UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE

Clinical and Experimental Sciences

Role of DNA Methylation in mediating the environmental and genetic effects in asthma

by

Veeresh K Patil

Thesis for the degree of Doctor of Philosophy

January 2022

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE
Clinical and Experimental Sciences

Thesis for the degree of Doctor of Philosophy

ROLE OF DNA METHYLATION IN MEDIATING ENVIRONMENTAL AND GENETIC EFFECTS IN ASTHMA

Veeresh K Patil

Genetics and environmental exposures at critical developmental windows, play an important role in the development of asthma. Genetically, asthma is a complex trait with known genes only explaining part of the heritability. Mechanisms by which environmental exposures lead to development of asthma are not clearly understood. Epigenetic marks, particularly DNA methylation can be influenced by both environmental exposures and genotype, thus may play a role in mediating the effects of genes and environment in asthma.

In 1989, the Isle of Wight Birth Cohort (IOWBC) was established to study the risk factors for the development of asthma and allergy. In this cohort, prenatal smoking exposure, *IL13* SNPs (single-nucleotide-polymorphism) and methylation were seen to interact in a two-step analysis; methylation at cg13566430 was influenced by interaction of maternal smoking during pregnancy and rs20541, in stage-2, genotype at rs1800925 interacted with methylation at cg13566430 significantly affecting airflow limitation and reactivity. Interaction of DNA Methylation at cg09791102 and rs3024685 in the *IL4R* influence the risk of asthma.

A multigenerational cohort was established by recruiting children of the IOWBC into the Third Generation Study with a focus on transgenerational epigenetics of asthma. An intergenerational effect of prenatal smoking exposure was seen; combined maternal and grandmaternal smoking during pregnancy was significantly associated with infant wheeze in grandchildren and methylation at cg04180046 (*MYO1G*) and cg05575921 (*AHRR*). An Epigenome-wide association study (EWAS) of cord blood DNA methylation in third generation showed differential methylation at 5 CpG sites (False-discovery-rate <0.1) in relation to maternal asthma. Pathway analysis using the top 100 sites from the EWAS showed possible enrichment of T cell differentiation and activation pathways.

Findings from the thesis provide support to the concept of role of DNA methylation in mediating genetic and environmental effects in asthma and the Third Generation Study will provide valuable information for future studies in the field.

Table of Contents

List of Tables

List of Figures

Academic Thesis: Declaration of Authorship

I, Dr Veeresh K Patil declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Title of Thesis: Role of DNA Methylation in mediating the environmental and hereditary effects in asthma

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Either none of this work has been published before submission, or parts of this work have been published as: [please list references below]:

Signed: ……

Date: ……15/01/2022……

My Role and contributions

I worked as research fellow at The David Hide Asthma and Allergy Research Centre, Isle of Wight for three and half years from Feb 2011. I supported the research projects at the centre under the supervision of Professor Hasan Arshad, Director of the research centre. Two main projects I was involved in two main projects, the Third Generation Study and 10 year follow up of Food Allergy and Intolerance Research birth cohort. The Third Generation cohort is the children of the original Isle of Wight Birth Cohort (IOWBC) which had completed the18 years follow up. In this thesis, I utilised available data from the 18 years follow up in the analysis of gene-environment-DNA methylation analysis (chapter 3). Rest of the result chapters are based on the Third Generation study.

As a fellow at the research centre, I was project managing the Third Generation study submitting amendments for ethics and R&D approvals, implementing the changes and supporting research nursing team. I was involved in recruiting potential mothers and fathers from antenatal clinics, collecting questionnaires, collection and storage of samples. I did second data entry after the first data entry by the study co-ordinator. Prof G Roberts, co-director of the David Hide centre provided support and guidance throughout the fellowship, particularly with paediatric allergy and asthma knowledge and in submissions of amendments for ethics approval.

I worked under supervision of Prof J W Holloway in the Respiratory Genetics lab, University of Southampton, co-ordinating transfer of samples from the Isle of Wight and extracting DNA and RNA in the lab on IOWBC 18 years whole blood samples and Third Generation cord blood samples along with the lab technician, Mrs Nikki Graham, and preparing samples for DNA methylation analysis. Methylation data was generated by an external service provider (University of Oxford). Gene expression data was generated by Prof. Susan Ewart at Michigan State University.

I also co-ordinated sharing of data between our bioinformatics collaborators from University of South Carolina and Memphis (Prof Karmaus and Prof Zhang team). I have analysed the data from the Third Generation in this thesis. Third generation data has also been utilised by the post graduate students and fellows from Bioinformatics department, University of Memphis. I joined a weekly meeting of PhD students at Memphis Biostatistical department using web conferencing (skype) providing clinical knowledge input, while benefitting from their statistical and bioinformatics know how. I have worked with them on data analyses. and contributed to drafting of manuscripts. For exploratory interaction studies I worked closely with Dr Soto-Ramirez who led the analysis of *IL4R,* and I led on *IL13* interaction study (chapters 3.2 and 3.3). I independently undertook all data analysis described in the result chapters of this thesis apart from the EWAS (chapter 6) for which I developed the analysis plans under the supervision of Prof. Holloway and was supported by Dr Rezwan (previously bioinformatics fellow with Prof Holloway) in running the analysis where I needed access to the university supercomputer.

During the fellowship I gained clinical experience in allergy and also completed MSc Allergy with distinction from University of Southampton (2015).

Publications arising directly from the thesis

Original Articles

- Patil VK, Holloway JW, Zhang H, Soto-Ramirez N, Ewart S, Arshad SH, Karmaus W: **Interaction of prenatal maternal smoking, interleukin 13 genetic variants and DNA methylation influencing airflow and airway reactivity**. Clin Epigenetics 2013, 5(1):22

- Soto-Ramirez N, Arshad SH, Holloway JW, Zhang H, Schauberger E, Ewart S, Patil V, Karmaus W: **The interaction of genetic variants and DNA methylation of the interleukin-4 receptor gene increase the risk of asthma at age 18 years**. *Clin Epigenetics* 2013, **5**(1):1.

- Arshad SH, Patil V, Mitchell F, Potter S, Zhang H, Ewart S, Mansfield L, Venter C, Holloway JW, Karmaus WJ. **Cohort profile update: the Isle of Wight whole population birth cohort (IOWBC).** Int J Epidemiology, 2020;10.1093

- Lockett GA, Patil VK, Soto-Ramirez N, Ziyab AH, Holloway JW, Karmaus W: **Epigenomics and allergic disease**. Epigenomics 2013, 5(6):685-699.

Conference Abstracts arising directly from the thesis

- Patil VK, Mukherjee N, Lockett G, Zhang H, Karmaus W, Holloway JW, Arshad SH: **Transgenerational epigenetics: Maternal and grandmaternal gestational smoking influences DNA methylation of the AHRR gene**. European Respiratory Journal 2016, 48(suppl 60).

- Patil VK, Karmaus W, Zhang H, Mitchell F, Ewart S, Arshad SH: **Third generation study: maternal and grand-maternal smoking in pregnancy and wheeze in children. Allergy.** In: Allergy. vol. 69: 326–453; 2014 (Conference abstract)

- Mitchell FA, Patil VK, Bryan RL, Potter S, Matthews SM, Arshad SH. **The Isle of Wight Birth Cohort- seeking the 3rd Generation**. EAACI-WAO;2013 June; Milan, Italy

- Patil VK, Holloway JW, Zhang H, Ewart S, Arshad SH, Karmaus W. **Interaction of DNA Methylation and Genetic Variants of IL13 Is Associated with FEV1/FVC and BHR**(Abstract). In World allergy organization international scientific conference (WISC); 2012 Dec 6-9; Hyderabad, India.

Publications utilising Third Generation Cohort data to which I contributed but not directly arising from analyses presented in this thesis

- Guthikonda K, Zhang H, Nolan VG, Soto-Ramirez N, Ziyab AH, Ewart S, Arshad HS, Patil V, Holloway JW, Lockett GA, Karamus W: Oral contraceptives modify the effect of GATA3 polymorphisms on the risk of asthma at the age of 18 years via DNA methylation. Clin Epigenetics 2014, **6**(1):17.

- Patil VK, Kurukulaaratchy RJ, Venter C, Grundy J, Roberts G, Dean T, Arshad SH: Changing prevalence of wheeze, rhinitis and allergic sensitisation in late childhood: findings from 2 Isle of Wight birth cohorts 12 years apart. Clin Exp Allergy 2015, 45(9):1430-1438

- Everson TM, Lyons G, Zhang H, Soto-Ramirez N, Lockett GA, Patil VK, Merid SK, Soderhall C, Melen E, Holloway JW, Arshad SH, Karmaus W.: DNA methylation loci associated with atopy and high serum IgE: a genome-wide application of recursive Random Forest feature selection. Genome Med 2015, 7:89.

- Lockett GA, Soto-Ramirez N, Ray MA, Everson TM, Xu CJ, Patil VK, Terry W, Kaushal A, Rezwan FI, Ewart SL, Gehring U, Postma DS, Koppelman GH, Arshad SH, Zhang H, Karmaus W, Holloway JW: Association of season of birth with DNA methylation and allergic disease. Allergy 2016, 71(9):1314-1324.

- Losol P, Rezwan FI, Patil VK, Venter C, Ewart S, Zhang H, Arshad SH, Karmaus W, Holloway JW. Effect of gestational oily fish intake on the risk of allergy in children may be influenced by FADS1/2, ELOVL5 expression and DNA methylation. Genes Nutr. 2019;14:20.

- Everson TM, Zhang H, Lockett GA, Kaushal A, Forthofer M, Ewart SL, Burrows K, Relton CL, Sharp GC, Henderson AJ, Patil VK, Rezwan FI, Arshad SH, Holloway JW, Karmaus W. Epigenome-wide association study of asthma and wheeze characterizes loci within HK1. Allergy Asthma Clin Immunol. 2019;15:43.

Publications utilising Third Generation Cohort data collected as part of this thesis subsequent to completion of my research fellowship

- Quraishi BM, Zhang H, Everson TM, Ray M, Lockett GA, Holloway JW, et al. Identifying CpG sites associated with eczema via random forest screening of epigenome-scale DNA methylation. Clin Epigenetics. 2015;7:68.

- Reese SE, Xu CJ, den Dekker HT, Lee MK, Sikdar S, Ruiz-Arenas C, et al. Epigenome-wide meta-analysis of DNA methylation and childhood asthma. J Allergy Clin Immunol. 2019;143(6):2062-74.

- Kupers LK, Monnereau C, Sharp GC, Yousefi P, Salas LA, Ghantous A, et al. Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight. Nat Commun. 2019;10(1):1893.

- Kazmi N, Sharp GC, Reese SE, Vehmeijer FO, Lahti J, Page CM, et al. Hypertensive Disorders of Pregnancy and DNA Methylation in Newborns. Hypertension. 2019;74(2):375-83.

- Merid SK, Novoloaca A, Sharp GC, Kupers LK, Kho AT, Roy R, et al. Epigenome-wide meta-analysis of blood DNA methylation in newborns and children identifies numerous loci related to gestational age. Genome Med. 2020;12(1):25.

- Kheirkhah Rahimabad P, Arshad SH, Holloway JW, Mukherjee N, Hedman A, Gruzieva O, et al. Association of Maternal DNA Methylation and Offspring Birthweight. Reprod Sci. 2021;28(1):218-27.

- Luo R, Mukherjee N, Chen S, Jiang Y, Arshad SH, Holloway JW, et al. Paternal DNA Methylation May Be Associated With Gestational Age at Birth. Epigenet Insights. 2020;13:2516865720930701.

Awards and Merits through the Research Fellow work

- Best research presentation, East of England Deanery Respiratory Training Day, Cambridge, 2015.
- Best Poster Prize, Asthma Diagnosis, EAACI (European Academy of Allergy and Clinical Immunology), Copenhagen 2014.
- Best Poster Prize, Risk factors for Asthma, EAACI (European Academy of Allergy and Clinical Immunology), Copenhagen 2014.
- Outstanding abstract award and First prize for Junior Member's Session poster; World Allergy Organisation, December 2012.

Acknowledgements

I would like to express my sincere appreciation to those whose support has made this project and completion of thesis possible. First of all, I would like to thank all the participants and their parents for their valuable time and involvement in the study.

I would like to express my sincere gratitude to my supervisors; Professor Hasan Arshad for providing the opportunity of research fellowship, constant support in and outside the research field and excellent supervision, Professor John Holloway for providing insight into epigenetics, motivating discussions, meticulous reviews, valuable time and supervision, and Professor Graham Roberts for valuable advice, support, time, and supervision.

My special thanks to the research nurses and study co-ordinators at The David Hide Asthma and Allergy Research Centre, Isle of Wight for their help. I would also like to thank Professor Wilfried Karmaus and his team at University of Memphis for their support with bioinformatics. Special thanks to all the fellows and technical staff in Professor Holloway`s laboratory.

I will always remain indebted to my mother for the unconditional love and grateful to my father and brother for their support. Lastly, in no way I can thank enough three people who complete me; my wife Dr Rajashri Patil for her patience and constant support and my daughters, Khushi and Inchara for their love.

Abbreviations

Abbreviation	Full designation
ADAM33	A disintegrin and metalloproteinase domain-containing protein 33
AHRR	Aryl hydrocarbon receptor repressor
BHR	Bronchial hyperresponsiveness
CATSPERB	Cation channel sperm associated auxiliary subunit beta
CD14	Cluster of differentiation 14
CNTNAP2	Contactin-associated protein-like 2
CpG	Cytosine phosphate guanine dinucleotide
CPT1A	Carnitine Palmitoyltransferase 1A
CYP1A1	Cytochrome P450 family 1 subfamily A member 1
DENND1B	DENN Domain Containing 1B
DNMT	DNA methyltransferase
DOHaD	The Developmental Origin of Health and Disease
DRS	Dose response slope
EWAS	Epigenome-wide association study
FDR	False discovery rate
FeNO	Fractional exhaled nitric oxide
FEV1	Forced expiratory volume in one second
FLG	Filaggrin
FVC	Forced vital capacity
GATA3	GATA binding protein 3
GF11	Growth differentiation factor 11
GFI1	Growth factor independent 1 transcriptional repressor
GINA	Global Intiative for Asthma
GSTM1	Glutathione- S-transferase mu 1

GWAS	Genome-wide association study
HLA	The human leukocyte antigen
IgE	Immunoglobulin E
IL13	Interleukin 13
IL18R1	Interleukin 18 receptor 1
IL1RL1	Interleukin 1 receptor like 1
IL1RN	Interleukin receptor 1 antagonist
IL2RB	Interleukin 2 receptor subunit beta
IL33	Interleukin 33
IL4R	Interleukin 4 receptor
IOWBC	The Isle of Wight Birth Cohort
ISAAC	International Study of Asthma and Allergies in Children
LINE1	long interspersed repetitive element-1
MBD	methyl CpG binding domain protein
MeCP2	methyl CpG binding protein 2
miRNA	Micro RNA
MOBA	The Norwegian Mother and Child Cohort Study
mRNA	Messanger RNA
MYO1G	Myosin 1G
ORMDL3	ORMDL sphingolipid biosynthesis regulator 3
PANK4	Pantothenate Kinase 4
PCP4	Purkinje Cell Protein 4
PDE4D	Phosphodiesterase 4D
PEFR	Peak expiratory flow rate
QTL	Quantitative trait locus
RAD50	RAD50 double strand break repair protein
RUNX1	Runt related transcription factor 1

SES	Socioeconomic status
SHISA3	shisa family member 3
SNP	single nucleotide polymorphisms
SPATA22	Spermatogenesis Associated 22
SPT	Skin Prick Test
TH1 cells	T helper 1 cells
Th2 cells	T helper 2 cells
TSLP	Thymic stromal lymphopoietin
ZNF75A	Zinc Finger Protein 75a

Chapter 1 Introduction

1.1 Asthma

1.1.1 Definition

Asthma is a chronic respiratory disease with characteristic features of chronic airway inflammation, airway airflow limitation and reactivity leading to characteristic symptoms of breathlessness, wheeze, chest tightness and cough, occurring recurrently with varied frequency and severity. Global Initiative for Asthma (GINA) defines asthma as "a heterogonous disease, usually characterized by chronic airway inflammation. It is defined by the history of respiratory symptoms such as wheeze, shortness of breath, chest tightness and cough that vary over time and in intensity, together with variable expiratory airflow limitation"(1) .

Symptoms and airflow limitation are often triggered by various factors, exposure to allergen or irritants, respiratory viral infections and exercise. Either symptoms or physiological abnormalities or a combination of both can be used to diagnose asthma in clinical and research settings (1-3). In recent years, increasing understanding of asthma has led to the concept that asthma represents a spectrum of phenotypes; it's a syndrome rather than a single disease entity (4). Multiple phenotypes are now recognised based on the observable characteristics, which can guide in management of asthma based on the phenotype (1).

1.1.2 Prevalence

A significant methodological limitation in epidemiological asthma research is the lack of standardization in the definition of asthma definition (5) and the heterogeneity seen in asthma phenotypes leading to difficulties in getting an accurate prevalence of asthma. Asthma is known to affect 1-18% of the population in different countries (1), it is estimated that asthma currently affects more than 300 million people globally (6). Asthma is the most common chronic disease in children (7) and affects 5.4 million people in the UK (8). Asthma prevalence is high in developed parts of the world. However, the prevalence

increased both in the developed (9) and the developing (10) parts of the world in the last two decades of the twentieth century. An increase has been seen in prevalence of diagnosed asthma, prevalence of asthma symptoms and also in airway hyperresponsiveness (10). Recent studies have suggested that prevalence has been plateauing in the developed world in the first decade of the twenty first century (9, 10). The time trends in the prevalence of asthma in 10 years old children comparing two unselected birth cohorts established in 1989 and 2001 respectively; the Isle of Wight Birth Cohort (IOWBC) and Food-Allergy-and-Intolerance-Research cohort (FAIR), showed that prevalence of reported and treated asthma was declining (11). Causes of the temporal increase in the prevalence of asthma have not been clearly understood. Asthma is a complex disease with both genetic and environmental exposures involved in the etiopathogenesis. The observed change in prevalence is too quick to be explained by genetic drift and it is most likely to be due to changes in lifestyle and other environmental factors and also with respect to clinical diagnosis, in changes to physician behaviour.

1.1.3 Allergy

The term “Allergy” was first used by Clemens von Priquet more than 100 years ago (1906) to describe “the changed or altered reactivity” induced by an external agent termed allergen (12). However over the last century, usage of the term allergy has changed to mainly focusing on hypersensitivity reactions and the word “Allergic disease” is used to describe the spectrum of clinical conditions with hypersensitivity being the main pathophysiological mechanism (13). The term allergic disease covers a long list of conditions, with the most common being asthma, allergic rhinitis, anaphylaxis and food allergy. Allergic sensitization is a common background to most allergic diseases and can be ascertained by skin prick test (SPT) or measurement of specific IgE in serum.

1.1.3.1 Atopy and its relationship with asthma

Atopy was a term previously used to loosely address allergic conditions like eczema, hay fever, food allergy and asthma. With evolving knowledge in the field since the description of IgE in 1966 (14) the term "Atopy" has been more specifically defined as a genetic predisposition to develop allergic sensitisation i.e. tendency to develop IgE response to specific allergen usually demonstrated either by positive skin prick test or allergen specific IgE in serum.

Atopy is known to be associated with asthma in childhood and adulthood, however the relationship is not clear (15, 16). Atopic march, also called allergic march, is a concept used to describe the natural progression of allergic conditions; eczema in early childhood to allergic rhinitis and food allergy in childhood and asthma in later life (17). Atopy remains the common underlying process in the march of these allergic conditions through childhood. There is still debate about the etiological significance of atopic march. Although there have been some longitudinal studies supporting this progression of allergic diseases, others concluded most children suffer from co-morbidities rather than the typical progression from one allergic disease to the other (18, 19).

Allergic sensitisation is associated with asthma during childhood and into adulthood, but there is no clear evidence of causal relation (20); longitudinal study of birth cohorts and cross-sectional studies show an increase in allergic sensitisation with age from childhood while asthma prevalence declines over the same time period (15, 21). Establishing this causal relationship needs large longitudinal epidemiological cohort studies but there are challenges of selection bias in these studies as participants who develop allergic sensitisation, an allergic condition or asthma are more likely to participate in follow up assessments (22). Another challenge is the broader definition of asthma in the studies, for example, based on physician diagnosed asthma or based on the responses to questionnaires, these methods may fail to capture different phenotypes of asthma (23), which might have different underlying pathophysiology (21, 24) and might have different relationship with atopy at different ages (15).

1.1.4 Immunology of Asthma

Asthma is a chronic inflammatory condition driven by immune response pathways initiated by various triggers including allergens, viruses, tobacco smoke exposure and exercise. Different endotypes (an endotype is a specific group characterised by specific pathophysiological pathway and immunological markers) have been recognised based on different immune response pathways involved in the pathophysiology of asthma (25). Eosinophilic and non-eosinophilic asthma remain the well-recognised broader endotypes at the extreme ends in the spectrum of asthma pathophysiology referred to as T-helper 2 lymphocyte (Th2) high and Th2 low (25-27).

1.1.4.1 Th2 high pathway

Eosinophilic asthma is driven by a Th2 immune response either following an allergic or non-allergic trigger. The allergic Th2 response is the better studied and understood pathway; following the exposure to allergen, the dendritic cells present the allergen to naïve T cells which leads to activation of Th2 cells. Th2 cytokines; IL-4 (IL; Interleukin), IL-5 and IL-13 are then released from activated Th2 cells. IL-4 and IL-13 cause class switching of the immunoglobulins production by B cells leading to increase in immunoglobulin E (IgE) which initiates degranulation of mast cells thus releasing histamine and leukotrienes causing bronchoconstriction (27). The non-allergic eosinophilic pathway, mainly triggered by pollutants and tobacco smoke exposure, involves activation of ILC2 (Innate Lymphocyte-2) cells which in turn produce IL-5, IL-9 and IL-13. These cytokines will cause eosinophilic inflammation, but due to lack of IL-4 production in this pathway there is no activation of class switching and hence lack of IgE production (26).

Role of Th2 cytokines

IL-4, IL-5 and IL-13 form the core of the Th2 response and activation of cytokines like IL-25, IL-33 and TSLP (Thymic stromal lymphopoietin) maintains the perpetual nature of the Th2 inflammation.

IL-5 promotes developments of eosinophils in bone marrow and by producing chemotactic molecules facilitates recruitment of eosinophils to the airway (27). Anti-IL-5 biological therapies are known to reduce eosinophilia resulting in clinical improvement. Currently Mepolizumab, Resluzimab and Benralizumab are the anti-IL-5 biological therapies in clinical use; Mepolizumab and Resluzimab bind IL-5 molecules whereas Benralizumab blocks the IL-5 receptor (25, 28).

IL-4 activates the class switching of immunoglobulin production by B cells (27); IgE production is enhanced, which in turn via priming of mast cells in the airways causes bronchoconstriction and typical allergic response following subsequent allergen exposure. Blocking IL-4 by targeted biological molecules like Dupilumab can lead to improvement in lung function and reduction in frequency of exacerbations in eosinophilic asthma patients (28).

IL-13 promotes goblet cell metaplasia leading to increased mucous production in eosinophilic asthma (27). IL-13 causes smooth muscles contraction and hypertrophy (bronchial constriction) which is the essential part of BHR (bronchial hyperreactivity), which in turn is the hallmark of asthma. IL-13 along with IL-4 also promotes class switching, which causes production of IgE by B cells. Anti-IL-13 therapy (lebrikizumab) is known to improve lung physiological parameters in Th2 high asthmatic patients (25, 28).

1.1.4.2 Th2 low pathway

Late onset asthma and steroid resistant asthma phenotypes do not show typical eosinophilic inflammation, rather demonstrate neutrophilic inflammation, which is usually driven by Th1 and Th17 pathways (25). A Th1 response is activated by respiratory viral infections which upregulates production of interferon- γ (IFN- γ) and IL-27. IFN-γ and IL-27, in addition to fighting the infection, are involved in airway inflammation (27).

Th17 response involves the cytokines IL-17 and IL-22. IL-17 activates neutrophilic inflammation and airway epithelial transformation while IL-22

promotes smooth muscle hyperplasia. Thus, together, they lead to airway remodelling in asthma (25, 27).

Overlap of these immunological responses is common in the pathophysiology of asthma, which has a wide spectrum of different inflammatory pathways. This poses challenges in clinical practice to clearly identify and select a targeted biological treatment.

1.1.5 Asthma: Physiological traits

Asthma is associated with airway inflammation and narrowing. Airway obstruction, which is usually reversible, airway hyperresponsiveness, inflammatory markers such as exhaled nitric oxide and high serum Immunoglobulin E (IgE) are asthma related physiological traits which can be quantified. These quantitative asthma traits are frequently studied as outcomes in asthma research. Physiological abnormalities associated with asthma provide objective measures of the condition, however they are neither exclusive to asthma nor provide a gold standard marker (test) to diagnose asthma; lung function can be normal between asthma attacks and bronchial hyperresponsiveness (BHR) can be present in the absence of asthma. Moreover, asthma and BHR are can be associated with different profiles of risk factors (29). Nonetheless, these measures are helpful and provide independent and complementary objective information to physicians in addition to symptoms (30).

1.1.5.1 Airway traits of asthma

Airway obstruction

The dynamic lung volumes form the basis for clinically diagnosing airway obstruction in asthma (30). Peak expiratory flow rate (PEFR) is a simple measure of maximum rate of expiration after a maximum inhalation. It is helpful for monitoring the variation in airflow but because of high inter-individual variability (>30%) it is not very useful in diagnosing the airway obstruction (31).

Spirometry is useful in measuring lung volumes and is used in diagnosing and monitoring asthma (32). Spirometric measurement of forced expiratory volume in 1 second (FEV_1), forced vital capacity (FVC), FEV_1/FVC ratio and pre- and post-bronchodilator FEV_1 are the core physiological measurements in asthma research particularly for characterisation of participants in prospective observational studies (33). Forced vital capacity (FVC) is the volume of total air exhaled forcefully after a maximum inhalation. Forced expiratory volume at 1 second (FEV_1) is the volume of air exhaled in the first second of FVC manoeuvre. FEV_1/FVC is the ratio of FEV_1 to FVC also termed as forced expiratory ratio (FER). FEV_1 is regarded as a measure of asthma control in addition to symptoms. In epidemiological studies, FEV_1 has been shown to predict asthma outcomes (34). Pre-bronchodilator FEV_1 is a commonly used endpoint in clinical trials and is sensitive to fluctuations in airflow limitation (30).

Bronchial hyperresponsiveness

Bronchial hyperresponsiveness (BHR) is an objective measurement of airway sensitivity to a stimulus. BHR is characteristic of airway smooth muscle's excessive response to a stimulus and measures variable airway obstruction. BHR is measured by provocation of airway with a stimulus and measuring the degree of reduction in FEV_1 as a measure of airway obstruction. BHR, like FEV_1, is weakly associated with asthma symptoms and airway inflammation (30). However studies have shown BHR to be associated with airway infection (35) suggesting the role of inflammation in BHR. Moderate to severe BHR is associated with significant morbidity in asthma (36), however BHR can be seen in healthy subjects without asthma and not all subjects with asthma have BHR (30). BHR decreases with increasing age; it is high in childhood and declines in adolescence (34, 37).

In a prospective epidemiological study, Gerritsen et al showed that initial BHR is associated with persistence of asthma (34). Asymptomatic BHR can be associated with asthma in later life (38), but other studies show contrasting results.

Therefore, it is not clear if non-asthmatic BHR predates the development of asthma (37, 39).

Terminologies used to describe the degree of bronchial responsiveness are, airway reactivity and bronchial reactivity. In this thesis, BHR will be used to refer to binomial outcome of bronchial challenge and airway reactivity for a continuous outcome of bronchial challenge.

1.1.5.2 Inflammatory traits of asthma

Induced sputum cell count

This is a non-invasive way of assessing airway inflammation and it also helps in determining the type of inflammation or the type of pathology occurring in the airways (40). Eosinophils and neutrophils are measured in sputum regularly for clinical follow up of asthma in severe and difficult asthma clinics. Differential inflammatory cell count from induced sputum is used to differentiate eosinophilic and neutrophilic inflammation. The differential counts are presented as percentage of total cell count. Disadvantages of induced sputum cell count are; lack of reversibility test, it is not suitable for children (<8 years) and it is not advisable in those with a low FEV_1 (<1.0 L) (30, 40). Induced sputum differential cell count improves phenotyping of asthma which allows for tailoring the treatment such as neutrophilic asthma being more steroid resistant (40).

Fractional exhaled nitric oxide

Fractional exhaled nitric oxide (FeNO) is another non-invasive surrogate marker of eosinophilic airway inflammation. Nitric oxide is derived from the epithelial nitric oxide synthase and FeNO has a high correlation to eosinophilic airway inflammation (41), suggesting that eosinophils trigger the induction of epithelial nitric oxide synthase, but the exact mechanism is not clear (42). In corticosteroid naïve patients FeNO indicates response to corticosteroid treatment and is also used in monitoring of severe asthma patients (30). FeNO is not a good marker of eosinophilic airway inflammation in current smokers (43). Atopy can be a confounding factor for FeNO in asthma, as studies have shown FeNO to be

associated with atopy (44). In mild asthma FeNO is inversely related to asthma control questionnaire (ACQ) scores, >40% reduction in FeNO had a positive predictive value of increase in ACQ by 83% (45).

Immunoglobulin E (IgE)/ Cord blood IgE

Allergy plays a significant role in asthma. Immunoglobulin E (IgE) is involved in allergic sensitisation and response. IgE is integral to most allergic conditions; it binds to the high affinity IgE receptor (FcεRI) on the surface of mast cells and on cross-linking by an allergen causes mast cell degranulation leading to allergic reaction. In case of asthma this leads to airway inflammation, bronchospasm and mucus production (46, 47). IgE has been detected in cord blood serum and has been studied as predictors for allergic diseases in childhood and adulthood. In IOWBC, cord blood IgE is associated with aeroallergen sensitisation at 4 years but not to airway symptoms (48). Cord IgE is associated with both allergen sensitization and asthma at 10 years of age (49).

In asthma epidemiological research clear characterisation of the outcome is vital and can be sometimes challenging. Apart from questionnaire-based outcomes, objective airway and inflammatory measurements can be applied, however these outcomes cannot be used in place of asthma but as a specific trait of asthma. For example, bronchodilator reversibility is weakly associated with airway inflammation and airway hyperresponsiveness, but significantly associated with steroid treatment response in asthma (30). And BHR is not a surrogate marker of asthma symptoms or airway inflammation, however BHR gives independent and complementary objective information in asthma (30). FeNO is increasingly used in management of asthma, however FeNO based treatment titration is shown to reduce exacerbations but not overall asthma control and its use in children (<6 years) is still not clear (50). IgE can be raised in other comorbidities which usually coexist along with asthma like eczema, thus IgE cannot be used as surrogate to asthma. None the less these airway and inflammatory traits provide objective outcome which might be more relevant in studies focusing on specific biological pathways or clinical parameter associated with asthma.

1.2 Environmental exposures and asthma

The increase in prevalence of asthma in recent decades that cannot be explained by natural genetic alterations, and the heterogeneity of asthma seen in populations of the same ethnicity but different environmental exposures (51), support a significant role of environment in its pathogenesis. Many hypotheses have been proposed for the environmental effect on the development of asthma; including “early life origins”, “hygiene hypothesis” and “artificial habitat hypothesis” which all involve the environmental exposure as the main focus (52).

1.2.1 Environment and Asthma: Hypotheses

Early Life Origins

Early life origin hypothesis also referred to as The Developmental Origin of Health and Disease (DOHaD) hypothesis, proposes that prenatal and infancy is a critical time in which the impacts on the child by environment lead to long term changes and determine the development of the disease in later life (53). Early life origin has been proposed to explore the aetiology of most complex diseases such as coronary artery disease, diabetes and asthma (54). Prenatal (in the womb through maternal exposures), perinatal (during delivery) and postnatal (exposures during infancy) are included in the early life origin concept. These early lifetime points correspond to the development and maturation of the immune system, and it is plausible that these early life exposures leave a mark on the developing immune system leading to the trajectory of adult asthma disease status. The foetal immune system is usually skewed to Th2 response to prevent maternal rejection of the “foreign foetus”. Maternal constraint can prevent the foetal immune system from maturation thus leading to increased risk of allergy or asthma later in life. Prenatally, foetal environment includes the maternal external environment, maternal gastrointestinal environment, placental structure and physiology. The foetal environment influences foetal growth and may pre-set the risk for developing disease later in life (55).

One of the proposed methods for the early life exposures leading to disease in adulthood is "developmental plasticity", learning from the cues from early life exposures. The developing organism establishes a phenotype suitable to cope with the forecasted future. But when there is a mismatch between the forecasted future and the actual demand in later life, this plasticity can increase the risk of disease (56). Early life exposures do play a part in the development of asthma i.e., asthma, at least in part, has its origins in early life. However, the aetiology of asthma is multifactorial with genetic predisposition, environmental exposures throughout life and gene-environmental interactions playing their roles.

Foetal lung development

Lung tissue is derived from all three germ layers: endoderm, mesoderm and ectoderm. Pulmonary epithelial cells are derived from endoderm, vascular, connective, cartilage tissues and muscles are derived from mesoderm and the neural tissue is derived from ectoderm (57). At 4 weeks of postmenstrual age lung bud appears from the foregut, from 5-9 weeks tracheobronchial branching to the level of segmental bronchi develops. Between 6 – 18 weeks glandular tissue appears along with elongation and repetitive branching of airways along with vascular development. The next stage is between 16- 26 weeks when differentiation of alveolar lining cells into type 1 and 2 cells happens along with formation of air blood barrier. Air sacs are formed between 24-38 weeks and then secondary septa are formed along with fusion and thinning of double-capillary network from 36 weeks up to 2 years of age (58). Thus, the phases of lung development start as early as 4 weeks gestation and last throughout prenatal life and into early postnatal life. Exposures like smoking and other factors leading to restricted intrauterine growth can have implication from early prenatal to postnatal life with effects persisting long after (57, 58).

Hygiene hypothesis

The hygiene hypothesis was first proposed by Strachan in 1989; *"Over the past century declining family size, improved household amenities and higher standards of* personal *cleanliness have reduced opportunities for cross- infection in young*

families. This may have resulted in more widespread clinical expression of atopic disease." It was proposed that the increasing prevalence of the allergic diseases in western world was because of decreasing exposure to early life infections, which are thought to be protective. This was supported by the Th1 and Th2 pathways of immune response; early life exposure to infections would switch on the Th1 response and thus suppressing the Th2 response, which plays a major role in allergic inflammation. Despite this plausible immunological mechanism to support the hypothesis; epidemiological studies, particularly of asthma, have failed to provide conclusive evidence to support the hypothesis. Some studies have demonstrated a protective effects of early childhood factors like day care attendance, presence of siblings at birth and farm exposure on asthma (59). However others show no effect of family size or birth order for allergen sensitization, allergic rhinitis or asthma (60). The hygiene hypothesis has failed to explain the increase of asthma prevalence in inner city populations and an increase in non-allergic asthma (61). This highlights that multiple factors are involved in asthma development or different factors induce the development of different forms of asthma.

Artificial habitat hypothesis

Maziak proposed that along with infections, non-infectious exposures should be included in the epidemiology of asthma and stated; *"Changes that occurred during the second half of the 20th century in industrialized nations with the spread of central heating/conditioning, building insulation, hygiene, TV/PC/games, manufactured food, indoor entertainment, cars, medical care, and sedentary lifestyles all seem to be depriving our children from the essential inputs needed to develop normal airway function (resistance). Asthma according to this view is a manifestation of our respiratory maladaptation to modern lifestyles, or in other words to our increasingly artificial habitats"* (61).

1.2.2 Early life exposures

Environmental exposures during the early life of humans can play a significant role in development of diseases later in life. Early life includes prenatal, birth and early childhood (up to 2 years). A number of such early life exposures have been suggested to play a role in asthma susceptibility, few examples are discussed below.

1.2.2.1 Prenatal Exposures

Maternal smoking during pregnancy

Prenatal environmental tobacco smoke exposure affects foetal lung parenchymal structure *in utero* (62) and is associated with impaired lung function and asthma in childhood (63) and in adolescence (64). Parental smoking is associated with increased wheezing incidence in children; a systematic review of fifty one longitudinal and case control studies concluded that maternal smoking was a risk for wheezing incidence in children (strongly up to 6 years and weakly linked after 6 years) and was associated with worse symptoms in children with diagnosed asthma (65). Though the general awareness regarding adverse effects of tobacco smoking on the fetus is growing, the incidence of smoking during pregnancy continues to be significant public health problem. For example, the BAMSE (Swedish abbreviation for Children, Allergy, Milieu, Stockholm, Epidemiology) Stockholm birth cohort study observed that 12 % of mothers smoked during the first or second trimesters and 8% smoked throughout pregnancy (66). All these findings highlight the significance of considering prenatal smoking as a risk for asthma.

Maternal diet and medications during pregnancy

Maternal dietary exposures during pregnancy have been associated with development of asthma in offspring. Vitamin D and E intake during pregnancy have been shown to be negatively associated with childhood asthma (67). A meta-analysis of 32 studies looking into maternal diet during pregnancy found protective effect of maternal intake of Vitamin D, Vitamin E and Zinc for early

childhood wheeze, however no consistent effect has been seen for asthma *per se* in offspring (68). Consumption of vegetables and fruits during pregnancy have been shown to be protective for asthma and allergic rhinitis (69, 70). Many prospective cohort studies have looked at maternal intake of fish, fruit and other dietary components during pregnancy with varying conclusions, however further randomized interventional studies are needed to confirm these associations.

Maternal use of antibiotics during pregnancy can increase the risk, and use of probiotics has been seen to be protective for asthma in offspring (71). Pooled analysis in meta-analyses has shown the use of acid suppression medications (histamine-2 receptor blockers or proton pump inhibitors) is associated with increased risk of childhood asthma in offspring (72, 73).

Other prenatal exposures (through maternal environment) such as traffic related pollution, house dust mite and viral infections have also been associated with increased risk of asthma in offspring (52).

Low birth weight

Adverse prenatal particularly intrauterine environment exposures can lead to lower birth weight and permanent changes in individual physiology and metabolism. Thus, increasing risk of development of diseases later in life like asthma. Potential mechanisms this association are not well understood. Various studies have shown that low birth weight is a risk factor for development of asthma (74-76) . Meta-analysis of 13 cohorts showed that infants born with low birth weight are had an approximately 16% higher risk of childhood asthma (77). Whether this association persists into the adulthood asthma has still not been conclusively established. However there is evidence that birth weight is associated with lower lung function in adult life, albeit reduced lung function is not always asthma (78).

1.2.2.2 Birth factors

Birth order

Hygiene hypothesis suggested sanitary environments in childhood may play a role in increasing susceptibility of allergic diseases in children after studying the relationship of birth order and allergic rhinitis (79). Exposure to older children at home is shown to be associated with lower risk of development of wheezing and development of asthma in later childhood (80). Children from later birth order (second or third born) get frequent exposure to pathogens which can shift the balance of Th1 and Th2 towards Th1. However, studies have shown contrasting findings depending on the age of children; presence of older siblings increasing risk of asthma in early childhood and protective for asthma in later childhood (81, 82). One of the reasons for this finding could be the heterogeneity of early childhood wheeze making asthma diagnosis difficult (83).

Mode of birth

Vaginal delivery is the first event for colonization of the gut microbiota for the newborn, caesarean section delivery on the other hand bypasses this opportunity. Gut microbiota differs in children born by Caesarean section compared to those born by vaginal delivery (84). Furthermore, a difference in gut microbiota (quantity and type) has been shown between children developing and those not developing atopy or allergic conditions (85)**.** Caesarian section delivery has been shown to be a risk factor for asthma (86).

Season of birth

The hypothesis for considering the month or the season of birth as a factor associated with the development of allergic conditions is partly based on the difference in the UV-B radiation exposure throughout the year and thus difference in vitamin D levels. This is supported by growing evidence of the link between vitamin D status and atopy or allergic diseases. The north-south gradient seen in adrenaline autoinjector and hypoallergenic formula prescriptions in the US (87) and Australia (88) to some extent suggests an effect of latitude on allergic

conditions. Low cord blood vitamin D levels and Interleukin10 (IL-10) have also been associated with birth in the months of low UV-B radiation (89). Although the results of vitamin D supplementation trials for primary prevention of asthma have proved mixed (90, 91).

Socioeconomic status

Several studies looking at socioeconomic status (SES), most often measured by parental occupation, education, number of bedrooms or family income, as an exposure have reported that children born in families with low SES have an increased risk of developing asthma (92-94). SES is not only a risk for developing asthma, but is also associated with asthma control; a low SES is associated with poor control of asthma in children (95). There is some evidence for the protective effect if families moved out of poverty (92).

1.2.2.3 Postnatal exposures

Early childhood infections

Early childhood airway infections have a contrasting effect on the development of asthma. Upper airway viral infections have been shown to be protective and lower airway/chest infection to be risk factors (59, 96). Repeated upper airway viral infections other than those with associated lower airway/chest infections are shown to be protective for development of asthma at 7 years (96). Similar results were seen in the Canadian early childhood development study, where a history of early childhood nose and throat infections was protective for asthma development in preschool children (59).

Chest infections in early years of childhood are risk factors for developing asthma, A German birth cohort study demonstrated a positive dose response association of early childhood chest infection to physician diagnosed asthma, the effect being strongest for more chest infections in the first three years of life (96). This study also showed that the total number of non-respiratory infections in the first three years of life was protective for asthma diagnosis at 7 years. When the different

types of infection were separated only viral infections was significant and the bacterial, fungal, urinary or gastro-infections were not significant (96).

Breast feeding

Breastfeeding as a factor associated with allergic diseases in general and asthma in particular has been studied in many studies, but the results are not conclusive (59). Several studies have shown breastfeeding to be protective for the development of asthma in childhood (59, 86, 97), while some studies have shown breastfeeding as a risk factor for asthma (98, 99). In the IOWBC, children breastfed for at least 4 months duration had better lung function at 10 years compared to those who were not breastfed (100). Meta-analysis of more than 100 studies has shown that children who are breastfed longer have a lower risk of asthma (101).

Pet exposure

Pet exposure particularly cat and dog exposure has been studied in association with asthma. In the Tuscon cohort indoor dog exposure in low risk children (no parental history of asthma) was protective (102) but a meta-analysis of thirty two studies revealed that dog exposure was a risk and in cohort studies cat exposure was protective (103) but this meta-analysis has been criticised for poor definitions of exposures, missing of some cohort results and not considering reverse causation (104). A Recent pooled analysis of 11 prospective European cohorts did not find any association of pet ownership and asthma in 6-10 year old children (105). From the above studies the effect of pet exposure on asthma is not clear. In the IOWBC, early life dog exposure was a risk and cat exposure was protective for childhood wheeze (106).

1.2.3 Adolescent exposures

Adolescence is a period of dynamic physiological changes marked by a growth spurt and reaching maximum lung function. Behavioral changes are also seen in this age such as starting smoking thus causing long term effects on lung function. After early life, adolescence could be another critical time for environmental influences on asthma.

Smoking

Epidemiological studies have shown that smoking exposure during adolescence is a risk factor for wheeze (107) and asthma in adolescence (108), (109). It is however important to untangle early life tobacco smoke exposure particularly prenatal exposure from adolescent smoking, but this is difficult and not often examined in studies looking into adolescent smoking and asthma. In a longitudinal follow up of 2,289 children in adolescent age, any (passive or active) smoking was associated with all three groups; current, persistent and late onset symptoms of wheeze or cough (110). In the IOWBC, analysis focusing on undiagnosed wheeze in adolescence, found adolescent smoking to be a risk for undiagnosed wheeze (107).

Medication use

Paracetamol (acetaminophen) is a common medication used at all ages and also used frequently during pregnancy. Paracetamol can predispose the individual to oxidative tissue injury by affecting the level of the antioxidant glutathione (111). It is plausible that paracetamol can predispose the lung to oxidative injury following environmental insults and use of paracetamol has been shown to be associated with asthma (112, 113). In the IOWBC, paracetamol use in adolescence was associated with undiagnosed wheeze in adolescence (107).

1.3 Challenges of epidemiological association studies

1.3.1 Characterisation of outcome

Asthma is a heterogeneous condition and studies investigating the risk factors associated with asthma face further challenges due to the multifactorial origin and even if the association is proven, it is difficult to prove causation. With growing knowledge about asthma being a syndrome encompassing multiple overlapping phenotypes and endotypes it is important that there is clear characterisation of asthma, however this remains challenging for large epidemiological studies.

One clear example of these challenges is the outcome measure of parent reported wheeze, commonly used in birth cohorts as a surrogate variable for childhood asthma. Diagnosing childhood asthma is difficult. Asthma has been described for many decades (or centuries) now but still remains a challenging diagnosis due to its subjective nature of symptoms and the physicians perspectives (114). Epidemiological studies of asthma mainly rely upon responses to the questionnaires to confirm the asthma outcomes. In young children this is based on parental report of symptoms like wheeze. Although this is a simple and straightforward way of collecting the information, parent reported wheeze has many limitations. Not all wheeze in childhood is asthma, and further information is needed on the frequency of symptoms and presence or absence of symptoms of viral infections to improve the quality of data, but this remains far from an accurate asthma diagnosis. Other reasons including parental understanding can influence the accuracy of reported wheeze in children (115). However, asthma is growingly recognized as a syndrome and childhood wheeze itself a phenotype associated with asthma and lung functions in later life. More so these epidemiological questions have been widely used in epidemiological studies, for example initially employed in large global studies like ISAAC (116, 117).

1.3.2 Defining and measuring the exposure

In large human observational studies, it's a challenge to clearly define and quantify the exposure of interest. For example, smoking exposure in most epidemiological studies is based on the responses from participants which can lead to reporting bias. Sometimes an objective measurement can be used to limit the reporting bias; Cotinine is a biomarker that can be used to objectively define nicotine exposure levels for smoking exposure. However its value is variable depending on cutoff levels used, tissues studied in (saliva, urine or serum) and the population studied (118). Other challenges related to exposures for analysing and interpreting association studies is confounders. Confounding means distortion of the association between an exposure and outcome by a different variable (could be exposure) called a confounder. For example, in association studies for asthma the effect of smoking on asthma outcome can be confounded by socioeconomical status of the participant, or association of use of paracetamol with asthma can be confounded by chest infections in childhood which are associated with development of asthma and also with paracetamol use (119, 120). In this thesis, smoking exposure is investigated as the main exposure in results chapter, hence will be discussed in further detail.

Tobacco smoking is known to be harmful to human general health. Smoking exposure is associated with asthma; early life exposure with development of childhood asthma, and personal smoking is associated with severity of asthma in adulthood (121). Smoking as a known exposure has been studied in epidemiological studies for a long time, however there is still debate about the validity of reported smoking exposure. Epidemiological studies capture this exposure mainly through questionnaires, and this self-reported smoking exposure may not be accurate. The sensitivity and specificity sometimes varies based on the population studied, for example self-reported smoking is known to be more underestimated in pregnant mothers (122) and COPD (chronic obstructive pulmonary disease) (123) populations compared to the general population. This is likely due to the social pressure and expectations which might

lead to participants concealing their smoking behavior. Nicotine is a chemical in tobacco, its main metabolite cotinine is used as an objective biomarker of smoking exposure (124) for both personal and passive smoking exposure. Cotinine can be measured in different samples like saliva, urine and blood. Due to its simplicity and noninvasive nature, urine cotinine is commonly used in large epidemiological studies. The accuracy of cotinine measurement as a marker of smoking exposure again depends on the type of sample, timing of the sample collection and there is still ongoing debate about the cutoff levels (118).

The accuracy of self-reported smoking in pregnancy can be affected further by the educational level and age of the participant (122, 125). These exposures, like socio-economic status and education levels, are associated with smoking exposure and hence can be confounding variables. These can affect the association of smoking exposure with the outcome.

Smoking exposure patterns vary in the population and with time and are sometimes challenging to accurately define and measure. Tobacco smoking exposure is generally classified into active smoking (also termed as personal smoking) where smoking exposure is through smoking cigarettes, cigars or pipes by the individual, and passive smoking, where a person who does not smoke her/himself is exposed to tobacco smoke from other persons smoking, usually at home or work. Passive smoking is also termed "secondhand smoking". Passive smoking exposure can be significant, particularly in early childhood if parents smoke, and is a risk factor for developing asthma (126) and chest infections (127). Passive smoking exposure is also associated with asthma exacerbations and severity of asthma (128). Variation in smoking behavior adds another layer of challenge in quantifying smoking exposure; for example daily or intermittent smoking patterns; measuring smoking exposure in intermittent smokers can be more challenging (129).

1.3.3 Reverse Causation

Reverse causation means an exposure studied could be affected by the outcome (disease) itself; for example, children with a family history of allergic diseases or with early presentation of allergic diseases like eczema are usually breastfed and for longer duration and this can affect the study looking at the effect of breast feeding on asthma in childhood. Observational studies also face another challenge called selection bias; this occurs when those in the target population who have both the exposure and outcome (or neither the risk factor nor the outcome) are more likely to be enrolled in the study or remain in the study than others (120).

In summary, interpretation and generalisation of epidemiological association studies of asthma have many challenges; characterisation of the outcome, defining the exposure, correcting for confounders, avoiding reverse causation and selection bias effects are some of the major factors to consider (119).

1.4 Multigenerational effects of environment on phenotype

It is known that environmental exposures impart effect on development of a disease; this is more so in a condition like asthma where exposures play a very significant role in development of the disease. Going further the question is can the effect of the environmental exposure be passed on to subsequent generations? This phenomenon is termed as multigenerational or transgenerational effects. Some of the examples of such multigenerational effects of environment are discussed here.

1.4.1 In invertebrate and vertebrate model organisms

There is good evidence that the effects of environmental exposures can be transmitted through multiple generations. For example, maternal high sugar caloric intake in *Drosophila* has been found to affect body composition and metabolism of at least two generations (130). Similar evidence is seen in mammals; in mice, early life traumatic stress in the paternal line has effects on behavioural and metabolic responses in progeny (131).

1.4.2 For asthma in animals

Using rat models Rehan et al. first showed that effect of nicotine exposure (F0) during pregnancy (surrogate to prenatal smoking exposure) on asthma, can be passed on to the F2 offspring (132), thus providing evidence for an intergenerational effect. Subsequently, using the same rat model, further studies showed that there was transgenerational transmission of the asthma phenotype to F3 offspring following perinatal nicotine exposure of F0 dams.

1.4.3 In humans

Some of the strongest evidence for transgenerational effects of environmental exposures in humans have been demonstrated in the Overkalix cohort; in this study paternal grandfather's food supply in pre-adolescence was linked to the mortality risk ratio of grandsons, while the paternal grandmother's food supply

was linked to the mortality risk ratio of the granddaughters (133). Furthermore, a surfeit supply of food during slow growth phase was associated with increased relative risk for mortality in grandchildren (134).

1.4.3.1 For asthma in humans

Trans-generational effects of tobacco smoking exposure were examined in a retrospective case-control study within the Children's Health Study in California. Grandmother smoking during pregnancy (maternal prenatal smoking exposure) was a risk factor for asthma in the child, even with no maternal smoking during pregnancy (child's prenatal smoking exposure) with odds ratio (OR) of 1.8 (95% CI 1.0 – 3.3). The risk was higher when both mother and child had *in utero* smoking exposure (both grandmother and mother smoking) with OR of 2.6 (95% CI 1.6 – 4.5)(135). However, intergenerational effects from maternal prenatal exposure were not observed in other studies (136), though observation of paternal prenatal exposure associated with asthma in granddaughters was observed (136). More recently, a large epidemiological study in the Swedish population did demonstrate an increased risk of asthma in the first 6 years of life for individuals whose grandmothers smoked during early pregnancy, independent of maternal smoking (137).

Terms like multigenerational and transgenerational effects are sometimes used interchangeably, Table 1-1 describes the terms multigenerational, intergenerational and transgenerational adopted from Arshad et al (138)

Table 1-1 Terms used in describing multigenerational inheritance
(adapted from (138)

Term	Definition
Multigenerational inheritance	Transmission of epigenetic information through the germline across generations in the absence of any direct environmental exposure or genetic manipulation
Intergenerational inheritance	Maternal exposure (F0) has direct effects on the developing fetus and potentially on the germline of the foetus, leading to altered phenotype of the child (F1) and grandchild (F2)
Transgenerational inheritance	Effect on subsequent generation persists in absence of the possibility of direct environmental exposure, such as effects persisting to the greatgrandchild (F3) on the maternal line or effects of prepubertal paternal or grandpaternal exposures in the male line

1.5 Genetics of asthma

Parental history of allergic disease has been the risk factor most consistently associated with allergic diseases. Findings of parental history increasing the risk of allergy and asthma in children suggest that genetics plays an important role in the etiology of asthma. This is further supported by the increased concordance of asthma in monozygotic twins compared to dizygotic twins (139). Asthma is a "complex genetic disease" or "complex trait". In complex genetic disease, in contrast to Mendelian disease, multiple genes in conjunction with environmental exposures are involved in determining risk of the disease, each of which may have only a small effect on the phenotype. Many genes involved in complex diseases like asthma display interaction with each other, with genetic interaction having bigger effect on the phenotype than individual gene effects. There have been numerous genetic studies leading to identification of many asthma genes; around 64 genes were reviewed in 2003 and by 2008 there were around 100 genes identified and the number has been increasing with high frequency (140-142).

One of the methods used to identify genetic variants associated with asthma is the candidate gene approach, which is hypothesis driven, where the genes studied are known to be involved in the pathophysiology of asthma. The 5q31 locus has been linked to asthma and includes the gene encoding IL-13 (*IL13*) (140). Some of the other candidate genes linked to asthma are *IL4RA*, Cysteinyl leukotriene receptor 1 (*CYSTR1*) (140) and *CD14 (143)*.

In contrast a Genome-wide association study (GWAS) is a hypothesis free approach where variants throughout the genome are examined and the those significantly associated with disease or phenotypes are identified. GWASs are based on the hypothesis of "common disease, common variant" which postulates that alleles that are common in the population (>1-5%) will be associated with the common disease phenotype (144, 145). The first GWAS study of asthma in a UK family cohort and German case-control sample showed strong association of single nucleotide polymorphisms (SNPs) in the 17q21 region with physician diagnosed asthma and transcript levels of *ORMDL3* (146). Another GWAS in a

well characterised asthma population showed polymorphisms in *RAD50* region on chromosome 5q31.1 and HLA-DR/DQ region on chromosome 6p21.3 were associated with asthma susceptibility (147). Subsequently many other genes have been associated with asthma using GWAS including *PDE4D* (148), *DENND1B* (149), *TLE4* (150), *IL18R1, IL33, IL2RB* (151).

Many asthma genes have been identified, however the outcome phenotypes need to be better characterised (141), studies need replication in other population samples and demonstration of the mechanistic path showing the role of the identified genetic variant / genes in asthma pathogenesis is needed (140). Recently GWAS for asthma in a large population sample (41 926 self-reported asthma cases and 239 773 controls) from UK Biobank identified 75 risk loci associated with asthma including 15 new loci not identified in previous GWAS studies (152).

Each gene associated with a complex disease such as asthma is likely not have a big effect on the phenotype, but it is plausible that multiple genes interact with each other and interaction of these multiple asthma genes leads to the more conducive background for the development of asthma (52) particularly upon exposure to the right environmental factors.

1.5.1 Missing Heritability

Missing heritability is the proportion of heritability of the phenotype not explained by known genetic variants. There have been more than 25 genome wide association studies in asthma and hundreds of genes have been identified linked to asthma in these and candidate gene studies (153, 154), however most of the genes identified for complex diseases confer only small increments in risk (155) and therefore the heritability of asthma is not well explained by all the identified genes (153, 156). Many reasons have been proposed for the missing heritability in complex diseases like asthma; not yet detected common variants with small effect on phenotype, rare variants with larger effect and structural variants in the genome all of which are not captured well by SNP arrays typically

utilized for GWAS studies (155). Non-coding inheritance and environmental influences could also add to the explanation of variation in phenotype. Epigenetic markers that can be established by environment and transmittable to one or more generations could contribute to the heritability in complex disease (156) particularly disease such as asthma and allergies which are strongly associated with environmental exposures.

1.5.2 Gene-environment interactions

As outlined above, asthma is a complex disease involving the effects of both multiple genes predisposing the risk for development of asthma and environment exposures particularly at critical periods of life. In complex diseases maternally mediated environmental modulation of gene expression in offspring may be more important than purely heritable genetic risk (135) as seen in the twins study of birth weight and osteoporosis (157). Impact gene-environment interactions on development of asthma has been focus of research for more than a decade (158). With the right genetic predisposition in the background, appropriate environmental exposures at a critical time window can lead to development of asthma. Many interactions of between a range of environmental exposures and asthma candidate genes have been reported (159, 160) and evidence for tobacco smoke exposure and genetic interactions is outlined below.

Tobacco smoke exposure and genetic interactions

Tobacco smoke exposure is associated with asthma (63-66, 110). Interaction between *CD14* polymorphism and smoking exposure has been shown to be associated with total IgE in Caucasians and baseline FVC and IgE in Latino population (161). *IL1RN* (interleukin receptor 1 antagonist gene) polymorphisms haven been linked to asthma in the presence of prenatal smoking exposure (47) suggesting the interaction of prenatal smoking exposure and *IL1RN*. An interaction between ADAM33 and prenatal smoking exposure has been shown to be associated with reduced lung function and development of BHR in children

(162). A genome wide linkage screen showed genetic and passive smoking exposure interactions associated with BHR (162).

Gene-environment interaction has also been repeatedly demonstrated in the IOWBC; *ILRN* was not associated with asthma but interaction of an *ILRN* SNP (rs2234678) and maternal smoking during pregnancy was associated with childhood asthma (163). Similarly, *IL13* and prenatal smoking exposure interaction has been associated with asthma (164). In both the above examples the genotypes were not significantly associated but the gene-environment interactions were, suggesting that genetic predisposition needs an environmental trigger to lead to the disease phenotype.

1.6 Epigenetics

The term “Epigenetics” was coined by Waddington in 1940s and is used to describe stable (through mitosis) changes in gene expression, which cannot be explained by changes in DNA sequence itself (165, 166). Originally these observations were noted while studying organisms such as Drosophila. This description suited Waddington`s original question - to explain how a single genome could lead to development of multicellular organism with cells of varied phenotype. Later he proposed that; “We certainly need to remember that between genotype and phenotype, and connecting them to each other, there lies a whole complex of developmental processes. It is convenient to have a name for this complex: ‘epigenotype’ seems suitable” (166).

Coded genetic information alone is not enough for development and maintenance of different tissues/organs of human body. All the cells in a blastocyst will have same coded genetic information, but for each cell to differentiate into different cell types with unique function more information, which is not included in the coded DNA is needed. All this information that is not coded in DNA but maintained through mitotic and cell meiotic divisions is termed as epigenetic. This non-coded information is not only important for development but also essential in maintaining the expression profiles of different cell types. Thus, epigenetics can be defined as “Heritable (mitotically or meiotically) changes in gene expression that do not involve change in the DNA sequence”. The epigenome refers to the entire epigenetic marks on the genome; these marks regulate chromatin structure and accessibility of the DNA to the machinery regulating gene expression (167).

1.6.1 Mechanisms of Epigenetic programming

The three main molecular epigenetic processes are histone modification, microRNA, and DNA methylation. Figure 1-1 describes the mechanisms involved in the epigenetic regulation of gene expression. Epigenetic modification modulates the rate of transcription to messenger RNA (mRNA) by controlling the

access of transcription factors to the DNA sequence. Active chromatin is in a more open state permitting access to transcription factors and the CpG sites of the promoter region are generally unmethylated, allowing for the binding of transcription factors. Inactive chromatin is more in a closed structural state and promoter CpG methylation leads to decreased binding of the transcription factors. Further epigenetic control is provided by microRNAs which reduce the rate of protein synthesis by binding to complementary sequences in the 3′ end of mRNA (168).

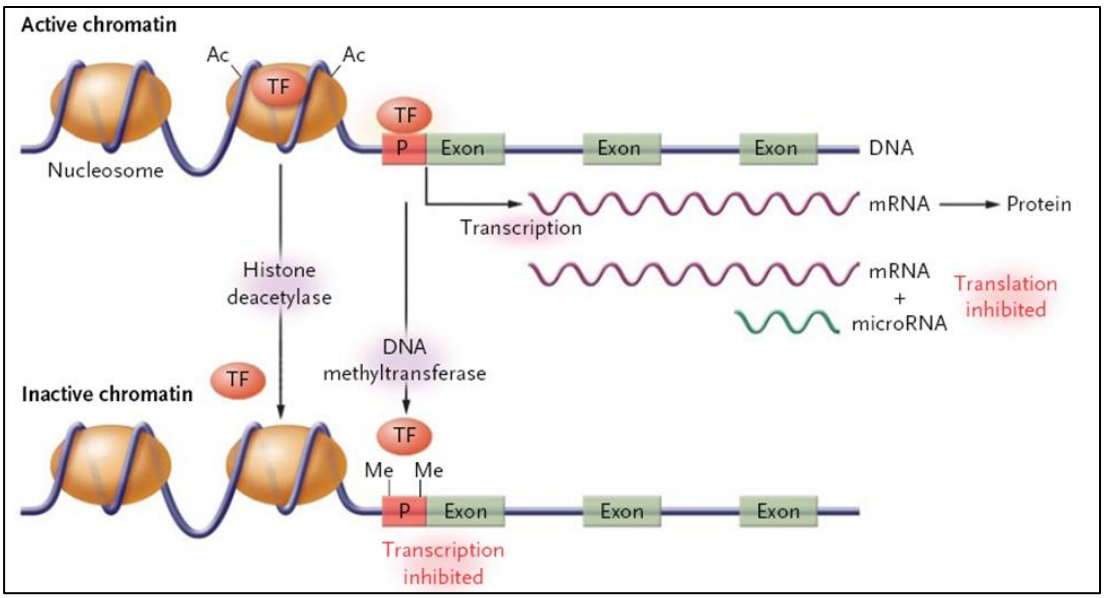


Figure 1-1 Regulation of Gene Expression through Epigenetic Processes
(Reproduced with permission from (168), Copyright Massachusetts Medical Society.)

Histone modification

Chromatin is the protein structure around which the DNA molecule is wrapped. The building block of the chromatin is nucleosome which is an octomer of eight histone molecules. N terminals of the histone molecules are modifiable by acetylation, methylation, phosphorylation and ubiquitination (169). These modifications control gene expression determining the phenotype and are reversible

Micro-RNAs (miRNAs)

miRNAs are short noncoding RNAs which regulate gene expression by binding to the 3' untranslated regions (3' UTRs) of messenger-RNA (mRNA) and suppressing protein expression post-transcriptionally (170). One miRNA can target multiple mRNAs and one target mRNA can be influenced by multiple miRNAs (171). miRNAs are generated from longer transcripts encoded in the genome (pre-miRNAs), which themselves can be epigenetically regulated.

These different epigenetic mechanisms act conjointly influencing the phenotype (171). However, while histone modification can be transmitted through mitosis, it is only DNA methylation that has been shown to be stable for long term and is easily measured at scale in epidemiological cohorts. Given this, the focus of this work will be on DNA methylation as an epigenetic mechanism.

1.6.2 DNA methylation

DNA methylation was the first recognised epigenetic mechanism (172) and refers to the reversible methylation of DNA at CpG (cytosine phosphate guanine) dinucleotides. The DNA molecule is modified by covalent methyl residues at 5' terminal of the cytosine in CpG dinucleotide (173). Generally methylation leads to gene silencing, most of the mammalian CpG sites (60% to 90%) are methylated (174) and some unmethylated CpG sites are concentrated in the regions termed "CpG islands", which are usually seen in the promoter regions of the gene (172). The DNA methylation state in the promoter region determines the gene expression state, with generally hypermethylated DNA leading to silencing of gene expression. There is now growing evidence that CpG sites away from CpG islands may also be important in regulating gene expression and these regions are termed as "CpG island shores" (175).

DNA methyltransferases (DNMTs) are the enzymes that catalyse DNA methylation. There are different forms of DNMTs performing two functions; maintenance of established methylation and *de novo* establishment of methylation (174). DNMT1 is involved in maintaining methylation during DNA

replication, after replication the daughter strand of the DNA is hypomethylated (only one strand is methylated) and DNMT1, having high preference for hemimethylated DNA, methylates the daughter DNA strand. DNMT3a and DNMT3b facilitate *de novo* methylation of CpGs and DNMT3L is not active and co-localizes with DNMT3a and DNMT3b (176). No action for DNMT2 has been clearly established, knockout of DNMT2 in mice has no effect on methylation suggesting no role for DNMT2 in DNA methylation (177).

There are two principle ways hypermethylation of DNA results in gene silencing. Firstly, is by direct interference in the recognition element of a transcription factor by a methyl residue. Secondly, indirectly by attracting methylated-DNA binding proteins (169, 178) like MeCP2 (methyl CpG binding protein 2), MBD (methyl CpG binding domain protein) 1, MBD2, MBD3 and MBD4, leading to chromatin remodeling.

1.6.2.1 Measurement of DNA methylation

DNA methylation can be assessed using different methods. The methods fall under three mechanistic categories: 1) Immunoprecipitation, where the methyl binding proteins or antibodies to methylated cytosine can be used to enrich methylated segments of DNA; 2) DNA cleavage using methylation sensitive restriction enzymes (endonucleases); and 3) reading C to T genotyping of bisulphite converted DNA (179, 180). In this thesis the Illumina Infinium HumanMethylation450 BeadChip (Illumina, Inc., CA) is utilised which is based on discriminating C to T transition in bisulphite treated DNA.

Each of these different methods has relative merits and de-merits in comparison to each other and a review Peter Laird has provided a comprehensive comparison of these different methods (181). Methods based on endonuclease cleavage usually require very high quality and high quantity of DNA and cannot cover all CpGs in the whole genome. Immunoprecipitation based methods need high quantities of template and the ratio of the reagent and input DNA sample needs careful attention, these methods are usually poor for single base level resolution. Both the above methods are more demanding in terms of costs, reagents and

labour. Illumina Infinium arrays, which are bisulfite treatment based, are compatible with degraded and low quantity of DNA, and has a favourable ratio of genome coverage and sample throughput. However, this method has technical issues including bias from incomplete bisulphite conversion and hybridisation bias. With relatively less labour and cost demanding, rapid processing and high sample throughput Illumina Infinium is increasingly used by researchers working on population-based studies. Laird reports in the review that bisulphite based methods are accurate and reproducible and a recent report demonstrated that the Infinium Human-Methylation450 array had strong reproducibility and high validity (182), with comparable methylation results for Infinium HumanMethylation450 when compared to whole genome bisulphite sequencing. The Infinium HumanMethylation450 array (Infinium Methylation 450K; Illumina, Inc. CA, USA) interrogates >484,000 CpG sites associated with approximately 24,000 genes.

The principle for bisulphite treatment of the genomic DNA is that sodium bisulphite chemically deaminates unmethylated cytosine much more rapidly than methylated cyotosine. Bisulphite treatment converts unmethylated cytosine to uracil and methylated cytosine (C) is mostly unchanged (183, 184). Uracil and thymine have similar hybridization behavior. So using bead arrays for genotyping assay for C/T polymorphisms, methylated and unmethylated sites can be identified (185).

1.6.3 Environment and DNA Methylation

DNA methylation can be altered by the environmental exposures leading to change in gene expression and hence disease susceptibility. In monozygotic twins (that have same coded genome) DNA methylation profiles differ with increasing age, which can be plausibly explained by different environmental exposure leading to differences in DNA methylation (186). Initial evidence of DNA methylation mediating the effect of environmental exposure on phenotype came from the agouti mouse model; Jirtle and colleagues and Cooney and colleagues work with the agouti mouse model revealed that dietary supplementation with

methyl donors (folic acid, vitamin B12 and other agents) during pregnancy influenced the coat colour of the agouti mouse and this was mediated through the methylation of CpG cites in the upstream of the agouti gene (187, 188). DNA methylation patterns can be programmed in early life and possibly, by a balance of methylation and demethylation, these patterns can be maintained through life. DNA methylation plays an important part in epigenetic programming during early life, providing a conduit for environmentally mediated changes to be retained well into adulthood.

One of the plausible mechanisms for the developmental origins hypothesis is the environmental exposures influencing the epigenome which controls gene expression. Studies on consequences of prenatal exposure to famine provided the first evidence that transient environmental conditions in human gestation can be recorded as persistent changes in the epigenome; people who were exposed to prenatal famine during Dutch Hunger Winter (1944-45) showed lower methylation of IGF2 gene 6 decades later (189) . The epigenome can serve as the *molecular archive of past environmental conditions* (190), which can be chemical or social. Mouse Agouti viable yellow gene studies show that epigenetic changes secondary to early life chemical exposure persist throughout life and have life-long phenotypic consequences (191). Maternal care in the rat has also been shown to induce programming of expression of the glucocorticoid receptor gene in the hippocampus that lasts well into the rat's maturity (192). Epigenetics is proposed to be the "*mechanism that interfaces between nurture and nature*" (169), nature representing the genetic background and nurture referring to the environment.

1.6.3.1 Tobacco Smoking and DNA methylation

Smoking can modify DNA methylation, as demonstrated by Launay et al; the promoter region of the monoamine oxidase B gene (*MAOB*) showed differential methylation among smokers (current and ex-smokers) compared to never smokers; methylation being higher in never smokers ($p<0.0001$) (193). This differential methylation seen even in ex-smokers (even after >10 years of

stopping smoking) highlights the long-term effects of the smoking on DNA methylation. Tobacco smoke exposure has been shown to modify DNA methylation of the lung cancer genes and these methylation marks are being investigated as targets for the lung cancer treatment (194). One of the proposed mechanisms is oxidative stress caused by smoking exposure interfering with DNMT`s binding to the DNA leading to hypomethylation (172, 195).

In a subset of 348 participants of Children's health study from Southern California, the effect of prenatal smoking exposure on DNA methylation was studied. By CpG loci screening 8 differentially methylated genes were identified out of which 2 genes (AXL receptor tyrosine kinase (AXL) and protein tyrosine phosphatase receptor type O (PTPRO)) were validated by pyrosequencing (196). An epigenome-wide scan of cord blood DNA in a US cohort, demonstrated methylation of different genes (AHRR, CYP1A1 and GF11) to be associated with maternal smoking during pregnancy (197).

A recent EWAS meta-analysis for maternal smoking during pregnancy by the Pregnancy And Childhood Epigenetics (PACE) consortium across 13 cohorts (n = 6,685) in newborns showed 6,073 statistically significant CpGs and 3,932 remained statistically significant after adjustment for cell type proportion (198). It is evident that smoking exposure can affect DNA methylation but it is not clear if these genes (differentially methylated) play any role in asthma (171). Further epigenetic studies focusing on asthma are needed to understand if the smoking exposure association with asthma is mediated through DNA methylation.

1.6.3.2 Diet and medication use

In the agouti mouse model, lack of methyl donors in the diet during pregnancy was a risk factor for many complex diseases in the offspring, and this risk was established through changes in DNA methylation (199). The impact of diet and medications on asthma related inflammation or on airways, if mediated through DNA methylation is not clearly understood (171).

1.6.4 Genotype dependent DNA methylation (DNA methylation and genetic interaction)

DNA methylation can vary at specific CpG sites depending on genetic variants (200, 201). DNA methylation mediates environment and genetic interactions, but in some specific genes environmental impact on DNA methylation could be dependent on genetic variants. For example, the effect of *in utero* smoking exposure on DNA methylation of long interspersed repetitive element-1 (LINE1) was only seen in the presence of common glutathione- S-transferase mu 1 (GSTM1) null genotype (196). Epigenome-wide study of DNA methylation profiles in ageing and age-related phenotypes (including lung function) has shown that in specific genes, DNA methylation varied depending on genetic variants. It was concluded that genetics and DNA methylation may impact on age related phenotypes, independently or by interaction with each other, with DNA methylation likely mediating genetic and phenotype association (200). These results suggest a significant role for genotype dependent DNA methylation in specific genes on disease status.

1.6.5 Trans-generational inheritance

As detailed in the section 1.4, the effect of environmental exposures can persist through multiple generations both in animals and humans. One of the plausible mechanisms for the effect of the prenatal environmental exposure of the mother transmitted to grandchildren could be that prenatal exposures through mother leaving marks on the fetal germ cells which will become ova, through which these marks will be passed onto the grandchild. As this is due to the direct effect of the environment on the germ cells that from the grandchild, it is not true transgenerational inheritance. However it is plausible that environmental exposures leave a mark in the epigenome which are then trans-generationally inherited, i.e. passed through meiosis in the absence of direct exposure (156, 202) and as discussed in section 1.4 there is some evidence for this in animals (132). Studying similar epigenetic inheritance in humans is challenging due to the long life cycle, difficulties in accurate measurement of exposures and characterisation

of the phenotype. A prospective multigenerational cohort can provide the platform for such multigenerational studies.

In summary, asthma is influenced by genotype and environmental exposures. The latter is known to influence DNA methylation, and DNA methylation has potential to explain some of the missing heritability of asthma. Thus DNA methylation is a growing focus of research in asthma (203). DNA methylation can be interrogated at hundreds of thousands of sites in multiple samples with new array platforms which makes it a practical epigenetic mechanism to study in epidemiological scale studies. At the same time there are challenges in studying DNA methylation, such as tissue specificity of epigenetic profiles, need for temporal DNA methylation assessment or else it is difficult to know if the disease is the result of DNA methylation changes secondary to environmental exposures or the DNA methylation itself is secondary to the disease phenotype, and that current technological platforms only assess a small fraction of the CpG site in the genome. It is not yet clear about the level of methylation at which gene expression is switched off or on. The same genes can be either hypermethylated or hypomethylated in different cells in the same tissue. For example, the asthma susceptibility gene ADAM33 was found to be hypermethylated in epithelial cells but hypomethylated in fibroblasts of asthmatics (204). Further merits and challenges of DNA methylation studies in asthma including the effect of cellular composition of the tissue being studied and sample size of studies are further discussed in result chapters and in chapter 7 (General Discussion).

1.7 The Isle of Wight Birth Cohort

The Isle of Wight is the largest island in England, located on the south coast of England separated from the mainland by the Solent and measures 380 square kilometers. The island has a reasonably stable population with net migration of just above 1000 in 2008. This is secondary to lack of a direct land link to UK or the rest of Europe, restricting free movement of people. The whole population is 140,200 out of which 16,000 are aged between 15 to 20 years, as reported in 2008 (205). The majority of the population is of Caucasian ethnicity.

The Isle of Wight birth cohort (IOWBC) was established first in 1989/90 by Drs David Hide and Hasan Arshad. This is a prospectively recruited birth cohort; children born between January 1, 1989, and February 28, 1990, were recruited. Parents were approached and the study explained; 1536 babies were enrolled at birth after obtaining parent's consent. This birth cohort has contributed enormously to the understanding of prevalence and identifying environmental and genetic risk factors for asthma and other allergic diseases (48, 49, 164, 206-212).

Cohort children have been followed up at 1, 2, 4, 10 and 18 years of age. Information on environmental exposures of the mother during pregnancy was collected at birth. Prospective information of environmental exposures of children, parental and child disease conditions was updated at each follow up. Information on maternal environmental exposures during pregnancy, and family history of allergic diseases were collected. Skin prick testing was carried out at 4, 10 and 18 years. Bronchial provocation testing with methacholine was undertaken at the 10 and 18-year follow-ups. At the 18-year follow up, exhaled nitric oxide (FeNO) and sputum induction were carried out in a subset of the cohort.

A summary of the cohort and the follow-ups is described in Table 1-2. The IOWBC is whole population based which limits selection bias and is also prospective, thus reducing recall-bias. The follow-ups have been successful with good follow up

rates ranging from the lowest of 80% at 2 years and the highest of 94% at 10 years.

The follow up at 18 years was undertaken between 2007 and 2009. The IOWBC had total participants of 1536 recruited at birth, out of which 1456 were available (excluding participants who had dropped out or died) for follow up at 18 years. 90.2% (1313/1456) of the available participants were followed up at 18 years.

Table 1-2 Summary of the Isle of Wight birth cohort

Birth (January 1989 – February 1990): Family history of allergy, household pets, parental smoking, socioeconomic status, gestational and parturient factors, cord blood IgE, maternal IgE	N = 1536
Age 1 year: Asthma and other allergic manifestations (physician diagnosis, symptoms and treatment), chest infections, household pets, parental smoking, method of feeding	86.9% (1316/1514)
Age 2 years: Asthma and other allergic manifestations, chest infections, household pets, parental smoking, method of feeding	79.8% (1205/1510)
Age 4 years: Asthma and other allergic manifestations, SPT, chest infections, household pets, parental smoking, housing characteristics (age, dampness, method of heating)	83.7% (1218/1456)
Age 10 years: Asthma and other allergic manifestations (additional: ISAAC questionnaire), SPT (n = 1,036), serum IgE and blood sampling for gene analysis (n = 945), baseline spirometry (n = 981), bronchial provocation test (n = 784), household pets, parental smoking	94.3% (1373/1456)
Age 18 years: Asthma and other allergic manifestations (ISAAC questionnaire, symptom questionnaire), SPT (n = 842), baseline spirometry (n = 839), bronchial provocation test (n = 585), Exhaled nitric oxide (n = 834), Sputum inductions (n=100, 80 good samples), serum IgE and blood sampling (n = 567)	90.2% (1313/1456)

Percentage determined with N = Total study participants available for follow-up, n = number of participants enrolled at each follow-up. SPT- skin prick test.

The IOWBC has contributed to understanding the prevalence, environmental risk factors, natural history and genetic factors of asthma up to the age of 18 years (213). Table 1-3 summarises the factors associated with different phenotypes related to asthma in the IOWBC.

Table 1-1-3 Factors associated with asthma and quantitative traits of asthma in the Isle of Wight birth cohort

Factor	Outcome	Effect (Risk or Protective)	Reference
Male gender	Wheeze at 4 yrs Currently diagnosed asthma at 10 yrs	risk	(214, 215)
Maternal asthma	Early onset persistent wheeze (4-10 yrs) Late onset persistent wheeze (>4-10 yrs) Wheeze at 10 yrs Currently diagnosed asthma at 10 yrs Symptomatic BHR at 10 yrs	risk	(214-217)
Paternal asthma	Late onset persistent wheeze (>4-10 yrs) Wheeze at 10 yrs Currently diagnosed asthma at 10 yrs	risk	(215, 216)
Sibling asthma	Late onset persistent wheeze (>4-10 yrs) Currently diagnosed asthma at 10 yrs.	risk	(215, 216)
Family history of asthma	Undiagnosed wheeze at 18 yrs.	risk	(107)
Prenatal smoking	Wheeze at 4 yrs. Early onset persistent wheeze (4-10y)	risk	(214)
Parental smoking at 1 yr.	Currently diagnosed asthma at 10y	risk	(215)
Parental smoking at 2 yrs.	Early onset persistent wheeze (4-10y)	risk	(216)
Parental smoking at 4 yrs.	Wheeze at 10 yrs. Symptomatic BHR	risk	(215, 217)
Low birth weight	Wheeze at 4 yrs.	risk	(214)
Cord serum IgE	Asthma at 10 yrs. Allergic sensitisation at 4 & 10 yrs.	risk	(49)

Factor	Outcome	Effect (Risk or Protective)	Reference
High social class at birth	BHR at 10 yrs.	risk	(215)
High social class at birth	Early onset persistent wheeze (4-10 yrs.)	protective	(216)
Breastfeeding for ≥4 months	Lung function at 10 yrs.	improved FVC, FEV_1 and PEF	(100)
Chest infections at 1 yr.	Early onset persistent wheeze (4-10 yrs.) Currently diagnosed asthma at 10 yrs.	risk	(215, 216)
Chest infections at 2 yrs.	Early onset persistent wheeze (4-10 yrs.) Wheeze at 10 yrs. Currently diagnosed asthma at 10 yrs.	risk	(215, 216)
Cat ownership	Early onset persistent wheeze (4-10 yrs.)	protective	(214)
Atopy at 4 yrs.	Wheeze at 10 yrs. Currently diagnosed asthma at 10 yrs. AHR at 10 yrs.	risk	(215)
Smoking at 18 yrs.	Undiagnosed wheeze at 18 yrs.	risk	(107)
Paracetamol use at 18 yrs.	Undiagnosed wheeze at 18 yrs. Adolescent onset asthma	risk	(107, 218)

Results extracted from published data on exposures and asthma outcomes in the Isle of Wight birth cohort. yrs - years of participant age, IgE; Immunoglobulin E, BHR; Bronchial hyperresponsiveness, FVC; Forced vital capacity, FEV_1; Forced expiratory volume in 1 second, PEF; Peak expiratory flow.

1.8 Hypotheses and objectives

In summary asthma is a syndrome with spectrum of multiple phenotypes and heterogeneity in susceptibility with genetics and environment playing varying roles in development of asthma. Epidemiological studies have explored and confirmed many environmental exposures associated with development of asthma and genetic studies in asthma, particularly genome wide association studies, have identified many genes linked to asthma but still both environmental and genetic factors fail to explain all the factors associated with development of asthma and the complex heritability of asthma. Epigenetics is the study of heritable changes in gene expression without any changes in the coding sequence of DNA. One of the recognized and better understood epigenetic mechanisms is DNA methylation, which is also amenable to modification by environmental exposures and regulates gene expression. It is plausible that DNA methylation mediates the gene-environment interactions and the multigenerational effects of environmental exposure.

DNA methylation can help to understand the missing heritability of asthma and to shed light on the biological mechanisms linking environmental exposure with susceptibility to disease. Parental asthma is a known risk factor for development of asthma in offspring. This risk of asthma in children is known to be higher for maternal asthma than for paternal asthma (219, 220). Maternal asthma is also a risk factor for early persistent wheeze (216) and asthma at 10 years in IOWBC (221). Epigenome-wide association studies on DNA methylation at birth associated with development of asthma show the association to be stronger in children with maternal asthma (222). Identifying DNA methylation profiles in cord blood associated with maternal asthma will add to the current understanding of the likely mechanism of asthma susceptibility.

1.8.1 Hypotheses

This thesis addresses four specific hypotheses:

Hypothesis 1. Environmental exposure, DNA methylation and genotype interact to influence asthma phenotype.

Hypothesis 2. Environmental exposure (prenatal smoking) alters risk of disease in multiple generations.

Hypothesis 3. Environmental exposure (prenatal smoking) influences DNA methylation across generations (intergeneration).

Hypothesis 4. Maternal history of asthma is associated with altered DNA methylation profile at birth in offspring.

Evidence to support or refute these hypotheses will be obtained through a combination of focussed candidate gene studies and epigenome-wide association studies as outlined below.

1.8.2 Objectives

1. To recruit the Third Generation Study, the children of the IOWBC. This will create a unique multigeneration cohort (across 3 generations) and thus facilitate study of development of asthma across generations (Chapter 4)

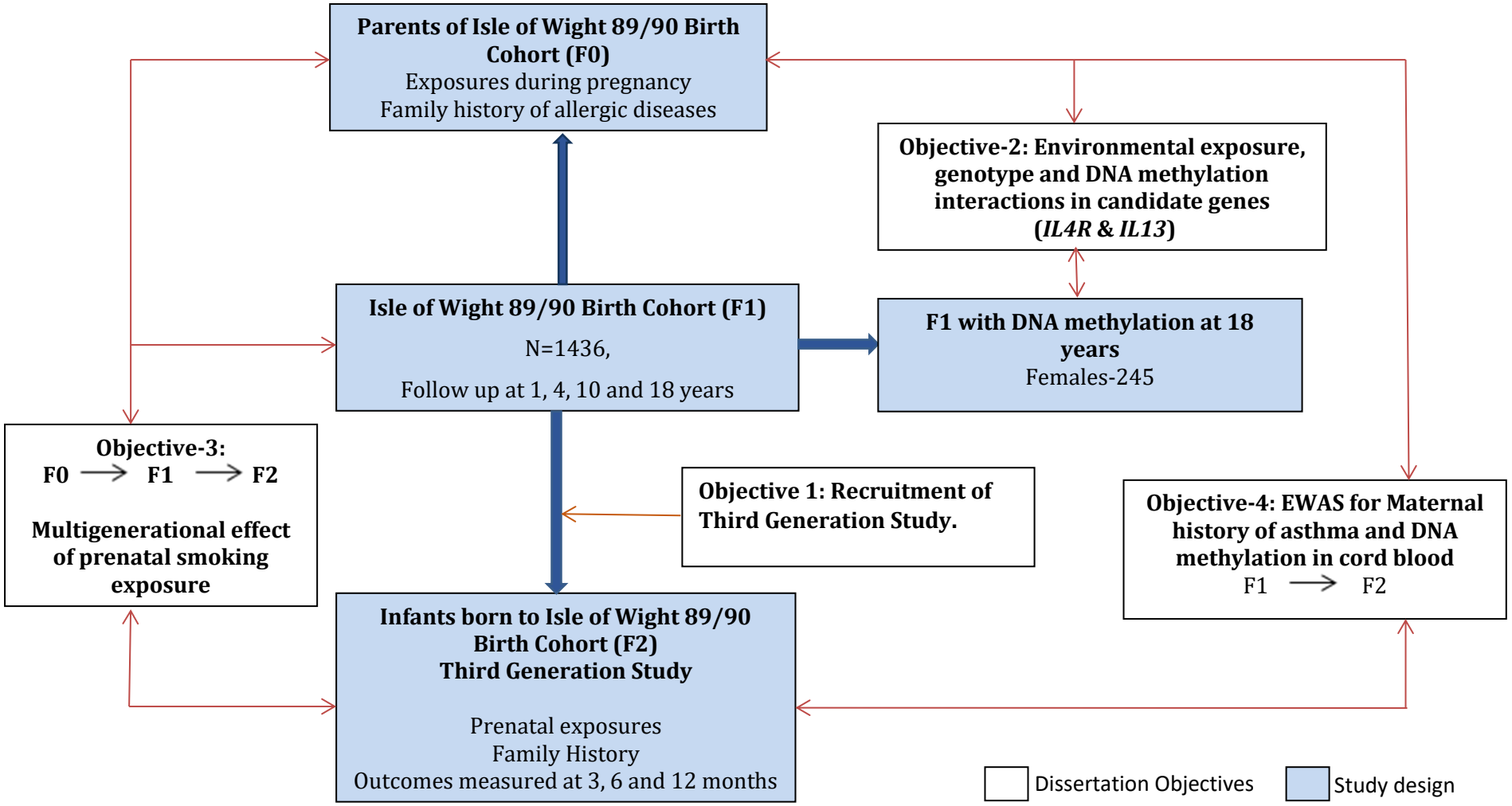
2. Hypotheses 1 is tested in two parts, first to explore the effect of the interaction of genotype and DNA methylation (*IL4R*) on risk of asthma (Chapter 3.2).
   In second part (Chapter 3.3) the objective is to identify CpG sites where the DNA methylation is influenced by interaction of genotype (*IL13*) and prenatal smoking exposure; this is to demonstrate that DNA methylation is influenced by both environment and genotype. Then to investigate if the identified CpG site(s) interact with genotype to influence asthma outcomes.

3. To test Hypothesis 2 and 3 using the multigeneration study (Chapter 5). To test Hypothesis 2 by investigating if smoking during pregnancy by grandmother (F0) and mother (F1) is associated with wheeze in grandchild (F2). Then Hypothesis 3 will be tested by exploring if maternal and grand-maternal smoking during pregnancy is associated with DNA methylation in cord blood of F2.

4. To test Hypothesis 4, an Epigenome-wide analysis (EWAS) of DNA methylation in cord blood of F2 generation in relation to maternal history of asthma will be carried out to identify differentially methylated loci based on maternal history of asthma (Chapter 6).

F0: parents of the Isle of Wight birth cohort, F1: Isle of Wight birth cohort participant and partner, F2: Third generation, children of F1.

The structure of the three generational study and how the objectives of this thesis are connected to the study are further depicted in Figure 1.2



F0: parent of Isle of Wight birth cohort, F1: Isle of Wight birth cohort participant and partner, F2: third generation/children of F1

Figure 1-2 Skeletal structure of the study design and the dissertation objectives

Chapter 2 General Methods

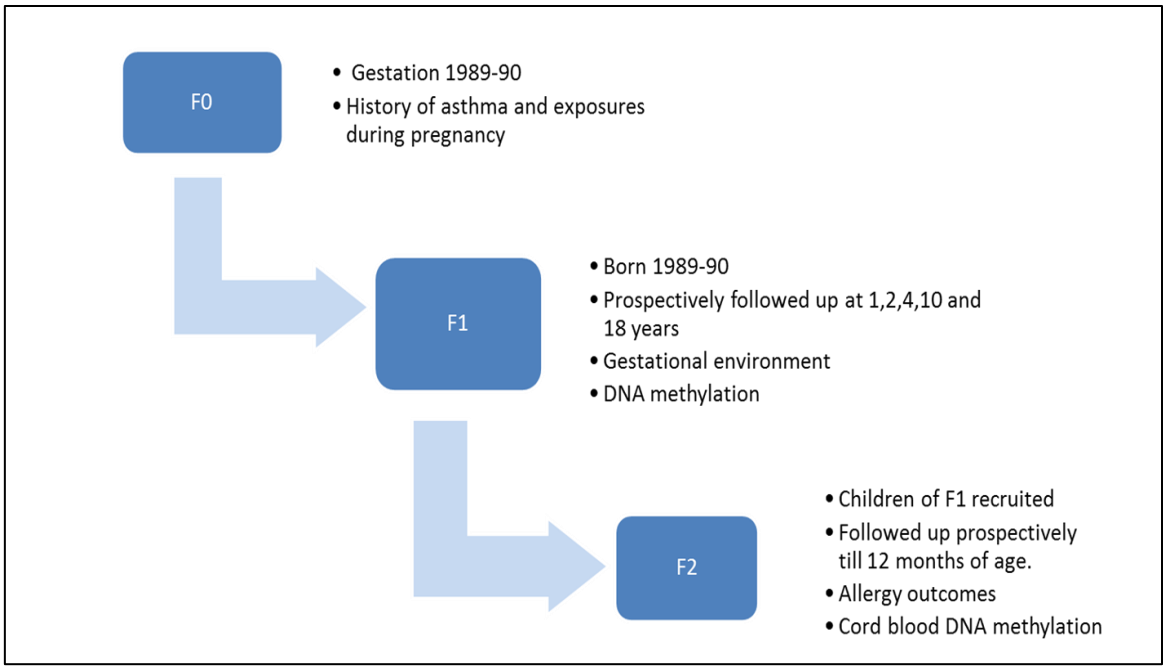
The main focus and the resource for this thesis is Third Generation Study which was established during the tenure of the research fellowship and recruitment is currently ongoing. Studies analysed and presented in chapters 4, 5 and 6 utilise the data from the Third Generation Study. Aspects relevant to specific result chapters are covered in the respective chapters. The thesis also includes data from the original IOWBC (including parents of the Third Generation Cohort) particularly from the 18 years follow up of the cohort. The gene-environment and epigenetic interaction study (chapter 3) is based on IOWBC 18 years follow up. General methodology of the Third Generation Study and IOWBC is described in this chapter.

2.1 Third Generation Study (F2)

2.1.1 Study design

The Third Generation Study (F2) is a prospective observational study involving establishment of a birth cohort of children that are born to participants of the IOWBC (F1). History of asthma, allergic diseases and exposures in parents of the Isle of Wight Birth Cohort (F0) have been collected. IOWBC participants (F1) themselves have been prospectively characterized for asthma, asthma related traits and allergic diseases.

The IOWBC is described in section 1.7. The recruitment and assessment of the Third Generation Study is described in detail below. Figure 2-1 summarises the multigenerational cohorts; IOWBC (F1) with prospective information from their parents (F0) and the Third Generation (F2)



F1; Isle of Wight birth cohort (1989-90), F2; third generation, children of F1, F0; parents of F1.

Figure 2-1 Third generation study design.

2.1.2 Study location

Clinical aspects of the study were based at the David Hide Asthma and Allergy Research Centre at St Mary's Hospital on the Isle of Wight. The DNA extraction was carried out in the Respiratory Genetics Laboratory, Faculty of Medicine, University of Southampton, UK.

2.1.3 Selection of study participants

General criteria for recruitment of Third Generation Study are:

Inclusion Criteria: Children of Isle of Wight Birth Cohort study participants, i.e., either mother or father is a member of the IOWBC.

Exclusion Criteria: There were no exclusion criteria except those who declined to participate in the study.

For Cord blood methylation studies (chapters 5 and 6) there was one pair of twins in the study which was excluded. Methylation changes in twins are

understood to become increasingly discordant with age which is likely due to different postnatal exposures (223). Given the exposures studied are prenatal exposures and the methylation was analysed at birth, twins were excluded in this study to avoid including two samples with shared prenatal exposures. Non twin siblings, for whom cord blood samples and questionnaire data were available, were included in the analysis. Non twin siblings do not share the same prenatal exposures. Studies of sibling pairs with discordant prenatal exposures have shown them to exhibit differential methylation profiles (224). To our knowledge, there were no participants with congenital abnormalities in the study subset.

Study subsets analysed in the results chapters of the study vary based on the focus of the study and the status of the recruitment at the time of the analysis. Relevant specific criteria for selection of subsets are described in detail in the methods section of the respective result chapters.

2.1.4 Recruitment of study participants

Participants of the F1 generation and their partners were recruited into the new study. The majority of the children in the 1989/90 IOWBC cohort remain on the Isle of Wight or in the UK. Contact details for the entire cohort are held at the David Hide Asthma and Allergy Centre. A newsletter was sent at the beginning of the study outlining the new research project. A study information sheet and a fridge magnet with a reminder about the study were issued together with a reply slip and stamped addressed envelope. Alternatively, the cohort members could inform us of their initial agreement in a text message or e-mail. In addition, a letter was sent every 6 months to all birth cohort members asking them to inform the study team if they or their partners plan to have a baby in the next year. Cohort mothers and partners of cohort boys are also identified from antenatal clinic bookings and invited to participate.

Table 2-1 Summary of Third Generation Study assessments

	Pregnancy				Postnatal		
Assessment	12 weeks	20 weeks	28 weeks	Birth	3 months	6 months	12 months
Questionnaires	√	√	√		√	√	√
Data from Hospital notes				√			
	Parental assessments/samples						
Maternal Blood sample	√	√	√				
Paternal blood sample	√¥						
Parental Saliva for DNA *	√*						
Maternal Urine	√	√	√				
Mother SPT	√¥						√
Father SPT	√¥						
Lung function	√¥						
	Child assessments/samples						
Cord blood				√			
Saliva for DNA							√
Child SPT							√
Examination: eczema (SCORAD)/wheezing					√		√

* Saliva sample collected from parents only if they refused to provide blood sample, ¥; Paternal samples and SPT were done at any point through pregnancy or sometimes soon after birth of the baby.

2.1.5 Questionnaires

Validated questionnaires from the International Study of Asthma and Allergies in Children (ISAAC) study (225) were used. The questionnaires are presented in Appendix 3 (Page 221). Smoking during pregnancy was based on questionnaires at 12-18 weeks, 20 weeks and 28 weeks and at birth. Based on the responses, exposure was catagorised as non-smoker, past smoker, smoking only in 1st trimester and persistent smoking throughout the pregnancy. These categorisation and frequencies are described in detail in chapter 5 (Table 5.3).

2.1.6 Skin Prick Test

Skin prick testing was performed using standard techniques and the protocol used at the 10 (209, 226) and 18 (44, 107) year follow-ups of the original IOWBC (F1 generation). A panel of 13 common allergens was tested in parents; house dust mite (*Dermatophagoides pteronyssinus*), grass pollen mix, tree pollen mix, cat and dog epithelia, *Alternaria alternata*, *Cladosporium herbarum*, milk, hen's egg, soya, cod, wheat and peanut plus histamine and physiological saline to act as positive and negative controls, respectively (Alk-Abello, Horsholm, Denmark). In children SPT was undertaken at the 12-month assessment and the panel had 10 common allergens: house dust mite (*Dermatophagoides pteronyssinus*), grass pollen mix, tree pollen mix, cat and dog epithelia, milk, hen's egg, sesame, cod, wheat and peanut.

Allergen skin test reaction with a mean wheal diameter of at ≥ 3 mm greater than the negative control was regarded as positive and the subject defined as atopic. Atopy was defined by at least one positive reaction to the panel of allergens tested.

2.1.7 Cord Blood

After delivery, a sample of whole blood was collected from the cord in a sterile manner (10 – 20 ml) by the midwife. Following delivery of the placenta, the external surface of the cord was wiped with alcohol wipes (or other material) to

avoid contamination with maternal cells and to ensure sterility. Cord blood was collected by puncturing the cord with a sterile needle to collect blood into a sterile syringe and then transferred into heparinized 5 ml tubes (Corning, non-pyrogenic sterile polypropylene tubes with heparin anticoagulant). The tube was gently shaken so that the anticoagulant mixed thoroughly with the blood. Cord blood serum was separated using a centrifuge by spinning at 3500 RPM for 15 minutes. Whole blood was used for DNA extraction, serum was stored at -80^{0}C for future metabolomic analysis (not part of this thesis).

2.1.8 DNA extraction

DNA was extracted from whole blood collected and frozen in EDTA vacutainers. A simple salting out method was used for DNA extraction. The underlying principle is dehydration and precipitation of the cellular proteins by using saturated sodium chloride as described by Miller et al. (11).

2.1.9 DNA Methylation

Measurement of DNA methylation using DNA methylation arrays was carried out in the Wellcome Trust Centre for Human Genetics, Oxford, UK following the manufacturer's protocol. In brief, 1 μg of DNA was bisulphite-converted using the EZ 96-DNA-methylation kit (Zymo Research, CA). Samples were processed with the Illumina Infinium HumanMethylation450k BeadChip which allows interrogation of over 485,000 CpG sites per sample or the Illumina MethylationEPIC Beadchip microarray, which assesses methylation at > 850, 000 CpGs (227). The BeadChips were scanned using a BeadStation, and the methylation level (beta value) calculated for each queried CpG locus using the Methylation Module of BeadStudio software (228). Multiple identical control samples were assigned to each bisulphite conversion batch to assess assay variability. Samples were randomly distributed on microarrays to control against batch effects.

2.1.9.1 Quantification and preprocessing of methylation

Following generation of raw intensity files from the bead array scanner, the methylation data underwent quantification and pre-processing before analysis. In this thesis two separate data processing pipelines were utilized in chapters 3, 5 and 6, due to the development of improved methods for data processing during the course of the thesis.

2.1.9.2 Quantification and preprocessing of methylation data using IMA

For the analyses presented in chapters 3 and 5, methylation data was preprocessed using Illumina Methylation Analyzer (IMA). IMA is an automated pipeline for Illumina Infinium HumanMethylation450k methylation data. Methylation levels for each CpG site interrogated from BeadStudio software were quantified using the "*beta*" value. The beta value is the ratio of the intensity of the fluorescent signal of methylated probe (M) and the overall intensity (Methylated (M) + unmethylated probe (U)) (229).

$$\text{Beta} = \max(M,0)/(\max(M,0)+\max(U,0)+C)$$

M: methylated, U: Unmethylated and C is the constant

A constant © is used to avoid fractioning by 0 when methylated (M) and unmethylated (U) values are very small. The beta value is a continuous variable ranging from 0 (unmethylated) to 1 (completely methylated). The beta value reported for each locus from the BeadStudio is an average beta value derived from the average M and U of approximately 30 bead replicates (229).

In some analyses, quantification of methylation level was undertaken by generating M-values. M-values are the ratio of the intensity of methylated (M) to the intensity of unmethylated probe in the base 2 logarithm

$$M\text{-}value = log2\ \{ (max(M,0)+\alpha)/(max(U,0)+\alpha)\ \}$$

$$M\text{-}value = log2\ \{ \frac{max(M,0)+\alpha}{max(U,0)+\alpha} \}$$

Where α is a constant.

M-values can range from negative infinity to positive infinity. M-value is 0 if the intensities of methylated and unmethylated are similar and therefore, M-value of 0 indicates 50% methylation. A positive M-value indicates more methylation, and the negative M-value indicates less methylation. M-values can be calculated from beta values. The logit transformation of the beta value will approximate the M-value (230).

$$M\text{-}value = log2(\frac{beta}{1-beta})$$

$$\text{M-value} = \log2(\ \text{beta}/(1\text{-beta}))$$

The beta value can be interpreted as the percentage methylation for a given site, whereas the M-value provides insight into the distribution of methylation across the genome. The beta value range is limited between 0 and 1 and this fails to comply with the Gaussian distribution assumption used by many statistical methods. But for the M-value, statistical analyses based on the Gaussian distribution assumption can be used (230). For linear regression analysis where DNA methylation is an independent variable, for example in gene expression and DNA methylation analysis, M-values can be used.

The data from BeadStudio along with the beta value also includes detection p values. Detection p values can be used as quality control measures for the probes. Detection p value indicates the difference between the intensity measured (M+U) at the CpG site and the intensity at the negative control.

The DNA methylation data gathered from the BeadStudio software from Illumina was pre-processed or cleaned to reduce the background noise and to correct for the potential technical effects such as array or plate effect. Also, there can be systemic errors introduced like batch effect which will be corrected. Preprocessing of methylation data was undertaken using IMA in Bioconductor (IMA is implemented in the R language and is freely available from http://www.rforge.net/IMA) (231), using a site-level p-value cut off of 0.05 with 10% sites, which means excluding sites with p-value > 0.05 in at least 10% of the sample was applied.

Normalisation

The quantified methylation data will need can be normalized to correct for any possible array or probe intensity effects. Normalisation needs to be done so that the results from different arrays can be comparable. The quantile normalisation method was used (228). Here the beta values are ranked from smallest to the largest. The smallest intensities from arrays are averaged and the smallest values in each array will be replaced by the average of the smallest values. This was repeated for the second smallest intensity and then third smallest, fourth smallest, fifth smallest and so on for all the intensities (232).

Peak correction

Infinium 450K uses two types of chemical arrays, Infinium type I and Infinium type II. Infinium I uses two different probes on two different beads for methylated and unmethylated and the intensities for both methylated and unmethylated are in same colour channel. Infinium II uses only one bead type to identify both methylated and unmethylated alleles, but the intensities will be in two different colours. The raw *beta* values from both show a bimodal distribution with the majority of CpGs sites either hypo- or hyper-methylated, however positions of these peaks can show a shift between the type I and type II Infinium assays. For the results from both assays to be comparable, peak correction needs to be applied for Infinium II results to match them to Infinium I results (233). This is also implemented in IMA.

2.1.9.3 Quantification and preprocessing of methylation data using CPACOR

For analyses presented in chapter 6, the CPACOR data processing pipeline, also implemented in R, was used (234). This is similar to IMA, but CPACOR has been demonstrated to provide better normalisation and batch correction in benchmarking tests. Probes were excluded if the detection p-values were larger than 1×10^{-16}. Any sample with less than 95% of the original probes remaining after filtering was removed. Intensity values were separated by probe colour and type, and quantile normalization was performed in respective groups. Intensity values were then converted to beta values.

2.1.9.4 Quantification Effect of Illumina platforms and preprocessing pipelines

In this thesis methylation data was analysed using both the 450K and EPIC platforms, as the 450K array was superseded by the EPIC array during the study. For the EWAS in this thesis samples from both platforms were analysed by limiting analysis to the CpGs present in both arrays, similar to the 450K platform, and the rest of the sites were not included. While this step maximises the sample size in analyses as samples assessed using both the 450K and EPIC arrays can be combined, it means that association between CpGs not common between the two arrays would not be detected.

Two different preprocessing pipelines were used in this thesis (IMA for chapters 3, 5 and CPACOR for chapter 6) due to developments in available analytical tools in the field during the course of this study. The pipeline used for preprocessing the DNA methylation data may influence the analysis, however there is growing evidence that this would not have had a significant effect. In a meta-analysis of EWAS from different study samples, using different preprocessing methods, no significant differences in the results were noted (198).

2.1.9.5 Batch effect correction

When the DNA methylation arrays of samples are done in different batches, there is a possibility of introducing a batch effect. Though batch effects can be reduced, they cannot be eliminated unless all the microarrays are processed in one batch. Batch effects were corrected using ComBat, an Empirical Bayes method (235) before merging data from different batches of DNA methylation arrays (236). For chapter 6, batch correction using ComBat also removed technical variability between Infinium 450k and Infinium MethylationEPIC beadchips.

2.1.9.6 SNP probe effect

There are around 60,000 to 70,000 array probes that may have SNPs lying within the probe annealing region. If this was the case, then the beta value may represent true methylation or a technical artefact due to the SNP effect. Probe SNPs were identified from the publicly available probe SNP information; http://www.rforge.net/IMA/snpsites.txt. If the identified CpG fells in probe SNP, then the distribution of methylation level at that site was be explored for any skewness. If the distribution was not normal or skewed, the CpG site was excluded. In addition, if the minor allele frequency for the SNP was large (>10% in the European population) the CpG assayed by the probe overlapping that probe SNP was excluded.

2.1.9.7 Gender effect

There are gender differences in methylation profiles. Differences in sex chromosome methylation account for most of the gender differences in DNA methylation (237), therefore probes on both X and Y chromosomes were excluded.

2.1.10 Guthrie Cards

Guthrie cards are used to collect blood spots from the pricked heel of new-borns to screen for diseases such as phenylketonuria, cystic fibrosis, and sickle cells. In the UK, four blood spots are taken, one of them is usually used for disease screening. With improving technologies in epigenomics extracting DNA and assessing methylation from blood spot on Guthrie cards has become possible (238). To address the issue of missing cord blood samples, amendments were made in the Third Generation Study to source the Guthrie cards for F2 participants. After ethical and R&D approval and with parental consent, cards were retrieved from regional NHS laboratory. Methodology outlined by Beyan et al. (238) was used to extract DNA from the blood spot on Guthrie cards. Analyses in this thesis are based on only cord blood DNA methylation. Subsequent studies of the comparability of cord blood and Guthrie card methylation (239) show some compatibility of cord blood and Guthrie card DNA methylation but not enough evidence to be able to combine the data.

2.2 Statistical Methods

2.2.1 Gene-Environment-Methylation interaction

Gene-environment-epigenome interaction studies in chapter 3 are candidate gene studies (*IL4* and *IL13*). Gene-methylation interaction studying methylation levels at CpG sites spanning the genomic region of the *IL4R* gene (chapter 3.2) were tested for association with asthma at age 18 years using Wilcoxon tests. Genotype and methylation interaction was assessed on a multiplicative scale by using a cutoff p-value = 0.05. All significant SNP × methylated CpG interactions effects on asthma were included in the final log-linear model. False discovery rate (p=0.05) was employed to correct for multiple testing.

In chapter 3.3, to explore the interplay of *IL13* genotype, prenatal smoking and DNA methylation affecting the airway obstruction and airway reactivity, a two-step model proposed by Karmaus et al. (240) was followed. The two-step approach is described further in detail in chapter 3.3 (section 3.3.2.1).

2.2.1.1 Sample Size for interaction studies

For both interaction studies the available sample size at time of analysis was N= 245. Power calculations were not available for these studies. These were one of the first studies assessing gene-environment-methylation interactions. For Gene-environment interactions, it is understood that power and sample size are determined by genotype allele frequency, outcome variable (binary or continuous) and magnitude of association between the exposure and the outcome. For analysis of gene-environment interactions with binary exposure and outcome, Hwang et al. presented a 2 x 2 x 2 table contrasting genetic susceptibility, exposure, and disease (241). They show that a modest sample size (200 cases and 400 controls) would be adequate when the frequency of exposure is moderate, and the susceptible genotype is common to detect Odds ratio of more than 4 with 80% significance. Foppa et al. extended this approach to be applicable to multiple level exposure and a binary outcome (242). Luan et al. provide an

alternative approach for power calculations and sample size when the outcome is a continuous variable outcome (243), they provide with scenarios with different frequency of rare allele which suggest requirement of large samples (around 5000 for 80% power). The large sample size requirement is mainly addressing the power calculations in genome wide interaction studies. Smaller, and likely underpowered, studies looking at gene-environment interactions carry risk of showing false negative or false positive results, however these studies are primarily exploratory and can still be valid first steps in the field (244).

2.2.2 Candidate gene DNA methylation analysis

Analysis using a candidate gene approach was carried on data with only the CpG sites available for the selected gene. Two steps were used to analyse genotype × methylation interactions. First, each interaction was tested individually (including the individual interaction and main effects) and then the results were corrected for multiple testing using false discovery rate (FDR) (245). In a second step, the significant interaction terms ($p<0.05$) along with all the main effects were entered into a multiple linear regression model in order to adjust the effects of each other. Once the interaction terms (SNP genotype × CpG methylation beta value) significantly associated with the outcomes in above steps were identified, they were then explored in detail with parameters of estimates to understand the effect of increase in DNA methylation on the outcome for each genotype.

2.2.3 Epigenome-wide association study (EWAS)

To assess the association between maternal asthma and genome-wide DNA methylation while controlling for confounding factors (maternal smoking, birth weight, birth order, gender and maternal socio-economic status.) was done by performing robust linear regression using R v3.3.2. Robust linear regression was preferred over linear regression as would be more data inclusive, through a more relaxed attention to normality and outliers. Further specific details of the EWAS analysis are provided in chapter 6.

For handling smaller datasets for prevalence and candidate gene approach analysis was performed using IBM SPSS Statistics for Windows, Version 21.0 (IBM Corp, Armonk, NY).

2.2.4 Socioeconomical status clusters

For EWAS analyses, Socioeconomic status (SES) was corrected for using SES clusters. PROC FASTCLUS was used to generate a cluster variable. Analyses were run for the number of clusters ranging from two to eight, and R^2, Pseudo-F and CCC statistics were used to determine the number of clusters. The clusters of SES are described in chapter 4.

2.3 Third Generation Study data management

A large volume of data was generated, making accurate and automated data management crucial. Data management followed a similar pattern to that successfully used previously for the IOWBC. Current name and address were entered into an Excel database. An Access database (Microsoft, USA), using study number only for identification, was established. Research participants were seen in The David Hide Asthma and Allergy Research Centre where all data was gathered, processed and stored in a secure manner. The IOWBC members (F1 generation) retained their original study numbers. Their offspring (F2 generation) were given a new number in numerical order. A unique study number was generated by employing "infant number – mother's number – father's number". This ensures the number remains unique even if same mother or father (F1 generation) had 2 or more children (F2).

Confidentiality of all samples, interviews, and medical records were assured by a) keeping all records under lock and key, b) separating data from names, c) keeping the linkage study numbers under lock and key, d) allowing only study staff members to have access to the data, e) keeping identifiers of individuals out of public material and reporting only aggregated data.

2.4 Isle of Wight Birth Cohort (IOWBC)

General methodology of the research is described in this chapter and the specific methods are further described in the respective result chapters. An overall description of the IOWBC has been provided in Chapter 1, here methods for specific clinical measurements related to this thesis are described.

2.4.1 Methacholine challenge test

The methacholine test protocol was a standardised protocol, as recommended by the American Thoracic Society (246). Testing was undertaken by a trained, research nurse and registrar. A computerised dosimeter system (Koko Digidoser,

PDS Instrumentation, Louisville, USA) was used with compressed air source at 8 L/minute and nebuliser output of 0.8 L/minute. A pre-test spirometry reading was obtained to ensure an FEV_1 of above 70% predicted for age and height. Initial inhalation of 0.9% saline was followed 1 minute later by spirometry recording to obtain a baseline value. If the FEV_1 dropped by more than 10% with saline, the challenge was to be deferred for an hour or rescheduled for another day; if it dropped by more than 15%, the test was postponed. Subsequently, incremental concentrations from 0.062 mg/mL to 16 mg/mL of methacholine were serially administered using the methods of Chai and co-workers (247). To perform this test, adolescents were required to be free from respiratory infection for 14 days, not taking oral steroids, not taken short acting β_2-agonist for 6 hours, and long acting β_2-agonist for 12 hours, and needed to have abstained from caffeine intake for at least 4 hours. A concentration causing a 20% fall in FEV_1 from the post-saline value was interpolated and expressed as PC_{20} FEV_1 (provocative concentration causing a 20% fall in FEV_1). The concentration of ≤ 8mg/mL was used as a cut-off for defining bronchial hyper-responsiveness (BHR) based on earlier population based studies (248-250). A continuous measure of dose-response-slope (DRS) was also applied so that all adolescents could be included in the analysis even if their FEV_1 did not drop by at least 20% with the final dose of methacholine.

2.4.2 Spirometry

The American Thoracic Society guidelines were followed to ensure spirometry validity and reproducibility (251). As recommended, the highest of three FEV_1 measurements within 5% of each other were used. The Koko system (Koko Spirometer manufactured by nSpire Health, Inc. 1830 Lefthand circle, Longmont, CO 80501 USA) was used. To perform this test study participant were required to be free from respiratory infection for 14 days, not taking short acting β_2-agonist medication for 6 hours, long acting β_2-agonist medication for 12 hours and abstain from caffeine intake for at least 4 hours. Forced expiratory volume in one second (FEV_1), forced vital capacity (FVC), mid expiratory flow (MEF), peak

expiratory flow (PEF) were recorded. Percent predicted for age, height, sex and ethnic origin were calculated for the above data and forced expiratory ratio (FEV_1/FVC).

2.4.3 Genotyping of candidate genes in IOWBC

Genotyping data for the F1 cohort had been previously generated for candidate genes including the IL13 locus at Michigan State University (16, 17). In brief genomic DNA was isolated from the samples using QIAamp DNA Blood Kits (Qiagen, Valencia, CA) or the ABI PRISM™ 6100 Nucleic Acid PrepStation (Applied Biosystems, Foster City, CA). Polymorphisms in the IL13 gene were examined using the SNPper and Applied Biosystems databases (18). Genotyping was conducted by fluorogenic 5' nuclease chemistry PCR using Assays on Demands kits cycled on a 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA), or biotin-streptavidin-based pyrosequencing performed on PSQ-96 instrumentation (Biotage AB, Uppsala, Sweden). SNPs spanning the genomic region of the IL13 gene were selected for genotyping using a haplotype tagging strategy implemented in Haploview using HapMap Caucasian data (Hapmap Data PhaseIII/Rel#2, Feb09, on NCBI B36 assembly, dbSNP b126) (19).

2.5 Ethics approval

Ethical approvals for both IOWBC and the Third Generation Study were obtained from the relevant authorities as follows:

Isle of Wight Birth Cohort 18-year follow up: Isle of Wight, Portsmouth & Southeast Hampshire Research Ethics Committee. REC reference number: 06/Q1701/34, Date: 21/06/2006.

Third Generation Study: National Research Ethics Service Southampton & Southwest Hampshire Research Ethics Committee. REC reference number: 09/H05041129, Date: 11/02/2010.

Chapter 3 Interaction of genotype, DNA methylation and environmental exposure.

3.1.1 3.0.1 Introduction

Asthma is a chronic airway disease characterised by airflow limitation and airway-reactivity. There is wide heterogeneity in disease susceptibly and phenotypes (252) . The aetiology of asthma and the reason for its heterogeneity are still not clearly understood, although genetics and environmental exposures are known to play their part. Multiple genes and gene-gene and gene-environment interactions have all been shown to play a role in determining susceptibility to asthma and associated phenotypes of lung function and airway-reactivity. The underlying process of these interactions is not clearly understood. DNA methylation represents a site of molecular interaction between the environment and genome. There is growing evidence that DNA methylation plays a role in complex diseases like asthma (253) and can be modified by environmental exposures such as tobacco smoke (254) as well as by disease-associated genotypes (255). DNA methylation likely plays a role in determining the influence of genetic variants and the gene-environment interactions on asthma.

To test hypothesis 1 (section 1.8.1) a targeted gene approach is employed in this chapter where known asthma genes (*IL4R* and *IL13*) are studied. First part of this chapter (3.1) looks at the association of interaction of genetic variants and DNA methylation of the IL4R gene on the risk of asthma. The second part (3.2) is a proof of concept study to assess the role of DNA methylation in gene-environment interactions; IL13 variants and maternal smoking on asthma related airway traits. In both these studies we focused on blood samples collected at 18 years of age from 245 female Isle of Wight cohort participants (F1).

3.1.2 Study subset of the Cohort

Both studies covered in this chapter were carried out in a subset of IOWBC female subjects who were randomly selected for measurement of DNA methylation in DNA extracted from peripheral blood collected at 18 years of age. The subset was based on those with information, samples and lung functions available at 18 years

follow up. DNA methylation was assessed in this subset as part of a pilot epigenetics project. Lung function measurements and methylation data were available in 245 females aged 18 years who were randomly selected from the cohort population for epigenetic association studies and multigenerational study because these were the only female samples analysed at the time of this study. The available epidemiological data up to 18 years and the genotyping and DNA methylation (at 18 years) data were utilised for these two exploratory studies. The description of the subset in comparison to the whole cohort females is provided in Table 3-1. To assess whether our analytic sample (245 DNA samples) was representative of the total cohort available at age 18 years, we compared the characteristics of these two subsets by using the chi-square test. There were no substantial differences between the female participants of the cohort and the subset. The general description of the IOWBC and the specific details of the methodology for spirometry and methacholine challenge test and the genotyping are provided in General Methods chapter 2. Maternal history of asthma and smoking during pregnancy was ascertained at birth. Birth weight was obtained from birth records. At ages 1, 2, 4, 10, and 18 years, the original questionnaire-based information was updated, and weight and height of the child were measured. Asthma at age 18 years was defined as subjects with a physician diagnosis of asthma plus current symptoms and/or asthma medication.

Table 3-1 Baseline characteristics for women with available methylation data compared to the female participants of the total cohort.

Factors	**Females with methylation data N=245**	**Total Females in the cohort N=750**	**P value** **Chi-square**
Maternal history of asthma	30/243 12.3%	80/742 10.8%	0.500
Maternal smoking during pregnancy	47/244 19.3%	188/743 25.3%	0.052
Low birth weight (≤2.5 kg)	9/237 3.8%	35/734 4.8%	0.53
Breastfeeding duration (weeks)	10.5 (0, 40)	8.0 (0, 40)	0.160
Asthma at 18 yrs	35/245 14.3%	128/659 19.4%	0.070
Atopy	78/242 32.2%	158/446 35.4%	0.400
Smoking at 18 yrs (active or passive)	98/237 41.4%	288/660 43.6%	0.542
Measurements	**n/N Mean**	**n/N Mean**	**t-test**
FEV_1/FVC	237/245 0.878	443/750 0.879	0.755
Log10 (DRS+10)	156/245 1.115	293/750 1.136	0.407

n= number of participants with the outcome or exposure, N= total number with the information available. Binomial variables analysed using Chi-square test, continuous variables analysed using t-test. FEV1/FVC: Ratio of Forced Expiratory Volume in first second over Forced vital capacity, DRS: Dose response slope of Methacholine challenge test.

3.2 Interaction of genotype and DNA methylation of *IL4R* and risk of asthma.

3.2.1 Introduction

The *IL4R* gene encodes the alpha chain of the IL-4R, a type I trans-membrane protein that can bind IL-4 and IL-13 to regulate IgE production(256). The encoded protein can bind IL-4 to promote differentiation of T helper (Th) 2 cells (257, 258). An animal study has shown that selective smooth muscle expression of *IL-4R* is sufficient for cytokine–dependent induction of airway hyperresponsiveness by increasing the expression of several genes related to smooth muscle differentiation, proliferation, migration, and contractility (259). Several single nucleotide polymorphisms (SNPs) of *IL4R,* have been associated with asthma (260-263). Rosa-Rosa et al., demonstrated that homozygosity of an *IL4R* variant allele (R576) increased in individuals with asthma (relative risk = 8.2) compared with controls (263). The *IL4R* variant rs3024622 has been shown to interact with of asthma resulting in a reduction of asthma exacerbations (264). However, the role of *IL4R* variants as a genetic risk factor for asthma remains unclear with some conflicting reports; a meta-analysis reveals a modest risk (261). Other genetic regulatory mechanisms beyond DNA sequence variation may aid in explaining the role of *IL4R* in asthma. This study was led by Dr Soto-Ramirez and I was part of the team exploring gene-environment-DNA methylation interactions, contributed to the study and in drafting of the manuscript.

3.2.2 Methods

Eight uncorrelated SNPs were identified after selecting one SNP that represents each LD block or an unlinked area following the estimates of estimates of linkage disequilibrium (LD) between SNPs being calculated (265). Details of the methodology were published by Soto-Ramirez et al (266).

After identifying eight uncorrelated *IL4R* SNPs (Table 3-2) and finding which CpG site (Table 3-3) was significantly associated with asthma, the statistical interaction was estimated between these SNPs and the methylation level of cg09791102 on the risk for asthma at age 18 years. The interaction was assessed

on a multiplicative scale (a total of 8 log-linear models) by using a cutoff p-value = 0.05. All SNP × methylated CpG interactions that showed significant effects on asthma at age 18 years were included in the final log-linear model.

Since a total of 8 crude SNP× methylation interactions were tested before selecting the full model, adjusted for multiple testing by applying false discovery rate (p=0.05) (267).

3.2.3 Results

A total of 12 CpG sites spanning the genomic region of the *IL4R* gene were analysed for association with asthma at age 18 years (Table 3-3). Only methylation levels of cg09791102 showed an association with asthma at age 18 years (Wilcoxon test: p-value = 0.01).

Table 3-2 Location and position of the single nucleotide polymorphisms in the *IL4R* (Chromosome 16)

SNP ID	Location of SNP	Chromosomal location
rs2057768	5'UTR	27322095
rs6498012	Intron	27331974
rs3024622	Intron	27365453
rs4787423	Intron	27367334
rs3024676	coding	27373558
rs3024685	3'UTR	27376910
rs12102586	3'UTR	27378053
rs16976728	3'UTR	27381712

Regarding the selection of SNPs in the *IL4R* gene, 8 out of 13 were chosen since they were uncorrelated. After testing for interaction between these SNPs and the methylation levels of cg09791102 on the risk for asthma at age 18 years, we observed that the interaction term of SNP rs3024685 × methylation levels of cg09791102 (continuous scale) was statistically significant (p-value = 0.0003; p-value after adjusting for false discovery rate=0.002; Table 3-4). In other words, the genetic risk of rs3024685 varies with the methylation level of cg09791102 (Figure 3-1).

Table 3-3 Location and position of the CpG sites in the *IL4R* gene.

CpG ID	Chromosomal location	Location of CpG site	Median	Percentiles	
				5%	95%
cg08932316	5:27324341	TSS1500	0.89	0.86	0.92
cg05729093	5: 27324953	TSS1500	0.06	0.04	0.08
cg03980304	5: 27325000	TSS1500	0.07	0.04	0.10
cg00090800	5: 27325237	TSS200	0.03	0.01	0.05
cg06641959	5: 27325254	5'UTR	0.08	0.05	0.11
cg01706029	5: 27325672	5'UTR	0.07	0.05	0.09
cg26937798	5: 27326054	5'UTR	0.09	0.06	0.12
cg16649560	5: 27338391	5'UTR	0.21	0.15	0.29
cg08317580	5: 27345891	5'UTR	0.90	0.87	0.93
cg09791102	5: 27353414	Body	0.88	0.85	0.91
cg01165142	5: 27367172	Body	0.58	0.51	0.65
cg05903710	5: 27375732	3'UTR	0.87	0.84	0.90

DNA-M level range for cg09791102 was 0.48 to 0.92 (blue bars in Figure 3-1). Since the number of participants at methylation levels of 0.85 or less were low, we categorized these methylation levels into ≤ 0.85 (n = 9). For descriptive purposes, 84 participants had methylation levels of 0.88 and less, 112 participants of 0.89-0.90, and 35 of 0.91- 0.92. We assumed an additive mode of inheritance and compared participants who had the "CC" and "CT" genotypes with those who were "TT" genotype. For the genotype "CC", compared to "TT" at rs3024685, we found that at methylation levels of 0.85, 0.86, 0.90, 0.91, and 0.92, the RRs of asthma were 0.01 (95% CI: 0.001, 0.40; p = 0.01), 0.04 (95% CI: 0.002, 0.74; p = 0.03), 4.65 (95% CI: 1.11, 19.33; p = 0.03), 14.76 (95% CI: 3.00, 72.52; p = 0.009) and 46.90 (95% CI: 6.65, 330.52; p-value = 0.0001, Figure 3-1), respectively. Similar results were found with "CT" genotype, however the interaction term was only marginally significant (p-value = 0.06). Compared to "TT", the genotype "CT" at methylation levels of 0.90, 0.91, and 0.92, the RRs were 2.82 (95% CI: 0.93, 8.53; p = 0.06), 4.37 (95% CI: 1.13, 16.93; p = 0.03), and 6.78 (95% CI: 1.24, 37.10; p = 0.02), respectively.

Descriptively, the interaction term translates into 13.2% and 14.3% of the participants having asthma at a methylation level of 0.88 and the genotype "CT" and "TT", respectively; and none of the "CC" genotype had asthma. Between 0.89 to 0.90 methylation, 15% of the "CC", 16.7% of the "CT", and 7.9% of the "TT" genotype had asthma. At methylation level larger than 90%, 54.6% of the "CC" and 16.7% of the "CT" genotype had asthma, and none of the "TT" genotype had the disease.

Table 3-4 Adjusted log-linear regression model of the interaction of genetic variants and DNA methylation of the *IL4R* gene on asthma at age 18 years.

Parameter		Estimate	95% CI		P-value
Intercept		19.59	-12.78	51.97	0.25
cg09791102		-26.97	-63.62	9.67	0.14
rs3024685	CC	-102.45	-158.51	-46.40	0.0003
	CT	-38.48	-80.71	3.75	0.07
	TT	Reference			
cg09791102*rs3024685	CC	115.54	53.18	177.91	0.0003
	CT	43.90	-3.45	91.27	0.06
	TT	Reference			
rs3024622	CC	-1.24	-3.45	0.95	0.26
	CG	-0.14	-1.00	0.72	0.74
	GG	Reference			
rs12102586	TT	2.41	0.29	4.53	0.02
	CT	0.65	-0.24	1.55	0.15
	CC	Reference			
rs16976728	TT	-0.53	-2.03	0.95	0.48
	CT	0.16	-0.78	1.11	0.72
	CC	Reference			
Maternal Smoking during pregnancy		0.43	-0.40	1.26	0.31
Maternal history of asthma		0.53	-0.41	1.49	0.26
Body mass index at age 18 years (kg/m^2)		0.05	-0.009	0.12	0.09
Breastfeeding duration (weeks)		0.02	-0.004	0.04	0.11

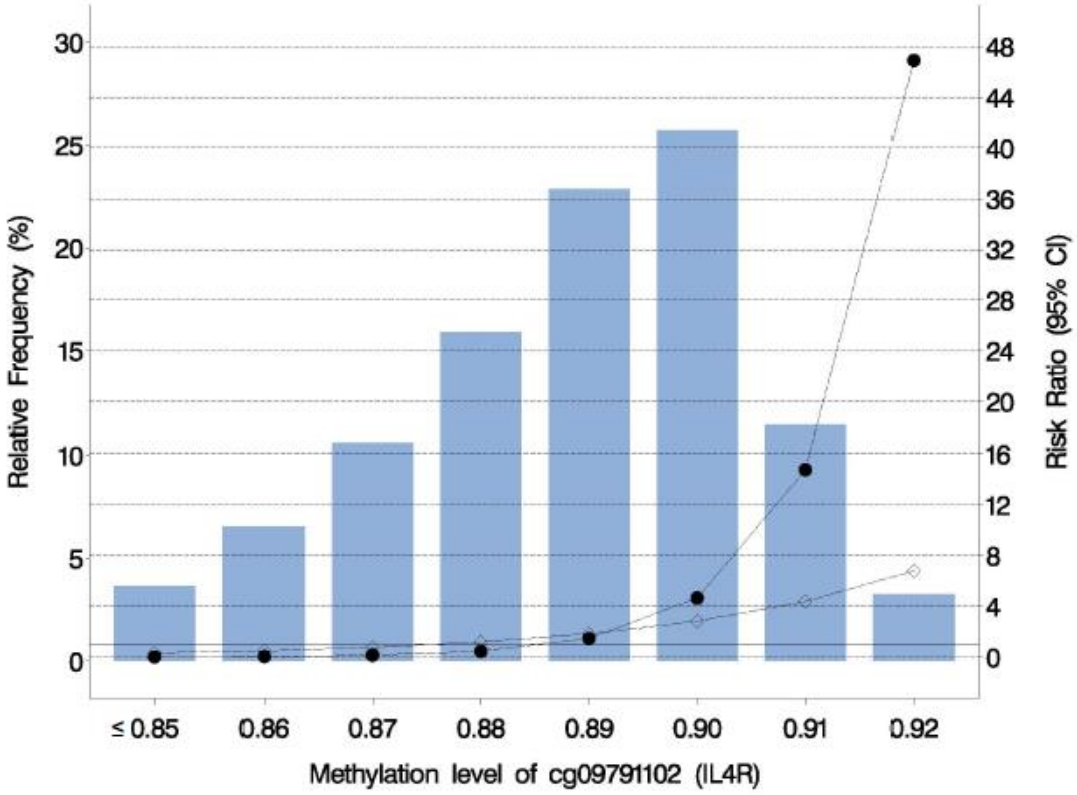


Figure 3-1 Risk Ratio of asthma at age 18 years versus methylation score at different genotypes of IL4Rrs3024685

The blue bars present the relative frequency of the DNA methylation levels. For instance, 87% methylation is found in 10% of the participants. The reference genotype is 'TT'. The solid horizontal line that indicates a risk ratio value of '1' shows the risk ratio of the reference 'TT' genotype. The black dot represents the 'CC' genotype, and the diamond is 'CT' genotype.

3.2.4 Discussion

This is one of the first pilot studies to determine the role of both genetic and epigenetic factors within the genomic region of the *IL4R* gene on the risk for asthma. Interaction of DNA Methylation at CpG site cg09791102 and SNP rs3024685 in the intragenic region of the *IL4R* gene modulate the risk of asthma even after adjusting for multiple testing.

The SNP rs3024685 in the 3'UTR region has no effect on asthma at age 18 years on its own; however, in interaction with the CpG site cg09791102, it is strongly associated with asthma in female participants. At 92% methylation level, rs3024685 ("CC" genotype compared to "TT") showed an increased risk for asthma (RR = 46.9); at a lower DNA-M level (≤85%) the risk of asthma decreased by 98%. Our observation of the gene-body methylation is further supported by the emerging evidence, which shows that methylation in intragenic regions is positively correlated with gene expression levels and phenotype variation (268), (269). However, the mechanism of intragenic DNA methylation is not clear. Theoretical assumption is that a higher DNA-M may mask an otherwise protective effect of rs3024685 and thus increases the risk of asthma. Our results indicate that considering both genetic variants and DNA methylation will significantly improve the explanation of asthma. The strong interaction of one SNP and DNA methylation is encouraging and provides a novel model how a joint effect of genetic variants and DNA methylation can explain asthma. There are several limitations; this study had a limited sample size, it only focused on female participants and utilised candidate gene approach. However, the results should motivate other larger studies to replicate interactions between genetic variants and DNA methylation, particularly for the *IL4R* gene and asthma.

3.3 Interaction of prenatal maternal smoking, IL13 genetic variants and DNA methylation influencing airflow and bronchial reactivity.

3.3.1 Introduction

Various early life environmental exposures are associated with development of asthma. Maternal smoking during pregnancy is a significant risk factor for developing asthma in offspring (270) and can also affect offspring lung function (271). IL-13 is a pleomorphic cytokine and is involved in the pathogenesis of asthma; in airway-inflammation, reactivity and remodelling. In severe asthma, monoclonal antibody to IL-13 use is associated with improvement in lung function in humans (272) and airway hyper-responsiveness in a murine model (273). *IL13* is located on chromosome 5q31 and has been consistently associated with asthma in both candidate gene and genome-wide studies (147, 274). Several functional genetic variants occur in *IL13* including rs1800925 and rs20541 which in a meta-analyses have been shown be associated with asthma risk (275, 276) and with Forced Expiratory Volume in 1 second (FEV_1) and Forced Expiratory Volume in 1 second / Forced vital capacity (FEV_1/FVC) in asthmatics (277, 278).

DNA methylation represents a site of molecular interaction between the environment and genome. There is growing evidence that DNA methylation plays a role in complex diseases like asthma (253) and can be modified by environmental exposures such as tobacco smoke (254) as well as by disease-associated genotypes (255). Study in IOWBC previously has shown that *IL13* polymorphism modifies the impact of prenatal smoking exposure on childhood asthma, suggesting a role for gene-environment interaction (164). It has also been shown that genetic variants in the IL-4 receptor interact with DNA methylation to determine the risk of asthma (266). DNA methylation is a potential integrator of different signals affecting disease susceptibility, with both environment and genotype influencing methylation levels. This makes DNA methylation a plausible mediator for the gene-environment interactions. The hypothesis is that DNA methylation at IL13 promoter region interacts with prenatal smoking exposure and genotype to influence the airway obstruction and airway reactivity.

3.3.2 Methods

The Forced Expiratory Ratio (FER), defined as / Forced Expiratory Volume in 1 second (FEV_1)/Forced Vital capacity (FVC), is used to measure airway limitation. Airway reactivity is measured using bronchial challenge test; PC (methacholine concentration causing 20% fall in FEV_1) is used as cut off to indicate airway reactivity. DRS (dose response slope) is measured as the gradient of fall in FEV_1 with increasing doses of methacholine during bronchial challenge and gives a quantitative measure of airway reactivity.

Regression analysis was employed to look for effect of interactions. First each interaction term was tested individually and then the results were corrected for multiple testing using false discovery rate (FDR)(245). To adjust for possible confounders in second step all the significant interaction terms ($p<0.05$) were entered into one model along with all main effects. SNP genotype dependent methylation was analysed using Kruskal-Wallis test.

3.3.2.1 Two-stage model

Karmaus et al. proposed a two-stage model to incorporate the role of genetic variants, environment and DNA methylation interactions in asthma (240). In stage 1, an environmental exposure and genetic variant interact to influence DNA methylation at a specific site in an adjacent locus. This stage identifies the conditional methQTL and the change in DNA methylation once established can differentially regulate gene activity. In stage 2, the phenotypic effects of sequence variants of the gene (modifiable genetic variants (modGVs)) can be modified by the pre-established methylation by the conditional methQTL. This two-stage model for asthma related lung function is depicted in Figure 3-2. Given the previous observation of the interaction between maternal smoke exposure and genotype in determining asthma, the hypothesis was that this interaction would occur through a mechanism involving methQTL and/or modGVs. A published manuscript of these results is included in Appendix 4 (279).

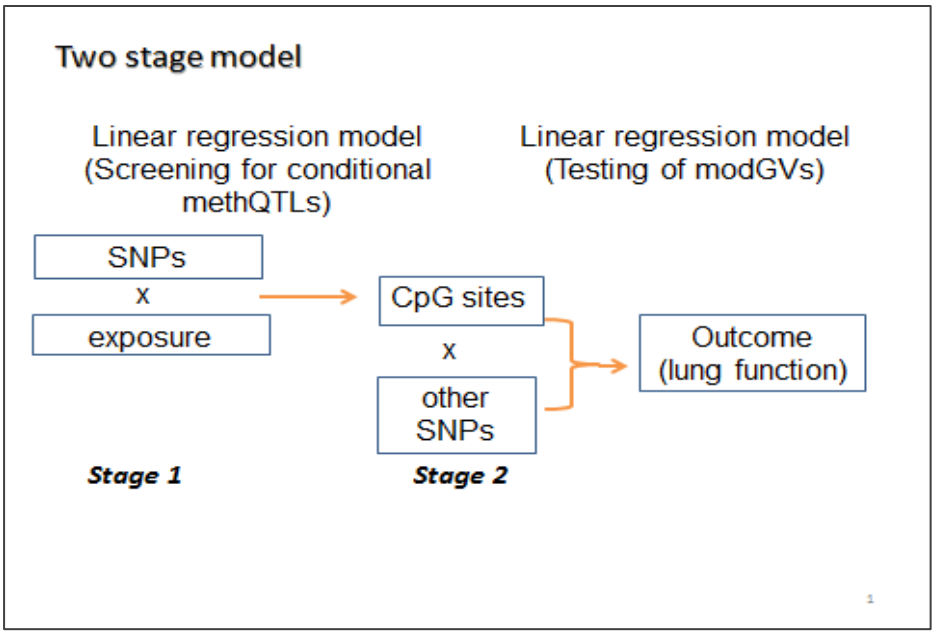
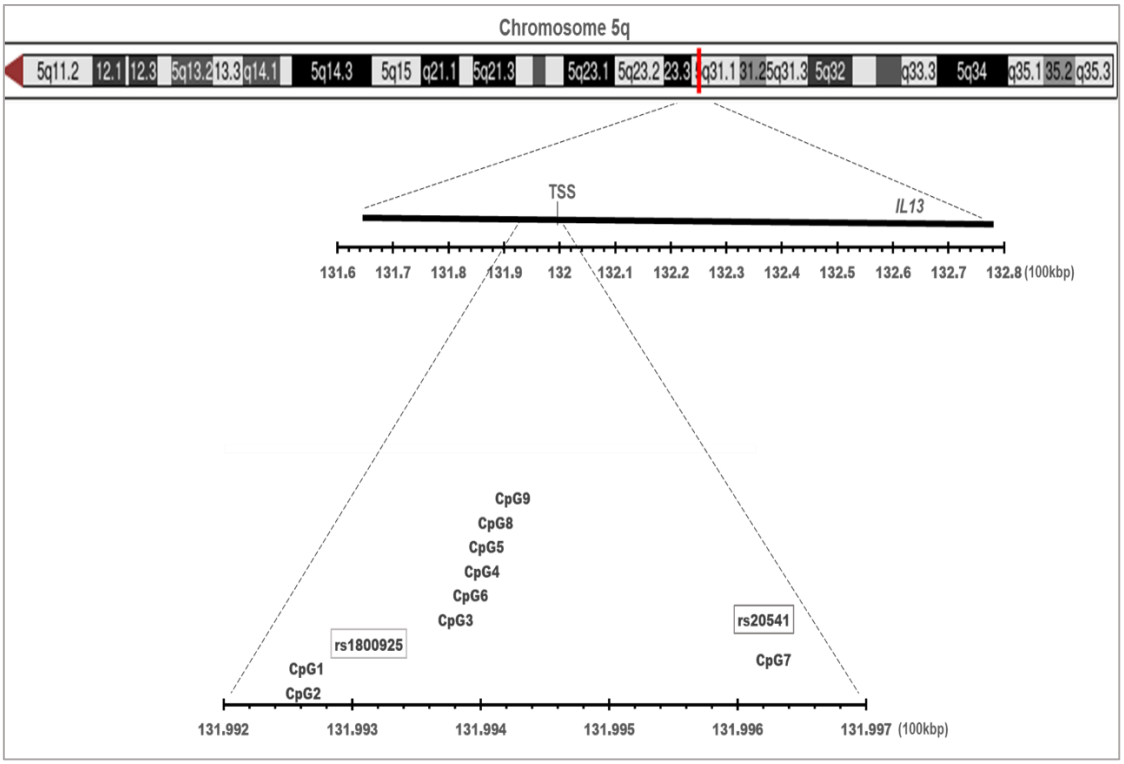


Figure 3-2 Two-stage model to investigate environment, SNP and DNA methylation interactions influencing asthma-related lung functions (Modified from Karmaus et al. (240).

3.3.3 Results

3.3.3.1 IL13 DNA methylation and selection of CpG cites

Nine CpG sites in the *IL13* genomic region are interrogated on the Illumina 450K Methylation array. Six CpG cites are in the promoter region and were included in the analysis. Table 3-5 gives the location and the description of methylation at the sites. Significant moderate correlation between CpG sites at the promoter region was determined using Spearman`s correlation test and a cut-off coefficient of 0.50 with a significance of p <0.001 was applied. Fig3.3 gives visual description of *IL13,* putative SNPs and CpG sites analysed in this study on Chromosome 5q.



CpG1; cg13566430, CpG2; cg04303330, CpG3; cg14523284, CpG4; cg06584121, CpG5; cg06967316, CpG6; cg15329179, CpG7; cg11798521, CpG8; cg07810967, CpG9; cg24580593

Figure 3-3 Pictorial description of the location of sites, SNPs and *IL13* on Chr 5.

Of the six CpG sites located in the promoter region of *IL13* present on the illumina methylation array, three are ≤200 base pairs (bps) from the transcription start site (TSS200) and remaining three are ≤1500 bps from the transcription start site (TSS1500). All 3 CpG sites (cg06584121, cg06967316, cg15329179) at TSS200 are located within a short region stretching 125 bps (131993853 to131993728) and methylation levels at these CpGs are significantly correlated (Table 3-6). Two (cg13566430 and cg04303330) of the 3 CpG sites at TSS1500 are located 25 bps apart and methylation levels are correlated, but cg14523284 is located 1184 bps away and methylation is not correlated with the cg13566430 and cg04303330. Methylation at any of the TSS200 CpG was not correlated with that of any CpG sites at TSS1500. One CpG (cg06584121) out of three correlated and closer sites from TSS200 and two un-correlated and distant CpG sites (cg13566430 and cg14523284) from TSS1500 were selected for further analysis.

Table 3-5 Location and description of IL13 CpG sites

CpG ID	Chromosomal location	Location of CpG site	Median methylation	Percentiles 5%	95%
cg13566430	5:131992455	TSS1500	0.18	0.14	0.23
cg04303330	5: 131992430	TSS1500	0.30	0.23	0.36
cg14523284	5: 131993614	TSS1500	0.86	0.83	0.89
cg06584121	5: 131993818	TSS200	0.80	0.730	0.84
cg06967316	5: 131993853	TSS200	0.74	0.66	0.80
cg15329179	5: 131993728	TSS200	0.87	0.81	0.90
cg11798521	5: 131996143	3'UTR	0.83	0.77	0.86
cg07810967	5: 131993925	Exon 1	0.86	0.83	0.90
cg24580593	5: 131994061	Body	0.88	0.85	0.90

TSS200- 200 bp from Transcription start site, TSS1500- 1500 bp from transcription start site, UTR- Un-translated region

Table 3-6 Promoter region CpG sites: genomic location and correlation of methylation

CpG site Chr Location Location		cg13566430 131992455 TSS1500	cg04303330 131992430 TSS1500	cg14523284 131993614 TSS1500	cg06584121 131993818 TSS200	cg06967316 131993853 TSS200	cg15329179 131993728 TSS200
cg13566430	Corr Coeff. Sig Distance(bp)		0.505 <0.001	-0.110 0.087	-0.243 <0.001	-0.238 <0.001	-0.182 0.004
cg04303330	Corr Coeff. Sig Distance(bp)	25		0.299 <0.001	-0.057 0.379	0.010 0.873	0.11 0.869
cg14523284	Corr Coeff. Sig Distance(bp)	1159	1184		0.406 <0.001	0.416 <0.001	0.363 <0.001
cg06584121	Corr Coeff. Sig Distance(bp)	1363	1388	204		0.765 <0.001	0.675 <0.001
cg06967316	Corr Coeff. Sig Distance(bp)	1398	1423	239	35		0.678 <0.001
cg15329179	Corr Coeff. Sig Distance (bp)	1273	1298	114	90	125	

Chr location: chromosomal location, Corr Coeff- Spearman`s correlation coefficient, Sig- two tailed significance level, bp-base pairs. Green box indicates non-correlation and red box indicates correlation. Blue boxes show the distance in base pairs.

SNP selection

In this candidate gene approach study two SNPs; rs1800925 and rs20541 of IL13 were included in the analysis of interaction. rs1800925 and rs20541 are two common functional SNPs of IL13 shown to be associated with asthma and asthma related lung function measures in multiple studies (275-277, 280). Chromosomal location and the description of the SNPs in the cohort are given in Table 3-7.

Table 3-7 Location and description of *IL13* SNPs

SNP	Chromosomal location	Location	Genotypes	n	%
rs1800925	5:131992809	5` Promoter Upstream	TT	8	3.4%
			CT	68	28.8%
			CC	158	67.8%
rs20541	5:131995964	Exon 4	AA	10	4.3%
			AG	66	28.1%
			GG	159	67.7%

3.3.3.2 Interplay of prenatal smoking, genotype and DNA methylation of IL13

All six CpG sites spanning the promoter region of *IL13* were analysed in Stage 1. The effect of prenatal smoking exposure interacting with both SNPs was explored independently for DNA methylation at each CpG site. DNA methylation at cg13566430 was influenced by the interaction of rs20541 and maternal smoking during pregnancy (p=0.043), this remained significant after correcting for personal smoking at 18 years (p=0.041) (Table 3-8).

Table 3-8 Screening for conditioned methQTLs in *IL13* promoter region (interaction of prenatal smoking exposure (PSm) and *IL13* SNPs affecting DNA methylation)

CpG	PSm* rs1800925		PSm* rs1800925 (corrected for current personal smoking)		PSm*rs20541		PSm* rs20541 (corrected for current personal smoking)	
	Estimate	p	Estimate	p	Estimate	p	Estimate	p
cg13566430	0.094	0.759	0.270	0.604	3.184	0.043	3.244	0.041
cg04303330	0.047	0.829	0.221	0.639	0.048	0.953	0.059	0.942
cg14523284	0.244	0.622	0.508	0.477	0.503	0.605	0.713	0.491
cg06584121	0.353	0.553	0.112	0.739	0.055	0.947	0.147	0.863
cg06967316	0.419	0.518	0.110	0.741	0.169	0.845	0.362	0.697
cg15329179	3.561	0.060	3.009	0.084	0.063	0.939	0.021	0.979

Psm; prenatal smoking, p; p value

In the second step, the interaction between DNA methylation at cg13566430 and rs1800925 genotype on lung function (FEV_1/FVC and DRS) was explored. DNA methylation at cg13566430 and rs1800925 genotype interaction was significantly associated to FEV_1/FVC and DRS (p=0.013 and 0.036 respectively) (Table 3-9). In subjects with the TT genotype, FEV_1/FVC increases with increasing methylation of cg13566430 (coefficient 3.274, p=0.042), for the CT genotype group the increase was smaller (coefficient 0.799, p=0.086) and CC was the reference genotype. Figure 3-4 provides a visual description of the effect of DNA methylation at cg13566430 and rs1800925 on FEV_1/FVC. DRS decreases quickly with increasing methylation of cg13566430 in TT genotype (coefficient -27.497, p=0.010), while the drop in DRS in CT is insignificant (coefficient -0.809, p=0.742).

Table 3-9 Interaction of methylation at cg13566430 and rs1800925 genotypes on FEV_1/FVC and DRS.

		FEV_1/FVC			DRS		
		Estimate	**95% CI**	**p**	**Estimate**	**95% CI**	**p**
Main effects							
cg13566430		0.175 -0.166 0.517	-0.166 0.517	0.313	-0.568	-2.361 1.226	0.533
rs20541	**AA**	0.020	-0.024 0.064	0.369	-0.075	-0.278 0.129	0.469
	AG	0.006	-0.014 0.026	0.548	0.036	-0.076 0.149	0.523
	GG		Reference			Reference	
rs1800925	**TT**	0.030	-0.019 0.079	0.231	0.299	0.045 0.554	0.021
	CT	-0.001	-0.021 0.019	-0.001	-0.026	-0.134 0.083	0.644
	CC		Reference			Reference	
Interaction							
rs1800925	**TT**	-0.412	-0.855 0.030	0.068	3.989	1.194 6.785	0.005
	CT	-0.134	-0.296 0.028	0.106	0.124	-0.733 0.982	0.775
	CC		Reference			Reference	
cg13566430		0.174	-0.301 0.648	0.472	0.624	-1.926 3.174	0.629
rs1800925 × cg13566430	**TT**	3.274	0.114 6.434	0.042	-27.497	-48.283 -6.710	0.010
	CT	0.799	-0.114 1.712	0.086	-0.809	-5.661 4.043	0.742
	CC		Reference			Reference	

DRS, dose response slope; FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity.

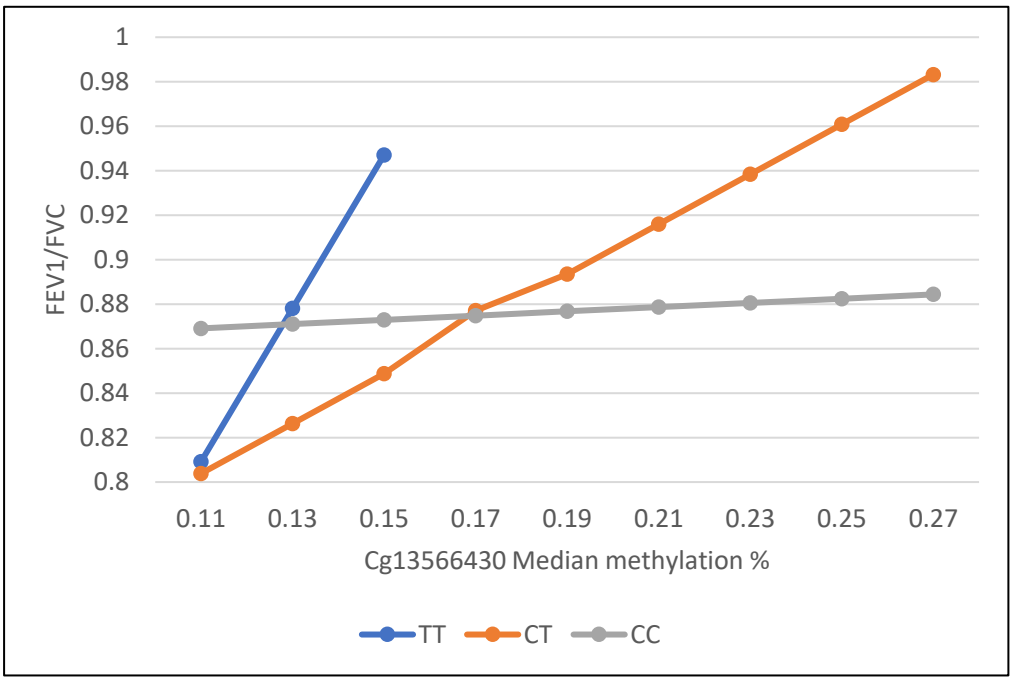


Figure 3-4.1 Interaction of rs1800925 and methylation at cg13566430 influencing FEV_1/FVC.

FEV_1, forced expiratory volume in 1 second; FVC, forced vital capacity.

3.3.3.3 Effect of interaction of genotype and DNA methylation on asthma

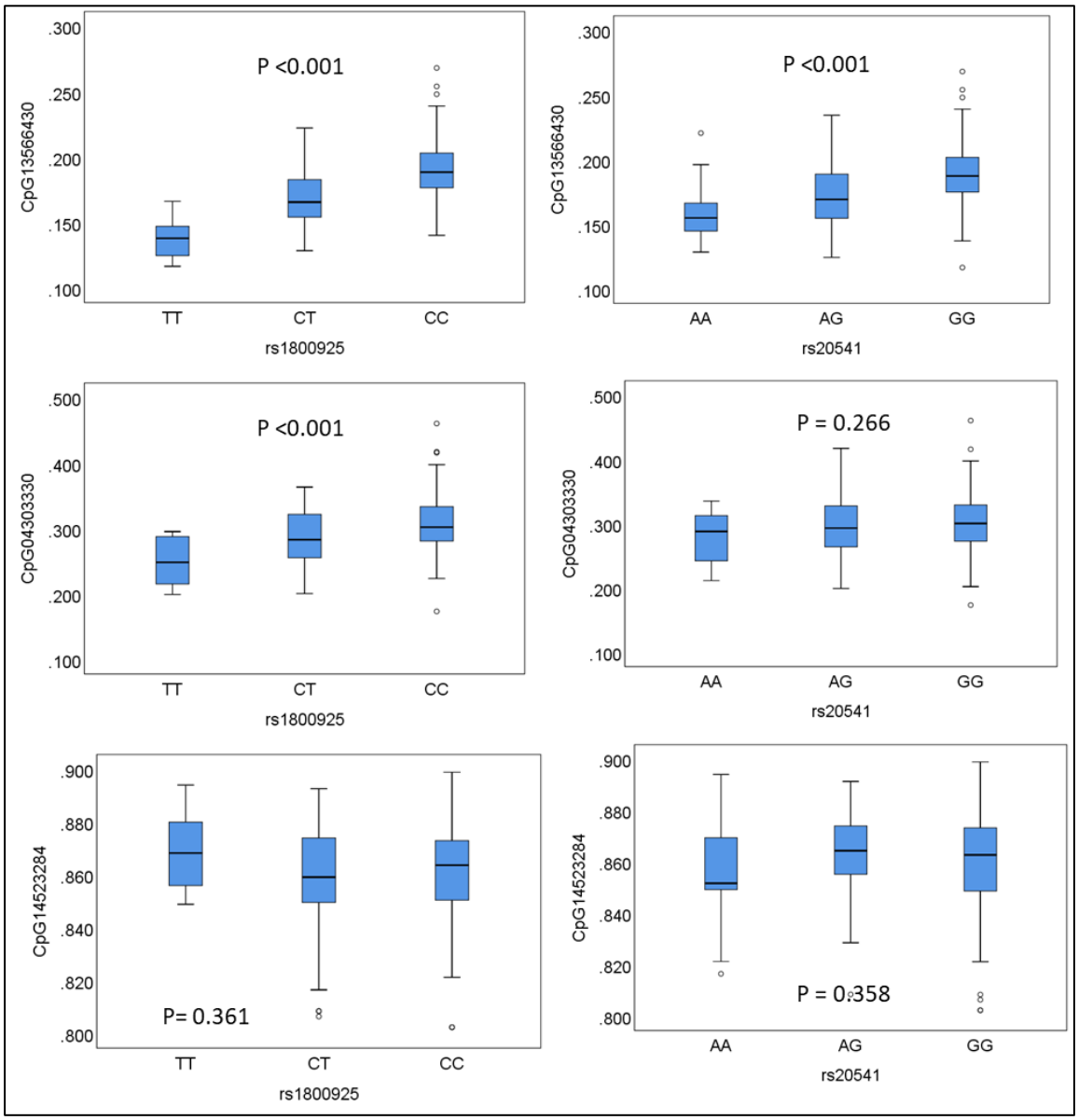
There was no significant effect observed for either the main effect of genotype of rs1800925 or DNA methylation at cg13566430 on Asthma at 18 years (defined as participants with a physician diagnosis of asthma plus current symptoms and/or asthma medication) as an outcome (Table 3-10). In addition, there was no significant interaction between genotype of rs1800925 or DNA methylation at cg13566430 on asthma.

Table 3-10 Interaction of methylation at cg13566430 and rs1800925 genotypes on Asthma at 18 years.

	Asthma at 18 years	
	Estimate	P
Main effects		
cg13566430	1.945	0.164
rs1800925	0.283	0.754
Interaction		
rs1800925	0.525	0.592
cg13566430	1.174	0.280
rs1800925* cg13566430	0.530	0.589

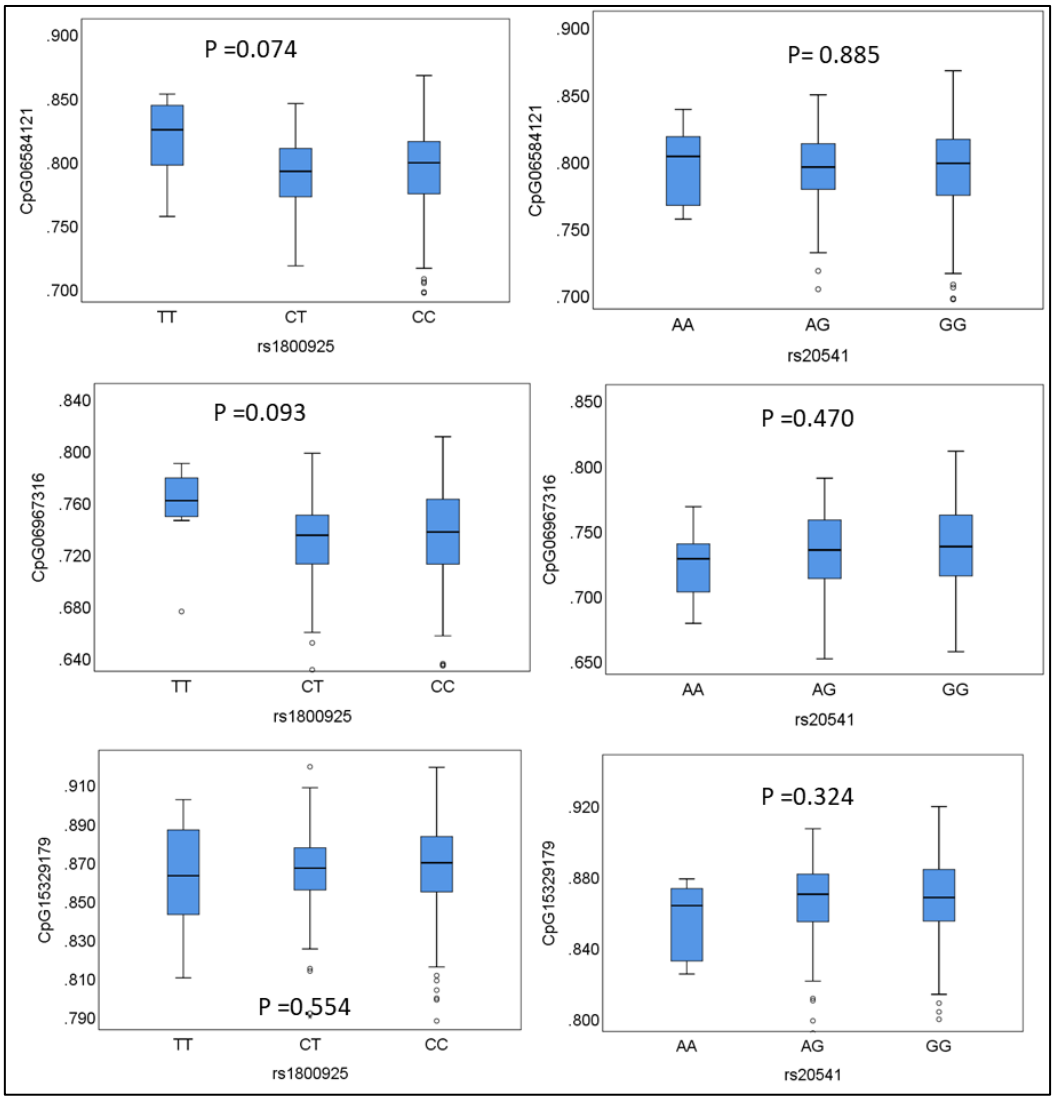
3.3.3.4 Genotype dependent DNA methylation

Methylation levels were significantly different ($p<0.001$) for cg13566430 across the genotypes of both rs20541and rs1800925 (Figures 3-5 and 3-6), and DNA methylation of cg0430330 was significantly different across rs1800925 only. Methylation at other CpG sites did not show genotype dependent difference.



P value is from Kruskal Wallis independent sample test

Figure 3-5 Methylation at CpGs (TSS1500) across rs1800925 and rs20541 genotypes



P value is from Kruskal Wallis independent sample test

Figure 3-6 Methylation at CpGs (TSS200) across rs1800925 and rs20541 genotype

3.3.4 Discussion

This study tested a two-stage model for integrating the interactions of maternal smoking during pregnancy, genetic variants and DNA-M for an asthma candidate gene *IL13*. Findings show that interaction of a functional *IL13* SNP, rs20541, and maternal smoking during pregnancy (an early life exposure known to be associated with asthma and lung functions) influenced DNA-M at cg13566430. Results also show that DNA-M at cg13566430 interacts with genotype of another functional SNP, rs1800925, to affect airflow limitation and airway reactivity. Michel et al (281) examined the effect of farm exposure on DNA-M of ten asthma candidate genes and found that DNA-M at one *IL13* site was more methylated in the exposed group compared to the non-exposed group. This differentially methylated site was the one spanning rs1800925; similarly, in the findings of this chapter, cg13566430 is close to rs1800925 in the promoter region of *IL13*. Michel et al did not see any significant differential methylation of *IL13* between asthmatic and non-asthmatic children but the interaction with genetic variants was not examined. The approach identifies the effect of environmental exposure and genetic variants on DNA-M and then also combines the interaction of other modifiable genetic variants with DNA-M on the outcomes. Results show genotype-dependent DNA-M in the promoter region of *IL13***;** methylation at cg13566430 revealed variation in levels dependent on the genotype of rs1800925. This genotype-dependent methylation has been shown in other genes (200). This varying distribution of DNA-methylation across genotypes supports the plausible role for DNA-methylation in the pathway between genotype and phenotype.

There was no significant effect seen on asthma at 18 years as an outcome; neither for main effects of genotype rs1800925 and DNA-M at cg13566430, or interaction of genotype and DNA-M. The results highlight that asthma is more of a complex syndrome with varied phenotypes than a binary diagnosis based on any clinical test. For example FEV1 can be normal in people with asthma diagnosis (282) particularly when the asthma is well controlled, also people with asthma

diagnosis might not exhibit airway reactivity and in epidemiological studies its known to have airway reactivity without asthma (283). These objective measures cannot be used interchangeably to diagnose asthma (284).

Only female participants were included in the analysis as methylation data was only available on female samples at the time of the analysis. However, the analysed samples were representative of female participants in the total cohort and there was no apparent selection bias. Results show significant interactions between polymorphisms and methylation at the promoter region. The levels of statistical significance are less than desired for a genetic association study, which is likely due to the small sample size.

This is a candidate gene approach where known gene and genomic variants (SNPs) are selectively studied which limits the generalisability of the results, however as one of the first proof of concept study the results add value. However, there is a need for further evaluation of these interactions in a large study. Non-wild type allele for rs1800925 (TT) was present in only 3.4% of the study sample adding to the generalisability of the findings and highlight the need for repetition of the study in a larger population.

While DNA methylation measurement using Illumina Infinium-HumanMethylation450 arrays has been shown to have reproducibility and high validity (182), technical replication of the DNA methylation measurements was not undertaken. DNA methylation was from peripheral blood and not airway tissue. it is now known that DNA methylation profiles can differ in different tissues; for example airway epithelial cells vs peripheral blood cells for asthma (285). Although the major source of IL-13 production in the airways is inflammatory cells, and methylation measurements in peripheral blood are likely to reflect asthma related immune mechanisms. Cell composition in peripheral blood can influence DNA methylation, however cell composition alone cannot explain the differential methylation observed (281).

In asthma, GWAS and epidemiological studies have not succeeded in explaining the heterogeneity in the phenotype. Because of its plasticity, methylation can be

established or modified by environmental exposure yet is stable enough to pass through mitotic cell divisions (286) and exert control on gene expression. Therefore, methylation might hold the key between the genotype and the environment interactions (52), bringing out the diversity in the phenotype. The effect of environment, genotype and DNA methylation was seen for two different, but equally important and objective characteristics of asthma, airway obstruction and airway reactivity. Differentially methylated regions for age-related phenotypes including lung function have been previously shown (200). Genotype-dependent methylation variation has also been shown by Crider et al. in an interventional study with folate supplementation in reproductive age females (201). Methylation of the promoter region can switch off transcription (287) and this differing distribution of methylation seen across genotypes supports the plausible role of methylation in the pathway between genotype and phenotype.

In summary, while requiring replication in an independent cohort, the results show the interplay of prenatal maternal smoking, genetic variants and DNA methylation of *IL13* influencing airway obstruction and airway reactivity. This highlights the need to consider environment, genotype and DNA methylation together when seeking to understand the pathogenesis of complex disease, as DNA methylation may play a role of integrator of multiple disease pathway signals. Future studies could examine the interaction between methylation and genotype at all genetic susceptibility loci identified for asthma, or even genome wide though this would require larger sample sizes.

Chapter 4 Third Generation Study

4.1 Introduction

The Isle of Wight Birth Cohort (Second generation, F1) was established in 1989 by contacting all pregnant women (First Generation, F0) between January 1989 and February 1990 and followed up through childhood, adolescent and adulthood (213), the cohort has been described further in Chapter 1. This well characterised prospective study with good follow up rates provided the platform for recruitment of the Third Generation Study (F2); children of the IOWBC.

The IOWBC has contributed towards understanding of the prevalence and the temporal changes in the allergic conditions from birth through childhood and early adulthood. There is growing evidence for early life origins of asthma and other allergic diseases which would require a prospective birth cohort with well measured prenatal exposures and accurate phenotyping of the parents. Another concept that has received attention in the field of asthma epidemiology is the concept of effects of environmental exposures being transmitted to next generations. The increasing in the understanding of the role for genetics in asthma, together with the evolution of epigenetics as a possible mechanism mediating environmental effects on asthma phenotype, requires further investigation. Establishing the Isle of Wight multigenerational study by recruiting the Third Generation will help to achieve this goal. Details of the Third Generation study methodology and assessments carried out are provided in chapter 2 and consents and questionnaires are in the Appendices 2 and 3.

4.1.1 Focus of the Third Generation Study

- Recruiting Participants of the F1 generation and their partners into the Third Generation study.
- Prospective collection of the exposure and disease status of the F1 generation (pregnant female participants and pregnant partners of the male participants) and prenatal exposures of the F2.
- Assessing genome wide DNA methylation at birth (F2)

- Follow up of the children at 3, 6 and 12 months with questionnaires and Skin prick test at 12 months. Now study has progressed to assess these children at 2, 3 and 6 years of age.

4.2 Recruitment

Third Generation study is a dynamic multigenerational cohort study, and recruitment is still on going. Children of the IOWBC have been recruited into the study since 2010. The description of the cohort presented in this chapter includes updated status until May 2017. In results chapters 5 and 6, a subset of this cohort has been used depending on the recruitment status and availability of data at the point of analyses.

F1 cohort participants, both men and women that already had children at the commencement of the study in 2010 were recruited into the Third Generation study as "postnatal recruits". Data on antenatal exposures were collected retrospectively in this group. Study recruitment started in May 2010 for postnatal recruitment where cohort participants had already informed the study team that they had children. Since July 2010, F1 cohort women, and partners of F_1 men who were pregnant, were recruited at 12 weeks of pregnancy (or as soon as possible during pregnancy after 12 weeks); "antenatal recruits". The recruitment is described in detail in the methods section. Briefly, pregnant F_1 cohort (participants or partners of the participants) were identified from antenatal clinics and additionally from responses to the regular study newsletters. The current recruitment status of the Third Generation study is described in Table 4-1.

A total of 436 children were recruited into the study; 315 (72.2%) recruited during pregnancy (antenatal) and 121 (27.8%) after the birth of the child (postnatal). Not having the DNA methylation at birth (from cord blood) is a limitation for postnatal recruits. To overcome this (after consenting the parents), access to the Guthrie cards was requested from the regional lab and Guthrie cards for 70/121 (57.9%) of postnatal recruits were able to be accessed. With the

availability of the methodology to extract DNA from the stored Guthrie cards (288), this will improve the numbers where epigenetic information is available, thus increasing the power of the study.

In the Isle IOWBC (F1), 48.8% (750/1536) were female and majority of recruitment into Third Generation study (F2) was from mothers belonging to this cohort. A quarter of Third Generation children were recruited from fathers who were members of the cohort, for rest of three quarters mother belonged to the cohort.

Table 4-1 Third Generation study recruitment

Total recruited	436
Recruitment Type	
Total recruited	436
Antenatal recruit	315/436 (72.2%)
Postnatal recruit	121/436 (27.8%)
Postnatal Recruit	
With Guthrie cards	70/121 (57.9%)
Without Guthrie cards	51/121 (42.1%)
Parent from IOWB cohort	
Only Mother	321/436 (73.6%)
Only Father	104/436 (23.9%)
Both Mother and Father	11/436 (2.5%)

4.3 Exposures

4.3.1 Prenatal factors

4.3.1.1 Home factors

At recruitment, mothers completed the Questionnaire (Appendix 3), which covered information on current living home. Location is the response to the question; "What term best describes where you live now?", air pollution exposure was derived from the response to; "How frequently are you exposed to outdoor air pollution (from traffic, industry etc) in your home if you keep the windows open?", and "How often do vehicles pass your house or on the street less than 100 meters away?" gave the information on exposure to vehicle pollution while inside the house.

Table 4-2 Location of home, exposure to air pollution and pets at home

Location of home	
Small Town	256/305 (83.9%)
City Centre	3/305 (1.0%)
Suburb of a city	14/305 (4.6%)
Rural Village	32/305 (10.5%)
Living in a farm	3/305 (1%)
Air pollution exposure	
Any air pollution exposure	173/305 (56.7%)
Everyday	125/173 (72.3%)
Once a week	29/173 (16.7%)
Once a month	10/173 (5.7%)
Once a year	9/173 9 (5.2%)
Passing by vehicles	
> 10 per hour	169/305 (55.4%)
1-9 per hour	84/305 (27.5%)
10 per day	36/305 (11.8%)
Seldom	16/305 (3.7%)
Pet at home	
Any pet at home	185/306 (60.5%)
Cat at home	98/185 (53%)
Dog at home	90/185 (48.6%)
Both cat and dog at home	33/185 (17.8%)
Other pet at home	65/185 (35.1%)

4.3.1.2 Socioeconomic status (SES)

Around half of the pregnant mothers were living in a privately rented house (49%), about a quarter in an owned house (27%) and rest were living in a council or housing association rented house. Table 4-3 gives the details of the socio-economic factors including family incomes of the pregnant females (F1).

Table 4-3 Maternal (F1) Socio-economical descriptions

House	
Owned	84/306 (27.4%)
Rented Privately	151/306 (49.2%)
Rented from Council/Housing association	57/306 (18.6%)
Other	15/307 (4.9%)
Number of bedrooms	
<2	27/307 (8.8%)
2-3	252/307 (82.1%)
>3	25/307 (8.1%)
Living with the Father of the baby	259/306 (84.6%)
Still in Education	26/355 (7.3%)
Currently Employed	216/354 (61%)
Family Income	
Less than £12,000	80/306 (26.1%)
£12,000 – £17,999	55/306 (18%)
£18,000 - £29,999	101/306 (33%)
£30,000 – £41,999	29/306 (9.5%)
Greater than £42,000	13/306 (4.2%)
Did not know/want to reveal	28/306 (9.2%)

In EWAS analysis (chapter 6) maternal SES was included as confounder and was described based on clusters. Five maternal SES clusters were identified with variables of family income, house: number of bedrooms and maternal education (statistical methods in chapter 2). Clusters are described in Table 4-4. The lowest SES cluster (=1) had low level of household income, and low number of rooms in the house and low level of maternal education. The second SES cluster (=2) had

low level of household income, low number of rooms and low-to-medium level of maternal education. The third SES cluster (=3) had low-to-medium level of household income, low number of rooms and high level of maternal education. The fourth SES cluster (=4) had medium level of household income, high number of rooms and low-to-medium level of maternal education. The highest SES cluster (=5) had high level of household income, high number of rooms in the house and medium level of maternal education.

Table 4-4 Maternal SES clusters

Cluster	Income	# rooms	Education	Frequency
1 (Low)	Low	Low	Low	77
2 (Medium 1)	Low	Low	Low to medium	97
3 (Medium 2)	Low to medium	Low	High	114
4 (Medium 3)	Medium	High	Low to medium	61
5 (High)	High	High	Medium	51
Missing	.	.	.	36

4.3.1.3 Prenatal smoking exposure

Maternal smoking during pregnancy is a prenatal exposure known to influence asthma outcomes in the child and it is the focus of interest in chapters 3 (F0 smoking during pregnancy) and 5 (F1 smoking during pregnancy) of this thesis. As seen in Table 4-5 the denominator is different for different trimesters based on the questionnaires completed at each point; at 12 weeks, 20 and 28 weeks of pregnancy. Approximately a quarter of mothers reported smoking at each time point and throughout pregnancy.

Table 4-5 Prenatal smoking exposure

Maternal smoking during pregnancy	
Smoking in Trimester 1	82/356 (23.0%)
Smoking in Trimester 2	92/379 (24.3%)
Smoking in Trimester 3	77/274 (28.1%)
Maternal smoking pattern during pregnancy	
Never Smoked	253/399 (63.4%)
Smoked during Pregnancy	146/399 (36.6%)
Early Pregnancy	22/146 (15.1%)
Transient	86/146 (58.9%)
Persistent	38/146 (26.0%)
Father Smoked during pregnancy	164/354 (46.3%)
Others smoked in the house during pregnancy	38/354 (10.8%)

4.3.1.4 Birth factors

Mode of delivery has been previously studied as an exposure influencing allergic conditions in later childhood but has regained more importance since an appreciation of the role of microbiome in asthma and allergic conditions. Passage through the genital tract of the mother during a normal vaginal birth exposes the foetus to the microbial environment of the mothers' birth canal thus establishing a varied microbiome on skin, gut and mucous membranes of the child at birth. Birth through caesarean section bypasses the microbial exposure of the mother's genital tract. The mode of delivery was collected from the hospital notes for which mothers consented. Table 4-6 gives the details of the mode of birth and the birth order for Third Generation infants.

Table 4-6 Mode of delivery

Delivery method	
Vaginal delivery	230/336 (68.5%)
Forceps or/and Vacuum assisted	36/336 (11.9%)
Planned C-section	18/336 (5.4%)
Emergency C-section	46/336 (13.7%)
Birth order	
1	233/409 (57%)
2	126/409 (30.8%)
>2	50/409 (12.2%)

4.3.1.5 Parental asthma

Asthma in parents (F1) was measured using ISAAC questions; "Have you had wheezing or whistling in the chest in the last 12 months?" Or/and "Have you ever had asthma?"

Table 4-7 Parental history of asthma in Third generation

Maternal asthma	
Current wheeze	41/255 (16.1%)
Ever asthma	37/255 (14.5%)
Current wheeze and ever asthma	28/255 (10.9%)
Paternal Asthma	
Current wheeze	65/242 (26.9%)
Ever Asthma	81/242 (33.5%)
Current wheeze and ever asthma	46/242 (19%)

4.3.2 Follow up of the Third Generation study

The initial protocol and the supporting National Institutes of Health grant was for the recruitment and follow up of the Third Generation population at 3,6 and 12 months. Subsequently ethic amendments were made to extend the follow up to 2, 3 and 6 years of age, which is currently ongoing. This thesis is based on data collected at 3, 6 and 12 months. The follow up rate is better at 3 months; this is mainly because the postnatal recruited children had only 3 and 12 months' assessments (Table 4-8).

Table 4-8 Status of follow up of the Third Generation

Born (by May 2017)	414/436 (94.4%)
>3 months	400
> 6 months	390
> 12 month	362
Three months follow up	354/400 (88.5%)
Six months follow up	194/390 (49.7%)
Twelve months follow up	228/362 (62.9%)

4.3.2.1 Infantile Eczema

Disease status was based on the parent's response to the validated ISAAC questions for eczema and wheeze. Reported eczema based on the response to "Has your child ever had a rash or eczema that has lasted for at least 7 days or more? (do not count typical nappy rash)" and "Has your child ever had eczema?".

Table 4-9 Parent reported eczema

	Eczema
At 3 months	28/353 (7.9%)
At 6 months	27/194 (13.9%)
At 12 months	32/228 (14%)
Any eczema in the first year	58/366 (15.8%)

4.3.2.2 Infantile wheeze

Wheeze was measured by the responses to validated ISAAC questionnaires (289); "Has your child had wheezing or whistling in the chest?" and if the response is yes then parents were asked; "Does your child wheeze in between cold or chest infection?" Wheeze during early life can be multifactorial and the common reasons are viral infections/bronchitis and asthma. In this thesis wheeze is defined as the wheeze reported in between colds or infections. The same definition is used in results chapter 5.

Table 4-10 Parent reported wheeze

Wheeze between cold/infections	
At 3 months	45/352 (12.8%)
At 6 months	27/194 (13.9%)
At 12 months	26/228 (11.4%)
Wheeze in the first year	67/365 (18.4%)

4.3.2.3 Allergen sensitivity

Allergen sensitivity was measured by skin prick test (SPT); methods are described in detail in chapter 2. SPT was carried out in 276 children at 12 months follow up, 10 of them had less than standard response to positive control (histamine) which had to be excluded. Three children underwent SPT for only aero-allergens and not for food allergens. Table 4-11 gives the detailed results of the SPT.

Table 4-11 Skin prick test results in the Third Generation

Number of Skin Prick Test	276	
	Cut off ≥ 2.5 mm	**Cut off ≥ 3 mm**
Inadequate reaction to Histamine	10/276 (3.6%)	31/276(11.2%)
Aeroallergen sensitivity	**27/276 (9.8%)**	**10/276 (3.6%)**
HDM	16/276 (5.8%)	5/276 (1.8%)
Cat	5/276 (1.8%)	2/276 (0.7%)
Dog	4/276 (1.4%%)	2/276 (0.7%)
Grass	6/276 (2.2%)	3/276 (1.1%)
Alternaria	1/276 (0.4%)	0
Tree	0	0
Food-allergen sensitivity	**13/273 (4.8%)***	**9/273 (3.3%)**
Milk	2/273 (0.7%)	2/273 (0.7%)
Egg	8/273 (2.9%)	4/273 (1.5%)
Peanut	4/273 (1.5%)	3/273 (1.1%)
Sesame	1/273 (0.4%)	1/273 (0.4%)
Wheat	0	0
Cod	0	0

4.4 Challenges and limitations

4.4.1 Recruitment and follow up

The study is dependent on the F1 cohort participants having children, thus we approached the whole cohort and informed them about the study. One year into the study, it was realised that significantly fewer children of F1 male participants were recruited. This issue was discussed with the team and discussed with the antenatal department. As a result of these discussions, the screening methods were changed. This led to improvement in recruitment of children from the fathers (male participants) of the original cohort, however majority of the recruitment remained from the mothers belonging to the IOWBC.

Some parents of Third Generation study offspring are very young and have varied family circumstances. Even though this was an observational study with no interventions involved, it became apparent that keeping these parents involved, and conducting all the assessments, was challenging. With the help of the research nursing team, a tailored approach was initiated, including giving the option of the research nurse undertaking home visits for the assessments; this helped with the retention of study participants with minimal dropouts.

Third Generation study data is dynamic, meaning it keeps increasing with time as new recruitment to the cohort and later age follow ups continue. Children will be of different ages and will be at different stages of the study at any one point in time. This poses the challenge of using a snapshot of the data at a particular time of the study. There will be some missing information as noted for parental history of asthma which had large missing information, this is because only the data where asthma questionnaire information was available for both father and mother is included. A major part of the missing information here is explained by the challenges of recruiting fathers. There is risk of self-selection bias in cohort studies, as noted here mothers are more likely to attend the follow up assessments than fathers. Similar patterns are also noted in the IOWBC where girls were more likely to attend the18 years follow up than boys (290). In

addition, among fathers that are already recruited it appears that asthma rate is higher, again adding to the likely self-selection. This modest bias should not have a major impact on the results as the numbers are expected to increase, and more fathers will be recruited into the study. Maintaining a high retention rate at the follow ups will also help to minimise the impact of self-selection. As mentioned above, mitigating actions have been taken to improve the recruitment and retention into the study. It is also possible that men might be having children later than women and we might see an increase in children of the fathers from IOWBC. Another bias inherent to the study is of younger parents particularly young mothers. When the recruitment started in 2011 mothers belonging to the cohort were 20-21 years. This must be considered when studying maternal and pregnancy related issues as they would differ with differing ages, younger aged mothers are more likely have lower education, socioeconomical status and higher smoking exposure. However, when studying effect of exposures across generations this should not have major impact. Also, as the recruitment of the study has continued the study will have mothers ranging from 20 to 30 years.

4.4.2 Measurement of the exposures and outcomes

Accurate measurement of exposures is essential for an epidemiological study of a condition like asthma and allergy. Various environmental factors can influence the outcome, thus in addition to the exposures in focus, a large number of concurrent exposures can confound the effect on the phenotype. Missing information can also be a limitation. As observed in the Third Generation study, some measurements were missing, particularly those related to paternal exposures. This may influence analyses of association of certain factors for instance, prenatal and postnatal exposures on the DNA methylation at birth.

Characterisation of the phenotype in a heterogeneous condition like asthma is important. Diagnosing asthma in early childhood is challenging, however, parent reported symptoms like recurring wheeze, known to increase the risk of the development of the condition later in life (291) are used as a surrogate phenotype in large asthma epidemiological studies (225, 289). However, more accurate

characterisation will be conducted in later childhood through follow up at 6-7 years. Objective measures like SPT and serum IgE will also help improve characterisation of the phenotypes. Another strength of the study is its prospective nature which will minimise the potential recall bias in questionnaire based information.

4.4.3 Generalisability of the results

The cohort is based on an island; the Isle of Wight is the largest island in England roughly 150 square miles in size. There might be limited diversity in environmental exposure, for example, most of the island is semirural in nature and there is no heavy industrial exposure, which might limit generalisability of the study results for such exposures. Prevalence of allergic conditions and time trends of the prevalence on the island have been comparable to that reported for the mainland (11) which supports the generalisability of the overall results. There is no evidence of genetic inbreeding as the genetic profile of IOWBC cohort participants is similar to that of the mainland populations and therefore, this should not affect the generalisability of the results of the study.

The Third Generation recruitment is steadily growing in numbers and will provide a unique opportunity to study the intergenerational effects of environmental exposures on disease outcomes. Establishing methylome data at birth and at later life will help in investigating if the observed DNA methylations are secondary to exposures or the disease status. Other samples like placenta, faeces, skin swabs, and serum will help in other studies looking into the role of microbiome and metabolome in the development of asthma.

Chapter 5 Multigenerational effect of smoking during pregnancy; grandmother to grandchild.

5.1 Introduction

Asthma is multifactorial; genetics, environmental exposures, and their interactions play an important role in development of asthma. It is known that environmental exposures, particularly prenatal exposures, influence development of asthma. However, it is unknown whether these effects of environmental exposures on health can be passed across generations. This phenomenon is termed 'multigenerational effects' or 'transgenerational effects'. While these terms are sometimes used interchangeably, they are referring to distinct biological processes. Arshad et al. (138) define the terms as shown in Table 1-1 in chapter 1.

Evidence for multigenerational effects of environment on phenotype in animals and humans is detailed in chapter 1 and here the relevant examples are summarized. Multigenerational effects for asthma and allergy was first demonstrated in a mouse model by Holingsworth et al. (292) , where supplementation of maternal diet influenced the severity of allergic airway disease in multiple generations. The strongest evidence for multigenerational and/or transgenerational effects of environmental exposures in humans has been demonstrated in the Overkalix cohort; in this study paternal grandfather's food supply in pre-adolescence was linked to the mortality risk ratio of grandsons, while the paternal grandmother's food supply was linked to the mortality risk ratio of the granddaughters (133). Furthermore, a surfeit supply of food during slow growth phase was associated with increased relative risk for mortality in grandchildren (134).

Tobacco smoke exposure is an important determinant associated with the development of asthma. It is well understood that personal smoking exposure is known to be associated with lung functions and asthma, there has been growing evidence that *in-utero* exposure to tobacco smoking is a risk factor for poor lung function and asthma in children. Gilliland et al. showed that maternal smoking during pregnancy is associated with physician diagnosed asthma and wheeze in childhood (293). In mice, Rehan et al. demonstrated that prenatal nicotine

exposure influenced lung function not only in next generation but the effect persisted in the following generation (grandchildren) as well, even though they were not directly exposed to nicotine (132). Results from animal models raise the question if similar multigenerational effects of in utero smoking exposure are seen in humans too. Li et al. in 2005 showed that not only child`s prenatal exposure but also mother's prenatal exposure is a risk factor for asthma in children (135), this was the first epidemiological evidence for the multigenerational effect of prenatal smoking on development of asthma in children. The multigenerational effects of prenatal smoking exposure on asthma seen both in animal and human studies raise the question of the possible mechanism involved.

There is evidence that epigenetics could be a plausible link between environment and genetics (52). Joubert et al. showed differential methylation across the genome in newborns in relation to maternal smoking during pregnancy (197). The study identified differential methylation at 26 CpG sites across 10 genes associated with maternal smoking during pregnancy. Rehan et al. (132) examined global DNA methylation and histone acetylation in their multigenerational smoke exposure model and proposed epigenetic mechanisms as possible conduit for the multigenerational effect of tobacco smoke.

Can epigenetic markers, specifically DNA methylation (which can be influenced by environmental exposure), play a role in mediating the influence of environmental exposure on phenotype and in transferring the effects through multiple generations? There is growing evidence, particularly in animal models, that epigenetics is the possible mechanism for this multigenerational inheritance and is termed as multigenerational epigenetic inheritance (294). To investigate this, the multigenerational birth cohort, a prospective longitudinal birth cohort study exploring the effect of maternal (F1) and grand-maternal (F0) smoking during pregnancy on non-infectious wheeze in infants (F2) was utilised. As a proof of concept study, DNA methylation at candidate loci in cord blood of the offspring is examined for the effect of smoking by mother and grandmother during pregnancy.

5.2 Methods

5.2.1 Study Sample

The IOWBC is described in section 1.5. The recruitment and assessment of the Third Generation Study is described in detail in section 2.2. In summary the multigenerational cohort involves information on three generations, IOWBC (F1) with prospective information from their parents (F0) and the Third Generation (F2).

The Third Generation study is an ongoing study with recruitment and follow up still on going. This poses a challenge, i.e. to define "a complete dataset" and "end of study". For the purposes of this thesis, data collected till end of April 2014 were analysed and results in this chapter are based on this data. Table 5-1 describes the study status in April 2014.

Table 5-1 Recruitment status by end of April 2014.

Total number	252
Male: Female	56.8% : 43.2%
Recruitment type	
Antenatal Recruitments	69% (174/252)
Babies born among Antenatal recruits	86.8% (151/174)
Postnatal Recruitments	31% (78/252)
Lost to follow up	2% (5/252)
Children of IOWBC girls	82.9% (209/252)
Children of IOWBC boys	18.2% 46/252
Both parents from IOWBC	3
Number of twins	3

IOWBC; The Isle of Wight Birth Cohort (1989-90)

When recruitment for the Third Generation study started, some of 1989 birth cohort participants already had children. Those with children already born were recruited and the pregnancy questionnaires were completed retrospectively. As the children were already born, cord blood samples were not available for postnatal recruited participants. In view of these limitations, only data from antenatally recruited samples is included in the analyses. For the multigenerational study, looking at the effect of maternal and grand-maternal smoking during pregnancy, the subset for analysis had to be restricted to the third generation children (F2) whose mothers belonged to the original cohort (F1); thus excluding children whose father belonged to the F1 cohort (born to Isle of Wight cohort (F1) boys), as the prospective data on non-cohort mothers' disease status was not available and additionally no information about grand-maternal smoking was available in this subset. The subset analysed is described in Table 5-2. The definition of exposure was persistent smoking during pregnancy and the parent reported wheeze in between infections or cold was used as the phenotype outcome for the purpose of this study, specific questions are given in chapter 4 and the questionnaires are attached to appendix 3.

5.3 Results

5.3.1 Study sample description

The IOWBC (1989-90) is a population based unselected birth cohort and thus question of selection bias does not arise. However, the multigenerational study was restricted to a subset of the Third Generation cohort; those with antenatally recruited third generation offspring with mothers belonging to the original cohort. Table 5-2 gives the description of the sample in comparison with the whole Third Generation cohort. The sample analysed is representative of the whole cohort in terms of sex of the child, maternal and paternal history of asthma and smoking exposures. There was also no significant difference in wheeze outcomes between the subset and the whole cohort.

5.3.2 Infant Wheeze

Parental reported wheeze in children was collected prospectively at the 3, 6 and 12 months follow up visits. Overall, 60.8% reported any wheeze during first year of life (Infant wheeze); 31.4% reported wheeze only with cold and infections and 29.4% reported wheeze without any cold or infections (Table 5-2). Diagnosing asthma is difficult in infants and in this study reported wheeze between cold/infections has been used as the outcome phenotype.

Table 5-2 Description of the multigenerational study subset and the whole Third Generation Cohort

Character	Antenatal and Cohort mother	All	p value
Sex (Boys)	50.8 (65/128)	56.8% (130/229)	0.319
Mother Ever had asthma	36.2 (46/127)	36.8% (81/220)	1
Father ever had asthma	31.0 (27/87)	33.8% (45/133)	0.769
Mother smoking	42.4 (50/118)	48% (97/202)	0.354
Grandmother smoking	38.1 (51/134)	37% (87/235)	0.911
Father smoking (F1)	47.5 (56/118)	50.75 (105/207)	0.645
Grandfather smoking (F0)	44.7 (38/85)	47.2% (60/127)	0.779
Postnatal parental smoking (F2)	51.0 (25/49)	44.9% (40/89)	0.593
Reported wheeze in first 12 months			
Wheeze only with infection/cold (F2)	31.4 (32/102)	28.9 (57/197)	0.690
Wheeze between infection/cold (F2)	29.4 (30/102)	25.9 (51/197)	0.583
Any wheeze (F2)	60.8 (62/102)	54.8 (108/197)	0.389

P value from chi square test

5.3.3 Smoking during pregnancy

Overall, 42.4% of F2 mothers smoked at any point during pregnancy; 17.8% smoked only in first trimester, 2.5% smoked in first and second trimester and 22% of mothers reported smoking throughout the pregnancy. 22% of mothers reported stopping smoking before pregnancy and 35.6% never smoked. For the analysis of multigenerational effect of prenatal smoking persistent smoking throughout the pregnancy was taken as the exposure. Table 5-3 gives the description of maternal and grand-maternal smoking during pregnancy; for 20.3% of F2 infants, both their mother and their grandmother smoked during pregnancy; for 17.8% only the grandmother smoked while pregnant with the mother, but the mother did not smoke while pregnant; and 22% only the mother and not the grandmother smoked during pregnancy. Combined maternal and grandmaternal smoking during pregnancy are described in Table 5-4.

Table 5-3 Maternal smoking patterns in Third Generation study

Smoking patterns	% (n/N)
Never smoked	35.6 (42/118)
Past smoker-stopped before pregnancy	22.0 (26/118)
Smoked only in 1st Trimester	17.8 (21/118)
Smoked in Trimesters 1 and 2	2.5 (3/118)
Persistent smoking throughout pregnancy	22.0 (26/118)
Any smoking in pregnancy	42.4 (50/118)

Table 5-4 Combined Maternal (F1) and Grand-maternal (F0) smoking during pregnancy

Smoking pattern	% (n/N)
No smoking during pregnancy-both F1 and F0	39.8 (47/118)
Only F0 smoked in pregnancy	17.8 (21/118)
Only F1 smoked in pregnancy	22.0 (26/118)
Both F0 and F1 smoked in pregnancy	20.3 (24/118)

5.3.4 Maternal and Grand-maternal smoking affecting wheeze

Univariate analysis for maternal and grand maternal smoking as exposure and infantile wheeze (between cold/infection) did not show significant association between maternal only or grand maternal only smoking during pregnancy with wheeze in F2. Combined F0 and F1 smoking during pregnancy was significantly associated with wheeze in F2; OR 3.4 (95% CI 1.1 to 10.9, P- 0.039). On multivariate analysis this association remained significant with Odds Ratio of 3.5 (95% CI 1 to 11.9), p value of 0.041 (Table 5-5).

Table 5-5 Combined maternal and Grand-maternal smoking affecting wheeze

	No wheeze	**Wheeze**	**Crude OR (95%CI)**	**p**	**Adjusted OR (95%CI)**	**p**
No smoking during pregnancy-both F1 and F0	47.9 (30/71)	32.1 (9/28)	R	R	R	R
Only F0 smoked in pregnancy	18.3 (13/71)	14.3 (4/28)	1.16 (0.30 – 4.44)	0.826	1.59 (0.39 – 6.53)	0.520
Only F1 smoked in pregnancy	19.7 (14/71)	21.4 (6/28)	1.62 (0.49 – 5.41)	0.434	2.07 (0.57 – 7.49)	0.266
Both F0 and F1 smoked in pregnancy	14.1 (10/71)	32.1 (9/28)	3.40 (1.06 – 10.87)	0.039	3.54 (1.05 – 11.94)	0.041

Logistic regression analysis. OR; Odds ratio, 95% CI; 95% confidence interval, R; reference category for the three F0 and F1 smoking variables. Adjusted OR; corrected of maternal, grand maternal smoking during pregnancy and parental history of asthma in F1 and F2

5.3.5 Effect of Maternal (F1) and Grandmaternal (F0) smoking on cord blood DNA methylation (F2)

Candidate loci were identified from the study by Joubert et al. showing differential methylation with maternal smoking during pregnancy (197). At the time of analysis, this was the largest study of the effect of maternal smoking on cord blood DNA methylation available. After preprocessing of the DNA methylation data (standardization, normalization and batch corrections as described in Chapter 2) data was available for 15 out of the top 26 CpG sites identified to be differentially methylated by Joubert et al. In cord blood of F2 16 CpG sites were explored for differential methylation in relation to maternal and grand-maternal smoking during pregnancy.

After categorising the maternal smoking into four groups; only maternal, only grandmaternal, both and neither smoked during pregnancy, with "neither maternal nor grand maternal smoking" as reference, there was no significant association seen for maternal only or grand-maternal only smoking with DNA methylation at any CpG sites. For both maternal and grand-maternal smoking 6 CpG sites showed significant association without correction for multiple testing. Using a multiple testing corrected p value (0.05/15(sites)=0.0033) only two sites remain significant: cg05575921 in *AHRR* and cg04180046 in *MYO1G* (Table 5-6).

Table 5-6 Combined maternal (F1) and grand-maternal (F0) smoking influencing DNA methylation in cord blood (F2)

		only mother smoking vs none			only grandmother smoking vs none			Both smoking vs none		
CpG	**Gene**	**Coef**	**SE**	**p**	**Coef**	**SE**	**p**	**Coef**	**SE**	**p**
cg21161138	AHRR	-0.004	0.010	0.667	-0.007	0.012	0.574	-0.25	0.012	0.035
cg23067299	AHRR	0.004	0.011	0.738	-0.016	0.014	0.237	0.012	0.013	0.379
cg05575921	AHRR	-0.013	0.014	0.378	-0.005	0.016	0.752	-0.055	0.017	0.002
cg03991871	AHRR	-0.008	0.013	0.550	-0.32	0.015	0.041	-0.006	0.010	0.571
cg12803068	MYO1G	0.019	0.026	0.470	0.025	0.026	0.334	0.071	0.027	0.012
cg04180046	MYO1G	0.016	0.018	0.386	0.023	0.019	0.238	0.081	0.023	0.001
cg05549655	CYP1A1	0.015	0.013	0.247	-0.004	0.012	0.760	0.016	0.014	0.263
cg11924019	CYP1A1	0.007	0.017	0.697	-0.007	0.017	0.692	0.009	0.018	0.614
cg14179389	GFI1	0.004	0.029	0.889	0.028	0.026	0.277	0.028	0.026	0.277
cg06338710	GFI1	0.033	0.044	0.452	0.022	0.047	0.644	-0.071	0.046	0.127
cg09662411	GFI1	0.016	0.038	0.667	0.028	0.038	0.465	-0.077	0.040	0.059
cg09935388	GFI1	0.017	0.048	0.729	0.048	0.045	0.293	-0.152	0.049	0.004
cg10399789	GFI1	0.028	0.034	0.406	0.020	0.035	0.575	-0.027	0.033	0.430
cg25949550	CNTNAP2	-0.005	0.006	0.407	0.000	0.005	0.947	-0.013	0.006	0.023
cg12477880	RUNX1	0.009	0.032	0.773	-0.001	0.035	0.970	0.027	0.035	0.450

List of CpG`s from Joubert et al. (197), Linear regression analysis. Coef; regression coefficient, SE; Standard error of regression coefficient, p; p value. Correction for multiple testing; 0.05/15(sites)=0.00

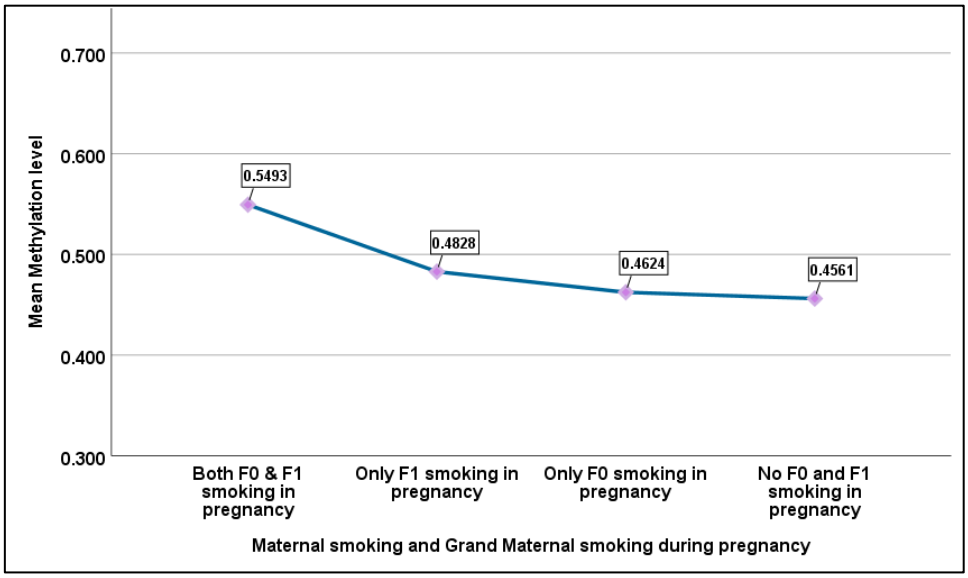


Figure 5-1 Differential methylation at cg04180046 in MYO1G in relation to maternal and grand-maternal smoking during pregnancy

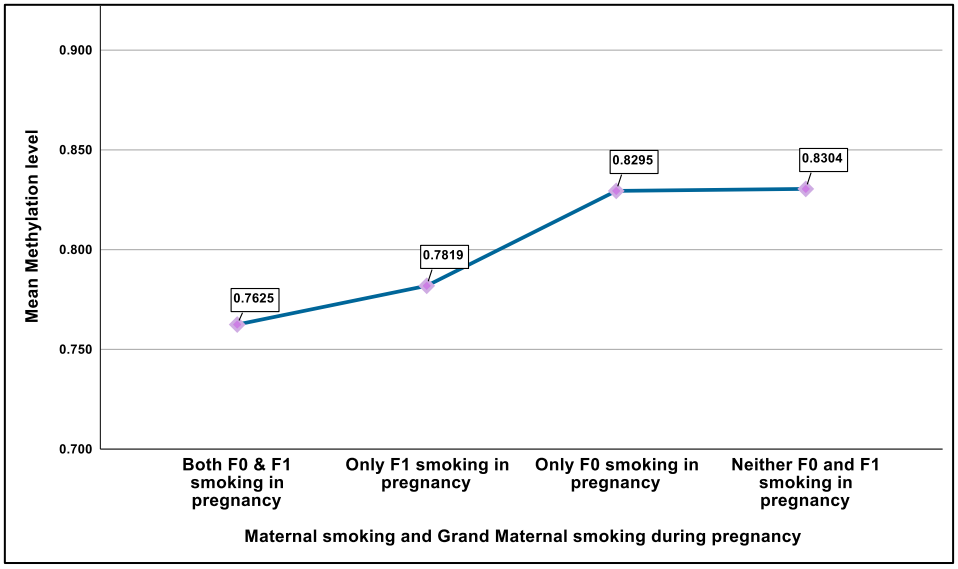


Figure 5-2 Differential methylation at cg05575921 in *AHRR* in relation to maternal and grand-maternal smoking during pregnancy

Mean methylation at *AHRR* cg05575921 was lowest when both mother and grandmother smoked during pregnancy at 0.76% compared to neither smoking 0.83%. As seen in Figure 5-1, a stepwise decrease in methylation was seen with grand-maternal, maternal and both smoking during pregnancy. Whereas for *MYO1G* cg04180046 methylation increased in steps from neither smoked (0.45%) through grand-maternal smoking, maternal smoking and being highest for both smoking during pregnancy (0.55%) (Figure 5-2).

For the two significant CpG sites, effects of any maternal and grandmaternal smoking were analysed for main effects and the interaction; interactions were not significant for either of the sites. For cg04180046 both main effects were significant and for cg05575921 only maternal smoking effect was significant (Table 5-7).

Table 5-7 Main effects and interaction of maternal and grand-maternal smoking during pregnancy on top two CpG sites

		Mother smoking		Grandmother		Interaction	
CpG	**Gene**	**Coef**	**p**	**Coef**	**p**	**Coef**	**p**
cg04180046	MYO1G	5.259	0.025	7.587	0.007	1.727	0.193
cg05575921	AHRR	6.269	0.014	3.124	0.081	2.120	0.150

5.3.6 Methylation at cg04180046 and cg05575921 in relation to reported wheeze

Methylation at cg04180046 in *MYO1G* and cg05575921 in *AHRR* were explored in relation to parent reported wheeze between infections or cold with linear regression analysis. Mean methylation at cg04180046 was lower (0.464) in the no wheeze group and higher (0.485) in the wheeze group. For cg05575921 mean methylation was higher (0.823) in the no wheeze group and lower (0.805) in the wheeze group. The differential methylation at both CpG sites in relation to wheeze was not statistically significant (Table 5-8).

Table 5-8 Differential methylation (non-significant) at cg04180046 and cg05575921in relation to reported wheeze

		Infant Wheeze*	
CpG	**Gene**	**Coef**	**p**
cg04180046	MYO1G	1.379	0.171
cg05575921	AHRR	-1.503	0.137

* Reported wheeze between infections/cold, Coef: regression coefficient.

5.4 Discussion

In a prospectively recruited longitudinal birth cohort (IOWBC, F1) and their children (Third Generation Study, F2), the effect of smoking during pregnancy by mother (F1) and grandmother (F0) when she was pregnant with mother of third generation (F2) child, on parent reported non infective infantile wheeze was explored. The study was undertaken in a subset of the cohort (F2) who were recruited prenatally and where the mother belonged to the F1 cohort. The results suggest an increased incidence of wheeze in the 1st year when both the mother and grandmother smoked during pregnancy. To explore the mechanism of this multigenerational effect of grand-maternal prenatal smoking, DNA methylation at candidate CpG sites previously identified to be differentially methylated with maternal prenatal smoking exposure were explored in relation to both maternal and grand-maternal smoking during pregnancy. This proof of concept study showed a trend of additive effects of combined maternal and grandmaternal smoking on DNA methylation at CpGs in *AHRR* and *MYO1G*. However, there was no significant change in methylation noted with only maternal smoking during pregnancy which is likely due to the limitation of the sample size resulting in low power.

In relation to wheeze, no significant differential methylation at CpGs in *AHRR* and *MYO1G* was observed. However, the direction of change in methylation was similar to that seen with maternal and grandmaternal smoking during pregnancy. At cg04180046 in *MYO1G* methylation level was lower in no smoking exposure group compared to smoking exposure group, similar direction of change in methylation was noted between wheeze (higher) and no wheeze (lower) groups. At cg05575921 in *AHRR* the direction of change was higher in smoking exposed and lower in no smoking exposure, similar direction of change noted in wheeze (lower) and no wheeze (higher) groups. However, no definitive conclusions can be drawn regarding the causal relationship between maternal smoking in pregnancy, DNA methylation and infant wheeze given the small sample size and the confounding between in utero and postnatal smoke exposure.

Wheeze in early childhood can be multifactorial, for example, viral bronchitis, chest infection and underlying asthma can result in wheeze, and it is difficult to diagnose asthma with confidence in this age group (83, 295). Early onset recurrent wheeze is associated with development of asthma and bronchial hyperresponsiveness in later life (291). Wheeze reported when child does not have a cold or infection was used to try and capture noninfectious early wheeze.

In a case control study within the Children`s Health Study in southern California, Li et al. showed for the first time that not only maternal smoking, but also grandmother smoking during pregnancy is a risk factor for asthma in children (135). The grandmaternal prenatal smoking exposure was collected retrospectively by asking the mother if her mother smoked during pregnancy which does bring in recollection bias. The Third Generation Study is prospective in nature where maternal prenatal smoking exposure was collected from the grandmother at the birth of mother, which limits recollection bias. Having only included the maternal line of the cohort limits the missing information from paternal grandmother smoking during pregnancy. However, this does mean not accounting for the father`s prenatal smoking exposure, which might have impact on infant wheeze (136). Miller et al. study did not show similar maternal and grand-maternal effects on childhood asthma, however, they did show that father`s prenatal exposure (paternal grandmother smoking during pregnancy) was significant particularly in girls (136). More recently paternal smoking before conception of their offspring has been shown to be associated with childhood asthma (296) and lung function (297) in a large multigenerational multi-center cohort.

Epigenetic mechanisms, particularly DNA methylation, have been postulated to account for the impact of prenatal smoking exposure on the development of asthma and allergies (298). First EWAS looking at genome wide differential methylation in cord blood related to prenatal exposure was in Norwegian Mother and Child Cohort Study (MoBa) using the Infinium HumanMethylation450 BeadChip (450K) (197), identified 26 CpG sites mapped to 10 genes to be differentially methylated. Methylation levels of *AHRR* cg05575921, the top hit

CpG, decreased in a dose-dependent manner with smoking exposure, and the trend was statistically significant. The aryl-hydrocarbon receptor repressor (*AHRR*) on chromosome 5 is a gene involved in a key pathway of response to tobacco smoke components. In this study, the same site was significant on univariate analysis and showed decrease in methylation in newborns where both mother and grandmother smoked during pregnancy compared to only mother smoking, which was less than only grandmother smoking, and neither smoking, during pregnancy. The top site showing significance and similar trend as previous study suggests this CpG site is influenced by maternal and grandmaternal smoking during pregnancy. The other CpG that was significant was in gene myosin 1G (*MYO1G*) cg04180046, which was again one of the top 10 sites previously identified as being associated with maternal smoking in pregnancy in an EWAS, showed increase in methylation with maternal and grandmaternal smoking during pregnancy. The directions of changes seen in methylation at these sites are consistent to that observed in the study of Joubert et al. on cord blood DNA methylation with prenatal smoking exposure (197). In a study of the MoBa cohort on DNA methylation on the candidate CpG sites showed only a minimal additional effect of grandmother smoking in her pregnancy with the mother in comparison to only mother smoking during pregnancy (299). Maternal smoking information was collected prospectively and based on cotinine levels whereas grand-maternal smoking data was collected retrospectively by asking the mother if her mother smoked during pregnancy; this may have introduced recall bias and the resulting inaccuracy might have had an impact on the results for minimal effect seen for grand-maternal smoking, in addition to maternal smoking. In the Third Generation Study, prenatal smoking was collected prospectively for both mother and grandmother, which may explain the effect seen for combined maternal and grand-maternal smoking to be more than only maternal smoking. A limitation of this study is that smoking exposure was not validated by cotinine levels at either time.

The observed results do not confirm a true transgenerational effect but rather suggest intergenerational inheritance. However, the study adds to the concept of

inter/multi-generational effect of environmental exposures and further explores the effect of grandparental environmental exposure on DNA methylation in cord blood of the grandchild, as a potential underlying mechanism. The results cannot be taken as robust evidence to support the concept, but the findings are novel and should stimulate further research using larger studies.

Chapter 6 Epigenome wide analysis of cord blood DNA methylation for maternal history of asthma

6.1 Introduction

History of asthma in immediate family, e.g. maternal, paternal or sibling, is a known risk factor for asthma (221). Parental history (both maternal and paternal) of asthma increases the risk of asthma. However maternal asthma carries a significantly higher risk than paternal asthma for development of asthma in offspring (220) and the risk increases with the severity of maternal asthma. A meta-analysis of 33 studies showed that the risk of asthma in children is greater from maternal asthma than paternal asthma (219). Children born to mothers with poorly controlled asthma are at higher risk of developing asthma or recurrent wheeze at 3 years compared to those born to asthmatic mothers with controlled asthma (300). A large Danish study also showed that the risk of early onset persistent asthma in children is higher in mild and moderate to severe uncontrolled maternal asthma compared to mild controlled asthma (301). In the IOWBC maternal asthma has been associated with early persistent childhood wheeze (216), and current wheeze and physician diagnosed asthma in later childhood (10 yrs) (221).

The precise biological mechanisms underlying the effect of maternal asthma on the risk of offspring asthma are unknown. However a range of mechanisms associated with the adverse effects of maternal asthma on pregnancy outcomes have been suggested, including inflammation, corticosteroid treatment, exacerbation of asthma, maternal hypoxia, and altered placental function (302). Most direct evidence centres on the stimulation/control of inflammatory responses and placental factors (303-305). A range of in utero exposures have been shown to result in epigenetic differences observable in offspring and to be risk factors for childhood asthma. These support the "developmental origins" of asthma theory.

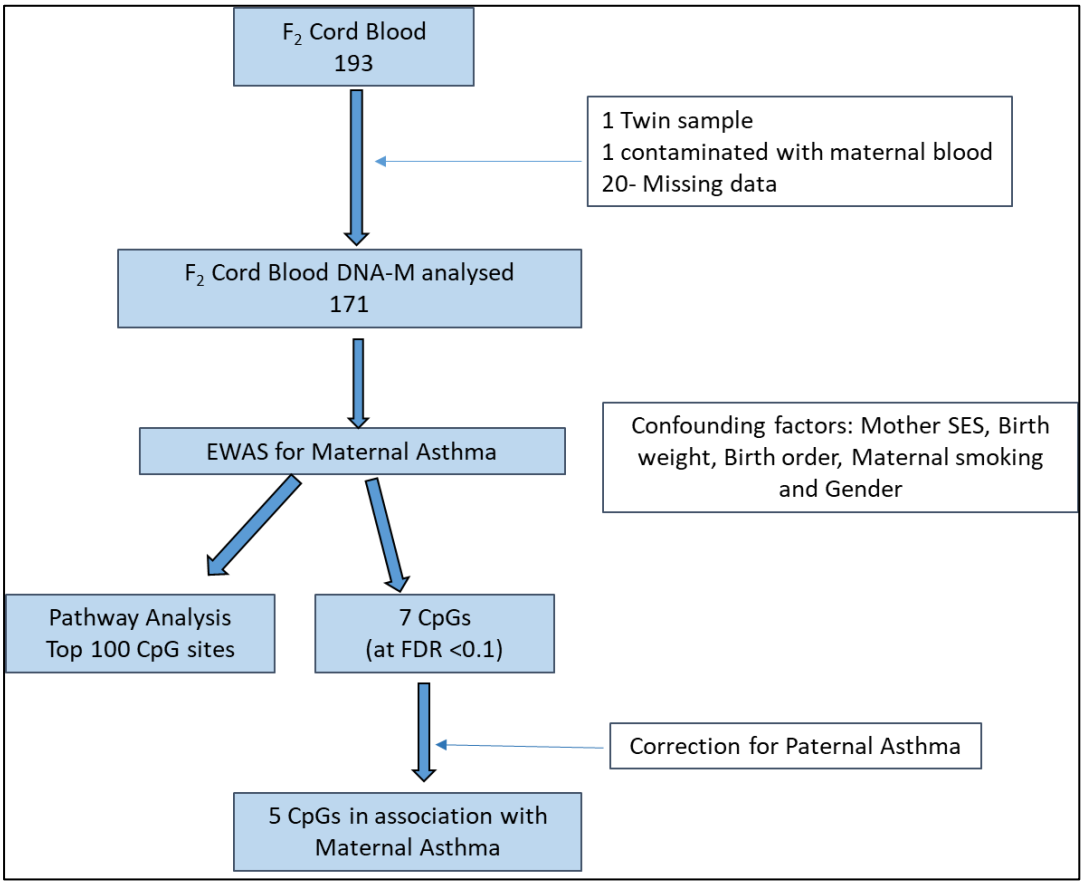
With respect to maternal asthma as an exposure, in a small study using the Illumina 27K methylation array, Gunawardhana et al. reported altered DNA methylation profiles in peripheral blood cells from 12-month-old infants with and without a mother with asthma (306). Twelve CpG loci (11 genes) showed greater

than 10% comparative difference in DNA methylation between the groups, though this was not significant after correction for multiple testing. DeVries et al. undertook a EWAS in a nested case control study of asthmatic children, with replication in independent cohorts. This study identified DNA methylation at birth associated with subsequent asthma, this association was stronger in children with maternal asthma (222). In a publicly available thesis of from university of Chicago (December 2020) Dr Kevin Magnaye looked at DNA methylation in airway epithelial cells among asthmatics with and without maternal asthma (307). Study identified 1,428 CpGs differentially methylated only in asthma cases with maternal asthma, but not in cases with no maternal asthma. All these studies support the concept of the potential role of DNA methylation mediating the risk of development of asthma imparted by maternal asthma.

The aim of this study is to carry out an EWAS of DNA methylation of third generation cord blood in relation to maternal asthma. Then to explore the enrichment with pathway analysis and EWAS database to identify potential biological pathways that may be altered because of exposure to maternal asthma. Furthermore, to understand the biological relevance of identified CpG sites in regulation of gene expression, the association of DNA methylation at identified CpG sites with gene expression will be assessed.

6.2 Methods

The study plan overview is provided in Figure 6-1; EWAS was conducted in 171 cord blood samples. Cord blood DNA methylation was available for 193 newborns, two had to be removed as one was a twin and one DNA sample was contaminated. Figure 6-1 lays out the analysis framework. Cord blood DNA methylation and maternal asthma status was available for 171 Third Generation new-borns, which were included in the study.



F2; Third Generation, DNA-M; DNA Methylation

Figure 6-1 study plan for EWAS

6.2.1 Statistical methods

To assess the association between maternal asthma and genome-wide DNA methylation while controlling for confounding factors (maternal smoking, birth weight, birth order, gender, and maternal socio-economic status) robust linear regression was performed using R v3.3.2. At the time of analysis, there was substantial missing data for paternal asthma
in
Generation as fathers were still being recruited, so paternal asthma was not added in the initial regression analysis. The CpGs identified in the first step were reanalysed adding paternal asthma as a confounder (sensitivity testing) to assess if methylation at CpGs associated with maternal asthma were sensitive to paternal asthma. Correction for multiple-hypothesis testing was done using false discovery rate (FDR) correction (308), where each CpG site was assigned adjusted p-values (Adjusted p-value <0.05 and <0.1). Boxplots were used to show the level of methylation in CpG sites between groups; boys and girls and paternal asthma and no paternal asthma, t-test of normalised beta values was used to test for means in two groups. Cord blood samples in this study were corrected for cell type distribution disparities using the reference panel described in Bakulski et al. (309). There were 399,383 CpGs remaining after quality control, which were included in the epigenome-wide analysis.

All identified CpGs were tested for association with expression levels of genes within a region of 250 kb upstream or downstream of the CpG (310). Corresponding annotated genes were identified from methylation label file (Infinium MethylationEPIC v1.0 B4 Manifest File), SNIPPER (https://csg.sph.umich.edu/boehnke/snipper/) (311) and the University of California Santa Cruz (UCSC) Genome Browser (https://genome.ucsc.edu/) (312). Linear regression analyses were used to look for the association of DNA methylation at identified CpG sites and gene expression.

6.2.2 Downstream and Pathway analysis

The DNA methylation results were ranked in order of statistical significance using the inferred adjusted p-values. CpG annotation was done using UCSC genome browser (313) provided by Illumina in the array manifest and using the program SNIPPER (314) (version 1.2, http://csg.sph.umich.edu/boehnke/snipper/ where a gene was not identified in the array manifest. The function and disease relation of each gene was researched using the Genecards Human Gene Database (315) and DisGeNET (IBA group, Barcelona) (316). The first 100 ranked CpGs were used for pathway analysis using Ingenuity® Pathway Analysis (IPA®,QIAGENInc.https://www.Qiagenbioinformatics.com/products/ingenuity-pathway-analysis) (317) for identification of global canonical pathways.

6.2.3 Gene expression level measurement

Gene expression in cord blood (n = 156) had previously been generated as described (318). Briefly, RNA was extracted from cord blood samples store in PAXgene tubes and RNA yield and the absence of DNA contamination was measured by Qubit 2.0 fluorometer (Life Technologies, Grand Island, NY, USA, https://www.thermofisher.com/uk/en/home/industrial/spectroscopy-elemental-isotope-analysis/molecular-spectroscopy/fluorometers/qubit.html). The RNA quality was evaluated by a Agilent Bioanalyzer 2100 via RNA 6000 Nano Chips, RNA was reverse- transcribed into complementary DNA (cDNA) and expression was measured using SurePrint G3 Human Gene Expression Microarrays (GeneSpring Technology). using Agilent's Single-Color Microarray-Based Gene Expression Analysis protocol version 6.0. Microarray data was processed using limma (319) in the R statistical computing environment, and background correction was performed using normal-exponential convolution (normexp) function (320). Data was converted to log2-transformed data for further analysis. Filtering was performed to remove low expressed probes that

are close to the background level. Negative control probes were also removed from the data.

6.3 Results

There was almost equal representation of gender; 86/171 male and 85/171 female newborns. 24.6% (42/171) had reported maternal asthma. Information on paternal asthma was available only for 61/171 newborns and 54% (33/61) had paternal asthma. Description of the study sample is provided in Table 6-1.

Table 6-1 Description of the study sample

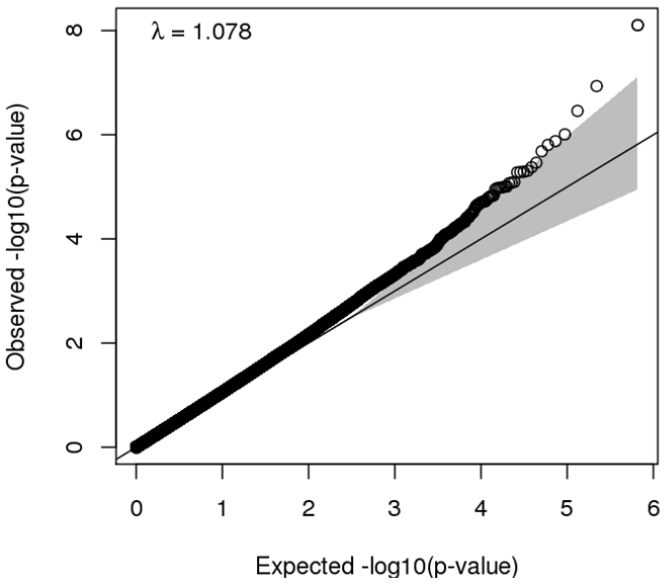
Characteristics	
Gender: Male	86 (50.3%)
Maternal age (mean) in years (SD)	23.7 (2.7)
Birth weight (mean) in grams (SD)	3425.6 (503.7)
Maternal Asthma	42/171 (24.6%)
Paternal Asthma	33/61 (54%)
Maternal smoking	63/171 (36.8%)
Birth Order	
1	85/171 (49.7%)
2	86/171 (50.3%)
Socioeconomic status	
1	25/171 (14.6%)
2	53/171 (31%)
3	52/171 (30.4%)
4	24/171 (14%)
5	17/171 (9.9%)

6.3.1 EWAS results

Figure 6-2 shows the Quantile-Quantile (QQ) plot of the observed p-values of the EWAS, which shows the lambda genomic inflation value of 1.078, which suggests the model is marginally inflated, meaning there is small risk of false positive results. An overview of the EWAS results are shown in the Manhattan plot; (Figure 6-3). Two CpGs satisfied the Bonferroni correction and 3 CpGs satisfied FDR <0.05. After easing the threshold to FDR <0.1, seven CpGs were identified to be differentially methylated in relation to maternal asthma while controlling for confounders (Mother SES, Birth weight, Birth order, Maternal smoking and Gender) (see table 6-2).

The 7 CpG sites associated with maternal asthma with FDR of 10% were cg10146326, cg11852794, cg13740698, cg16228090, cg24156983, cg25421566 and cg20201954. With cg10146326, cg25421566 and cg20201954 significant at FDR of 5%. In the second model where 7 CpGs were reanalyzed adding paternal asthma to the above confounders showed only five out of seven; cg10146326 (*PANK4*), *cg*11852794 (*CATSPERB*), cg13740698 (*SHISA3*), cg16228090 (*ZNF75A*), and cg24156983 (*SPATA22*) remained significantly associated with Maternal asthma, however due to missing data, this reduce the sample size (N=61) in the sensitivity analysis.

Known associations of the CpGs with an exposure of condition were explored using the EWAS Atlas (321), which is an open epigenome wide association studies knowledgebase platform. DNA methylation at cg10146326 and cg13740698 are known to be associated with prenatal arsenic exposure and cg11852794 with long term air pollution exposure. cg24156983 has been associated with asthma. Full description of ID, location, gene mapped to, effect size, p-value and adjusted p-values of and the known associations from EWAS Atlas for the 7 CpGs are provided in Table 6-2.



The QQ plot (left) indicates the expected and observed –log10 p-values and the lambda genomic inflation factor of 1.078 is shown.

Figure 6-2 Quantile-Quantile (QQ) plot for EWAS Maternal asthma and DNA methylation in the cord blood of Third Generation study

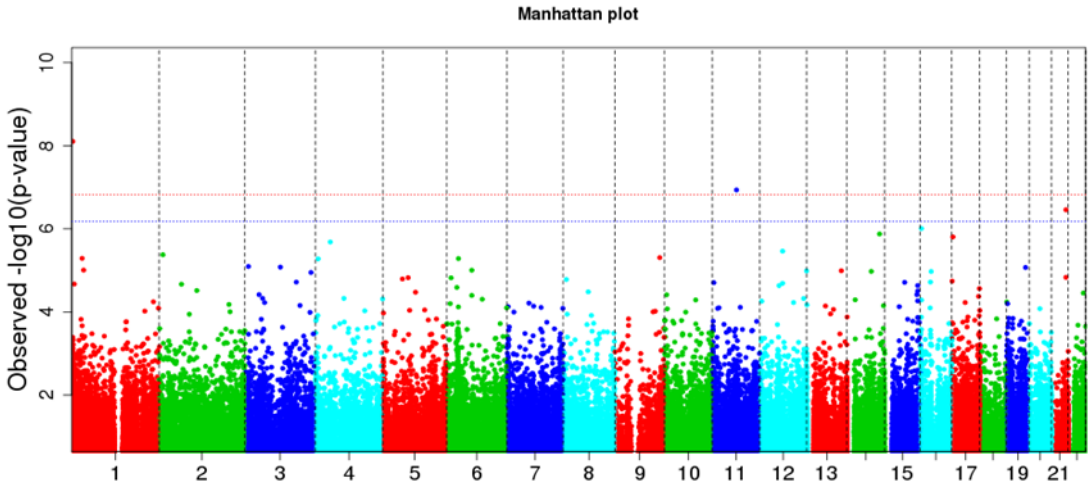


Figure 6-3 Manhattan plot of EWAS for maternal asthma in cord blood.

The Y axis represents the observed p-values for EWAS of maternal asthma and DNA methylation, X axis indicates the chromosome on which the CpGs are located. Blue horizontal line indicates FDR 0.05 correction, and the red line indicates the Bonferroni threshold

Table 6-2 Differentially methylated CpG sites in association with maternal asthma (at threshold <0.1)

CpG	CHR	MAPINFO	UCSC Ref Gene Name	BETA	SE	PVAL	ADJ_P	EWAS ATLAS
cg10146326	1	2439901	*PANK4*	0.016047	0.006567	0.014536	0.020351	Prenatal arsenic exposure
cg11852794	14	92198970	*CATSPERB*	0.030419	0.010421	0.00351	0.006143	Long term Air pollution (NO2) Systemic lupus erythematosus
cg13740698	4	42399384	*SHISA3*	0.011149	0.002788	6.35E-05	0.000258	Prenatal arsenic exposure Acute B Lymphoblastic leukaemia Gingivo-buccal oral cancer
cg16228090	16	3356011	*ZNF75A*	0.026613	0.008251	0.001257	0.002933	Aging Risk of Breast Cancer
cg24156983	17	3343601	*SPATA22*	0.01768	0.004461	7.38E-05	0.000258	Asthma Adenoma
Two sites not significant after sensitivity analysis for paternal asthma								
cg25421566	21	41239488	*PCP4*	0.014447	0.009576	0.131387	0.153285	N/A
cg20201954	11	68609518	*CPT1A*	-0.00405	0.002905	0.162906	0.162906	N/A

Columns: CpG site ID number, Chr; chromosomal location, Map info; the genomic location of the CpG site, BETA; effect size of methylation (<0; hypomethylated, >0; hypermethylated), p-value of association with maternal asthma (vs no maternal asthma) without multiple testing correction, Adjusted p-value; after FDR correction. N/A; not available

6.3.2 Maternal Asthma

Figure 6-4 shows the direction of differential methylation in relation to maternal asthma. Box plots showing the methylation differences at CpG sites (top five most significant CpG sites from maternal asthma EWAS; cg10146326, cg11852794, cg13740698, cg16228090 and cg24156983) in cord blood DNA of newborns with and without maternal asthma. They all show hypermethylation in maternal asthma group compared to those with no maternal asthma.

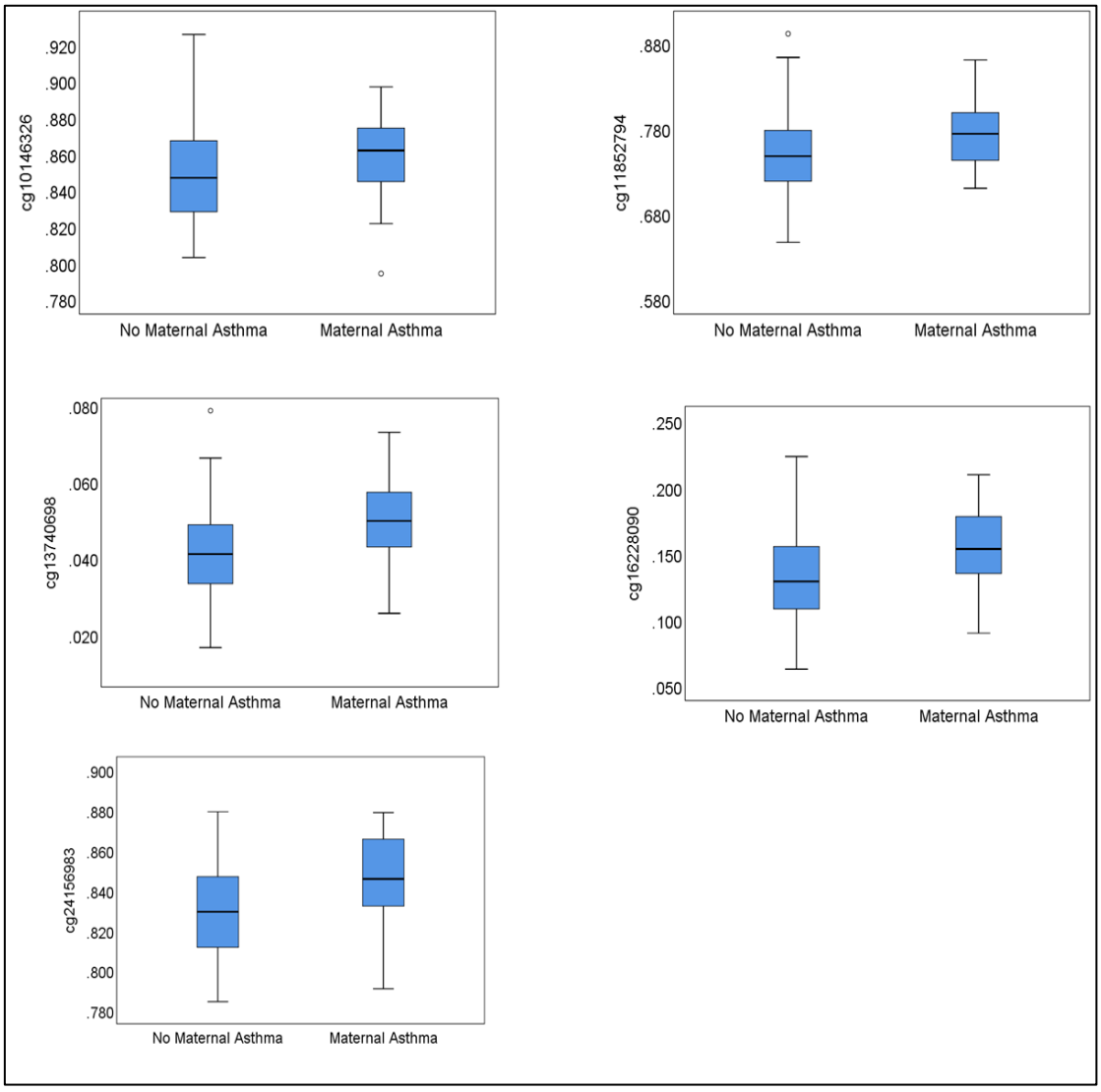


Figure 6-4 Direction of methylation at 5 CpGs in association with maternal asthma.

6.3.3 Paternal Asthma

Methylation level at the five CpGs was also examined in subgroups of newborns with paternal and without paternal asthma. There was significant difference in methylation only at cg13740698 (p=0.02) mapped to SPATA22. There was no difference in methylation at other four CpG sites in relation to paternal asthma status.

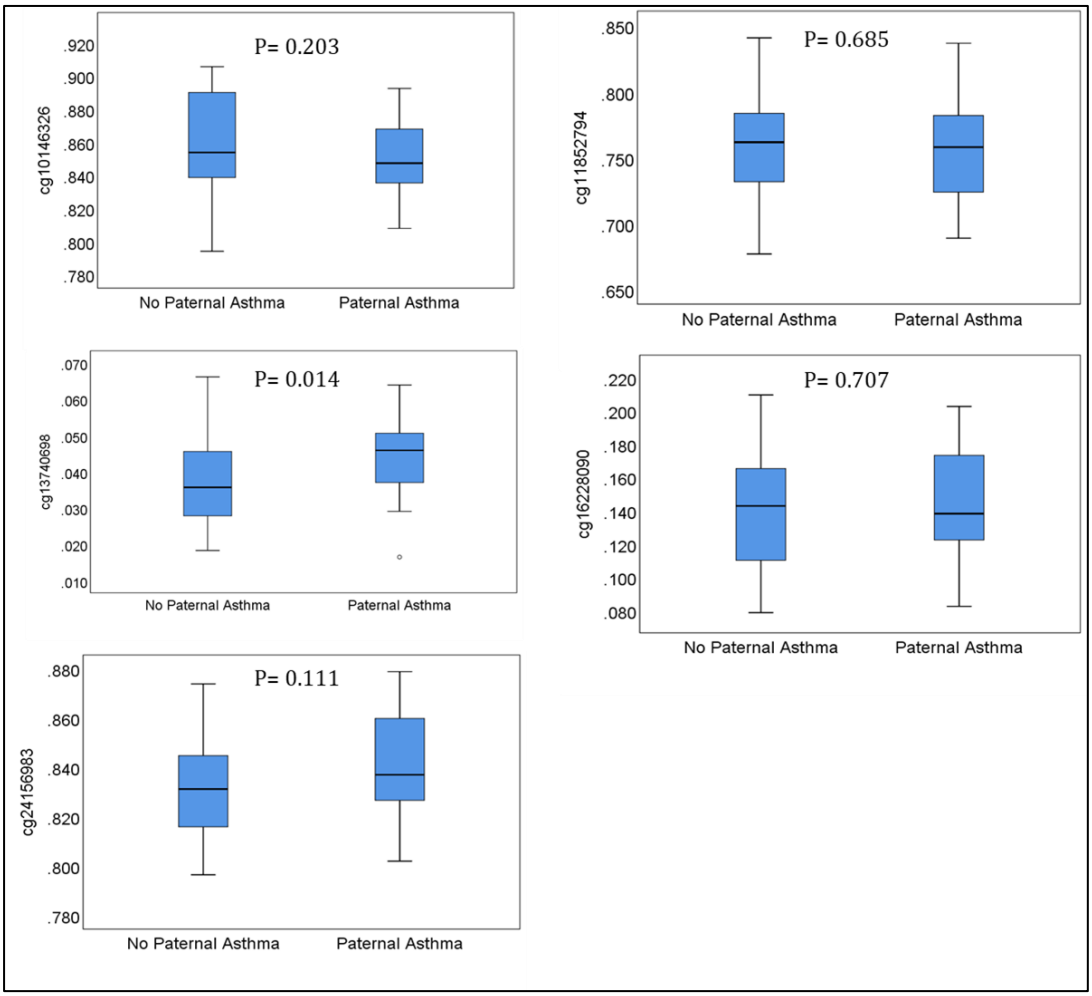


Figure 6-5 DNA methylation at 5 CpG sites in newborns with and without paternal asthma.

Box plots showing the methylation differences at CpG sites (top five most significant CpG sites from maternal asthma EWAS; cg10146326, cg11852794, cg13740698, cg16228090 and cg24156983) in cord blood DNA of newborns with and without paternal asthma. No significant difference at any CpG site observed. Significant difference in methylation was observed only at cg13740698 (p=0.02) mapped to SPATA22.

6.3.4 Pathway analysis

Canonical pathway analysis using the top 100 differentially methylated CpGs assigned to the closest gene in the main model constructed using IPA showed enrichment for 15 biological pathways pathways with p-value<0.05.

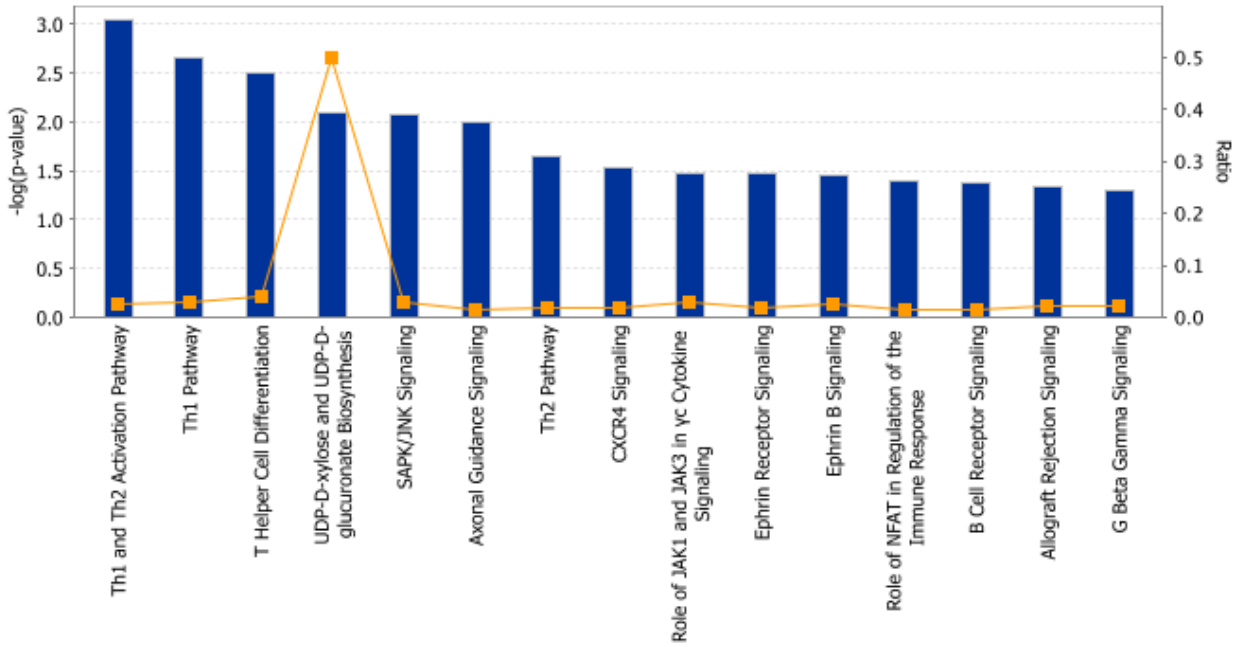


Figure 6-6 Canonical pathway

Figure 6-6 shows the pathways in a bar chart ranked according to the p-values. Bar chart representing canonical pathways from IPA®: x axis shows significant canonical pathways (p-value ≤ 0.05); y axis displays -log of p-value calculated by Fisher's exact test right-tailed (blue bars on chart). Yellow line with square markers represents the ratio of each pathway, which is calculated by dividing the number of molecules in the pathway that meet the cut-off criteria, by the total number of molecules that make up the pathway and that are in the reference set (ratio describes amount of overlap).

6.3.5 Network analysis

Network analysis of the top 100 CpGs most strongly associated with maternal asthma using IPA revealed 9 networks of 254 molecules with different functions as described in Table 6-3. Cell-mediated immune response was enriched in network 4, which is depicted in the network diagram. A list of the 100 CpGs is given in appendix 1.

Table 6-3 Networks from IPA using top 100 differentially methylated CpGs

ID	Score	Focus Molecules	Top Diseases and Functions
1	29	15	Cell Death and Survival, DNA Replication, Recombination, and Repair, Auditory and Vestibular System Development and Function
2	29	15	Cell Death and Survival, Embryonic Development, Connective Tissue Disorders
3	27	14	Cellular Development, Cell Morphology, Cellular Function and Maintenance
4	24	13	Cell-mediated Immune Response, Cellular Development, Cellular Function and Maintenance
5	20	11	Cell Death and Survival, Neurological Disease, Organismal Injury and Abnormalities
6	17	10	Cellular Assembly and Organization, Cellular Function and Maintenance, Amino Acid Metabolism
7	9	6	Cellular Development, Cellular Growth and Proliferation, Embryonic Development
8	2	1	Cell Cycle, Gene Expression, Protein Synthesis
9	2	1	Cancer, Organismal Injury and Abnormalities, Skeletal and Muscular Disorders

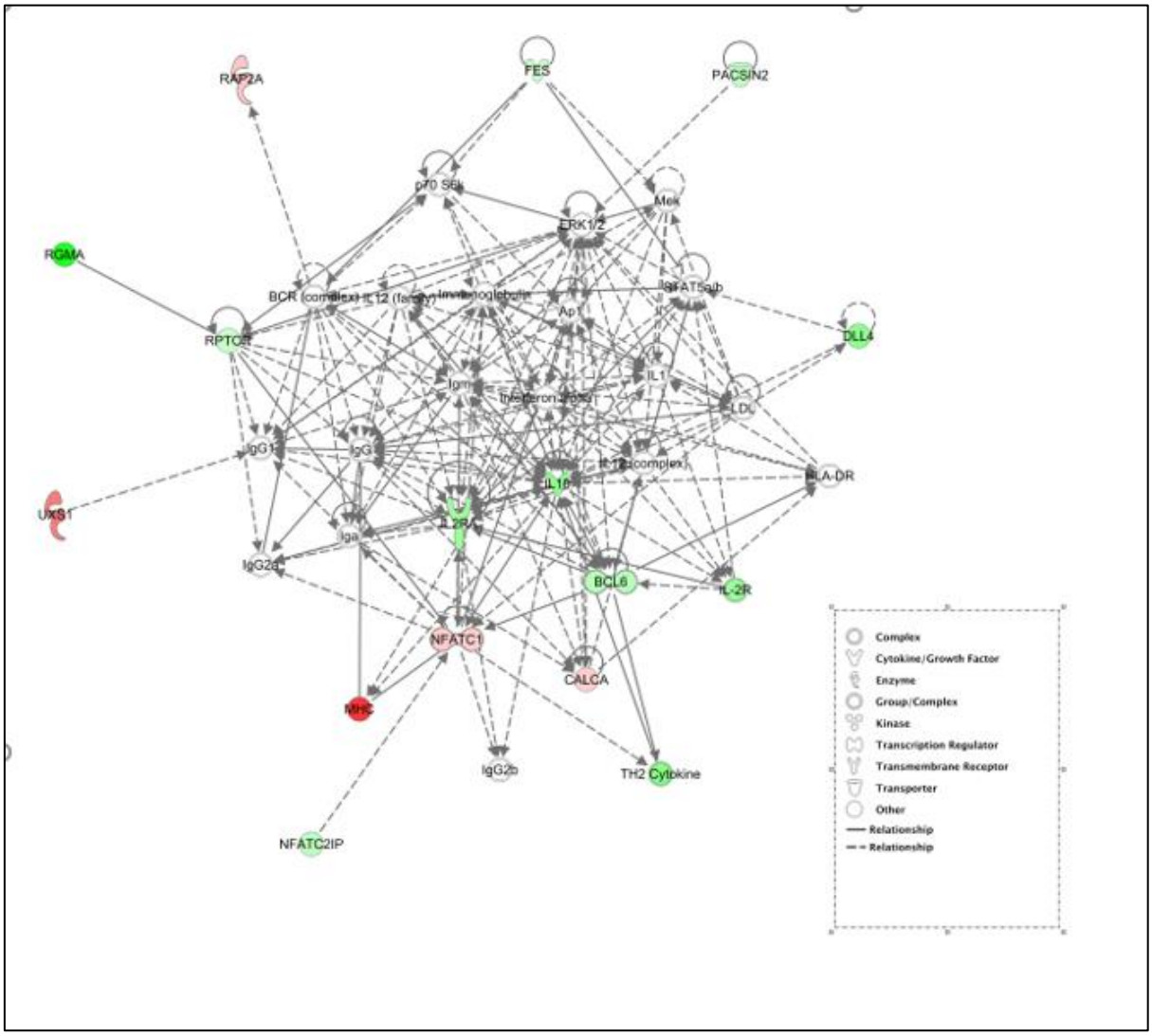


Figure 6-7 Network 4, highlighting Cell-mediated immune response

Figure 6-7 gives a visual representation of Network 4 from IPA analysis; interaction between components encoded by genes significantly associated with maternal asthma and DNA methylation. The shapes represent the type of molecule (e.g. enzyme, ion channel, transporter) and the red colour indicates hyp0methylation, with the intensity representing the degree of significance. Full lines indicate direct interaction and dotted lines indirect interaction. Arrows are present when a molecule is acting on another, whilst no arrow represents binding only.

6.3.6 DNA methylation and Gene expression

Gene expression and DNA methylation in cord blood were available for 156 children. Association of DNA methylation at CpG sites and gene expression was analysed using linear regression. Five CpG sites identified to be associated with maternal asthma are mapped to 5 genes and 7 probes. cg16228090 is mapped to two probes in *ZNF75A.* cg24156983 is mapped to probes A_21_P0000078 & A_23_P101084 on *SPATA22.* There was no significant correlation between DNA methylation at these sites and gene expression in respective probes. DNA methylation at both these sites were analysed with mean expression at the respective two probes these sites are mapped to, cg24156983 showed stronger association, however it was not statistically significant.

Table 6-4 Association between DNA methylation at 5 CpG sites and annotated gene expression level.

CpG sites	Gene	Probe Name	B	95% CI for B	t	*p*-value
cg10146326	*PANK4*	A_23_P149690	-0.214	-0.560 to 0.132	-1.223	0.223
cg11852794	*CATSPERB*	A_23_P77043	-0.056	-0.477 to 0.365	-0.263	0.793
cg13740698	*SHISA3*	A_23_P41476	-0.35	-0.449 to 0.379	-0.169	0.866
cg16228090	*ZNF75A*	A_33_P3279109	-0.34	-0.225 to 0.158	-0.348	0.729
cg16228090	*ZNF75A*	A_33_P3285047	0.145	-0.226 to 0.515	0.773	0.441
cg16228090	*Mean of two probes of ZNF75A*	A_33_P3279109 & A_33_P3285047	0.059	-0.169 to 0.287	0.511	0.610
cg24156983	*SPATA22*	A_21_P0000078	0.333	-0.421 to 1.087	0.873	0.384
cg24156983	*SPATA22*	A_23_P101084	0.190	-0.568 to 0.949	0.496	0.621
cg24156983	Mean of two probes for *SPATA22*	A_21_P0000078 & A_23_P101084	0.262	-0.233 to 0.756	1.046	0.297

B: Beta coefficient, 95% CI: 95% confidence interval, t: t statistics from regression analysis.

6.4 Discussion

The Epigenome-wide association study of cord blood DNA methylation in relation to maternal asthma was carried out in the Third Generation Study identifying differentially methylated CpG sites. Pathway analysis with the methylation at top 100 CpG sites has showed top three pathways being related to T helper cell differentiation and activation which are known to play an important role in asthma pathophysiology.

From the EWAS Atlas database we can see that 3 of the 5 CpG sites have been previously associated with prenatal exposures. DNA methylation at cg10146326 (mapped to *PANK4)* and cg13740698 (mapped to *SHISA)* in cord blood are known to be associated with prenatal arsenic exposure in a Mexican EWAS study (322). An EWAS study in Dutch adults showed DNA methylation at cg11852794 (mapped to *CATSPERB*) in whole blood to be associated with long term air pollution exposure (323). One of the CpG sites identified to be differentially methylated in relation to maternal asthma is known to be associated with asthma. An EWAS on airway epithelial cells in 74 asthmatics and 41 non asthmatics showed that DNA methylation in locus 17q12-21 was associated with asthma risk (324). cg24156983 (mapped to *SPATA22*) identified in my EWAS study falls in the 17q12-21 locus and has been to be associated with asthma identified in the above study (324).

The genes mapped to top 5 CpG sites are mapped to genes in biological pathways not previously been known to be related to asthma. However, some of these genes are known to be differentially methylated in relation to asthma and related exposures at other CpG sites than ones identified in the current study. *PANK4* (Pantothenate Kinase 4) is a protein coding gene and is related to pathways of metabolism. DNA methylation of *PANK4* has been associated with smoking in buccal samples of adults (325). The CpG site (cg25386426) identified in that study is different from the one mapped to *PANK4* in the current study. A metanalysis of the association between sustained maternal smoking during pregnancy and DNA methylation in cord blood showed *PANK4* to be differentially

methylated (198) (Supplementary Table S3). In the current EWAS we adjusted for maternal smoking as a potential confounder even then CpG linked to this PANK4 was identified in relation to maternal asthma, in the metanalysis by Joubert et al, there was no correction for maternal asthma included in the analytical model.

CATSPERB (Cation channel sperm associated auxiliary subunit beta) is probably involved in sperm hyperactivation. In an EWAS of Der p allergen-specific immunotherapy in children with allergic asthma, methylation of cg19000611 (a different site from one identified in the current EWAS) in peripheral blood was significantly associated with immunotherapy (326). *SHISA3* (Shisa family member 3) a single-transmembrane protein coding gene known to be differentially methylated both in peripheral blood and saliva between asthmatics and controls at cg11065575 (Supplementary Table 3) (327). *ZNF75A* (Zinc Finger Protein 75a) is related to nucleic acid binding and DNA binding transcription factor activity. Its methylation has been shown to be associated with poverty status in a middle aged white population at the same CpG site identified in relation to maternal asthma (Table S3) (328). Methylation of *SPATA22* (Spermatogenesis Associated 22) has been associated with asthma as described above. Its methylation is also associated with cord blood Leptin levels at cg06862644 (Table S1) (329). Methylation of *SPATA22* was associated with maternal asthma even when birth weight was controlled for in EWAS analysis, these findings should stimulate further studies of DNA methylation of this gene for its role in inheritance of asthma and the potential effects of the maternal metabolome on offspring epigenetic profile.

Pathway analysis showed that T helper cell differentiation and activation, Th1 and Th2 pathways were enriched with genes covering the top 100 CpGs, suggesting a role in asthma etiopathology. This has to be interpreted with caution as these top 100 CpGs did not remain significant following multiple testing correction, however these are likely to be enriched for true positive associations and it is not possible to undertake pathway enrichment with only 5 or 7 genes. One of the reasons for the lack of statistically significant results could be stringent correction for the potential confounders in EWAS analysis which might weaken

the strength of associations. However, the key limitation is the relatively small sample size. Even then, both the sample size (N=171 vs N=40) and numbers of CpG sites assessed (450K vs. 27K) are an order of magnitude larger than the only previous study of the effect of maternal asthma on the offspring epigenome (306). None-the-less my findings provide proof of concept that maternal asthma influence methylation at loci that are involved in immune response pathways.

Although there was no statistically significant association seen between DNA methylation at the differentially methylated sites and the gene expression at the respective mapped probes, there is consistent and positive correlation for cg24156983 and gene expression at the two probes mapped to *SPATA22.* As discussed above DNA methylation at this same site is known to be associated with asthma.

DNA methylation associated with asthma status in later childhood could be established at birth itself. Importantly, it has been shown in 3 birth cohorts that methylation at SMAD3 promoter is associated with childhood asthma selectively in neonates with a maternal history of asthma (222). The differentially methylated region at SMAD3 covered 24 CpGs (region Map info; 67356706 - 67357027, from Table E4 of the supplementary data), in the current EWAS study SMAD3 methylation was not significantly associated with maternal asthma (p-value of 1.81E-02 and adjusted p-value of 0.747). Likely reasons for this finding are; one that although both are in cord blood, directions of analyses are opposite, EWAS in this thesis was for maternal asthma in a birth cohort whereas the other study started with EWAS for a disease status. Second reason could be the different confounding factors that are corrected for in both studies.

DNA methylation can be influenced by exposures as well as disease status; running the EWAS in cord blood DNA does minimize the chances of the effects of disease status, but prenatal exposures can influence the DNA methylation in cord blood. Prenatal exposures were corrected for including mother's SES, birth weight, birth order, maternal smoking and gender. Practically it is difficult to control for all the possible exposures that might influence DNA methylation. In

the Third Generation Study data has been collected prospectively, but the number of subjects is smaller given the nature of the Third Generation Cohort. A larger sample would be ideal; however current results do provide initial results which need further validation through replication. In contrast to the availability of good prospective data from mothers, information from fathers of the Third Generation Cohort was lacking for many samples. As a result, paternal asthma was not included in the initial genome wide analysis, rather it was incorporated in the second model of linear regression for the 7 identified CpGs along with the initial confounders. In this secondary analysis, two of the top three CpG sites lost significance. This probably means there are common loci where both paternal and maternal asthma affect methylation significantly. However, lack of paternal data can be seen as a limitation of the study.

Epigenetics is a plausible mechanism that could explain the effect of maternal asthma on offspring risk of asthma. However, DNA methylation is only one part and including other epigenetic markers is needed to answer the question if epigenetics can explain some part of developmental programming of a complex disease like asthma.

In conclusion, an EWAS of DNA methylation in cord blood in relation to maternal asthma identified five CPG sites with differential methylation. Replication in an independent cohort is needed and, as association does not prove causality, further work exploring the effect of methylation on the expression of the genes identified and functional relevance to asthma is needed. Also, as the third generation children grow, and the disease is established, the effect of DNA methylation at the identified CpG sites on the disease status in later childhood will provide further evidence about the role of methylation at these identified loci in development of asthma.

Chapter 7 General Discussion

The initial objective of this study was to recruit children into the Third Generation birth cohort of the IOWBC, which along with the original cohort, provides a unique, prospective, well characterised, multigenerational cohort for studying the development of asthma and allergies. Longitudinal cohorts such as the Isle of Wight multigenerational birth cohort, where exposures, phenotyping and biological samples are collected prospectively, will contribute to the understanding of the role of inter – and transgenerational effects (including epigenetic mechanisms) in the evolution of complex diseases like asthma (330). The second objective of this study was to explore the role of the epigenetic marker DNA methylation in asthma in the multigenerational birth cohort. The findings provide support for the role of DNA methylation in mediating gene-environment interaction of *IL13* and prenatal smoking exposure on measurable asthma traits, intergenerational effects of tobacco smoke exposure influencing differential methylation in candidate loci known to be influenced by prenatal smoking and finally results show differential methylation established at birth in offspring associated with maternal asthma.

7.1 DNA methylation and asthma

Among different epigenetic markers DNA methylation has been the focus in several asthma studies. This is mainly because of the stable nature of DNA methylation, availability of platforms to survey the DNA methylation genome-wide, and the close functional relationship between DNA methylation and gene expression (331). Additionally specific changes in DNA methylation have been associated with early life exposures known to affect asthma and the phenotypes of asthma (332, 333). Figure 7-1 visualises the potential role of DNA methylation in asthma. Prenatal exposures such as prenatal smoking can influence the methylation at birth directly and also through interaction with the genotype, and the DNA methylation alterations influence the development of asthma through regulation of gene activity. Genetic susceptibility also increases the risk of developing asthma and this may be dependent on interacting with the same or different pre- and postnatal environmental exposures.

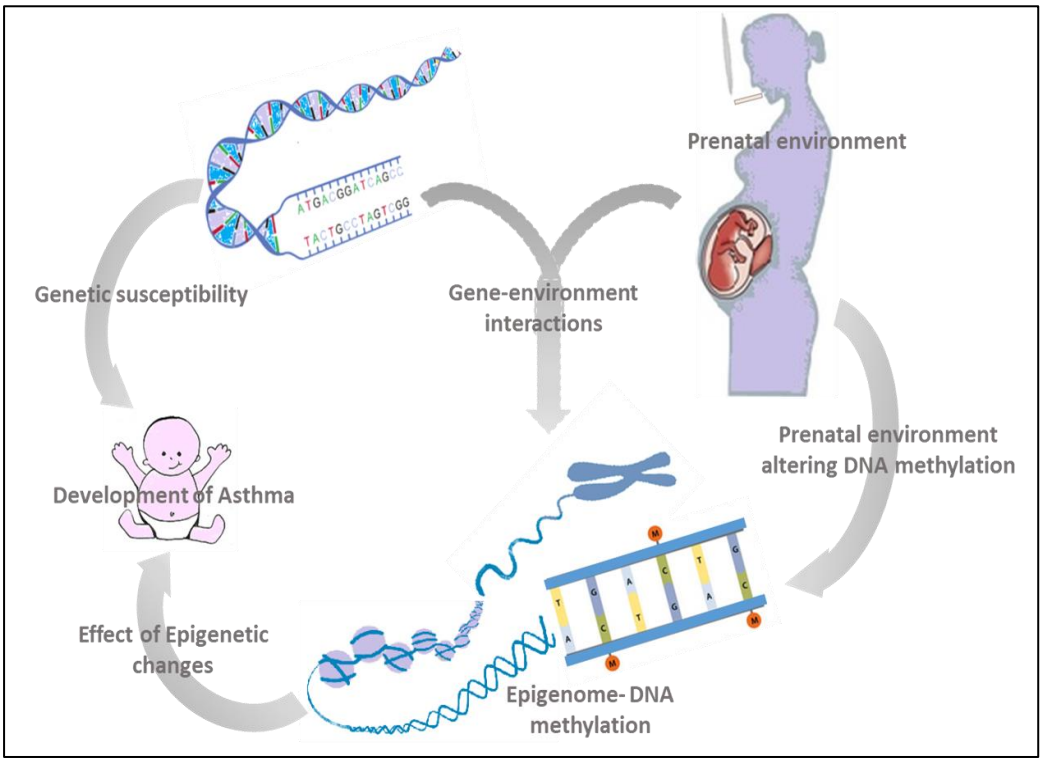


Figure 7-1 Role of DNA methylation in mediating effects of genetics and environment on development of asthma

7.1.1 Role of DNA methylation in gene-environment interaction

Nature and Nurture; genetic and environmental exposures, both play an important role in the development of asthma. The interaction between genetics and environmental exposures is known to influence asthma susceptibility (334, 335). The mechanistic conduit for such interactions influencing the phenotype is not clearly understood and epigenetic modifications are one of the plausible mechanisms (336). In chapter 3 this interaction was further explored by including the DNA methylation in the interplay, first interaction of *IL4R* genotype and methylation influences asthma risk in 18 years old females. Then in a two-stage model, exposure (prenatal smoking) and *IL13* genotype interaction influenced DNA methylation (conditional methQTL), and then the interaction of methylation and the modifiable genetic variant influencing the outcome

(FEV1/FVC and dose response slope of FEV1 in methacholine challenge). The interaction of a functional *IL13* SNP, rs20541, and maternal smoking during pregnancy influenced methylation at cg13566430, and DNA-M at cg13566430 interacts with genotype of functional SNP, rs1800925, to affect airflow limitation and airway reactivity (279). There was no significant effect on reported asthma observed. This highlights the heterogeneity of asthma and need for consideration of using specific phenotypes and traits of asthma as outcome in interaction association studies (336). In a case control study of 60 patients with dust mite sensitised allergic rhinitis and 65 controls, Li et a. show that DNA hypomethylation at *IL4* CpG5 and at *IL13* CpG38 were significantly associated with a higher risk of HDM-sensitised allergic rhinitis susceptibility (337). Methylation at *IL13* CpG38 was inversely proportion to gene expression, showing the biological relevance of the association. Similar to our findings, they also report genotype dependent DNA methylation for the same two SNPs, rs20541 and rs1800925 on DNA methylation at *IL13* CpG38 site.

Further exploratory analysis in IOWBC have looked at the interplay of a range of genetic variants, environmental exposures, and DNA methylation in different candidate genes with asthma and other outcomes. These are summarised in Table 7-1.

Table 7-1 Published analyses of interplay of genetic variants, environment exposure and DNA methylation in the Isle of Wight birth cohort.

Gene Polymorphism	Exposure	CpG site	Effect
GATA3	Oral contraceptive (OCP)	cg17124583	OCP × GATA3 polymorphisms increase risk of asthma via DNA-M (338)
Th2 Genes (*IL4, IL4R, IL13, GATA3,* and *STAT6*)	NA	cg12405139 (GATA3)	(GATA3 and IL4R) SNP × DNA-M contribute to asthma risk at 18 yrs.
		cg09791102	
		cg26937798 (IL4R)	Change in DNA-M can influence asthma transition from 10 to 18 yrs. (339).
FLG (Filaggrin) variants	NA	cg07548383	Combined effect of FLG loss of function and DNA-M increased risk of eczema at 18years (340).
LEP (Leptin) rs11763517	NA	cg00666422	SNP × DNA-M was associated with spirometry at 18 yrs, deceased with increasing level of methylation (341).

NA- not applicable, ×; indicates interaction SNP; single nucleotide polymorphism, DNA-M; DNA methylation, CpG; cytosine phosphate guanine nucleotide.

These findings from a candidate gene approach in a subset of the cohort with DNA methylation data available, support the concept of the role of DNA methylation in mediating the effect of gene-environment interactions but are far from conclusive. It is problematic to test for all possible gene-gene interactions due to the scale of multiple comparisons that are required. Gene-environment interactions are known to be important determinants of risk to asthma, studying genome-wide gene-environment interaction will require very large samples with uniformly measured exposure data (154) and will have to include rare variants as common genetic polymorphisms might not explain interaction effects on the phenotype (342). Czamara et al. conducted an integrated (main effects, additive and

interaction effects) analysis of genotype and prenatal environment on DNA methylation at differentially methylated regions in cord blood from four cohorts (343). The results show that differential methylation is best explained by genotype and environment interactions and additive effects than the main effects. Findings highlight the importance of including both genotype and environment when assessing differentially methylated profiles associated with complex disease traits like asthma.

7.1.2 Role of DNA methylation in multigenerational epigenetic inheritance

Environmental exposures have significant impact in a condition like asthma where exposures play a very significant role in development of the disease. Question is, can the effects of the environmental exposures be passed on to the next generations?

In asthma, prenatal smoking exposure has been studied across two generations; Li et al. in a case control study reported association between grand-maternal smoking during pregnancy and childhood asthma. They showed an increased risk of child developing asthma if both, mother smoked during pregnancy and also grandmother smoked when pregnant with mother (135). More recently, a large epidemiological study in the Swedish population did demonstrate an increased risk of asthma in the first 6 years of life for individuals whose grandmothers smoked during early pregnancy, independent of maternal smoking (137).

An increased risk of parent reported non-infectious wheeze in the first year of life, was observed when there was combined infant`s prenatal and mother`s prenatal smoking exposure (chapter 5). The exposures which could be potential confounding factors could not be corrected such as father`s smoking before conception and/or during pubertal development period when the male gametes are formed, and other post-natal passive smoking exposures could not be included. Whether there is an effect on clinical asthma diagnosis awaits further follow up of the Third Generation Cohort at age 6-7 years when asthma diagnosis

can be more confidently made. Other studies have shown similar association between grandmaternal smoking and grandchild asthma, Accordini et al. show effects of smoking along the maternal line in a large study. European Community Respiratory Health Survey (ECRHS) involving 2233 mothers and 1964 fathers showed that grandmothers' smoking during pregnancy was associated with asthma in their daughters and with asthma in presence of nasal allergies in their grandchildren within the maternal line (344). Magnus et al. in the Norwegian Mother and Child Cohort Study (MoBa) study also reported a positive association of grandmaternal smoking when pregnant with mother with asthma in grandchildren at 3 and 7 years (345). This study had potential recall bias as the grandmaternal smoking exposure was collected retrospectively when mother was recruited during pregnancy. Contrasting results were seen in a study of the UK birth cohort Avon Longitudinal Study of Parents and Children (ALSPAC), where no association between grandmaternal smoking when pregnant with mother and asthma in grandchild at 7 years of age was observed. However, there was an association between grandmaternal smoking when pregnant with the father and asthma in grandchild (136). The likely explanations for these contrasting results are the recall bias in retrospectively collected exposure information and more importantly could be due to unmeasured confounding factors. For example, smoking during pregnancy is usually associated with lower socioeconomic status. Measuring and controlling for these factors has to be considered. The ALSPAC study did control for maternal socioeconomic status. Birth weight can be another confounder, maternal smoking during pregnancy is associated with lower birth weight in babies (346) and lower birth weight in turn is associated with wheeze (347) and asthma in childhood (348). Measuring for all the potential confounders and then controlling for them in analysis can be challenging and will demand larger samples to see any associations.

To overcome such challenge, particularly with respect to sample size, the Pregnancy and Childhood Epigenetics (PACE) consortium which, to date, has brought together over 40 studies with >30,000 samples and DNA methylation data in pregnant women, new-borns and/or children (349). Alongside

significantly increased statistical power to detect associations, bringing studies together in the PACE Consortium for meta-analysis, greatly decreases the risk of false-positive associations. The larger power also enables the use of statistical approaches such as mediation analysis and mendelian randomisation to assess the causality of identified differentially methylated CpG sites on the phenotypic outcomes of interest. In addition, a few of the participating studies (including the IOWBC) have measured DNA methylation at multiple time points from birth through childhood and/or in adolescence, which enables investigation into the persistence of differential DNA methylation signals over time. Finally, the availability of data from studies with participants from various backgrounds in terms of ethnicity / genetic background, geographical location and living environment enables testing of identified associations across different settings and evaluation of heterogeneity of effects across study populations. The PACE consortium has now undertaken a number of meta-analyses across cohorts including for maternal smoking (198), and childhood asthma (350). Since completion of this study, a PACE study of maternal asthma has been proposed to which the IOWBC and IOW Third Generation Cohort intend to participate (Personal communication from Prof J W Holloway).

7.1.3 Evidence for DNA methylation as the mechanism of multigenerational effects of environment on phenotype

There is growing interest in the possibility that epigenetic programming as a result of environmental exposures such as prenatal smoke not only represents the mechanism by which effects of exposure modify development and lead to risk of disease in later life, but may also be transmitted from one generation to the next, however there is no direct evidence for this in humans (351, 352).

In animals

Animal studies have shown that environmental exposure can impact DNA methylation and in these changes can be passed on to future generations (353). In agouti mouse studies nutritional changes influencing DNA methylation, and thus the gene function, resulted in a change in coat colour and increase in bodyweight (354). In a mouse asthma model epigenetic markers established from prenatal smoke exposure have been shown to be transmitted to the next two generations through germline epigenetic alterations (132).

In humans

It is now widely accepted that environmental exposure does have an impact on the DNA methylation (298) in humans, which brings on the interesting question; whether these changes in methylation are transferred through meiosis to subsequent generations? DNA methylation is widely studied in epidemiological studies, however it is not the only epigenetic mechanism and is likely working in concert with other mechanisms such as histone modifications and non-coding RNAs to regulate gene function (352).

In chapter 5, it was observed that there is a small impact of combined grand-maternal and maternal smoking during pregnancy on cord blood DNA methylation, methylation levels of *AHRR* at cg05575921 decreased with maternal smoking and further decrease with combined grand-maternal and maternal

smoking during pregnancy. Methylation at both candidate sites explored in intergeneration study (cg04180046 mapped to *MYO1G* and cg05575921 mapped to *AHRR*) have been shown associated with nicotine metabolites in a separate study (355). The impact of combined prenatal and mother's prenatal smoking exposure on wheeze and DNA methylation seen in Third Generation study has to be interpreted with caution. This is not a definite proof for DNA methylation mediating the transgenerational effect of an environmental exposure. In this study, the impact of prenatal smoking is only examined for the subsequent two generations. This by definition cannot provide proof of true transgenerational inheritance; as when grandmother smoked during pregnancy, the mother was directly exposed, along with the gametes that go on to be grandchildren. Therefore, there is a possibility that the impact seen could be just from the direct exposure, i.e., intergenerational effects and not transgenerational inheritance. However, the association observed of smoking during pregnancy with wheeze and methylation in grandchildren in a prospective longitudinal cohort study, provides additional evidence that transgenerational epigenetic processes may play a role in determine asthma susceptibility.

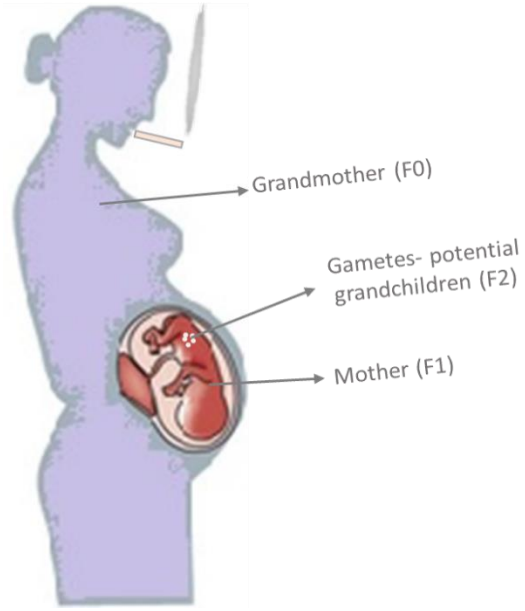


Figure 7-2 Grandmother (F0) smoking during pregnancy can lead to direct exposure for the mother who is the fetus (F1) and the fetal gametes, which are going to be grandchildren (F2)

When a female is smoking during pregnancy the foetus can be affected in utero (F1), as can the germline of the foetus (the future F2). These may not be true transgenerational inheritance and may be only intergenerational epigenetic inheritance (138). True transgenerational effect would be when epigenetic information is transmitted across generations to F3 and possibly future generations in maternal line, in the absence of further environmental exposure or germline mutations. There is no substantive evidence in humans that transgenerational inheritance is actually epigenetic inheritance (351, 352, 356). Further larger epidemiological studies with prospective data, covering all the confounders, pre and post exposure methylation status and ideally including other epigenetic markers will help in better understanding of transgenerational epigenetic inheritance. Transgenerational studies in humans will be challenging due to the long life cycle, measuring exposures accurately and defining a specific phenotype in a heterogeneous condition like asthma add further challenges.

7.1.4 Isle of Wight multigenerational birth cohort

The multigenerational cohort established on the Isle of Wight provides a unique opportunity to study intergenerational inheritance of effect of environmental exposures and intergenerational inheritance of DNA methylation changes.

Strengths

The original IOWBC is an unselected birth cohort established on the Isle of Wight of mainly (>95%) Caucasian descent. Retention rates have been high at all follow ups throughout childhood and adolescence (Table 1-2). Exposures have been collected prospectively before the onset of the outcomes and the asthma and allergy phenotypes have been well characterised. Information on grandparental disease status and exposures during pregnancy have been collected prospectively at the time of recruitment of the original IOWBC in 1989. In the third generation, samples and information are being collected prospectively for genome wide genotyping, DNA methylation, metabolomics (355) and extensive phenotyping of

asthma and allergic conditions. Prenatal exposures are well characterised through assessment at three times during pregnancy.

Weakness

Selection bias is a challenge through the follow ups of cohort as noted in original IOWBC, where females showed higher retention rates than males. However, with over all high retention rates throughout (70 – 90%), the bias is minimal and should not affect the generalisability of the results. Similar challenges are noted in recruitment of the third generation, where again mothers belonging to IOWBC were higher in numbers compared to the fathers from IOWBC (chapter 4). Another challenge is that there were children already born to the IOWBC participants before 2010 when the third generation study recruitment began. These children were postnatally recruited, and information has been collected retrospectively and biological samples from pregnancy and birth (e.g., cord blood) are not available. This should not influence the results in this thesis as only antenatal recruitments with cord blood samples were studied.

The generalisability of results from a cohort based on the Isle of Wight to a wider population can be questioned. The Isle of Wight being a modest sized island with no large urban conurbations or heavy industry may have less diversity in environmental exposures. The island population is mainly Caucasian in ethnicity, however there is no genetic inbreeding and there is frequent movement between mainland UK and the island. Also, as the majority of the population is from the same ethnicity, this reduces heterogeneity with a lower sample size required for studies in genetics and epigenetics.

Other multigenerational studies either mainly focusing on asthma, or those with available information on asthma, are summarised by Arshad et al. (138). The ALSPAC study was established in 1991 with 14,00 pregnant mothers in southwest England and their children and now grandchildren have been or are being recruited into the study providing a multigenerational cohort (357). The cohort has reported data of asthma and DNA methylation assessed at 5-7 and 15- 17 years of age. This data has been utilised in combination with IOWBC data for

EWAS of asthma (358). The IOWBC and ALSPAC, together with the Manchester Asthma and Allergy Study (MAAS), the Study of Eczema and Asthma To Observe the influence of Nutrition (SEATON) cohort and the Ashford birth cohort from part of the UNICORN (UNIfied COhort Resource Network) that is combing genomic and phenotypic information across five UK birth cohorts to enable harmonisation and secure co-analysis of these studies (359). For example, the UNICON investigators recently used latent class analysis to derive wheeze phenotypes among 7,719 participants from the five birth cohorts with complete report of wheeze at five time periods to explore the relationship of these phenotypes with asthma outcomes and lung function (23). The German Infant Study on the Influence of Nutrition Intervention plus Air pollution (GINIplus) and Influence of Life-style factors on Development of the Immune System and Allergies (LISAplus) studies with 5,991 and 3,097 participants recruited in 1995 to 1998 (360) are other studies with potential intergenerational information.

The Norwegian Mother and Child Cohort Study (MoBa) cohort established between 1999 and 2008, children linked to the national birth registry in 2014, have follow up data at 3 and 7 years. Similar to our intergenerational results, this study also reported a positive association of grandmaternal smoking when pregnant with mother with asthma in grandchildren at 3 and 7 years (345). However, the MoBa study had recall bias as the grandmaternal smoking exposure was collected retrospectively when mother was recruited during pregnancy.

7.1.5 Role of epigenetics in developmental programming of asthma: EWAS for maternal asthma

One of the strongest risk factors for asthma is maternal asthma (219). In the past this was assumed to be because of the transmission of inherited genetic variants from parent to child. However inherited genetic factors identified to date do not appear to account for a large proportion of the heritability of asthma. Epigenetic programming represents a plausible mechanism to explain, as least in part, the link between maternal asthma and childhood asthma. Maternal asthma would cause epigenetic changes through two main potential ways. Firstly, by indirectly

affecting placental function. Maternal asthma and use of inhaled glucocorticoids are associated with reductions in fetoplacental vascularity and placental blood flow (305, 361) which could lead to restricted foetal growth and low birthweight which have been shown to be associated with childhood asthma. Secondly, by direct effects on foetus. This could be via: (i) Inflammatory mediators crossing placenta (362, 363), (ii) initiating pre-natal immune responses similar to those seen in relation to maternal allergy (364), or (iii) via the potential effects of asthma medications during pregnancy. Currently there is no convincing evidence of asthma medications crossing placental barrier.

What is known, however, is that differential DNA methylation has been observed at birth between children who develop asthma and who do not later in life (222), suggesting that the trajectory of asthma may begin at birth, if not before. In the study of de Vries et al., cord blood methylation from 36 children (18 non-asthmatic and 18 asthmatic subjects by age 9 years) from an unselected birth cohort were analysed, which showed 589 differentially methylated CpG sites. Replication of the analysis in three different cohorts showed that the methylation at a single significant site was associated with childhood asthma selectively in neonates with a maternal history of asthma (222). These findings further support the hypothesis that DNA methylation may provide a mechanism by which maternal asthma increases offspring asthma risk.

An EWAS of DNA methylation concerning maternal asthma has been previously reported. Gunawardhana et al. looked at whole blood DNA methylation in infants with (n=25) and without (n=15) maternal asthma during pregnancy (306). The EWAS identified 12 CpG sites significantly associated with maternal asthma. They further report that DNA methylation at one of these CpG sites differed in relation to mothers with asthma being on inhaled steroid treatment. The 5 CpG sites identified in my study were not identified in this study. The reasons for this could include that DNA methylation was in infants which can be influenced by other postnatal exposures or setting of the early disease itself. Unlike our study, there was no correction for the confounders in this study, there might be population

differences in terms of asthma comorbidities and severity. Finally, the small sample size of this study suggests that these may be false positive findings.

It is interesting that among the 5 CpG sites identified in cord blood of third generation study in relation to maternal asthma there were sites known to be associated with exposures and asthma as outcome. Three sites are known to be associated with exposures, cg10146326 (*PANK4)* and cg13740698 (*SHISA)* to prenatal arsenic exposure and cg11852794 (*CATSPERB*) to long term air pollution. Airway epithelial cell DNA methylation at cg24156983 (*SPATA22*) is known to be associated with asthma. The above show the potential role of DNA methylation at the identified sites to be amenable to modification by the exposures related to asthma and to be associated with disease status as well. The EWAS for maternal asthma in cord blood DNA of the Third Generation Cohort (chapter 6) identified differentially methylated CpG sites. While this suggests that maternal asthma results in epigenetic programming of the offspring's genome, it does not provide evidence of a casual mechanism linking maternal asthma with increased asthma susceptibility in the offspring. Further follow up of the cohort (currently ongoing at age 6-7 years) will allow assessment of the association between methylation of these CpG sites with asthma in the offspring. Nonetheless the current results, particularly the pathway analysis which show the immune response pathway methylation is influenced by maternal asthma and known association of the sites with asthma related exposure and disease status, should stimulate replication in larger samples and functional studies to explore the biological pathways.

7.2 Limitations

The limitations of the thesis have been discussed at length in the result chapters and in the general discussion sections above; some of the important limitations are further discussed in this section.

7.2.1 Study sample

7.2.1.1 Power of the study

Sample size will affect the power of the analyses. While the association analyses were done in a well phenotyped prospective multigenerational birth cohort there were limited by the number of samples where complete exposure, phenotype and methylation data were available. The analyses presented in this thesis may have been underpowered to detect smaller effects in the background of larger effects of confounding factors. With the limitations of the appropriate sample size the results of the exploratory studies presented in this thesis provide direction for future larger confirmatory studies and can set new hypothesis for functional analyses exploring the biological pathways (244). Power calculations and more appropriate models are now being explored to accurately assess gene-methylation interactions, Romanowska et al. have explored the approaches of studying genotype and DNA methylation interactions by using summarised methylation levels over nearby genomic region which can simplify the biological interpretation (365). Similarly for gene-environment- epigenome interactions, as mentioned earlier Czamara et al. have explored alternative approaches (343). Some of the other ways the problem of sample size can be addressed are further discussed in section 7.3 (Future directions).

7.2.1.2 Confounders and reverse causation

Parent reported wheeze is used as an outcome in chapter 5. This outcome definition has limitations including not all wheeze in childhood being asthma and being subjective to parental understanding. Diagnosing asthma in early childhood is challenging, and for an epidemiological study, parent reported wheeze

continues to be widely used as an outcome. Although the questions used appear simplified, they have been validated in global asthma studies by ISAAC (225, 366). Moreover, asthma is a syndrome with heterogenous phenotypes, and early childhood wheeze in itself is being recognised as a phenotype with long term consequences, including lower lung function later in life. Thus, parent reported wheeze can be used as an outcome, as long as caution is exercised not to generalise this to mean definite asthma.

Self-reported smoking is underestimated, particularly in pregnant mothers. Cotinine levels can be used as a surrogate marker, but this has many limitations. Accuracy of cotinine level depends on the type of sample; measured cotinine in saliva likely to have a better accuracy than urine cotinine. Cotinine level at the time of measurement can be affected by variability in individual's metabolism, inhalation patterns and brand of cigarette, and the levels can also be increased in people using transdermal patches (118). Cotinine is used as it's the metabolite of nicotine with longer half-life, but this is reduced in pregnant women further limiting its role in objectively measuring smoking during pregnancy (367). Collecting repeated samples during pregnancy is challenging and can affect the retaining of participants when multiple assessments need to be carried out. Self-reported smoking will remain a limitation in this study. It is likely that the underreporting might dilute the effect seen on the outcomes. This further enhances the significance of the small effects noted.

The accuracy of measuring environmental exposures significantly impacts epigenetic studies. Although the cohorts used in this thesis are prospectively recruited, missing information has been a problem in the Third Generation Cohort, in particular paternal exposures and disease status information has been limited. This limits the confounding exposures that can be corrected for and can affect the sample size. This was the case in the EWAS for maternal asthma; paternal asthma status was available only for a portion of the sample 61/171 (35.7%, in Table 6-1). This problem was overcome by correcting for paternal asthma after initial EWAS in sensitivity analyses, the details of this are given in chapter 6. Similarly, the phenotype has to be measured accurately. Parent report

of non-infective wheeze was used as the outcome for intergenerational effect of prenatal smoking exposure, which may not accurately mirror asthma in infants. Although wheeze is often observed as a precursor symptom for developing asthma in children (368), it is important to note that wheeze is a broader and heterogeneous phenotype that can be caused by asthma and other respiratory conditions (369). Using wheeze, as a surrogate for asthma in early childhood has to be interpreted with caution but due to difficulties in diagnosing asthma accurately in early childhood, wheeze is commonly used as an outcome in this age group in epidemiological studies of asthma.

Is the epigenetic change secondary to the exposure or the disease status? Reverse causation can be difficult to distinguish from the effect of the exposure. The intergenerational effect of maternal and grand-maternal smoking and EWAS for maternal asthma was carried out in cord blood DNA methylation, thus avoiding the effects of postnatal exposures, disease, and treatment on methylation. Cord blood is more appropriate for studying the association of prenatal exposures. The interplay of prenatal maternal smoking, genetic variants, and DNA methylation of *IL13* influencing asthma related lung function was analysed in peripheral blood at 18 years of age in 1989 birth cohort (F1). The main limitation here is that it is difficult to separate the effect of disease status and other postnatal exposures through childhood. However, the results described in chapter 3 provide plausible support to the idea that methylation might hold the key to gene-environment interactions.

7.2.1.3 Cell type/Tissue

Asthma is a chronic airway disease and ideally airway samples, such as bronchial biopsies or sputum cells should be investigated in association studies of asthma. However, there are ethical and practical problems in large epidemiological studies to collect such samples involving invasive procedures, especially in children. The results of the thesis are based on the DNA from peripheral blood cells, which arguably is not the direct site of disease pathogenesis. However, a range of in utero exposures has been shown to result in epigenetic marks in cord

blood (198, 370), and cord blood methylation has been shown to be associated with a range of postnatal phenotypes (222, 371). Thus, it is plausible that exposures, particularly during early development, may result in epigenetic changes in multiple tissue types for which cord blood acts as a biomarker. Further, peripheral blood is a relevant tissue to study epigenetics that regulates abnormal immune responses leading to asthma.

Peripheral blood can be influenced by other factors including the disease status and treatment; cell counts in peripheral blood will differ based on the disease status, for example, in asthma cell count can be skewed towards Th2 profile. This challenge can be addressed by estimating the cell proportions from the methylation data the using validated methods like Housman`s cell correction panel (372). In cord blood samples the cell proportions are different and a validated cell type estimating panel for cord blood was used to correct for cell type when studying cord blood DNA methylation (309). The small changes seen in peripheral blood methylation related to exposure can represent either large change in a small cell proportion or small change in large cell proportions, which is difficult to distinguish. Thus, the study is limited by not having the sorted cell count from the original blood sample and instead using the surrogate panels for cell distribution. Future studies using sorted cell populations or even single cell epigenetic profiling will be needed to resolve this.

7.2.1.4 Measurement and interpretation of DNA Methylation

With rapidly advancing technology, the resolution at which DNA methylation can be assessed in large epidemiological samples has improved significantly; 27K gave access to 27,000 CpGs, 450K can interrogate 480,000 and EPIC platform can assess methylation at 850,000 CpG sites. This poses a challenge when analysing the samples from different platforms. Although data from both the 450K and EPIC platforms, for the EWAS samples from both platforms were analysed by limiting to the CpGs present in both arrays which would be similar to 450K platform. Similarly the pipelines used for preprocessing the DNA methylation data keep improving and may influence the analysis, however studies looking at samples

using different preprocessing approaches show no significant differences in the results noted (198).

If DNA methylation can mediate the development of asthma, then it is important that methylation is shown to be stable overtime and this needs to be tested in serial samples later in life. There is evidence for the stability of methylation changes resulting from prenatal smoke exposure till later childhood (18 years) in the IOWBC (198) and up to 31 years of age in the Northern Finnish Birth Cohort (373).

Association does not mean causation. A number of approaches have been used to provide support for epigenetic associations in causation of phenotype. Functional analysis of the identified loci using gene expression would add value to the results of EWAS (gene expression data was not available for this thesis analysis), pathway analysis, animal models particularly for controlled exposures to reduce confounders and for transgenerational studies.

Other approaches are being used in epigenetic epidemiology studies to potentially strengthen the causal inference such as "Mendelian Randomisation" (374). In Mendelian randomisation, a genetic variant that is associated with an exposure can be used as an instrumental variable to estimate the causal effect of the exposure on a phenotype of interest (375). Such methods are now being applied in exploring the causal role of DNA methylation in relationship between a prenatal exposure and phenotype in later life, for example a recent study supports a the causal effect of prenatal vitamin B12 responsive DNA methylation on cognition of children (376). Although these approaches can be considered to strengthen the causal inference of epigenetic association studies (376), the limitations of the assumptions, possibility of pleiotropy (a genetic locus affecting more than one phenotype) of the instrumental variable and availability of genetic variants that are robustly associated with the exposure of interest are limitations of the approach (375, 377).

7.3 Future Directions

One approach that will undoubtedly contribute to the understanding of the role of epigenetic mechanisms in the evolution of complex disease is the ability to utilise longitudinal cohort studies (331). Resources that are particularly valuable in this regard are studies that have collected and stored biological samples prospectively, and thus allow for the analysis of epigenetic profiles well before the onset of disease. The serial sampling of biological samples, at multiple time points across the life course, will provide further value, allowing insights into the temporal variation in epigenetic signatures over time. Recruitment and continued follow up of a multigenerational cohort can be challenging needing long time period and increased resources (138) but they provide the best platform for epigenetic studies in asthma.

Limitations of the power of epigenetic studies can be addressed by collaborating with other epidemiological cohorts. The Isle of Wight multigenerational cohort has collaborated in two ways; one is by forming a large consortium of many birth cohorts for epigenetic studies (349), allowing replication of observations in independent cohorts. Secondly, this has enabled the use of meta-analysis of EWAS across cohorts for increased power. For example the Isle of Wight Cohort was one of the 13 cohorts used for meta-analysis of epigenome-wide association for maternal smoking in cord blood DNA methylation (198).

Another potential approach to improve sample size is using the methylation of DNA extracted from blood spots on Guthrie cards collected for neonatal screening. There has been growing use of Guthrie cards in epidemiological epigenetic studies. Although the techniques for DNA extraction from Guthrie cards are more laborious and there are still issues that needs to be addressed such as standardisation of processing and storage of the cards and ethical/consenting issues around using the stored Guthrie cards for research (378, 379). None-the-less use of Guthrie card derived DNA allows assessment of early life epigenetic profiles where biological samples (e.g., cord blood) were not

collected prospectively.

For example, since the completion of this study, use of Guthrie cards collected greater than two decades previously in the IOWBC allowed identification of early life DNA methylation markers of lung function trajectories throughout childhood and adolescence (380). However, there are limitations with this approach, not least that there is evidence that Guthrie card and cord blood derived methylation profiles may not be directly comparable. Recent analysis by Jiang et al. comparing DNA methylation profiles between 34 paired cord blood and Guthrie cards of Third Generation infants shows only around 70% agreement evaluated with the difference in methylation values between the two sources demonstrated via paired t-tests at each CpG site and relatively low consistency measured by correlation (7.7%) and stability measured (6.0%) (239). When further looking at the differentially methylated regions associated with prenatal smoking exposure, interestingly the proportion of sites showing high correlation (correlation ≥0.5) or high ICC (ICC ≥0.5) proportion increased dramatically between the two sources. These findings suggest that the DNA methylation may be comparable in cord blood, particularly with respect to CpG sites modified by in utero exposures. However, these results need replication in larger samples.

In future, it is likely that there will be multiple large epigenome data with ever increasing resolution which will bring challenges of computational analysis and complex approaches in assessing the gene-environment-epigenome interplay in a genetically complex trait like asthma. Various public consortia such as the Encyclopedia of DNA Elements (ENCODE) (381) initiative and the International Human Epigenome Consortium (382) with huge methylome data, will help in building reference maps of human methylation profiles across different cell types and conditions (383).

For future studies, it might boil down to striking a balance between conducting larger but less precise vs smaller but more precise studies. Larger studies will have limitations of less precise characterisation of exposures and outcomes and less disease specific tissue samples demanding less extensive contribution from

the participants and less expensive. In comparison the smaller studies can aim for more precise characterisation of exposures and outcomes with more disease specific tissue collection demanding more extensive contribution from the participants and more expensive.

The ongoing follow up of the Third Generation children will provide a unique opportunity to explore the results from this thesis further. The differentially methylated CpG sites at birth in relation to maternal asthma can be explored further in later childhood to see if the association persists with increasing age and if methylation at these sites is associated with asthma in later childhood.

Further analysis can consider approaches to prove the causality of the associations seen in this thesis by incorporating genotype and gene expression data. Further studies to explore the biological pathways explaining the observed statistical interaction effects may also aid in ascertaining causation. There is great interest in DNA methylation in asthma research studies. Future studies should focus on studying specific measurable exposures with well-defined outcomes (specific asthma phenotype or trait) in appropriate sample size and including the timing of the exposures and epigenetic analysis, while simultaneously considering following factors; impact of cell type studied, role of cross talk of different epigenetic mechanisms and genotype in studies of environment related DNA methylation. DNA methylation can help in predicting resistance to some specific treatment agents and as therapeutic targets in cancer (384). Further expansion of confirmatory studies can help in discovering DNA methylation-based prevention strategies, diagnostic and therapeutic targets in asthma (385).

7.4 Conclusion

Does epigenetics play a role in the development of asthma? Knowledge in this field is ever increasing and parallels can be drawn between the etio-pathophysiology of asthma and patterns of epigenetic modifications. Environment exposures play an important role in asthma, both the development of asthma as well as trigger of symptoms and exacerbation, various such environmental exposures are known to influence DNA methylation. Prenatal exposures play an important role in development of asthma and these exposures are known to influence DNA methylation in cord blood. Asthma is a heritable disease; likewise, there is increasing evidence that epigenetic markers like DNA methylation are also heritable. These common features in etiopathology of asthma and the DNA methylation throw some light on the concept of the role of DNA methylation in asthma.

The project on which this thesis is based on established a third generation cohort leading to a valuable prospective multigenerational cohort, which will help in multigenerational epigenetic study of asthma. This thesis provides exploratory proof of concept analyses and EWAS showing (i) possible role of DNA methylation in mediating gene environment interactions of prenatal smoking exposure and *IL13*, (ii) intergenerational effect of environmental exposure such as maternal and grand-maternal smoking during pregnancy on (a) wheeze in grandchild and (b) DNA methylation levels at birth, and finally (iii) maternal asthma influencing methylation at birth supporting the plausible role of DNA methylation in developmental programming. Replication in other cohorts and integrating models to provide evidence of causality and biological pathways will further strengthen the evidence for the role of DNA methylation in the development of asthma.

References

1. Global initiative for asthma. Global Strategy for Asthma Management and Prevention, 2019. U.S. Dept. of Health and Human Services, Public Health Service. Available from www.ginasthma.org.

2. Pekkanen J, Pearce N. Defining asthma in epidemiological studies. The European respiratory journal. 1999;14(4):951-7.

3. Toelle BG, Peat JK, Salome CM, Mellis CM, Woolcock AJ. Toward a definition of asthma for epidemiology. The American review of respiratory disease. 1992;146(3):633-7.

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

5. Tang EA, Matsui, E., Wiesch, G.D., Samet, J.M.,. Epidemiology of Asthma and Allergic Diseases. Middleton`s Allergy, Principles & Practice. 2. Seventh ed: Elsevier; 2009. p. 715-67.

6. Masoli M, Fabian D, Holt S, Beasley R, Global Initiative for Asthma P. The global burden of asthma: executive summary of the GINA Dissemination Committee report. Allergy. 2004;59(5):469-78.

7. WHO. Asthma 2012 [Available from: http://www.who.int/mediacentre/factsheets/fs307/en/index.html.

8. UK A. Asthma facts [Available from: http://www.asthma.org.uk/.

9. Moorman JE, Rudd RA, Johnson CA, King M, Minor P, Bailey C, et al. National surveillance for asthma--United States, 1980-2004. Morbidity and mortality weekly report Surveillance summaries. 2007;56(8):1-54.

10. Eder W, Ege MJ, von Mutius E. The asthma epidemic. The New England journal of medicine. 2006;355(21):2226-35.

11. Patil VK, Kurukulaaratchy RJ, Venter C, Grundy J, Roberts G, Dean T, et al. Changing prevalence of wheeze, rhinitis and allergic sensitisation in late childhood: findings from 2 Isle of Wight birth cohorts 12 years apart. Clin Exp Allergy. 2015;45(9):1430-8.

12. von Pirquet C. Allergie. Munchen Med Wchnschr. 1906;53:1457

13. Kay AB. 100 years of 'Allergy': can von Pirquet's word be rescued? Clinical & Experimental Allergy. 2006;36(5):555-9.

14. Ishizaka K, Ishizaka T, Hornbrook MM. Physicochemical properties of reaginic antibody. V. Correlation of reaginic activity wth gamma-E-globulin antibody. J Immunol. 1966;97(6):840-53.

15. Owens L, Laing IA, Zhang G, Turner S, Le Souef PN. Prevalence of allergic sensitization, hay fever, eczema, and asthma in a longitudinal birth cohort. J Asthma Allergy. 2018;11:173-80.

16. Pearce N, Pekkanen J, Beasley R. How much asthma is really attributable to atopy? Thorax. 1999;54(3):268-72.

17. Hill DA, Spergel JM. The atopic march: Critical evidence and clinical relevance. Ann Allergy Asthma Immunol. 2018;120(2):131-7.

18. Belgrave DC, Granell R, Simpson A, Guiver J, Bishop C, Buchan I, et al. Developmental profiles of eczema, wheeze, and rhinitis: two population-based birth cohort studies. PLoS Med. 2014;11(10):e1001748.

19. Custovic A, Custovic D, Kljaic Bukvic B, Fontanella S, Haider S. Atopic phenotypes and their implication in the atopic march. Expert Rev Clin Immunol. 2020;16(9):873-81.

20. Illi S, von Mutius E, Lau S, Niggemann B, Gruber C, Wahn U, et al. Perennial allergen sensitisation early in life and chronic asthma in children: a birth cohort study. Lancet. 2006;368(9537):763-70.

21. Ronchetti R, Jesenak M, Rennerova Z, Barreto M, Ronchetti F, Villa MP. Relationship between atopic asthma and the population prevalence rates for asthma or atopy in children: atopic and nonatopic asthma in epidemiology. Allergy Asthma Proc. 2009;30(1):55-63.

22. Turner S. Birth Cohort Studies: Their Next Coming of Age. Am J Respir Crit Care Med. 2020;202(12):1612-4.

23. Oksel C, Granell R, Haider S, Fontanella S, Simpson A, Turner S, et al. Distinguishing Wheezing Phenotypes from Infancy to Adolescence. A Pooled Analysis of Five Birth Cohorts. Ann Am Thorac Soc. 2019;16(7):868-76.

24. Pavord ID, Beasley R, Agusti A, Anderson GP, Bel E, Brusselle G, et al. After asthma: redefining airways diseases. Lancet. 2018;391(10118):350-400.

25. Lambrecht BN, Hammad H. The immunology of asthma. Nat Immunol. 2015;16(1):45-56.

26. Borish L. The immunology of asthma: Asthma phenotypes and their implications for personalized treatment. Ann Allergy Asthma Immunol. 2016;117(2):108-14.

27. Gans MD, Gavrilova T. Understanding the immunology of asthma: Pathophysiology, biomarkers, and treatments for asthma endotypes. Paediatr Respir Rev. 2019.

28. McGregor MC, Krings JG, Nair P, Castro M. Role of Biologics in Asthma. Am J Respir Crit Care Med. 2019;199(4):433-45.

29. Kurukulaaratchy RJ, Matthews S, Waterhouse L, Arshad SH. Factors influencing symptom expression in children with bronchial hyperresponsiveness at 10 years of age. J Allergy Clin Immunol. 2003;112(2):311-6.

30. Reddel HK, Taylor DR, Bateman ED, Boulet LP, Boushey HA, Busse WW, et al. An official American Thoracic Society/European Respiratory Society statement: asthma control and exacerbations: standardizing endpoints for clinical asthma trials and clinical practice. Am J Respir Crit Care Med. 2009;180(1):59-99.

31. Enright PL, Lebowitz MD, Cockroft DW. Physiologic measures: pulmonary function tests. Asthma outcome. Am J Respir Crit Care Med. 1994;149(2 Pt 2):S9-18; discussion S9-20.

32. 2021 N. Asthma: diagnosis, monitoring and chronic asthma management. NICE2021 [Available from: nice.org.uk/ng80.

33. Tepper RS, Wise RS, Covar R, Irvin CG, Kercsmar CM, Kraft M, et al. Asthma outcomes: pulmonary physiology. J Allergy Clin Immunol. 2012;129(3 Suppl):S65-87.

34. Gerritsen J, Koeter GH, Postma DS, Schouten JP, Knol K. Prognosis of asthma from childhood to adulthood. The American review of respiratory disease. 1989;140(5):1325-30.

35. Empey DW, Laitinen LA, Jacobs L, Gold WM, Nadel JA. Mechanisms of bronchial hyperreactivity in normal subjects after upper respiratory tract infection. The American review of respiratory disease. 1976;113(2):131-9.

36. Woolcock AJ, Peat JK. Epidemiology of bronchial hyperresponsiveness. Clinical reviews in allergy. 1989;7(3):245-56.

37. Burrows B, Sears MR, Flannery EM, Herbison GP, Holdaway MD, Silva PA. Relation of the course of bronchial responsiveness from age 9 to age 15 to allergy. Am J Respir Crit Care Med. 1995;152(4 Pt 1):1302-8.

38. Rasmussen F, Taylor DR, Flannery EM, Cowan JO, Greene JM, Herbison GP, et al. Outcome in adulthood of asymptomatic airway hyperresponsiveness in childhood: a longitudinal population study. Pediatric pulmonology. 2002;34(3):164-71.

39. Siwik JP, Johnson CC, Havstad SL, Peterson EL, Ownby DR, Zoratti EM. Airway hyperresponsiveness to methacholine at age 6 to 8 years in nonasthmatic patients is not related to increased health-care utilization for asthma in the ensuing 5 years: a longitudinal study of a birth cohort. Chest.128(4):2420-6.

40. Wadsworth S, Sin D, Dorscheid D. Clinical update on the use of biomarkers of airway inflammation in the management of asthma. Journal of asthma and allergy. 2011;4:77-86.

41. Taylor DR, Pijnenburg MW, Smith AD, De Jongste JC. Exhaled nitric oxide measurements: clinical application and interpretation. Thorax. 2006;61(9):817-27.

42. Jiang J, Malavia N, Suresh V, George SC. Nitric oxide gas phase release in human small airway epithelial cells. Respiratory research. 2009;10:3.

43. Malinovschi A, Janson C, Holmkvist T, Norback D, Merilainen P, Hogman M. Effect of smoking on exhaled nitric oxide and flow-independent nitric oxide exchange parameters. Eur Respir J. 2006;28(2):339-45.

44. Scott M, Raza A, Karmaus W, Mitchell F, Grundy J, Kurukulaaratchy RJ, et al. Influence of atopy and asthma on exhaled nitric oxide in an unselected birth cohort study. Thorax.65(3):258-62.

45. Michils A, Baldassarre S, Van Muylem A. Exhaled nitric oxide and asthma control: a longitudinal study in unselected patients. The European respiratory journal. 2008;31(3):539-46.

46. M MS. The biology of asthma In: P FA, editor. Fishman`s Pulmonary diseases and disorders. 12008. p. 773-85.

47. Frew AJ. The immunology of respiratory allergies. Toxicol Lett. 1996;86(2-3):65-72.

48. Tariq SM, Arshad SH, Matthews SM, Hakim EA. Elevated cord serum IgE increases the risk of aeroallergen sensitization without increasing respiratory allergic symptoms in early childhood. Clinical and Experimental Allergy. 1999;29(8):1042-8.

49. Sadeghnejad A, Karmaus W, Davis S, Kurukulaaratchy RJ, Matthews S, Arshad SH. Raised cord serum immunoglobulin E increases the risk of allergic sensitisation at ages 4 and 10 and asthma at age 10. Thorax. 2004;59(11):936-42.

50. Turner SW, Chang AB, Yang IA. Clinical utility of exhaled nitric oxide fraction in the management of asthma and COPD. Breathe (Sheff). 2019;15(4):306-16.

51. Wong GW, Chow CM. Childhood asthma epidemiology: insights from comparative studies of rural and urban populations. Pediatric pulmonology. 2008;43(2):107-16.

52. Ho SM. Environmental epigenetics of asthma: an update. J Allergy Clin Immunol. 2010;126(3):453-65.

53. Shaheen SO, Adcock IM. The developmental origins of asthma: does epigenetics hold the key? Am J Respir Crit Care Med. 2009;180(8):690-1.

54. Barker DJ. The origins of the developmental origins theory. Journal of internal medicine. 2007;261(5):412-7.

55. Gluckman PD, Hanson MA. Developmental origins of disease paradigm: a mechanistic and evolutionary perspective. Pediatric research. 2004;56(3):311-7.

56. Gluckman PD, Hanson MA, Low FM. The role of developmental plasticity and epigenetics in human health. Birth defects research Part C, Embryo today : reviews. 2011;93(1):12-8.

57. A. Jobe JW, & S. Abman Lung Growth, Development, and Disease. . Fetal and Neonatal Lung Development: Clinical Correlates and Technologies for the Future: Cambridge: Cambridge University Press.; 2016.

58. B. Matthews LJ, S.Lakshminrusimha. Pathophysiology of fetal lung development. Essential of Neonatal Ventilation, E-book: Elsevier Health Sciences; 2018.

59. Midodzi WK, Rowe BH, Majaesic CM, Saunders LD, Senthilselvan A. Early Life Factors Associated with Incidence of Physician-diagnosed Asthma in Preschool Children: Results from the Canadian Early Childhood Development Cohort Study. J Asthma. 2010;47(1):7-13.

60. Strachan DP. Family size, infection and atopy: the first decade of the "hygiene hypothesis". Thorax. 2000;55 Suppl 1:S2-10.

61. Maziak W. The asthma epidemic and our artificial habitats. BMC pulmonary medicine. 2005;5:5.

62. Elliot JG, Carroll NG, James AL, Robinson PJ. Airway alveolar attachment points and exposure to cigarette smoke in utero. Am J Respir Crit Care Med. 2003;167(1):45-9.

63. Magnusson LL, Olesen AB, Wennborg H, Olsen J. Wheezing, asthma, hayfever, and atopic eczema in childhood following exposure to tobacco smoke in fetal life. Clin Exp Allergy. 2005;35(12):1550-6.

64. Alati R, Al Mamun A, O'Callaghan M, Najman JM, Williams GM. In utero and postnatal maternal smoking and asthma in adolescence. Epidemiology. 2006;17(2):138-44.

65. Strachan DP, Cook DG. Health effects of passive smoking. 6. Parental smoking and childhood asthma: longitudinal and case-control studies. Thorax. 1998;53(3):204-12.

66. Lannero E, Wickman M, van Hage M, Bergstrom A, Pershagen G, Nordvall L. Exposure to environmental tobacco smoke and sensitisation in children. Thorax. 2008;63(2):172-6.

67. Allan KM, Prabhu N, Craig LC, McNeill G, Kirby B, McLay J, et al. Maternal vitamin D and E intakes during pregnancy are associated with asthma in children. Eur Respir J. 2015;45(4):1027-36.

68. Beckhaus AA, Garcia-Marcos L, Forno E, Pacheco-Gonzalez RM, Celedon JC, Castro-Rodriguez JA. Maternal nutrition during pregnancy and risk of asthma, wheeze, and atopic diseases during childhood: a systematic review and meta-analysis. Allergy. 2015;70(12):1588-604.

69. Baiz N, Just J, Chastang J, Forhan A, de Lauzon-Guillain B, Magnier AM, et al. Maternal diet before and during pregnancy and risk of asthma and allergic rhinitis in children. Allergy Asthma Clin Immunol. 2019;15:40.

70. Castro-Rodriguez JA, Ramirez-Hernandez M, Padilla O, Pacheco-Gonzalez RM, Perez-Fernandez V, Garcia-Marcos L. Effect of foods and Mediterranean diet during pregnancy and first years of life on wheezing, rhinitis and dermatitis in preschoolers. Allergol Immunopathol (Madr). 2016;44(5):400-9.

71. Miller RL, Ho SM. Environmental epigenetics and asthma: current concepts and call for studies. Am J Respir Crit Care Med. 2008;177(6):567-73.

72. Lai T, Wu M, Liu J, Luo M, He L, Wang X, et al. Acid-Suppressive Drug Use During Pregnancy and the Risk of Childhood Asthma: A Meta-analysis. Pediatrics. 2018;141(2).

73. Devine RE, McCleary N, Sheikh A, Nwaru BI. Acid-suppressive medications during pregnancy and risk of asthma and allergy in children: A systematic review and meta-analysis. J Allergy Clin Immunol. 2017;139(6):1985-8 e12.

74. Hack M, Schluchter M, Cartar L, Rahman M, Cuttler L, Borawski E. Growth of very low birth weight infants to age 20 years. Pediatrics.112(1 Pt 1):e30-8.

75. Brooks AM, Byrd RS, Weitzman M, Auinger P, McBride JT. Impact of low birth weight on early childhood asthma in the United States. Arch Pediat Adol Med. 2001;155(3):401-6.

76. Seidman DS, Laor A, Gale R, Stevenson DK, Danon YL. Is low birth weight a risk factor for asthma during adolescence? Arch Dis Child. 1991;66(5):584-7.

77. Xu XF, Li YJ, Sheng YJ, Liu JL, Tang LF, Chen ZM. Effect of low birth weight on childhood asthma: a meta-analysis. BMC Pediatr. 2014;14:275.

78. Saad NJ, Patel J, Burney P, Minelli C. Birth Weight and Lung Function in Adulthood: A Systematic Review and Meta-analysis. Ann Am Thorac Soc. 2017;14(6):994-1004.

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

80. Ball TM, Castro-Rodriguez JA, Griffith KA, Holberg CJ, Martinez FD, Wright AL. Siblings, day-care attendance, and the risk of asthma and wheezing during childhood. N Engl J Med. 2000;343(8):538-43.

81. McKeever TM, Lewis SA, Smith C, Collins J, Heatlie H, Frischer M, et al. Siblings, multiple births, and the incidence of allergic disease: a birth cohort study using the West Midlands general practice research database. Thorax. 2001;56(10):758-62.

82. Kikkawa T, Yorifuji T, Fujii Y, Yashiro M, Okada A, Ikeda M, et al. Birth order and paediatric allergic disease: A nationwide longitudinal survey. Clin Exp Allergy. 2018;48(5):577-85.

83. Silverman M, Wilson N. Wheezing phenotypes in childhood. Thorax. 1997;52(11):936-7.

84. Gronlund MM, Lehtonen OP, Eerola E, Kero P. Fecal microflora in healthy infants born by different methods of delivery: permanent changes in intestinal flora after cesarean delivery. Journal of Pediatric Gastroenterology & Nutrition. 1999;28(1):19-25.

85. Penders J, Thijs C, van den Brandt PA, Kummeling I, Snijders B, Stelma F, et al. Gut microbiota composition and development of atopic manifestations in infancy: the KOALA Birth Cohort Study. Gut. 2007;56(5):661-7.

86. Subbarao P, Becker A, Brook JR, Daley D, Mandhane PJ, Miller GE, et al. Epidemiology of asthma: risk factors for development. Expert Rev Clin Immu. 2009;5(1):77-95.

87. Camargo CA, Jr., Clark S, Kaplan MS, Lieberman P, Wood RA. Regional differences in EpiPen prescriptions in the United States: the potential role of vitamin D. Journal of Allergy & Clinical Immunology. 2007;120(1):131-6.

88. Mullins RJ, Clark S, Camargo CA, Jr. Regional variation in epinephrine autoinjector prescriptions in Australia: more evidence for the vitamin D-anaphylaxis hypothesis. Annals of Allergy, Asthma, & Immunology. 2009;103(6):488-95.

89. Zittermann A, Dembinski J, Stehle P. Low vitamin D status is associated with low cord blood levels of the immunosuppressive cytokine interleukin-10. Pediatric Allergy & Immunology. 2004;15(3):242-6.

90. Tareke AA, Hadgu AA, Ayana AM, Zerfu TA. Prenatal vitamin D supplementation and child respiratory health: A systematic review and meta-analysis of randomized controlled trials. World Allergy Organ J. 2020;13(12):100486.

91. Litonjua AA. Vitamin D and childhood asthma: causation and contribution to disease activity. Curr Opin Allergy Clin Immunol. 2019;19(2):126-31.

92. Kozyrskyj AL, Kendall GE, Jacoby P, Sly PD, Zubrick SR. Association between socioeconomic status and the development of asthma: analyses of income trajectories. Am J Public Health. 2010;100(3):540-6.

93. Lindbaek M, Wefring KW, Grangard EH, Ovsthus K. Socioeconomical conditions as risk factors for bronchial asthma in children aged 4-5 yrs. Eur Respir J. 2003;21(1):105-8.

94. Almqvist C, Pershagen G, Wickman M. Low socioeconomic status as a risk factor for asthma, rhinitis and sensitization at 4 years in a birth cohort. Clin Exp Allergy. 2005;35(5):612-8.

95. Bloomberg GR, Banister C, Sterkel R, Epstein J, Bruns J, Swerczek L, et al. Socioeconomic, family, and pediatric practice factors that affect level of asthma control. Pediatrics. 2009;123(3):829-35.

96. Illi S, von Mutius E, Lau S, Bergmann R, Niggemann B, Sommerfeld C, et al. Early childhood infectious diseases and the development of asthma up to school age: a birth cohort study. BMJ. 2001;322(7283):390-5.

97. Oddy WH, Holt PG, Sly PD, Read AW, Landau LI, Stanley FJ, et al. Association between breast feeding and asthma in 6 year old children: findings of a prospective birth cohort study. Brit Med J. 1999;319(7213):815-9.

98. Sears MR, Greene JM, Willan AR, Wiecek EM, Taylor DR, Flannery EM, et al. A longitudinal, population-based, cohort study of childhood asthma followed to adulthood. New Engl J Med. 2003;349(15):1414-22.

99. Lima RD, Victora CG, Menezes AMB, Barros FC. Do risk factors for childhood infections and malnutrition protect against asthma? A study of Brazilian male adolescents. Am J Public Health. 2003;93(11):1858-64.

100. Ogbuanu IU, Karmaus W, Arshad SH, Kurukulaaratchy RJ, Ewart S. Effect of breastfeeding duration on lung function at age 10 years: a prospective birth cohort study. Thorax. 2009;64(1):62-6.

101. Dogaru CM, Nyffenegger D, Pescatore AM, Spycher BD, Kuehni CE. Breastfeeding and childhood asthma: systematic review and meta-analysis. Am J Epidemiol. 2014;179(10):1153-67.

102. Taussig LM, Wright AL, Holberg CJ, Halonen M, Morgan WJ, Martinez FD. Tucson Children's Respiratory Study: 1980 to present. J Allergy Clin Immunol. 2003;111(4):661-75; quiz 76.

103. Takkouche B, Gonzalez-Barcala FJ, Etminan M, Fitzgerald M. Exposure to furry pets and the risk of asthma and allergic rhinitis: a meta-analysis. Allergy. 2008;63(7):857-64.

104. Chen CM, Heinrich J. Re: Exposure to furry pets and the risk of asthma and allergic rhinitis: a meta-analysis. Allergy. 2009;64(3):494-5.

105. Lødrup Carlsen KC RS, Carlsen KH, Mowinckel P, Wijga AH, Brunekreef B, Torrent M, Roberts G, Arshad SH, Kull I, Krämer U, von Berg A, Eller E, Høst A, Kuehni C, Spycher B, Sunyer J, Chen CM, Reich A, Asarnoj A, Puig C, Herbarth O, Mahachie John JM, Van Steen K, Willich SN, Wahn U, Lau S, Keil T; GALEN WP 1.5 'Birth Cohorts' working group. Does pet ownership in infancy lead to asthma or allergy at school age? Pooled analysis of individual participant data from 11 European birth cohorts. PLoS ONE [Electronic Resource]. 2012;7(8):e43214.

106. Kurukulaaratchy RJ, Matthews S, Arshad SH. Relationship between childhood atopy and wheeze: what mediates wheezing in atopic phenotypes? Annals of Allergy, Asthma, & Immunology.97(1):84-91.

107. Raza A KR, Grundy JD, Clayton CB, Mitchell FA, Roberts G, Ewart S, Sadeghnejad A, Arshad SH. What does adolescent undiagnosed wheeze represent? Findings from the Isle of Wight Cohort. Eur Respir J. 2012;Sep;40(3):.

108. Genuneit J, Weinmayr G, Radon K, Dressel H, Windstetter D, Rzehak P, et al. Smoking and the incidence of asthma during adolescence: results of a large cohort study in Germany. Thorax. 2006;61(7):572-8.

109. Larsson L. Incidence of asthma in Swedish teenagers: relation to sex and smoking habits. Thorax. 1995;50(3):260-4.

110. Withers NJ, Low L, Holgate ST, Clough JB. The natural history of respiratory symptoms in a cohort of adolescents. Am J Respir Crit Care Med. 1998;158(2):352-7.

111. Dimova S, Hoet PH, Nemery B. Paracetamol (acetaminophen) cytotoxicity in rat type II pneumocytes and alveolar macrophages in vitro. Biochemical pharmacology. 2000;59(11):1467-75.

112. Shaheen S, Potts J, Gnatiuc L, Makowska J, Kowalski ML, Joos G, et al. The relation between paracetamol use and asthma: a GA2LEN European case-control study. The European respiratory journal. 2008;32(5):1231-6.

113. Beasley R, Clayton T, Crane J, von Mutius E, Lai CKW, Montefort S, et al. Association between paracetamol use in infancy and childhood, and risk of asthma, rhinoconjunctivitis, and eczema in children aged 6-7 years: analysis from Phase Three of the ISAAC programme. Lancet. 2008;372(9643):1039-48.

114. Turner S, Upham JW. Editorial: Asthma in Children and Adults - What Are the Differences and What Can They Tell Us About Asthma? Front Pediatr. 2020;8:141.

115. Michel G, Silverman M, Strippoli MP, Zwahlen M, Brooke AM, Grigg J, et al. Parental understanding of wheeze and its impact on asthma prevalence estimates. Eur Respir J. 2006;28(6):1124-30.

116. Mallol J, Crane J, von Mutius E, Odhiambo J, Keil U, Stewart A, et al. The International Study of Asthma and Allergies in Childhood (ISAAC) Phase Three: a global synthesis. Allergol Immunopathol (Madr). 2013;41(2):73-85.

117. Ait-Khaled N, Pearce N, Anderson HR, Ellwood P, Montefort S, Shah J, et al. Global map of the prevalence of symptoms of rhinoconjunctivitis in children: The International Study of Asthma and Allergies in Childhood (ISAAC) Phase Three. Allergy. 2009;64(1):123-48.

118. Connor Gorber S, Schofield-Hurwitz S, Hardt J, Levasseur G, Tremblay M. The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. Nicotine Tob Res. 2009;11(1):12-24.

119. The Lancet Respiratory M. Asthma: looking beyond the headlines. Lancet Respir Med. 2019;7(5):365.

120. Matsui EC, Keet CA. Weighing the evidence: Bias and confounding in epidemiologic studies in allergy/immunology. J Allergy Clin Immunol. 2017;139(2):448-50.

121. Tiotiu A, Ioan I, Wirth N, Romero-Fernandez R, Gonzalez-Barcala FJ. The Impact of Tobacco Smoking on Adult Asthma Outcomes. Int J Environ Res Public Health. 2021;18(3).

122. Aurrekoetxea JJ, Murcia M, Rebagliato M, Lopez MJ, Castilla AM, Santa-Marina L, et al. Determinants of self-reported smoking and misclassification during pregnancy, and analysis of optimal cut-off points for urinary cotinine: a cross-sectional study. Bmj Open. 2013;3(1).

123. Hilberink SR, Jacobs JE, van Opstal S, van der Weijden T, Keegstra J, Kempers P, et al. Validation of smoking cessation self-reported by patients with chronic obstructive pulmonary disease. Int J Gen Med. 2011;4:85-90.

124. Benowitz NL. Cotinine as a biomarker of environmental tobacco smoke exposure. Epidemiol Rev. 1996;18(2):188-204.

125. Dietz PM, Homa D, England LJ, Burley K, Tong VT, Dube SR, et al. Estimates of nondisclosure of cigarette smoking among pregnant and nonpregnant women of reproductive age in the United States. Am J Epidemiol. 2011;173(3):355-9.

126. Gopal SH, Mukherjee S, Das SK. Direct and Second Hand Cigarette Smoke Exposure and Development of Childhood Asthma. J Environ Health Sci. 2016;2(6).

127. Cao S, Yang C, Gan Y, Lu Z. The Health Effects of Passive Smoking: An Overview of Systematic Reviews Based on Observational Epidemiological Evidence. PLoS One. 2015;10(10):e0139907.

128. Thomson NC, Chaudhuri R, Livingston E. Asthma and cigarette smoking. Eur Respir J. 2004;24(5):822-33.

129. Shiffman S, Tindle H, Li X, Scholl S, Dunbar M, Mitchell-Miland C. Characteristics and smoking patterns of intermittent smokers. Exp Clin Psychopharmacol. 2012;20(4):264-77.

130. Buescher JL, Musselman LP, Wilson CA, Lang T, Keleher M, Baranski TJ, et al. Evidence for transgenerational metabolic programming in Drosophila. Dis Model Mech. 2013;6(5):1123-32.

131. Gapp K, Jawaid A, Sarkies P, Bohacek J, Pelczar P, Prados J, et al. Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. Nat Neurosci. 2014;17(5):667-9.

132. Rehan VK, Liu J, Naeem E, Tian J, Sakurai R, Kwong K, et al. Perinatal nicotine exposure induces asthma in second generation offspring. Bmc Med. 2012;10:129.

133. Pembrey ME, Bygren LO, Kaati G, Edvinsson S, Northstone K, Sjostrom M, et al. Sex-specific, male-line transgenerational responses in humans. Eur J Hum Genet. 2006;14(2):159-66.

134. Kaati G, Bygren LO, Edvinsson S. Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period. Eur J Hum Genet. 2002;10(11):682-8.

135. Li YF, Langholz B, Salam MT, Gilliland FD. Maternal and grandmaternal smoking patterns are associated with early childhood asthma. Chest. 2005;127(4):1232-41.

136. Miller LL, Henderson J, Northstone K, Pembrey M, Golding J. Do grandmaternal smoking patterns influence the etiology of childhood asthma? Chest. 2014;145(6):1213-8.

137. Lodge CJ, Braback L, Lowe AJ, Dharmage SC, Olsson D, Forsberg B. Grandmaternal smoking increases asthma risk in grandchildren: A nationwide Swedish cohort. Clin Exp Allergy. 2018;48(2):167-74.

138. Arshad SH, Karmaus W, Zhang H, Holloway JW. Multigenerational cohorts in patients with asthma and allergy. J Allergy Clin Immunol. 2017;139(2):415-21.

139. Los H, Postmus PE, Boomsma DI. Asthma genetics and intermediate phenotypes: a review from twin studies. Twin research : the official journal of the International Society for Twin Studies. 2001;4(2):81-93.

140. Zhang J, Pare PD, Sandford AJ. Recent advances in asthma genetics. Respiratory research. 2008;9:4.

141. Meyers DA. Genetics of asthma and allergy: what have we learned? J Allergy Clin Immunol. 2010;126(3):439-46; quiz 47-8.

142. Akhabir L, Sandford AJ. Genome-wide association studies for discovery of genes involved in asthma. Respirology. 2011;16(3):396-406.

143. Zhao L, Bracken MB. Association of CD14 -260 (-159) C>T and asthma: a systematic review and meta-analysis. BMC medical genetics. 2011;12:93.

144. Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? Am J Hum Genet. 2001;69(1):124-37.

145. Reich DE, Lander ES. On the allelic spectrum of human disease. Trends in genetics : TIG. 2001;17(9):502-10.

146. Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. Nature. 2007;448(7152):470-3.

147. Li X, Howard TD, Zheng SL, Haselkorn T, Peters SP, Meyers DA, et al. Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions. J Allergy Clin Immunol. 2010;125(2):328-35 e11.

148. Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP, et al. Genome-wide Association Analysis Identifies PDE4D as an Asthma-Susceptibility Gene. Am J Hum Genet. 2009;84(5):581-93.

149. Sleiman PMA, Flory J, Imielinski M, Bradfield JP, Annaiah K, Willis-Owen SAG, et al. Variants of DENND1B Associated with Asthma in Children. New Engl J Med. 2010;362(1):36-44.

150. Hancock DB, Romieu I, Shi M, Sienra-Monge JJ, Wu H, Chiu GY, et al. Genome-Wide Association Study Implicates Chromosome 9q21.31 as a Susceptibility Locus for Asthma in Mexican Children. Plos Genet. 2009;5(8).

151. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. The New England journal of medicine. 2010;363(13):1211-21.

152. Johansson A, Rask-Andersen M, Karlsson T, Ek WE. Genome-wide association analysis of 350 000 Caucasians from the UK Biobank identifies novel loci for asthma, hay fever and eczema. Hum Mol Genet. 2019;28(23):4022-41.

153. Dewan AT, Egan KB, Hellenbrand K, Sorrentino K, Pizzoferrato N, Walsh KM, et al. Whole-exome sequencing of a pedigree segregating asthma. BMC medical genetics. 2012;13(1):95.

154. Ober C, Yao TC. The genetics of asthma and allergic disease: a 21st century perspective. Immunol Rev. 2011;242(1):10-30.

155. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature. 2009;461(7265):747-53.

156. Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, et al. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet. 2010;11(6):446-50.

157. Antoniades L, MacGregor AJ, Andrew T, Spector TD. Association of birth weight with osteoporosis and osteoarthritis in adult twins. Rheumatology (Oxford). 2003;42(6):791-6.

158. McLeish S, Turner SW. Gene-environment interactions in asthma. Arch Dis Child. 2007;92(11):1032-5.

159. Kabesch M. Gene by environment interactions and the development of asthma and allergy. Toxicol Lett. 2006;162(1):43-8.

160. Su MW, Tung KY, Liang PH, Tsai CH, Kuo NW, Lee YL. Gene-gene and gene-environmental interactions of childhood asthma: a multifactor dimension reduction approach. PloS one. 2012;7(2):e30694.

161. Choudhry S, Avila PC, Nazario S, Ung N, Kho J, Rodriguez-Santana JR, et al. CD14 tobacco gene-environment interaction modifies asthma severity and immunoglobulin E levels in Latinos with asthma. Am J Respir Crit Care Med. 2005;172(2):173-82.

162. Reijmerink NE, Kerkhof M, Koppelman GH, Gerritsen J, de Jongste JC, Smit HA, et al. Smoke exposure interacts with ADAM33 polymorphisms in the development of lung function and hyperresponsiveness. Allergy. 2009;64(6):898-904.

163. Ramadas RA, Sadeghnejad A, Karmaus W, Arshad SH, Matthews S, Huebner M, et al. Interleukin-1R antagonist gene and pre-natal smoke exposure are associated with childhood asthma. The European respiratory journal. 2007;29(3):502-8.

164. Sadeghnejad A, Karmaus W, Arshad SH, Kurukulaaratchy R, Huebner M, Ewart S. IL13 gene polymorphisms modify the effect of exposure to tobacco smoke on persistent wheeze and asthma in childhood, a longitudinal study. Respir Res. 2008;9:2.

165. Baedke J. The epigenetic landscape in the course of time: Conrad Hal Waddington's methodological impact on the life sciences. Stud Hist Philos Biol Biomed Sci. 2013;44(4 Pt B):756-73.

166. Waddington CH. The epigenotype. 1942. Int J Epidemiol. 2012;41(1):10-3.

167. Heijmans BT, Tobi EW, Lumey LH, Slagboom PE. The epigenome: archive of the prenatal environment. Epigenetics : official journal of the DNA Methylation Society. 2009;4(8):526-31.

168. Gluckman PD, Hanson MA, Cooper C, Thornburg KL. Effect of in utero and early-life conditions on adult health and disease. The New England journal of medicine. 2008;359(1):61-73.

169. Szyf M. The early life environment and the epigenome☆. Biochimica et Biophysica Acta (BBA) - General Subjects. 2009;1790(9):878-85.

170. Cannell IG, Kong YW, Bushell M. How do microRNAs regulate gene expression? Biochem Soc T. 2008;36:1224-31.

171. Ho S-M. Environmental epigenetics of asthma: An update. Journal of Allergy and Clinical Immunology. 2010;126(3):453-65.

172. Durham AL, Wiegman C, Adcock IM. Epigenetics of asthma. Biochimica et biophysica acta. 2011;1810(11):1103-9.

173. Razin A. CpG methylation,chromatin structure and gene silencing--a three-way connection. The EMBO Journal. 1998;17.

174. Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. Nature reviews Molecular cell biology. 2010;11(9):607-20.

175. Irizarry RA, Ladd-Acosta C, Wen B, Wu ZJ, Montano C, Onyango P, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet. 2009;41(2):178-86.

176. Siedlecki P, Zielenkiewicz P. Mammalian DNA methyltransferases. Acta biochimica Polonica. 2006;53(2):245-56.

177. Freitag M, Selker EU. Controlling DNA methylation: many roads to one modification. Current opinion in genetics & development. 2005;15(2):191-9.

178. Comb M, Goodman HM. CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. Nucleic Acids Res. 1990;18(13):3975-82.

179. Bibikova M, Fan JB. GoldenGate assay for DNA methylation profiling. Methods in molecular biology. 2009;507:149-63.

180. Bibikova M, Fan JB. Genome-wide DNA methylation profiling. Wiley interdisciplinary reviews Systems biology and medicine. 2010;2(2):210-23.

181. Laird PW. Principles and challenges of genomewide DNA methylation analysis. Nat Rev Genet. 2010;11(3):191-203.

182. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. Genomics. 2011;98(4):288-95.

183. Wang RY, Gehrke CW, Ehrlich M. Comparison of bisulfite modification of 5-methyldeoxycytidine and deoxycytidine residues. Nucleic Acids Res. 1980;8(20):4777-90.

184. Hayatsu H. Discovery of bisulfite-mediated cytosine conversion to uracil, the key reaction for DNA methylation analysis--a personal account. Proceedings of the Japan Academy Series B, Physical and biological sciences. 2008;84(8):321-30.

185. Bibikova M, Lin ZW, Zhou LX, Chudin E, Garcia EW, Wu B, et al. High-throughput DNA methylation profiling using universal bead arrays. Genome Res. 2006;16(3):383-93.

186. Poulsen P, Esteller M, Vaag A, Fraga MF. The epigenetic basis of twin discordance in age-related diseases. Pediatric research. 2007;61(5 Pt 2):38R-42R.

187. Cooney CA, Dave AA, Wolff GL. Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. The Journal of nutrition. 2002;132(8 Suppl):2393S-400S.

188. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Molecular and cellular biology. 2003;23(15):5293-300.

189. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci U S A. 2008;105(44):17046-9.

190. Heijmans. The epigenome - Archive of the prenatal environment 2009.

191. Waterland RA, Jirtle RL. Transposable Elements: Targets for Early Nutritional Effects on Epigenetic Gene Regulation. Molecular and cellular biology. 2003;23(15):5293-300.

192. Weaver IC, La Plante P, Weaver S, Parent A, Sharma S, Diorio J, et al. Early environmental regulation of hippocampal glucocorticoid receptor gene expression:

characterization of intracellular mediators and potential genomic target sites. Mol Cell Endocrinol. 2001;185(1-2):205-18.

193. Launay JM, Del Pino M, Chironi G, Callebert J, Peoc'h K, Megnien JL, et al. Smoking Induces Long-Lasting Effects through a Monoamine-Oxidase Epigenetic Regulation. PloS one. 2009;4(11).

194. Digel W, Lubbert M. DNA methylation disturbances as novel therapeutic target in lung cancer: Preclinical and clinical results. Crit Rev Oncol Hemat. 2005;55(1):1-11.

195. Franco R, Schoneveld O, Georgakilas AG, Panayiotidis MI. Oxidative stress, DNA methylation and carcinogenesis. Cancer letters. 2008;266(1):6-11.

196. Breton CV, Byun HM, Wenten M, Pan F, Yang A, Gilliland FD. Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. Am J Respir Crit Care Med. 2009;180(5):462-7.

197. Joubert BR, Haberg SE, Nilsen RM, Wang X, Vollset SE, Murphy SK, et al. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. Environ Health Perspect. 2012;120(10):1425-31.

198. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, et al. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. Am J Hum Genet. 2016;98(4):680-96.

199. Doherty AS, Mann MR, Tremblay KD, Bartolomei MS, Schultz RM. Differential effects of culture on imprinted H19 expression in the preimplantation mouse embryo. Biol Reprod. 2000;62(6):1526-35.

200. Bell JT, Tsai PC, Yang TP, Pidsley R, Nisbet J, Glass D, et al. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. Plos Genet. 2012;8(4):e1002629.

201. Crider KS, Quinlivan EP, Berry RJ, Hao L, Li Z, Maneval D, et al. Genomic DNA Methylation Changes in Response to Folic Acid Supplementation in a Population-Based Intervention Study among Women of Reproductive Age. PloS one. 2011;6(12).

202. Nadeau JH. Transgenerational genetic effects on phenotypic variation and disease risk. Human molecular genetics. 2009;18(R2):R202-10.

203. Bae DJ, Jun JA, Chang HS, Park JS, Park CS. Epigenetic Changes in Asthma: Role of DNA CpG Methylation. Tuberc Respir Dis (Seoul). 2020;83(1):1-13.

204. Yang Y, Haitchi HM, Cakebread J, Sammut D, Harvey A, Powell RM, et al. Epigenetic mechanisms silence a disintegrin and metalloprotease 33 expression in bronchial epithelial cells. J Allergy Clin Immunol. 2008;121(6):1393-9, 9 e1-14.

205. Council I. The Island's Demographics and Population. In: council IoW, editor. 2008.

206. Ziyab AH, Raza A, Karmaus W, Tongue N, Zhang H, Matthews S, et al. Trends in eczema in the first 18 years of life: results from the Isle of Wight 1989 birth cohort study. Clinical & Experimental Allergy.40(12):1776-84.

207. Sadeghnejad A, Karmaus W, Hasan Arshad S, Ewart S. IL13 gene polymorphism association with cord serum immunoglobulin E. Pediatric Allergy & Immunology.18(4):288-92.

208. Kurukulaaratchy RJ, Raza A, Scott M, Williams P, Ewart S, Matthews S, et al. Characterisation of asthma that develops during adolescence; findings from the Isle of Wight Birth Cohort. Respiratory Medicine.106(3):329-37.

209. Kurukulaaratchy RJ, Fenn M, Twiselton R, Matthews S, Arshad SH. The prevalence of asthma and wheezing illnesses amongst 10-year-old schoolchildren. Respiratory Medicine.96(3):163-9.

210. Karmaus W, Dobai AL, Ogbuanu I, Arshard SH, Matthews S, Ewart S. Long-term effects of breastfeeding, maternal smoking during pregnancy, and recurrent lower respiratory tract infections on asthma in children. J Asthma.45(8):688-95.

211. Arshad SH, Karmaus W, Kurukulaaratchy R, Sadeghnejad A, Huebner M, Ewart S. Polymorphisms in the interleukin 13 and GATA binding protein 3 genes and the development of eczema during childhood. British Journal of Dermatology.158(6):1315-22.

212. Arshad SH, Karmaus W, Raza A, Kurukulaaratchy RJ, Matthews SM, Holloway JW, et al. The effect of parental allergy on childhood allergic diseases depends on the sex of the child. J Allergy Clin Immunol. 2012;130(2):427-34 e6.

213. Arshad SH, Holloway JW, Karmaus W, Zhang H, Ewart S, Mansfield L, et al. Cohort Profile: The Isle Of Wight Whole Population Birth Cohort (IOWBC). Int J Epidemiol. 2018.

214. Kurukulaaratchy RJ, Waterhouse L, Matthews SM, Arshad SH. Are influences during pregnancy associated with wheezing phenotypes during the first decade of life? Acta Paediatrica.94(5):553-8.

215. Arshad SH, Kurukulaaratchy RJ, Fenn M, Matthews S. Early life risk factors for current wheeze, asthma, and bronchial hyperresponsiveness at 10 years of age. Chest. 2005;127(2):502-8.

216. Kurukulaaratchy RJ, Matthews S, Arshad SH. Does environment mediate earlier onset of the persistent childhood asthma phenotype? Pediatrics. 2004;113(2):345-50.

217. Kurukulaaratchy RJ, Matthews S, Waterhouse L, Arshad SH. Factors influencing symptom expression in children with bronchial hyperresponsiveness at 10 years of age. J Allergy Clin Immun. 2003;112(2):311-6.

218. Kurukulaaratchy RJ, Raza A, Scott M, Williams P, Ewart S, Matthews S, et al. Characterisation of asthma that develops during adolescence; findings from the Isle of Wight Birth Cohort. Respir Med. 2012;106(3):329-37.

219. Lim RH, Kobzik L, Dahl M. Risk for asthma in offspring of asthmatic mothers versus fathers: a meta-analysis. PLoS One. 2010;5(4):e10134.

220. Litonjua AA, Carey VJ, Burge HA, Weiss ST, Gold DR. Parental history and the risk for childhood asthma. Does mother confer more risk than father? Am J Respir Crit Care Med. 1998;158(1):176-81.

221. Arshad SH, Kurukulaaratchy RJ, Fenn M, Matthews S. Early life risk factors for current wheeze, asthma, and bronchial hyperresponsiveness at 10 years of age. Chest. 2005;127(2):502-8.

222. DeVries A, Wlasiuk G, Miller SJ, Bosco A, Stern DA, Lohman IC, et al. Epigenome-wide analysis links SMAD3 methylation at birth to asthma in children of asthmatic mothers. J Allergy Clin Immunol. 2017;140(2):534-42.

223. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A. 2005;102(30):10604-9.

224. Kim E, Kwak SH, Chung HR, Ohn JH, Bae JH, Choi SH, et al. DNA methylation profiles in sibling pairs discordant for intrauterine exposure to maternal gestational diabetes. Epigenetics. 2017;12(10):825-32.

225. Asher MI, Keil U, Anderson HR, Beasley R, Crane J, Martinez F, et al. International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. Eur Respir J. 1995;8(3):483-91.

226. Kurukulaaratchy RJ. The natural history of asthma and atopic disease in childhood and the risk factors contributing to their development.: University of Southampton; 2002.

227. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome Biol. 2016;17(1):208.

228. Illumina. BeadStudio Gene Expression User Guide. Illumina.

229. Kuan PF, Wang S, Zhou X, Chu H. A statistical framework for Illumina DNA methylation arrays. Bioinformatics. 2010;26(22):2849-55.

230. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC bioinformatics. 2010;11:587.

231. Wang D, Yan L, Hu Q, Sucheston LE, Higgins MJ, Ambrosone CB, et al. IMA: an R package for high-throughput analysis of Illumina's 450K Infinium methylation data. Bioinformatics. 2012;28(5):729-30.

232. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics. 2003;19(2):185-93.

233. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. Epigenomics. 2011;3(6):771-84.

234. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. Genome Biol. 2015;16:37.

235. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8(1):118-27.

236. Chen C, Grennan K, Badner J, Zhang D, Gershon E, Jin L, et al. Removing Batch Effects in Analysis of Expression Microarray Data: An Evaluation of Six Batch Adjustment Methods. PloS one. 2011;6(2):e17238.

237. Liu J, Morgan M, Hutchison K, Calhoun VD. A study of the influence of sex on genome wide methylation. PloS one. 2010;5(4):e10028.

238. Beyan H, Down TA, Ramagopalan SV, Uvebrant K, Nilsson A, Holland ML, et al. Guthrie card methylomics identifies temporally stable epialleles that are present at birth in humans. Genome Res. 2012;22(11):2138-45.

239. Jiang Y, Wei J, Zhang H, Ewart S, Rezwan FI, Holloway JW, et al. Epigenome wide comparison of DNA methylation profile between paired umbilical cord blood and neonatal blood on Guthrie cards. Epigenetics. 2020;15(5):454-61.

240. Karmaus W, Ziyab AH, Everson T, Holloway JW. Epigenetic mechanisms and models in the origins of asthma. Current opinion in allergy and clinical immunology. 2013;13(1):63-9.

241. Hwang SJ, Beaty TH, Liang KY, Coresh J, Khoury MJ. Minimum sample size estimation to detect gene-environment interaction in case-control designs. Am J Epidemiol. 1994;140(11):1029-37.

242. Foppa I, Spiegelman D. Power and sample size calculations for case-control studies of gene-environment interactions with a polytomous exposure variable. Am J Epidemiol. 1997;146(7):596-604.

243. Luan JA, Wong MY, Day NE, Wareham NJ. Sample size determination for studies of gene-environment interaction. Int J Epidemiol. 2001;30(5):1035-40.

244. Dempfle A, Scherag A, Hein R, Beckmann L, Chang-Claude J, Schafer H. Gene-environment interactions for complex traits: definitions, methodological requirements and challenges. Eur J Hum Genet. 2008;16(10):1164-72.

245. Benjamini Y HY. Controlling the False Discovery Rate- A Practical and Powerful Approach to Multiple Testing. Benjamini 1995. Journal of the Royal Statistical Society. 1995;Series B (Methodological), Vol. 57,:289-300.

246. Crapo RO, Casaburi R, Coates AL, Enright PL, Hankinson JL, Irvin CG, et al. Guidelines for methacholine and exercise challenge testing-1999. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, July 1999. American journal of respiratory and critical care medicine. 2000;161(1):309-29.

247. Chai H, Farr RS, Froehlich LA, Mathison DA, McLean JA, Rosenthal RR, et al. Standardization of bronchial inhalation challenge procedures. J Allergy Clin Immunol. 1975;56(4):323-7.

248. Hancox RJ, Milne BJ, Poulton R, Taylor DR, Greene JM, McLachlan CR, et al. Sex differences in the relation between body mass index and asthma and atopy in a birth cohort. Am J Respir Crit Care Med. 2005;171(5):440-5.

249. Sears MR, Greene JM, Willan AR, Wiecek EM, Taylor DR, Flannery EM, et al. A longitudinal, population-based, cohort study of childhood asthma followed to adulthood. N Engl J Med. 2003;349(15):1414-22.

250. Arshad SH, Bateman B, Matthews SM. Primary prevention of asthma and atopy during childhood by allergen avoidance in infancy: a randomised controlled study. Thorax. 2003;58(6):489-93.

251. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. Eur Respir J. 2005;26(2):319-38.

252. Agache I, Akdis C, Jutel M, Virchow JC. Untangling asthma phenotypes and endotypes. Allergy. 2012;67(7):835-46.

253. Kabesch M, Michel S, Tost J. Epigenetic mechanisms and the relationship to childhood asthma. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology. 2010;36(4):950-61.

254. Joubert BR, Haberg SE, Nilsen RM, Wang X, Vollset SE, Murphy SK, et al. 450K Epigenome-Wide Scan Identifies Differential DNA Methylation in Newborns Related to Maternal Smoking During Pregnancy. Environmental health perspectives. 2012.

255. Bell CG, Finer S, Lindgren CM, Wilson GA, Rakyan VK, Teschendorff AE, et al. Integrated genetic and epigenetic analysis identifies haplotype-specific methylation in the FTO type 2 diabetes and obesity susceptibility locus. PloS one. 2010;5(11):e14040.

256. Hanson LA, Korotkova M, Haversen L, Mattsby-Baltzer I, Hahn-Zoric M, Silfverdal SA, et al. Breast-feeding, a complex support system for the offspring. Pediatr Int. 2002;44(4):347-52.

257. Romagnani S. Th1/Th2 cells. Inflamm Bowel Dis. 1999;5(4):285-94.

258. Gould HJ, Sutton BJ, Beavil AJ, Beavil RL, McCloskey N, Coker HA, et al. The biology of IGE and the basis of allergic disease. Annu Rev Immunol. 2003;21:579-628.

259. Perkins C, Yanase N, Smulian G, Gildea L, Orekov T, Potter C, et al. Selective stimulation of IL-4 receptor on smooth muscle induces airway hyperresponsiveness in mice. J Exp Med.208(4):853-67.

260. Izuhara K, Shirakawa T. Signal transduction via the interleukin-4 receptor and its correlation with atopy. Int J Mol Med. 1999;3(1):3-10.

261. Loza MJ, Chang BL. Association between Q551R IL4R genetic variants and atopic asthma risk demonstrated by meta-analysis. J Allergy Clin Immunol. 2007;120(3):578-85.

262. Michel S, Liang L, Depner M, Klopp N, Ruether A, Kumar A, et al. Unifying candidate gene and GWAS Approaches in Asthma. PLoS One.5(11):e13894.

263. Rosa-Rosa L, Zimmermann N, Bernstein JA, Rothenberg ME, Khurana Hershey GK. The R576 IL-4 receptor alpha allele correlates with asthma severity. J Allergy Clin Immunol. 1999;104(5):1008-14.

264. R. E. Slager, G. Hawkins, B. Otulana, Y. Yen, S. E. Wenzel, D. A. Meyers, et al. Interleukin 4 Receptor Polymorphisms Predict Therapeutic Pitrakinra Treatment

Response In Moderate To Severe Asthma. American Thoracic Society; Denver, Colorado: Am J Respir Crit Care Med; 2011.

265. Hill WG, Robertson A. The effect of linkage on limits to artificial selection. Genet Res. 1966;8(3):269-94.

266. Soto-Ramirez N, Arshad SH, Holloway JW, Zhang H, Schauberger E, Ewart S, et al. The interaction of genetic variants and DNA methylation of the interleukin-4 receptor gene increase the risk of asthma at age 18 years. Clin Epigenetics. 2013;5(1):1.

267. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing

J Roy Stat Soc Ser B Meth 1995;57:289-300.

268. Jjingo D, Conley AB, Yi SV, Lunyak VV, Jordan IK. On the presence and role of human gene-body DNA methylation. Oncotarget.3(4):462-74.

269. Shenker N, Flanagan JM. Intragenic DNA methylation: implications of this epigenetic mechanism for cancer research. Br J Cancer.106(2):248-53.

270. Karmaus W, Dobai AL, Ogbuanu I, Arshard SH, Matthews S, Ewart S. Long-term effects of breastfeeding, maternal smoking during pregnancy, and recurrent lower respiratory tract infections on asthma in children. The Journal of asthma : official journal of the Association for the Care of Asthma. 2008;45(8):688-95.

271. Wang L, Pinkerton KE. Detrimental effects of tobacco smoke exposure during development on postnatal lung function and asthma. Birth defects research Part C, Embryo today : reviews. 2008;84(1):54-60.

272. Corren J, Lemanske RF, Hanania NA, Korenblat PE, Parsey MV, Arron JR, et al. Lebrikizumab treatment in adults with asthma. The New England journal of medicine. 2011;365(12):1088-98.

273. Yang G, Volk A, Petley T, Emmell E, Giles-Komar J, Shang X, et al. Anti-IL-13 monoclonal antibody inhibits airway hyperresponsiveness, inflammation and airway remodeling. Cytokine. 2004;28(6):224-32.

274. Holloway JW, Yang IA, Holgate ST. Genetics of allergic disease. The Journal of allergy and clinical immunology. 2010;125(2 Suppl 2):S81-94.

275. Nie W, Liu Y, Bian J, Li B, Xiu Q. Effects of polymorphisms -1112C/T and +2044A/G in interleukin-13 gene on asthma risk: a meta-analysis. PLoS One. 2013;8(2):e56065.

276. Cui L, Jia J, Ma CF, Li SY, Wang YP, Guo XM, et al. IL-13 polymorphisms contribute to the risk of asthma: a meta-analysis. Clin Biochem. 2012;45(4-5):285-8.

277. Beghe B, Hall IP, Parker SG, Moffatt MF, Wardlaw A, Connolly MJ, et al. Polymorphisms in IL13 pathway genes in asthma and chronic obstructive pulmonary disease. Allergy. 2010;65(4):474-81.

278. Park H-W, Lee J-E, Kim S-H, Kim Y-K, Min K-U, Kim Y-Y, et al. Genetic variation of IL13 as a risk factor of reduced lung function in children and adolescents: a cross-sectional population-based study in Korea. Respiratory Medicine. 2009;103(2):284-8.

279. Patil VK, Holloway JW, Zhang H, Soto-Ramirez N, Ewart S, Arshad SH, et al. Interaction of prenatal maternal smoking, interleukin 13 genetic variants and DNA methylation influencing airflow and airway reactivity. Clinical epigenetics. 2013;5(1):22.

280. Park HW, Lee JE, Kim SH, Kim YK, Min KU, Kim YY, et al. Genetic variation of IL13 as a risk factor of reduced lung function in children and adolescents: a cross-sectional population-based study in Korea. Respir Med. 2009;103(2):284-8.

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

282. Macy E, Schatz M, Gibbons C, Zeiger R. The prevalence of reversible airflow obstruction and/or methacholine hyperreactivity in random adult asthma patients identified by administrative data. J Asthma. 2005;42(3):213-20.

283. Cockcroft DW, Murdock KY, Berscheid BA, Gore BP. Sensitivity and specificity of histamine PC20 determination in a random selection of young college students. J Allergy Clin Immunol. 1992;89(1 Pt 1):23-30.

284. Ulrik CS, Postma DS, Backer V. Recognition of asthma in adolescents and young adults: which objective measure is best? J Asthma. 2005;42(7):549-54.

285. Lin PI, Shu H, Mersha TB. Comparing DNA methylation profiles across different tissues associated with the diagnosis of pediatric asthma. Sci Rep. 2020;10(1):151.

286. Bergman Y, Cedar H. Epigenetic control of recombination in the immune system. Seminars in immunology. 2010;22(6):323-9.

287. Santos KF, Mazzola TN, Carvalho HF. The prima donna of epigenetics: the regulation of gene expression by DNA methylation. Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas / Sociedade Brasileira de Biofisica [et al]. 2005;38(10):1531-41.

288. Beyan H, Down TA, Ramagopalan SV, Uvebrant K, Nilsson A, Holland ML, et al. Guthrie card methylomics identifies temporally stable epialleles that are present at birth in humans. Genome Research. 2012;22(11):2138-45.

289. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee. Lancet. 1998;351(9111):1225-32.

290. Arshad SH, Holloway JW, Karmaus W, Zhang H, Ewart S, Mansfield L, et al. Cohort Profile: The Isle Of Wight Whole Population Birth Cohort (IOWBC). Int J Epidemiol. 2018;47(4):1043-4i.

291. Granell R, Henderson AJ, Sterne JA. Associations of wheezing phenotypes with late asthma outcomes in the Avon Longitudinal Study of Parents and Children: A population-based birth cohort. J Allergy Clin Immunol. 2016;138(4):1060-70 e11.

292. Hollingsworth JW, Maruoka S, Boon K, Garantziotis S, Li Z, Tomfohr J, et al. In utero supplementation with methyl donors enhances allergic airway disease in mice. J Clin Invest. 2008;118(10):3462-9.

293. Gilliland FD, Li YF, Peters JM. Effects of maternal smoking during pregnancy and environmental tobacco smoke on asthma and wheezing in children. Am J Respir Crit Care Med. 2001;163(2):429-36.

294. Daxinger L, Whitelaw E. Understanding transgenerational epigenetic inheritance via the gametes in mammals. Nat Rev Genet. 2012;13(3):153-62.

295. Ducharme FM, Tse SM, Chauhan B. Diagnosis, management, and prognosis of preschool wheeze. Lancet. 2014;383(9928):1593-604.

296. Accordini S, Calciano L, Johannessen A, Portas L, Benediktsdottir B, Bertelsen RJ, et al. A three-generation study on the association of tobacco smoking with asthma. Int J Epidemiol. 2018;47(4):1106-17.

297. Accordini S, Calciano L, Johannessen A, Benediktsdottir B, Bertelsen RJ, Braback L, et al. Prenatal and prepubertal exposures to tobacco smoke in men may cause lower lung function in future offspring: a three-generation study using a causal modelling approach. Eur Respir J. 2021;58(4).

298. Tezza G, Mazzei F, Boner A. Epigenetics of allergy. Early Hum Dev. 2013;89 Suppl 1:S20-1.

299. Joubert BR, Haberg SE, Bell DA, Nilsen RM, Vollset SE, Midttun O, et al. Maternal smoking and DNA methylation in newborns: in utero effect or epigenetic inheritance? Cancer Epidemiol Biomarkers Prev. 2014;23(6):1007-17.

300. Mirzakhani H, Carey VJ, Zeiger R, Bacharier LB, O'Connor GT, Schatz MX, et al. Impact of parental asthma, prenatal maternal asthma control, and vitamin D status on risk of asthma and recurrent wheeze in 3-year-old children. Clin Exp Allergy. 2019;49(4):419-29.

301. Liu X, Agerbo E, Schlunssen V, Wright RJ, Li J, Munk-Olsen T. Maternal asthma severity and control during pregnancy and risk of offspring asthma. J Allergy Clin Immunol. 2018;141(3):886-92 e3.

302. Barrett EG. Maternal influence in the transmission of asthma susceptibility. Pulm Pharmacol Ther. 2008;21(3):474-84.

303. Hagendorens MM, Ebo DG, Bridts CH, Van de Water L, De Clerck LS, Stevens WJ. Prenatal exposure to house dust mite allergen (Der p 1), cord blood T cell phenotype and cytokine production and atopic dermatitis during the first year of life. Pediatr Allergy Immunol. 2004;15(4):308-15.

304. Jones AC, Miles EA, Warner JO, Colwell BM, Bryant TN, Warner JA. Fetal peripheral blood mononuclear cell proliferative responses to mitogenic and allergenic stimuli during gestation. Pediatr Allergy Immunol. 1996;7(3):109-16.

305. Clifton VL, Giles WB, Smith R, Bisits AT, Hempenstall PA, Kessell CG, et al. Alterations of placental vascular function in asthmatic pregnancies. Am J Respir Crit Care Med. 2001;164(4):546-53.

306. Gunawardhana LP, Baines KJ, Mattes J, Murphy VE, Simpson JL, Gibson PG. Differential DNA methylation profiles of infants exposed to maternal asthma during pregnancy. Pediatr Pulmonol. 2014;49(9):852-62.

307. Magnaye K. Genetic and Epigenetic Mechanisms Underlying Maternal Transmission of Asthma Risk: The University of Chicago; 2020.

308. Benjamini Y, Hochberg Y. CONTROLLING THE FALSE DISCOVERY RATE - A PRACTICAL AND POWERFUL APPROACH TO MULTIPLE TESTING. Journal of the Royal Statistical Society Series B-Methodological. 1995;57(1):289-300.

309. Bakulski KM, Feinberg JI, Andrews SV, Yang J, Brown S, McKenney SL, et al. DNA methylation of cord blood cell types: Applications for mixed cell birth studies. Epigenetics. 2016;11(5):354-62.

310. Zhang H, Wang F, Kranzler HR, Yang C, Xu H, Wang Z, et al. Identification of methylation quantitative trait loci (mQTLs) influencing promoter DNA methylation of alcohol dependence risk genes. Hum Genet. 2014;133(9):1093-104.

311. Welch RP, Willer CJ, J. SL, Boehnke M. Snipper: a research tool for extracting and searching biological annotations on genes near SNPs: University of Michigan; 2013.

312. Karolchik D, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, et al. The UCSC Genome Browser Database. Nucleic Acids Res. 2003;31(1):51-4.

313. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Research. 2002;12(6):996-1006.

314. Snipper. 1.2 ed: Center for Statistical Genetics, University of Michigan.

315. Science, WIo. GeneCards®: The Human Gene Database [Available from: http://www.genecards.org/.

316. Pinero J, Queralt-Rosinach N, Bravo A, Deu-Pons J, Bauer-Mehren A, Baron M, et al. DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. Database (Oxford). 2015;2015:bav028.

317. Kramer A, Green J, Pollard J, Jr., Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics (Oxford, England). 2014;30(4):523-30.

318. Losol P, Rezwan FI, Patil VK, Venter C, Ewart S, Zhang H, et al. Effect of gestational oily fish intake on the risk of allergy in children may be influenced by FADS1/2, ELOVL5 expression and DNA methylation. Genes Nutr. 2019;14:20.

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

320. Ritchie ME, Silver J, Oshlack A, Holmes M, Diyagama D, Holloway A, et al. A comparison of background correction methods for two-colour microarrays. Bioinformatics. 2007;23(20):2700-7.

321. Li M, Zou D, Li Z, Gao R, Sang J, Zhang Y, et al. EWAS Atlas: a curated knowledgebase of epigenome-wide association studies. Nucleic Acids Res. 2019;47(D1):D983-D8.

322. Rojas D, Rager JE, Smeester L, Bailey KA, Drobna Z, Rubio-Andrade M, et al. Prenatal arsenic exposure and the epigenome: identifying sites of 5-methylcytosine

alterations that predict functional changes in gene expression in newborn cord blood and subsequent birth outcomes. Toxicol Sci. 2015;143(1):97-106.

323. de FCLAJ, van der Plaat DA, de Jong K, van Diemen CC, Postma DS, Nedeljkovic I, et al. Long-term Air Pollution Exposure, Genome-wide DNA Methylation and Lung Function in the LifeLines Cohort Study. Environ Health Perspect. 2018;126(2):027004.

324. Nicodemus-Johnson J, Myers RA, Sakabe NJ, Sobreira DR, Hogarth DK, Naureckas ET, et al. DNA methylation in lung cells is associated with asthma endotypes and genetic risk. JCI Insight. 2016;1(20):e90151.

325. Teschendorff AE, Yang Z, Wong A, Pipinikas CP, Jiao Y, Jones A, et al. Correlation of Smoking-Associated DNA Methylation Changes in Buccal Cells With DNA Methylation Changes in Epithelial Cancer. JAMA Oncol. 2015;1(4):476-85.

326. Wang CM, Chang CB, Lee SP, M WYC, Wu SF. Differential DNA methylation profiles of peripheral blood mononuclear cells in allergic asthmatic children following dust mite immunotherapy. J Microbiol Immunol Infect. 2020;53(6):986-95.

327. Langie SA, Szarc Vel Szic K, Declerck K, Traen S, Koppen G, Van Camp G, et al. Whole-Genome Saliva and Blood DNA Methylation Profiling in Individuals with a Respiratory Allergy. PLoS One. 2016;11(3):e0151109.

328. Chitrala KN, Hernandez DG, Nalls MA, Mode NA, Zonderman AB, Ezike N, et al. Race-specific alterations in DNA methylation among middle-aged African Americans and Whites with metabolic syndrome. Epigenetics. 2020;15(5):462-82.

329. Kadakia R, Zheng Y, Zhang Z, Zhang W, Josefson JL, Hou L. Association of cord blood methylation with neonatal leptin: An epigenome wide association study. PLoS One. 2019;14(12):e0226555.

330. Relton CL, Davey Smith G. Is epidemiology ready for epigenetics? Int J Epidemiol. 2012;41(1):5-9.

331. DeVries A, Vercelli D. Epigenetic Mechanisms in Asthma. Ann Am Thorac Soc. 2016;13 Suppl 1:S48-50.

332. Martino D, Prescott S. Epigenetics and prenatal influences on asthma and allergic airways disease. Chest. 2011;139(3):640-7.

333. Kabesch M. Epigenetics in asthma and allergy. Curr Opin Allergy Clin Immunol. 2014;14(1):62-8.

334. Lau MY, Dharmage SC, Burgess JA, Lowe AJ, Lodge CJ, Campbell B, et al. CD14 polymorphisms, microbial exposure and allergic diseases: a systematic review of gene-environment interactions. Allergy. 2014;69(11):1440-53.

335. Simpson A, John SL, Jury F, Niven R, Woodcock A, Ollier WE, et al. Endotoxin exposure, CD14, and allergic disease: an interaction between genes and the environment. Am J Respir Crit Care Med. 2006;174(4):386-92.

336. Lee JU, Kim JD, Park CS. Gene-Environment Interactions in Asthma: Genetic and Epigenetic Effects. Yonsei Med J. 2015;56(4):877-86.

337. Li JY, Zhang Y, Lin XP, Ruan Y, Wang Y, Wang CS, et al. Association between DNA hypomethylation at IL13 gene and allergic rhinitis in house dust mite-sensitized subjects. Clin Exp Allergy. 2016;46(2):298-307.

338. Guthikonda K, Zhang H, Nolan VG, Soto-Ramirez N, Ziyab AH, Ewart S, et al. Oral contraceptives modify the effect of GATA3 polymorphisms on the risk of asthma at the age of 18 years via DNA methylation. Clinical epigenetics. 2014;6(1):17.

339. Zhang H, Tong X, Holloway JW, Rezwan FI, Lockett GA, Patil V, et al. The interplay of DNA methylation over time with Th2 pathway genetic variants on asthma risk and temporal asthma transition. Clin Epigenetics. 2014;6(1):8.

340. Ziyab AH, Karmaus W, Holloway JW, Zhang H, Ewart S, Arshad SH. DNA methylation of the filaggrin gene adds to the risk of eczema associated with loss-of-function variants. J Eur Acad Dermatol Venereol. 2013;27(3):e420-3.

341. Mukherjee N, Lockett GA, Merid SK, Melen E, Pershagen G, Holloway JW, et al. DNA methylation and genetic polymorphisms of the Leptin gene interact to influence lung function outcomes and asthma at 18 years of age. Int J Mol Epidemiol Genet. 2016;7(1):1-17.

342. Ege MJ, Strachan DP, Cookson WO, Moffatt MF, Gut I, Lathrop M, et al. Gene-environment interaction for childhood asthma and exposure to farming in Central Europe. J Allergy Clin Immunol. 2011;127(1):138-44, 44 e1-4.

343. Czamara D, Eraslan G, Page CM, Lahti J, Lahti-Pulkkinen M, Hamalainen E, et al. Integrated analysis of environmental and genetic influences on cord blood DNA methylation in new-borns. Nat Commun. 2019;10(1):2548.

344. Accordini S, Calciano L, Johannessen A, Portas L, Benediktsdottir B, Bertelsen RJ, et al. A three-generation study on the association of tobacco smoking with asthma. International journal of epidemiology. 2018.

345. Magnus MC, Haberg SE, Karlstad O, Nafstad P, London SJ, Nystad W. Grandmother's smoking when pregnant with the mother and asthma in the grandchild: the Norwegian Mother and Child Cohort Study. Thorax. 2015;70(3):237-43.

346. Wilcox AJ. Birth weight and perinatal mortality: the effect of maternal smoking. Am J Epidemiol. 1993;137(10):1098-104.

347. Kurukulaaratchy RJ, Waterhouse L, Matthews SM, Arshad SH. Are influences during pregnancy associated with wheezing phenotypes during the first decade of life? Acta Paediatr. 2005;94(5):553-8.

348. Bjerg A, Hedman L, Perzanowski M, Lundback B, Ronmark E. A strong synergism of low birth weight and prenatal smoking on asthma in schoolchildren. Pediatrics. 2011;127(4):e905-12.

349. Felix JF, Joubert BR, Baccarelli AA, Sharp GC, Almqvist C, Annesi-Maesano I, et al. Cohort Profile: Pregnancy And Childhood Epigenetics (PACE) Consortium. Int J Epidemiol. 2018;47(1):22-3u.

350. Reese SE, Xu CJ, den Dekker HT, Lee MK, Sikdar S, Ruiz-Arenas C, et al. Epigenome-wide meta-analysis of DNA methylation and childhood asthma. J Allergy Clin Immunol. 2019;143(6):2062-74.

351. Daxinger L, Whitelaw E. Transgenerational epigenetic inheritance: more questions than answers. Genome Res. 2010;20(12):1623-8.

352. Grossniklaus U, Kelly WG, Kelly B, Ferguson-Smith AC, Pembrey M, Lindquist S. Transgenerational epigenetic inheritance: how important is it? Nat Rev Genet. 2013;14(3):228-35.

353. Jirtle RL, Skinner MK. Environmental epigenomics and disease susceptibility. Nature reviews. 2007;8(4):253-62.

354. Dolinoy DC, Weidman JR, Waterland RA, Jirtle RL. Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. Environ Health Perspect. 2006;114(4):567-72.

355. Kheirkhah Rahimabad P, Anthony TM, Jones AD, Eslamimehr S, Mukherjee N, Ewart S, et al. Nicotine and Its Downstream Metabolites in Maternal and Cord Sera: Biomarkers of Prenatal Smoking Exposure Associated with Offspring DNA Methylation. Int J Environ Res Public Health. 2020;17(24).

356. McRae AF, Powell JE, Henders AK, Bowdler L, Hemani G, Shah S, et al. Contribution of genetic variation to transgenerational inheritance of DNA methylation. Genome biology. 2014;15(5):R73.

357. Lawlor DA, Lewcock M, Rena-Jones L, Rollings C, Yip V, Smith D, et al. The second generation of The Avon Longitudinal Study of Parents and Children (ALSPAC-G2): a cohort profile. Wellcome Open Res. 2019;4:36.

358. Everson TM, Zhang H, Lockett GA, Kaushal A, Forthofer M, Ewart SL, et al. Epigenome-wide association study of asthma and wheeze characterizes loci within HK1. Allergy Asthma Clin Immunol. 2019;15:43.

359. Custovic A, Ainsworth J, Arshad H, Bishop C, Buchan I, Cullinan P, et al. The Study Team for Early Life Asthma Research (STELAR) consortium 'Asthma e-lab': team science bringing data, methods and investigators together. Thorax. 2015;70(8):799-801.

360. Heinrich J, Bruske I, Cramer C, Hoffmann U, Schnappinger M, Schaaf B, et al. GINIplus and LISAplus - Design and selected results of two German birth cohorts about natural course of atopic diseases and their determinants. Allergol Select. 2017;1(1):85-95.

361. Mayhew TM, Jenkins H, Todd B, Clifton VL. Maternal asthma and placental morphometry: effects of severity, treatment and fetal sex. Placenta. 2008;29(4):366-73.

362. Aaltonen R, Heikkinen T, Hakala K, Laine K, Alanen A. Transfer of proinflammatory cytokines across term placenta. Obstet Gynecol. 2005;106(4):802-7.

363. Zaretsky MV, Alexander JM, Byrd W, Bawdon RE. Transfer of inflammatory cytokines across the placenta. Obstet Gynecol. 2004;103(3):546-50.

364. Fu Y, Lou H, Wang C, Lou W, Wang Y, Zheng T, et al. T cell subsets in cord blood are influenced by maternal allergy and associated with atopic dermatitis. Pediatr Allergy Immunol. 2013;24(2):178-86.

365. Romanowska J, Haaland OA, Jugessur A, Gjerdevik M, Xu Z, Taylor J, et al. Gene-methylation interactions: discovering region-wise DNA methylation levels that modify SNP-associated disease risk. Clin Epigenetics. 2020;12(1):109.

366. Weiland SK, Bjorksten B, Brunekreef B, Cookson WO, von Mutius E, Strachan DP, et al. Phase II of the International Study of Asthma and Allergies in Childhood (ISAAC II): rationale and methods. Eur Respir J. 2004;24(3):406-12.

367. Rebagliato M, Bolumar F, Florey Cdu V, Jarvis MJ, Perez-Hoyos S, Hernandez-Aguado I, et al. Variations in cotinine levels in smokers during and after pregnancy. Am J Obstet Gynecol. 1998;178(3):568-71.

368. Castro-Rodriguez JA, Holberg CJ, Wright AL, Martinez FD. A clinical index to define risk of asthma in young children with recurrent wheezing. Am J Respir Crit Care Med. 2000;162(4 Pt 1):1403-6.

369. Martinez JA. Not all that wheezes is asthma! J Bras Pneumol. 2013;39(4):518-20.

370. Cao-Lei L, Massart R, Suderman MJ, Machnes Z, Elgbeili G, Laplante DP, et al. DNA methylation signatures triggered by prenatal maternal stress exposure to a natural disaster: Project Ice Storm. PLoS One. 2014;9(9):e107653.

371. Liang L, Willis-Owen SAG, Laprise C, Wong KCC, Davies GA, Hudson TJ, et al. An epigenome-wide association study of total serum immunoglobulin E concentration. Nature. 2015;520(7549):670-4.

372. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13:86.

373. Wiklund P, Karhunen V, Richmond RC, Parmar P, Rodriguez A, De Silva M, et al. DNA methylation links prenatal smoking exposure to later life health outcomes in offspring. Clin Epigenetics. 2019;11(1):97.

374. Relton CL, Davey Smith G. Mendelian randomization: applications and limitations in epigenetic studies. Epigenomics. 2015;7(8):1239-43.

375. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet. 2014;23(R1):R89-98.

376. Caramaschi D, Sharp GC, Nohr EA, Berryman K, Lewis SJ, Davey Smith G, et al. Exploring a causal role of DNA methylation in the relationship between maternal vitamin B12 during pregnancy and child's IQ at age 8, cognitive performance and educational attainment: a two-step Mendelian randomization study. Hum Mol Genet. 2017;26(15):3001-13.

377. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27(8):1133-63.

378. Carmichael M. Newborn screening: a spot of trouble. Nature. 2011;475(7355):156-8.

379. Ramagopalan SV, Rakyan VK. The promise and challenges of blood spot methylomics. Epigenetics. 2013;8(8):775-7.

380. Mukherjee N, Arathimos R, Chen S, Kheirkhah Rahimabad P, Han L, Zhang H, et al. DNA methylation at birth is associated with lung function development until age 26 years. Eur Respir J. 2021;57(4).

381. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74.

382. Bujold D, Morais DAL, Gauthier C, Cote C, Caron M, Kwan T, et al. The International Human Epigenome Consortium Data Portal. Cell Syst. 2016;3(5):496-9 e2.

383. Ong ML, Lin X, Holbrook JD. Measuring epigenetics as the mediator of gene/environment interactions in DOHaD. J Dev Orig Health Dis. 2015;6(1):10-6.

384. Cheng Y, He C, Wang M, Ma X, Mo F, Yang S, et al. Targeting epigenetic regulators for cancer therapy: mechanisms and advances in clinical trials. Signal Transduct Target Ther. 2019;4:62.

385. de Planell-Saguer M, Lovinsky-Desir S, Miller RL. Epigenetic regulation: the interface between prenatal and early-life exposure and asthma susceptibility. Environ Mol Mutagen. 2014;55(3):231-43.

Appendices

Appendix 1 Supplementary Data

Table-A1 Top 100 differentially methylated CpG sites from EWAS for maternal asthma used for pathway analysis in chapter 6.

ProbeID	CHR	Mapinfo	UCSC RefGene Name	BETA	P value	Adjusted p
cg02505293	2	2.2E+08	IHH	5.33E-03	1.34E-09	0.000535
cg16228090	1	2.26E+08	ENAH	3.20E-02	1.65E-07	0.030841
cg10146326	6	32188822	NOTCH4	1.65E-02	3.39E-07	0.030841
cg11441533	10	1.35E+08	CFAP46	1.03E-02	3.71E-07	0.030841
cg12482231	19	44763334	ZNF233	2.85E-02	3.86E-07	0.030841
cg19832529	1	2.38E+08	RYR2	-1.50E-02	4.77E-07	0.031777
cg27419217	6	1.02E+08	GRIK2	-9.08E-03	6.97E-07	0.039764
cg20361600	1	29063076	YTHDF2	-4.99E-03	8.01E-07	0.040005
cg11825603	8	1.02E+08	YWHAZ	1.70E-02	9.71E-07	0.043084
cg03179291	17	81021751	B3GNTL1	1.66E-02	1.81E-06	0.069334
cg13800769	9	90112515	DAPK1	3.61E-02	1.91E-06	0.069334
cg08304120	6	39281885	KCNK17	5.25E-03	2.45E-06	0.081607
cg20043649	19	59031425	ZBTB45	-3.20E-02	3.15E-06	0.096633
cg03179419	10	661319	DIP2C	-4.35E-02	3.93E-06	0.110764
cg12264626	15	59279556	RNF111	2.00E-02	4.39E-06	0.110764
cg01826959	16	87871160	SLC7A5	-5.32E-03	4.44E-06	0.110764
cg05593669	11	250989	PSMD13	-1.07E-02	6.18E-06	0.141823
cg03752965	19	4302846	TMIGD2	-1.90E-02	6.39E-06	0.141823
cg17160391	15	1.02E+08	TM2D3	1.13E-02	6.95E-06	0.146027
cg02336802	21	33984905	C21orf59	-1.10E-02	7.75E-06	0.154755
cg18222853	7	75399336	CCL26	-1.42E-02	9.36E-06	0.175935

ProbeID	CHR	Mapinfo	UCSC RefGene Name	BETA	P value	Adjusted p
cg01785514	1	2.45E+08	KIF26B	3.39E-02	9.69E-06	0.175935
cg04560534	12	54779175	ZNF385A	-1.27E-02	1.08E-05	0.186253
cg12606485	5	88185768	MEF2C	8.20E-03	1.12E-05	0.186253
cg21590729	1	1150085	TNFRSF4	1.06E-02	1.25E-05	0.188415
cg09020697	16	89436838	ANKRD11	1.41E-02	1.26E-05	0.188415
cg24156983	11	1321530	TOLLIP	1.51E-02	1.33E-05	0.188415
cg04437407	3	39509081	MOBP	1.94E-02	1.45E-05	0.188415
cg15046927	3	1.9E+08	CLDN16	-9.40E-03	1.49E-05	0.188415
cg18730318	13	45995831	SLC25A30	-9.52E-03	1.51E-05	0.188415
cg15810996	6	1.37E+08	PDE7B	-1.18E-02	1.54E-05	0.188415
cg20201954	3	1.02E+08	ZPLD1	-5.39E-03	1.57E-05	0.188415
cg03250842	13	1.13E+08	SOX1	-5.98E-03	1.63E-05	0.188415
cg10639029	12	66524411	LLPH	-1.86E-02	1.65E-05	0.188415
cg08204939	6	1.67E+08	T	-1.05E-02	1.68E-05	0.188415
cg13128305	15	80853140	ARNT2	-4.09E-03	1.70E-05	0.188415
cg02376282	11	1.3E+08	ST14	1.85E-02	1.87E-05	0.193181
cg19918821	17	80884963	TBCD	1.17E-02	1.88E-05	0.193181
cg13685951	9	97841956	C9orf3	-6.62E-03	1.89E-05	0.193181
cg16666115	19	48991098	LMTK3	4.90E-03	1.94E-05	0.193804
cg13977597	12	1.28E+08		2.19E-02	2.05E-05	0.196103
cg07173760	11	1.03E+08	DCUN1D5	2.39E-02	2.08E-05	0.196103
cg18469292	17	48845722	C17orf73	-1.29E-02	2.11E-05	0.196103
cg03314473	16	89355071	ANKRD11	1.10E-02	2.22E-05	0.201469
cg25421566	10	13933021	FRMD4A	2.18E-02	2.37E-05	0.205693

ProbeID	CHR	Mapinfo	UCSC RefGene Name	BETA	P value	Adjusted p
cg04234063	7	1.59E+08	WDR60	3.31E-02	2.47E-05	0.209643
cg11853320	17	5322940	RPAIN	-1.22E-02	2.55E-05	0.212455
cg15766075	6	44234175	NFKBIE	-7.59E-03	2.61E-05	0.212455
cg04168137	5	1.52E+08		1.64E-02	2.66E-05	0.212715
cg09784168	14	1.05E+08	KIAA0284	2.06E-02	2.75E-05	0.213688
cg08463681	1	2.36E+08	GPR137B	4.24E-03	2.79E-05	0.213688
cg07424889	7	30170466	MTURN	-7.20E-03	2.85E-05	0.213688
cg09369792	14	1.04E+08	C14orf73	-1.79E-02	2.89E-05	0.213688
cg00713401	2	1.36E+08	RAB3GAP1	-1.02E-02	2.97E-05	0.215427
cg21616047	11	62358814	TUT1	2.91E-02	3.06E-05	0.218435
cg13929433	17	34245111	RDM1	1.14E-02	3.14E-05	0.2198
cg10453390	6	1.68E+08	CCR6	-8.64E-03	3.19E-05	0.2198
cg18681028	3	87729320		7.26E-03	3.30E-05	0.223439
cg24580071	11	18610297	UEVLD	3.05E-02	3.46E-05	0.229684
cg02471132	19	58570454	ZNF135	-1.24E-02	3.51E-05	0.229684
cg19190900	3	48617067	MIR711	-1.73E-02	3.57E-05	0.229684
cg02325128	12	82749915	CCDC59	-8.88E-03	3.75E-05	0.237624
cg11434468	11	1.18E+08	ATP5L	4.88E-03	3.92E-05	0.239872
cg20371378	18	48346350	MRO	-2.98E-02	4.00E-05	0.239872
cg11241498	1	42288825	HIVEP3	1.17E-02	4.10E-05	0.239872
cg01427435	19	33167718	RGS9BP	4.88E-03	4.15E-05	0.239872
cg06541760	4	1.66E+08	CPE	-1.10E-02	4.18E-05	0.239872
cg10451200	16	85262708	FAM92B	1.30E-02	4.20E-05	0.239872
cg21621645	2	18300908	KCNS3	-1.19E-02	4.23E-05	0.239872

ProbeID	CHR	Mapinfo	UCSC RefGene Name	BETA	P value	Adjusted p
cg15550572	1	24528833	LOC284632	1.20E-02	4.57E-05	0.239872
cg21322819	2	2.39E+08	ASB1	-8.44E-03	4.60E-05	0.239872
cg00949446	1	2.41E+08	GREM2	-9.88E-03	4.66E-05	0.239872
cg16685608	3	1.84E+08	AP2M1	2.49E-02	4.70E-05	0.239872
cg06361667	19	49503178	RUVBL2	-1.01E-02	4.86E-05	0.239872
cg19634247	10	1.03E+08	BTRC	7.15E-03	4.90E-05	0.239872
cg00837860	15	35049448	GJD2	-1.09E-02	4.95E-05	0.239872
cg20544562	14	50100181	C14orf104	-1.58E-02	5.00E-05	0.239872
cg26705408	16	83195465	CDH13	2.46E-02	5.02E-05	0.239872
cg12599641	7	1535941	INTS1	1.98E-02	5.07E-05	0.239872
cg17515462	12	56325730	DGKA	-9.64E-03	5.11E-05	0.239872
cg04877966	5	61699957	DIMT1L	8.02E-03	5.11E-05	0.239872
cg10498682	20	31407542	MAPRE1	9.09E-03	5.19E-05	0.239872
cg11067412	17	77769065	CBX8	-7.85E-03	5.23E-05	0.239872
cg11829416	14	21737252	HNRNPC	1.66E-02	5.23E-05	0.239872
cg03958208	1	1.11E+08	SLC16A4	-3.60E-03	5.31E-05	0.239872
cg11410404	5	31777552	PDZD2	2.75E-02	5.45E-05	0.239872
cg20485084	3	1.6E+08	IL12A	-1.84E-02	5.46E-05	0.239872
cg07034329	4	207357	ZNF876P	1.91E-02	5.46E-05	0.239872
cg03714619	6	10415290	TFAP2A	-6.39E-03	5.58E-05	0.239872
cg05529874	6	1.17E+08	FAM26D	2.97E-02	5.61E-05	0.239872
cg07841848	10	1.2E+08	CASC2	-1.50E-02	5.68E-05	0.239872
cg14509623	20	21283801	XRN2	6.03E-03	5.72E-05	0.239872
cg07726423	5	52319150	ITGA2	1.14E-02	5.73E-05	0.239872

ProbeID	CHR	Mapinfo	UCSC RefGene Name	BETA	P value	Adjusted p
cg19682367	14	1E+08	CYP46A1	1.09E-02	5.83E-05	0.239872
cg16792014	21	31813075	KRTAP15-1	-2.60E-02	6.13E-05	0.248868
cg10867327	2	72372819	CYP26B1	5.32E-03	6.18E-05	0.248868
cg11628055	8	29210373	DUSP4	1.59E-02	6.23E-05	0.248868