Elucidation of the role of obesity and air pollution in the asthma etiology in adults

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-

Naturwissenschaftlichen Fakultät

der Universität Basel

von

Ayoung Jeong

aus Südkorea

Basel, 2019

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Dr. Nicole Probst-Hensch und Prof. Dr. John Holloway

Basel, 26. Juni 2018

Prof. Dr. Martin Spiess

Dekan der Philosophisch-Naturwissenschaftlichen Fakultät

Acknowledgements

My biggest thanks go to my supervisor Nicole Probst-Hensch for putting her trust in me to undertake this PhD project. I still remember vividly how excited it all sounded but overwhelming at the same time when you first described your ideas on the exposomics research. I learnt so much from you, not only from your insights and experience in research but also from your perspectives, broader view embracing research and public health, and team spirit.

I would also like to thank Christian and Gianfranco. Your lectures I attended in my early years and all our discussions afterwards shaped my entire knowledge in biostatistics. It has always been a great pleasure to discuss the research ideas and plans with you, and also reassuring to know that I can always turn to you to discuss statistical issues.

I deeply thank Medea for her support throughout this PhD project. It has been a huge comfort to have someone like you, with seemingly endless tolerance, capacity, and experience.

I would also like to thank my co-referee John Holloway. Your support and collaborative spirit are much appreciated. I am looking forward to our further collaboration.

To everyone in my PhD committee, I would like you to know that I really appreciate all the discussions and exchanges we had. Every single meeting left me lighter minded and excited with new ideas. I am so lucky to work with you, who excel in and love what they are doing.

I would like to take this opportunity to thank the SAPALDIA team and all the participants. Without their commitment this PhD project would not have been possible.

I would also like to thank all the co-authors and partners in SAPALDIA and EXPOsOMICS, which made this PhD project possible.

A big thank goes to all my colleagues at Swiss TPH, especially to Nora Bauer, Christine

Mensch, and Dagmar Batra, whose support are critical for many of us including me. I would also like to thank all my fellow students, especially the wonderful ladies — Joelle, Sofie, Bettina, and Mari. I miss the days we used to work in the same old creaking room.

Last but not least I would like to thank my family and friends. My sister So Young, you have always been my motivation to be a better person and the endless source of laughter. And my oldest friends — the members of "eagle five" — I am so lucky to have you as my friends. I have been, and will always be, proud to be your friend. And of course my partner Taeuk, you always see the best in me. I cannot imagine anybody better than you to have in my life.

Contents

List of Figures

List of Tables

Abbreviations

27k array	Illumina Infinium HumanMethylation27 BeadChip
450k array	Illumina Infinium HumanMethylation450 BeadChip
ACS	acute coronary syndromes
AHR	airway hyperresponsiveness
AKT	protein kinase B
AMI	amyotrophic myocardial infarction
BMI	body mass index
BMIQ	beta-mixture quantile normalization
CCVD	cardio- and cerebrovascular diseases
CVD	cardiovascular diseases
CpG	cytosine-guanine dinucleotide
DALY	disability-adjusted life-years
DMR	differentially methylated region
ERK	extracellular-regulated kinase
ESCAPE	European Study of Cohorts for Air Pollution Effects
EWAS	epigenome-wide association study
EWIS	epigenome-wide interaction study
GAN	Global Asthma Network
GINA	Global Initiative for Asthma
GLM	generalized linear model
GNC	German National Cohort
GWAS	genome-wide association study
HIC	high-income country
IARC	International Agency for Research on Cancer
ICD	International Classification of Diseases
IL17	interleukin-17
IL1B	interleukin-1β
ILC3	innate lymphoid cell group 3

LMIC	low- and middle-income country
MITM	meet-in-the-middle
MS	mass spectrometry
NCD	non-communicable disease
NLRP3	nucleotide-binding domain, leucine-rich repeats-containing family, pyrin domain-containing-3
NMR	nuclear magnetic resonance
NO_2	nitrogen dioxide
PBF	percent body fat
PI3K	phosphatidylinositol-3-kinases
PM_{10}	particulate matter with <10 μm in diameter
$PM_{2.5}$	particulate matter with <2.5 μm in diameter
PPAR	peroxisome proliferator-activated receptor
QTOF	quadrupole time-of-flight
RXRα	retinoid X receptor alpha
SAPALDIA	Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults
SNP	single nucleotide polymorphism
T_h2	type 2 helper T cell
TLR	toll-like receptor
UFP	ultrafine particles
UHPLC	ultrahigh performance liquid chromatography
WC	waist circumference
WHO	World Health Organization
WHR	waist-hip ratio
WHtR	waist-height ratio

Summary

Non-communicable disease (NCD) epidemic threatens public health in all regions of the world. Asthma is one of the major NCDs along with cardiovascular diseases, cancer, diabetes, and other chronic respiratory diseases. Asthma etiology is poorly understood, hindering the efficient primary prevention. Recent findings indicate that asthma is a mixture of various phenotypes with potentially different mechanism. While obesity and air pollution have been indicated as risk factors for asthma, it is not clear yet whether they contribute to the development of asthma rather than exacerbation of already existing disease and through which mechanisms they exert the effects on asthma development. Elucidation of such mechanism, especially if it is shared by multiple exposures and/or multiple diseases, will critically benefit primary prevention. The research efforts for mechanistic understanding can be contextualized as part of exposome — the entirety of the exposures an individual experiences throughout the life course — and aging phenome — the diseases and morbidities often accompanied with aging — research, where systems approach e.g. omics analysis finds a critical usage.

The Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA) is an on-going population cohort since 1991. With its detailed information on the participants' health, life style, and exposure, SAPALDIA offers a unique opportunity to investigate the NCD etiology. This PhD project aimed to contribute to better understanding the role of obesity and air pollution exposure in asthma etiology, taking the heterogeneity of the disease phenotype into account.

We identified four asthma phenotypes using latent class analysis, which showed heterogeneity in the association with obesity. We conducted epigenomics — assessments of genome-wide DNA methylation — and metabolomics — assessments of the entirety of small molecules — on the blood samples taken from the adult-onset asthma cases and con-

trols. Epigenomics pathway analysis revealed that DNA methylation on the inflammation-related genes modifies the effect of BMI on non-atopic adult-onset asthma. This pathway analysis also provided evidence that the NLRP3-IL1B-IL17 axis, a component of innate immunity, plays a role in the asthma etiology in humans, confirming the previous research findings in mice experiments. Metabolomics pathway analysis pointed to the perturbation of inflammatory pathways as a potentially shared mechanism through which long-term air pollution exposure affects adult-onset asthma and cardio- and cerebrovascular diseases.

Despite the cross-sectional study design and the limited statistical power, this PhD project achieved to demonstrate the importance of distinguishing asthma phenotypes to study etiology; to exemplify the usefulness of cohort studies with biobanks in exposome research and the applicability of systems approach in cohort studies; and to provide a proof-of-concept evidence of the disease mechanism shared by multiple NCDs. Our findings can be considered as the first step of the translational approach — innovation, validation, and application. Once validated by future research including replication in other populations and consolidation of causality using Mendelian randomization, the pursuit of mechanistic understanding can guide prevention strategies to efficiently tackle the NCD epidemic.

Zusammenfassung

Die Epidemie der nicht übertragbaren Krankheiten (non-communicable diseases; NCD) bedroht die Gesundheit der Bevölkerung in allen Regionen der Welt. Asthma zählt neben Herzkreislauferkrankungen, Krebs, Diabetes, und anderen chronischen Atemwegserkrankungen, zu den Hauptkrankheiten der NCDs. Die Ätiologie von Asthma ist noch weitgehend unbekannt, was die effiziente Primärprävention erschwert. Neuere Befunde deuten darauf hin, dass Asthma aus verschiedenen Phänotypen besteht, mit unterschiedlichen Ursachen. Obwohl Fettleibigkeit und Luftverschmutzung als Risikofaktoren von Asthma gelten, ist es noch nicht klar ob sie zur Entwicklung von Asthma beitragen oder bestehendes Asthma nur verschlimmern und welche Mechanismen einen Effekt auf die Asthmaentwicklung haben. Die Aufklärung solcher Mechanismen, insbesondere wenn mehrere Risikofaktoren und Krankheiten diese teilen, wird der Primärprävention zugutekommen. Die Forschungsarbeiten zur mechanistischen Aufklärung kann man als einen Teil von Exposom- und Alterungsphänomforschung verstehen, welche versucht die Gesamtheit aller Expositionen die ein Individuum über den Lebensverlauf erfährt bzw. die Gesamtheit altersbedingter Krankheiten zu untersuchen. Solche Forschung benötigt einen Systemansatz einschliesslich Omics-Analysen.

Die Schweizer Kohortenstudie über Luftverschmutzung und Atemwegs- und Herzerkrankungen bei Erwachsenen (SAPALDIA) ist eine seit 1991 laufende Bevölkerungskohortenstudie. Mit ihren detaillierten Informationen zur Gesuntheit, Lebensstil, und anderen Expositionen, bietet SAPALDIA eine einzigartige Gelegenheit die Ätiologie der NCDs zu untersuchen. Ziel dieser Dissertation ist es, einen Beitrag zu leisten zum Verständnis der Rolle von Fettleibigkeit und Luftverschmutzung in der Asthmaentwicklung, unter Berücksichtigung der Heterogenität der Asthmaphänotypen.

Mit Hilfe von Latent-Class-Analysen identifizierten wir vier Asthmaphänotypen, die

eine Heterogenität im Zusammenhang mit Fettleibigkeit zeigen. Epigenom- und Metabolomanalysen wurden auf die Blutproben der spät einsetzenden Asthmafälle und Kontrollgruppen durchgeführt, womit man das umfassende DNA-Methylierungsprofil bzw. die Gesamtheit aller kleinen Moleküle beurteilt. Die "Epigenomics" Pathway-Analyse zeigte dass die DNA-Methylierung der Gene, welche mit Entzündungen zusammenhängen, den Effekt von BMI auf nicht atopischem, spät einsetzendem Asthma modifiziert. Die Pathway-Analyse erbrachte ebenfalls den Nachweis dafür, dass die NLRP3-IL1B-IL17-Achse, ein Teil der angeborenen Immunität, eine Rolle bei Asthmaentwicklung am Menschen spielt, wie die früheren Forschungsbefunde bei Mäusen festgestellt haben. Die "Metabolomics" Pathway-Analyse zeigte, dass Stoffwechselwegstörungen, durch langfristige Luftverschmutzung, als einen möglicherweise gemeinsamen Mechanismus zur Entwicklung von Asthma und Herzkreislauferkrankungen beitragen.

Gleichwohl des Querschnittsdesign und der limitierten statistischen Power veranschaulicht dieses PhD Projekt die Wichtichkeit bei Ätiologie-Forschung, Asthmaphänotypen zu unterscheiden; die Wichtigkeit der Kohortenstudien mit assoziierten Biobanken wie SAPALDIA für die Omics-Analyse zur Exposomforschung zu nützen; und erbrachten einen "Proof-of-Concept", dass ein eventueller gemeinsamer Mechanismus mehrere NCDs beeinflusst. Die Befunde dieser Dissertation können als erster Schritt eines translationalen Ansatzes — Innovation, Validierung, und Anwendung — angesehen werden. Nach Validierung durch zukünftige Forschung, einschliesslich Replikation in weiteren Bevölkerungen und Kausalitätsfeststellung z.B. mit Hilfe von Mendelian Randomization, können die Forschungsarbeiten zur mechanistischen Aufklärung, Präventionsstrategien leiten um die NCD-Epidemie effizient zu bekämpfen.

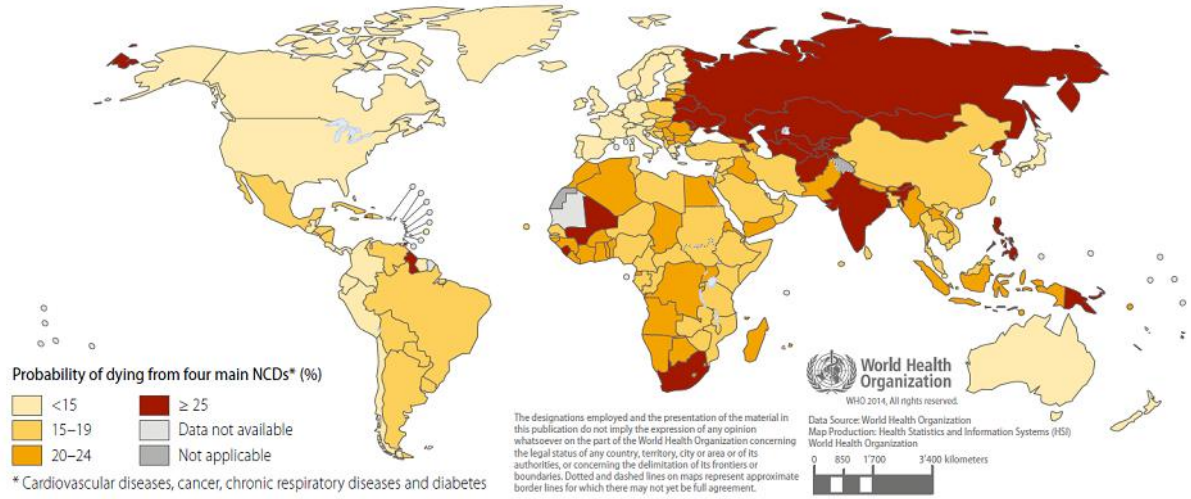
1. INTRODUCTION

1.1 Non-communicable disease (NCD) epidemic

In the early days of epidemiology, infectious diseases – smallpox, cholera, influenza, among others – were the killers. Industrialization leading to improved sanitation and nutrition accompanied by development of antibiotics and vaccines moved the battlefield from the infectious diseases to the NCDs. According to the recent World Health Organization (WHO) report, NCDs caused twice as many deaths as caused by all other causes combined in 2012 (WHO, 2014). The Global Burden of Diseases, Injuries, and Risk Factors Study 2015 reported that NCDs caused the highest burden measured in disability-adjusted life-years (DALYs) (Forouzanfar et al., 2016). This NCD epidemic affects not only the high-income countries (HICs) but all regions in the world. The low- and middle-income countries (LMICs) have been experiencing rapid increase in NCD mortality and morbidity due to globalization and urbanization (**Figure 1.1**). In 2012, 74% of the 38 million deaths due to NCDs and 82% of the 16 million premature deaths due to NCDs occurred in LMICs (WHO, 2014). Economic losses due to NCDs including premature deaths and productivity loss alongside with the costs to the health systems prevent the development and progress of the LMICs (The Global Asthma Network (GAN) 2014).

The NCD epidemic is dominated by cardiovascular diseases, cancer, diabetes, and chronic respiratory diseases including asthma, all of which can be considered as aging-

Figure 1.1. Probability of dying from the four main non-communicable diseases between the ages of 30 and 70 years, comparable estimates, 2012 (WHO, 2014, Global status report on noncommunicable diseases 2014 ©World Health Organization (2014), all rights reserved, used with permission)



related diseases. Elderly often suffer from multiple morbidities, while centenarians never develop any such NCDs, indicating the potential clustering of the NCDs. It is conceivable that there are common mechanisms shared in the aging-related diseases, causing or preventing multiple, clustered NCDs. Meta-analyses of genome-wide association studies (GWAS) identified pleiotropic loci associated with multiple NCDs, supporting the hypothesis of shared mechanism in various NCDs (He et al., 2016, Jeck et al., 2012). Elucidation of the shared mechanism leading to the multiple, clustered NCDs associated with aging – aging phenome – will facilitate successful prevention and control of the NCDs.

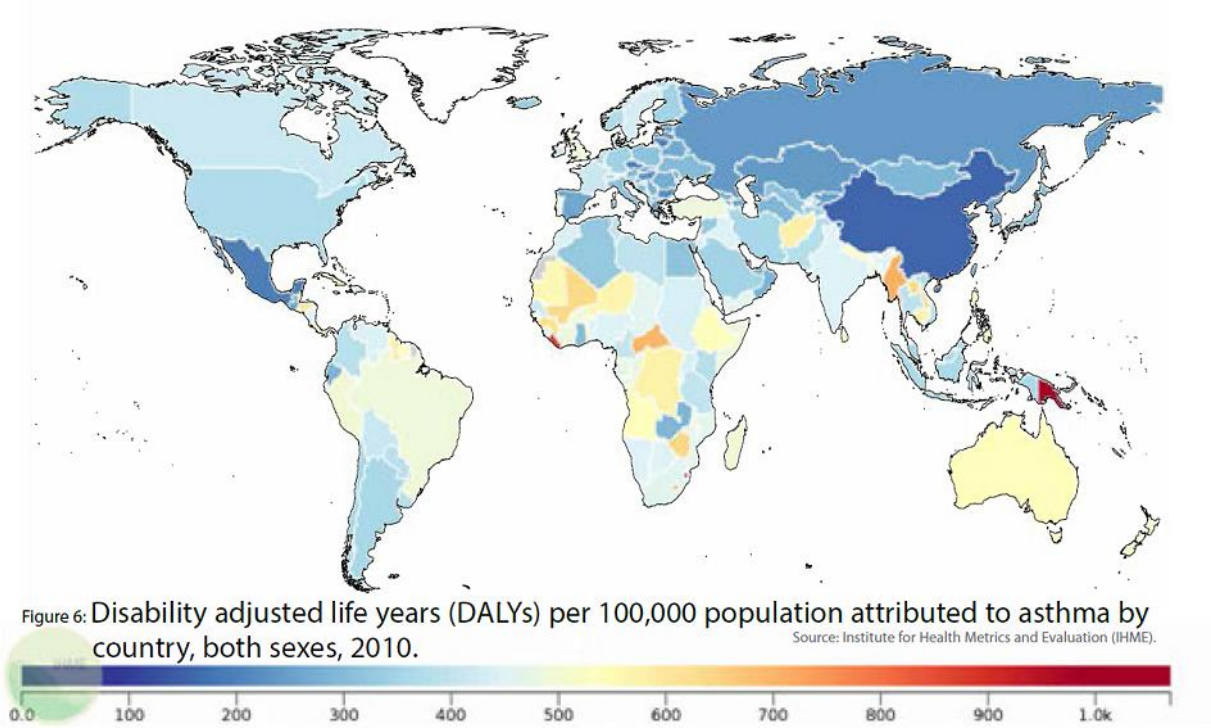
1.2 Asthma as part of the NCD epidemic

Asthma is a chronic airway disease often characterized by chronic airway inflammation, airway obstruction, and airway hyperresponsiveness, but these characteristics are neither necessary nor sufficient to define asthma (Global Initiative for Asthma (GINA) 2018).

Asthma is controllable with appropriate medication and if well controlled, rarely affects daily life of the patients. Poorly controlled asthma, however, decreases quality of life and imposes economic burden due to direct costs as well as indirect costs due to productivity loss (GAN, 2014).

Approximately 340 million people live with asthma worldwide, affecting all age groups and all regions (Forouzanfar et al., 2016) as depicted in the map of DALYs (**Figure 1.2**). Asthma caused 24 million DALYs worldwide in 2016, one third occurred in LMICs. On the other hand, more than half of the 0.42 million deaths due to asthma occurred in LMICs in 2016, indicating the poor control of the disease in LMICs.

Figure 1.2. DALYs per 100,000 population attributed to asthma by country in 2010 (GAN, 2014, The Global Asthma Report 2014 ©The Global Asthma Network (2014), all rights reserved, used with permission)



As priority interventions against NCD epidemic, the WHO report focused on tobacco control, salt intake reduction, healthier diet, physical activity, reduction in harmful alcohol intake, and availability and affordability of essential medicines and technology (WHO,

2014). Asthma is included in the WHO agenda of NCDs as part of chronic respiratory diseases. However, it is important to distinguish the measures to prevent asthma attacks and improve the control of the disease from the measures to prevent the incidence of the disease. The priority interventions set by the WHO hardly address the latter, considering that the etiology of the disease is largely unknown and the proposed target risk factors have small population attributable fractions for asthma (Pearce et al., 2013). Better understanding of the disease mechanism will facilitate successful prevention and control of asthma.

1.3 Causes and risk factors of asthma

Although asthma often runs in families and heritability of childhood asthma is indeed estimated up to 82% (Ullemar et al., 2016), the genetic variants identified by GWAS account for little of the disease prevalence (Demenais et al., 2018, Moffatt et al., 2010). There are non-genetic factors considered as risk factors for asthma including allergy, respiratory infections, smoking, air pollution, physical activity, female sex, and obesity. Their causal effects are largely inconclusive yet. Asthma had been considered as an allergic disease but allergy does not accompany all asthma cases. As will be elaborated later, asthma is now believed to be a mixture of heterogeneous phenotypes including allergic and non-allergic asthma, rather than a single disease (GINA, 2018, Wenzel, 2012). While respiratory infections, exercise, and exposure to tobacco smoke and air pollution can trigger asthma attacks, it is unclear if they cause the disease progress. Asthma prevalence is higher in boys than in girls but the ratio reverses around puberty (Carey et al., 2007). Reasons for the sex difference in asthma are yet to be elucidated. Obesity has been well associated with asthma in children and in adults (Beuther and Sutherland,

2007, Egan et al., 2013). The hypothesized mechanisms to explain the obesity-asthma association will be described later in this chapter.

1.4 Heterogeneity of asthma phenotype

As reviewed by Wenzel (Wenzel, 2012), various asthma phenotypes can be distinguished in terms of disease history (early-onset or adult-onset), clinical and physiological features (allergic, non-allergic, exercise-induced, or obesity-related), biomarkers (eosinophilic, neutrophilic, presence or absence of the type 2 helper T cell (T_h2) signature cytokines), and response to therapy. GINA also acknowledged the heterogeneity of asthma phenotypes and distinguished allergic, non-allergic, late-onset, obesity-related asthma, and long-standing asthma with fixed airflow limitation (GINA, 2018).

Early-onset asthma is the most common asthma phenotype and often allergic and responsive to corticosteroid therapy. A large GWAS identified genetic determinants specific for early-onset asthma (Moffatt et al., 2010). Late-onset asthma is less studied compared to early-onset asthma and often refractory to corticosteroid therapy. Non-allergic asthma is often adult-onset and more likely to refractory to corticosteroid therapy compared to allergic asthma. Obesity-related asthma tends to be female predominant, adult-onset, non-allergic, lacking T_h2 signature cytokines, non-eosinophilic, highly symptomatic, and refractory to corticosteroid therapy. It has not been established yet how to categorize asthma phenotypes and how to relate them to treatment options or clinical benefits. Clustering methods have been applied with some success (Haldar et al., 2008, Moore et al., 2010, Siroux et al., 2014, Boudier et al., 2013) but consensus is yet to be achieved. In this PhD project we aimed to fill this gap by investigating asthma heterogeneity and etiology with special attention to late-onset asthma.

1.5 Obesity as a risk factor for asthma

Obesity has repeatedly been associated with incident asthma in children (Egan et al., 2013) and in adults (Beuther and Sutherland, 2007). Obesity-related asthma has also been considered as a distinct asthma phenotype as described above. Recent Mendelian randomization studies provided causal reasoning of the obesity effect on asthma (Granell et al., 2014, Skaaby et al., 2017). However, the biological mechanism through which obesity exerts the effects on asthma is yet poorly understood. It should be borne in mind that the obesity-asthma relationship may differ across asthma phenotypes. Obesity was reported to have stronger effects on adult-onset asthma than early-onset asthma (Haldar et al., 2008, Moore et al., 2010) and on non-atopic asthma than atopic asthma (Castro-Giner et al., 2009, Fenger et al., 2012).

The potential mechanisms explaining the obesity-asthma association include: Obesity can reduce lung volume and promote airway narrowing; Obesity can increase the work of breathing leading to misdiagnosis of asthma; Obesity-related hormones, i.e. adipokines, might play a role in the asthma development; Comorbidities of obesity – dyslipidemia, gastroesophageal reflux disease, sleep dyspnea, type 2 diabetes, etc. – may exacerbate asthma; Or the same genetic and/or environmental factors cause both conditions (Shore, 2008). The most likely hypothesis is that obesity-induced low-grade systemic inflammation causes asthma development. Adipose tissue in obese individuals is known to produce abnormal amount of pro-inflammatory cytokines (Weisberg et al., 2003). In obesity, macrophages are known to infiltrate into adipose tissue and differentiate predominantly into M1, the pro-inflammatory type of macrophages, leading to low-grade systemic inflammation (Castoldi et al., 2016, Engin, 2017).

A recent mice study provided a convincing evidence of innate immunity as a link between obesity and asthma (Kim et al., 2014). They demonstrated that obesity-induced

airway hyperresponsiveness (AHR) was dependent on the NLRP3 (nucleotide-binding domain, leucine-rich repeats-containing family, pyrin domain-containing 3) inflammasome and its downstream activity via interleukin-1β(IL1B) and interleukin-17 (IL17) produced by innate lymphoid cell group 3 (ILC3) cells by showing that obese mice did not develop AHR when the NLRP3-IL1B-IL17 axis was blocked either by knockout or administration of antagonists. It was the keen interest of this PhD project if this NLRP3-IL1B-IL17 axis also explains the obesity-asthma association in humans.

1.6 Air pollution exposure as a risk factor for asthma

The acute effect of air pollution exposure on exacerbation of pre-existing asthma has been established over decades (Schwartz et al., 1993, Weinmayr et al., 2010). The long-term effect of air pollution exposure – whether air pollution exposure contributes to asthma development – is less clear. The largest study to date with individual exposure estimates for over 600,000 adults reported cross-sectional association of annual mean exposure to nitrogen dioxide (NO_2) and particulate matter (PM) with <10 μm in diameter (PM_{10}) with asthma prevalence (Cai et al., 2017). The long-term effect of air pollution exposure on asthma incidence is less consistent. The European Study of Cohorts for Air Pollution Effects (ESCAPE) study reported positive albeit not statistically significant association of annual mean exposure to various air pollutants including NO_2, PM_{10}, and $PM_{2.5}$ with asthma incidence (Jacquemin et al., 2015). Compared to PM_{10} and $PM_{2.5}$, ultrafine particles (UFP; <0.1 μm in diameter) has been less studied for the effects on asthma. Various air pollutants have different source and composition, and therefore can have different toxicity (Kumar et al., 2015, Schwarze et al., 2007). Smaller particles are believed to have more hazardous effects on respiratory system because they can reach easily into

the alveoli and the higher ratio of surface to mass can harbor larger amount of toxic substances (Li et al., 2016). On the other hand, larger particles with higher iron content can be more hazardous (Kumar et al., 2015).

The mechanism by which air pollution exposure exerts the effect either on exacerbation or new onset of asthma is not clear. The potential mechanisms include: high concentration air pollutants irritates lung epithelium and causes acute inflammation; exposure to specific air pollutants enhance airway sensitization; chronic, low level exposure to air pollution induces oxidative stress and inflammation thereof (Guarnieri and Balmes, 2014)

1.7 Omics analysis: a new epidemiology tool in the era of NCD epidemic

Success of the human genome project followed by fast-paced development of high-throughput technologies began a new era of omics analysis. The term "-omics" refers to a comprehensive study of a totality of biological molecules – "genome" as for the entirety of genetic variants, "methylome" as for the genome-wide DNA methylation, etc. It is now possible to measure vast number of biological molecules simultaneously at a relatively low cost. Cohort studies, if accompanied with biobanks, can therefore afford systems approach, which is crucial to better understand disease etiology.

1.7.1 Epigenomics

Epigenetics refers to the processes by which gene expression is regulated without changing DNA sequence. Various epigenetic mechanisms are known, including DNA methylation, histone modification, chromatin remodeling, and non-coding RNAs. DNA methylation is the most widely studied epigenetic mechanism. This PhD project, as many epidemiologi-

cal studies usually do, investigated DNA methylation, because technology for cost effective measurements is not available for other epigenetic mechanisms. DNA methylation occurs in the cytosine-guanine (CpG) dinucleotides via covalent bonding of a methyl group to the 5-carbon of the pyrimidine ring of the cytosine residue to form a 5-methylcytosine. The unmethylated cytosine can be distinguished from the 5-methylcytosine by bisulfite conversion, where only the unmethylated cytosine converts to uracil, and quantified by e.g. Illumina Infinium array chips. The methylation level can regulate gene expression. Hypermethylation in CpG islands – 200 bp or larger sequence with high contents of CpG dinucleotides – embedded in the promoter region is known to silence the gene. Function of methylation in other loci than CpG islands within promoter regions – CpG islands in gene body, intergenic CpG islands, and CpG sites outside of CpG islands – is less understood but recently being recognized (Jones, 2012).

Unlike genome, in principle, methylome is dynamic, i.e. DNA methylation level changes over time in response to environment in a tissue-specific manner. Therefore DNA methylation can serve as an excellent biomarker for exposures. For example, epigenome-wide association studies (EWAS) identified numerous CpG sites whose methylation levels were strongly associated with tobacco smoking (Joehanes et al., 2016). For some of the CpG sites, their smoking-induced methylation changes did not restore to the normal level even several decades after smoking cessation (Guida et al., 2015). Such persistent biomarkers offer a new opportunity to better characterize the exposure as well as to contribute to better understanding the mechanism related to the exposure. As the epigenetic markers are modifiable unlike genetic variants, causal understanding of the mechanism can lead to novel therapeutic options (Heerboth et al., 2014, Tough et al., 2016). Moreover, considering the importance of epigenetic reprogramming in embryonic development and the possibility of transgenerational epigenetic inheritance via primordial germ cells of the em-

bryo, it is also conceivable that epigenetic markers – probably involving other epigenetic processes than DNA methylation – represent a crucial channel through which multiple hazards affect various organs.

1.7.2 Metabolomics

Small molecules in the bodily fluid – generally referred as "metabolites" – can be considered as the product of genetics, endogenous processes governed by genetics and epigenetics, and environmental exposures including diet and behavior. Comprehensive investigation of such metabolites allows an access to unique information on disease etiology. The measurements of metabolites are conducted either by using nuclear magnetic resonance (NMR) spectroscopy or mass spectrometry (MS) in combination of liquid chromatography (LC). MS-based methods are more sensitive and detect broader spectrum of chemicals in terms of chemical class and concentration range compared to NMR-based methods. NMR-based methods have strengths in structure elucidation and reproducibility (Bictash et al., 2010, Tzoulaki et al., 2014). Both methods can be applied in untargeted or targeted metabolomics. The untargeted metabolomics aims to profile the full spectrum of chemicals in the samples as allowed by the method used. The targeted metabolomics restricts the profiling to a specific class, e.g. lipids, leukotrienes, volatile organic compounds, etc. Identification of the metabolites poses a unique challenge in untargeted metabolomics. Metabolite databases with information on their chemical and clinical information such as The Human Metabolome Database (HMDB; http://www.hmdb.ca) (Wishart et al., 2007) and METLIN (https://metlin.scripps.edu) (Smith et al., 2005) or bioinformatics tools to automate the chemical annotation such as xMSannotator (Uppal et al., 2017) and Mummichog (Li et al., 2013) can be consulted to tackle the challenge (Rattray et al., 2018).

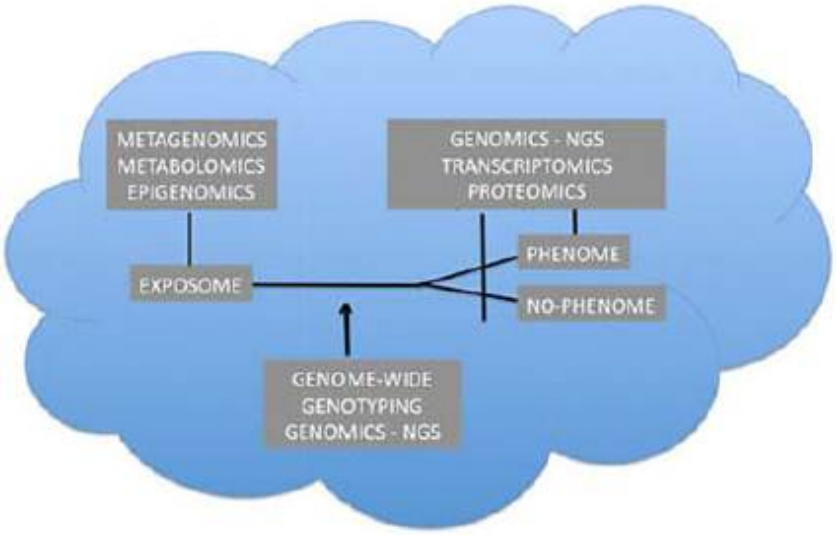
1.7.3 Exposomics

The term "exposome" was first coined by Wild as a match for "genome" to point out the critical needs of more accurate and comprehensive exposure assessment to catch up with that of genetics (Wild, 2005). The exposome is defined as every exposure to which an individual is subjected throughout the life course. It does not only include the usual environmental hazards such as air pollution and water contamination but also more general environmental exposures such as socioeconomic status and urban structure, as well as internal exposures such as endogenous processes and microbiota (Wild, 2012). Such broad spectrum exposures cannot be captured by the classic assessment of single exposures. Therefore omics analyses constitute an inevitable component of exposome research in epidemiology (Lopez de Maturana et al., 2016). Integration of multi-omics information benefits the mechanistic understanding of the exposome (**Figure 1.3**). The EXPOsOMICS was one of the first large-scale implementation of exposome research funded by European Union (Vineis et al., 2016, Turner et al., 2018). In the EXPOsOMICS project, various omics analyses in combination with classic exposure assessment were conducted to provide mechanistic understanding of the exposure-disease association, paying special attention to water and air pollution (Vineis et al., 2016). This PhD project was in part conducted in the context of the EXPOsOMICS.

1.7.4 Meet-in-the-middle (MITM)

Vineis and Perera proposed the MITM approach as to strengthen the causal interpretation of the intermediate biomarkers (Vineis and Perera, 2007). The MITM can be implemented by prospective search for the intermediate biomarkers associated with both exposure and disease (**Figure 1.4**). Such intermediate biomarkers are more likely on the causal pathway linking the exposure to the disease. Recent studies applied the MITM

Figure 1.3. Omics integrative epidemiology (Lopez de Maturana et al, 2016, Toward the integration of Omics data in epidemiological studies: still a "long and winding road". Genet Epidemiol, 40, 558-569. ©John Wiley and Sons (2016), all rights reserved, used with permission)

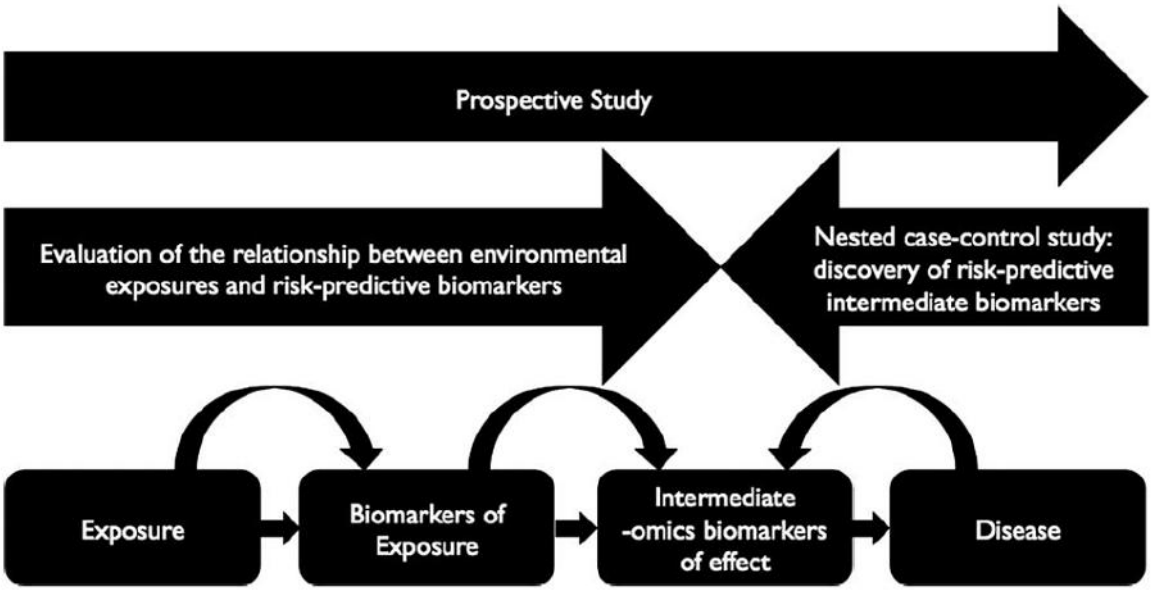


approach successfully to consolidate the causal interpretation of the effect of smoking on lung cancer (Fasanelli et al., 2015) using the DNA methylation markers previously reported to be strongly associated with smoking (Guida et al., 2015) as the intermediate biomarkers; of the effect of lifestyle on liver cancer using metabolites identified from untargeted metabolomics as the intermediate biomarkers (Assi et al., 2015); and of the effect of air pollution exposure on cardio- and cerebrovascular diseases using the DNA methylation markers on the inflammation-related genes as the intermediate biomarkers (Fiorito et al., 2017). As part of the EXPOsOMICS, this PhD project applied the MITM approach to interpret the association between air pollution exposure and asthma using metabolomics biomarkers.

1.7.5 Statistics in omics analysis

Omics analysis inevitably involves high dimension data, causing the "$n \ll p$" issue, i.e. the number of parameters p greatly exceeds the number of observation n. The most widely ap-

Figure 1.4. The MITM approach (Vineis et al, 2013, Advancing the application of omics-based biomarkers in environmental epidemiology. Environ Mol Mutagen, 54, 461-7. ©John Wiley and Sons (2013), all rights reserved, used with permission)



plied approach is the univariate analysis followed by multiple testing corrections (Balding, 2006). This approach is referred as ome-wide association analysis, e.g. GWAS, EWAS, etc. The simplest method for multiple testing corrections is to adjust the significance level α by α/n, so called Bonferroni correction. However, the tests should not be considered independent because of their correlation structure: genome-wide data measured as SNPs can have correlation mainly driven by physical distance (linkage disequilibrium); methylome data can have clustered correlations (Lovkvist et al., 2016); untargeted metabolomics data can have a very complex correlation reflecting biological relationship between the molecules. In the situation of such correlation, the Bonferroni correction can be too stringent leading to increased false negatives. Alternatives include the permutation-based estimation of the effective number of tests (Chadeau-Hyam et al., 2013) and the false discovery rate approach (Benjamini and Hochberg, 1995).

Regardless of the choice for the multiple testing corrections, the univariate methods bound to overlook the relationship between molecules, which may provide better insight on the biological mechanism (Agier et al., 2016). Although a variety of multivariate approaches have been developed and applied to omics analyses, including principal component analysis, (sparse) partial least squares regression, Lasso, Elastic net, among others, there is no established pipeline for all types of molecules and platforms. Specifically in epigenomics, approaches to search for differentially methylated regions (DMRs) have been developed and applied in acknowledgement of the clustering nature of the data (Clifford et al., 2018, Perry et al., 2018).

1.7.6 Bioinformatics in omics analysis

Besides the issues caused by the high dimension, statistical analyses of omics data face another level of challenge posed by the biological nature of the data. Proper design of the analysis or interpretation of the results require each parameter – SNPs in genomics, CpG sites in epigenomics, metabolites in metabolomics – to be contextualized in terms of genes, regulatory elements, and biological pathways. Bioinformatics resources have been growing at an unprecedented pace (Chen, 2015). Prominent resources include: the Reference Sequence (RefSeq) (Pruitt et al., 2005) and the Single Nucleotide Polymorphism Database (dbSNP) (Sayers et al., 2011) are the DNA sequence databases housed by the National Center for Biotechnology Information (NCBI); Ensembl offers a genome browser jointly built by European Bioinformatics Institute (EBI) and Wellcome Sanger Institute (Hubbard et al., 2002); the Encyclopedia of DNA Elements (ENCODE) provides functional information on human genome (Hong et al., 2016); and Kyoto Encyclopedia of Genes and Genomes (KEGG) is a collection of databases of pathways, genes, and chemicals (Kanehisa and Goto, 2000). The journal Nucleic Acids Research publishes a

special issue in January every year devoted to update the available molecular biology databases. The Nucleic Acids Research Database Issues provide a good overview.

1.7.7 Personalized health research from a public health and primary prevention perspective

Since the success of the human genome project, omics analysis has been burgeoning but often understood as a servant for personalized medicine. However, personalized medicine and omics analysis in that sense do not necessarily contradict public health benefit. Risk factors of NCDs often have small effect size and work in concert, and therefore systems approach is imperative. Omics analysis contextualized in a well-designed cohort study can contribute to better understanding of disease etiology, which in turn contributes to better primary prevention. Multiple correlated exposures challenge the classic assessment of single exposures. Systems approach accompanied by advancement of statistical methodologies to address correlation structure and dynamics of the exposures can provide a novel way of risk assessment (National Academies of Sciences and Medicine, 2017). Mechanistic understanding between exposome and aging phenome can potentially provide a novel window of intervention; improve burden of disease estimates; and best inform the primary prevention strategies.

2. OBJECTIVES

In this PhD project we aimed to contribute to better understanding the role of obesity and air pollution exposure in asthma etiology, taking the heterogeneity of the disease phenotypes into account. Specifically, the objectives of this PhD project were as follows:

1. Identify asthma phenotypes and assess heterogeneity in association with obesity across asthma phenotypes

2. Assess association of obesity with asthma and its heterogeneity across various obesity measures

3. Improve understanding the biological mechanism mediating the effects of obesity on asthma

4. Assess association of air pollution exposure with asthma

5. Improve understanding the biological mechanism mediating the effects of air pollution exposure on asthma

The results of the objectives 1 and 2 are presented in Article I; the results of the objective 3 are in Article II; and the results of the objectives 4 and 5 are presented in Article III.

3. METHODS

3.1 Study samples

3.1.1 The Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA)

SAPALDIA is a population-based adult cohort in Switzerland. SAPALDIA was initiated in 1991 (SAPALDIA1), recruiting 9651 adults aged 18-62 years (Martin et al., 1997). Participants were randomly selected from eight areas representing the diverse geography, meteorology, and degrees of urbanization of the country: Basel, Wald, Davos, Lugano, Montana, Payerne, Aarau, and Geneva (**Figure 3.1**). 8047 and 6088 out of the 9651 initial participants were followed up in 2001-3 (SAPALDIA2) (Ackermann-Liebrich et al., 2005) and in 2010-11 (SAPALDIA3), respectively. Information on respiratory health and various risk factors including life style was collected by questionnaires and in-person interviews as well as on-site physical examinations including spirometry. At SAPALDIA2 and SAPALDIA3, blood samples were also collected and stored in a biobank. All participants provided informed consent and ethical approval was obtained from the Swiss Academy of Medical Sciences and the regional committees for each study area. In this PhD project we analyzed mainly SAPALDIA3 data.

Figure 3.1. SAPALDIA study areas

3.1.2 Nested case-control study

In the context of EXPOsOMICS, a nested case-control study was conducted from SAPALDIA3. Cases were selected among the ever-asthma cases who had not smoked at least 10 years before SAPALDIA3 based on availability of blood samples in the biobank and non-missing information on covariates including atopy and age of asthma onset. Controls were randomly selected among the participants who had not smoked at least 10 years before SAPALDIA3 and never reported the following from SAPALDIA1 to SAPALDIA3: ever-asthma; doctor-diagnosed asthma; current asthma; wheezing without cold in the last 12 months; three or more asthma-related symptoms in the last 12 months (symptoms considered: breathless while wheezing; woken up with a feeling of chest tightness; attack of shortness of breath after exercise; attack of shortness of breath while at rest; woken up by attack of shortness of breath). Article II and III relied on the case-control samples after exclusion of cases with age of asthma onset earlier than 16 years, in order to specifically investigate adult-onset asthma as a distinct phenotype.

3.2 Asthma phenotypes

Ever asthma is defined if the question "Have you ever had asthma?" was answered "yes" at least once from SAPALDIA1 to SAPALDIA3. Doctor-diagnosed asthma is defined if both questions "Have you ever had asthma?" and "Was this confirmed by a doctor?" were answered "yes" at least once from SAPALDIA1 to SAPALDIA3. Adult-onset asthma was defined if the self-reported age of asthma onset was 16 years or older among ever-asthma cases. Current asthma was defined if either the question "Have you had an attack of asthma in the last 12 months?" or the question "Are you currently taking any medicines including inhalers, aerosols, or tablets for asthma?" was answered "yes" at SAPALDIA3. Atopy was defined if the skin prick test at baseline showed an adjusted mean wheal diameter $\geq$3 mm to at least one of eight common allergens including cat fur, dog epithelia, house dust mite (*Dermatophagoides pteronyssinus*), timothy grass pollen, birch pollen, *Parietaria* pollen, and the moulds *Alternaria* and *Cladosporium*. In this PhD project we investigated prevalent asthma rather than incident asthma, acknowledging the challenge in defining incident asthma cases. Asthma can go unnoticed for long time, grow out, and resurface, all of which increase the misclassification risk for incident asthma.

3.3 Obesity metrics

Body mass index (BMI), derived from anthropometric measurements as weight in kilograms divided by the square of height in meters, was used as the main obesity metric. While BMI is widely used as obesity metric, its limitation is also well acknowledged that it is ignorant of fat distribution and incapable of distinguishing fat from lean mass. In Article I, alternative metrics were included in the analysis and compared with BMI, including: percent body fat, derived as (weight – fat-free mass)/weight in percentage with the fat-

free mass was estimated from bioelectric impedance measurement; waist circumference; waist-hip ratio; waist-height ratio. In Article II, BMI change defined as the difference between SAPALDIA3 and SAPALDIA2 was also used as better proxy for accumulating excess fat in late adulthood.

3.4 Air pollution exposure estimates

The SAPALDIA participants were assigned annual mean exposure to $PM_{2.5}$ and NO_2, and biennial mean exposure to UFP by geocoding their home address. $PM_{2.5}$ estimates in 2010 were derived from PolluMap, a nation-wide, high resolution dispersion model in Switzerland (FOEN, 2013). Lagged estimates up to 7 years before SAPALDIA3 were derived by Meteotest on behalf of the federal office for the environment (FOEN, 2014). UFP estimates in 2011/2012 were derived from the multi-area land use regression (LUR) models developed in SAPALDIA covering 4 out of the 8 SAPALDIA study areas (Basel, Wald, Lugano, and Geneva) (Eeftens et al., 2016). Lung deposited surface area (LDSA), in addition to particle number concentration (PNC), was estimated in acknowledgement of the previous experimental studies reporting stronger effects of the surface area than the mass of the UFP (Schwarze et al., 2006, Stoeger et al., 2006). NO_2 estimates were derived from a European LUR model developed in the context of EXPOsOMICS (de Hoogh et al., 2016).

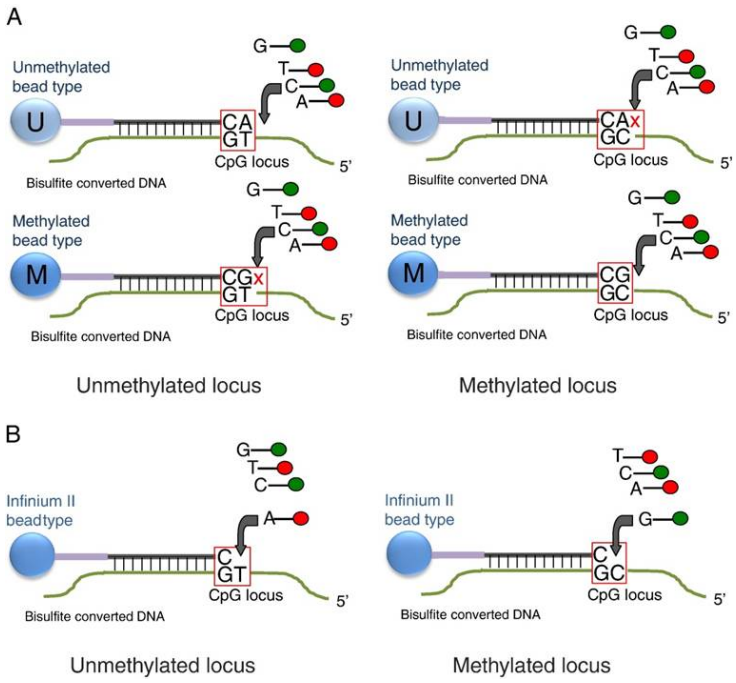
3.5 Epigenomics

Buffy coat fraction collected and stored at SAPALDIA3 was analyzed for the nested case-control samples. DNA was extracted from buffy coat fraction and treated with bisulfite so that unmethylated cytosine is converted to uracil while methylated cytosine (5-

methylcytosine) remains unaffected. Bisulfite-converted DNA was analyzed with Illumina Infinium HumanMethylation450 BeadChip (450k array) in a carefully designed order to minimize batch effect. All the laboratory analysis was conducted in the International Agency for Research on Cancer (IARC). A detailed description of the laboratory process is presented in Article II.

Raw data acquired from the chip measurements were fluorescence intensities, from which β values were derived as the ratio of methylated intensity over total intensity. Preprocessing and quality control were conducted mainly using R package *minfi* (Aryee et al., 2014). In addition to the *minfi* pipeline, beta-mixture quantile normalization (BMIQ) (Teschendorff et al., 2013) was applied to correct for the Illumina probe design bias. The 450k array is a mixture of two types of assays: type I assay consists of two bead types, one for methylated and the other for unmethylated loci, from which the fluorescence was measured by a single color channel; type II assay uses one bead type only, where methylated and unmethylated loci fluoresce at different wavelengths and therefore need to be measured by two different color channels (**Figure 3.2**). Approximately 28% of the 450k array probes are measured by type I assay and 72% by type II. This probe design is known to possibly cause bias (Dedeurwaerder et al., 2011). BMIQ is one of the widely applied techniques to mitigate this bias by rescaling the type II probe values to follow the distribution obtained from the type I probes. In order to correct for batch effect, a principal component analysis (PCA) was conducted on the 220 control probes incorporated into the 450k array and designed to assist quality control for bisulfite conversion, staining, hybridization, etc. The components derived from the PCA represent the summary measure of technical variation. The residuals out of the regression of β values on the first 30 components were considered to represent the batch effect-free methylation level.

Figure 3.2. Illumina Infinium HumanMethylation450 BeadChip array design (Bibikova et al, 2011, High density DNA methylation array with single CpG site resolution. Genomics, 98, 288-95. ©Elsevier (2011), all rights reserved, used with permission)



3.6 Metabolomics

Serum samples collected and stored at SAPALDIA3 were analyzed for the nested case-control samples with an ultrahigh performance liquid chromatography/quadrupole time-of-flight/mass spectrometry (UHPCL-QTOF-MS) system in a randomized order as a single batch. The peak heights were preprocessed to identify features defined by mass-to-charge ratio and retention time. After preprocessing, 12,003 features were identified and 7089 features were ready for statistical analyses after filtering out the features with >40% missing values. The chemical identity of the features was unknown. Given the high work load involved in the annotation of the features, only the features found interesting out of the statistical analyses were further examined for the peak validity and annotation. Of note is that some features identified by the preprocessing could still be artifacts or false findings from the feature finding algorithm. The additional inspection of the selected

features sorted out such false features. A true identity of the remaining features could be confirmed by comparing them with the standard substance, if the feature was measured in a detectable concentration and the standard substance was available. All the laboratory analysis including quality control and annotation was conducted in the IARC. A detailed description of the laboratory analysis is presented in Article III.

4. Article I

Heterogeneity of obesity-asthma association disentangled by latent class analysis, the SAPALDIA cohort

Ayoung Jeong[a,b], Medea Imboden[a,b], Sofie Hansen[a,b], Elisabeth Zemp[a,b], Pierre-Olivier Bridevaux[c,d], Gianfranco Lovison[e], Christian Schindler[a,b], and Nicole Probst-Hensch[a,b,*]

Repiratory Medicine 125 (2017): 25-32

[*]Corresponding author. Socinstrasse 57, 4051, Basel, Switzerland. E-mail address: Nicole.Probst@unibas.ch (N. Probst-Hensch).

[a]Swiss Tropical and Public Health Institute, Basel, Switzerland

[b]University of Basel, Basel, Switzerland

[c]Service de Pneumologie, Hôpital du Valais, Sion, Switzerland

[d]University of Geneva, Geneva, Switzerland

[e]University of Palermo, Palermo, Italy

Abstract

Although evidence for the heterogeneity of asthma accumulated, consensus for definitions of asthma phenotypes is still lacking. Obesity may have heterogeneous effects on various asthma phenotypes. We aimed to distinguish asthma phenotypes by latent class analysis and to investigate their associations with different obesity parameters in adults using a population-based Swiss cohort (SAPALDIA).

We applied latent class analysis to 959 self-reported asthmatics using information on disease activity, atopy, and age of onset. Associations with obesity were examined by multinomial logistic regression, after adjustments for age, sex, smoking status, educational level, and study centre. Body mass index, percent body fat, waist hip ratio, waist height ratio, and waist circumference were used as obesity measure.

Four asthma classes were identified, including persistent multiple symptom-presenting asthma (n = 122), symptom-presenting asthma (n = 290), symptom-free atopic asthma (n = 294), and symptom-free non-atopic asthma (n = 253). Obesity was positively associated with symptom-presenting asthma classes but not with symptom-free ones. Percent body fat showed the strongest association with the persistent multiple symptom-presenting asthma.

We observed heterogeneity of associations with obesity across asthma classes, indicating different asthma aetiologies.

4.1 Introduction

Asthma is a highly heterogeneous disease with common pathophysiological features including airway hyperresponsiveness and airway inflammation but also with divergent features distinctive of asthma subtypes (Wenzel, 2012). Non-eosinophilic asthma, characterized by an absence of eosinophils in the airway inflammation, differs from eosinophilic asthma in many aspects (Haldar and Pavord, 2007). Non-eosinophilic asthma is more likely to be refractory to corticosteroid therapy and to be non-atopic, whereas epithelial hyperplasia or hypertrophy occurs only in the eosinophilic subtype. This indicates that the variable phenotypes presumably have distinct aetiologies. Recent findings from the Genome Wide Association Studies (GWAS) also suggest that early-onset asthma has distinct genetic risk factors in comparison to the late-onset subtype (Moffatt et al., 2010). Distinguishing asthma phenotypes allows for the examination of the aetiology and pathobiology of the disease and may also contribute to a better prediction of disease progression and more targeted therapies.

Previous studies reported association between obesity and incident asthma (Beuther and Sutherland, 2007, Brumpton et al., 2013, Hjellvik et al., 2010, Egan et al., 2013). However, few studies were designed so that obesity preceded true asthma onset. Asthma can often be unnoticed or undiagnosed for a while. This hinders ensuring that obesity precedes the true incidence of asthma. Therefore, despite the accumulated reports on the association, causality remains inconclusive.

While body mass index (BMI) is the most widely used obesity measure, it might not be the optimal measure regarding its role in pathophysiology for respiratory diseases such as asthma. BMI cannot distinguish fat mass from muscular mass, and hence cannot capture one of the most important features of obesity – body fat distribution. Moreover, the relationship between obesity and asthma may be heterogeneous across different asthma

phenotypes (Fenger et al., 2012, Holguin et al., 2011, Castro-Giner et al., 2009, Ma and Xiao, 2013).

Latent class analysis (LCA) has been successfully applied to distinguish asthma phenotypes (Henderson et al., 2008, Siroux et al., 2011, Siroux et al., 2014, Spycher et al., 2008). LCA is a method to analyse the relationships among manifest variables, assuming some unobserved categorical variables (Hagenaars and McCutcheon, 2002). In this study, we applied LCA to distinguish asthma phenotypes.

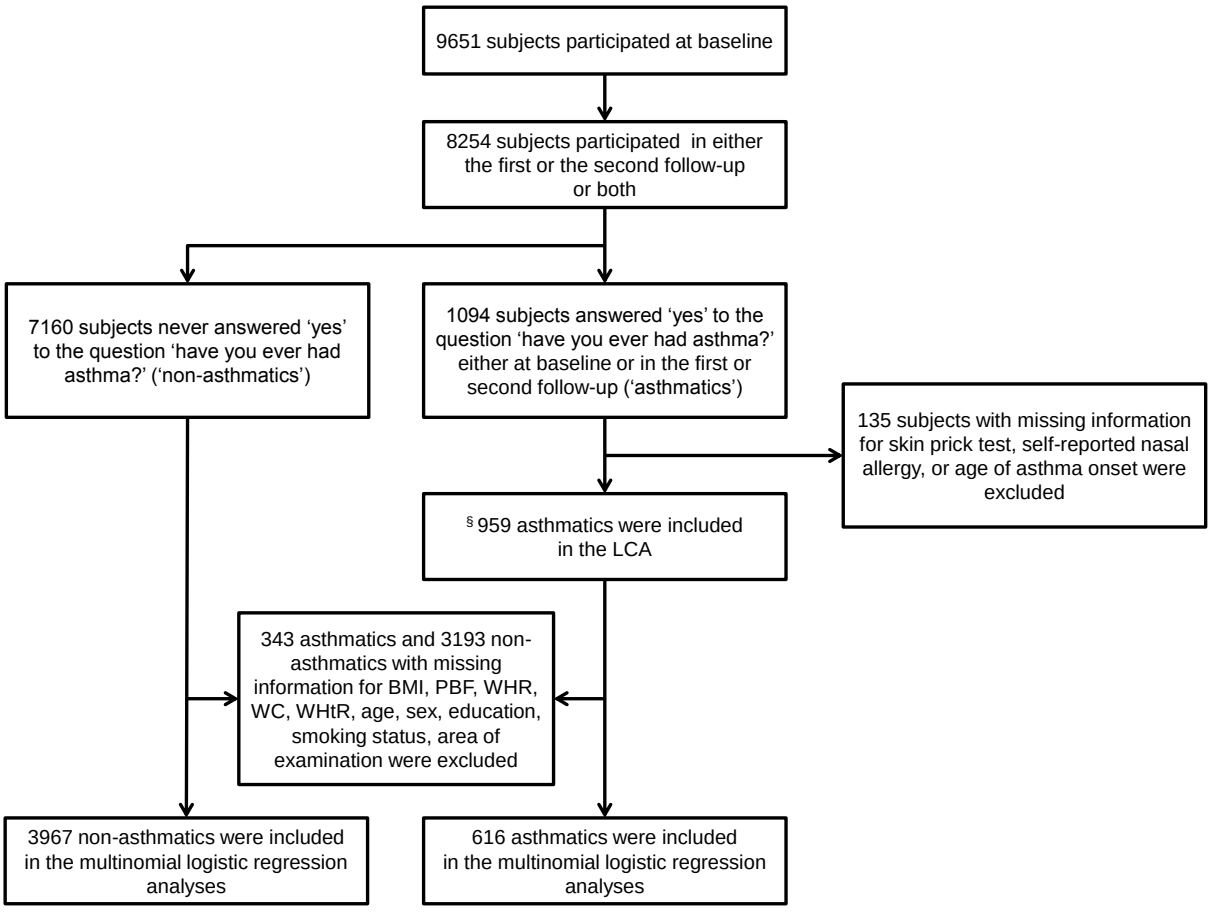
We examined the association between a variety of obesity measures – BMI, percent body fat (PBF), waist circumference (WC), waist hip ratio (WHR), waist height ratio (WHtR) – and different asthma classes found by LCA, utilizing the Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA).

4.2 Methods

4.2.1 Study population

The Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA) was initiated in 1991 (SAPALDIA1), recruiting 9651 adults aged 18 to 62 years (Ackermann-Liebrich et al., 2005). 8047 subjects from the initial cohort participated in the first follow-up in 2001-3 (SAPALDIA2) (Martin et al., 1997) and 6088 subjects in the second follow-up in 2010-11 (SAPALDIA3). At each survey, participants underwent a spirometry examination and a detailed in-person interview on respiratory health and risk factors. The subjects who participated at baseline and at least in one follow-up were included in this study (**Figure 4.1**). Ethical approval was obtained from the Swiss Academy of Medical Sciences and the regional committees for each study centre.

Figure 4.1. Flow chart of inclusion and exclusion criteria



§As sensitivity analyses, LCA applied to 677 physician-diagnosed asthmatics instead of 959 self-reported asthmatics or to 472 asthmatics who reported either asthma attack in the last 12 months or current asthma medication at least once from baseline to the second follow-up.

4.2.2 Asthma definition

Subjects were considered to be asthmatic if they answered "yes" to the question "Have you ever had asthma?" either at baseline or in the first or the second follow-up (n = 1094). After exclusion of asthmatics with missing information for skin prick test, self-reported nasal allergy, or age of asthma onset (n = 135), LCA was applied to 959 asthmatics. As a sensitivity analysis, we used physician-diagnosed asthma, restricting the sample to 677 asthmatics if they answered "yes" to both questions "Have you ever had asthma?" and "Was this confirmed by a doctor?" either at baseline or in the first or the second follow-up. In an additional sensitivity analysis, we restricted LCA to those who reported either asthma attack in the last 12 months or current asthma medication at least once from baseline to the second follow-up (n = 472).

4.2.3 Obesity measures

We examined five obesity measures including body mass index (BMI; weight in kilograms divided by the square of height in meters), percent body fat (PBF), waist hip ratio (WHR), waist circumference (WC), and waist height ratio (WHtR) in SAPALDIA3. Height was measured in SAPALDIA1, 2, and 3. Weight was asked in SAPALDIA1 and measured in SAPALDIA2 and 3. Waist and hip circumference were measured in SAPALDIA3. Bioelectric impedance was measured in SAPALDIA3 using the device Helios (Helios, Forana, Frankfurt, Germany). Fat-free mass was derived from the measured resistance and reactance using the formula of Kyle et al (Kyle et al., 2001). Fat mass was then computed as the difference between body weight and estimated fat-free mass. PBF was defined as the ratio of fat mass to body weight in percent.

4.2.4 Clustering asthma classes using LCA

Seven variables were chosen as manifest variables to reflect different aspects of asthma phenotypes: 1) asthma attack in the last 12 months (yes or no). SAPALDIA3 information on current asthma attack, current asthma medication, and current asthma symptoms was given priority and then complemented with the information from SAPALDIA2 for those who did not participate in SAPALDIA3; 2) current asthma medication (yes or no); 3) number of asthma symptoms in the last 12 months (no symptoms, one or two symptoms, or more than two symptoms). Five typical respiratory symptoms were considered: breathless while wheezing, chest tightness, shortness of breath at rest, shortness of breath after exercise, and woken by shortness of breath at night. The asthma symptom variables were constructed by counting positive answers across five symptoms and throughout study follow-ups, regardless the number of non-missing answers; 4) number of asthma symptoms repeatedly reported from baseline to the second follow-up (no persistent symptoms, one or two persistent symptoms, or more than two persistent symptoms); 5) atopy defined by positive skin prick test at baseline (yes or no), identified by an adjusted mean wheal diameter $\geq$3 mm to at least one of eight common allergens (cat fur, dog epithelia, house dust mite (*Dermatophagoides pteronyssinus*), timothy grass pollen, birch pollen, *Parietaria pollen*, and the moulds *Alternaria* and *Cladosporium*) (Martin et al., 1997, Wüthrich et al., 1995); 6) nasal allergy including hay fever reported at least once from baseline to the second follow-up (yes or no); 7) age of asthma onset $\geq$16 or $<$16 years (late or early onset), following Moffatt et al (2010). The cut-off of 16 years is the time around which boys and girls attain puberty and around puberty gender disproportionate incidence rates reverse from male to female preponderance.

LCA was applied to asthmatics with non-missing information on allergy and age of onset (n = 959). For asthma attack in the last 12 months and current asthma medication,

subjects with missing information were assumed to be negative. In order to find the appropriate number of latent classes, models were fitted with 2 to 8 latent classes. The best number was selected primarily based on the Bayesian information criterion (BIC) while the prevalence of classes was also considered. Without compromising too much on BIC, the number of latent classes resulting in more evenly distributed classes was chosen. Each subject was assigned to the latent class with the highest posterior probability.

A descriptive analysis was conducted by examining distributions across LCA-derived asthma classes of age, sex, obesity, education level, smoking status, physical activity, high-sensitive C-reactive protein (hs-CRP) level, airway obstruction, and lung function at baseline including forced expiratory volume in one second (FEV_1) as percentage of the predicted, forced vital capacity (FVC) as percentage of the predicted, FEV_1/FVC, forced expiratory flow between 25% and 75% of FVC (FEF_{25-75}) as percentage of the predicted, and bronchial hyperresponsiveness (BHR). hs-CRP was measured at SAPALDIA2. Extreme hs-CRP values, i.e. higher than 10 mg/L, were excluded. Airway obstruction was defined as $FEF_1/FVC < 0.7$ according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) (Pauwels et al., 2001). BHR was defined by 20% decline in FEV_1 on methacholine challenge, taking saline as reference. Lung function measurements were obtained using pre-bronchodilator spirometry as previously described (Ackermann-Liebrich et al., 2005). The predicted values for FEV_1, FVC, and FEF_{25-75} were obtained using Brändli et al equations (Brandli et al., 1996, Brandli et al., 2000).

4.2.5 Obesity-asthma association examined by multinomial logistic regression

LCA-derived asthma classes and non-asthmatics as reference were regressed on one of the five different obesity measures, adjusting for age, sex, smoking status, education level

and study centre. To enable comparison across different obesity measures, odds ratios (OR) were computed for 1 standard deviation (SD) increase. For interpretation purposes, we also reported ORs for overweight or obesity, following commonly used categorisation (**Table 4.S2**). Men were classified as obese if BMI $\geq$ 30 kg/m^2, WHR $\geq$ 1.0, WC $\geq$ 102 cm, or WHtR $\geq$ 0.6 and as overweight if BMI $\geq$ 25 kg/m^2, PBF $>$ 25%, WHR $\geq$ 0.9, WC $\geq$ 94 cm, or WHtR $\geq$ 0.5 but not obese. Women were classified as obese if BMI $\geq$ 30 kg/m^2, WHR $\geq$ 0.85, WC $\geq$ 88 cm, or WHtR $\geq$ 0.6 and as overweight if BMI $\geq$ 25 kg/m^2, PBF $>$ 32%, WHR $\geq$ 0.8, WC $\geq$ 80 cm, or WHtR $\geq$ 0.5 but not obese. Although PBF higher than 25% for men and 32% for women is generally considered overweight, the consensus for optimal cut-offs of PBF is lacking.

4.2.6 Additional analyses

In an attempt to examine the effect of chronic exposure to obesity, a multinomial logistic regression model was fitted to the stably overweight participants defined as being overweight (BMI $\geq$ 25 kg/m^2) from baseline to the second follow-up. Another sensitivity analysis was conducted, restricting to physically active participants. Subjects were defined as physically active if they reported either moderate physical activity $\geq$ 150 minutes/week, vigorous physical activity $\geq$ 60 minutes/week, or combined duration (duration of moderate physical activity + 2 $\times$ duration of vigorous physical activity) $\geq$ 150 minutes/week. Information on physical activity was obtained from four questions assessing frequency and duration of moderate and vigorous activities (Federal Statistical Office. Schweizerische Gesundheitsbefragung).

4.2.7 Statistical software

All analyses were conducted using R 3.1.3 (Team, 2015). In particular, R packages poLCA (Linzer and Lewis, 2011) and nnet (Ripley, 2002) were used for the LCA and multinomial logistic regression, respectively.

4.3 Results

4.3.1 Four asthma classes identified by LCA

Although five classes resulted in slightly better BIC, the model with four classes was chosen due to more evenly distributed class membership (**Table 4.S1**). The LCA with four classes distinguished persistent multiple symptom-presenting asthma (class 1, n = 122), symptom-presenting asthma (class 2, n = 290), symptom-free atopic asthma (class 3, n = 294), and symptom-free non-atopic asthma (class 4, n = 253). class 1 was characterized by a high probability of experiencing an asthma attack in the last 12 months, currently being on asthma medication, and having persistent asthma symptoms (**Table 4.1**). class 1 subjects were more likely to have late-onset asthma. class 2 was characterized by having one or two persistent or current asthma symptoms. class 3 and class 4 were characterized by experiencing neither current nor persistent asthma symptoms and were distinguished mainly by atopy and nasal allergy: class 3 subjects were more likely to have atopy and nasal allergy, whereas class 4 subjects were predominantly non-atopic and less likely to have nasal allergy. Contrasts in skin prick test were stronger than contrasts in nasal allergy self-report.

The distribution of age, sex, obesity, education level, smoking status, and physical activity did not differ much between the four classes, except that women are over-represented

Table 4.1. Class-conditional probabilities for each of the manifest variables

		Class 1	Class 2	Class 3	Class 4
Asthma attack in the last 12 months		58.6	29.4	5.1	4.1
Current asthma medication		56.2	38.2	5.7	7.3
Number of asthma symptoms	1 - 2 symptoms	2.1	84.4	19.2	9.2
in the last 12 months	$>$ 2 symptoms	96.3	0.0	0.7	0.0
Number of asthma symptoms	1 - 2 symptoms	36.0	72.1	5.5	10.6
reported at least twice	$>$ 2 symptoms	57.2	8.0	1.5	1.2
Positive skin prick test at baseline		44.9	48.2	100.0	7.2
Nasal allergy including hay fever		64.6	61.0	85.5	31.4
Age of asthma onset $\geq$ 16 years		75.2	66.8	51.7	57.2

All values are presented in per cent. class 1: persistent multiple symptom-presenting asthma; class 2: symptom-presenting asthma; class 3: symptom-free atopic asthma; class 4: symptom-free non-atopic asthma.

in class 1 (**Tables 4.2 and 4.S3**). Bronchial hyperresponsiveness (BHR) at baseline was more prevalent in classes 1, 2 and 3 than in class 4.

Notably, classes 1 and 2 showed higher prevalence of airway obstruction. For class 1 and 2, airway obstruction was already observed at baseline. FEV_1% predicted, FEV_1/FVC ratio, and FEF_{25-75}% predicted were lower in comparison to classes 3 and 4. FVC% predicted did not differ much by asthma classes.

The sensitivity analyses, applying LCA to 768 asthmatics who participated in the second follow-up, or restricting LCA to 677 physician-diagnosed asthmatics, resulted in similar class membership (**Table 4.S4**; Kappa $>$ 0.9 for both). When restricted to 472 asthmatics who ever reported either asthma attack in the last 12 months or current asthma medication, LCA could not distinguish atopic and non-atopic classes among the symptom-free asthmatics (**Table 4.S4**; Kappa $>$ 0.3). Instead, the symptom-presenting asthma (class 2) was further differentiated into atopic and non-atopic classes. In any case, the class with highest probability of multiple persistent symptoms similar to the class 1

Table 4.2. Characteristics of four LCA-derived asthma classes

	Persistent multiple symptom-presenting asthma (class 1)		Symptom-presenting asthma (class 2)		Symptom-free atopic asthma (class 3)		Symptom-free non-atopic asthma (class 4)		Non-asthmatics	
	Men	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
N	43 (35.2)	79 (64.8)	132 (45.5)	158 (54.5)	164 (55.8)	130 (44.2)	109 (43.1)	144 (56.9)	3458 (48.3)	3702 (51.7)
Age at baseline [years]	40.8 ± 13.5	38.9 ± 10.9	39.9 ± 11.8	41.5 ± 11.6	37.9 ± 11.7	37.3 ± 11.6	41.6 ± 11.2	41.0 ± 11.1	40.8 ± 11.6	41.6 ± 11.5
Education level										
Low	4 (9.3)	11 (13.9)	7 (5.3)	13 (8.2)	8 (4.9)	11 (8.5)	5 (4.6)	12 (8.4)	193 (5.6)	451 (12.2)
Middle	25 (58.1)	50 (63.3)	74 (56.1)	114 (72.2)	77 (47.0)	81 (62.3)	67 (61.5)	106 (74.1)	2098 (60.7)	2617 (70.8)
High	14 (32.6)	18 (22.8)	51 (38.6)	31 (19.6)	79 (48.2)	38 (29.2)	37 (33.9)	25 (17.5)	1166 (33.7)	630 (17.0)
Smoking status										
Never smoker	18 (41.9)	31 (39.7)	42 (31.8)	74 (46.8)	68 (41.5)	66 (50.8)	36 (33.0)	66 (45.8)	1145 (33.2)	1823 (49.6)
Former smoker	18 (41.9)	30 (38.5)	68 (51.5)	62 (39.2)	62 (37.8)	48 (36.9)	57 (52.3)	52 (36.1)	1420 (41.2)	1114 (30.3)
Current smoker	7 (16.3)	17 (21.8)	22 (16.7)	22 (13.9)	34 (20.7)	16 (12.3)	16 (14.7)	26 (18.1)	885 (25.7)	742 (20.2)
Physical activity										
Insufficient	11 (32.4)	21 (28.8)	28 (24.3)	42 (29.8)	25 (19.8)	26 (27.7)	13 (14.4)	34 (28.8)	732 (26.6)	851 (31.2)
Sufficient	23 (67.6)	52 (71.2)	87 (75.7)	99 (70.2)	101 (80.2)	68 (72.3)	77 (85.6)	84 (71.2)	2023 (73.4)	1876 (68.8)
hs-CRP* [mg/L]	2.2 ± 2.2	2.1 ± 2.1	1.4 ± 1.5	2.4 ± 2.3	1.5 ± 1.6	1.5 ± 1.5	1.3 ± 1.4	1.9 ± 2.0	1.5 ± 1.7	1.8 ± 1.9
Airway obstruction (GOLD)	16 (57.1)	30 (52.6)	53 (50.5)	60 (48.4)	44 (38.6)	26 (28.6)	35 (43.2)	41 (38.0)	683 (27.9)	500 (21.3)
Lung function at baseline										
FEV_1% pred.	88.0 ± 17.6	91.6 ± 18.3	86.1 ± 19.5	92.0 ± 14.0	94.8 ± 14.3	97.5 ± 11.6	97.8 ± 14.3	99.0 ± 15.4	99.9 ± 12.9	100.8 ± 13.4
FVC% pred.	96.4 ± 11.1	97.0 ± 14.0	95.5 ± 14.1	97.1 ± 13.4	98.5 ± 11.7	99.9 ± 11.7	101.5 ± 11.7	101.3 ± 14.7	100.6 ± 12.4	100.5 ± 13.2
FEV_1/FVC	0.73 ± 0.11	0.78 ± 0.11	0.73 ± 0.12	0.77 ± 0.08	0.77 ± 0.09	0.80 ± 0.07	0.77 ± 0.08	0.79 ± 0.07	0.79 ± 0.07	0.82 ± 0.07
FEF_{25-75}% pred.	76.1 ± 38.5	84.8 ± 32.8	76.4 ± 31.0	81.8 ± 29.5	90.3 ± 29.7	91.5 ± 25.4	93.0 ± 33.5	94.0 ± 28.7	100.4 ± 28.8	103.1 ± 28.7
BHR	13 (68.4)	23 (51.1)	34 (41.5)	58 (55.2)	51 (37.2)	50 (54.9)	18 (22.0)	35 (32.4)	236 (8.6)	506 (18.5)

Data are presented as mean ± standard deviation or number of subjects (%). Unless otherwise noted, information was retrieved from SAPALDIA3 but complemented from SAPALDIA2 for those who did not participate in SAPALDIA3. *Only available in SAPALDIA2. hs-CRP: high-sensitive C-reactive protein; FEV_1: forced expiratory volume in one second; FVC: forced vital capacity; FEF_{25-75}: forced expiratory flow between 25% and 75% of FVC, BHR: bronchial hyperresponsiveness.

again showed a stronger association with obesity compared to any other classes (data not shown).

4.3.2 Heterogeneity of obesity-asthma association

Multinomial logistic regression models were fitted to the four LCA-derived asthma classes with non-asthmatics as reference. Participants with any missing values in the five obesity measures were excluded (**Figure 4.1**). Among the five obesity measures examined as continuous determinants, BMI, PBF, WC and WHtR showed a significant association with class 1 (**Table 6.3**). PBF showed the strongest association (OR = 1.63 (95% confidence interval (CI): 1.21 - 2.20) for 1 SD increase) and further adjustment for BMI did not attenuate this (OR = 1.57 (95% CI: 0.96 - 2.56)). These results imply that in our sample 1% higher PBF is associated with a 6.1% increased risk of having the class 1 if BMI remains the same. For class 2, all five obesity measures showed a significant positive association. Interestingly, the associations of PBF, WC and WHtR to class 2 became stronger when adjusted for BMI. None of the five obesity measures showed a significant positive association to symptom-free asthma (classes 3 and 4). WHR was even negatively associated with class 4. Interaction analyses suggested a gender difference in the positive association of obesity with class 1 and the association to be stronger in men, but the results were inconsistent across different obesity measures (data not shown).

Being obese showed a positive association with classes 1 and 2 irrespective of the parameter used for classification (BMI, WHR, WC or WHtR) (**Table 4.S5**). Being overweight defined by PBF showed strong positive associations with classes 1 and 2, in comparison with being overweight defined by other obesity measures.

Table 4.3. Odds ratio for 1 SD increase in each of five obesity measures after adjustment for age, sex, smoking status, educational level, and area of examination

	Class 1	Class 2	Class 3	Class 4
BMI	1.32 [1.09, 1.60]	1.23 [1.08, 1.41]	1.01 [0.85, 1.19]	1.04 [0.88, 1.21]
PBF	1.63 [1.21, 2.20]	1.47 [1.21, 1.78]	0.96 [0.78, 1.19]	0.98 [0.79, 1.21]
adjusted for BMI	1.57 [0.96, 2.56]	1.49 [1.09, 2.04]	0.88 [0.63, 1.23]	0.86 [0.62, 1.21]
WHR	1.29 [0.98, 1.71]	1.46 [1.23, 1.75]	0.98 [0.78, 1.22]	0.79 [0.64, 0.98]
adjusted for BMI	1.13 [0.82, 1.55]	1.38 [1.14, 1.68]	0.96 [0.75, 1.23]	0.73 [0.58, 0.93]
WC	1.40 [1.10, 1.77]	1.42 [1.22, 1.66]	1.01 [0.83, 1.21]	0.93 [0.77, 1.13]
adjusted for BMI	1.21 [0.74, 1.97]	1.79 [1.30, 2.46]	0.99 [0.69, 1.41]	0.69 [0.48, 0.98]
WHtR	1.41 [1.14, 1.75]	1.38 [1.19, 1.59]	1.03 [0.86, 1.23]	0.97 [0.82, 1.16]
adjusted for BMI	1.41 [0.87, 2.26]	1.73 [1.26, 2.38]	1.09 [0.76, 1.57]	0.78 [0.55, 1.11]

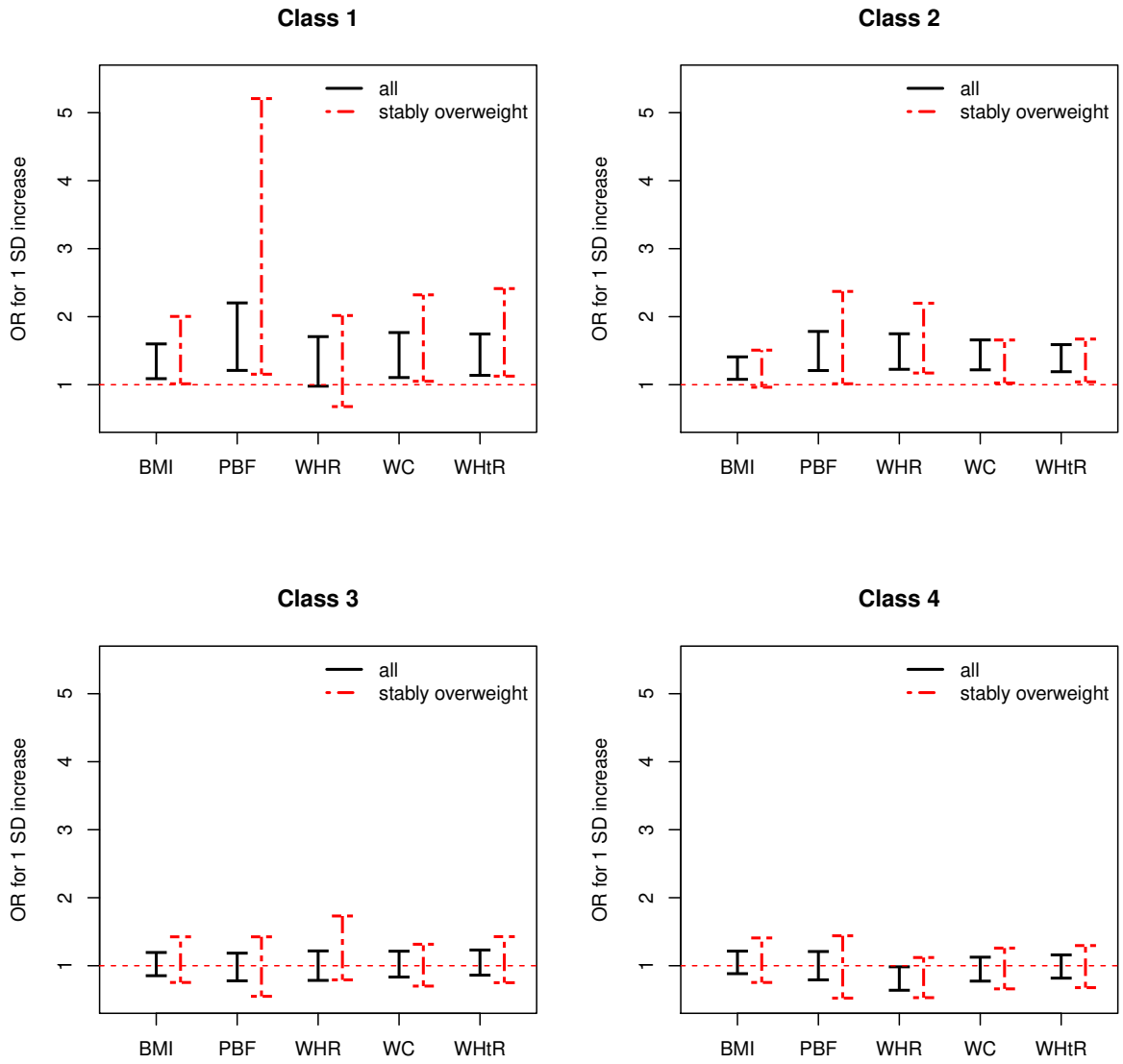
95% confidence intervals are in square brackets. Note that the odds ratios are obtained from multinomial logistic regression with non-asthmatics as reference category, and hence they are conditional on either being non-asthmatic or respective class. class 1: persistent multiple symptom-presenting asthma; class 2: symptom-presenting asthma; class 3: symptom-free atopic asthma; class 4: symptom-free non-atopic asthma.

4.3.3 Stronger association among the stably overweight

When the analysis was restricted to participants who were stably overweight (BMI $\geq$ 25 kg/m^2) from baseline to the second follow-up, the association of PBF with persistent multiple symptom-presenting asthma increased (OR = 2.45 (95% CI 1.15 - 5.21)) (**Figure 4.2 and Table 4.S6**). This corresponds to saying that among the stably overweight, 1% higher PBF is associated with a 12.4% increased risk of having class 1. BMI, WC and WHtR also showed a stronger association to class 1 when restricted to the stably overweight, but not as pronounced as for PBF. This restricted analysis did not lead to much increase in ORs for class 2.

When the analyses were restricted to physically active participants, the associations were not altered (**Table 4.S7**).

Figure 4.2. Odds ratio for 1 SD increase in each of five obesity measures before and after restriction to the stably overweight participants, adjusted for age, sex, smoking status, educational level, and area of examination



4.4 Discussion

LCA enabled us to identify asthma sub-phenotypes in an agnostic way, with *a priori* selected relevant characteristics taken into consideration. Simple classification, for example by creating a contingency table, would suffer from low power, given the large number of characteristics to consider. Unlike such simple classification, LCA reveals the co-occurrence and importance in distinguishing classes over multiple characteristics. The LCA-derived asthma classes were distinguished mostly by disease activity and atopic status. Our multinomial logistic regression analyses showed that obesity was associated with symptom-presenting asthma classes but not with symptom-free ones, indicating they may indeed have different aetiologies. Associations were consistently strongest for PBF and the highest odds ratios were observed for the association between PBF and class 1 asthma sub-phenotype.

class 1 represented relatively severe and presumably poorly controlled asthma. Subjects of this class are also more likely to have late-onset, non-atopic asthma and to be female. This finding is in line with results from earlier studies aiming to identify asthma sub-phenotypes by applying various clustering methods (Haldar et al., 2008, Moore et al., 2010, Siroux et al., 2011, Newby et al., 2014). In contrast to the previous clustering studies, we did not identify age of disease onset to be a key differentiating factor. However, categorization of age-of-onset by 16 years cut-off may not be the optimal way to assess. A recent SAPALDIA study showed that gender difference in asthma incidence attenuated in late adulthood (Hansen et al., 2015) and menopause has been associated with asthma phenotypes (van den Berge et al., 2009). It would be interesting to investigate asthma that manifests later in adulthood as potentially a separate phenotype or to examine if the association to obesity changes around menopause, but limited number of observations did not allow such additional analysis. Our analyses revealed the strongest

association of obesity with class 1, pointing to a distinct asthma entity both from a clinical and an aetiological perspective. Although this study assessed self-reported ever asthma, possibly including the asthmatics whose childhood asthma had grown out, class 1 was also identified when LCA was restricted to those who reported either asthma attack or medication during the time of SAPALDIA follow-up and showed the strongest association with obesity.

Most obesity measures examined in this study showed a positive association with the symptom-presenting asthma classes. Comparing the OR for 1 SD increase, PBF had the strongest association with class 1, suggesting that PBF captures the effects of adiposity on respiratory health better than BMI, confirming the limitation of BMI to be used as health-relevant obesity measure. In recognition of the limitation of BMI, Fenger et al examined various obesity measures in relation to asthma (Fenger et al., 2012) and lung function (Fenger et al., 2014), although they did not report any specific measure being superior to BMI. Wang et al showed stronger association of asthma to PBF than to BMI among children (Wang et al., 2014). Alternatively, this strong association between PBF and symptom-presenting asthma classes might be in part attributed to reverse causation, i.e. asthmatics tend to lack physical activity and lose muscle mass, which then associates with higher PBF.

One of the most favoured hypotheses explaining the obesity-asthma association is that low-grade chronic inflammation induced by visceral adipose tissue leads to airway inflammation. In fact, we did observe higher serum levels of high-sensitive C-reactive protein (hs-CRP) in severe asthma classes (**Table 4.2**). While a positive association between hs-CRP and BMI was observed among SAPALDIA participants, ANCOVA with LCA-derived asthma classes as factor and BMI and sex as covariates did not identify asthma classes as a statistically significant determinant of hs-CRP (data not shown). Obese

asthmatics have often shown a dissociation between symptoms and biomarkers of airway inflammation such as sputum eosinophil count or exhaled nitric oxide (McLachlan et al., 2007, Todd et al., 2007), suggesting a distinct underlying inflammatory mechanism. A recent study also reported that airway inflammation was not elevated in obese asthmatics (Sideleva et al., 2012). Elucidation of the pathophysiology linking obesity to asthma requires further studies paying attention to the heterogeneity of asthma phenotypes.

Our results might also be biased due to the fact that obese individuals may be over-diagnosed with asthma. Obesity is thought to cause physiological impairments in lung function such as reduced lung volumes and chest wall restriction (Steier et al., 2014) and dyspnoea caused by obesity-related impairments may be mistaken for asthma (Beuther and Sutherland, 2005). However, in our study, PBF showed a strong association to symptom-presenting asthma phenotypes even if adjusted for BMI. This suggests that the obesity-asthma relationship is not solely attributed to the impaired lung function caused by obesity. Moreover, we also observed decrease in FEF_{25-75}% predicted, but not in FVC% predicted, in symptom-presenting asthma classes, suggesting that obesity-asthma association is likely due to the airway inflammation rather than mechanical impairments. Independent evidence also showed that the risk of asthma over-diagnosis is not higher among obese than non-obese (Aaron et al., 2008).

Nevertheless, reverse causation remains a plausible explanation for the obesity-asthma association. One can suspect that asthmatics gain weight as a side effect of systemic corticosteroids, higher systemic inflammation, or sedentary life style. However, the commonly used asthma treatment, an inhaler, is not generally known to cause systemic side effects (Hedberg and Rossner, 2000). A more obvious hypothesis would be that respiratory symptoms hinder asthmatics from being physically active and hence lead to weight gain. Due to our study design, we cannot demonstrate that obesity preceded true asthma onset.

However, the obesity effect observed in this study did not attenuate when the analysis was restricted to physically active participants, suggesting that the observed association cannot entirely be explained by reverse causation. Interaction analyses also showed that physical activity did not modify the effect of obesity on the severe asthma classes, regardless of obesity metrics used (data not shown).

The effects of all five obesity measures became stronger when the analyses were restricted to stably overweight participants. This seems to support the causality of the association between obesity and persistent multiple symptom-presenting asthma. Recent findings from a Mendelian randomisation approach point to the causality of the association in childhood asthma (Granell et al., 2014). However, in order for a conclusive causal inference, further biological and epidemiological studies are required.

4.5 Conclusion

We demonstrated that LCA is a useful tool to disentangle the heterogeneity of asthma phenotypes. Four LCA-derived asthma classes were distinguished mainly by disease activity and atopic status. We observed heterogeneous associations with obesity across LCA-derived classes, indicating possible aetiological differences. Most obesity measures showed a positive association with symptom-presenting asthma classes but not with symptom-free ones. PBF was better than BMI in explaining persistent multiple symptom-presenting asthma class. The obesity-asthma association was stronger among the stably overweight.

Author Contributions

AJ, CS, and NPH developed the research question and designed the study. AJ, CS, and GL conducted the statistical analyses. AJ, MI, SH, EZ, PB, and NPH contributed to the draft of the manuscript. All authors read and approved the final manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Research Support

This work was supported by The Swiss National Science Foundation (grants no 33CS30-148470/1&2, 33CSCO-134276/1, 33CSCO-108796, 324730_135673, 3247BO-104283, 3247BO-104288, 3247BO-104284, 3247-065896, 3100-059302, 3200-052720, 3200-042532, 4026-028099, PMPDP3_129021/1, PMPDP3_141671/1), the Federal Office for the Environment, the Federal Office of Public Health, the Federal Office of Roads and Transport, the canton's government of Aargau, Basel-Stadt, Basel-Land, Geneva, Luzern, Ticino, Valais, and Zürich, the Swiss Lung League, the canton's Lung League of Basel Stadt/ Basel Landschaft, Geneva, Ticino, Valais, Graubünden and Zurich, Stiftung ehemals Bündner Heilstätten, SUVA, Freiwillige Akademische Gesellschaft, UBS Wealth Foundation, Talecris Biotherapeutics GmbH, Abbott Diagnostics, European Commission 018996 (GABRIEL), Wellcome Trust WT 084703MA, Exposomics EC FP7 grant(Grant agreement No: 308610). The funders had no role in study design; in the collection, analysis, and interpretation of data; in decision to publish; and in preparation of the manuscript.

Acknowledgements

The study could not have been done without the help of the study participants, technical and administrative support and the medical teams and field workers at the local study sites as well as the entire SAPALDIA team.

SAPALDIA study directorate: N. Probst-Hensch (principal investigator; epidemiology/genetic and molecular biology), T. Rochat (pneumology), C. Schindler (statistics), N. Künzli (epidemiology/exposure) and J.M. Gaspoz (cardiology).

Scientific team: J.C. Barthélémy (cardiology), W. Berger (genetic and molecular biology), R. Bettschart (pneumology), A. Bircher (allergology), C. Brombach (nutrition), P.O. Bridevaux (pneumology), L. Burdet (pneumology), D. Felber Dietrich (epidemiology), M. Frey (pneumology), U. Frey (paediatrics), M.W. Gerbase (pneumology), D. Gold (epidemiology), E. de Groot (cardiology), W. Karrer (pneumology), F. Kronenberg (genetic and molecular biology), B. Martin (physical activity), A. Mehta (epidemiology), D. Miedinger (occupational health), M. Pons (pneumology), F. Roche (cardiology), T. Rothe (pneumology), P. Schmid-Grendelmeyer (allergology), D. Stolz (pneumology), A. Schmidt-Trucksäss (physical activity), J. Schwartz (epidemiology), A. Turk (pneumology), A. von Eckardstein (clinical chemistry) and E. Zemp (epidemiology).

Scientific team at coordinating centres: M. Adam (epidemiology), I. Aguilera (exposure), S. Braun (statistics), D. Carballo (cardiology), S. Caviezel (physical activity), I. Curjuric (epidemiology), A. Di Pascale (statistics), J. Dratva (epidemiology), R. Ducret (statistics), E. Dupuis Lozeron (statistics), M. Eeftens (exposure), I. Eze (epidemiology), E. Fischer (genetic and molecular biology), M. Foraster (epidemiology), M. Germond (statistics), L. Grize (statistics), S. Hansen (epidemiology), A. Hensel (statistics), M. Imboden (genetic and molecular biology), A. Ineichen (exposure), A. Jeong (genetic and molecular biology), D. Keidel (statistics), A. Kumar (genetic and molecular biology), N. Maire (statistics), A. Mehta (epidemiology), R. Meier (exposure), E. Schaffner (statistics), T. Schikowski (epidemiology) and M. Tsai (exposure).

Local fieldworkers: Aarau: S. Brun, G. Giger, M. Sperisen and M. Stahel; Basel: C. Bürli, C. Dahler, N. Oertli, I. Harreh, F. Karrer, G. Novicic and N. Wyttenbacher; Davos: A. Saner, P. Senn and R. Winzeler; Geneva: F. Bonfils, B. Blicharz, C. Landolt and J. Rochat; Lugano: S. Boccia, E. Gehrig, M.T. Mandia, G. Solari and B. Viscardi; Montana: A.P. Bieri, C. Darioly and M. Maire; Payerne: F. Ding, P. Danieli and A. Vonnez; Wald: D. Bodmer, E. Hochstrasser, R. Kunz, C. Meier, J. Rakic, U. Schafroth and A. Walder.

Administrative staff: N. Bauer Ott, C. Gabriel and R. Gutknecht.

4.6 Supplementary Material

Table 4.S1. Summary of LCA results after 100 repetitions

Number of classes	BIC[a]	Size of the smallest class[b]	Size of the largest class[b]
2	8965 ± 3e-11	392	567
3	8883 ± 2e-9	268	354
4	8823 ± 2e-9	122	294
5	8815 ± 2e-10	94	278
6	8842 ± 2e-7	76	290
7	8887 ± 9e-1	34 [27, 52]	279 [279, 322]
8	8935 ± 1e-1	31 [31, 47]	309 [257, 309]

[a]Data are presented as mean ± standard deviation. [b]In case the results varied over 100 repetitions, data are presented as median [min, max].

Table 4.S2. Categorization used for different obesity measures

	Men			Women		
	Normal	Overweight	Obese	Normal	Overweight	Obese
BMI [kg/m^2]	< 25	≥ 25 and < 30	≥ 30	< 25	≥ 25 and < 30	≥ 30
PBF [%]	≤ 25	> 25	-	≤ 32	> 32	-
WHR	< 0.9	≥ 0.9 and < 1.0	≥ 1.0	< 0.8	≥ 0.8 and < 0.85	≥ 0.85
WC [cm]	< 94	≥ 94 and < 102	≥ 102	< 80	≥ 80 and < 88	≥ 88
WHtR	< 0.5	≥ 0.5 and < 0.6	≥ 0.6	< 0.5	≥ 0.5 and < 0.6	≥ 0.6

BMI: body mass index; PBF: percent body fat; WHR: waist hip ratio; WC: waist circumference; WHtR: waist height ratio.

Table 4.S3. Distribution of obesity measures in the LCA-derived asthma classes and non-asthmatic participants

	Persistent multiple symptom-presenting asthma (class 1)		**Symptom-presenting asthma (class 2)**		**Symptom-free atopic asthma (class 3)**		**Symptom-free non-atopic asthma (class 4)**		**Non-asthmatics**	
	Men	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
BMI [kg/m^2]	28.3 ± 3.7 (n = 34)	26.5 ± 6.1 (n = 73)	27.3 ± 4.7 (n = 115)	27.5 ± 6.8 (n = 138)	26.3 ± 3.6 (n = 127)	25.4 ± 5.4 (n = 97)	27.2 ± 3.6 (n = 89)	26.1 ± 5.1 (n = 118)	27.1 ± 3.9 (n = 2744)	25.6 ± 5.1 (n = 2691)
overweight	13 (38.2)	26 (36.1)	54 (47.0)	45 (33.1)	67 (52.8)	20 (20.6)	42 (47.2)	33 (28.0)	1343 (48.9)	803 (29.9)
obese	13 (38.2)	14 (19.4)	25 (21.7)	38 (27.9)	16 (12.6)	20 (20.6)	19 (21.3)	29 (24.6)	540 (19.7)	458 (17.0)
PBF[a] [%]	28.7 ± 6.4 (n = 21)	36.6 ± 7.2 (n = 60)	27.0 ± 5.6 (n = 87)	37.7 ± 6.7 (n = 115)	25.4 ± 6.0 (n = 97)	35.1 ± 6.0 (n = 72)	26.1 ± 5.4 (n = 70)	35.6 ± 6.3 (n = 96)	26.1 ± 5.8 (n = 2011)	35.4 ± 6.2 (n = 1963)
overweight	15 (71.4)	44 (73.3)	58 (66.7)	91 (79.1)	61 (62.9)	47 (65.3)	36 (51.4)	72 (75.0)	1166 (58.0)	1373 (69.9)
WHR[a]	0.99 ± 0.09 (n = 23)	0.84 ± 0.07 (n = 61)	0.97 ± 0.07 (n = 93)	0.86 ± 0.08 (n = 120)	0.95 ± 0.07 (n = 101)	0.82 ± 0.07 (n = 76)	0.95 ± 0.06 (n = 73)	0.82 ± 0.08 (n = 101)	0.95 ± 0.07 (n = 2133)	0.84 ± 0.08 (n = 2068)
overweight	6 (26.1)	25 (41.0)	44 (47.3)	17 (14.2)	59 (58.4)	19 (25.0)	37 (50.7)	38 (37.6)	1202 (56.4)	534 (25.8)
obese	13 (56.5)	21 (34.4)	34 (36.6)	75 (62.5)	20 (19.8)	26 (34.2)	16 (21.9)	31 (30.7)	508 (23.8)	922 (44.6)
WC[a] [cm]	102.9 ± 10.8 (n = 23)	86.0 ± 15.2 (n = 61)	99.6 ± 12.2 (n = 93)	89.5 ± 15.5 (n = 120)	96.1 ± 10.2 (n = 101)	82.1 ± 12.8 (n = 76)	96.1 ± 11.3 (n = 73)	83.7 ± 12.8 (n = 101)	96.7 ± 11.2 (n = 2134)	84.6 ± 13.0 (n = 2070)
overweight	5 (21.7)	12 (19.7)	19 (20.4)	21 (17.5)	30 (29.7)	19 (25.0)	18 (24.7)	25 (24.8)	593 (27.8)	447 (21.6)
obese	13 (56.5)	25 (41.0)	38 (40.9)	64 (53.3)	30 (29.7)	19 (25.0)	23 (31.5)	33 (32.7)	642 (30.1)	769 (37.1)
WHtR[a]	0.59 ± 0.08 (n = 23)	0.53 ± 0.10 (n = 61)	0.57 ± 0.08 (n = 93)	0.55 ± 0.10 (n = 120)	0.55 ± 0.06 (n = 101)	0.51 ± 0.08 (n = 76)	0.55 ± 0.07 (n = 73)	0.52 ± 0.08 (n = 101)	0.55 ± 0.07 (n = 2134)	0.52 ± 0.08 (n = 2070)
overweight	9 (39.1)	16 (26.2)	49 (52.7)	47 (39.2)	63 (62.4)	21 (27.6)	39 (53.4)	41 (40.6)	1238 (58.0)	793 (38.3)
obese	11 (47.8)	13 (21.3)	28 (30.1)	32 (26.7)	17 (16.8)	11 (14.5)	17 (23.3)	15 (14.9)	453 (21.2)	352 (17.0)

Data are presented as mean ± standard deviation or number of subjects (%). Unless otherwise noted, information was retrieved from SAPALDIA3 but complemented from SAPALDIA2 for those who did not participate in SAPALDIA3. [a]Only available in SAPALDIA3. Men were classified as obese if BMI $\geq$ 30 kg/m^2, WHR $\geq$ 1.0, WC $\geq$ 102 cm, or WHtR $\geq$ 0.6 and as overweight if BMI $\geq$ 25 kg/m^2, PBF $>$ 25%, WHR $\geq$ 0.9, WC $\geq$ 94 cm, or WHtR $\geq$ 0.5. Women were classified as obese if BMI $\geq$ 30 kg/m^2, WHR $\geq$ 0.85, WC $\geq$ 88 cm, or WHtR $\geq$ 0.6 and as overweight if BMI $\geq$ 25 kg/m^2, PBF $>$ 32%, WHR $\geq$ 0.8, WC $\geq$ 80 cm, or WHtR $\geq$ 0.5. BMI: body mass index; PBF: percent body fat; WHR: waist hip ratio; WC: waist circumference; WHtR: waist height ratio.

Table 4.S4. Agreement of class membership between the original and the sensitivity analyses, either by applying LCA to asthmatics who participated in SAPALDIA3, by applying LCA to physician-diagnosed asthmatics, or by applying LCA to asthmatics who ever reported either asthma attack in the last 12 months or current asthma medication

	SAPALDIA3					**Physician-diagnosed asthma**					**Ever current asthma**				
	Class 1	Class 2	Class 3	Class 4	NA[a]	Class 1	Class 2	Class 3	Class 4	NA[b]	Class 1	Class 2	Class 3	Class 4	NA[c]
Original class 1	91	0	0	0	31	88	0	0	0	34	73	35	0	0	14
Original class 2	0	228	6	5	51	0	204	5	12	69	0	102	97	8	83
Original class 3	0	2	185	40	67	0	0	180	15	99	0	1	12	78	203
Original class 4	0	1	0	210	42	0	0	0	173	80	0	6	0	60	187

NA[a]: not assessed at the second follow-up (SAPALDIA3) due to non-participation and thus was not included in the sensitivity LCA. NA[b]: not included in the sensitivity LCA due to negative or missing answer to the second question "Was this confirmed by a doctor?" with the first question being "Have you ever had asthma?" NA[c]: not included in the sensitivity LCA because neither asthma attack in the last 12 months nor current asthma medication was reported from baseline to the second follow-up. class 1: persistent multiple symptom-presenting asthma; class 2: symptom-presenting asthma; class 3: symptom-free atopic asthma; class 4: symptom-free non-atopic asthma.

Table 4.S5. Association[a] of obesity with LCA-derived asthma classes: odds ratio for overweight or obesity, based on the cut-offs specific for each of five obesity measures

		Class 1	Class 2	Class 3	Class 4
BMI	overweight	1.51 [0.89, 2.58]	1.17 [0.83, 1.64]	0.94 [0.66, 1.34]	0.99 [0.68, 1.42]
	obese	2.25 [1.22, 4.15]	1.73 [1.17, 2.55]	1.05 [0.67, 1.66]	1.40 [0.92, 2.15]
PBF	overweight	1.67 [0.99, 2.83]	1.55 [1.11, 2.16]	1.15 [0.82, 1.60]	0.94 [0.67, 1.33]
adjusted for BMI	overweight	1.26 [0.69, 2.30]	1.29 [0.88, 1.90]	1.19 [0.80, 1.77]	0.87 [0.58, 1.31]
WHR	overweight	1.57 [0.84, 2.92]	0.97 [0.63, 1.48]	0.94 [0.64, 1.39]	1.03 [0.69, 1.52]
	obese	1.80 [0.96, 3.37]	2.34 [1.57, 3.48]	0.92 [0.59, 1.42]	0.64 [0.41, 0.99]
adjusted for BMI	overweight	1.42 [0.76, 2.66]	0.94 [0.61, 1.44]	0.93 [0.62, 1.38]	0.97 [0.65, 1.44]
	obese	1.35 [0.68, 2.66]	2.15 [1.39, 3.33]	0.89 [0.54, 1.45]	0.55 [0.33, 0.89]
WC	overweight	1.08 [0.57, 2.06]	1.07 [0.71, 1.60]	1.25 [0.85, 1.83]	0.95 [0.64, 1.43]
	obese	2.06 [1.21, 3.50]	1.91 [1.36, 2.69]	1.05 [0.71, 1.55]	0.93 [0.64, 1.37]
adjusted for BMI	overweight	0.97 [0.50, 1.91]	1.04 [0.68, 1.60]	1.26 [0.83, 1.90]	0.88 [0.58, 1.36]
	obese	1.57 [0.75, 3.29]	1.81 [1.12, 2.91]	1.06 [0.61, 1.86]	0.77 [0.45, 1.34]
WHtR	overweight	0.73 [0.41, 1.29]	1.37 [0.95, 1.98]	1.01 [0.69, 1.46]	0.99 [0.68, 1.45]
	obese	2.12 [1.15, 3.91]	2.21 [1.44, 3.39]	1.05 [0.63, 1.73]	0.97 [0.59, 1.58]
adjusted for BMI	overweight	0.58 [0.30, 1.11]	1.33 [0.87, 2.04]	1.01 [0.65, 1.58]	0.90 [0.57, 1.40]
	obese	1.21 [0.44, 3.30]	2.06 [1.05, 4.04]	1.06 [0.49, 2.30]	0.76 [0.35, 1.62]

[a]Adjusted for age, sex, smoking status, educational level, and study centre. 95% confidence intervals are in square brackets. Men were classified as obese if BMI $\geq$ 30 kg/m^2, WHR $\geq$ 1.0, WC $\geq$ 102 cm, or WHtR $\geq$ 0.6 and as overweight if BMI $\geq$ 25 kg/m^2, PBF $>$ 25%, WHR $\geq$ 0.9, WC $\geq$ 94 cm, or WHtR $\geq$ 0.5. Women were classified as obese if BMI $\geq$ 30 kg/m^2, WHR $\geq$ 0.85, WC $\geq$ 88 cm, or WHtR $\geq$ 0.6 and as overweight if BMI $\geq$ 25 kg/m^2, PBF $>$ 32%, WHR $\geq$ 0.8, WC $\geq$ 80 cm, or WHtR $\geq$ 0.5. BMI: body mass index; PBF: percent body fat; WHR: waist-hip ratio; WC: waist circumference; WHtR: waist-height ratio. Odds ratios are obtained from multinomial logistic regression with non-asthmatics as reference category, and hence are conditional on either being non-asthmatic or respective class. Class 1: persistent multiple symptom-presenting asthma; Class 2: symptom-presenting asthma; Class 3: symptom-free atopic asthma; Class 4: symptom-free non-atopic asthma.

Table 4.S6. Association[a] of obesity with LCA-derived asthma classes: odds ratio for 1 standard deviation (SD) increase in each of five obesity measures among the stably overweight participants

	Class 1	**Class 2**	**Class 3**	**Class 4**
BMI	1.42 [1.01, 2.00]	1.20 [0.96, 1.51]	1.04 [0.75, 1.42]	1.03 [0.75, 1.41]
PBF	2.45 [1.15, 5.21]	1.55 [1.01, 2.37]	0.89 [0.55, 1.43]	0.87 [0.52, 1.44]
adjusted for BMI	2.15 [0.75, 6.16]	1.47 [0.83, 2.62]	0.76 [0.40, 1.42]	0.74 [0.38, 1.44]
WHR	1.17 [0.68, 2.02]	1.60 [1.17, 2.20]	1.17 [0.79, 1.73]	0.77 [0.53, 1.12]
adjusted for BMI	1.01 [0.56, 1.81]	1.55 [1.12, 2.14]	1.17 [0.78, 1.77]	0.75 [0.51, 1.10]
WC	1.56 [1.05, 2.32]	1.30 [1.03, 1.66]	0.96 [0.70, 1.32]	0.91 [0.66, 1.26]
adjusted for BMI	1.45 [0.67, 3.13]	1.37 [0.90, 2.11]	0.83 [0.50, 1.38]	0.75 [0.46, 1.24]
WHtR	1.65 [1.13, 2.41]	1.32 [1.04, 1.67]	1.04 [0.75, 1.43]	0.94 [0.68, 1.30]
adjusted for BMI	1.93 [0.86, 4.33]	1.47 [0.94, 2.30]	1.01 [0.58, 1.79]	0.79 [0.46, 1.33]

[a]Adjusted for age, sex, smoking status, educational level, and study centre. 95% confidence intervals are in square brackets. Being stably overweight was defined as being overweight (BMI $\geq$ 25 kg/m^2) from baseline to the second follow-up. Odds ratios are obtained from multinomial logistic regression with non-asthmatics as reference category, and hence are conditional on either being non-asthmatic or respective class. class 1: persistent multiple symptom-presenting asthma; class 2: symptom-presenting asthma; class 3: symptom-free atopic asthma; class 4: symptom-free non-atopic asthma.

Table 4.S7. Association[a] of obesity with LCA-derived asthma classes: odds ratio for 1 SD increase in each of five obesity measures among the physically active participants

	Class 1	**Class 2**	**Class 3**	**Class 4**
BMI	1.34 [1.08, 1.67]	1.27 [1.09, 1.48]	0.98 [0.81, 1.19]	1.07 [0.90, 1.28]
PBF	1.73 [1.24, 2.41]	1.59 [1.27, 2.00]	0.99 [0.78, 1.26]	1.06 [0.83, 1.34]
WHR	1.31 [0.94, 1.84]	1.52 [1.22, 1.89]	1.00 [0.77, 1.28]	0.86 [0.67, 1.10]
WC	1.45 [1.11, 1.90]	1.47 [1.23, 1.77]	1.04 [0.84, 1.29]	0.99 [0.80, 1.22]
WHtR	1.42 [1.11, 1.81]	1.43 [1.20, 1.69]	1.04 [0.85, 1.27]	1.01 [0.83, 1.22]

[a]Adjusted for age, sex, smoking status, educational level, and study centre. 95% confidence intervals are in square brackets. Odds ratios are obtained from multinomial logistic regression with non-asthmatics as reference category, and hence are conditional on either being non-asthmatic or respective class. Class 1: persistent multiple symptom-presenting asthma; Class 2: symptom-presenting asthma; Class 3: symptom-free atopic asthma; Class 4: symptom-free non-atopic asthma.

5. Article II

DNA methylation in inflammatory pathways modifies the association between BMI and adult-onset non-atopic asthma

Ayoung Jeong[a,b], Medea Imboden[a,b], Akram Ghantous[c], Alexei Novoloaca[c], Anne-Elie Carsin[d,e,f], Manolis Kogevinas[d,e,f], Christian Schindler[a,b], Gianfranco Lovison[g], Zdenko Herceg[c], Cyrille Cuenin[c], Roel Vermeulen[h], Deborah Jarvis[i], André F S Amaral[i], Florian Kronenberg[j], Paolo Vineis[k,l], and Nicole Probst-Hensch[a,b,*]

International Journal of Environmental Research and Public Health 16.4

(2019): 600

*Correspondence to: Nicole Probst-Hensch, Swiss Tropical and Public Health Institute, Socinstrasse 57, 4051 Basel, Switzerland; Tel: +41 (0)61 284 8378; Mobile: +41 (0)79 280 3414; Email: nicole.probst@swisstph.ch

[a]Swiss Tropical and Public Health Institute, Basel, Switzerland

[b]University of Basel, Basel, Switzerland

[c]International Agency for Research on Cancer, Lyon, France

[d]ISGlobal, Barcelona Institute for Global Health, Barcelona, Spain

[e]Universitat Pompeu Fabra (UPF), Barcelona, Spain

[f]CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain

[g]University of Palermo, Palermo, Italy

[h]Utrecht University, Institute for Risk Assessment Sciences, Environmental Epidemiology Division, Utrecht, Netherlands

[i]Population Health and Occupational Disease, National Heart and Lung Institute, Imperial College, London, UK

[j]Division of Genetic Epidemiology, Medical University of Innsbruck, Austria

[k]MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, UK

[l]Italian Institute for Genomic Medicine (IIGM), Turin, Italy

Abstract

A high body mass (BMI) index has repeatedly been associated with non-atopic asthma, but the biological mechanism linking obesity to asthma is still poorly understood. We aimed to test the hypothesis that inflammation and/or innate immunity plays a role in the obesity-asthma link. DNA methylome was measured in blood samples of 61 non-atopic participants with asthma and 146 non-atopic participants without asthma (non-smokers for at least 10 years) taking part in the Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA) study. Modification by DNA methylation of the association of BMI or BMI change over 10 years with adult-onset asthma was examined at each CpG site and differentially methylated region. Pathway enrichment tests were conducted for genes in *a priori* curated inflammatory pathways and the NLRP3-IL1B-IL17 axis. The latter was chosen on the basis of previous work in mice. Inflammatory pathways including glucocorticoid/PPAR signaling ($p = 0.0023$), MAPK signaling ($p = 0.013$), NF-$\varkappa$B signaling ($p = 0.031$), and PI3K/AKT signaling ($p = 0.031$) were enriched for the effect modification of BMI, while NLRP3-IL1B-IL17 axis was enriched for the effect modification of BMI change over 10 years ($p = 0.046$). DNA methylation measured in peripheral blood is consistent with inflammation as a link between BMI and adult-onset asthma and with the NLRP3-IL1B-IL17 axis as a link between BMI change over 10 years and adult-onset asthma in non-atopic participants.

5.1 Introduction

Obesity and overweight have repeatedly been linked to asthma (Beuther and Sutherland, 2007, Egan et al., 2013), with several studies reporting a stronger association of obesity or overweight with non-atopic as compared to atopic asthma (Castro-Giner et al., 2009, Fenger et al., 2012) and with late-onset asthma compared to early-onset asthma (Haldar et al., 2008, Moore et al., 2010). We have previously observed heterogeneity of the overweight-asthma association across asthma classes identified by latent class analysis (Jeong et al., 2017).

The biological mechanism linking obesity and overweight with asthma is yet poorly understood. Excessive adipose tissue may increase the work associated with breathing, reduce lung volume, and promote airway hyperresponsiveness and airway narrowing (Shore, 2008, Steier et al., 2014). However, the more likely hypothesis is that the obesity-asthma association is not entirely mechanical, but that obesity-related chronic inflammation contributes to asthma development. Adiposity is characterized by dysregulated production of pro-inflammatory cytokines and infiltration and activation of macrophages (Suganami et al., 2005, Weisberg et al., 2003). While M2 macrophages are predominant in non-obese adipose tissue, pro-inflammatory M1 macrophages increase in obese adipose tissue, leading to low-grade chronic systemic inflammation (Castoldi et al., 2016). Whether and how obesity and overweight leads to airway inflammation is controversial. An interesting finding in mice experiments pointed to NLRP3 (nucleotide-binding domain, leucine-rich repeats-containing family, pyrin domain-containing-3) inflammasome and interleukin-17 (IL17) producing innate lymphoid cell group 3 (ILC3) cells as a link between obesity and airway hyperresponsiveness (AHR) (Kim et al., 2014). On recognition of various danger signals, NLRP3 inflammasome produces interleukin-1β(IL1B) via caspase-1. IL1B, in turn, activates ILC3 cells to produce interleukin-17 (IL17), leading to AHR. Kim and her

colleagues demonstrated that the NLRP3-IL1B-IL17 axis is crucial in AHR development in obese mice (Kim et al., 2014).

High-throughput arrays allow cost-effective genome-wide quantification of DNA methylation. The epigenome-wide association study (EWAS) design has been successfully applied to identify methylation markers measured in peripheral blood related to a variety of endogenous and environmental insults as well as health outcomes. Recently, the largest EWAS on asthma identified DNA methylation at several immunity and inflammation related CpG sites to be associated with asthma in children (Xu et al., 2018, Forno et al., 2018, Reese et al., 2018). Several studies reported methylation markers of obesity measured in peripheral blood (Dick et al., 2014, Wahl et al., 2017, Xu et al., 2013). Several studies reported methylation markers of obesity and overweight measured in peripheral blood, which in part reflected inflammatory pathways (Dick et al., 2014, Wahl et al., 2017, Xu et al., 2013). Rastogi and her colleagues reported differential DNA methylation in obese children with non-atopic asthma that was consistent with a role of inflammation (Rastogi et al., 2013). However, differential DNA methylation related to obesity-associated asthma in adults is largely unexplored. An earlier EWAS study in adults showed heterogeneity in differential DNA methylation patterns across inflammatory sub-phenotypes of asthma (Gunawardhana et al., 2014).

In the light of suggestive evidence for inflammation as a mediator in the overweight/obesity-asthma link, we formally explored whether interaction signals between BMI or BMI change and DNA methylation in peripheral blood on non-atopic adult-onset asthma are enriched for signals mapping to inflammatory pathways. Specifically we tested the hypothesis by conducting an epigenome-wide interaction study (EWIS) followed by candidate pathway enrichment analysis for a priori curated inflammatory pathways and the NLRP3-IL1B-IL17 axis, making use of the information from the Swiss Cohort Study on

Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA). Identification of differential DNA methylation enriched in the candidate pathways would add further support that inflammation and/or innate immunity play a role in overweight-asthma link, although the inflammation was not directly measured. This hypothesis-driven approach was corroborated by agnostic pathway enrichment analysis in combination with differentially methylated region (DMR) analysis.

5.2 Materials and methods

5.2.1 Study samples

The Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA) was initiated in 1991 (SAPALDIA1), recruiting 9651 participants in eight regions representing various meteorological and geographical environments in Switzerland. 8047 and 6088 out of the 9651 participants were followed-up in the second and the third survey, respectively (SAPALDIA2 in 2001-3 and SAPALDIA3 in 2010-11). The detailed study protocol was reported previously (Ackermann-Liebrich et al., 2005, Martin et al., 1997).

We conducted a nested case-control study of adult-onset asthma among the non-atopic SAPALDIA3 participants, all of whom were non-smokers for at least 10 years before blood draw and interview. Cases were selected among the participants with self-reported asthma and self-reported age of onset later than 16 years, based on the availability of archived blood samples and covariate information. Controls were randomly selected among the participants who never reported the following throughout the surveys: self-reported asthma; physician-diagnosed asthma; asthma attack in the last 12 months; current asthma medication; wheezing without cold in the last 12 months; three or more asthma-related symptoms in the last 12 months (symptoms considered: breathless while wheezing; woken up with a

feeling of chest tightness; attack of shortness of breath after exercise; attack of shortness of breath while at rest; woken by attack of shortness of breath). Cases and controls with positive skin prick test at baseline defined as an adjusted mean wheal diameter ≥3 mm to at least one of eight common respiratory allergens were excluded (allergens considered: cat fur, dog epithelia, house dust mite (*Dermatophagoides pteronyssinus*), timothy grass pollen, birch pollen, *Parietaria* pollen, and the molds *Alternaria* and *Cladosporium*). In total 61 cases and 146 controls were examined in the EWIS followed by the pathway enrichment tests. Study samples' characteristics are summarized in **Table 5.1**. All participants gave written informed consent and ethical approval was obtained from the Swiss Academy of Medical Sciences and the regional committees for each study center.

5.2.2 Covariates

Weight and height were measured and body mass index (BMI) was computed as weight in kilograms divided by the square of height in meters. BMI change was defined as the difference in BMI between SAPALDIA3 and SAPALDIA2. Negative values of the BMI change mean reduction in BMI. Educational level was categorized from self-reported highest education into "low" (primary school), "middle" (secondary/middle school or apprenticeship), and "high" (college or university). Pack-years of cigarettes smoked in life were computed from self-reported number of cigarettes smoked per day and smoking history. Physical activity was dichotomized from self-reported frequency and duration of moderate and vigorous physical activity into "sufficiently active" (either moderate physical activity $\geq$ 150 min/week, vigorous physical activity $\geq$ 60 min/week, or combined duration (duration of moderate physical activity + 2 $\times$ duration of vigorous physical activity) $\geq$ 150 min/week) and "insufficiently active" (otherwise).

In order to confirm that BMI is related to chronic inflammation, we examined as-

Table 5.1. Study samples' characteristics by adult-onset asthma status at SAPALDIA 3

	Cases	**Controls**
N	61	146
Age [year]	60.8 (15.6)	57.4 (15.0)
Female	43 (70%)	82 (56%)
BMI[a] [kg/m^2]	25.7 (5.8)	24.5 (4.8)
BMI change[b] [kg/m^2]	0.4 (2.0)	0.5 (1.6)
Smoking[c]		
Former	27 (44%)	50 (34%)
Never	34 (56%)	96 (66%)
Pack-years[d]	7.8 (13.3)	6.8 (11.6)
Education level[e]		
Low	0 (0%)	2 (1%)
Middle	43 (70%)	94 (64%)
High	18 (30%)	50 (34%)
Physical activity[f]		
Insufficiently active	18 (30%)	30 (21%)
Sufficiently active	42 (69%)	113 (77%)
N/A	1 (2%)	3 (2%)
Bench time[g] [min]	80.0 (34.0)	82.5 (32.5)
hs-CRP[h] [min]	1.3 (1.4)	0.7 (1.2)

Data are presented as count (%) or median (interquartile range). [a]Measured at SAPALDIA3. [b]Change in BMI between SAPALDIA2 and SAPALDIA3. [c]Former smokers had not smoked for at least 10 years before blood was drawn. [d]Only computed in former smokers (pack-years were set to zero for never smokers). [e]Low = primary school; middle = secondary/middle school or apprenticeship; high = college or university. [f]Sufficiently active at SAPALDIA3= either moderate physical activity $\geq$ 150 min/week, vigorous physical activity $\geq$ 60 min/week, or combined duration (duration of moderate physical activity + 2 $\times$ duration of vigorous physical activity) $\geq$ 150 min/week; insufficiently active = otherwise. N/A = not available. [g] Time elapsed between blood draw and storage in freezer. [h] Measured at SAPALDIA2.

sociation between BMI and high-sensitive C-reactive protein (hs-CRP) within the study subjects (n = 206; one subject was excluded due to missing information on hs-CRP). We used information on both BMI and hs-CRP at SAPALDIA2 because hs-CRP was measured only at SAPALDIA2. Log-transformed hs-CRP was regressed on BMI after adjustment for age, sex, education level, study area, and pack-years of cigarettes smoked up to SAPALDIA2.

5.2.3 Methylome

Peripheral blood samples had been collected at SAPALDIA3, the second follow-up visit in 2010 of the cohort study. Pre-analytically, the blood samples were processed and the buffy coat fraction was archived at $-80°C$ for five years until DNA extraction using the QIAamp Blood Mini Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions (Hebels et al., 2013). A small number of samples yielded limited DNA quantity and were replaced by DNA extracted (using the Gentra Puregene Blood Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions) from whole blood of the same venipuncture as used for buffy coats. Bisulfite conversion of 600 ng of each sample was performed using the EZ-96 DNA Methylation-GoldT Kit according to the manufacturer's protocol (Zymo Research, Orange, CA, USA). Then, 200 ng of bisulfite-converted DNA was used for hybridization on the Illumina Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA, USA), following the Illumina Infinium HD Methylation protocol. Each array consisted of 96-samples distributed equally among 8 chips. The arrays were designed such that batch effects (e.g. sample position and intra- and inter-variability in arrays and chips) do not completely confound with biological covariates. This design allows the retention of biological variation (including the variable of interest) after correction for technical variation. Specifically, each chip incorporated proportional amounts of

samples representing the different centers, confounding factors and cases-control status. Cases and controls were also placed on the chips (not following a specific sequence) in order to minimize technical variation between them. Raw fluorescence intensities were retrieved and preprocessed using the R package "*minfi*" (Aryee et al., 2014). One sample with sex mismatch was excluded. Background correction and dye bias correction were performed using Noob (normal-exponential out-of-band) procedure (Triche et al., 2013). DNA methylation levels were expressed as β values, defined as the ratio of methylated intensity over total intensity with offset = 100. β values were set to missing if the detection p-value was higher than 10^{-16}. Probes on sex chromosome were excluded. Probes were then filtered by call rate < 0.95. All samples had call rate > 0.95. Beta-mixture quantile normalization (BMIQ) procedure was conducted to correct for the Illumina probe design bias (Teschendorff et al., 2013). The probes known to hybridize with multiple genomic locations or to target CpG sites overlapping known SNPs with minor allele frequency greater than 1% in Europeans were excluded (Chen et al., 2013). Finally 430,591 CpGs were ready for analysis. In addition, principal component analysis (PCA) was conducted on the 220 control probes incorporated on the Illumina chip following Lehne and Drong (Lehne et al., 2015) and β values were regressed on the first 30 components. All the statistical analyses used the resulting residuals in place of the β values to account for batch effects.

5.2.4 EWIS of DNA methylation and BMI on adult-onset asthma

Logistic regression models were fitted for adult-onset asthma status on BMI at SAPALDIA3, residual of the β value at each CpG site, and their multiplicative interaction upon adjustment for age (in years), sex, education level, study area, pack-years of cigarettes smoked in life, bench time (in minutes), and Houseman estimates (Houseman et al., 2012)

of white blood cell composition for B cells, CD4 T cells, CD8 T cells, natural killer cells, monocytes, and eosinophils.

$$Asthma \sim BMI \times Residual_i + Age + Sex + Education + Area + Packyear + Benchtime + Bcell + CD4T + CD8T + NK + Mono + Eos \quad (i \in [1, \ldots, 430591])$$

We did not adjust for neutrophils because immune response in non-atopic participants was possibly driven by neutrophil proliferation (Annunziato et al., 2015, Linden and Dahlen, 2014) and therefore adjustment for neutrophils could obscure the association of interest. Despite the female preponderance in cases compared to controls (**Table 5.1**), we did not consider stratification or effect modification by sex based on our observation that the association between BMI and non-atopic adult-onset asthma did not differ by sex (**Table 5.S3**). The interaction was considered genome-wide significant when the p-value from the interaction term was smaller than 0.1 after the Benjamini-Hochberg correction for multiple testing. As a sensitivity analysis, the same EWIS was repeated after further adjustment either for physical activity or for neutrophil estimates.

5.2.5 EWIS of DNA methylation and BMI change on adult-onset asthma

Logistic regression models were fitted for adult-onset asthma status on BMI change, residual of the β value at each CpG site, and their multiplicative interaction after adjustment for the same set of covariates as above and additionally for BMI at SAPALDIA2.

$$Asthma \sim (BMI_{S3} - BMI_{S2}) \times Residual_i + BMI_{S2} + Age + Sex + Education + Area + Packyear + Benchtime + Bcell + CD4T + CD8T + NK + Mono + Eos \quad (i \in [1, \ldots, 430591])$$

The same sensitivity analyses were conducted as above.

5.2.6 Candidate pathway enrichment analyses using Weighted Kolmogorov-Smirnov (WKS) method

Genes relevant to the NLRP3-IL1B-IL17 axis were curated based on Kim et al (Kim et al., 2014). Inflammation-related genes were curated previously by Loza et al into 17 mutually exclusive pathways (Loza et al., 2007). The complete list of the genes assigned to each pathway can be found in **Table 5.S1**. CpG sites were then assigned to the pathway if the CpG sites reside within 200 bp upstream or downstream of the genes included in each pathway.

We tested if the pathways are over-represented in the EWIS results by applying the Weighted Kolmogorov-Smirnov (WKS) enrichment test (Charmpi and Ycart, 2015). Using this algorithm, the absolute Z-statistics of the CpG sites assigned to each pathway (e.g. 219 CpG sites assigned to NLRP3-IL1B-IL17 axis) were compared with the null distribution created by 10000 Monte-Carlo simulations of the absolute Z-statistics from the entire 430 591 CpG sites. In this approach, Z-statistics from all CpGs mapped to a pathway were compared to the null distribution without selection based on EWIS-derived p-values. Over-representation of the pathway was determined by Kolmogorov-Smirnov tests. Pathways with WKS p-value < 0.05 were declared as enriched. The procedure includes permutation-based multiple testing correction (Charmpi and Ycart, 2015, van der Laan et al., 2005).

5.2.7 Identification of differentially methylated regions (DMR)

We used the R package DMRcate to identify DMRs (Peters et al., 2015). The Z-statistics from EWIS were squared and smoothed using a Gaussian kernel with a bandwidth of 1000 bp and scaling factor of 2, which is equivalent to the kernel standard deviation of

500 bp. P-values were computed for each CpG site by comparison to the null distribution of the smoothed estimates. The regions containing at least one CpG site with Benjamini-Hochberg adjusted p-value < 0.05 were declared as significant. The significant DMRs were annotated to the genes whose promoter region, defined as 2000 bp from the transcription start site, overlaps with the DMRs.

5.2.8 Agnostic pathway enrichment analyses using Ingenuity Pathway Analysis (IPA)

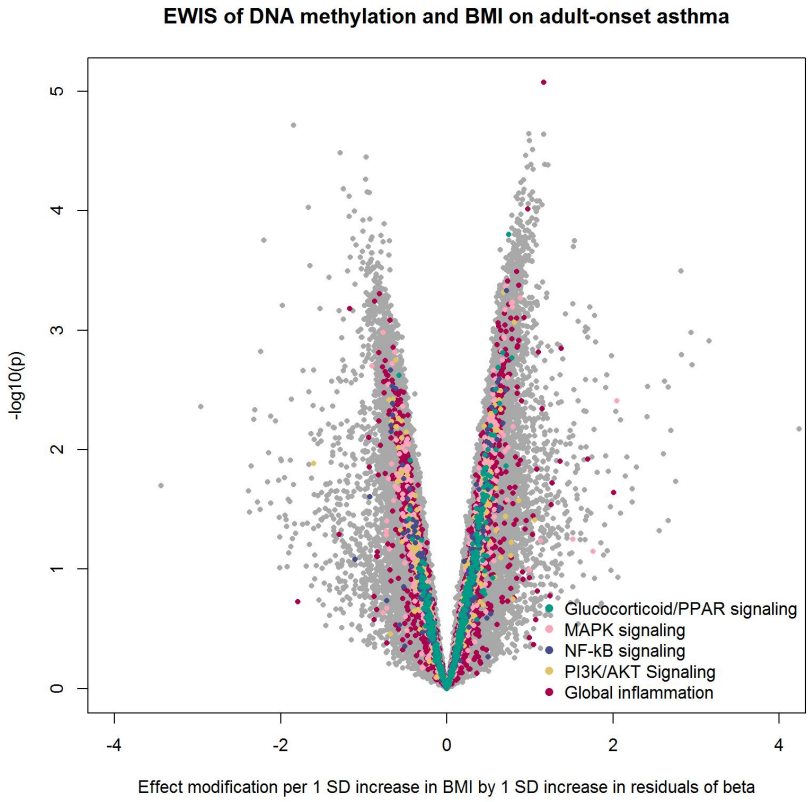
The 1305 genes annotated to the 1131 DMRs identified as significant effect modification of BMI on adult-onset asthma were tested for over-representation using IPA (http://www.ingenuity.com/; QIAGEN, Redwood City, CA, USA) canonical pathway analysis. In brief, the maximum effect modification estimate and the minimum Benjamini-Hochberg adjusted p-value for each DMR were assumed to represent the expression level and the p-value, respectively, for the gene annotated to the DMR. The DMRs annotated multiple genes constituted multiple entries each annotated a single gene. The DMRs with no gene annotation were excluded (n = 114). The 20 genes annotated to the 18 DMRs for BMI change were too few to conduct the same pathway analysis.

5.3 Results

From association analysis using SAPALDIA2 information, we confirmed a positive association between BMI and hs-CRP. One unit increase in BMI was associated with 0.1 unit increase in log-transformed hs-CRP (95% confidence interval [0.07, 0.14]; $p < 10^{-8}$). We conducted an EWIS of DNA methylation and BMI or BMI change over 10 years on adult-onset asthma among non-atopic, non-smoking SAPALDIA3 participants (**Table 5.1**). We

found no epigenome-wide significant effect modification after multiple testing corrections. Sensitivity analyses with additional adjustment for physical activity or neutrophil estimates also resulted in no epigenome-wide significant CpG sites. **Figures 5.1, 5.2, and 5.S1-5.S4** summarize the EWIS results.

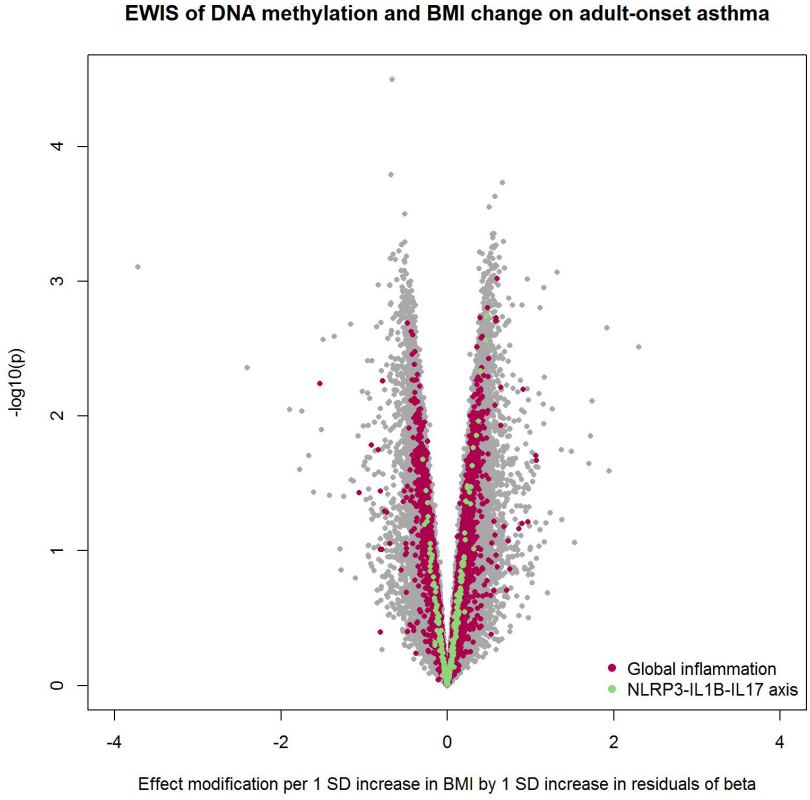
Figure 5.1. Volcano plot from the EWIS of DNA methylation and BMI on adult-onset asthma



The EWIS fitted logistic regression models of adult-onset asthma on BMI, residuals of DNA methylation at a single CpG site, and their multiplicative interaction, upon adjustment for age, sex, education level, study area, pack-years of cigarettes smoked in life, bench time, and white blood cell composition estimates for B cells, CD4 T cells, CD8 T cells, natural killer cells, monocytes, and eosinophils. The CpGs assigned to the pathway enriched with $p < 0.05$ are highlighted in colors. No line of significance was drawn as no CpG reached genome-wide significance after multiple testing corrections.

After pathway enrichment analysis of 17 *a priori* curated inflammatory pathways (Loza et al., 2007), we found an over-representation of effect modification by DNA methylation of BMI on adult-onset asthma in several pathways: Glucocorticoid/PPAR (peroxi-

Figure 5.2. Volcano plot from the EWIS of DNA methylation and BMI change on adult-onset asthma



The EWIS fitted logistic regression models of adult-onset asthma on BMI change, residuals of DNA methylation at a single CpG site, and their multiplicative interaction, upon adjustment for age, sex, education level, study area, pack-years of cigarettes smoked in life, bench time, and white blood cell composition estimates for B cells, CD4 T cells, CD8 T cells, natural killer cells, monocytes, and eosinophils. The CpGs assigned to the pathway enriched with $p < 0.05$ are highlighted in colors. No line of significance was drawn as no CpG reached genome-wide significance after multiple testing corrections.

some proliferator-activated receptor) signaling, MAPK (mitogen-activated protein kinase) signaling, NF-κB (nuclear factor kappa-B) signaling, and PI3K/AKT (phosphatidylinositol-3-kinases/protein kinase B) signaling (**Table 5.2**). The pathway "global inflammation", defined as the entirety of the 1027 genes assigned to the 17 inflammation pathways, also showed enrichment. In the sensitivity analyses, the enrichment of PI3K/AKT signaling disappeared after adjustment for physical activity, while the enrichment of NF-κB signaling and PI3K/AKT signaling disappeared after adjustment for neutrophil estimates (**Table 5.2**).

Table 5.2. EWIS of DNA methylation and BMI on adult-onset asthma: enrichment test results for 17 inflammation pathways and NLRP3-IL1B-IL17 axis

Pathway	**#Genes**	**#CpGs**	**Enrichment p-value**		
			Basic model	**Adjusted for physical activity**	**Adjusted for neutrophil counts**
Adhesion-extravasation-migration	142	1737	0.48	0.30	0.37
Apoptosis signaling	68	1210	0.22	0.34	0.32
Calcium signaling	14	413	0.81	0.72	0.70
Complement cascade	40	483	0.92	0.73	0.96
Cytokine signaling	172	1883	0.070	0.053	0.067
Eicosanoid signaling	39	450	0.58	0.78	0.55
Glucocorticoid/PPAR signaling	21	404	**0.0023**	**0.0053**	**0.0039**
G-Protein coupled receptor signaling	42	1133	0.74	0.49	0.66
Innate pathogen detection	50	515	0.89	0.72	0.88
Leukocyte signaling	121	1429	0.14	0.059	0.090
MAPK signaling	118	2682	**0.013**	**0.0036**	**0.018**
Natural killer cell signaling	31	368	0.54	0.41	0.51
NF-κB signaling	33	654	**0.031**	**0.0028**	0.054
Phagocytosis-Ag presentation	39	1058	0.81	0.72	0.66
PI3K/AKT signaling	37	907	**0.031**	0.23	0.053
ROS/glutathione/cytotoxic granules	22	190	0.58	0.45	0.53
TNF superfamily signaling	38	537	0.78	0.69	0.73
Global inflammation[§]	1027	15985	**0.0026**	**0.011**	**0.0057**
NLRP3-IL1B-IL17 axis	11	219	1.00	0.99	1.00

The basic model regressed adult-onset asthma on BMI, residuals of DNA methylation at a single CpG site, and their multiplicative interaction, upon adjustment for age, sex, education level, study area, pack-years of cigarettes smoked in life, bench time, and white blood cell composition estimates for B cells, CD4 T cells, CD8 T cells, natural killer cells, monocytes, and eosinophils. [§]Total of the 17 inflammation pathways; the number of CpG in this pathway (15985) is smaller than the sum of the CpGs assigned to 17 pathways because there are CpGs assigned to multiple pathways, although the 17 pathways are mutually exclusive at gene level. Enrichment p-values are in bold if $p < 0.05$.

When the EWIS was conducted using BMI change instead of BMI, NLRP3-IL1B-IL17

axis and global inflammation were enriched (**Table 5.3**). No enrichment was found after additional adjustment for physical activity. Global inflammation remained enriched after adjustment for neutrophil estimates while the NLRP3-IL1B-IL17 axis did not. **Tables 5.2 and 5.3** summarize the WKS enrichment test results.

Table 5.3. EWIS of DNA methylation and BMI change on adult-onset asthma: enrichment test results for 17 inflammation pathways and NLRP3-IL1B-IL17 axis

Pathway	#Genes	#CpGs	Enrichment p-value		
			Basic model	Adjusted for physical activity	Adjusted for neutrophil counts
Adhesion-extravasation-migration	142	1737	0.67	0.60	0.39
Apoptosis signaling	68	1210	0.50	0.37	0.22
Calcium signaling	14	413	0.29	0.34	0.21
Complement cascade	40	483	0.45	0.64	0.34
Cytokine signaling	172	1883	0.26	0.35	0.21
Eicosanoid signaling	39	450	0.48	0.17	0.61
Glucocorticoid/PPAR signaling	21	404	0.063	0.15	0.072
G-Protein coupled receptor signaling	42	1133	0.47	0.88	0.46
Innate pathogen detection	50	515	0.059	0.12	0.13
Leukocyte signaling	121	1429	0.35	0.49	0.34
MAPK signaling	118	2682	0.13	0.33	0.24
Natural killer cell signaling	31	368	0.91	0.75	0.91
NF-κB signaling	33	654	0.70	0.49	0.62
Phagocytosis-Ag presentation	39	1058	0.51	0.89	0.71
PI3K/AKT signaling	37	907	0.98	0.98	0.89
ROS/glutathione/cytotoxic granules	22	190	0.24	0.55	0.14
TNF superfamily signaling	38	537	0.085	0.33	0.065
Global inflammation[§]	1027	15985	**0.048**	0.23	**0.028**
NLRP3-IL1B-IL17 axis	11	219	**0.046**	0.13	0.15

The basic model regressed adult-onset asthma on BMI change, residuals of DNA methylation at a single CpG site, and their multiplicative interaction, upon adjustment for BMI at SAPALDIA2, age, sex, education level, study area, pack-years of cigarettes smoked in life, bench time, and white blood cell composition estimates for B cells, CD4 T cells, CD8 T cells, natural killer cells, monocytes, and eosinophils. [§]Total of the 17 inflammation pathways; the number of CpG in this pathway (15985) is smaller than the sum of the CpGs assigned to 17 pathways because there are CpGs assigned to multiple pathways, although the 17 pathways are mutually exclusive at gene level. Enrichment p-values are in bold if $p < 0.05$.

Our study is likely underpowered to identify differential methylation markers from the EWAS approach, considering the large dimension of the methylome data, the relatively small sample size, and the investigation of effect modification instead of main effects. Acknowledging these issues, we additionally searched for differentially methylated

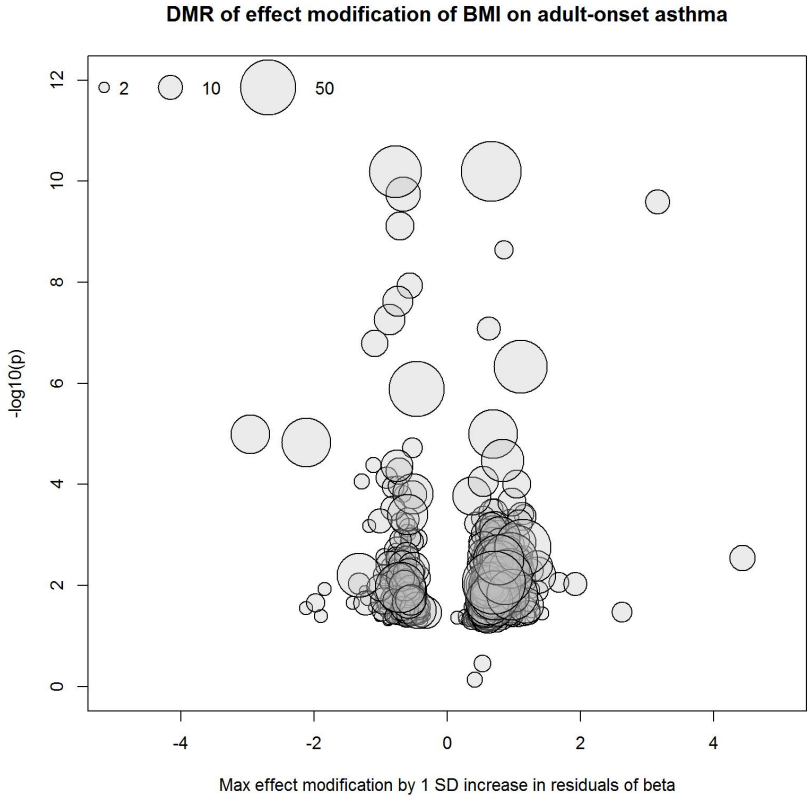
regions (DMR) using the R package DMRcate (Peters et al., 2015). Based on the EWIS, we identified 1131 DMRs that modify the association of BMI with non-atopic asthma as well as 18 DMRs that interact with BMI change affecting its association with non-atopic asthma. **Figures 5.3 and 5.4** summarize the DMRs. Each circle represents one DMR, whose x- and y-coordinates depict maximum effect modification by 1 SD increase in residuals within the region and minimum Benjamini-Hochberg adjusted p-value within the region, respectively. The 1131 and 18 DMRs were annotated to 1305 and 20 genes, respectively, and there were two overlapping genes. In an agnostic pathway enrichment analysis, using Ingenuity Pathway Analysis (IPA), we found PPARα/RXRα (retinoid X receptor alpha) activation ($p = 0.015$), ERK (extracellular-regulated kinase)/MAPK signaling ($p = 0.038$), and glucocorticoid receptor signaling ($p = 0.038$), among others, enriched for the 1305 genes annotated to the 1131 DMRs. **Figure 5.S5 and Table 5.S2** summarize the IPA pathway analysis results.

5.4 Discussion

We found no single CpG sites of genome-wide significant effect modification, however, we did find DMRs and pathway enrichments. DNA methylation markers usually act in concert at neighboring CpG sites (Hansen et al., 2011, Irizarry et al., 2009) and therefore EWIS alone may fail to identify true differential methylation markers (Li et al., 2015).

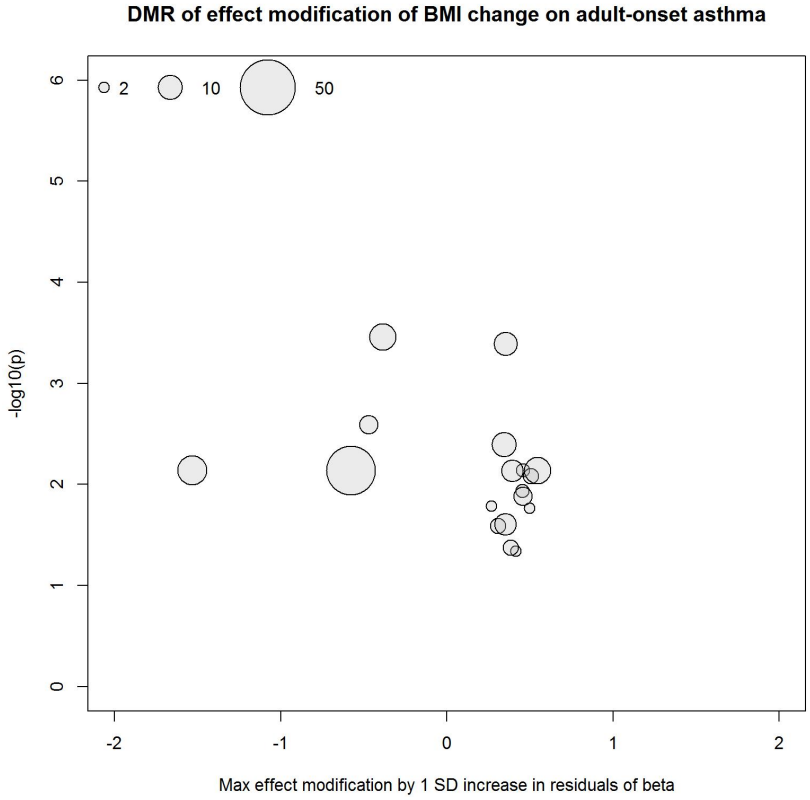
Global inflammation, defined as the entirety of the 1027 inflammation-related genes according to the classification proposed by Loza et al (Loza et al., 2007), was over-represented in the DNA methylation signals modifying the BMI–adult-onset asthma association. The agnostic search for the pathway enrichment of the DMRs also revealed several relevant pathways.Our study results are consistent with inflammation modify-

Figure 5.3. DMRs derived from the EWIS of DNA methylation and BMI on adult-onset asthma



Circle size represents the number of CpG sites in the region.

Figure 5.4. DMRs derived from the EWIS of DNA methylation and BMI change on adult-onset asthma



Circle size represents the number of CpG sites in the region.

ing the effect of BMI on adult-onset non-atopic asthma. Adiposity is believed to induce chronic systemic inflammation via dysregulated production of pro-inflammatory cytokines and immune cells infiltrated into adipose tissue (Galic et al., 2010, Weisberg et al., 2003). Our findings suggest that altered methylation in pro-inflammatory gene networks potentially mediate the link between overweight and non-atopic adult-onset asthma. This is in line with previous findings in children. Rastogi and her colleagues reported hypomethylation in the promoter of genes involved in innate immunity and non-atopic inflammation in obese children with asthma (Rastogi et al., 2013).

Among the 17 inflammation pathways curated by Loza and his colleagues (Loza et al., 2007), glucocorticoid/PPAR signaling showed the strongest enrichment. The agnostic search for the pathway enrichment in the DMRs also found enrichment of Glucocorticoid receptor signaling in addition to PPARα/RXRα activation. The glucocorticoid/PPAR signaling includes the genes coding for the nuclear receptors for glucocorticoids, PPARs, and associated proteins. PPARs have been associated with asthma and PPAR agonists are considered as a new asthma treatment (Banno et al., 2018). While we cannot rule out the possibility that asthma medication led to DNA methylation on the genes involved in glucocorticoid receptor signaling, it is not likely that this would explain the interaction of methylation signals in this pathway with BMI or BMI change.

MAPK signaling, NF-ϰB signaling, and PI3K/AKT signaling, are all involved in signal transduction downstream to the detection of insults e.g. by TLR (toll-like receptor). The enrichment signals for PI3K/AKT signaling disappeared when the model was additionally adjusted for physical activity, suggesting that the effect modification of BMI on adult-onset asthma might be confounded or mediated by physical activity. The enrichment of NF-ϰB signaling and PI3K/AKT signaling disappeared when the model was further adjusted for neutrophil estimates, suggesting that the effect modification in this pathway

might be modulated by neutrophil proliferation and hence disguised by the adjustment for neutrophil estimates. ERK/MAPK signaling was also over-represented in the DMRs from the agnostic pathway analysis.

Interestingly, NLRP3-IL1B-IL17 axis was enriched in the EWIS using BMI change, suggesting that BMI change represents a phenotype distinct from BMI. The well-known limitation of BMI is that it cannot distinguish fat from lean mass. Weight change in late adulthood is more likely attributable to change in fat than in lean mass (Newman et al., 2005) and people tend to lose lean mass while aging (Kim et al., 2017, Santanasto et al., 2017). Therefore, fat composition could be better reflected in BMI change than in BMI. Our finding that the enrichment of NLRP3-IL1B-IL17 axis disappeared upon adjustment for neutrophil estimates is consistent with the growing evidence of IL17 playing a role in recruitment, accumulation, and survival of neutrophils in asthma (Annunziato et al., 2015, Linden and Dahlen, 2014).

The NLRP3 inflammasome and downstream activity have been associated with both asthma (Esser et al., 2013) and obesity (Vandanmagsar et al., 2011) in humans. To the best of our knowledge, however, this study is the first to provide evidence of the NLRP3-IL1B-IL17 axis as a link between overweight and non-atopic adult-onset asthma in humans. This study is also the first to provide evidence that inflammation represented in the DNA methylation profile may play a role in the link between overweight and non-atopic adult-onset asthma.

Pathway enrichment analyses have often been applied to interpret genome-wide patterns of differential methylation. Widely used tools for pathway enrichment analyses include GSEA (gene set enrichment analysis) (Subramanian et al., 2005), DAVID (the database for annotation, visualization and integrated discovery) (Huang da et al., 2009), and IPA (http://www.ingenuity.com; QIAGEN, Redwood City, CA, USA), which were

originally developed to analyze differential expression of genes. In order to apply these tools to epigenetics, differential methylation signals first need to be translated from CpGs to genes. This can lead to a bias, e.g. that large genes with multiple CpGs are more likely to be represented. In this study, we applied the WKS method to test pathway enrichment of the EWIS results. The WKS method works in a similar way as GSEA but the enrichment is quantified by using CpG (and not gene) level statistics. This method also supports examination of custom-curated pathways, allowing straightforward interpretation.

We were underpowered to identify differential DNA methylation as effect modifiers, although the problem was partly overcome by applying integrative approaches, i.e. DMR and pathway enrichment analysis. The fact that DMR analysis resulted in more than 1000 signals while EWIS identified no signals consolidates that multivariate approach suits better than univariate approach to study epigenetic marks that function in clusters. We applied the WKS enrichment analyses to the absolute Z-statistics. Therefore, the direction of the effect modification, meaning whether hypo- or hyper-methylation was associated with increased effect of BMI, was not taken into consideration. The IPA pathway analysis results may have been biased by transforming the DMRs into gene-level statistics. Moreover, we annotated CpG sites simply based on the location, regardless of their functional information i.e. whether they resided in promoter, gene body, or intergenic region, CpG islands or not, etc. The cross-sectional design of our study, along with the effect modification being studied without mediation analysis, prevents inference of causal mediation. However, our findings of the enriched pathways using BMI change cannot be driven entirely by reverse causality, because BMI change preceded the DNA methylation measurements. Given the recent Mendelian randomization studies, reporting a causal effect of BMI on childhood asthma (Granell et al., 2014) and a causal effect of BMI on lifetime asthma (Skaaby et al., 2017), the overweight-asthma association may potentially be

causal. In this study we observed differential DNA methylation enriched in inflammatory pathways but did not measure chronic inflammation directly in the study subjects. However, we confirmed that the study subjects showed a strong positive association between BMI and hs-CRP at SAPALDIA2. In order to elucidate if overweight-induced inflammation causes asthma, further studies, including two-step Mendelian randomization studies, are warranted. Taking transcriptomics and proteomics study of blood, lung, and adipose tissue with asthma phenotype heterogeneity into consideration will be crucial.

5.5 Conclusion

DNA methylation measured in peripheral blood is consistent with inflammation as a potential link between BMI and adult-onset asthma, and to the NLRP3-IL1B-IL17 axis as a potential link between BMI change over 10 years and adult-onset asthma, in non-atopic non-smokers.

Acknowledgements

The **EXPOsOMICS Consortium** also includes (in alphabetic order): Andre F.S. Amaral, Toby Athersuch, Sabrina Bertinetti, Leda Chatzi, Theo De Kok, Almudena Espin Perez, Mario Fernandez, Claudia Galassi, Akram Ghantous, Hans Gmuender, John Henderson, Zdenko Herceg, Gerard Hoek, Medea Imboden, Pooja Jain, Debbie Jarvis, Maryam Karimi, Frank Kelly, Pekka Keski-Rahkonen, Jos Kleinjans, Julian Krauskopf, David Morley, Nahid Mostafavi Montazeri, Alessio Naccarati, Georgiadis Panos, Michelle Plusquin, George Preston, Andrea Ranzi, Stephen Rappaport, Laia Font Ribeira, Lorenzo Richiardi, Susan M. Ring, Oliver Robinson, Alberto Rodriguez, Augustin Scalbert, Terrence Simmons, Martyn T. Smith, Jordi Sunyer, Sonia Tarallo, Veronique Terrasse, Ming Tsai, Erik van Nunen, Karin van Veldhoven, Cristina M. Villanueva, Timo Wittenberger.

The **SAPALDIA consortium** includes:

Study directorate: NM Probst-Hensch (PI); D Stolz; C Schindler; N Künzli

Scientific team: JC Barthélémy, W Berger, R Bettschart, A Bircher, C Brombach, PO Bridevaux, L Burdet, Felber Dietrich D, T Sigrist, U Frey, MW Gerbase, D Gold, E de Groot, W Karrer, F Kronenberg, B Martin, A Mehta, D Miedinger, M Pons, F Roche, T Rothe, P Schmid-Grendelmeyer, P Gasche, A Schmidt-Trucksäss, J Schwartz, A Turk, A von Eckardstein, E Zemp Stutz. Scientific team at coordinating centers: M Adam, I Aguilera, A Beckmeyer-Borowko, S Brunner, D Carballo, S Caviezel, I Curjuric, A Di Pascale, J Dratva, R Ducret, E Dupuis Lozeron, M Eeftens, I Eze, E Fischer, M Foraster, M Germond, L Grize, S Hansen, A Hensel, M Imboden, A Ineichen, A Jeong, D Keidel, A Kumar, N Maire, A Mehta, R Meier, E Schaffner, T Schikowski, M Tsai

Administrative staff: N Bauer Ott, C Gabriel

The study could not have been done without the help of the study participants, technical and administrative support and the medical teams and field workers at the local study sites.

Local fieldworkers : Aarau: S Brun, G Giger, M Sperisen, M Stahel, Basel: C Bürli, C Dahler, N Oertli, I Harreh, F Karrer, G Novicic, N Wyttenbacher, Davos: A Saner, P Senn, R Winzeler, Geneva: F Bonfils, B Blicharz, C Landolt, J Rochat, Lugano: S Boccia, E Gehrig, MT Mandia, G Solari, B Viscardi, Montana: AP Bieri, C Darioly, M Maire, Payerne: F Ding, P Danieli A Vonnez, Wald: D Bodmer, E Hochstrasser, R Kunz, C Meier, J Rakic, U Schafroth, A Walder.

Funding

This work has been supported by the grant FP7 of the European Commission "Enhanced exposure 36 assessment and omic profiling for high priority environmental exposures in Europe" (EXPOsOMICS grant 308610 to P.V.).

The SAPALDIA study is supported by the Swiss National Science Foundation (grants no 33CS30-148470/1&2, 33CSCO-134276/1, 33CSCO-108796, 324730_135673, 3247BO-104283, 3247BO-104288, 3247BO-104284, 3247-065896, 3100-059302, 3200-052720, 3200-042532, 4026-028099, PMPDP3_129021/1, PMPDP3_141671/1), the Federal Office for the Environment, the Federal Office of Public Health, the Federal Office of Roads and Transport, the canton's government of Aargau, Basel-Stadt, Basel-Land, Geneva, Luzern, Ticino, Valais, and Zürich, the Swiss Lung League, the canton's Lung League of Basel Stadt/ Basel Landschaft, Geneva, Ticino, Valais, Graubünden and Zurich, Stiftung ehemals Bündner Heilstätten, SUVA, Freiwillige Akademische Gesellschaft, Klinik Barmelweid, Hirslanden Klinik Aarau, the European Commission (Grant 018996 GABRIEL to W. Cookson), and the Wellcome Trust WT 084703MA (W. Cookson).

Author contribution

A.J. and N.P.-H. designed the study; M.I., A.G., A.N., and C.C. analyzed the methylome. A.J. conducted the statistical analyses in consultation with C.S., G.L., A.G., and A.N.; A.J. wrote the manuscript; M.I., A.G., A.-E.C., C.S., P.V., and N.P.-H. revised the manuscript; M.K., G.L., Z.H., R.V., D.J., A.F.S.A., and F.K. helped development of the manuscript; All authors reviewed and approved the final manuscript.

Competing interests

The authors declare no conflict of interests.

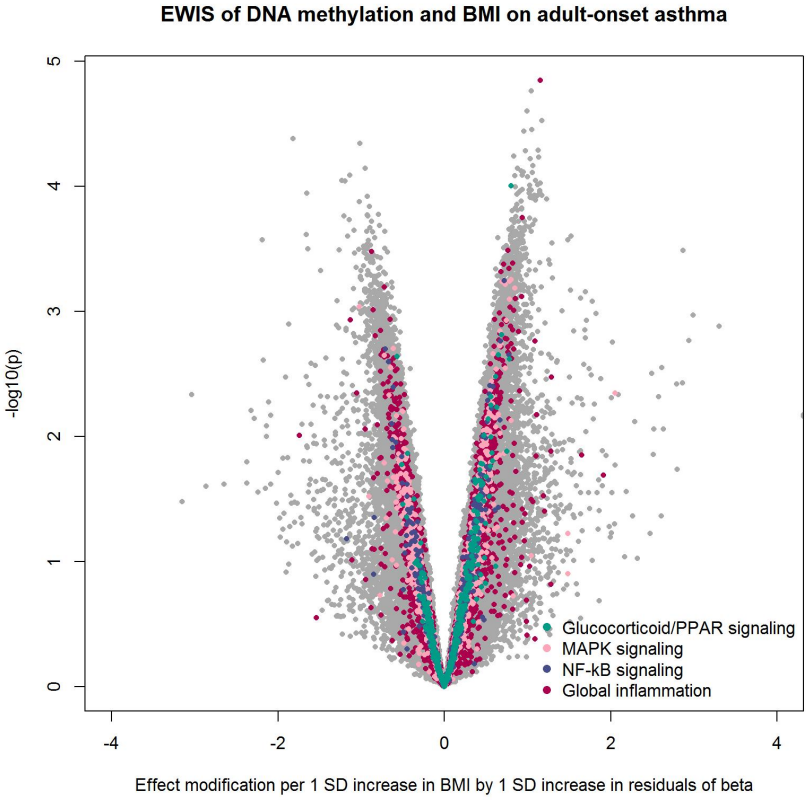
Data and materials availability

EXPOsOMICS data can be made available to external researchers on the basis of a written project that will be examined by the Steering Board. Interested colleagues are encouraged to contact P.V. Colleagues interested in data for this current study should contact N.P.-H., corresponding author of this paper and principle investigator of the SAPALDIA cohort.

5.6 Supplementary Material

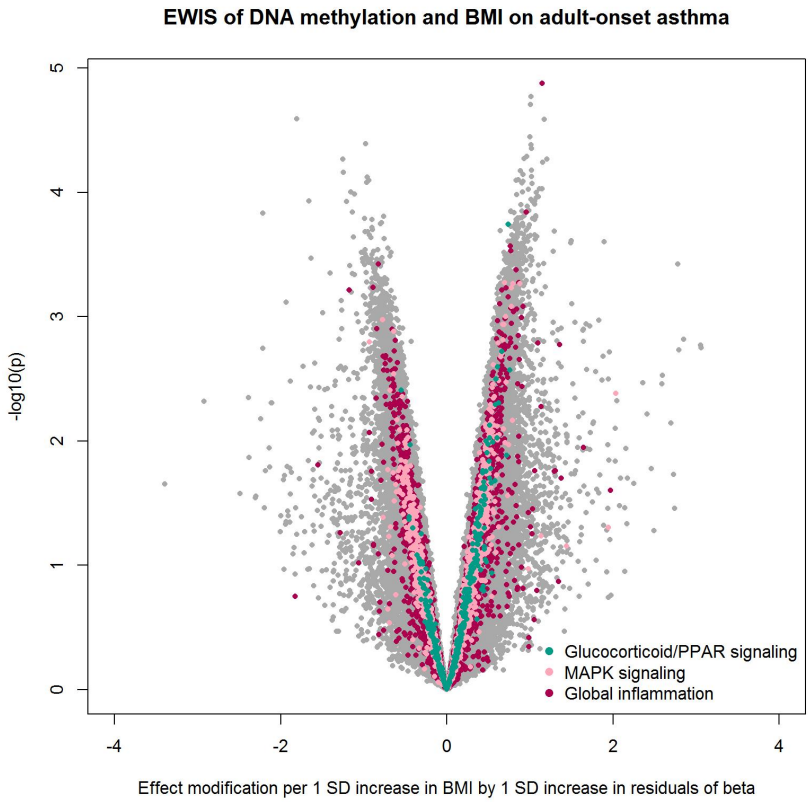
5.6.1 Supplementary figures

Figure 5.S1. Volcano plot from the EWIS of DNA methylation and BMI on adult-onset asthma, further adjusted for physical activity



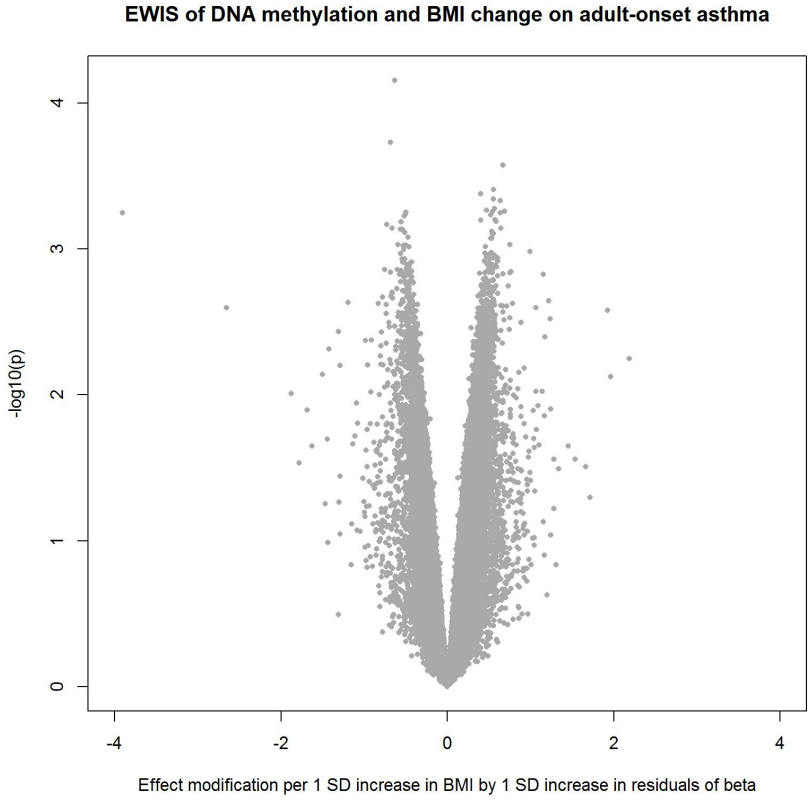
The EWIS fitted logistic regression models of adult-onset asthma on BMI, residuals of DNA methylation at a single CpG site, and their multiplicative interaction, upon adjustment for age, sex, education level, study area, pack-years of cigarettes smoked in life, physical activity, bench time, and white blood cell composition estimates for B cells, CD4 T cells, CD8 T cells, natural killer cells, monocytes, and eosinophils. The CpGs assigned to the pathway enriched with $p < 0.05$ are highlighted in colors.

Figure 5.S2. Volcano plot from the EWIS of DNA methylation and BMI on adult-onset asthma, further adjusted for neutrophil estimates



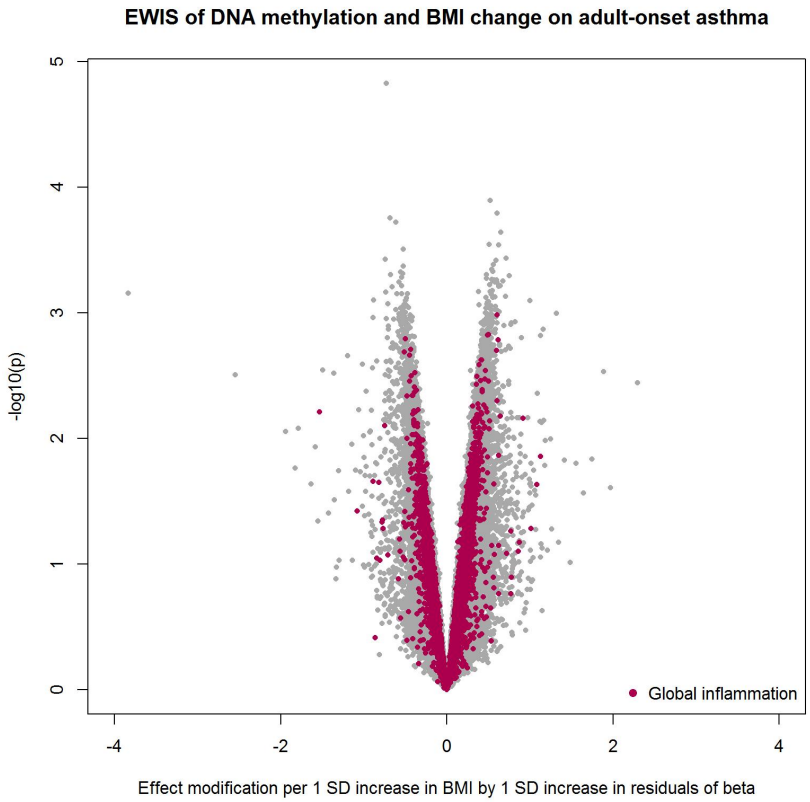
The EWIS fitted logistic regression models of adult-onset asthma on BMI, residuals of DNA methylation at a single CpG site, and their multiplicative interaction, upon adjustment for age, sex, education level, study area, pack-years of cigarettes smoked in life, bench time, and white blood cell composition estimates for B cells, CD4 T cells, CD8 T cells, natural killer cells, monocytes, eosinophils, and neutrophils. The CpGs assigned to the pathway enriched with $p < 0.05$ are highlighted in colors.

Figure 5.S3. Volcano plot from the EWIS of DNA methylation and BMI change on adult-onset asthma, further adjusted for physical activity



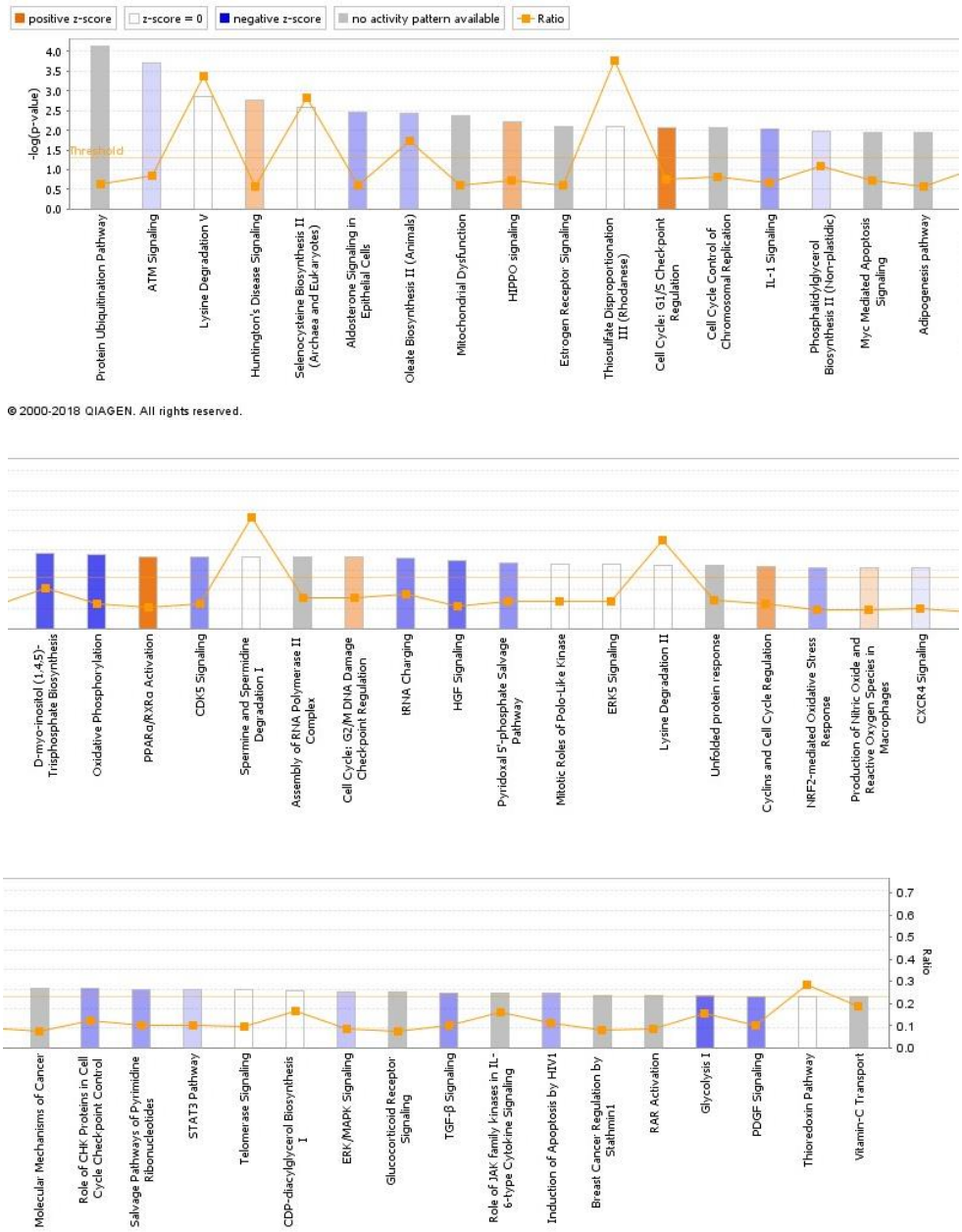
The EWIS fitted logistic regression models of adult-onset asthma on BMI change, residuals of DNA methylation at a single CpG site, and their multiplicative interaction, upon adjustment for age, sex, education level, study area, pack-years of cigarettes smoked in life, physical activity, bench time, and white blood cell composition estimates for B cells, CD4 T cells, CD8 T cells, natural killer cells, monocytes, and eosinophils.

Figure 5.S4. Volcano plot from the EWIS of DNA methylation and BMI change on adult-onset asthma, further adjusted for neutrophil estimates



The EWIS fitted logistic regression models of adult-onset asthma on BMI change, residuals of DNA methylation at a single CpG site, and their multiplicative interaction, upon adjustment for age, sex, education level, study area, pack-years of cigarettes smoked in life, bench time, and white blood cell composition estimates for B cells, CD4 T cells, CD8 T cells, natural killer cells, monocytes, eosinophils, and neutrophils. The CpGs assigned to the pathway enriched with $p < 0.05$ are highlighted in colors.

Figure 5.S5. Agnostic pathway enrichment results of the DMRs identified from the EWIS of DNA methylation and BMI on adult-onset asthma



5.6.2 Supplementary tables

Table 5.S1. Genes curated to 17 inflammatory pathways and NLRP3-IL1B-IL17 axis

Pathway	#Genes	Genes
Adhesion-extravasation-migration	142	PTPRU, VCAM1, CD58, CD2, MUC1, DARC, CD48, F11R, XCL1, SELP, SELL, SELE, RASSF5, CD34, MENA, ITGB1, CXCL12, VCL, ADAM8, ARHGAP1, CTNND1, SIPA1, CTTN, FUT4, MMP7, MMP10, MMP1, MMP12, BLR1, THY1, JAM3, CD9, ITGB7, ITGA5, MMP19, MYL6, SELPLG, PXN, MMP14, SPN, ITGAL, ITGAM, ITGAX, ITGAD, MLCK, MMP2, CCL22, CX3CL1, CCL17, CDH5, CKLF, CRK, ITGAE, CXCL16, MYH10, CCL2, CCL7, CCL11, CCL8, CCL13, CCL1, CCL5, CCL16, CCL15, CCL23, CCL18, CCL4, CCR7, ITGA2B, ITGB3, ITGA3, ICAM2, PECAM1, ITGB4, CD226, ICAM1, ICAM5, ICAM3, CD97, CEACAM5, CEACAM6, CEACAM3, CEACAM1, CEACAM8, PLAUR, VASP, CD33, ROCK2, ITGA6, ITGA4, ITGAV, ALS2, CCL20, SIGLEC1, MMP9, JAM2, ITGB2, MIF, MYH9, CCR4, CX3CR1, CCR8, CCBP2, CXCR6, XCR1, CCR1, CCR3, CCR2, CCR5, CCRL2, RHOA, ALCAM, CD47, CD96, RHOH, CXCL1, PPBP, CXCL5, CXCL3, CXCL2, CXCL9, CXCL10, CXCL11, CXCL13, FYB, CCL28, ITGA1, ITGA2, CXCL14, LECT2, HMMR, NT5E, VIL2, CCR6, CCL26, CCL24, PTP-PEST, CD36, CCL27, CCL19, CCL21, CD99
Apoptosis signaling	68	DFFB, DFFA, CASP9, MCL1, DAP3, LMNA, FASLG, CAPN2, PARP1, CDC2, FAS, CASP7, GAS2, BAD, CAPN1, FADD, BIRC3, BIRC2, CASP12, BCL2L14, CRADD, APAF1, DIABLO, ACIN1, BCL2L2, RIPK3, BCL2L10, BCL2A1, TP53, ROCK1, BCL2, CASP14, CAPNS1, BBC3, BAX, BCL2L12, HTRA2, BCL2L11, TANK, CFLAR, CASP10, CASP8, CAPN10, BCL2L1, BCL2L13, BID, A4GALT, BIK, FAIM, TNFSF10, PTPN13, CASP6, CASP3, DAP, DAXX, BAK1, TNFRSF21, CASP8AP2, CYCS, CASP2, TNFRSF10B, TNFRSF10C, TNFRSF10D, TNFRSF10A, RIPK2, DAPK1, SPTAN1, ENDOG
Calcium signaling	14	CAMK1D, PPP3CB, NFATC4, AKAP5, NFATC3, NFAT5, NFATC1, PPP3R1, NFATC2, CABIN1, PPP3CA, CAMK4, PPP3CC, PPP3R2
Complement cascade	40	MASP2, C1QA, C1QC, C1QB, C8A, C8B, SERPINC1, CFH, C4BPB, C4BPA, CD55, CR2, CR1, CD46, MBL2, CD59, SERPING1, C1S, C1R, C3AR1, SERPINA1, SERPINA5, SERPINF2, CFD, C3, CPAMD8, C5AR1, SERPIND1, MASP1, CFI, C9, C7, C6, C2, CFB, C4B, SERPINE1, C5, C8G, PFC
Cytokine signaling	172	IL22RA1, TXLNA, CSF3R, JAK1, IL23R, IL12RB2, CSF1, PIAS3, S100A9, S100A12, S100A8, ILF2, IL6R, CRP, IL10, IL19, IL20, IL24, IRF6, TGFB2, IL15RA, IL2RA, GATA3, BMPR1A, IRF7, CASP1, IL18, IL10RA, STAT2, STAT6, IFNG, IL26, IL22, SOCS2, IL31, IL17D, HMGB1, IL25, ISGF3G, BMP4, SOCS4, TGFB3, PIAS1, IL16, IL32, SOCS1, IL4R, IL21R, IRF8, IL17C, CSF3, STAT5B, STAT5A, STAT3, TBX21, ACE, SOCS3, TGIF, PTPN2, PIAS2, SOCS6, PIAS4, EBI3, C19ORF10, TYK2, IL27RA, JAK3, IL12RB1, IL28B, IL28A, IL29, TGFB1, IRF3, IL11, ADAM17, SOCS5, TGFA, IL1R2, IL1R1, IL1RL2, IL1RL1, IL18R1, IL18RAP, IL1A, IL1B, IL1F7, IL1F9, IL1F6, IL1F8, IL1F5, IL1F10, IL1RN, NMI, STAT1, STAT4, BMPR2, IL8RB, IL8RA, BMP2, TGIF2, CEBPB, PTPN1, IFNAR2, IL10RB, IFNAR1, IFNGR2, IL17RA, LIF, OSM, CSF2RB, IL2RB, PDGFB, IL5RA, TGFBR2, CISH, IL17RB, IL12A, THPO, IL1RAP, PDGFRA, IL8, BMPR1B, IL2, IL21, IL15, IRF2, IL7R, LIFR, OSMR, IL31RA, IL6ST, IL3, CSF2, IRF1, IL5, IL13, IL4, IL9, IL17B, CSF1R, PDGFRB, IL12B, IRF4, AGER, VEGF, IL17A, IL17F, IFNGR1, IL6, IRF5, IL7, JAK2, C9ORF26, IFNB1, IFNW1, IFNA21, IFNA4, IFNA10, IFNA5, IFNA6, IFNA2, IFNA8, IFNA1, IFNK, NFIL3, TGFBR1, ENG, IL3RA, CRSP2, IL2RG, IL13RA2, IL13RA1

Table 5.S1. (Cont.) Genes curated to 17 inflammatory pathways and NLRP3-IL1B-IL17 axis

Pathway	#Genes	Genes
Eicosanoid signaling	39	PLA2G2A, PLA2G2D, PTGER3, PTGFR, MGST3, PTGS2, AKR1C3, ALOX5, GPR44, PTGES3, LTA4H, ALOX5AP, CYSLTR2, PTGDR, PTGER2, DPEP3, DPEP2, DPEP1, ALOX15, ALOX12, ALOX15B, ALOX12B, TBXA2R, PTGER1, PTGIR, FPRL1, FPRL2, PTGIS, GGT1, PGDS, MGST2, PTGER4, LTC4S, TBXAS1, PTGS1, PTGES2, PTGES, PTGDS, CYSLTR1
Glucocorticoid/ PPAR signaling	21	GMEB1, FAF1, FKBP4, NR4A1, PPARBP, NCOA1, NR4A2, SDPR, HSPD1, GMEB2, NRIP1, PPARA, KPNA1, PPARGC1A , NR2F1, NR3C1, FOXO3A, SGK, CITED2, GLCCI1, NCOA2
G-Protein coupled receptor signaling	42	PDE4B, PRKACB, ADORA3, RGS1, ADORA1, CREM, ADRA2A, PDE3B, PLCB3, ADRBK1, PDE2A, PDE3A, ADCY6, PDE1B, ADCY4, PLCB2, CREBBP, ADORA2B, PRKAR1A, PDE4A, PRKACA, PDE4C, PDE1A, CREB1, PLCB1, PLCB4, ADORA2A, HRH1, PRKAR2A, ADCY5, GRK4, ADRA2C, ADCY2, PDE4D, ADRB2, HRH2, SYNGAP1, PDE1C, ADCY1, PRKAR2B, PTK2B, PRKACG
Innate pathogen detection	50	PGLYRP3, PGLYRP4, CIAS1, NALP6, TOLLIP, CASP5, COP1, ICEBERG, TIRAP, IRAK4, IRAK3, HSP90B1, OAS1, OAS2, PYCARD, CARD15, NALP1, SARM1, TICAM1, SITPEC, PGLYRP2, PGLYRP1, CARD8, NALP12, NALP2, CARD12, IFIH1, PRKRA, KIAA1271, LBP, IRAK2, MYD88, TLR9, TLR10, TLR1, TLR6, TLR2, TLR3, TLR5, CARD6, CD180, TICAM2, CD14, CARD4, LY96, DDX58, TLR4, TLR7, TLR8, IRAK1
Leukocyte signaling	121	CD52, PTAFR, LCK, CD53, PTPN22, IGSF3, IGSF2, VTCN1, FCGR1C, FCER1A, SLAMF9, SLAMF6, SLAMF1, SLAMF7, FCER1G, FCGR2A, FCGR2B, SH2D1B, CD3Z, PTPRC, PTPN7, PIGR, TCF8, BLNK, PIK3AP1, CD44, RAG1, PTPRJ, MS4A2, MS4A1, CD5, SCGB1A1, SLC3A2, TCIRG1, CD3E, CBL, CD4, LRRC23, KITLG, FLT3, EDNRB, IGHA1, CSK, LAT, SLC7A5, SCARF1, CD68, CD79B, GRB2, SECTM1, VAV1, FCER2, CD22, CD79A, CD37, FLT3LG, SIGLEC10, SIGLEC5, LILRB3, LILRA6, LILRB2, LILRA3, LILRA5, LILRA4, LILRA2, LILRA1, LILRB4, LILRP2, FCAR, TACR1, CD8A, CD8B1, MAL, ZAP70, MARCO, DPP4, CD28, CTLA4, ICOS, SIRPB1, SIRPG, PTPNS1, CST7, SLA2, CD40, ICOSLG, SCARF2, VPREB1, IGLL1, GRAP2, MST1R, CD80, CD86, SH3BP2, CD38, TXK, TEC, DAPP1, EDNRA, TCF7, ITK, LCP2, SCGB3A1, TREM2, TREM1, TFEB, FYN, TRGV9, PILRB, PBEF1, MSR1, PAG1, PDCD1LG2, PAX5, SHB, SEMA4D, SYK, ABL1, FOXP3, BTK, CD40LG
MAPK signaling	118	PRKCZ, RAP1GAP, RPS6KA1, MAP3K6, HDAC1, MKNK1, JUN, RAP1A, NRAS, SHC1, MEF2D, IFI16, PLA2G4A, MAPKAPK2, ATF3, DUSP10, PRKCQ, MAPK8, HRAS, INS, MADD, RPS6KA4, FOSL1, PPP1CA, PAK1, PPP2R1B, HSPB2, ETS1, KRAS, RAPGEF3, ATF1, DDIT3, DUSP6, PPP1CC, MAPKAPK5, SOS2, PPM1A, MAX, FOS, RPS6KA5, RASGRP1, TLN2, MAP2K1, MEF2A, EEF2K, PRKCB1, MAPK3, BCAR1, MAP2K4, MAP2K3, KSR1, PRKCA, MAP2K6, MAP2K2, JUND, MEF2B, MAP4K1, PPP2R1A, YWHAQ, MYCN, PPP1CB, SOS1, PRKCE, DUSP2, ATF2, PPP1R7, SRC, PLCG1, YWHAB, ETS2, HMGN1, MAPK1, YWHAH, RAC2, MAP3K7IP1, ATF4, EP300, MAPK12, MAPK11, PPARG, RAF1, KCNH8, MAPKAPK3, PRKCD, PPP2R3A, EGF, MAP3K1, RASA1, MEF2C, HINT1, PPP2CA, PPP2R2B, DUSP1, MAPK9, MAPK14, MAPK13, MAP3K7, HDAC2, PTPRK, MAP3K5, MAP3K7IP2, ESR1, RAC1, EGFR, HSPB1, YWHAG, BRAF, DUSP4, PPP2CB, LYN, YWHAZ, MYC, PTK2, TLN1, RAPGEF1, ARAF, ELK1, DUSP9

Table 5.S1. (Cont.) Genes curated to 17 inflammatory pathways and NLRP3-IL1B-IL17 axis

Pathway	#Genes	Genes
Natural killer cell signaling	31	CD160, CD244, FCGR3A, NCAM1, B3GAT1, PTPN6, KLRB1, KLRD1, KLRK1, KLRC4, KLRC2, KLRC1, PTPN11, CD300A, TYROBP, SIGLEC7, LAIR1, LILRB1, KIR3DL3, KIR2DS4, KIR3DL2, NCR1, HLA-G, HLA-E, MICA, MICB, NCR3, NCR2, RAET1E , ULBP3, SH2D1A
NF-κB signaling	33	BCL10, CHUK, BTRC, NFKB2, RELA, NFRKB, TBK1, UBE2N, NFKBIA, MEFV, CSNK2A2, MAP3K14, MAP3K3, CARD14, MALT1, MAP2K7, NFKBIB, BCL3, RELB, EIF2AK2, REL, CSNK2A1, UBE2V1, CARD10, BCL6, NFKB1, RIPK1, CSNK2B, NFKBIE, CARD11, IKBKB, CARD9, IKBKG
Phagocytosis-Ag presentation	39	CTSS, CD1D, CD1A, CD1C, CD1B, CD1E, PSMA1, LAG3, RFX4, PSMB5, PSME1, PSME2, CIITA, CD209, RFX1, IFI30, RFXANK, LILRB5, CD207, XBP1, CD74, PRSS16, HLA-A, HLA-C, HLA-B, HLA-DRA, HLA-DQA1, HLA-DQA2, HLA-DQB2, TAP2, TAP1, PSMB9, HLA-DMB, HLA-DMA, HLA-DOA, HLA-DPA1, HLA-DPB1, TAPBP, NFX1
PI3K/AKT signaling	37	PIK3CD, FRAP1, PIK3R3, THEM4, AKT3, MAP3K8, PTEN, ILK, RPS6KB2, CCND1, INPPL1, CDKN1B, MDM2, FOXO1A, HSP90AA1, AKT1, TSC2, PDPK1, PIK3R5, MYH4, RPS6KB1, CDC37, PIK3R2, AKT2, LIMS1, INPP5D, CTNNB1, PIK3CB, PIK3CA, EIF4E, GAB1, PIK3R1, CDKN1A, HSP90AB1, NOS3, RHEB, TSC1
ROS/glutathione/ cytotoxic granules	22	PRDX1, NCF2, PRF1, CAT, PRG2, CMA1, GZMH, GZMB, GPX2, ANPEP, NOS2A, GZMM, PRTN3, ELA2, PRDX2, BPI, SOD1, GZMA, GPX3, SOD2, PRDX4, CYBB
TNF superfamily signaling	38	TNFRSF4, TNFRSF14, TNFRSF25, TNFRSF9, TNFRSF1B, TNFSF18, TNFSF4, TRAF5, TRAF6, TNFRSF1A, LTBR, TNFRSF7, TNFRSF19, TNFSF11, TNFSF13B, TRAF3, TNFRSF12A, TNFRSF17, TRADD, TNFSF12, TNFRSF13B, TNFRSF11A, TNFSF9, TNFSF7, TNFSF14, TNFAIP6, TNFRSF13C, PTX3, TNIP1, TTRAP, LTA, TNF, TNFAIP3, TNFRSF11B, TNFSF15, TNFSF8, TRAF1, TRAF2
NLRP3-IL1B-IL17 axis	11	NLRP3, PYCARD, CASP1, IL1B, IL1R1, RORC, CCR6, ATXN1, THY1, CD44, IL17A

Table 5.S2. Agnostic pathway enrichment results of the DMRs identified from the EWIS of DNA methylation and BMI on adult-onset asthma

Pathway	P-value	Ratio[a]	Genes[b]
Protein Ubiquitination Pathway	0.000076	0.11	USP35, PSMA7, UBR2, HSPA1A/HSPA1B, CDC23, UBE2W, SKP1, UBE2O, HSPA1L, HSPA4, USP7, UBE2B, STUB1, HSPE1, UCHL5, DNAJC30, BIRC3, UBE2Q1, UBE2M, USP19, DNAJC1, HSPD1, UBE3A, UBE2G2, CBL, HSCB, PSMA5, DNAJC18, PSMD1, HSPB1
ATM Signaling	0.00019	0.15	MAP2K4, PPP2R2A, TRIM28, MAPK9, TDP1, KAT5, PPM1D, USP7, PPP2R1A, BRAT1, RAD17, H2AFX, SMC1B, PPP2R5C, CHEK2
Lysine Degradation V	0.0014	0.60	AASDHPPT, PIPOX, ALDH7A1
Huntington's Disease Signaling	0.0018	0.10	MAP2K4, SGK1, PACSIN1, HSPA1A/HSPA1B, HDAC10, HSPA1L, PRKCZ, HSPA4, CDK5, NTRK1, HDAC7, NCOR1, GNB1L, BET1L, NAPB, MAP2K7, HDAC1, APAF1, MAPK9, HIP1, ATP5F1C, CAPNS1, TAF4, IRS1, NCOR2
Selenocysteine Biosynthesis II (Archaea and Eukaryotes)	0.0026	0.50	SEPHS1, SARS2, SEPSECS
Aldosterone Signaling in Epithelial Cells	0.0035	0.11	SGK1, HSPA1A/HSPA1B, PLCG1, DNAJC1, HSPD1, SLC9A1, PRKCZ, HSPA1L, HSPA4, PIP5K1A, DUSP1, HSCB, IRS1, HSPE1, DNAJC18, DNAJC30, PI4KA, HSPB1
Oleate Biosynthesis II (Animals)	0.0038	0.31	SCD, UFSP2, FADS2, FADS1
Mitochondrial Dysfunction	0.0043	0.11	MAP2K4, NDUFV1, COX4I2, NDUFS7, ACO2, MAPK9, ATP5MG, NDUFB1, NDUFA13, UQCRB, VPS9D1, ATP5F1C, APH1A, NDUFA6, NDUFS6, ATP5MF, GPX4, TXNRD2
HIPPO signaling	0.0062	0.13	YWHAQ, PPP2R1A, YWHAH, PPP2R2A, PPP1R7, SMAD3, NF2, PPP1R14A, PPP2R5C, SKP1, PRKCZ
Estrogen Receptor Signaling	0.0079	0.11	TAF9, MED23, GTF2F2, TAF5L, ERCC2, GTF2A1, TAF4, MED15, ERCC3, MED21, SPEN, TAF3, NCOR1, NCOR2
Thiosulfate Disproportionation III (Rhodanese)	0.0081	0.67	MPST, TST
Cell Cycle: G1/S Checkpoint Regulation	0.0085	0.13	MYC, E2F4, MAX, TGFB1, SMAD3, HDAC7, HDAC1, HDAC10, SKP1
Cell Cycle Control of Chromosomal Replication	0.0089	0.14	MCM5, CDK13, CDK5, CDK11B, ORC6, DNA2, POLA2, CHEK2
IL-1 Signaling	0.0093	0.12	ECSIT, MAP2K4, ADCY9, MAP2K7, TOLLIP, GNAS, GNA12, MAP3K7, MAPK9, GNB1L, ADCY7
Phosphatidylglycerol Biosynthesis II (Non-plastidic)	0.011	0.19	AGPAT5, AGPAT1, MBOAT1, PGS1, MBOAT7
Myc Mediated Apoptosis Signaling	0.011	0.13	FADD, MAP2K4, MYC, YWHAQ, YWHAH, IRS1, APAF1, MAPK9, PRKCZ
Adipogenesis pathway	0.012	0.10	SAP18, LEP, SMAD3, HDAC1, HDAC10, ERCC2, CDK5, TGFB1, ERCC3, HDAC7, CLOCK, FGFRL1, TBL1XR1, FZD7
D-myo-inositol (1,4,5)-Trisphosphate Biosynthesis	0.013	0.19	PIP5K1A, PI4K2A, PLCG1, PI4K2B, PI4KA
Oxidative Phosphorylation	0.013	0.11	VPS9D1, ATP5F1C, NDUFV1, COX4I2, NDUFS7, NDUFA6, NDUFS6, ATP5MF, ATP5MG, NDUFB1, NDUFA13, UQCRB

[a]Ratio of the number of genes in the DMR to the number of genes in the pathway. [b]Genes overlapping between the DMR and the pathway.

Table 5.S2. (Cont.) Agnostic pathway enrichment results of the DMRs identified from the EWAS of effect modification of BMI on adult-onset asthma

Pathway	P-value	Ratio[a]	Genes[b]
PPARα/RXRαActivation	0.015	0.094	MAP2K4, MAP2K7, MED23, GNAS, SMAD3, PLCG1, AIP, ADCY9, TGFB1, IRS1, MAP3K7, CLOCK, NCOR1, SLC27A1, NCOR2, ADCY7, ACVR2A
CDK5 Signaling	0.015	0.11	ADCY9, PPP2R1A, GNAS, CDK5, PPP2R2A, PPP1R7, EGR1, MAPK9, PPP1R14A, PPP2R5C, ADCY7
Spermine and Spermidine Degradation I	0.016	0.50	PAOX, SAT2
Assembly of RNA Polymerase II Complex	0.016	0.14	TAF9, TAF4, ERCC3, TAF5L, TAF3, GTF2A1, ERCC2
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	0.016	0.14	YWHAQ, YWHAH, BORA, SKP1, PRKCZ, CHEK2, PPM1D
tRNA Charging	0.016	0.15	CARS2, HARS, EARS2, HARS2, SARS2, QARS
HGF Signaling	0.019	0.10	MAP2K4, ELF2, PXN, MAP2K7, IRS1, MAP3K7, MAPK9, PLCG1, MAP3K8, STAT3, ELK3, PRKCZ
Pyridoxal 5'-phosphate Salvage Pathway	0.021	0.12	MAP2K4, PNPO, CDK5, SGK1, MAPK9, MAP3K8, HIPK1, ACVR2A
Mitotic Roles of Polo-Like Kinase	0.023	0.12	PLK4, PPP2R1A, PPP2R2A, TGFB1, FBXO5, CDC23, PPP2R5C, CHEK2
ERK5 Signaling	0.023	0.12	MYC, YWHAQ, YWHAH, SGK1, GNA12, NTRK1, MAP3K8, PRKCZ
Lysine Degradation II	0.025	0.40	AASDHPPT, ALDH7A1
Unfolded protein response	0.026	0.13	HSPA4, MAP2K7, SREBF2, HSPA1A/HSPA1B, CANX, NFE2L2, HSPA1L
Cyclins and Cell Cycle Regulation	0.028	0.11	PPP2R1A, E2F4, PPP2R2A, TGFB1, HDAC7, HDAC1, HDAC10, PPP2R5C, SKP1
NRF2-mediated Oxidative Stress Response	0.028	0.088	MAP2K4, MAP2K7, MAPK9, DNAJC1, MAFK, PRKCZ, TXNRD1, BACH1, FTL, KEAP1, IRS1, STIP1, MAP3K7, DNAJC18, GSTO2, FKBP5, NFE2L2
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	0.030	0.088	MAP2K4, MAP2K7, APOB, PPP2R2A, MAPK9, PLCG1, PPP1R14A, SPI1, PRKCZ, PON1, RHOV, PPP2R1A, PPP1R7, IRS1, MAP3K7, PPP2R5C, MAP3K8
CXCR4 Signaling	0.030	0.091	MAP2K4, PXN, GNAS, MYL2, GNA12, EGR1, MAPK9, PRKCZ, ADCY9, ELMO3, RHOV, IRS1, GNB1L, ELMO1, ADCY7
Molecular Mechanisms of Cancer	0.030	0.076	MAP2K4, SMAD3, CTNNA1, PRKCZ, MYC, SYNGAP1, CDK5, TGFB1, MAP3K7, BIRC3, CHEK2, CDK13, E2F4, GNAS, GNA12, BMP8B, APAF1, MAPK9, ARHGEF17, FADD, ADCY9, RHOV, CBL, MAX, CDK11B, APH1A, IRS1, ADCY7, WNT1, FZD7
Role of CHK Proteins in Cell Cycle Checkpoint Control	0.030	0.12	PPP2R1A, E2F4, PPP2R2A, RAD17, PPP2R5C, RFC5, CHEK2
Salvage Pathways of Pyrimidine Ribonucleotides	0.032	0.10	MAP2K4, CDK5, SGK1, MAPK9, AK4, UCK1, CMPK1, MAP3K8, HIPK1, ACVR2A
STAT3 Pathway	0.032	0.10	MAP2K4, MYC, SOCS1, TGFB1, NTRK1, MAPK9, FGFRL1, IL27RA, STAT3, NDUFA13
Telomerase Signaling	0.033	0.099	MYC, ELF2, PPP2R1A, PPP2R2A, IRS1, HDAC7, HDAC1, TERT, HDAC10, PPP2R5C, ELK3
CDP-diacylglycerol Biosynthesis I	0.035	0.17	AGPAT5, AGPAT1, MBOAT1, MBOAT7

[a]Ratio of the number of genes in the DMR to the number of genes in the pathway. [b]Genes overlapping between the DMR and the pathway.

Table 5.S2. (Cont.) Agnostic pathway enrichment results of the DMRs identified from the EWAS of effect modification of BMI on adult-onset asthma

Pathway	P-value	Ratio[a]	Genes[b]
ERK/MAPK Signaling	0.037	0.085	PXN, YWHAH, PPP2R2A, SRF, PLCG1, PPP1R14A, STAT3, YWHAQ, MYC, PPP2R1A, ELF2, DUSP1, PPP1R7, IRS1, PPP2R5C, ELK3, HSPB1
Glucocorticoid Receptor Signaling	0.038	0.077	CD247, MAP2K4, YWHAH, GTF2F2, SGK1, SMAD3, HSPA1A/HSPA1B, GTF2A1, HSPA1L, HSPA4, TGFB1, MAP3K7, TAF3, NCOR1, FKBP5, TAF9, MAP2K7, TAF5L, MAPK9, STAT3, ERCC2, TAF4, DUSP1, IRS1, ERCC3, NCOR2
TGF-αSignaling	0.041	0.10	MAP2K4, TGFB1, RNF111, SMAD3, HDAC1, SKI, MAP3K7, MAPK9, ACVR2A
Role of JAK family kinases in IL-6-type Cytokine Signaling	0.041	0.16	MAP2K4, SOCS1, MAPK9, STAT3
Induction of Apoptosis by HIV1	0.042	0.11	FADD, MAP2K4, MAP2K7, APAF1, MAPK9, SLC25A3, BIRC3
Breast Cancer Regulation by Stathmin1	0.046	0.083	E2F4, GNAS, CAMK4, PPP2R2A, PPP1R14A, ARHGEF17, TUBB, PRKCZ, ADCY9, PPP2R1A, PPP1R7, IRS1, UHMK1, TUBA1C, PPP2R5C, GNB1L, ADCY7
RAR Activation	0.046	0.084	MAP2K4, SMAD3, MAPK9, ERCC2, PRKCZ, PTEN, ADCY9, TAF4, TGFB1, DUSP1, ERCC3, NCOR1, NCOR2, CSNK2B, CRABP2, ADCY7
Glycolysis I	0.047	0.15	ENO1, PGAM1, PKM, ALDOA
PDGF Signaling	0.049	0.10	MAP2K4, MYC, PDGFA, IRS1, SPHK2, SRF, PLCG1, CSNK2B, STAT3
Thioredoxin Pathway	0.049	0.29	TXNRD2, TXNRD1

[a]Ratio of the number of genes in the DMR to the number of genes in the pathway. [b]Genes overlapping between the DMR and the pathway.

Table 5.S3. Sex difference in BMI effect on adult-onset asthma

		#Cases	#Controls	Coefficient for BMI [95% CI]	Coefficient for BMI:female [95% CI]
All subjects[a]		61	146	0.44 [0.09, 0.79]	-
Stratification analysis[b]	Female	43	82	0.39 [−0.05, 0.83]	-
	Male	18	64	0.51 [−0.13, 1.15]	-
Interaction analysis[c]		61	146	0.70 [0.06, 1.35]	−0.37 [−1.13, 0.38]

All subjects are non-smoking for at least 10 years and non-atopic. [a]Logistic regression of adult-onset asthma on BMI adjusted for sex, age, education level, study area, and pack-years of cigarettes smoked in life. [b]Logistic regression of adult-onset asthma on BMI adjusted for age, education level, study area, and pack-years of cigarettes smoked in life, stratified by sex. [c]Logistic regression of adult-onset asthma on BMI, sex, and the interaction between BMI and sex, after adjustment for age, education level, study area, and pack-years of cigarettes smoked in life.

6. Article III

Perturbation of metabolic pathways mediates the association of air pollutants with asthma and cardiovascular diseases

Ayoung Jeong[a,b,*], Giovanni Fiorito[c,d,*], Pekka Keski-Rahkonen[e], Medea Imboden[a,b], Agneta Kiss[e], Nivonirina Robinot[e], Hans Gmuender[f], Jelle Vlaanderen[g], Roel Vermeulen[g], Soterios Kyrtopoulos[h], Zdenko Herceg[e], Akram Ghantous[e], Gianfranco Lovison[i], Claudia Galassi[j], Andrea Ranzi[k], Vittorio Krogh[l], Sara Grioni[l], Claudia Agnoli[l], Carlotta Sacerdote[m], Nahid Mostafavi[g], Alessio Naccarati[c], Augustin Scalbert[e], Paolo Vineis[c,n,**,1], Nicole Probst-Hensch[a,b,**,1] for the Exposomics consortium[2]

Environment International 119 (2018): 334-345

*equal first authorship; **equal senior authorship

[1]Correspondence to: Nicole Probst-Hensch, Swiss Tropical and Public Health Institute, Socinstrasse 57, 4051 Basel, Switzerland; Tel: +41 (0)61 284 8378; Mobile: +41 (0)79 280 3414; Email: nicole.probst@unibas.ch and to Paolo Vineis, Imperial College

[a]Swiss Tropical and Public Health Institute, Basel, Switzerland

[b]University of Basel, Basel, Switzerland

[c]Italian Institute for Genomic Medicine (IIGM) Turin, Italy

[d]Department of Medical Sciences – University of Turin, Italy

[e]International Agency for Research on Cancer, Lyon, France

[f]Genedata AG, Basel, Switzerland

[g]Utrecht University, Institute for Risk Assessment Sciences, Environmental Epidemiology Division, Utrecht, Netherlands

[h]National Hellenic Research Foundation, Athens, Greece

[i]University of Palermo, Palermo, Italy

[j]Unit of Cancer Epidemiology, Città della Salute e della Scienza University-Hospital and Center for Cancer Prevention (CPO), Turin, Italy

[k]Environmental Health Reference Center, Regional Agency for Prevention, Environment and Energy of Emilia-Romagna, Modena, Italy

[l]Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

[m]Piedmont Reference Center for Epidemiology and Cancer Prevention (CPO Piemonte), Via Santena 7, 10126 Turin, Italy

[n]MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, UK

London, St Mary's Campus, Norfolk Place W2 1PG, London, UK; Tel: +44 (0)20 7594 3372; Mobile: +44 (0)7478 544 464; p.vineis@imperial.ac.uk

[2]The EXPOsOMICS Consortium includes (in alphabetic order): Andre F.S. Amaral, Toby Athersuch, Sabrina Bertinetti, Marc Chadeau-Hyam, Theo Dekok, Michaela Dijmarescu, Almudena Espin Perez, Mario Fernandez, Claudia Galassi, Akram Ghantous, Hans Gmuender, John Gulliver, John Henderson, Zdenko Herceg, Gerard Hoek, Medea Imboden, Pooja Jain, Debbie Jarvis, Frank Kelly, Pekka Keski-Rahkonen, Jos Kleinjans, Manolis Kogevinas, Julian Krauskopf, Soterios Kyrtopoulos, David Morley, Nahid Mostafavi Montazeri, Alessio Naccarati, Tim Nawrot, Mark Nieuwenhuijsen, Georgiadis Panos, David Phillips, Michelle Plusquin, George Preston, Nicole Probst-Hensch, Andrea Ranzi, Stephen Rappaport, Laia Font Ribeira, Lorenzo Richiardi, Susan M. Ring, Oliver Robinson, Albert Rodriguez, Augustin Scalbert, Terrence Simmons, Martyn T. Smith, Sonia Tarallo, Veronique Terrasse, Ming Tsai, Erik van Nunen, Catharina van Veldhoven, Roel C.H. Vermeulen, Cristina M. Villanueva, Paolo Vineis, Jelle Vlaanderen, Christopher P. Wild, Timo Wittenberger.

Abstract

Background: Epidemiologic evidence indicates common risk factors, including air pollution exposure, for respiratory and cardiovascular diseases, suggesting the involvement of common altered molecular pathways.

Objectives: The goal was to find intermediate metabolites or metabolic pathways that could be associated with both air pollutants and health outcomes ("meeting-in-the-middle"), thus shedding light on mechanisms and reinforcing causality.

Methods: We applied a statistical approach named "meet-in-the-middle" to untargeted metabolomics in two independent case-control studies nested in cohorts on adult-onset asthma (AOA) and cardio-cerebrovascular diseases (CCVD). We compared the results to identify both common and disease-specific altered metabolic pathways.

Results: A novel finding was a strong association of AOA with ultrafine particles (UFP; odds ratio 1.80 [1.26, 2.55] per increase by 5,000 particles/cm^3). Further, we have identified several metabolic pathways that potentially mediate the effect of air pollution on health outcomes. Among those, perturbation of Linoleate metabolism pathway was associated with air pollution exposure, AOA and CCVD.

Conclusions: Our results suggest common pathway perturbations may occur as a consequence of chronic exposure to air pollution leading to increased risk for both AOA and CCVD.

6.1 Introduction

Asthmatics often suffer from comorbidities including cardiovascular diseases. Comorbidity influences the disease prognosis and control. Refractory asthma is more likely to manifest with cardiovascular comorbidity than controlled asthma (Hekking et al., 2017). Asthma and cardiovascular disease share common risk factors such as smoking, obesity, aging and air pollution exposure, consistent with common molecular pathways altered in the etiology of diseases.

Short-term effects of air pollution exposure on asthma exacerbation have long been established in adults and in children (Peel et al., 2005; Schwartz et al., 1993; Sunyer et al., 1997). The role of air pollution in asthma onset is less conclusive, particularly in adults (Anderson et al., 2013; Jacquemin et al., 2012). Only a few studies used individually assigned exposure estimates to study the effects of ambient air pollution on adult-onset asthma. The largest study sample was based on over 600,000 subjects, including 27,000 asthmatics, and demonstrated an association of PM_{10} exposure – derived from a pan-European land use regression model – with asthma prevalence (Cai et al., 2017). The European Study of Cohorts for Air Pollution Effects (ESCAPE) reported a positive but not statistically significant association with asthma incidence in adults for all air pollution metrics (NO_2, NO, PM_{10}, $PM_{2.5}$, traffic load; traffic intensity) except PM_{coarse} (Jacquemin et al., 2015). In the Swiss SAPALDIA cohort, long term improvement in air pollution levels was associated with an attenuated age-related lung function decline (Downs et al., 2007), with a decreased prevalence of respiratory symptoms including wheezing and breathlessness (Schindler et al., 2009), and with a decreased onset of asthma in adults (Kunzli et al., 2009).

In addition, a growing number of epidemiological studies showed that air pollution is associated with coronary artery disease (McGuinn et al., 2016; Wolf et al., 2015),

cardiovascular diseases (Brook et al., 2010; Franklin et al., 2015), and cerebrovascular diseases (Stafoggia et al., 2014) including ischemic stroke (Chung et al., 2017; Cox 2017). A recent meta-analysis within ESCAPE showed that increases in $PM_{2.5}$ and PM_{10} were associated with risks of fatal and total coronary events, respectively (Cesaroni et al., 2014), and increased risk for cerebrovascular diseases was reported for higher exposure to $PM_{2.5}$ and NO_2 (Stafoggia et al., 2014).

Ultrafine particles (UFP) exposure has been less studied than exposure to larger particles, and no regulatory agencies have established guidelines for UFP so far. Compared to larger particulate matter, UFP have distinctive characteristics that may lead to higher toxicity: their extremely small size allows them to reach deeper into the tissues and evade clearance, and higher surface-to-mass ratio facilitates adhesion of larger amounts of hazardous materials. Whether this indeed translates into a higher risk of respiratory or cardiovascular diseases in humans remains to be ascertained (Herbert and Kumar 2017).

The biological mechanisms explaining the effects of air pollution on asthma and its phenotypes and cardio- and cerebrovascular disease (CCVD) are still poorly understood. The best studied putative biological mechanism is oxidative stress caused by air pollutants, followed by pulmonary and systemic inflammation (Guarnieri and Balmes 2014; Herbert and Kumar 2017; Newby et al., 2015; Uzoigwe et al., 2013). Previous studies investigating the association between long-term exposure to air pollution and various inflammatory blood biomarkers reported inconsistent results, concerning specific cytokines and pro- or anti-inflammatory effects (Chuang et al., 2011; Fiorito et al., 2017; Mostafavi et al., 2015).

Large-scale profiling of small molecules in biological samples has become available recently, opening the door to the agnostic interrogation of disease processes at the molecular level in epidemiological settings. The metabolome reflects endogenous processes as well as the influences from environment and behaviors, and therefore metabolomics provides

a unique opportunity to link genome, exposome, and disease. Metabolomics has been increasingly applied to investigate asthma and major adverse cardiovascular events (Kelly et al., 2017; Kordalewska and Markuszewski 2015; Shah et al., 2012; Wurtz et al., 2015). However, few studies conducted an untargeted search for blood biomarkers of air pollution exposure (Vlaanderen et al., 2017) or asthma in adults, and none investigated the link between CCVD, asthma and air pollution.

This study was conducted in the framework of EXPOsOMICS, an EU-funded project to investigate the air- and water-borne exposome (Vineis et al., 2016). One of the research questions EXPOsOMICS addresses is the applicability of the "meet-in-the-middle (MITM)" concept, i.e. intermediate biomarkers as evidence of causality (Vineis et al., 2013). We have applied the MITM approach within two independent case-control studies nested in cohorts: one on adult-onset asthma (AOA) within the SAPALDIA cohort, the other on CCVD within EPIC Italy cohort, and we compared the results to identify both common and disease-specific altered metabolic pathways.

6.2 Methods

6.2.1 Study population

This study consisted of two independent case-control studies in two different cohorts. All data were collected and analyzed independently in each cohort and the results were compared to identify metabolites or metabolic pathways mediating the effect of air pollution exposure to AOA and CCVD.

Asthma in SAPALDIA

Adult-onset asthma (AOA) metabolomics was studied in a nested case-control study from the Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA). A total of 9,651 adults were recruited in eight cities representing different geographical and meteorological environments in Switzerland in 1991 (SAPALDIA1); 8,047 and 6,088 of them participated in the first follow-up in 2001-3 (SAPALDIA2) and in the second follow-up in 2010-11 (SAPALDIA3), respectively. The study protocol was described in detail previously (Ackermann-Liebrich et al., 2005; Martin et al., 1997). The present study examined blood samples from SAPALDIA3. A detailed description of the population cohort and of the study protocol was described in detail previously (Ackermann-Liebrich et al., 2005; Martin et al., 1997). Briefly, asthma cases were selected among the self-reported diagnosis of asthma occurred later than 16 years of age (n = 141) (Siroux et al., 2014) and with archived blood sample available. Controls were randomly sampled among the participants who never reported the following since SAPALDIA1: self-reported asthma; physician-diagnosed asthma; asthma attack in the last 12 months; current asthma medication; wheezing without cold in the last 12 months; three or more asthma-related symptoms in the last 12 months (symptoms considered: breathless while wheezing; woken up with a feeling of chest tightness; attack of shortness of breath after exercise; attack of shortness of breath while at rest; woken by attack of shortness of breath) (Jacquemin et al., 2015) All cases and controls had not smoked for at least 10 years before blood was drawn. Study participants were non-fasted at the time of blood collection and bench time was less than 2 hours for all but ten cases and five controls. Subjects characteristics are summarized in **Table 6.1** and compared to the entire cases and controls among SAPALDIA3 participants in **Table 6.S1**. The cases in this study were comparable to the entire SAPALDIA3

cases. The controls in this study were younger, thinner, and less exposed to air pollution compared to the entire SAPALDIA3 controls.

Cardio-cerebrovascular diseases in EPIC Italy

Study participants were part of the Italian component (Turin and Varese centers) of the EPICOR study (Bendinelli et al., 2011), which is the cardiovascular section of the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort (Palli et al., 2003). In the period 1993-1998, EPIC Italy completed the recruitment of 47,749 volunteers. The Turin and Varese cohorts include 10,604 and 12,083 participants respectively, all aged 35-65 years. We designed a case-control study nested in the cohort including 386 samples (193 matched case-control pairs), using the incident density sampling method (Richardson 2004). Criteria for cases and controls selection and matching, outcome classification and relevant covariates acquisition were described previously (Fiorito et al., 2017). Briefly, we selected all the incident CCVD cases which arose in the cohort during the follow-up (until December 2010) among non-smokers (never or former smokers for at least one year) with available blood sample archived and stored in liquid nitrogen and with at least one matched control. Matching criteria were smoking status (never/time since quitting), gender, age (±2.5 years), season and year of recruitment in the cohort. CCVDs include amyotrophic myocardial infarctions (AMI), acute coronary syndromes (ACS), coronary angioplasties, carotid angioplasties, fatal coronary events and ischemic strokes, identified after merging hospital discharge records with the EPIC Italy database, according to the criteria of the International Classification of Diseases – Ninth Revision (ICD-9). Suspected cardiovascular events or ischemic strokes were verified by a medical doctor, supported by information on onset symptoms, levels of cardiac enzymes and troponins, and electrocardiographic data. We treated CCVD as a single outcome because

Table 6.1. SAPALDIA sample characteristics – adult-onset asthma

	AOA cases	**Controls**	**AOA cases**[a]	**Controls**[a]
N	139	196	73	115
Age [year]	59.4 (19.4)	57.1 (15.8)	60.3 (19.1)	54.8 (15.5)
Female	87 (63%)	101 (52%)	47 (64%)	62 (54%)
BMI [kg/m^2]	25.7 (6.4)	24.4 (4.8)	27.0 (6.8)	24.7 (4.8)
Smoking[b]				
Former	54 (39%)	62 (32%)	34 (47%)	37 (32%)
Never	85 (61%)	134 (68%)	39 (53%)	78 (68%)
Education level[c]				
Low	3 (2%)	2 (1%)	1 (1%)	2 (2%)
Middle	86 (62%)	121 (62%)	46 (63%)	72 (63%)
High	50 (36%)	73 (37%)	26 (36%)	41 (36%)
Fasting time [hour]	2.7 (1.2)	2.8 (1.7)	2.9 (1.8)	2.7 (1.8)
Bench time [min]	80.0 (34.5)	80.0 (28.2)	80.0 (30.0)	80.0 (28.0)
PM$_{2.5}$[d] [μg/m^3]				
t = 1	14.6 (1.9)	14.3 (1.7)	15.4 (1.5)	14.7 (2.0)
t = 2	14.7 (2.2)	14.4(1.8)	15.7 (2.3)	14.8 (2.4)
t = 3	14.6 (2.8)	14.3 (2.2)	16.0 (2.3)	14.7 (2.4)
t = 4	16.0 (2.8)	15.6 (2.2)	16.7 (1.7)	16.2 (1.9)
t = 5	17.3 (2.3)	17.1 (2.1)	17.8 (1.8)	17.4 (1.9)
t = 6	16.5 (2.4)	16.0 (2.2)	17.2 (1.8)	16.4 (2.3)
t = 7	16.8 (3.4)	16.2 (3.0)	17.6 (3.2)	16.8 (3.2)
PNC [particles/cm^3]	-	-	13418 (6376)	9660 (7970)
LDSA [μ m^2/cm^3]	-	-	33.9 (16.1)	27.1 (16.3)
NO$_2$ [μ g/m^3]	25.0 (14.3)	21.6 (10.9)	29.3 (11.9)	23.7 (15.0)
Current asthma[e]	73 (53%)	-	40 (55%)	-
Current medication	45 (32%)	-	27 (37%)	-

Data are presented as count (%) or median (interquartile range). $PM_{2.5}$: annual mean estimates derived from the PolluMap in 2010; PNC and LDSA: biennial mean estimates derived from a SAPALDIA multi-area LUR in 2011/2012; NO_2: annual mean estimates derived from a European LUR in 2010. [a]Data set used for UFP MWASs, number of observation smaller due to limited availability of UFP estimates. [b]Former smokers had not smoked for at least 10 years before blood was drawn. [c]Education level low: primary school; middle: secondary/middle school or apprenticeship; high: college or university. [d]365 days average t-1 years before the examination. [e]Current asthma was defined as either having reported asthma attack in the last 12 months or currently taking asthma medication.

they share common risk factors and etiology, mainly originating from blood vessels disorders caused by atherosclerosis (Simons et al., 2009; Soler and Ruiz 2010). All subjects were fasting at the time of blood collection and bench time was always lower than two hours for cases and controls. **Table 6.2** summarizes the subjects' characteristics and **Table 6.S2** summarizes their comparison to the entire EPIC subjects. Compared to the whole EPIC cohort, this study subjects were more likely treated for hypertension; CCVD cases were more likely female, since we selected non-smokers cases only and in EPIC Italy women less likely were smokers; controls were older (due to the matching by age), had higher BMI and a higher proportion of "middle" educated individuals.

This study complies with the Declaration of Helsinki principles, and conforms to ethical requirements. All volunteers signed an informed consent form at enrolment. The study protocol of SAPALDIA was approved by the Swiss Academy of Medical Sciences and the regional committees for each study center and the one of EPIC by the Ethics Committees at the International Agency for Research on Cancer (Lyon, France) and at the Human Genetics Foundation (now IIGM, Turin, Italy) for EPIC.

6.2.2 Metabolome analyses

Serum samples were analyzed with a UHPLC-QTOF-MS system (Agilent Technologies, Palo Alto, CA, USA) in randomized order as a single batch within study. The total number of molecular features was 12,003 and 5290 for SAPALDIA and EPIC Italy respectively. A detailed description of laboratory and preprocessing procedures can be found in Supplementary Material. The features with non-missing values for at least 60% of the total sample were retained. The final dataset contained 7089 and 2790 features for SAPALDIA and EPIC Italy respectively (1,452 were in common). In EPIC Italy, additional missing values were imputed using the procedure implemented in the R package imputeLCMD.

Table 6.2. EPIC Italy sample characteristics – cardio-cerebrovascular diseases[§]

	CCVD cases	**Controls**	**CCVD cases**[a]	**Controls**[a]
N	166 [d]	155	71	73
Center				
Turin	71 (43%)	73 (47%)	71 (100%)	73 (100%)
Varese	95 (57%)	82 (53%)	-	-
Age [years]	56.16 (9.56)	56.55 (9.44)	58.01 (8.85)	57.27 (10.11)
Female	107 (64%)	95 (61%)	12 (17%)	13 (18%)
BMI [kg/m^2]	26.34 (4.91)	26.09 (4.91)	26.04 (3.82)	25.89 (4.05)
Smoking[b]				
Former	52 (31%)	54 (35%)	38 (53%)	38 (52%)
Never	114 (69%)	101 (65%)	33 (47%)	35 (48%)
Education level[c]				
Low	103 (69%)	84 (56%)	32 (45%)	22 (30%)
Middle	48 (32%)	44 (29%)	29 (41%)	31 (43%)
High	12 (8%)	22 (15%)	10 (14%)	20 (27%)
Treatments				
Hypertension	75 (45%)	47 (30%)	26 (36%)	26 (36%)
Hyperlipidaemia	69 (41%)	44 (28%)	34 (48%)	26 (36%)
PM$_{2.5}$ [μg/m^3]	21.27 (2.19)	21.27 (2.16)		
PNC [particles/cm^3]			13,283 (2,335)	13,150 (2,497)
NO$_2$ [μg/m^3]	55.15 (14.95)	54.67 (16.48)		

Data are presented as count (%) or median (interquartile range). $PM_{2.5}$: annual mean estimates derived a European LUR in 2010; PNC: annual mean estimates derived from a local LUR in 2014/2015; NO_2: annual mean estimates derived from a European LUR in 2010. [a]Data set used for UFP MWASs, number of observation smaller due to limited availability of UFP estimates. [b]Former smokers had not smoked for at least 1 year before blood was drawn. [c]Education level: low (primary school or none), middle (vocational or another secondary school), and high (university or vocational postsecondary school). [§]AMI/ACS 20%; coronary angioplasties 19%; AMI/ACS + coronary angioplasties 23%, carotid angioplasties 5%, fatal coronary events 4%; ischemic strokes 29%.

6.2.3 Air pollution exposure estimates

For $PM_{2.5}$ and UFP, the exposure estimated derived from different models were used in the two cohorts as an attempt to make the best use of available data. In SAPALDIA, annual mean exposure to $PM_{2.5}$ in 2010 (SAPALDIA3 survey) of study participants was estimated by using PolluMap, a national air pollution dispersion model for Switzerland (FOEN 2013). Lagged estimates up to 7 years before SAPALDIA3 were obtained by interpolation from Meteotest (FOEN 2014). Biennial mean exposure to UFP was estimated based on multi-area land use regression (LUR) models derived from SAPALDIA specific-measurement campaigns conducted in 2011/2012 and covering 4 out of 8 SAPALDIA study areas (Eeftens et al., 2016). In EPIC Italy, $PM_{2.5}$ exposure was estimated by a newly developed European LUR model derived from measurements in 2010 (de Hoogh et al., 2016). UFP exposure in Turin was estimated by a local LUR model derived from measurements in 2014/2015 (van Nunen et al., 2017). Both SAPALDIA and EPIC Italy used the NO_2 exposure estimates provided by the aforementioned European LUR model (de Hoogh et al., 2016). In addition to particle number concentration (PNC), lung deposited surface area (LDSA) was used as UFP metric in SAPALDIA. The air pollution model performance varied across air pollutants and the models: the cross-validation R^2 was 0.54 for $PM_{2.5}$ derived from the European LUR (de Hoogh et al., 2016); 0.82 and 0.87 for PNC and LDSA, respectively, from the Swiss local LUR (Eeftens et al., 2016); 0.33 for PNC from the European local LUR (van Nunen et al., 2017); 0.58 for NO_2 from the European LUR (de Hoogh et al., 2016). As we relied on LUR models developed to cover limited areas, UFP estimates were available for a subset of samples, 75 AOA cases and 115 controls, and 71 CCVD cases and 73 controls. Each subject was assigned air pollution exposure estimates by geocoding the residential address. In the case of SAPALDIA this

was the address at the time point of the SAPALDIA3 survey. In the case of EPIC Italy this was the address at the time of blood sample collection.

6.2.4 Statistical analyses

Statistical analyses were conducted independently in the two cohorts, applying slightly different models and covariates to accommodate the discrepancy in the data availability between the two cohorts.

Association of air pollution exposure with AOA

We assessed the effect of air pollution exposure on AOA by fitting logistic regression models. AOA was regressed, with non-asthmatics as the reference, on air pollution exposure after adjustment for age, sex, education level, body mass index (BMI), and study area as random effect. For $PM_{2.5}$, the main predictors were two polynomial lag terms defined as $u_0 = \sum_{t=1}^{7} PM_{2.5}(t)$ and $u_1 = \sum_{t=1}^{7} t \cdot PM_{2.5}(t)$, where $PM_{2.5}(t)$ is average exposure to $PM_{2.5}$ of 365 days $t-1$ years SAPALDIA3 examination. For UFP and NO_2, the main predictors were biennial and annual mean estimates respectively. The association was also assessed in the entire SAPALDIA subjects (N=3,011; 272 AOA cases). In the analysis of the entire SAPALDIA subjects, a binary indicator for perfect geocoding quality was additionally included as a potential modifier of the effect of air pollution exposure on the metabolite level. Geocoding was declared perfect if the matching was possible at the level of residential address. In the analysis of the nested case-control samples, the observations with non-perfect geocoding quality were excluded because the models with the effect modifier did not converge.

Association of air pollution exposure with CCVD

The association of exposure to air pollution with CCVD was assessed in the nested case-control study by logistic regression models adjusting for age at recruitment, center of recruitment, sex, BMI, smoking status, and education level (see Supplementary Material for details). In addition, we conducted Cox proportional hazard regression to assess the association between air pollution exposure and the risk of future CCVD among all EPIC subjects (Turin and Varese centers; N=18,982; 948 CCVD events).

In both studies, odds ratios (OR), hazard ratios (HR), and 95% confidence intervals (CI) refer to an increase of 5 μg/m^3 $PM_{2.5}$, 5,000 particles/cm^3 PNC, 10 m^2/cm^3 LDSA, and 10 μg/m^3 NO_2.

Metabolome-wide association study (MWAS) on AOA

We conducted logistic regression analyses of AOA on each of the 7089 features after adjustment for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 hours, and their multiplicative interaction terms with fasting time. We did not adjust for smoking because all subjects were non-smokers since 10 years. Feature intensity, age, bench time, and fasting time were scaled to have mean equal 0 and standard deviation equal 1. We applied the Firth's bias-reduction method (Firth 1993; Perry 2016) to obtain less biased estimates and the Benjamini-Hochberg method to correct for multiple testing (Benjamini and Hochberg 1995). Acknowledging the cross-sectional nature of the study of the AOA MWAS, we conducted a sensitivity analysis by repeating the AOA MWAS after further adjustment for current asthma medication.

MWAS on CCVD

For each of the 2790 features, we tested for their association with incident CCVD by logistic regression models adjusting for age at recruitment, center of recruitment, sex, BMI, smoking status, and education level. Sensitivity analyses were conducted on cardiovascular cases only (i.e. excluding ischemic strokes).

MWAS on air pollution

In SAPALDIA and EPIC Italy separately, each feature was regressed on $PM_{2.5}$, UFP, or NO_2 after adjustment for the same covariates as in AOA MWAS and in CCVD MWAS, respectively. In SAPALDIA, a binary indicator for perfect geocoding quality was additionally included as a potential modifier of the effect of air pollution exposure on the metabolite level. Geocoding was declared perfect if the matching was possible at the level of residential address. As in the association of air pollution with AOA, first and second order polynomial lag terms were used for $PM_{2.5}$ while biennial and annual mean exposures were used for UFP and NO_2, respectively. In EPIC Italy, annual average exposure was used as the proxy for long-term exposure for each pollutant.

Link and variance functions

In EPIC Italy, feature intensities were Box-Cox transformed before regression (Han and Kronmal 2004). In SAPALDIA, the best link and variance were sought for each feature and semi-partial pseudo-R^2 was computed as a measure of effect size (see Supplementary Material for details).

6.2.5 Meet-in-the-middle (MITM) approach

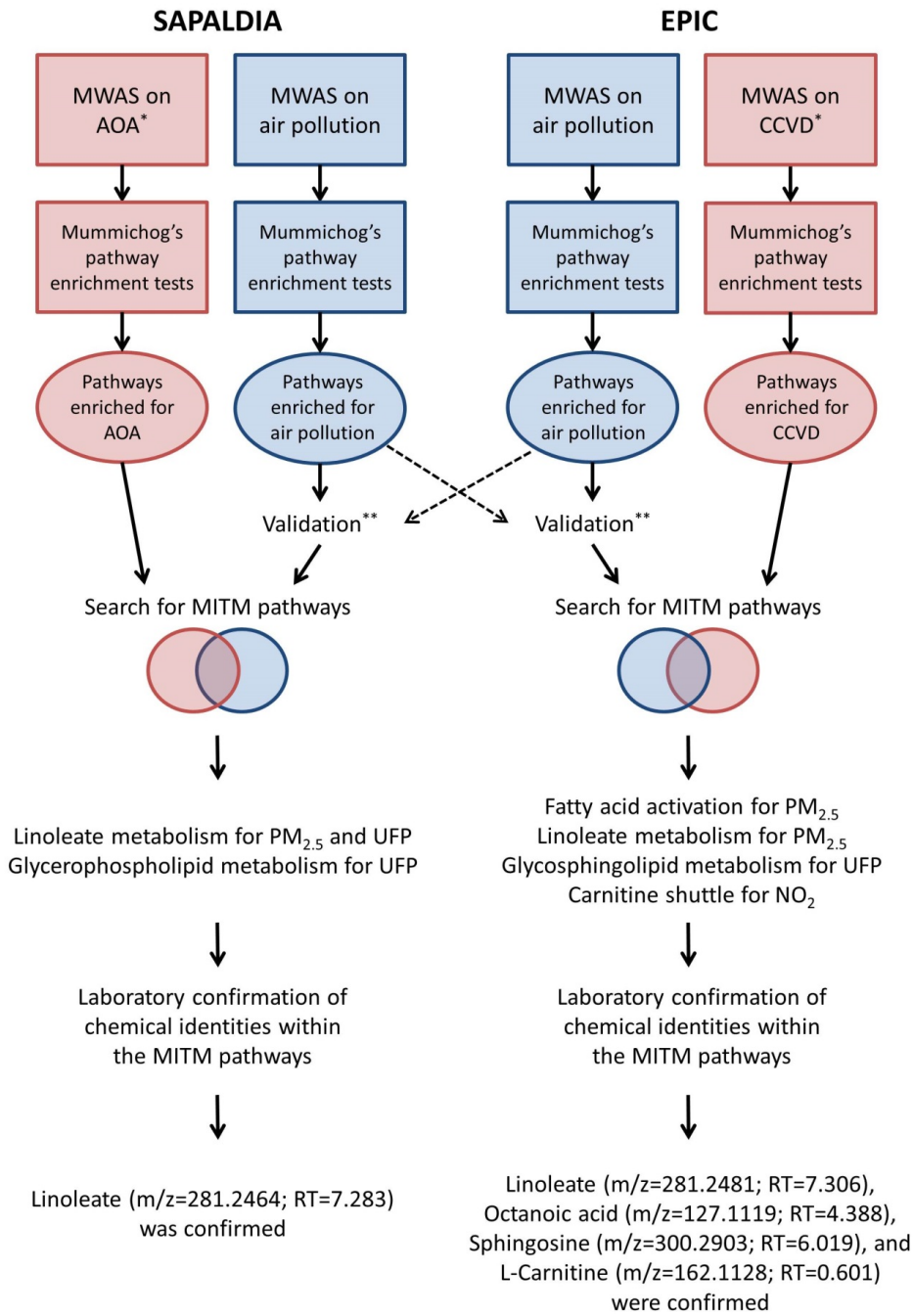
Search for MITM features

We examined if any of the features associated with air pollution overlapped with the features associated with AOA or CCVD as an attempt to search for MITM features. As no single feature showed metabolome-wide significant association with AOA or CCVD, we found no single MITM features. Instead, we searched for MITM pathways as described below. The history of our analyses in this study is summarized as flowcharts in supplementary materials (**Figure 6.S1**: MITM features; **Figure 6.1**: MITM pathways).

Functional annotation and pathway enrichment tests using Mummichog

Mummichog is an algorithm developed to predict functional activities of metabolites (Li et al., 2013). Taking untargeted MWAS results as input, Mummichog searches for chemical identities by matching the measured mass (m/z) of the features to a reference metabolic model, integrated from KEGG (Kanehisa et al., 2006), UCSD BiGG (Duarte et al., 2007), and Edinburgh human metabolic network (Ma et al., 2007). Based on this putative annotation, it conducts pathway enrichment tests using Fisher's exact test. The statistical significance of pathway enrichment is estimated by permutation, where the features are randomly selected and mapped to each of the possible annotations to produce null distribution. We customized the types of ions that Mummichog searches for chemical identities, to match with the UHPLC-QTOF-MS method used. Cut-off p-value was chosen to have a reasonable number of significant features to ensure for the algorithm to conduct pathway enrichment analysis. We first used the 10th percentile of the p-values from each MWAS result as the cut-off and then the 5th percentile as a sensitivity analysis (**Table 6.S3**).

Figure 6.1. Search for the MITM pathways



*Adjusted for the corresponding air pollutant; **by excluding the pathways not enriched in the other cohort.

Search for MITM pathways

Pathways found enriched (empirical p-value < 0.05) from Mummichog were listed. The pathways with overlap size – the number of features that contributed to the enrichment – smaller than 4 were ignored. This is an attempt to reduce the false positive findings as Mummichog annotates features only by matching m/z and hence matches are subject to error. The pathways that were not enriched for the same air pollution metric in both SAPALDIA and EPIC Italy were excluded. If the pathway enriched for air pollution metric was also enriched for AOA or CCVD after adjustment for the same metric, they were declared as "MITM" pathways (**Figures 6.S3-6.S5**). The MITM pathways were evaluated by confirmation of the putative annotation which Mummichog used to compute pathway enrichment (see Supplementary Material for details).

6.3 Results

6.3.1 Exposure to UFP is associated with AOA

From logistic regression of AOA (n = 73) with non-asthmatics as the reference group (n = 115), we found a strong association of UFP exposure with AOA (**Table 6.3**). The odds ratios were 1.80 [95% CI 1.26, 2.55] for an increase in particle number concentration (PNC) by 5,000 particles/cm^3, and 1.73 [95% CI 1.27, 2.36] for an increase in lung deposited surface area (LDSA) by 10 μm^2/cm^3. On the contrary, $PM_{2.5}$ and NO_2 did not show a significant association with AOA. The estimated risk for AOA due to UFP exposure is still significant after the inclusion of either $PM_{2.5}$ or NO_2 in the regression model. LDSA showed a stronger, significant association with AOA in the multipollutant model, supporting the independence of the effect (**Tables 6.S4-6.S5**). The ORs were lower when estimated in the whole cohort, which may be partly attributed to the dis-

crepancy in the samples: While the cases in this study were comparable to the entire SAPALDIA3 cases, the controls in this study had higher exposure to air pollutants than the entire SAPALDIA3 controls (**Table 6.S1**).

Table 6.3. Association of air pollution with AOA and CCVD

Air pollution metric	**AOA**		**CCVD**	
	OR[a] [95% CI]	**OR**[b] [95% CI]	**OR**[c] [95% CI]	**HR**[d] [95% CI]
$PM_{2.5}$[e]	1.05 [0.57, 1.95]	1.00 [0.65, 1.56]	1.34 [0.72, 2.52]	1.29 [1.08, 1.55]
PNC[f]	1.80 [1.26, 2.55]	1.39 [1.03, 1.87]	1.09 [0.60, 2.00]	1.16 [0.97, 1.39]
LDSA[g]	1.73 [1.27, 2.36]	1.36 [1.04, 1.79]	-	-
NO_2[h]	1.12 [0.81, 1.55]	1.16 [0.94, 1.43]	1.03 [0.89, 1.18]	1.12 [0.99, 1.27]

Note: Odds ratios are reported for all cross-sectional analyses (meet-in-the-middle/metabolome subsample) while hazard ratio is reported for the longitudinal analysis on larger CCVD samples; Sample size is smaller for UFP than other pollutants because the LUR models were derived only for 4 out of 8 study areas in SAPALDIA and for Turin but not for Varese in EPIC Italy. [a]odds ratio adjusted for age, sex, education level, BMI, and study area as random effect (N=335 for $PM_{2.5}$ and NO_2; N=188 for UFP). [b]odds ratio adjusted for age, sex, education level, BMI, and study area as random effect (N=3,011 for $PM_{2.5}$ and NO_2; N=1,555 for UFP). [c]odds ratio adjusted for age, center of recruitment, sex, BMI, smoking status, and educational level (N=321 for $PM_{2.5}$ and NO_2; N=144 for UFP). [d]hazard ratio adjusted for age, center of recruitment, sex, BMI, smoking status, and educational level (N=18,982 for $PM_{2.5}$ and NO_2; N=8,753 for UFP). [e]per 5 μg/m3 increase in biennial (SAPALDIA) or annual (EPIC Italy) mean $PM_{2.5}$. [f]per increase by 5000 particles/cm^3 in biennial (SAPALDIA) or annual (EPIC Italy) mean PNC. [g]per increase by 10 $\mu m^2/cm^3$ in biennial mean LDSA. [h]per increase by 10 $\mu g/m^3$ in annual mean NO_2.

6.3.2 Weak but consistent association of air pollution with CCVD

We have observed a positive association of exposure to $PM_{2.5}$, PNC, and NO_2 with the risk of CCVD (OR = 1.34 [95% CI 0.72, 2.52] for 10 $\mu g/cm^3$ increase in $PM_{2.5}$; OR = 1.09 [95% CI 0.60, 2.00] for 5,000 particles/cm^3 increase in PNC; OR = 1.03 [95% CI 0.89, 1.18] for $\mu g/cm^3$ increase in NO_2), though the associations did not reach statistical significance (**Table 6.3**). However, when we expanded the analyses to the whole EPIC Turin-Varese subjects (N=18,982; 948 CCVD events), the associations became stronger

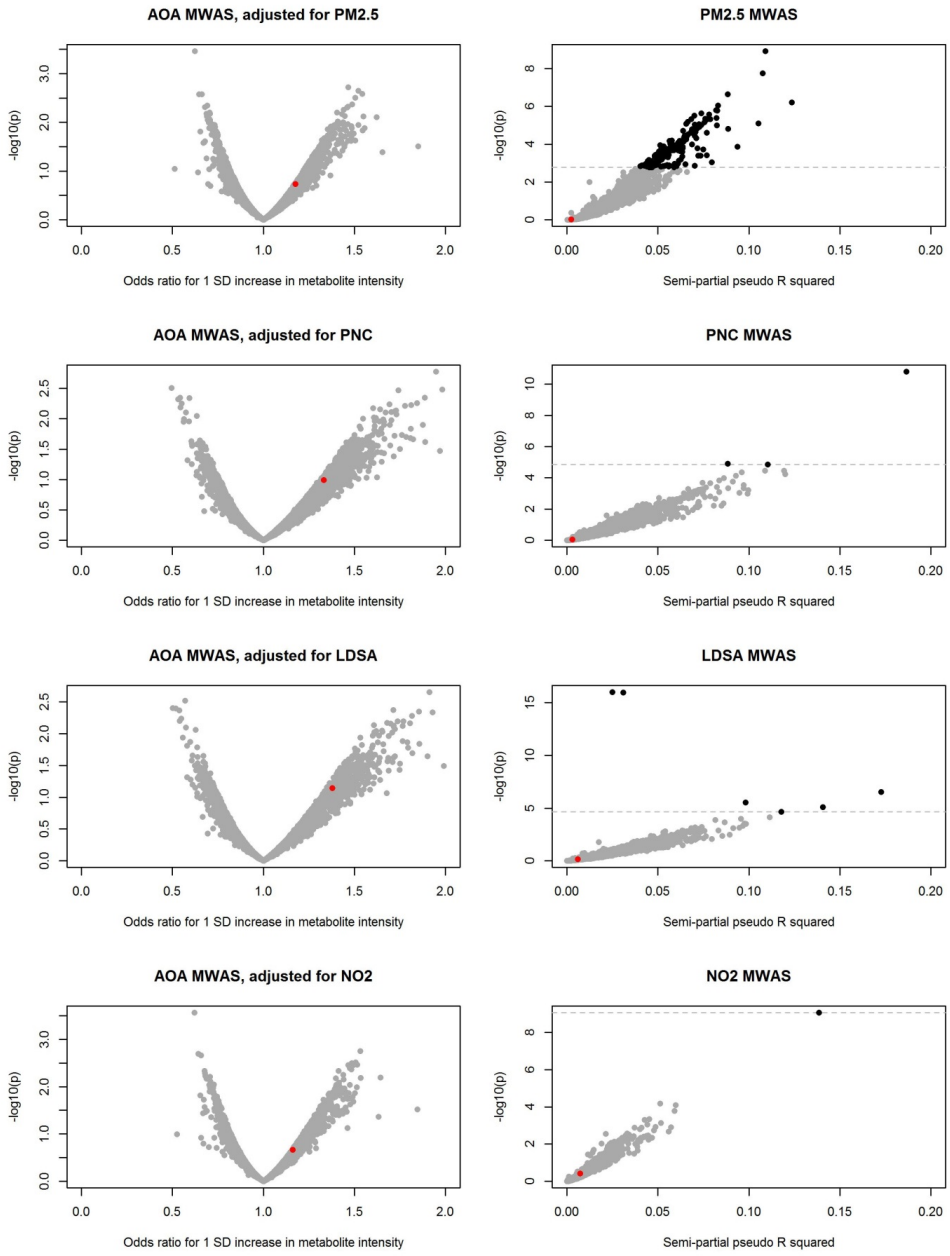
and significant (HR = 1.29 [95% CI 1.08, 1.55] for 10 μg/m^3 increase in $PM_{2.5}$; HR = 1.16 [95% CI 0.97, 1.39] for 5,000 particles/cm^3 increase in PNC (Turin subjects; N=8,753); HR = 1.12 [95% CI 0.99, 1.27] for μg/cm^3 increase in NO_2). In the multipollutant model, the OR for PNC was higher than that estimated in the single pollutant model. On the contrary the OR for $PM_{2.5}$ dramatically decreased when including other pollutants in the regression model. These results suggest that the association of air pollution with CCVD could be mainly driven by PNC.

6.3.3 MWAS: no single metabolites are associated with both air pollution and AOA or CCVD

None of the 7089 features in SAPALDIA or 2790 features in EPIC Italy showed a significant association with AOA or CCVD after multiple testing corrections, respectively (**Figure 6.S1**). The air pollution MWAS in SAPALDIA showed 237, three, six and one features significantly associated with $PM_{2.5}$, PNC, LDSA, and NO_2, respectively (**Figure 6.2**). One of the three PNC associated features coincided with the LDSA associated features. Five out of the eight UFP associated features were not associated with any other air pollutant. The only NO_2 associated feature was also associated with $PM_{2.5}$ (**Figure 6.S2**). In EPIC Italy, no single feature showed a significant association with air pollution exposure, nor with CCVD after multiple testing corrections (**Figure 6.3**). The top 100 signals from each of the air pollution MWASs in both cohorts are summarized in Supplementary Material[a] with putative annotation.

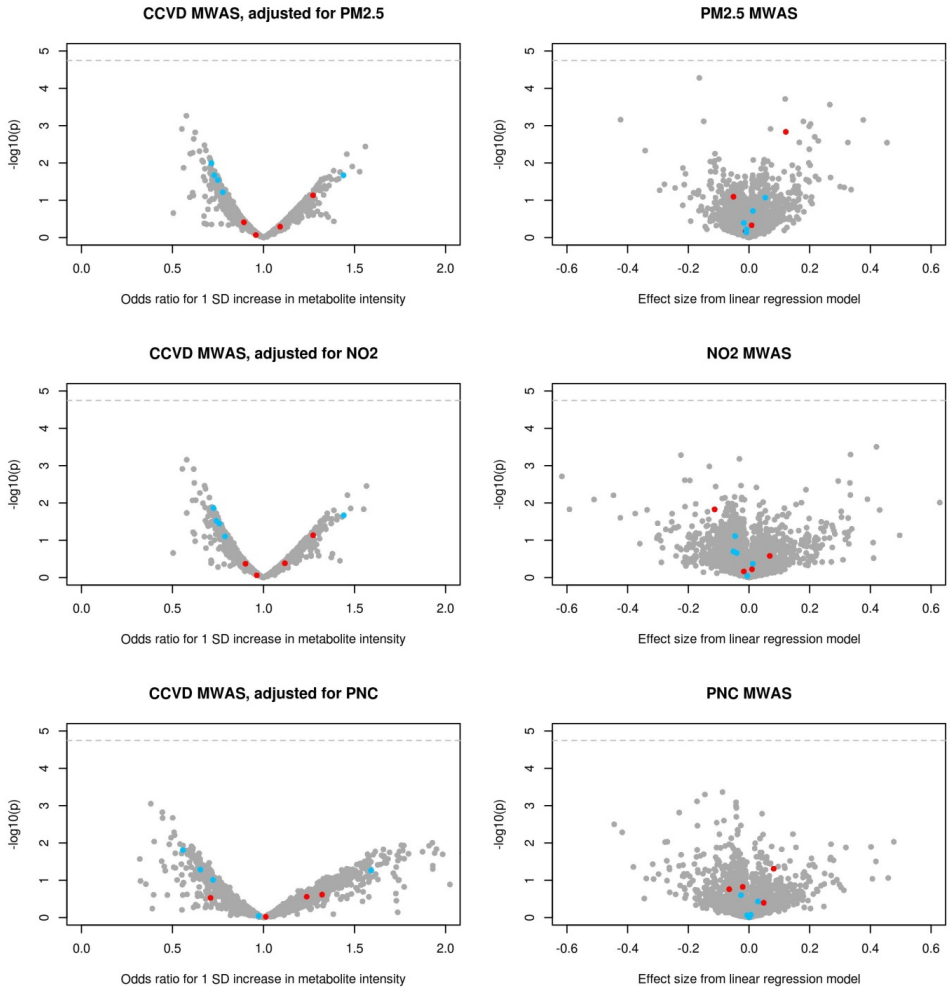
[a]This table is available in online supplement and not included in this dissertation.

Figure 6.2. Volcano plots of MWAS results in SAPALDIA



Note the asymmetric distribution of points in air pollution MWASs due to the positive nature of semi-partial pseudo-R^2 used as a measure of effect size. Linoleate (m/z = 281.2464; RT = 7.283) whose annotation was confirmed with confidence level 1 is highlighted in red; Metabolome-wide signals after Benjamini-Hochberg correction in black. Dotted line depicts Benjamini-Hochberg adjusted p = 0.05.

Figure 6.3. Volcano plots of MWAS results in EPIC Italy



Metabolites whose annotation was confirmed with confidence level 1 are highlighted in red: Linoleate (m/z = 281.2481; RT = 7.306), Octanoic acid (m/z = 127.1119; RT = 4.388), Sphingosine (m/z = 300.2903; RT = 6.019), and L-carnitine (m/z = 162.1128; RT = 0.601); Metabolites whose annotation was confirmed with confidence level 3 are in blue: α-Linolenic acid (m/z = 279.2321; RT = 7.166), D-Glucose (m/z = 145.0495; RT = 0.646), Linoelaidyl carnitine (m/z = 424.3428; RT = 6.199), Octadecenoyl carnitine (m/z = 426.3590; RT = 6.337), and Stearoylcarnitine (m/z = 428.373; RT = 6.479). No metabolome-wide signals after Benjamini-Hochberg correction.

6.3.4 Several metabolic pathways are commonly associated with air pollution in both cohorts

Various pathways were associated with air pollution varying with the air pollutant and the cohort examined (**Figure 6.1, Tables 6.S6-6.S11**). The pathways that were enriched for the same air pollutant in both cohorts are summarized in **Table 6.4** and **Figures 6.S3-6.S5**: Linoleate metabolism and Fatty acid activation were enriched for $PM_{2.5}$; Linoleate metabolism, Glycerophospholipid metabolism, and Glycosphingolipid metabolism for UFP; Carnitine shuttle and Pyrimidine metabolism for NO_2. No overlap was found looking at the list of features that contributed to the enrichment in the two studies (**Table 6.S13**). We then repeated the same enrichment analysis using the 5th percentile p-value as the cut-off, as a sensitivity analysis. Linoleate metabolism and Glycerophospholipid metabolism, associated to UFP, were confirmed in both cohorts. All the pathways associated to NO_2, Carnitine shuttle and Pyrimidine metabolism, were also confirmed.

Table 6.4. Pathways associated to air pollution in both SAPALDIA and EPIC Italy

		SAPALDIA			EPIC Italy		
Air pollutant	**Pathway**	**Overlap size**	**Pathway size**	**p-value**	**Overlap size**	**Pathway size**	**p-value**
$PM_{2.5}$	Linoleate metabolism[b,c]	17	21	0.0007	6	20	0.0249
	Fatty acid activation[c]	10	21	0.0054	5	15	0.0180
UFP[a]	Linoleate metabolism[b]	12	21	0.0007	7	20	0.0084
	Glycerophospholipid metabolism[b]	12	36	0.0023	13	35	0.0022
	Glycosphingolipid metabolism[c]	8	26	0.0079	6	21	0.0367
NO_2	Carnitine shuttle[c]	10	26	0.0063	6	19	0.0040
	Pyrimidine metabolism	12	33	0.0074	8	28	0.0035

[a]Either PNC or LDSA in SAPALDIA and PNC in EPIC Italy. [b]also enriched for AOA after further adjustment for the corresponding air pollutant. [c]also enriched for CCVD after further adjustment for the corresponding air pollutant.

6.3.5 Pathways enrichment and MITM analysis for AOA and CCVD

We found various altered metabolic pathways associated with AOA and CCVD (**Figure 6.1**, **Tables 6.5 and 6.6**). The majority of the enriched pathways did not overlap between AOA and CCVD. Pathways associated with AOA and CCVD, respectively, after adjustment for single air pollution metrics to identify MITM pathways are presented in **Tables 6.S14-6.S20**.

Table 6.5. Pathways associated to AOA unadjusted for air pollution exposure

Pathway	Overlap size	Pathway size	p-value
Tryptophan metabolism	20	54	0.0009
Vitamin B6 (pyridoxine) metabolism	4	6	0.0017
Biopterin metabolism	6	13	0.0021
TCA cycle	4	8	0.0041
Hexose phosphorylation	5	12	0.0048
Fatty Acid Metabolism	5	14	0.0101
De novo fatty acid biosynthesis	7	22	0.0102
Drug metabolism - cytochrome P450	12	42	0.0102
Valine, leucine and isoleucine degradation	7	23	0.0137
Urea cycle/amino group metabolism	9	32	0.0169
Fatty acid activation	6	21	0.0258
Leukotriene metabolism	13	51	0.0278
Butanoate metabolism	5	17	0.0284
Glycosphingolipid metabolism	7	26	0.0312
Lysine metabolism	6	22	0.0342
Drug metabolism - other enzymes	5	18	0.0387
Arginine and Proline Metabolism	6	23	0.0447
Starch and Sucrose Metabolism	4	14	0.0458
Pentose and Glucuronate Interconversions	4	14	0.0458
Vitamin E metabolism	8	32	0.0461

Mummichog pathway enrichment test on the results from AOA MWAS adjusted for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 hours, and their multiplicative interaction terms with fasting time.

Table 6.6. Pathways associated to CCVD unadjusted for air pollution exposure

Pathway	Overlap size	Pathway size	p-value
De novo fatty acid biosynthesis	9	14	0.0011
Hexose phosphorylation	8	12	0.0012
Phosphatidylinositol phosphate metabolism	6	10	0.0031
Carnitine shuttle	9	19	0.0047
Starch and Sucrose Metabolism	6	11	0.0051
Linoleate metabolism	9	20	0.0070
Glycosphingolipid metabolism	9	21	0.0105
Glutamate metabolism	5	10	0.0139
Caffeine metabolism	5	11	0.0249
Fatty acid activation	6	15	0.0398
Glycolysis and Gluconeogenesis	4	9	0.0479
Fructose and mannose metabolism	4	9	0.0479

Mummichog pathway enrichment test on the results from CCVD MWAS adjusted for age at recruitment, center of recruitment, sex, BMI, smoking status, and education level.

6.3.6 Linoleate metabolism is a common MITM pathway linking air pollution to AOA and CCVD

Linoleate metabolism was enriched for $PM_{2.5}$ and UFP in both cohorts and for AOA after adjustment for $PM_{2.5}$ or UFP (**Tables 6.S14-6.S16**) as well as for CCVD after adjustment for $PM_{2.5}$ (**Table 6.S18**). Therefore, we considered Linoleate metabolism as MITM linking $PM_{2.5}$ and UFP to AOA and $PM_{2.5}$ to CCVD. Similarly, we considered Glycerophospholipid metabolism as MITM linking UFP to AOA (**Table 6.S16**); Fatty acid activation, Glycosphingolipid metabolism, and Carnitine shuttle as MITM linking $PM_{2.5}$, UFP, or NO_2 to CCVD, respectively (**Tables 6.S18-6.S20**).

Linoleate metabolism and Glycerophospholipid metabolism were confirmed as MITM pathways linking UFP to AOA after the sensitivity analysis (5^{th} percentile of p-values as the cut-off), as well as Glycosphingolipid metabolism linking UFP to CCVD, and Carnitine shuttle linking NO_2 to CCVD. In SAPALDIA, Linoleate metabolism and Glycerophospholipid metabolism remained significant after adjustment for current asthma

medication, indicating that these MITM pathways were not driven by reverse causation (**Tables 6.S21-6.S22**). In EPIC, Glycosphingolipid metabolism and Carnitine shuttle were still significant as the MITM pathways for UFP and NO_2 respectively in the sensitivity analyses after excluding ischemic strokes (**Tables 6.S23-6.S24**).

6.3.7 Confirmed annotation of metabolites in MITM pathways

A total of 108 features mapping to the aforementioned MITM pathways were selected for confirmation of the putative annotation. **Table 6.7** summarizes all the features whose annotation was confirmed using chemical standards and fragmentation spectra. Linoleate was confirmed in both cohorts with confidence level 1 according to the classification of the Chemical Analysis Working Group (CAWG) (Sumner et al., 2007). In SAPALDIA, linoleate was considered as a signal for the AOA MWAS further adjusted for UFP and contributed to the enrichment of Linoleate metabolism and Glycerophospholipid metabolism. In EPIC Italy, linoleate was considered as a signal for the $PM_{2.5}$ MWAS and contributed to the enrichment of Linoleate metabolism. Also confirmed were octanoic acid, sphingosine, and L-carnitine, contributing in EPIC Italy to the enrichment of Fatty acid activation for $PM_{2.5}$, Glycosphingolipid metabolism for UFP, and Carnitine shuttle for CCVD adjusted for NO_2, respectively. Five additional features were confirmed for their chemical classes with confidence level 3 for the CAWG (Sumner et al., 2007).

6.3.8 Additional sensitivity analyses

For consistency between the two studies, we performed further sensitivity analyses on AOA. Additional adjustment for education level resulted in a non-relevant change of the results, while adjustment for BMI slightly changed the results (**Table 6.S25**). In the pathway enrichment analyses, Glycerophospholipid metabolism remained as MITM link-

Table 6.7. MWAS results for features with confirmed annotation

Metabolite	Putative annotation from Mummichog	Level of confidence	Pathway	MWAS	Regression model	Coefficient	p-value	Pseudo-$R^{2§}$
m/z = 281.2464 RT = 7.283	Linoleate	Level 1	Linoleate metabolism; Glycerophospholipid metabolism	AOA, PNC adjusted AOA, LDSA adjusted PNC in SAPALDIA LDSA in SAPALDIA	Logistic Logistic Gamma with log link Gamma with log link	0.29 0.32 7.9e-7 -0.00026	0.10 0.071 0.86 0.69	- - 0.0030 0.0059
m/z = 281.2481 RT = 7.306	Linoleate	Level 1	Fatty acid activation; Linoleate metabolism	CCVD, $PM_{2.5}$ adjusted $PM_{2.5}$ in EPIC Italy	Logistic Linear	1.05 0.06	0.40 0.001	- -
m/z = 127.1119 RT = 4.388	Octanoic acid	Level 1	Fatty acid activation	CCVD, $PM_{2.5}$ adjusted $PM_{2.5}$ in EPIC Italy	Logistic Linear	0.93 -0.03	0.17 0.08	- -
m/z = 300.2903 RT = 6.019	Sphingosine	Level 1	Glycosphingolipid metabolism	CCVD, PNC adjusted PNC in EPIC Italy	Logistic Linear	0.65 0.0003	0.30 0.08	- -
m/z = 162.1128 RT = 0.601	L-carnitine	Level 1	Carnitine shuttle	CCVD, NO_2 adjusted NO_2 in EPIC Italy	Logistic Linear	3.24 0.001	0.07 0.59	- -
m/z = 279.2321 RT = 7.166	α-Linolenic acid; γ-Linolenic acid	Level 3	Linoleate metabolism Fatty acid activation	CCVD, $PM_{2.5}$ adjusted $PM_{2.5}$ in EPIC Italy	Logistic Linear	0.51 0.03	0.01 0.08	- -
m/z = 145.0495 RT = 0.646	D-Glucose; Galactose	Level 3	Glycosphingolipid metabolism	CCVD, PNC adjusted PNC in EPIC Italy	Logistic Linear	1.59 0.0001	0.05 0.97	- -
m/z = 424.3428 RT = 6.199	Linoelaidyl carnitine; Linoleyl carnitine	Level 3	Carnitine shuttle	CCVD, NO_2 adjusted NO_2 in EPIC Italy	Logistic Linear	0.79 -0.007	0.08 0.20	- -
m/z = 426.3590 RT = 6.337	Octadecenoyl carnitine; Vaccenyl carnitine; Elaidic carnitine	Level 3	Carnitine shuttle	CCVD, NO_2 adjusted NO_2 in EPIC Italy	Logistic Linear	0.75 -0.005	0.03 0.22	- -
m/z = 428.373 RT = 6.479	Stearoylcarnitine	Level 3	Carnitine shuttle	CCVD, NO_2 adjusted NO_2 in EPIC Italy	Logistic Linear	0.34 -0.006	0.03 0.08	- -

§ Semi-partial pseudo R^2 as a measure of effect size for air pollution MWASs in SAPALDIA: for details see supplementary information. Level 1: identified compounds; Level 3: putatively characterized compound class (Sumner et al., 2007).

ing UFP to AOA after adjustment for BMI or for education level. Linoleate metabolism remained as MITM linking UFP to AOA after adjustment for education level but not after adjustment for BMI.

6.4 Discussion

In short-term studies, UFP exposure has been reported to have cardio-respiratory effects that were stronger than for larger particles. Peters et al. reported that UFP exposure had a stronger effect on peak expiratory flow than larger particles (Peters et al., 1997). Exposure to UFP but not to larger particles was associated with asthma exacerbations in children (Evans et al., 2014). However, a recent *in vitro* study showed that coarse particles might have stronger effects on airway epithelium, possibly due to the higher iron content in coarse particles (Kumar et al., 2015). Studies investigating the long-term cardio-respiratory effects of UFP exposure remain very limited. In the California Teachers Study cohort, UFP exposure derived from a chemical transport model was associated with all-cause and ischemic heart disease mortality (Ostro et al., 2015). In the SAPALDIA cohort, UFP exposure was associated with carotid-intima media thickness, a marker of subclinical atherosclerosis (Aguilera et al., 2016). UFP exposure derived from a city-specific LUR model in Toronto linked to health registry data of 1.1 million adult city residents found no positive association of UFP exposure with respiratory disease incidence including AOA (Weichenthal et al., 2017). This is in contrast to our findings, which are based on individual reports of asthma and which provide evidence of UFP effects being stronger than, and independent of, those of larger particles. The results for the multipollutant model suggest that ultrafine particles instead of particulate matters could mainly drive associations with AOA and CCVD. In fact, the risks conferred by PNC estimated

in the multipollutant models were higher than those determined in the single pollutant models. On the contrary, the ORs for $PM_{2.5}$ dramatically decrease when including PNC and NO_2 in the regression models. However, it is not easy to address this question and to correctly interpret these results due to the strong correlation among various pollutants. To investigate whether ultrafine particles or particulate matters confer the main risk goes beyond the scope of this study and needs a more specific study design.

Traffic-related pollutants contribute mainly to the fine or ultrafine particles, while specks of dust of geological origin including metals link to the coarse particles (Kelly and Fussell, 2012, Yamada et al., 2005). Particulates of various sizes may have different toxicity dependent on their composition (Kumar et al., 2015, Schwarze et al., 2007). However, we cannot rule out the possibility of the residual confounding by unaddressed exposures potentially correlated with air pollution e.g. transportation noise.

6.4.1 Meet-in-the-middle (MITM) approach

We applied the "meet-in-the-middle (MITM)" approach, which helps in developing a causal hypothesis and improve biological understanding for air pollution-cardio-respiratory health associations, making use of high-resolution metabolomic data. In the MITM approach, one searches for intermediate biomarkers that are associated with both the exposure and the outcome (Vineis et al., 2013). Ideally, this applies to longitudinal studies where the exposure precedes the biomarker measurement, and the biomarker measurement precedes the outcome, e.g. incidence of cardiovascular events, as we did for CCVD in EPIC Italy. It is much less straightforward to define incident cases for asthma than for CCVD. Asthma is a complex chronic disease phenotype that develops over a long period of time, can go unnoticed for years if not for decades, and can also disappear as well as resurface. This difficulty inherent to asthma research is complicating the assessment of

causality to identified risks such as air pollution. Realizing this difficulty, we pursued the MITM approach for asthma even though our study is by design cross-sectional. For all these reasons, we restricted the outcome to adult (after the 16 years of age) onset of asthma which is less susceptible to reverse causation bias and exposure misclassification.

6.4.2 MWAS analyses

At the level of single metabolites, we found no intermediate biomarkers among the 7089 and 2790 features investigated in SAPALDIA and EPIC Italy respectively, due to lacking metabolome-wide significant associations. Multiple testing corrections can be too stringent, given the highly inter-correlated nature of the metabolome. The effective number of tests (ENT) computed for the SAPALDIA metabolome was 2728, indicating a high degree of dependency in the data. Given this highly correlated, high dimensional data structure, our study likely suffers from low power to detect subtle differences related to chronic diseases, and in particular to asthma, where distinguishing sub-phenotypes may be essential for understanding risk and etiology of the disease (Jeong et al., 2017, Siroux et al., 2014, Wenzel, 2012). Therefore, heterogeneity and misclassification might have attenuated the associations with biomarkers. Distinguishing further sub-phenotypes requires larger data in future metabolome studies. Given the above, we focused on pathway enrichment analyses.

6.4.3 Pathway enrichment analyses

Metabolomics, given the high dimensionality and high dependency, benefits much from multivariate systems approaches like pathway enrichment tests. Yet, the challenge unique to metabolomics in this context is annotation. Unlike other omics, annotation of the features obtained from untargeted metabolomics requires laborious manual work. The

Mummichog software offers an opportunity to bypass this step and to conduct pathway enrichment tests directly from untargeted MWAS results. Using Mummichog, we found various pathways enriched for AOA, CCVD, and air pollution exposures. Air pollution MWASs and pathway enrichment tests conducted in two cohorts served as each other's validation. Although we found no single overlapping features between the two cohorts when comparing validated pathways, lack of such overlap does not exclude the possibility that the pathways truly reflect air pollution-induced metabolic changes, involving different molecules. The specific molecules affected in a pathway may, for example, depend on the particle composition which can vary across different areas (Kelly and Fussell, 2012).

However, the untargeted metabolomics using UHPLC-QTOF-MS system detects a limited range of molecular weight and concentration, making exogenous molecules hard to be detected. And Mummichog relies on the knowledge based databases, where the metabolic studies were concentrated on the endogenous molecules (Rappaport et al., 2014). This could have brought bias into our pathway enrichment findings.

6.4.4 Linoleate metabolism is a common MITM pathway for AOA and CCVD

AOA and CCVD were mostly associated with different sets of pathways and hence MITM pathways linking air pollution exposure to both chronic diseases differed. The two chronic diseases may involve different biological mechanisms and the same environmental insults may act through different pathways. One exception was Linoleate metabolism pathway, which was found not only as MITM pathway linking $PM_{2.5}$ and UFP to AOA but also linking $PM_{2.5}$ to CCVD. Laboratory analysis confirmed the annotation of linoleate in both cohorts. The feature confirmed as linoleate showed a positive association with AOA, while it did not show statistically significant association with UFPs exposure and did

not contribute to the pathway enrichment for UFPs. Still, the Lineolate MITM-pathway finding seems biologically interesting. Linoleate was reported in an *in vitro* experiment to regulate the pro-inflammatory cytokine IL8 (Maruyama et al., 2014) and induce smooth muscle contraction via the free fatty acid receptor 1 (FFAR1) (Mizuta et al., 2015). Another *in vitro* study demonstrated that α1-antitrypsin bound to linoleate reduced the expression and secretion of IL1β in LPS-stimulated neutrophils, while free α1-antitrypin did not (Aggarwal et al., 2016). In observational studies in children, eczema was positively associated with linoleate intake (Miyake et al., 2011) and atopy with circulating linoleate (Yen et al., 2008). A recent targeted metabolomic study investigated 64 lipid metabolites and reported Linoleate metabolism and Arachidonic acid metabolism as the top pathways albeit not statistically significantly associated with asthma control (McGeachie et al., 2015). Few studies associated linoleate with CCVD, although in general ω-6 fatty acids have long been believed to have pro-inflammatory effects in the cardiovascular system. An early *in vitro* study suggested that linoleate may lead to atherogenesis by NFκB signaling mediated vascular adhesion molecule-1 (VCAM-1) expression (Dichtl et al., 2002). The finding of linoleate metabolism was driven by our untargeted approach and confirmatory existing evidence. But for the aforementioned limitations in this approach, we may have missed other pathways, where evidence also exists.

6.4.5 CCVD specific MITM pathways

Glycosphingolipid metabolism was found as MITM pathway linking exposure to UFP and CCVD and annotation of sphingosine was confirmed as one of the modulated metabolites in this pathway. Sphingolipids are structural components of cell membrane but known to play a crucial role in apoptosis, cell growth, senescence, and cell cycle control (Yang et al., 2004). Sphingolipids in blood have been associated with cardiovascular diseases including

acute coronary syndrome (Pan et al., 2014) and myocardial infarction (Park et al., 2015). A recent clinical trial reported a strong association between blood sphingolipids and incident cardiovascular diseases (Wang et al., 2017). Sphingolipids have also been associated with asthma (Petrache and Berdyshev, 2016) in contrast to our findings. Perturbation of sphingolipid metabolism may be more relevant for allergic or child-onset asthma (Ono et al., 2015).

Carnitine shuttle pathway was identified as a MITM pathway linking exposure to NO_2 and CCVD. Carnitines facilitate the transport of long-chain fatty acids from the cytosol into the mitochondria and play an important role in fatty acid metabolism and carbohydrate utilization. The role of L-carnitine in CCVD has been extensively described, reporting protective effects of L-carnitine administration for various cardiovascular diseases including coronary artery disease, congestive heart failure, and hypertension (Ferrari et al., 2004). A recent meta-analysis of randomized controlled trials demonstrated the efficacy of L-carnitine against chronic heart failure (Song et al., 2017). In an experimental study in rats, inflammation accompanied with hypertension was attenuated by L-carnitine administration (Miguel-Carrasco et al., 2008). In this study, however, L-carnitine was associated with increased risk of CCVD.

6.4.6 Strengths and limitations

Strengths of our study include its prospective nature (nested in longitudinal cohorts), the individual assessment of exposure to air pollution, the accurate diagnoses for the outcomes, the agnostic nature of our metabolome-wide measurements, and the application of 'meet-in-the-middle' as a novel approach helping in the causal interpretation of the results. We focused on biological pathways that were associated with air pollution (mostly UFP) in both studies, supporting the robustness and replicability of our findings. Limita-

tions include the small sample size for metabolome-wide analyses; we focused on pathways enrichment but we were not able to identify single features associated with both air pollution and at least one disease due to the lack of statistical power. The demonstration of the MITM approach in the context of exposome research faces several challenges, one of which is the access to sufficiently powered, harmonized cohort data, and the other the availability of co-morbid phenotype information in the same study participants. These challenges limit the causal inference of the results presented. Yet, the added value of the current study is its role in capitalizing on the full potential of exposome research, namely identifying public health relevant pathways broadly associated with chronic diseases. Also, we used slightly different statistical methods (including the set of confounders) in the two studies, mainly due to the nature of the outcomes and the quite different estimation of exposure in the two studies. PNC and $PM_{2.5}$ exposure in EPIC and in Basel could not be derived from the same exposure measurement campaigns and models. With regard to PNC the Swiss model was more accurate than the Italian model. The inaccuracy of exposure measures limited our ability to find statistically significant associations and overlapping signals between the two studies and phenotypes. Unlike in SAPALDIA, we could not take into consideration geocoding quality in EPIC Italy as for the source of potential exposure misclassification. Unlike in EPIC Italy, we did not adjust for BMI in SAPALDIA. Air pollution exposure can increase the risk of obesity (Eze et al., 2015, Wei et al., 2016) and obesity may have a causal effect on asthma (Wenzel, 2012), therefore adjustment for BMI can lead to missing some signals. Given the smaller sample size and expected subtle effects, parsimony was more strongly sought in AOA MWAS. And a previous study observed less strong association between socioeconomic status and air pollution exposure in Switzerland than in Italy (Temam et al., 2017). Sensitivity analysis showed that the additional adjustment did not affect the results. Similarly, we group together

cardio- and cerebrovascular diseases and treated as a single outcome (CCVD) because they share common risk factors and etiology. The sensitivity analyses on cardiovascular outcomes only, confirmed the main results of this study. Finally, we did not consider the indoor air pollution, which may also have effects on AOA and CCVD. However, we speculated that the indoor air pollution would rather be an effect modifier than a confounder of ambient air pollution, i.e. we would conceivably have observed stronger effects of ambient air pollution in subjects less exposed to indoor air pollution.

6.5 Conclusions

In summary, we successfully applied a MITM approach in untargeted metabolomics to produce evidence of common and disease-specific pathway perturbations in the etiological relationship between air pollution exposure, AOA, and CCVD. Our findings need to be confirmed in future targeted and untargeted studies.

Acknowledgements

This work was supported by the grant FP7 of the European Commission "Enhanced exposure 36 assessment and omic profiling for high priority environmental exposures in Europe" (EXPOsOMICS grant 308610 to PV). The SAPALDIA cohort and biobank is funded by the Swiss National Science Foundation grant no 33CS30-148470/1&2 to NP. EPIC-Italy was financially supported by the Italian Association for Cancer Research (AIRC). We thank Mr. Vincent Cahais for coordinating the data storage and transfer. The authors declare they have no actual or potential competing financial interests.

Data availability

Raw metabolomic data that support the findings of this study are available from EXPOsOMICS but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of EXPOsOMICS.

6.6 Supplementary Material

6.6.1 Covariates acquisition

SAPALDIA

Height and weight were measured at SAPALDIA3 and body mass index (BMI) was calculated as weight in kg divided by squared height in meters. Education level was categorized into primary school ('low'), secondary or middle school or apprenticeship ('medium'), and college or university ('high'), based on the self-reported highest education at SAPALDIA3.

EPIC Italy

Smoking habits data were collected at study enrolment through the use of a questionnaire, and participants were categorized as 'never', 'former' and 'current' smokers. Height and weight were measured at enrolment with a standardized protocol, and body mass index (BMI) was calculated as the ratio between weight in kg and squared height in meters, treated as a continuous variable. The self-reported highest educational level was categorized as primary school or none ('low'), vocational or another secondary school ('medium'), and university or vocational postsecondary school ('high') and used as a proxy for the socio-economic condition.

6.6.2 Metabolome analyses: laboratory and pre-processing procedures

Sample preparation

Serum samples were prepared by mixing 20 μL aliquot of a sample with 200 μL of acetonitrile, and filtering the precipitate with 0.2 μm Captiva ND plates (Agilent Technologies). The filtrate was collected into a polypropylene well plate that was sealed with a Rapid

EPS well plate sealing tape (BioChromato) and kept frozen until analysis. Quality control (QC) samples were prepared using a sample pool that was prepared by combining small aliquots of the study samples.

Sample analysis

Samples were analyzed in randomized order as a single batch with a UHPLC-QTOF-MS system (Agilent Technologies) consisting of a 1290 Binary LC system, a Jet Stream electrospray ionization (ESI) source, and a 6550 QTOF mass spectrometer. Autosampler tray was kept refrigerated and 2 μL of the sample solution was injected on an ACQUITY UPLC HSS T3 column (2.1 $\times$ 100mm, 1.8 μm; Waters). Column temperature was 45 °C and mobile phase flow rate 0.4 ml/min, consisting of ultrapure water and LC-MS grade methanol, both containing 0.05 % (v/v) of formic acid. The gradient profile was as follows: 0-6 min: 5% $\rightarrow$ 100% methanol, 6-10.5 min: 100% methanol, 10.5-13 min: 5% methanol.

The mass spectrometer was operated in positive polarity using the following conditions: drying gas (nitrogen) temperature 175 °C and flow 12 L/min, sheath gas temperature 350 °C and flow 11 L/min, nebulizer pressure 45 psi, capillary voltage 3500 V, nozzle voltage 300 V, and fragmentor voltage 175 V. Data acquisition was performed using 2 GHz extended dynamic range mode across a mass range of 50-1000. Scan rate was 1.67 Hz and data acquisition was in centroid mode. Continuous mass axis calibration was performed by monitoring two reference ions from an infusion solution throughout the runs (m/z 121.050873 and m/z 922.009798). Data was acquired using MassHunter Acquisition B.05.01 (Agilent Technologies). The analytical run included all study samples in random order, intervened after every 12 injections with a QC sample to monitor instrument performance and sample stability.

Data preprocessing

Preprocessing of the acquired data was performed using Qualitative Analysis B.06.00, DA Reprocessor, and Mass Profiler Professional 12.1 software (Agilent Technologies). Recursive feature finding was employed to find compounds as singly charged proton adducts $[M+H]^+$, by using data from all study samples. The initial processing of the data was performed using Qualitative Analysis with MFE algorithm set to small molecules. Threshold values for mass and chromatographic peak heights were 1500 and 8000 counts, respectively. A single mass peak was considered a feature if neutral mass could be calculated and peak spacing tolerance for isotope peaks was 0.0025 m/z plus 7 ppm, with the isotope model set to common organic molecules. Only singly charged ions were included.

After the initial feature finding, the compounds that existed in at least 2% of all the samples were combined into a single list, using 0.1 min retention time and 15 ppm +2 mDa mass windows for alignment. The resulting list was used as a target for the recursive feature extraction of the data, which was performed using an Agilent FBF algorithm with match tolerance for the compound mass and retention time set at ±10 ppm and ±0.05 min. Multiply charged ions were excluded, ion species was limited to $[M+H]^+$, and chromatographic peak height threshold was 2000 counts. Any number of ions associated with each compound was allowed.

6.6.3 Link and variance function in GLM

In EPIC Italy, feature intensities were Box-Cox transformed before regression. In SAPALDIA, the best link was sought for each feature by searching for a link function in the power family leading to the highest log-likelihood among six links (power(2), identity, sqrt, log, inverse, and 1/mu^2). The best variance function was chosen among Gaussian, inverse Gaussian, and Gamma, based on AIC. P-values were obtained from likelihood-ratio tests

comparing the model with the air pollution related variables to the model without them. A concern about this approach was that it may have caused underestimation of p-values because link and variance functions were selected using the same data as for inference, rather than set *a priori*, and this additional uncertainty was not taken into account in p-value computation. First, we conducted simulation studies where we compared the p-values computed from models fitted with the best link and variance function with the p-values from models fitted with identity link and Gaussian distribution, respectively. We observed few additional false positives due to selection of link or variance functions (data not shown). Second, even though this approach leads to underestimation of p-values, it does not affect the comparability between models. And increase in false positives is of less concern in this study as we do not aim to report robust biomarkers but to seek MITM evidence.

Using various link and variance functions made the resulting coefficients incomparable. Therefore we computed semi-partial pseudo-R^2 as a measure of effect size.

6.6.4 Semi-partial pseudo-R^2 coefficient

Suppose we are interested in comparing two nested models, a full model $M(\beta_F) : \mu_F = X_F\beta_F$ and a reduced model $M(\beta_R) : \mu_R = X_R\beta_R$, where β_F is a regression parameter vector of size k and β_R of size$r$, $k > r$, and $\beta_R \subset \beta_F$. We can partition β_F as $\beta_F = [\begin{smallmatrix}\beta_R \\ \beta_C\end{smallmatrix}]$ and X_F as $X_F = [X_R\ X_C]$.

Let us define the following quadratic forms (or sum of squares, if in scalar form):

Total Sum of Squares:

$$SST = y^T[I - \frac{1}{n}J]y = \sum_{i=1}^{n}(y_i - \bar{y})^2 \tag{1}$$

Sum of Squares due to (marginal) Regression on X_R:

$$SSReg_R = y^T[H_R - \frac{1}{n}J]y = \sum_{i=1}^{n}(\hat{\mu}_{Ri} - \bar{y})^2 \tag{2}$$

Sum of Squares due to (conditional) Regression on X_C adjusting for X_R:

$$SSReg_{R|F} = y^T[H_F - H_R]y = \sum_{i=1}^{n}(\hat{\mu}_{Fi} - \hat{\mu}_{Ri})^2 \tag{3}$$

Residual Sum of Squares (of the full model):

$$SSE_F = y^T[I - H_F]y = \sum_{i=1}^{n}(y_i - \hat{\mu}_{Fi})^2 \tag{4}$$

where $H_F = X_F(X_F^T X_F)^{-1}X_F^T$ is the hat-matrix of the full model and $H_R = X_R(X_R^T X_R)^{-1}X_R^T$ is the hat-matrix of the reduced model.

The fundamental theorem of regression becomes:

$$SST = SSReg_R + SSReg_{R|F} + SSE_F \tag{5}$$

These quadratic forms and the identity (5) can be used to compute the *coefficient of semi-partial squared correlation* as a measure of "fraction of variation explained" by X_C:

$$R^2_{F|R;sp} = \frac{y^T[H_F - H_R]y}{y^T[I - \frac{1}{n}J]y} = \frac{SSReg_{R|F}}{SST} = 1 - \frac{SSE + SSReg_R}{SST} \tag{6}$$

Extending $R^2_{F|R;sp}$ to generalized linear models (GLMs) is difficult because there is no exact decomposition like (5) for GLMs in general. This lack of obvious extension of $R^2_{F|R;sp}$ to GLMs has opened the way to many different generalizations, which collectively go under the term of "*pseudo-*R^2" measures.

To introduce notation, let $M(\beta_F) : \mu_F = h(X_F\beta_F)$ and $M(\beta_R) : \mu_R = h(X_R\beta_R)$ be a full and a reduced generalized linear model respectively, with $\mu = E[y]$, and let β_F and X_F be partitioned as before. The response variable Y is assumed to belong to the Natural Exponential Class of distributions $Y|x \sim NEC(\mu, \phi)$ with ϕ a dispersion parameter.

Furthermore, let:

$$l(\beta_F; y, X_F) = \log(L(\beta_F; y, X_F)) \text{ the log-likelihood of the full model} \quad (7)$$

$$l(\beta_R; y, X_R) = \log(L(\beta_R; y, X_R)) \text{ the log-likelihood of the reduced model} \quad (8)$$

$$l(\beta_0; y, 1_n) = \log(L(\beta_F; y, 1_n)) \text{ the log-likelihood of the null model} \quad (9)$$

$$l(\beta_{sat}; y, X_{sat}) = \log(L(\beta_{sat}; y, X_{sat})) \text{ the log-likelihood of the saturated model} \quad (10)$$

Estimation of all models is carried out through Maximum Likelihood, and therefore substitution of the unknown parameters β_F, β_R, β_0, and β_{sat} with their MLE's $\hat{\beta}_F$, $\hat{\beta}_R$, $\hat{\beta}_0$, and $\hat{\beta}_{sat}$ provides the maximized versions of the log-likelihoods (7), (8), (9), and (10): $l(\beta_F; y, X_F)$, $l(\beta_R; y, X_R)$, $l(\beta_0; y, 1_n)$, and $l(\beta_{sat}; y, X_{sat})$. With these maximized log-likelihoods, it is possible to build up the following quantities:

$$D_0 = 2(l(\hat{\beta}_{sat}; y, X_{sat}) - l(\hat{\beta}_0; y, 1_n)) \text{ deviance of the null model} \quad (11)$$

$$D_F = 2(l(\hat{\beta}_{sat}; y, X_{sat}) - l(\hat{\beta}_F; y, X_F)) \text{ deviance of the full model} \quad (12)$$

$$D_{0|F} = 2(l(\hat{\beta}_F; y, X_F) - l(\hat{\beta}_0; y, 1_n)) \text{ extra-deviance of the null vs full model} \quad (13)$$

$$D_{0|R} = 2(l(\hat{\beta}_R; y, X_R) - l(\hat{\beta}_0; y, 1_n)) \text{ extra-deviance of the null vs reduced model} \quad (14)$$

$$D_{R|F} = 2(l(\hat{\beta}_F; y, X_F) - l(\hat{\beta}_R; y, X_R)) \text{ extra-deviance of the reduced vs full model} \quad (15)$$

As correctly observed by Cameron and Windmeijr (Cameron and Windmeijer, 1997), (11), (12), (13), (14), and (15) are the exact analog for GLMs of SST, SSE_F, $SSReg_F$, $SSReg_R$, and $SSReg_{R|F}$ in linear models. It is trivial to show that exact additive decomposition, analog to (5), holds for (11), (12), (13), (14), and (15):

$$D_0 = D_{0|F} + D_F \quad (16)$$

$$D_0 = D_{0|R} + D_{R|F} + D_F \quad (17)$$

Cameron and Windmeijr(Cameron and Windmeijer, 1997) used (16) to propose their *multiple pseudo-R^2*:

$$R^2_{CW} = \frac{D_{0|F}}{D_0} = 1 - \frac{D_F}{D_0} \tag{18}$$

Although Cameron and Windmeijr did not propose explicitly generalization of the semi-partial R^2 coefficients used in linear models, decomposition (17) immediately provides such generalization; by analogy with (6), it is natural to define the *semi-partial pseudo-R^2 coefficient*:

$$R^2_{CW;F|R;sp} = \frac{D_{R|F}}{D_0} = 1 - \frac{D_{0|R} + D_F}{D_0} \tag{19}$$

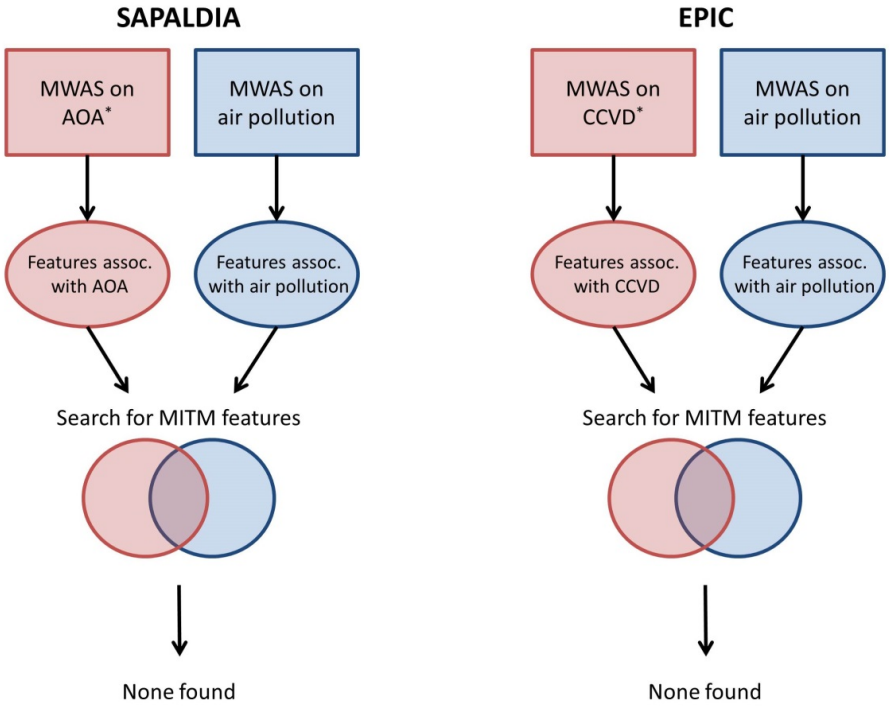
This *semi-partial pseudo-R^2 coefficient* measures the relative contribution of the variables in X_C, air pollution exposure variables in this study, to the reduction in variation of the response Y; it equals the fraction of total variation in Y which is explained by X_C, when its variables are added to those in X_R, already included in the model.

6.6.5 Confirmation of chemical identities from Mummichog

We examined if the ion Mummichog assigned to a metabolite was theoretically feasible, if the proposed charge state and retention time were plausible, and if the chromatographic peaks were of adequate quality. Final confirmation was based on comparisons of the retention time and fragmentation spectra against those of an authentic chemical standard.

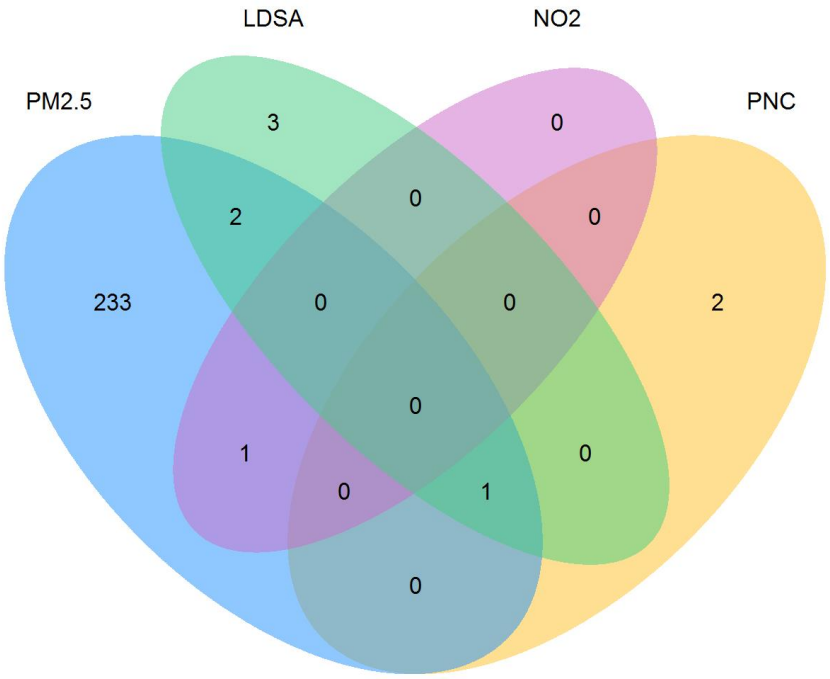
6.6.6 Supplementary figures

Figure 6.S1. Search for the MITM pathways



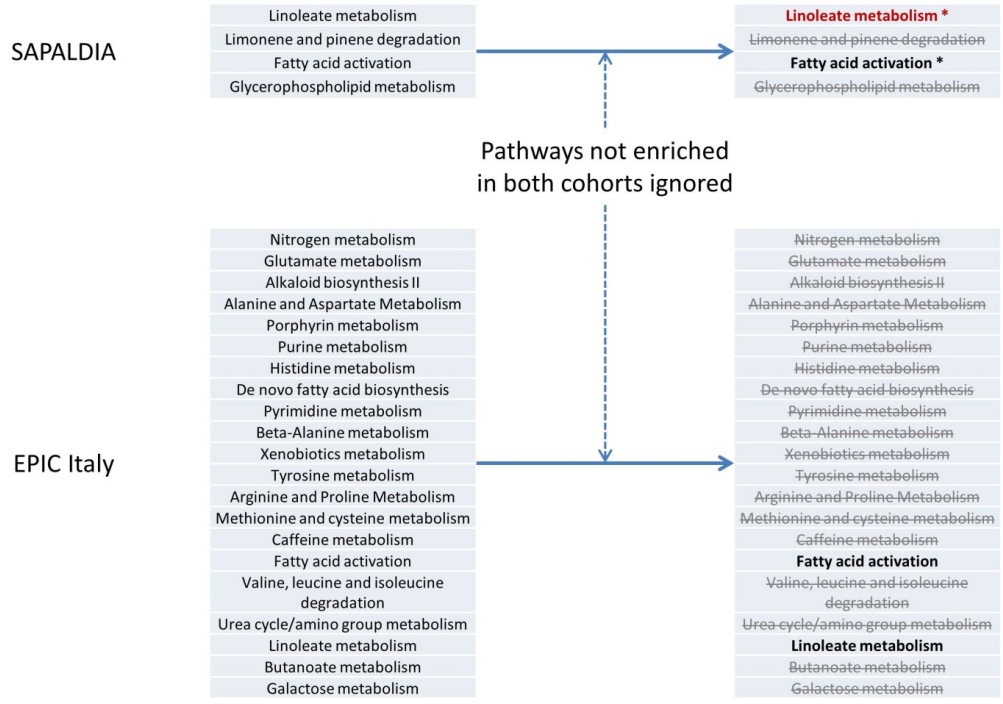
*Adjusted for the corresponding air pollutant; **by excluding the pathways not enriched in the other cohort.

Figure 6.S2. Venn diagram of air pollution MWAS signals in SAPALDIA



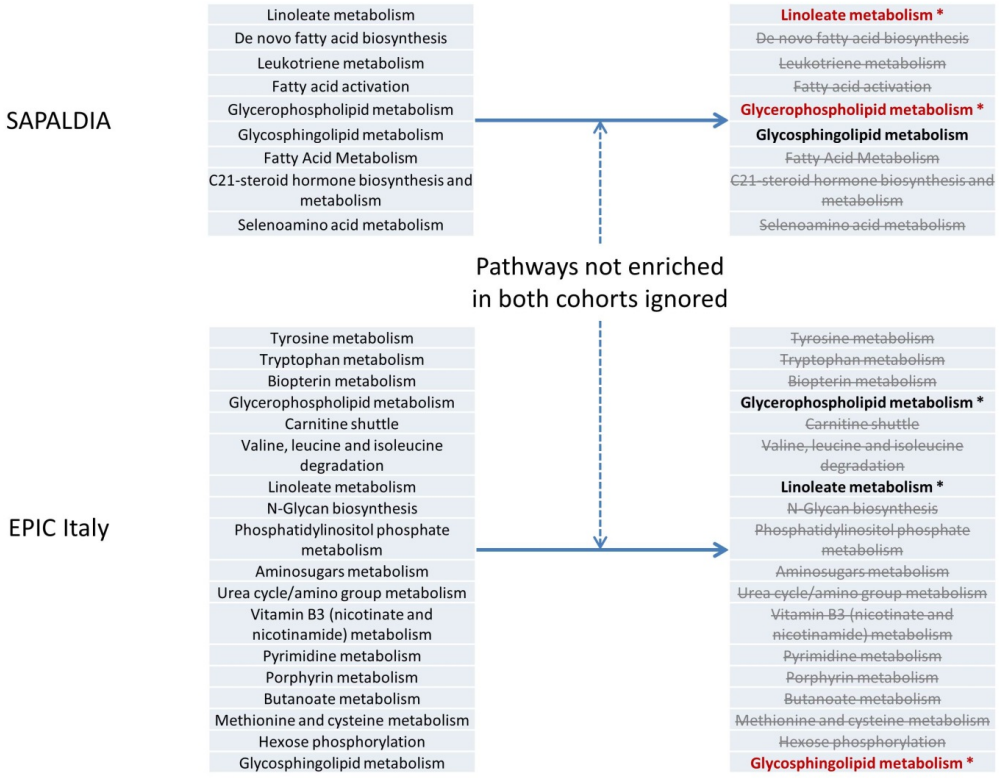
The number of signals with p-value smaller than 0.05 after Benjamini-Hochberg correction from MWAS either on $PM_{2.5}$, PNC, LDSA, or NO_2 in SAPALDIA, adjusted for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 hours, their multiplicative interaction terms with fasting time, and geocoding quality.

Figure 6.S3. List of pathways enriched by Mummichog on $PM_{2.5}$ MWAS results



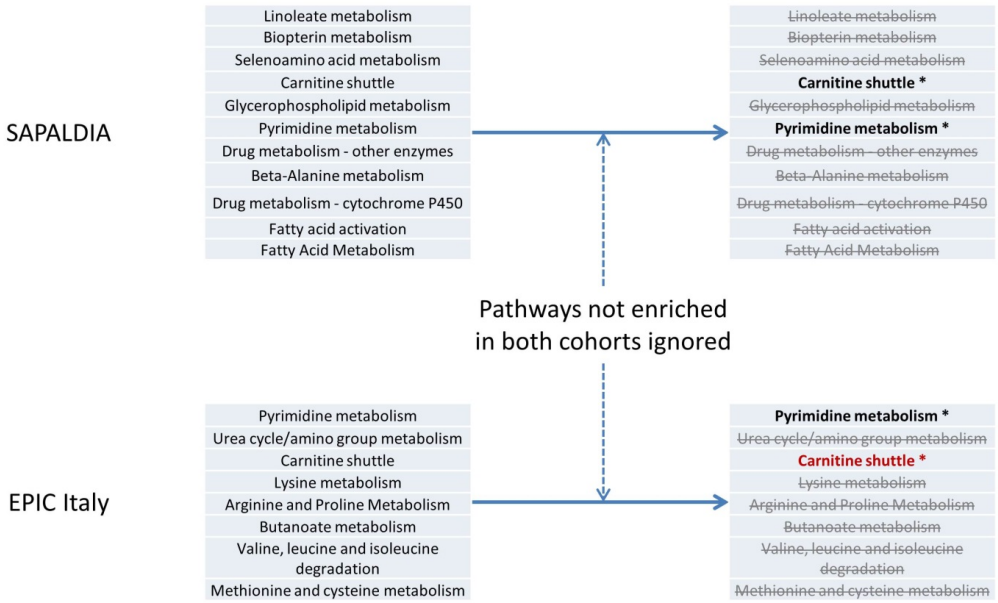
The pathways enriched for $PM_{2.5}$ in both cohorts are in bold; MITM pathways in red. * depicts the pathways confirmed in a sensitivity analysis using the 5^{th} percentile p-value as cut-off.

Figure 6.S4. List of pathways enriched by Mummichog on UFP MWAS results



The pathways enriched for UFP in both cohorts are in bold; MITM pathways in red. * depicts the pathways confirmed in a sensitivity analysis using the 5^{th} percentile p-value as cut-off. In SAPALDIA, the pathways enriched either for PNC or LDSA are listed.

Figure 6.S5. List of pathways enriched by Mummichog on NO_2 MWAS results



The pathways enriched for NO_2 in both cohorts are in bold; MITM pathways in red. * depicts the pathways confirmed in a sensitivity analysis using the 5^{th} percentile p-value as cut-off.

6.6.7 Supplementary tables

Table 6.S1. SAPALDIA sample characteristics – AOA

	AOA cases			Controls		
	Nested study	**SAPALDIA3**	**p**[a]	**Nested study**	**SAPALDIA3**	**p**[a]
N	139	374	-	196	3755	-
Age [year]	59.4 (19.4)	58.8 (18.2)	0.98	57.1 (15.8)	60.4 (17.6)	< 0.001
Female	87 (63%)	230 (61%)	0.90	101 (52%)	1920 (51%)	0.97
BMI [kg/m^2	25.7 (6.4)	26.1 (6.3)	0.31	24.4 (4.8)	25.7 (5.6)	< 0.001
Smoking						
Current[b]	-	51 (14%)	-	-	582 (15%)	-
Former[c]	54 (39%)	168 (45%)		62 (32%)	1401 (37%)	
Never	85 (61%)	155 (41%)		134 (68%)	1772 (47%)	
Education level[d]						
Low	3 (2%)	20 (5%)		2 (1%)	205 (5%)	
Middle	86 (62%)	239 (64%)	0.20	121 (62%)	2463 (66%)	0.0027
High	50 (36%)	115 (31%)		73 (37%)	1086 (29%)	
PM$_{2.5}$[e] [μg/m^3]						
t = 1	14.6 (1.9)	14.7 (1.4)	0.49	14.3 (1.7)	14.5 (1.6)	0.048
t = 2	14.7 (2.2)	14.9(1.6)	0.40	14.4 (1.8)	14.6 (1.8)	0.034
t = 3	14.6 (2.8)	14.7 (1.9)	0.44	14.3 (2.2)	14.5 (2.1)	0.022
t = 4	16.0 (2.8)	16.1 (2.4)	0.30	15.6 (2.2)	15.9 (2.4)	0.091
t = 5	17.3 (2.3)	17.5 (1.7)	0.53	17.1 (2.1)	17.2 (2.0)	0.13
t = 6	16.5 (2.4)	16.5 (1.7)	0.98	16.0 (2.2)	16.2 (2.0)	0.051
t = 7	16.8 (3.4)	17.0 (2.3)	0.47	16.2 (3.0)	16.6 (3.0)	0.097
PNC[f] [particles/cm^3]	13,418 (6,376)	12,463 (5,969)	0.40	9,660 (7,970)	10,585 (8,540)	0.021
LDSA[f] [μ m^2/cm^3]	33.9 (16.1)	32.0 (12.8)	0.20	27.1 (16.3)	29.6 (19.9)	0.0078
NO$_2$ [μ g/m^3]	25.0 (14.3)	25.4 (12.2)	0.57	21.6 (10.9)	23.5 (12.2)	0.054
Current asthma[g]	73 (53%)	198 (53%)	0.92	-	-	-
Current medication	45 (32%)	135 (36%)	0.45	-	-	-

Data are presented as count (%) or median (interquartile range). $PM_{2.5}$: annual mean estimates derived from the PolluMap in 2010; PNC and LDSA: biennial mean estimates derived from a SAPALDIA multi-area LUR in 2011/2012; NO_2: annual mean estimates derived from a European LUR in 2010. [a]For categorical variables p-values derived from χ^2 test and for continuous variables from Wilcoxon rank sum tests. [b]Current smokers were excluded for the metabolomics analysis. [c]Former smokers had not smoked for at least 10 years before blood was drawn. [d]Education level low: primary school; middle: secondary/middle school or apprenticeship; high: college or university. [e]365 days average t-1 years before the examination. [f]Number of observation smaller due to limited availability of UFP estimates. [g]Current asthma was defined as either having reported asthma attack in the last 12 months or currently taking asthma medication.

Table 6.S2. EPIC Italy sample characteristics – CCVD

	CCVD cases			Controls		
	Nested study	EPIC Turin/Varese	p[a]	Nested study	EPIC Turin/Varese	p[a]
N	166	948	-	155	18,034	-
Center						
Turin	71 (43%)	418 (44%)	0.81	73 (47%)	8,335 (46%)	0.89
Varese	95 (57%)	530 (56%)		82 (53%)	9,699 (54%)	
Age [years]	56.16 (9.56)	56.54 (10.95)	0.92	56.55 (9.44)	50.37 (12.01)	< 0.001
Female	107 (64%)	450 (47%)	<0.001	95 (61%)	11,566 (64%)	0.51
BMI [kg/m^2]	26.34 (4.91)	26.22 (4.89)	0.27	26.09 (4.91)	24.99 (4.81)	0.002
Smoking						
Current[b]	-	287 (30%)	-	-	3,983 (22%)	-
Former[c]	52 (31%)	260 (28%)		54 (35%)	4,811 (27%)	
Never	114 (69%)	401 (42%)		101 (65%)	9,240 (51%)	
Education level[d]						
Low	103 (69%)	581 (61%)	0.86	84 (56%)	9,329 (52%)	< 0.001
Middle	48 (32%)	298 (32%)		44 (29%)	6,810 (38%)	
High	12 (8%)	66 (7%)		22 (15%)	1,877 (10%)	
Treatments						
Hypertension	75 (45%)	346 (36%)	0.04	47 (30%)	3885 (21%)	0.01
Hyperlipidaemia	69 (41%)	365 (38%)	0.51	44 (28%)	4,457 (25%)	0.33
$PM_{2.5}$ [μg/m^3]	21.27 (2.19)	21.76 (2.57)	0.39	21.27 (2.16)	21.29 (1.82)	0.93
PNC[e] [particles/cm^3]	13,283 (2,335)	13,244 (1,646)	0.55	13,150 (2,497)	13,212 (1,910)	0.91
NO_2 [μg/m^3]	55.15 (14.95)	53.55 (16.54)	0.38	54.67 (16.48)	52.28 (20.33)	0.11

Data are presented as count (%) or median (interquartile range). $PM_{2.5}$: annual mean estimates derived a European LUR in 2010; PNC: annual mean estimates derived from a local LUR in 2014/2015; NO_2: annual mean estimates derived from a European LUR in 2010. [a]For categorical variables p-values derived from χ^2 test and for continuous variables from Wilcoxon rank sum tests. [b]Current smokers were excluded for the metabolomics analysis. [c]Former smokers had not smoked for at least 1 year before blood was drawn. [d]Education level: low (primary school or none), middle (vocational or another secondary school), and high (university or vocational postsecondary school). [e]Number of observation smaller due to limited availability of UFP estimates.

Table 6.S3. 10^{th} and 5^{th} percentile of the nominal p-values used as cut-off for Mummichog

	MWAS	10^{th} percentile of p-values	5^{th} percentile of p-values
SAPALDIA	**$PM_{2.5}$**	0.014	0.0041
	PNC	0.046	0.020
	LDSA	0.086	0.039
	NO_2	0.17	0.088
	AOA	0.12	0.062
	AOA, adj. for $PM_{2.5}$	0.11	0.058
	AOA, adj. for PNC	0.11	0.065
	AOA, adj. for LDSA	0.11	0.065
	AOA, adj. for NO_2	0.12	0.062
EPIC Italy	**$PM_{2.5}$**	0.10	0.05
	PNC	0.11	0.06
	NO_2	0.08	0.04
	CCVD	0.14	0.07
	CCVD, adj. for $PM_{2.5}$	0.12	0.07
	CCVD, adj. for PNC	0.10	0.06
	CCVD, adj. for NO_2	0.13	0.08

Table 6.S4. Association of UFP with AOA or CCVD, independent of other air pollutants

UFP metric		**AOA**		**CCVD**	
		OR[a] [95% CI]	**OR**[b] [95% CI]	**OR**[c] [95% CI]	**HR**[d] [95% CI]
PNC[e]	adj. for $PM_{2.5}$	1.75 [0.86, 3.55]	1.34 [0.80, 2.25]	1.08 [0.58, 1.99]	1.21 [1.00, 1.47]
	adj. for NO_2	2.21 [1.15, 4.26]	1.13 [0.78, 1.64]	1.11 [0.60, 2.04]	1.20 [0.98, 1.46]
LDSA[f]	adj. for $PM_{2.5}$	2.67 [1.13, 6.33]	1.58 [0.95, 2.63]	-	-
	adj. for NO_2	1.88 [1.18, 3.01]	1.15 [0.78, 1.69]	-	-

Note: Odds ratios are reported for all cross-sectional analyses (meet-in-the-middle/metabolome subsample) while hazard ratio is reported for the longitudinal analysis on larger CCVD samples. [a]odds ratio adjusted for age, sex, education level, BMI, and study area as random effect (N=188). [b]odds ratio adjusted for age, sex, education level, BMI, and study area as random effect (N=1,555). [c]odds ratio adjusted for age, center of recruitment, sex, BMI, smoking status, and educational level (N=144). [d]hazard ratio adjusted for age, center of recruitment, sex, BMI, smoking status, and educational level (N=8,753). [e]per increase by 5000 particles/cm^3 in biennial (SAPALDIA) or annual (EPIC Italy) mean PNC. [f]per increase by 10 $\mu m^2/cm^3$ in biannual mean LDSA.

Table 6.S5. Multipollutant model results for association of air pollutants with AOA and CCVD

Air pollution metric		AOA		CCVD	
		OR [95% CI] (N = 188)	**OR** [95% CI] (N = 1555)	**OR** [95% CI] (N = 144)	**HR** [95% CI] (N = 8753)
Model using PNC	$PM_{2.5}$[a]	1.03 [0.11, 9.86]	0.90 [0.25, 3.25]	0.77 [0.45, 1.33]	0.42 [0.15, 1.14]
	PNC[b]	2.21 [0.86, 5.66]	1.17 [0.69, 1.99]	1.46 [0.71, 3.00]	1.24 [1.03, 1.50]
	NO_2[c]	0.76 [0.37, 1.56]	1.39 [0.92, 2.11]	0.82 [0.54, 1.25]	1.12 [0.97, 1.29]
Model using LDSA	$PM_{2.5}$[a]	0.19 [0.01, 4.30]	0.55 [0.10, 2.98]	-	-
	LDSA[d]	2.92 [1.13, 7.52]	1.38 [0.79, 2.43]	-	-
	NO_2[c]	0.87 [0.48, 1.57]	1.27 [0.75, 2.15]	-	-

Note: Odds ratios are reported for all cross-sectional analyses (meet-in-the-middle/metabolome subsample) while hazard ratio is reported for the longitudinal analysis on larger CCVD samples. AOA status was regressed on biennial mean $PM_{2.5}$, biennial mean UFP, and annual mean NO_2, after adjustment for age, sex, education level, BMI, and study area as random effect; Geocoding quality was either included as effect modifier or the observations with non-perfect geocoding quality were excluded when the model failed to converge. CCVD status was regressed on annual mean $PM_{2.5}$, UFP, and NO_2, after adjustment for age, center of recruitment, sex, BMI, smoking status, and educational level. [a]per 5 μg/m3 increase in biennial (SAPALDIA) or annual (EPIC Italy) mean $PM_{2.5}$. [b]per increase by 5000 particles/cm^3 in biennial (SAPALDIA) or annual (EPIC Italy) mean PNC. [c]per increase by 10 $\mu g/m^3$ in annual mean NO_2. [d]per increase by 10 $\mu m^2/cm^3$ in biennial mean LDSA.

Table 6.S6. Pathways associated to $PM_{2.5}$ in SAPALDIA

Pathway	Overlap size	Pathway size	p-value
Linoleate metabolism	17	21	0.0007
Limonene and pinene degradation	6	6	0.0008
Fatty acid activation	10	21	0.0054
Glycerophospholipid metabolism	14	36	0.0239

Mummichog pathway enrichment test on the results from $PM_{2.5}$ MWAS adjusted for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 hours, their multiplicative interaction terms with fasting time, and geocoding quality.

Table 6.S7. Pathways associated to $PM_{2.5}$ in EPIC

Pathway	Overlap size	Pathway size	p-value
Nitrogen metabolism	4	4	0.0009
Glutamate metabolism	6	10	0.0011
Alkaloid biosynthesis II	4	6	0.0018
Alanine and Aspartate Metabolism	7	16	0.0022
Porphyrin metabolism	8	20	0.0026
Purine metabolism	10	27	0.0027
Histidine metabolism	7	17	0.0029
De novo fatty acid biosynthesis	6	14	0.0033
Pyrimidine metabolism	10	28	0.0035
Beta-Alanine metabolism	5	12	0.0057
Xenobiotics metabolism	14	45	0.0061
Tyrosine metabolism	22	76	0.0078
Arginine and Proline Metabolism	9	29	0.0114
Methionine and cysteine metabolism	10	33	0.0120
Caffeine metabolism	4	11	0.0178
Fatty acid activation	5	15	0.0180
Valine, leucine and isoleucine degradation	6	19	0.0182
Urea cycle/amino group metabolism	12	43	0.0204
Linoleate metabolism	6	20	0.0249
Butanoate metabolism	5	16	0.0255
Vitamin B9 (folate) metabolism	3	8	0.0292
Galactose metabolism	7	25	0.0324

Mummichog pathway enrichment test on the results from $PM_{2.5}$ MWAS adjusted for age at recruitment, center of recruitment, sex, BMI, smoking status, and education level.

Table 6.S8. Pathways associated to PNC in SAPALDIA

Pathway	Overlap size	Pathway size	p-value
Linoleate metabolism	12	21	0.0007
De novo fatty acid biosynthesis	9	22	0.0014
Leukotriene metabolism	17	51	0.0014
Fatty acid activation	8	21	0.0023
Glycerophospholipid metabolism	12	36	0.0023
Glycosphingolipid metabolism	8	26	0.0079
Fatty Acid Metabolism	5	14	0.0083
C21-steroid hormone biosynthesis and metabolism	20	86	0.0396

Mummichog pathway enrichment test on the results from PNC MWAS adjusted for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 hours, their multiplicative interaction terms with fasting time, and geocoding quality.

Table 6.S9. Pathways associated to PNC in EPIC

Pathway	Overlap size	Pathway size	p-value
Tyrosine metabolism	35	76	0.0010
Tryptophan metabolism	24	58	0.0011
Biopterin metabolism	7	11	0.0012
Glycerophospholipid metabolism	13	35	0.0022
Carnitine shuttle	7	19	0.0063
Valine, leucine and isoleucine degradation	7	19	0.0063
Linoleate metabolism	7	20	0.0084
N-Glycan biosynthesis	4	10	0.0142
Phosphatidylinositol phosphate metabolism	4	10	0.0142
Aminosugars metabolism	5	14	0.0146
Urea cycle/amino group metabolism	12	43	0.0193
Vitamin B3 (nicotinate and nicotinamide) metabolism	5	15	0.0208
Pyrimidine metabolism	8	28	0.0256
Porphyrin metabolism	6	20	0.0275
Butanoate metabolism	5	16	0.0291
Methionine and cysteine metabolism	9	33	0.0315
Hexose phosphorylation	4	12	0.0316
Glycosphingolipid metabolism	6	21	0.0367

Mummichog pathway enrichment test on the results from PNC MWAS adjusted for age at recruitment, center of recruitment, sex, BMI, smoking status, and education level.

Table 6.S10. Pathways associated to LDSA in SAPALDIA

Pathway	Overlap size	Pathway size	p-value
Linoleate metabolism	11	21	0.0009
Fatty acid activation	11	21	0.0009
Glycerophospholipid metabolism	15	36	0.0011
De novo fatty acid biosynthesis	9	22	0.0024
Leukotriene metabolism	15	51	0.0144
Selenoamino acid metabolism	4	10	0.0151
Fatty Acid Metabolism	5	14	0.0165

Mummichog pathway enrichment test on the results from LDSA MWAS adjusted for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 hours, their multiplicative interaction terms with fasting time, and geocoding quality.

Table 6.S11. Pathways associated to NO_2 in SAPALDIA

Pathway	Overlap size	Pathway size	p-value
Linoleate metabolism	10	21	0.0016
Biopterin metabolism	6	13	0.0056
Selenoamino acid metabolism	5	10	0.0059
Carnitine shuttle	10	26	0.0063
Glycerophospholipid metabolism	13	36	0.0068
Pyrimidine metabolism	12	33	0.0074
Drug metabolism - other enzymes	7	18	0.0119
Beta-Alanine metabolism	4	9	0.0199
Drug metabolism - cytochrome P450	13	42	0.0346
Fatty acid activation	7	21	0.0371

Mummichog pathway enrichment test on the results from NO_2 MWAS adjusted for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 hours, their multiplicative interaction terms with fasting time, and geocoding quality.

Table 6.S12. Pathways associated to NO_2 in EPIC

Pathway	Overlap size	Pathway size	p-value
Pyrimidine metabolism	8	28	0.0035
Urea cycle/amino group metabolism	11	43	0.0036
Carnitine shuttle	6	19	0.0040
Lysine metabolism	7	24	0.0041
Arginine and Proline Metabolism	8	29	0.0042
Butanoate metabolism	5	16	0.0065
Valine, leucine and isoleucine degradation	5	19	0.0158
Methionine and cysteine metabolism	7	33	0.0326

Mummichog pathway enrichment test on the results from NO_2 MWAS adjusted for age at recruitment, center of recruitment, sex, BMI, smoking status, and education level.

Table 6.S13. Overlapping features in the pathways associated with air pollution in both cohorts

Air pollution	Pathway	Overlap size		#Features with Δm/z < 10 ppm	
		SAPALDIA	EPIC Italy	Total[a]	ΔRT < 0.1 min[b]
$PM_{2.5}$	Linoleate metabolism	17	6	1	0
	Fatty acid activation	10	5	1	0
PNC	Linoleate metabolism	12	7	1	0
	Glycerophospholipid metabolism	12	13	3	1[d]
	Glycosphingolipid metabolism	8	6	0	0
LDSA[c]	Linoleate metabolism	11	7	1	0
	Glycerophospholipid metabolism	15	13	8	2[d]
NO_2	Carnitine shuttle	10	6	1	0
	Pyrimidine metabolism	12	8	2	0

[a]total number of features with absolute difference in m/z smaller than 10 ppm between SAPALDIA and EPIC Italy. [b]out of the number of features with Δm/z < 10 ppm, the number of features with absolute difference in retention time smaller than 0.1 min between SAPALDIA and EPIC Italy. [c]LDSA in SAPALDIA compared with PNC in EPIC Italy. [d]Incorrect annotation by Mummichog.

Table 6.S14. Pathways associated to AOA adjusted for $PM_{2.5}$

Pathway	Overlap size	Pathway size	p-value
Tryptophan metabolism	25	54	0.0007
Biopterin metabolism	9	13	0.0007
Vitamin B6 (pyridoxine) metabolism	4	6	0.0018
TCA cycle	4	8	0.0043
Hexose phosphorylation	5	12	0.0051
Glutathione Metabolism	4	10	0.0104
Lysine metabolism	7	22	0.0109
Leukotriene metabolism	14	51	0.0136
Aminosugars metabolism	8	29	0.0243
Drug metabolism - cytochrome P450	11	42	0.0268
Linoleate metabolism	**6**	**21**	**0.0275**
Butanoate metabolism	5	17	0.0302
De novo fatty acid biosynthesis	6	22	0.0365
Aspartate and asparagine metabolism	11	44	0.0403
Drug metabolism - other enzymes	5	18	0.0411
Fatty Acid Metabolism	4	14	0.0484
Urea cycle/amino group metabolism	8	32	0.0495
Vitamin E metabolism	8	32	0.0495

Mummichog pathway enrichment test on the results from AOA MWAS adjusted for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 hours, their multiplicative interaction terms with fasting time, and $PM_{2.5}$. The MITM pathway is in bold.

Table 6.S15. Pathways associated to AOA adjusted for PNC

Pathway	Overlap size	Pathway size	p-value
Biopterin metabolism	8	13	0.0008
Selenoamino acid metabolism	5	10	0.0027
Linoleate metabolism	**8**	**21**	**0.0032**
Fatty acid activation	8	21	0.0032
De novo fatty acid biosynthesis	8	22	0.0042
Xenobiotics metabolism	16	57	0.0104
Leukotriene metabolism	14	51	0.0159
Glycine, serine, alanine and threonine metabolism	8	28	0.0238
Tryptophan metabolism	14	54	0.0299
Drug metabolism - cytochrome P450	11	42	0.0340
C21-steroid hormone biosynthesis and metabolism	21	86	0.0472

Mummichog pathway enrichment test on the results from AOA MWAS adjusted for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 hours, their multiplicative interaction terms with fasting time, and PNC. The MITM pathway is in bold.

Table 6.S16. Pathways associated to AOA adjusted for LDSA

Pathway	Overlap size	Pathway size	p-value
Biopterin metabolism	8	13	0.0008
Glycine, serine, alanine and threonine metabolism	9	28	0.0053
Fatty acid activation	7	21	0.0066
Xenobiotics metabolism	16	57	0.0067
De novo fatty acid biosynthesis	7	22	0.0087
Squalene and cholesterol biosynthesis	11	41	0.0161
Tryptophan metabolism	14	54	0.0174
Linoleate metabolism	**6**	**21**	**0.0221**
C21-steroid hormone biosynthesis and metabolism	21	86	0.0258
Carnitine shuttle	7	26	0.0264
Drug metabolism - other enzymes	5	18	0.0338
Glycerophospholipid metabolism	**9**	**36**	**0.0351**
Fatty Acid Metabolism	4	14	0.0407
Drug metabolism - cytochrome P450	10	42	0.0481
Leukotriene metabolism	12	51	0.0488
Porphyrin metabolism	6	24	0.0496

Mummichog pathway enrichment test on the results from AOA MWAS adjusted for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 hours, their multiplicative interaction terms with fasting time, and LDSA. The MITM pathways are in bold.

Table 6.S17. Pathways associated to AOA adjusted for NO_2

Pathway	Overlap size	Pathway size	p-value
Biopterin metabolism	9	13	0.0007
Tryptophan metabolism	20	54	0.0009
Vitamin B6 (pyridoxine) metabolism	4	6	0.0019
Drug metabolism - cytochrome P450	13	42	0.0042
Hexose phosphorylation	5	12	0.0052
De novo fatty acid biosynthesis	6	22	0.0415
Drug metabolism - other enzymes	5	18	0.0488

Mummichog pathway enrichment test on the results from AOA MWAS adjusted for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 hours, their multiplicative interaction terms with fasting time, and NO_2.

Table 6.S18. Pathways associated to CCVD adjusted for $PM_{2.5}$

Pathway	Overlap size	Pathway size	p-value
De novo fatty acid biosynthesis	10	14	0.0004
Carnitine shuttle	11	19	0.0004
Starch and Sucrose Metabolism	7	11	0.0006
Hexose phosphorylation	7	12	0.0008
Phosphatidylinositol phosphate metabolism	6	10	0.0011
Tryptophan metabolism	22	58	0.0020
Fatty acid activation	**7**	**15**	**0.0029**
Lysine metabolism	10	24	0.0031
Caffeine metabolism	5	11	0.0077
Ascorbate (Vitamin C) and Aldarate Metabolism	7	18	0.0110
Glycolysis and Gluconeogenesis	4	9	0.0154
Fructose and mannose metabolism	4	9	0.0154
Glycerophospholipid metabolism	12	35	0.0161
Sialic acid metabolism	7	19	0.0169
Arginine and Proline Metabolism	10	29	0.0186
Linoleate metabolism	**7**	**20**	**0.0253**
Porphyrin metabolism	7	20	0.0253
Glutamate metabolism	4	10	0.0270
Glycosphingolipid metabolism	7	21	0.0367
Fatty Acid Metabolism	4	11	0.0445

Mummichog pathway enrichment test on the results from CCVD MWAS adjusted for age at recruitment, center of recruitment, sex, BMI, smoking status, education level, and $PM_{2.5}$. The MITM pathways are in bold.

Table 6.S19. Pathways associated to CCVD adjusted for PNC

Pathway	Overlap size	Pathway size	p-value
Tyrosine metabolism	35	76	0.0010
Tryptophan metabolism	24	58	0.0011
Biopterin metabolism	7	11	0.0012
Glycerophospholipid metabolism	13	35	0.0022
Carnitine shuttle	7	19	0.0063
Valine, leucine and isoleucine degradation	7	19	0.0063
Linoleate metabolism	7	20	0.0084
N-Glycan biosynthesis	4	10	0.0142
Phosphatidylinositol phosphate metabolism	4	10	0.0142
Aminosugars metabolism	5	14	0.0146
Urea cycle/amino group metabolism	12	43	0.0193
Vitamin B3 (nicotinate and nicotinamide) metabolism	5	15	0.0208
Pyrimidine metabolism	8	28	0.0256
Porphyrin metabolism	6	20	0.0275
Butanoate metabolism	5	16	0.0291
Methionine and cysteine metabolism	9	33	0.0315
Hexose phosphorylation	4	12	0.0316
Glycosphingolipid metabolism	**6**	**21**	**0.0367**

Mummichog pathway enrichment test on the results from CCVD MWAS adjusted for age at recruitment, center of recruitment, sex, BMI, smoking status, education level, and PNC. The MITM pathway is in bold.

Table 6.S20. Pathways associated to CCVD adjusted for NO_2

Pathway	Overlap size	Pathway size	p-value
De novo fatty acid biosynthesis	10	14	0.0006
Starch and Sucrose Metabolism	7	11	0.0011
Hexose phosphorylation	7	12	0.0015
Phosphatidylinositol phosphate metabolism	6	10	0.0019
Carnitine shuttle	**9**	**19**	**0.0026**
Tryptophan metabolism	22	58	0.0035
Fatty acid activation	7	15	0.0050
Lysine metabolism	10	24	0.0054
Linoleate metabolism	8	20	0.0119
Caffeine metabolism	5	11	0.0136
Ascorbate (Vitamin C) and Aldarate Metabolism	7	18	0.0192
Glycolysis and Gluconeogenesis	4	9	0.0270
Fructose and mannose metabolism	4	9	0.0270
Glycerophospholipid metabolism	12	35	0.0274
Sialic acid metabolism	7	19	0.0292
Porphyrin metabolism	7	20	0.0433

Mummichog pathway enrichment test on the results from CCVD MWAS adjusted for age at recruitment, center of recruitment, sex, BMI, smoking status, education level, and NO_2. The MITM pathway is in bold.

Table 6.S21. Pathways associated to AOA adjusted for PNC and in addition for current asthma medication — sensitivity analysis

Pathway	Overlap size	Pathway size	p-value
Biopterin metabolism	9	13	0.0007
Linoleate metabolism	**10**	**21**	**0.0010**
Drug metabolism - cytochrome P450	14	42	0.0028
Xenobiotics metabolism	18	57	0.0029
Fatty acid activation	8	21	0.0031
De novo fatty acid biosynthesis	8	22	0.0040
Selenoamino acid metabolism	4	10	0.0111
Fatty Acid Metabolism	5	14	0.0115
Drug metabolism - other enzymes	6	18	0.0118
Tryptophan metabolism	15	54	0.0131
Glycerophospholipid metabolism	**10**	**36**	**0.0202**
Carnitine shuttle	7	26	0.0370

Mummichog pathway enrichment test on the results from AOA MWAS adjusted for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 hours, their multiplicative interaction terms with fasting time, PNC, and current asthma medication. The MITM pathways found from the main analysis are in bold.

Table 6.S22. Pathways associated to AOA adjusted for LDSA and in addition for current asthma medication — sensitivity analysis

Pathway	Overlap size	Pathway size	p-value
Linoleate metabolism	**11**	**21**	**0.0012**
Biopterin metabolism	8	13	0.0012
Xenobiotics metabolism	19	57	0.0062
Fatty acid activation	8	21	0.0082
De novo fatty acid biosynthesis	8	22	0.0113
Drug metabolism - cytochrome P450	13	42	0.0217
Tryptophan metabolism	16	54	0.0283
Fatty Acid Metabolism	5	14	0.0292
Drug metabolism - other enzymes	6	18	0.0325

Mummichog pathway enrichment test on the results from AOA MWAS adjusted for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 hours, their multiplicative interaction terms with fasting time, LDSA, and current asthma medication. The MITM pathway found from the main analysis is in bold.

Table 6.S23. Pathways associated to CVD adjusted for PNC — sensitivity analysis excluding cerebrovascular disease cases

Pathway	Overlap size	Pathway size	p-value
Tyrosine metabolism	18	76	0.0009
Biopterin metabolism	5	11	0.001
Tryptophan metabolism	13	58	0.001
Carnitine shuttle	6	19	0.001
Linoleate metabolism	6	20	0.002
Porphyrin metabolism	4	20	0.02
Glycerophospholipid metabolism	6	35	0.02
Glycosphingolipid metabolism	**4**	**21**	**0.02**
Xenobiotics metabolism	7	45	0.03

Mummichog pathway enrichment test on the results from CVD MWAS (cerebrovascular disease cases excluded) adjusted for age at recruitment, centre of recruitment, sex, BMI, smoking status, education level and PNC. The MITM pathway is in bold.

Table 6.S24. Pathways associated to CVD adjusted for NO_2 — sensitivity analysis excluding cerebrovascular disease cases

Pathway	Overlap size	Pathway size	p-value
Starch and Sucrose Metabolism	7	11	0.0009
Phosphatidylinositol phosphate metabolism	6	10	0.0012
Hexose phosphorylation	6	12	0.0019
De novo fatty acid biosynthesis	6	14	0.0035
Caffeine metabolism	5	11	0.0041
Linoleate metabolism	7	20	0.0074
Arginine and Proline Metabolism	9	29	0.0109
Tryptophan metabolism	16	58	0.0158
Carnitine shuttle	**6**	**19**	**0.0179**
Lysine metabolism	7	24	0.0237
Galactose metabolism	7	25	0.0311
Glycosphingolipid metabolism	6	21	0.0326
Glycerophospholipid metabolism	9	35	0.0457

Mummichog pathway enrichment test on the results from CVD MWAS (cerebrovascular disease cases excluded) adjusted for age at recruitment, centre of recruitment, sex, BMI, smoking status, education level and NO_2. The MITM pathway is in bold.

Table 6.S25. Comparison of the MWAS results with or without further adjustment for BMI or education level — AOA analysis

	Adjusted for BMI		Adjusted for education level	
MWAS	**Pearson correlation of coefficients**	**Spearman correlation of p-values**	**Pearson correlation of coefficients**	**corr. correlation of p-values**
$PM_{2.5}$	1.000 (u0) 1.000 (u1)	0.996	1.000 (u0) 1.000 (u1)	0.996
PNC	1.000	0.995	1.000	0.991
LDSA	0.997	0.995	1.000	0.992
NO_2	0.995	0.995	0.997	0.996
AOA	0.983	0.933	1.000	0.998
AOA, adj. for $PM_{2.5}$	0.984	0.937	1.000	0.998
AOA, adj. for PNC	0.970	0.901	0.997	0.991
AOA, adj. for LDSA	0.971	0.907	0.997	0.988
AOA, adj. for NO_2	0.982	0.930	1.000	0.998

7. DISCUSSION

7.1 Main findings

The aim of this PhD project was to contribute to better understanding the role of obesity and air pollution exposure in asthma etiology, taking into account asthma heterogeneity. Following sections will discuss the contributions this PhD project made towards the aim.

7.1.1 Importance of distinguishing asthma phenotypes

All the findings presented in this PhD project support the importance of distinguishing asthma phenotypes. Among the self-reported ever asthma cases in SAPALDIA, LCA identified various asthma phenotypes which differed by disease activity, atopy, and age of disease onset. The relatively severe phenotypes but not the mild phenotypes showed association with obesity (Article I). The relatively severe phenotypes were characterized by multiple persistent asthma symptoms, lacking atopy, and late-onset of the disease. In the pathway analysis of DNA methylation as effect modifier of obesity on asthma, the DNA methylation in inflammation-related genes modified the effect of BMI on non-atopic adult-onset asthma (Article II). In the metabolomics pathway analysis, fatty acid metabolism-related pathways appeared to be meeting-in-the-middle, i.e. mediating the air pollution exposure to adult-onset asthma (Article III).

The finding that the association with risk factors and the pathways involved in the

association differed by asthma phenotypes indicates that the different phenotypes are potentially distinct disease entities with different etiology. It also emphasizes the importance of refined phenotype information to take into account asthma phenotypes, although there is yet no consensus how to classify them.

7.1.2 Asthma-obesity association

Confirming the accumulating evidence in literature (Beuther and Sutherland, 2007, Egan et al., 2013, Peters et al., 2018), a strong positive association between obesity and the relatively severe asthma phenotypes identified by LCA was observed. Although the cross-sectional design hinders causal interpretation of the association, this work provided evidence against potential reverse causation, i.e. asthma symptoms lead to physical inactivity and in turn to obesity. The association between obesity and the LCA-identified severe phenotypes remained significant when the analysis was restricted to the subjects with sufficient physical activity. The association became stronger in the subjects who had been overweight or obese over 20 years, which can hardly be explained by reverse causation. Considering the recent Mendelian randomization evidence (Granell et al., 2014, Skaaby et al., 2017), it seems plausible to consider obesity leading to asthma rather than the other way around. Understanding the mechanism behind the association could also strengthen the causal interpretation. The pathway level evidence from Article III provided plausible explanation of the mechanism linking obesity to asthma, strengthening the causal interpretation.

7.1.3 BMI not the optimal metric of obesity

Out of the five obesity measures examined, i.e. BMI, percent body fat (PBF), waist circumference (WC), waist-hip ratio (WHR), and waist-height ratio (WHtR), the strongest

association to LCA-identified severe asthma phenotypes was observed with PBF, and the association was independent of BMI (Article I). The pathway enrichment analysis found the DNA methylation of NLRP3-IL1B-IL17 axis as effect modifier of BMI change over 10 years on non-atopic adult-onset asthma, while did not as effect modifier of BMI. The findings indicated that compared to BMI, PBF and BMI change over 10 years captured distinct features of obesity as a risk factor for asthma. As a metric of obesity, BMI has limitations: it cannot distinguish fat mass from lean mass, cannot take into consideration fat distribution and body shape. WHR is considered better than BMI to capture fat distribution. However, fat distribution varies widely across ethnic groups and therefore WHR might not work for all ethnic groups. PBF, estimated based on bioelectric impedance measurements, relates more directly to adiposity than BMI. Longitudinal assessment of BMI, e.g. BMI change over 10 years, possibly captures accumulating excess fat in late adulthood better than cross-sectional BMI. BMI tends to increase as people age and the weight gain in late adulthood is more likely due to fat accumulation (Newman et al., 2005). It is also not uncommon for elderly to lose weight and it is more likely due to muscle loss (Kim et al., 2017, Santanasto et al., 2017).

7.1.4 Asthma-air pollution association

A strong positive association was observed between UFP exposure and adult-onset asthma among non-smoking subsamples of SAPALDIA3. $PM_{2.5}$ or NO_2 exposure did not show a strong association with adult-onset asthma. Multipollutant models provided evidence that the UFP association was independent of other pollutants. This finding is believed to contribute to the literature where the conclusive evidence of the UFP effects on adult-onset asthma has been lacking. Although cross-sectional design of the study did not allow causal interpretation of the association, untargeted metabolomics analysis of the peripheral blood

in the same subsamples found Linoleate metabolism and Glycerophospholipid metabolism as MITM pathways, providing evidence towards causality.

7.1.5 Inflammation: shared mechanism linking various exposures, asthma, and comorbidities

The candidate pathway analysis of genome-wide DNA methylation demonstrated that the inflammation-related genes modify the effects of BMI on non-atopic adult-onset asthma. This finding was further supported by the agnostic pathway analysis of differentially methylated regions. The agnostic pathway analysis of untargeted blood metabolomics identified Linoleate metabolism as MITM pathway linking air pollution exposure not only to adult-onset asthma but also to cardio- and cerebrovascular diseases. Previous in vitro studies demonstrated the pro-inflammatory role of linoleate (Maruyama et al., 2014, Mizuta et al., 2015), although its role in the disease etiology is yet inconclusive. It is conceivable that inflammation conveys the effects of various insults, endogenous (e.g. obesity) and exogenous (e.g. air pollution) alike. The inflammation can be considered as the shared mechanism of the aging phenome, in line with the previous findings from meta-analyses of GWAS reporting pleiotropic loci associated with multiple NCDs around inflammation-related genes (Jeck et al 2012; He et al 2016). This finding is also relevant to LMICs suffering from the dual burden of diseases, i.e. the continuing high burden of infectious disease including tuberculosis, HIV/AIDS, and parasitic diseases in combination with the rapidly increasing burden of NCDs (Young et al., 2009, Remais et al., 2013). The associations between the infectious diseases and the NCDs are in many cases potentially bidirectional (Oni and Unwin, 2015). Tuberculosis and diabetes, for example, have long been associated and while the association is often interpreted that diabetes increases the risk of tuberculosis infection, it is also conceivable that tuberculosis infection increases

the risk of diabetes via inflammation or more broadly innate immune response (Pickup, 2004, Young et al., 2009).

7.1.6 Innate immunity: a plausible mechanism behind the link between obesity and non-atopic asthma

The finding from mice experiments that obesity-induced airway hyperresponsiveness occurred via the NLRP3-IL1B-IL17 axis situated in the center from the beginning of this PhD project. Obesity-related asthma has been considered as a distinct asthma phenotype and reported to be more likely non-atopic and to have neutrophilic airway inflammation (Fenger et al., 2012, Zheng et al., 2016). The candidate pathway analysis of DNA methylation on the NLRP3-IL1B-IL17 axis-related genes provided novel evidence that this axis may also play a role as a link between obesity and non-atopic adult-onset asthma. Interestingly, the NLRP3-IL1B-IL17 axis was found enriched only when the model was not adjusted for neutrophil counts. This was in line with the literature that the effect of IL17 on asthma was mediated by recruitment and survival of neutrophils (Annunziato et al., 2015, Linden and Dahlen, 2014).

7.2 Strengths and limitations

This PhD project relied on the high-quality, rich data from SAPALDIA: three surveys on the same subjects over two decades; refined information on respiratory health, allowing to investigate various asthma phenotypes; dense information on risk factors including obesity and life style; individual level estimates of various air pollutants from high resolution models; multiple omics measurements on the same subsamples. The omics analyses were most likely underpowered as they were conducted only for the subsamples of relatively

small size. The limited statistical power may have led to false negative findings in the omics analyses. The limited statistical power was also one of the reasons unable to investigate broader spectrum exposures. However, the low power was at least in part overcome by applying multivariate approaches such as pathway enrichment tests and DMR approach in methylome analysis. Thanks to the collaboration with partners of a variety of expertise, we could have made use of state-of-the-art technologies and statistical methods, and allowed investigation of broader phenome as done in the metabolomics study (Article III). At the same time, this PhD project has been the incubator of such collaboration, which will serve future research.

Although SAPALDIA offers longitudinal data, we used cross-sectional designs mainly because of the difficulties in defining incident cases of asthma. This made it difficult to draw causal inference. Throughout this PhD project, therefore, the effort has been made not to overly interpret the findings as causal. However, it should still be said that the findings contributed to strengthening the causal reasoning. Causality in observational studies can be best conceptualized by counterfactual model. Risk factors of NCDs such as asthma are neither necessary nor sufficient and often have small effect size. Implementing counterfactual models for NCDs requires infeasibly large sample size, if not impossible. Therefore causal inference in the NCD epidemiology has to rely on the pieces of evidence coming from various fields – in vitro studies, animal studies, clinical trials, and cohort studies. The MITM concept applied in Article III, for example, is only associational *per se* but can still be understood as a tool to strengthen causal interpretation on the basis of prior knowledge. It should be noted that cohort studies are often the only source of evidence for causal effects of certain exposures, e.g. air pollution, for which experimental studies are not ethically acceptable or feasible. Such comprehensive consideration for causality has been already practiced by the IARC in the evaluation process of carcinogen

classification: for evaluation of the evidence, a variety of factors are taken into consideration, including study size, sound methodology, and replication across various study designs with emphasis on epidemiological studies (IARC, 2006).

7.3 Research implications

We confirmed the findings from previous studies that obesity-induced asthma is more likely to be non-atopic and adult-onset and provided novel evidence of inflammation as a potential mechanism behind the obesity-induced non-atopic adult-onset asthma. The findings demonstrated the importance of distinguishing asthma phenotypes in understanding of the etiology and the necessity of refined information collected from cohort participants to allow such investigation. This PhD project exemplified the applicability of omics analysis to obtain insights into the NCD etiology and provided a proof-of-concept for the investigation of common pathways shared in aging phenome. Further research, including Mendelian randomization studies, is warranted to consolidate the causal inference of the mechanistic findings reported here. Thanks to the unprecedented genetic data, for the first time we epidemiologists now possess a tool for causal inference in observational study settings. Mendelian randomization has been indeed increasingly applied in epidemiology (Timpson et al., 2011, Granell et al., 2014, Richmond et al., 2014, Millard et al., 2015, Skaaby et al., 2017), and the literature will only grow. However, it should be borne in mind that instrument variables are not available for all the risk factors of public health relevance. For example, air pollution exposure cannot be expected to have genetic variants as instrument variables. Such lack of instrument variables and hence inapplicability of Mendelian randomization should not discourage the research efforts to elucidate the causal effects of air pollution exposure. Moreover, future research should

note that people are inevitably exposed to multiple exposures. Refined and validated exposure assessments, possibly accompanied with multiple omics analyses, will help study such a mixture of exposures.

7.3.1 Lessons learnt from EXPOsOMICS

The large exposome research consortium, EXPOsOMICS, exemplified how to implement exposome research as well as what challenges such research faces. The biggest challenge was the limited availability of harmonized data with sufficient power (Vineis et al., 2016). Considering the small effect size of the NCD risk factors and the broad spectrum of exposures that the exposome research aims to cover, large sample size is required. Besides the budding initiatives for mega-cohorts – the All of Us in the United States (NIH, 2018), the National Cohort (NAKO) in Germany (GNC Consortium, 2014), the CONSTANCES in France (Zins and Goldberg, 2015), the China Kadoorie Biobank (Chen et al., 2011), among others – will facilitate future research, for the time being the exposome research needs to utilize consortium comprising of multiple cohorts. The exposome concept does not only cover the full spectrum of exposures but also spans the life course of the exposures from in utero to late adulthood. In order to capture the multiple key stages of life, albeit not the full life course, a consortium comprising multiple cohorts of various age groups – birth cohorts, children cohorts, adult cohorts, etc. – is inevitable (Wild, 2012). Even though multiple cohorts participated in the EXPOsOMICS, it was not always straightforward to integrate the data, which had been collected from the different cohorts before the consortium. Efforts were made to ensure that the data newly collected within the consortium were standardized across the participating cohorts, for example by developing West-European LUR models for $PM_{2.5}$ and NO_2 (de Hoogh et al., 2016, van Nunen et al., 2017). Careful design, management, and communication – in particular the

interdisciplinary collaboration between epidemiologists, fieldworkers, and laboratory scientists at each step of data collection including storage, processing, storage, and sharing – were of paramount importance to ensure the harmonized, high quality data in such a large, multi-center, international consortium.

7.4 Public health implications

Given the escalating prevalence of obesity, the obesity-asthma association implies public health concerns not only in HICs like Switzerland but also in LMICs. Over the last decades obesity has been growing rapidly in LMICs, contributing to huge burden of NCDs (Ford et al., 2017, Forouzanfar et al., 2016). Considering that the air pollution is often poorly controlled in LMICs, the strong association between UFP and asthma we observed casts serious concerns in LMIC settings. NCD epidemic can form a vicious circle in the resource-deficient settings, i.e. poverty drives unhealthy life style leading to higher risk of NCDs and productivity loss and health costs due to the NCDs in turn aggravate the poverty (WHO, 2014).

NCDs often require life-long, expensive treatments, posing serious economic burden not only to the affected individuals and households but also to the societies. Treatment-focused strategies would not be affordable in the HICs and the LMICs alike, given the alarmingly increasing health costs in the HICs and the already stretched, fragile health systems in the LMICs. The most cost-effective strategy against NCDs is primary prevention. And the primary prevention must be implemented at the society level, because many NCD risk factors are under influence of the infrastructure, law system, and culture within the society. Air pollution exposure can be little modified at the individual level. Obesity is often considered as a result of genetics and behavior but societal factors

should not be ignored: public transport, urban structure, physical activity education in school, regulation on food industry, labor law, etc. Tailoring such systemic factors should be informed by the state-of-the-art evidence of the disease etiology. Considering that most NCD risk factors have small effect size and work in concert, systems approach is imperative for NCD etiology research.

Omics analysis alongside with investigation of high quality information other than omics can provide a new winder to capture the patterns of multiple exposures and contribute to better understanding of the disease etiology. Elucidation of the shared mechanism for aging phenome can critically benefit prevention and intervention strategies to combat the NCD epidemic. Moreover, if there is a shared mechanism in the association between multiple exposures and multiple diseases, the burden of disease estimated must be informed accordingly, as the burden of disease could have been underestimated in case of synergistic effects of various exposures on comorbidities or could have been overestimated in case of compensating/saturating effects. Omics analysis can also offer novel therapy options as exemplified by several epigenetic drugs (Heerboth et al., 2014, Tough et al., 2016).

7.5 Conclusion and future work

Making use of the rich, high quality data in SAPALDIA, we demonstrated as an example how to take a step forward to mechanistic understanding of asthma, and at the same time exemplified the usefulness of cohort studies with refined phenotype information accompanied with biobank. We also provided a proof-of-concept for inflammation as the shared mechanism of asthma and cardiovascular diseases, major components of NCD epidemic. Before these findings actually benefit global public health, future research should

replicate and refine the mechanistic understanding; consolidate its causality by Mendelian randomization studies; and expand to broader spectrum of exposures taking into account their correlation structure and interactions. Importantly, exposures that societies rather than individuals have leverage on, e.g. urban structure and socioeconomic status should be further investigated as NCD risk factors, considering that poverty is a consequence as well as a cause of NCDs. This mandates the cohort studies in various populations including LMICs because each population is positioned in a specific setting with different spectrum of exposures and hence findings from one specific population are unlikely generalized into other populations. The pursuit of mechanistic understanding of NCDs, as this PhD project aimed, can be considered as the first step of translational approaches we all do in every sector in public health: innovations shall be validated and applied in real-world settings to improve public health.

References

AARON, S. D., VANDEMHEEN, K. L., BOULET, L. P., MCIVOR, R. A., FITZGERALD, J. M., HERNANDEZ, P., LEMIERE, C., SHARMA, S., FIELD, S. K., ALVAREZ, G. G., DALES, R. E., DOUCETTE, S. & FERGUSSON, D. 2008. Overdiagnosis of asthma in obese and nonobese adults. *CMAJ,* 179, 1121-31.

ACKERMANN-LIEBRICH, U., KUNA-DIBBERT, B., PROBST-HENSCH, N. M., SCHINDLER, C., DIETRICH, D. F., STUTZ, E. Z., BAYER-OGLESBY, L., BAUM, F., BRÄNDLI, O., BRUTSCHE, M., DOWNS, S. H., KEIDEL, D., GERBASE, M. W., IMBODEN, M., KELLER, R., KNÖPFLI, B., KÜNZLI, N., NICOD, L., PONS, M., STAEDELE, P., TSCHOPP, J.-M., ZELLWEGER, J.-P. & LEUENBERGER, P. 2005. Follow-up of the Swiss Cohort Study on Air Pollution and Lung Diseases in Adults (SAPALDIA 2) 19912003: methods and characterization of participants. *Sozial- und Präventivmedizin*, 50, 245-263.

AGGARWAL, N., KORENBAUM, E., MAHADEVA, R., IMMENSCHUH, S., GRAU, V., DINARELLO, C. A., WELTE, T. & JANCIAUSKIENE, S. 2016. α-Linoleic Acid Enhances the Capacity of α1-Antitrypsin to Inhibit Lipopolysaccharide-Induced IL-1αin Human Blood Neutrophils. *Molecular Medicine,* 22, 680-693.

AGIER, L., PORTENGEN, L., CHADEAU-HYAM, M., BASAGANA, X., GIORGIS-ALLEMAND, L., SIROUX, V., ROBINSON, O., VLAANDEREN, J., GONZALEZ, J. R., NIEUWENHUIJSEN, M. J., VINEIS, P., VRIJHEID, M., SLAMA, R. & VERMEULEN, R. 2016. A Systematic Comparison of Linear Regression-Based Statistical Methods to Assess Exposome-Health Associations. *Environ Health Perspect,* 124, 1848-1856.

AGUILERA, I., DRATVA, J., CAVIEZEL, S., BURDET, L., DE GROOT, E., DUCRET-STICH, R. E., EEFTENS, M., KEIDEL, D., MEIER, R., PEREZ, L., ROTHE, T., SCHAFFNER, E., SCHMIT-TRUCKSASS, A., TSAI, M. Y., SCHINDLER, C., KUNZLI, N. & PROBST-HENSCH, N. 2016. Particulate Matter and Subclinical Atherosclerosis: Associations between Different Particle Sizes and Sources with Carotid Intima-Media Thickness in the SAPALDIA Study. *Environ Health Perspect,* 124, 1700-1706.

ANDERSON, H. R., FAVARATO, G. & ATKINSON, R. W. 2013. Long-term exposure to air pollution and the incidence of asthma: meta-analysis of cohort studies. *Air Quality, Atmosphere & Health,* 6 **,** 47-56.

ANNUNZIATO, F., ROMAGNANI, C. & ROMAGNANI, S. 2015. The 3 major types of innate and adaptive cell-mediated effector immunity. *J Allergy Clin Immunol,* 135, 626-35.

ARYEE, M. J., JAFFE, A. E., CORRADA-BRAVO, H., LADD-ACOSTA, C., FEINBERG, A. P., HANSEN, K. D. & IRIZARRY, R. A. 2014. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics,* 30 , 1363-9.

ASSI, N., FAGES, A., VINEIS, P., CHADEAU-HYAM, M., STEPIEN, M., DUARTE-SALLES, T., BYRNES, G., BOUMAZA, H., KNUPPEL, S., KUHN, T., PALLI, D., BAMIA, C., BOSHUIZEN, H., BONET, C., OVERVAD, K., JOHANSSON, M., TRAVIS, R., GUNTER, M. J., LUND, E., DOSSUS, L., ELENA-HERRMANN, B., RIBOLI, E., JENAB, M., VIALLON, V. & FERRARI, P. 2015. A statistical framework to model the meeting-in-the-middle principle using metabolomic data: application to hepatocellular carcinoma in the EPIC study. *Mutagenesis,* 30, 743-53.

BALDING, D. J. 2006. A tutorial on statistical methods for population association studies. *Nat Rev Genet,* 7, 781-91.

BANNO, A., REDDY, A. T., LAKSHMI, S. P. & REDDY, R. C. 2018. PPARs: Key Regulators of Airway Inflammation and Potential Therapeutic Targets in Asthma. *Nucl Receptor Res,* 5.

BENDINELLI, B., MASALA, G., SAIEVA, C., SALVINI, S., CALONICO, C., SACERDOTE, C., AGNOLI, C., GRIONI, S., FRASCA, G., MATTIELLO, A., CHIODINI, P., TUMINO, R., VINEIS, P., PALLI, D. & PANICO, S. 2011. Fruit, vegetables, and olive oil and risk of coronary heart disease in Italian women: the EPICOR Study. *Am J Clin Nutr,* 93 , 275-83.

BENJAMINI, Y. & HOCHBERG, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)* , 289-300.

BEUTHER, D. A. & SUTHERLAND, E. R. 2005. Obesity and pulmonary function testing. *J Allergy Clin Immunol,* 115, 1100-1.

BEUTHER, D. A. & SUTHERLAND, E. R. 2007. Overweight, obesity, and incident asthma: a meta-analysis of prospective epidemiologic studies. *Am J Respir Crit Care Med,* 175, 661-6.

BIBIKOVA, M., BARNES, B., TSAN, C., HO, V., KLOTZLE, B., LE, J. M., DELANO, D., ZHANG, L., SCHROTH, G. P., GUNDERSON, K. L., FAN, J. B. & SHEN, R. 2011. High density DNA methylation array with single CpG site resolution. *Genomics,* 98, 288-95.

BICTASH, M., EBBELS, T. M., CHAN, Q., LOO, R. L., YAP, I. K., BROWN, I. J., DE IORIO, M., DAVIGLUS, M. L., HOLMES, E. & STAMLER, J. 2010. Opening up the" Black Box": Metabolic phenotyping and metabolome-wide association studies in epidemiology. *Journal of clinical epidemiology,* 63, 970-979.

BOUDIER, A., CURJURIC, I., BASAGANA, X., HAZGUI, H., ANTO, J. M., BOUSQUET, J., BRIDEVAUX, P. O., DUPUIS-LOZERON, E., GARCIA-AYMERICH, J., HEINRICH, J., JANSON, C., KUNZLI, N., LEYNAERT, B., DE MARCO, R., ROCHAT, T., SCHINDLER, C., VARRASO, R., PIN, I., PROBST-HENSCH, N., SUNYER, J., KAUFFMANN, F. & SIROUX, V. 2013. Ten-year follow-up of cluster-based asthma phenotypes in adults. A pooled analysis of three cohorts. *Am J Respir Crit Care Med,* 188, 550-60.

BRANDLI, O., SCHINDLER, C., KUNZLI, N., KELLER, R. & PERRUCHOUD, A. P. 1996. Lung function in

healthy never smoking adults: reference values and lower limits of normal of a Swiss population. *Thorax,* 51 , 277-83.

BRANDLI, O., SCHINDLER, C., LEUENBERGER, P. H., BAUR, X., DEGENS, P., KUNZLI, N., KELLER, R. & PERRUCHOUD, A. P. 2000. Re-estimated equations for 5th percentiles of lung function variables. *Thorax,* 55 , 173-4.

BROOK, R. D., RAJAGOPALAN, S., POPE, C. A., 3RD, BROOK, J. R., BHATNAGAR, A., DIEZ-ROUX, A. V., HOLGUIN, F., HONG, Y., LUEPKER, R. V., MITTLEMAN, M. A., PETERS, A., SISCOVICK, D., SMITH, S. C., JR., WHITSEL, L. & KAUFMAN, J. D. 2010. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation,* 121, 2331-78.

BRUMPTON, B. M., CAMARGO, C. A., JR., ROMUNDSTAD, P. R., LANGHAMMER, A., CHEN, Y. & MAI, X. M. 2013. Metabolic syndrome and incidence of asthma in adults: the HUNT study. *Eur Respir J,* 42, 1495-502.

CAI, Y., ZIJLEMA, W. L., DOIRON, D., BLANGIARDO, M., BURTON, P. R., FORTIER, I., GAYE, A., GULLIVER, J., DE HOOGH, K., HVEEM, K., MBATCHOU, S., MORLEY, D. W., STOLK, R. P., ELLIOTT, P., HANSELL, A. L. & HODGSON, S. 2017. Ambient air pollution, traffic noise and adult asthma prevalence: a BioSHaRE approach. *Eur Respir J,* 49.

CAMERON, A. C. & WINDMEIJER, F. A. 1997. An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics,* 77, 329-342.

CAREY, M. A., CARD, J. W., VOLTZ, J. W., ARBES, S. J., JR., GERMOLEC, D. R., KORACH, K. S. & ZELDIN, D. C. 2007. It's all about sex: gender, lung development and lung disease. *Trends Endocrinol Metab,* 18, 308-13.

CASTOLDI, A., NAFFAH DE SOUZA, C., CMARA, N. O. S. & MORAES-VIEIRA, P. M. 2016. The macrophage switch in obesity development. *Frontiers in immunology,* 6, 637.

CASTRO-GINER, F., KOGEVINAS, M., IMBODEN, M., DE CID, R., JARVIS, D., MACHLER, M., BERGER, W., BURNEY, P., FRANKLIN, K. A., GONZALEZ, J. R., HEINRICH, J., JANSON, C., OMENAAS, E., PIN, I., ROCHAT, T., SUNYER, J., WJST, M., ANTO, J. M., ESTIVILL, X. & PROBST-HENSCH, N. M. 2009. Joint effect of obesity and TNFA variability on asthma: two international cohort studies. *Eur Respir J,* 33, 1003-9.

CESARONI, G., FORASTIERE, F., STAFOGGIA, M., ANDERSEN, Z. J., BADALONI, C., BEELEN, R., CARACCIOLO, B., DE FAIRE, U., ERBEL, R., ERIKSEN, K. T., FRATIGLIONI, L., GALASSI, C., HAMPEL, R., HEIER, M., HENNIG, F., HILDING, A., HOFFMANN, B., HOUTHUIJS, D., JOCKEL, K. H., KOREK, M., LANKI, T., LEANDER, K., MAGNUSSON, P. K., MIGLIORE, E., OSTENSON, C. G., OVERVAD, K., PEDERSEN, N. L., J, J. P., PENELL, J., PERSHAGEN, G., PYKO, A., RAASCHOU-NIELSEN, O., RANZI, A., RICCERI, F., SACERDOTE, C., SALOMAA, V., SWART, W., TURUNEN, A. W., VINEIS, P., WEINMAYR, G., WOLF, K., DE HOOGH, K., HOEK, G., BRUNEKREEF, B. &

PETERS, A. 2014. Long term exposure to ambient air pollution and incidence of acute coronary events: prospective cohort study and meta-analysis in 11 European cohorts from the ESCAPE Project. *Bmj,* 348, f7412.

CHADEAU-HYAM, M., CAMPANELLA, G., JOMBART, T., BOTTOLO, L., PORTENGEN, L., VINEIS, P., LIQUET, B. & VERMEULEN, R. C. 2013. Deciphering the complex: methodological overview of statistical models to derive OMICS-based biomarkers. *Environ Mol Mutagen,* 54, 542-57.

CHARMPI, K. & YCART, B. 2015. Weighted Kolmogorov Smirnov testing: an alternative for Gene Set Enrichment Analysis. *Stat Appl Genet Mol Biol,* 14, 279-93.

CHEN, R. 2015. On bioinformatic resources. *Genomics Proteomics Bioinformatics,* 13, 1-3.

CHEN, Y. A., LEMIRE, M., CHOUFANI, S., BUTCHER, D. T., GRAFODATSKAYA, D., ZANKE, B. W., GALLINGER, S., HUDSON, T. J. & WEKSBERG, R. 2013. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics,* 8 , 203-9.

CHEN, Z., CHEN, J., COLLINS, R., GUO, Y., PETO, R., WU, F. & LI, L. 2011. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *International journal of epidemiology,* 40, 1652-1666.

CHUANG, K. J., YAN, Y. H., CHIU, S. Y. & CHENG, T. J. 2011. Long-term air pollution exposure and risk factors for cardiovascular diseases among the elderly in Taiwan. *Occup Environ Med,* 68, 64-8.

CHUNG, J. W., BANG, O. Y., AHN, K., PARK, S. S., PARK, T. H., KIM, J. G., KO, Y., LEE, S., LEE, K. B., LEE, J., KANG, K., PARK, J. M., CHO, Y. J., HONG, K. S., NAH, H. W., KIM, D. H., CHA, J. K., RYU, W. S., KIM, D. E., KIM, J. T., CHOI, J. C., OH, M. S., YU, K. H., LEE, B. C., LEE, J. S., LEE, J., PARK, H. K., KIM, B. J., HAN, M. K. & BAE, H. J. 2017. Air Pollution Is Associated With Ischemic Stroke via Cardiogenic Embolism. *Stroke,* 48, 17-23.

CLIFFORD, R. L., FISHBANE, N., PATEL, J., MACISAAC, J. L., MCEWEN, L. M., FISHER, A. J., BRANDSMA, C. A., NAIR, P., KOBOR, M. S., HACKETT, T. L. & KNOX, A. J. 2018. Altered DNA methylation is associated with aberrant gene expression in parenchymal but not airway fibroblasts isolated from individuals with COPD. *Clin Epigenetics,* 10 , 32.

COX, L. A., JR. 2017. Socioeconomic and air pollution correlates of adult asthma, heart attack, and stroke risks in the United States, 2010-2013. *Environ Res,* 155, 92-107.

DE HOOGH, K., GULLIVER, J., DONKELAAR, A. V., MARTIN, R. V., MARSHALL, J. D., BECHLE, M. J., CESARONI, G., PRADAS, M. C., DEDELE, A., EEFTENS, M., FORSBERG, B., GALASSI, C., HEINRICH, J., HOFFMANN, B., JACQUEMIN, B., KATSOUYANNI, K., KOREK, M., KUNZLI, N., LINDLEY, S. J., LEPEULE, J., MELEUX, F., DE NAZELLE, A., NIEUWENHUIJSEN, M., NYSTAD, W., RAASCHOU-NIELSEN, O., PETERS, A., PEUCH, V. H., ROUIL, L., UDVARDY, O., SLAMA, R., STEMPFELET, M., STEPHANOU, E. G., TSAI, M. Y., YLI-TUOMI, T., WEINMAYR, G., BRUNEKREEF, B., VIENNEAU, D. & HOEK, G. 2016. Development of West-European PM2.5 and NO2

land use regression models incorporating satellite-derived and chemical transport modelling data. *Environ Res,* 151, 1-10.

DEDEURWAERDER, S., DEFRANCE, M., CALONNE, E., DENIS, H., SOTIRIOU, C. & FUKS, F. 2011. Evaluation of the Infinium Methylation 450K technology. *Epigenomics,* 3, 771-84.

DEMENAIS, F., MARGARITTE-JEANNIN, P., BARNES, K. C., COOKSON, W. O. C., ALTMULLER, J., ANG, W., BARR, R. G., BEATY, T. H., BECKER, A. B., BEILBY, J., BISGAARD, H., BJORNSDOTTIR, U. S., BLEECKER, E., BONNELYKKE, K., BOOMSMA, D. I., BOUZIGON, E., BRIGHTLING, C. E., BROSSARD, M., BRUSSELLE, G. G., BURCHARD, E., BURKART, K. M., BUSH, A., CHAN-YEUNG, M., CHUNG, K. F., COUTO ALVES, A., CURTIN, J. A., CUSTOVIC, A., DALEY, D., DE JONGSTE, J. C., DEL-RIO-NAVARRO, B. E., DONOHUE, K. M., DUIJTS, L., ENG, C., ERIKSSON, J. G., FARRALL, M., FEDOROVA, Y., FEENSTRA, B., FERREIRA, M. A., FREIDIN, M. B., GAJDOS, Z., GAUDERMAN, J., GEHRING, U., GELLER, F., GENUNEIT, J., GHARIB, S. A., GILLILAND, F., GRANELL, R., GRAVES, P. E., GUDBJARTSSON, D. F., HAAHTELA, T., HECKBERT, S. R., HEEDERIK, D., HEINRICH, J., HELIOVAARA, M., HENDERSON, J., HIMES, B. E., HIROSE, H., HIRSCHHORN, J. N., HOFMAN, A., HOLT, P., HOTTENGA, J., HUDSON, T. J., HUI, J., IMBODEN, M., IVANOV, V., JADDOE, V. W. V., JAMES, A., JANSON, C., JARVELIN, M. R., JARVIS, D., JONES, G., JONSDOTTIR, I., JOUSILAHTI, P., KABESCH, M., KAHONEN, M., KANTOR, D. B., KARUNAS, A. S., KHUSNUTDINOVA, E., KOPPELMAN, G. H., KOZYRSKYJ, A. L., KREINER, E., KUBO, M., KUMAR, R., KUMAR, A., KUOKKANEN, M., LAHOUSSE, L., LAITINEN, T., LAPRISE, C., LATHROP, M., LAU, S., LEE, Y. A., LEHTIMAKI, T., LETORT, S., LEVIN, A. M., LI, G., LIANG, L., LOEHR, L. R., LONDON, S. J., LOTH, D. W., MANICHAIKUL, A., et al. 2018. Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. *Nat Genet,* 50 , 42-53.

DICHTL, W., ARES, M. P., JONSON, A. N., JOVINGE, S., PACHINGER, O., GIACHELLI, C. M., HAMSTEN, A., ERIKSSON, P. & NILSSON, J. 2002. Linoleic acid-stimulated vascular adhesion molecule-1 expression in endothelial cells depends on nuclear factor-kappaB activation. *Metabolism,* 51, 327-33.

DICK, K. J., NELSON, C. P., TSAPROUNI, L., SANDLING, J. K., AISSI, D., WAHL, S., MEDURI, E., MORANGE, P.-E., GAGNON, F., GRALLERT, H., WALDENBERGER, M., PETERS, A., ERDMANN, J., HENGSTENBERG, C., CAMBIEN, F., GOODALL, A. H., OUWEHAND, W. H., SCHUNKERT, H., THOMPSON, J. R., SPECTOR, T. D., GIEGER, C., TRGOUT, D.-A., DELOUKAS, P. & SAMANI, N. J. 2014. DNA methylation and body-mass index: a genome-wide analysis. *The Lancet,* 383, 1990-1998.

DOWNS, S. H., SCHINDLER, C., LIU, L. J., KEIDEL, D., BAYER-OGLESBY, L., BRUTSCHE, M. H., GERBASE, M. W., KELLER, R., KUNZLI, N., LEUENBERGER, P., PROBST-HENSCH, N. M., TSCHOPP, J. M., ZELLWEGER, J. P., ROCHAT, T., SCHWARTZ, J. & ACKERMANN-LIEBRICH, U. 2007. Reduced exposure to PM10 and attenuated age-related decline in lung function. *N Engl J Med,* 357, 2338-47.

DUARTE, N. C., BECKER, S. A., JAMSHIDI, N., THIELE, I., MO, M. L., VO, T. D., SRIVAS, R. &

PALSSON, B. O. 2007. Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc Natl Acad Sci U S A,* 104, 1777-82.

EEFTENS, M., MEIER, R., SCHINDLER, C., AGUILERA, I., PHULERIA, H., INEICHEN, A., DAVEY, M., DUCRET-STICH, R., KEIDEL, D., PROBST-HENSCH, N., KÜNZLI, N. & TSAI, M.-Y. 2016. Development of land use regression models for nitrogen dioxide, ultrafine particles, lung deposited surface area, and four other markers of particulate matter pollution in the Swiss SAPALDIA regions. *Environmental Health,* 15, 1-14.

EGAN, K. B., ETTINGER, A. S. & BRACKEN, M. B. 2013. Childhood body mass index and subsequent physician-diagnosed asthma: a systematic review and meta-analysis of prospective cohort studies. *BMC Pediatr,* 13 , 121.

ENGIN, A. B. 2017. Adipocyte-Macrophage Cross-Talk in Obesity. *Adv Exp Med Biol,* 960, 327-343.

ESSER, N., L'HOMME, L., DE ROOVER, A., KOHNEN, L., SCHEEN, A. J., MOUTSCHEN, M., PIETTE, J., LEGRAND-POELS, S. & PAQUOT, N. 2013. Obesity phenotype is related to NLRP3 inflammasome activity and immunological profile of visceral adipose tissue. *Diabetologia,* 56 **,** 2487-97.

EVANS, K. A., HALTERMAN, J. S., HOPKE, P. K., FAGNANO, M. & RICH, D. Q. 2014. Increased ultrafine particles and carbon monoxide concentrations are associated with asthma exacerbation among urban children. *Environmental research,* 129, 11-19.

EZE, I. C., SCHAFFNER, E., FORASTER, M., IMBODEN, M., VON ECKARDSTEIN, A., GERBASE, M. W., ROTHE, T., ROCHAT, T., KUNZLI, N., SCHINDLER, C. & PROBST-HENSCH, N. 2015. Long-Term Exposure to Ambient Air Pollution and Metabolic Syndrome in Adults. *PLoS One,* 10, e0130337.

FASANELLI, F., BAGLIETTO, L., PONZI, E., GUIDA, F., CAMPANELLA, G., JOHANSSON, M., GRANKVIST, K., JOHANSSON, M., ASSUMMA, M. B. & NACCARATI, A. 2015. Hypomethylation of smoking-related genes is associated with future lung cancer in four prospective cohorts. *Nature communications,* 6, 10192.

FEDERAL STATISTICAL OFFICE *Schweizerische Gesundheitsbefragung* [Online]. Available: http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen___quellen/blank/blank/ess/04.html.

FENGER, R. V., GONZALEZ-QUINTELA, A., VIDAL, C., GUDE, F., HUSEMOEN, L. L., AADAHL, M., BERG, N. D. & LINNEBERG, A. 2012. Exploring the obesity-asthma link: do all types of adiposity increase the risk of asthma? *Clin Exp Allergy,* 42, 1237-45.

FENGER, R. V., GONZALEZ-QUINTELA, A., VIDAL, C., HUSEMOEN, L. L., SKAABY, T., THUESEN, B. H., AADAHL, M., MADSEN, F. & LINNEBERG, A. 2014. The longitudinal relationship of changes of adiposity to changes in pulmonary function and risk of asthma in a general adult population. *BMC Pulm Med,* 14, 208.

FERRARI, R., MERLI, E., CICCHITELLI, G., MELE, D., FUCILI, A. & CECONI, C. 2004. Therapeutic effects of L-carnitine and propionyl-L-carnitine on cardiovascular diseases: a review. *Ann N Y Acad Sci,* 1033 , 79-91.

FIORITO, G., VLAANDEREN, J., POLIDORO, S., GULLIVER, J., GALASSI, C., RANZI, A., KROGH, V., GRIONI, S., AGNOLI, C., SACERDOTE, C., PANICO, S., TSAI, M. Y., PROBST-HENSCH, N., HOEK, G., HERCEG, Z., VERMEULEN, R., GHANTOUS, A., VINEIS, P. & NACCARATI, A. 2017. Oxidative stress and inflammation mediate the effect of air pollution on cardio- and cerebrovascular disease: A prospective study in nonsmokers. *Environ Mol Mutagen.*

FIRTH, D. 1993. Bias reduction of maximum likelihood estimates. *Biometrika,* 80, 27-38.

FOEN 2013. PM10 and PM2.5 ambient concentrations in Switzerland. Modelling results for 2005, 2010 and 2020. Federal Office for the Environment, Bern. Environmental studies no. 1304: 83 pp.

FOEN 2014. Karten der luftbelastung schweiz. Dokumentation zu Datengrundlagen, Berechnungsverfahren und Resultaten. Aktualisierung 2013/2014.

FORD, N. D., PATEL, S. A. & NARAYAN, K. M. 2017. Obesity in Low- and Middle-Income Countries: Burden, Drivers, and Emerging Challenges. *Annu Rev Public Health,* 38, 145-164.

FORNO, E., WANG, T., QI, C., YAN, Q., XU, C. J., BOUTAOUI, N., HAN, Y. Y., WEEKS, D. E., JIANG, Y., ROSSER, F., VONK, J. M., BROUWER, S., ACOSTA-PEREZ, E., COLON-SEMIDEY, A., ALVAREZ, M., CANINO, G., KOPPELMAN, G. H., CHEN, W. & CELEDON, J. C. 2018. DNA methylation in nasal epithelium, atopy, and atopic asthma in children: a genome-wide study. *Lancet Respir Med.*

FOROUZANFAR, M., AFSHIN, A., ALEXANDER, L., ANDERSON, H., BHUTTA, Z., BIRYUKOV, S., BRAUER, M., BURNETT, R., CERCY, K., CHARLSON, F., COHEN, A., DANDONA, L., ESTEP, K., FERRARI, A., FROSTAD, J., FULLMAN, N., GETHING, P., GODWIN, W., GRISWOLD, M., KINFU, Y., KYU, H., LARSON, H., LIANG, X., LIM, S., LIU, P., LOPEZ, A., LOZANO, R., MARCZAK, L., MENSAH, G., MOKDAD, A., MORADI-LAKEH, M., NAGHAVI, M., NEAL, B., REITSMA, M., ROTH, G., SALOMON, J., SUR, P., VOS, T., WAGNER, J., WANG, H., ZHAO, Y., ZHOU, M., AASVANG, G., ABAJOBIR, A., ABATE, K., ABBAFATI, C., ABBAS, K., ABD-ALLAH, F., ABDULLE, A., ABERA, S., ABRAHAM, B., ABU-RADDAD, L., ABYU, G., ADEBIYI, A., ADEDEJI, I., ADEMI, Z., ADOU, A., ADSUAR, J., AGARDH, E., AGARWAL, A., AGRAWAL, A., KIADALIRI, A., AJALA, O., AKINYEMIJU, T., AL-ALY, Z., ALAM, K., ALAM, N., ALDHAHRI, S., ALDRIDGE, R., ALEMU, Z., ALI, R., ALKERWI, A., ALLA, F., ALLEBECK, P., ALSHARIF, U., ALTIRKAWI, K., MARTIN, E., ALVIS-GUZMAN, N., AMARE, A., AMBERBIR, A., AMEGAH, A., AMINI, H., AMMAR, W., AMROCK, S., ANDERSEN, H., ANDERSON, B., ANTONIO, C., ANWARI, P., ÄRNLÖV, J., ARTAMAN, A., ASAYESH, H., ASGHAR, R., ASSADI, R., ATIQUE, S., AVOKPAHO, E., AWASTHI, A., QUINTANILLA, B., AZZOPARDI, P., BACHA, U., BADAWI, A., et al. 2016. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet,* 388, 1659-1724.

FRANKLIN, B. A., BROOK, R. & ARDEN POPE, C., 3RD 2015. Air pollution and cardiovascular disease. *Curr Probl Cardiol,* 40**,** 207-38.

GALIC, S., OAKHILL, J. S. & STEINBERG, G. R. 2010. Adipose tissue as an endocrine organ. *Mol Cell Endocrinol,* 316, 129-39.

GINA 2018. Global strategy for asthma management and prevention.

GLOBAL ASTHMA NETWORK. 2014. *The Global Asthma Report 2014,* Auckland, New Zealand.

GNC CONSORTIUM 2014. The German National Cohort: aims, study design and organization. *Eur J Epidemiol,* 29, 371-82.

GRANELL, R., HENDERSON, A. J., EVANS, D. M., SMITH, G. D., NESS, A. R., LEWIS, S., PALMER, T. M. & STERNE, J. A. 2014. Effects of BMI, fat mass, and lean mass on asthma in childhood: a Mendelian randomization study. *PLoS Med,* 11, e1001669.

GUARNIERI, M. & BALMES, J. R. 2014. Outdoor air pollution and asthma. *Lancet,* 383, 1581-92.

GUIDA, F., SANDANGER, T. M., CASTAGNÉ, R., CAMPANELLA, G., POLIDORO, S., PALLI, D., KROGH, V., TUMINO, R., SACERDOTE, C. & PANICO, S. 2015. Dynamics of smoking-induced genome-wide methylation changes with time since smoking cessation. *Human molecular genetics,* 24 , 2349-2359.

HAGENAARS, J. A. & MCCUTCHEON, A. L. 2002. *Applied latent class analysis,* Cambridge ; New York, Cambridge University Press.

HALDAR, P. & PAVORD, I. D. 2007. Noneosinophilic asthma: a distinct clinical and pathologic phenotype. *J Allergy Clin Immunol,* 119, 1043-52; quiz 1053-4.

HALDAR, P., PAVORD, I. D., SHAW, D. E., BERRY, M. A., THOMAS, M., BRIGHTLING, C. E., WARDLAW, A. J. & GREEN, R. H. 2008. Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med,* 178, 218-24.

HAN, C. & KRONMAL, R. 2004. Box-Cox transformation of left-censored data with application to the analysis of coronary artery calcification and pharmacokinetic data. *Stat Med,* 23, 3671-9.

HANSEN, K. D., TIMP, W., BRAVO, H. C., SABUNCIYAN, S., LANGMEAD, B., MCDONALD, O. G., WEN, B., WU, H., LIU, Y., DIEP, D., BRIEM, E., ZHANG, K., IRIZARRY, R. A. & FEINBERG, A. P. 2011. Increased methylation variation in epigenetic domains across cancer types. *Nat Genet,* 43, 768-75.

HANSEN, S., PROBST-HENSCH, N., KEIDEL, D., DRATVA, J., BETTSCHART, R., PONS, M., BURDET, L., BRIDEVAUX, P.-O., SCHIKOWSKI, T., SCHINDLER, C., ROCHAT, T. & ZEMP, E. 2015. Gender differences in adult-onset asthma: results from the Swiss SAPALDIA cohort study. *European Respiratory Journal*.

HE, L., KERNOGITSKI, Y., KULMINSKAYA, I., LOIKA, Y., ARBEEV, K. G., LOIKO, E., BAGLEY, O., DUAN, M., YASHKIN, A., UKRAINTSEVA, S. V., KOVTUN, M., YASHIN, A. I. & KULMINSKI, A. M. 2016. Pleiotropic Meta-Analyses of Longitudinal Studies Discover Novel Genetic Variants Associated with Age-Related Diseases. *Front Genet,* 7 , 179.

HEBELS, D. G., GEORGIADIS, P., KEUN, H. C., ATHERSUCH, T. J., VINEIS, P., VERMEULEN, R., PORTENGEN, L., BERGDAHL, I. A., HALLMANS, G., PALLI, D., BENDINELLI, B., KROGH, V.,

TUMINO, R., SACERDOTE, C., PANICO, S., KLEINJANS, J. C., DE KOK, T. M., SMITH, M. T. & KYRTOPOULOS, S. A. 2013. Performance in omics analyses of blood samples in long-term storage: opportunities for the exploitation of existing biobanks in environmental health research. *Environ Health Perspect,* 121 , 480-7.

HEDBERG, A. & ROSSNER, S. 2000. Body weight characteristics of subjects on asthma medication. *Int J Obes Relat Metab Disord,* 24 , 1217-25.

HEERBOTH, S., LAPINSKA, K., SNYDER, N., LEARY, M., ROLLINSON, S. & SARKAR, S. 2014. Use of epigenetic drugs in disease: an overview. *Genet Epigenet,* 6, 9-19.

HEKKING, P. W., AMELINK, M., WENER, R. R., BOUVY, M. L. & BEL, E. H. 2017. Comorbidities in Difficult-to-Control Asthma. *J Allergy Clin Immunol Pract.*

HENDERSON, J., GRANELL, R., HERON, J., SHERRIFF, A., SIMPSON, A., WOODCOCK, A., STRACHAN, D. P., SHAHEEN, S. O. & STERNE, J. A. 2008. Associations of wheezing phenotypes in the first 6 years of life with atopy, lung function and airway responsiveness in mid-childhood. *Thorax,* 63, 974-80.

HERBERT, C. & KUMAR, R. K. 2017. Ambient air pollution and asthma. *Eur Respir J,* 49.

HJELLVIK, V., TVERDAL, A. & FURU, K. 2010. Body mass index as predictor for asthma: a cohort study of 118,723 males and females. *Eur Respir J,* 35, 1235-42.

HOLGUIN, F., BLEECKER, E. R., BUSSE, W. W., CALHOUN, W. J., CASTRO, M., ERZURUM, S. C., FITZPATRICK, A. M., GASTON, B., ISRAEL, E., JARJOUR, N. N., MOORE, W. C., PETERS, S. P., YONAS, M., TEAGUE, W. G. & WENZEL, S. E. 2011. Obesity and asthma: an association modified by age of asthma onset. *J Allergy Clin Immunol,* 127, 1486-93 e2.

HONG, E. L., SLOAN, C. A., CHAN, E. T., DAVIDSON, J. M., MALLADI, V. S., STRATTAN, J. S., HITZ, B. C., GABDANK, I., NARAYANAN, A. K., HO, M., LEE, B. T., ROWE, L. D., DRESZER, T. R., ROE, G. R., PODDUTURI, N. R., TANAKA, F., HILTON, J. A. & CHERRY, J. M. 2016. Principles of metadata organization at the ENCODE data coordination center. *Database (Oxford),* 2016.

HOUSEMAN, E. A., ACCOMANDO, W. P., KOESTLER, D. C., CHRISTENSEN, B. C., MARSIT, C. J., NELSON, H. H., WIENCKE, J. K. & KELSEY, K. T. 2012. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics,* 13, 86.

HUANG DA, W., SHERMAN, B. T. & LEMPICKI, R. A. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc,* 4, 44-57.

HUBBARD, T., BARKER, D., BIRNEY, E., CAMERON, G., CHEN, Y., CLARK, L., COX, T., CUFF, J., CURWEN, V., DOWN, T., DURBIN, R., EYRAS, E., GILBERT, J., HAMMOND, M., HUMINIECKI, L., KASPRZYK, A., LEHVASLAIHO, H., LIJNZAAD, P., MELSOPP, C., MONGIN, E., PETTETT, R., POCOCK, M., POTTER, S., RUST, A., SCHMIDT, E., SEARLE, S., SLATER, G., SMITH, J., SPOONER, W., STABENAU, A., STALKER, J., STUPKA, E., URETA-VIDAL, A., VASTRIK, I. & CLAMP, M. 2002. The Ensembl genome database project. *Nucleic Acids Res,* 30, 38-41.

IARC 2006. Preamble to the IARC Monographs. *Scientific Review and Evaluation. http://monographs. iarc. fr/ENG/Preamble/currentbscientificintro0706. php. Published.*

IRIZARRY, R. A., LADD-ACOSTA, C., WEN, B., WU, Z., MONTANO, C., ONYANGO, P., CUI, H., GABO, K., RONGIONE, M., WEBSTER, M., JI, H., POTASH, J., SABUNCIYAN, S. & FEINBERG, A. P. 2009. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet,* 41, 178-186.

JACQUEMIN, B., SCHIKOWSKI, T., CARSIN, A. E., HANSELL, A., KRAMER, U., SUNYER, J., PROBST-HENSCH, N., KAUFFMANN, F. & KUNZLI, N. 2012. The role of air pollution in adult-onset asthma: a review of the current evidence. *Semin Respir Crit Care Med,* 33, 606-19.

JACQUEMIN, B., SIROUX, V., SANCHEZ, M., CARSIN, A. E., SCHIKOWSKI, T., ADAM, M., BELLIS-ARIO, V., BUSCHKA, A., BONO, R., BRUNEKREEF, B., CAI, Y., CIRACH, M., CLAVEL-CHAPELON, F., DECLERCQ, C., DE MARCO, R., DE NAZELLE, A., DUCRET-STICH, R. E., FERRETTI, V. V., GERBASE, M. W., HARDY, R., HEINRICH, J., JANSON, C., JARVIS, D., AL KANAANI, Z., KEI-DEL, D., KUH, D., LE MOUAL, N., NIEUWENHUIJSEN, M. J., MARCON, A., MODIG, L., PIN, I., ROCHAT, T., SCHINDLER, C., SUGIRI, D., STEMPFELET, M., TEMAM, S., TSAI, M. Y., VARRASO, R., VIENNEAU, D., VIERKOTTER, A., HANSELL, A. L., KRAMER, U., PROBST-HENSCH, N. M., SUNYER, J., KUNZLI, N. & KAUFFMANN, F. 2015. Ambient air pollution and adult asthma incidence in six European cohorts (ESCAPE). *Environ Health Perspect,* 123 , 613-21.

JECK, W. R., SIEBOLD, A. P. & SHARPLESS, N. E. 2012. Review: a meta-analysis of GWAS and age-associated diseases. *Aging Cell,* 11, 727-31.

JEONG, A., IMBODEN, M., HANSEN, S., ZEMP, E., BRIDEVAUX, P. O., LOVISON, G., SCHINDLER, C. & PROBST-HENSCH, N. 2017. Heterogeneity of obesity-asthma association disentangled by latent class analysis, the SAPALDIA cohort. *Respir Med,* 125, 25-32.

JOEHANES, R., JUST, A. C., MARIONI, R. E., PILLING, L. C., REYNOLDS, L. M., MANDAVIYA, P. R., GUAN, W., XU, T., ELKS, C. E., ASLIBEKYAN, S., MORENO-MACIAS, H., SMITH, J. A., BRODY, J. A., DHINGRA, R., YOUSEFI, P., PANKOW, J. S., KUNZE, S., SHAH, S. H., MCRAE, A. F., LOHMAN, K., SHA, J., ABSHER, D. M., FERRUCCI, L., ZHAO, W., DEMERATH, E. W., BRESSLER, J., GROVE, M. L., HUAN, T., LIU, C., MENDELSON, M. M., YAO, C., KIEL, D. P., PETERS, A., WANG-SATTLER, R., VISSCHER, P. M., WRAY, N. R., STARR, J. M., DING, J., RODRIGUEZ, C. J., WAREHAM, N. J., IRVIN, M. R., ZHI, D., BARRDAHL, M., VINEIS, P., AMBATIPUDI, S., UITTERLINDEN, A. G., HOFMAN, A., SCHWARTZ, J., COLICINO, E., HOU, L., VOKONAS, P. S., HERNANDEZ, D. G., SINGLETON, A. B., BANDINELLI, S., TURNER, S. T., WARE, E. B., SMITH, A. K., KLENGEL, T., BINDER, E. B., PSATY, B. M., TAYLOR, K. D., GHARIB, S. A., SWENSON, B. R., LIANG, L., DEMEO, D. L., O'CONNOR, G. T., HERCEG, Z., RESSLER, K. J., CONNEELY, K. N., SOTOODEHNIA, N., KARDIA, S. L., MELZER, D., BACCARELLI, A. A., VAN MEURS, J. B., ROMIEU, I., ARNETT, D. K., ONG, K. K., LIU, Y., WALDENBERGER, M., DEARY, I. J., FORNAGE, M., LEVY, D. & LONDON, S. J. 2016. Epigenetic Signatures of Cigarette Smoking. *Circ Cardiovasc Genet,* 9, 436-447.

JONES, P. A. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics,* 13 , 484.

KANEHISA, M. & GOTO, S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res,* 28, 27-30.

KANEHISA, M., GOTO, S., HATTORI, M., AOKI-KINOSHITA, K. F., ITOH, M., KAWASHIMA, S., KATAYAMA, T., ARAKI, M. & HIRAKAWA, M. 2006. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res,* 34, D354-7.

KELLY, F. J. & FUSSELL, J. C. 2012. Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter. *Atmospheric environment,* 60, 504-526.

KELLY, R. S., DAHLIN, A., MCGEACHIE, M. J., QIU, W., SORDILLO, J., WAN, E. S., WU, A. C. & LASKY-SU, J. 2017. Asthma Metabolomics and the Potential for Integrative Omics in Research and the Clinic. *Chest,* 151, 262-277.

KIM, H. Y., LEE, H. J., CHANG, Y. J., PICHAVANT, M., SHORE, S. A., FITZGERALD, K. A., IWAKURA, Y., ISRAEL, E., BOLGER, K., FAUL, J., DEKRUYFF, R. H. & UMETSU, D. T. 2014. Interleukin-17-producing innate lymphoid cells and the NLRP3 inflammasome facilitate obesity-associated airway hyperreactivity. *Nat Med,* 20, 54-61.

KIM, S. K., KWON, Y. H., CHO, J. H., LEE, D. Y., PARK, S. E., OH, H. G., PARK, C. Y., LEE, W. Y., OH, K. W., PARK, S. W. & RHEE, E. J. 2017. Changes in Body Composition According to Age and Sex among Young Non-Diabetic Korean Adults: The Kangbuk Samsung Health Study. *Endocrinol Metab (Seoul),* 32, 442-450.

KORDALEWSKA, M. & MARKUSZEWSKI, M. J. 2015. Metabolomics in cardiovascular diseases. *J Pharm Biomed Anal,* 113, 121-36.

KUMAR, R. K., SHADIE, A. M., BUCKNALL, M. P., RUTLIDGE, H., GARTHWAITE, L., HERBERT, C., HALLIBURTON, B., PARSONS, K. S. & WARK, P. A. 2015. Differential injurious effects of ambient and traffic-derived particulate matter on airway epithelial cells. *Respirology,* 20, 73-79.

KUNZLI, N., BRIDEVAUX, P. O., LIU, L. J., GARCIA-ESTEBAN, R., SCHINDLER, C., GERBASE, M. W., SUNYER, J., KEIDEL, D. & ROCHAT, T. 2009. Traffic-related air pollution correlates with adult-onset asthma among never-smokers. *Thorax,* 64, 664-70.

KYLE, U. G., GENTON, L., KARSEGARD, L., SLOSMAN, D. O. & PICHARD, C. 2001. Single prediction equation for bioelectrical impedance analysis in adults aged 20–94 years. *Nutrition,* 17, 248-53.

LEHNE, B., DRONG, A. W., LOH, M., ZHANG, W., SCOTT, W. R., TAN, S. T., AFZAL, U., SCOTT, J., JARVELIN, M. R., ELLIOTT, P., MCCARTHY, M. I., KOONER, J. S. & CHAMBERS, J. C. 2015. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol,* 16, 37.

LI, D., XIE, Z., PAPE, M. L. & DYE, T. 2015. An evaluation of statistical methods for DNA methylation microarray data analysis. *BMC Bioinformatics,* 16, 217.

LI, N., GEORAS, S., ALEXIS, N., FRITZ, P., XIA, T., WILLIAMS, M. A., HORNER, E. & NEL, A. 2016. A work group report on ultrafine particles (American Academy of Allergy, Asthma & Immunology): Why ambient ultrafine and engineered nanoparticles should receive special attention for possible adverse health outcomes in human subjects. *J Allergy Clin Immunol,* 138, 386-96.

LI, S., PARK, Y., DURAISINGHAM, S., STROBEL, F. H., KHAN, N., SOLTOW, Q. A., JONES, D. P. & PULENDRAN, B. 2013. Predicting network activity from high throughput metabolomics. *PLoS Comput Biol,* 9**,** e1003123.

LINDEN, A. & DAHLEN, B. 2014. Interleukin-17 cytokine signalling in patients with asthma. *Eur Respir J,* 44, 1319-31.

LINZER, D. A. & LEWIS, J. B. 2011. poLCA: An R Package for Polytomous Variable Latent Class Analysis. *Journal of Statistical Software,* 42, 1-29.

LOPEZ DE MATURANA, E., PINEDA, S., BRAND, A., VAN STEEN, K. & MALATS, N. 2016. Toward the integration of Omics data in epidemiological studies: still a "long and winding road". *Genet Epidemiol,* 40 , 558-569.

LOVKVIST, C., DODD, I. B., SNEPPEN, K. & HAERTER, J. O. 2016. DNA methylation in human epigenomes depends on local topology of CpG sites. *Nucleic Acids Res,* 44, 5123-32.

LOZA, M. J., MCCALL, C. E., LI, L., ISAACS, W. B., XU, J. & CHANG, B. L. 2007. Assembly of inflammation-related genes for pathway-focused genetic analysis. *PLoS One,* 2, e1035.

MA, H., SOROKIN, A., MAZEIN, A., SELKOV, A., SELKOV, E., DEMIN, O. & GORYANIN, I. 2007. The Edinburgh human metabolic network reconstruction and its functional analysis. *Mol Syst Biol,* 3, 135.

MA, J. & XIAO, L. 2013. Association of general and central obesity and atopic and nonatopic asthma in US adults. *J Asthma,* 50 , 395-402.

MARTIN, B. W., ACKERMANN-LIEBRICH, U., LEUENBERGER, P., KUNZLI, N., STUTZ, E. Z., KELLER, R., ZELLWEGER, J. P., WUTHRICH, B., MONN, C., BLASER, K., BOLOGNINI, G., BONGARD, J. P., BRANDLI, O., BRAUN, P., DEFILA, C., DOMENIGHETTI, G., GRIZE, L., KARRER, W., KELLER-WOSSIDLO, H., MEDICI, T. C., PEETERS, A., PERRUCHOUD, A. P., SCHINDLER, C., SCHOENI, M. H., VILLIGER, B. & ET AL. 1997. SAPALDIA: methods and participation in the cross-sectional part of the Swiss Study on Air Pollution and Lung Diseases in Adults. *Soz Praventivmed,* 42 , 67-84.

MARUYAMA, H., TAKAHASHI, M., SEKIMOTO, T., SHIMADA, T. & YOKOSUKA, O. 2014. Linoleate appears to protect against palmitate-induced inflammation in Huh7 cells. *Lipids Health Dis,* 13**,** 78.

MCGEACHIE, M. J., DAHLIN, A., QIU, W., CROTEAU-CHONKA, D. C., SAVAGE, J., WU, A. C., WAN, E. S., SORDILLO, J. E., AL-GARAWI, A., MARTINEZ, F. D., STRUNK, R. C., LEMANSKE, R. F.,

JR., LIU, A. H., RABY, B. A., WEISS, S., CLISH, C. B. & LASKY-SU, J. A. 2015. The metabolomics of asthma control: a promising link between genetics and disease. *Immun Inflamm Dis,* 3, 224-38.

MCGUINN, L. A., WARD-CAVINESS, C. K., NEAS, L. M., SCHNEIDER, A., DIAZ-SANCHEZ, D., CASCIO, W. E., KRAUS, W. E., HAUSER, E., DOWDY, E., HAYNES, C., CHUDNOVSKY, A., KOUTRAKIS, P. & DEVLIN, R. B. 2016. Association between satellite-based estimates of long-term PM2.5 exposure and coronary artery disease. *Environ Res,* 145 , 9-17.

MCLACHLAN, C. R., POULTON, R., CAR, G., COWAN, J., FILSELL, S., GREENE, J. M., TAYLOR, D. R., WELCH, D., WILLIAMSON, A., SEARS, M. R. & HANCOX, R. J. 2007. Adiposity, asthma, and airway inflammation. *J Allergy Clin Immunol,* 119, 634-9.

MIGUEL-CARRASCO, J. L., MATE, A., MONSERRAT, M. T., ARIAS, J. L., ARAMBURU, O. & VAZQUEZ, C. M. 2008. The role of inflammatory markers in the cardioprotective effect of L-carnitine in L-NAME-induced hypertension. *Am J Hypertens,* 21, 1231-7.

MILLARD, L. A., DAVIES, N. M., TIMPSON, N. J., TILLING, K., FLACH, P. A. & DAVEY SMITH, G. 2015. MR-PheWAS: hypothesis prioritization among potential causal effects of body mass index on many outcomes, using Mendelian randomization. *Sci Rep,* 5, 16645.

MIYAKE, Y., TANAKA, K., SASAKI, S. & ARAKAWA, M. 2011. Polyunsaturated fatty acid intake and prevalence of eczema and rhinoconjunctivitis in Japanese children: the Ryukyus Child Health Study. *BMC Public Health,* 11, 358.

MIZUTA, K., ZHANG, Y., MIZUTA, F., HOSHIJIMA, H., SHIGA, T., MASAKI, E. & EMALA, C. W., SR. 2015. Novel identification of the free fatty acid receptor FFAR1 that promotes contraction in airway smooth muscle. *Am J Physiol Lung Cell Mol Physiol,* 309, L970-82.

MOFFATT, M. F., GUT, I. G., DEMENAIS, F., STRACHAN, D. P., BOUZIGON, E., HEATH, S., VON MUTIUS, E., FARRALL, M., LATHROP, M. & COOKSON, W. O. 2010. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med,* 363, 1211-21.

MOORE, W. C., MEYERS, D. A., WENZEL, S. E., TEAGUE, W. G., LI, H., LI, X., D'AGOSTINO, R., JR., CASTRO, M., CURRAN-EVERETT, D., FITZPATRICK, A. M., GASTON, B., JARJOUR, N. N., SORKNESS, R., CALHOUN, W. J., CHUNG, K. F., COMHAIR, S. A., DWEIK, R. A., ISRAEL, E., PETERS, S. P., BUSSE, W. W., ERZURUM, S. C. & BLEECKER, E. R. 2010. Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med,* 181, 315-23.

MOSTAFAVI, N., VLAANDEREN, J., CHADEAU-HYAM, M., BEELEN, R., MODIG, L., PALLI, D., BERGDAHL, I. A., VINEIS, P., HOEK, G., KYRTOPOULOS, S. & VERMEULEN, R. 2015. Inflammatory markers in relation to long-term air pollution. *Environ Int,* 81, 1-7.

NATIONAL ACADEMIES OF SCIENCES, E. & MEDICINE 2017. *Using 21st century science to improve risk-related evaluations*, National Academies Press.

NEWBY, C., HEANEY, L. G., MENZIES-GOW, A., NIVEN, R. M., MANSUR, A., BUCKNALL, C., CHAUDHURI, R., THOMPSON, J., BURTON, P., BRIGHTLING, C. & BRITISH THORACIC SOCIETY SEVERE REFRACTORY ASTHMA, N. 2014. Statistical cluster analysis of the British Thoracic Society Severe refractory Asthma Registry: clinical outcomes and phenotype stability. *PLoS One,* 9, e102987.

NEWBY, D. E., MANNUCCI, P. M., TELL, G. S., BACCARELLI, A. A., BROOK, R. D., DONALDSON, K., FORASTIERE, F., FRANCHINI, M., FRANCO, O. H., GRAHAM, I., HOEK, G., HOFFMANN, B., HOYLAERTS, M. F., KUNZLI, N., MILLS, N., PEKKANEN, J., PETERS, A., PIEPOLI, M. F., RAJAGOPALAN, S. & STOREY, R. F. 2015. Expert position paper on air pollution and cardiovascular disease. *Eur Heart J,* 36, 83-93b.

NEWMAN, A. B., LEE, J. S., VISSER, M., GOODPASTER, B. H., KRITCHEVSKY, S. B., TYLAVSKY, F. A., NEVITT, M. & HARRIS, T. B. 2005. Weight change and the conservation of lean mass in old age: the Health, Aging and Body Composition Study. *Am J Clin Nutr,* 82, 872-8; quiz 915-6.

NIH 2018. All of Us Research Program.

ONI, T. & UNWIN, N. 2015. Why the communicable/non-communicable disease dichotomy is problematic for public health control strategies: implications of multimorbidity for health systems in an era of health transition. *Int Health,* 7, 390-9.

ONO, J. G., WORGALL, T. S. & WORGALL, S. 2015. Airway reactivity and sphingolipids-implications for childhood asthma. *Mol Cell Pediatr,* 2, 13.

OSTRO, B., HU, J., GOLDBERG, D., REYNOLDS, P., HERTZ, A., BERNSTEIN, L. & KLEEMAN, M. J. 2015. Associations of mortality with long-term exposures to fine and ultrafine particles, species and sources: results from the California Teachers Study Cohort. *Environ Health Perspect,* 123, 549-56.

PALLI, D., BERRINO, F., VINEIS, P., TUMINO, R., PANICO, S., MASALA, G., SAIEVA, C., SALVINI, S., CEROTI, M., PALA, V., SIERI, S., FRASCA, G., GIURDANELLA, M. C., SACERDOTE, C., FIORINI, L., CELENTANO, E., GALASSO, R., DECARLI, A. & KROGH, V. 2003. A molecular epidemiology project on diet and cancer: the EPIC-Italy Prospective Study. Design and baseline characteristics of participants. *Tumori,* 89, 586-93.

PAN, W., YU, J., SHI, R., YAN, L., YANG, T., LI, Y., ZHANG, Z., YU, G., BAI, Y., SCHUCHMAN, E. H., HE, X. & ZHANG, G. 2014. Elevation of ceramide and activation of secretory acid sphingomyelinase in patients with acute coronary syndromes. *Coron Artery Dis,* 25 , 230-5.

PARK, J. Y., LEE, S. H., SHIN, M. J. & HWANG, G. S. 2015. Alteration in metabolic signature and lipid metabolism in patients with angina pectoris and myocardial infarction. *PLoS One,* 10, e0135228.

PAUWELS, R. A., BUIST, A. S., CALVERLEY, P. M., JENKINS, C. R. & HURD, S. S. 2001. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. *Am J Respir Crit Care Med,* 163, 1256-76.

PEARCE, N., ASHER, I., BILLO, N., BISSELL, K., ELLWOOD, P., EL SONY, A., GARCIA-MARCOS, L., CHIANG, C. Y., MALLOL, J., MARKS, G. & STRACHAN, D. 2013. Asthma in the global NCD agenda: a neglected epidemic. *Lancet Respir Med,* 1, 96-8.

PEEL, J. L., TOLBERT, P. E., KLEIN, M., METZGER, K. B., FLANDERS, W. D., TODD, K., MULHOLLAND, J. A., RYAN, P. B. & FRUMKIN, H. 2005. Ambient air pollution and respiratory emergency department visits. *Epidemiology,* 16, 164-74.

PERRY, M. M., LAVENDER, P., KUO, C. S., GALEA, F., MICHAELOUDES, C., FLANAGAN, J. M., FAN CHUNG, K. & ADCOCK, I. M. 2018. DNA methylation modules in airway smooth muscle are associated with asthma severity. *Eur Respir J,* 51.

PERRY, P. O. 2016. Fast moment-based estimation for hierarchical models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology).*

PETERS, A., WICHMANN, H. E., TUCH, T., HEINRICH, J. & HEYDER, J. 1997. Respiratory effects are associated with the number of ultrafine particles. *American journal of respiratory and critical care medicine,* 155, 1376-1383.

PETERS, T. J., BUCKLEY, M. J., STATHAM, A. L., PIDSLEY, R., SAMARAS, K., V LORD, R., CLARK, S. J. & MOLLOY, P. L. 2015. De novo identification of differentially methylated regions in the human genome. *Epigenetics & Chromatin,* 8, 6.

PETERS, U., DIXON, A. E. & FORNO, E. 2018. Obesity and asthma. *J Allergy Clin Immunol,* 141, 1169-1179.

PETRACHE, I. & BERDYSHEV, E. V. 2016. Ceramide Signaling and Metabolism in Pathophysiological States of the Lung. *Annu Rev Physiol,* 78, 463-80.

PICKUP, J. C. 2004. Inflammation and activated innate immunity in the pathogenesis of type 2 diabetes. *Diabetes Care,* 27**,** 813-23.

PRUITT, K. D., TATUSOVA, T. & MAGLOTT, D. R. 2005. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res,* 33, D501-4.

RAPPAPORT, S. M., BARUPAL, D. K., WISHART, D., VINEIS, P. & SCALBERT, A. 2014. The blood exposome and its role in discovering causes of disease. *Environ Health Perspect,* 122, 769-74.

RATTRAY, N. J. W., DEZIEL, N. C., WALLACH, J. D., KHAN, S. A., VASILIOU, V., IOANNIDIS, J. P. A. & JOHNSON, C. H. 2018. Beyond genomics: understanding exposotypes through metabolomics. *Hum Genomics,* 12, 4.

R Core Team. 2015. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2013. *Document freely available on the internet at: http://www. r-project. org.*

REESE, S. E., XU, C. J., DEN DEKKER, H. T., LEE, M. K., SIKDAR, S., RUIZ-ARENAS, C., MERID, S. K., REZWAN, F. I., PAGE, C. M., ULLEMAR, V., MELTON, P. E., OH, S. S., YANG, I. V., BURROWS,

K., SODERHALL, C., JIMA, D. D., GAO, L., ARATHIMOS, R., KUPERS, L. K., WIELSCHER, M., RZEHAK, P., LAHTI, J., LAPRISE, C., MADORE, A. M., WARD, J., BENNETT, B. D., WANG, T., BELL, D. A., VONK, J. M., HABERG, S. E., ZHAO, S., KARLSSON, R., HOLLAMS, E., HU, D., RICHARDS, A. J., BERGSTROM, A., SHARP, G. C., FELIX, J. F., BUSTAMANTE, M., GRUZIEVA, O., MAGUIRE, R. L., GILLILAND, F., BAIZ, N., NOHR, E. A., CORPELEIJN, E., SEBERT, S., KARMAUS, W., GROTE, V., KAJANTIE, E., MAGNUS, M. C., ORTQVIST, A. K., ENG, C., LIU, A. H., KULL, I., JADDOE, V. W. V., SUNYER, J., KERE, J., HOYO, C., ANNESI-MAESANO, I., ARSHAD, S. H., KOLETZKO, B., BRUNEKREEF, B., BINDER, E. B., RAIKKONEN, K., REISCHL, E., HOLLOWAY, J. W., JARVELIN, M. R., SNIEDER, H., KAZMI, N., BRETON, C. V., MURPHY, S. K., PERSHAGEN, G., ANTO, J. M., RELTON, C. L., SCHWARTZ, D. A., BURCHARD, E. G., HUANG, R. C., NYSTAD, W., ALMQVIST, C., HENDERSON, A. J., MELEN, E., DUIJTS, L., KOPPELMAN, G. H. & LONDON, S. J. 2018. Epigenome-wide meta-analysis of DNA methylation and childhood asthma. *J Allergy Clin Immunol*.

REMAIS, J. V., ZENG, G., LI, G., TIAN, L. & ENGELGAU, M. M. 2013. Convergence of non-communicable and infectious diseases in low- and middle-income countries. *Int J Epidemiol,* 42, 221-7.

RICHARDSON, D. B. 2004. An incidence density sampling program for nested case-control analyses. *Occup Environ Med,* 61, e59.

RICHMOND, R. C., DAVEY SMITH, G., NESS, A. R., DEN HOED, M., MCMAHON, G. & TIMPSON, N. J. 2014. Assessing causality in the association between child adiposity and physical activity levels: a Mendelian randomization analysis. *PLoS Med,* 11, e1001618.

RIPLEY, B. D. 2002. *Modern applied statistics with S*, Springer.

SANTANASTO, A. J., GOODPASTER, B. H., KRITCHEVSKY, S. B., MILJKOVIC, I., SATTERFIELD, S., SCHWARTZ, A. V., CUMMINGS, S. R., BOUDREAU, R. M., HARRIS, T. B. & NEWMAN, A. B. 2017. Body Composition Remodeling and Mortality: The Health Aging and Body Composition Study. *J Gerontol A Biol Sci Med Sci,* 72, 513-519.

SAYERS, E. W., BARRETT, T., BENSON, D. A., BOLTON, E., BRYANT, S. H., CANESE, K., CHETVERNIN, V., CHURCH, D. M., DICUCCIO, M., FEDERHEN, S., FEOLO, M., FINGERMAN, I. M., GEER, L. Y., HELMBERG, W., KAPUSTIN, Y., LANDSMAN, D., LIPMAN, D. J., LU, Z., MADDEN, T. L., MADEJ, T., MAGLOTT, D. R., MARCHLER-BAUER, A., MILLER, V., MIZRACHI, I., OSTELL, J., PANCHENKO, A., PHAN, L., PRUITT, K. D., SCHULER, G. D., SEQUEIRA, E., SHERRY, S. T., SHUMWAY, M., SIROTKIN, K., SLOTTA, D., SOUVOROV, A., STARCHENKO, G., TATUSOVA, T. A., WAGNER, L., WANG, Y., WILBUR, W. J., YASCHENKO, E. & YE, J. 2011. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res,* 39 **,** D38-51.

SCHINDLER, C., KEIDEL, D., GERBASE, M. W., ZEMP, E., BETTSCHART, R., BRANDLI, O., BRUTSCHE, M. H., BURDET, L., KARRER, W., KNOPFLI, B., PONS, M., RAPP, R., BAYER-OGLESBY, L., KUNZLI, N., SCHWARTZ, J., LIU, L. J., ACKERMANN-LIEBRICH, U. & ROCHAT, T. 2009. Improvements

in PM10 exposure and reduced rates of respiratory symptoms in a cohort of Swiss adults (SAPALDIA). *Am J Respir Crit Care Med,* 179, 579-87.

SCHWARTZ, J., SLATER, D., LARSON, T. V., PIERSON, W. E. & KOENIG, J. Q. 1993. Particulate air pollution and hospital emergency room visits for asthma in Seattle. *Am Rev Respir Dis,* 147, 826-31.

SCHWARZE, P. E., OVREVIK, J., HETLAND, R. B., BECHER, R., CASSEE, F. R., LAG, M., LOVIK, M., DYBING, E. & REFSNES, M. 2007. Importance of size and composition of particles for effects on cells in vitro. *Inhal Toxicol,* 19 Suppl 1, 17-22.

SCHWARZE, P. E., OVREVIK, J., LAG, M., REFSNES, M., NAFSTAD, P., HETLAND, R. B. & DYBING, E. 2006. Particulate matter properties and health effects: consistency of epidemiological and toxicological studies. *Hum Exp Toxicol,* 25, 559-79.

SHAH, S. H., KRAUS, W. E. & NEWGARD, C. B. 2012. Metabolomic profiling for the identification of novel biomarkers and mechanisms related to common cardiovascular diseases: form and function. *Circulation,* 126, 1110-20.

SHORE, S. A. 2008. Obesity and asthma: possible mechanisms. *J Allergy Clin Immunol,* 121, 1087-93; quiz 1094-5.

SIDELEVA, O., SURATT, B. T., BLACK, K. E., THARP, W. G., PRATLEY, R. E., FORGIONE, P., DIENZ, O., IRVIN, C. G. & DIXON, A. E. 2012. Obesity and asthma: an inflammatory disease of adipose tissue not the airway. *Am J Respir Crit Care Med,* 186, 598-605.

SIMONS, L. A., SIMONS, J., FRIEDLANDER, Y. & MCCALLUM, J. 2009. A comparison of risk factors for coronary heart disease and ischaemic stroke: the Dubbo study of Australian elderly. *Heart Lung Circ,* 18, 330-3.

SIROUX, V., BASAGANA, X., BOUDIER, A., PIN, I., GARCIA-AYMERICH, J., VESIN, A., SLAMA, R., JARVIS, D., ANTO, J. M., KAUFFMANN, F. & SUNYER, J. 2011. Identifying adult asthma phenotypes using a clustering approach. *Eur Respir J,* 38, 310-7.

SIROUX, V., GONZALEZ, J. R., BOUZIGON, E., CURJURIC, I., BOUDIER, A., IMBODEN, M., ANTO, J. M., GUT, I., JARVIS, D., LATHROP, M., OMENAAS, E. R., PIN, I., WJST, M., DEMENAIS, F., PROBST-HENSCH, N., KOGEVINAS, M. & KAUFFMANN, F. 2014. Genetic heterogeneity of asthma phenotypes identified by a clustering approach. *Eur Respir J,* 43 , 439-52.

SKAABY, T., TAYLOR, A. E., THUESEN, B. H., JACOBSEN, R. K., FRIEDRICH, N., MOLLEHAVE, L. T., HANSEN, S., LARSEN, S. C., VOLKER, U., NAUCK, M., VOLZKE, H., HANSEN, T., PEDERSEN, O., JORGENSEN, T., PATERNOSTER, L., MUNAFO, M., GRARUP, N. & LINNEBERG, A. 2017. Estimating the causal effect of body mass index on hay fever, asthma and lung function using Mendelian randomization. *Allergy.*

SMITH, C. A., O'MAILLE, G., WANT, E. J., QIN, C., TRAUGER, S. A., BRANDON, T. R., CUSTODIO, D. E., ABAGYAN, R. & SIUZDAK, G. 2005. METLIN: a metabolite mass spectral database. *Ther Drug Monit,* 27, 747-51.

SOLER, E. P. & RUIZ, V. C. 2010. Epidemiology and risk factors of cerebral ischemia and ischemic heart diseases: similarities and differences. *Curr Cardiol Rev,* 6, 138-49.

SONG, X., QU, H., YANG, Z., RONG, J., CAI, W. & ZHOU, H. 2017. Efficacy and Safety of L-Carnitine Treatment for Chronic Heart Failure: A Meta-Analysis of Randomized Controlled Trials. *Biomed Res Int,* 2017, 6274854.

SPYCHER, B. D., SILVERMAN, M., BROOKE, A. M., MINDER, C. E. & KUEHNI, C. E. 2008. Distinguishing phenotypes of childhood wheeze and cough using latent class analysis. *Eur Respir J,* 31, 974-81.

STAFOGGIA, M., CESARONI, G., PETERS, A., ANDERSEN, Z. J., BADALONI, C., BEELEN, R., CARACCIOLO, B., CYRYS, J., DE FAIRE, U., DE HOOGH, K., ERIKSEN, K. T., FRATIGLIONI, L., GALASSI, C., GIGANTE, B., HAVULINNA, A. S., HENNIG, F., HILDING, A., HOEK, G., HOFFMANN, B., HOUTHUIJS, D., KOREK, M., LANKI, T., LEANDER, K., MAGNUSSON, P. K., MEISINGER, C., MIGLIORE, E., OVERVAD, K., OSTENSON, C. G., PEDERSEN, N. L., PEKKANEN, J., PENELL, J., PERSHAGEN, G., PUNDT, N., PYKO, A., RAASCHOU-NIELSEN, O., RANZI, A., RICCERI, F., SACERDOTE, C., SWART, W. J., TURUNEN, A. W., VINEIS, P., WEIMAR, C., WEINMAYR, G., WOLF, K., BRUNEKREEF, B. & FORASTIERE, F. 2014. Long-term exposure to ambient air pollution and incidence of cerebrovascular events: results from 11 European cohorts within the ESCAPE project. *Environ Health Perspect,* 122 , 919-25.

STEIER, J., LUNT, A., HART, N., POLKEY, M. I. & MOXHAM, J. 2014. Observational study of the effect of obesity on lung volumes. *Thorax,* 69, 752-9.

STOEGER, T., REINHARD, C., TAKENAKA, S., SCHROEPPEL, A., KARG, E., RITTER, B., HEYDER, J. & SCHULZ, H. 2006. Instillation of six different ultrafine carbon particles indicates a surface area threshold dose for acute lung inflammation in mice. *Environ Health Perspect,* 114 , 328-33.

SUBRAMANIAN, A., TAMAYO, P., MOOTHA, V. K., MUKHERJEE, S., EBERT, B. L., GILLETTE, M. A., PAULOVICH, A., POMEROY, S. L., GOLUB, T. R., LANDER, E. S. & MESIROV, J. P. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A,* 102, 15545-50.

SUGANAMI, T., NISHIDA, J. & OGAWA, Y. 2005. A paracrine loop between adipocytes and macrophages aggravates inflammatory changes: role of free fatty acids and tumor necrosis factor alpha. *Arterioscler Thromb Vasc Biol,* 25, 2062-8.

SUMNER, L. W., AMBERG, A., BARRETT, D., BEALE, M. H., BEGER, R., DAYKIN, C. A., FAN, T. W., FIEHN, O., GOODACRE, R., GRIFFIN, J. L., HANKEMEIER, T., HARDY, N., HARNLY, J., HIGASHI, R., KOPKA, J., LANE, A. N., LINDON, J. C., MARRIOTT, P., NICHOLLS, A. W., REILY, M. D., THADEN, J. J. & VIANT, M. R. 2007. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics,* 3, 211-221.

SUNYER, J., SPIX, C., QUENEL, P., PONCE-DE-LEON, A., PONKA, A., BARUMANDZADEH, T., TOULOUMI, G., BACHAROVA, L., WOJTYNIAK, B., VONK, J., BISANTI, L., SCHWARTZ, J. & KATSOUYANNI, K. 1997. Urban air pollution and emergency admissions for asthma in four European cities: the APHEA Project. *Thorax,* 52, 760-5.

TEMAM, S., BURTE, E., ADAM, M., ANTO, J. M., BASAGANA, X., BOUSQUET, J., CARSIN, A. E., GALOBARDES, B., KEIDEL, D., KUNZLI, N., LE MOUAL, N., SANCHEZ, M., SUNYER, J., BONO, R., BRUNEKREEF, B., HEINRICH, J., DE HOOGH, K., JARVIS, D., MARCON, A., MODIG, L., NADIF, R., NIEUWENHUIJSEN, M., PIN, I., SIROUX, V., STEMPFELET, M., TSAI, M. Y., PROBST-HENSCH, N. & JACQUEMIN, B. 2017. Socioeconomic position and outdoor nitrogen dioxide (NO2) exposure in Western Europe: A multi-city analysis. *Environ Int,* 101, 117-124.

TESCHENDORFF, A. E., MARABITA, F., LECHNER, M., BARTLETT, T., TEGNER, J., GOMEZ-CABRERO, D. & BECK, S. 2013. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics,* 29 , 189-96.

TIMPSON, N. J., NORDESTGAARD, B. G., HARBORD, R. M., ZACHO, J., FRAYLING, T. M., TYBJARG-HANSEN, A. & SMITH, G. D. 2011. C-reactive protein levels and body mass index: elucidating direction of causation through reciprocal Mendelian randomization. *International journal of obesity,* 35, 300-308.

TODD, D. C., ARMSTRONG, S., D'SILVA, L., ALLEN, C. J., HARGREAVE, F. E. & PARAMESWARAN, K. 2007. Effect of obesity on airway inflammation: a cross-sectional analysis of body mass index and sputum cell counts. *Clin Exp Allergy,* 37, 1049-54.

TOUGH, D. F., TAK, P. P., TARAKHOVSKY, A. & PRINJHA, R. K. 2016. Epigenetic drug discovery: breaking through the immune barrier. *Nat Rev Drug Discov,* 15, 835-853.

TRICHE, T. J., JR., WEISENBERGER, D. J., VAN DEN BERG, D., LAIRD, P. W. & SIEGMUND, K. D. 2013. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res,* 41, e90.

TURNER, M. C., VINEIS, P., SELEIRO, E., DIJMARESCU, M., BALSHAW, D., BERTOLLINI, R., CHADEAU-HYAM, M., GANT, T., GULLIVER, J., JEONG, A., KYRTOPOULOS, S., MARTUZZI, M., MILLER, G. W., NAWROT, T., NIEUWENHUIJSEN, M., PHILLIPS, D. H., PROBST-HENSCH, N., SAMET, J., VERMEULEN, R., VLAANDEREN, J., VRIJHEID, M., WILD, C. & KOGEVINAS, M. 2018. EXPOsOMICS: final policy workshop and stakeholder consultation. *BMC Public Health,* 18, 260.

TZOULAKI, I., EBBELS, T. M., VALDES, A., ELLIOTT, P. & IOANNIDIS, J. P. 2014. Design and analysis of metabolomics studies in epidemiologic research: a primer on-omic technologies. *American journal of epidemiology,* 180, 129-139.

ULLEMAR, V., MAGNUSSON, P. K., LUNDHOLM, C., ZETTERGREN, A., MELEN, E., LICHTENSTEIN, P. & ALMQVIST, C. 2016. Heritability and confirmation of genetic association studies for childhood asthma in twins. *Allergy,* 71, 230-8.

UPPAL, K., WALKER, D. I. & JONES, D. P. 2017. xMSannotator: An R Package for Network-Based Annotation of High-Resolution Metabolomics Data. *Anal Chem,* 89, 1063-1067.

UZOIGWE, J. C., PRUM, T., BRESNAHAN, E. & GARELNABI, M. 2013. The emerging role of outdoor and indoor air pollution in cardiovascular disease. *N Am J Med Sci,* 5, 445-53.

VAN DEN BERGE, M., HEIJINK, H. I., VAN OOSTERHOUT, A. J. & POSTMA, D. S. 2009. The role of female sex hormones in the development and severity of allergic and non-allergic asthma. *Clin Exp Allergy,* 39, 1477-81.

VAN DER LAAN, M. J., BIRKNER, M. D. & HUBBARD, A. E. 2005. Empirical Bayes and resampling based multiple testing procedure controlling tail probability of the proportion of false positives. *Stat Appl Genet Mol Biol,* 4, 29.

VAN NUNEN, E., VERMEULEN, R., TSAI, M. Y., PROBST-HENSCH, N., INEICHEN, A., DAVEY, M., IMBODEN, M., DUCRET-STICH, R., NACCARATI, A., RAFFAELE, D., RANZI, A., IVALDI, C., GALASSI, C., NIEUWENHUIJSEN, M., CURTO, A., DONAIRE-GONZALEZ, D., CIRACH, M., CHATZI, L., KAMPOURI, M., VLAANDEREN, J., MELIEFSTE, K., BUIJTENHUIJS, D., BRUNEKREEF, B., MORLEY, D., VINEIS, P., GULLIVER, J. & HOEK, G. 2017. Land Use Regression Models for Ultrafine Particles in Six European Areas. *Environ Sci Technol,* 51, 3336-3345.

VANDANMAGSAR, B., YOUM, Y. H., RAVUSSIN, A., GALGANI, J. E., STADLER, K., MYNATT, R. L., RAVUSSIN, E., STEPHENS, J. M. & DIXIT, V. D. 2011. The NLRP3 inflammasome instigates obesity-induced inflammation and insulin resistance. *Nat Med,* 17, 179-88.

VINEIS, P., CHADEAU-HYAM, M., GMUENDER, H., GULLIVER, J., HERCEG, Z., KLEINJANS, J., KOGEVINAS, M., KYRTOPOULOS, S., NIEUWENHUIJSEN, M., PHILLIPS, D. H., PROBST-HENSCH, N., SCALBERT, A., VERMEULEN, R. & WILD, C. P. 2016. The exposome in practice: Design of the EXPOsOMICS project. *Int J Hyg Environ Health.*

VINEIS, P. & PERERA, F. 2007. Molecular epidemiology and biomarkers in etiologic cancer research: the new in light of the old. *Cancer Epidemiol Biomarkers Prev,* 16, 1954-65.

VINEIS, P., VAN VELDHOVEN, K., CHADEAU-HYAM, M. & ATHERSUCH, T. J. 2013. Advancing the application of omics-based biomarkers in environmental epidemiology. *Environ Mol Mutagen,* 54 **,** 461-7.

VLAANDEREN, J. J., JANSSEN, N. A., HOEK, G., KESKI-RAHKONEN, P., BARUPAL, D. K., CASSEE, F. R., GOSENS, I., STRAK, M., STEENHOF, M., LAN, Q., BRUNEKREEF, B., SCALBERT, A. & VERMEULEN, R. C. H. 2017. The impact of ambient air pollution on the human blood metabolome. *Environ Res,* 156, 341-348.

WAHL, S., DRONG, A., LEHNE, B., LOH, M., SCOTT, W. R., KUNZE, S., TSAI, P. C., RIED, J. S., ZHANG, W., YANG, Y., TAN, S., FIORITO, G., FRANKE, L., GUARRERA, S., KASELA, S., KRIEBEL, J., RICHMOND, R. C., ADAMO, M., AFZAL, U., ALA-KORPELA, M., ALBETTI, B., AMMERPOHL, O., APPERLEY, J. F., BEEKMAN, M., BERTAZZI, P. A., BLACK, S. L., BLANCHER, C., BONDER, M. J., BROSCH, M., CARSTENSEN-KIRBERG, M., DE CRAEN, A. J., DE LUSIGNAN, S., DEHGHAN, A., ELKALAAWY, M., FISCHER, K., FRANCO, O. H., GAUNT, T. R., HAMPE, J., HASHEMI, M., ISAACS, A., JENKINSON, A., JHA, S., KATO, N., KROGH, V., LAFFAN, M., MEISINGER, C.,

MEITINGER, T., MOK, Z. Y., MOTTA, V., NG, H. K., NIKOLAKOPOULOU, Z., NTELIOPOULOS, G., PANICO, S., PERVJAKOVA, N., PROKISCH, H., RATHMANN, W., RODEN, M., ROTA, F., ROZARIO, M. A., SANDLING, J. K., SCHAFMAYER, C., SCHRAMM, K., SIEBERT, R., SLAGBOOM, P. E., SOININEN, P., STOLK, L., STRAUCH, K., TAI, E. S., TARANTINI, L., THORAND, B., TIGCHELAAR, E. F., TUMINO, R., UITTERLINDEN, A. G., VAN DUIJN, C., VAN MEURS, J. B., VINEIS, P., WICKREMASINGHE, A. R., WIJMENGA, C., YANG, T. P., YUAN, W., ZHERNAKOVA, A., BATTERHAM, R. L., SMITH, G. D., DELOUKAS, P., HEIJMANS, B. T., HERDER, C., HOFMAN, A., LINDGREN, C. M., MILANI, L., VAN DER HARST, P., PETERS, A., ILLIG, T., RELTON, C. L., WALDENBERGER, M., JARVELIN, M. R., BOLLATI, V., SOONG, R., SPECTOR, T. D., SCOTT, J., MCCARTHY, M. I., et al. 2017. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature,* 541, 81-86.

WANG, D. D., TOLEDO, E., HRUBY, A., ROSNER, B. A., WILLETT, W. C., SUN, Q., RAZQUIN, C., ZHENG, Y., RUIZ-CANELA, M., GUASCH-FERRE, M., CORELLA, D., GOMEZ-GRACIA, E., FIOL, M., ESTRUCH, R., ROS, E., LAPETRA, J., FITO, M., AROS, F., SERRA-MAJEM, L., LEE, C. H., CLISH, C. B., LIANG, L., SALAS-SALVADO, J., MARTINEZ-GONZALEZ, M. A. & HU, F. B. 2017. Plasma Ceramides, Mediterranean Diet, and Incident Cardiovascular Disease in the PREDIMED Trial (Prevencion con Dieta Mediterranea). *Circulation,* 135, 2028-2040.

WANG, R., CUSTOVIC, A., SIMPSON, A., BELGRAVE, D. C., LOWE, L. A. & MURRAY, C. S. 2014. Differing associations of BMI and body fat with asthma and lung function in children. *Pediatr Pulmonol,* 49 , 1049-57.

WEI, Y., ZHANG, J. J., LI, Z., GOW, A., CHUNG, K. F., HU, M., SUN, Z., ZENG, L., ZHU, T., JIA, G., LI, X., DUARTE, M. & TANG, X. 2016. Chronic exposure to air pollution particles increases the risk of obesity and metabolic syndrome: findings from a natural experiment in Beijing. *Faseb j,* 30, 2115-22.

WEICHENTHAL, S., BAI, L., HATZOPOULOU, M., VAN RYSWYK, K., KWONG, J. C., JERRETT, M., VAN DONKELAAR, A., MARTIN, R. V., BURNETT, R. T., LU, H. & CHEN, H. 2017. Long-term exposure to ambient ultrafine particles and respiratory disease incidence in in Toronto, Canada: a cohort study. *Environ Health,* 16, 64.

WEINMAYR, G., ROMEO, E., DE SARIO, M., WEILAND, S. K. & FORASTIERE, F. 2010. Short-term effects of PM10 and NO2 on respiratory health among children with asthma or asthma-like symptoms: a systematic review and meta-analysis. *Environ Health Perspect,* 118, 449-57.

WEISBERG, S. P., MCCANN, D., DESAI, M., ROSENBAUM, M., LEIBEL, R. L. & FERRANTE, A. W., JR. 2003. Obesity is associated with macrophage accumulation in adipose tissue. *J Clin Invest,* 112**,** 1796-808.

WENZEL, S. E. 2012. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat Med,* 18, 716-25.

WHO 2014. *Global status report on noncommunicable diseases 2014* .

WILD, C. P. 2005. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev,* 14**,** 1847-50.

WILD, C. P. 2012. The exposome: from concept to utility. *Int J Epidemiol,* 41, 24-32.

WISHART, D. S., TZUR, D., KNOX, C., EISNER, R., GUO, A. C., YOUNG, N., CHENG, D., JEWELL, K., ARNDT, D., SAWHNEY, S., FUNG, C., NIKOLAI, L., LEWIS, M., COUTOULY, M. A., FORSYTHE, I., TANG, P., SHRIVASTAVA, S., JERONCIC, K., STOTHARD, P., AMEGBEY, G., BLOCK, D., HAU, D. D., WAGNER, J., MINIACI, J., CLEMENTS, M., GEBREMEDHIN, M., GUO, N., ZHANG, Y., DUGGAN, G. E., MACINNIS, G. D., WELJIE, A. M., DOWLATABADI, R., BAMFORTH, F., CLIVE, D., GREINER, R., LI, L., MARRIE, T., SYKES, B. D., VOGEL, H. J. & QUERENGESSER, L. 2007. HMDB: the Human Metabolome Database. *Nucleic Acids Res,* 35, D521-6.

WOLF, K., STAFOGGIA, M., CESARONI, G., ANDERSEN, Z. J., BEELEN, R., GALASSI, C., HENNIG, F., MIGLIORE, E., PENELL, J., RICCERI, F., SORENSEN, M., TURUNEN, A. W., HAMPEL, R., HOFFMANN, B., KALSCH, H., LAATIKAINEN, T., PERSHAGEN, G., RAASCHOU-NIELSEN, O., SACERDOTE, C., VINEIS, P., BADALONI, C., CYRYS, J., DE HOOGH, K., ERIKSEN, K. T., JEDYNSKA, A., KEUKEN, M., KOOTER, I., LANKI, T., RANZI, A., SUGIRI, D., TSAI, M. Y., WANG, M., HOEK, G., BRUNEKREEF, B., PETERS, A. & FORASTIERE, F. 2015. Long-term Exposure to Particulate Matter Constituents and the Incidence of Coronary Events in 11 European Cohorts. *Epidemiology,* 26, 565-74.

WURTZ, P., HAVULINNA, A. S., SOININEN, P., TYNKKYNEN, T., PRIETO-MERINO, D., TILLIN, T., GHORBANI, A., ARTATI, A., WANG, Q., TIAINEN, M., KANGAS, A. J., KETTUNEN, J., KAIKKONEN, J., MIKKILA, V., JULA, A., KAHONEN, M., LEHTIMAKI, T., LAWLOR, D. A., GAUNT, T. R., HUGHES, A. D., SATTAR, N., ILLIG, T., ADAMSKI, J., WANG, T. J., PEROLA, M., RIPATTI, S., VASAN, R. S., RAITAKARI, O. T., GERSZTEN, R. E., CASAS, J. P., CHATURVEDI, N., ALA-KORPELA, M. & SALOMAA, V. 2015. Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. *Circulation,* 131, 774-85.

WÜTHRICH, B., SCHINDLER, C., LEUENBERGER, P. & ACKERMANN-LIEBRICH, U. 1995. Prevalence of atopy and pollinosis in the adult population of Switzerland (SAPALDIA study). *International archives of allergy and immunology,* 106, 149-156.

XU, C. J., SODERHALL, C., BUSTAMANTE, M., BAIZ, N., GRUZIEVA, O., GEHRING, U., MASON, D., CHATZI, L., BASTERRECHEA, M., LLOP, S., TORRENT, M., FORASTIERE, F., FANTINI, M. P., CARLSEN, K. C. L., HAAHTELA, T., MORIN, A., KERKHOF, M., MERID, S. K., VAN RIJKOM, B., JANKIPERSADSING, S. A., BONDER, M. J., BALLEREAU, S., VERMEULEN, C. J., AGUIRRE-GAMBOA, R., DE JONGSTE, J. C., SMIT, H. A., KUMAR, A., PERSHAGEN, G., GUERRA, S., GARCIA-AYMERICH, J., GRECO, D., REINIUS, L., MCEACHAN, R. R. C., AZAD, R., HOVLAND, V., MOWINCKEL, P., ALENIUS, H., FYHRQUIST, N., LEMONNIER, N., PELLET, J., AUFFRAY, C., VAN DER VLIES, P., VAN DIEMEN, C. C., LI, Y., WIJMENGA, C., NETEA, M. G., MOFFATT, M. F., COOKSON, W., ANTO, J. M., BOUSQUET, J., LAATIKAINEN, T., LAPRISE, C., CARLSEN, K. H.,

GORI, D., PORTA, D., INIGUEZ, C., BILBAO, J. R., KOGEVINAS, M., WRIGHT, J., BRUNEKREEF, B., KERE, J., NAWIJN, M. C., ANNESI-MAESANO, I., SUNYER, J., MELEN, E. & KOPPELMAN, G. H. 2018. DNA methylation in childhood asthma: an epigenome-wide meta-analysis. *Lancet Respir Med,* 6 , 379-388.

XU, X., SU, S., BARNES, V. A., DE MIGUEL, C., POLLOCK, J., OWNBY, D., SHI, H., ZHU, H., SNIEDER, H. & WANG, X. 2013. A genome-wide methylation study on obesity: differential variability and differential methylation. *Epigenetics,* 8, 522-33.

YAMADA, E., FUNOKI, S., ABE, Y., UMEMURA, S., YAMAGUCHI, D. & FUSE, Y. 2005. Size distribution and characteristics of chemical components in ambient particulate matter. *Anal Sci,* 21, 89-94.

YANG, J., YU, Y., SUN, S. & DUERKSEN-HUGHES, P. J. 2004. Ceramide and other sphingolipids in cellular responses. *Cell Biochem Biophys,* 40, 323-50.

YEN, C. H., DAI, Y. S., YANG, Y. H., WANG, L. C., LEE, J. H. & CHIANG, B. L. 2008. Linoleic acid metabolite levels and transepidermal water loss in children with atopic dermatitis. *Ann Allergy Asthma Immunol,* 100, 66-73.

YOUNG, F., CRITCHLEY, J. A., JOHNSTONE, L. K. & UNWIN, N. C. 2009. A review of co-morbidity between infectious and chronic disease in Sub Saharan Africa: TB and diabetes mellitus, HIV and metabolic syndrome, and the impact of globalization. *Global Health,* 5**,** 9.

ZHENG, J., ZHANG, X., ZHANG, L., ZHANG, H. P., WANG, L. & WANG, G. 2016. Interactive effects between obesity and atopy on inflammation: A pilot study for asthma phenotypic overlap. *Annals of Allergy, Asthma & Immunology,* 117, 716-717.

ZINS, M. & GOLDBERG, M. 2015. The French CONSTANCES population-based cohort: design, inclusion and follow-up. *Eur J Epidemiol,* 30 , 1317-28.