04	8 / 30	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, CD40, TNF*
83120	Asthma	BioSystems: KEGG	1.63E-07	3.00E-04	8 / 31	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, CD40, TNF*
M10765	Lck and Fyn tyrosine kinases in initiation of TCR Activation	MSigDB C2 BIOCARTA (v7.1)	1.99E-07	3.66E-04	6 / 14	*HLA-DRA, HLA-DRB5, CD3D, CD3E, CD3G, CD247*
M6856	Hematopoietic cell lineage	MSigDB C2 BIOCARTA (v7.1)	2.24E-07	4.13E-04	12 / 87	*HLA-DRA, HLA-DRB5, CD3D, CD3E, CD3G, CD5, CSF1, CSF1R, CD8B, CSF3, CD9, TNF*
M16004	Antigen processing and presentation	MSigDB C2 BIOCARTA (v7.1)	2.55E-07	4.70E-04	12 / 88	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, B2M, HLA-DRB5, NFYA, HSPA1A, HSPA1B, CD8B, CD74*
83069	Cell adhesion molecules (CAMs)	BioSystems: KEGG	3.32E-07	6.13E-04	15 / 145	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, ITGA9, ITGB2, ITGB7, SELPLG, ICAM1, CD8B, TIGIT, CD40, PDCD1LG2*
P00053	T cell activation	PantherDB	4.39E-07	8.09E-04	11 / 76	*HLA-DPA1, HLA-DQA1, HLA-DRA, B2M, GRAP2, LCP2, CD3D, CD3E, CD3G, CD247, CD74*
169642	Toxoplasmosis	BioSystems: KEGG	6.04E-07	1.11E-03	13 / 113	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, STAT1, JAK2, CCR5, HSPA1A, HSPA1B, CD40, TNF*

Table S3.1. Significant pathways for the genes that were correlated with the blue and magenta maternal asthma-associated co-methylation modules (continued).

M16476	Cell adhesion molecules (CAMs)	MSigDB C2 BIOCARTA (v7.1)	6.67E-07	1.23E-03	14 / 133	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, ITGA9, ITGB2, ITGB7, SELPLG, ICAM1, CD8B, CD40, PDCD1LG2*
138018	Downstream signaling in naive CD8+ T cells	BioSystems: Pathway Interaction Database	7.49E-07	1.38E-03	9 / 50	*B2M, PTPN7, FASLG, CD3D, CD3E, CD3G, CD247, CD8B, TNF*
1269201	Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell	BioSystems: REACTOME	8.77E-07	1.62E-03	14 / 136	*LILRB2, B2M, ITGB2, ITGB7, LILRB4, ICAM1, CD3D, CD3E, CD3G, CD247, CD8B, CRTAM, CD40, CD96*
213780	Tuberculosis	BioSystems: KEGG	9.87E-07	1.82E-03	16 / 179	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, IRAK2, ITGB2, CD209, STAT1, JAK2, CARD9, NFYA, CALML3, CD74, TNF*
137936	IL12 signaling mediated by STAT4	BioSystems: Pathway Interaction Database	1.26E-06	2.32E-03	7 / 28	*HLA-DRA, ETV5, IRF1, CD3D, CD3E, CD3G, CD247*
1269176	Downstream TCR signaling	BioSystems: REACTOME	1.44E-06	2.65E-03	12 / 103	*HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB2, HLA-DRA, INPP5D, HLA-DRB5, UBA52, CD3D, CD3E, CD3G, CD247*
1269312	Interferon alpha/beta signaling	BioSystems: REACTOME	1.46E-06	2.70E-03	10 / 69	*IRF1, IRF4, IRF6, ADAR, STAT1, OAS2, IFIT2, IFIT3, ISG15, MX1*
M13519	Graft-versus-host disease	MSigDB C2 BIOCARTA (v7.1)	1.65E-06	3.04E-03	8 / 41	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, FASLG, TNF*
83124	Graft-versus-host disease	BioSystems: KEGG	1.65E-06	3.04E-03	8 / 41	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, FASLG, TNF*
M3126	Leishmania infection	MSigDB C2 BIOCARTA (v7.1)	2.18E-06	4.03E-03	10 / 72	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, ITGB2, STAT1, JAK2, TNF*
M12617	Type I diabetes mellitus	MSigDB C2 BIOCARTA (v7.1)	2.41E-06	4.45E-03	8 / 43	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, FASLG, TNF*
83095	Type I diabetes mellitus	BioSystems: KEGG	2.41E-06	4.45E-03	8 / 43	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, FASLG, TNF*
200309	Rheumatoid arthritis	BioSystems: KEGG	2.47E-06	4.55E-03	11 / 90	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, ITGB2, ICAM1, CSF1, LTB, TNF*
144181	Leishmaniasis	BioSystems: KEGG	2.48E-06	4.58E-03	10 / 73	*HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DRA, HLA-DRB5, ITGB2, STAT1, JAK2, TNF*

Table S3.1. Significant pathways for the genes that were correlated with the blue and magenta maternal asthma-associated co-methylation modules (continued).

M9809	Cytokine-cytokine receptor interaction	MSigDB C2 BIOCARTA (v7.1)	2.74E-06	5.05E-03	19 / 265	*IL12RB1, IL12RB2, TNFSF14, TNFSF10, CXCL10, CCR2, BMP2, BMP7, CXCL9, CCR5, FASLG, CSF1, CSF1R, CSF3, IL21R, CD40, LTB, TNF, TNFRSF1B*
Magenta module (703 correlated genes)						
ID	Name	Source	P-value	Bonferroni-adjusted P-value	# of input genes / # of total genes in annotation	Genes from input
1339156	Selenocysteine synthesis	BioSystems: REACTOME	4.17E-18	8.99E-15	28 / 96	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, SARS1, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1268690	Eukaryotic Translation Elongation	BioSystems: REACTOME	7.66E-18	1.65E-14	28 / 98	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, RPL14, RPS27L, EEF1B2, RPL13A, RPL6, RPL11, RPL18*
1268681	Formation of a pool of free 40S subunits	BioSystems: REACTOME	9.62E-18	2.08E-14	29 / 107	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, EIF3D, EIF3K, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1268691	Peptide chain elongation	BioSystems: REACTOME	1.91E-17	4.11E-14	27 / 93	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1269120	Viral mRNA Translation	BioSystems: REACTOME	1.91E-17	4.11E-14	27 / 93	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
M189	Ribosome	MSigDB C2 BIOCARTA (v7.1)	4.69E-17	1.01E-13	26 / 88	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1268692	Eukaryotic Translation Termination	BioSystems: REACTOME	6.28E-17	1.36E-13	27 / 97	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1268689	SRP-dependent cotranslational protein targeting to membrane	BioSystems: REACTOME	1.05E-16	2.27E-13	29 / 116	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, SEC61A1, RPL14, RPS27L, SEC11A, RPL13A, RPL6, RPL11, RPL18,*

Table S3.1. Significant pathways for the genes that were correlated with the blue and magenta maternal asthma-associated co-methylation modules (continued).

1269718	Nonsense Mediated Decay (NMD) independent of the Exon Junction Complex (EJC)	BioSystems: REACTOME	1.47E-16	3.18E-13	27 / 100	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1268686	GTP hydrolysis and joining of the 60S ribosomal subunit	BioSystems: REACTOME	2.21E-16	4.77E-13	29 / 119	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, EIF3D, EIF3K, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1268688	L13a-mediated translational silencing of Ceruloplasmin expression	BioSystems: REACTOME	2.21E-16	4.77E-13	29 / 119	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, EIF3D, EIF3K, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1269716	Nonsense-Mediated Decay (NMD)	BioSystems: REACTOME	3.58E-16	7.73E-13	29 / 121	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, SMG9, SMG6, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1269717	Nonsense Mediated Decay (NMD) enhanced by the Exon Junction Complex (EJC)	BioSystems: REACTOME	3.58E-16	7.73E-13	29 / 121	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, SMG9, SMG6, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
83036	Ribosome	BioSystems: KEGG	8.67E-16	1.87E-12	32 / 154	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, MRPL11, MRPL9, MRPL16, RPL14, RPS27L, MRPS7, MRPS2, RPL13A, RPL6, RPL11, RPL18*
1268680	Cap-dependent Translation Initiation	BioSystems: REACTOME	1.43E-15	3.09E-12	29 / 127	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, EIF3D, EIF3K, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1268679	Eukaryotic Translation Initiation	BioSystems: REACTOME	1.43E-15	3.09E-12	29 / 127	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, EIF3D, EIF3K, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1339149	Selenoamino acid metabolism	BioSystems: REACTOME	4.87E-15	1.05E-11	28 / 123	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, SARS1, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*

Table S3.1. Significant pathways for the genes that were correlated with the blue and magenta maternal asthma-associated co-methylation modules (continued).

1268678	Translation	BioSystems: REACTOME	6.95E-15	1.50E-11	32 / 165	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, SEC61A1, EIF3D, EIF3K, RPL14, RPS27L, EEF1B2, SEC11A, RPL13A, RPL6, RPL11, RPL18*
1269115	Influenza Viral RNA Transcription and Replication	BioSystems: REACTOME	1.08E-13	2.33E-10	28 / 138	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, POLR2C, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1269109	Influenza Life Cycle	BioSystems: REACTOME	5.65E-13	1.22E-09	28 / 147	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, POLR2C, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1269108	Influenza Infection	BioSystems: REACTOME	3.59E-12	7.75E-09	28 / 158	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, POLR2C, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1383086	Major pathway of rRNA processing in the nucleolus and cytosol	BioSystems: REACTOME	1.97E-11	4.24E-08	29 / 181	*DDX47, RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, DHX37, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1427846	rRNA processing in the nucleus and cytosol	BioSystems: REACTOME	9.80E-11	2.12E-07	29 / 193	*DDX47, RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, DHX37, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1383085	rRNA processing	BioSystems: REACTOME	3.38E-10	7.29E-07	29 / 203	*DDX47, RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, DHX37, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1270158	Metabolism of amino acids and derivatives	BioSystems: REACTOME	2.40E-09	5.18E-06	39 / 367	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, RPL23, SARS1, PSMF1, ARG2, ACAD8, ASS1, ALDH4A1, PSMB8, PSMB9, MCCC2, PSPH, DIO1, QDPR, RPL14, RPS27L, RPL13A, RPL6, RPL11, RPL18*
1269056	Infectious disease	BioSystems: REACTOME	5.19E-09	1.12E-05	40 / 393	*RPL24, RPL27, RPL30, RPL32, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, APOBEC3G, RPS17, RPS21, RPS25, RPS27, RPL23, PSMF1, POLR2C, SH3GL3, FYN, MAP2K6, PSMB8, PSMB9, NMT1, NELFE, RPL14, LCK, RPS27L, CD4, CD28, RPL13A, RPL6, RPL11, RPL18*

Table S3.1. Significant pathways for the genes that were correlated with the blue and magenta maternal asthma-associated co-methylation modules (continued).

1268682	Formation of the ternary complex, and subsequently, the 43S complex	BioSystems: REACTOME	7.62E-09	1.64E-05	14 / 55	*RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, EIF3D, EIF3K, RPS27L*
1268685	Ribosomal scanning and start codon recognition	BioSystems: REACTOME	4.99E-08	1.08E-04	14 / 63	*RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, EIF3D, EIF3K, RPS27L*
1268684	Translation initiation complex formation	BioSystems: REACTOME	6.17E-08	1.33E-04	14 / 64	*RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, EIF3D, EIF3K, RPS27L*
1268683	Activation of the mRNA upon binding of the cap-binding complex and eIFs, and subsequent binding to 43S	BioSystems: REACTOME	7.61E-08	1.64E-04	14 / 65	*RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, RPS17, RPS21, RPS25, RPS27, EIF3D, EIF3K, RPS27L*
1268854	Disease	BioSystems: REACTOME	6.68E-07	1.44E-03	61 / 867	*RPL24, RPL27, RPL30, RPL32, SFTA3, RPL34, RPL35A, RPL37, RPL37A, RPLP2, RPS5, RPS9, RPS11, RPS12, RPS13, RPS14, RPS15, APOBEC3G, RPS17, RPS21, RPS25, RPS27, RPL23, ZFYVE9, FGFR1, PLCG1, NEURL1B, RASAL2, KLB, PSMF1, FMOD, POLR2C, SH3GL3, MUC13, RASAL3, CSPG4, SDC3, IKBKB, FYN, MAP2K6, HDAC9, PSMB8, PSMB9, MCCC2, GAB2, NFKB2, GPC2, NMT1, NELFE, KSR2, RPL14, LCK, RPS27L, CD4, AKAP9, CD28, TLR2, RPL13A, RPL6, RPL11, RPL18*
137998	TCR signaling in naive CD4+ T cells	BioSystems: Pathway Interaction Database	1.52E-06	3.28E-03	12 / 60	*PLCG1, MAP3K8, ZAP70, CARD11, IKBKB, FYN, GAB2, LCK, RASGRP1, CD4, CD28, MAP4K1*
M34	TCR signaling in naïve CD4+ T cells	MSigDB C2 BIOCARTA (v7.1)	3.70E-06	7.99E-03	12 / 65	*PLCG1, MAP3K8, ZAP70, CARD11, IKBKB, FYN, GAB2, LCK, RASGRP1, CD4, CD28, MAP4K1*

Table S3.2. Twenty-five differentially expressed (DE) genes among those that were correlated with the eigenvector of only one maternal asthma-associated co-methylation module at |rho| > 0.20. DE analysis was performed between asthma cases without a mother with asthma (N = 51) and asthma cases with a mother with asthma (N = 30) for each of the genes uniquely correlated with one module. One and 25 DE genes were detected for the black and blue modules, respectively (FDR < 10%). None of the uniquely correlated genes of the green, magenta, red or turquoise modules were significantly DE. Representative studies are listed for genes that have previously been implicated in asthma or allergies.

Black module (1 DE gene)					
Gene	Description	Coefficient	P-value	FDR-adjusted p-value	Ref.
SAMSN1	SAM Domain, SH3 Domain and Nuclear Localization Signals 1	-0.88	4.59E-4	0.042	(*106*)
Blue module (25 DE genes)					
Gene	Description	Coefficient	P-value	FDR-adjusted p-value	Ref.
TMEM273	Transmembrane Protein 273	-0.88	6.35E-05	0.024	(*108, 127*)
CD209	CD209 Molecule	-0.98	1.72E-04	0.032	(*128*)
CD84	CD84 Molecule	-0.60	2.60E-04	0.032	(*129*)
SOCS2	Suppressor of Cytokine Signaling 2	-0.72	3.65E-04	0.034	(*108, 130*)
SNX20	Sorting Nexin 20	-0.75	7.45E-04	0.056	(*43*)
FAM183A	Family with Sequence Similarity 183 Member A	0.26	2.27E-03	0.079	(*108*)
WIPF1	WAS/WASL Interacting Protein Family Member 1	-0.42	2.36E-03	0.079	(*107, 131*)
DOCK10	Dedicator of Cytokinesis 10	-0.49	1.55E-03	0.079	(*132*)
LCP2	Lymphocyte Cytosolic Protein 2	-0.56	3.36E-03	0.079	(*133*)
IRF4	Interferon Regulatory Factor 4	-0.80	2.94E-03	0.079	(*43*)
CGAS	Cyclic GMP-AMP Synthase	-0.44	1.98E-03	0.079	(*134*)
MYO1G	Myosin IG	-0.64	1.64E-03	0.079	
IKZF1	IKAROS Family Zinc Finger 1	-0.48	2.80E-03	0.079	(*107, 135*)
WDFY4	WDFY Family Member 4	-0.66	2.70E-03	0.079	
ZNF511	Zinc Finger Protein 511	0.17	3.22E-03	0.079	(*108, 136*)
SCGB2A1	Secretoglobin Family 2A Member 1	0.41	2.99E-03	0.079	(*108*)
CSF1	Colony Stimulating Factor 1	-0.47	4.24E-03	0.088	(*137*)
GNG2	G Protein Subunit Gamma 2	-0.64	4.11E-03	0.088	(*107*)
POP5	POP5 Homolog, Ribonuclease P-MRP Subunit	0.13	4.59E-03	0.090	(*108*)
PTPN22	Protein Tyrosine Phosphatase Non-Receptor Type 22	-0.53	5.97E-03	0.093	(*108, 138*)
PTPN7	Protein Tyrosine Phosphatase Non-Receptor Type 7	-0.51	5.96E-03	0.093	(*139*)
DOCK2	Dedicator of Cytokinesis 2	-0.56	5.19E-03	0.093	(*107*)
DCTN3	Dynactin Subunit 3	0.14	5.33E-03	0.093	(*108*)
LILRB4	Leukocyte Immunoglobulin Like Receptor B4	-0.81	5.81E-03	0.093	(*140*)
APBB1IP	Amyloid Beta Precursor Protein Binding Family B Member 1 Interacting Protein	-0.51	6.58E-03	0.098	(*141*)

Table S3.3. Nine differentially expressed (DE) genes of the enriched pathways of the blue module. Of 374 genes that were correlated only with the blue module eigenvector (|rho| > 0.20), 25 genes were DE between asthma cases with (MA; N = 30) and without an asthmatic mother (No MA; N = 51; FDR < 10%). Of these, nine genes were DE **(**and present in one to five of the enriched pathways of the blue module).

Gene	Name	Number of pathways	Blue module pathways
IRF4	Interferon Regulatory Factor 4	5	"Interferon signaling", "Cytokine signaling in immune system", "Interferon gamma signaling", "Th17 cell differentiation", "Interferon alpha/beta signaling"
CSF1	Colony Stimulating Factor 1	4	"Cytokine signaling in immune system", "Hematopoietic cell lineage", "Rheumatoid arthritis", "Cytokine-cytokine receptor interaction"
LCP2	Lymphocyte Cytosolic Protein 2	3	"Generation of second messenger molecules", "TCR signaling", "T cell activation"
PTPN22	Protein Tyrosine Phosphatase Non-Receptor Type 22	3	"Translocation of ZAP-70 to immunological synpase", "Phosphorylation of CD3 and TCR zeta chains", "TCR signaling"
CD209	CD209 Molecule	2	"Measles", "Tuberculosis"
SOCS2	Suppressor of Cytokine Signaling 2	1	"Cytokine signaling in immune system"
LILRB4	Leukocyte Immunoglobulin Like Receptor B4	1	"Immunoregulatory interactions between a lymphoid and a non-lymphoid cell"
PTPN7	Protein Tyrosine Phosphatase Non-Receptor Type 7	1	"Downstream signaling in naïve CD8+ T cells"
APBB1IP	Amyloid Beta Precursor Protein Binding Family B Member 1 Interacting Protein	1	"Cytokine signaling in immune system"

3.6.2 Supplementary figures

Figure S3.1. Scatterplot of beta coefficients for the 9,756 differentially methylated CpGs (DMCs). Three comparisons of differential methylation were performed: non-asthma controls (N=42) to all asthma cases (N = 88), to cases with no maternal asthma (NMA) (N = 56), and to cases with maternal asthma (MA) (N = 32). Overall, 9,756 DMCs were significant (FDR < 10%) in the analysis including all cases but not in the other two comparisons (FDR > 10%). The beta coefficients of the 9,756 DMCs are shown for NMA cases (x-axis) and MA cases (y-axis). The 6,522 (66.85%) DMCs with larger effects in MA cases compared to NMA cases are shown as blue points, and the 3,234 DMCs (33.15%) with larger effects in NMA cases are shown as orange points. Paired signed-rank test; $P<2.2x10^{-16}$.

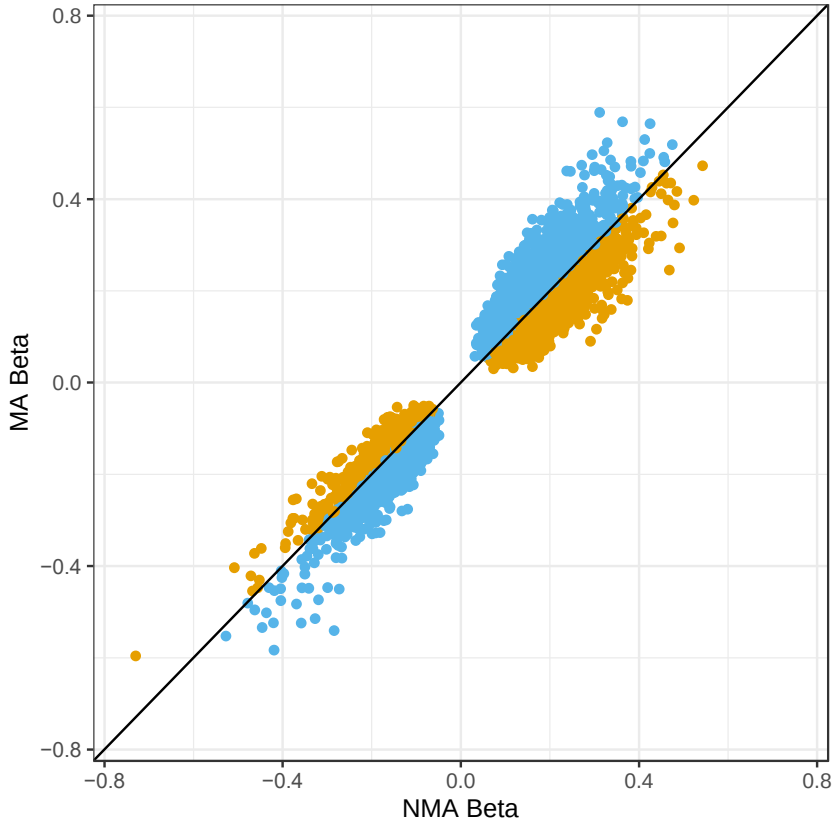


Figure S3.2. Cluster dendrogram of 1,428 maternal asthma-associated differentially methylated CpGs (MA-DMCs). Of the 1,428 MA-DMCs, 935 (65.48%) formed nine co-methylation modules. After merging closely related modules using the MergeCloseModules function, as implemented in WGCNA (*99*) (default settings), there were six co-methylation modules (32 to 344 CpGs in each module). There were 493 (34.52%) uncorrelated CpGs assigned to a gray module.

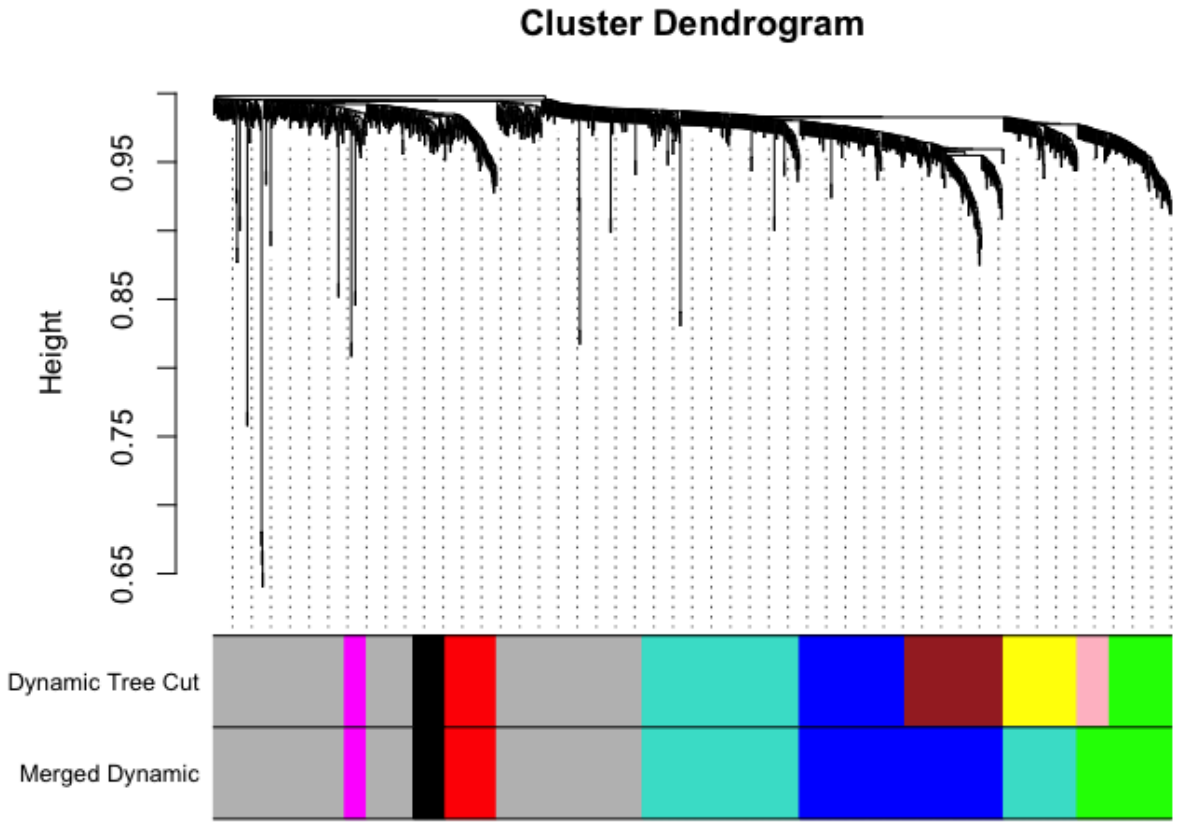


Figure S3.3. Twenty-five differentially expressed (DE) genes between cases with and without an asthmatic mother. Twenty-five DE genes between asthma cases with maternal asthma (MA; N = 30) and with no maternal asthma (NMA; N = 51; FDR < 10%) among 374 of the correlated genes of the blue module eigenvector (|rho| > 0.20), 25 were DE. Among the 25 DE genes, (**A**) 20 were downregulated and (**B**) 5 were upregulated. Normalized and covariate-adjusted gene expression is plotted on the y-axis. P-values were determined using a linear model. The FDR-corrected p-values of the 25 DE genes are reported. Boxplots include the upper (75^{th}) and lower (25^{th}) quartiles and the median (dark line). Data falling outside the interquartile range are plotted as outliers of the data. Figure starts on the next page.

Figure S3.3. Twenty-five differentially expressed (DE) genes between cases with and without an asthmatic mother (continued).

A

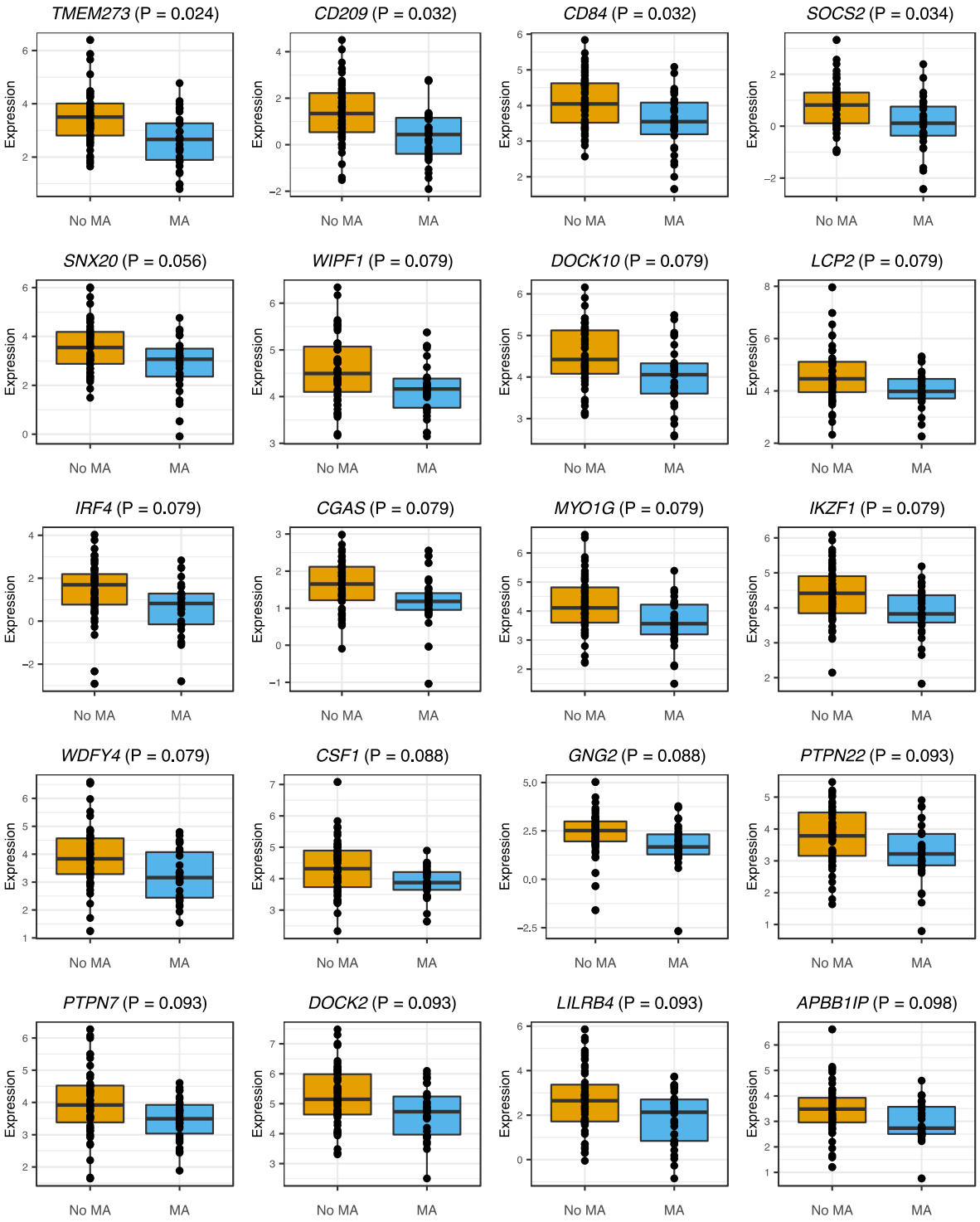


Figure S3.3. Twenty-five differentially expressed (DE) genes between cases with and without an asthmatic mother (continued).

B

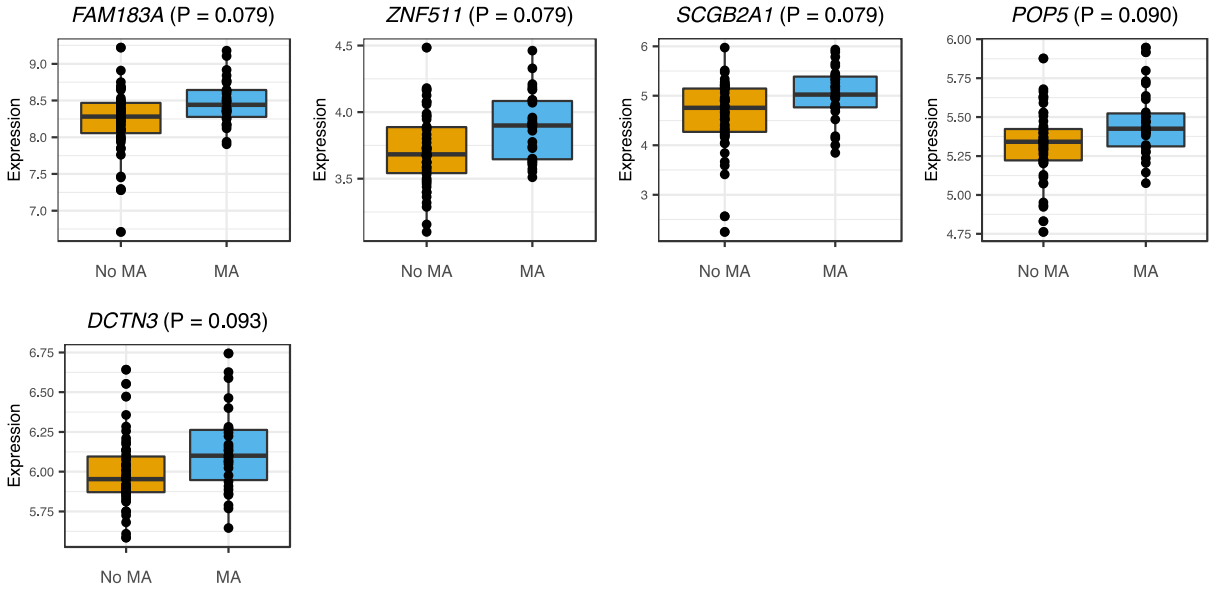
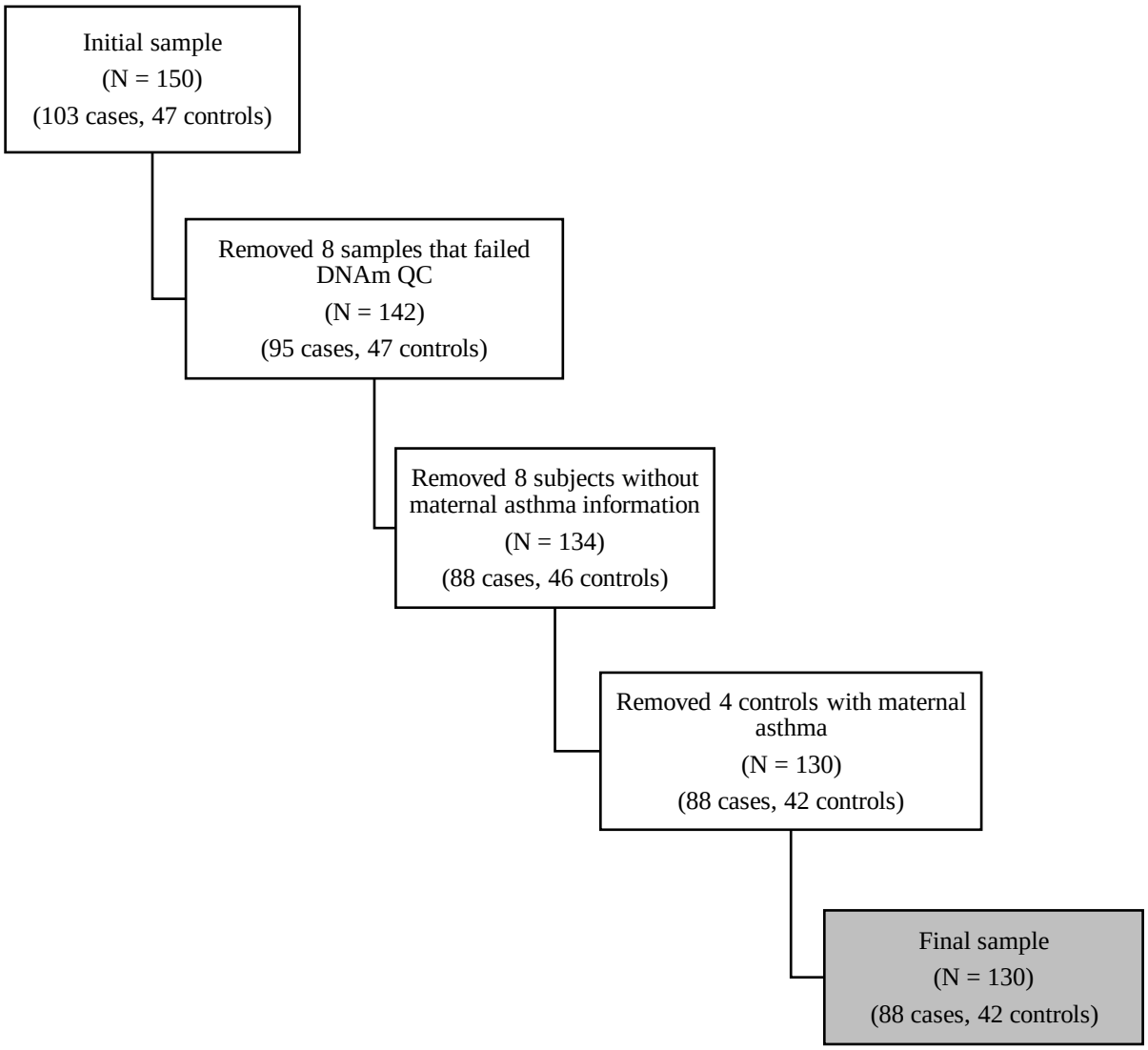


Figure S3.4. Flow chart of sample selection. All analyses included the 130 subjects shown in the shaded box, as described in **Table 3.1**.



CHAPTER 4

Associations between mitochondrial variants and haplogroups with asthma in ethnically diverse mother-child dyads

4.1 Abstract

One of the most consistent risk factors for childhood-onset asthma is having a mother with asthma. A possible explanation for this risk is the transmission of asthma-associated variation encoded in the mitochondrial (mt) genome, which is maternally-inherited. I hypothesized, therefore, that associations between asthma and mt single nucleotide polymorphisms (mtSNPs) would be more pronounced among asthmatic children of asthmatic mothers compared to asthmatic children of non-asthmatic mothers, or to children of asthmatic fathers. My study included 3,155 mother-child and 2,960 father-child dyads from nine ethnically diverse cohorts. Overall, 20 mtSNPs were associated with asthma in maternal asthma dyads in three of the cohorts (African American, Brazilian, Mexican American), but only one mtSNP was associated with asthma in paternal asthma dyads in a different African American cohort. All significant associations were cohort-specific. One missense mutation, A93G (P=0.033; OR=1.37; MAF=0.03), in the D-loop and one L3e3'4 haplogroup-defining missense mutation, G5262A (P=0.016; OR=1.62; MAF=0.05), in the NADH dehydrogenase subunit 2 (*MT-ND2*) gene were associated with asthma risk in African American mother-child dyads. In all three cohorts with associations, 11 mtSNPs were in genes encoding NADH dehydrogenase subunits of the respiratory complex I. Taken together, my study suggests that low frequency ethnic-specific mtSNPs are associated with asthma only among asthmatic children of asthmatic mothers.

4.2 Introduction

Asthma is a common chronic inflammatory disease of the airways, affecting over 300 million people worldwide (*1*). Genome-wide association studies (GWAS) have identified many asthma-associated SNPs, but these variants have modest effects and account for little of the heritability (*43-47*). This "missing heritability" suggests that there are remaining genetic contributions to asthma risk to be discovered. Among the many epidemiologic risk factors for asthma, asthma in the mother is one of the most reproducible risks for asthma in her child (*49*). One possible explanation for this maternal risk may be the presence of asthma-associated variation in the maternally-inherited mitochondrial genome.

The mitochondrial respiratory chain is a major source of reactive oxygen species (ROS), which are important signaling molecules for metabolic homeostasis. Overproduction of ROS has been associated with the development of asthma (*142*, *143*). When airway epithelial cells are exposed to excess ROS, their mitochondria can become damaged and apoptosis can result (*144*). Similarly, mitochondrial dysfunction can produce a large amount of ROS, which attacks cellular contents and boosts inflammation (*145*). Mitochondrial dysfunction is also consistent with the higher abundance and structural changes of mitochondria in the airway epithelium of mice (*146*). These studies support a role for aberrant mitochondrial function contributing to asthma pathogenesis.

The lack of recombination and the intrinsically high mutation rates of mtDNA have facilitated the construction of a human mtDNA phylogenetic tree (*147*). Populations who live in close geographical proximity and have a recent shared ancestry have similar mitochondrial haplotypes, which have been grouped by sequence similarity into mtDNA haplogroups. One European mtDNA haplogroup, called U, was associated with asthma in one study (*148*) and

elevated IgE in a second study (*149*); both in subjects of European ancestry. Another study from Germany reported associations between mitochondrial variants and asthma that were enriched in the D-loop (non-coding) and in the NADH-dehydrogenase subunit (*150*). Although these findings suggested a possible role for mtSNPs in asthma risk, they all focused primarily on individuals of European ancestry. Importantly, these studies did not consider maternal asthma.

Here, I performed a mitochondrial-wide association study (MWAS) of mitochondrial single nucleotide polymorphisms (mtSNPs) and haplogroups and asthma in asthmatic children with and without an asthmatic mother (maternal asthma dyads). As a control for my hypothesis, I performed a MWAS in asthmatic children with and without an asthmatic father (paternal asthma dyads). I report a greater number of associations of mtSNPs and haplogroups with asthma only in the maternal asthma dyads compared to the paternal asthma dyads. None of the maternal asthma-associated mtSNPs and haplogroups were associated in the paternal asthma dyads. Overall, I demonstrated that some of the risk for asthma in children of asthmatic mothers may be due to variation inherited in the maternally-inherited mitochondrial genome.

4.3 Results

4.3.1 mtDNA SNP associations

Previous MWAS tested associations between mtSNPs and asthma in a standard case-control framework. However, the transmission of the mitochondrial genome from mother to offspring suggests that mtSNPs may only be associated with, or more strongly associated with, asthma in maternal asthma dyads compared to paternal asthma dyads. To test this hypothesis, I first performed a MWAS between mtSNPs and asthma in 3,155 maternal asthma dyads from nine cohorts of African American, Brazilian, European American, Honduran, Mexican American, and Puerto Rican ancestry (**Table 4.1**).

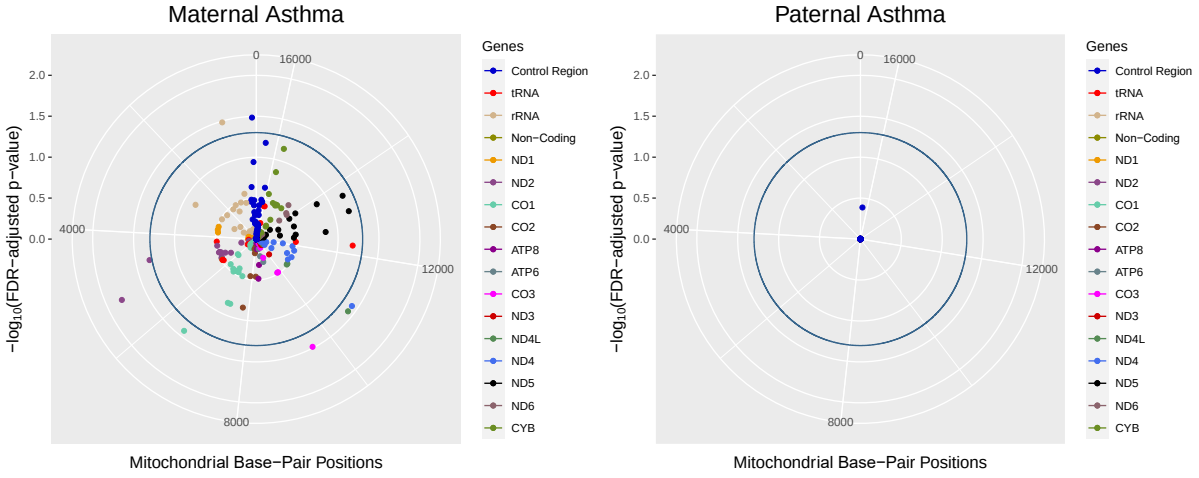
Table 4.1. Sample composition of nine ethnically diverse cohorts. MA, Maternal Asthma. PA, Paternal Asthma. APIC, Asthma Phenotypes in the Inner City. CAG, Chicago Asthma Genetics. SAPPHIRE, Study of Asthma Phenotypes and Pharmacogenomic Interactions by Race-Ethnicity. URECA, Urban Environment and Childhood Asthma. BIAS, Brazilian Immunogenetics of Asthma and Schistosomiasis. CAMP, Childhood Asthma Management Program. HONDAS, Honduran subjects who self-reported as Garifuna. GALA, Genes-environments and Admixture in Latino Americans.

Ethnicity	**Study**	**Male / Female**	**Cases with MA / without MA / Unknown**	**Cases with PA / without PA / Unknown**	**Total**
African American	APIC (*151*)	186/132	114/204/0	83/213/22	318
	CAG (*135*)	66/91	48/108/1	32/119/6	157
	SAPPHIRE (*135*)	193/353	123/410/13	88/414/44	546
	URECA (*152*)	74/49	69/54/0	35/78/10	123
Brazilian	BIAS (*47*)	57/56	40/66/7	20/32/61	113
European American	CAMP (*135*)	233/139	94/276/2	82/283/7	372
Honduran	HONDAS (*47*)	62/110	53/118/1	44/126/2	172
Mexican American	GALA II (*135*)	310/250	98/458/4	39/485/36	560
Puerto Rican	GALA II (*135*)	461/368	378/444/7	261/526/42	829
	Total	**1642/1548**	**1017/2138/35**	**684/2276/230**	**3190**

I tested low-frequency and common variants (MAF > 0.01) within each cohort (**Table S4.1**). Three of ten cohorts had significant associations between mtSNPs and asthma in the maternal asthma dyads (FDR-adjusted P-value < 0.05). Eight, six and six significant associations were detected in an African American (APIC) (**Figure 4.1A-C**), Brazilian (BIAS) (**Figure 4.1D-F**), and Mexican American (GALA II) (**Figure 4.1G-I**) cohort, respectively. Within each cohort, many of the variants had similar or identical summary statistics, likely reflecting the LD between mtSNPs on the same or related haplogroups. As examples, T6524C, T10667C and A10816G (FDR-adjusted P = 0.037; OR = 1.55 [1.19, 2.02]) in the African American cohort (APIC), and T4977C and C15535T (FDR-adjusted P = 0.016; OR = 0.68 [0.55, 0.85]) and G6962A and T12477C (FDR-adjusted P = 0.042; OR = 0.74 [0.61, 0.90]) in the Brazilian cohort (BIAS) likely reflect cohort-specific haplogroups associated with asthma in mother-child dyads.

Figure 4.1. Summary of MWAS results of parent-child dyads. Results are presented for three cohorts: (**A-C**) African Americans (APIC), (**D-F**) Brazilians (BIAS), and (**G-I**) Mexican Americans (GALA II). Solar plots of cohort-specific FDR-adjusted P-values for maternal and paternal asthma analyses, summary statistics of the significant mtSNPs, and summary statistics for the significant mtSNPs in the other cohorts are shown for APIC (**A-C**), BIAS (**D-F**), and GALA II (**G-I**). Only variants that passed QC and had MAF > 0.01 are shown for each cohort. Each point in the solar plot represents the mtSNP association, color coded by mitochondrial gene and plotted by its relative base pair position in the circular mtDNA genome based on the revised Cambridge Reference Sequence (rCRS). The outer blue circle shows the cohort-specific FDR-adjusted P-value of 0.05. Solar plots for the other six cohorts are shown in **Figure S4.1**.

A. Solar Plots in African Americans (APIC)



B. Significant mtSNPs in African Americans (APIC)

				Maternal Asthma Results		Paternal Asthma Results	
mtSNP	**Gene**	**Major / Minor Allele**	**MAF**	**Adjusted P-value**	**OR (95% CI)**	**Adjusted P-value**	**OR (95% CI)**
G5262A	*MT-ND2*	G/A	0.03	0.016	1.62 [1.28, 2.06]	1.00	0.90 [0.71, 1.14]
A750G	*MT-RNR1*	A/G	0.05	0.033	1.53 [1.20, 1.94]	1.00	0.97 [0.77, 1.23]
A93G	D-loop	A/G	0.05	0.033	1.37 [1.15, 1.64]	1.00	1.00 [0.84, 1.19]
G9554A	*MT-CO3*	G/A	0.03	0.033	1.52 [1.20, 1.94]	1.00	0.87 [0.68, 1.12]
T6524C	*MT-CO1*	T/C	0.02	0.037	1.55 [1.19, 2.02]	1.00	0.88 [0.68, 1.14]
T10667C	*MT-ND4L*	T/C	0.02	0.037	1.55 [1.19, 2.02]	1.00	0.88 [0.68, 1.14]
A10816G	*MT-ND4*	A/G	0.02	0.037	1.55 [1.19, 2.02]	1.00	0.88 [0.68, 1.14]
G4655A	*MT-ND2*	G/A	0.02	0.047	1.53 [1.18, 1.99]	1.00	0.90 [0.69, 1.16]

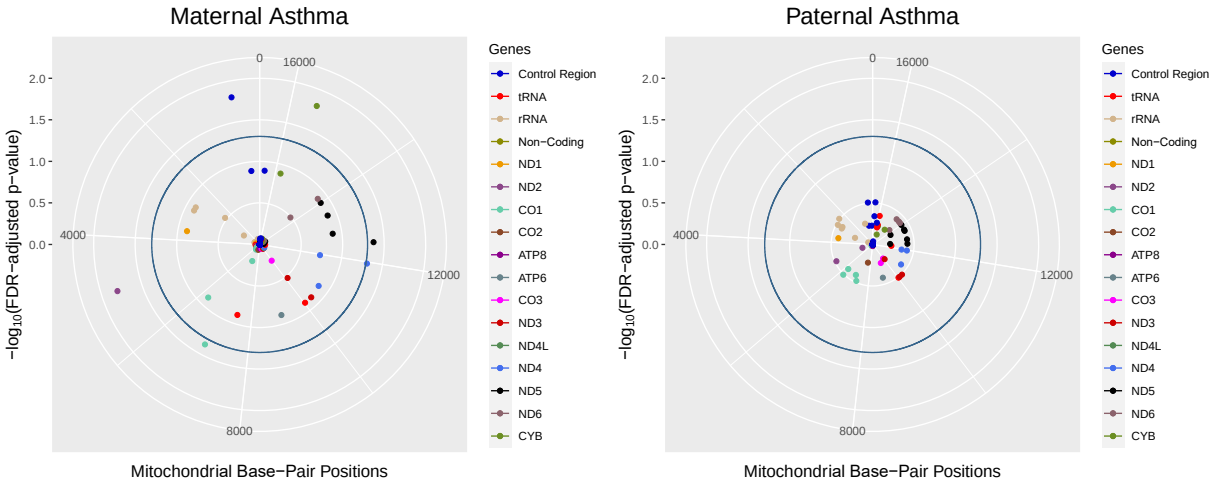
C. APIC mtSNPs in Other Cohorts

			Maternal Asthma Results		Paternal Asthma Results	
Ethnicity	**Study**	**MAF**	**Adjusted P-value**	**OR (95% CI)**	**Adjusted P-value**	**OR (95% CI)**
G5262A (*MT-ND2*)						
African Americans	URECA	0.02	0.91	0.66 [0.40, 1.13]	0.93	1.07 [0.66, 1.73]
A750G (*MT-RNR1*)						
African Americans	URECA	0.03	0.91	0.60 [0.40, 0.92]	0.93	0.95 [0.64, 1.42]
A93G (D-loop)						
African Americans	URECA	0.07	0.91	1.24 [0.96, 1.60]	0.93	0.88 [0.69, 1.13]

Figure 4.1. Summary of MWAS results of parent-child dyads (continued).

Brazilian	BIAS	0.03	0.84	1.18 [0.84, 1.65]	0.59	0.83 [0.49, 1.41]
Honduran	HONDAS	0.1	1.00	1.00 [0.94, 1.06]	0.97	0.97 [0.87, 1.09]
G9554A (*MT-CO3*)						
African Americans	URECA	0.02	0.96	0.90 [0.54, 1.51]	0.93	0.89 [0.51, 1.56]
T6524C (*MT-CO1*)						
African Americans	URECA	0.01	0.94	0.74 [0.40, 1.34]	0.93	0.89 [0.51, 1.56]
T10667C (*MT-ND4L*)						
African Americans	URECA	0.01	0.94	0.74 [0.40, 1.34]	0.93	0.89 [0.51, 1.56]
A10816G (*MT-ND4*)						
African Americans	URECA	0.01	0.94	0.74 [0.40, 1.34]	0.93	0.89 [0.51, 1.56]
G4655A (*MT-ND2*)						
African Americans	URECA	0.01	0.94	0.74 [0.40, 1.34]	0.93	0.89 [0.51, 1.56]

D. Solar Plots in Brazilians (BIAS)



E. Significant mtSNPs in Brazilians (BIAS)

				Maternal Asthma Results		Paternal Asthma Results	
mtSNP	**Gene**	**Major / Minor Allele**	**MAF**	**Adjusted P-value**	**OR (95% CI)**	**Adjusted P-value**	**OR (95% CI)**
T4977C	*MT-ND2*	T/C	0.07	0.016	0.68 [0.55, 0.85]	0.74	0.92 [0.64, 1.34]
C15535T	*MT-CYB*	C/T	0.07	0.016	0.68 [0.55, 0.85]	0.74	0.92 [0.64, 1.34]
G499A	D-loop	G/A	0.08	0.016	0.71 [0.58, 0.86]	0.59	0.90 [0.68, 1.18]
G6962A	*MT-CO3*	G/A	0.08	0.042	0.74 [0.61, 0.90]	0.38	1.30 [0.93, 1.80]
T12477C	*MT-ND5*	T/C	0.08	0.042	0.74 [0.61, 0.90]	0.38	1.30 [0.93, 1.80]
A11959G	*MT-ND4*	A/G	0.07	0.049	0.73 [0.59, 0.90]	0.38	1.30 [0.93, 1.80]

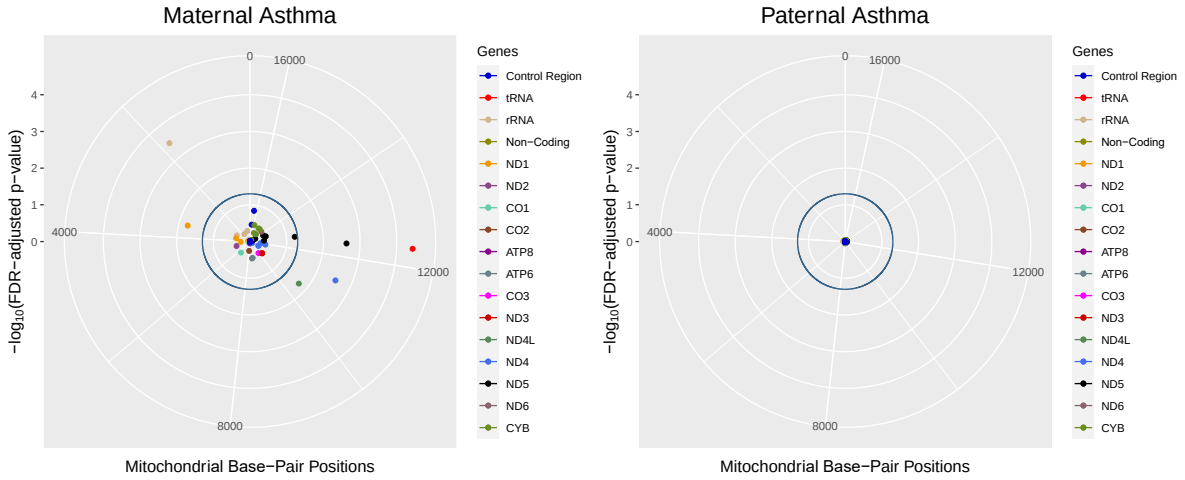
F. BIAS mtSNPs in Other Cohorts

			Maternal Asthma Results		Paternal Asthma Results	
Ethnicity	**Study**	**MAF**	**Adjusted P-value**	**OR (95% CI)**	**Adjusted P-value**	**OR (95% CI)**
T4977C (*MT-ND2*)						
Honduran	HONDAS	0.01	1.00	1.27 [0.83, 1.94]	0.97	0.92 [0.60, 1.41]
Mexican American	GALA II (MEX)	0.22	0.41	1.03 [0.99, 1.07]	0.99	1.00 [0.97, 1.03]
Puerto Rican	GALA II (PR)	0.05	0.94	1.01 [0.93, 1.09]	0.58	0.93 [0.86, 1.00]

Figure 4.1. Summary of MWAS results of parent-child dyads (continued).

C15535T (*MT-CYB*)						
Honduran	HONDAS	0.01	1.00	1.27 [0.83, 1.94]	0.97	0.92 [0.60, 1.41]
G499A (D-loop)						
African American	SAPPHIRE	0.02	0.97	0.96 [0.84, 1.11]	0.019	1.27 [1.11, 1.45]
	URECA	0.02	0.96	0.86 [0.54, 1.36]	0.64	0.70 [0.46, 1.07]
European American	CAMP	0.03	0.97	0.99 [0.87, 1.13]	0.83	1.04 [0.92, 1.18]
Honduran	HONDAS	0.02	1.00	1.05 [0.79, 1.39]	0.87	0.88 [0.65, 1.19]

G. Solar Plots in Mexican Americans (GALA II)



H. Significant mtSNPs in Mexican Americans (GALA II)

				Maternal Asthma Results		Paternal Asthma Results	
mtSNP	**Gene**	**Major / Minor Allele**	**MAF**	**Adjusted P-value**	**OR (95% CI)**	**Adjusted P-value**	**OR (95% CI)**
A12308G	*MT-TL2*	A/G	0.02	3.7×10^{-5}	1.30 [1.17, 1.44]	0.90	0.96 [0.89, 1.03]
A1811G	*MT-RNR2*	A/G	0.02	3.4×10^{-4}	1.30 [1.16, 1.47]	0.90	0.95 [0.88, 1.03]
G12372A	*MT-ND5*	G/A	0.04	0.0028	1.17 [1.08, 1.26]	0.99	1.01 [0.95, 1.06]
T11299C	*MT-ND4*	T/C	0.01	0.0028	1.31 [1.13, 1.51]	0.90	0.96 [0.87, 1.06]
A3480G	*MT-ND1*	A/G	0.01	0.018	1.28 [1.09, 1.49]	0.90	0.95 [0.85, 1.05]
A10550G	*MT-ND4L*	A/G	0.01	0.018	1.28 [1.09, 1.49]	0.90	0.95 [0.85, 1.05]

I. GALA II mtSNPs in Other Cohorts

			Maternal Asthma Results		Paternal Asthma Results	
Ethnicity	**Study**	**MAF**	**Adjusted P-value**	**OR (95% CI)**	**Adjusted P-value**	**OR (95% CI)**
A12308G (*MT-TL2*)						
African American	SAPPHIRE	0.02	0.97	1.08 [0.94, 1.23]	0.80	0.96 [0.84, 1.09]
	URECA	0.02	0.91	0.71 [0.43, 1.18]	0.93	1.17 [0.73, 1.88]
European American	CAMP	0.19	0.73	1.02 [0.96, 1.08]	0.82	0.98 [0.93, 1.04]
Puerto Rican	GALA II (PR)	0.02	0.99	1.00 [0.90, 1.12]	0.93	0.96 [0.87, 1.07]
A1811G (*MT-RNR2*)						
European American	CAMP	0.10	0.42	1.09 [1.01, 1.17]	0.81	0.95 [0.88, 1.02]
G12372A (*MT-ND5*)						
Puerto Rican	GALA II (PR)	0.05	0.94	1.03 [0.95, 1.12]	0.93	0.98 [0.90, 1.06]

Nearly all of the significant associations were with low frequency alleles, and all were cohort- and ethnic-specific. This specificity may be partially explained by the fact that not all mtSNPs were typed in every cohort and that most of the associated mtSNPs were rare in the other cohorts (**Table S4.2**). The latter is consistent with my previous findings regarding low frequency exome variants and asthma in some of these same cohorts (*135*), in which associations with low frequency variants were population-specific. However, there was heterogeneity within the African American cohorts themselves, in whom I observed associations in only one of the four cohorts. None of the associated mtSNPs in APIC were available in two other African American cohorts genotyped by arrays (SAPPHIRE and CAG), but one other cohort (URECA) also had whole genome sequencing available and did not replicate any of the eight associations. In fact, whereas the point estimates of the odds ratios (ORs) in APIC were >1.0, all but one of the variants had ORs <1.0, despite both these cohorts being comprised of childhood-onset asthma cases (*151, 152*) (**Figure 4.1C**). Whether this is due to the smaller number of URECA maternal asthma dyads (153 in URECA vs. 318 in APIC) impacting statistical power or to other underlying differences between these children cannot be assessed in these data.

In contrast, the A93G mtSNP in the D-loop that was associated with asthma in maternal asthma dyads in APIC (OR = 1.37) had ORs, although not significant, in the same direction (≥1.0) in URECA, BIAS and HONDAS. Similarly, another variant in the D-loop, G499A, was associated with asthma in the BIAS (Brazilian) maternal asthma dyads (OR = 0.71), and the ORs, although also not significant, were <1.0 (lower frequency in maternal asthma dyads) in three of four additional cohorts: African Americans in both SAPPHIRE and URECA and Europeans in CAMP (**Figure 4.1D-F**). An OR <1.0 suggests that reduced maternal transmission of the variant to children of mothers with asthma is associated with asthma. In Mexican Americans from GALA II,

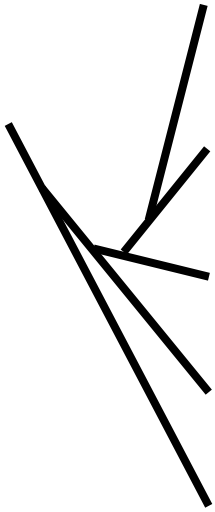
the most significant variant, A12308G, in the tRNA gene *MT-TL2*, was associated with asthma in maternal asthma dyads (OR = 1.3); although not significant, three of four cohorts also had ORs >1.0: African Americans in SAPPHIRE, European Americans in CAMP and Puerto Ricans in GALA II (**Figure 4.1G-I**). Overall, all of the significant maternal asthma-associated variants were cohort-specific. However, the two associated variants in the D-loop (one each in APIC and BIAS) showed similar directions of effects in maternal asthma dyads from multiple cohorts, suggesting a potentially important role for variants in the D-loop contributing to asthma risk across populations.

To further test my hypothesis that the effects of mtSNPs on asthma risk will be confined to or more pronounced in children of asthmatic mothers, I next repeated the same MWAS comparing asthma cases with or without an asthmatic father. Among the nine cohorts, only one mtSNP, G499A, in one African American cohort (SAPPHIRE), was associated with asthma in the paternal asthma dyads (FDR-adjusted P = 0.019; OR = 1.27 [1.11, 1.45]) (**Figure S4.1B; Figure 4.1F**). Neither G499A nor any other mtSNPs were associated with asthma in the paternal asthma dyads in the other eight cohorts. Curiously, this same mtSNP was associated with asthma in the BIAS (Brazilian) maternal asthma dyads, but the effects were in opposite directions. This result suggests complex mechanisms of associations of the same mtSNP with asthma in children with either a mother or father with asthma. Overall, in support of my hypothesis, I identified a greater number of mtSNPs associated with asthma (20 vs. one) in maternal asthma dyads compared to paternal asthma dyads.

4.3.2 mtDNA haplogroup associations

Because there is no recombination in the mitochondrial genome, SNPs are inherited together and can be assigned to common haplotypes, or haplogroups. For the two African American cohorts (APIC and URECA) with whole genome sequence data, I could assign haplogroups with little

Figure 4.2. Summary of significant mtDNA haplogroup associations with maternal asthma. (**A**) The three mtDNA haplogroups associated with asthma in maternal asthma dyads in the APIC cohort are in bold. Two mtDNA haplogroups (L3e and L3e3b) highlight the convergence of the association signal in APIC on L3e3'4. Seven significant APIC mtSNPs (cohort-specific FDR-adjusted P-value < 0.05 in the single mtSNP MWAS) are shown with their corresponding haplogroup (HG). Summary statistics for the APIC and URECA cohorts are shown. A phylogeny based on the relationships between the haplogroups is displayed to the left side of the table.



				Maternal Asthma Results		Paternal Asthma Results	
HG	**Study**	**Significant mtSNPs**	**HG Freq.**	**Adj. P**	**OR (95% CI)**	**Adj. P**	**OR (95% CI)**
L3e3b	APIC	G4655A	0.04	0.070	1.23 [1.07, 1.41]	0.90	0.95 [0.83, 1.09]
	URECA	-	0.02	0.77	0.86 [0.64, 1.16]	0.92	0.95 [0.72, 1.25]
L3e3	**APIC**	**T6524C, G9554A, T10667C, A10816G**	**0.04**	**0.037**	**1.25 [1.09,1.42]**	**0.90**	**0.94 [0.82,1.07]**
	URECA	-	0.02	0.77	0.86 [0.64, 1.16]	0.92	0.95 [0.72, 1.25]
L3e3'4	**APIC**	**G5262A**	**0.05**	**0.011**	**1.26 [1.12,1.43]**	**0.90**	**0.96 [0.85,1.08]**
	URECA	-	0.03	0.71	0.82 [0.63, 1.06]	0.96	1.03 [0.81, 1.32]
L3e3'4'5	**APIC**	**A750G**	**0.05**	**0.011**	**1.26 [1.12,1.43]**	**0.90**	**0.96 [0.85,1.08]**
	URECA	-	0.03	0.71	0.77 [0.61, 0.97]	0.97	0.99 [0.8, 1.24]
L3e	APIC	-	0.17	0.70	1.05 [0.97, 1.12]	0.90	0.96 [0.89, 1.03]
	URECA	-	0.23	0.71	0.94 [0.84, 1.05]	0.92	1.03 [0.93, 1.15]

ambiguity. Therefore, I tested each low-frequency and common haplogroup (frequency > 0.01) in these two cohorts. I considered haplogroups at every tip and node of the mtDNA phylogenetic tree from Phylotree (*153*), which yielded 86 and 81 low-frequency and common haplogroups in APIC and URECA, respectively; 75 haplogroups overlapped between the two cohorts (see **Methods in 4.5.3**; see **Table S4.3** for haplogroup frequencies in each cohort). APIC and URECA included 76 of 86 (88.4%) and 73 of 81 (90.1%) canonical African (L) haplogroups, respectively. In APIC, I identified an L3e3'4'5 lineage of related mtDNA haplogroups that was associated with asthma in maternal asthma dyads (FDR-adjusted P-value < 0.05), but not in paternal asthma dyads (**Figure 4.2**). Seven of the eight maternal asthma-associated mtSNPs defined four African haplogroups that

were related to each other in the L3e3'4'5 lineage; three haplogroups were significantly associated with asthma in maternal asthma dyads. Among the three significant haplogroups, the haplogroup that was both the most significant and recently diverged in the lineage, L3e3'4 (FDR-adjusted P = 0.011; OR = 1.26 [1.12, 1.43]), was defined by one variant in Phylotree, G5262A mutation in the *MT-ND2* gene, which was itself the most significant mtSNP in APIC (FDR-adjusted P = 0.016; OR = 1.62 [1.28, 2.06]). Therefore, I suggest that the signal of association in the L3e3'4'5 lineage is at the L3e3'4 haplogroup-defining variant, G5262A. The A93G variant in the D-loop that was associated with asthma in maternal asthma dyads in the APIC cohort did not define any haplogroups in the L3e3'4 lineage or in the L3 lineage more generally (see **Table S4.4** for a list of haplogroups that are defined by A93G). No associations between haplogroups and asthma were identified in URECA. These results suggest that two individual mtSNPs, G5262A and A93G, contribute independently to asthma risk in African American maternal asthma dyads.

4.4 Discussion

I performed the largest study to date of mtSNP and haplogroup effects on asthma risk in ethnically diverse asthmatic individuals. Indeed, I identified 20 mtSNPs that were associated with asthma among maternal asthma dyads. These maternal asthma-associated mtSNPs were not associated with asthma in the paternal asthma dyads, supporting my hypothesis that some of the maternal asthma-associated risk for asthma in her children is due to transmission of mtDNA variation.

I identified associations in only three of the nine cohorts, and all significant associations were cohort-specific. However, three mtSNPs had consistent directions of effect in multiple cohorts, and two were in the D-loop, the non-coding region of mtDNA that controls DNA and RNA synthesis (*154*): A93G in African Americans and G499A in Brazilians. Neither have been previously associated with asthma, although variants in the D-loop were associated with asthma

in a study from Germany (*150*). Collectively, these results suggest that perturbations of mitochondrial replication and transcription may contribute to asthma risk. In contrast, most of the maternal asthma-associated variants were in protein-coding genes, particularly genes that encode subunits of NADH dehydrogenase, implicating variation in respiratory chain processes in asthma risk, discussed in more detail below.

In one African American cohort (APIC), I used whole-genome sequences to identify a mitochondrial haplogroup associated with asthma in maternal asthma dyads. By leveraging the tree structure of the mtDNA genome, I localized the association signal of seven associated variants to one African haplogroup L3e3'4, which is itself defined by only one variant, G5262A. G5262A is a missense mutation in the NADH dehydrogenase subunit 2 gene (*MT-ND2*), a core subunit of NADH dehydrogenase, or respiratory complex I. G5262A has not been previously reported to be associated with asthma, although variants in *MT-ND2* were significantly associated with asthma in the study from Germany, mentioned above (*150*). Thus, I suggest that the L3e3'4 haplogroup-defining variant, G5262A, in *MT-ND2* confers risk for asthma in African American maternal asthma dyads in the APIC cohort.

I could not assign mitochondrial haplogroups for the two cohorts with genotyping array data (GALA II and BIAS); however, I also observed maternal asthma-associated variants in genes that encode subunits of respiratory complex I (*MT-ND1, MT-ND2, MT-ND4, MT-ND4L,* and *MT-ND5*). Although the number of associated variants was too small to test for enrichments in specific mitochondrial annotations, there were greater proportions of maternal asthma-associated mitochondrial variants compared to all tested variants in respiratory complex I genes in both cohorts: 3/6 (50%) vs. 29/70 (41%) in Brazilians (BIAS), and 4/6 (67%) vs. 24/48 (50%) in Mexican Americans (GALA II). These mutations may be associated with perturbations in the

function of respiratory complex I, leading to overproduction of superoxide (*155*, *156*), the proximal mitochondrial reactive oxygen species, and subsequent oxidative stress (*157*, *158*), aberrant immune responses (*159*), and airway inflammation (*160*) in asthmatic individuals.

Although the results of my study overall support the hypothesis that maternally-inherited mtDNA variation is associated with asthma among children with an asthmatic mother, there are limitations. First, I did not identify associated variants in six of the cohorts. This may be due in part to the varying number of mtSNPs on each genotyping platform. However, ascertainment strategies and clinical heterogeneity between cohorts, due to variable ages of asthma onset, disease severity, and co-occurrence with allergies, may account for some of these differences. For example, I reported significant associations in one African American cohort, APIC, that could not be replicated in another African American cohort (URECA). Although both cohorts were comprised of childhood-onset asthma cases and had whole-genome sequences for assigning mtSNPs and haplogroups, the APIC cohort was nearly 3-times larger than the URECA cohort and included more maternal asthma dyads (318 vs. 123, respectively). Moreover, URECA is a birth cohort whereas APIC was a one-year longitudinal study. Due to differences in ascertainment, the APIC subjects also had more moderate-to-severe asthma (*151*), whereas the URECA subjects had more mild asthma (*152*). Any of these differences could account for the lack of replication between these two cohorts. None of the eight significant mtSNPs in APIC were genotyped in the other two African American cohorts, but these cohorts also differed with respect to ascertainment (both were case-control studies), and SAPPHIRE was comprised of mainly adults whereas CAG was comprised mostly of children (*161*). Finally, both mitochondrial and autosomal genes encode proteins of the respiratory chain and therefore variants in mitochondrial genes on different autosomal backgrounds may differ between cohorts and potentially explain some of the discrepant

results. Additional asthma cohorts with larger sample sizes and harmonized phenotype and genotype data are needed to address these outstanding questions.

To my knowledge, this is the first study to interrogate the relationship between maternal asthma and mtDNA variation on childhood asthma risk in her children. my data indicate that asthma-associated mitochondrial variation that is inherited from asthmatic mothers may occur at low frequencies and be ethnic- and even cohort-specific. Overall, these data support my hypothesis that some of the risk for asthma in offspring of asthmatic mothers may be transmitted via genetic variation inherited in the mitochondrial genome.

4.5 Methods

4.5.1 Sample composition

I used autosomal and mitochondrial genotype data from previously published studies for seven cohorts (*47, 135, 162*), and unpublished whole-genome sequence data for two cohorts. I included in my study only cohorts with ≥100 asthma cases and ≥20 in each of the four comparison groups: cases with an asthmatic mother, cases with an asthmatic father, cases without an asthmatic mother, and cases without an asthmatic father. Nine cohorts met these criteria; four were comprised of African American subjects, and one each were comprised of subjects of Brazilian, European American, Honduran, Mexican American, and Puerto Rican subjects. Nine additional cohorts did not satisfy these criteria and were excluded from my studies (**Table S4.5**). My study was comprised of 3,190 asthma cases, which included 3,155 and 2,960 cases with maternal and paternal asthma information, respectively (**Table 4.1**). I used each cohort's definition of asthma, all of which included at least self- or parent-reported doctor's diagnosis of asthma. Descriptions of each of the study cohorts can be found in prior published work (*47, 135, 151, 152, 161, 162*) and in the

Supplementary Note in 4.6.3. These studies were approved by the local institutional research ethics committees for each study, and at the University of Chicago.

4.5.2 Genotypes and quality control

Genotypes for seven studies were obtained from three different genotyping arrays: the Illumina Infinium HumanExome BeadChip, the Axiom Genome-Wide LAT 1, and/or the Illumina Omni2.5 BeadChip arrays (**Table S4.6**). Autosomal SNPs were used to calculate ancestry PCs, which were provided to us for each cohort. These arrays interrogated 200, 120 and 521 mtSNPs, respectively. For the two studies with whole-genome sequence data (unpublished), mtSNPs were called using a custom GATK4 mitochondria pipeline (https://github.com/gatk-workflows/gatk4-mitochondria-pipeline).

Quality control of autosomal genotypes was performed previously in each cohort, as described (*47, 135, 162*). Briefly, samples were excluded for gender and ancestry discordances with self-reported race or ethnicity, using sex chromosomal SNPs and principal components analysis, respectively; one of each duplicate sample as well as individuals with missing asthma or covariate information were removed. One cohort comprised of trios further excluded those with Mendelian errors. For each study, I excluded mtSNPs with genotype call rates < 0.95. MtSNPs annotated as problematic on the Illumina Infinium HumanExome BeadChip array design group were also removed, as previously reported (*135*). For mtSNPs called from whole-genome sequences, sites with low read depth (<20 reads) were removed. Each study had between 81 and 1,388 mtSNPs that passed quality control (**Table S4.6**). Finally, for each cohort, I excluded subjects with <95% of mtSNPs passing quality control. I retained between 113 and 829 asthma cases in the nine cohorts for single-variant association testing (**Table S4.1**).

4.5.3 Haplogroup assignments

For the two cohorts with whole-genome sequencing data (APIC and URECA), all mtSNPs that passed QC were used to predict mitochondrial haplogroups using HaploGrep2 (*163*). HaploGrep2 assigns mtDNA haplogroups for each subject by traversing the phylogenetic tree in Phylotree version 17 (*153*), which is comprised of 5,437 haplogroups from 24,275 mtDNA sequences. In APIC and URECA, 1,388 and 805 mtSNPs, respectively, were used to assign mitochondrial haplogroups. Of the 320 and 124 asthma cases in APIC and URECA, respectively, two and one subjects with haplogroup quality scores < 0.80 were removed.

I then assigned all possible mtDNA haplogroups for each individual that was present at a frequency ≥0.01 in both APIC and URECA. For example, an individual who was assigned by HaploGrep2 as L0a2 was also assigned to L0a and L0, the broader haplogroup assignments at the nodes in Phylotree. After repeating this process for all subjects in APIC and URECA, there were 120 mtDNA haplogroups (see **Table S4.7** for a list of mtDNA haplogroups). For the 120 haplogroups, all APIC and URECA subjects were assigned a 1 or 0 for the presence or absence of the haplogroup, respectively. Finally, only those mtDNA haplogroups that were present at a frequency ≥0.01 in APIC and URECA, separately, were kept for haplogroup association testing. This yielded 86 and 81 haplogroups in APIC and URECA, respectively; 75 haplogroups overlapped between both cohorts, which included the five haplogroups in the L3e lineage (L3e, L3e3'4'5, L3e3'4, L3e3, and L3e3b). (**Table S4.1; Table S4.3**).

4.5.4 Statistical analyses

I included low-frequency and common mtSNPs and haplogroups (≥0.01) in each MWAS. Associations with maternal asthma or paternal asthma among children with asthma were tested for each mtSNP and haplogroup, using a linear mixed model, as implemented in GEMMA (*164*),

adjusting for age, sex, five autosomal ancestry principal components, and an autosomal genetic relatedness matrix. As a control for my hypothesis, I repeated the same analysis between asthmatic children with and without a father with asthma. Because the mtDNA genome is one non-recombining "locus", many mtSNPs are in strong LD. I therefore used an FDR correction, a less conservative multiple testing procedure, and considered cohort-specific FDR-corrected p-values <0.05 as significant.

4.6 Supplementary information

4.6.1 Supplementary tables

Table S4.1. Number of variants tested in each parental asthma-specific analysis and in both analyses. The number of haplogroups tested in the APIC and URECA cohorts are included in parentheses.

Ethnicity	Study	Maternal asthma	Paternal asthma	Tested in both analyses
African American	APIC	236 (86)	234 (86)	234 (86)
	CAG	39	39	39
	SAPPHIRE	45	45	44
	URECA	214 (82)	212 (81)	212 (81)
Brazilian	BIAS	85	73	70
European American	CAMP	35	35	35
Honduran	HONDAS	65	65	65
Mexican American	GALA II	53	48	48
Puerto Rican	GALA II	93	84	82

Table S4.2. Minor allele frequencies of the significant mtSNPs in the nine cohorts. SNPs that were not genotyped or did not pass QC are indicated with a "-". AA, African American. B, Brazilian. EA, European American. H, Honduran. MA, Mexican American. PR, Puerto Rican.

mtSNP	APIC (AA)	CAG (AA)	SAPPHIRE (AA)	URECA (AA)	BIAS (B)	CAMP (EA)	HONDAS (H)	GALAII (MA)	GALAII (PR)
A93G	0.047	-	-	0.073	0.073	-	0.154	-	-
G499A	0.008	0.012	0.018	0.020	0.128	0.026	0.015	-	0.013
A750G	0.475	-	-	0.476	-	-	-	-	-
A1811G	0.005	0	0.003	0.004	0	0.140	0	0.019	0.016
A3480G	0	-	-	0	0	-	0	0.010	0.004
G4655A	0.020	-	-	0.012	-	-	-	0.002	0.001
T4977C	0.003	-	-	0.004	0.117	-	0.011	0.232	0.043
G5262A	0.025	-	-	0.016	-	-	-	-	-
T6524C	0.020	-	-	0.012	-	-	-	-	-
G6962A	0	-	-	0	0.061	-	0	-	-
G9554A	0.025	-	-	0.016	-	-	-	-	-
A10550G	0	-	-	0	0	-	0	0.010	0.001
T10667C	0.020	-	-	0.012	-	-	-	-	-
A10816G	0.020	-	-	0.012	-	-	-	-	-
T11299C	0.002	-	-	0	-	-	-	0.011	0.002
A11959G	0	-	-	0.004	0.055	-	0	-	-
A12308G	0.009	0.018	0.022	0.016	0	0.237	0	0.030	0.046
G12372A	-	-	-	-	0.012	-	0	0.034	0.047
T12477C	0.003	-	-	0	0.061	-	0	-	-
C15535T	0.002	-	-	0.004	0.122	-	0.011	-	-

Table S4.3. Haplogroup frequencies in APIC and URECA. The 86 and 81 low-frequency and common (≥0.01) haplogroups that were tested for associations in APIC and URECA, respectively, are shown. Haplogroups that had a frequency <0.01 are indicated with a "-".

Haplogroup	APIC Frequency	URECA Frequency
H	0.044	0.027
H1	0.024	-
HV	0.047	0.035
JT	-	0.018
L0	0.057	0.062
L0a	0.057	0.062
L0a1	0.041	0.062
L0a1a	0.027	0.035
L0a1a2	0.02	0.027
L0a2	0.017	-
L1	0.189	0.115
L1'2'3'4'5'6	0.943	0.938
L1b	0.088	0.08
L1b1	0.081	0.062
L1b1a	0.081	0.062
L1b1a3	0.017	-
L1c	0.101	0.035
L1c1	0.027	0.035
L1c1'2'4'5'6	0.064	0.035

Table S4.3. Haplogroup frequencies in APIC and URECA (continued).

L1c1'2'4'6	0.054	0.035
L1c1a	0.014	0.027
L1c1a'b'd	0.02	0.027
L1c1a+@198	0.014	0.027
L1c2	0.027	-
L1c2'4	0.03	-
L1c3	0.037	-
L1c3a	0.024	-
L2	0.267	0.327
L2'3'4'5'6	0.753	0.823
L2'3'4'6	0.753	0.823
L2a	0.176	0.159
L2a'b'c'd	0.264	0.319
L2a1	0.176	0.159
L2a1'2'3'4	0.176	0.159
L2a1+143	0.084	0.062
L2a1+143+16189	0.03	0.018
L2a1a	0.037	0.062
L2a1a2	0.017	0.035
L2a1c	0.037	0.027
L2a1f	0.041	0.018
L2a1i	0.017	0.018
L2b	0.051	0.097
L2b'c	0.088	0.159
L2b'c'd	0.088	0.159
L2b1	0.037	0.071
L2b1a	0.03	0.044
L2b1a3	0.017	0.027
L2b2	-	0.018
L2c	0.037	0.062
L2c2	0.014	0.027
L3	0.476	0.496
L3'4	0.486	0.496
L3'4'6	0.486	0.496
L3b	0.105	0.133
L3b'f	0.132	0.097
L3b1	0.044	0.053
L3b1a	0.037	0.053
L3b1a1	-	0.027
L3b1a1a	-	0.027
L3c'd	0.047	0.027
L3d	0.047	0.027
L3d1	0.03	-
L3d1b	0.017	-
L3e	0.162	0.23
L3e'i'k'x	0.162	0.23
L3e1	0.03	0.088
L3e1e	0.017	0.053
L3e2	0.084	0.097
L3e2a	0.027	0.053

Table S4.3. Haplogroup frequencies in APIC and URECA (continued).

L3e2a1	0.027	0.053
L3e2a1b	0.027	0.053
L3e2b	0.057	0.044
L3e2b+152	-	0.035
L3e2b1	0.027	-
L3e2b1a	0.027	-
L3e3	0.041	0.027
L3e3'4	0.047	0.035
L3e3'4'5	0.047	0.044
L3e3b	0.037	0.027
L3f	0.061	0.035
L3f1	0.054	0.035
L3f1b	0.054	0.035
L3f1b+16292	0.054	0.035
L3f1b+16292+150	0.027	0.018
L3f1b1	0.024	0.018
L3f1b1a	0.014	0.018
M	0.027	0.018
N	0.091	0.115
R	0.084	0.106
R0	0.047	0.035
R2'JT	-	0.018
U	0.02	0.035

Table S4.4. All assigned Phylotree haplogroups for A93G (D-loop).

B5a2a	H1ao	H1b1d	H1e1a1	H1e3	H29	H2a2a1f
H3h7	H49a2	H7a1a	HV1a2b	K2b1a4	L0a'b'g	L0d2a2
L1c1a2	L1c3c	L2c	R0a1a2	R0b	R5a1	V7

Table S4.5. Excluded cohorts. Nine cohorts were excluded from my study because they had <100 asthma cases and/or <20 cases in each of the four comparison groups: cases with maternal asthma (MA), cases with paternal asthma (PA), cases without MA, and cases without PA. The total number of cases refers to those cases with asthma information for at least one parent.

Ethnicity	**Study**	**Male / Female**	**Cases with MA / without MA / Unknown**	**Cases with PA / without PA / Unknown**	**Total**
African Americans	ABRIDGE	29/50	24/53/2	22/42/15	79
	GRAAD	13/16	6/18/5	7/22/0	29
Barbadian	BAGS	39/54	37/48/8	12/34/47	93
European	COPSAC	52/38	65/25/0	26/64/0	90
European American	CAG	54/42	23/58/15	36/57/3	96
	COAST	36/21	17/38/2	25/32/0	57
Mexican American	GALA II	30/23	13/40/0	2/48/3	53
Mexican	MCCAS	123/91	8/194/12	15/191/8	214
Puerto Rican	GALA II	43/38	36/44/1	21/53/7	81

Table S4.6. Number of variants that passed quality control in both analyses of 3,155 maternal asthma and 2,960 paternal asthma dyads for each minor allele frequency category. 1. Illumina Infinium HumanExome BeadChip (*135*) 2. Axiom Genome-Wide LAT (*162*) 1 3. Illumina Omni2.5 BeadChip (*47*) 4. Whole-genome sequencing

Ethnicity	Study	Platform	All SNPs	Rare MAF<0.01	Low Frequency 0.01≤MAF≤0.05	Common MAF>0.05
African American	APIC	4	1388	1105	204	79
	CAG	1	100	48	31	21
	SAPPHIRE	1	123	73	26	24
	URECA	4	805	526	201	78
Brazilian	BIAS	3	99	9	47	43
European American	CAMP	1	83	50	18	15
Honduran	HONDAS	3	81	16	11	54
Mexican American	GALA II	1, 2	198	136	36	26
Puerto Rican	GALA II	1, 2	191	78	70	43

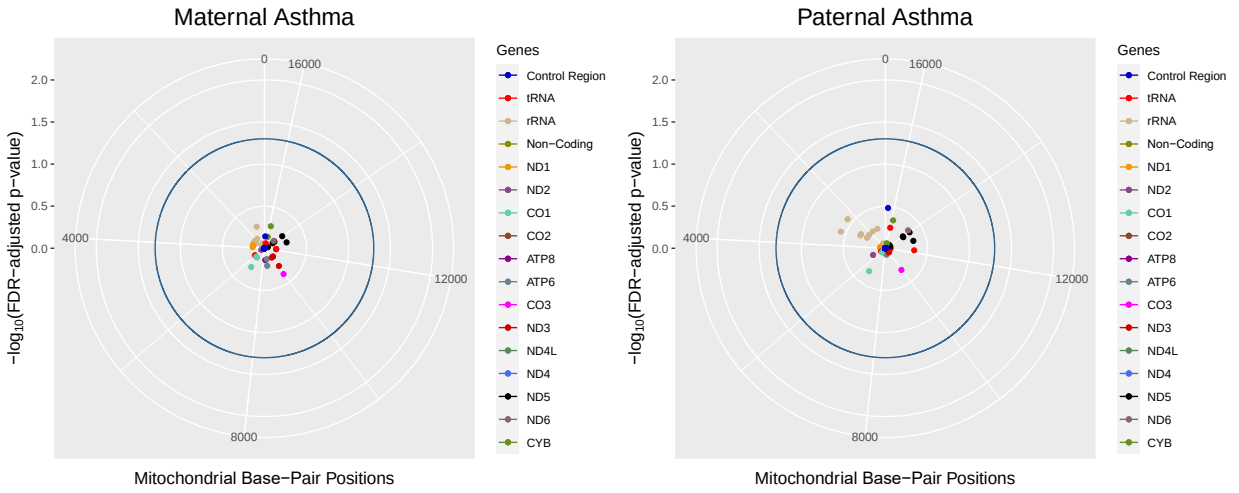
Table S4.7. All 120 low frequency and common haplogroups (>1%) at every node and tip of the phylogenetic tree in HaploGrep2 for the APIC and URECA cohorts.

A	K	L1c2	L2b'c	L3d1	L3f1
A2	L0	L1c2'4	L2b'c'd	L3d1b	L3f1b
A2+(64)	L0a	L1c3	L2b1	L3e	L3f1b+16292
A2k	L0a1	L1c3a	L2b1a	L3e'i'k'x	L3f1b+16292+150
A2k1	L0a1a	L2	L2b1a3	L3e1	L3f1b1
A2k1a	L0a1a2	L2'3'4'5'6	L2b2	L3e1a	L3f1b1a
B	L0a2	L2'3'4'6	L2c	L3e1e	M
B2	L1	L2a	L2c2	L3e2	M8
C	L1'2'3'4'5'6	L2a'b'c'd	L3	L3e2a	M80'D
C1	L1b	L2a1	L3'4	L3e2a1	N
C1b	L1b1	L2a1'2'3'4	L3'4'6	L3e2a1b	R
C1b2	L1b1a	L2a1+143	L3b	L3e2b	R0
CZ	L1b1a3	L2a1+143+16189	L3b'f	L3e2b+152	R2'JT
D	L1c	L2a1a	L3b1	L3e2b1	T
D1	L1c1	L2a1a2	L3b1a	L3e2b1a	T2
D4	L1c1'2'4'5'6	L2a1c	L3b1a+@16124	L3e3	U
H	L1c1'2'4'6	L2a1c4	L3b1a1	L3e3'4	U2'3'4'7'8'9
H1	L1c1a	L2a1f	L3b1a1a	L3e3'4'5	U8
HV	L1c1a'b'd	L2a1i	L3c'd	L3e3b	U8b
JT	L1c1a+@198	L2b	L3d	L3f	U8b'c

4.6.2 Supplementary figures

Figure S4.1. MWAS results in the six cohorts with no maternal asthma-associated mtSNPs. Solar plots for maternal and paternal asthma MWAS are shown for African Americans in CAG (A), SAPPHIRE (B), and URECA (C); European Americans in CAMP (D), Hondurans in HONDAS (E), and Puerto Ricans from GALA I (F).

A. African American (CAG)



B. African American (SAPPHIRE)

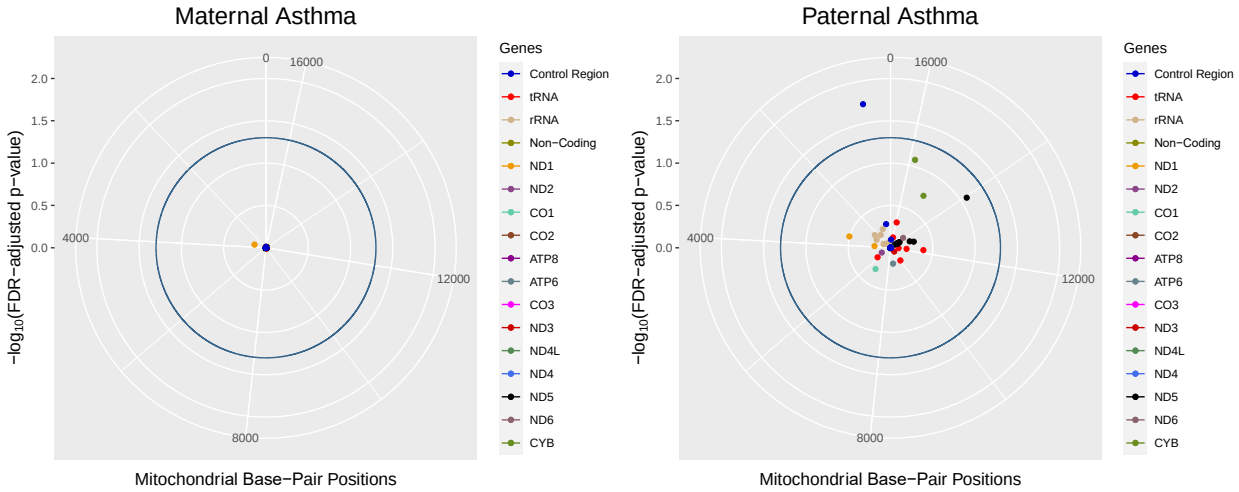
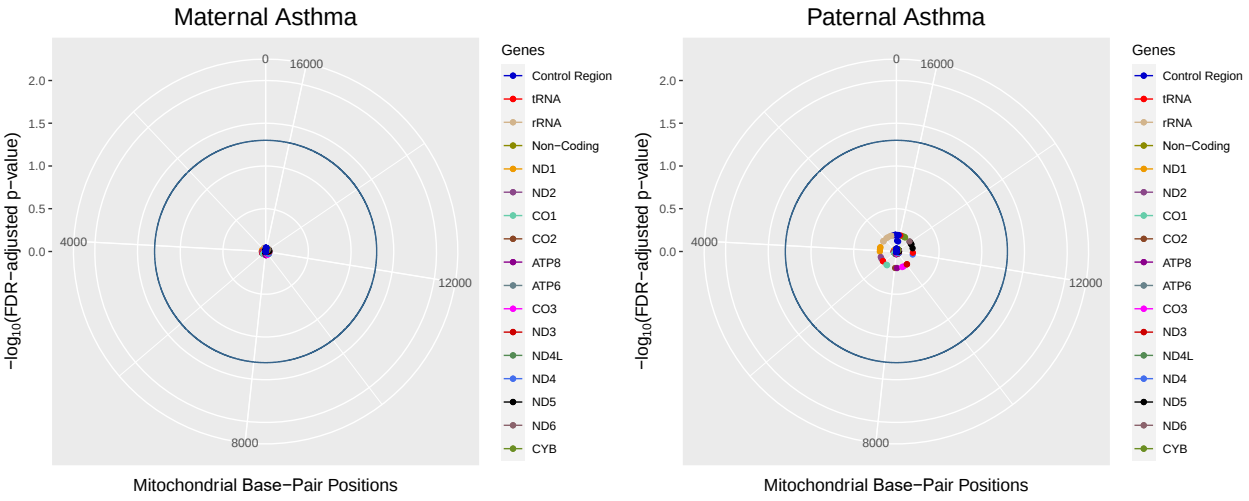


Figure S4.1. MWAS results in the six cohorts showing no maternal asthma-associated mtSNPs (continued).

C. African American (URECA)



D. European American (CAMP)

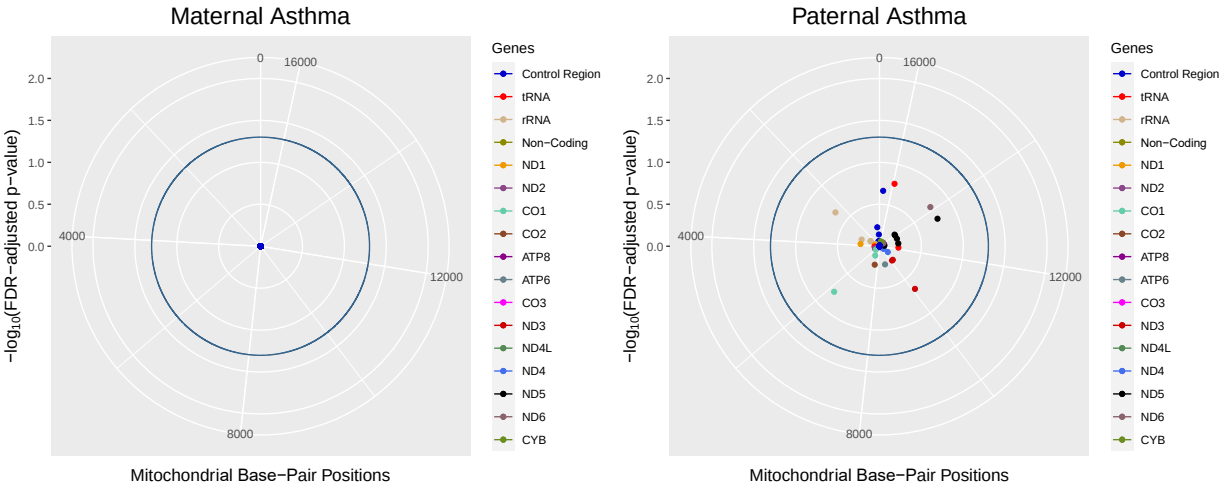
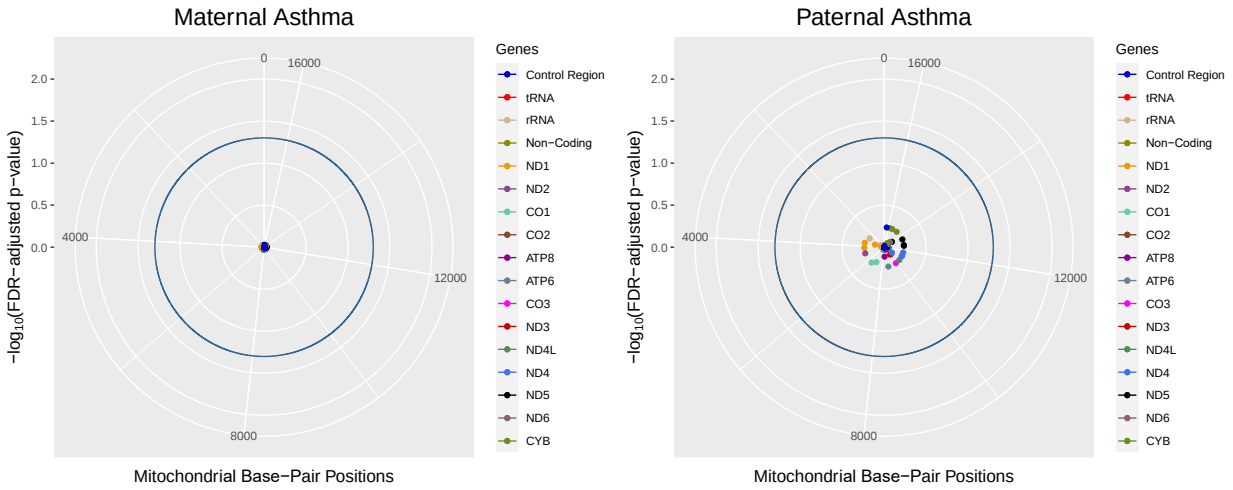
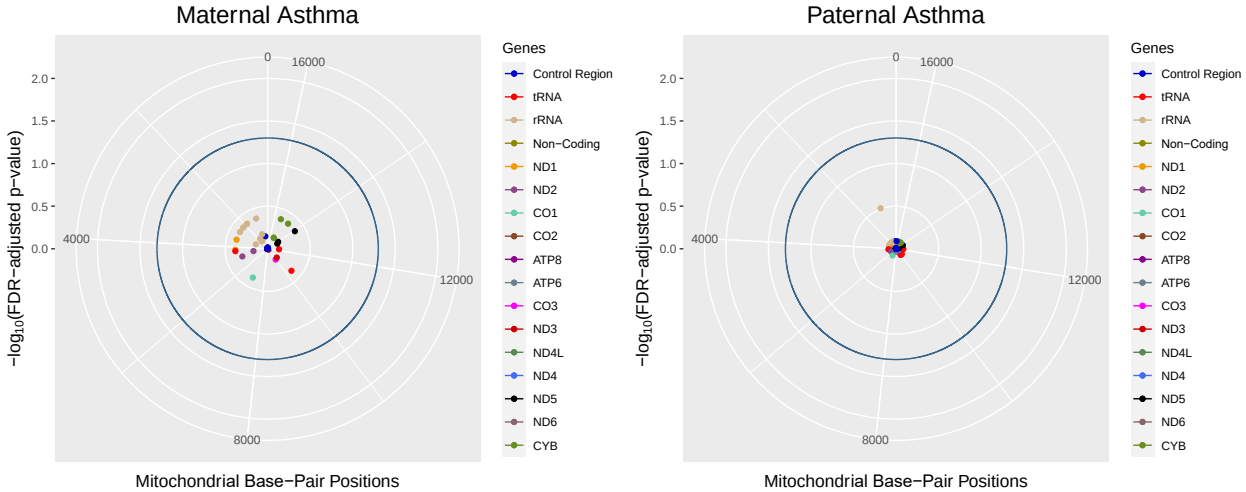


Figure S4.1. MWAS results in the six cohorts showing no maternal asthma-associated mtSNPs (continued).

E. Honduras (HONDAS)



F. Puerto Rican (GALA II)



4.6.3 Descriptions of study cohorts

Asthma Phenotypes in the Inner City (APIC) (*151*)

APIC included 6-17-year-old children and adolescents with a physician- or parent-reported diagnosis of asthma, from low-income areas of 9 U.S cities (Baltimore MD, Boston MA, Chicago IL, Cincinnati OH, Dallas TX, Denver CO, Detroit MI, New York NY, Washington DC). Research protocols were approved by each center's institutional review boards, and written informed assent was obtained from each participant and consent legal guardians of the participants. All children had physician's diagnosis of asthma. Parents provided self-reported diagnosis of asthma. See (*151*) for additional criteria.

Chicago Asthma Genetics (CAG) (*161*)

CAG included both European American and African American unrelated cases and control subjects collected at the University of Chicago from a) families ascertained through affected sib pairs, b) affected children and their parents, c) adults and children attending asthma clinics. Individuals in a-c were recruited in the adult and/or pediatric asthma clinics at University of Chicago Hospital. All children had physician's diagnosis of asthma. Parents provided self-reported diagnosis of asthma. See (*161*) for additional criteria.

Study of Asthma Phenotypes and Pharmacogenomic Interactions by Race-Ethnicity (SAPPHIRE) (*161*)

Study participants were ascertained through a large health system serving southeast Michigan. All individuals were reported as African American. SAPPHIRE cases met the following criteria: age 12-56 years, had a diagnosis of asthma (based on both patient report and documentation in the medical record). Parents provided self-reported diagnosis of asthma. See (*161*) for additional criteria. The Institutional Review Board of Henry Ford Health System approved all of the components of this study.

Urban Environment and Childhood Asthma (URECA) (*152*)

URECA is an observational birth cohort of children from birth to age 7. Subjects were recruited from low-income areas from four cities (Baltimore, MD, Boston, MA, New York City, NY, and St. Louis, MO). Children and their parents had a physician's diagnosis of asthma. See (*152*) for additional criteria.

Brazilian Immunogenetics of Asthma and Schistosomiasis (BIAS) (47)

BIAS is a population ascertainment designed study of asthma and schistosomiasis in the rural district of Conde, Bahia, located 200 km north of Salvador, Brazil. Subjects were from five communities (Buri, Camarao, Genipapo, Sempre Viva, and Cobo). The data used in this study were from 2 large pedigrees with 535 and 310 members comprising 318 nuclear families (ages 5-85 years). Child and their parents had either physician or self-reported diagnosis of asthma. See (*47*) for additional criteria.

Childhood Asthma Management Program (CAMP) (*161*)

CAMP is comprised of non-Hispanic white subjects from a multi-center clinical trial that followed 1,041 children with asthma for four years and 84% of the original participants for 12 years. All participants had mild to moderate asthma. Children and their parents had a physician's diagnosis of asthma. See (*161*) for additional criteria.

Honduran subjects who self-reported as Garifuna (HONDAS) (47)

HONDAS is a population-based study of asthma in the Garífuna (Black Carib) people from the northern coast of Honduras, a population of African and Red Carib Native Amerindian ancestry; and of Honduran autochthonous Amerindian populations of Mesoamerican or South American ancestry. Subjects were recruited from 12 villages (Bajamar, Travesía, Corozal, Sambo Creek, Alfonzo Lacayo, Belén Gualcho, San Juan, Tornabé, Triunfo de la Cruz, Cristales, Río Negro, and Santa Fe; aged 5-85). Participants responded to a standardized, interviewer administered questionnaire that includes a modified version of the 1987 American Thoracic Society Division of Lung Disease Epidemiology Questionnaire to collect information on asthma and allergy symptoms. The questionnaire was used to determine whether a subject has a history of both self-reported and physician diagnosed asthma (asthma case), or had no history of asthma (healthy control).

Genes-environments and Admixture in Latino Americans (GALA) (*161*)

GALA includes children and their biological parents recruited from schools, clinics and hospitals that cared for Latino patients at four sites in San Francisco, CA, New York City, NY, Puerto Rico, and Mexico City, Mexico. Medical records were reviewed to identify patients with physician-diagnosed mild or moderate-severe asthma based on medical billing records (ICD 9 codes). Based on interviews and questionnaire data, children were included in the study if they were between the ages of 8-40 with physician diagnosed mild to moderate-severe asthma and had experienced two or more symptoms in the previous two years at time of recruitment (including wheezing, coughing and/or shortness of breath).

CHAPTER 5

CONCLUSION

Genetic studies of asthma have largely focused on genome-wide association studies, which have identified >150 independent loci associated with risk of asthma. Yet, the variants at these loci explain little of the heritability (*43-47*). Genetic variation that lies outside the autosomal genome and epigenetic variation throughout the genome may provide more clues to the etiology of asthma. Advances in next-generation sequencing technologies have allowed for investigations into non-canonical mechanisms of disease (*165, 166*). Here, I leveraged multi-omic approaches to explore one of the most consistent risk factors for childhood-onset asthma, the presence of asthma in the mother (*49*). Specifically, I found genetic and epigenetic mechanisms underlying maternal transmission of asthma risk to her children.

5.1 ADAR-mediated editing of miR-200b-3p is associated with asthma severity in children of asthmatic mothers

In chapter 2, I describe the first study to investigate the role of ADAR-mediated miRNA editing in bronchial epithelial cells on asthma risk. To do this, I used small RNA-seq data and performed a genome-wide scan to identify A-to-I edited sites (*88*). Overall, I identified 19 A-to-I edited sites in miRNAs. Similar to another study of miRNA editing in cancer (*64*), most of these sites were in seed sequences, suggesting altered miRNA-mediated gene silencing of downstream target genes. Although most sites in my study were rare (present in <25% of subjects), one site at the 5th position of miR-200b-3p was present in nearly all subjects and was significantly decreased in cases with moderate and severe asthma compared to controls. The predicted target genes of the edited miR-200b-3p was specifically enriched for IL-4 and interferon gamma signaling pathways and included the *SOCS1* (suppressor of cytokine signaling) gene. *SOCS1*, a centrally important negative

regulator of Type 2 cytokine and interferon signaling, was also highly expressed only among moderate and severe asthma cases. Both miR-200b-3p editing and *SOCS1* expression were significantly associated with bronchoalveolar lavage (BAL) eosinophil levels and an epithelial cell gene signature of Type 2 asthma (*75*), characteristic of a steroid-responsive endotype of severe asthma. Collectively, these results suggest for the first time miRNA editing as an epigenetic mechanism underlying the features of moderate-to-severe asthma and identify potential new therapeutic targets for treatment of severe symptoms.

I also found that the expression of the primary *ADAR* enzyme, an interferon-inducible gene (*65*), was among the top 2% of genes expressed in bronchial epithelial cells and was significantly decreased in asthma cases with an asthmatic mother compared to both asthma cases without an asthmatic mother and controls. Overall, maternal asthma may alter the gene regulatory landscape at the *ADAR* locus in bronchial epithelial cells *in utero*, with long-lasting effects on ADAR-mediated editing of the 5th position of miR-200b-3p. I further suggest that this "exposure" results in reduced suppression of *SOCS1* in bronchial epithelial cells from asthmatics, leading to more symptoms and more severe asthma in these individuals.

5.2 Maternal asthma-associated DNA methylation signatures reflect endotypes of severe asthma

In chapter 3, I characterized the genome-wide DNA methylation profiles of bronchial epithelial cells from asthma cases with an asthmatic mother. More CpGs were differentially methylated in cases with a mother with asthma compared to cases without a mother with asthma. I also identified six DNA co-methylation signatures that were associated with distinct genetic pathways and clinical measures of asthma. All six signatures were associated with severe asthma and lung function measures; five signatures were also associated with features of atopic asthma (*86*),

including total serum IgE and BAL eosinophil levels. One module was also correlated with genes enriched for selenocysteine synthesis, a pathway that produces antioxidant selenoproteins that may alleviate inflammation (*103, 104*), suggesting that selenium supplementation in early life may be a potential treatment for asthma among children with an asthmatic mother (*122*). Only one module was not associated with atopic asthma features, and the genes that were correlated with this module were enriched for diverse immune-related processes and diseases, with the top pathway being "interferon signaling" (*101*). One of the correlated genes of this module was *ADAR*, which suggests that maternal asthma-associated DNA methylation profiles leads to reduced *ADAR* levels, and links the findings in Chapter 2 to those in Chapter 3. These findings indicate that maternal asthma *in utero* or in early life alters diverse epigenetically-mediated developmental pathways that lead to subtypes of severe asthma in adulthood, including hard-to-treat, non-atopic asthma.

5.3 Associations between mitochondrial variants and haplogroups with asthma in ethnically diverse mother-child dyads

In Chapter 4, I describe the largest and most ethnically diverse study to date examining the effects of maternally-inherited mitochondrial variation on asthma risk. To date, no previous studies have interrogated the effects of mitochondrial variation on asthma risk in children of asthmatic mothers or have included populations of non-European descent (*148-150*).

Overall, I identified 20 mitochondrial variants that were associated with asthma only in maternal asthma dyads in three study cohorts of African American, Brazilian and Mexican American ancestry. These 20 maternal asthma-associated mitochondrial variants were not associated with asthma in paternal asthma dyads. In fact, across all tested variants and cohorts, only one variant was associated with asthma in paternal asthma dyads, which was not among those associated in maternal asthma dyads. The 20 maternal asthma-associated mitochondrial variants

were low-frequency and ethnic-specific. Remarkably, none of the variants overlapped between the three study cohorts, which can partly be explained by the population-specificity of the variant allele frequencies, as previously described in a rare-variant exome study in these same subjects (*135*), as well as the fact that not all variants were present in all populations due to utilization of different genotyping platforms.

By leveraging the phylogeny of known mitochondrial haplogroups (*153*), I also identified one individual variant (A93G) and one haplogroup-defining variant (G5262A of haplogroup L3e3'4) that were associated with asthma in African American maternal asthma dyads. G5262A and many other maternal asthma-associated variants in the Brazilian and Mexican American cohorts were in genes of the NADH dehydrogenase complex, indicating that defects in respiratory complex I may contribute to features of asthma. Although the effects were modest, similar to autosomal variants, my study suggested that some of the risk of maternal asthma may be due to variation inherited in the mitochondrial genome.

5.4 Future directions

In this research, I explored three mechanisms that may explain maternal risk of asthma: ADAR-mediated microRNA editing, DNA methylation, and mitochondrial DNA. I have provided evidence for each, but all require further in-depth investigations. I propose the following approaches to characterize these mechanisms, as well as experiments to discern their clinical significance.

In chapter 2, I identified a novel epigenetic mechanism, A-to-I editing at the 5th position of miR-200b-3p, to be decreased in subjects with moderate-to-severe asthma. I also found the expression of the primary editing enzyme, *ADAR*, to be downregulated only in asthma cases with an asthmatic mother. These findings need to be replicated in independent datasets, and in other

asthma-associated cell types, such as nasal epithelial cells and immune cells. This same A-to-I site was also increased in lung cancer cells (*66, 67*), promoting the idea of an optimal range of A-to-I editing at this site to prevent both cancer and asthma. Thus, a microRNA therapeutic may require balanced A-to-I editing at this site to promote a healthy bronchial epithelium. Additionally, a thorough characterization of the functional consequences of this A-to-I edited site is necessary in both healthy and diseased bronchial epithelium. Downregulation of *ADAR* in cases with an asthmatic mother also suggests decreased A-to-I editing at sites throughout the genome. Future studies should investigate altered A-to-I editing across the genome in asthma and in other immune-mediated diseases.

In chapter 3, I identified genome-wide DNA methylation patterns in bronchial epithelial cells from asthma cases with an asthmatic mother that reflected distinct endotypes of severe asthma. As proposed by others (*56, 59, 95*), these epigenetic changes are thought to occur in a maternal asthma *in utero* or early life environment and persist into adulthood. There was an insufficient number of controls with a mother with asthma in my study, so I could not fully disentangle the effects of maternal asthma and asthma on DNA methylation profiles. In future investigations, large cohorts of asthmatic and non-asthmatic children with and without an asthmatic mother should be examined. To fully understand the impacts of maternal asthma on the child's epigenome, longitudinal studies of mother-child dyads are required to directly assess maternal asthma effects on epigenetic and transcriptional landscapes and their associations with clinical features of asthma severity.

In chapter 4, I found that variation inherited in the mitochondrial genome contributed to asthma risk only in asthmatic children with an asthmatic mother. Specifically, I identified ethnic-specific variants that confer risk for asthma only in maternal asthma dyads of diverse ancestries.

Because these variants were population and cohort-specific, they may be associated with asthma risk via interactions with disease-promoting environmental exposures or with population-specific variants in autosomal genes. Importantly, genes that encode different components of the respiratory chain are in both the autosomal and mitochondrial genomes (*167*). It is possible that variants and haplotypes in these genes may together increase risk for asthma. Additionally, admixed populations of African ancestry (i.e. Puerto Ricans, African Americans) have the highest prevalence of asthma (*9*). One possible explanation for this epidemiological phenomenon is the interaction between African-associated mitochondrial variation and non-African-associated autosomal variation or vice versa. Larger population cohorts of mother-child dyads in ethnically diverse populations are needed to further investigate the complex relationship between the autosomal and mitochondrial genomes.

5.5 Concluding remarks

I describe three research studies, focused on distinct genetic or epigenetic mechanisms, that address an important epidemiological phenomenon – maternal asthma as one of the strongest risk factors for childhood-onset asthma. As sequencing becomes cheaper and advanced statistical and computational methods become available, I can further characterize these genetic and epigenetic mechanisms in larger population cohorts and in relevant asthma-related tissues and cell types. Ultimately, the discoveries made in this thesis serve as a basis for future studies and enhance potential diagnostics and therapeutics for the treatment of asthma.

BIBLIOGRAPHY

1. M. Masoli, D. Fabian, S. Holt, R. Beasley, G. I. f. A. G. Program, The global burden of asthma: executive summary of the GINA Dissemination Committee report. *Allergy* **59**, 469-478 (2004).

2. S. E. Wenzel, Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat Med* **18**, 716-725 (2012).

3. K. F. Chung, S. E. Wenzel, J. L. Brozek, A. Bush, M. Castro, P. J. Sterk, I. M. Adcock, E. D. Bateman, E. H. Bel, E. R. Bleecker, L. P. Boulet, C. Brightling, P. Chanez, S. E. Dahlen, R. Djukanovic, U. Frey, M. Gaga, P. Gibson, Q. Hamid, N. N. Jajour, T. Mauad, R. L. Sorkness, W. G. Teague, International ERS/ATS guidelines on definition, evaluation and treatment of severe asthma. *Eur Respir J* **43**, 343-373 (2014).

4. L. J. Akinbami, A. E. Simon, L. M. Rossen, Changing Trends in Asthma Prevalence Among Children. *Pediatrics* **137**, (2016).

5. K. G. Tantisira, R. Colvin, J. Tonascia, R. C. Strunk, S. T. Weiss, A. L. Fuhlbrigge, C. A. M. P. R. Group, Airway responsiveness in mild to moderate childhood asthma: sex influences on the natural history. *Am J Respir Crit Care Med* **178**, 325-331 (2008).

6. D. S. Postma, Gender differences in asthma development and progression. *Gend Med* **4 Suppl B**, S133-146 (2007).

7. J. G. Zein, S. C. Erzurum, Asthma is Different in Women. *Curr Allergy Asthma Rep* **15**, 28 (2015).

8. S. J. Uekert, G. Akan, M. D. Evans, Z. Li, K. Roberg, C. Tisler, D. Dasilva, E. Anderson, R. Gangnon, D. B. Allen, J. E. Gern, R. F. Lemanske, Jr., Sex-related differences in immune development and the expression of atopy in early childhood. *J Allergy Clin Immunol* **118**, 1375-1381 (2006).

9. J. E. Moorman, R. A. Rudd, C. A. Johnson, M. King, P. Minor, C. Bailey, M. R. Scalia, L. J. Akinbami, C. f. D. C. a. P. (CDC), National surveillance for asthma--United States, 1980-2004. *MMWR Surveill Summ* **56**, 1-54 (2007).

10. C. Flores, S. F. Ma, M. Pino-Yanes, M. S. Wade, L. Perez-Mendez, R. A. Kittles, D. Wang, S. Papaiahgari, J. G. Ford, R. Kumar, J. G. Garcia, African ancestry is associated with asthma risk in African Americans. *PLoS One* **7**, e26807 (2012).

11. S. Choudhry, E. G. Burchard, L. N. Borrell, H. Tang, I. Gomez, M. Naqvi, S. Nazario, A. Torres, J. Casal, J. C. Martinez-Cruzado, E. Ziv, P. C. Avila, W. Rodriguez-Cintron, N. J.

Risch, Ancestry-environment interactions and asthma risk among Puerto Ricans. *Am J Respir Crit Care Med* **174**, 1088-1093 (2006).

12. C. Vergara, L. Caraballo, D. Mercado, S. Jimenez, W. Rojas, N. Rafaels, T. Hand, M. Campbell, Y. J. Tsai, L. Gao, C. Duque, S. Lopez, G. Bedoya, A. Ruiz-Linares, K. C. Barnes, African ancestry is associated with risk of asthma and high total serum IgE in a population from the Caribbean Coast of Colombia. *Hum Genet* **125**, 565-579 (2009).

13. C. Vergara, T. Murray, N. Rafaels, R. Lewis, M. Campbell, C. Foster, L. Gao, M. Faruque, R. R. Oliveira, E. Carvalho, M. I. Araujo, A. A. Cruz, H. Watson, D. Mercado, J. Knight-Madden, I. Ruczinski, G. Dunston, J. Ford, L. Caraballo, T. H. Beaty, R. A. Mathias, K. C. Barnes, African ancestry is a risk factor for asthma and high total IgE levels in African admixed populations. *Genet Epidemiol* **37**, 393-401 (2013).

14. J. E. Gern, R. F. Lemanske, W. W. Busse, Early life origins of asthma. *J Clin Invest* **104**, 837-843 (1999).

15. K. G. Tournoy, S. Provoost, C. Van Hove, G. Joos, The role of immune tolerance in asthma pathogenesis. *Curr Allergy Asthma Rep* **6**, 437-443 (2006).

16. L. Chatzi, M. Torrent, I. Romieu, R. Garcia-Esteban, C. Ferrer, J. Vioque, M. Kogevinas, J. Sunyer, Mediterranean diet in pregnancy is protective for wheeze and atopy in childhood. *Thorax* **63**, 507-513 (2008).

17. J. A. Dunstan, T. A. Mori, A. Barden, L. J. Beilin, A. L. Taylor, P. G. Holt, S. L. Prescott, Fish oil supplementation in pregnancy modifies neonatal allergen-specific immune responses and clinical outcomes in infants at high risk of atopy: a randomized, controlled trial. *J Allergy Clin Immunol* **112**, 1178-1184 (2003).

18. H. Bisgaard, J. Stokholm, B. L. Chawes, N. H. Vissing, E. Bjarnadóttir, A. M. Schoos, H. M. Wolsk, T. M. Pedersen, R. K. Vinding, S. Thorsteinsdóttir, N. V. Følsgaard, N. R. Fink, J. Thorsen, A. G. Pedersen, J. Waage, M. A. Rasmussen, K. D. Stark, S. F. Olsen, K. Bønnelykke, Fish Oil-Derived Fatty Acids in Pregnancy and Wheeze and Asthma in Offspring. *N Engl J Med* **375**, 2530-2539 (2016).

19. C. T. McEvoy, E. R. Spindel, Pulmonary Effects of Maternal Smoking on the Fetus and Child: Effects on Lung Development, Respiratory Morbidities, and Life Long Lung Health. *Paediatr Respir Rev* **21**, 27-33 (2017).

20. Å. Neuman, C. Hohmann, N. Orsini, G. Pershagen, E. Eller, H. F. Kjaer, U. Gehring, R. Granell, J. Henderson, J. Heinrich, S. Lau, M. Nieuwenhuijsen, J. Sunyer, C. Tischer, M. Torrent, U. Wahn, A. H. Wijga, M. Wickman, T. Keil, A. Bergström, E. Consortium, Maternal smoking in pregnancy and asthma in preschool children: a pooled analysis of eight birth cohorts. *Am J Respir Crit Care Med* **186**, 1037-1043 (2012).

21. M. Cheraghi, S. Salvi, Environmental tobacco smoke (ETS) and respiratory health in children. *Eur J Pediatr* **168**, 897-905 (2009).

22. H. M. Salihu, M. H. Aliyu, B. J. Pierre-Louis, G. R. Alexander, Levels of excess infant deaths attributable to maternal smoking during pregnancy in the United States. *Matern Child Health J* **7**, 219-227 (2003).

23. J. Stocks, A. Hislop, S. Sonnappa, Early lung development: lifelong effect on respiratory health and disease. *Lancet Respir Med* **1**, 728-742 (2013).

24. X. F. Xu, Y. J. Li, Y. J. Sheng, J. L. Liu, L. F. Tang, Z. M. Chen, Effect of low birth weight on childhood asthma: a meta-analysis. *BMC Pediatr* **14**, 275 (2014).

25. J. Zhang, C. Ma, A. Yang, R. Zhang, J. Gong, F. Mo, Is preterm birth associated with asthma among children from birth to 17 years old? -A study based on 2011-2012 US National Survey of Children's Health. *Ital J Pediatr* **44**, 151 (2018).

26. C. H. ANDREWES, D. M. CHAPRONIERE, A. E. GOMPELS, H. G. PEREIRA, A. T. RODEN, Propagation of common-cold virus in tissue cultures. *Lancet* **265**, 546-547 (1953).

27. S. L. Johnston, P. K. Pattemore, G. Sanderson, S. Smith, F. Lampe, L. Josephs, P. Symington, S. O'Toole, S. H. Myint, D. A. Tyrrell, Community study of role of viral infections in exacerbations of asthma in 9-11 year old children. *BMJ* **310**, 1225-1229 (1995).

28. J. E. Gern, The ABCs of rhinoviruses, wheezing, and asthma. *J Virol* **84**, 7418-7426 (2010).

29. D. J. Jackson, R. E. Gangnon, M. D. Evans, K. A. Roberg, E. L. Anderson, T. E. Pappas, M. C. Printz, W. M. Lee, P. A. Shult, E. Reisdorf, K. T. Carlson-Dakes, L. P. Salazar, D. F. DaSilva, C. J. Tisler, J. E. Gern, R. F. Lemanske, Jr., Wheezing rhinovirus illnesses in early life predict asthma development in high-risk children. *Am J Respir Crit Care Med* **178**, 667-672 (2008).

30. M. M. Stein, E. E. Thompson, N. Schoettler, B. A. Helling, K. M. Magnaye, C. Stanhope, C. Igartua, A. Morin, C. Washington, D. Nicolae, K. Bønnelykke, C. Ober, A decade of research on the 17q12-21 asthma locus: Piecing together the puzzle. *J Allergy Clin Immunol* **142**, 749-764.e743 (2018).

31. M. C. Arrieta, L. T. Stiemsma, P. A. Dimitriu, L. Thorson, S. Russell, S. Yurist-Doutsch, B. Kuzeljevic, M. J. Gold, H. M. Britton, D. L. Lefebvre, P. Subbarao, P. Mandhane, A. Becker, K. M. McNagny, M. R. Sears, T. Kollmann, W. W. Mohn, S. E. Turvey, B. B. Finlay, C. S. Investigators, Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci Transl Med* **7**, 307ra152 (2015).

32. K. E. Fujimura, A. R. Sitarik, S. Havstad, D. L. Lin, S. Levan, D. Fadrosh, A. R. Panzer, B. LaMere, E. Rackaityte, N. W. Lukacs, G. Wegienka, H. A. Boushey, D. R. Ownby, E. M. Zoratti, A. M. Levin, C. C. Johnson, S. V. Lynch, Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat Med* **22**, 1187-1191 (2016).

33. J. Durack, N. E. Kimes, D. L. Lin, M. Rauch, M. McKean, K. McCauley, A. R. Panzer, J. S. Mar, M. D. Cabana, S. V. Lynch, Delayed gut microbiota development in high-risk for asthma infants is temporarily modifiable by Lactobacillus supplementation. *Nat Commun* **9**, 707 (2018).

34. J. Thorsen, M. A. Rasmussen, J. Waage, M. Mortensen, A. Brejnrod, K. Bønnelykke, B. L. Chawes, S. Brix, S. J. Sørensen, J. Stokholm, H. Bisgaard, Infant airway microbiota and topical immune perturbations in the origins of childhood asthma. *Nat Commun* **10**, 5001 (2019).

35. H. H. F. Tang, A. Lang, S. M. Teo, L. M. Judd, R. Gangnon, M. D. Evans, K. E. Lee, R. Vrtis, P. G. Holt, R. F. Lemanske, D. J. Jackson, K. E. Holt, M. Inouye, J. E. Gern, Developmental patterns in the nasopharyngeal microbiome during infancy are associated with asthma risk. *J Allergy Clin Immunol*, (2020).

36. S. R. Levan, K. A. Stamnes, D. L. Lin, A. R. Panzer, E. Fukui, K. McCauley, K. E. Fujimura, M. McKean, D. R. Ownby, E. M. Zoratti, H. A. Boushey, M. D. Cabana, C. C. Johnson, S. V. Lynch, Elevated faecal 12,13-diHOME concentration in neonates at high risk for asthma is produced by gut bacteria and impedes immune tolerance. *Nat Microbiol* **4**, 1851-1861 (2019).

37. D. J. Barker, C. Osmond, Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. *Lancet* **1**, 1077-1081 (1986).

38. S. O. Shaheen, I. M. Adcock, The developmental origins of asthma: does epigenetics hold the key? *Am J Respir Crit Care Med* **180**, 690-691 (2009).

39. J. W. Hollingsworth, S. Maruoka, K. Boon, S. Garantziotis, Z. Li, J. Tomfohr, N. Bailey, E. N. Potts, G. Whitehead, D. M. Brass, D. A. Schwartz, In utero supplementation with methyl donors enhances allergic airway disease in mice. *J Clin Invest* **126**, 2012 (2016).

40. C. V. Breton, H. M. Byun, M. Wenten, F. Pan, A. Yang, F. D. Gilliland, Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. *Am J Respir Crit Care Med* **180**, 462-467 (2009).

41. Y. A. Bobetsis, S. P. Barros, D. M. Lin, J. R. Weidman, D. C. Dolinoy, R. L. Jirtle, K. A. Boggess, J. D. Beck, S. Offenbacher, Bacterial infection promotes DNA hypermethylation. *J Dent Res* **86**, 169-174 (2007).

42. C. Ober, T. C. Yao, The genetics of asthma and allergic disease: a 21st century perspective. *Immunol Rev* **242**, 10-30 (2011).

43. M. Pividori, N. Schoettler, D. L. Nicolae, C. Ober, H. K. Im, Shared and distinct genetic risk factors for childhood-onset and adult-onset asthma: genome-wide and transcriptome-wide studies. *Lancet Respir Med* **7**, 509-522 (2019).

44. M. A. R. Ferreira, R. Mathur, J. M. Vonk, A. Szwajda, B. Brumpton, R. Granell, B. K. Brew, V. Ullemar, Y. Lu, Y. Jiang, P. K. E. Magnusson, R. Karlsson, D. A. Hinds, L. Paternoster, G. H. Koppelman, C. Almqvist, a. R. Team, e. Consortium, B. Consortium, Genetic Architectures of Childhood- and Adult-Onset Asthma Are Partly Distinct. *Am J Hum Genet* **104**, 665-684 (2019).

45. T. A. Olafsdottir, F. Theodors, K. Bjarnadottir, U. S. Bjornsdottir, A. B. Agustsdottir, O. A. Stefansson, E. V. Ivarsdottir, J. K. Sigurdsson, S. Benonisdottir, G. I. Eyjolfsson, D. Gislason, T. Gislason, S. Guðmundsdóttir, A. Gylfason, B. V. Halldorsson, G. H. Halldorsson, T. Juliusdottir, A. M. Kristinsdottir, D. Ludviksdottir, B. R. Ludviksson, G. Masson, K. Norland, P. T. Onundarson, I. Olafsson, O. Sigurdardottir, L. Stefansdottir, G. Sveinbjornsson, V. Tragante, D. F. Gudbjartsson, G. Thorleifsson, P. Sulem, U. Thorsteinsdottir, G. L. Norddahl, I. Jonsdottir, K. Stefansson, Eighty-eight variants highlight the role of T cell regulation and airway remodeling in asthma pathogenesis. *Nat Commun* **11**, 393 (2020).

46. F. Demenais, P. Margaritte-Jeannin, K. C. Barnes, W. O. C. Cookson, J. Altmüller, W. Ang, R. G. Barr, T. H. Beaty, A. B. Becker, J. Beilby, H. Bisgaard, U. S. Bjornsdottir, E. Bleecker, K. Bønnelykke, D. I. Boomsma, E. Bouzigon, C. E. Brightling, M. Brossard, G. G. Brusselle, E. Burchard, K. M. Burkart, A. Bush, M. Chan-Yeung, K. F. Chung, A. Couto Alves, J. A. Curtin, A. Custovic, D. Daley, J. C. de Jongste, B. E. Del-Rio-Navarro, K. M. Donohue, L. Duijts, C. Eng, J. G. Eriksson, M. Farrall, Y. Fedorova, B. Feenstra, M. A. Ferreira, M. B. Freidin, Z. Gajdos, J. Gauderman, U. Gehring, F. Geller, J. Genuneit, S. A. Gharib, F. Gilliland, R. Granell, P. E. Graves, D. F. Gudbjartsson, T. Haahtela, S. R. Heckbert, D. Heederik, J. Heinrich, M. Heliövaara, J. Henderson, B. E. Himes, H. Hirose, J. N. Hirschhorn, A. Hofman, P. Holt, J. Hottenga, T. J. Hudson, J. Hui, M. Imboden, V. Ivanov, V. W. V. Jaddoe, A. James, C. Janson, M. R. Jarvelin, D. Jarvis, G. Jones, I. Jonsdottir, P. Jousilahti, M. Kabesch, M. Kähönen, D. B. Kantor, A. S. Karunas, E. Khusnutdinova, G. H. Koppelman, A. L. Kozyrskyj, E. Kreiner, M. Kubo, R. Kumar, A. Kumar, M. Kuokkanen, L. Lahousse, T. Laitinen, C. Laprise, M. Lathrop, S. Lau, Y. A. Lee, T. Lehtimäki, S. Letort, A. M. Levin, G. Li, L. Liang, L. R. Loehr, S. J. London, D. W. Loth, A. Manichaikul, I. Marenholz, F. J. Martinez, M. C. Matheson, R. A. Mathias, K. Matsumoto, H. Mbarek, W. L. McArdle, M. Melbye, E. Melén, D. Meyers, S. Michel, H. Mohamdi, A. W. Musk, R. A. Myers, M. A. E. Nieuwenhuis, E. Noguchi, G. T. O'Connor, L. M. Ogorodova, C. D. Palmer, A. Palotie, J. E. Park, C. E. Pennell, G. Pershagen, A. Polonikov, D. S. Postma, N. Probst-Hensch, V. P. Puzyrev, B. A. Raby, O. T. Raitakari, A. Ramasamy, S. S. Rich, C. F. Robertson, I. Romieu, M. T. Salam, V. Salomaa, V. Schlünssen, R. Scott, P. A. Selivanova, T. Sigsgaard, A. Simpson, V. Siroux, L. J. Smith, M. Solodilova, M. Standl, K. Stefansson,

D. P. Strachan, B. H. Stricker, A. Takahashi, P. J. Thompson, G. Thorleifsson, U. Thorsteinsdottir, C. M. T. Tiesler, D. G. Torgerson, T. Tsunoda, A. G. Uitterlinden, R. J. P. van der Valk, A. Vaysse, S. Vedantam, A. von Berg, E. von Mutius, J. M. Vonk, J. Waage, N. J. Wareham, S. T. Weiss, W. B. White, M. Wickman, E. Widén, G. Willemsen, L. K. Williams, I. M. Wouters, J. J. Yang, J. H. Zhao, M. F. Moffatt, C. Ober, D. L. Nicolae, A. A. G. C. A. collaborators, Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. *Nat Genet* **50**, 42-53 (2018).

47. M. Daya, N. Rafaels, T. M. Brunetti, S. Chavan, A. M. Levin, A. Shetty, C. R. Gignoux, M. P. Boorgula, G. Wojcik, M. Campbell, C. Vergara, D. G. Torgerson, V. E. Ortega, A. Doumatey, H. R. Johnston, N. Acevedo, M. I. Araujo, P. C. Avila, G. Belbin, E. Bleecker, C. Bustamante, L. Caraballo, A. Cruz, G. M. Dunston, C. Eng, M. U. Faruque, T. S. Ferguson, C. Figueiredo, J. G. Ford, W. Gan, P. A. Gourraud, N. N. Hansel, R. D. Hernandez, E. F. Herrera-Paz, S. Jiménez, E. E. Kenny, J. Knight-Madden, R. Kumar, L. A. Lange, E. M. Lange, A. Lizee, P. Maul, T. Maul, A. Mayorga, D. Meyers, D. L. Nicolae, T. D. O'Connor, R. R. Oliveira, C. O. Olopade, O. Olopade, Z. S. Qin, C. Rotimi, N. Vince, H. Watson, R. J. Wilks, J. G. Wilson, S. Salzberg, C. Ober, E. G. Burchard, L. K. Williams, T. H. Beaty, M. A. Taub, I. Ruczinski, R. A. Mathias, K. C. Barnes, CAAPA, Association study in African-admixed populations across the Americas recapitulates asthma risk loci in non-African populations. *Nat Commun* **10**, 880 (2019).

48. C. Bycroft, C. Freeman, D. Petkova, G. Band, L. T. Elliott, K. Sharp, A. Motyer, D. Vukcevic, O. Delaneau, J. O'Connell, A. Cortes, S. Welsh, A. Young, M. Effingham, G. McVean, S. Leslie, N. Allen, P. Donnelly, J. Marchini, The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203-209 (2018).

49. R. H. Lim, L. Kobzik, M. Dahl, Risk for asthma in offspring of asthmatic mothers versus fathers: a meta-analysis. *PLoS One* **5**, e10134 (2010).

50. M. E. Soto-Quiros, E. K. Silverman, L. A. Hanson, S. T. Weiss, J. C. Celedón, Maternal history, sensitization to allergens, and current wheezing, rhinitis, and eczema among children in Costa Rica. *Pediatr Pulmonol* **33**, 237-243 (2002).

51. R. Enriquez, M. R. Griffin, K. N. Carroll, P. Wu, W. O. Cooper, T. Gebretsadik, W. D. Dupont, E. F. Mitchel, T. V. Hartert, Effect of maternal asthma and asthma control on pregnancy and perinatal outcomes. *J Allergy Clin Immunol* **120**, 625-630 (2007).

52. T. K. Sorensen, J. C. Dempsey, R. Xiao, I. O. Frederick, D. A. Luthy, M. A. Williams, Maternal asthma and risk of preterm delivery. *Ann Epidemiol* **13**, 267-272 (2003).

53. D. Nicolae, N. J. Cox, L. A. Lester, D. Schneider, Z. Tan, C. Billstrand, S. Kuldanek, J. Donfack, P. Kogut, N. M. Patel, J. Goodenbour, T. Howard, R. Wolf, G. H. Koppelman, S. R. White, R. Parry, D. S. Postma, D. Meyers, E. R. Bleecker, J. S. Hunt, J. Solway, C. Ober,

Fine mapping and positional candidate studies identify HLA-G as an asthma susceptibility gene on chromosome 6p21. *Am J Hum Genet* **76**, 349-357 (2005).

54. Z. Tan, G. Randall, J. Fan, B. Camoretti-Mercado, R. Brockman-Schneider, L. Pan, J. Solway, J. E. Gern, R. F. Lemanske, D. Nicolae, C. Ober, Allele-specific targeting of microRNAs to HLA-G and risk of asthma. *Am J Hum Genet* **81**, 829-834 (2007).

55. J. Nicodemus-Johnson, B. Laxman, R. K. Stern, J. Sudi, C. N. Tierney, L. Norwick, D. K. Hogarth, J. F. McConville, E. T. Naureckas, A. I. Sperling, J. Solway, J. A. Krishnan, D. L. Nicolae, S. R. White, C. Ober, Maternal asthma and microRNA regulation of soluble HLA-G in the airway. *J Allergy Clin Immunol* **131**, 1496-1503 (2013).

56. L. P. Gunawardhana, K. J. Baines, J. Mattes, V. E. Murphy, J. L. Simpson, P. G. Gibson, Differential DNA methylation profiles of infants exposed to maternal asthma during pregnancy. *Pediatr Pulmonol* **49**, 852-862 (2014).

57. L. Kashima, M. Toyota, H. Mita, H. Suzuki, M. Idogawa, K. Ogi, Y. Sasaki, T. Tokino, CHFR, a potential tumor suppressor, downregulates interleukin-8 through the inhibition of NF-kappaB. *Oncogene* **28**, 2643-2653 (2009).

58. E. Morales, M. Bustamante, N. Vilahur, G. Escaramis, M. Montfort, R. de Cid, R. Garcia-Esteban, M. Torrent, X. Estivill, J. O. Grimalt, J. Sunyer, DNA hypomethylation at ALOX12 is associated with persistent wheezing in childhood. *Am J Respir Crit Care Med* **185**, 937-943 (2012).

59. A. DeVries, G. Wlasiuk, S. J. Miller, A. Bosco, D. A. Stern, I. C. Lohman, J. Rothers, A. C. Jones, J. Nicodemus-Johnson, M. M. Vasquez, J. A. Curtin, A. Simpson, A. Custovic, D. J. Jackson, J. E. Gern, R. F. Lemanske, S. Guerra, A. L. Wright, C. Ober, M. Halonen, D. Vercelli, Epigenome-wide analysis links SMAD3 methylation at birth to asthma in children of asthmatic mothers. *J Allergy Clin Immunol* **140**, 534-542 (2017).

60. N. Schoettler, C. Ober, Genetic architecture of moderate-to-severe asthma mirrors that of mild asthma. *J Allergy Clin Immunol* **144**, 1521-1523 (2019).

61. O. D. Solberg, E. J. Ostrin, M. I. Love, J. C. Peng, N. R. Bhakta, L. Hou, C. Nguyen, M. Solon, A. J. Barczak, L. T. Zlock, D. P. Blagev, W. E. Finkbeiner, K. M. Ansel, J. R. Arron, D. J. Erle, P. G. Woodruff, Airway epithelial miRNA expression is altered in asthma. *Am J Respir Crit Care Med* **186**, 965-974 (2012).

62. L. J. Simpson, S. Patel, N. R. Bhakta, D. F. Choy, H. D. Brightbill, X. Ren, Y. Wang, H. H. Pua, D. Baumjohann, M. M. Montoya, M. Panduro, K. A. Remedios, X. Huang, J. V. Fahy, J. R. Arron, P. G. Woodruff, K. M. Ansel, A microRNA upregulated in asthma airway T cells promotes TH2 cytokine production. *Nat Immunol* **15**, 1162-1170 (2014).

63. J. H. Bahn, J. Ahn, X. Lin, Q. Zhang, J. H. Lee, M. Civelek, X. Xiao, Genomic analysis of ADAR1 binding and its involvement in multiple RNA processing pathways. *Nat Commun* **6**, 6355 (2015).

64. D. Paul, A. N. Sinha, A. Ray, M. Lal, S. Nayak, A. Sharma, B. Mehani, D. Mukherjee, S. V. Laddha, A. Suri, C. Sarkar, A. Mukhopadhyay, A-to-I editing in human miRNAs is enriched in seed sequence, influenced by sequence contexts and significantly hypoedited in glioblastoma multiforme. *Sci Rep* **7**, 2466 (2017).

65. J. B. Patterson, D. C. Thomis, S. L. Hans, C. E. Samuel, Mechanism of interferon action: double-stranded RNA-specific adenosine deaminase from human cells is inducible by alpha and gamma interferons. *Virology* **210**, 508-511 (1995).

66. P. J. Teoh, O. An, T. H. Chung, J. Y. Chooi, S. H. M. Toh, S. Fan, W. Wang, B. T. H. Koh, M. J. Fullwood, M. G. Ooi, S. de Mel, C. Y. Soekojo, L. Chen, S. B. Ng, H. Yang, W. J. Chng, Aberrant hyperediting of the myeloma transcriptome by ADAR1 confers oncogenicity and is a marker of poor prognosis. *Blood* **132**, 1304-1317 (2018).

67. Y. Wang, X. Xu, S. Yu, K. J. Jeong, Z. Zhou, L. Han, Y. H. Tsang, J. Li, H. Chen, L. S. Mangala, Y. Yuan, A. K. Eterovic, Y. Lu, A. K. Sood, K. L. Scott, G. B. Mills, H. Liang, Systematic characterization of A-to-I RNA editing hotspots in microRNAs across human cancers. *Genome Res* **27**, 1112-1125 (2017).

68. J. H. Yang, X. Luo, Y. Nie, Y. Su, Q. Zhao, K. Kabir, D. Zhang, R. Rabinovici, Widespread inosine-containing mRNA in lymphocytes regulated by ADAR1 in response to inflammation. *Immunology* **109**, 15-23 (2003).

69. V. Agarwal, G. W. Bell, J. W. Nam, D. P. Bartel, Predicting effective microRNA target sites in mammalian mRNAs. *Elife* **4**, (2015).

70. J. Chen, E. E. Bardes, B. J. Aronow, A. G. Jegga, ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* **37**, W305-311 (2009).

71. C. Lee, T. B. Kolesnik, I. Caminschi, A. Chakravorty, W. Carter, W. S. Alexander, J. Jones, G. P. Anderson, S. E. Nicholson, Suppressor of cytokine signalling 1 (SOCS1) is a physiological regulator of the asthma response. *Clin Exp Allergy* **39**, 897-907 (2009).

72. H. Dickensheets, N. Vazquez, F. Sheikh, S. Gingras, P. J. Murray, J. J. Ryan, R. P. Donnelly, Suppressor of cytokine signaling-1 is an IL-4-inducible gene in macrophages and feedback inhibits IL-4 signaling. *Genes Immun* **8**, 21-27 (2007).

73. J. Pothlichet, M. Chignard, M. Si-Tahar, Cutting edge: innate immune response triggered by influenza A virus is negatively regulated by SOCS1 and SOCS3 through a RIG-I/IFNAR1-dependent pathway. *J Immunol* **180**, 2034-2038 (2008).

74. R. A. Piganis, N. A. De Weerd, J. A. Gould, C. W. Schindler, A. Mansell, S. E. Nicholson, P. J. Hertzog, Suppressor of cytokine signaling (SOCS) 1 inhibits type I interferon (IFN) signaling via the interferon alpha receptor (IFNAR1)-associated tyrosine kinase Tyk2. *J Biol Chem* **286**, 33811-33818 (2011).

75. P. G. Woodruff, B. Modrek, D. F. Choy, G. Jia, A. R. Abbas, A. Ellwanger, L. L. Koth, J. R. Arron, J. V. Fahy, T-helper type 2-driven inflammation defines major subphenotypes of asthma. *Am J Respir Crit Care Med* **180**, 388-395 (2009).

76. N. A. E. a. P. Program, Expert Panel Report 3 (EPR-3): Guidelines for the Diagnosis and Management of Asthma-Summary Report 2007. *J Allergy Clin Immunol* **120**, S94-138 (2007).

77. N. Hernandez-Pacheco, N. Farzan, B. Francis, L. Karimi, K. Repnik, S. J. Vijverberg, P. Soares, M. Schieck, M. Gorenjak, E. Forno, C. Eng, S. S. Oh, L. Pérez-Méndez, V. Berce, R. Tavendale, L. A. Samedy, S. Hunstman, D. Hu, K. Meade, H. J. Farber, P. C. Avila, D. Serebrisky, S. M. Thyne, E. Brigino-Buenaventura, W. Rodriguez-Cintron, S. Sen, R. Kumar, M. Lenoir, J. R. Rodriguez-Santana, J. C. Celedón, S. Mukhopadhyay, U. Potočnik, M. Pirmohamed, K. M. Verhamme, M. Kabesch, C. N. A. Palmer, D. B. Hawcutt, C. Flores, A. H. Maitland-van der Zee, E. G. Burchard, M. Pino-Yanes, Genome-wide association study of inhaled corticosteroid response in admixed children with asthma. *Clin Exp Allergy* **49**, 789-798 (2019).

78. A. C. Y. Mak, M. J. White, W. L. Eckalbar, Z. A. Szpiech, S. S. Oh, M. Pino-Yanes, D. Hu, P. Goddard, S. Huntsman, J. Galanter, A. C. Wu, B. E. Himes, S. Germer, J. M. Vogel, K. L. Bunting, C. Eng, S. Salazar, K. L. Keys, J. Liberto, T. J. Nuckton, T. A. Nguyen, D. G. Torgerson, P. Y. Kwok, A. M. Levin, J. C. Celedón, E. Forno, H. Hakonarson, P. M. Sleiman, A. Dahlin, K. G. Tantisira, S. T. Weiss, D. Serebrisky, E. Brigino-Buenaventura, H. J. Farber, K. Meade, M. A. Lenoir, P. C. Avila, S. Sen, S. M. Thyne, W. Rodriguez-Cintron, C. A. Winkler, A. Moreno-Estrada, K. Sandoval, J. R. Rodriguez-Santana, R. Kumar, L. K. Williams, N. Ahituv, E. Ziv, M. A. Seibold, R. B. Darnell, N. Zaitlen, R. D. Hernandez, E. G. Burchard, N. T.-O. f. P. M. T. Consortium, Whole-Genome Sequencing of Pharmacogenetic Drug Response in Racially Diverse Children with Asthma. *Am J Respir Crit Care Med* **197**, 1552-1564 (2018).

79. N. Shrine, M. A. Portelli, C. John, M. Soler Artigas, N. Bennett, R. Hall, J. Lewis, A. P. Henry, C. K. Billington, A. Ahmad, R. J. Packer, D. Shaw, Z. E. K. Pogson, A. Fogarty, T. M. McKeever, A. Singapuri, L. G. Heaney, A. H. Mansur, R. Chaudhuri, N. C. Thomson, J. W. Holloway, G. A. Lockett, P. H. Howarth, R. Djukanovic, J. Hankinson, R. Niven, A. Simpson, K. F. Chung, P. J. Sterk, J. D. Blakey, I. M. Adcock, S. Hu, Y. Guo, M. Obeidat, D. D. Sin, M. van den Berge, D. C. Nickle, Y. Bossé, M. D. Tobin, I. P. Hall, C. E.

Brightling, L. V. Wain, I. Sayers, Moderate-to-severe asthma in individuals of European ancestry: a genome-wide association study. *Lancet Respir Med* **7**, 20-34 (2019).

80. V. Cesarini, D. A. Silvestris, V. Tassinari, S. Tomaselli, S. Alon, E. Eisenberg, F. Locatelli, A. Gallo, ADAR2/miR-589-3p axis controls glioblastoma cell migration/invasion. *Nucleic Acids Res* **46**, 2045-2059 (2018).

81. K. Maemura, K. Watanabe, T. Ando, N. Hiyama, T. Sakatani, Y. Amano, H. Kage, J. Nakajima, Y. Yatomi, T. Nagase, D. Takai, Altered editing level of microRNAs is a potential biomarker in lung adenocarcinoma. *Cancer Sci* **109**, 3326-3335 (2018).

82. V. Gielen, A. Sykes, J. Zhu, B. Chan, J. Macintyre, N. Regamey, E. Kieninger, A. Gupta, A. Shoemark, C. Bossley, J. Davies, S. Saglani, P. Walker, S. E. Nicholson, A. H. Dalpke, O. M. Kon, A. Bush, S. L. Johnston, M. R. Edwards, Increased nuclear suppressor of cytokine signaling 1 in asthmatic bronchial epithelium suppresses rhinovirus induction of innate interferons. *J Allergy Clin Immunol* **136**, 177-188.e111 (2015).

83. M. Harada, K. Nakashima, T. Hirota, M. Shimizu, S. Doi, K. Fujita, T. Shirakawa, T. Enomoto, M. Yoshikawa, H. Moriyama, K. Matsumoto, H. Saito, Y. Suzuki, Y. Nakamura, M. Tamari, Functional polymorphism in the suppressor of cytokine signaling 1 gene associated with adult asthma. *Am J Respir Cell Mol Biol* **36**, 491-496 (2007).

84. Y. H. Tsai, J. S. Parker, I. V. Yang, S. N. P. Kelada, Meta-analysis of airway epithelium gene expression in asthma. *Eur Respir J* **51**, (2018).

85. E. Doran, D. F. Choy, A. Shikotra, C. A. Butler, D. M. O'Rourke, J. A. Johnston, A. Kissenpfennig, P. Bradding, J. R. Arron, L. G. Heaney, Reduced epithelial suppressor of cytokine signalling 1 in severe eosinophilic asthma. *Eur Respir J* **48**, 715-725 (2016).

86. J. V. Fahy, Type 2 inflammation in asthma--present in most, absent in many. *Nat Rev Immunol* **15**, 57-65 (2015).

87. G. Jun, M. Flickinger, K. N. Hetrick, J. M. Romm, K. F. Doheny, G. R. Abecasis, M. Boehnke, H. M. Kang, Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. *Am J Hum Genet* **91**, 839-848 (2012).

88. Y. Lu, A. S. Baras, M. K. Halushka, miRge 2.0 for comprehensive analysis of microRNA sequencing data. *BMC Bioinformatics* **19**, 275 (2018).

89. A. Dobin, T. R. Gingeras, Mapping RNA-seq Reads with STAR. *Curr Protoc Bioinformatics* **51**, 11.14.11-11.14.19 (2015).

90. C. W. Law, Y. Chen, W. Shi, G. K. Smyth, voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29 (2014).

91. J. T. Leek, J. D. Storey, Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet* **3**, 1724-1735 (2007).

92. S. Alon, E. Mor, F. Vigneault, G. M. Church, F. Locatelli, F. Galeano, A. Gallo, N. Shomron, E. Eisenberg, Systematic identification of edited microRNAs in the human brain. *Genome Res* **22**, 1533-1540 (2012).

93. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, G. K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).

94. T. Nurmagambetov, R. Kuwahara, P. Garbe, The Economic Burden of Asthma in the United States, 2008-2013. *Ann Am Thorac Soc* **15**, 348-356 (2018).

95. A. DeVries, D. Vercelli, The neonatal methylome as a gatekeeper in the trajectory to childhood asthma. *Epigenomics* **9**, 585-593 (2017).

96. J. Rothers, D. A. Stern, I. C. Lohman, A. Spangenberg, A. L. Wright, A. DeVries, D. Vercelli, M. Halonen, Maternal Cytokine Profiles during Pregnancy Predict Asthma in Children of Mothers without Asthma. *Am J Respir Cell Mol Biol* **59**, 592-600 (2018).

97. N. Schoettler, E. Rodríguez, S. Weidinger, C. Ober, Advances in asthma and allergic disease genetics: Is bigger always better? *J Allergy Clin Immunol* **144**, 1495-1506 (2019).

98. H. J. Gould, B. J. Sutton, IgE in allergy and asthma today. *Nat Rev Immunol* **8**, 205-217 (2008).

99. P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).

100. C. X. George, C. E. Samuel, Human RNA-specific adenosine deaminase ADAR1 transcripts possess alternative exon 1 structures that initiate from different promoters, one constitutively active and the other interferon inducible. *Proc Natl Acad Sci U S A* **96**, 4621-4626 (1999).

101. M. R. Edwards, K. Strong, A. Cameron, R. P. Walton, D. J. Jackson, S. L. Johnston, Viral infections in allergy and immunology: How allergic inflammation influences viral infections and illness. *J Allergy Clin Immunol* **140**, 909-920 (2017).

102. S. J. Fairweather-Tait, Y. Bao, M. R. Broadley, R. Collings, D. Ford, J. E. Hesketh, R. Hurst, Selenium in human health and disease. *Antioxid Redox Signal* **14**, 1337-1383 (2011).

103. J. C. Avery, P. R. Hoffmann, Selenium, Selenoproteins, and Immunity. *Nutrients* **10**, (2018).

104. P. R. Hoffmann, Selenium and asthma: a complex relationship. *Allergy* **63**, 854-856 (2008).

105. J. Wei, R. J. Kishton, M. Angel, C. S. Conn, N. Dalla-Venezia, V. Marcel, A. Vincent, F. Catez, S. Ferré, L. Ayadi, V. Marchand, D. Dersh, J. S. Gibbs, I. P. Ivanov, N. Fridlyand, Y. Couté, J. J. Diaz, S. B. Qian, L. M. Staudt, N. P. Restifo, J. W. Yewdell, Ribosomal Proteins Regulate MHC Class I Peptide Generation for Immunosurveillance. *Mol Cell* **73**, 1162-1173.e1165 (2019).

106. A. Poole, C. Urbanek, C. Eng, J. Schageman, S. Jacobson, B. P. O'Connor, J. M. Galanter, C. R. Gignoux, L. A. Roth, R. Kumar, S. Lutz, A. H. Liu, T. E. Fingerlin, R. A. Setterquist, E. G. Burchard, J. Rodriguez-Santana, M. A. Seibold, Dissecting childhood asthma with nasal transcriptomics distinguishes subphenotypes of disease. *J Allergy Clin Immunol* **133**, 670-678.e612 (2014).

107. H. Persson, A. T. Kwon, J. A. Ramilowski, G. Silberberg, C. Söderhäll, C. Orsmark-Pietras, B. Nordlund, J. R. Konradsen, M. J. de Hoon, E. Melén, Y. Hayashizaki, G. Hedlin, J. Kere, C. O. Daub, Transcriptome analysis of controlled and therapy-resistant childhood asthma reveals distinct gene expression profiles. *J Allergy Clin Immunol* **136**, 638-648 (2015).

108. Z. Jevnikar, J. Östling, E. Ax, J. Calvén, K. Thörn, E. Israelsson, L. Öberg, A. Singhania, L. C. K. Lau, S. J. Wilson, J. A. Ward, A. Chauhan, A. R. Sousa, B. De Meulder, M. J. Loza, F. Baribaud, P. J. Sterk, K. F. Chung, K. Sun, Y. Guo, I. M. Adcock, D. Payne, B. Dahlen, P. Chanez, D. E. Shaw, N. Krug, J. M. Hohlfeld, T. Sandström, R. Djukanovic, A. James, T. S. C. Hinks, P. H. Howarth, O. Vaarala, M. van Geest, H. Olsson, U. B. i. P. o. R. D. O. s. group, Epithelial IL-6 trans-signaling defines a new asthma phenotype with increased airway inflammation. *J Allergy Clin Immunol* **143**, 577-590 (2019).

109. F. D. Gilliland, Y. F. Li, J. M. Peters, Effects of maternal smoking during pregnancy and environmental tobacco smoke on asthma and wheezing in children. *Am J Respir Crit Care Med* **163**, 429-436 (2001).

110. Z. Hehua, C. Qing, G. Shanyan, W. Qijun, Z. Yuhong, The impact of prenatal exposure to air pollution on childhood wheezing and asthma: A systematic review. *Environ Res* **159**, 519-530 (2017).

111. J. Riedler, C. Braun-Fahrländer, W. Eder, M. Schreuer, M. Waser, S. Maisch, D. Carr, R. Schierl, D. Nowak, E. von Mutius, A. S. Team, Exposure to farming in early life and development of asthma and allergy: a cross-sectional survey. *Lancet* **358**, 1129-1133 (2001).

112. W. Murk, K. R. Risnes, M. B. Bracken, Prenatal or early-life exposure to antibiotics and risk of childhood asthma: a systematic review. *Pediatrics* **127**, 1125-1138 (2011).

113. H. Sbihi, R. C. Boutin, C. Cutler, M. Suen, B. B. Finlay, S. E. Turvey, Thinking bigger: How early-life environmental exposures shape the gut microbiome and influence the development of asthma and allergic disease. *Allergy* **74**, 2103-2115 (2019).

114. A. H. Liu, Revisiting the hygiene hypothesis for allergy and asthma. *J Allergy Clin Immunol* **136**, 860-865 (2015).

115. R. L. Norton, P. R. Hoffmann, Selenium and asthma. *Mol Aspects Med* **33**, 98-106 (2012).

116. R. N. Rubin, L. Navon, P. A. Cassano, Relationship of serum antioxidants to asthma prevalence in youth. *Am J Respir Crit Care Med* **169**, 393-398 (2004).

117. A. Kocyigit, F. Armutcu, A. Gurel, B. Ermis, Alterations in plasma essential trace elements selenium, manganese, zinc, copper, and iron concentrations and the possible role of these elements on oxidative status in patients with childhood asthma. *Biol Trace Elem Res* **97**, 31-41 (2004).

118. C. Picado, R. Deulofeu, R. Lleonart, M. Agustí, J. Mullol, L. Quintó, M. Torra, Dietary micronutrients/antioxidants and their relationship with bronchial asthma severity. *Allergy* **56**, 43-49 (2001).

119. S. O. Shaheen, J. A. Sterne, R. L. Thompson, C. E. Songhurst, B. M. Margetts, P. G. Burney, Dietary antioxidants and asthma in adults: population-based case-control study. *Am J Respir Crit Care Med* **164**, 1823-1828 (2001).

120. S. O. Shaheen, R. B. Newson, A. J. Henderson, P. M. Emmett, A. Sherriff, M. Cooke, A. S. Team, Umbilical cord trace elements and minerals and risk of early childhood wheezing and eczema. *Eur Respir J* **24**, 292-297 (2004).

121. G. Devereux, G. McNeill, G. Newman, S. Turner, L. Craig, S. Martindale, P. Helms, A. Seaton, Early childhood wheezing symptoms in relation to plasma selenium in pregnant mothers and neonates. *Clin Exp Allergy* **37**, 1000-1008 (2007).

122. S. O. Shaheen, C. M. Rutterford, S. J. Lewis, S. M. Ring, J. W. Holloway, J. Golding, A. J. Henderson, Maternal selenium status in pregnancy, offspring glutathione peroxidase 4 genotype, and childhood asthma. *J Allergy Clin Immunol* **135**, 1083-1085.e1083 (2015).

123. J. Nicodemus-Johnson, R. A. Myers, N. J. Sakabe, D. R. Sobreira, D. K. Hogarth, E. T. Naureckas, A. I. Sperling, J. Solway, S. R. White, M. A. Nobrega, D. L. Nicolae, Y. Gilad, C. Ober, DNA methylation in lung cells is associated with asthma endotypes and genetic risk. *JCI Insight* **1**, e90151 (2016).

124. M. J. Aryee, A. E. Jaffe, H. Corrada-Bravo, C. Ladd-Acosta, A. P. Feinberg, K. D. Hansen, R. A. Irizarry, Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363-1369 (2014).

125. T. J. Peters, M. J. Buckley, A. L. Statham, R. Pidsley, K. Samaras, R. V Lord, S. J. Clark, P. L. Molloy, De novo identification of differentially methylated regions in the human genome. *Epigenetics Chromatin* **8**, 6 (2015).

126. J. P. Fortin, T. J. Triche, Jr., K. D. Hansen, Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics* **33**, 558-560 (2017).

127. M. C. Peters, L. Ringel, N. Dyjack, R. Herrin, P. G. Woodruff, C. Rios, B. O'Connor, J. V. Fahy, M. A. Seibold, A Transcriptomic Method to Determine Airway Immune Dysfunction in T2-High and T2-Low Asthma. *Am J Respir Crit Care Med* **199**, 465-477 (2019).

128. K. Yuryeva, I. Saltykova, L. Ogorodova, N. Kirillova, E. Kulikov, E. Korotkaya, Y. Iakovleva, I. Feoktistov, A. Sazonov, S. Ryzhov, Expression of adenosine receptors in monocytes from patients with bronchial asthma. *Biochem Biophys Res Commun* **464**, 1314-1320 (2015).

129. T. A. Doherty, D. Scott, H. H. Walford, N. Khorram, S. Lund, R. Baum, J. Chang, P. Rosenthal, A. Beppu, M. Miller, D. H. Broide, Allergen challenge in allergic rhinitis rapidly induces increased peripheral blood type 2 innate lymphoid cells that express CD84. *J Allergy Clin Immunol* **133**, 1203-1205 (2014).

130. C. A. Knosp, H. P. Carroll, J. Elliott, S. P. Saunders, H. J. Nel, S. Amu, J. C. Pratt, S. Spence, E. Doran, N. Cooke, R. Jackson, J. Swift, D. C. Fitzgerald, L. G. Heaney, P. G. Fallon, A. Kissenpfennig, J. A. Johnston, SOCS2 regulates T helper type 2 differentiation and the generation of type 2 allergic responses. *J Exp Med* **208**, 1523-1531 (2011).

131. S. Katayama, K. Stenberg Hammar, K. Krjutškov, E. Einarsdottir, G. Hedlin, J. Kere, C. Söderhäll, Acute wheeze-specific gene module shows correlation with vitamin D and asthma medication. *Eur Respir J* **55**, (2020).

132. J. H. Kim, H. S. Cheong, J. S. Park, A. S. Jang, S. T. Uh, Y. H. Kim, M. K. Kim, I. S. Choi, S. H. Cho, B. W. Choi, J. S. Bae, C. S. Park, H. D. Shin, A genome-wide association study of total serum and mite-specific IgEs in asthma patients. *PLoS One* **8**, e71958 (2013).

133. K. J. Baines, J. L. Simpson, L. G. Wood, R. J. Scott, P. G. Gibson, Transcriptional phenotypes of asthma defined by gene expression profiling of induced sputum samples. *J Allergy Clin Immunol* **127**, 153-160, 160.e151-159 (2011).

134. Y. Han, L. Chen, H. Liu, Z. Jin, Y. Wu, W. Li, S. Ying, Z. Chen, H. Shen, F. Yan, Airway Epithelial cGAS Is Critical for Induction of Experimental Allergic Airway Inflammation. *J Immunol* **204**, 1437-1447 (2020).

135. C. Igartua, R. A. Myers, R. A. Mathias, M. Pino-Yanes, C. Eng, P. E. Graves, A. M. Levin, B. E. Del-Rio-Navarro, D. J. Jackson, O. E. Livne, N. Rafaels, C. K. Edlund, J. J. Yang, S. Huntsman, M. T. Salam, I. Romieu, R. Mourad, J. E. Gern, R. F. Lemanske, A. Wyss, J. A. Hoppin, K. C. Barnes, E. G. Burchard, W. J. Gauderman, F. D. Martinez, B. A. Raby, S. T. Weiss, L. K. Williams, S. J. London, F. D. Gilliland, D. L. Nicolae, C. Ober, Ethnic-specific associations of rare and low-frequency DNA sequence variants with asthma. *Nat Commun* **6**, 5965 (2015).

136. H. T. den Dekker, K. Burrows, J. F. Felix, L. A. Salas, I. Nedeljkovic, J. Yao, S. L. Rifas-Shiman, C. Ruiz-Arenas, N. Amin, M. Bustamante, D. L. DeMeo, A. J. Henderson, C. G. Howe, M. F. Hivert, M. A. Ikram, J. C. de Jongste, L. Lahousse, P. R. Mandaviya, J. B. van Meurs, M. Pinart, G. C. Sharp, L. Stolk, A. G. Uitterlinden, J. M. Anto, A. A. Litonjua, C. V. Breton, G. G. Brusselle, J. Sunyer, G. D. Smith, C. L. Relton, V. W. V. Jaddoe, L. Duijts, Newborn DNA-methylation, childhood lung function, and the risks of asthma and COPD across the life course. *Eur Respir J* **53**, (2019).

137. H. G. Moon, S. J. Kim, M. K. Lee, H. Kang, H. S. Choi, A. Harijith, J. Ren, V. Natarajan, J. W. Christman, S. J. Ackerman, G. Y. Park, Colony-stimulating factor 1 and its receptor are new potential therapeutic targets for allergic asthma. *Allergy* **75**, 357-369 (2020).

138. E. Majorczyk, M. Jasek, R. Płoski, M. Wagner, A. Kosior, A. Pawlik, A. Obojski, W. Luszczek, I. Nowak, A. Wiśniewski, P. Kuśnierczyk, Association of PTPN22 single nucleotide polymorphism with rheumatoid arthritis but not with allergic asthma. *Eur J Hum Genet* **15**, 1043-1048 (2007).

139. K. J. Baines, M. Fricker, V. M. McDonald, J. L. Simpson, L. G. Wood, P. A. B. Wark, H. E. Macdonald, A. Reid, P. G. Gibson, Sputum transcriptomics implicates increased p38 signalling activity in severe asthma. *Respirology* **25**, 709-718 (2020).

140. J. Liu, Q. Wu, J. Shi, W. Guo, X. Jiang, B. Zhou, C. Ren, LILRB4, from the immune system to the disease target. *Am J Transl Res* **12**, 3149-3166 (2020).

141. M. Maghsoudloo, S. Azimzadeh Jamalkandi, A. Najafi, A. Masoudi-Nejad, An efficient hybrid feature selection method to identify potential biomarkers in common chronic lung inflammatory diseases. *Genomics* **112**, 3284-3293 (2020).

142. K. Nakamoto, M. Watanabe, M. Sada, T. Inui, M. Nakamura, K. Honda, H. Wada, Y. Mikami, H. Matsuzaki, M. Horie, S. Noguchi, Y. Yamauchi, H. Koyama, T. Kogane, T. Kohyama, H. Takizawa, Serum Reactive Oxygen Metabolite Levels Predict Severe Exacerbations of Asthma. *PLoS One* **11**, e0164948 (2016).

143. L. Zuo, N. P. Otenbaker, B. A. Rose, K. S. Salisbury, Molecular mechanisms of reactive oxygen species-related pulmonary inflammation and asthma. *Mol Immunol* **56**, 57-63 (2013).

144. M. A. Riedl, A. E. Nel, Importance of oxidative stress in the pathogenesis and treatment of asthma. *Curr Opin Allergy Clin Immunol* **8**, 49-56 (2008).

145. P. H. Reddy, Mitochondrial Dysfunction and Oxidative Stress in Asthma: Implications for Mitochondria-Targeted Antioxidant Therapeutics. *Pharmaceuticals (Basel)* **4**, 429-456 (2011).

146. U. Mabalirajan, A. K. Dinda, S. Kumar, R. Roshan, P. Gupta, S. K. Sharma, B. Ghosh, Mitochondrial Structural Changes and Dysfunction Are Associated with Experimental Allergic Asthma. *The Journal of Immunology* **181**, 3540-3548 (2008).

147. P. A. Underhill, T. Kivisild, Use of y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu Rev Genet* **41**, 539-564 (2007).

148. E. Zifa, Z. Daniil, E. Skoumi, M. Stavrou, K. Papadimitriou, M. Terzenidou, K. Kostikas, V. Bagiatis, K. I. Gourgoulianis, Z. Mamuris, Mitochondrial genetic background plays a role in increasing risk to asthma. *Mol Biol Rep* **39**, 4697-4708 (2012).

149. B. A. Raby, B. Klanderman, A. Murphy, S. Mazza, C. A. Camargo, Jr., E. K. Silverman, S. T. Weiss, A common mitochondrial haplogroup is associated with elevated total serum IgE levels. *J Allergy Clin Immunol* **120**, 351-358 (2007).

150. A. Flaquer, A. Heinzmann, S. Rospleszcz, B. Mailaparambil, H. Dietrich, K. Strauch, R. Grychtol, Association study of mitochondrial genetic polymorphisms in asthmatic children. *Mitochondrion* **14**, 49-53 (2014).

151. E. M. Zoratti, R. Z. Krouse, D. C. Babineau, J. A. Pongracic, G. T. O'Connor, R. A. Wood, G. K. Khurana Hershey, C. M. Kercsmar, R. S. Gruchalla, M. Kattan, S. J. Teach, S. M. Sigelman, P. J. Gergen, A. Togias, C. M. Visness, W. W. Busse, A. H. Liu, Asthma phenotypes in inner-city children. *J Allergy Clin Immunol* **138**, 1016-1029 (2016).

152. J. E. Gern, The Urban Environment and Childhood Asthma study. *J Allergy Clin Immunol* **125**, 545-549 (2010).

153. M. van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* **30**, E386-394 (2009).

154. D. A. Clayton, Transcription and replication of mitochondrial DNA. *Hum Reprod* **15 Suppl 2**, 11-17 (2000).

155. M. P. Murphy, How mitochondria produce reactive oxygen species. *Biochem J* **417**, 1-13 (2009).

156. V. Larosa, C. Remacle, Insights into the respiratory chain and oxidative stress. *Biosci Rep* **38**, (2018).

157. U. M. Sahiner, E. Birben, S. Erzurum, C. Sackesen, Ö. Kalayci, Oxidative stress in asthma: Part of the puzzle. *Pediatr Allergy Immunol* **29**, 789-800 (2018).

158. A. Nadeem, S. K. Chhabra, A. Masood, H. G. Raj, Increased oxidative stress and altered levels of antioxidants in asthma. *J Allergy Clin Immunol* **111**, 72-78 (2003).

159. T. J. Guzik, R. Korbut, T. Adamek-Guzik, Nitric oxide and superoxide in inflammation and immune regulation. *J Physiol Pharmacol* **54**, 469-487 (2003).

160. S. Loukides, D. Bouros, G. Papatheodorou, P. Panagou, N. M. Siafakas, The relationships among hydrogen peroxide in expired breath condensate, airway inflammation, and asthma severity. *Chest* **121**, 338-346 (2002).

161. D. G. Torgerson, E. J. Ampleford, G. Y. Chiu, W. J. Gauderman, C. R. Gignoux, P. E. Graves, B. E. Himes, A. M. Levin, R. A. Mathias, D. B. Hancock, J. W. Baurley, C. Eng, D. A. Stern, J. C. Celedon, N. Rafaels, D. Capurso, D. V. Conti, L. A. Roth, M. Soto-Quiros, A. Togias, X. Li, R. A. Myers, I. Romieu, D. J. Van Den Berg, D. Hu, N. N. Hansel, R. D. Hernandez, E. Israel, M. T. Salam, J. Galanter, P. C. Avila, L. Avila, J. R. Rodriquez-Santana, R. Chapela, W. Rodriguez-Cintron, G. B. Diette, N. F. Adkinson, R. A. Abel, K. D. Ross, M. Shi, M. U. Faruque, G. M. Dunston, H. R. Watson, V. J. Mantese, S. C. Ezurum, L. Liang, I. Ruczinski, J. G. Ford, S. Huntsman, K. F. Chung, H. Vora, X. Li, W. J. Calhoun, M. Castro, J. J. Sienra-Monge, B. del Rio-Navarro, K. A. Deichmann, A. Heinzmann, S. E. Wenzel, W. W. Busse, J. E. Gern, R. F. Lemanske, Jr., T. H. Beaty, E. R. Bleecker, B. A. Raby, D. A. Meyers, S. J. London, S. Mexico City Childhood Asthma, F. D. Gilliland, S. Children's Health, H. study, E. G. Burchard, S. o. G.-E. Genetics of Asthma in Latino Americans Study, A. Admixture in Latino, A. G. Study of African Americans, Environments, F. D. Martinez, R. Childhood Asthma, N. Education, S. T. Weiss, P. Childhood Asthma Management, L. K. Williams, P. Study of Asthma, R.-E. Pharmacogenomic Interactions by, K. C. Barnes, S. Genetic Research on Asthma in African Diaspora, C. Ober, D. L. Nicolae,

Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat Genet* **43**, 887-892 (2011).

162. K. K. Nishimura, J. M. Galanter, L. A. Roth, S. S. Oh, N. Thakur, E. A. Nguyen, S. Thyne, H. J. Farber, D. Serebrisky, R. Kumar, E. Brigino-Buenaventura, A. Davis, M. A. LeNoir, K. Meade, W. Rodriguez-Cintron, P. C. Avila, L. N. Borrell, K. Bibbins-Domingo, J. R. Rodriguez-Santana, Ś. Sen, F. Lurmann, J. R. Balmes, E. G. Burchard, Early-life air pollution and asthma risk in minority children. The GALA II and SAGE II studies. *Am J Respir Crit Care Med* **188**, 309-318 (2013).

163. H. Weissensteiner, D. Pacher, A. Kloss-Brandstätter, L. Forer, G. Specht, H. J. Bandelt, F. Kronenberg, A. Salas, S. Schönherr, HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res* **44**, W58-63 (2016).

164. X. Zhou, M. Stephens, Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* **44**, 821-824 (2012).

165. Q. Huang, Genetic study of complex diseases in the post-GWAS era. *J Genet Genomics* **42**, 87-98 (2015).

166. E. Meaburn, R. Schulz, Next generation sequencing in epigenetics: insights and challenges. *Semin Cell Dev Biol* **23**, 192-199 (2012).

167. D. Bar-Yaacov, A. Blumberg, D. Mishmar, Mitochondrial-nuclear co-evolution and its effects on OXPHOS activity and regulation. *Biochim Biophys Acta* **1819**, 1107-1111 (2012).