Role of epigenetic mechanisms in childhood asthma development

Dissertation
zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät III
Agrar- und Ernährungswissenschaften,
Geowissenschaften und Informatik

der Martin-Luther-Universität Halle-Wittenberg

vorgelegt von

Herrn M. Sc. Matthias Klös
Geb. am 08.11.1987 in Koblenz

Gutachter: Prof. Dr. Gabriele I. Stangl

Prof. Dr. Irina Lehmann

Prof. Dr. Anke Tönjes

Verteidigung am 22.11.2021

TABLE OF CONTENT

I. List of figures

II. List of tables

III. Abbreviation

ACOT7	*Acyl-CoA Thioesterase 7*
ADRB2	*Adrenoceptor Beta 2*
ALOX12	*Arachidonate 12-Lipoxygenase, 12S Type*
ALOX15	*Arachidonate 15-Lipoxygenase, 15S Type*
APC	Antigen presenting cells
AXL	*AXL Receptor Tyrosine Kinase*
bp	Base pairs
CAMP	Childhood asthma management program
ChiA-PET seq.	Chromatin Interaction Analysis by Paired-End Tag Sequencing
ChIP-Seq	Chromatin immunoprecipitation sequencing
Chr	Chromosome
CoA	Coenzyme A
CpG	Cytosine-phosphat-Guanine
DHSs	DNase I hypersensitive site
DMC	Differentially methylated cytosine
DMR	Differentially methylated region
DNA	Deoxyribonucleic acid
DNMT	*DNA methyltransferase*
dNTPs	Deoxynukleosidtriphosphate
EPX	*Eosinophil Peroxidase*
eQTL	Expression quantitative trait locus
et al.	*Et alii*
EVL	*Enah/Vasp-Like*
EWAS	Epigenome wide association study

FcεR1	*Fc epsilon receptor*
FLG	*Fillagrin*
GATA3	*GATA Binding Protein 3*
gDMR	Genotype associated differentially methylated region
GEE	Generalized estimating equation
GWAS	Genome wide association study
HAT	*Histoneacetylase*
HDAC	*Histone deacetylase*
IFNGR2	*Interferon Gamma Receptor 2*
IgE	Immunoglobulin E
IL13	*Interleukin 13*
IL25	*Interleukin 25*
IL33	*Interleukin 33*
IL4	*Interleukin 4*
IL5	*Interleukin 5*
IL5RA	*Interleukin 5 receptor alpha*
IRF1	*Interferon Regulatory Factor 1*
KIF3A	*Kinesin Family Member 3A*
kU/l	Kilo Unit per liter
LINA cohort	Lifestyle and Environmental Factors and their Influence on Newborns Allergy risk cohort
LINE	Long interspersed nuclear elements
LISA cohort	Influences of Lifestyle- related factors on the Immune System and the Development of Allergies in Childhood cohort
µM	Micromol
MALDI-TOF	Matrix-assisted laser desorption/ionization- time of flight
meQTL	Methylation quantitative trait locus

microRNA	Micro ribonucleic acid
mRNA	Messenger Ribonucleic acid
NFATC1	*Nuclear Factor Of Activated T Cells 1*
ngDMR	Non genotype associated differentially methylated region
No.	Number
ORMDL	*ORMDL Sphingolipid Biosynthesis Regulator*
PASTURE cohort	Protection against Allergy--Study in Rural Environments cohort
pg/ml	Picogram per milliliter
PPARGC1B	*Peroxisome proliferator activated receptor gamma Coactivator 1 Beta*
qPCR	Real time quantitative polymerase chain reaction
RAD50	*RAD50 Double Strand Break Repair Protein*
RELA	*RELA Proto-Oncogene, NF-KB Subunit*
ROI	Regions of interest
RPTOR	*Regulatory Associated Protein Of MTOR Complex 1*
RUNX3	*RUNX family transcription factor 3*
SMAD3	*SMAD Family Member 3*
SNP	Single nucleotide polymorphism
SPINK5	*Serine Peptidase Inhibitor Kazal Type 5*
SPT	Skin prick test
Th1 cells	T helper cell typ 1
Th17 cells	T helper cell typ 17
Th2 cells	T helper cell typ 2
TNFSF10	*TNF Superfamily Member 10*
TRAF1/3	*TNF Receptor Associated Factor 1/3*
TSLP	*Thymic Stromal Lymphopoietin*

TSS	Transcription start site
TTS	Transcription termination site
YWHAZ	*Tyrosine 3-Monooxygenase/Tryptophan 5-Monooxygenase Activation Protein Zeta*
YY1	*YY1 Transcription Factor*
ZFPM1	*Zinc finger protein, FOG family member 1*

1 INTRODUCTION

1.1 Asthma aetiology

The global asthma report, published in 2018, classifies asthma as one of the most common non-communicable chronic inflammatory diseases worldwide with more than 339 million people being affected (1). Although decades ago, asthma was known as highly prevalent in industrialized western countries, it is now on the rise in developing and emerging countries (2). In Germany, the Robert Koch Institute estimated the general lifetime prevalence of asthma as about 6.3 % for children between 0-17 years. The observed difference in prevalence between boys (7.4%) and girls (5.2%) before adulthood (3) reverses with increasing age (4). Asthma, which is not curable so far, is a worldwide health burden with enormous costs for the health care system (2). Accordingly, efforts are being made to understand the underlying mechanisms of asthma development and to establish reliable instruments for prevention, early diagnosis and therapy.

Generally, asthma is characterized by a chronic inflammation which affects the respiratory tract leading to elevated expression of inflammatory proteins such as cytokines, chemokines, adhesion molecules, inflammatory enzymes and receptors (5). This chronic inflammation of the airways is accompanied by an activation and infiltration of inflammatory immune cells, such as mast cells, dendritic cells, lymphocytes and eosinophils into the lung epithelium promoting bronchoconstriction, vasodilatation, airway oedema and the activation of sensory nerve endings. Clinically common asthma symptoms are shortness of breath, wheezing, tightness in the chest and coughing (6). These symptoms are caused by the difficulty to exhale air out of the lungs due to bronchoconstriction and associated airway narrowing accompanied with airway wall thickening and an increased amount of mucus (7). An acute asthma attack is triggered by various conditions including allergens, air pollution, viral infections, exposure to tobacco smoke, physical exercise and stress (8).

1.2 Pathomechanisms of childhood asthma

1.2.1 Asthma phenotypes

Asthma is categorized in two basic phenotypes: the allergic (atopic) or extrinsic asthma, and the non-atopic or intrinsic asthma, respectively (9). In early childhood and adolescence allergic asthma is more prevalent, whereas non atopic asthma is common in adult patients. Over the last decade asthma phenotypes were further classified into different endotypes based on their distinct mechanistic pathways. The type 2 high endotype is primarily characterized by an

increased number of T-helper 2 (Th2) cells. Other cell types are highly prevalent in this endotype like eosinophils, mast cells and basophils. This increase in immune cells goes along with elevated levels of inflammatory cytokines (interleukin 4 (IL4); interleukin 5 (IL5) and interleukin 13 (IL13)) and activated proteins, so called alarmins (Thymic Stromal Lymphopoietin (TSLP), interleukin 25 (IL25), interleukin 33 (IL33)) (9-11). One of the most important factors in immune response in atopic asthma is elevated Immunoglobulin E (IgE). IgE initiates the response to allergens on dendritic cells, driving the T-helper 2 cell inflammatory process (12, 13). Elevated IgE concentration in the blood of asthmatic children suggests IgE as an important marker for inflammation and immune response to allergens or infections. Elevated IgE levels and Th2 driven inflammation of the lungs were observed in the type 2 high group with an increased activation and infiltration of eosinophils in the airways. Related phenotypes are early onset allergic asthma and late onset eosinophilic asthma (9). The type 2 low endotype is more closely related to an increase in neutrophils and an activation of T-helper 1 (Th1) and or T-helper 17 (Th17) cells. This endotype shows an association to obesity, smoking and is more common among very late onset asthma (9).

1.2.2 Inflammatory process of asthma induced by allergens

In early childhood allergic asthma is very common and often the inflammatory process of the airways is initiated by allergens (e.g. pollens). A schematic overview of the inflammatory response in allergic asthma is shown in Figure 1.1. Antigen presenting cells (APC) (e.g. dendritic cells) at the epithelial surface of the airways perceive inhaled allergens and induce a signal to initiate differentiation of naïve T cells into Th2 cells (14). The Th2 cells secrete cytokines like IL-5, which subsequently leads to an activation of eosinophils by binding to the IL5 receptor (IL5RA). Furthermore the activation of Th2 cells can induce immunoglobulin class switching of B cells promoting IgE production. IgE activates mast cells by binding to the IgE-specific Fc epsilon receptor (FcεR1) (15). Activated mast cells then release histamines and leukotrienes, which induce bronchoconstriction and activate goblet cells to produce mucus (16).

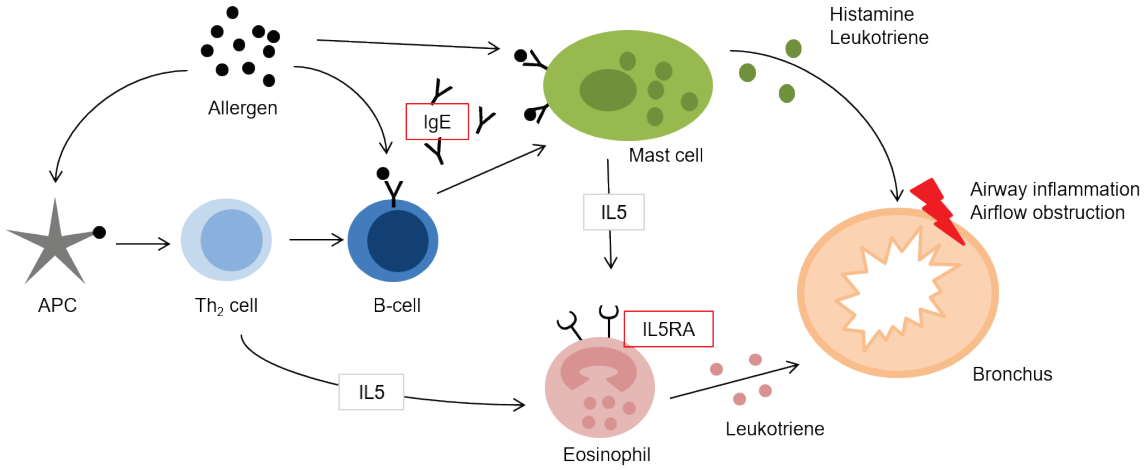


Figure 1.1: Schematic overview of the inflammatory process of asthma induced by allergens (adapted from KEGG asthma pathway (17)). Allergens presented by allergen presenting cells (APC) induce Th2 cells and B-cells to release cytokines (e.g. IL5) and IgE leading to an activation of mast cells and eosinophils by binding to receptors (FcεR1 and IL5RA). Activated mast cells and eosinophils secrete histamines and leukotrienes to induce inflammation and obstruction of the airways.

In conjunction with the inflammation of the airways, structural changes of the lung epithelium are observed in asthmatic patients. These changes known as airway remodelling include subepithelial fibrosis, increased smooth muscle mass, epithelial alteration and goblet and mucous gland hyperplasia (18, 19). The sub-epithelial fibrosis and smooth muscle hyperplasia have both been associated with severe asthma (18). An increase of smooth muscle mass contributes to a rise in proinflammatory cytokines and promotes airflow constriction (19), which can be further exacerbated by an increased mucus production and airway thickening due to goblet and mucous gland hyperplasia (18). Airway remodelling can impair barrier functions of the epithelium and thus increase vulnerability to e.g. virus infections (14). A variety of factors can promote the asthma-related activation of immune response. The following sections provide brief overview of commonly discussed risk factors for childhood asthma.

1.3 Risk factors of asthma development

1.3.1 Genetic predisposition

Although the symptoms and manifestation of asthma are commonly recognized, the underlying pathomechanisms are still not fully understand. The genetic background is one important risk factor of asthma (20). During the last two decades, a variety of studies was conducted in order to identify genetic loci which may explain asthma heritability. Initially, candidate gene studies were applied to identify genetic variants associated with asthma or with an inflammation process by targeted analyses (21, 22). Genetic variants such as single nucleotide

polymorphism (SNP), which occur in a frequency of approximately one in 1000bp in the genome, have impact on promoter activity and therefore influence transcription and translation of genes (23). Numerous loci or risk variants were identified for asthma including loci of immune system-related genes like interleukin 4 (IL4) or interleukin 13 (IL13) or those involved in lung function (e.g. adrenoceptor beta 2 (ADRB2) and filaggrin (FLG)) (20).
A more comprehensive approach to identify risk alleles, loci and genes related to asthma are genome-wide association studies (GWAS) which are unbiased and the current gold standard discovering new disease associated loci (24). Research on adult twins revealed that genetic factors may determine up to 24% of the phenotype variability in asthma severity (25). A recent review by Ober *et al*. estimated heritability of asthma between 35-95% (20). However, these estimates should be considered with caution as they are most likely inflated. Large GWAS studies pool individuals with different environmental backgrounds and phenotypes, for which heritability might be very different (20, 26, 27). In particular, heritability estimates from twin studies might lead to inflation as twins not only share nearly the same genetic background but also a similar living environment (21).

In 2007, Moffatt *et al.* (28) published the first asthma GWAS. In total, more than 317.000 single nucleotide polymorphisms were analyzed in 994 children with asthma and 1243 controls with European ethnicity. Several loci on chromosome 17q21 were significantly associated with childhood asthma and could be replicated in an independent cohort. Another major finding was an expression quantitative trait locus (eQTL) for *rs7216389* located between the *ORMDL Sphingolipid Biosynthesis Regulator (ORDML)* and *gasdermin-like (GSMDL)* region affecting the *ORMDL* gene expression in asthmatic children. An eQTL is a genetic variant, which regulate gene expression (29). The *ORMDL* eQTL is a commonly investigated and validated asthma locus (30) and *ORMDL* was indeed shown to be the target gene involved in airway inflammation and remodeling (31). A variety of different genomic loci have been associated with an increased asthma risk so far including loci inter alia 17q21 (ORMDL Sphingolipid Biosynthesis Regulator 3 (ORMDL3)), 5q31 *(RAD50 Double Strand Break Repair Protein (RAD50)) 9q24 (interleukin 33 (IL33))* 15q22 *(SMAD Family Member 3 (SMAD3))* (28, 32, 33).

However, it is of note that so far, only a small variance of asthma heritability could be explained with genomic variants identified by GWAS (20, 34). Loci identified are all common variants, while rare variants with larger effect size have not been found (30). In addition, most common variants were found in studies including cohorts of European ancestry. Furthermore, GWAS did not consider gene-gene interactions or gene-environment interactions (21). In addition, a clear diagnosis of complex asthma disease is often difficult due to poor clinical data (34).

1.3.2 Environmental factors

Obviously the impact of genetic predisposition does not fully explain causes and consequences of asthma. However, a substantial part of the asthma risk seems to be independent of genetic predisposition but rather attributable to environmental factors. Apparently, the environment has an impact on the perturbation of genetic patterns. Recently, researchers discovered a variety of environmental risk factors which act already during prenatal development during a crucial phase of early development, and thus might have an impact on asthma development later in life (35) including maternal stress (36-38) or maternal tobacco smoke exposure during pregnancy (39-41). In addition, further risk factors in early childhood are the socio-economic status (42, 43), nutrition (44), obesity (45), lower respiratory infections including virus infections (46), the exposition to air pollution (47) and allergens (48, 49). As the genetic predisposition only explains part of the risk for asthma development, environmental factors are likely to add to inherited risk. In recent years epigenetic modifications came into focus of research as likely mediators between the genetic background and environment-related factors (50).

1.4 Epigenetic modifications

The term epigenetics describes the study of heritable modifications which affect gene expression without changing the underlying DNA sequence (51). The different levels of epigenetic regulation are comprised of DNA methylation, histone modifications and microRNAs (52). These epigenetic modifications are responsible for the alteration of DNA accessibility for transcription factors and chromatin structure, which further results in different gene activity (53). An overview of the epigenetic modifications is shown in Figure 1.2.

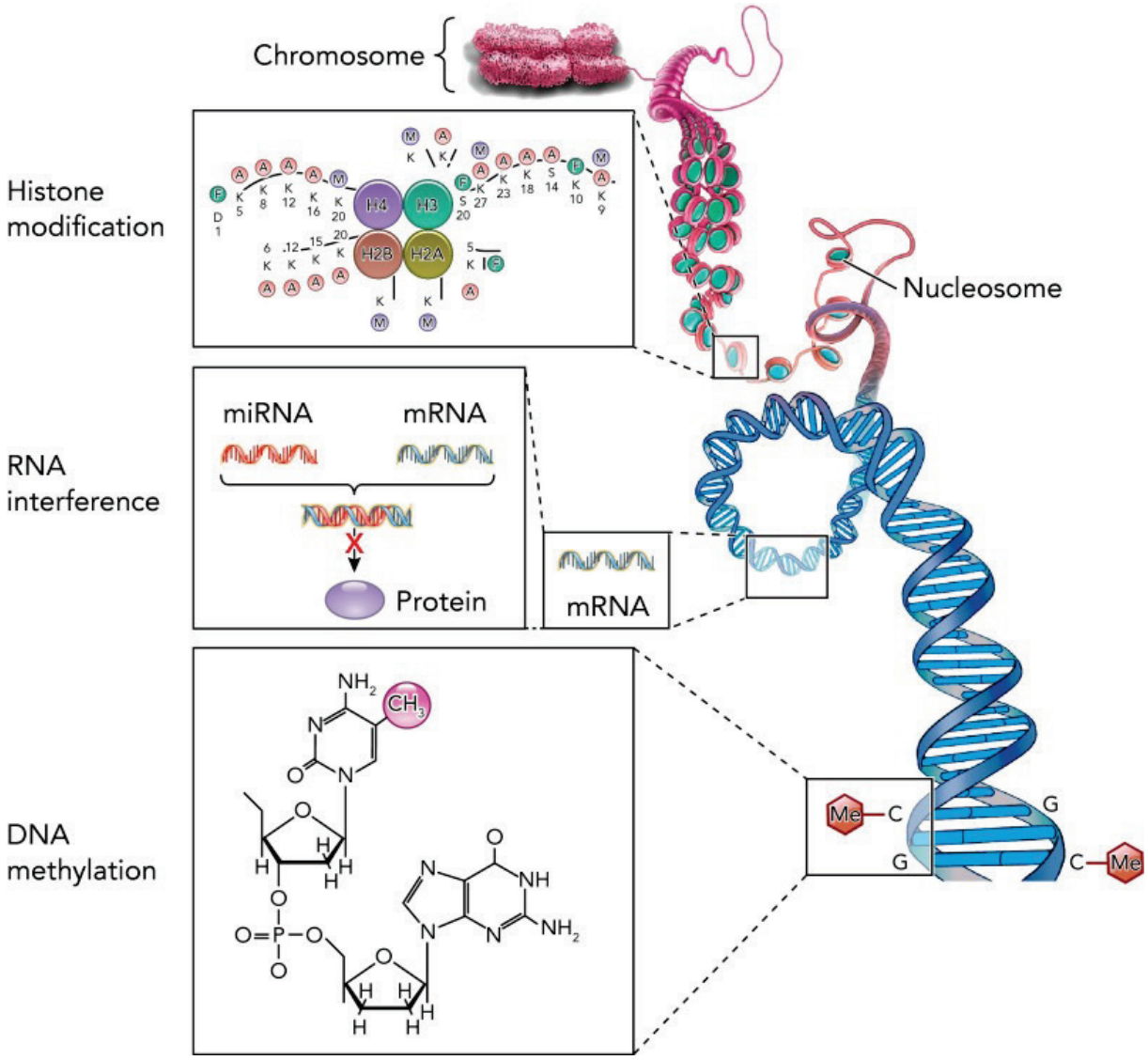


Figure 1.2: Epigenetic mechanism including histone modification, RNA interference and DNA methylation (54)

1.4.1 DNA methylation

The best studied epigenetic modification and stable long term epigenetic modification is DNA methylation (55, 56). In humans, DNA methylation mainly affects cytosines in the CpG (5'-cytosine-phosphat group-guanine-3') dinucleotide context (57). Different DNA methyltransferases (DNMT) can mediate the covalent binding of a methyl group (-CH3) to the cytosine, resulting in 5-methylcytosine (53). CpG sites are not randomly distributed across the human genome. They are condensed in CpG-rich regions, which are known as CpG-islands (58). These regions are usually unmethylated. Exceptions are subsets of promoter islands that are nearly fully methylated and are associated with specific germline genes, imprinted genes (e.g. only expressed from either maternal or paternal copy) or genes that undergo X-chromosome inactivation in females (58). However, DNA methylation can also occur at

transcription start sites (TSS) within or without CpG islands, in gene bodies or repetitive sequences as well as at regulatory elements in particular enhancer regions (59). DNA methylation at CpG islands is associated with gene silencing (60) by blocking DNA recognition and binding of transcription factors (53). The chromatin structure of methylated DNA is more compact. This can lead to a reduced DNA-binding capacity of transcription factors and the transcription machinery in general can be reduced (61). As a consequence gene expression might be repressed or completely silenced (59).

1.4.2 Histone modification and micro RNAs

The second of the three major epigenetic modification is known as histone modification (62). Histone modifications are flexible short term post-translational modifications (53) and have an impact on the DNA accessibility for gene transcription (62). The basic structure of chromatin is the nucleosome consisting of 145-147bp and the DNA sequence is wrapped around (53). The nucleosome consisting of a protein octamer formed by four dimers of two histone 2A and 2B (H2A-H2B) and two histone 3 and histone 4 (H3-H4) (63). These proteins can be modified via phosphorylation, acetylation, methylation, ubiquitination, SUMOylation and ADP-ribosylation (53).

The third epigenetic mechanism which potentially affects gene expression are microRNAs. These small non-coding single stranded RNAs (ssRNAs) occur post-transcriptional and interfere with transcribed mRNA (64) and thus can inhibit translation (65).

1.4.3 Epigenetic modifications in genomic regulatory elements and gene expression

Regulatory elements such as enhancers and promoters are important components of the genome, which interact with transcription factors to regulate gene expression (66). These regions are epigenetically characterized by histone modifications and DNA methylation (66). For example, an enhancer is associated with histone modifications such as mono- and dimethylation of histone H3 lysine 4 (H3K4me1/2) as well as histone H3 lysine 27 acetylation (H3K27Ac) (67). Other histone modification such as the methylation of histone H3 lysine 9 and histone H3 lysine 27 lead to a more condensed chromatin and thus to transcriptional repression (66). High DNA methylation in promoter regions is also associated with a decreased gene expression, as transcription factors are inhibit to bind at the DNA (68). Although, DNA methylation in gene bodies is positively correlated with gene expression levels (59, 69). Therefore, epigenetic modification especially in genomic regulatory elements and their resulting consequences for gene transcription might play a crucial role for metabolic downstream processes and finally disease development.

1.4.4 Epigenetic modifications related to asthma disease

In the last years, research on such epigenetic modifications related to asthma disease evolved. Although, compared to DNA methylation, histone modifications and miRNA are less explored in the context of asthma, there is rising evidence that e.g. altered histone modifications may have an impact on asthma development and related traits (70-73) and miRNA expression profiles seem to be associated with asthma disease progression (74). As mentioned before, DNA methylation is the most stable long term epigenetic modification and therefore very important for disease development. So far, there are two main approaches to investigate DNA methylation pattern of asthmatic and healthy patients, namely candidate gene studies and epigenome wide association studies (EWAS). These epigenome wide studies are commonly based on arrays (e.g. Illumina 27K CpGs and 450K CpGs Array) and use different biological material such as whole blood or nasal epithelial cells (75-77). DNA methylation measurement (single CpGs) using 27K or 450K arrays is based on fluorescence detection of bisulfite-converted DNA captured by beads (78). These strategy already revealed several genes associated with childhood asthma and responsible immune mechanisms (79) as shown by Morales and colleagues (80), who identified DNA hypomethylation at one CpG site of *arachidonate 12-lipoxygenase (ALOX12*) gene, which is associated with a higher risk of persistent wheeze compared to children who have never wheezed. *ALOX12* is involved in the metabolism of arachidonic acid, resulting to the formation of the inflammatory eicosanoids (80). There are several other studies that have discovered DNA methylation differences in asthmatics compared to healthy controls (76, 81-83), such as DNA hypomethylation of *IL13* (involved in B-cell maturation), *RUNX3* (a transcription factor involved in T-cell differentiation) (83) and *T-cell specific genes (TIGITs*) which are associated with higher serum IgE levels and the amount of predicted percent forced expiratory pressure in one second (FEV1), which is a marker for lung function (84).

Another 450K array based study (85) observed 36 differentially methylated CpGs associated with IgE serum levels in peripheral blood leukocytes. In addition, hypomethylated CpG sites of *IL5RA* and *IL4* were identified in asthmatic patients with high serum IgE levels compared to asthmatics with low IgE levels and healthy controls in isolated eosinophils (85). Two studies of Gao *et al*. and Barton *et al*. tried to determine DNA methylation pattern at the time of birth and its impact on risk stratification for development of wheezing and asthma in early childhood. They identified DNA hypermethylation of a CpG in *AXL Receptor Tyrosine Kinase (AXL)* gene at time of birth in combination with a higher risk of wheezing and a CpG in *GATA Binding Protein 3 (GATA3)* gene in association with a lower risk of asthma in early childhood (86, 87).

Unlike most studies that investigated altered DNA methylation in whole blood, there are only a few studies that focused on DNA methylation changes in lung and respiratory cells. In a study,

Murphy *et al*. investigated DNA methylation differences in monozygotic twin pairs and found an association of a single CpG methylation (cg23603194 downstream of the *H2.0 like homeobox (HLX)* gene) and asthma in buccal cell samples (75).

Furthermore, it was reported in literature, that DNA methylation differences were observed already in nasal epithelial cells between asthmatics and controls (88, 89). Although, the majority of the studies was focussed on methylation of single CpG sites some further reported changes in differentially methylated regions (DMRs). For instance, a publication by Yang *et al.* identified DMCs (n=118) but also DMRs (n=119) within regions of *Acyl-CoA Thioesterase 7 (ACOT7), interferon gamma receptor 2 (IFNGR2) and arachidonate 15-lipoxygenase (ALOX15)* in nasal epithelial cells (77).

So far, all of the recent studies had in common, that they are array based EWAS (epigenome wide association studies) with the ability to screen only fragments of epigenome. Thereby only 5% of all CpG sites could be covered, mostly in intragenic regions. In addition, most of the identified DNA methylation perturbation were located in promoters, introns or exons of genes but not in regulatory regions away from transcription start site of the genes, in particular enhancers. Moreover, research has illuminated that no study to date has provided a comprehensive genome wide view on epigenetic pattern of asthmatic children using whole genome bisulfite sequencing.

1.5 Aim of the study

It is highly challenging to accurately diagnose asthma in early childhood. Other conditions that cause asthma-like symptoms, such as respiratory infections, can potentially mask an asthma condition or lead to a misdiagnosis. The physician needs to consider parental reported asthma-like symptoms e.g. wheeze and their frequency, as well as other lung function tests (spirometry) or allergies, in order to diagnose asthma. Due to their age, these tests are difficult to conduct in infants. The diagnosis of asthma in adolescent children is more accurate when the asthma phenotype is manifest and side effects of the symptoms (e.g. IgE levels, wheezing symptoms) can be better classified. Obviously, it is important to diagnose asthma in early childhood using validated diagnostic markers to counteract asthma development from very early on.

Wheezing episodes in general are common during childhood affecting one of three infants before the age of three and usually caused by viral lower respiratory tract infections in early life (90, 91). Whereas the asthma development and airway hyperresponsiveness seems to be facilitated by Immunoglobulin E (IgE)-mediated lung inflammation (92-94). Early childhood asthma is mostly associated with allergic hypersensitivity (atopic status), while non-allergic

(intrinsic) asthma is a rather rare condition (92, 93). IgE is an important driver in the development of impaired lung function in early childhood, therefore several studies tried to differentiate children with transient wheezing episodes from those with persistent wheeze based on early childhood serum IgE levels. These studies provide evidence that persistently wheezing children may have increased IgE levels from very early on (95). However, those studies disregard the fact that not every child with persistent or late-onset wheeze will develop asthma later on. None of the previous implemented studies concentrating on IgE, wheeze, and asthma distinguished between persistent wheeze and asthma (93, 95).

Thus, **the first aim of this thesis was to evaluate the longitudinal trends of serum IgE concentrations in wheeze and asthma considering wheeze and asthma as independent endotypes from two German birth cohorts, LINA and LISA**. In addition to IgE, the transcription of *interleukin 5 receptor α (IL5RA)* was included, as it is known to be closely linked to both IgE and asthma development. The IL5 receptor contributes together with IgE to eosinophilic and regulatory T cell pathways and thus, is an important factor for allergic inflammatory response (96, 97). Further, first studies observed epigenetic (76, 85) and transcriptional (98) associations with total serum IgE levels.

The development of asthma is strongly influenced by environmental factors (e.g. exposure to tobacco smoke or air pollution) and genetic predispositions, both contributing to altered gene regulation of asthma relevant genes most likely mediated via epigenetic mechanism. So far, the knowledge of epigenetic modification involved in asthma development is still limited. DNA methylation changes have previously been described based on selected CpG sites as they can be observed by 27K and 450K arrays or target- specific analyses. A more precise genome wide DNA methylation analysis would help to understand the development of asthma better, and thus have new entry points for future causative therapies. **Therefore, the second aim of this thesis was to investigate genome wide DNA methylation pattern of asthmatic children and healthy controls by performing whole genome bisulfite sequencing (WGBS) from 82 children of three German birth cohorts, LINA, LISA and PASTURE.** In accordance with the first part of the thesis, the influence of DNA methylation on transcription and functional regulation was explored.

The third objective of this thesis was to investigate longitudinal DNA methylation pattern and stability of asthmatic children and healthy controls at the time of birth by performing WGBS from 48 samples of the LINA and PASTURE cohort. The influence of genetic and environmental factors on DNA methylation perturbation were evaluated.

Taken together, this thesis aims to add new data to the asthma research field 1) to improve

early diagnosis of asthma with clinical relevant view of different entities between persistent wheeze and asthma; 2) to better understand the role of epigenetic perturbation in early asthma development and so, to help in identification of new biomarkers highly desirable for risk stratification, asthma prevention and therapy.

2 MATERIAL AND METHODS

2.1 Material

Table 2.1.1: Overview applied materials and software

	catalog No/ID	manufacturer
Kits		
EZ-96 DNA Methylation (Lightning) Kit	D5002/D5003	Zymo Research Europe, Freiburg, Germany
QIAamp DNA Blood Mini Kit	51106	Qiagen, Hilden, Germany
EpiTYPER® Reagent And SpectroCHIP® Kit	11377	Sequenom/Agena Bioscience, Hamburg
TruSeq-Methyl Capture EPIC Library Prep Kit	FC-151-1003	Illumina, San Diego, USA
PAXgene Blood RNA kit	762164	Qiagen, Hilden, Germany
Reagents		
Gibco™ RPMI 1640 Medium	11530586	Thermo Fisher Scientific, Massachusetts; USA
HotStartTaq DNA Polymerase	203205	Qiagen, Hilden, Germany
dNTP mix 4 x 25 µmol (250 µl), 100 mM	178.1	Carl Roth, Karlsruhe, Germany
mitogen phytohemagglutinin (PHA) 50 µg/ml	11249738001	Sigma Aldrich, Hamburg, Germany
Agarose	A8963.0500	Applichem, Darmstadt, Germany
100bp DNA ladder 50µg	T835.1	Carl Roth, Karlsruhe, Germany
Loading buffer	A3481.0010	Applichem, Darmstadt, Germany
Ethidium bromide 1%	A1152.0010	Applichem, Darmstadt, Germany
FastStart Universal Probe Master (Rox)	04913949001	Roche Applied Science, Mannheim, Germany
peqGold RNA Pure	30-1010	peqlab, Erlangen, Germany
ImProm-IITM Reverse Transcription System	A3800	Promega, Mannheim, Germany

Others		
UPL probe # 72	04688953001	Roche Applied Science, Mannheim, Germany
UPL probe # 60	04688589001	Roche Applied Science, Mannheim, Germany
UPL probe # 57	04688546001	Roche Applied Science, Mannheim, Germany
UPL probe # 65	04688643001	Roche Applied Science, Mannheim, Germany
UPL probe # 38	04687965001	Roche Applied Science, Mannheim, Germany
GE 96.96 Dynamic Array™ DNA Binding Dye Sample & Assay Loading Reagent Kit	100-3415-R	Fluidigm, San Francisco, USA
96.96 Dynamic Array™ IFC for Gene Expression	BMK-M-96.96	Fluidigm, San Francisco, USA

Instruments	
MassARRAY system	Sequenom/Agena Bioscience, Hamburg
BioMark HD Reader	Fluidigm, San Francisco, USA
IFC controller HX (96.96)	Fluidigm, San Francisco, USA
Biometra TAdvanced	Analytik Jena, Jena, Germany
HiSeq 2500/ HiSeq X	Illumina, San Diego, USA

Software	
IBM SPSS Statistics version 25	IBM Corps., Chicago, USA
GraphPad Prism	GraphPad Software Inc., La Jolla, USA
Cytoscape version 3.6.1	U.S. National Institute of General Medical Sciences (NIGMS), Bethesda, USA
Statistica, version 12/13	TIBCO Software Inc., Palo Alto, USA
RStudio version 1.2.5001	RStudio: Integrated Development for R. RStudio, Inc., Boston, USA
R version 3.6.1	R Foundation for Statistical Computing, Vienna, Austria

2.1.1 Databases

Table 2.1.2: Overview applied databases

Database	Source of supply
Ensembl Genome Browser	www.ensembl.org
National Center for Biotechnology Information	www.ncbi.nlm.nih.gov
GREAT: Genomic Regions Enrichment of Annotations Tool	http://bejerano.stanford.edu/great/public/html/
UCSC Genome Browser Gateway	https://genome.ucsc.edu/
Primer Design and Search Tool	http://bisearch.enzim.hu/
GeneCards Human Gene Database	http://www.genecards.org/
GSEA: Gene Set Enrichment Analysis	http://software.broadinstitute.org/gsea/index.jsp
ENCODE: Encyclopedia of DNA Elements	https://www.encodeproject.org/
Motif Analysis of Large Nucleotide Datasets	http://meme-suite.org/tools/meme-chip
Roadmap Epigenomics Project	http://www.roadmapepigenomics.org/

2.1.2 List of primer

Table 2.1.3: List of qPCR primer

Gene	Forward primer	Reverse primer	UPL#
PGK1	5'-tgcaaaggccttggagag-3'	5'-tggatcttgtctgcaactttagc-3'	72
GAPD	5'-gctctctgctcctcctgttc-3'	5'-acgaccaaatccgttgactc-3'	60
GUSB	5'-cgccctgcctatctgtattc-3'	5'-tccccacagggagtgtgtag-3'	57
IL5RA	5'-cagcaccaaaaagtaatatcaaagat-3'	5'-ccaaagtgacagtcaaaacacag-3'	65
EPX	5'-ccctgtctcctcaccaacc-3'	5'-gtttccgttgatcgggtgt-3'	38
IL4	5'-agctgatccgattcctgaaa-3'	5'-agctgcttgtgcctgtggaactg-3'	57

Table 2.1.4: List of MassARRAY primer

Gene	Forward primer	Reverse primer	Annealing Temperature
IL4	5'-x-GATGTTATATTAGTGAAAGGAG-3'	5'-y-CAAAAAATAAACCAAAACACTC-3'	56°C
IL5RA	5'-x-TATTTTTTGTTAATTGTATATGGTG-3'	5'-y-ACTCATTCCTCTATTAAATTTC-3'	52°C
EPX	5'-x-GTGGGGTTAGGGAGTTTATG-3'	5'-y-CAAACAACTCCTTAAAAATAATAC-3'	52°C

10Dimer forward tag X=aggaagagag; T7-reverse primer tag y-cagtaatacgactcactatagggagaaggct

2.2 Methods

2.2.1 Epidemiologic studies

For the presented studies in this thesis, samples and data were derived from the three independent prospective birth cohorts LINA (Lifestyle and Environmental Factors and their Influence on Newborns Allergy risk), LISA (Influences of Lifestyle- related factors on the Immune System and the Development of Allergies in Childhood) and PASTURE (Protection against Allergy: Study in Rural Environments). Personal contributions in the epidemiological parts of this thesis were the calculation of the respiratory phenotypes within the LINA and LISA questionnaires as well as all main statistical data analysis.

2.2.1.1 LINA cohort

The LINA (Lifestyle and Environmental Factors and their Influence on Newborns Allergy risk) cohort is a prospective mother-child study. 629 mother-child pairs (622 mothers and 629 children, 7 twin pairs) were recruited between May 2006 and December 2008 in Leipzig, Germany. The aim of the study was to investigate the impact of lifestyle and environmental factors on the newborn's risk for allergy development. Whole blood samples were obtained annually at clinical visit. Comprehensive standardized questionnaires regarding lifestyle factors and children's health outcomes were answered by the parents at the corresponding time points. Participation in the study was voluntary and informed consent was given by all participants (99, 100). The study was approved by local Ethics Committees (046-2006, 160-2008, 160b/2008, EK-BR-02/13-1, 169/13ff, 150/14ff).

2.2.1.2 LISA cohort

The LISA study (Influences of Lifestyle- related factors on the Immune System and the Development of Allergies in Childhood) is a prospective birth cohort, for which 3097 newborns were recruited from November 1997 until January 1999 in eastern and western Germany (Leipzig, Munich, Wesel and Bad Honnef). The study was designed to investigate the influence of lifestyle and environmental factors on the immune system and the development of allergic diseases. The study design was similar and comparable to the LINA study including study aims and questionnaires. Comprehensive standardized questionnaire were answered by parents and at later age additionally by the children themselves (101). Whole blood samples were collected at the time of birth (cord blood) and thereafter at the age around two, six, ten and 15 years during clinical visits. Participation in the study was voluntary and informed consent was received from all participants. The study was approved by local Ethics Committees (398-12-05112012).

2.2.1.3 PASTURE cohort– EFRAIM study group

The PASTURE cohort (Protection against Allergy: Study in Rural Environments) from the EFRAIM (Mechanisms of early protective exposures on allergy development) study group is a prospective birth cohort study focusing on rural areas within Europe. Five study centers in Austria, Germany, Finland, France and Switzerland recruited 1133 children born to farm and of non-farm women between September 2002 and May 2005. The objective of the study was to assess the effect of farming-related factors on the development of childhood asthma and allergies (102). Whole blood samples were collected at the time of birth (cord blood) and thereafter at the age around one, four, five and six years during clinical visits. Comprehensive questionnaires were answered by parents to gather information on health outcomes and environmental exposures. Participation was voluntary and informed consent was obtained by at least one parent prior study begin. The study was ethical approved by local authorities at all five study centers (103).

2.2.2 Definition of respiratory phenotypes

Information on children's respiratory outcomes was based on questionnaires. Asthma was defined based on the positive answer to the question: "Has a physician diagnosed your child with asthma during the past 12 months?", which was repeated at each follow-up examination. Wheezing was defined by positively answering the question: "Had your child whistling or wheezing sounds in the chest while breathing during the past 12 months?". Children with wheezing symptoms were separated in three different subtypes according to the classification by Martinez *et al*. (104): *Transient wheeze* was defined as wheezing symptoms occurring during the first three years of life, but not thereafter. *Persistent wheeze* was defined as wheezing at least once up to the age of three and at least once during the following years up to the age of eight. *Late-onset wheeze* was defined as wheezing after the age of three. Healthy controls never had asthma or obstructive bronchitis, or wheezing symptoms. For an overview of study participants and sample numbers of the study "The Role of IgE in asthma and wheeze" (chapter 3.1) see Table 3.1.1; for "The role of epigenetic modifications in childhood asthma" (chapter 3.2) see Table 3.2.1; for "Longitudinal DNA methylation pattern in asthma" (chapter 3.3) see Table 3.3.1.

2.2.3 IgE and cytokine measurement

Total IgE (immunoglobulin E) and specific IgE (sIgE) concentrations in sera of children were measured by the group of Prof Dr. Ulrich Sack at the University of Leipzig, Medical Faculty, Institute for Clinical Immunology using Pharmacia CAP System Phadia GmbH, Freiburg, Germany. The measurement of IL-4 concentration was performed by the group of Dr. Gunda Herberth in the Department of Environmental Immunology at the Helmholtz Centre for

Environmental Research, Leipzig, Germany. Personal contribution included the statistical analyses of the provided data.

Briefly, specific IgEs (aeroallergens rx1/sx1) were assessed by using a multiple allergen panel. rx1 mixture includes allergens of timothy, mugwort, ribwort, glass herb and birch, while sx1 panel includes allergens of house dust mite (Der p 1 dermatophagoides pteronyssinus), cat, dog, timothy, rye, birch, mugwort and cladosporium herbarum. Samples with sIgE concentrations >0.35 kU/l were regarded as positive.

To assess interleukin-4 (IL-4) cytokine concentration 500 µl heparinized whole blood samples were stimulated for 4 h at 37°C with the mitogen phytohemagglutinin (PHA, 50 µg/ml; Sigma Aldrich, Hamburg, Germany). Samples were diluted 1:1 with RPMI1640 medium without supplements. After centrifugation the supernatants were collected and stored at -80°C until further analysis. IL-4 concentrations were quantified by flow cytometry using the cytometric bead array (BD CBA Human Soluble Flex Set system, Becton Dickinson, Heidelberg, Germany) according to manufacturer's recommendation. For samples with a concentration below the detection limit of 3 pg/ml a value according to half of detection limit was set (99).

2.2.4 Gene expression

Measurement of relative mRNA expression was performed by the group of Dr. Mario Bauer in the Department of Environmental Immunology at the Helmholtz Centre for Environmental Research, Leipzig, Germany. Personal contribution included the statistical analyses of the provided data.

Total RNA was extracted by peqGold RNA Pure (peqlab, Erlangen, Germany) and the PAXgene Blood RNA kit (Qiagen, Hilden, Germany) according to the manufacture´s protocols. For the cDNA synthesis 1 µg of RNA were reverse transcribed by using the ImProm-IITM Reverse Transcription System (Promega, Mannheim, Germany). Gene expression was measured by applying 96.96 Dynamic Arrays (Fluidigm, San Francisco, CA, USA) with the FastStart Universal Probe Master (Rox) mix (Roche Applied Science, Mannheim, Germany). Intron-spanning primers were designed (Table 2.1.3). All reactions were performed in triplicates. Gene expression values were determined by the 2-ΔΔCT method (149) with GAPD, PGK1, and GUSB, as reference genes and normalized to the lowest measured value.

2.2.5 DNA extraction

Genomic DNA (gDNA) extraction was performed by the group of Dr. Mario Bauer in the Department of Environmental Immunology at the Helmholtz Centre for Environmental Research, Leipzig, Germany.

Briefly, DNA was isolated from whole blood using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's recommendation and stored at 4°C until further analyses as previously described (105).

2.2.6 Whole-genome bisulfite sequencing

Preparation and measurement of whole blood samples by WGBS were performed in the Genomics & Proteomics Core Facilities (headed by Dr. Stephan Wolf) at the German Cancer Research Center in Heidelberg, Germany. Data processing was performed by Matthias Bieg, Digital Health Center (headed by Prof. Dr. Roland Eils), Berlin Institute of Health, Berlin, Germany. Personal contribution included the calling and annotation of DMRs as well as subsequent data analysis.

Whole blood DNA samples (n=130) of the three cohorts (see Table 3.2.1 and Table 3.3.1) were subjected to whole genome wide bisulfite sequencing (WGBS) as previously described in Bauer *et al.* 2016 (106). In brief, libraries were prepared using the TruSeq DNA Sample Prep Kit v2-Set A (Illumina Inc., San Diego, CA, USA) (and EpiTect II TruSeq DNA (Illumina Inc., San Diego, CA, USA)) according to the manufacturer's instructions. Adapter-ligated libraries were treated with bisulfite and PCR-amplified. Whole genome sequencing was performed on HiSeq2000 (three lanes, 101-bp paired-end) and Illumina HiSeq X Ten V2.5 (two lanes;150bp) using standard Illumina protocols and the 200-cycle TruSeq SBS Kit v3 and HiSeq X Ten Reagent Kit v2.5 (Illumina Inc., San Diego, CA, USA). Reads were aligned against the phase II reference sequence of the 1000 genomes project including decoy sequences (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/phase2_reference_assembly_sequence/hs37d5.fa.gz). Since DNA sequences were bisulfite converted, a special alignment protocol was adapted for WGBS data (using methylCTools: https://github.com/hovestadt/methylCtools). In brief, the reference sequence was in-silico bisulfite converted (all cytosines to thymines on the forward strand, and all guanines to adenines on the reverse strand). The same procedure was performed for the sequencing reads, but the original bases were stored. Afterwards, the reads were aligned against the converted forward and reverse reference-strands, using BWA mem (version:0.7.8) (107). Duplicate reads were removed using sambamba (version:0.5.9) (108). The aligned reads were back transformed into their original state and methylation ratios were called per reference CpG site.

2.2.6.1 Determination of genome-wide differentially methylated regions (DMR) of whole blood and cord blood DNA

To determine differentially methylated regions between 40 asthma and 42 control children between the age of five and 15 (whole blood, chapter 3.2) and 23 asthma vs. 25 control newborns (cord blood, chapter 3.3) two independent statistical algorithms were used: DSS (**D**ispersion **S**hrinkage for **S**equencing data) (109) and metilene (110). A DMR was defined as three consecutive CpG sites with a test statistic *p*-value<0.01 (DSS) or *q*-value≤0.05 (metilene). Only chromosomes 1-22 were considered for analysis and gonosomes were omitted.

2.2.6.2 DSS

DSS version v2.12.0 was used to estimate DMRs. DSS is a Bioconductor package for analysis of high-throughput sequencing data. DSS uses a Bayesian hierarchical model to estimate and shrink gene- or CpG site-specific dispersions, followed by Wald tests for detecting differential methylation (109). Here at first, methylation differences of single CpG sites are defined, followed by the determination of differentially methylated regions (DMRs) based on the parameters presented in Table 2.2.1.

Table 2.2.1: DSS parameters used for DMR calling

parameter value	parameter
0	delta.m *[delta methylation threshold]*
0.01	p.threshold *[p-value threshold]*
50	minlen *[minimum length in bp]*
3	minCG *[minimum number of CpG sites]*
100	dis.merge *[merge DMRs if closer than 100bp]*
0.3	Pct.sig *[percentage of CG sites with significant p-values]*

2.2.6.3 Metilene

The second used algorithm for the determination of DMRs was metilene, developed by the group of Steve Hoffmann and Peter F. Stadler at the University of Leipzig (110). Metilene is a bioinformatical tool, which uses a segmentation algorithm to determine DMRs directly, in contrast to DSS which first determine DMCs. It investigates the maximum difference of the cumulative sum of the mean methylation value in a specific region by an approximation to the maximum difference (110). Metilene version v0.2-6 with the parameters given in Table 2.2.2 was used to determine DMRs.

Table 2.2.2: Metilene set up conditions for DMR calling

parameter value	parameter
100	M *[maximal distance in bp between CpGs within a DMR]*
3	m *[minimum number of CpGs sites]*
0	d *[delta methylation threshold]*
5	t *[minimum number of CpGs sites]*
≤0.05	*q*-value

2.2.6.4 DMR annotation and definition of enhancers

HOMER (Hypergeometric Optimization of Motif EnRichment) tools software package was used to determine initial genomic annotation to the nearest TSS (transcription start site) in order to identified DMR applying genome version hg19 (111). Predicted target genes of DMRs were defined either as an intersection with chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) data (112-114) or promoter DNaseI hypersensitive sites (DHSs) data from ENCODE (Encyclopedia of DNA Elements) (115-117).

DMRs were defined as enhancers, if their genomic location intersected (at least 1bp overlap) with previously identified enhancer regions according to ENCODE cell line sets (n = 16) (118, 119) or ROADMAP hematopoietic stem cells (HSC) or B-cell sets (n=9) and blood or T-cell sets (n=14) (120) data. All data sets considered for this analysis are summarized in Supplemental Table 8.1.1. In addition, DMRs overlapping with an active regulatory element as previously identified in an earlier LINA study by Bauer e*t al.* (106) based on chromatin immunoprecipitation DNA-sequencing (ChIP-Seq) experiments were considered as enhancer associated DMRs. Active regulatory elements were defined by merging three active states (state 1: H3K27Ac and H3K4me1 active, state 2: H3K27Ac active and state 3: H3K4me poised).

2.2.6.5 Enhancer enrichment of DMRs

Enrichment of enhancer was conducted in cooperation with Matthias Bieg, Digital Health Center (headed by Prof. Dr. Roland Eils), Berlin Institute of Health, Berlin, Germany. Personal contribution included the preparation of the data for the enrichment analysis.

To investigate the enrichment of DMRs within known enhancer regions, the algorithm/R package Locus Overlap Analysis (LOLA) (version 1.16.0) was utilized (121). LOLA is based on a Fisher's exact test, to evaluate whether the ratio of foreground features (in particular case DMRs) that overlap with a list of regions of interest (ROIs) is enriched against the ratio of background features that overlap with a list of ROIs. For ROIs merged enhancers were used from HSC or B-cell sets (n=9), from blood or T-cell sets (n=14), from ENCODE cell line sets

(n=16) and active regulatory elements of a subset of LINA children (106). Enhancer regions were merged from a set, if their distance is smaller or equal to 1000 bp. Set of background features were calculated by merging CpG positions among the whole genome that have a distance of smaller or equal to 100 bp. In a second step, all merged regions containing less than three CpG sites were removed from further analysis. The definition of the background-set of features reflects the parameters for DMR calling (refer to Table 2.2.1 and Table 2.2.2) and ensures that the foreground set of features is a subset of the background set of features.

2.2.6.6 Intersection of DMRs between the cohorts

The intersection of identified DMRs were determined using the R package Genomic Ranges between the two statistical approaches (dss and metilene), the cohorts (LINA, LISA and Pasture) and the time points within the established phenotype and at the time of birth (122). Only DMRs with a coverage (number of reads) above five were considered for further analyses according to Ziller et al. (123). DMRs were individually assessed for significance in each cohort at the established asthma phenotype time point (chapter 3.2) and at the time of birth (chapter 3.3) before intersection analyses. To this end a factorial analysis of variance (ANOVA) followed by a Bonferroni multiple test correction (p-value <0.00012) was applied.

2.2.7 Targeted DNA methylation analyses

The EpiTYPER- MassARRAY (Agena Bioscience, Hamburg, Germany) is a target specific approach to quantify DNA methylation of multiple CpG sites following bisulfite treatment of genomic DNA. During the conversion reaction unmethylated CpGs were converted to uracils while methylated CpGs remain. For targeted DNA methylation analyses 500ng gDNA from whole blood of LINA and LISA samples was bisulfite-converted using the EZ DNA Methylation Kit (Zymo Research, Freiburg, Germany) and the EZ DNA Methylation-Lighting Kit (Zymo Research, Freiburg, Germany) according to manufactures instructions. Bisulfite-converted genomic DNA was stored at -80°C until further analyses.

DNA sequences of interest were amplified by PCR with a T7-promoter tagged reverse primer, which is necessary for the following *in vitro* transcription by a T7 RNA polymerase. Primer sequences are shown in Table 2.1.4 in section Materials. For that, 2 µl bisulfite treated genomic DNA was PCR amplified using HotStarTaq DNA polymerase (Qiagen, Hilden, Germany) (final concentration/ 5µl reaction: 1x reaction buffer, 200µM dNTPs, 0.2U Polymerase, 200nM forward primer, 200nM reverse primer, H_2O) with the following cycling program.

Step	Temperature (°C)	Duration	Number of cycles
Initial denaturation	95	15 min	
Denaturation	94	30 s	45
Annealing	52(56/60)*	30 s	
Extension	72	30 s	
Final extension	72	5 min	

*depending on Amplicon see primer Table 2.1.4

To proof a successful amplification, 2 d gel electrophoresis with 1 µl PCR product was performed. DNA methylation standards (0%, 20%, 40%, 60%, 80%, 100% methylated genomic DNA) were used on each 384 well plate to control for potential PCR bias. Further, 4µl of the amplified PCR products were phosphorylated by a shrimp alkaline phosphatase (1U/µl) degrading any incorporated phosphorylated nucleotides to avoid disturbance of later reaction (20 min incubation at 37°C and inactivation at 85°C for 5 min). After the dephosphorylation step, 2 µl of dephosphorylated PCR product was transcribed *in vitro* into RNA and was cleaved uracil base-specifically by an endoribonuclease RNase A. Epityper T Complete Reagent set (Agena Bioscience, Hamburg, Germany) was used for RNA transcription (final concentration/ 7 µl reaction: 0.64x T7 Polymerase Buffer, 3.14mM DTT (Dithiothreitol), 22 U T7RNA & DNA Polymerase, 0.09mg/ml RNase A, add H_20 and T Cleavage Mix; incubation 3h at 37°C). Cleaved RNA fragments were analyzed by MALDI-TOF mass spectrometry (matrix assisted laser desorption ionization - time of flight). The methylation status of each fragment was quantified due to the mass shift of methylated against unmethylated fragments (124). The following figure shows the workflow of the MassARRAY platform (Figure 2.2.1).

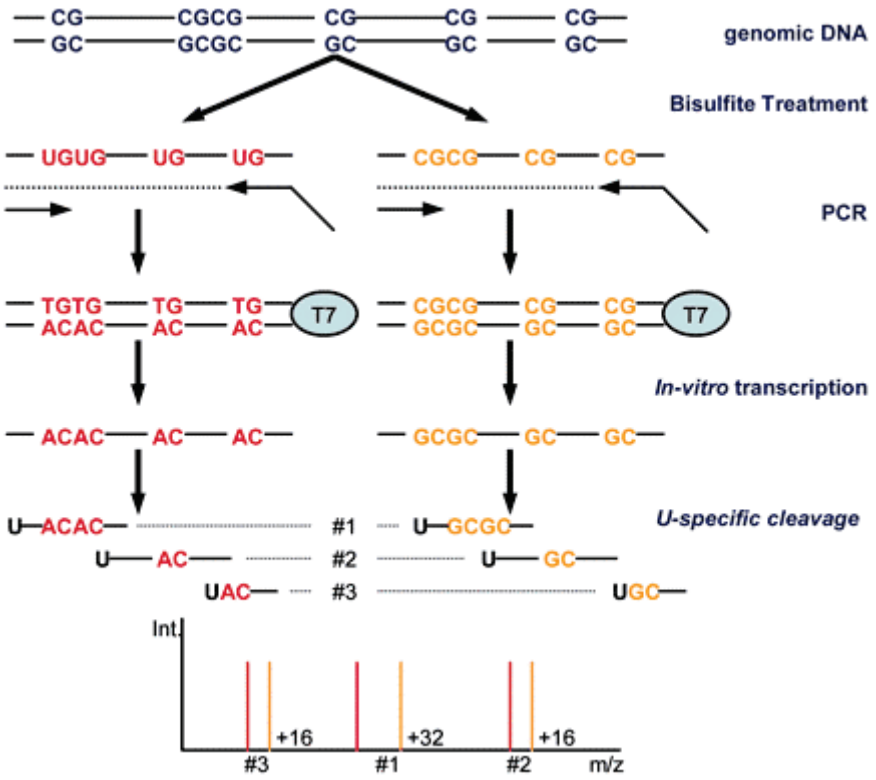


Figure 2.2.1: Analysis of DNA methylation using MassARRAY platform (124) Yellow indicated methylated CpGs and red indicated unmethylated CpGs

Personal contributions included the bisulfite conversion of the gDNA, the primer design, assay and MassArray samples preparation as well as final analysis. Subjection to mass spectrometry analysis and initial quality control were conducted by Oliver Mücke and Dr. Dieter Weichenhan at the DKFZ, Division Epigenomics and Cancer Risk factors (headed by Prof. Dr. Christoph Plass), Heidelberg, Germany. Oliver Mücke provided methylation standards.

2.2.8 SNP calling of WGBS data

SNP-calling was performed by Matthias Bieg, Digital Health Center (headed by Prof. Dr. Roland Eils), Berlin Institute of Health, Berlin, Germany.

Briefly, SNPs were called using Bis-SNP (version 0.81.2) using standard parameters (125). The derived SNP information was matched against known SNPs from dbSNP (version: 141) (126). Only SNPs matched with dbSNPs were retained for subsequent analyses.

2.2.9 Definition of genotype associated DMRs

Definition of genotype associated DMRs was performed by Matthias Bieg, Digital Health Center (headed by Prof. Dr. Roland Eils), Berlin Institute of Health, Berlin, Germany. Personal contribution included the statistical analyses of the provided data.

To identify genotype associated DMR (gDMR) a Spearman correlation test was performed to

correlate SNPs to the mean methylation values of DMRs. For each DMR, known SNPs were considered from dbSNP (version: 141) (126) within the DMR, and in 5kb vicinity around the DMR endpoints. For each SNP and sample the genotype was encoded as 0 (homozygous reference alleles), 1 (heterozygous alleles), and 2 (homozygous alternative alleles). This enabled to calculate the Spearman correlation coefficient of the genotypes versus the mean methylation values within the DMR. Furthermore, a correlation test was performed (H0: No correlation between the genotype and the mean methylation within the DMR could be identified), and rejected H0, if the corrected *p*-value was smaller or equal to 0.1. *P*-value correction was performed according to the Benjamini-Hochberg, controlling a false discovery rate (FDR) of 10 %. Therefore, DMRs were defined either as genotype associated (gDMR), if a correlation with a SNP was observed or as non genotype associated (ngDMR) if no correlation with a SNP was present.

2.2.10 Motif Analysis of DMRs

MEME Suite version 5.0.5 (Motif-based sequence analysis tools) (127, 128) were used in order to identify motifs within the DMR related genomic regions. For analysis, the MEME-ChIP tool (http://meme-suite.org/tools/meme-chip) was applied with DMRs elongated by 20bp at the start and at the end of every DMR to ensure an intersection with motifs. Only motifs with a width between six and twelve nucleotides and an *E*-value<0.05 (estimate of the statistical significance of each motif) were considered.

2.2.11 Gene enrichment and module analysis of DMRs

Gene enrichment analysis with the genomic sequences of the identified DMRs was performed using GREAT version 3.3.0 (Genomic Regions Enrichment of Annotations Tool) (129) using the whole human genome as background and a significance level of *p*-value<0.05.

Cytoscape version 3.6.1 (130) was utilized for network and module analysis of predicted target genes of these DMR associated enhancer regions and, if not an enhancer region, genes closest to the next TSS. Reactome Functional Interaction (FI) v 6.1.0 with FI network version 2016 (131) was implemented to determine network patterns of familiar and predicted interactions. These predicted interactions were estimated via Naïve Bayes Classifier. Only the original set of genes was considered without additional linker genes (132). Cluster FI network was applied to decipher modules (cluster of genes) (133). Subsequently module pathway enrichment was performed using the underlying databases CellMap (134), Reactome (135), Kyoto Encyclopedia of Genes and Genomes (KEGG) (136), US National Cancer Institute pathway interaction database (NCI PID) (137), Panther (138) and BioCarta (139). However, solely pathway enrichment results with a cut off FDR less than *p*-value<0.001 were considered.

2.2.12 Natural Language Processing

Natural language processing was performed by Dr. Sascha Schäuble and Erik Fässler of the Jena University Language & Information Engineering (JULIE) Lab, Friedrich-Schiller-University Jena, Jena, Germany and Systems Biology and Bioinformatics, Leibniz Institute for Natural Product Research and Infection Biology - Hans Knöll Institute, Jena, Germany. Personal contribution included evaluation and analyses of the provided natural language processing data.

To elucidate whether DMR related genes were previously described in association with asthma natural language processing was applied. Target gene names were mapped to human NCBI gene identifiers by symbol, name and synonym match. In case this procedure led to more than one possible NCBI Gene ID candidate target gene names were matched to GeneRIF (140) data, which are brief descriptions of gene functions frequently containing the gene name. Matches were scored with Lucene's (141) BM25 (142) implementation. The match with the highest score was used as the correct candidate target gene.

To identify associations of gene names with asthma related outcomes, the following terms were used for the search in the PubMed and the PubmedCentral literature database: asthma, asthmatic, asthmatics, wheeze, bronchial hyperreactivity, airway hyperreactivity, bronchial hyperresponsiveness, and hyperreactive airway disease.

The BANNER (143) gene tagger was employed to identify gene mentions in the texts. Mapped NCBI Gene IDs of the target genes were matched with the identified gene mentions in the literature. All documents matching an asthma related outcome term together with a target gene were extracted.

2.2.13 Literature search of DNA methylation studies related to environmental exposures

For the investigation of the impact of environmental factors and prenatal exposures on the DNA methylation at the time of birth a literature search at Pubmed (query date: 10.10.2020) was applied. Briefly, studies were identified using the search terms ("DNA methylation"[All Fields] AND "prenatal exposure"[All Fields]) OR "450K"[All Fields] OR "WGBS"[All Fields] OR "850K"[All Fields] OR "EWAS"[All Fields]. Only studies were used, which were conducted in human newborns (birth-1 month of age) or infants (birth-23 months of age) in MEDLINE, in English language and published from the year 2007 to 2020. Further, only studies with DNA methylation measurement in cord blood or blood were considered using array based (Illumina HumanMethylation27/450/850K BeadChip) or WGBS approach in order to reach comparable DNA methylation data also used in this thesis. Altogether, 32 studies were included in the

analysis. Further, four publications in the LINA cohort regarding environmental exposure and DNA methylation in cord blood were included in addition.

2.2.14 Statistical Analysis

Used software is listed at Table 2.1.1. For the study of longitudinal trends of serum IgE in asthma and wheeze (chapter 3.1) analyses were based on quantitative information for specific (sIgE) and total serum IgE (tIgE) concentrations (kU/l). Both variables were ln transformed for general estimating equations (GEE) and mediation analysis. *IL5RA* expression was considered based on the normalized values described in section 2.2.4 gene expression.

X^2- test was applied to evaluate distribution between analyzed subcohorts and entire LINA and LISA cohort. One-way repeated measurement ANOVA followed by Dunnett's post-hoc test was conducted to decipher the relationship between sIgE, tIgE and *IL5RA* expression in the different subgroups of wheezing and asthma. Fisher exact test was applied to determine differences in frequencies of elevated serum IgE (sIgE>0.35kU/l; tIgE>100kU/l) between asthmatic children compared to healthy controls or wheezing endotypes.

The longitudinal association of tIgE, *IL5RA* expression and asthma was investigated based on GEE. For this analysis a binary logistic model with a binomial distribution and logit as the link function was used. A first order autoregressive relationship (AR(1)) was assumed as the structure of the working correlation matrix. Mediation analysis with LINA and LISA data together was performed using the PROCESS macro v2.16.03 (6) for SPSS. A bootstrap approach based on 5000 samples was applied to estimate the statistical significance of the indirect effect as bias-corrected 95% confidence intervals. The following confounding factors were considered for wheezing and asthma respectively in these analyses: mode of delivery, gender, maternal history of atopy, parental education, prenatal tobacco smoke exposure, birth year.

For the study "the role of epigenetic modifications in childhood asthma" (in chapter 3.2) X^2- test was applied to evaluate the gender distribution between asthmatics and controls. A factorial ANOVA was used to assess significance of DMRs in each of the three individual cohorts. DNA methylation values were ln transformed in order to reach normal distribution.

Adjusted logistic regression analyses were applied to determine the methylation patterns of asthmatic and healthy children based on MassARRAY data from the LINA cohort (age six and/or eight) and the LISA cohort (age 15). Mann-Whitney-U test was used to compare DNA methylation and relative mRNA expression levels between asthmatics and healthy controls in six to eight year old LINA children and 15-year-old LISA children. Spearman correlation was

used to determine the association between DNA methylation and relative mRNA expression. Adjusted multiple regression model was applied to examine the effect of DNA methylation and asthma outcome on relative mRNA expression.

Mediation analysis were performed using the PROCESS macro version v3.4 (24) for SPSS. A bootstrap approach based on 5000 samples was applied to estimate the statistical significance of the indirect effect as bias-corrected 95% confidence intervals. Confounding factors used in the analyses were gender and family history of atopy. Additional confounding factor affiliation to cohort was applied in multiple regression model.

For the chapter 3.3 “longitudinal DNA methylation pattern in childhood asthma” one factorial ANOVA was applied to identify significant DMRs in LINA and PASTURE cohort separately. Intersection analysis using R package GenomicRanges version 1.40.0 was applied to compare the identified DMRs with the DMCs/DMRs of DNA methylation studies related to environmental exposures.

3 RESULTS

3.1 The role of IgE in asthma and wheeze

Note: Results of the following chapter have been published in Klös *et al.* (2019) *"Longitudinal trends of serum IgE and IL5RA expression throughout childhood are associated with asthma but not with persistent wheeze"* in Allergy (DOI: 10.1111/all.13837). Personal contributions included the preparation of all figures and tables and the initial manuscript, as well as the statistical analyses regarding the longitudinal association between serum IgE, asthma and the wheezing phenotypes. Contribution by others are highlighted in the corresponding method descriptions.

3.1.1 Study characteristics and phenotype classification

This study is based on samples and data derived from the two German birth cohorts LINA (38) and LISA (144). For the LINA study (discovery cohort) annual, longitudinal blood samples until the age of eight years were available, including information on total serum IgE (tIgE), specific IgE (sIgE sensitization against aeroallergens), *IL5RA* mRNA expression, wheezing symptoms, and asthma onset. For the validation, blood samples of the LISA study (validation cohort) from two, six, ten and 15 year-old children were used. For subsequent analyses, only children with longitudinal blood samples and available RNA (LINA: n=98; LISA: n=453) were included (Figure 3.1.1).

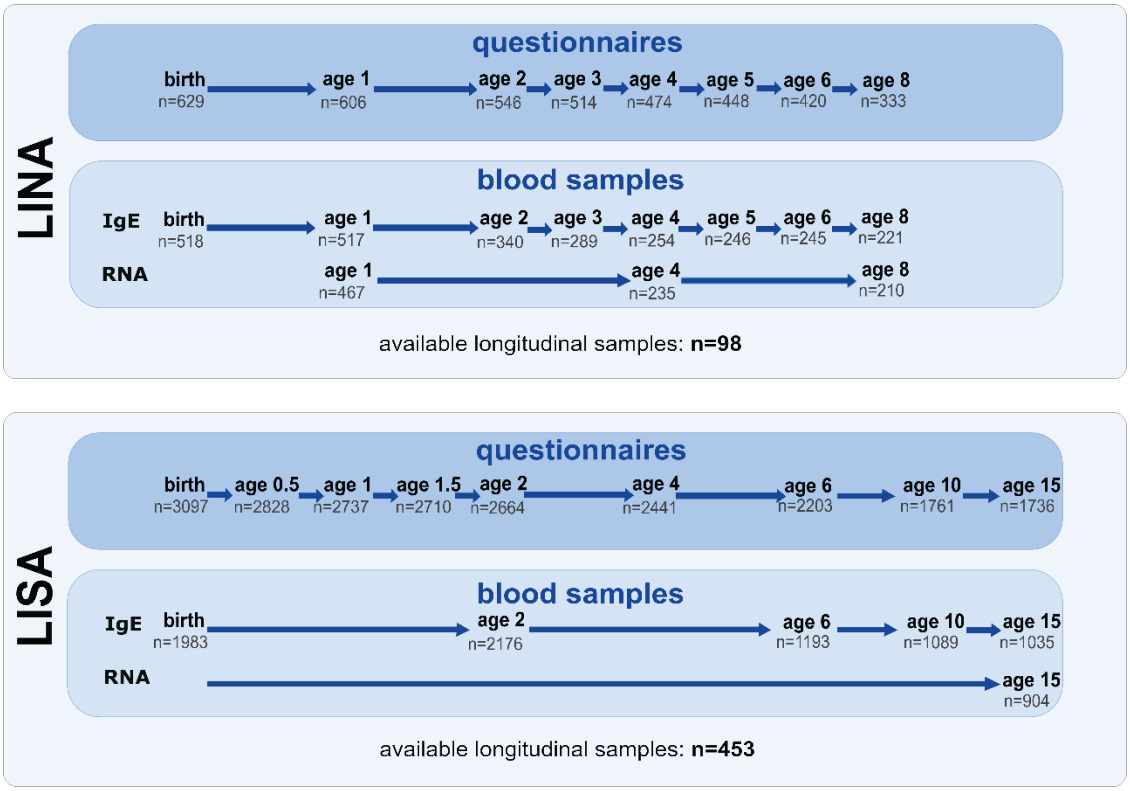


Figure 3.1.1: Overview of the subcohorts with complete longitudinal questionnaires, RNA availability and IgE measurements at all time points.

There were no differences between the analysed subcohorts and the entire cohorts (LINA age

8 n=333 and LISA age 15 n=1736) regarding gender, parental history of atopy, smoking during pregnancy, parental educational level and mode of delivery, neither in LINA nor in LISA (Table 3.1.1).

Table 3.1.1: General study population characteristics of LINA and LISA cohorts. Chi-Square test (*p*-value) was applied to compare analysed subcohorts with the entire cohorts of LINA (age of eight) and LISA (age of 15), respectively.

	LINA (Discovery cohort)			**LISA (Validation cohort)**		
	Analyzed sub-cohort n = 98[a] n(%)	**Total cohort age 8** n = 333[e] n(%)	***X²*-Test** [*p*-value]	**Analyzed sub-cohort** n = 453[b] n(%)	**Total cohort age 15** n = 1736[f] n(%)	***X²*-Test** [*p*-value]
Gender of the child			0.96			0.10
Female	48(49.0)	164(50.7)		198(43.7)	833(48.0)	
Male	50(51.0)	169(49.3)		255(56.3)	93(52.0)	
Parental history of atopy[c]			0.99			0.93
No	33(33.7)	109(32.7)		201(44.5)	799(46.0)	
Mother	47(48.0)	162(48.7)		166(36.7)	631(36.4)	
Father	36(36.7)	123(37.0)		148(32.7)	544(31.3)	
Both	18(18.4)	61(18.4)		63(14.0)	243(14.0)	
Smoking during pregnancy			0.80			0.67
No	88(89.8)	296(88.9)		383(87.2)	1449(83.5)	
Yes	10(10.2)	37(6.9)		56(12.8)	227(13.1)	
Parental educational level[d]			0.78			0.33
Low	1(1.0)	4(1.2)		10(2.2)	60(3.5)	
Intermediate	22(22.5)	64(19.2)		116(25.7)	413(23.8)	
High	75(76.5)	265(79.6)		326(72.1)	1249(72.0)	
Mode of delivery			0.61			0.42
Spontaneous	77(78.6)	254(76.4)		324(73.3)	1221(50.5)	
Caesarean	21(21.4)	75(22.6)		84(19.0)	305(17.6)	
Others	0(0.0)	3(0.9)		34(7.7)	164(9.5)	
Lung outcomes			0.74			0.10
Never[g]	51(52.0)	163(48.9)		237(52.3)	877(50.5)	
Wheeze	37(37.8)	128(38.4)		167(36.9)	495(28.5)	
Asthma	10(10.2)	25(7.5)		49(10.8)	141(8.1)	

[a] n may be different from 98 due to missing data;

[b] n may be different from 453 due to missing data;

[c] defined as: occurrence of asthma, atopic dermatitis, hay fever or food allergy; assessed at the time of pregnancy

[d] low = 8 years of schooling ('Hauptschulabschluss`); intermediate = 10 years of schooling (`Mittlere Reife`); high = 12 years of schooling or more (`(Fach-)hochschulreife'); assessed at birth

[e] LINA study participants at the 8-year-follow-up

[f] LISA study participants at the 15-year-follow-up

[g] no asthma, wheezing, or obstructive bronchitis during observational period

For analyses, four different lung dysfunction phenotypes were distinguished. The first group is composed of children with a physician diagnosed asthma, according to the question: “Has a physician diagnosed your child with asthma during the past 12 months?”. Within the studied LINA and LISA subcohorts 10.2% (10/98) and 10.8% (49/453) of the children were ever diagnosed with asthma, respectively. Children with parent-reported wheezing episodes but without an asthma diagnosis were classified according to their wheezing endotypes: reported wheezing up to the age of three but not thereafter (“transient”; 22.5% (22/98) and 19.9% (90/453)), wheezing reported after age three (“late-onset”; 7.1% (7/98) and 10.2% (46/453)), and wheezing during the first three years of life and at least one time thereafter (“persistent wheeze”; 8.2% (8/98) and 6.8% (31/453)). The children without wheezing symptoms, 52.0% (51/98) and 52.3% (237/453) respectively, were defined as controls and were compared to children with wheezing symptoms or to children diagnosed with asthma.

3.1.2 Longitudinal pattern of specific aeroallergen and total serum IgE

Immunoglobulin E (IgE) is known as one of the most important factors for the immune-related response of atopic asthma (12). Both cohorts showed significantly different specific IgE concentrations (sIgE) against aeroallergens between groups over time using one-way repeated measurement ANOVA (Figure 3.1.2, left panel). The specific IgE test against aeroallergens included timothy, mugwort, ribwort, glass herb and birch (Rx1; only LISA year two), while Sx1 includes house dust mite (Der p 1 dermatophagoides pteronyssinus), cat, dog, timothy, rye, birch, mugwort and cladosporium herbarum (all other time points in LINA and LISA). Highest concentrations were observed in asthmatic children (mean concentration LINA (age1-8)=19.4 kU/l; LISA (age2-15)=25.3 kU/l) compared to healthy controls (mean concentration LINA (age1-8)=3.0 kU/l; LISA (age2-15)=5.3 kU/l).

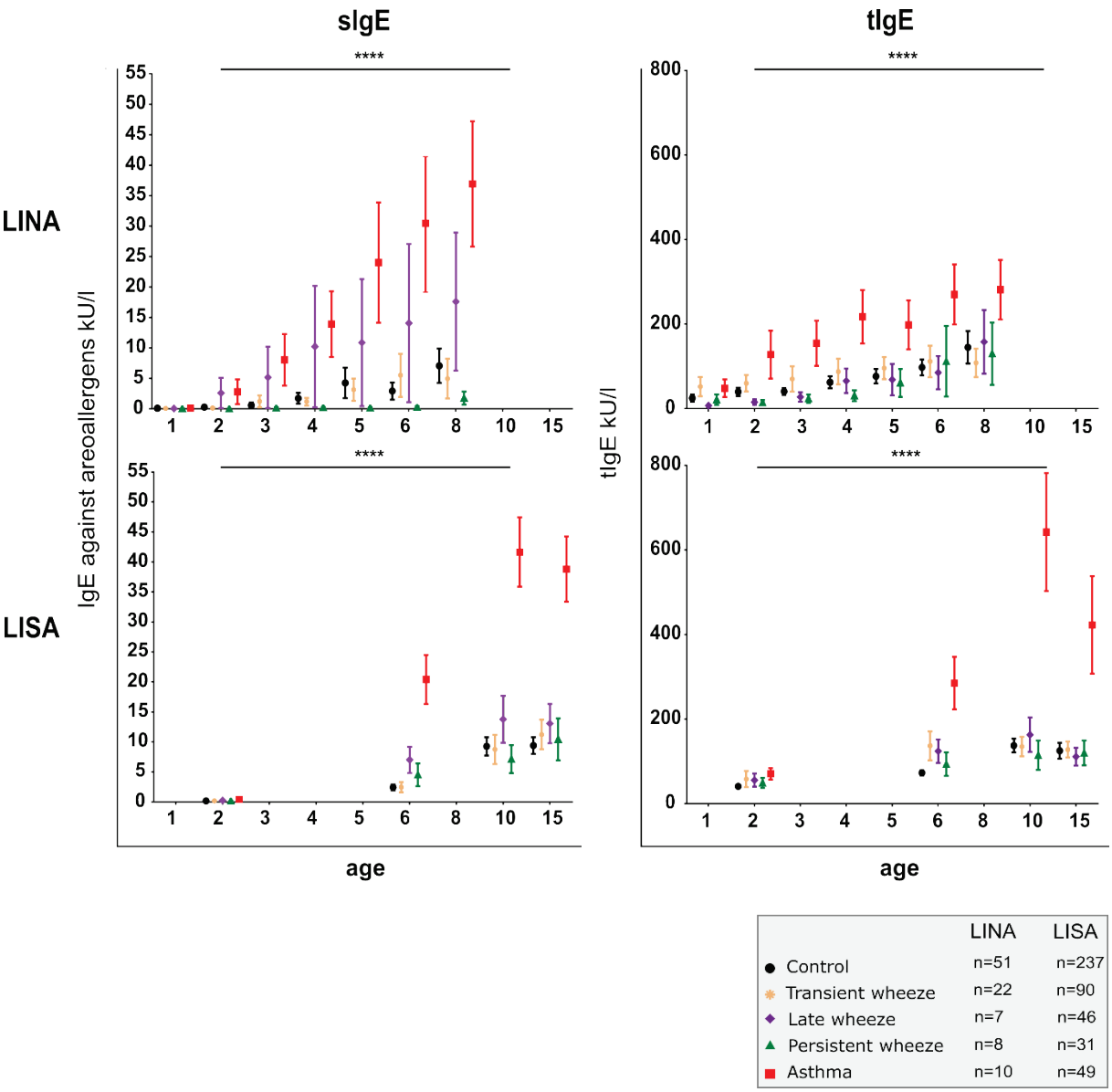


Figure 3.1.2: Longitudinal pattern of specific (sIgE) against aeroallergens and total serum IgE (tIgE) concentrations in the LINA (n=98) and LISA (n=453) cohorts**.** Given are mean +/- SEM. A repeated measurement one-way ANOVA was applied to determine the difference between groups over time. *p*-value<0.05 *; *p*-value<0.01**; *p*-value<0.001***; *p*-value<0.0001****

LINA children diagnosed with asthma later in life showed significantly higher sIgE levels already starting at age four (*p*-value=0.028 from post hoc comparison by Dunnett's test, Table 3.1.2). From age five onwards, the sIgE concentration significantly differed between persistent wheeze and asthma. This difference lasted until the age of eight years. In line, in the validation cohort LISA from age six onward – the earliest available time-point after the age of two – asthmatic children showed significantly higher sIgE concentrations compared to all wheezing endotypes including persistent wheeze (Table 3.1.2; Table 3.1.3).

Table 3.1.2: Comparison of specific IgE against aeroallergens (sx1/rx1) between asthmatic and non-asthmatic children in the LINA cohort (n=98).

compared phenotypes			age 1	age 2	age 3	age 4	age 5	age 6	age 8
asthma	vs	controls	1.0	0.952	0.291	**0.028**	**0.0001**	**0.0001**	**0.0001**
		persistent	1.0	0.978	0.525	0.095	**0.0006**	**0.0001**	**0.0001**
		late	1.0	1.000	0.977	0.946	0.138	**0.044**	**0.011**
		transient	1.0	0.958	0.460	**0.040**	**0.0002**	**0.0001**	**0.0001**

p-value < 0.05 are indicated in bold; Dunnett´s test

Table 3.1.3: Comparison of specific IgE against aeroallergens (sx1/rx1) between asthmatic and non-asthmatic children in the LISA cohort (n=453).

compared phenotypes			age 2	age 6	age 10	age 15
asthma	vs	controls	1.0	**0.0001**	**0.0001**	**0.0001**
		persistent	1.0	**0.0009**	**0.0001**	**0.0001**
		late	1.0	**0.0019**	**0.0001**	**0.0001**
		transient	1.0	**0.0001**	**0.0001**	**0.0001**

p-value < 0.05 are indicated in bold; Dunnett´s test

Similarly, total IgE (tIgE) was significantly different between groups over time (Figure 3.1.2, right panel). Again, highest concentrations were observed in asthmatic children (mean concentration LINA (age1-8)=185.4 kU/l; LISA (age2-15)=355.5 kU/l) compared to healthy controls (mean concentration LINA (age1-8)=69.6 kU/l; LISA (age2-15)=94.5 kU/l). However, differences in tIgE between wheezing endotypes and asthmatic children were less consistent compared to sIgE (Table 3.1.4; Table 3.1.5).

Table 3.1.4: Comparison of total IgE between asthmatic and non-asthmatic children in the LINA cohort (n=98).

compared phenotypes			age 1	age 2	age 3	age 4	age 5	age 6	age 8
asthma	vs	controls	0.969	0.204	0.062	**0.005**	**0.041**	**0.001**	**0.017**
		persistent	0.983	0.253	0.150	**0.017**	0.123	0.058	0.074
		late	0.936	0.292	0.196	0.090	0.182	**0.026**	0.218
		transient	1.000	0.515	0.318	0.051	0.169	**0.011**	**0.004**

p-value < 0.05 are indicated in bold; Dunnett´s test

Table 3.1.5: Comparison of total IgE between asthmatic and non-asthmatic children of controls, and wheezing endotypes in the LISA cohort (n=453).

compared phenotypes			age 2	age 6	age 10	age 15
asthma	vs	controls	0.917	**0.0001**	**0.0001**	**0.0001**
		persistent	0.998	**0.015**	**0.0001**	**0.0001**
		late	0.997	**0.024**	**0.0001**	**0.0001**
		transient	0.993	**0.015**	**0.0001**	**0.0001**

p-value < 0.05 are indicated in bold; Dunnett´s test

In summary, total and specific IgE against aeroallergens concentrations showed a significant difference in the five groups over time in the LINA and LISA cohorts. The highest sIgE and tIgE levels were observed in the asthmatics and differed significantly from the healthy controls and the wheezing groups starting already at the age of four in the LINA cohort, respectively at the age of six in the LISA cohort.

3.1.3 Frequency of enhanced total IgE and sensitization against aeroallergens

Higher sensitization (sIgE>0.35 kU/l) frequencies against aeroallergens (sx1/rx1) and enhanced total serum IgE concentrations (tIgE>100 kU/l) were found in asthmatic children compared to all other wheezing endotypes. Asthmatic children had enhanced total IgE level twice as often compared to healthy controls or any of the wheezing endotypes (Figure 3.1.3 and for detailed percentage numbers of frequency see supplemental Table 8.1.2 and Table 8.1.3). Only in LINA, late-onset wheezers showed higher frequency of elevated sIgE (sx1) compared to tIgE, which most likely was related to those children not yet diagnosed with asthma during the observation period until the age of eight years. Noteworthy, the frequency of enhanced specific and total IgE in the persistent wheezing group was more similar to the healthy controls than to the asthma group in both cohorts.

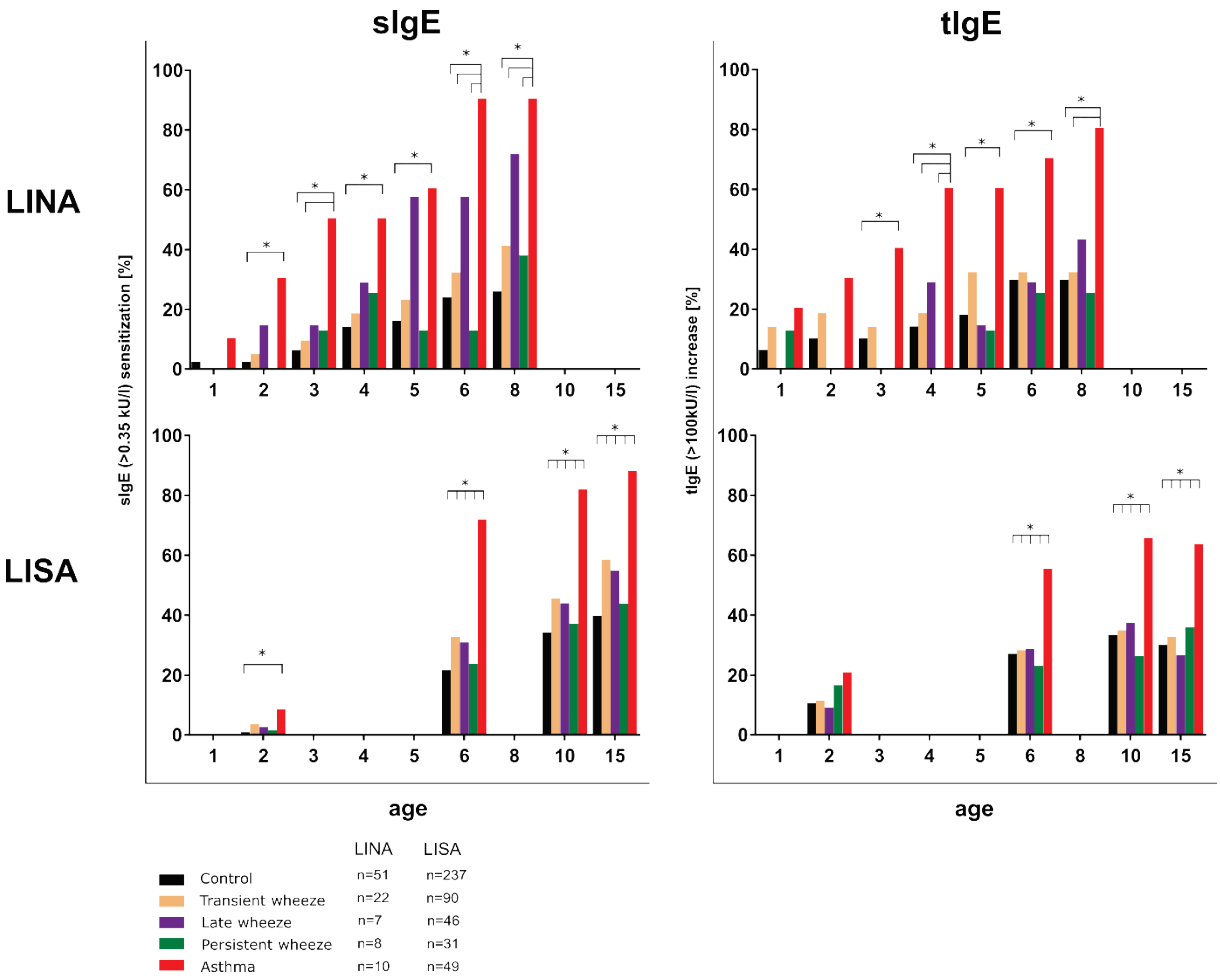


Figure 3.1.3: Percentages of children with a positive sensitization to aeroallergens (>0.35 kU/l) or elevated total serum IgE (>100 kU/l). Fisher exact test was applied to determine differences between asthmatic children, controls and wheezing endotypes (* *p*-value< 0.05).

In summary, asthmatic children were twice as likely to show elevated specific IgE against aeroallergens and increased total IgE compared to healthy controls and all wheezing groups in the LINA and in the LISA cohort. The sensitization pattern for persistent wheeze resembled that of healthy controls rather than asthmatics.

3.1.4 Association between specific and total serum IgE, and *IL5RA*

Recently interleukin 5 receptor α (*IL5RA*), which is a crucial driver for the allergic inflammatory response (96, 97), has been epigenetically (85) and transcriptionally (98) associated with total serum IgE levels. To elucidate, whether transcriptional changes in *IL5RA* were linked to IgE levels and asthma in the present study, the *IL5RA* mRNA expression was measured using a 96.96 Dynamic Array (Fluidigm, San Francisco, CA, USA).

Similarly to IgE, *IL5RA* mRNA expression was significantly increased only in children with asthma compared to controls and all wheezing endotypes starting at the age of four (Figure 3.1.4). In addition, a significant difference (*p*-value<0.01) in *IL5RA* mRNA expression was found between asthma, controls and transient wheezing children at the age of 8 years for LINA and at the age of 15 years for LISA. There was also a significant difference (*p*-value<0.01) between asthma and persistent wheeze in 8 year old children in the LINA cohort (Figure 3.1.4).

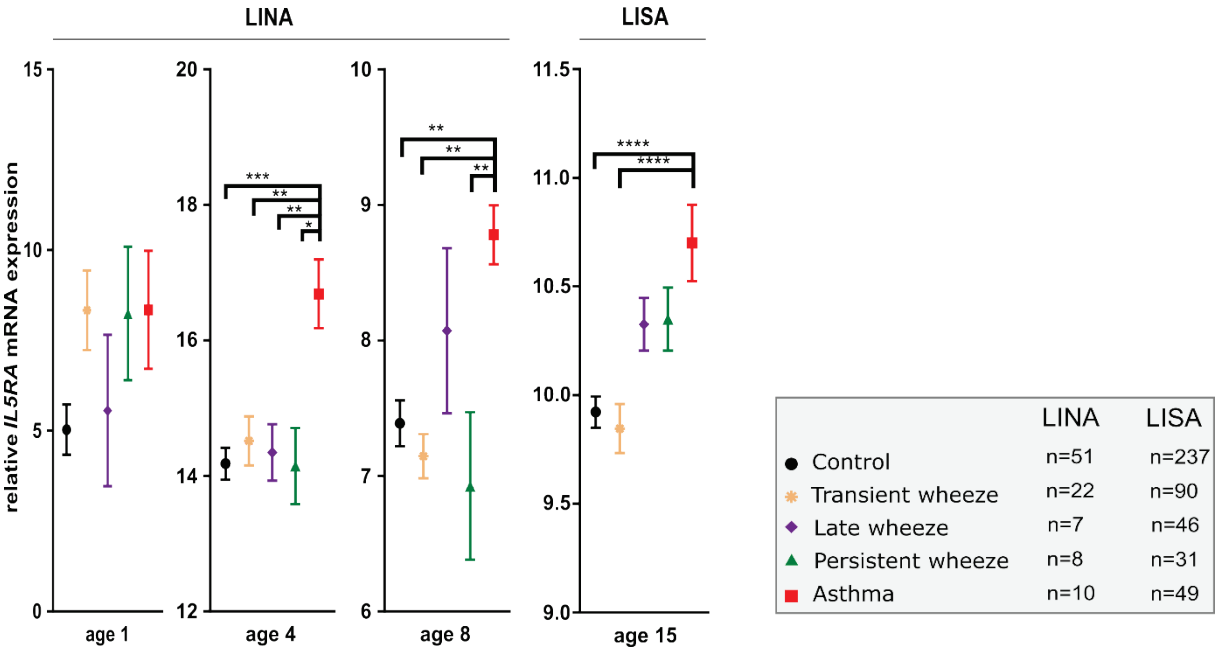


Figure 3.1.4: Comparison of relative *IL5RA* mRNA expression in asthma, different wheezing endotypes, and controls in the LINA (n=98) and the LISA cohort (n=453), respectively. One-way ANOVA followed by Dunnett´s post hoc test was applied to determine the difference between asthma and the other groups. *p*-value< 0.05*; *p*-value<0.01**; *p*-value<0.001***; *p*-value<0.0001****

To elucidate whether the increase of *IL5RA* mRNA expression in asthmatic children occurs before the disease onset, four-year-old children of the LINA study who were subsequently diagnosed with asthma were examined in comparison to all other wheezing groups and healthy controls. Children who were diagnosed with asthma later in their life, had already at the age of four a significantly increased *IL5RA* mRNA expression compared to all wheezing groups and healthy controls (Figure 3.1.5).

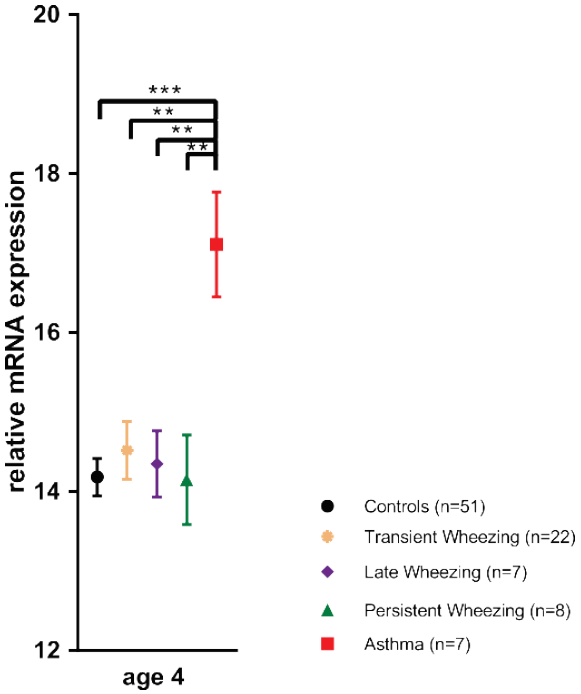


Figure 3.1.5: Differential expression of IL5RA in 4-year-old children considering in the asthma group only those children who were first diagnosed for asthma later in life (n=7).: *p*-value< 0.05 *; *p*-value<0.01**; *p*-value<0.001***; *p*-value<0.0001****

Taken together, a significantly increased *IL5RA* mRNA expression was observed in the asthmatic children compared to healthy controls and all wheezing groups in the LINA cohort from the age of four and in the LISA cohort at the age of 15. In addition, four-year-old children of the LINA study who were subsequently (after age 4) diagnosed with asthma showed increased *IL5RA* mRNA expression already before disease onset.

To determine the longitudinal association between specific and total IgE concentrations, *IL5RA* mRNA expression and asthma development, general estimation equations (GEE) were applied adjusted for mode of delivery, gender, maternal history of atopy, parental education, prenatal tobacco smoke exposure and birth year. The advantage of this model is the assessment of the correlation between variables over time in contrast to cross-sectional design models, which have to consider every point in time individually. In this adjusted GEE model both total and specific IgE against aeroallergens showed a significant longitudinal association with the development of asthma (Table 3.1.6).

Table 3.1.6: Results of longitudinal association of specific IgE (aeroallergens Sx1/Rx1), total IgE, and *IL5RA* expression of asthmatic children compared to healthy controls in the LINA and LISA cohort (wheezing children were excluded from these analyses). Given are ORs with upper and lower 95% CI and *p*-values from adjusted* general estimation equations.

	LINA			LISA		
	OR	**95%CI**	***p*-value**	**OR**	**95%CI**	***p*-value**
sIgE	1.13	1.03-1.22	0.006	1.46	1.33-1.59	<0.0001
tIgE	1.20	1.06-1.35	0.003	1.71	1.38-2.11	<0.0001
IL5RA	1.04	1.03-1.06	0.0001		---#	

*adjusted for: gender, maternal history of atopy, parental educational level, mode of delivery and prenatal tobacco smoke exposure (cotinine level)
RNA only available in the LISA cohort for age 15

In addition, *IL5RA* mRNA expression was significantly longitudinally associated with the development of asthma in LINA. In the LISA cohort, RNA was only available at age of 15 years. Therefore, the longitudinal influence of *IL5RA* mRNA expression on asthma development could not be investigated. Neither in LINA nor in LISA the *IL5RA* expression was associated with persistent wheezing or any other of the wheezing endotypes (Figure 3.1.6 A and B, data for *IL5RA* were not shown).

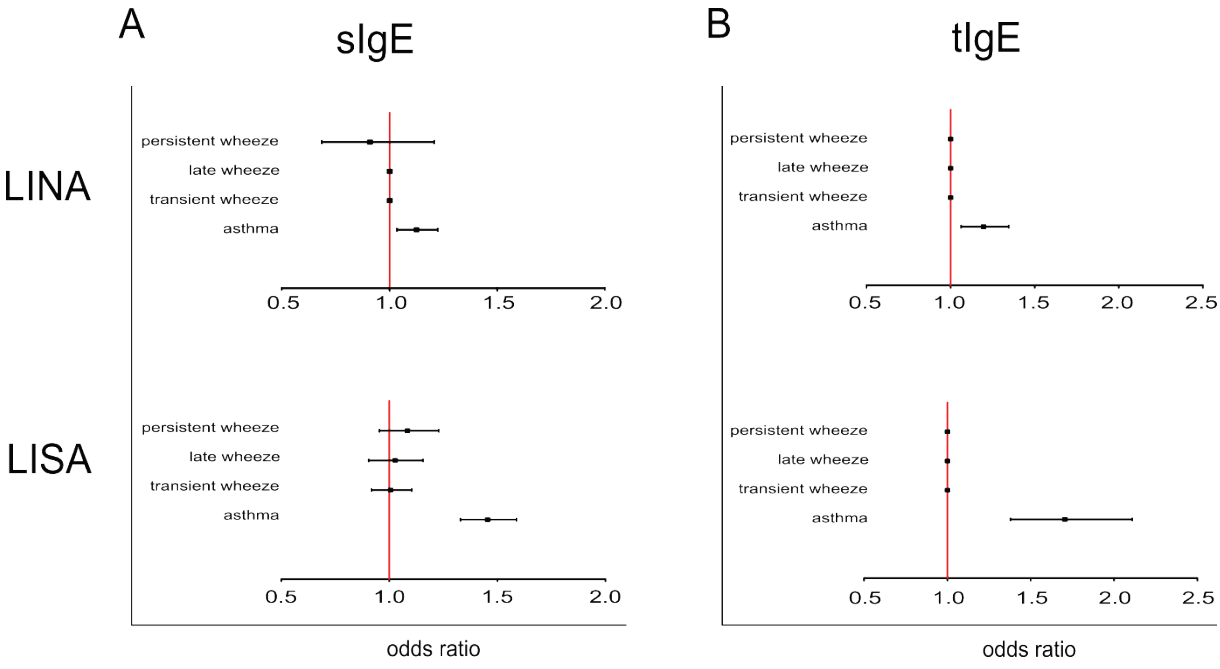


Figure 3.1.6: Longitudinal association of specific serum IgE against aeroallergens **(A)** and total IgE **(B)** to respiratory outcomes in the LINA and LISA cohorts. GEEs considered IgE concentrations determined at age of one to six, and age of eight for LINA, for LISA IgE concentrations from age of two, six, ten and 15. Given are odds ratios with upper and lower 95% CIs from GEEs adjusted for gender, maternal history of atopy, parental educational level, mode of delivery, prenatal tobacco smoke exposure and birth year.

Taken together, elevated serum concentrations of specific IgE against aeroallergens or of total IgE and an increased *IL5RA* gene expression were longitudinally associated with an increased risk of developing asthma. The wheezing outcomes, however, showed no significant longitudinal correlation.

3.1.5 Association between total serum IgE, *IL5RA* and asthma

The results presented so far, considered total serum IgE levels and *IL5RA* mRNA expression individually in their relationship to asthma development. Mediation models allow to assess the relationship between several variables and separates their impact on an outcome variable as direct or indirect. A macro for SPSS developed by Andrew F. Hayes (145) was used to test whether elevated IgE levels had a direct influence on the development of asthma or whether this effect was rather indirect and mediated by changes in *IL5RA* mRNA expression. For this analysis, the eight-year-old children of the LINA and the fifteen-year-old children of the LISA cohort were combined in a meta-analysis so that a higher number of cases could be considered. The model was adjusted for gender, maternal history of atopy, mode of delivery and prenatal tobacco smoke exposure. The mediation analyses revealed a direct effect of tIgE (unstandardized b=0.41, CI=0.19-0.64) as well as an indirect effect of tIgE mediated by *IL5RA* mRNA expression (unstandardized b=0.16, CI=0.06-0.30) on asthma (Figure 3.1.7).

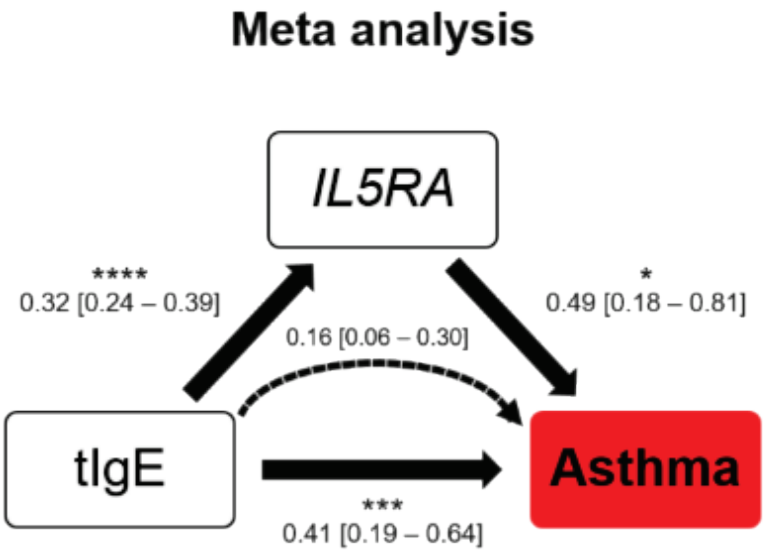


Figure 3.1.7: Mediation meta-analysis for the relationship of tIgE, *IL5RA* expression and asthma development for eight-year-old children in LINA and 15-year-old children in LISA together. Models were adjusted for gender, maternal history of atopy, mode of delivery and prenatal tobacco smoke exposure. Given are unstandardized b-values and lower and upper CI. *p*-value< 0.05*; *p*-value<0.01**; *p*-value<0.001***; *p*-value<0.0001****

In summary, the interaction of total serum IgE and *IL5RA* expression has a crucial influence on the development of asthma.

The presented study showed a clear separation of asthma and wheezing endotypes using longitudinal total serum IgE and specific IgE (against aeroallergens) concentrations together with the mRNA expression of *interleukin 5 receptor α (IL5RA)* in two German birth cohorts, LINA and LISA.

3.2 The role of epigenetic modifications in childhood asthma

Note: The following results are currently in preparation for submission in Klös *et al.* (2021) *"Role of epigenetic mechanism in asthma development in early childhood"* for the Journal of Allergy and Clinical Immunology. Personal contributions include the preparation of the manuscript, DMR calling and all subsequent downstream analyses and preparation of all figures and tables. Contribution of others include initial data processing of WGBS data and relative gene expression analysis as indicated in the corresponding method descriptions as well as work on the manuscript.

3.2.1 Study characteristics

For this study, whole genome bisulfite sequencing (WGBS) was used to measure the DNA methylation pattern in whole blood samples of children with asthma and age-matched controls (median age=6). Overall, 82 whole blood samples from three independent German birth cohorts (LINA, age 5-8; LISA, age 15 and PASTURE/EFRAIM, age 6) were used for the determination of differentially methylated regions (DMRs) in asthmatic children vs. healthy controls. The 82 samples were composed of n=25 samples from LINA, n=29 from LISA and n=28 from PASTURE/EFRAIM. There was no selection bias with regard to gender (Table 3.2.1). In total, 40 asthmatic children were compared to 42 healthy controls (= WGBS cohort).

Table 3.2.1: Characteristics of the samples used for whole genome bisulfite sequencing

	LINA (N=25) n(%)	**LISA** (N=29) n(%)	**PASTURE/EFRAIM** (N=28) n(%)	**WGBS cohort** (N=82) n(%)
age				
	5-8	15	6	6[a]
gender				
female	13(48)	18(62.1)	5(17.9)	36 (43.9)[b]
male	12(52)	11(37.9)	23(82.1)	46 (56.1)[b]
lung phenotype				
healthy	13(48)	15(51.7)	14(50)	42 (51.2)
asthma	12(52)	14(48.3)	14(50)	40 (48.2)

[a]age matched between asthma and control group, median age; [b]χ^2-test *p*-value=0.26

3.2.2 Determination of DMRs and further workflow

By applying two independent algorithms, metilene and DSS, 824 and 1182 DMRs were identified, respectively. In order to obtain statistically reliable results, the intersection between the DMRs identified by the two independent DMR calling approaches was determined in the total cohort. The resulting 423 DMRs were further evaluated by applying a one factorial ANOVA including a multiple test correction by Bonferroni (*p*-value<0.00013) to test, whether these regions were significantly differently methylated in each of the three separate cohorts. Out of the 423 concordant DMRs, 341 DMRs in LINA, 311 DMRs in LISA and 267 DMRs in PASTURE/EFRAIM remained statistically significant, the intersect of all three cohorts included

161 DMRs (Figure 3.2.1). All further down-stream analyses were conducted on this set of 161 DMRs (see supplemental Table 8.1.4 for DMR List). First, the distribution of the DMRs across the genome was assessed and their association to the genotype was evaluated. Next, gene-set enrichment and transcription factor binding motif analyses were conducted to obtain further information on potentially affected biological processes.

For further validation, the DNA methylation changes observed in the WGBS samples, selected regions associated with asthma-relevant genes, were validated in the entire LINA and LISA cohort using a targeted DNA-methylation analysis approach based on mass spectrometry (MassARRAY). Subsequently, it was investigated whether genes or proteins associated with these DMRs showed altered expression.

To identify regulatory networks affected by DNA methylation changes a regulatory network analysis was performed and natural language processing was applied to assess whether genes in this network had been previously associated with an asthma-related phenotype.

To investigate whether the altered DNA methylation patterns observed in relation to the established asthma phenotype were already present before the onset of the disease or even at the time of birth, longitudinal WGBS data of the LINA and PASTURE/EFRAIM cohort were analysed (Chapter 3.3). As single nucleotide polymorphisms (SNP) can persistently influence DNA methylation, individuals were genotyped based on the WGBS data, providing the opportunity to distinguish SNP-associated from prenatal exposure-related DNA methylation changes (Figure 3.2.1).

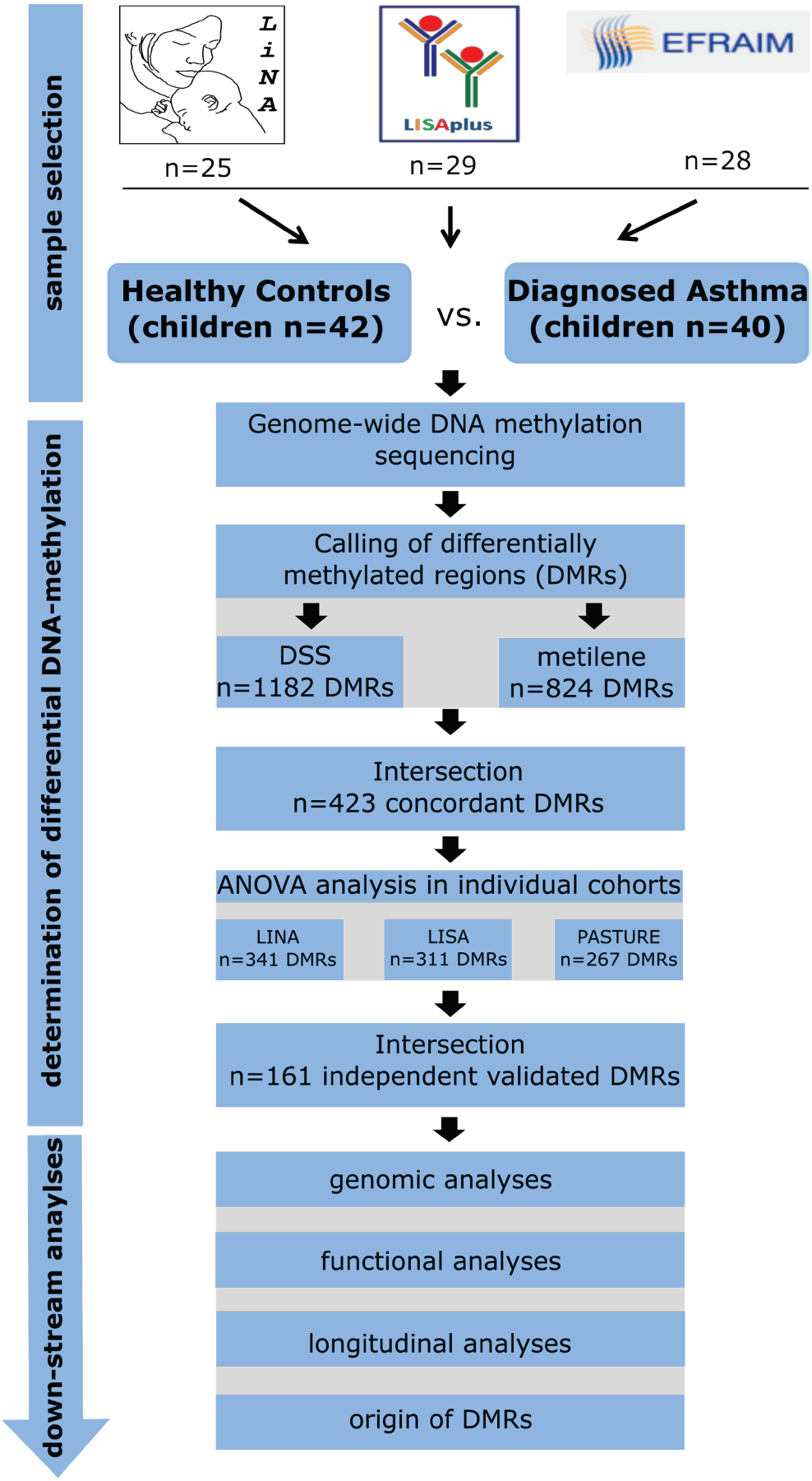


Figure 3.2.1: Overview of analysis workflow to determine DMRs in asthmatic children and subsequent functional analyses.

3.2.3 Characteristics and genomic distribution of 161 asthma-related DMRs

The 161 DMRs identified in the three cohorts showed an average sequencing depth of 32 reads, which aligned to known human reference DNA sequences. The length of the DMRs varied between 15 bp and 1376 bp. Each DMR contained between three to 135 CpG sites. Only DMRs allocated on autosomes (Figure 3.2.2) were considered, 58.4% of these DMRs (n=94) were located in intragenic and 41.6% (n=67) in intergenic genomic regions. Overall, 160 DMRs were hypomethylated (99.4%) in asthmatics. Only one DMR located in first exon of the *TET3* gene was hypermethyated in asthmatic children compared to healthy controls. A detailed list of all 161 DMRs is provided in the supplements (Table 8.1.4).

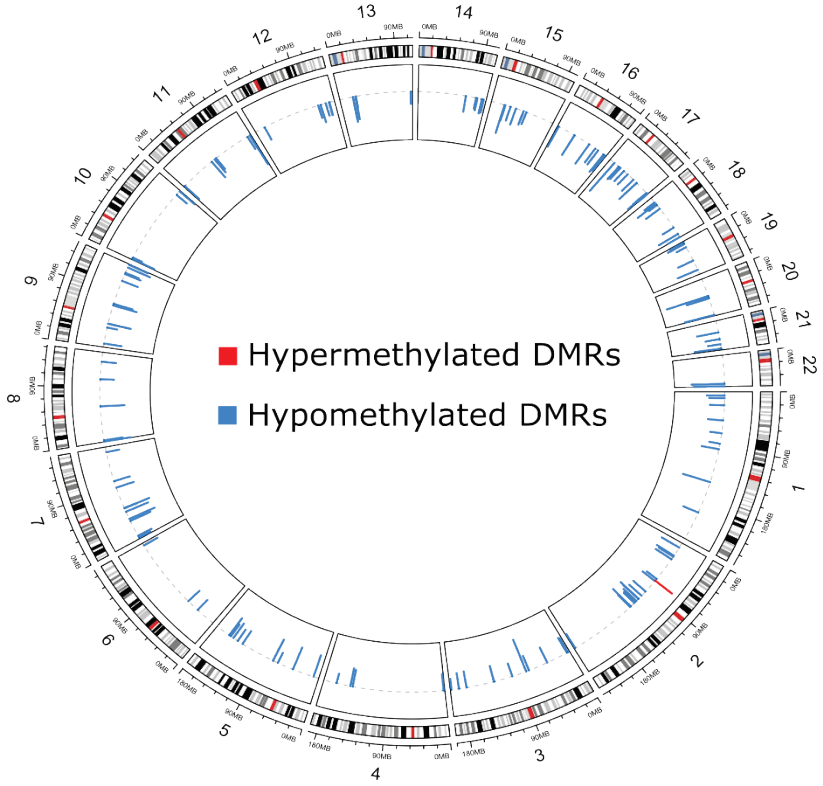


Figure 3.2.2: Genome-wide distribution of the 161 differentially methylated regions (DMRs) identified in asthmatic children compared to healthy controls. The vast majority of these DMRs were hypomethylated in asthmatic children

The DMRs were especially enriched in enhancer regions (p-value<1.70^{-12}). As such 62.7% of the 161 DMRs (n=101) overlapped with either an enhancer region according to the ENCODE database (119), or the ROADMAP database (120) for T- and B-cell- enhancers, or overlapped with active enhancer marks determined in an earlier LINA study by Bauer *et al*. (106) based on ChIP-Seq experiments (H3K27Ac and H3K4me1 marks). Only 1.2% of the DMRs were located in promoter regions (Figure 3.2.3).

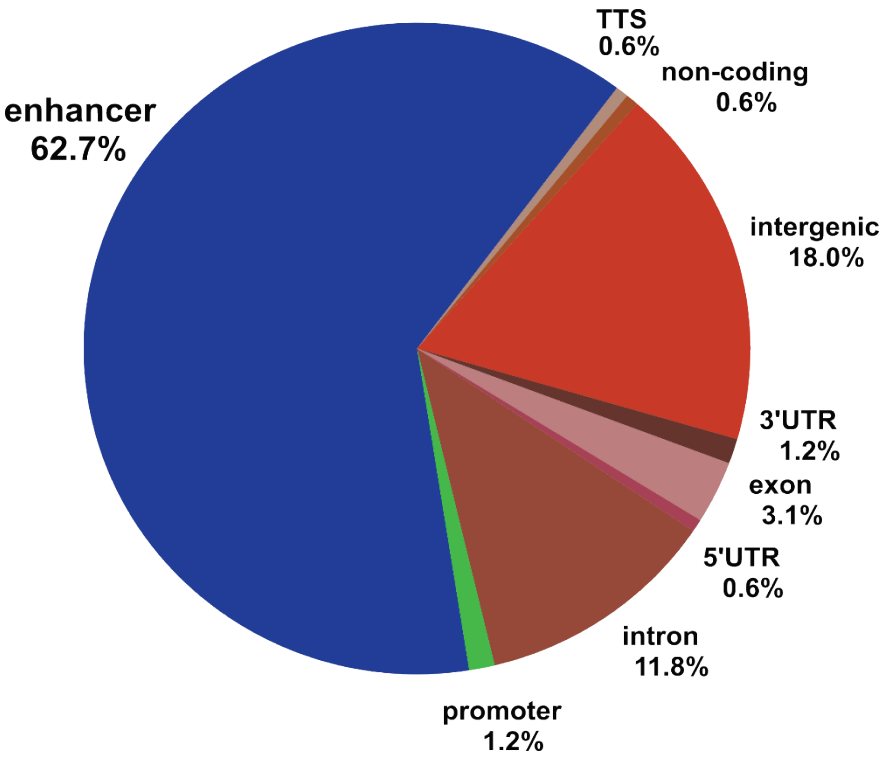


Figure 3.2.3: Localization of 161 DMRs among the genome.

Taken together, the 161 almost exclusively hypomethylated DMRs present in all three cohorts were distributed across all autosomes without any enrichment for particular chromosomes. However, DNA methylation changes preferentially occurred in enhancer regions.

3.2.4 Impact of the genotype on DNA methylation differences

DNA methylation can be strongly influenced by the genotype. A genetic variation, i.e. a SNP, which influences the methylation status of particular CpGs, is referred to as a methylation qualitative trait loci (meQTL) (146, 147). Changes in the DNA methylation patterns can also be shaped by genetic variants that either directly influence epigenetic enzymes or indirectly by affecting the binding of transcription factors (147, 148). To identify genotype associated DMRs, SNP calling was performed for each of the sequenced individuals using the bisulfite-conversion-aware SNP caller Bis-SNPs (125). SNPs within a range of 5 kb up- and downstream of the DMR, which were significantly (FDR 10%) correlated with DNA-methylation, were defined as meQTLs and corresponding DMRs defined accordingly as genotype-related DMRs (gDMRs). All DMRs not significantly associated with a meQTL-SNP and likely not influenced by the genotype were defined as non-genetically influenced DMRs (ngDMRs). For 24 out of the 161 DMRs a meQTL was found (gDMRs, 14.9%). The remaining 137 DMRs were not related to the genotype (ngDMRs, 85.1%). All SNPs, which were associated with a DNA

methylation change in this study are listed in supplemental Table 8.1.5.

Taken together, the majority of identified asthma-related DMRs was independent of the genotype (85.1% of the identified DMRs).

3.2.5 Pathway enrichment and transcription factor binding site motif analysis

The genomic sequences of the 161 DMRs were subjected to gene enrichment pathway analysis using the online browser software GREAT (129) (Figure 3.2.4A). The top enriched KEGG pathway was “Asthma” (*p*-value=8.34x10^{-6}), including the genes *interleukin 4* (*IL4)*, *eosinophil peroxidase* (*EPX)* and *interleukin 13* (*IL13)* related to four different DMRs, two of which associated to *EPX*. Further enriched pathways were related to immune response or inflammation such as adhesion, diapedesis of granulocytes, Wnt-signalling, *interleukin 5*-signaling pathway and *TNF receptor associated factor 3 (TRAF3)*-dependent *interferon regulatory factor (IRF)* activation pathway. For a detailed list of pathways and the related genes, see Table 3.2.2.

Table 3.2.2: Top enriched pathways and corresponding observed genes

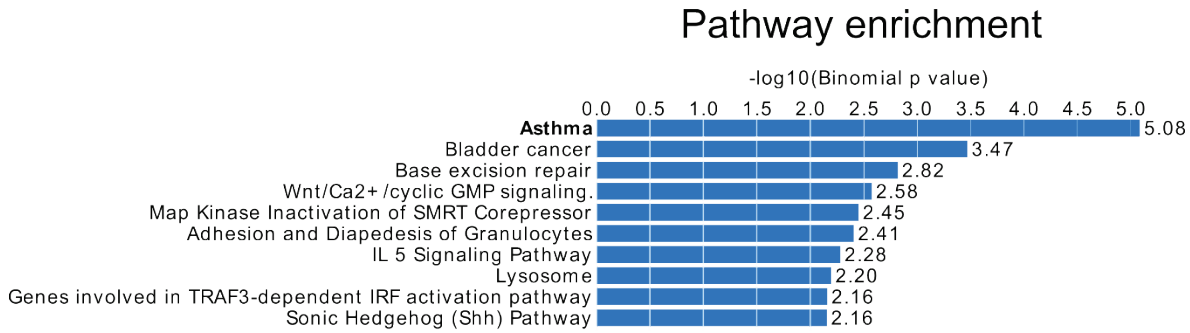
Pathway	*p* -value	Observed genes (total genes) in pathway	Number DMRs	Genes
Asthma	8.23E-06	3(28)	4	EPX,IL13,IL4
Bladder cancer	3.38E-04	4(42)	5	EGFR,RPS6KA5,THBS1,TYMP
Base excision repair	1.51E-03	3(33)	3	NTHL1,PARP4,POLE4
Wnt/Ca2+/cyclic GMP signaling.	2.65E-03	3(20)	3	ITPR1,NFAT5,PDE6A
Map Kinase Inactivation of SMRT Corepressor	3.52E-03	2(11)	3	EGFR,NCOR2
Adhesion and Diapedesis of Granulocytes	3.92E-03	2(13)	2	C5,ITGAL
IL 5 Signaling Pathway	5.22E-03	2(10)	2	IL4,IL5RA
Lysosome	6.36E-03	4(121)	5	AP3B1,CTSD,DNASE2B,LGMN
Genes involved in TRAF3-dependent IRF activation pathway	6.93E-03	2(14)	2	DDX58,TRAF3
Sonic Hedgehog (Shh) Pathway	6.99E-03	3(16)	3	DYRK1A,GLI2,PRKACB

To evaluate whether the 161 identified asthma-related DMRs may affect transcription factor binding sites (TFBS), a *de novo* binding-motif analysis using the MEME Suite sequence analysis tool (MEME-ChIP) was performed (127, 128). Within all DMRs, six transcription factor (TF) binding motifs were enriched (*p*-value<0.05) (Figure 3.2.4B), including motifs for *early growth response 1 (EGR1)*, *GATA binding protein 3 (GATA3)*, *nuclear factor of activated T cells 1 (NFATC1)*, *signal transducer and activator of transcription 2 (STAT2)*, *interferon regulatory factor 1 (IRF1)* and *thrombospondin 1 (THBS1)*. These TFs are known to regulate genes associated with lung development, lung function or have previously been associated with asthma (149-153) (Figure 3.2.4B). In this study, for example a putative binding site for *GATA3* (motif 3) was hypomethylated in asthmatic children and associated with the known *GATA3* target gene *peroxisome proliferator-activated receptor gamma, coactivator 1 beta*

(PPARGC1B) (154). This gene has previously been related to a risk increase for asthma development (155-157). Another DMR overlapping with the binding motif 3 was located in an enhancer region of the *suppressor of cytokine signalling 7 (SOCS7). SOCS7* is a known target gene of *GATA3/GATA2* and is involved in the JAK-STAT-signalling pathway important for the activation of inflammatory processes (158).

Furthermore, an enhancer DMR associated with *tumour necrosis factor (ligand) superfamily, member 10 (TNFSF10*) intersected with the observed binding motif 5. Although *TNFSF10* has not yet been identified as a target gene of the motif 5 related transcription factors, *TNFSF10* itself is a known regulator of acute allergic airways inflammation (159). For detailed information of the identified motifs and overlapping DMRs see supplemental Table 8.1.6.

A



B

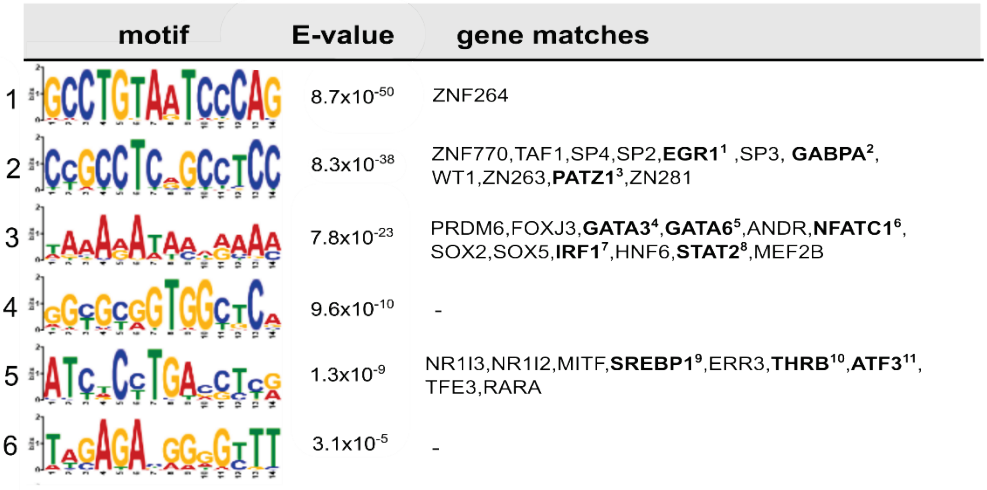


Figure 3.2.4: (A) Genomic Regions Enrichment of Annotation Tool (GREAT, university of Stanford (129)) reveals an enrichment of DMRs in the asthma pathway (KEGG Pathway database, -log10 transformed p-values are shown). **(B)** Transcription factor binding sites motifs overlapping with the 161 DMRs as determined by MEME-Suite (127, 128) software. Transcription factors associated with lung development, function or asthma are bold and index numbers referred to references in supplemental Table 8.1.7

Taken together, enrichment analysis showed that genes related to the identified DMRs were enriched in the KEGG “Asthma” pathway and further pathways related to immune response. In addition, DMRs were enriched in transcription factor binding motifs known to regulate genes previously associated with asthma and inflammation.

3.2.6 Validation of candidate DMRs in the entire LINA and LISA cohort

To validate the DMRs determined in the smaller WGBS cohort (n=82) in a larger sample set a targeted DNA methylation analysis was performed based on the MassARRAY platform (124). Candidate DMRs associated with genes, which were enriched in the asthma pathway and the closely linked *IL5* signalling pathway, were selected. The *IL5* signalling is an important asthma related pathway, as *IL5* is a major regulator for eosinophil differentiation, growth and immigration to the airways (160). Only samples from asthmatic children and healthy controls, who never had asthma, wheezing symptoms or obstructive bronchitis were included in the validation step. In total, n=127 samples of the LINA (age six and or eight) and n=142 of the LISA (age 15) cohort were available for this analysis. Target specific methylation analysis was performed for three DMRs associated to *EPX (chr17: 56272299-56272501)*, *IL5RA (chr3:3150228-3150424)* and *IL4 (chr5:132002374-132002506).*

The first validation analysis was conducted in the region associated to *EPX*. This particular region (chr17: 56272299-56272501) was located in the sixth exon of the *eosinophil peroxidase gene (EPX)* and consisted of 10 CpG sites. Overall, for nine of the ten CpGs covered by MassARRAY amplicon sufficient data quality was achieved. These methylation values were summarized as the mean methylation of the whole region. The DNA methylation difference of the *EPX*-related DMR observed by WGBS between asthmatic children and healthy controls was confirmed by the targeted MassARRAY analysis when comparing the same individuals from the WGBS-cohort and also in the entire LINA and LISA cohort using a logistic regression model adjusted for gender and family history of atopy (Table 3.2.3).

Table 3.2.3: Validation of DMRs by MassARRAY in the LINA and LISA cohort.

			LINA			LISA			PASTURE/EFRAIM		
Gene	**method**	**CpGs**	**A/C[a] N**	**ΔM [ß-value][b]**	***p*-value**	**A/C N**	**ΔM [ß-value]**	**p-value**	**A/C N**	**ΔM [ß-value]**	**p-value**
EPX	WGBS	10	12/13	0.04	**<0.001**	14/15	0.07	**<0.001**	14/14	0.04	**<0.001**
	MassARRAY	9	19/108	0.03	**0.04[c]**	25/117	0.04	**0.02[c]**		-	
IL5RA	WGBS	9	12/13	0.06	**<0.001**	14/15	0.06	**<0.001**	14/14	0.04	**<0.001**
	MassARRAY	8	19/108	0.02	**0.006[c]**	25/117	0.02	**0.04[c]**		-	
IL4	WGBS	5	12/13	0.05	**<0.001**	14/15	0.06	**<0.001**	14/14	0.05	**<0.001**
	MassARRAY	7/3	19/108	0.03	**0.02[c]**	25/117	0.02	**0.02[c]**		-	

[a] A=asthma; C=control [b]ΔM= Δmethylation [c]logistic regression model adjusted for gender, family history of atopy *EPX* (refers to DMR No.62) *IL5RA* (refers to DMR No.115) *IL4* (refers to DMR No.127 in supplemental Table 8.1.4)

A further validated region was located in the second exon of the *interleukin 5 receptor α (IL5RA)* gene (chr3:3150228-3150424), 1731bp downstream to transcription start site (TSS). The DMR identified in the WGBS samples spanned nine CpG sites and could be covered by MassARRAY amplicon. For eight of nine CpGs sufficient data quality was achieved in the MassARRAY. Again, the methylation change observed in the WGBS samples could be reproduced by this independent method with a significant DNA methylation difference between asthmatic children and controls in both LINA (*p*-value<0.001) and LISA (*p*-value<0.001). Logistic regression models adjusted for gender and family history of atopy for the entire LINA and LISA cohort confirmed a significant hypomethylation of the *IL5RA* region (Table 3.2.3).

The third DMR selected for validation was intergenically located on chromosome 5 (chr5:132002374-132002506), upstream of *the interleukin 4 (IL4)* promoter. The WGBS DMR included five CpGs, while the MassARRAY amplicon included seven CpGs, three of which overlapped with the WGBS region. ChiA-PET data depicted in the UCSC browser (https://genome.ucsc.edu/) showed an interaction of this region with the *IL4* promoter (Figure 3.2.6A), which makes this region appear as a probable enhancer for *IL4*. Validation by MassARRAY again confirmed the significant asthma-related hypomethylation of the *IL4* enhancer region not only in the same individuals of the WGBS study but also in both entire LINA and LISA sample sets. The adjusted logistic regression analyses showed a significant relationship between the DNA methylation difference and the asthma phenotype in the entire LINA (*p*-value=0.02) and LISA (*p*-value=0.02) cohort, respectively (Table 3.2.3).

In summary, the results obtained by methylation sequencing in 82 samples could be replicated for selected regions in a larger sample set of the LINA (n=127) and the LISA cohort (n=142) using a mass spectrometry based targeted approach.

3.2.7 Functional validation of DMRs

To elucidate whether the differential DNA-methylation affects gene expression, the mRNA expression of *IL5RA*, *EPX* and *IL4* was analysed by qPCR.

A significant difference of *IL5RA* mRNA expression was identified in asthmatics compared to healthy controls in LINA (*p*-value<0.05) and LISA (*p*-value<0.01) (Figure 3.2.5B). By combining the available expression data of both cohorts a significant correlation between *IL5RA* methylation and mRNA expression was observed (r=-0.35 and *p*-value=$3x10^{-15}$) (Figure 3.2.5C). Applying a multiple regression model adjusted for gender, parental history of atopy and study cohort demonstrated that both *IL5RA* DNA hypomethylation and asthma affect the *IL5RA* mRNA expression. (Figure 3.2.5D).

Similarly expression of *EPX* showed a significant increase in asthmatics compared to the healthy controls in the LISA cohort (*p*-value<0.05) but not in the LINA cohort (*p*-value>0.05) (Figure 3.2.5F). However, when combining both datasets a decreased DNA methylation in the *EPX* related DMR was significantly correlated with an increased *EPX* mRNA expression (r=-0.41 and *p*-value=$4.5x10^{-16}$; Figure 3.2.5G). Also the adjusted multiple regression analysis (adjusted for gender, family history of atopy and study cohort) revealed a significant association of DNA methylation and asthma with *EPX* mRNA expression (Figure 3.2.5H).

Taken together, these data indicate a functional translation of the asthma-related methylation change associated to *IL5RA* and *EPX* to the transcriptional level. Both genes showed a significantly different mRNA expression in asthmatic children compared to healthy controls correlating with the DNA methylation status both in the LINA and the LISA cohort.

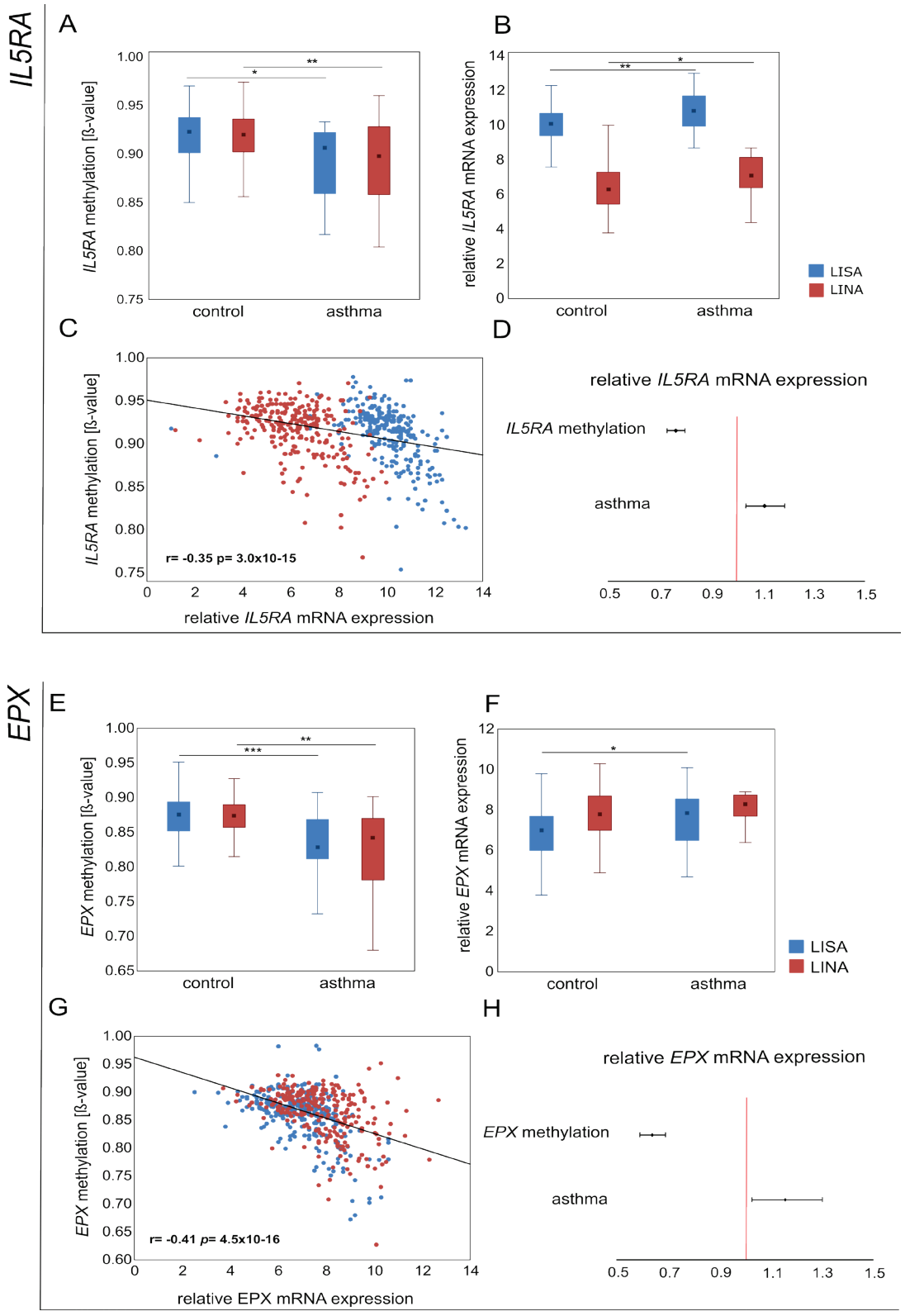


Figure 3.2.5: Man-Whitney-U-Test showed a significant difference between asthmatic children and healthy controls in LINA 6- to 8-year-old children (red) and LISA 15 year-old-children (blue) for *IL5RA* **(A)** *EPX* **(E)** methylation (MassARRAY) and *IL5RA* **(B)** and *EPX* **(F)** transcription (*p*-value < 0.05*, *p*-value < 0.01**, *p*-value < 0.001***, *p*-value < 0.0001****). **(C)** Spearman correlation of LINA and LISA children together determined association between *IL5RA* and *EPX* **(G)** methylation and RNA expression. **(D)** Multiple regression models was applied to demonstrate the effect of *IL5RA* methylation and asthma outcome on relative *IL5RA* and *EPX* **(H)** mRNA expression. Given are log transformed effect sizes with upper and lower 95% CIs from multiple regression model adjusted for gender, parental history of atopy and study cohort.

The DMR associated with *IL4* was located in a putative enhancer region as a ChiA-PET interaction to the promoter was observed in the cancer blood cell line K-562 (Figure 3.2.6A). Hypomethylation of this enhancer region was negatively correlated with the *IL4* mRNA expression (p<0.0001 R=-0.30) (Figure 3.2.6B). Available blood protein concentrations of IL4 from six-year-old children of the LINA cohort positively correlated with the *IL4* gene expression (Spearman: p-value<0.001 R=0.28) (Figure 3.2.5B). In addition, a mediation analysis was applied to elucidate the relationship between DNA methylation, transcription, and IL4 protein concentration in six-year-old children of the LINA cohort (n=220). The mediation model, adjusted for gender and family history of atopy, showed no direct effect of DNA methylation on the protein concentration, but rather an indirect effect of the *IL4* enhancer methylation on IL4-protein mediated by *IL4* mRNA expression (Figure 3.2.6C).

In summary, these data show a functional translation of the perturbed *IL4* enhancer methylation in asthmatic children into an altered gene and protein expression. Hypomethylation of the *IL4* enhancer region was significantly associated with an increased *IL4* mRNA expression and enhanced IL4 protein levels. A mediation analysis revealed an indirect effect of DNA methylation on the protein concentration transmitted by *IL4* mRNA expression.

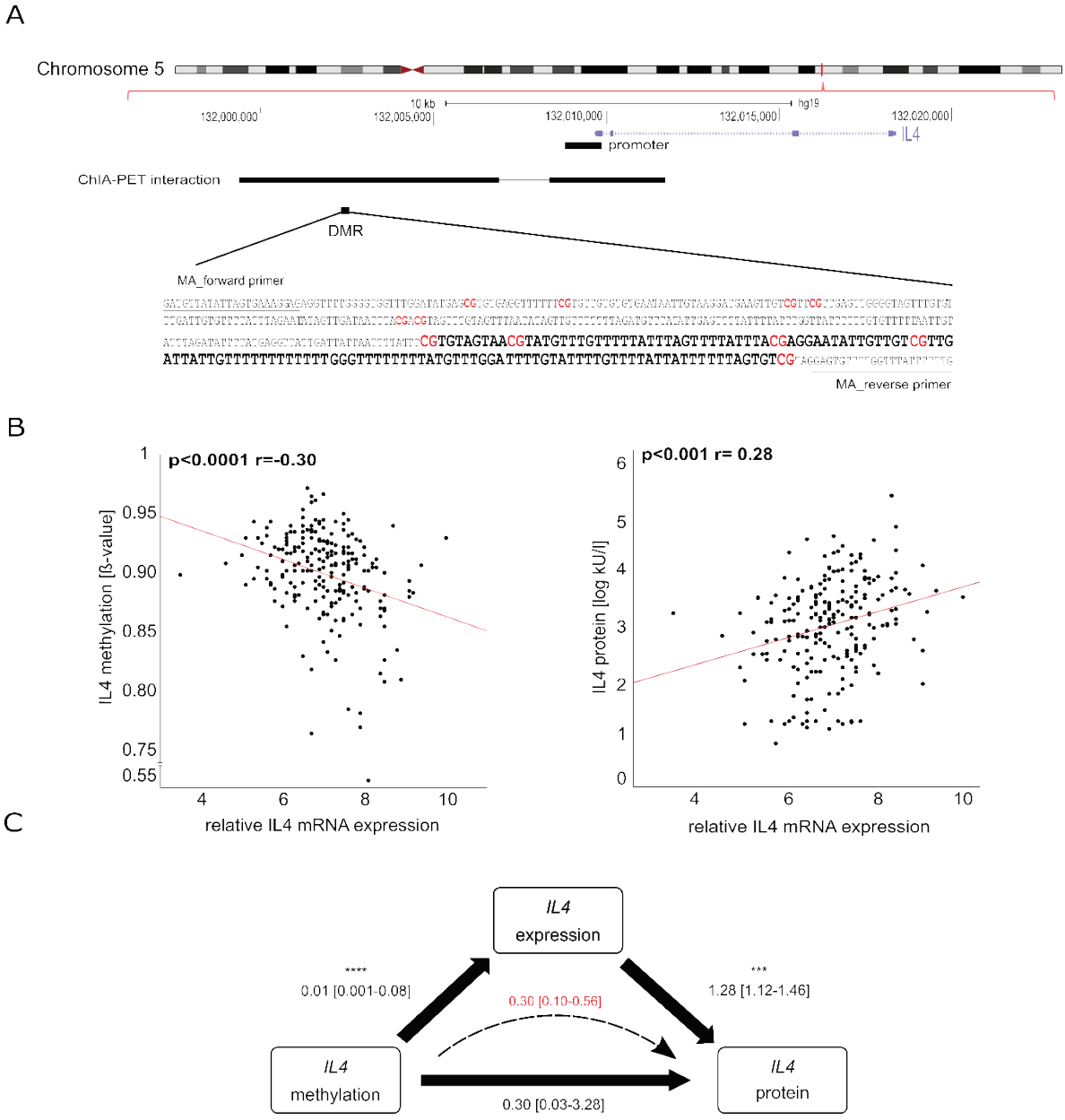


Figure 3.2.6: (A) Genomic location of the *IL4* associated DMR (bold nucleotides) and the MassARRAY amplicon used for targeted validation. **(B)** Spearman correlation of methylation, transcription and protein level of *IL4* gene in six year-old children of the LINA cohort (n=220) **(C)** Mediation analysis for the relationship between *IL4* DNA-methylation, transcription, and protein expression of six-year-old children of the LINA cohort (n=220). Models were adjusted for gender and family history of atopy. The effect sizes for each path of the mediation analysis are given as unstandardized B-values with lower and upper confidence intervals (P < 0.05*, P < 0.01**, P < 0.001***, P< 0.0001****)

3.2.8 Network analysis of DMR associated genes

As described above, an enrichment of enhancer regions was observed among the asthma specific DMRs. More than half of the DMRs (63% n=101 DMRs) were intersecting with enhancer elements. Based on the predicted target genes of these enhancer regions and the

genes closest to the next TSS for all other DMRs (if not an enhancer region) a network analysis was performed using Cytoscape version 3.6.1 (131, 132). Cytoscape, considers functional, genetic, physical and regulatory interactions of defined gene sets based on different databases. Such interactions are visualized in a network of modules represented by interacting genes, which are involved in the same biological pathways. Based on the 311 genes, used as input for this analysis 11 modules were identified. These modules included 47 genes connected to at least two other genes (Figure 3.2.7). The eleven identified modules, were subjected to a pathway enrichment analysis with a FDR <0.001. Module 1 was related to immune response and inflammation including genes of the NF-kappa B pathway or tumour necrosis factor (TNF) alpha signalling. Cellular stress response was reflected in module 4 containing genes among others belonging to the mTOR-signalling pathway. Furthermore JAK-STAT signalling was enriched in module 7 with *IL4, IL5RA* and *interleukin 17 subunit D (IL17D*). In addition, modules related to epigenetic modifications (module 6, HDAC signalling) and genomic regulation (module 8/9, DNA damage control and RNA degradation) were found. For detailed information on the eleven modules, see supplemental Table 8.1.9.

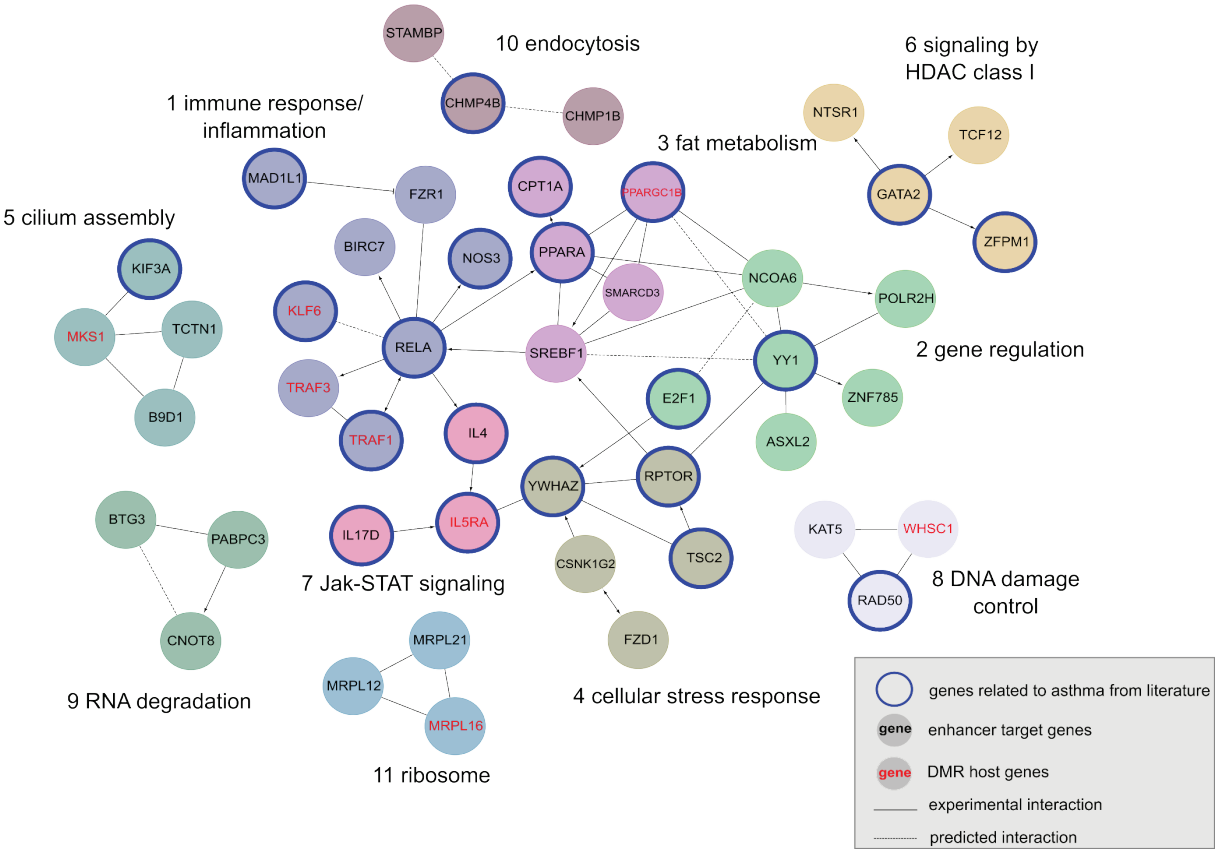


Figure 3.2.7: Network module analysis based on enhancer target genes (black font) and DMR host genes (red font). Only modules with at least two connecting genes are shown with either experimental (solid line) or predicted interaction (dashed line). Genes previously associated with asthma were determined by natural language processing of abstracts and full texts available in PubMed or PMC (highlighted by blue outline circle).

To assess whether the genes of the identified network (Figure 3.2.7) have indeed been previously shown to be related to asthma, a comprehensive literature search of the Pubmed and PMC database was performed based on natural language processing (140-143). This analysis has been conducted by collaborators at University of Jena (see also Chapter 2 section 2.13). In the available literature all co-occurrences of either the genes included in the network (n=47) with one of the following terms: asthma, asthmatic, asthmatics, wheeze, bronchial hyperreactivity, airway hyperreactivity, bronchial hyperresponsiveness, and hyperreactive airway disease were determined. In total, 44.7% (n=21) of the genes of the eleven modules were associated with an asthma related disease term (Figure 3.2.7 labelled with blue circle, and for according publication see supplemental Table 8.1.8). The vast majority of modules included genes with a confirmed relationship to asthma based on natural language processing (n=9 out of 11).

Taken together, DMRs identified in asthmatic children are related to a network of genes involved in immune response and inflammation. A major part of these genes has already been described in the literature in association with asthma. The strong overlap of current literature knowledge with DMR-related genes might suggest that also other genes of the network like *SREBF1*, *NCOA6* or *KAT5* so far not described as associated with asthma, could play a role in asthma aetiology.

The presented study showed clear differences in DNA methylation patterns of blood cells in children with asthma compared to controls, particularly in enhancer regions. In addition to the identification of DNA methylation changes of known genes related to asthma, including *IL4, EPX* and *IL5RA*, previously unknown regions of DNA methylation have been identified.

3.3 Longitudinal DNA methylation pattern in asthma

In the previous chapter, 161 differentially methylated regions (DMRs) between asthmatic children and healthy controls were described. The following chapter explores, whether this aberrant DNA methylation pattern observed in children with an established asthmatic phenotype was already present at the time of birth.

To this end DNA methylation was assessed by WGBS in cord blood samples of children who developed asthma later in their life compared to healthy controls. Only children with WGBS DNA methylation data in the asthma phenotype (see chapter 3.2) were considered for this longitudinal study. In total, 48 cord blood samples from the LINA (n=20) and the PASTURE/EFRAIM (n=28) cohort were selected (Table 3.3.1).

Table 3.3.1: Study characteristics of available cord blood samples used for whole genome bisulfite sequencing.

	LINA N = 20 n(%)	**PASTURE/EFRAIM** N = 28 n(%)	**entire WGBS cohort birth** N = 48 n(%)
Age			
	birth	birth	birth
Gender			
Female	11(55)	5(17.9)	16(33.3)
Male	9(45)	23(82.1)	32(66.7)
Phenotype			
Control	11(55)	14(50)	25(47.9)
Asthma	9(45)	14(50)	23(52.1)

To investigate the longitudinal stability of asthma-related methylation changes, two independent approaches (top-down and bottom-up approach) were applied. The two approaches and the corresponding results are explained in more detail in the following sections.

Further, this chapter elucidates, whether the presumably stable asthma -related DMRs can be related to genetic variation or can be linked to the environmental exposure such as prenatal tobacco smoke exposure, maternal stress, maternal depression, air pollution, viral infections, maternal asthma, and maternal vitamin D intake during pregnancy.

3.3.1 Top-down approach

For the top-down approach it was tested whether methylation changes observed in established disease (161 DMRs) can already be found in cord blood. A factorial ANOVA, multiple-test

corrected by Bonferroni (p<0.00031), was conducted separately for the LINA and PASTURE/EFRAIM cohorts (Figure 3.3.1).

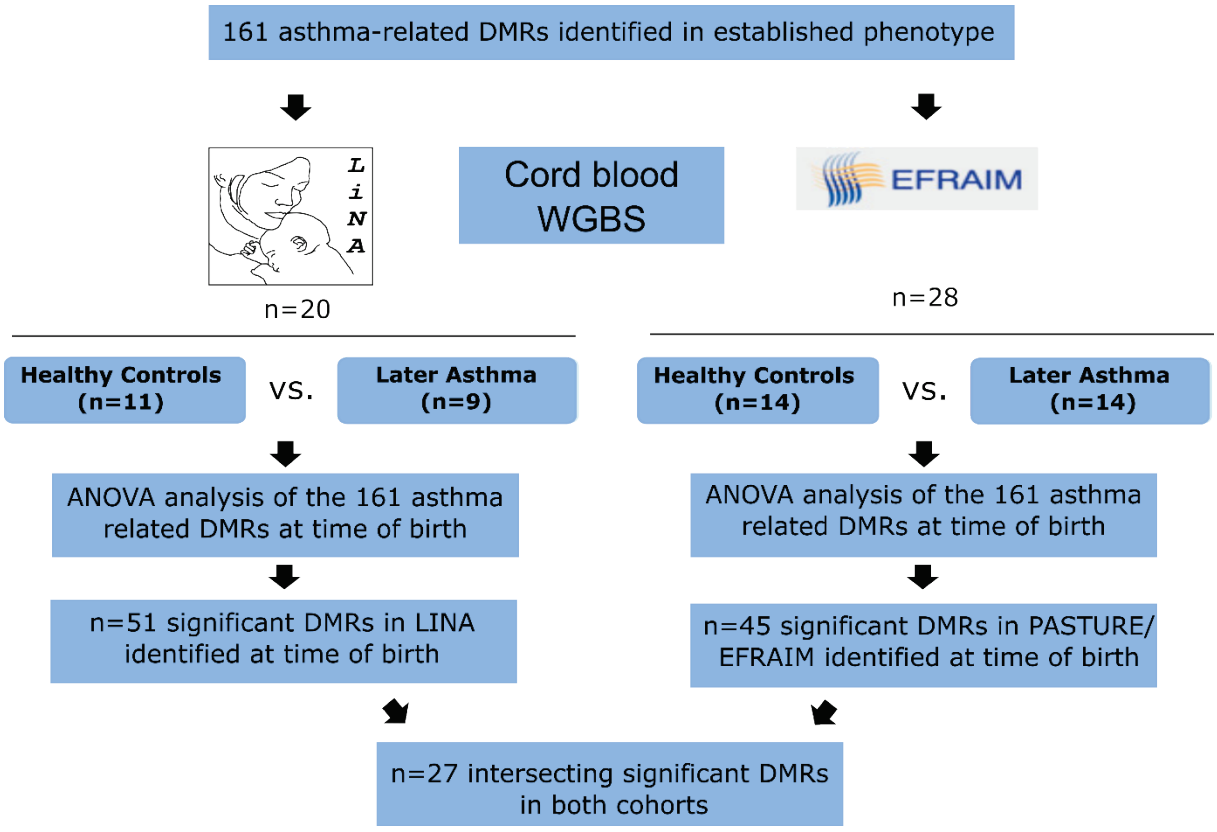


Figure 3.3.1: Workflow of the top down approach for the identification of DMRs at the time of birth based on the 161 asthma-related DMRs identified in the established asthma phenotype.

Overall, out of the 161 asthma-related DMRs, 51 regions in LINA and 45 regions in PASTURE were significantly different between asthmatic children compared to healthy controls at the time of birth (Bonferroni multiple test corrected p-value<0.00031). Of these 27 DMRs were independently identified in both cohorts indicating persistent DNA hypomethylation from birth until asthma development (Table 3.3.2). The majority of these 27 intersecting DMRs were not associated to the genotype (n=22 ngDMRs; 81.5%) and were located in enhancer regions (n=19; 70.3%). Thus, only five (18.5%) DMRs were linked to the genotype (=gDMRs) (Table 3.3.2).

Table 3.3.2: Overlap of differentially methylated regions (DMRs) in cord blood and the asthma phenotype based on the DMRs identified within the asthma phenotype and a subsequent ANOVA analysis of the methylation level of these regions in the cord blood WGBS data.

Chr	Start	End	Gene	Predicted Target Gene	ΔMeth. birth [ß-value]	ΔMeth. asthma phenotype [ß-value]	Enhancer	ng DMR
10	1404992	1406101	ADARB2-AS1	LARP4B;GTPBP4;IDI1;WDR37;ADARB2	-0.023	-0.051	-	YES
9	69500968	69501069	ANKRD20A4	ANKRD20A4	-0.058	-0.053	-	YES
4	170695782	170696267	C4orf27	CLCN3;MFAP3L	-0.006	-0.045	YES	YES
2	31154684	31155156	CAPN13	CAPN13	-0.098	-0.127	YES	YES
2	118617605	118618119	DDX18	DDX18	-0.058	-0.063	-	YES
3	75445436	75445698	FAM86DP	FAM86DP	-0.051	-0.065	YES	-
9	38487906	38488164	FAM95C	FAM95C	-0.013	-0.057	-	YES
7	90895326	90896701	FZD1	FZD1	-0.040	-0.042	YES	YES
1	6341136	6341791	GPR153	ICMT;HES3;ACOT7	-0.040	-0.056	YES	YES
22	50985261	50985924	KLHDC7B	TYMP;SCO2;ODF3B;CHKB-AS1;CHKB;CHKB-CPT1B	-0.037	-0.055	YES	YES
13	24914323	24914904	LINC00566	MIPEP;SPATA13;PARP4	-0.059	-0.061	YES	YES
8	58192499	58193337	LINC00588	LINC00588	-0.042	-0.067	-	-
20	29550781	29551738	LINC01598	LINC01598	-0.037	-0.045	-	YES
8	2585088	2586090	LOC101927815	LOC101927815	-0.035	-0.039	-	YES
7	36007074	36007281	LOC101928618	HERPUD2;ANLN	-0.068	-0.085	YES	YES
5	77146478	77147360	LOC101929154	AP3B1	-0.067	-0.118	YES	-
5	77142381	77142898	LOC101929154	OTP;AP3B1	-0.030	-0.029	YES	-
2	24233608	24234116	MFSD2B	UBXN2A;ATAD2B;FAM228B;SF3B14;FKBP1B;C2orf44;TP53I3;MFSD2B	-0.063	-0.082	YES	YES
13	20968573	20969084	MIR4499	GJA3;GJB6;CRYL1;IL17D	-0.068	-0.074	YES	YES
7	1914009	1914392	MIR4655	MAD1L1	-0.049	-0.042	YES	-
16	88540019	88540525	MIR5189	PIEZO1	-0.064	-0.106	YES	YES
16	88579452	88580071	MIR5189	ZFPM1	-0.029	-0.038	YES	YES
11	68517647	68517834	MTL5	GAL;C11orf24;TPCN2;CPT1A;GAL;MRPL21;IGHMBP2	-0.096	-0.084	YES	YES
9	135114516	135114648	NTNG2	C9orf171	-0.097	-0.078	YES	YES
11	132951692	132952491	OPCML	OPCML	-0.114	-0.080	YES	YES
17	19627951	19628165	SLC47A2	ALDH3A2;ALDH3A1;B9D1;SPECC1	-0.044	-0.069	YES	YES
2	121816270	121816884	TFCP2L1	GLI2	-0.003	-0.072	-	YES

Notably, seven target genes of the stable 19 enhancer DMRs, such as *Zinc Finger Protein, FOG Family Member 1 (ZFPM1)* (161), *ACOT7* (76, 162), *chloride voltage-gated channel 3 (CLCN3)* (163), *FKBP prolyl isomerase 1B (FKBP1B)* (164), *aldehyde dehydrogenase 3 family member A1 (ALDH3A1)* (165), *interleukin 17D (IL17D)* (166) and *Mitotic Arrest Deficient 1 Like 1 (MAD1L1)* (167) have previously been described in relation to asthma using natural language processing as described in chapter 3.2.8 .

Taken together, the aberrant DNA-methylation pattern in children with an established asthma phenotype can at least partially be seen already in cord blood. A subset of 27 out of the 161 asthma-related DMRs was observed already at the time of birth in two independent cohorts.

3.3.2 Bottom-up approach

For a comprehensive detection of longitudinally stable asthma-related DMRs a second approach (bottom-up) not restricted to the 161 DMRs identified in the disease phenotype, was applied. Based on cord blood WGBS data of 48 children of the LINA and PASTURE/EFRAIM cohort, DMRs were determined by comparing children known to develop asthma (n=23) to healthy controls (n=25). The same settings as described in chapter 3.2 were used for DMR calling (Figure 3.3.2).

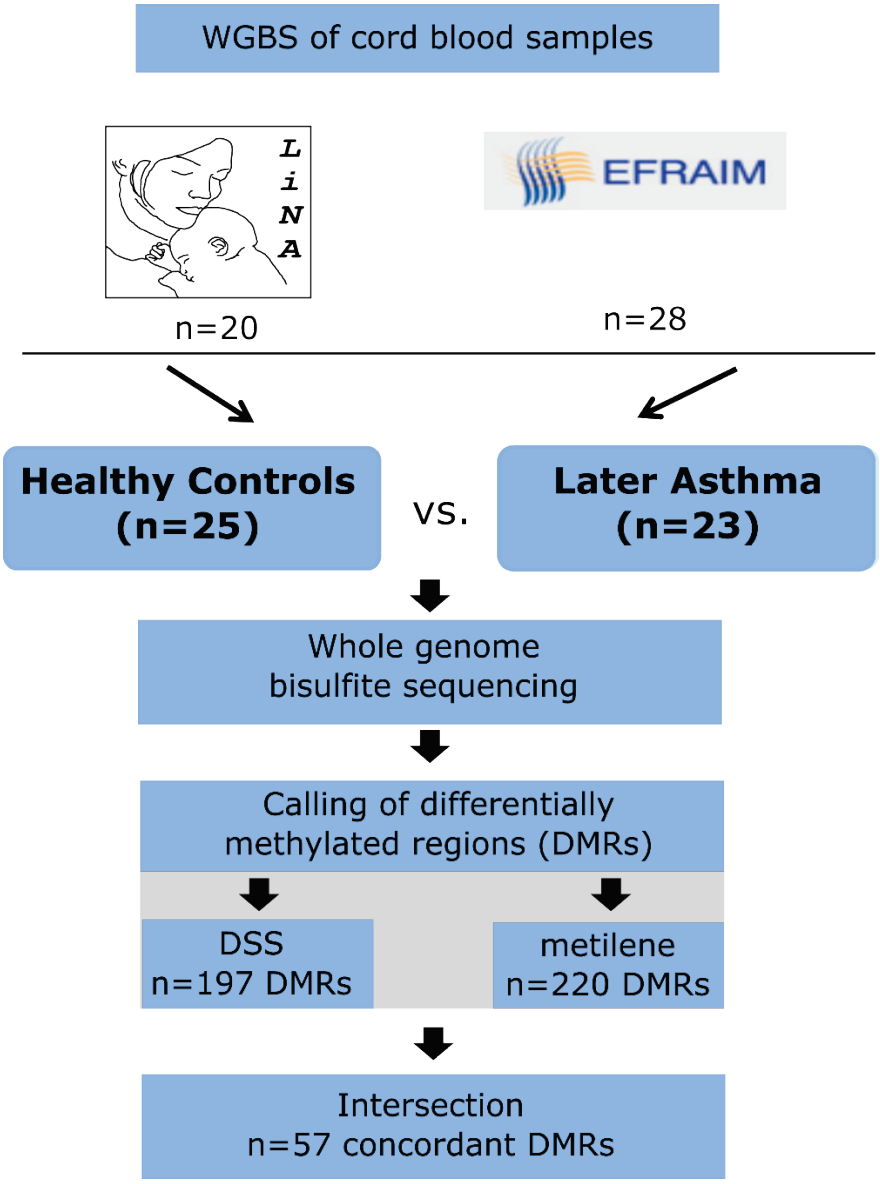


Figure 3.3.2: Workflow of the bottom-up approach for the determination of asthma-related DMRs at the time of birth based on cord blood WGBS data comparing 23 children who later develop asthma to 25 healthy controls.

DSS and metilene identified 197 and 220 DMRs, respectively. In order to obtain statistically reliable results, the intersection between DMRs found by the two DMR calling approaches independently was determined and consisted of 57 concordant DMRs, including 16 hyper- and 41 hypomethylated regions in the asthmatic children (Figure 3.3.2 and in the supplements Table 8.1.10). Nearly all chromosomes were affected by methylation changes except chromosome 6, 14 and 15 (Figure 3.3.3A). Out of the 57 DMRs, 35 DMRs (61.4%) intersected with enhancer regions (Figure 3.3.3B) and 35 (61.4%) DMRs were genotype associated.

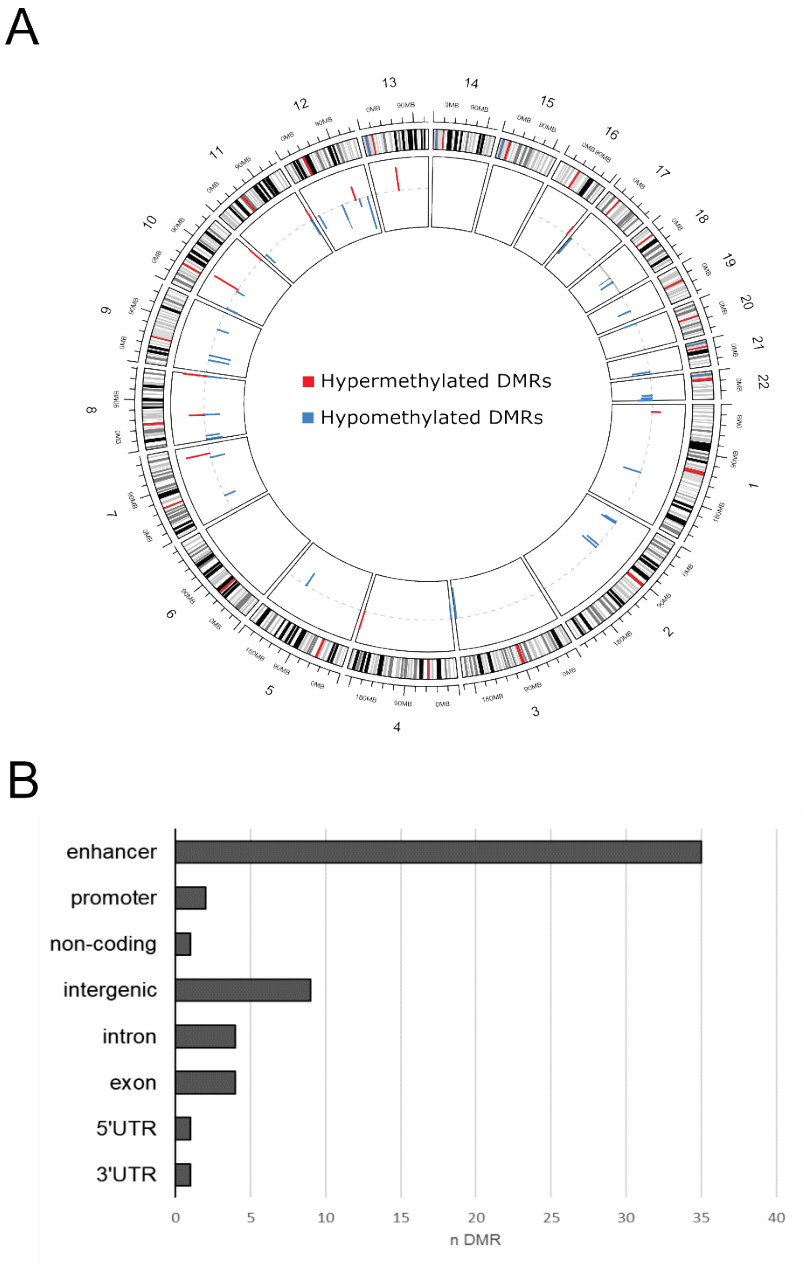


Figure 3.3.3: (A) Genome-wide distribution of the 57 DMRs identified by DSS/metilene based on cord blood WGBS data in asthmatic children compared to healthy controls. Most of the 57 DMRs were hypomethylated in asthmatic children. **(B)** Localization of the 57 DMRs in the genome.

Taken together, by the bottom-up approach 57 DMRs were identified at the time of birth. The majority of the DMRs were located in enhancer regions and 22 (38.6%) DMRs were not associated with the genotype.

3.3.3 Intersection of bottom-up and top-down approach

In order to determine reliable, longitudinally stable regions of aberrant DNA methylation pattern that are related to asthma, the 27 DMRs identified using the top-down (one factorial ANOVA) and the 57 DMRs of the bottom-up (DMR calling) approach were compared. In total, six DMRs intersected with at least 1 bp between the two approaches (Table 3.3.3).

Table 3.3.3: Stable asthma-related DMRs (annotated by phenotype regions) determined by two independent approaches.

Chr	Start	End	Direction	Mean methylation [ß-value] Asthma	Mean methylation [ß-value] Controls	∆Methy-lation [ß-value]	Gene	Predicted Targetgene	Enhancer	ngDMR
10	1404992	1406101	hypo	0.40	0.48	-0.08	ADARB2-AS1	LARP4B;GTPBP4;IDI1;WDR37;ADARB2	-	YES
11	132951692	132952491	hypo	0.48	0.59	-0.11	OPCML	OPCML	YES	YES
22	50985261	50985924	hypo	0.57	0.67	-0.10	KLHDC7B	TYMP;SCO2;ODF3B;CHKB-AS1;CHKB;CHKB-CPT1B	YES	YES
7	36007074	36007281	hypo	0.36	0.45	-0.09	LOC101928618	HERPUD2;ANLN	YES	YES
8	2585088	2586090	hypo	0.14	0.20	-0.06	LOC101927815	LOC101927815	-	YES
8	58192499	58193337	hypo	0.50	0.58	-0.08	LINC00588	LINC00588	-	-

All six DMRs showed a significant differentially methylation pattern between asthmatic children and healthy controls at the time of birth and were stable until phenotype development. Five of these six DMRs (83%) were not associated to the genotype and three DMRs (50%) were located in enhancer regions. Since the majority of these DMRs were not associated with the genotype, the next section (3.3.4) evaluates which other prenatal environmental factors - such as prenatal tobacco smoke exposure - might contribute to the altered DNA-methylation pattern.

3.3.4 Impact of environmental factors on longitudinal stable asthma-related DMRs

To test, whether the longitudinally stable cord blood DMRs (five ngDMRs excluding the gDMRs) were influenced by environmental factors a literature search for potential external factors (e.g. stress, depression, air pollution viral infections or prenatal maternal smoke exposure) associated with these regions was performed (for a detailed description of search criteria please refer to section 2.2.14).

In total, 36 studies were found reporting DNA methylation changes in newborns (birth-one month of age) or infants (birth-23 months of age) in relation to relevant prenatal factors and were compared with the five longitudinally stable ngDMRs (Table 3.3.4). Among them, four studies were conducted in the LINA cohort and evaluated the influence of a prenatal tobacco smoke exposure (106), maternal vitamin D intake (168), prenatal-maternal stress exposure (38) and bisphenol a (BPA) exposure (169) during pregnancy on cord blood's DNA methylation pattern. In addition, for a complete overview of the overlap results of DNA methylation studies related to environmental exposures and the 27 and 57 DMRs identified with the top-down and bottom up approach, respectively, see supplemental table (Table 8.1.11 and Table 8.1.12).

Table 3.3.4: Overview of DNA methylation studies related to environmental exposures and the five longitudinally stable ngDMRs.

Environmental factor	Epigenetic studies	Approach	Intersected regions
Prenatal maternal smoke exposure	Bauer *et al(106)****	WGBS	chr11:132951692-132952491 (*OPCML*)
	Joubert *et al.(170)*		
	Miyake *et al.(171)*		
	deVocht *et al.(172)*		
	Küpers *et al.(173)*		
	Lee *et al.(174)*	450K	---
	Markunas *et al.(175)*		
	Reese *et al.(176)*		
	Xu *et al.(177)*		
Maternal stress	Vangeel *et al.(178)*	450K	---
	Trump *et al.(38)****	WGBS	
	Kertes *et al.(179)*	450K	
Maternal depression	Viuff *et al.(180)*	450K	---
	Non *et al.(181)*		
	Cardenas *et al.(182)*		
Child´s ADHD	Walton *et al.(183)*	450K	---
Air pollution during pregnancy	Gruzieva *et al.(184)*	450K	---
	Peng *et al.(185)*		
	Ladd-Acosta *et*		

	al.(186)		
Maternal asthma	deVries *et al.(187)*	450K	---
Prenatal maternal arsenic exposure	Bozack *et al.(188)* Kaushal *et al.(189)* Rojas *et al.(190)*	450K	---
Prenatal maternal mercury and arsenic exposure	Cardenas *et al.(191)*	450K	---
Prenatal maternal lead exposure	Sen *et al.(192)* Wu *et al.(193)*	450K	---
Prenatal maternal cadmium exposure	Kippler *et al.(194)*	450K	---
Maternal gestational diabetes mellitus	Howe *et al.(195)*	450K	---
Preterm birth	Goede *et al.(196)* Wu *et al.(197)*	450K	---
Prenatal maternal phthalate exposure	Solomon *et al.(198)*	450K	---
Prenatal maternal BPA exposure	Junge *et al.(169)****	450K	---
Birthweight of neonates	Küpers *et al.(199)*	450K	---
Maternal BMI during pregnancy	Sharp *et al.(200)*	450K	---
Prenatal DHA supplementation	van Dijk *et al.(201)*	450K	---
Vitamin D intake during pregnancy	Junge *et al.(168)*	WGBS	---

***studies in the LINA cohort, WGBS= whole genome bisulfite sequencing BeadChip 450K= Illumina HumanMethylation450 BeadChip; ADHD=attention deficit hyperactivity disorder; BPA= Bisphenol A; DHA=docosahexaenoic acid

Among the environmental associated DMRs or DMPs identified by using array-based approaches, no intersection to the five stable asthma-related ngDMRs was observed. However, one of the stable asthma-related DMRs overlapped with prenatal tobacco smoke related DNA methylation changes previously identified by WGBS (106). This regions within the *Opioid Binding Protein/Cell Adhesion Molecule Like (OPCML)* gene is located in a regulatory region. Figure 3.3.4 shows the stable DNA methylation profile of this particular region in asthma children and healthy controls in the two cohorts at the time of birth and the phenotype time point.

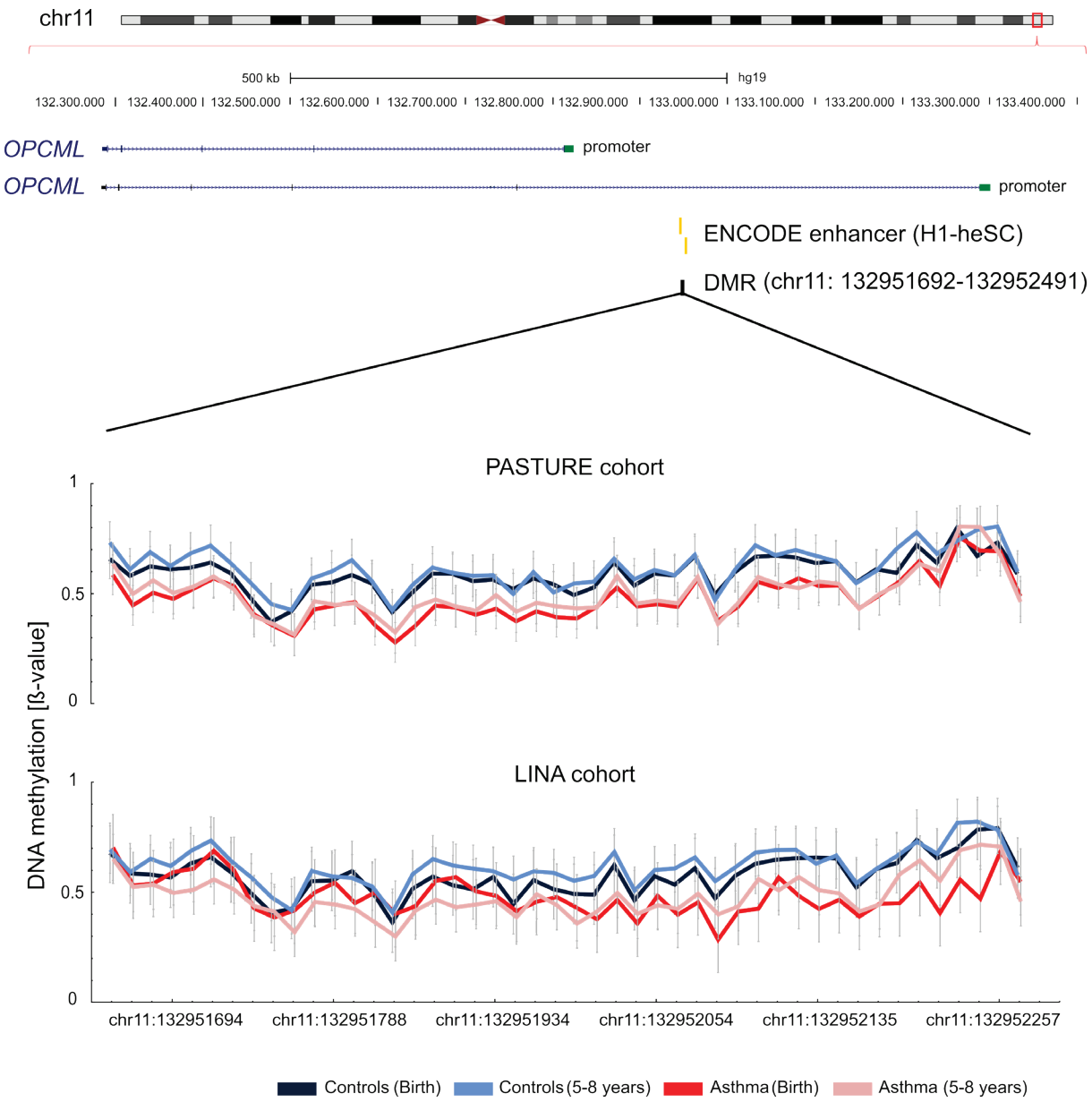


Figure 3.3.4: Longitudinal comparison of the methylation pattern in the *OPCML* DMR between healthy and asthmatic children at the time of birth and age 5-8 years in PASTURE and LINA cohorts (bottom panels). Adapted from the UCSC browser (top panel). The enhancer definition based on the H1-HeSC Chromatin state Segmentation by HMM track from ENCODE/Broad (202, 203). Shown are mean DNA methylation *(ß-value)* and +/- SEM for each CpG in the DMR.

The presented study demonstrated that known DNA methylation pattern for asthma disease were present long time before the disease onset in cord blood of children. One of these DMRs located in the *OPCML* gene has previously been associated to prenatal tobacco smoke exposure.

4 DISCUSSION

4.1 Longitudinal trends of serum IgE and asthma

The relationship between early sensitization represented by total (tIgE) or specific IgE (sIgE) and the risk for the development of wheeze or asthma in early childhood remains incompletely understood. In particular, whether the tIgE/sIgE-pattern could help to identify wheezing children who are at risk for a subsequent asthma development has not been sufficiently studied. In addition to IgE, *IL5RA* (interleukin 5 receptor subunit alpha) is an important mediator between IgE and asthma and thus could potentially be used as an additional risk indicator for asthma development.

Diagnosis of asthma in early childhood is difficult and often wheezing, in particular persistent wheeze, is used as a surrogate for asthma in young children (204, 205). However, not all children who show wheezing symptoms subsequently develop asthma (90). Efficient treatment of childhood asthma would strongly benefit from early indicators helping to identify wheezing children at risk for asthma manifestation. Therefore, one aim of this thesis was to address the role of IgE and *IL5RA* as an early marker for asthma development by using longitudinal data of tIgE, sIgE and *IL5RA* expression of the prospective mother-child cohorts LINA and LISA.

In the current study longitudinal sIgE/tIgE profiles were evaluated in children with different wheezing endotypes (transient, late-onset and persistent), as previously defined by Martinez et al. (95), and compared to those of children diagnosed with asthma or to healthy controls respectively. This approach provided the opportunity to identify early indicators that might aid differentiating childhood asthma from childhood wheezing.

The present study showed that the longitudinal patterns of total and specific IgE against aeroallergens were able to discriminate children with a subsequent asthma diagnosis from healthy controls, but more importantly also from the different wheezing endotypes, including persistent wheeze. Differences in the longitudinal course of specific and total IgE between persistently wheezing children and asthmatic children were observed in the LINA cohort for children up to the age of eight years and could be independently confirmed in the LISA cohort for children up to 15 years of age. This separation was more pronounced in the LISA cohort than in the LINA cohort, which might have been related to the fact that some of the eight-year-old children with wheezing symptoms of the LINA cohort had not been diagnosed with asthma at the time of questionnaire evaluation. As such, these children have potentially masked the difference between wheeze and asthma. With increasing age, differences in IgE concentrations between asthmatic, wheezing and healthy children became more pronounced.

Previous studies already aimed to differentiate children with transient wheezing episodes from those with persistent wheeze based on early childhood total serum IgE levels (93, 95).
In a ten-year follow up study Peat *et al.* investigated the relationship of lung function and atopy. School children aged eight-ten were recruited and their airway hyperresponsiveness and sensitization determined over a ten-year period in two-year intervals. Skin-prick tests (SPT) were performed every other year while blood samples were only taken at the ten-year follow-up. Based on total IgE levels they found a dose-response relationship between allergic sensitization and hyperresponsiveness of the airways in the young adults (93). This is in line with the findings of this thesis that indeed elevated IgE levels might be an indication of asthma development. However, whether children with wheezing symptoms who were not diagnosed with asthma showed different patterns from children with a diagnosed asthma was not assessed by Peat *et al*.

In a landmark study for the characterization of different wheezing endotypes Martinez and colleagues investigated wheezing episodes in 826 young infants starting at birth until the age of six years. From this longitudinal assessment the different wheezing endotypes, as mentioned earlier, were derived. They found no increased risk for asthma for those children with transient wheezing symptoms, while persistently wheezing children had an increased risk for development of asthma (95), which might be one of the reasons why many studies use persistent wheeze as a substitute for childhood asthma. However, the study by Martinez *et al.* was one of the first that found an association with persistent wheezing episodes and elevated total serum IgE levels at nine months after birth and at six years of age compared to other wheezing endotypes and those children who never wheezed (95).
This is in contrast to the results presented in this thesis. While children who were diagnosed with asthma indeed showed significantly increased IgE levels, persistent wheeze was not associated with elevated IgE levels. As Martinez *et al.* did not separate children with an asthma diagnosis from the different wheezing endotypes, high IgE levels might have been driven by children with an existing asthma diagnosis. The results presented in this thesis support the evidence that persistent wheezing cannot be considered as a substitute for childhood asthma. Therefore, the major advantages of the here presented study were, that the children were not only monitored until the age of six, but rather until the age of 15. Children with an asthma diagnosis were considered as a separate group from the children with wheezing without an asthma diagnosis. Due to the longer observation period it was possible to use a reliable asthma diagnosis for retrospective data analysis. This made it possible to reduce the false-positive rate in the wheezing groups, which led to an even clearer separation in the IgE concentration of the asthma children from the wheezers.

Interestingly, this difference between asthmatic children and healthy controls was more prominent for sIgE than for tIgE concentrations. This might be due to the fact that sensitization to aeroallergens as evaluated in this study (e.g. pollens or dust mites) is indicative of the bronchial mucosae mast cell response, leading to a direct immune reaction such as the release of IL4 or IL5 cytokines for inflammatory processes (206). In contrast, total serum IgE concentrations reflect the systemic concentration of IgEs against a variety of different allergens and exposures (207, 208).

The increased sIgE levels in asthmatic children compared to healthy controls were further supported by different studies, reporting that wheezing and asthma are associated with impaired lung function and sensitization to IgE levels against aeroallergens (209-212). Thus, Arroyave *et al.* studied 351 children (1-17 years) and 390 adults (18+ years) of the 2005-2006 National Health and Nutrition survey (NHANES) in the United States including sIgE (indoor allergens e.g. house dust mites). They found sensitization to particular indoor environmental allergens and conclude that to be a risk factor for wheeze and asthma visits in the clinic (211). In the German Multicentre Allergy Study (MAS) 1,314 children (birth to seven years) with, e.g. wheezing symptoms and sIgE values against food and inhalation allergens were longitudinally examined including lung function measurement at the age of seven years. They found a link between asthmatic children and impaired lung function (209). In a review of 17 examined studies about risk factors on asthma by Boa *et al*, increased sIgE was identified as an risk factor among others for developing asthma in preschool children (212). However, these studies do not distinguish between individual wheezing endotypes and asthma.

This thesis did not only identify differences in the longitudinal pattern of serum IgEs in children with asthma compared to those with different wheezing endotypes or healthy controls, but also found that the percentage of sensitized children (tIgE>100kU/l or sIgE>0.35kU/l) was twice as high in asthmatics than in the other groups. It should be particularly emphasised that the percentage of sensitization in persistently wheezing children in the LINA study was similar to healthy controls rather than to asthmatic children. This difference in percentage of sensitized children could be confirmed in the LISA cohort with an even stronger separation between asthmatics and all other groups, potentially driven by a higher sample size and a more sophisticated age.

This is in line with a variety of previous studies reporting early sensitization in asthmatic children (213-215). Thus, Boersma *et al.* investigated 116 children (age >6 years) and suggested that sensitization to inhalant allergens is a strong predictor of school age asthma (213). In the Copenhagen Prospective Studies on Asthma in Childhood$_{2000}$ (COPSAC$_{2000}$), Schoos *et al*. longitudinally (6, 18 months, 4, 6 and 13 years) assessed sIgE against food- and aeroallergens in 411 asthmatic children. They observed that children with high levels of sIgE

or large SPT wheel sizes during childhood were associated with an increased risk of asthma at the age of 13 years (214). In a meta-analysis including five population-based studies including 14.849 adults with measurements of sIgE against inhalant allergens, Skaaby *et al.* compared sIgE sensitized with non-sensitized adults. A significant increased risk for asthma was observed in the group of sensitized adults (215). These studies mainly focussed on the association of sensitization and the risk for asthma and did not further investigate wheezing endotypes.

In contrast to the mentioned studies before there are others reporting early sensitization in wheezing children (216, 217). In the Southampton Women´s Survey Collins *et al.* investigated the different wheezing endotypes according to Martinez *et al.* in 926 children from birth until the age of 6 using longitudinal airway function and atopic sensitization (SPT). They could differentiate persistent and late-onset wheezers compared to children never wheezed using atopic sensitization (216). In a study on the relationship between tIgE levels and the development of wheezing and allergic sensitisation in 540 children (up to the age of 11 years) from the Tucson Children's Respiratory Study, Sherrill *et al.* (217) reported that early sensitization to aeroallergens and increased serum IgE levels are associated with wheezing endotypes. They suggested that both early allergic sensitization and elevated serum IgE levels as risk factors for asthma development but not differentiate separately the wheezing groups and asthma (217). Thereby, the here presented results suggest that persistent wheeze and asthma should be considered as independent groups.

In conclusion, specific IgE levels against aeroallergens and/or total IgE levels should be considered in conjunction with clinical features to allow a more accurate diagnosis of asthma.

In addition to IgE, this thesis reports an increased expression of *IL5RA* in asthmatics compared to healthy controls and to all wheezing endotypes in both investigated cohorts (LINA and LISA). By binding to its receptor (encoded by *IL5RA*) Interleukin 5 mediates important pathways closely linked to both IgE and asthma development (98) as it is involved in eosinophilic (96) and regulatory T cell pathways (160), which play a crucial role in lung inflammation in asthma (218). The result of increased *IL5RA* expression in asthmatic children in this thesis might suggest, that *IL5RA* expression might be an early predictor for asthma development and perturbed IL5 signalling might already be present before the clinical manifestation of asthma.

This close interaction between IgE and *IL5RA* was mirrored by the result of the mediation analysis suggesting that the development of asthma require both, high IgE levels as well as high *IL5RA* transcriptional activities. In line differential gene expression analysis of 326 children´s blood samples (age six to 12 years) and 610 young adult´s blood samples (age 16 to 25 years) from two Costa Rican cohorts of Virkud *et al.* revealed an association between *IL5RA* expression and total IgE levels in childhood asthma (98).

IL5RA transcription is highly cell type specific and is mainly expressed by eosinophil granulocytes (219). Thus a higher expression of *IL5RA* might be a result of increased eosinophil counts commonly found in asthmatics. Interestingly, in the LINA study four-year-old children already showed an increased *IL5RA* expression prior to their initial asthma diagnosis, suggesting that *IL5RA* activation may occur already before the onset of asthma and could constitute a potential cause of asthma development rather than a consequence. IL-5RA is already therapeutically targeted to directly decrease the number of eosinophils as an adjunct to a standard anti-IgE therapy in severe uncontrolled asthma (220). As such blood *IL5RA* mRNA levels could aid stratification of children at risk for asthma development. Further studies are required in order to confirm the usage of *IL5RA* expression levels for a precise early asthma prediction.

It should be mentioned that the present study has some potential limitations. As the definition of asthma was based on questionnaire information it might be less accurate than an asthma diagnosis made in a clinical setting. As such the possibility that children were either wrongfully classified as asthmatics or asthmatics were regarded as wheezing children cannot be fully excluded. *IL5RA* expression is most likely strongly influenced by the number of eosinophil granulocytes. Unfortunately, both cohorts lacked information on individual blood cell type composition. Whether increased *IL5RA* expression only reflected increased eosinophil counts is therefore unknown. It would have been interesting to see if the transcriptional changes also translate to the protein level. Limitations in biomaterial availability rendered it impossible to address this issue. Although this study was based on a rather small case number and therefore has to be interpreted with caution, it shows a distinct discrimination between asthmatic children and children with different wheezing endotypes or healthy controls. Further research will be needed to elucidate the underlying mechanisms of *IL5RA* activation prior to asthma development.

The strengths of this presented study consists of dense long term observation and assessment of investigated parameters. Moreover, while the majority of birth cohorts dispose of only few blood sampling time points, the LINA study with annual clinical visits and blood sampling until the age of eight offered the opportunity to study the timing of sensitization, wheezing symptoms, and asthma onset in a tightly time-resolved manner. The LISA study even allowed verifying and validating these parameters up to the age of 15.
Taken together, this study provides interesting new data of potential clinical relevance for asthma development. The study could show that the longitudinal pattern of serum tIgE and sIgE are able to distinguish children with asthma from wheezing children and healthy controls. In particular, also from children with persistent wheeze, which is often used as a surrogate for

asthma in early childhood, underlining the necessity to consider both as different disease entities. In addition, mediation analysis suggested that asthma development is not only driven by high IgE levels but potentially also by elevated *IL5RA* mRNA expression. The presented results give evidence that the activation of the *IL5RA* may occur already before the onset of asthma. In early life, high IgE levels and the concomitant increase in *IL5RA* mRNA expression might help to distinguish children developing asthma from those with wheezing symptoms, but which will never suffer from asthma.

4.2 Aberrant DNA methylation pattern in childhood asthma

During the last decade epigenetic mechanisms moved into the focus of asthma research as a modifiable layer mediating the influence of the environment and the genetic predisposition thereby affecting transcription and translation processes (229). In particular, DNA methylation as the best studied epigenetic modification so far (230) is receiving increasing attention in asthma research. Most DNA methylation changes in asthma have been studied using array based technologies with limited access to all CpGs in the genome. Therefore, the recent study in this thesis was aimed to study asthma-related DNA methylation changes by a more comprehensive approach. Genome-wide DNA-methylation sequencing (WGBS) was used, as it is an unbiased approach assessing all CpG sites in the genome, to minimize limitations of current array based approaches. In total 40 asthmatic children were compared to 42 age-matched healthy controls from three German birth cohorts (LINA, LISA and PASTURE). Since differentially methylated regions (DMRs) are believed to be more likely to impact transcription than single loci, DMRs comprising of at least three consecutive CpG sites, were determined (221, 222).

To the best of my knowledge, this is the first truly genome-wide study of DNA methylation on a single CpG level for childhood asthma based on a next generation sequencing approach. In total, 161 asthma-related DMRs were detected. Interestingly, nearly all identified DMRs were hypomethylated and 63% of them are located in enhancer regions, potentially suggesting an activation of the identified pathways in asthmatic children. This DNA hypomethylation observed in asthmatic children is in line with a recent study by Xu and colleagues. In their epigenome-wide meta-analysis, they assessed methylation changes between 392 children with asthma and 1156 controls based on 450K methylation arrays (88). They identified 27 differentially methylated CpGs (= differentially methylated position, DMP) in total, of which 14 were successfully validated in six additional cohorts with 247 cases and 2949 controls. Similar to the methylation pattern described in this thesis all of these 14 loci were hypomethylated in asthmatic children (88). Xu et al. further assessed the genomic location of the methylation

changes and found 13 out of the 14 DMPs were located in putative enhancer regions (88, 125). Their results support the finding of this study that DNA methylation changes related to childhood asthma preferentially affect enhancer regions. The targeted validation analyses that was performed in much larger sample set, confirmed the hypomethylation of selected candidate DMRs. This did not only support the genome-wide approach by WGBS applied in this study but strengthened the validity of a discovered DMR as a previously unknown enhancer region likely related to IL-4.

Further supporting results of this study is the overlap found with CpGs described in earlier studies in relation to childhood asthma (75-77, 81). In the set of DMRs, 18 CpGs had previously been discovered by 450K studies affecting genes functionally linked to pathophysiological processes of asthma (75-77, 81, 162). In the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort Arathimos *et al.* assessed DNA methylation in peripheral blood of approximately 1000 children (age: 7.5 - 16.5 years of age) in association with asthma and wheeze outcomes (81). In addition to the findings of Arathimos and colleagues, Reese *et al.* identified 179 CpGs and 36 differentially methylated regions in the blood of 631 asthmatic children using Illumina 450K arrays (76). Among the genes affected by methylation changes were e*osinophil peroxidase* (*EPX)*, *zinc finger protein, FOG family member 1* (*ZFPM1)*, *GATA binding protein 2* (*GATA2)* and *acyl-CoA thioesterase 7* (*ACOT7).* In this thesis, DMRs were also identified in these asthma related genes. For instance, *EPX* is predominantly expressed by eosinophils and therefore important for inflammation and asthma (223). The transcription factor *Zinc Finger Protein, FOG Family Member 1*, is involved in the downregulation of Th2-cytokines (e.g. IL4), and therefore might be an important regulator in allergic disease, which usually leads to an increase in Th2 cytokines (224). Furthermore, another transcription factor the *GATA Binding Protein 2* is involved in allergic response via basophils and mast cells regulation (225). The *Acyl-CoA Thioesterase 7* is involved in inflammation through the production of arachidonic acid (226) which is used to induce an inflammatory reaction by, e.g. the formation of mediators. Two pathways are central for an inflammatory reaction, the cyclooxygenase (COX)- and the 5-lipoxygnease-pathway (226). Arachidonic acid plays a central role especially in the 5-lipoxygnease-pathway since its converted into prostaglandins, thromboxanes and leukotrienes, which are causal mediators for lung inflammation and the development of asthma (226). Furthermore, EWAS studies associated DNA methylation of *ACOT7* with serum IgE (161, 227), which have a crucial role in asthma as already described in the previous chapter.

Of note, DNA methylation changes related to *ZFPM1, EVL, ACOT7 and EPX* and were not only observed in the blood of asthma patients, but also were detected in nasal or bronchial epithelial cells (81, 168). This overlap to asthma specific methylation pattern of nasal or

bronchial epithelium indicates that even blood could potentially be used as surrogate to predict methylation changes in target tissues of asthma disease and further might help to understand disease development to improve risk stratification, asthma prevention and therapy.

The methylation changes of the current study did not only show an overlap with already described differentially methylated CpGs but affected similar target genes, although related to DNA methylation changes at different sites of the genome. Out of the genes associated with the 161 DMRs 43 have already been described as being affected in asthmatic children in previous 450k studies (187, 228-230). Among these were *interleukin 4 (IL4), interleukin 5 receptor α (IL5RA), regulatory associated protein of MTOR complex 1 (RPTOR) and Enah/Vasp-Like* (*EVL). IL4* and *IL5RA* are commonly known as drivers of allergic inflammatory processes (231, 232). It has been suggested that activation of mTOR signalling is involved in the onset of asthma, by *RAPTOR.* This gene is an important mediator in mTOR signalling and pharmaceutical targeting of RPTOR by rapamycin has been shown in animal studies to alleviate asthma symptoms (233). Remodelling of the airways is one pathological change of asthma. Cytoskeleton remodelling including *EVL* is important for this airway alterations (234) and was previously associated with asthma via epigenome-wide association studies (76, 81). The observed DNA methylation changes might be involved in impaired functional translation of the identified genes. Furthermore, the presented study identified genes which were found in the NHGRI-EBI Catalog of human genome-wide association studies (235) with asthma traits such as *RAD50 Double Strand Break Repair Protein (RAD50) (236), TNF Receptor Associated Factor 1 and 3 (TRAF1 (237) /TRAF3 (238)), and Kinesin Family Member 3A (KIF3A) (239)*.

The vast majority of epigenetic studies associated with childhood asthma did not assess whether the observed methylation changes were associated with transcriptional changes. However, it is an essential question whether the altered DNA methylation pattern contributes to RNA expression changes observed in asthma. In one recent meta-analysis Reese *et al.* showed a correlation between DNA methylation and RNA expression among others for *IL5RA* and *EPX* in asthmatic children (76). Indeed also in the current study the hypomethylation in DMRs related to *IL5RA* and *EPX* correlated with an increase in transcription of *IL5RA* and *EPX,* respectively. In addition, asthmatic children showed a higher expression of these genes than the controls. For IL4, the decreased DNA methylation in the newly identified enhancer region was not only related to a significant increase of *IL4* mRNA expression but also to an elevated IL4 blood protein concentration. As such, DNA methylation changes seem to play an important role in the perturbed expression of asthma related mediators.

One of the drawbacks of 450k arrays is the poor coverage of intragenic regions as these arrays mainly cover promoter areas or intergenic CpGs. However, intragenic regions have high regulatory potential since enhancers are often located away from TSS or gene-rich regions.

Sequencing approaches like WGBS in principle could cover all CpGs in the genome and therefore also intragenic regions. By applying WGBS the current study could identify a strong enrichment of DNA methylation changes in enhancer regions. Sixty three percent of the identified DMRs intersected with known activated chromatin regions or enhancers. The enhancer target genes affected by asthma related DNA methylation changes spanned a network of genes involved in immune response and inflammation. In addition, for several of these network genes a relationship with asthma has already been described in the literature found by natural language processing analysis, including *RELA Proto-Oncogene, NF-KB Subunit (RELA) (240), YY1 Transcription Factor (YY1) (241) and Tyrosine 3-Monooxygenase/Tryptophan 5-Monooxygenase Activation Protein Zeta (YWHAZ*) (242). This highlights the notion that not a single gene or loci affect asthma development, but rather the interplay of several genes contributes to the disease aetiology most likely mediated by epigenetic modifications of enhancers.

The majority of the identified DMRs was independent of the genotype. Only 24 of the 161 DMRs showed an association to a nearby SNP in a window of 5kb up or downstream of the DMR. A comparison of the SNP associated DMRs with the NHGRI-EBI Catalog of human genome-wide association studies (235) revealed that none of these SNPs had been described in earlier genome-wide association studies (GWAS) in relation to asthma. Of note, as the associations to the genotype were assessed up to a defined distance to the DMR, genetic variants outside this investigated windows could have been missed. However, as for the majority of DMRs no association with the genotype was observed this study further emphasizes the importance of life style and environmental factors contributing to asthma development (243, 244). These results are in line with Bauer *et al.* (106), who showed that prenatal tobacco smoke exposure was associated with a change in DNA methylation, which was partially independent of the genotype. In addition, a link between tobacco smoke-associated DNA methylation changes and later wheezing could be shown.

DNA methylation changes can alter transcription by affecting the ability of the transcription machinery to interact with the DNA. Transcription factors are important mediators of gene regulation (245) and their binding preferentially occurs in regulatory regions like in enhancers or silencers (246). Interestingly, asthma-related DMRs were enriched in binding sites for transcription factors for which a role in asthma has already been suggested like *GATA Binding Protein 3 (GATA3) (247), Nuclear Factor Of Activated T Cells 1 (NFATC1) (248) or Interferon Regulatory Factor 1 (IRF1) (249)*. Accordingly, the genes associated to these DMRs are functionally involved in asthma and immune related pathways such as *PPARG Coactivator 1 Beta (PPARGC1B) (155) or TNF Superfamily Member 10 (TNFSF10) (250).* For instance*, TNFSF10* is known to induce apoptosis of transformed cells. This gene is involved in airway

remodelling process and expressed in cells involved in asthma such as eosinophils, mast cells or airway epithelial cells (250). The DNA methylation changes observed in this study therefore seem to affect regulatory hubs namely enhancer regions preferentially bound by immune modulating transcription factors.

Although this study had the particular advantage to assess DNA methylation at a truly genome-wide level some limitations should also be noted. The higher costs of sequencing usually lead to a small number of samples that can be assessed impeding the inclusion of potential confounders including cell-type distribution. This is in contrast to most array-based analyses in which case numbers are higher and such confounding factors can more easily be included in regression analyses. To counteract this disadvantage samples were age-matched and equally distributed for gender when comparing asthmatics and healthy controls. Two independent algorithms for the identification of DMRs were applied and only the overlap was further validated by targeted DNA methylation approach in a larger sample size. This approach revealed perturbed DNA methylation in regulatory elements outside of regions commonly covered by array-based studies not only at single CpGs positions but also in larger genomic regions.

The results of this study suggest dysregulation of enhancer methylation as an important feature of asthma that in the future may be exploited for therapeutic purposes. Further studies are certainly necessary to assess the functional role of perturbations in the DNA methylation pattern in enhancer regions as well as other regulatory elements. As such integrating transcriptome and proteome data analysis in future methylome studies would help to achieve a more comprehensive picture of the functional relevance of DNA methylation changes. In addition, other more dynamic epigenetic mechanisms such as histone modifications that also affect transcription and translation should also be considered.

Taken together, the results of the study in this thesis showed that the development of asthma is associated with changes in the DNA methylation pattern of blood cells, particularly in enhancer regions. In addition to the identification of DNA methylation changes of known genes related to asthma, including *IL4, EPX* and *IL5RA*, DNA methylation changes have also been identified in genes previously not linked to asthma development. The study expands the knowledge of the perturbed DNA methylation pattern between asthmatic children and controls, which might help to understand disease development.

4.3 Longitudinal DNA methylation pattern in asthma

Longitudinally stable DNA methylation pattern might serve as marker regions for stratification of children at risk for of asthma development. Therefore, the analysis conducted in children

with a developed asthma phenotype in comparison to controls was extended to the cord blood methylome. A top-down and bottom-up approach were followed to determine stable asthma-related DMRs. In the top-down approach it was assessed whether the differential methylation of the 161 DMRs identified within the already established asthma phenotype was already present at time of birth (Chapter 3.2). Indeed, 27 of these DMRs showed a significant DNA methylation difference in children who later in life develop asthma compared to healthy controls already in cord blood. Of note, the majority of these DMRs showed no association to the genotype underlining the notion that the prenatal environment essentially contributes to the asthma risk of a child (55).

In the bottom-up approach cord-blood WGBS data were compared between individuals who were diagnosed with asthma later in their life and healthy controls. Almost 60 DMRs were identified at the time of birth, the majority of them overlapping with enhancer regions, suggesting a potential regulatory function in regard to the disease onset. The overlap of both approaches was comprised of six regions of which five were not associated to the genotype. Interestingly, one of these regions, a DMR located in an enhancer region of the opioid binding protein/cell adhesion molecule like (*OPCML*) gene, have previously been linked to prenatal tobacco smoke exposure (106). Bauer *et al.* studied epigenetic changes associated with maternal smoking in 16 children at three different time points (birth, year one and four) in the LINA cohort using WGBS (106). Remarkably, this particular DMR was identified in cord blood and remained stable until the age of four. Another study by Joubert *et al.* showed aberrant DNA methylation pattern in the *OPCML* gene in cord blood of children in association with maternal folate intake during pregnancy (258).

These DNA methylation changes might be of value as potential early biomarker to stratify children at risk for asthma development. Further studies are warranted to validate these findings in other populations.

So far, only few studies have been published on longitudinal DNA methylation changes with regards to childhood asthma (76, 87, 187). A 450K array based study by deVries and colleagues (187) observed a methylation change in the promoter of the *SMAD3* gene in cord blood, which was associated with subsequent asthma development. *SMAD3* is known to be involved in regulation of T-cell differentiation (251) and thereby may indirectly affect inflammation processes inherent to asthma (252). DeVries *et al*. hypothesised that maternal asthma contributes to this epigenetic change in *SMAD3* thereby mediating the risk increase for childhood asthma (187). This is in line with the results of this thesis, showing that epigenetic modifications are already present at time of birth distinguishing children developing asthma later in life from healthy children. In a targeted study that investigated promoter methylation of Th2 lineage genes at time of birth Barton *et al.* observed a negative relationship of *GATA3*

methylation and development of asthma in later childhood (87). No DMR in the *GATA3* promoter was observed in this study most likely related to the different approaches used as targeted assessments of methylation changes do not require *p*-value adjustment like genome-wide approaches do. Nevertheless, the enrichment of *GATA3* binding motifs observed in the set of almost exclusively hypomethylated DMRs suggests a critical involvement of GATA3 mediated signalling in asthma development that is affected by epigenetic modification. An EWAS investigating cord blood DNA methylation of children with a later asthma onset by Reese et al. (80) included a meta-analysis and successfully identified nine differentially methylated CpGs and 35 regions at the time of birth between 668 cases and 2904 controls. These CpGs and regions were not described with asthma before and they suggested these as markers for disease development. However, these findings could not be confirmed in the current study. Reasons for those lacks in confirming previous results might be driven by the different approaches to detect DNA methylation. The here presented study used WGBS in contrast to DNA methylation arrays and the applied DMR definition of three consecutive CpGs instead of single CpG positions. However, in general the results from LINA and PASTURE at birth confirm that DNA methylation changes between asthma and controls can be identified already before disease onset.

A potential limitation of this study is the missing information on cell type compositions of the blood samples. Therefore, it can be not ruled out, that differences in cell composition might have contributed to the observed DNA methylation changes (16). Especially inflammation processes activating specific cell types such as eosinophilic granulocytes potentially can trigger differences in cell type composition (253). However, it was suggested that cell composition of cord blood samples of children with developing asthma later in life and those who would remained healthy are very similar so that cell composition differences are potentially not such a strong confounder (168). As already discussed, WGBS studies are limited in the number of sample that can be assessed due to comparatively higher costs. To overcome this limitation, two different approaches (bottom up and top down) were used to determine longitudinal stable asthma-related DNA methylation changes. Nevertheless, further research is warranted to validate the identified longitudinally stable DMRs in larger cohorts.

The current study identified six previously unknown longitudinally stable DMRs. These DNA methylation changes might be of value as potential early biomarkers to stratify children at risk for asthma development. The findings of the current study might contribute to a better understanding of mechanism and pathways potentially involved in driving the disease onset. Further interesting research questions remain to be clarified e.g. which additional prenatal influencing factors (e.g. prenatal smoking exposure) could drive DNA methylation pattern and could help to improve asthma diagnosis and prediction.

5 CONCLUSION AND PERSPECTIVE

The World Health Organization declares asthma as one of the most common inflammatory diseases in children worldwide. Consequently, intense research is required to understand the pathomechanism behind the disease development and manifestation in order to further establish reliable tools for early diagnosis, prevention and therapy. This is especially important in early childhood to counteract asthma development from very early on. However, diagnosing asthma at this stage of life is highly challenging. So far, predicting markers for early risk stratification of asthma are insufficient or missing. Thus, the major aim of this thesis was the identification of reliable makers for the risk of asthma development in early childhood based on clinical parameters and epigenetic regulating features.

Wheezing is a common symptom of asthma and is still used as a clinical surrogate for asthma diagnosis, although not every child with wheezing symptoms is also developing asthma. This work successfully demonstrated that wheezing children who later developed asthma can be distinguished from children without asthma manifestation by using a longitudinal IgE course and early mRNA expression profiles of *IL5RA*.

In addition to IgE and *IL5RA* expression levels, disease specific DNA methylation pattern revealed to be another promising tool to identify children at high risk for later asthma development. Differentially methylated regions were identified distinguishing between children with and without asthma. Some of these phenotype specific methylation changes could already be found in cord blood right after the children were born.

These newly identified longitudinal stable regions could potentially be used in the future as early biomarkers to stratify children at high risk for asthma development. In summary, these findings might contribute to a better understanding of mechanism and pathways potentially driving the disease onset. However, some questions remain to be clarified for example which early exposures such as prenatal tobacco smoke could drive such differences in DNA methylation pattern.

Although a comprehensive assessment of transcriptome and proteome data could be considered to further strengthen the validity of the here presented data, the investigation of longitudinal epigenetic marks involved in early asthma development is a promising approach to understand the disease onset and putative influencing factors.

6 SUMMARY

Asthma is one of the most common chronic inflammatory diseases, affecting more than 339 million people worldwide including especially children with a prevalence of 6.3% in Germany. Since an early diagnosis is crucial to inhibit disease progression, asthma is being intensively researched during the last decades to understand disease driving mechanism in order to find reliable tools for early diagnosis, therapy and prevention. So far, early asthma diagnosis in children is still challenging thus the lung, as major target organ is still growing and similar symptoms such as wheezing are therefore difficult to distinguish from asthma itself. Nevertheless, wheezing episodes in general are common during childhood and usually caused by viral infections. Whereas the asthma development and airway hyperresponsiveness seems to be facilitated by immunoglobulin E (IgE)-mediated lung inflammation. Early childhood asthma is mostly associated with allergic hypersensitivity (atopy), while non-allergic (intrinsic) asthma is a rather rare condition. IgE is an important driver of impaired lung function in early childhood, therefore several studies tried to differentiate children with transient wheezing episodes from those with persistent wheezing based on early childhood serum IgE levels. These studies provided evidence that persistently wheezing children may have increased IgE levels from very early on. However, those studies disregard the fact that not every child with persistent wheeze will develop asthma later on. Based on this, the first aim of this thesis was to evaluate the longitudinal trends of serum IgE concentrations in wheeze and asthma endotypes of two German birth cohorts: LINA (Lifestyle and Environmental Factors and their Influence on Newborns Allergy risk) and LISA (Influences of Lifestyle- related factors on the Immune System and the Development of Allergies in Childhood).

This thesis successfully showed that children with asthma could indeed be distinguished from persistent wheezing children based on elevated longitudinal IgE data. This result underlines the necessity to consider persistent wheeze and asthma as different disease entities. Furthermore, since previous research also indicate a close relation between the *IL-5 receptor alpha* (*IL5RA*) and IgE mediated asthma development, *IL5RA* mRNA expression levels were investigated, aiming to identify another stable marker for disease progression. The here presented results could provide further evidence that IL5RA activation may occur already before the onset of asthma as indicated by increased *IL5RA* mRNA expression in children with later developing asthma compared to wheezing children and controls. In early life, high IgE levels and the concomitant increase in *IL5RA* mRNA expression might help to distinguish children developing asthma later in life from those with wheezing symptoms who will never suffer from asthma.

Besides the problem of an early asthma risk detection, the underlying mechanisms of asthma development are still poorly understood. It is well known that environmental factors (e.g. exposure to tobacco smoke or allergens) combined with a genetic predisposition strongly influence the disease risk potentially mediated via epigenetic mechanisms such as DNA methylation. Understanding these epigenetic mechanisms might contribute to a broader knowledge of early asthma development. Therefore, the second major aim of this thesis was to investigate whole blood DNA methylation pattern of asthmatic children and healthy controls by performing whole genome bisulfite sequencing (WGBS) from 82 children of three German birth cohorts, LINA, LISA and PASTURE (Protection against Allergy: Study in Rural Environments).

This thesis was able to clearly demonstrate differences in DNA methylation pattern of blood cells between asthmatic children and healthy controls, especially in regulatory hubs, such as enhancer regions. In addition to the identification of epigenetic modifications in genes known to be related to asthma, including *IL4, EPX* and *IL5RA*, DNA methylation changes have also been identified in previously unknown genes. Thus, these newly identified genes may also influence asthma development, the here presented study further contribute towards a better understanding of asthma development in early life.

DNA methylation changes occurring before the development of the clinical phenotype might serve as marker for the prediction of later disease development. Therefore, the third aim of this dissertation focussed on cord blood DNA methylation pattern of children who developed asthma later in life compared to healthy controls by performing WGBS from 48 samples of the LINA and PASTURE cohort. By doing so this thesis successfully demonstrated that some of the DNA methylation changes identified in asthmatic children were already present in cord blood long time before the onset of the disease. These longitudinal stable differentially methylated regions might be used in the future as potential predictive biomarkers to stratify the risk for asthma development for children later in life.

Altogether, this dissertation could provide important contributions to the identification of children at high risk for asthma development and could furthermore demonstrate the involvement of DNA methylation changes within enhancer regions in this early progress. Finally this study investigated genome-wide DNA methylation data from a whole genome sequencing approach, which for the first time enabled the possibility to examine regulatory regions not covered by array-based methods so far.

Zusammenfassung

Asthma ist eine der häufigsten chronisch entzündlichen Erkrankungen, an der weltweit mehr als 339 Millionen Menschen leiden, darunter vor allem auch Kinder mit einer Prävalenz von 6.3% in Deutschland. Da vor allem eine frühe Diagnose wichtig ist für einen verbesserten Krankheitsverlauf wird an der Entstehung von Asthma sowie zugrundeliegender Mechanismen in den letzten Jahrzehnten intensiv geforscht, mit dem Ziel neue und zuverlässige Instrumente für die frühe Diagnostik, Therapie und Prävention zu finden oder stetig zu verbessern. Aktuell ist jedoch die Asthma-Frühdiagnostik bei Kindern noch eine große Herausforderung, da die hauptsächlich von der Erkrankung betroffene Lunge sich noch im Wachstum befindet. Daher lassen sich Symptome wie z.B. das „Wheezen“ (Keuchen oder Giemen) von Asthma direkt schwierig unterscheiden. Solche Keuchanfälle treten bei jungen Kindern häufig auf und werden meist durch virale Infektionen der unteren Atemwege ausgelöst. Wohingegen die Entwicklung von Asthma sowie eine Überempfindlichkeit der Atemwege eher durch Immunglobulin E (IgE)-vermittelte Entzündungen in der Lunge vermittelt wird. Frühkindliches Asthma ist meist mit einer allergischen Überempfindlichkeit (Atopie) assoziiert, während nicht allergisches (intrinsisches) Asthma eher selten ist. Dem IgE kommt bei der Entwicklung einer gestörten Lungenfunktion in der frühen Kindheit eine wichtige Rolle zu. Daher versuchten mehrere Studien, Kinder mit vorübergehenden Keuchepisoden von Kindern mit anhaltendem Keuchen anhand der frühkindlichen Serum-IgE-Spiegel zu unterscheiden. Diese lieferten Hinweise darauf, dass bei Kindern mit anhaltendem Keuchen die IgE-Spiegel bereits sehr früh im Leben erhöht sein können. Allerdings erkrankt nicht jedes Kind mit persistierendem oder spät einsetzenden Keuchen an Asthma, was aber in den meisten Studien nicht weiter unterschieden wurde. Daher war das erste Ziel dieser Dissertation die longitudinalen Verläufe der Serum IgE Konzentrationen bei Kindern mit Keuchen und Asthma zu untersuchen. Dabei wurden Asthma und persistierendes Keuchen als unabhängige Entitäten in den zwei deutschen Geburtskohorten LINA und LISA betrachtet.

Die hier dargelegte Dissertation konnte erfolgreich zeigen, dass Kinder mit Asthma von Kindern mit persistierendem Keuchen durchaus anhand von longitudinalen IgE Daten weiter unterschieden werden können. Dies unterstreicht die Notwendigkeit, beide als unterschiedliche Krankheitsentitäten zu betrachten. Neben den IgE Spiegeln wurde die mRNA Expression des IL-5 Rezeptors untersucht, da dieser eng mit IgE und Asthma verbunden ist. Es konnte gezeigt werden, dass die Entwicklung von Asthma nicht nur durch hohe IgE-Spiegel, sondern möglicherweise auch durch eine erhöhte *IL5RA*-mRNA-Expression angetrieben wird. Die hier vorgestellten Ergebnisse geben Hinweise darauf, dass die Aktivierung des IL-5 Rezeptors bereits vor Ausbruch der Erkrankung erfolgen kann, wie die erhöhte *IL5RA*-mRNA-Expression bei Kindern mit später entwickelndem Asthma im Vergleich zu keuchenden Kindern

und Kontrollen zeigt. Im frühen Kindesalter könnten hohe IgE-Spiegel und die damit einhergehende erhöhte *IL5RA*-mRNA-Expression dazu beitragen, Kinder, die Asthma entwickeln, von solchen mit Keuchsymptomen, die nie an Asthma leiden werden, zu unterscheiden.

Neben dem Problem der frühen Erkennung von Kindern mit einem hohen Asthma Risiko sind die zugrundeliegenden Mechanismen der Krankheitsentstehung bisher noch wenig bekannt. Es ist jedoch klar, dass das Krankheitsrisiko vor allem durch das Zusammenspiel von Umweltfaktoren (z.B. die Exposition gegenüber Tabakrauch oder Allergenen) mit genetischen Prädispositionen beeinflusst wird. Des Weiteren werden epigenetische Mechanismen, wie die DNA-Methylierung, als mögliche Mediatoren zwischen diesen Einflussfaktoren in Betracht gezogen. Ein besseres Verständnis dieser epigenetischen Mechanismen kann dazu beitragen, die frühe Entstehung von Asthma besser zu verstehen. Daher war das zweite Ziel dieser Arbeit die DNA-Methylierung im Blut von asthmatischen und gesunden Kindern vergleichend zu untersuchen. Dafür wurden Blutproben von 82 Kindern aus drei deutschen Geburtskohorten (LINA, LISA und PASTURE) genutzt und mittels genomweiter Bisulfit-Sequenzierung (WGBS) analysiert.

Diese Dissertation konnte feststellen, dass Kinder mit Asthma deutliche Unterschiede in den DNA-Methylierungsmustern der Blutzellen im Vergleich zu den gesunden Kontrollen aufwiesen, insbesondere in regulatorischen Bereichen wie den Enhancer (Verstärker)-Regionen. Neben der Identifizierung von epigenetischen Modifikationen in bereits mit Asthma in Verbindung stehenden Genen, wie *IL4, EPX und IL5RA*, wurden auch DNA-Methylierungsänderungen in bisher unbekannten Regionen bzw. Genen identifiziert. Diese Gene könnten ebenfalls an der Entstehung von Asthma beteiligt sein. Somit leistet die hier dargestellte Studie einen wichtigen Beitrag um die Asthmaentwicklung im frühen Kindesalter besser zu verstehen und erweitert das Wissen über die veränderten DNA-Methylierungsmuster zwischen asthmatischen Kindern und gesunden Kontrollen.

DNA Methylierungsveränderungen können bereits vor der Entstehung eines klinischen Phänotyps auftreten und somit als ein wichtiger Marker für die Krankheitsvorhersage genutzt werden. Daher war das dritte Ziel dieser Dissertation die DNA-Methylierung aus dem Nabelschnurblut von später an Asthma erkrankten Kindern mit dem von gesunden Kontrollen zu vergleichen und im Hinblick auf DNA-Methylierungsmuster zum Zeitpunkt des bereits etablierten Asthmas zu untersuchen. Insgesamt wurde das Nabelschnurblut von 48 Kindern der LINA und PASTURE Kohorte mit genomweiter Bisulfit-Sequenzierung (WGBS) untersucht. Dabei konnte gezeigt werden, dass im etablierten Phänotypen identifizierte asthmaspezifische DNA-Methylierungsmuster zum Teil bereits im Nabelschnurblut präsent sind, lange vor dem Ausbruch der Krankheit. Diese longitudinal stabilen differentiell methylierten Regionen könnten

in Zukunft möglicherweise als frühe prädiktive Biomarker zur Stratifizierung des Risikos für die Entwicklung von Asthma im späteren Leben verwendet werden.

Zusammenfassend konnte diese Dissertation einen wichtigen Beitrag zur Früherkennung von Kindern mit erhöhtem Asthma Risiko liefern und zeigen, dass DNA-Methylierungsveränderungen insbesondere in Enhancer-Regionen an der frühkindlichen Asthmaentwicklung beteiligt sind. Hervorzuheben ist, dass diese Studie zum ersten Mal genomweite DNA-Methylierungsdaten mit Hilfe eines ganzheitlichen Sequenzierungsansatzes untersuchte, welcher es ermöglichte neue regulatorische Regionen zu identifizieren, die von den bisher verwendeten Array-basierten Methoden nicht abgedeckt wurden.

7 LITERATURE

1. The Global Asthma Report 2018. Auckland NZG, Asthma Network. Global Asthma Report 2018. 2018.

2. WHO. WHO, editor. Global surveillance, prevention and control of chronic respiratory diseases: a comprehensive approach. Switzerland: WHO; 2007. 2007.

3. Schmitz R, Thamm M, Ellert U, Kalcklosch M, Schlaud M, Ki GGSSG. [Prevalence of common allergies in children and adolescents in Germany: results of the KiGGS study: first follow-up (KiGGS Wave 1)]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2014;57(7):771-8.

4. Almqvist C, Worm M, Leynaert B. Impact of gender on asthma in childhood and adolescence: a GA2LEN review. Allergy. 2008;63(1):47-57.

5. Barnes PJ. Immunology of asthma and chronic obstructive pulmonary disease. Nature reviews Immunology. 2008;8.

6. Haktanir Abul M, Phipatanakul W. Severe asthma in children: Evaluation and management. Allergology international : official journal of the Japanese Society of Allergology. 2019;68(2):150-7.

7. (GINA) GIfA. Global strategy for asthma management and prevention. 2019.

8. Gautier C, Charpin D. Environmental triggers and avoidance in the management of asthma. Journal of asthma and allergy. 2017;10:47-56.

9. Kuruvilla ME, Lee FE-H, Lee GB. Understanding Asthma Phenotypes, Endotypes, and Mechanisms of Disease. Clin Rev Allergy Immunol. 2019;56(2):219-33.

10. Wang W, Li Y, Lv Z, Chen Y, Li Y, Huang K, et al. Bronchial Allergen Challenge of Patients with Atopic Asthma Triggers an Alarmin (IL-33, TSLP, and IL-25) Response in the Airways Epithelium and Submucosa. J Immunol. 2018;201(8):2221-31.

11. Porsbjerg CM, Sverrild A, Lloyd CM, Menzies-Gow AN, Bel EH. Anti-alarmins in asthma: targeting the airway epithelium with next-generation biologics. European Respiratory Journal. 2020;56(5):2000260.

12. Djukanovic R, Hanania N, Busse W, Price D. IgE-mediated asthma: New revelations and future insights. Respiratory medicine. 2016;112:128-9.

13. Geha RS, Jabara HH, Brodeur SR. The regulation of immunoglobulin E class-switch recombination. Nature Reviews Immunology. 2003;3(9):721-32.

14. Holgate ST. A brief history of asthma and its mechanisms to modern concepts of disease pathogenesis. Allergy, asthma & immunology research. 2010;2(3):165-71.

15. Barnes PJ. Immunology of asthma and chronic obstructive pulmonary disease. Nature Reviews Immunology. 2008;8(3):183-92.

16. Hudon Thibeault AA, Laprise C. Cell-Specific DNA Methylation Signatures in Asthma. Genes (Basel). 2019;10(11).

17. Kyoto Encyclopedia of Genes and Genomes Kd. KEGG PATHWAY: Asthma - Homo sapiens (human) 2020 [Available from: https://www.genome.jp/dbget-bin/www_bget?pathway+hsa05310.

18. Bergeron C, Tulic MK, Hamid Q. Airway remodelling in asthma: from benchside to clinical practice. Can Respir J. 2010;17(4):e85-e93.

19. Fehrenbach H, Wagner C, Wegmann M. Airway remodeling in asthma: what really matters. Cell Tissue Res. 2017;367(3):551-69.

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

21. Ober C, Yao T-C. The genetics of asthma and allergic disease: a 21st century perspective. Immunological reviews. 2011;242(1):10-30.

22. Vercelli D. Discovering susceptibility genes for asthma and allergy. Nature reviews Immunology. 2008;8(3):169-82.

23. Shastry BS. SNPs: impact on gene function and phenotype. Methods Mol Biol. 2009;578:3-22.

24. Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS Results: A Review of Statistical

Methods and Recommendations for Their Application. The American Journal of Human Genetics. 2010;86(1):6-22.

25. Thomsen SF, van der Sluis S, Kyvik KO, Backer V. A study of asthma severity in adult twins. Clin Respir J. 2012;6(4):228-37.

26. van der Sluis S, Verhage M, Posthuma D, Dolan CV. Phenotypic complexity, measurement bias, and poor phenotypic resolution contribute to the missing heritability problem in genetic association studies. PloS one. 2010;5(11):e13929-e.

27. Vineis P, Pearce N. Missing heritability in genome-wide association study research. Nat Rev Genet. 2010;11(8):589.

28. 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.

29. Nica AC, Dermitzakis ET. Expression quantitative trait loci: present and future. Philos Trans R Soc Lond B Biol Sci. 2013;368(1620):20120362-.

30. Kim KW, Ober C. Lessons Learned From GWAS of Asthma. Allergy, asthma & immunology research. 2019;11(2):170-87.

31. Yang R, Tan M, Xu J, Zhao X. Investigating the regulatory role of ORMDL3 in airway barrier dysfunction using in vivo and in vitro models. International journal of molecular medicine. 2019;44(2):535-48.

32. Sleiman PM, Flory J, Imielinski M, Bradfield JP, Annaiah K, Willis-Owen SA, et al. Variants of DENND1B associated with asthma in children. N Engl J Med. 2010;362(1):36-44.

33. Ferreira MA, Matheson MC, Duffy DL, Marks GB, Hui J, Le Souef P, et al. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. Lancet (London, England). 2011;378(9795):1006-14.

34. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. Nature Reviews Genetics. 2019;20(8):467-84.

35. Durham AL, Wiegman C, Adcock IM. Epigenetics of asthma. Biochim Biophys Acta.

2011;1810(11):1103-9.

36. Douros K, Moustaki M, Tsabouri S, Papadopoulou A, Papadopoulos M, Priftis KN. Prenatal Maternal Stress and the Risk of Asthma in Children. Front Pediatr. 2017;5:202-.

37. Rosa MJ, Lee AG, Wright RJ. Evidence establishing a link between prenatal and early-life stress and asthma development. Curr Opin Allergy Clin Immunol. 2018;18(2):148-58.

38. Trump S, Bieg M, Gu Z, Thurmann L, Bauer T, Bauer M, et al. Prenatal maternal stress and wheeze in children: novel insights into epigenetic regulation. Sci Rep. 2016;6:28616.

39. Thacher JD, Gehring U, Gruzieva O, Standl M, Pershagen G, Bauer CP, et al. Maternal Smoking during Pregnancy and Early Childhood and Development of Asthma and Rhinoconjunctivitis - a MeDALL Project. Environmental health perspectives. 2018;126(4):047005.

40. Braback L, Lodge CJ, Lowe AJ, Dharmage SC, Olsson D, Forsberg B. Childhood asthma and smoking exposures before conception-A three-generational cohort study. Pediatr Allergy Immunol. 2018;29(4):361-8.

41. Kim A, Lim G, Oh I, Kim Y, Lee T, Lee J. Perinatal factors and the development of childhood asthma. Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology. 2018;120(3):292-9.

42. Thakur N, Oh SS, Nguyen EA, Martin M, Roth LA, Galanter J, et al. Socioeconomic status and childhood asthma in urban minority youths. The GALA II and SAGE II studies. Am J Respir Crit Care Med. 2013;188(10):1202-9.

43. Chen JH. Asthma and child behavioral skills: does family socioeconomic status matter? Social science & medicine (1982). 2014;115:38-48.

44. Lumia M, Takkinen HM, Luukkainen P, Kaila M, Lehtinen-Jacks S, Nwaru BI, et al. Food consumption and risk of childhood asthma. Pediatr Allergy Immunol. 2015;26(8):789-96.

45. Lang JE, Bunnell HT, Hossain MJ, Wysocki T, Lima JJ, Finkel TH, et al. Being Overweight or Obese and the Development of Asthma. Pediatrics. 2018;142(6).

46. van Meel ER, den Dekker HT, Elbert NJ, Jansen PW, Moll HA, Reiss IK, et al. A

population-based prospective cohort study examining the influence of early-life respiratory tract infections on school-age lung function and asthma. Thorax. 2018;73(2):167.

47. Guarnieri M, Balmes JR. Outdoor air pollution and asthma. Lancet (London, England). 2014;383(9928):1581-92.

48. Sheehan WJ, Phipatanakul W. Indoor allergen exposure and asthma outcomes. Current opinion in pediatrics. 2016;28(6):772-7.

49. O'Connor GT, Lynch SV, Bloomberg GR, Kattan M, Wood RA, Gergen PJ, et al. Early-life home environment and risk of asthma among inner-city children. The Journal of allergy and clinical immunology. 2018;141(4):1468-75.

50. 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.

51. Eccleston A, DeWitt N, Gunter C, Marte B, Nath D. Epigenetics. Nature. 2007;447(7143):395-.

52. Tammen SA, Friso S, Choi S-W. Epigenetics: the link between nature and nurture. Mol Aspects Med. 2013;34(4):753-64.

53. Handy DE, Castro R, Loscalzo J. Epigenetic modifications: basic mechanisms and role in cardiovascular disease. Circulation. 2011;123(19):2145-56.

54. Hagood JS. Beyond the genome: epigenetic mechanisms in lung remodeling. Physiology (Bethesda). 2014;29(3):177-85.

55. Martin EM, Fry RC. Environmental Influences on the Epigenome: Exposure- Associated DNA Methylation in Human Populations. Annual Review of Public Health. 2018;39(1):309-33.

56. Chen Z-x, Riggs AD. DNA methylation and demethylation in mammals. The Journal of biological chemistry. 2011;286(21):18347-53.

57. Bernstein BE, Meissner A, Lander ES. The Mammalian Epigenome. Cell. 2007;128(4):669-81.

58. Esteller M. Epigenetics in evolution and disease. The Lancet. 2008;372:S90-S6.

59. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13(7):484-92.

60. Portela A, Esteller M. Epigenetic modifications and human disease. Nat Biotechnol. 2010;28.

61. Choy JS, Wei S, Lee JY, Tan S, Chu S, Lee T-H. DNA methylation increases nucleosome compaction and rigidity. J Am Chem Soc. 2010;132(6):1782-3.

62. Alaskhar Alhamwe B, Khalaila R, Wolf J, von Bülow V, Harb H, Alhamdan F, et al. Histone modifications and their role in epigenetics of atopy and allergic diseases. Allergy, Asthma & Clinical Immunology. 2018;14(1):39.

63. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997;389(6648):251-60.

64. Salam MT, Zhang Y, Begum K. Epigenetics and childhood asthma: current evidence and future research directions. Epigenomics. 2012;4(4):415-29.

65. Hausser J, Syed AP, Bilen B, Zavolan M. Analysis of CDS-located miRNA target sites suggests that they can effectively inhibit translation. Genome Res. 2013;23(4):604-15.

66. Doane AS, Elemento O. Regulatory elements in molecular networks. Wiley Interdiscip Rev Syst Biol Med. 2017;9(3):10.1002/wsbm.374.

67. Spicuglia S, Vanhille L. Chromatin signatures of active enhancers. Nucleus. 2012;3(2):126-31.

68. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013;38(1):23-38.

69. Jjingo D, Conley AB, Yi SV, Lunyak VV, Jordan IK. On the presence and role of human gene-body DNA methylation. Oncotarget. 2012;3(4):462-74.

70. Stefanowicz D, Lee JY, Lee K, Shaheen F, Koo H-K, Booth S, et al. Elevated H3K18 acetylation in airway epithelial cells of asthmatic subjects. Respiratory research. 2015;16(1):95.

71. Yang Y, Wicks J, Haitchi HM, Powell RM, Manuyakorn W, Howarth PH, et al. Regulation of a disintegrin and metalloprotease-33 expression by transforming growth factor-β. Am J Respir Cell Mol Biol. 2012;46(5):633-40.

72. Seumois G, Chavez L, Gerasimova A, Lienhard M, Omran N, Kalinke L, et al. Epigenomic analysis of primary human T cells reveals enhancers associated with TH2 memory cell differentiation and asthma susceptibility. Nature immunology. 2014;15(8):777-88.

73. Harb H, Raedler D, Ballenberger N, Bock A, Kesper DA, Renz H. Childhood allergic asthma is associated with increased IL-13 and FOXP3 histone acetylation. The Journal of allergy and clinical immunology. 2015;136.

74. Maneechotesuwan K. Role of microRNA in severe asthma. Respiratory Investigation. 2019;57(1):9-19.

75. Murphy TM, Wong CC, Arseneault L, Burrage J, Macdonald R, Hannon E, et al. Methylomic markers of persistent childhood asthma: a longitudinal study of asthma-discordant monozygotic twins. Clinical epigenetics. 2015;7:130.

76. 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. The Journal of allergy and clinical immunology. 2018.

77. Yang IV, Pedersen BS, Liu AH, O'Connor GT, Pillai D, Kattan M, et al. The nasal methylome and childhood atopic asthma. The Journal of allergy and clinical immunology. 2017;139(5):1478-88.

78. Wilhelm-Benartzi CS, Koestler DC, Karagas MR, Flanagan JM, Christensen BC, Kelsey KT, et al. Review of processing and analysis methods for DNA methylation array data. British Journal of Cancer. 2013;109(6):1394-402.

79. Gomez JL. Epigenetics in Asthma. Current allergy and asthma reports. 2019;19(12):56.

80. Morales E, Bustamante M, Vilahur N, Escaramis G, Montfort M, de Cid R, et al. DNA hypomethylation at ALOX12 is associated with persistent wheezing in childhood. Am J Respir Crit Care Med. 2012;185(9):937-43.

81. Arathimos R, Suderman M, Sharp GC, Burrows K, Granell R, Tilling K, et al.

Epigenome-wide association study of asthma and wheeze in childhood and adolescence. Clinical epigenetics. 2017;9:112.

82. Xu CJ, Soderhall C, Bustamante M, Baiz N, Gruzieva O, Gehring U, et al. DNA methylation in childhood asthma: an epigenome-wide meta-analysis. The Lancet Respiratory medicine. 2018;6(5):379-88.

83. Yang IV, Pedersen BS, Liu A, O'Connor GT, Teach SJ, Kattan M, et al. DNA methylation and childhood asthma in the inner city. Journal of Allergy and Clinical Immunology. 2015;136(1):69-80.

84. Tepper RS, Wise RS, Covar R, Irvin CG, Kercsmar CM, Kraft M, et al. Asthma outcomes: pulmonary physiology. The Journal of allergy and clinical immunology. 2012;129(3 Suppl):S65-87.

85. Liang L, Willis-Owen SA, Laprise C, Wong KC, Davies GA, Hudson TJ, et al. An epigenome-wide association study of total serum immunoglobulin E concentration. Nature. 2015;520(7549):670-4.

86. Gao L, Millstein J, Siegmund KD, Dubeau L, Maguire R, Gilliland FD, et al. Epigenetic regulation of AXL and risk of childhood asthma symptoms. Clinical epigenetics. 2017;9:121.

87. Barton SJ, Ngo S, Costello P, Garratt E, El-Heis S, Antoun E, et al. DNA methylation of Th2 lineage determination genes at birth is associated with allergic outcomes in childhood. Clin Exp Allergy. 2017;47(12):1599-608.

88. Zhang X, Biagini Myers JM, Burleson JD, Ulm A, Bryan KS, Chen X, et al. Nasal DNA methylation is associated with childhood asthma. Epigenomics. 2018;10(5):629-41.

89. Forno E, Wang T, Qi C, Yan Q, Xu CJ, Boutaoui N, et al. DNA methylation in nasal epithelium, atopy, and atopic asthma in children: a genome-wide study. The Lancet Respiratory medicine. 2019;7(4):336-46.

90. Taussig LM, Wright AL, Holberg CJ, Halonen M, Morgan WJ, Martinez FD. Tucson Children's Respiratory Study: 1980 to present. The Journal of allergy and clinical immunology. 2003;111(4):661-75; quiz 76.

91. Bloomberg GR. Recurrent wheezing illness in preschool-aged children: assessment

and management in primary care practice. Postgraduate medicine. 2009;121(5):48-55.

92. Just J, Bourgoin-Heck M, Amat F. Clinical phenotypes in asthma during childhood. Clin Exp Allergy. 2017;47(7):848-55.

93. Peat JK, Toelle BG, Dermand J, van den Berg R, Britton WJ, Woolcock AJ. Serum IgE levels, atopy, and asthma in young adults: results from a longitudinal cohort study. Allergy. 1996;51(11):804-10.

94. Kusel MM, de Klerk NH, Kebadze T, Vohma V, Holt PG, Johnston SL, et al. Early-life respiratory viral infections, atopic sensitization, and risk of subsequent development of persistent asthma. J Allergy Clin Immunol. 2007;119(5):1105-10.

95. Martinez FD, Wright AL, Taussig LM, Holberg CJ, Halonen M, Morgan WJ. Asthma and wheezing in the first six years of life. The Group Health Medical Associates. N Engl J Med. 1995;332(3):133-8.

96. Cheong HS, Kim LH, Park BL, Choi YH, Park HS, Hong SJ, et al. Association analysis of interleukin 5 receptor alpha subunit (IL5RA) polymorphisms and asthma. J Hum Genet. 2005;50(12):628-34.

97. Bagnasco D, Ferrando M, Varricchi G, Puggioni F, Passalacqua G, Canonica GW. Anti-Interleukin 5 (IL-5) and IL-5Ra Biological Drugs: Efficacy, Safety, and Future Perspectives in Severe Eosinophilic Asthma. Front Med (Lausanne). 2017;4:135-.

98. Virkud YV, Kelly RS, Croteau-Chonka DC, Celedon JC, Dahlin A, Avila L, et al. Novel eosinophilic gene expression networks associated with IgE in two distinct asthma populations. Clin Exp Allergy. 2018;48(12):1654-64.

99. Herberth G, Hinz D, Roder S, Schlink U, Sack U, Diez U, et al. Maternal immune status in pregnancy is related to offspring's immune responses and atopy risk. Allergy. 2011;66(8):1065-74.

100. Weisse K, Lehmann I, Heroux D, Kohajda T, Herberth G, Roder S, et al. The LINA cohort: indoor chemical exposure, circulating eosinophil/basophil (Eo/B) progenitors and early life skin manifestations. Clin Exp Allergy. 2012;42(9):1337-46.

101. Brockow I, Zutavern A, Franke K, Schaaf B, von Berg A, Krämer U, et al. Einfluss von

Lebensbedingungen und Verhaltensweisen auf die Entwicklung von Immunsystem und Allergien im Ost-West-Vergleich (LISA). Monatsschrift KinderheilkundeZeitschrift für Kinder- und Jugendmedizin. 2008;156(3):249-55.

102. von Mutius E, Schmid S. The PASTURE project: EU support for the improvement of knowledge about risk factors and preventive factors for atopy in Europe. Allergy. 2006;61(4):407-13.

103. CORDIS EU research results. Mechanisms of early protective exposures on allergy development | EFRAIM Project | FP7 | CORDIS | European Commission 2019, July 16 [Available from: https://cordis.europa.eu/project/id/211911.

104. Martinez FD, Wright AL, Taussig LM, Holberg CJ, Halonen M, Morgan WJ. Asthma and wheezing in the first six years of life. The Group Health Medical Associates. N Engl J Med. 1995;332(3):133-8. .

105. Hinz D, Bauer M, Roder S, Olek S, Huehn J, Sack U, et al. Cord blood Tregs with stable FOXP3 expression are influenced by prenatal environment and associated with atopic dermatitis at the age of one year. Allergy. 2012;67(3):380-9.

106. Bauer T, Trump S, Ishaque N, Thurmann L, Gu L, Bauer M, et al. Environment-induced epigenetic reprogramming in genomic regulatory elements in smoking mothers and their children. Mol Syst Biol. 2016;12(3):861.

107. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60.

108. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. Sambamba: fast processing of NGS alignment formats. Bioinformatics. 2015;31(12):2032-4.

109. Wu H, Xu T, Feng H, Chen L, Li B, Yao B, et al. Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates. Nucleic Acids Res. 2015;43(21):e141.

110. Juhling F, Kretzmer H, Bernhart SH, Otto C, Stadler PF, Hoffmann S. metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. Genome Res. 2016;26(2):256-62.

111. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38(4):576-89.

112. Fullwood MJ, Han Y, Wei CL, Ruan X, Ruan Y. Chromatin interaction analysis using paired-end tag sequencing. Curr Protoc Mol Biol. 2010;Chapter 21:Unit 21.15.1-25.

113. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, et al. An oestrogen-receptor-alpha-bound human chromatin interactome. Nature. 2009;462(7269):58-64.

114. Li G, Cai L, Chang H, Hong P, Zhou Q, Kulakova EV, et al. Chromatin Interaction Analysis with Paired-End Tag (ChIA-PET) sequencing technology and application. BMC Genomics. 2014;15(12):S11.

115. Ho L, Crabtree GR. Chromatin remodelling during development. Nature. 2010;463(7280):474-84.

116. Geiman TM, Robertson KD. Chromatin remodeling, histone modifications, and DNA methylation-how does it all fit together? Journal of cellular biochemistry. 2002;87(2):117-25.

117. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, et al. The accessible chromatin landscape of the human genome. Nature. 2012;489(7414):75-82.

118. Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res. 2012;22(9):1813-31.

119. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74.

120. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539):317-30.

121. Sheffield NC, Bock C. LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. Bioinformatics. 2016;32(4):587-9.

122. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, et al. Software for Computing and Annotating Genomic Ranges. PLOS Computational Biology.

2013;9(8):e1003118.

123. Ziller MJ, Hansen KD, Meissner A, Aryee MJ. Coverage recommendations for methylation analysis by whole-genome bisulfite sequencing. Nat Methods. 2015;12(3):230-2, 1 p following 2.

124. Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc Natl Acad Sci U S A. 2005;102(44):15785-90.

125. Liu Y, Siegmund KD, Laird PW, Berman BP. Bis-SNP: combined DNA methylation and SNP calling for Bisulfite-seq data. Genome biology. 2012;13(7):R61.

126. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 2001;29(1):308-11.

127. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME Suite: tools for motif discovery and searching. Nucleic Acids Research. 2009;37(suppl_2):W202-W8.

128. Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics. 2011;27(12):1696-7.

129. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010;28(5):495-501.

130. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498-504.

131. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research. 2003;13.

132. Wu G, Haw R. Functional Interaction Network Construction and Analysis for Disease Discovery. Methods Mol Biol. 2017;1558:235-53.

133. Newman MEJ. Modularity and community structure in networks. Proceedings of the National Academy of Sciences. 2006;103(23):8577-82.

134. Dallago C, Goldberg T, Andrade-Navarro MA, Alanis-Lobato G, Rost B. CellMap visualizes protein-protein interactions and subcellular localization. F1000Res. 2017;6:1824.

135. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, et al. The reactome pathway knowledgebase. Nucleic Acids Res. 2020;48(D1):D498-d503.

136. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017;45(D1):D353-d61.

137. Schaefer CF, Anthony K, Krupa S, Buchoff J, Day M, Hannay T, et al. PID: the Pathway Interaction Database. Nucleic acids research. 2009;37(Database issue):D674-D9.

138. Mi H, Thomas P. PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. Methods in molecular biology (Clifton, NJ). 2009;563:123-40.

139. Nishimura D. BioCarta. Biotech Software & Internet Report. 2001;2(3):117-20.

140. NCBI. [Available from: https://www.ncbi.nlm.nih.gov/gene/about-generif.

141. Apache. [Available from: https://lucene.apache.org/.

142. Robertson SE, Walker S, Jones S, Hancock-Beaulieu M, Gatford M. Okapi at TREC-3. TREC. 1994;500-225:109-26.

143. R L, G G. - BANNER: an executable survey of advances in biomedical named entity recognition. Pac Symp Biocomput. 2008:652-63.

144. Tischer C, Standl M, Lehmann I, Schaaf B, von Berg A, Heinrich J. Sleeping on animal fur is related to asthma outcomes in later childhood. Eur Respir J. 2015;46(1):107-14.

145. Hayes AF. Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. New York, NY, US: Guilford Press; 2013. xvii, 507-xvii, p.

146. Lim IY, Lin X, Karnani N. Implications of Genotype and Environment on Variation in DNA Methylation. In: Patel V, Preedy V, editors. Handbook of Nutrition, Diet, and Epigenetics. Cham: Springer International Publishing; 2017. p. 1-20.

147. Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, et al. DNA

methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome biology. 2011;12(1):R10.

148. Wang H, Lou D, Wang Z. Crosstalk of Genetic Variants, Allele-Specific DNA Methylation, and Environmental Factors for Complex Disease Risk. Frontiers in genetics. 2019;9:695-.

149. Chu L, Wang T, Hu Y, Gu Y, Su Z, Jiang H. Activation of Egr-1 in human lung epithelial cells exposed to silica through MAPKs signaling pathways. PloS one. 2013;8(7):e68943.

150. Nakamura Y, Ghaffar O, Olivenstein R, Taha RA, Soussi-Gounni A, Zhang D-H, et al. Gene expression of the GATA-3 transcription factor is increased in atopic asthma. Journal of Allergy and Clinical Immunology. 1999;103(2):215-22.

151. Diehl S, Krahl T, Rinaldi L, Norton R, Irvin CG, Rincón M. Inhibition of NFAT Specifically in T Cells Prevents Allergic Pulmonary Inflammation. The Journal of Immunology. 2004;172(6):3597-603.

152. Bergauer A, Sopel N, Kroß B, Vuorinen T, Xepapadaki P, Weiss ST, et al. IFN-α/IFN-λ responses to respiratory viruses in paediatric asthma. Eur Respir J. 2017;49(2).

153. Zhao Y, Xiong Z, Lechner EJ, Klenotic PA, Hamburg BJ, Hulver M, et al. Thrombospondin-1 triggers macrophage IL-10 production and promotes resolution of experimental lung injury. Mucosal Immunology. 2014;7(2):440-8.

154. Rouillard AD, Gundersen GW, Fernandez NF, Wang Z, Monteiro CD, McDermott MG, et al. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database. 2016;2016.

155. Temesi G, Virág V, Hadadi E, Ungvári I, Fodor LE, Bikov A, et al. Novel genes in Human Asthma Based on a Mouse Model of Allergic Airway Inflammation and Human Investigations. Allergy, asthma & immunology research. 2014;6(6):496-503.

156. Lee SH, Jang AS, Woo Park S, Park JS, Kim YK, Uh ST, et al. Genetic effect of single-nucleotide polymorphisms in the PPARGC1B gene on airway hyperreactivity in asthmatic patients. Clin Exp Allergy. 2011;41(11):1533-44.

157. Dong X, Xu M, Ren Z, Gu J, Lu M, Lu Q, et al. Regulation of CBL and ESR1 expression

by microRNA-22-3p, 513a-5p and 625-5p may impact the pathogenesis of dust mite-induced pediatric asthma. International journal of molecular medicine. 2016;38(2):446-56.

158. Sasi W, Sharma AK, Mokbel K. The Role of Suppressors of Cytokine Signalling in Human Neoplasms. Molecular Biology International. 2014;2014:630797.

159. Collison A, Li J, Pereira de Siqueira A, Zhang J, Toop HD, Morris JC, et al. Tumor necrosis factor-related apoptosis-inducing ligand regulates hallmark features of airways remodeling in allergic airways disease. Am J Respir Cell Mol Biol. 2014;51(1):86-93.

160. Pelaia C, Paoletti G, Puggioni F, Racca F, Pelaia G, Canonica GW, et al. Interleukin-5 in the Pathophysiology of Severe Asthma. Frontiers in physiology. 2019;10:1514-.

161. Chen W, Wang T, Pino-Yanes M, Forno E, Liang L, Yan Q, et al. An epigenome-wide association study of total serum IgE in Hispanic children. The Journal of allergy and clinical immunology. 2017;140(2):571-7.

162. Cardenas A, Sordillo JE, Rifas-Shiman SL, Chung W, Liang L, Coull BA, et al. The nasal methylome as a biomarker of asthma and airway inflammation in children. Nat Commun. 2019;10(1):3095.

163. Gaurav R, Bewtra AK, Agrawal DK. Novel CLC3 transcript variants in blood eosinophils and increased CLC3 expression in nasal lavage and blood eosinophils of asthmatics. Immunity, inflammation and disease. 2014;2(4):205-13.

164. Du Y, Zhao J, Li X, Jin S, Ma WL, Mu Q, et al. Dissociation of FK506-binding protein 12.6 kD from ryanodine receptor in bronchial smooth muscle cells in airway hyperresponsiveness in asthma. Am J Respir Cell Mol Biol. 2014;50(2):398-408.

165. Song W, Zheng S, Li M, Zhang X, Cao R, Ye C, et al. Linking endotypes to omics profiles in difficult-to-control asthma using the diagnostic Chinese medicine syndrome differentiation algorithm. J Asthma. 2020;57(5):532-42.

166. Quan-San Z, Xiaohong X, Ying L, Zhaojia S. Role of Th17-cell related cytokines in geriatric asthma. J Int Med Res. 2019;47(2):580-90.

167. McErlean P, Favoreto S, Jr., Costa FF, Shen J, Quraishi J, Biyasheva A, et al. Human rhinovirus infection causes different DNA methylation changes in nasal epithelial cells from

healthy and asthmatic subjects. BMC Med Genomics. 2014;7:37-.

168. Junge KM, Bauer T, Geissler S, Hirche F, Thürmann L, Bauer M, et al. Increased vitamin D levels at birth and in early infancy increase offspring allergy risk—evidence for involvement of epigenetic mechanisms. Journal of Allergy and Clinical Immunology. 2016;137(2):610-3.

169. Junge KM, Leppert B, Jahreis S, Wissenbach DK, Feltens R, Grützmann K, et al. MEST mediates the impact of prenatal bisphenol A exposure on long-term body weight development. Clinical epigenetics. 2018;10:58.

170. 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.

171. Miyake K, Kawaguchi A, Miura R, Kobayashi S, Tran NQV, Kobayashi S, et al. Association between DNA methylation in cord blood and maternal smoking: The Hokkaido Study on Environment and Children's Health. Scientific Reports. 2018;8(1):5654.

172. de Vocht F, Simpkin AJ, Richmond RC, Relton C, Tilling K. Assessment of Offspring DNA Methylation across the Lifecourse Associated with Prenatal Maternal Smoking Using Bayesian Mixture Modelling. Int J Environ Res Public Health. 2015;12(11):14461-76.

173. Küpers LK, Xu X, Jankipersadsing SA, Vaez A, la Bastide-van Gemert S, Scholtens S, et al. DNA methylation mediates the effect of maternal smoking during pregnancy on birthweight of the offspring. International journal of epidemiology. 2015;44(4):1224-37.

174. Lee KW, Richmond R, Hu P, French L, Shin J, Bourdon C, et al. Prenatal exposure to maternal cigarette smoking and DNA methylation: epigenome-wide association in a discovery sample of adolescents and replication in an independent cohort at birth through 17 years of age. Environmental health perspectives. 2015;123(2):193-9.

175. Markunas CA, Xu Z, Harlid S, Wade PA, Lie RT, Taylor JA, et al. Identification of DNA methylation changes in newborns related to maternal smoking during pregnancy. Environmental health perspectives. 2014;122(10):1147-53.

176. Reese SE, Zhao S, Wu MC, Joubert BR, Parr CL, Håberg SE, et al. DNA Methylation Score as a Biomarker in Newborns for Sustained Maternal Smoking during Pregnancy. Environmental health perspectives. 2017;125(4):760-6.

177. Xu CJ, Bonder MJ, Söderhäll C, Bustamante M, Baïz N, Gehring U, et al. The emerging landscape of dynamic DNA methylation in early childhood. BMC Genomics. 2017;18(1):25.

178. Vangeel EB, Pishva E, Hompes T, van den Hove D, Lambrechts D, Allegaert K, et al. Newborn genome-wide DNA methylation in association with pregnancy anxiety reveals a potential role for GABBR1. Clinical epigenetics. 2017;9(1):107.

179. Kertes DA, Bhatt SS, Kamin HS, Hughes DA, Rodney NC, Mulligan CJ. BNDF methylation in mothers and newborns is associated with maternal exposure to war trauma. Clinical epigenetics. 2017;9:68.

180. Viuff AC, Sharp GC, Rai D, Henriksen TB, Pedersen LH, Kyng KJ, et al. Maternal depression during pregnancy and cord blood DNA methylation: findings from the Avon Longitudinal Study of Parents and Children. Translational Psychiatry. 2018;8(1):244.

181. Non AL, Binder AM, Kubzansky LD, Michels KB. Genome-wide DNA methylation in neonates exposed to maternal depression, anxiety, or SSRI medication during pregnancy. Epigenetics. 2014;9(7):964-72.

182. Cardenas A, Faleschini S, Cortes Hidalgo A, Rifas-Shiman SL, Baccarelli AA, DeMeo DL, et al. Prenatal maternal antidepressants, anxiety, and depression and offspring DNA methylation: epigenome-wide associations at birth and persistence into early childhood. Clinical epigenetics. 2019;11(1):56.

183. Walton E, Pingault JB, Cecil CA, Gaunt TR, Relton CL, Mill J, et al. Epigenetic profiling of ADHD symptoms trajectories: a prospective, methylome-wide study. Molecular psychiatry. 2017;22(2):250-6.

184. Gruzieva O, Xu CJ, Breton CV, Annesi-Maesano I, Anto JM, Auffray C, et al. Epigenome-Wide Meta-Analysis of Methylation in Children Related to Prenatal NO2 Air Pollution Exposure. Environmental health perspectives. 2017;125(1):104-10.

185. Peng C, den Dekker M, Cardenas A, Rifas-Shiman SL, Gibson H, Agha G, et al. Residential Proximity to Major Roadways at Birth, DNA Methylation at Birth and Midchildhood, and Childhood Cognitive Test Scores: Project Viva(Massachusetts, USA). Environmental health perspectives. 2018;126(9):97006.

186. Ladd-Acosta C, Feinberg JI, Brown SC, Lurmann FW, Croen LA, Hertz-Picciotto I, et

al. Epigenetic marks of prenatal air pollution exposure found in multiple tissues relevant for child health. Environ Int. 2019;126:363-76.

187. 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. Journal of Allergy and Clinical Immunology. 2017.

188. Bozack AK, Cardenas A, Quamruzzaman Q, Rahman M, Mostofa G, Christiani DC, et al. DNA methylation in cord blood as mediator of the association between prenatal arsenic exposure and gestational age. Epigenetics. 2018;13(9):923-40.

189. Kaushal A, Zhang H, Karmaus WJJ, Everson TM, Marsit CJ, Karagas MR, et al. Genome-wide DNA methylation at birth in relation to in utero arsenic exposure and the associated health in later life. Environmental health : a global access science source. 2017;16(1):50.

190. Rojas D, Rager JE, Smeester L, Bailey KA, Drobná 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. Toxicological sciences : an official journal of the Society of Toxicology. 2015;143(1):97-106.

191. Cardenas A, Koestler DC, Houseman EA, Jackson BP, Kile ML, Karagas MR, et al. Differential DNA methylation in umbilical cord blood of infants exposed to mercury and arsenic in utero. Epigenetics. 2015;10(6):508-15.

192. Sen A, Heredia N, Senut MC, Land S, Hollocher K, Lu X, et al. Multigenerational epigenetic inheritance in humans: DNA methylation changes associated with maternal exposure to lead can be transmitted to the grandchildren. Sci Rep. 2015;5:14466.

193. Wu S, Hivert MF, Cardenas A, Zhong J, Rifas-Shiman SL, Agha G, et al. Exposure to Low Levels of Lead in Utero and Umbilical Cord Blood DNA Methylation in Project Viva: An Epigenome-Wide Association Study. Environmental health perspectives. 2017;125(8):087019.

194. Kippler M, Engström K, Mlakar SJ, Bottai M, Ahmed S, Hossain MB, et al. Sex-specific effects of early life cadmium exposure on DNA methylation and implications for birth weight. Epigenetics. 2013;8(5):494-503.

195. Howe CG, Cox B, Fore R, Jungius J, Kvist T, Lent S, et al. Maternal Gestational Diabetes Mellitus and Newborn DNA Methylation: Findings From the Pregnancy and Childhood Epigenetics Consortium. Diabetes care. 2020;43(1):98-105.

196. de Goede OM, Lavoie PM, Robinson WP. Cord blood hematopoietic cells from preterm infants display altered DNA methylation patterns. Clinical epigenetics. 2017;9:39.

197. Wu Y, Lin X, Lim IY, Chen L, Teh AL, MacIsaac JL, et al. Analysis of two birth tissues provides new insights into the epigenetic landscape of neonates born preterm. Clinical epigenetics. 2019;11(1):26.

198. Solomon O, Yousefi P, Huen K, Gunier RB, Escudero-Fung M, Barcellos LF, et al. Prenatal phthalate exposure and altered patterns of DNA methylation in cord blood. Environmental and molecular mutagenesis. 2017;58(6):398-410.

199. Küpers 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.

200. Sharp GC, Salas LA, Monnereau C, Allard C, Yousefi P, Everson TM, et al. Maternal BMI at the start of pregnancy and offspring epigenome-wide DNA methylation: findings from the pregnancy and childhood epigenetics (PACE) consortium. Human molecular genetics. 2017;26(20):4067-85.

201. van Dijk SJ, Zhou J, Peters TJ, Buckley M, Sutcliffe B, Oytam Y, et al. Effect of prenatal DHA supplementation on the infant epigenome: results from a randomized controlled trial. Clinical epigenetics. 2016;8:114.

202. Ernst J, Kellis M. Discovery and characterization of chromatin states for systematic annotation of the human genome. Nature Biotechnology. 2010;28(8):817-25.

203. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature. 2011;473(7345):43-9.

204. Castro-Rodriguez JA, Cifuentes L, Martinez FD. Predicting Asthma Using Clinical Indexes. Front Pediatr. 2019;7(320).

205. Bush A, Grigg J, Saglani S. Managing wheeze in preschool children. BMJ. 2014;348:g15.

206. Passalacqua G, Ciprandi G. Allergy and the lung. Clin Exp Immunol. 2008;153 Suppl 1(Suppl 1):12-6.

207. Hamilton RG, Oppenheimer J. Serological IgE Analyses in the Diagnostic Algorithm for Allergic Disease. The journal of allergy and clinical immunology In practice. 2015;3(6):833-40; quiz 41-2.

208. Ansotegui IJ, Melioli G, Canonica GW, Caraballo L, Villa E, Ebisawa M, et al. IgE allergy diagnostics and other relevant tests in allergy, a World Allergy Organization position paper. World Allergy Organization Journal. 2020;13(2).

209. Lau S, Illi S, Sommerfeld C, Niggemann B, Völkel K, Madloch C, et al. Transient early wheeze is not associated with impaired lung function in 7-yr-old children. European Respiratory Journal. 2003;21(5):834-41.

210. Froidure A, Mouthuy J, Durham SR, Chanez P, Sibille Y, Pilette C. Asthma phenotypes and IgE responses. European Respiratory Journal. 2016;47(1):304-19.

211. Arroyave WD, Rabito FA, Carlson JC. The Relationship Between a Specific IgE Level and Asthma Outcomes: Results from the 2005-2006 National Health and Nutrition Examination Survey. The Journal of Allergy and Clinical Immunology: In Practice. 2013;1(5):501-8.

212. Bao Y, Chen Z, Liu E, Xiang L, Zhao D, Hong J. Risk Factors in Preschool Children for Predicting Asthma During the Preschool Age and the Early School Age: a Systematic Review and Meta-Analysis. Current allergy and asthma reports. 2017;17(12):85.

213. Boersma NA, Meijneke RWH, Kelder JC, van der Ent CK, Balemans WAF. Sensitization predicts asthma development among wheezing toddlers in secondary healthcare. Pediatric pulmonology. 2017;52(6):729-36.

214. Schoos AM, Jelding-Dannemand E, Stokholm J, Bønnelykke K, Bisgaard H, Chawes BL. Single and multiple time-point allergic sensitization during childhood and risk of asthma by age 13. Pediatr Allergy Immunol. 2019;30(7):716-23.

215. Skaaby T, Husemoen LLN, Thuesen BH, Fenger RV, Linneberg A. IgE sensitization to

inhalant allergens and the risk of airway infection and disease: A population-based study. PloS one. 2017;12(2):e0171525.

216. Collins SA, Pike KC, Inskip HM, Godfrey KM, Roberts G, Holloway JW, et al. Validation of novel wheeze phenotypes using longitudinal airway function and atopic sensitization data in the first 6 years of life: evidence from the Southampton Women's survey. Pediatric pulmonology. 2013;48(7):683-92.

217. Sherrill DL, Stein R, Halonen M, Holberg CJ, Wright A, Martinez FD. Total serum IgE and its association with asthma symptoms and allergic sensitization among children. Journal of Allergy and Clinical Immunology. 1999;104(1):28-36.

218. Kanda A, Driss V, Hornez N, Abdallah M, Roumier T, Abboud G, et al. Eosinophil-derived IFN-γ induces airway hyperresponsiveness and lung inflammation in the absence of lymphocytes. Journal of Allergy and Clinical Immunology. 2009;124(3):573-82.e9.

219. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419.

220. Matucci A, Maggi E, Vultaggio A. Eosinophils, the IL-5/IL-5Rα axis, and the biologic effects of benralizumab in severe asthma. Respiratory medicine. 2019;160:105819.

221. Jaffe AE, Murakami P, Lee H, Leek JT, Fallin MD, Feinberg AP, et al. Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. International journal of epidemiology. 2012;41(1):200-9.

222. Robinson MD, Kahraman A, Law CW, Lindsay H, Nowicka M, Weber LM, et al. Statistical methods for detecting differentially methylated loci and regions. Frontiers in genetics. 2014;5:324.

223. Scher H, Berman D, Weinberg EG, Schinkel M, Peper B, Chalton DO, et al. Granulocyte proteins in serum in childhood asthma: relation to spirometry and therapy. Clinical & Experimental Allergy. 1996;26(10):1131-41.

224. Kitamura N, Mori A, Tatsumi H, Nemoto S, Hiroi T, Kaminuma O. Zinc Finger Protein, Multitype 1, Suppresses Human Th2 Development via Downregulation of IL-4. International Archives of Allergy and Immunology. 2011;155(suppl 1)(Suppl. 1):53-6.

225. Sasaki H, Kurotaki D, Tamura T. Regulation of basophil and mast cell development by transcription factors. Allergology International. 2016;65(2):127-34.

226. Swarbrick C, Román N, Forwood J, editors. Role of ACOT7 in Arachidonic Acid Production and Inflammation2011.

227. Peng C, Cardenas A, Rifas-Shiman SL, Hivert MF, Gold DR, Platts-Mills TA, et al. Epigenome-wide association study of total serum immunoglobulin E in children: a life course approach. Clinical epigenetics. 2018;10:55.

228. Yang IV, Pedersen BS, Liu A, O'Connor GT, Teach SJ, Kattan M, et al. DNA methylation and childhood asthma in the inner city. The Journal of allergy and clinical immunology. 2015;136(1):69-80.

229. 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.

230. Kogan V, Millstein J, London SJ, Ober C, White SR, Naureckas ET, et al. Genetic-Epigenetic Interactions in Asthma Revealed by a Genome-Wide Gene-Centric Search. Human Heredity. 2018;83(3):130-52.

231. Junttila IS. Tuning the Cytokine Responses: An Update on Interleukin (IL)-4 and IL-13 Receptor Complexes. Frontiers in immunology. 2018;9:888.

232. Losol P, Kim S-H, Seob Shin Y, Min Ye Y, Park H-S. A genetic effect of IL-5 receptor α polymorphism in patients with aspirin-exacerbated respiratory disease. Experimental & Molecular Medicine. 2013;45(3):e14-e.

233. Zhang Y, Jing Y, Qiao J, Luan B, Wang X, Wang L, et al. Activation of the mTOR signaling pathway is required for asthma onset. Scientific Reports. 2017;7(1):4532.

234. Mouneimne G, Hansen SD, Selfors LM, Petrak L, Hickey MM, Gallegos LL, et al. Differential remodeling of actin cytoskeleton architecture by profilin isoforms leads to distinct effects on cell migration and invasion. Cancer Cell. 2012;22(5):615-30.

235. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays

and summary statistics 2019. Nucleic Acids Res. 2019;47(D1):D1005-d12.

236. Zhu Z, Guo Y, Shi H, Liu CL, Panganiban RA, Chung W, et al. Shared genetic and experimental links between obesity-related traits and asthma subtypes in UK Biobank. The Journal of allergy and clinical immunology. 2020;145(2):537-49.

237. Han Y, Jia Q, Jahani PS, Hurrell BP, Pan C, Huang P, et al. Genome-wide analysis highlights contribution of immune system pathways to the genetic architecture of asthma. Nat Commun. 2020;11(1):1776.

238. Ferreira MAR, Mathur R, Vonk JM, Szwajda A, Brumpton B, Granell R, et al. Genetic Architectures of Childhood- and Adult-Onset Asthma Are Partly Distinct. Am J Hum Genet. 2019;104(4):665-84.

239. Pividori M, Schoettler N, Nicolae DL, Ober C, Im HK. Shared and distinct genetic risk factors for childhood-onset and adult-onset asthma: genome-wide and transcriptome-wide studies. The Lancet Respiratory medicine. 2019;7(6):509-22.

240. Gagliardo R, Chanez P, Mathieu M, Bruno A, Costanzo G, Gougat C, et al. Persistent activation of nuclear factor-kappaB signaling pathway in severe uncontrolled asthma. Am J Respir Crit Care Med. 2003;168(10):1190-8.

241. Guo J, Lin X, Williams MA, Hamid Q, Georas SN. Yin-Yang 1 regulates effector cytokine gene expression and T(H)2 immune responses. The Journal of allergy and clinical immunology. 2008;122(1):195-201, .e1-5.

242. Garbacki N, Di Valentin E, Huynh-Thu VA, Geurts P, Irrthum A, Crahay C, et al. MicroRNAs profiling in murine models of acute and chronic asthma: a relationship with mRNAs targets. PloS one. 2011;6(1):e16509-e.

243. Murrison LB, Brandt EB, Myers JB, Hershey GKK. Environmental exposures and mechanisms in allergy and asthma development. The Journal of Clinical Investigation. 2019;129(4):1504-15.

244. Burbank AJ, Sood AK, Kesic MJ, Peden DB, Hernandez ML. Environmental determinants of allergy and asthma in early life. Journal of Allergy and Clinical Immunology. 2017;140(1):1-12.

245. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The Human Transcription Factors. Cell. 2018;172(4):650-65.

246. Caramori G, Casolari P, Adcock I. Role of Transcription Factors in the Pathogenesis of Asthma and COPD. Cell Communication & Adhesion. 2013;20(1-2):21-40.

247. Shrine N, Portelli MA, John C, Soler Artigas M, Bennett N, Hall R, et al. Moderate-to-severe asthma in individuals of European ancestry: a genome-wide association study. The Lancet Respiratory medicine. 2019;7(1):20-34.

248. Koch S, Graser A, Mirzakhani H, Zimmermann T, Melichar VO, Wolfel M, et al. Increased expression of nuclear factor of activated T cells 1 drives IL-9-mediated allergic asthma. The Journal of allergy and clinical immunology. 2016;137(6):1898-902.e7.

249. Landgraf-Rauf K, Boeck A, Siemens D, Klucker E, Vogelsang V, Schmidt S, et al. IRF-1 SNPs influence the risk for childhood allergic asthma: A critical role for pro-inflammatory immune regulation. Pediatr Allergy Immunol. 2018;29(1):34-41.

250. Chaudhari BR, Murphy RF, Agrawal DK. Following the TRAIL to apoptosis. Immunologic Research. 2006;35(3):249-62.

251. Martinez GJ, Zhang Z, Chung Y, Reynolds JM, Lin X, Jetten AM, et al. Smad3 differentially regulates the induction of regulatory and inflammatory T cell differentiation. J Biol Chem. 2009;284(51):35283-6.

252. Lloyd CM, Hessel EM. Functions of T cells in asthma: more than just T(H)2 cells. Nature reviews Immunology. 2010;10(12):838-48.

253. Chen L, Deng H, Cui H, Fang J, Zuo Z, Deng J, et al. Inflammatory responses and inflammation-associated diseases in organs. Oncotarget. 2017;9(6):7204-18.

254. Chan IH, Tang NL, Leung TF, Huang W, Lam YY, Wong GW, et al. Association of early growth response-1 gene polymorphisms with total IgE and atopy in asthmatic children. Pediatr Allergy Immunol. 2009;20(2):142-50.

255. Xue HH, Bollenbacher-Reilley J, Wu Z, Spolski R, Jing X, Zhang YC, et al. The transcription factor GABP is a critical regulator of B lymphocyte development. Immunity. 2007;26(4):421-31.

256. Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, et al. Leveraging Polygenic Functional Enrichment to Improve GWAS Power. Am J Hum Genet. 2019;104(1):65-75.

257. Zhang Y, Goss AM, Cohen ED, Kadzik R, Lepore JJ, Muthukumaraswamy K, et al. A Gata6-Wnt pathway required for epithelial stem cell development and airway regeneration. Nat Genet. 2008;40(7):862-70.

258. Hsieh YY, Wan L, Chang CC, Tsai CH, Tsai FJ. STAT2*C related genotypes and allele but not TLR4 and CD40 gene polymorphisms are associated with higher susceptibility for asthma. International journal of biological sciences. 2009;5(1):74-81.

259. Smet M, Van Hoecke L, De Beuckelaer A, Vander Beken S, Naessens T, Vergote K, et al. Cholesterol-sensing liver X receptors stimulate Th2-driven allergic eosinophilic asthma in mice. Immun Inflamm Dis. 2016;4(3):350-61.

260. Duan QL, Du R, Lasky-Su J, Klanderman BJ, Partch AB, Peters SP, et al. A polymorphism in the thyroid hormone receptor gene is associated with bronchodilator response in asthmatics. The pharmacogenomics journal. 2013;13(2):130-6.

261. Roussel L, Robins S, Schachter A, Berube J, Hamid Q, Rousseau S. Steroids and extracellular signal-regulated kinase 1/2 activity suppress activating transcription factor 3 expression in patients with severe asthma. The Journal of allergy and clinical immunology. 2011;127(6):1632-4.

262. Cui A-H, Zhao J, Liu S-X, Hao Y-S. Associations of IL-4, IL-6, and IL-12 levels in peripheral blood with lung function, cellular immune function, and quality of life in children with moderate-to-severe asthma. Medicine. 2017;96(12):e6265-e.

263. Trifilieff A, Bench A, Hanley M, Bayley D, Campbell E, Whittaker P. PPAR-alpha and -gamma but not -delta agonists inhibit airway inflammation in a murine model of asthma: in vitro evidence for an NF-kappaB-independent effect. British journal of pharmacology. 2003;139(1):163-71.

264. Dahgam S, Nyberg F, Modig L, Naluai AT, Olin AC. Single nucleotide polymorphisms in the NOS2 and NOS3 genes are associated with exhaled nitric oxide. Journal of medical genetics. 2012;49(3):200-5.

265. Johansson E, Biagini Myers JM, Martin LJ, He H, Pilipenko V, Mersha T, et al. KIF3A genetic variation is associated with pediatric asthma in the presence of eczema independent of allergic rhinitis. The Journal of allergy and clinical immunology. 2017;140(2):595-8.e5.

266. Li X, Hawkins GA, Moore WC, Hastie AT, Ampleford EJ, Milosevic J, et al. Expression of asthma susceptibility genes in bronchial epithelial cells and bronchial alveolar lavage in the Severe Asthma Research Program (SARP) cohort. J Asthma. 2016;53(8):775-82.

267. Jahreis S, Trump S, Bauer M, Bauer T, Thurmann L, Feltens R, et al. Maternal phthalate exposure promotes allergic airway inflammation over 2 generations through epigenetic modifications. The Journal of allergy and clinical immunology. 2018;141(2):741-53.

268. Xu W. Expression data analysis to identify biomarkers associated with asthma in children. International journal of genomics. 2014;2014:165175.

269. Yang IV, Tomfohr J, Singh J, Foss CM, Marshall HE, Que LG, et al. The clinical and environmental determinants of airway transcriptional profiles in allergic asthma. Am J Respir Crit Care Med. 2012;185(6):620-7.

270. Byersdorfer CA. The role of Fatty Acid oxidation in the metabolic reprograming of activated t-cells. Frontiers in immunology. 2014;5:641-.

271. Yick CY, Zwinderman AH, Kunst PW, Grunberg K, Mauad T, Chowdhury S, et al. Gene expression profiling of laser microdissected airway smooth muscle tissue in asthma and atopy. Allergy. 2014;69(9):1233-40.

272. Duan QL, Lasky-Su J, Himes BE, Qiu W, Litonjua AA, Damask A, et al. A genome-wide association study of bronchodilator response in asthmatics. The pharmacogenomics journal. 2014;14(1):41-7.

273. Pedros C, Altman A, Kong K-F. Role of TRAFs in Signaling Pathways Controlling T Follicular Helper Cell Differentiation and T Cell-Dependent Antibody Responses. Frontiers in immunology. 2018;9:2412-.

274. Junge KM, Lehmann I, Borte M. Can vitamin D intake during pregnancy affect the risk of allergy in children? Clin Immunol. 2013;9(8):699–701.

8 SUPPLEMENTARY MATERIAL

8.1 Supplemental tables

Table 8.1.1: ENCODE and ROADMAP enhancer sets used for enhancer analyses

Enhancer Roadmap Blood and T-cell
E062 Primary mononuclear cells from peripheral blood
E034 Primary T cells from peripheral blood
E045 Primary T cells effector memory enriched from peripheral blood
E033 Primary T cells from cord blood
E044 Primary T regulatory cells from peripheral blood
E043 Primary T helper cells from peripheral blood
E039 Primary T helper naïve cells from peripheral blood
E041 Primary T helper cells PMA-I stimulated
E042 Primary T helper 17 cells PMA-I stimulated
E040 Primary T helper memory cells from peripheral blood 1
E037 Primary T helper memory cells from peripheral blood 2
E048 Primary T CD8+ memory cells from peripheral blood
E038 Primary T helper naïve cells from peripheral blood
E047 Primary T CD8+ naïve cells from peripheral blood
Enhancer Roadmap HSC and B-cell
E029 Primary monocytes from peripheral blood
E031 Primary B cells from cord blood
E035 Primary hematopoietic stem cells
E051 Primary hematopoietic stem cells G-CSF-mobilized Male
E050 Primary hematopoietic stem cells G-CSF-mobilized Female
E036 Primary hematopoietic stem cells short term culture
E032 Primary B cells from peripheral blood
E046 Primary Natural Killer cells from peripheral blood
E030 Primary neutrophils from peripheral blood
Enhancer Encode
E114 A549 EtOH 0.02pct lung carcinoma
E115 Dnd41 T cell leukaemia
E116 GM12878 lymphoblastoid
E117 HeLa-S3 cervical carcinoma
E118 HepG2 hepatocellular carcinoma
E119 HMEC mammary epithelial
E120 HSMM skeletal muscle myoblasts
E121 HSMM-derived skeletal muscle myotubes
E122 HUVEC umbilical vein endothelial
E123 K562 leukaemia
E124 Monocytes-CD14+ RO01746
E125 NH-A astrocyte
E126 NHDF-ad adult dermal fibroblast
E127 NHEK-epidermal keratinocyte
E128 NHLF lung fibroblast
E129 Osteoblast

Table 8.1.2: Percentages of children with a positive sensitization to aeroallergens (>0.35 kU/l)

	Age	Controls	Transient wheeze	Late wheeze	Persistent wheeze	Asthma
LINA	1	2.0	0.0	0.0	0.0	10.0
	2	2.0	4.6	14.3	0.0	30.0
	3	5.9	9.1	14.3	12.5	50.0
	4	13.7	18.2	28.6	25.0	50.0
	5	15.7	22.7	57.1	12.5	60.0
	6	23.5	31.8	57.1	12.5	90.0
	8	25.5	40.9	71.4	37.5	90.0
LISA	2	0.4	3.2	2.2	1.1	8.2
	6	21.1	32.3	30.4	23.3	71.4
	10	33.8	45.2	43.5	36.7	81.6
	15	39.2	58.1	54.4	43.3	87.8

Table 8.1.3: Percentages of children with an elevated total serum IgE (>100 kU/l)

	Age	Controls	Transient wheeze	Late wheeze	Persistent wheeze	Asthma
LINA	1	5.9	13.6	0.0	12.5	20.0
	2	9.8	18.2	0.0	0.0	30.0
	3	9.8	13.6	0.0	0.0	40.0
	4	13.7	18.2	28.6	0.0	60.0
	5	17.7	31.8	14.3	12.5	60.0
	6	29.4	31.8	28.6	25.0	70.0
	8	29.4	31.8	42.9	25.0	80.0
LISA	2	10.1	11.1	8.7	16.1	20.4
	6	26.6	27.8	28.3	22.6	55.1
	10	32.9	34.4	37.0	25.8	65.3
	15	29.5	32.2	26.1	35.5	63.3

Table 8.1.4: List of 161 concordant differentially methylated regions (DMRs) detected in the LINA, LISA, and PASTURE cohort by whole genome bisulfit sequencing (WGBS).

No.	Chr	Start	End	Mean Methylation Difference	Mean methylation (*b*-value) Asthma	Mean methylation (*b*-value) Controls	Nearest Gene	Predicted Target Gene	Enhancer	ngDMR
1	chr1	67600310	67600658	-0.04	0.06	0.10	C1orf141	C1orf141	YES	-
2	chr1	46770750	46771041	-0.05	0.79	0.84	UQCRH	NSUN4;LRRC41;UQCRH;MAST2;POMGNT1;RAD54L;FAAH;MKNK1	YES	YES
3	chr1	93970706	93970976	-0.06	0.83	0.89	FNBP1L	CCDC18;TMED5	YES	YES
4	chr1	6341136	6341791	-0.06	0.80	0.85	GPR153	ICMT;HES3;ACOT7	YES	YES
5	chr1	84744687	84744807	-0.04	0.92	0.95	SAMD13	SAMD13	YES	YES
6	chr1	202121664	202121814	-0.05	0.86	0.91	PTPN7	PTPN7	YES	YES
7	chr1	22097059	22097226	-0.05	0.88	0.93	USP48	USP48	YES	YES
8	chr1	149162004	149162462	-0.09	0.46	0.55	NBPF25P	NBPF25P	-	YES
9	chr1	10436586	10436850	-0.05	0.86	0.91	PGD	APITD1-CORT;APITD1;PGD	-	YES
10	chr10	1404992	1406101	-0.08	0.40	0.48	ADARB2-AS1	LARP4B;GTPBP4;IDI1;WDR37;ADARB2	-	YES
11	chr10	3625837	3626735	-0.03	0.86	0.89	KLF6	KLF6	-	YES
12	chr10	134139414	134139778	-0.07	0.87	0.94	LRRC27	STK32C;LRRC27	-	YES
13	chr10	126593720	126593912	-0.04	0.87	0.91	ZRANB1	ZRANB1	-	-
14	chr11	128694096	128694401	-0.05	0.19	0.24	KCNJ1	KCNJ5;ARHGAP32	YES	-
15	chr11	65477123	65477451	-0.06	0.81	0.87	KAT5	KAT5;MUS81;CFL1;CCDC85B;FIBP;RELA;	YES	YES
16	chr11	132951692	132952491	-0.11	0.48	0.59	OPCML	OPCML	YES	YES
17	chr11	1828650	1828782	-0.06	0.79	0.86	SYT8	SYT8	YES	YES
18	chr11	68517647	68517834	-0.05	0.15	0.20	MTL5	C11orf24;TPCN2;CPT1A;MRPL21;IGHMBP2;GAL;	YES	YES
19	chr11	59560470	59560548	-0.07	0.85	0.92	MRPL16	MRPL16	-	YES
20	chr12	111137400	111137595	-0.06	0.87	0.93	HVCN1	TCTN1	YES	YES
21	chr12	16161553	16161814	-0.05	0.87	0.91	DERA	DERA	YES	YES
22	chr12	119591663	119592118	-0.04	0.17	0.21	HSPB8	SRRM4	YES	YES
23	chr12	125482583	125482828	-0.04	0.86	0.90	BRI3BP	BRI3BP	-	-
24	chr12	124905467	124905758	-0.04	0.88	0.92	MIR6880	MIR6880	-	YES
25	chr12	107273279	107273680	-0.04	0.90	0.94	TMEM263	TMEM263	-	YES
26	chr13	20988857	20989414	-0.09	0.52	0.60	MIR4499	XPO4	YES	YES

No.	Chr	Start	End	Mean Methylation Difference	Mean methylation (*b*-value) Asthma	Mean methylation (*b*-value) Controls	Nearest Gene	Predicted Target Gene	Enhancer	ngDMR
27	chr13	20968573	20969084	-0.08	0.50	0.58	MIR4499	GJA3;GJB6;CRYL1;IL17D	YES	YES
28	chr13	24914323	24914904	-0.08	0.67	0.75	LINC00566	MIPEP;SPATA13;PARP4	YES	YES
29	chr13	25670023	25670238	-0.04	0.11	0.15	PABPC3	PABPC3	YES	YES
30	chr13	111948367	111948558	-0.04	0.86	0.90	TEX29	ARHGEF7	YES	YES
31	chr14	75153156	75153307	-0.05	0.87	0.92	AREL1	FCF1;AREL1	YES	YES
32	chr14	91272379	91272909	-0.04	0.87	0.91	TTC7B	TTC7B	YES	YES
33	chr14	100610169	100610683	-0.05	0.85	0.90	DEGS2	SLC25A29;MIR345;DEGS2;YY1;EVL	YES	YES
34	chr14	93212312	93212486	-0.04	0.90	0.94	LGMN	LGMN	-	YES
35	chr14	103200841	103201127	-0.05	0.88	0.93	TRAF3	TRAF3	-	YES
36	chr15	57511786	57512215	-0.04	0.90	0.94	TCF12	TCF12	YES	YES
37	chr15	52707259	52707362	-0.05	0.89	0.94	MYO5A	MYO5A	YES	YES
38	chr15	31134409	31134667	-0.05	0.87	0.92	HERC2P10	HERC2P10	YES	YES
39	chr15	52872030	52872159	-0.04	0.91	0.95	ARPP19	ARPP19	-	YES
40	chr15	40093789	40094022	-0.06	0.88	0.94	FSIP1	EIF2AK4;FSIP1	-	YES
41	chr15	21359928	21360042	-0.10	0.67	0.77	LINC01193	LINC01193	-	YES
42	chr15	22646998	22647395	-0.05	0.13	0.18	MIR4509-1	MIR4509-1	-	YES
43	chr16	30464225	30464522	-0.08	0.59	0.66	SEPHS2	SEPHS2;ZNF785;ZNF688	YES	-
44	chr16	57831974	57832179	-0.08	0.53	0.61	KIFC3	LOC388282	YES	YES
45	chr16	2141333	2141613	-0.08	0.59	0.67	MIR1225	SNORD60;TRAF7;MIR1225;IGFALS;FAHD1;HAGH;NTHL1;TSC2;PKD1	YES	YES
46	chr16	88540019	88540525	-0.06	0.84	0.90	MIR5189	PIEZO1	YES	YES
47	chr16	30552372	30552612	-0.05	0.88	0.93	ZNF747	ZNF688	YES	YES
48	chr16	85654108	85654323	-0.07	0.83	0.90	GSE1	GSE1	YES	YES
49	chr16	69489543	69489664	-0.06	0.88	0.94	CYB5B	CYB5B	YES	YES
50	chr16	88579452	88580071	-0.04	0.84	0.88	MIR5189	ZFPM1	YES	YES
51	chr16	75026217	75026456	-0.06	0.80	0.86	ZNRF1	WDR59;ZNRF1	YES	YES
52	chr17	19627951	19628165	-0.05	0.10	0.15	SLC47A2	ALDH3A1;ALDH3A2;B9D1;ALDH3A2;SPECC1	YES	YES
53	chr17	36572579	36572896	-0.05	0.85	0.90	ARHGAP23	SOCS7;HNF1B;MLLT6	YES	YES

No.	Chr.	Start	End	Mean Methylation Difference	Mean methylation (*b*-value) Asthma	Mean methylation (*b*-value) Controls	Nearest Gene	Predicted Target Gene	Enhancer	ngDMR
54	chr17	17946397	17946584	-0.06	0.87	0.93	GID4	SREBF1;GID4;ATPAF2;SMCR7;LGALS9C	YES	YES
55	chr17	28580392	28580613	-0.05	0.88	0.93	SLC6A4	GOSR1	YES	YES
56	chr17	49057182	49057238	-0.05	0.88	0.94	SPAG9	SPAG9	YES	YES
57	chr17	8769570	8769883	-0.03	0.90	0.93	PIK3R6	PIK3R6	YES	YES
58	chr17	78569835	78569887	-0.06	0.85	0.91	RPTOR	RPTOR	YES	YES
59	chr17	21119605	21119864	-0.05	0.85	0.90	TMEM11	USP22;C17orf103;USP22	YES	YES
60	chr17	79466120	79466418	-0.08	0.60	0.68	ACTG1	SLC25A10;MRPL12	YES	YES
61	chr17	8702637	8702755	-0.12	0.26	0.38	MFSD6L	ODF4;NTN1	-	-
62	chr17	56272299	56272501	-0.05	0.85	0.91	EPX	EPX;MKS1	-	YES
63	chr17	56274149	56274597	-0.04	0.87	0.91	EPX	MKS1;BZRAP1;EPX	-	YES
64	chr17	81093746	81093809	-0.04	0.91	0.95	METRNL	METRNL	-	YES
65	chr17	56283478	56283522	-0.08	0.87	0.95	MKS1	DYNLL2	-	YES
66	chr18	12076364	12076621	-0.08	0.23	0.30	ANKRD62	CHMP1B;IMPA2	YES	YES
67	chr18	14458381	14458936	-0.03	0.04	0.07	CXADRP3	POTEC	YES	-
68	chr18	71910027	71910088	-0.07	0.83	0.90	CYB5A	CYB5A	YES	YES
69	chr18	8754970	8755342	-0.04	0.83	0.87	MTCL1	MTCL1	YES	YES
70	chr18	77703283	77703531	-0.05	0.90	0.94	PQLC1	PQLC1	YES	YES
71	chr18	22016574	22016799	-0.06	0.87	0.93	IMPACT	IMPACT	-	YES
72	chr18	59588681	59588813	-0.04	0.90	0.94	RNF152	RNF152	-	YES
73	chr19	3520495	3521153	-0.05	0.81	0.85	FZR1	CACTIN;C19orf71;FZR1;MFSD12	YES	YES
74	chr19	1854531	1854765	-0.05	0.81	0.86	REXO1	NDUFS7;TCF3;ATP8B3;MIR1909;DOT1L;TCF3;MKNK2;CSNK1G2;IZUMO4	YES	YES
75	chr19	51373740	51373966	-0.07	0.55	0.62	KLK2	ACPT;SNORD88A;SNORD88B;SNORD88C;C19orf48;KLK2;SHANK1;KLK12	YES	-
76	chr19	34859991	34860409	-0.05	0.81	0.86	GPI	PDCD2L;UBA2	YES	YES
77	chr19	4382715	4382767	-0.06	0.85	0.91	SH3GL1	CHAF1A;SH3GL1	YES	YES
78	chr19	10404092	10405284	-0.04	0.19	0.22	ICAM5	ICAM3;MIR1181;CDC37;ZGLP1;FDX1L;CDC37;ICAM1;ICAM4;ICAM5;RAVER1;COL5A3;PPAN;ZGLP1;TYK2	-	YES
79	chr2	8676830	8676988	-0.07	0.83	0.90	LOC101929567	ID2;HMGB1P25	YES	YES

No.	Chr.	Start	End	Mean Methylation Difference	Mean methylation (*b*-value) Asthma	Mean methylation (*b*-value) Controls	Nearest Gene	Predicted Target Gene	Enhancer	ngDMR
80	chr2	24233608	24234116	-0.07	0.54	0.60	MFSD2B	UBXN2A;ATAD2B;FAM228B;SF3B14;FKBP1B;C2orf44;TP53I3;MFSD2B	YES	YES
81	chr2	241459177	241460046	-0.05	0.32	0.37	ANKMY1	ANKMY1	YES	-
82	chr2	31154684	31155156	-0.06	0.87	0.93	CAPN13	CAPN13	YES	YES
83	chr2	113956545	113956672	-0.067031	0.22234	0.28937	PSD4	AHSA2;PSD4;LOC654433;NT5DC4;PSD4;CBWD2;RABL2A;RPL23AP7	YES	YES
84	chr2	74213621	74213840	0.07	0.70	0.63	TET3	MTHFD2;BOLA3-AS1;BOLA3;STAMBP;DGUOK;MTHFD2;	YES	YES
85	chr2	132404284	132404978	-0.06	0.26	0.31	LINC01087	NF1P8;POTEE;RHOQP2;MZT2A;POTEKP;C2orf27B	YES	YES
86	chr2	75089515	75089818	-0.05	0.85	0.90	HK2	C2orf81;HMGA1P8	YES	YES
87	chr2	26094522	26094609	-0.06	0.88	0.93	ASXL2	ASXL2;	YES	YES
88	chr2	106795368	106795456	-0.05	0.87	0.93	UXS1	UXS1	YES	YES
89	chr2	118617605	118618119	-0.06	0.39	0.45	DDX18	DDX18	-	YES
90	chr2	97401278	97401371	-0.05	0.88	0.93	LMAN2L	CNNM4;LMAN2L	-	YES
91	chr2	113426404	113426418	-0.08	0.85	0.93	SLC20A1	SLC20A1	-	YES
92	chr2	121816270	121816884	-0.05	0.81	0.85	TFCP2L1	GLI2	-	YES
93	chr2	70734255	70734340	-0.04	0.91	0.95	TGFA-IT1	TGFA	-	YES
94	chr2	122984491	122984673	-0.05	0.88	0.93	TSN	TSN	-	YES
95	chr2	130986715	130986827	-0.08	0.49	0.57	TUBA3E	MZT2B;SMPD4;TUBA3E;IMP4;PTPN18	-	YES
96	chr20	29525072	29525474	-0.09	0.19	0.28	LINC01598	LINC01598	YES	YES
97	chr20	33416638	33416741	-0.08	0.84	0.92	NCOA6	NCOA6;MMP24-AS1;EDEM2	YES	YES
98	chr20	61629182	61629312	-0.10	0.15	0.24	BHLHE23	BHLHE23;LOC63930;NTSR1;OGFR;COL9A3;SLC17A9;BIRC7;COL20A1	YES	-
99	chr20	32232295	32232457	-0.05	0.85	0.90	C20orf144	SOCS2P1;SNTA1;CBFA2T2;NECAB3;C20orf144;E2F1;PXMP4;RPL31P2;CHMP4B;TPM3P2;RALY;EIF2S2	YES	-
100	chr20	29515851	29515953	-0.14	0.48	0.62	LINC01598	LINC01598	-	YES
101	chr20	29550781	29551738	-0.05	0.16	0.21	LINC01598	LINC01598	-	YES
102	chr21	45705600	45705880	-0.09	0.33	0.42	AIRE	PFKL	YES	YES
103	chr21	19184847	19184908	-0.07	0.82	0.89	C21orf91	BTG3	YES	YES
104	chr21	38750599	38750876	-0.04	0.85	0.90	DYRK1A	DSCR3	YES	YES
105	chr21	30298129	30298293	-0.05	0.90	0.94	N6AMT1	N6AMT1	-	YES

No.	Chr.	Start	End	Mean Methylation Difference	Mean methylation (*b*-value) Asthma	Mean methylation (*b*-value) Controls	Nearest Gene	Predicted Target Gene	Enhancer	ngDMR
106	chr22	50985261	50985924	-0.10	0.57	0.67	KLHDC7B	TYMP;ODF3B;SCO2;CHKB-AS1;CHKB;CHKB-CPT1B	YES	YES
107	chr22	46762433	46763143	-0.11	0.61	0.72	TRMU	PPARA;GTSE1;CELSR1;CERK	YES	-
108	chr3	4790490	4790704	-0.06	0.75	0.81	EGOT	EGOT	YES	YES
109	chr3	184243755	184244148	-0.05	0.17	0.22	EIF2B5-AS1	CLCN2;POLR2H	YES	-
110	chr3	75445436	75445698	-0.11	0.36	0.47	FAM86DP	FAM86DP	YES	-
111	chr3	172243109	172243330	-0.06	0.84	0.90	TNFSF10	TNFSF10	YES	YES
112	chr3	128134844	128135028	-0.05	0.86	0.91	DNAJB8-AS1	GATA2	YES	YES
113	chr3	128317561	128317754	-0.05	0.88	0.93	LINC01565	GATA2	YES	YES
114	chr3	39395430	39395804	-0.06	0.85	0.91	CCR8	TTC21A;CX3CR1;CCR8;EEF1A1P24;SLC25A38;MOBP;MYRIP	-	YES
115	chr3	3150228	3150424	-0.05	0.88	0.93	IL5RA	TRNT1	-	YES
116	chr3	70560282	70560338	-0.05	0.88	0.93	LINC01212	LINC01212	-	YES
117	chr3	195964960	195965369	-0.04	0.89	0.93	SLC51A	RNF168;PCYT1A	-	YES
118	chr3	98476467	98476656	-0.05	0.89	0.93	ST3GAL6	ST3GAL6	-	YES
119	chr4	170695782	170696267	-0.03	0.14	0.18	C4orf27	CLCN3;MFAP3L	YES	YES
120	chr4	148634323	148634373	-0.06	0.88	0.94	ARHGAP10	ARHGAP10	-	YES
121	chr4	144833125	144833345	-0.06	0.14	0.19	LOC101927636	LOC101927636	-	YES
122	chr4	1908638	1908945	-0.06	0.86	0.92	WHSC1	WHSC1	-	YES
123	chr5	42943969	42944683	-0.08	0.33	0.41	FLJ32255	LOC153684;ANXA2R	YES	YES
124	chr5	8457869	8457979	-0.08	0.16	0.23	MIR4458HG	MIR4458	YES	YES
125	chr5	157117442	157117958	-0.07	0.42	0.49	C5orf52	C5orf52	YES	-
126	chr5	154224429	154224646	-0.05	0.88	0.93	FAXDC2	CNOT8;LARP1	YES	YES
127	chr5	132002374	132002506	-0.05	0.86	0.91	IL4	IL4;GDF9;UQCRQ;SEPT8;SHROOM1;RAD50;KIF3A;LEAP2;HSPA4	YES	YES
128	chr5	77146478	77147360	-0.08	0.56	0.64	LOC101929154	AP3B1	YES	-
129	chr5	68700315	68700723	-0.04	0.92	0.95	MARVELD2	MARVELD2	YES	YES
130	chr5	77142381	77142898	-0.08	0.51	0.59	LOC101929154	OTP;AP3B1	YES	-
131	chr5	140961391	140961495	-0.05	0.85	0.90	LOC100505658	LOC100505658	-	YES

No.	Chr.	Start	End	Mean Methylation Difference	Mean methylation (*b*-value) Asthma	Mean methylation (*b*-value) Controls	Nearest Gene	Predicted Target Gene	Enhancer	ngDMR
132	chr5	149145166	149145196	-0.06	0.88	0.94	PPARGC1B	PPARGC1B	-	YES
133	chr6	32063911	32064191	-0.05	0.21	0.26	TNXB	GPSM3	YES	-
134	chr6	166674955	166675082	-0.05	0.88	0.93	LOC101929297	LOC101929297	-	YES
135	chr6	52385312	52385483	-0.05	0.89	0.94	TRAM2	TRAM2	-	YES
136	chr7	1914009	1914392	-0.04	0.87	0.91	MIR4655	MAD1L1	YES	-
137	chr7	90895326	90896701	-0.06	0.67	0.73	FZD1	FZD1	YES	YES
138	chr7	102003600	102003766	-0.053187	0.82712	0.8803	LOC100289561	SH2B2	YES	YES
139	chr7	36007074	36007281	-0.09	0.36	0.45	LOC101928618	HERPUD2;ANLN	YES	YES
140	chr7	150647915	150648062	-0.04	0.89	0.94	KCNH2	STRADBP1;TMEM176B;TMEM176A;KCNH2;NOS3;ABCB8;SLC4A2;CDK5;FASTK;AGAP3;ASB10;ABCF2;SMARCD3;WDR86	YES	YES
141	chr7	48887537	48887890	-0.08	0.28	0.36	CDC14C	CDC14C	-	YES
142	chr7	55412705	55412995	-0.09	0.37	0.46	LANCL2	LANCL2	-	-
143	chr7	54900863	54901102	-0.04	0.84	0.89	LOC100996654	LOC100996654	-	YES
144	chr7	32357921	32358754	-0.07	0.21	0.28	PDE1C	PDE1C	-	YES
145	chr7	5382633	5382782	-0.05	0.90	0.95	SLC29A4	SLC29A4	-	YES
146	chr8	102236403	102236617	-0.04	0.12	0.16	ZNF706	ZNF706;YWHAZ	YES	-
147	chr8	131047175	131047344	-0.05	0.90	0.95	FAM49B	FAM49B;GSDMC	YES	YES
148	chr8	2585088	2586090	-0.06	0.14	0.20	LOC101927815	LOC101927815	-	YES
149	chr8	599524	600397	-0.09	0.83	0.91	ERICH1	ZNF596;ERICH1	-	-
150	chr8	58192499	58193337	-0.08	0.50	0.58	LINC00588	LINC00588	-	-
151	chr9	38687681	38687991	-0.07	0.36	0.43	FAM201A	FAM201A	YES	YES
152	chr9	140113368	140113558	-0.07	0.80	0.87	RNF208	DPP7;MAN1B1-AS1;MAN1B1	YES	YES
153	chr9	111605072	111605551	-0.05	0.10	0.15	ACTL7B	ACTL7B	YES	-
154	chr9	32430999	32431302	-0.03	0.91	0.95	ACO1	ACO1	YES	YES
155	chr9	5819260	5819333	-0.06	0.89	0.95	ERMP1	ERMP1	YES	YES
156	chr9	135114516	135114648	-0.08	0.32	0.39	NTNG2	C9orf171	YES	YES

No.	Chr.	Start	End	Mean Methylation Difference	Mean methylation (*b*-value) Asthma	Mean methylation (*b*-value) Controls	Nearest Gene	Predicted Target Gene	Enhancer	ngDMR
157	chr9	69500968	69501069	-0.08	0.72	0.80	ANKRD20A4	ANKRD20A4	-	YES
158	chr9	72131654	72131850	-0.06	0.78	0.84	APBA1	APBA1	-	YES
159	chr9	38487906	38488164	-0.08	0.48	0.56	FAM95C	FAM95C	-	YES
160	chr9	125879001	125879079	-0.05	0.88	0.93	MIR600HG	MIR600HG	-	YES
161	chr9	123744449	123744761	-0.06	0.83	0.89	TRAF1	PSMD5-AS1;PSMD5	-	YES

No. = Number of DMRs; Chr: chromosome, ng: non genotype dependent differentially methylated region

Table 8.1.5: SNP information for the 24 out of 161 DMRs, which were classified as genotype associated

Chr	Start	End	Gene	ngDMR	SNPID
chr16	30464225	30464522	SEPHS2	-	rs111186170;rs11646954;rs12443981;rs8044131;rs67575585
chr22	46762433	46763143	TRMU	-	rs35364389;rs12165943;rs75097576;rs9627428;rs9615965;rs12169526;rs73448958;rs11912939;rs9615352;rs11703059;rs9615351;rs76955646;rs73889254;rs9627423;rs75999278;rs77823653;rs11913988;rs9627426;rs9615966;rs9626857;rs111066975;rs9615964;rs12172608;rs79846114;rs7285194
chr8	102236403	102236617	ZNF706	-	rs1431007;rs2163400;rs1447017;rs877051
chr7	1914009	1914392	MIR4655	-	rs10950415;rs10269191;rs112425367;rs4719330;rs35349665;rs12699415;rs4721143;rs79610527;rs4639400;rs10266703;rs4719336;rs117125814;rs10950413;rs12669758;rs148725722;rs4721135;rs13224015;rs10265944;rs111357851;rs71523270
chr1	67600310	67600658	C1orf141	-	rs12044149;rs4655679;rs12068633;rs11209003;rs12060309;rs12065558;rs10789224;rs6588243;rs12569203;rs6588242;rs10489631;rs7543257;rs12069782;rs2024825;rs12095536
chr11	128694096	128694401	KCNJ1	-	rs2155549;rs636312;rs645601;rs571856;rs654169
chr2	241459177	241460046	ANKMY1	-	rs4676354;rs112154518;rs13394744;rs3821348;rs75898640
chr3	184243755	184244148	EIF2B5-AS1	-	rs8180000;rs13080490;rs13081033;rs35131513;rs13096674;rs62287408;rs11707574;rs11707620;rs11715352;rs62287379;rs148579807;rs28435810;rs12374080;rs62287380
chr9	111605072	111605551	ACTL7B	-	rs11789195;rs12552409;rs7034770;rs12552444;rs13296942;rs9299161;rs4978367;rs7039374;rs7024768;rs13297186;rs9299162;rs4542002;rs1006375;rs13302048;rs4358865;rs11789215;rs11789217;rs2000129;rs10979541;rs10979542;rs78878948;rs10979548;rs7857412;rs2000131;
chr18	14458381	14458936	CXADRP3	-	rs12606292;rs6505905;rs12604275;rs11080761;rs77810523;rs145868807;rs6505906;rs62080636;rs11080760;rs11080763;rs80054821;rs62080629;rs11080762;rs59482664;rs111762042;rs62080637;rs62080630
chr19	51373740	51373966	KLK2	-	rs2569741;rs3760728;rs3760730;rs11084039;rs2739470;rs2739464;rs2569742;rs2739477;rs2569738;rs2739476;rs11670728;rs2739473;rs12984666;rs12984214;rs62115062;rs2739459;rs2739466;rs1506684;rs1997563;rs2739460;rs2569739;rs2739462;rs2739475;rs2739461;rs2664156;rs2739469;rs965537;rs11084038;rs73592831
chr3	75445436	75445698	FAM86DP	-	rs7373315;rs12486482;rs13099914;rs35054081;rs7372634;rs7373724;rs11705831;
chr5	157117442	157117958	C5orf52	-	rs11134902;rs11750200;rs11743418;rs1133684;rs10066203;rs10065255;rs6887040;rs11134900;rs2133732;rs9313707;
chr5	77146478	77147360	LOC101929154	-	rs4262083;rs10942834;rs12658675;rs6880703;rs6877994;rs13162500;rs13180688;rs12697875;rs13187880;rs2361312;rs13159821;rs2361313;rs7700998;rs7717920;rs7708151;rs13165722;rs7732015;rs6881736;rs10079675;rs6886200;rs13172056;rs6889917;rs7728693;rs386488634;rs7704525;rs13183228;rs13162947;
chr6	32063911	32064191	TNXB	-	rs204897;rs204896;rs17201602;rs41270461
chr20	61629182	61629312	BHLHE23	-	rs76985826;rs113178859;rs75770746;rs6122341;rs12625084;rs113990986;rs62197287;rs4809447
chr20	32232295	32232457	C20orf144	-	rs12625805;rs2075734

Chr	Start	End	Gene	ngDMR	SNPID
chr5	77142381	77142898	LOC101929154	-	rs6877994;rs7700998;rs13180688;rs12697875;rs13159821;rs7717920;rs7708151;rs13165722;rs7732015;rs6881736;rs6886200;rs10079675;rs6887636;rs17185061;rs7728693;rs386488634;rs13171927;rs7704525;rs72633994;rs75420090;rs75762999;rs111778721;rs16874790
chr17	8702637	8702755	MFSD6L	-	rs904168;rs12950996;rs1826925;rs1826924;rs9915911;rs9916087;rs9916352;rs9894980;rs6503162;rs9900380;rs9900202;rs9909368;rs9916348;rs34110722;rs9907726;rs9894755;rs9895434;rs34911988;rs9900175;rs34856299;rs9899703;rs9916347;rs34184531;rs7209143;rs4791757;rs2242373
chr10	126593720	126593912	ZRANB1	-	rs4478950;rs12569669;rs61873263;rs61873264;rs61873265;rs57130385;rs12359621;rs12359031;rs7089014;rs7083298;rs7083336;rs7079199;rs55674450;rs61873268;rs7089131;rs12569940;rs17152377;rs17152381;rs7089281;rs7082846;rs7097282;rs7086447;rs1807099;rs2364522;rs733402;rs733401;rs733400;rs61873270;rs12571196;rs77088732;rs17152374;rs1863828;rs7093604;rs7081640;rs7078652
chr12	125482583	125482828	BRI3BP	-	rs4765198;rs7303487;rs4418906;rs4765014;rs10161214;rs10161415;rs7307530;rs6488973;
chr7	55412705	55412995	LANCL2	-	rs6970274;rs2331066;rs4948014
chr8	599524	600397	ERICH1	-	rs4735895;rs13277578;rs4593520;rs11991053;rs1669718;rs1703882;rs2335877;rs1703945;rs1703938;rs17751994;rs67918799;rs4735894;rs58792201;rs10087796;rs2878547;rs6559040;rs113381654;rs1669720;rs10107345;rs1669721;rs1669723;
chr8	58192499	58193337	LINC00588	-	rs67344620;rs66731170;rs68112919;rs66479724;rs72652905;rs72652906;rs58947041;rs73609747;rs60786705;rs58720297;rs59383954;rs59265461;rs66840104;rs9650139;rs112793785;rs3814486;rs68076606;rs113745779;rs72652908;rs111915672;rs67105789;rs73609760;rs75585481;rs73609764;rs76270388;rs67966192;rs67042991;rs66477954;rs59341413;rs58982816;rs61998258;rs61998259;rs61638902;rs61733801;rs73591706;rs59843034;rs68114352;rs55912204;rs55977795;rs58371676;rs56130194;rs66886949;rs113857287;rs73609762;rs78138632;rs112284078;rs76944716;rs76023408;rs59708460;rs2270610;rs2270609;rs2270607;rs10504229;rs6999878;rs2270608;rs1390411;rs57314710;rs67389108;rs58608483;rs10504230;rs59275611;rs58957714;rs79311869;rs6981914;rs60646469;

Chr= Chromosome; ng= non genotype associated differentially methylated region

Table 8.1.6: Results of the Motif analysis and the intersecting DMRs

Motif	Hits	Chr	Start	End	Gene Name	Predicted Target gene	Enhancer	ngDMR
1	1	chr17	36572579	36572896	ARHGAP23	SOCS7;SOCS7;HNF1B;MLLT6	YES	YES
1	2	chr17	8769570	8769883	PIK3R6	PIK3R6	YES	YES
1	3	chr17	21119605	21119864	TMEM11	USP22;C17orf103;USP22	YES	YES
1	4	chr16	69489543	69489664	CYB5B	CYB5B	YES	YES
1	5	chr19	34859991	34860409	GPI	PDCD2L;UBA2;UBA2	YES	YES
1	6	chr16	75026217	75026456	ZNRF1	WDR59;ZNRF1;WDR59;WDR59;ZNRF1	YES	YES
1	7	chr1	22097059	22097226	USP48	USP48	YES	YES
1	8	chr2	26094522	26094609	ASXL2	ASXL2;ASXL2	YES	YES
1	9	chr21	19184847	19184908	C21orf91	BTG3	YES	YES
1	10	chr21	38750599	38750876	DYRK1A	DSCR3;DSCR3;DSCR3	YES	YES
1	11	chr10	126593720	126593912	ZRANB1	ZRANB1	-	-
1	12	chr12	125482583	125482828	BRI3BP	BRI3BP;BRI3BP	-	-
1	13	chr14	93212312	93212486	LGMN	LGMN	-	YES
1	14	chr18	22016574	22016799	IMPACT	IMPACT;IMPACT	-	YES
1	15	chr2	97401278	97401371	LMAN2L	CNNM4;LMAN2L	-	YES
1	16	chr3	39395430	39395804	CCR8	SLC25A38;TTC21A;CX3CR1;CCR8;EEF1A1P24;SLC25A38;MOBP;MYRIP	-	YES
1	17	chr4	148634323	148634373	ARHGAP10	ARHGAP10	-	YES
1	18	chr6	166674955	166675082	LOC101929297	LOC101929297	-	YES
1	19	chr7	5382633	5382782	SLC29A4	SLC29A4	-	YES
2	1	chr8	102236403	102236617	ZNF706	ZNF706;ZNF706;YWHAZ	YES	-
2	2	chr19	3520495	3521153	FZR1	CACTIN;C19orf71;FZR1;MFSD12;FZR1;MFSD12	YES	YES
2	3	chr15	57511786	57512215	TCF12	TCF12	YES	YES
2	4	chr2	113956545	113956672	PSD4	AHSA2;PSD4;LOC654433;NT5DC4;PSD4;CBWD2;RABL2A;RPL23AP7	YES	YES
2	5	chr11	132951692	132952491	OPCML	OPCML	YES	YES
2	6	chr20	29525072	29525474	LINC01598	LINC01598	YES	YES
2	7	chr20	33416638	33416741	NCOA6	NCOA6;MMP24-AS1;EDEM2	YES	YES
2	8	chr5	68700315	68700723	MARVELD2	MARVELD2	YES	YES
2	9	chr17	8769570	8769883	PIK3R6	PIK3R6	YES	YES
2	10	chr17	21119605	21119864	TMEM11	USP22;C17orf103;USP22	YES	YES
2	11	chr7	102003600	102003766	LOC100289561	SH2B2	YES	YES
2	12	chr16	75026217	75026456	ZNRF1	WDR59;ZNRF1;WDR59;WDR59;ZNRF1	YES	YES
2	13	chr2	26094522	26094609	ASXL2	ASXL2;ASXL2	YES	YES
2	14	chr21	19184847	19184908	C21orf91	BTG3	YES	YES

Motif	Hits	Chr	Start	End	Gene Name	Predicted Targetgene	Enhancer	ngDMR
2	15	chr21	38750599	38750876	DYRK1A	DSCR3;DSCR3;DSCR3	YES	YES
2	16	chr17	8702637	8702755	MFSD6L	ODF4;NTN1	-	-
2	17	chr10	126593720	126593912	ZRANB1	ZRANB1	-	-
2	18	chr12	125482583	125482828	BRI3BP	BRI3BP;BRI3BP	-	-
2	19	chr14	93212312	93212486	LGMN	LGMN	-	YES
2	20	chr15	52872030	52872159	ARPP19	ARPP19;ARPP19	-	YES
2	21	chr2	97401278	97401371	LMAN2L	CNNM4;LMAN2L	-	YES
2	22	chr3	39395430	39395804	CCR8	SLC25A38;TTC21A;CX3CR1;CCR8;EEF1A1P24;SLC25A38;MOBP;MYRIP	-	YES
2	23	chr4	148634323	148634373	ARHGAP10	ARHGAP10	-	YES
2	24	chr4	144833125	144833345	LOC101927636	LOC101927636	-	YES
2	25	chr5	149145166	149145196	PPARGC1B	PPARGC1B	-	YES
2	26	chr7	5382633	5382782	SLC29A4	SLC29A4	-	YES
3	1	chr7	1914009	1914392	MIR4655	MAD1L1	YES	-
3	2	chr17	36572579	36572896	ARHGAP23	SOCS7;SOCS7;HNF1B;MLLT6	YES	YES
3	3	chr11	128694096	128694401	KCNJ1	KCNJ5;ARHGAP32	YES	-
3	4	chr2	74213621	74213840	TET3	STAMBP;DGUOK;MTHFD2;BOLA3-AS1;BOLA3;STAMBP;DGUOK;MTHFD2;STAMBP;DGUOK;MTHFD2;STAMBP;DGUOK;MTHFD2	YES	YES
3	5	chr1	93970706	93970976	FNBP1L	CCDC18;TMED5	YES	YES
3	6	chr17	28580392	28580613	SLC6A4	GOSR1	YES	YES
3	7	chr9	32430999	32431302	ACO1	ACO1	YES	YES
3	8	chr16	85654108	85654323	GSE1	GSE1	YES	YES
3	9	chr17	8769570	8769883	PIK3R6	PIK3R6	YES	YES
3	10	chr7	102003600	102003766	LOC100289561	SH2B2	YES	YES
3	11	chr16	69489543	69489664	CYB5B	CYB5B	YES	YES
3	12	chr16	75026217	75026456	ZNRF1	WDR59;ZNRF1;WDR59;WDR59;ZNRF1	YES	YES
3	13	chr1	22097059	22097226	USP48	USP48	YES	YES
3	14	chr1	149162004	149162462	NBPF25P	NBPF25P	-	YES
3	15	chr1	10436586	10436850	PGD	APITD1-CORT;APITD1;PGD;APITD1-CORT;APITD1;PGD	-	YES
3	16	chr11	59560470	59560548	MRPL16	MRPL16	-	YES
3	17	chr12	125482583	125482828	BRI3BP	BRI3BP;BRI3BP	-	-
3	18	chr12	107273279	107273680	TMEM263	TMEM263	-	YES
3	19	chr18	22016574	22016799	IMPACT	IMPACT;IMPACT	-	YES
3	20	chr18	59588681	59588813	RNF152	RNF152	-	YES
3	21	chr2	113426404	113426418	SLC20A1	SLC20A1	-	YES

Motif	Hits	Chr	Start	End	Gene Name	Predicted Targetgene	Enhancer	ngDMR
3	22	chr2	122984491	122984673	TSN	TSN	-	YES
3	23	chr3	39395430	39395804	CCR8	SLC25A38;TTC21A;CX3CR1;CCR8;EEF1A1P24;SLC25A38;MOBP;MYRIP	-	YES
3	24	chr3	98476467	98476656	ST3GAL6	ST3GAL6	-	YES
3	25	chr5	140961391	140961495	LOC100505658	LOC100505658	-	YES
3	26	chr5	149145166	149145196	PPARGC1B	PPARGC1B	-	YES
3	27	chr6	52385312	52385483	TRAM2	TRAM2	-	YES
3	28	chr8	58192499	58193337	LINC00588	LINC00588	-	-
4	1	chr17	36572579	36572896	ARHGAP23	SOCS7;SOCS7;HNF1B;MLLT6	YES	YES
4	2	chr11	128694096	128694401	KCNJ1	KCNJ5;ARHGAP32	YES	-
4	3	chr15	57511786	57512215	TCF12	TCF12	YES	YES
4	4	chr20	33416638	33416741	NCOA6	NCOA6;MMP24-AS1;EDEM2	YES	YES
4	5	chr14	100610169	100610683	DEGS2	YY1;SLC25A29;MIR345;DEGS2;YY1;EVL	YES	YES
4	6	chr17	21119605	21119864	TMEM11	USP22;C17orf103;USP22	YES	YES
4	7	chr16	69489543	69489664	CYB5B	CYB5B	YES	YES
4	8	chr19	34859991	34860409	GPI	PDCD2L;UBA2;UBA2	YES	YES
4	9	chr21	19184847	19184908	C21orf91	BTG3	YES	YES
4	10	chr21	38750599	38750876	DYRK1A	DSCR3;DSCR3;DSCR3	YES	YES
4	11	chr10	126593720	126593912	ZRANB1	ZRANB1	-	-
4	12	chr12	125482583	125482828	BRI3BP	BRI3BP;BRI3BP	-	-
4	13	chr18	22016574	22016799	IMPACT	IMPACT;IMPACT	-	YES
4	14	chr2	97401278	97401371	LMAN2L	CNNM4;LMAN2L	-	YES
4	15	chr4	148634323	148634373	ARHGAP10	ARHGAP10	-	YES
4	16	chr6	166674955	166675082	LOC101929297	LOC101929297	-	YES
4	17	chr7	5382633	5382782	SLC29A4	SLC29A4	-	YES
4	18	chr9	69500968	69501069	ANKRD20A4	ANKRD20A4	-	YES
4	19	chr9	123744449	123744761	TRAF1	PSMD5-AS1;PSMD5	-	YES
5	1	chr1	67600310	67600658	C1orf141	C1orf141	YES	-
5	2	chr15	57511786	57512215	TCF12	TCF12	YES	YES
5	3	chr5	68700315	68700723	MARVELD2	MARVELD2	YES	YES
5	4	chr3	172243109	172243330	TNFSF10	TNFSF10	YES	YES
5	5	chr17	21119605	21119864	TMEM11	USP22;C17orf103;USP22	YES	YES
5	6	chr7	102003600	102003766	LOC100289561	SH2B2	YES	YES
5	7	chr19	34859991	34860409	GPI	PDCD2L;UBA2;UBA2	YES	YES
5	8	chr19	4382715	4382767	SH3GL1	CHAF1A;SH3GL1	YES	YES
5	9	chr11	68517647	68517834	MTL5	GAL;C11orf24;GAL;TPCN2;GAL;CPT1A;MRPL21;IGHMBP2;GAL;CPT1A;MRPL21;IGHMBP2	YES	YES

Motif	Hits	Chr	Start	End	Gene Name	Predicted Targetgene	Enhancer	ngDMR
5	10	chr2	26094522	26094609	ASXL2	ASXL2;ASXL2	YES	YES
5	11	chr21	19184847	19184908	C21orf91	BTG3	YES	YES
5	12	chr21	38750599	38750876	DYRK1A	DSCR3;DSCR3;DSCR3	YES	YES
5	13	chr10	126593720	126593912	ZRANB1	ZRANB1	-	-
5	14	chr14	93212312	93212486	LGMN	LGMN	-	YES
5	15	chr2	97401278	97401371	LMAN2L	CNNM4;LMAN2L	-	YES
5	16	chr4	148634323	148634373	ARHGAP10	ARHGAP10	-	YES
5	17	chr5	140961391	140961495	LOC100505658	LOC100505658	-	YES
5	18	chr5	149145166	149145196	PPARGC1B	PPARGC1B	-	YES
6	1	chr4	170695782	170696267	C4orf27	CLCN3;MFAP3L	YES	YES
6	2	chr5	42943969	42944683	FLJ32255	LOC153684;ANXA2R	YES	YES
6	3	chr17	49057182	49057238	SPAG9	SPAG9	YES	YES
6	4	chr7	102003600	102003766	LOC100289561	SH2B2	YES	YES
6	5	chr16	75026217	75026456	ZNRF1	WDR59;ZNRF1;WDR59;WDR59;ZNRF1	YES	YES
6	6	chr1	22097059	22097226	USP48	USP48	YES	YES
6	7	chr3	128317561	128317754	LINC01565	GATA2	YES	YES
6	8	chr10	3625837	3626735	KLF6	KLF6;KLF6	-	YES
6	9	chr11	59560470	59560548	MRPL16	MRPL16	-	YES
6	10	chr12	125482583	125482828	BRI3BP	BRI3BP;BRI3BP	-	-
6	11	chr14	93212312	93212486	LGMN	LGMN	-	YES
6	12	chr17	81093746	81093809	METRNL	METRNL	-	YES
6	13	chr18	22016574	22016799	IMPACT	IMPACT;IMPACT	-	YES
6	14	chr5	140961391	140961495	LOC100505658	LOC100505658	-	YES

Chr= chromosome; ngDMR= non genotype associated differentially methylated region

Table 8.1.7: Transcription factors related to asthma according to Figure 3.2.4

Transcription factor	Related reference
EGR	Chan *et al.* (254)
GABPA	Xue *et al.(255)*
PATZ1	Kichaev *et al.(256)*
GATA3	Shrine *et al.(247)*
GATA6	Zhang *et al.(257)*
NFAC1	Koch *et al.(248)*
IRF	Landgraf-Rauf *et al.(249)*
STAT2	Hsieh *et al.(258)*
SRBP1	Smet *et al.(259)*
THB	Duan *et al.(260)*
ATF3	Roussel *et al.(261)*

Table 8.1.8: Genes related to asthma determined by natural language processing tool highlighted in Figure 3.2.7 by blue outline circle

Asthma-Gene	Related reference
IL4	Cui *et al.* (262)
PPARA	Trifilieff *et al.*(263)
RELA	Gagliardo *et al.* (240)
NOS3	Dahgam *et al.*(264)
KIF3A	Johansson *et al*.(265)
RAD50	Li *et al.* (266)
ZFPM1	Jahreis *et al*.(267)
E2F1	Xu *et al*. (268)
IL17D	Quan-San *et al*.(166)
GATA2	Yang *et al*.(269)
TSC2	Landgraf-Rauf *et al*. (249)
YY1	Guo *et al*. (241)
CPT1A	Byersdorfer *et al*. (270)
RPTOR	Yick *et al.* (271)
MAD1L1	McErlean *et al.* (167)
YWHAZ	Garbacki *et al*.(242)
CHMP4B	Kogan *et al.* (230)
KLF6	Duan *et al*.(272)
PPARGC1B	Temesi *et al.* (155)
TRAF1	Pedros *et al.* (273)
IL5RA	Liang *et al.* (85)

Table 8.1.9: Module analysis and gene enrichment for all enhancer target genes and DMR associated genes (FDR <0.001). Colour indicated corresponding module shown in Figure 3.2.7.

Module		pathways	genes observed/ in pathway	*p*-value	FDR	Genes
1	immune response/ inflammation	Small cell lung cancer(K)	(86/4)	1.35E-07	2.11E-05	TRAF3,TRAF1,BIRC7,RELA
		CD40/CD40L signaling(N)	(29/3)	6.60E-07	5.15E-05	TRAF3,TRAF1,RELA
		Viral carcibogenesis(K)	(203/4)	4.07E-06	2.12E-04	TRAF3,TRAF1,MAD1L1,RELA
		tnfr2 signaling pathway(B)	(9/2)	1.44E-05	5.60E-04	TRAF3,TRAF1
		NF-kappa B signaling pathway(K)	(95/3)	2.28E-05	7.06E-04	TRAF3,TRAF1,RELA
		TNF signaling pathway(K)	(110/3)	3.52E-05	9.16E-04	TRAF3,TRAF1,RELA
2	gene regula-tion	Activation of anterior HOX genes in hindbrain development during early embryogenesis(R)	(66/3)	2.22E-06	1.40E-04	YY1,NCOA6,POLR2H
		Generic Transcription Pathway(R)	(494/4)	2.06E-05	6.38E-04	YY1,NCOA6,E2F1,POLR2H
3	fat metabo-lism	Fatty acid, triacylglycerol, and ketone body metabolism(R)	(195/5)	1.86E-09	5.95E-08	SREBF1,CPT1A,SMARCD3,PPARA,PPARGC1B
		Circadian Clock(R)	(67/4)	7.19E-09	1.15E-07	SREBF1,CPT1A,SMARCD3,PPARA
		Insulin resistance(K)	(109/4)	5.02E-08	5.02E-07	SREBF1,CPT1A,PPARA,PPARGC1B
		Mitochondrial biogenesis(R)	(52/3)	1.09E-06	6.91E-06	SMARCD3,PPARA,PPARGC1B
		Regulation of cholesterol biosynthesis by SREBP (SREBF)(R)	(53/3)	1.15E-06	6.91E-06	SREBF1,SMARCD3,PPARA
		Transcriptional regulation of white adipocyte differentiation(R)	(84/3)	4.57E-06	2.28E-05	SREBF1,SMARCD3,PPARA
		RXR and RAR heterodimerization with other nuclear receptor(N)	(26/2)	5.70E-05	2.28E-04	SREBF1,PPARA
		mechanism of gene regulation by peroxisome proliferators via ppara(B)	(36/2)	1.09E-04	4.36E-04	CPT1A,PPARA
4	cellular stress response	LKB1 signaling events(N)	(43/3)	6.16E-07	5.11E-05	RPTOR,TSC2,YWHAZ
		mTOR signaling pathway(N)	(64/3)	2.02E-06	8.30E-05	RPTOR,TSC2,YWHAZ
		mTOR signaling pathway(K)	(154/3)	2.79E-05	5.95E-04	RPTOR,FZD1,TSC2
		TCF dependent signaling in response to WNT(R)	(159/3)	3.06E-05	5.95E-04	FZD1,YWHAZ,CSNK1G2
		p38 signaling mediated by MAPKAP kinases(N)	(21/2)	3.72E-05	5.95E-04	TSC2,YWHAZ
5	cilium assembly	Cilium Assembly(R)	(179/4)	7.36E-08	6.63E-07	KIF3A,TCTN1,MKS1,B9D1
6	signaling by HDAC class I	Signaling events mediated by HDAC Class I(N)	(56/2)	7.94E-05	8.73E-04	GATA2,ZFPM1
7	Jak-Stat signaling	Jak-STAT signaling pathway(K)	(158/3)	3.07E-06	1.35E-04	IL4,IL5RA,IL17D

Module		pathways	genes observed/ in pathway	P-value	FDR	Genes
8	DNA damage control	Nonhomologous End-Joining (NHEJ)(R)	(36/3)	3.64E-08	9.94E-07	RAD50,KAT5, WHSC1
		DNA Double Strand Break Response(R)	(45/3)	7.10E-08	9.94E-07	RAD50,KAT5, WHSC1
		HDR through Homologous Recombination (HR) or Single Strand Annealing (SSA)(R)	(85/3)	4.79E-07	4.31E-06	RAD50,KAT5, WHSC1
		Cell Cycle Checkpoints(R)	(164/3)	3.44E-06	2.41E-05	RAD50,KAT5, WHSC1
		DNA Damage/Telomere Stress Induced Senescence(R)	(27/2)	1.85E-05	9.24E-05	RAD50,KAT5
		ATM pathway(N)	(34/2)	2.93E-05	1.17E-04	RAD50,KAT5
9	RNA degrada-tion	RNA degradation(K)	(77/3)	3.56E-07	1.78E-06	BTG3,PABPC 3,C-T8
10	Endo-cytosis	Endocytosis(K)	(260/3)	1.37E-05	6.85E-05	CHMP1B,CH MP4B,STAM BP
11	Ribo-some	Mitochondrial translation(R)	(92/3)	6.07E-07	1.21E-06	MRPL16,MR PL12,MRPL2 1
		Ribosome(K)	(154/3)	2.85E-06	2.85E-06	MRPL16,MR PL12,MRPL2 1

Databases: (B) Biocarta; (R) Reactome, (N) Pathway Interaction Database, (K) KEGG

Table 8.1.10: Differentially methylated regions (DMRs) between asthmatic children (n=23) and healthy controls (n=25) based on whole genome bisulfite sequencing of cord blood DNA samples

Chr	Start	End	Gene	Predicted target gene	Enhancer	ngDMR
1	157248949	157249331	ETV3	ARHGEF11	YES	-
1	17198854	17199360	LOC112267871	CROCCP3;CROCC	YES	-
2	95537378	95537563	TEKT4	TEKT4	-	YES
2	88125240	88125526	RGPD2	RGPD2	-	YES
2	27038161	27038368	CENPA	GPR113	YES	-
2	23851888	23852088	KLHL29	KLHL29;ATAD2B;UBXN2A;MFSD2B;C2orf44;FKBP1B;TP53I3;PFN4	YES	YES
3	195576617	195576929	LINC01983	PPP1R2;ACAP2;TFRC;TCTEX1D2	-	-
4	191014262	191014425	DBET	DBET	-	YES
4	124092	124382	ZNF718	ZNF595;ZNF718;ZNF595	YES	-
4	124988	125374	ZNF718	ZNF595;ZNF718;	YES	-
4	125411	125759	ZNF718	ZNF595;ZNF718	YES	-
5	131607157	131607757	PDLIM4	RAD50;PDLIM4;SLC22A4;LOC553103;SLC22A5;FNIP1;SLC22A4;IRF1	YES	YES
7	127911773	127911846	LEP	LEP	YES	-
7	36007150	36007678	LOC101928618	HERPUD2;ANLN	YES	YES
7	137831625	137832414	MIR4468	SVOPL	YES	-
8	58055837	58056142	LINC01606	LINC01606	-	-
8	58192325	58193091	LINC00588	LINC00588	-	-
8	145240404	145240582	MROH1	PUF60;PLEC;PARP10;OPLAH;SHARPIN;KIAA1875;SCRT1;VPS28	YES	YES
8	145755797	145756044	C8orf82	OPLAH;SCRT1;TMEM249;ZNF517;RPL8;ZNF34;ZNF16;VPS28;ARHGAP39;ZNF251;BOP1;HSF1;DGAT1;ADCK5;CPSF1;CYHR1;KIFC2;PPP1R16A;GPT;MFSD3;RECQL4;LRRC14;C8orf82;ZNF250	YES	YES
8	2591089	2591237	LOC101927815	LOC101927815	YES	-
8	145164208	145164352	WDR97	GRINA;SPATC1;PARP10;FBXL6;SLC52A2;TONSL;MIR937;SCRIB;PUF60;PLEC;BOP1;HSF1;EEF1D;TIGD5;TSTA3;ZNF707;MAPK15;FAM83H;NRBP2;EPPK1;OPLAH;GPAA1;SHARPIN;KIAA1875;SCXA;DGAT1;SCRT1;ADCK5;CPSF1;VPS28;	YES	-
8	10583876	10584071	LOC102723313	C8orf74;SOX7	YES	YES
8	145164383	145164473	WDR97	GRINA;SPATC1;PARP10;FBXL6;SLC52A2;TONSL;MIR937;SCRIB;PUF60;PLEC;BOP1;HSF1;TSTA3;ZNF707;MAPK15;NRBP2;SHARPIN;KIAA1875;SCXA;DGAT1;CPSF1	YES	-

Chr	Start	End	Gene	Predicted target gene	Enhancer	ngDMR
8	145164483	145165157	WDR97	GRINA;SPATC1;DGAT1;GRINA;PARP10;FBXL6;SLC52A2;TONSL;MIR937;SCRIB;PUF60;GRINA;PLEC;PARP10;BOP1;HSF1	YES	-
8	58054992	58055599	LINC01606	LINC01606	YES	-
8	2585296	2586090	LOC101927815	LOC101927815	-	-
8	145163229	145163397	WDR97	FBXL6;SLC52A2;MIR937;SCRIB;PUF60;GRINA;PLEC;PARP10;TONSL;BOP1;HSF1	YES	-
9	27690402	27691023	C9orf72	C9orf72	-	-
9	38526666	38526756	FAM95C	FAM95C	-	-
9	99999691	100000160	SUGT1P4-STRA6LP-CCDC180	SUGT1P4-STRA6LP-CCDC180	YES	-
10	1405996	1406154	ADARB2-AS1	LARP4B;GTPBP4;IDI1;WDR37;ADARB2	-	YES
10	50244174	50244378	MIR4294	WDFY4	-	-
10	1404927	1405391	ADARB2-AS1	LARP4B;GTPBP4;IDI1;WDR37;ADARB2	-	YES
10	134778338	134779072	LINC01166	INPP5A;CALY	YES	YES
10	45694300	45694677	MIR3156-1	RASSF4;ANKRD54P1;ALOX5;MARCH8	YES	-
11	7110247	7110453	RBMXL2	RBMXL2	-	-
11	134393902	134394000	LOC283177	LOC283177	-	-
11	132951860	132952259	OPCML	OPCML	YES	YES
12	64936842	64936990	RASSF3	RASSF3	-	YES
12	64936996	64937193	RASSF3	RASSF3	-	-
12	104697654	104697775	EID3	EID3	-	YES
12	96709740	96710084	ELK3	CDK17	YES	YES
12	123010390	123010570	RSRC2	LOC100507066;ARL6IP4;DIABLO;CLIP1;ZCCHC8;DENR	YES	YES
12	9217441	9217685	LINC00612	TRAFD1;PHC1;M6PR	-	-
13	50700925	50701002	DLEU2	MIR3613;TRIM13;DLEU1;RNASEH2B-AS1;RNASEH2B;EBPL;CTAGE10P	YES	-
13	50701146	50701365	DLEU2	MIR3613;TRIM13;DLEU1;RNASEH2B-AS1;RNASEH2B;	YES	-
16	89258991	89259585	SLC22A31	CDH15;ZNF778;SLC22A31;ACSF3;LOC400558;SPATA2L;APRT;SPG7;GALNS;TRAPPC2L;RPL13	YES	-
17	3807783	3808443	P2RX1	OR3A3	YES	-
18	48255938	48256056	MRO	MRO	-	YES
18	37379452	37379720	LINC01902	LINC01902	-	-
18	37379842	37379925	LINC01902	LINC01902	-	YES
19	37266756	37267020	LOC728485	ZNF567	-	-

Chr	Start	End	Gene	Predicted target gene	Enhancer	ngDMR
20	259654	260174	C20orf96	C20orf96;ZCCHC3	YES	YES
21	43243499	43244020	PRDM15	C2CD2;UMODL1;C21orf128	YES	-
22	50985318	50985541	KLHDC7B	TYMP;ODF3B;SCO2;CHKB-AS1;CHKB;CHKB-CPT1B	YES	YES
22	45832576	45833136	RIBC2	KIAA0930	YES	-
22	39884874	39884999	MGAT3	SMCR7L;S-RD43;RPL3;ATF4;RPS19BP1;APOBEC3G;MGAT3;GRAP2	YES	YES

Chr= chromosome; ngDMR= non genotype associated differentially methylated region

Table 8.1.11: Overview of included DNA methylation studies related to environmental exposures and the 27 identified DMRs by the top down approach (factorial ANOVA)

Environmental factor	Epigenetic studies	Approach	Intersected regions
Prenatal maternal smoke exposure	Bauer *et al(106)****	WGBS	Chr3:75445436-75445698 *(FAM86DP)* Chr13:20968573-209690084 *(MIR4499)* Chr8:2585088-2586090 *(LOC101927815)* Chr8:58192499-58193337 *(LINC00588)* Chr11:132951692-132952491 *(OPCML)*
	Joubert *et al.(170)* Miyake *et al.(171)* deVocht *et al.(172)* Küpers *et al.(173)* Lee *et al.(174)* Markunas *et al.(175)* Reese *et al.(176)*	450K	
	Xu *et al.(177)*	450K	Chr5:77142381-77142898*(LOC101929154)* Chr7:90895326-90896701 *(FZD1)*
Maternal stress	Vangeel *et al.(178)*	450K	
	Trump *et al.(38)****	WGBS	
	Kertes *et al.(179)*	450K	
Maternal depression	Viuff *et al.(180)* Non *et al.(181)* Cardenas *et al.(182)*	450K	
Child´s ADHD	Walton *et al.(183)*	450K	
Air pollution during pregnancy	Gruzieva *et al.(184)* Peng *et al.(185)* Ladd-Acosta *et al.(186)*	450K	
Maternal asthma	deVries *et al.(187)*	450K	
Prenatal maternal arsenic exposure	Bozack *et al.(188)* Kaushal *et al.(189)* Rojas *et al.(190)*	450K	
Prenatal maternal mercury and arsenic exposure	Cardenas *et al.(191)*	450K	
Prenatal	Sen *et al.(192)*	450K	

maternal lead exposure	Wu *et al.(193)*		Chr13:24914323-24914904 *(LINC00566)* Chr13:20968573-20969084 *(MIR4499)* Chr13:24914323-24914904 *(LINC00566)*
Prenatal maternal cadmium exposure	Kippler *et al.(194)*	450K	
Maternal gestational diabetes mellitus	Howe *et al.(195)*	450K	
Preterm birth	Goede *et al.(196)*	450K	Chr11:68517647-68517834 *(MTL5) CT1* Chr2:24233608-24234116 *(MFSD2B) CT4* Chr16:88540019-88540525 *(MIR5189) CT4* Chr2:24233608-24234116 *(MFSD2B) CT8* Chr7:1914009-1914392 *(MIR4655) CT8* Chr16:88540019-88540525 *(MIR5189) CT8*
	Wu *et al.(197)*		Chr2:24233608-24234116 (MFSD2B) pt2 Chr2:24233608-24234116 (MFSD2B) pt4
Prenatal maternal phthalate exposure	Solomon *et al.(198)*	450K	
Prenatal maternal BPA exposure	Junge *et al.(169)****	450K	
Birthweight of neonates	Küpers *et al.(199)*	450K	chr1:6341136-6341791 (GPR153) chr16:88540019-88540525 (MIR5189) chr7:90895326-90896701 (FZD1) chr16:88540019-88540525 (MIR5189) chr7:90895326-90896701 (FZD1) chr1:6341136-6341791 (GPR153)
Maternal BMI during pregnancy	Sharp *et al.(200)*	450K	chr16:88540019-88540525(MIR5189) chr16:88540019-88540525 (MIR5189)
Prenatal DHA supplementation	van Dijk *et al.(201)*	450K	chr2:118617605-118618119 DDX18 chr9:38487906-38488164 FAM95C chr13:20968573-20969084 (MIR4499) chr17:19627951-19628165 (SLC47A2)
Vitamin D intake during pregnancy	Junge *et al.(274)*	WGBS	chr5:77146478-77147360 (LOC101929154)

***studies in the LINA cohort, WGBS= whole genome bisulfite sequencing BeadChip 450K= Illumina HumanMethylation450 BeadChip; ADHD=attention deficit hyperactivity disorder; BPA= Bisphenol A; DHA=docosahexaenoic acid

Table 8.1.12: Overview of included DNA methylation studies related to environmental exposures and the 57 identified DMRs by the bottom up approach (DMR calling)

Environmental factor	Epigenetic studies	Approach	Intersected regions
Prenatal maternal smoke exposure	Bauer *et al(106)****	WGBS	Chr19:37266756-37267020 (*LOC728485*) Chr8:58054992-58055599 (*LINC01606*) Chr8:58055837-58056142 (*LINC01606*) Chr18:37379452-37379720 (*LINC01902*) Chr8:2585088-2586090 (*LOC101927815*); Chr8:58192499-58193337 (*LINC00588*); Chr11:132951692-132952491 (*OPCML*
	Joubert *et al.(170)* Miyake *et al.(171)* deVocht *et al.(172)* Küpers *et al.(173)* Lee *et al.(174)* Markunas *et al.(175)* Reese *et al.(176)*	450K	
	Xu *et al.(177)*	450K	Chr12:9217441-9217685 (LINC00612) Chr13:50701146-50701365 (DLEU2)
Maternal stress	Vangeel *et al.(178)*	450K	
	Trump *et al.(38)****	WGBS	Chr7:137831625-137832414 (*MIR4468*) Chr8:145164208-145164352 (*WDR97*) Chr8:145164383-145164473 (*WDR97*) Chr8:145164483-145165157 (*WDR97*) Chr12:64936996-64937193 (*RASSF3*) Chr9:27690402-27691023 (*C9orf72*)
	Kertes *et al.(179)*	450K	
Maternal depression	Viuff *et al.(180)* Non *et al.(181)* Cardenas *et al.(182)*	450K	
Child´s ADHD	Walton *et al.(183)*	450K	
Air pollution during pregnancy	Gruzieva *et al.(184)* Peng *et al.(185)* Ladd-Acosta *et al.(186)*	450K	Chr18:48255938-48256056 *(MRO)*
Maternal asthma	deVries *et al.(187)*	450K	Chr4:124988-125374 (*ZNF718*)
Prenatal maternal arsenic exposure	Bozack *et al.(188)* Kaushal *et al.(189)* Rojas *et al.(190)*	450K	
Prenatal maternal mercury and arsenic exposure	Cardenas *et al.(191)*	450K	
Prenatal maternal lead exposure	Sen *et al.(192)* Wu *et al.(193)*	450K	Chr12:9217441-9217685 *(LINC00612)* Chr8:2591089-2591237*(LOC101927815)* Chr10:134778338-134779072 *(LINC01166)*
Prenatal maternal	Kippler *et al.(194)*	450K	

cadmium exposure			
Maternal gestational diabetes mellitus	Howe *et al.(195)*	450K	
Preterm birth	Goede *et al.(196)* Wu *et al.(197)*	450K	
Prenatal maternal phthalate exposure	Solomon *et al.(198)*	450K	
Prenatal maternal BPA exposure	Junge *et al.(169)****	450K	
Birthweight of neonates	Küpers *et al.(199)*	450K	Chr18:48255938-48256056 *(MRO)*
Maternal BMI during pregnancy	Sharp *et al.(200)*	450K	Chr16:88540019-88540525*(MIR5189)* S3 Chr16:88540019-88540525 *(MIR5189)* S5
Prenatal DHA supplementation	van Dijk *et al.(201)*	450K	Chr19:37266756-37267020 *(LOC728485)* Chr11:7110247-7110453 *(RBMXL2)* Chr18:48255938-48256056 *(MRO)* Chr10:1405996-1406154 *(ADARB2-AS1)* Chr2:95537378-95537563 *(TEKT4)* Chr8:58055837-58056142 *(LINC01606)* Chr10:1404927-1405391 *(ADARB2-AS1)* Chr3:195576617-195576929 *(LINC01983)* Chr8:58192325-58193091 *(LINC00588)* Chr12:104697654-104697775 *(EID3)* Chr8:2585296-2586090 *(LOC101927815)* Chr12:9217441-9217685 *(LINC00612)* Chr8:145755797-145756044 *(C8orf82)* Chr20:259654-260174 *(C20orf96)* Chr7:127911773-127911846 *(LEP)* Chr8:2591089-2591237 *(LOC101927815)* Chr16:89258991-89259585 *(SLC22A31)* Chr13:50700925-50701002 *(DLEU2)* Chr13:50701146-50701365 *(DLEU2)* Chr4:124092-124382 *(ZNF718)* Chr10:134778338-134779072 *(LINC01166)* Chr8:145164208-145164352 *(WDR97)* Chr5:131607157-131607757 *(PDLIM4)* Chr22:50985318-50985541 *(KLHDC7B)* Chr11:132951860-132952259 *(OPCML)* Chr8:145164383-145164473 *(WDR97)* Chr8:145164483-145165157 *(WDR97)* Chr8:58054992-58055599 *(LINC01606)* Chr4:124988-125374 *(ZNF718)* Chr4:125411-125759 *(ZNF718)* Chr9:99999691-100000160 *(SUGT1P4-*

			STRA6LP-CCDC180) Chr8:145163229-145163397 *(WDR97)* Chr2:118617605-118618119 *(DDX18)* Chr9:38487906-38488164 *(FAM95C)* Chr13:20968573-20969084 *(MIR4499)* Chr17:19627951-19628165 *(SLC47A2)*
Vitamin D intake during pregnancy	Junge *et al.(274)*	WGBS	Chr20:259654-260174 (*C20orf96*)

***studies in the LINA cohort, WGBS= whole genome bisulfite sequencing BeadChip 450K= Illumina HumanMethylation450 BeadChip; ADHD=attention deficit hyperactivity disorder; BPA= Bisphenol A; DHA=docosahexaenoic acid

Curriculum Vitae

Name: Matthias Klös

Birth: 08.11.1987, Koblenz

Nationality: German

03/2019-10/2019	Parental leave
Since 02/2018	PhD student at Molecular Epidemiology Unit (Prof. Dr. Irina Lehmann) at the Berlin Institute of Health and Charité Universitätsmedizin Berlin, Berlin, Germany
11/2014-01/2018	PhD student at the department of Environmental Immunology (Prof. Dr. Irina Lehmann) at the Helmholtz-Center for Environmental Research - UFZ, Leipzig, Germany Thesis: **"Role of epigenetic mechanisms in childhood asthma development"**
10/2011-10/2014	Student of Nutritional Sciences, Martin-Luther-University Halle-Wittenberg, Faculty of Natural Sciences III, Institute for Agricultural and Nutritional Sciences, Halle Graduation: Master of Science (1.6) Master thesis, JRG Functional Genetics (Dr. Y. Böttcher), Integrated Research and Treatment Center (IFB) AdiposityDiseases, Leipzig Thesis: **"Differential Methylation of the candidate gene *SSPN* in human subcutaneous and visceral adipose tissue"**
02/2014-05/2014	Research assistant, University Leipzig, Integrated Research and Treatment Center (IFB) AdiposityDiseases, Leipzig
02/2013-04/2013	Research internship (MS Pro), Integrated Research and

	Treatment Center (IFB) AdiposityDiseases, Leipzig
10/2008-9/2011	Student of Nutrition and Food Sciences, Friedrich-Wilhelm University, Bonn Graduation: Bachelor of Science (2.5)

_____ _____

Date /Datum Matthias Klös

List of publications

Klös, M., Thurmann, L., Bauer, M., Heinrich, J., Standl, M., von Berg, A., Schaaf, B., Weichenhan, D., Mucke, O., Plass, C., Roder, S., Herberth, G., Sack, U., Borte, M., Stangl, G. I., Eils, R., Trump, S. and Lehmann, I. (2019) : **"Longitudinal trends of serum IgE and IL5RA expression throughout childhood are associated with asthma but not with persistent wheeze"** Allergy, 74**,** 2002-2006.

Keller, M., **Klös, M**., Rohde, K., Krüger, J., Kurze, T., Dietrich, A., Schön, M. R.,Gärtner,D.,Lohmann, T., Dreßler, M., Stumvoll, M., Blüher, M., Kovacs P. & Böttcher, Y. (2018). **DNA methylation of SSPN is linked to adipose tissue distribution and glucose metabolism**. The FASEB Journal, 32, 6898-6910.

Thürmann, L., Grützmann, K., **Klös, M**., Bieg, M., Winter, M., Polte, T., Bauer, T., Schick, M., Bewerunge-Hudler, M., Roeder, S., Bauer, M., Wissenbach, D.K., Sack, U., Weichenhan, D., Mücke, O., Plass, C., Borte, M., von Bergen, M., Lehmann, I., Eils, R., Trump, S., (2018): **"Early-onset childhood atopic dermatitis is related to NLRP2 repression"** J. Allergy Clin. Immunol., 141, 1482-1485 e16.

Rohde, K., **Klös, M**., Hopp, L., Liu, X., Keller, M., Stumvoll, M., Dietrich, A., Schon, M.R., Gartner, D., Lohmann, T., Dressler, M., Kovacs, P., Binder, H., Bluher, M., and Bottcher, Y. (2017). **"IRS1 DNA promoter methylation and expression in human adipose tissue are related to fat distribution and metabolic traits."** Sci Rep 7(1): 12369.

Keller, M., Hopp, L., Liu, X., Wohland, T., Rohde, K., Cancello, R., **Klös, M**., Bacos, K., Kern, M., Eichelmann, F., Dietrich, A., Schon, M.R., Gartner, D., Lohmann, T., Dressler, M., Stumvoll, M., Kovacs, P., DiBlasio, A.M., Ling, C., Binder, H., Bluher, M., and Bottcher, Y. (2017). **"Genome-wide DNA promoter methylation and transcriptome analysis in human adipose tissue unravels novel candidate genes for obesity"**. Molecular metabolism 6(1): 86-100.

Rohde K, Keller M, **Klös M**, Schleinitz D, Dietrich A, Schön MR, Gärtner D, Lohmann T, Dreßler M, Stumvoll M, Kovacs P, Blüher M, Böttcher Y (2014). **"Adipose tissue depot specific promoter methylation of TMEM18."** J Mol Med (Berl) **92**(8): 881-888.

Published abstracts

Trump, S., **Klös, M.**, Gu, L., Bauer, T., Ishaque, N., Thurmann, L., Bieg, M., Bauer, M., Gu, Z., Borte, M., Mallm, J.-P., Roder, S., Herrmann, C., Rippe, K., Lehmann, I., Eils, R.,

(2017):**“Asthma development in early childhood is associated with enhancer reprogramming”**

Am. J. Respir. Crit. Care Med. **195** , A4910

List of oral presentation

Trump, S., **Klös, M.**, Schmidt, M., Bauer, T., Ishaque, N., Thürmann, L., Bieg, M., Herrmann, C., Röder, S., Bauer, M.,, Weichenhan, D., Mücke, O., Plass, C., Borte, M., von Mutius, E., Kabesch, M., Stangl, GI, Lauener, R., Pekkanen, J., Dalphin, J-C., Riedler, J., Eils, R., Lehmann, I. (2018): **"Development of early childhood asthma goes along with massive enhancer activation in blood cells"** *European Academy of Allergy and Clinical Immunology (EAACI) conference, München, Germany*

List of poster presentation

Trump, S., **Klös, M.,** Schmidt, M., Gu, L., Bauer, T., Ishaque, N., Thürmann, L., Bieg, M., Röder, S., Bauer, M., Borte, M., Kabesch, M., Eils,R., Lehmann, I. (2017): **"Asthma development in early childhood is associated with epigenetic changes in enhancer regions"** *Systems Biology of Human Disease (SBHD), Heidelberg, Germany*

Klös, M., Trump, S., Borte, M., Kabesch, M., Eils, R., Stangl, G., Lehmann, I. (2017) **"Breathless through childhood - What has the epigenome got to do with it?"** *HIGRADE Conference 2017 | UFZ-PhD Conference, Leipzig Germany*

Klös, M., Trump, S., Thürmann, L., Bauer, M., Borte, M., Weichenhan, D., Mücke, O., Herberth, G., Eils, R., Stangl, G., Lehmann,I. (2016) **"Role of IgE and IL5RA regulation in the development of childhood asthma"** *Annual Retreat of the International Graduate Programme "Molecular Biology and Medicine of the Lung, Rauischholzhausen, Germany*

Klös, M., Trump, S., Bauer, T., Ishaque, N., Thürmann, L., Bieg, M., Röder, S., Bauer, M., Borte, M., Grützmann, K., Herrmann, C., Eils, R., Lehmann, I. (2016) **"Epigenetic modifications in cord blood associated with asthma development later in children's life"** *14th ERS European Respiratory Society Lung Science Conference, Estoril, Portugal*

Declaration under Oath/ Eidesstattliche Erklärung

I declare under penalty of perjury that this thesis is my own work entirely and has been written without any help from other people. I used only the sources mentioned and included all the citations correctly both in word or content.

Ich erkläre an Eides statt, dass ich die Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Date /Datum	Signature of the applicant/ Unterschrift des Antragstellers

Acknowledgement

Attention will be drawn here to say thank you to all people supporting me and participated to finalize this thesis.

First, I would like to thank and express my appreciation to Prof Dr. Gabriele Stangl and Prof. Dr. Irina Lehmann for being my thesis referees and for providing me the opportunity to conduct my PhD at Martin-Luther-University Halle-Wittenberg.

Foremost, I especially would like to thank Prof Dr. Irina Lehmann and Dr. Saskia Trump who guided me during my PhD in an excellent research environment. I really appreciated the fruitful discussions, critical questions and helpful comments during various meetings and discussion rounds.

I would thank all the current and former members of the Department of Environmental Immunology at the UFZ in Leipzig and the Molecular Epidemiology Unit at the Berlin Institute of Health and the Charité Universitätsmedizin Berlin who supported me with help and assistance in the lab and daily life.

Dr. Loreen Thürmann was instrumental for transmitting me their knowledge about MassARRAY technique. For this, I am extremely grateful. Additionally, I would like to thank all the members of Eils lab and Berlin Hub, especially Dr. Carl Herrmann and Matthias Bieg for the excellent bioinformatics analyses and transmitting their knowledge to me in the here presented thesis.

Additionally, I would like to express my gratitude to all groups that further conducted measurements for the analyses within the thesis, namely the groups Prof Dr. Christoph Plass for MassARRAY, Dr. Stefan Wolf for sequencing analyses, Dr. Gunda Herberth, Immunoparameter analyses.

Finally, I would like to thank my family and friends for their great support and patience regarding this process to finalize this thesis.