Thesis for doctoral degree (Ph.D.) 2011

AIR POLLUTION, GENETIC SUSCEPTIBILITY AND INFLAMMATION Sviatlana Panasevich

Thesis for doctoral degree (Ph.D.)
2011

AIR POLLUTION, GENETIC SUSCEPTIBILITY AND INFLAMMATION

Focusing on cardiovascular effects in adults and respiratory effects in children

Sviatlana Panasevich

Karolinska Institutet

From THE INSTITUTE OF ENVIRONMENTAL MEDICINE,
Karolinska Institutet, Stockholm, Sweden

AIR POLLUTION, GENETIC SUSCEPTIBILITY AND INFLAMMATION

Focusing on cardiovascular effects in adults and respiratory effects in children

Sviatlana Panasevich

Karolinska Institutet

Stockholm 2011

Published by Karolinska Institutet.

ISBN 978-91-7457-351-0

Printed by
REPROPRINT AB
Stockholm 2011
www.reproprint.se
Gårdsvägen 4, 169 70 Solna

To my parents,
Lidziya and Dzmitry

ABSTRACT

Air pollution exposure can induce low-grade systemic inflammation with consequences for both cardiovascular and respiratory systems. The overall aim of this thesis was to investigate the effects of air pollutants on the development of complex inflammatory diseases (myocardial infarction and respiratory disease), and genetically determined susceptibility for these effects, using epidemiologic methodology.

The two study populations were drawn from a case-control study of myocardial infarction (SHEEP) and a birth cohort (BAMSE). From SHEEP, the present study population included 1192 first-time myocardial infarction (MI) cases aged 45-70 years identified in Stockholm County during 1992-1994, and 1536 matched population controls from the study base. Participants completed questionnaires and underwent medical examination as well as blood sampling. Their air pollution exposure was assessed retrospectively both long-term (1-30 years) and short-term (12 h - 5 days). NO_2 was used as an indicator of emissions from road traffic and SO_2 as an indicator of emissions from residential heating. From the BAMSE cohort, which recruited 4089 new-born children during 1994-1996 in four municipalities in Stockholm County, this study included 497 wheezers and 485 non-wheezing controls at 4 years, and 198 asthma cases and 192 non-asthma controls at 8 years. Questionnaires were completed by the parents when the children were 2 months, 1, 2, 4, and 8 years of age. The children were also invited for medical examination and blood sampling at 4 and 8 years.

In adults, long-term exposure to both traffic- NO_2 and heating-SO_2 emissions showed an association with IL-6 levels. For instance, 30-year traffic-NO_2 exposure was associated with a 64.5% (95%CI 6.7-153.8%) increase in serum IL-6 per 28.8 $\mu g/m^3$ (corresponding to the difference between the 5th and the 95th percentile exposure value). There was also suggested association between short-term exposure to traffic-related air pollutants and inflammatory markers (IL-6, TNF-α). Gene-environment interaction was observed for several *IL6* and *TNF* single nucleotide polymorphisms (SNPs) in relation to inflammation blood marker levels. For example, 1-year traffic-NO_2 exposure interacted with *IL6* -174G/C in an additive way, where each additional *IL6*-174C allele was associated with an increased air pollution effect on IL-6 levels, and 1-year heating-SO_2 exposure was associated with higher TNF-α levels in *TNF*-308AA homozygotes but not in -308G carriers. Also short-term air pollution exposure interacted with *IL6* and *TNF* SNPs in relation to marker levels. The risk of MI followed the pattern of effect on blood markers across genotype groups.

In children, interaction with early maternal smoking was seen for 3 *TNF* SNPs with respect to early wheeze. The odds ratio for developing early wheeze related to maternal smoking was 2.4 (95%CI 1.6-3.7) in *TNF* -857CC homozygote children, while no tobacco-related risk was seen in children with the rare -857T allele. Suggestive interaction with early maternal smoking was also seen for 3 *GSTP1* SNPs with respect to transient wheeze. SNPs in *TNS1*, *ADAM19*, *THSD4* and *ADCY2* identified through genome-wide analyses on lung function in adults showed association also with lung function in children; *DAAM2* rs2395730 showed suggestive interaction with current tobacco smoke exposure at 8 years.

In summary, the results indicate that air pollutants affect levels of inflammatory blood markers, and this effect appears to be modified by genetic variants, affecting both

blood marker levels and consequent MI risk. Polymorphisms in genes related to inflammation (*TNF*) and antioxidant defense (*GSTP1*) seem to modify the effect of early tobacco smoke exposure on childhood wheezing. Several gene variants of importance for lung function in adults also seem to affect lung function in children.

LIST OF PUBLICATIONS

Papers will be referred to in the text by their roman numbers.

I. **Panasevich S**, Leander K, Rosenlund M, Ljungman P, Bellander T, de Faire U, Pershagen G, Nyberg F.
Associations of long- and short-term air pollution exposure with markers of inflammation and coagulation in a population sample.
Occup Environ Med 2009; 66: 747-753

II. **Panasevich S**, Leander K, Ljungman P, Bellander T, de Faire U, Pershagen G, Nyberg F.
Interaction between air pollution exposure and genes in relation to levels of inflammatory markers and risk of myocardial infarction.
Manuscript

III. **Panasevich S**, Lindgren C, Kere J, Wickman M, Pershagen G, Nyberg F, Melén E.
Interaction between early maternal smoking and variants in TNF and GSTP1 in childhood wheezing.
Clin Exp Allergy 2010; 40: 458-467

IV. **Panasevich S**, Melén E, Hallberg J, Bergström A, Pershagen G, Nyberg F.
Investigation of novel genes for lung function in children, and their potential interaction with early tobacco smoke exposure.
Manuscript

CONTENTS

LIST OF ABBREVIATIONS

ADAM19	ADAM metallopeptidase domain 19 gene
ADCY2	Adenylate cyclase 2 gene
ADRB2	Beta-2-adrenergic receptor gene
BAMSE	Children (Barn), Allergy, Milieu, Stockholm, Epidemiological survey
BMI	Body mass index
DAAM2	Disheveled associated activator of morphogenesis 2 gene
DNA	Deoxyribonucleic acid
ETS	Environmental tobacco smoke
FAM13A	Family with sequence familiarity 13, member A gene
FEV_1	Forced expiratory volume in one second
FVC	Forced vital capacity
GSTP1	Glutathione S-transferase P1
GSTP1	Glutathione S-transferase P1 gene
GWAS	Genome-wide association study
GWIS	Genome-wide interaction study
HHIP	Hedgehog interacting protein gene
IL6	Interleukin-6 gene
IL-6	Interleukin-6
LD	Linkage disequilibrium
LRT	Likelihood ratio test
MI	Myocardial infarction
NO_2	Nitrogen dioxide
O_3	Ozone
PM_{10}	Particulate matter with an aerodynamic diameter of up to 10μm
SHEEP	Stockholm Heart Epidemiology Program
SNP	Single nucleotide polymorphism
SO_2	Sulphur dioxide
THSD4	Thrombospondin, type1, domain-containing protein 4 gene
TNF	Tumor necrosis factor alpha gene
TNF-α	Tumor necrosis factor alpha
TNS1	Tensin1 gene

1 BACKGROUND

1.1 AIR POLLUTION

1.1.1 Outdoor air pollution

Outdoor air pollutants include gaseous compounds, e.g. nitrogen dioxide (NO_2), sulfur dioxide (SO_2), ozone (O_3), and particulate matter (PM) of various sizes and composition. These pollutants can all be generated naturally, but today anthropogenic sources of air pollution are a major concern. Sources and WHO guidelines for the pollutants investigated in this study, of importance for urban populations, are presented in Table 1.

Table 1 Sources of ambient air pollutants and WHO air quality guidelines.

Air pollutant	Source[1 2]	World Health Organization guidelines[3]
NO_2	Motor vehicles and other fossil fuel combustion (heating, power generation)	1-hour average: 200 μg/m^3 Annual average: 40 μg/m^3
SO_2	Combustion of fossil fuel (power plants, diesel engines), oil refineries	10-minute average: 500μg/m^3 24-hour average: 20 μg/m^3
O_3	Formed in the troposphere as a result of chemical reactions between nitrogen dioxide and hydrocarbons in the presence of sunlight	8-hour average: 100 μg/m^3
Particulates	Road dust, tire fragmentation, motor vehicle emissions, burning of fossil fuel, construction works, wood burning, windblown soil, pollens and molds.	PM_{10}: Annual average: 20 μg/m^3 24-hour average: 50 μg/m^3

When the levels are elevated, air pollutants may cause respiratory symptoms, impair lung function, and exacerbate existing respiratory and cardiovascular diseases.[1] Therefore, air pollutants including NO_2, SO_2, O_3, and PM_{10} are commonly monitored on a continuous basis in urban areas since a few decades back, and these measurements are often available also for scientific studies. Monitoring data are used to develop environmental policy measures for controlling and reducing air pollution, as well as for warning susceptible populations to avoid being outdoors during high pollution periods.

Air pollutants are often generated as a mixture and it is difficult to disentangle which of them have the most pronounced detrimental effect on health. Certain pollutants are most characteristic for a particular source, and may then also be used as indicators for that source-related pollution. For example, NO_2 is commonly used as a marker of motor

exhaust emissions, while PM with an aerodynamic diameter of up to 10μm (PM_{10}) represents more coarse particles, like road dust. It is believed that both gaseous and particulate air pollutants can induce irritation in the airways and cause oxidative damage.[4] However, particulates, due to their diversity of sources and composition, include a variety of components, such as nitrates, sulfates, elemental and organic carbon, organic compounds (polycyclic aromatic hydrocarbons), biological compounds (endotoxin), and a variety of metals (iron, copper, nickel, zinc, and vanadium). [2] Certain compounds may exert adverse effects of air pollutants in the lungs and trigger an immune response. The finer fraction of PM can penetrate into the lower respiratory tract, and ultrafine particles are also believed to translocate into the circulatory system.[5]

While measurement-based data for short-term exposure to air pollution is widely available, long-term exposures often require retrospective modeling. In this case, source-specific air pollution indicators can be assigned to the appropriate emission sources, in an emission-based approach. In the present study two indicator pollutants were used: NO_2 from local traffic emissions (traffic-NO_2) as a marker of the mixture of vehicle emissions, also including exhaust particles, and SO_2 from local residential combustion heating (heating-SO_2) as a marker mainly reflecting oil combustion. In this emission-based approach, both indicators characterize only local sources of pollution and do not cover long-range air pollution contributions.

1.1.2 Tobacco smoke

Another important air pollutant, more characteristic of the indoor environment, is tobacco smoke (TS), which affects both active smokers and persons exposed passively. TS is a complex mixture of more than 4800 different compounds.[6] Many of them are known carcinogens and mutagens, or possess cytotoxic and irritant properties. TS constituents include polycyclic aromatic hydrocarbons and N-nitrosamines, free radicals, aromatic amines, aldehydes, and metals such as nickel, chromium, and cadmium.[6] The chemical composition of side-stream TS is qualitatively similar to that of main-stream smoke but differs quantitatively with respect to the levels of individual constituents, and the average particle size of side-stream TS is smaller.[6]

The prevalence of smoking in women of child-bearing age ranges between 17 and 35% around the world. [7] Although the smoking prevalence during pregnancy has declined over the past years, a significant proportion (13-16%) of women continues to smoke during pregnancy in the Nordic countries, and this is influenced by maternal age, ethnicity, education and socioeconomic level.[7-9] Data for the Swedish BAMSE study participants showed similar prevalence rates: 12 % of mothers smoked during pregnancy, and 25% of the children were exposed to environmental tobacco smoke during the first two years of life.[10] These data refer to the 1990s (1994-1996), while today the situation is gradually improving. The overall adult daily smoking prevalence in Sweden was 15% in 2006.[11]

1.2 HEALTH EFFECTS OF AIR POLLUTANTS

1.2.1 Induction of inflammation

Inflammation is a physiological reaction of tissue response to injury. Tissue injury can be caused by microorganisms and chemical or physical agents, including oxidative

substances. As a result of tissue damage, both damaged cells (e.g. epithelial cells) and activated tissue macrophages produce proinflammatory mediators (including interleukin 1 and 6 (IL-1, IL-6) and tumor necrosis factor alpha (TNF-α)). Vascular permeability increases and leads to leakage of fluid to the inflammation site. Proinflammatory mediators released into the blood stream have a systemic effect and stimulate recruitment of neutrophils to the site of inflammation, as well as production of acute phase proteins (C-reactive protein [CRP] and fibrinogen). Activated neutrophils gather at the site of inflammation to destroy and eliminate the causative agent. These processes induce the symptoms of inflammation, characterized by four Latin words: *calor* (heat), *dolor* (pain), *rubor* (redness), and *tumor* (swelling). Normally, when the causative agent is eliminated, anti-inflammatory agents act to diminish inflammation, limit the damage of surrounding tissues and promote tissue healing. Acute inflammation is a rapid, short-lived (minutes to days) response to acute injury.[12]

When the inflammation persists, due to incomplete clearance of a causative agent or as a result of multiple acute or subacute events occurring in the same location, it becomes a chronic inflammation.[12] An important role in the transition from acute to chronic inflammation is played by IL-6.[13] Chronic inflammation leads to accumulation of macrophages and lymphocytes at the site, as well as to the growth of fibroblasts and vascular tissue, resulting in tissue scarring.[12]

Oxidative damage in the lungs caused by exposure to air pollutants is capable of inducing low-grade systemic inflammation, affecting the lung tissue and, due to the systemic release of inflammatory mediators, with wider effect also on the cardiovascular system.[5] There are several mechanisms that are potentially involved in cardiovascular effects of lung exposures. On the one hand, systemic cytokine-mediated inflammatory response induced in the lungs[2 4] can contribute to atherosclerosis. [14-16] On the other hand, ultrafine particulate matter can directly enter the circulation and affect vascular endothelium, also potentially contributing to atherosclerotic changes.[5 16] Atherosclerosis is considered to be an inflammatory disease, where inflammation occurs in the developing atherosclerotic plaque,[14 17] and consequently, the additional load of proinflammatory cytokines and ultrafine potentially oxidizing particles in the bloodstream may augment plaque development. Yet another mechanism involves effects of air pollutants on cardiac autonomic control and increased risk of arrhythmia, either as a result of airway receptor stimulation or indeed as a consequence of the inflammatory response. [4] Changes in blood markers, including inflammation and coagulation parameters, after exposure to air pollutants may therefore help to elucidate the pathways of disease development caused by air pollution exposure. [2 4] Such effects on blood markers have been studied in both epidemiological and experimental studies for short-term exposure to air pollution,[18-22] but there is a lack of studies investigating long-term effects.

When air pollution, such as tobacco smoke, enters the lung in children, local inflammatory response may impair lung function through epithelial injury and fibrosis as well as small airway remodeling. [6] Systemic inflammation may also ensue, as described above. Particular attention is warranted for intrauterine exposure to tobacco smoke, where harmful effects are mediated by systemic factors transmitted directly from mother to foetus. [23] It has been shown that exposure to tobacco smoke *in utero* leads to oxidative stress, as indicated in blood analysis in newborns and children at 3 months of age. [24 25] Tobacco-exposed children have been shown to have lower birth

weight, impaired lung function, and changed immune function parameters as well as a higher risk of developing wheeze and asthma. [26-28] The mechanisms leading to these effects are not fully clarified. Hypothetically, the response capacity of both antioxidant and inflammatory systems in the developing child, defined by gene-environment interactions, determine the variability in an individual's phenotype.

1.2.2 Effects on the cardiovascular system by ambient air pollution

Both long- and short-term exposure to air pollution have been associated with cardiovascular morbidity and mortality. [4 5 29-31] A previous study based on the SHEEP study population reported an association between long-term exposure to air pollution and fatal myocardial infarction, especially out-of hospital deaths. [32] Epidemiologic studies have also shown an association between ambient air pollution and atherosclerosis, assessing carotid intima-media thickness. [33 34] The combination of high intima-media thickness and high levels of CRP has been shown to substantially increase the risk of heart failure.[35] Nonetheless, there is limited evidence on possible mechanisms behind these effects.

Both the inflammation and coagulation systems are involved in mechanisms underlying cardiovascular effects. Thus, the risk of developing myocardial infarction has been shown to be associated with increased levels of blood markers of inflammation (IL-6, TNF-α, CRP) and coagulation (fibrinogen, PAI-1).[36-41] These markers have also been linked to air pollution exposure, e.g. in a study on myocardial infarction survivors in six European cities, where short-term exposure to particulate matter was found to be associated with increased levels of IL-6 and fibrinogen. [21 42] However, epidemiologic studies have often reported inconsistent results regarding the associations of air pollution exposure with IL-6, CRP, or fibrinogen.[21 22 43-47] Data are also still lacking to characterize the full link from air pollution to blood markers to cardiovascular risk in the same study population.

1.2.3 Effects on the respiratory system by tobacco smoke

Children raised in smoker homes have a higher incidence of respiratory infections, recurrent wheezing, bronchitis, nocturnal cough, and asthma. [6] Maternal smoking and passive exposure of children to cigarette smoke are also associated with impaired lung function, wheezing and illness. [6 28 48] While investigating exposure occurring very early in life, it is mostly difficult to disentangle prenatal and early postnatal exposures, which tend to be quite highly correlated. [48] Some studies suggest that in utero exposure to heavy smoking has a stronger effect on the development of asthma and decreased lung function than postnatal exposure to environmental TS. [49] Development of lung function in children is of particular importance, since evidence suggests that impaired lung function in childhood has an impact also in adulthood. [50] Lung function is important for the functional capacity of the human body, [51] and is also a reasonably good predictor of mortality in general population. [52] Both forced expiratory volume in one second (FEV_1) and forced vital capacity (FVC) are influenced by lung size, but FEV_1 is the primary parameter influenced by airway obstruction, and the ratio FEV_1/FVC is thus used for defining obstructive defects. [51 53]

1.3 GENES IN COMPLEX DISEASES

Complex diseases are characterized by multiple genetic associations with small-to-moderate effects. Following the steady progress in genotyping methods, genetic study design and genetic understanding, genetic effects have been investigated in linkage studies (identifying chromosome areas of interest), candidate gene studies (genes hypothetically involved in pathogenesis), and most recently broader scope hypothesis-free investigation of genetic effects in genome-wide association studies (GWAS) yielding new, often previously unknown, genes of interest. It is nowadays feasible to investigate up to a million or more SNPs using current GWAS chips. By design, these SNPs cover the majority of common SNP genetic variation by the use of so-called tagging SNPs across the whole genome, and provide an indication of the gene or gene region containing the putative functionally relevant SNP or SNPs with high correlation to the significant tag SNP. Future prospects include targeting also rare variants in GWAS studies using even denser GWAS chips or whole-genome sequencing, as well as investigation of epigenetic modifications. On a broader functional perspective, the candidate gene approach may be extended to systems genetics, where whole networks of genes are studied.

The present work (except paper IV) was initiated before the advent of extensive GWAS research. Nevertheless, investigation of the interaction between heredity (genes) and environment may potentially utilize all of the approaches mentioned above, including GWAS. There are, however, numerous issues to consider in employing these methods. For example, in spite of the attractive wide coverage of the genome, genome-wide studies of gene-environment interaction (so-called GWIS) have serious limitations. Firstly, they require extremely large sample size to reach sufficient power, and even at large size would still generally fail for rare genetic variants.[54] Secondly, they can only identify interactions with consistent effects across different genetic and environmental exposure background, which is often difficult to achieve.[54] Candidate gene interaction studies, which are based on previous knowledge of the biological mechanisms involved and where the environment is likely to play a pivotal role, thus remain an attractive design for identifying and characterizing gene-environment interactions. [54]

1.3.1 Genetic associations with cardiovascular disease

The present work has focused on genes which are related to cardiovascular pathology (inflammation, coagulation), and at the same time are believed to be involved in the response of the organism to air pollution. Variants in all selected candidate genes (*IL6*, *TNF*, fibrinogen Bβ gene (*FGB*), *PAI-1*) have shown overall association with MI risk in the SHEEP study and in other studies. [37 39 55 56] Recent GWAS studies have identified a number of variants in genes to be associated with coronary artery disease and other cardiovascular and metabolic phenotypes. [57] Among the newly identified genes, there are genes functionally involved in lipid metabolism, collagen processing, apoptosis, adhesion signaling, coronary calcification, lipoprotein(a), inflammation, as well as a number of genes with unknown function. [58]

1.3.2 Genetic associations with asthma and lung function

In the present investigation, the focus of interest was on genes previously shown to be associated with respiratory disease, [59] and also potentially modify the response to tobacco smoke exposure: on the level of antioxidant defense (*GSTP1*), inflammation (*TNF*), and local lung receptor dynamics (*ADRB2*).

Based on recent genome-wide findings, a number of specific groups of genes have received strong support for being involved in determining susceptibility to allergic disease. These include genes involved in immune responses, barrier function, tissue response, eosinophil activations, and the interplay between the environmental factors and cellular responses. [60] To date, genome-wide association studies have identified variants in the following genes to be associated with asthma and blood eosinophil counts: *IL1RL1*, *WDR36*, *MYB*, *IL33*, *CTNNA3*, *ORMDL3*, *PDE4D*, *TLE4*. [60] Larger future genome-wide studies are very likely to find further genes to be involved. Recent publications have also identified a number of SNPs from a range of genetic loci to be associated with lung function in adults. [53 61] However, there are as yet no studies investigating genome-wide associations with lung function in children.

1.4 INTERACTIONS OF AIR POLLUTION WITH GENES

1.4.1 Cardiovascular effects in adults

1.4.1.1 Inflammation

Given the effects observed on cardiovascular disease by air pollution on the one hand and genes on the other, it is of interest to understand whether individual genetic makeup can modify an individual's response to air pollution, both in terms of intermediate biomarkers related to inflammation and coagulation, and also on the development or acute onset of myocardial infarction (MI). One example of a study looking at such interaction on an inflammatory marker is the study of MI survivors conducted in six European cities, where *IL6* (rs2069832, intron) and *Fibrinogen Bβ* (rs1800790, promoter) variants were seen to modify the effect of air pollution exposure on the levels of IL-6. [62]

1.4.1.2 Myocardial infarction

Several studies have investigated the interaction of genes and air pollution in relation to the health outcomes associated with cardiovascular disease (blood pressure, QT intervals, homocysteine levels), focusing on genes related to oxidative stress, inflammation and lipid metabolism, and reported significant associations for *GSTM1*, *GSTT1*, *HFE*, *HMOX-1*, *cSHMT*, *TF C2*, *PHF11*, *MMP1*, *ITPR2*, *NQO1*, *GSS*, and a number of other genes.[63] Two studies have investigated interactions between tobacco smoke exposure and the *IL6*-174G/C variant. The variant *IL6*-174C allele was associated with increased all-cause mortality risk among 80-year-old subjects, but only in non-smokers. [64] Another study reported -174C allele association in middle-aged men with increased MI risk only in smokers.[65]

1.4.2 Respiratory effects in children

1.4.2.1 Wheezing

A range of single nucleotide polymorphisms (SNPs) in xenobiotic-metabolizing enzyme genes (e.g. *CYP1A1*, *CYP1B1*, *GSTM1*) have been found to significantly interact with smoking exposure in relation to asthma susceptibility. [66 67] Most recently, SNPs in other genes with yet unknown function such as *GSDML*/*ORMDL3* in the 17q21 region have also been suggested to modify the effect of passive smoking. [68] Several studies have investigated interactions of genes with tobacco smoke exposure, focusing mainly on genes involved in antioxidant protection and inflammation, such as the glutathione S transferase M1, T1 or P1 (*GSTM1*, *GSTT1*, *GSTP1*) and tumor necrosis factor alpha (*TNF*) genes. [67 69-72] The *GSTP1* coding SNP Ile105Val had a significant interaction with *in utero* tobacco smoke exposure and exposure to air pollution in relation to wheeze and asthma. [70 73] A number of *TNF* SNPs, e.g. *TNF*-308 and *TNF*-238, have been investigated in relation to modifying the effect of tobacco smoke exposure on asthma phenotypes in school-age children. [71 72] *ADRB2 Arg*16*Gly* was reported to modify the effect of *in utero* and childhood tobacco smoke exposures on childhood wheezing. [74] Recently, results on gene-environment interactions with early exposure to air pollution have been reported from BAMSE, which support an effect modification by *GSTP1 Ile*105*Val* and *TNF*-308 variants on the development of childhood allergy. [75 76] Previous studies have thus pointed to the importance of the inflammation and antioxidative pathways for modulating the response to environmental stimuli, including tobacco smoke, but the evidence is not conclusive in the development of respiratory disease in children.

1.4.2.2 Lung function

Few studies investigated interactions of genes with tobacco smoke exposure in relation to lung function in children. *In utero* exposure to tobacco smoke in *GSTT1* null children was associated with decreased lung function at 9-11 years compared with *GSTT1* positive children not exposed to tobacco smoke. [69] Also *ADAM33* gene has been shown to interact with *in utero* TS exposure affecting lung function in children. [77] Investigating potential interaction of novel lung function associated SNPs with tobacco smoke exposure, including different time windows of exposure, may elucidate new pathways of lung function deterioration in children.

2 AIMS

The overall objective of the thesis was to investigate the effects of air pollutants (ambient air pollution, tobacco smoke) on the development of complex inflammatory diseases (myocardial infarction and respiratory disease), and to identify possible subgroups of individuals with a genetic susceptibility based on investigation of important candidate genes and their importance for gene-environment interaction with these exposures.

The specific aims, each addressed in one of the constituent papers, were:

I. To investigate if long- and short-term air pollution exposures affect the levels of certain blood markers of interest for cardiovascular disease in adults.

II. To explore potential gene-environment interactions between genetic variants related to inflammation and coagulation and long- and short-term exposure to air pollution with respect to both blood marker levels and the risk of myocardial infarction

III. To investigate whether variants in certain asthma candidate genes modify the effect of early maternal smoking (during pregnancy and/or postnatally) on the development of asthma, wheeze and allergic sensitization in children.

IV. To attempt a replication of recent adult GWAS associations for lung function also in children, as well as to investigate their potential interaction with early exposure to tobacco smoke in relation to lung function.

3 METHODS

3.1 STUDY POPULATIONS

The present studies were conducted in two different populations, one of adults and one of children. Cardiovascular outcomes were investigated in adults in the Stockholm Heart Epidemiology Program (SHEEP), and respiratory effects were explored in children within the BAMSE birth cohort (Children (Barn), Allergy, Milieu, Stockholm, Epidemiological survey).

3.1.1 The SHEEP case-control study

The Stockholm Heart Epidemiology Program (SHEEP) is a large population-based case-control study. SHEEP included all first-time myocardial infarction cases (n=2246) aged 45-70 years identified in Stockholm County during 1992-1994. Controls (n=3206) were selected randomly from the study base (Stockholm County population, 1992-1994), matched on sex, age (±5 years) and hospital catchment area. [78] The participation rate was 83% for cases and 73% for controls; blood sampling was performed in 77% of surviving MI patients and in 67% controls. [39] The study was approved by the Ethical Committee at Karolinska Institutet.

For the analysis of air pollution effects on blood marker levels (Paper I), all population control subjects (n=1536) with available data on the blood markers were included. The investigation of genetic modification of air pollution effects on blood marker levels and MI risk (Paper II) was performed in 1506 controls with available data on genotypes and blood marker levels, as well as in 1192 first time non-fatal myocardial infarction cases with available data on genotypes and time of MI symptoms onset. Gene-environment interactions for short-and long-term air pollution on blood marker levels were studied in the controls, for long-term exposure on MI risk using case-control design, and for short-term exposure on MI onset using case-crossover design.

3.1.2 The BAMSE birth cohort

The prospective birth cohort BAMSE was recruited during the years 1994-1996, and included 4089 newborn infants (75% of all eligible children in the selected areas of Stockholm county, representing both urban and suburban districts).[79] At inclusion, when the children were approximately 2 months of age, parents completed detailed questionnaires with information on living conditions, environmental exposures at home, parental allergy and smoking habits of mothers and fathers. Subsequent questionnaires were completed when the children were 1, 2, 4 and 8 years of age with response rates of 96%, 94%, 92% and 84%, respectively. The children were invited for clinical testing, including blood sampling and lung function measurements, at 4 and 8 years. Blood sampling was performed in 2614 children at 4 years and in 2480 children at 8 years. The study was approved by the Ethical Committee at Karolinska Institutet.

To investigate effect modification by the *TNF*, *GSTP1* and *ADRB2* genes on early maternal smoking in relation to wheezing at 4 years (Paper III), genetic analyses were performed in selected groups of wheezing cases (n=542) and their random controls (n=542), sampled from the full cohort with a case-cohort type of procedure. From the

2614 blood samples, 2298 were suitable for further genetic analysis (after exclusion of samples with too little blood, lack of questionnaire data and with no parental consent for genetic analysis), forming the so called “genetics cohort” (Fig. 1). From the “genetics cohort”, 709 children were randomly selected, 542 of them being non-wheezers, which yielded a random sample of controls. The 167 wheezers from the random sample were complemented by all other identified wheezers (n=375) in the cohort, yielding a total group of 542 wheezing cases (Fig. 1). After excluding subjects with failure in DNA extraction (n=29) or >50% missing genetic data from genotype failures, indicating poor DNA quality (n=73), the final genetic analysis group included 982 children (497 wheezers and 485 controls).

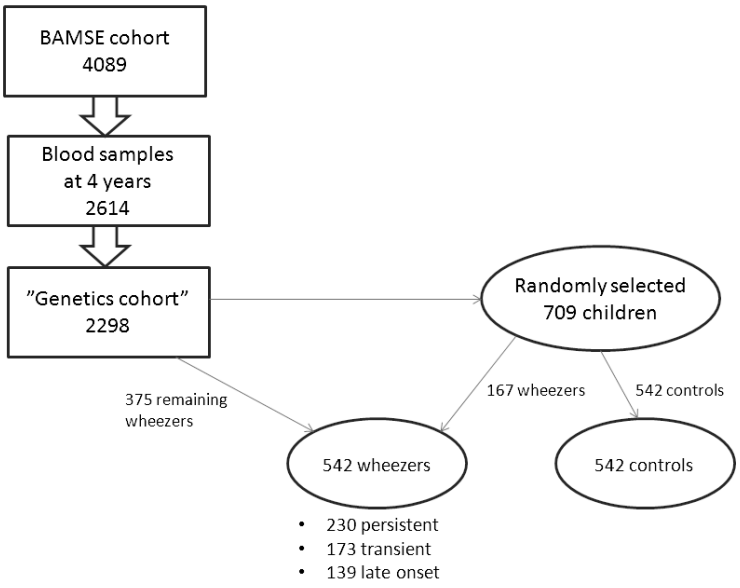

Figure 1. Selection of BAMSE children at 4 years based on wheezing phenotype (Paper III)

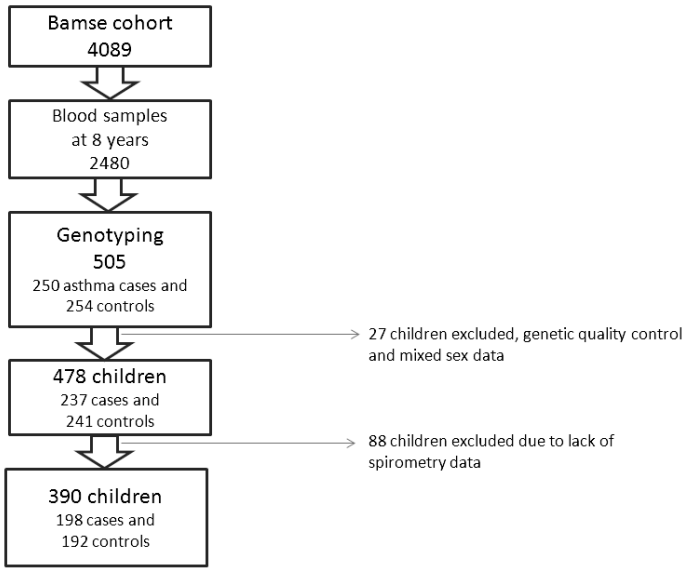

Figure 2. Selection of BAMSE children at 8 years based on asthma phenotype (Paper IV)

The genome-wide study of genetic effects on lung function in children at 8 years (Paper IV) was based on a sample of asthma cases and their controls selected for genome-wide genotyping. From the BAMSE cohort (n=4089), blood samples were obtained from 2480 children during the medical examination at 8 years. DNA was extracted from samples from 2033 children (after exclusion of samples with too little blood, lack of questionnaire data, or where parental consent to genetic analysis was not obtained). Of these, all children with a doctor's diagnosis of asthma (at 1, 2, 4 or 8 years) were selected as cases (n=250) and an equivalent number of children with no history of asthma or other allergic diseases as controls (n=254). Not all children had lung function data; the final group of children with GWAS data that passed quality control and with available lung function measurements consisted of 390 children (198 cases and 192 controls) (Fig. 2).

3.2 EXPOSURE ASSESSMENT

3.2.1 Ambient air pollution exposure

In the SHEEP study, individual exposure to air pollution was estimated in relation to various time windows, both long-term (years) and short-term (days).

Long-term exposure to source-specific air pollution was assessed at the participant's historical home addresses using a geographical information system. The detailed methods of assessment have been described previously.[32 80] Briefly, source-specific emission databases were established for each decade since 1960, providing information for dispersion modeling of annual mean levels of locally emitted air pollutants. [80] All addresses inhabited by study participants for more than 2 years were transformed into

geographical coordinates, and calibrated dispersion models for air pollution were used to estimate the annual mean level at each address.[32] Using these, average exposures for the preceding 1, 5, and 30 years before the year of study inclusion were calculated. At most 5 of 30 years and 1 of 5 years of address information was allowed to be missing for calculating 30-year and 5-year means, respectively. The missing air pollution data that resulted for years of unknown residency were replaced by the mean among the controls for that calendar year.[32] For the 1-year mean, no missing address information was allowed. Two air pollutants were studied, characterizing different exposure sources: NO_2 from local traffic emissions (traffic-NO_2) as a marker of the mixture of vehicle emissions, also including exhaust particles, and SO_2 from local residential combustion heating (heating-SO_2) as a marker mainly reflecting oil combustion.

Short-term air pollution exposure measures were based on routine hourly rooftop measurements of total ambient NO_2, PM_{10}, O_3 and SO_2 in Stockholm. PM_{10} measurements started in spring 1994, and PM_{10} exposure can thus only be calculated for subjects included in the study after that point in time. Average exposures were calculated for intervals 0-12h, 12-24h, 48h (2 days) and 120h (5 days) before the index time point. The index time point was the hour of blood sampling for assessment of effects on blood markers, and the hour of MI symptoms onset or corresponding control time point for assessment of the risk of MI onset in the case-crossover analysis. Control intervals were estimated for each case using a time-stratified approach: within the calendar month of MI occurrence, all time points on weekdays corresponding to the time and weekday of MI onset were used to define control periods for a case period.[81] Exposure intervals with 75% or more of hourly measurements available were included and the interval mean was calculated from the available values.

3.2.2 Tobacco smoke exposure

The studies of children (BAMSE) were focused on exploring the effect of tobacco smoke exposure on respiratory health. The tobacco smoke exposure information was based on the questionnaires answers by the parents.

Early maternal smoking was considered present if the mother smoked at least one cigarette per day at the time of the first questionnaire, when the child was approximately 2 months old (postnatal exposure), and/or smoked at least one cigarette per day at any time during pregnancy (*in utero* exposure). Furthermore, three categories were defined for the exposure to tobacco smoke during pregnancy according to the highest reported number of cigarettes smoked by the mother per day during any of the pregnancy trimesters: 0 (reference group), 1-9 and >9 cigarettes per day. *Current tobacco smoke exposure at 8 years of age* was present if any of the parents smoked daily at the time of completing the 8-year questionnaire.

3.2.3 Genotyping

Genetic factors were explored in papers II- IV.

In the candidate gene analysis in SHEEP (Paper II), selected genes and SNPs were candidates for involvement in individual susceptibility to cardiovascular disease. They included *IL6* with SNPs -598G/A (rs180797), -573G/C (rs1800796), -174G/C

(rs1800795); *TNF* with SNPs -1031T/C (rs1799964), -863C/A (rs1800630), -857C/T (rs1799724), -308G/A (rs1800629), -238G/A (rs3615525); fibrinogen Bβ (*FGB*) with SNP -455G/A; and *PAI-1* with SNP -675 4G/5G. Genotyping for *IL6* and *TNF* were performed using dynamic allele specific hybridization, [36 37] and *FGB* and *PAI-1* polymorphisms were genotyped using the polymerase chain reaction method. [38 39]

In the candidate gene analysis in BAMSE (Paper III), analyses of SNPs in selected genes were performed with matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry (Sequenom, San Diego, CA, USA). [76] The selected genetic variants for analysis included five *TNF* (rs1799964, rs1799724, rs1800629, rs1800610, rs3093664), six *GSTP1* (rs762803, rs1695, rs749174, rs1138272, rs1871042, rs4891) and three *ADRB2* (rs1042714, rs1042717, rs1042718) SNPs.

In the study of lung function genes, relying on extraction of data from the genome-wide genetic data on a subset of BAMSE (Paper IV), genotyping was conducted in collaboration with the GABRIEL consortium. [82] DNA samples were genotyped using the Illumina Human 610 quad array in the genotyping laboratory at the Centre de Genotypage, Evry, France.

3.3 ASSESSMENT OF OUTCOMES

3.3.1 Blood markers of inflammation and coagulation

SHEEP (Papers I, II): Serum IL-6 was measured using an enzyme-linked immunosorbent assay (IL-6 Eli-pair, Diaclone Research, Besancon, France).[36] TNF-α was detected in serum with the Quantikine HS human TNF-alfa kit (R&D Systems, Minneapolis, MN).[37] C-reactive protein (CRP) was measured on EDTA plasma samples using a high sensitivity immunonephelometric assay (Dade-Behring, Marburg, Germany) on the automated BN II system from Dade Behring.[36] An assay fibrinogen fibrin polymerization time (FPT test) according to the method described by Vermylen *et al*.[83] was used to measure plasma fibrinogen levels,[38] and the Spectrolyze PAI-1 kit (Biopool AB, Umea, Sweden) was used for determining PAI-1 activity in citrated plasma samples.[39]

3.3.2 Myocardial infarction

SHEEP (Paper II): First-time MI cases were identified using standard diagnostic criteria: certain symptoms according to case history information, specified changes in blood levels of the enzymes CK and LD, specified ECG-changes and autopsy findings.[78] Only cases who survived 28 days after their MI were included. The time of MI symptoms onset was determined primarily using information from the medical clinic about the symptom onset or admission, or alternatively the admission date was obtained from the National Hospital Discharge Registry.

3.3.3 Wheezing, asthma and allergic sensitization

BAMSE (Paper III): Wheeze phenotypes were determined in children followed up to 4 years based on questionnaire information. *Early-onset wheeze* was defined as wheeze up to 2 years of age, and was further subdivided into *transient wheeze* (≥3 episodes of wheeze between 3 months and 2 years of age, but no episode in the last 12 months at 4

years) or *persistent wheeze* (≥1 episode of wheeze between 3 months and 2 years of age and ≥1 episode in the last 12 months at 4 years). *Late-onset wheeze* was defined as no episode of wheeze between 3 months and 2 years of age, but ≥1 episode in the last 12 months at 4 years.[84] *Current asthma* was defined as reported physician-diagnosed asthma up to 4 years of age and one or more episodes of wheezing in the last 12 months at 4 years. The reference group for analyses of asthma consisted of non-wheezing children without asthma diagnosis at 4 years.

Allergic sensitization was assessed with serum IgE tests for 2614 of the children at 4 years of age. It was analysed as specific IgE antibodies to inhalant and food allergens (Phadia AB, Uppsala, Sweden). The mix of inhalant allergens (Phadiatop®) included allergens of cat, dog, horse, birch, timothy, mugwort, *D. pteronyssinus* and *Cladosporium*, and the mix of food allergens (Fx5®) included allergens of milk, egg white, peanut, wheat, soy bean and fish. Sensitization to each category of allergens was defined as serum IgE antibody level ≥0.35 kilounits of antibody per litre (kU_A/L).

3.3.4 Lung function

BAMSE (Paper IV): Lung function measurements were performed in children at 8 years. Maximum expiratory flow volume tests were performed using a spirometer (2200 Pulmonary Function Laboratory; Sensormedics, Anaheim, CA, USA).[85] The highest values of forced expiratory volume in 1 sec (FEV_1) and forced vital capacity (FVC) were extracted and used in the analysis, provided that the child's effort was coded as maximal by the test leader, the test curves passed visual quality inspection, and that the two highest readings were reproducible according to the American Thoracic Society and European Respiratory Society (ATS/ERS) criteria.[86]

3.4 STATISTICAL ANALYSES

In the SHEEP study (Papers I, II), air pollution effect estimates are given per unit change in the air pollutant, where one unit change corresponds to the difference between the 5th and 95th percentile of the exposure distribution among controls (for long-term effects on blood markers and MI), controls with no missing data (for short-term effects on blood markers), or all exposure intervals with no missing data (for short-term effects on MI, case-crossover analyses).

In the BAMSE study (Papers III, IV), tobacco smoke exposure was coded as a dichotomous variable (0/1) in most analyses, and additionally as a categorical variable based on the amount of cigarettes smoked per day by the mother in some analyses (Paper III).

In both studies, SHEEP and BAMSE, genetic exposure was assessed either in additive (effect per variant allele, i.e. coded 0-1-2), dominant (0, 1 – with 1 for heterozygote and homozygote variant allele carriers), recessive (0, 1 – with 1 for homozygote variant allele carriers), or unconstrained genotype-specific (two indicator variables, one for variant heterozygote versus reference homozygote and one for variant homozygote versus reference homozygote) genetic models for SNP variants in studied genes.

Analyses were performed using STATA 9.0 (StataCorp, College Station, TX, USA), apart from the analyses of genome-wide data (Paper IV), where the PLINK-1.07 [87] program was used.

3.4.1 Linear regression analyses for continuous outcomes

Linear regression models were used to estimate the effects of environmental factors or gene-environment interactions on outcomes that were measured as continuous variables. These were levels of blood markers IL-6, TNF-α, CRP, fibrinogen and PAI-1 in the SHEEP study (Papers I, II), and lung function parameters (FEV_1, FVC, FEV_1/FVC%) in the BAMSE study (Paper IV).

SHEEP (Papers I, II): Levels of blood markers IL-6, TNF-α, CRP, fibrinogen and PAI-1 were log-transformed to achieve a more normal distribution. Estimated effects for both long-term source-specific as well as short-term ambient air pollution exposures were calculated as the change in the log level of an outcome blood parameter (presented as per cent change), per exposure increment in each respective air pollution exposure corresponding to the difference between the 5th and the 95th percentile exposure value. For the analysis of air pollution effects on blood marker levels (Paper I), models with full sets of predictors specific for each blood marker were used: age, sex, physical inactivity and level of HDL cholesterol for IL-6; age and sex for TNF-α; age, sex, body mass index (BMI), smoking status, coffee consumption, vegetables consumption and level of insulin for CRP; age, sex, smoking status, BMI, coffee consumption and level of LDL cholesterol for fibrinogen; age, sex, BMI, levels of triglycerides and insulin for PAI-1.

For the analysis of gene-environment interactions (Paper II), the following adjustments were used for the basic model: age and sex for long-term air pollution effects; age, sex and ambient temperature for short-term air pollution effects on blood markers. Introducing additional adjustment covariates into the models from the relevant sets of predictor variables determined for each blood marker in Paper I did not produce any further substantial changes that altered interpretation of results.

BAMSE (Paper IV): Linear regression models were used to assess the genetic effect in children at 8 years of top-lung function associated SNPs obtained in genome-wide association studies in adults,[53 61] as well as interaction of these SNPs with early tobacco smoke exposure. Effect estimates are presented as the change in a lung parameter (FEV_1, FVC) in mL, or as per cent unit change for FEV_1/FVC%. Of the previously reported 37 top SNPs associated with lung function in adults[53 61], 16 were directly available in the BAMSE GWAS dataset, and for 5 additional SNPs proxy SNPs in high linkage disequilibrium (r^2>0.9) were identified, that were present in the GWAS dataset. Some of those proxy SNPs were already included among the 16 directly available SNPs and thus, 18 individual SNPs in 14 genes were investigated in total (Table 2). Genetic models were adjusted for age, sex, height at 8 years, ever doctor's diagnosed asthma (at 1, 2, 4 or 8 years) based on the sample selection study design, and informative principal components (eigenvectors 3 and 4) from the ancestry-informative principal components analysis.

Table 2. Previously reported SNPs associated with lung function in adults, and availability of SNPs for analysis in the BAMSE GWAS dataset.

Reference	SNPs associated with lung function	Chromosome	Gene	Associated with*	Availability in BAMSE dataset	Final list of SNPs for the present analysis
Repapi et al. 2010 [61]	rs2571445	2	*TNS1*	FEV_1	Proxy, r^2=0.96	rs918949
	rs10516526	4	*GSTCD*	FEV_1	Yes	rs10516526
	rs12504628	4	*HHIP*	FEV_1/FVC	No†	
	rs3995090	5	*HTR4*	FEV_1	Yes	rs3995090
	rs6889822	5	*HTR4*	FEV_1	No†	
	rs2070600	6	*AGER*	FEV_1/FVC	Yes	rs2070600
	rs2395730	6	*DAAM2*	FEV_1/FVC	Yes	rs2395730
	rs12899618	15	*THSD4*	FEV_1/FVC	Yes	rs12899618
Hancock et al. 2010 [53]	rs1980057	4	*HHIP*	FEV_1/FVC	No†	
	rs1032295	4	*HHIP*	FEV_1/FVC	Yes	rs1032295
	rs3817928	6	*GPR126*	FEV_1/FVC	Proxy, r^2=1.0	rs11155242‡
	rs7776375	6	*GPR126*	FEV_1/FVC	No†	
	rs6937121	6	*GPR126*	FEV_1/FVC	Proxy, r^2=0.94	rs9496346
	rs11155242	6	*GPR126*	FEV_1/FVC	Yes	rs11155242
	rs2277027	5	*ADAM19*	FEV_1/FVC	Yes	rs2277027
	rs1422795	5	*ADAM19*	FEV_1/FVC	Yes	rs1422795
	rs2070600	6	*AGER*	FEV_1/FVC	Yes	rs2070600
	rs10947233	6	*PPT2*	FEV_1/FVC	No†	
	rs2869967	4	*FAM13A*	FEV_1/FVC	Yes	rs2869967
	rs6830970	4	*FAM13A*	FEV_1/FVC	Yes	rs6830970
	rs16909898	9	*PTCH1*	FEV_1/FVC	Proxy, r^2=1.0	rs10512249‡
	rs10512249	9	*PTCH1*	FEV_1/FVC	Yes	rs10512249
	rs1435867	2	*PID1*	FEV_1/FVC	Yes	rs1435867
	rs10498230	2	*PID1*	FEV_1/FVC	No†	
	rs11168048	5	*HTR4*	FEV_1/FVC	No†	
	rs7735184	5	*HTR4*	FEV_1/FVC	No†	
	rs17331332	4	*NPNT*	FEV_1	No†	
	rs17036341	4	*NPNT*	FEV_1	No†	
	rs11727189	4	*INTS12*	FEV_1	No†	
	rs17036090	4	*INTS12*	FEV_1	No†	
	rs17036052	4	*FLJ20184*	FEV_1	No†	
	rs17035960	4	*FLJ20184*	FEV_1	Yes	rs17035960
	rs11097901	4	*GSTCD*	FEV_1	Proxy, r^2=1.0	rs10516526‡
	rs11728716	4	*GSTCD*	FEV_1	No†	
	rs3749893	6	*TSPYL4*	FEV_1	No†	
	rs1052443	6	*NT5DC1*	FEV_1	No†	
	rs6555465	5	*ADCY2*	FEV_1	Yes	rs6555465
	rs7710510	5	*ADCY2*	FEV_1	Yes	rs7710510

*Associations found in the studies of Repapi *et al*. 2010 [61] and Hancock *et al*. 2010 [53].
†SNP not available, nor proxy at r^2>0.9. ‡Proxy was already included in the final list.

3.4.2 Logistic regression analyses for dichotomous outcomes

Logistic regression models were used to estimate odds ratios (OR) with 95% confidence intervals (CI) for the effects of environmental factors or gene-environmental interactions on the binary outcomes. These were development of MI and acute onset of MI in the SHEEP study (Paper II), and asthma, wheezing and sensitization in the BAMSE study (Paper III).

SHEEP (Paper II): Logistic regression analysis was used to estimate the risk of developing MI in relation to air pollution exposure and genetic variants. For long-term air pollution effects in the case-control study, logistic regression models were adjusted for age, sex and hospital catchment area. Further adjustment for previously identified covariates for MI risk[32] did not introduce any substantial changes in the results. Conditional logistic regression was used in the case-crossover analysis for matched sets of case and control intervals to estimate the acute risk of MI onset in relation to short-term air pollution exposure. Models in the case-crossover analysis were adjusted for ambient temperature.

BAMSE (Paper III): Logistic regression analysis was performed to estimate ORs for current asthma, wheeze and sensitization in relation to genetic factors and tobacco smoke exposure. Final results are presented in crude (unadjusted) models, since further adjustment for relevant covariates including sex, mother's age, breast feeding, socioeconomic status, environmental tobacco smoke exposure at 1, 2, and 4 years did not change the results appreciably.

3.4.3 Gene-environment interactions

SHEEP (Paper II): The analysis of gene-environment interactions was performed in two steps. First a likelihood ratio test was used to assess the significance of models with interaction, compared to models containing just air pollution and genetic variables without the interaction term. Each SNP was assessed in additive, dominant and recessive genetic models. This assessment provided the basis for identifying interaction patterns and selecting best-fitting models. Secondly, a graphical assessment was performed to confirm the interaction pattern and best-fitting model, by plotting results for the unconstrained genetic models (genotype-specific risk), except for PM_{10} exposure where due to lower statistical power the dominant model for the variant allele was used.

BAMSE (Paper III): Likelihood ratio testing was performed to assess the significance of gene-environment interaction parameters, as described for SHEEP above. The interaction effects of gene polymorphisms and exposure were tested with both a dominant and an unconstrained (genotype-specific) model. For the TNF interaction analyses, only dominant models were tested because of the few children with rare homozygote genotypes (due to a low rare allele frequency).

BAMSE (Paper IV): Assessment of gene-environment interaction for selected top SNPs in the GWAS data was performed by evaluating the significance of the interaction parameter coefficient in a regression model that included SNPs coded in the

additive genetic model and an interaction parameter ($P<0.05$ as standard significance level, and $P<0.1$ for a suggested interaction).

4 RESULTS

4.1 CARDIOVASCULAR EFFECTS IN ADULTS

4.1.1 Air pollution and inflammation

Significantly higher IL-6 levels were found after long-term exposure to elevated residential levels of traffic-related NO_2 as well as combustion-heating-related SO_2 (Fig. 3). Similarly positive effect estimates were also seen for CRP levels, although with more marginal significance. 30-year exposure to traffic-NO_2 was associated with an IL-6 level increased by 64.5% (95% CI 6.7% to 153.8%) per 28.8 $\mu g/m^3$ traffic-NO_2 (corresponding to the difference between the 5^{th} and the 95^{th} percentile exposure value), and 30-year exposure to heating emissions (measured as heating-SO_2) increased the level of IL-6 by 67.6% (95% CI 7.1% to 162.2%) per 39.4 $\mu g/m^3$ heating-SO_2 (5^{th} to 95^{th} percentile value difference). The traffic-NO_2 effects appeared to be more pronounced in non-smokers, physically active people, and in hypertensive individuals, although the differences in effect across subgroups were not statistically significant. The pattern of effect modification was similar for heating-SO_2. Long-term exposure to source-specific air pollution was not associated with significant differences in the levels of TNF-α, fibrinogen or PAI-1.

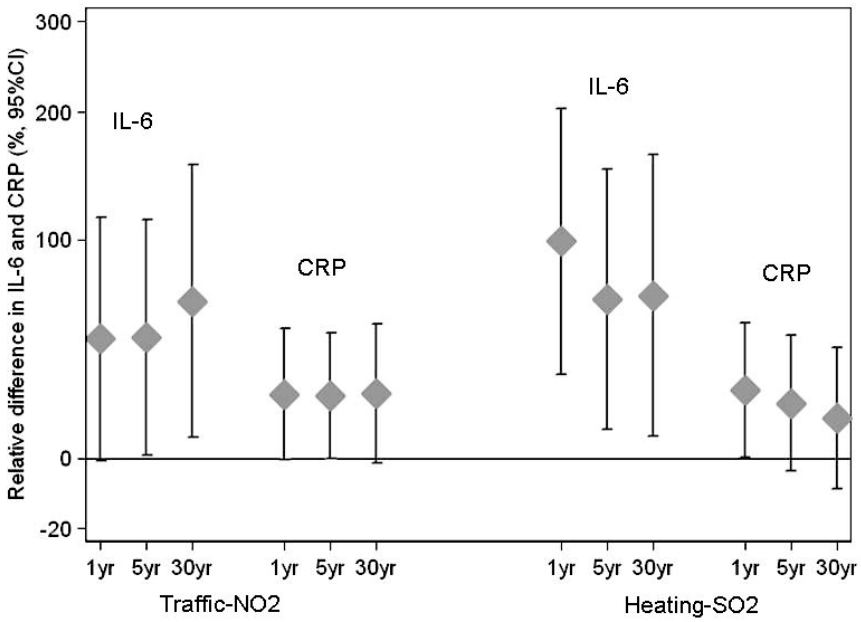

Figure 3 Effects of long-term exposure to traffic-NO_2 and heating-SO_2 on IL-6 and CRP levels in population control subjects from the SHEEP study.

Analysis of the effects of short-term exposure to elevated urban background levels of ambient air pollution on blood markers suggested a possible, albeit not statistically significant, association of increased levels of the inflammatory markers IL-6 and TNF-α with exposure to NO_2 and PM_{10} 0-12 h and 12-24 h before blood sampling, and CRP levels related to exposure 2 and 5 days before (Fig. 4).

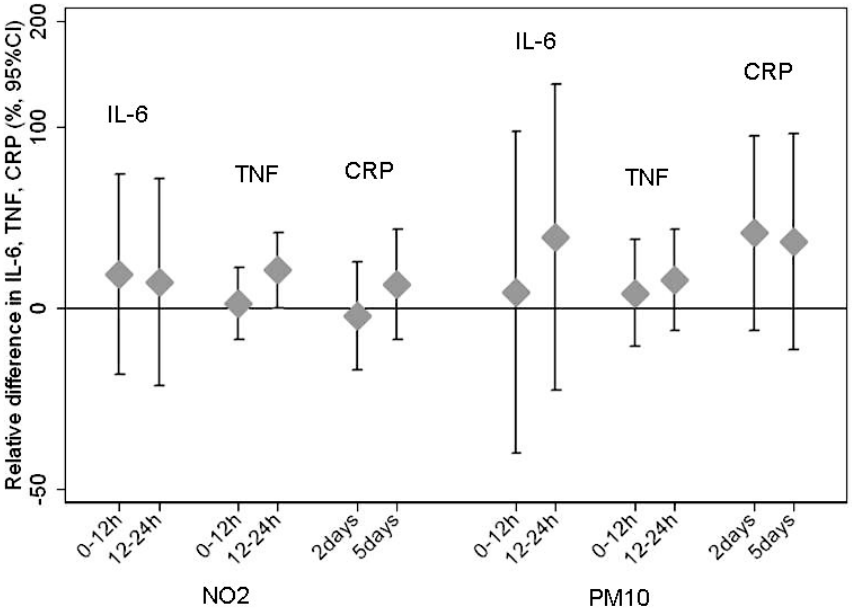

Figure 4 Effects of short-term exposure to ambient NO_2 and PM_{10} on IL-6, TNF-α and CRP levels in population control subjects from the SHEEP study.

4.1.2 Air pollution and genetic effects on inflammation and myocardial infarction

Long-term air pollution showed interactions with both *IL6* and *TNF* polymorphisms in relation to the levels of the respective inflammatory blood markers, and in addition a consistent tendency for the same pattern was seen for MI risk. For example, long-term exposure to traffic-NO_2 interacted with *IL6*-174G/C in an additive pattern (Fig. 5, left panel), where each additional variant C allele was associated with incremental increase in the air pollution effect on IL-6 levels. The MI risk analogously increased with each additional variant C allele, for all studied exposure windows (Fig. 5, right panel). A similar interaction pattern was seen for the *IL6*-598G/A SNP which was in high LD with *IL6*-174G/C (r^2=0.93). For TNF, interaction with long-term air pollution was seen for *TNF*-308G/A, with the rare homozygote -308AA being associated with elevated TNF-α levels after long-term exposure both to traffic-NO_2 and heating-SO_2 (recessive genetic pattern), and a consistent pattern seen for MI risk, although the statistical power

of this analysis was limited (Fig. 6). No interactions with long-term air pollution exposure were seen for coagulation gene variants in *FGB* and *PAI-1*.

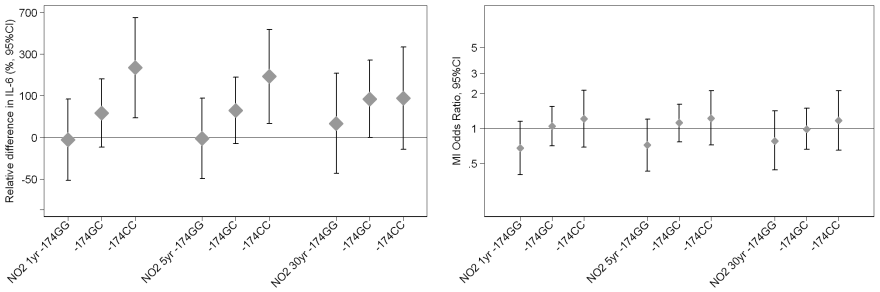

Figure 5 Gene-environment interactions of the *IL6*-174G/C variant with long-term traffic-NO_2 exposure in relation to IL-6 levels (in population controls) and MI risk (case-control analysis) in the SHEEP study.

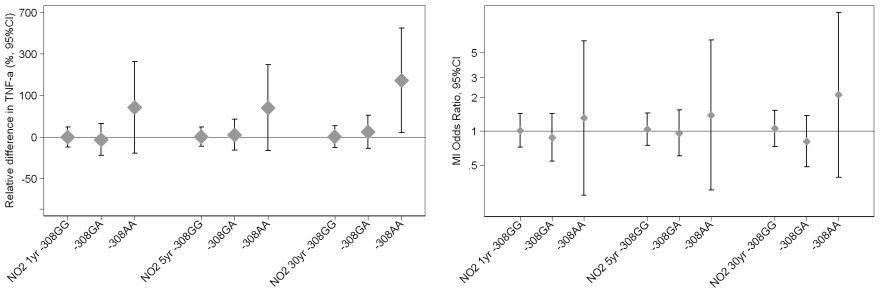

Figure 6 Gene-environment interactions of the *TNF*-308G/A variant with long-term traffic-NO_2 exposure in relation to TNF-α levels (in population controls) and MI risk (case-control analysis) in the SHEEP study.

The analysis of short-term air pollution exposure interacting with *IL6*-174G/C and -598G/A resulted in a reverse pattern, where the variant C or A allele was associated with decreased IL-6 levels related to air pollution (Fig. 7) The corresponding pattern for the risk of MI onset across genotypes was relatively concordant, reflecting higher MI risks from air pollution among subjects lacking the variant allele (common -174GG and -598GG homozygotes). For these interactions, the best-fitting one degree of freedom genetic model was the dominant model, especially for 2- and 5-day intervals (Fig. 7). Similarly, the *TNF*-308AA genotype was associated with lower levels of TNF-α after short-term exposure to ambient SO_2 (Fig. 8); again, the pattern for MI onset was similar. Two other *TNF* SNPs, -863C/A and -1031T/C (in moderate LD, r^2=0.69) showed interactions with short-term PM_{10} exposure, with common homozygotes having a positive association of air pollution with TNF-α levels while variant allele

carriers did not. This pattern was similar for MI onset. No consistent interaction patterns with short-term air pollution exposure were found for coagulation gene variants in *FGB* and *PAI-1* in relation to the levels of respective blood markers and MI risk.

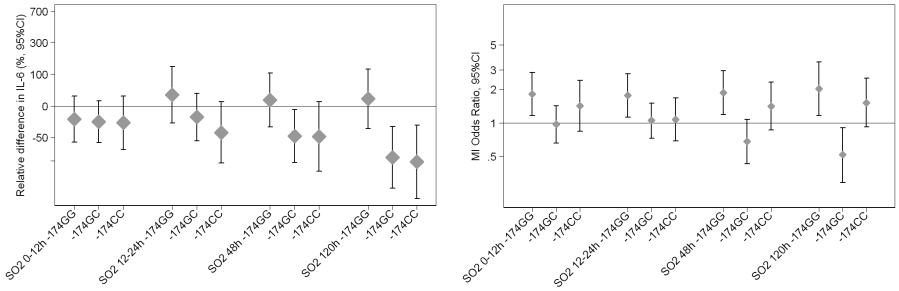

Figure 7 Gene-environment interactions of the *IL6*-174G/C variant with short-term ambient SO_2 exposure in relation to IL-6 levels (in population controls) and risk of MI onset (case-crossover analysis) in the SHEEP study.

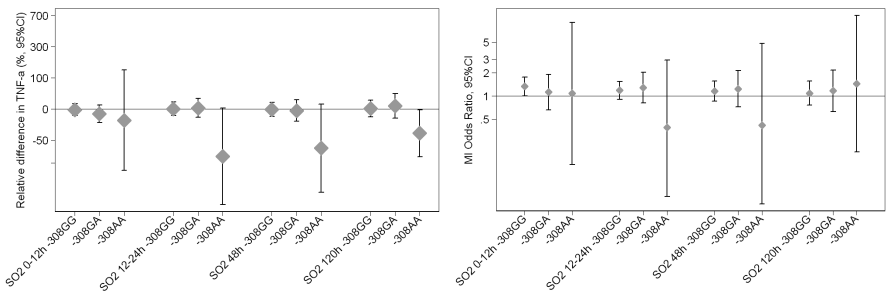

Figure 8 Gene-environment interactions of the *TNF*-308G/A variant with short-term ambient SO_2 exposure in relation to TNF-α levels (in population controls) and risk of MI onset (case-crossover analysis) in the SHEEP study.

4.2 RESPIRATORY EFFECTS IN CHILDREN

4.2.1 Tobacco smoke and genetic effects on wheezing

Main genetic effects were identified for the *TNF*-308G/A, *GSTP1* Intron 5 and *GSTP1* Ala114Val SNPs in relation to current asthma, the *ADRB2* Exon 1 SNP in relation to early-onset wheeze and the *TNF*-308G/A in relation to allergic sensitization.

Analysis of gene-environment interactions between the selected SNPs and early maternal smoking identified interactions for *TNF* SNPs in the promoter region (-857C/T) and in introns 1 and 3 for early-onset wheeze. The interaction models indicated higher risk for developing early-onset wheeze related to early maternal

smoking in wild-type CC homozygotes of the *TNF* SNPs -857C/T and Intron 1 (OR=2.4, 95% CI 1.6 to 3.7 and OR=2.5, 95% CI 1.6 to 3.8, respectively), while carriers of the rare T allele of these SNPs did not experience an elevated risk. For carriers of the rare G allele in the *TNF* Intron 3 SNP, the suggested increased risk was even greater (OR=6.1, 95% CI 1.9 to 19.6), while common AA homozygotes had a less clear increase in risk.

In the dose-response analysis based on the highest number of cigarettes smoked by the mother per day during any of the pregnancy trimesters (0 [reference], 1-9, and >9 cigarettes), carriers of the common *TNF* -857 and Intron1 CC genotype showed a dose-dependent risk pattern (P<0.001 for trend for both SNPs) and the highest risk of developing wheeze was detected in children exposed to >9 cigarettes per day (OR=3.4, 95% CI 1.9 to 6.2 (Fig. 9), and OR=3.4, 95% CI 1.9 to 6.2, respectively).

A statistically significant interaction with early maternal smoking was also noted for some *GSTP1* SNPs on transient wheeze. Carriers of the rare T allele in the Introns 5 and 6 SNPs, or the Ile105Val G allele (coding for Val), who were exposed to early tobacco smoke, had an increased risk of developing transient wheeze, with the highest OR in children with a rare homozygote genotype (e.g. OR=8.5, 95% CI 1.9 to 37.2 for *GSTP1* Ile105Val rare homozygote Val/Val carriers).

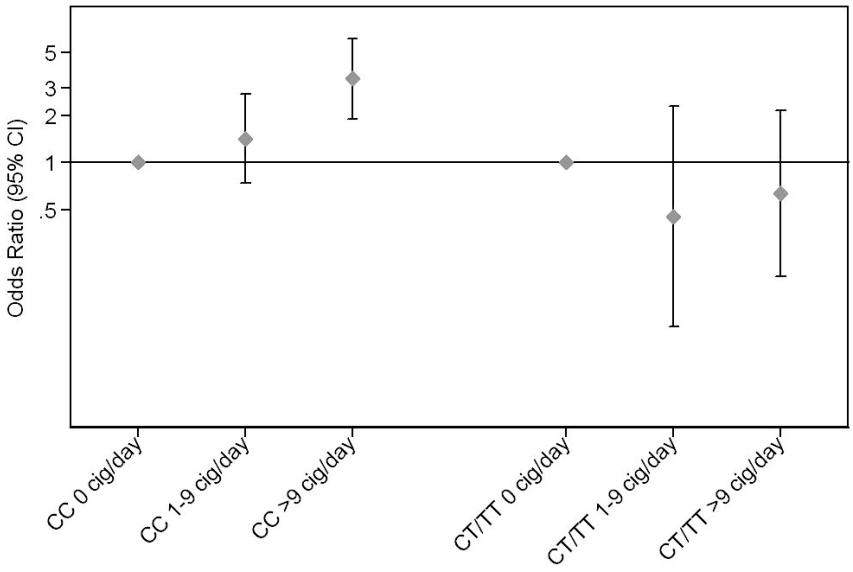

Figure 9 Odds ratios for early onset wheeze in children followed up to 4 years in relation to the extent of early maternal smoking during pregnancy (no, moderate: 1-9 cigarettes per day, and high: >9 cigarettes per day), stratified by genetic variation in the *TNF*-857C/T SNP.

4.2.2 Tobacco smoke and genetic effects on lung function

When attempting to replicate a main genetic effect on lung function in children, SNPs in the genes *TNS1*, *ADCY2*, *ADAM19* showed the strongest and most significant associations with FEV_1/FVC% (Fig. 10). On the other hand, rs12899618 in *THSD4* was consistently associated with both FEV_1 and FVC separately, but not with FEV_1/FVC%. For *HHIP* a suggestive effect on FVC was seen. The minor allele frequencies were quite similar to what has been reported in the previous studies for all the investigated SNPs [53 61]. Out of 14 combinations of gene and lung function parameters (FEV_1, FEV_1/FVC) reported previously as significant findings[53 61] that were investigated here, 11 were in the same direction as the original finding. Among the genes with opposite direction of effect was *ADCY2*, where positive associations were observed with both FEV_1/FVC% (significant), and FEV_1 (non-significant), while Hancock *et al*. reported a negative significant association with FEV_1 as the discovery finding. [53]

Early maternal smoking and current tobacco smoke exposure were associated with lower FEV_1/FVC% and FEV_1, although statistically significant only for current exposure and FEV_1/FVC%. Separate analysis of tobacco smoke exposure effects in asthma cases and in controls did not show major differences.

Since clearer associations were observed for both SNPs and tobacco smoke exposure with FEV_1/FVC%, gene-environment interactions were only tested with this outcome. Recent exposure at 8 years seemed to interact with the *DAAM2* rs2395730, where variant C allele was associated with 2.1% units lower FEV_1/FVC% per allele in exposed children, but not in unexposed children (P for interaction 0.05). In addition, there was some suggestion of a positive interaction for *ADCY2* rs7710510 and rs6555465, with variant alleles T and A associated with 2.3% and 2.4% unit higher FEV_1/FVC% in children with recent exposure, but not in children without exposure (P for interaction 0.1 and 0.08, respectively). Both of these SNPs, as noted above, also showed a significant main effect in relation to the same phenotype. Another potential interaction was noted for *FAM13A* rs6830970 with early maternal smoking, where the variant G allele was associated with a 2.3% unit lower FEV_1/FVC% in exposed children, but not in unexposed (P for interaction 0.07).

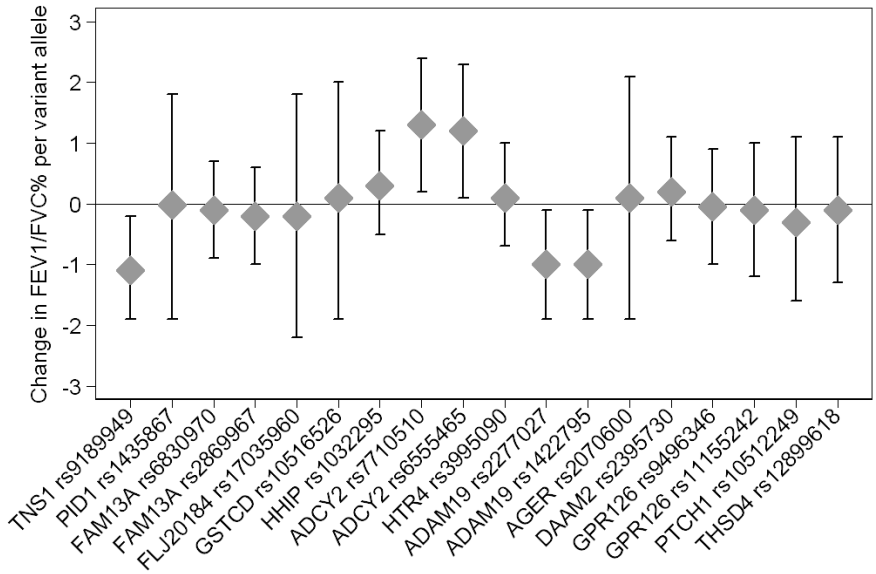

Figure 10 Genetic effect per variant allele on FEV_1/FVC (%) in children at 8 years from the BAMSE study.

5 DISCUSSION

5.1 CARDIOVASCULAR EFFECTS IN ADULTS

The present findings support the view that air pollution can affect the levels of certain inflammatory markers after both long- and short-term exposures to air pollution. At the same time, these effects can be modified by genetic polymorphisms in inflammatory genes, determining not only the levels of inflammatory blood markers, but also the consequent MI risk.

Previous epidemiological and experimental studies have reported contradictory results for the effects of short-term exposure to air pollutants on levels of inflammatory blood markers. Experimental exposure of volunteers to PM resulted in increased fibrinogen levels but showed no association with levels of IL-6, TNF-α, CRP or PAI-1. [18 19] In a study exposing men with prior myocardial infarction to diluted diesel exhaust, no changes were detected in the levels of CRP and PAI-1. [20] Short-term exposure to particulate matter has been associated with an increase in IL-6 levels [21 43] or shown no relationship. [22] Similarly, some epidemiological studies of short-term PM exposure reported positive association with CRP levels [22 45] or no association. [21 44] Based on previous evidence and the present results (Paper I), it is possible that a response to short-term air pollution exposure has considerable variation, depending on genetic factors, lifestyle characteristics, or on the nature (chemical composition) of the pollutant mixture. Potentially, continuous exposure may lead to sustained increased levels of inflammatory markers and induction of chronic low-grade inflammation which may modify short-term effects.

The apparent stronger effects of long-term exposure to traffic-NO_2 in non-smokers might be explained by a lack of competing influences from tobacco smoke in inducing low-grade inflammatory responses.[88] Thus, the pulmonary and systemic inflammatory reactivity to an additional air pollution load may be more difficult to capture clearly in smokers as compared to non-smokers.

Association of elevated IL-6 levels with coronary heart disease (CHD) and MI has been reported in several studies, including SHEEP. [36 89] At the same time both air pollution exposure [21] and genetic variants [42] have been linked to IL-6 levels. However, even for the widely studied *IL6*-174G/C variant, there is conflicting evidence regarding its association with MI. [90-93] In Paper II, a gene-environment interaction was observed for the *IL6*-174C allele, with increased IL-6 levels and a consistent trend of increased MI risk after long-term traffic-NO_2 exposure. On the other hand, protective effects in terms of lower IL-6 levels and risk of MI onset were noted in carriers of the -174C allele after short-term ambient SO_2 exposure. Similar interaction results were seen for *IL6*-598G/A, in accordance with its high LD with *IL6*-174G/C. The studies that have investigated interaction of *IL6* variants with air pollution have shown mixed results, but overall the existence of some kind of interaction is suggested. For instance, *IL6* SNP rs2069832 modified the effect of air pollution on IL-6 levels in a study of MI survivors. [62] A population-based study found the variant -174C allele to be associated with increased all-cause mortality risk among 80-year-old subjects only in non-smokers, [64]

while another study in middle-aged men reported -174C allele to be associated with increased MI risk only in smokers. [65]

The present study also identified interactions with *TNF* gene variants. There have been conflicting results regarding the association of the *TNF*-308G/A polymorphism with MI [94-96]. A recent meta-analysis of 24 studies reported an association of the variant *TNF*-308A allele with developing CHD (OR 1.5, 95% CI: 1.23-1.77 for -308A allele carriers compared to common -308GG homozygotes) in Caucasian populations, but not in other ethnic groups. [97] The present investigation showed increased levels of TNF-α in *TNF*-308AA homozygotes associated with long-term exposure to traffic-NO_2 and heating-SO_2, while -308G allele carriers did not show any air pollution effect. However, after short-term exposure to ambient SO_2 this pattern was reversed, and *TNF*-308AA homozygotes had slightly decreased TNF-α levels as well as an indication of decreased risk of MI onset. The variant -308A allele has been associated with higher constitutive and inducible TNF-α levels, [98] and possibly this balance differs after long-term and short-term environmental exposures affecting TNF-α levels.

Interestingly, both *IL6* and *TNF* SNPs showed a reverse pattern of genetic modification of associations for blood marker levels and MI for long- vs. short-term exposure to air pollution. One potential explanation might be that when particular genotypes contribute to an elevation of inflammatory marker levels over a long period of time, certain saturation occurs. In this case additional stimulus may not result in further increase in blood marker level. An analogous effect trend was observed for the overall associations of IL-6 levels with air pollution exposure (Paper I), where long-term air pollution exposure did not contribute to an increase of IL-6 levels in current smokers, whereas never-smokers responded with elevated IL-6. In a healthy organism IL-6 is expressed in low levels and is controlled by a variety of mechanisms, while its expression is rapidly induced by infection, trauma, or other stress conditions. [99] During the immediate response to stressors, IL-6 mediates an acute phase response, whereas persistent activity of IL-6 contributes to a switch from acute to chronic inflammation. [13]

The presented results thus confirm the hypothesis of induction of systemic inflammation due to exposure to air pollution. Several genetic variants in inflammatory genes were identified that define vulnerable groups of people, where exposure to elevated levels of air pollution affects both levels of inflammatory markers and produces a suggested increase in the risk of MI. Further studies are needed to confirm this genetic effect modification, clarifying also the mechanisms that may underline varying response to long-term vs. short-term air pollution exposure. Future research in this direction may form a basis for advice regarding preventive measures (e.g. avoidance of air pollution exposure) for susceptible groups, as well as for developing therapeutic interventions.

5.2 RESPIRATORY EFFECTS IN CHILDREN

The present results indicate the presence of gene-environment interactions between early maternal smoking and *TNF* and *GSTP1* polymorphisms in relation to development of early onset wheeze in children up to 4 years of age. Among the investigated SNPs previously associated with lung function in adults, several associations were also reproduced in children, with identifying suggestive interaction of *DAAM2* variant with tobacco smoke exposure.

Previous studies of interaction between tobacco smoke exposure and *TNF* polymorphisms in relation to respiratory symptoms in children focused mainly on the *TNF*-308G/A SNP. This polymorphism modified the risk for respiratory illness-related school absences associated with second-hand smoke exposure, with -308A carriers exposed to tobacco smoke having higher risk than unexposed homozygotes for the common -308G allele. [71] An increased risk of childhood asthma for the rare *TNF* -308A and -238A alleles was seen for children of non-smoking parents, while among children with smoking parents none of the studied *TNF* polymorphisms were associated with asthma. [72]

In the present study, a greater range of variants across the *TNF* gene were investigated. The basis for investigating the interaction of *TNF* SNPs in Paper III was primarily the biological hypothesis of involvement of inflammation (with induction of pro-inflammatory TNF-α) in the pathogenesis of respiratory symptoms caused by tobacco smoke exposure. Notably, a main genetic effect of the *TNF* -308G/A SNP was detected on current asthma and sensitization, whereas interaction with early maternal smoking was seen for other *TNF* variants, which, to our knowledge, were not previously investigated in terms of an interaction with *in utero* tobacco smoke exposure. The selection of SNPs for interaction analysis without considering their main genetic effects has been discussed, and is based on the view that SNPs with modest genetic effects under the influence of an environmental factor have more profound, and potentially better detected, effects.[100] In the present study, a clear dose dependent effect was also seen in children with the risk genotypes, which supports the biological relevance of findings.

Potential explanations for detecting interactions with the *TNF* -308G/A variant in other studies, but not in the present study, may be related to the study design: different outcomes, age of participants, and consideration of various time windows of tobacco smoke exposure. The results in the present study were highly concordant for the *TNF*-857C/T and Intron 1 SNPs due to a high linkage disequilibrium between them (r^2=0.93). It is also possible that the *TNF* -857C/T, Intron 1, and -308G/A SNPs may be linked to other functionally important polymorphisms within the gene, with different linkage disequilibrium patterns in different populations.

Differences in host detoxification ability, including antioxidant defense (where GSTP1 is an integral part), may determine the balance between resolution or progression of inflammation after exposure to an environmental trigger. [101] Most studies have focused on the coding *GSTP1* Ile105Val SNP. Conflicting results have been reported regarding a possible protective effect of the variant homozygote *GSTP1* Val105Val genotype for allergy and asthma. [76 102 103] Children carrying a Val105 allele exposed to *in utero* tobacco smoke have even been reported to have about twice the risk of current wheezing and early-onset asthma as compared to Ile105Ile homozygote children without TS exposure. [70] Significant gene-environment interaction has also been found for the Ile105Val polymorphism in Taiwanese children, where Ile105 homozygotes exposed to high levels of air pollution had significantly higher risk of asthma compared to Val105 carriers exposed to low levels. [73] A previous report from the BAMSE cohort identified an increased risk of allergic sensitization in children with the *GSTP1* Val105

allele exposed to elevated levels of traffic-related air pollution compared to the risk in exposed Ile105 homozygotes. [76] In the present study there were indications of an interaction between early maternal smoking and three *GSTP1* SNPs, including the Ile105Val SNP (Val carriers having the highest risk also in this study), in relation to the development of wheeze, which supports a role of *GSTP1* in relation to environment exposures and asthma-related traits in children. Further research could extend these findings by addressing issues of the functional activity of *TNF* and *GSTP1* SNPs, as well as the type and critical time intervals of environmental exposures that may be of particular importance for various asthma-related phenotypes.

As opposed to identifying gene-environment interactions in the first place, another view suggests instead first identifying or exploring the genetic variants with strong main genetic effect, and then testing them for interactions. Given the availability of emerging data from GWAS studies, it seems tempting to use this opportunity for identifying interactions, using the GWAS findings as candidate gene input for detailed gene-environment analysis. At the same time, full-scale genome-wide interaction studies are not numerous at present and they face challenges in reliably detecting significant interactions, as discussed in more detail under 5.3 below.

In Paper IV, top SNP findings from adult GWAS studies associated with lung function, were thus tested in BAMSE data in terms of replication of main genetic effects, as well as for interaction with TS exposure in relation to lung function in children. Variants in the genes *TNS1*, *ADCY2*, *ADAM19* and *THSD4* showed the strongest and most consistent association with lung function in children, with possible association also for *HHIP*. Functionally, apart from the reported association with lung function, [53 61] the mentioned genes are involved in the processes of inflammation, morphogenesis, and healing of injury. The ADAM metallopeptidase domain 19 gene (*ADAM19*) is expressed in the apical part of the bronchial epithelium, as well as in smooth muscle and inflammatory cells, which suggests involvement in early immune defense mechanisms and perpetuation of the inflammatory process. [104] The *THSD4* gene codes the thrombospondin, type 1, domain-containing protein 4, which shows homology with other members of the thorombospondin gene family implicated in wound healing, inflammation, and angiogenesis. [105] Both tensin 1(*TNS1*) and *HHIP* genes showed an association with adult height in a recent meta-analysis. [106] The hedgehog interacting protein gene (*HHIP*) is involved in the hedgehog signaling pathway, which influences lung development [107] and is activated in the airway epithelium during repair of acute injury. [108] A genome-wide association study of pulmonary function in the Framingham Heart Study identified a region on chromosome 4q31 near the *HHIP* gene to be associated with percent predicted FEV_1/FVC. [109] In another study the *HHIP* variant rs13118928 was associated with FEV_1/FVC, fat-free body mass and chronic obstructive pulmonary disease (COPD) exacerbations, but not with smoking intensity. [110] In the present study, the rare allele of HHIP rs1032295 showed a negative association mainly with FVC, suggesting an early influence on development of lung volume, with potential adult consequences.

In terms of interactions investigated in the present study, children with recent exposure to tobacco smoke at 8 years showed a potential interaction of TS exposure with variants in the disheveled associated activator of morphogenesis 2 (*DAAM2*) and adenylate

cyclase 2 (*ADCY2*) genes in relation to lung function at 8 years. Suggestive interaction was also shown for early maternal smoking with variants in *FAM13A* gene in relation to lung function at 8 years. The *ADCY2* gene has been identified in a set of genes reported to affect smoker's abilities to quit smoking. [111] The family with sequence familiarity 13, member A gene (*FAM13A*) contains two linked SNPs (rs1903003 and rs7671167) which were associated with COPD, but not with pack-years of cigarette smoking in three cohorts of current or former smokers. [112] This finding was replicated in two COPD cohorts, which confirmed the association of the *FAM13A* SNP rs7671167 with lung function. [110]

Gene-environment interactions identified in the present study confirm the hypothesis that tobacco smoke exposure induces inflammatory response with negative consequences for respiratory system both in terms of clinical symptoms of wheezing, but also in terms of lung function in children. When studying environmental exposures in children, it is very important to identify the most sensitive exposure time windows in order to adequately support any suggested measures to protect children's health. The present study suggests negative effects on respiratory system of both early (peri-natal) and later (at 8 years) childhood TS exposures. Children carrying certain genotypes were suggested to be more susceptible to the negative effects of tobacco smoke exposure, which, after further confirmation in epidemiologic and functional studies, may be useful for reinforcing counseling to parents aimed for avoidance of harmful exposures.

5.3 METHODOLOGICAL CONSIDERATIONS

Sample size

The studied populations in both SHEEP and BAMSE studies were large and both were characterized by relatively high participation rates. There was quite good statistical power to detect gene-environment interactions, especially in the SHEEP study with blood marker levels as an outcome (Paper II). However, there was more of an issue with statistical power when investigating interactions with genetic variants of low minor allele frequency, e.g. some in the *TNF* and *IL6* genes (Papers II and III). In the lung function replication study (Paper IV), the study population included asthma cases and controls on whom there was genome-wide genetic data available. Although this sample size had rather low power to study genetic associations, it was nevertheless possible to replicate several of the top adult SNP findings, and consistency of direction was quite encouraging. Most of these SNPs were relatively common due to the selection of tagging SNPs for GWAS chips, which was helpful for power. For investigating gene-environment interactions with these SNPs, however, the lack of power became more obvious. Larger studies are necessary to investigate the effects of rare genetic variants interacting with environmental factors, which may include international collaborations. However, such studies should include populations with rather similar environmental factors, and aim for as much ethnic and genetic homogeneity as possible in order to successfully address the gene-environment interactions.[54]

Exposure assessment

The adult studies in SHEEP presented here are unique in addressing in the same population effects on inflammation and coagulation blood markers (Paper I) and gene-

environment interactions (Paper II) of both long and short-term exposures to air pollution. Consequently, it was feasible to identify exposure time windows of higher risk, compare the differences between various time periods and between individual pollutants (short-term) or source-related pollution mixes (long-term).

Long-term air pollution exposure was assessed using spatial modeling of emissions at each subject's residential addresses up to 30 years retrospectively, taking into account changes of residence. This method relies on retrospective emissions estimates over a long period of time, which can contribute to exposure measurement error. Nevertheless, this is likely to be a non-differential misclassification, resulting in attenuation on average of estimated effects. The variables used were designed as indicators for different pollution mixes (traffic-related exhaust or heating-related emissions). Therefore, it is difficult to disentangle the specific inflammation-inducing agent, given that both gaseous and particulate compounds of air pollution may be capable of inducing an inflammatory response in the lungs. The estimates of air pollution exposure were based on residential addresses, and do not address exposures at work, travel and leisure activities. However, some of the most relevant exposures, such as occupational exposure to diesel and motor exhausts, were assessed separately for their potential confounding effect. No strong confounding was identified by these exposures. Since the traffic-NO_2 and heating-SO_2 variables are emission-based indicators for source-related pollutant mixtures, they represent local emission sources and do not incorporate, for example, long-range or transboundary air pollution. Long-range transport would not normally, however, be expected to contribute significantly to inter-individual variation within a defined geographical area. More sophisticated models of retrospective estimation of air pollution exposure could possibly be elaborated based on developing monitoring procedures and growing international collaboration in this field.

Short-term exposure to air pollution was estimated based on arithmetic means of hourly inner-city ambient rooftop measurements. Such estimates mainly reflect temporal, but not geographical variation, since the pollutant levels vary over the relatively large geographical study area and, for example, can be lower in suburban areas. This exposure assessment methodology is commonly used for studies of short-term variation in air pollution levels and its usefulness relies on the assumption that although absolute levels may vary across the study area, the central measurements reflect well the variation in exposure levels in different locations. This argument is stronger for a time-series design, however, and more misclassification of exposure might be expected in a cross-sectional study, but such misclassification of exposure would likely be independent of the measured outcomes, resulting in some loss of power and attenuation on average of estimated effects. For a better estimation of short-term air pollution exposure, a more targeted sampling of subjects focusing on the vicinity of monitoring stations as well as modeling of levels more distant from the stations could be ways forward.

Exposure to tobacco smoke was assessed based on the information from questionnaires in the BAMSE study. It was collected repeatedly, also prior to any adverse health symptoms (wheeze or asthma), which enabled non-biased exposure estimates, especially for very early (*in utero*) exposures. The available tobacco smoke exposure information is detailed on both extent and timing, from pregnancy and up to 8 years of

age. Such data potentially provide a basis for identifying the most important exposure time windows (intrauterine exposure, postnatal, current). However, as often noted in studies of early tobacco smoke exposure, it is difficult to disentangle exposure during pregnancy with early postnatal exposure, since smoking mothers usually smoke during both periods,[113] and it is hard to get enough statistical power for specific pre-natal and post-natal exposures. Information on passive smoking of pregnant mothers was also not available in the present study. It could also be advantageous in future studies to include objective measurements of tobacco smoke exposure, for example cotinine levels in blood, urine or saliva of a pregnant mother and a child, although this only reflects relatively short-term exposure.

Outcome definitions

Data were available in the SHEEP study to assess the effects of air pollution and gene-environment interactions on both intermediate outcomes reflecting the risk of disease development (IL-6, TNF-α and others), as well as on the final outcome, myocardial infarction.

Blood samples were analyzed at the Department of Clinical Chemistry, Karolinska Hospital, in a random, blinded manner to reduce any bias. The assessment of blood marker levels was based only on one reading per subject, and will be affected by measurement error and intra-individual variation related to circadian rhythms, inflammatory conditions, stress levels, etc. Again, such measurement error should result in some power loss and tendency to attenuation of estimated effects. In future research it could be advisable to reduce such effects by increasing the study population, as well as by performing multiple sampling in the same individual accompanied by information on any relevant characteristics affecting short-term variation.

MI cases were identified using standard diagnostic criteria accepted by the Swedish Association of Cardiologists in 1991.[78] The exact time of MI symptoms onset, necessary for the case-crossover analysis (Paper II) was determined primarily using the information from medical clinics. The study of genetic factors was limited to non-fatal MI cases who had provided a blood sample, leaving open the question of whether genetic variants affecting susceptibility to air pollution effects may be differently represented in fatal MI cases. For testing the causative pathway of gene-environment interactions affecting inflammatory marker levels with consequent effect on the risk of MI, it would be advantageous to separately investigate both fatal and not fatal MI as an outcome.

Wheezing up to 4 years assessed in children in the BAMSE study was based on information from repeated detailed questionnaires. Specifically for the assessment of effects of early maternal smoking exposure (during pregnancy), it could be useful to measure the levels of inflammatory markers (e.g. TNF-α) in children at birth. Such measurements were not available in the present study. Lung function was tested in children at 8 years. All children performed several maximum expiratory flow volume measurements (MEFV), the MEFV curves passed visual quality inspection, and the two highest readings were reproducible according to ATS/ERS criteria.[85]

Multiple testing

In the exploration of gene-environment interactions of candidate genes, in both the SHEEP and BAMSE studies, adjustment for multiple testing was not applied. All genes were selected based on a strong prior biological hypothesis, reflecting inflammatory, antioxidant or coagulation responses of the body to air pollutants, and the focus of most analyses was on characterization of an expected association and estimation of effects. Moreover, the number of tests actually performed is extremely difficult to ascertain. The conservative Bonferroni correction, for example, requires independent observations, which does not apply to the studies presented here on several counts: firstly, the air pollution acts as a mix, not as individual pollutants, secondly, SNPs in the genes are in linkage disequilibrium with more or less correlation, and thirdly, the studied outcomes are often inter-related (for example, CRP is produced in response to IL-6 increase in the blood, or wheezing and asthma phenotypes can be mutually inclusive). In genetic analyses, multiple correction is applied in GWAS studies, where exploration of genetic associations with an outcome is completely hypothesis-free and most tested associations are known at outset to be null, a situation that is quite different from the present studies. Even in Paper IV, the SNPs included for replication in children with respect to lung function had high prior probability of yielding true associations, having been both discovered and replicated in adults in the original publications.[53 61] However, in genome-wide interaction studies, multiple correction may also hinder the discovery of true interactions, as noted above.[54]

Gene-environment interactions

Many studies have explored causative environmental factors, as well as genetic associations in complex multifactorial diseases. However, there are still missing pieces of the puzzle, such as the often discussed "missing heritability" that even GWAS studies do not seem to be able to find,[114] or the fact that not everyone exposed to particular environmental agent develops a disease. It seems that focusing on studying gene-environmental interactions in the development of a disease may be a suitable way forward. [74 100]

As mentioned above, different approaches to the study of gene-environment interactions have been suggested. One approach advocates that first disease-associated genes (or their variants) should be identified, and then studied in terms of their interaction with environmental exposure. This approach was applied in the study of lung function in children (Paper IV). Several SNPs associated previously with lung function in adults were replicated. These SNPs fall quite well into expected biological models, since some of them are involved in inflammatory processes and wound healing. However, testing interactions of these SNPs with TS exposure resulted in much more modest results, reflecting lack of power and possibly also that the most obvious gene-environment interactions are not necessarily related to the genes with strongest average overall effects.

Another approach therefore suggests that the marginal effects of genetic variants may often be overpowered by the influence of environmental factors, leading to more pronounced effects to be seen when gene-environment interaction is considered. [73 74 100] Here, searching for interactions becomes a first and primary goal. This approach was applied in the present study (Papers II, III), with interesting results which confirmed

underlying biological hypotheses and even highlighted unexpected trends in gene-environment interactions (opposite direction of effect on inflammatory blood marker levels in specific genotypes after long-term versus short-term exposure to air pollution, with similar opposite trends for long- and short-term MI risk). Supposedly, the success of candidate gene interaction studies is based on studying relatively homogenous environmental exposure as well as ethnic groups (reflecting genetic homogeneity).[54] In addition, the environmental exposure assessment may be better optimized in such a homogenous setting. These pre-requisites are often not present in genome-wide interaction studies, where due to the requirement of very high statistical power, large groups of individuals are recruited, calling for international collaborations across heterogeneous populations, exposures and studies. Therefore, there are not many GWIS published to date, and so far they have not been successful in identifying significant gene-environmental interactions, including even those reported previously from candidate gene interaction studies.[54]

6 CONCLUSIONS

- Exposure to moderate levels of air pollution appears to increase serum levels of inflammatory markers in adults. For instance, long-term exposure to local traffic- or residential heating-related air pollution was associated with increased serum IL-6 levels, and a positive association was suggested also for short-term exposure to NO_2 and PM_{10} for IL-6, TNF-α and CRP.

- Variants in inflammatory genes can modify the intensity of inflammatory response to air pollution, with likely consequences for the MI risk. In the present study, SNPs in *IL6* and *TNF* were seen to modify the effect of long- and short-term exposure to air pollution on the levels of inflammatory markers in healthy adult subjects. Effects by long- and short-term air pollution, respectively, on long-term MI risk and acute MI onset closely followed patterns of effects on inflammatory markers across genotype groups for these modifier SNPs.

- Variants in genes related to inflammation and antioxidant defense appeared to affect respiratory symptoms in children exposed to tobacco smoke. The risk of early childhood wheeze associated with early maternal smoking was seen to be modified by *TNF* and possibly also *GSTP1* polymorphisms. Several SNPs in both genes (coding, promotor, intronic) showed interaction.

- Adult lung function genes are likely to be important already for lung function in childhood. Their interaction with early tobacco smoke exposure requires further study. Among a set of recently identified genes from GWAS of adult lung function, *TNS1*, *ADAM19*, *THSD4* and *ADCY2* variants were significantly associated also with lung function in children, with a consistent direction of effect replicated for most genes. Variants in *DAAM2* may modify the effect of exposure to tobacco smoke on lung function in children.

- It is possible that the identified genetic variants that appear to predispose individuals to adverse cardiovascular and respiratory effects of exposure to ambient air pollutants or tobacco smoke, together with genetic variants identified in other studies, may at a future point in time provide a basis for preventive testing and advice regarding avoidance of pollutants (such as stopping tobacco smoke exposure, moving to cleaner area). Achieving a sufficient predictive value beyond present clinical evaluation is, however, likely to be very challenging, and may only become relevant in settings of expected high exposures, e.g. occupational settings. The cornerstone of prevention is likely to remain a broad general advice and measures to minimize exposures in society.

- More importantly, the results also contribute to the knowledge regarding mechanisms of adverse effects, in illustrating the interaction between environment and the human organism, as represented by the genetic

background. This is likely to contribute in guiding further research on the topics of health effects of ambient air pollution and tobacco smoke exposure. Such information may also potentially be useful for designing therapeutic interventions.

7 SAMMANFATTNING PÅ SVENSKA

Exponering för luftföroreningar kan ge upphov till låggradig systemisk inflammation med konsekvenser för både det kardiovaskulära och det respiratoriska systemet. Det övergripande syftet med denna avhandling var att undersöka effekter av luftföroreningar för att utveckla komplexa inflammatoriska sjukdomar (hjärtinfarkt och luftvägssjukdom), och att identifiera genetiskt känsliga grupper för dessa effekter.

De två studiepopulationerna hämtades dels från en fall-kontrollstudie av hjärtinfarkt (SHEEP) och dels från en födelsekohort (BAMSE). Från SHEEP inkluderades i denna studie 1192 förstagångs fall av hjärtinfarkt i åldrarna 45-70 år från Stockholms län 1992-1994, samt 1536 matchade populations kontroller från samma studiebas. Deltagarna besvarade frågeformulär och genomgick läkarundersökning samt blodprovstagning. Deras exponering för luftföroreningar skattades retroaktivt både på lång (1-30 år) och kort sikt (12 timmar - 5 dagar). NO_2 användes som indikator för utsläpp från vägtrafiken och SO_2 som indikator för utsläpp från bostadsuppvärmning. Från BAMSE-kohorten, som rekryterade totalt 4089 nyfödda barn under perioden 1994-1996 i fyra kommuner i Stockholms län, inkluderades i denna studie 497 barn med astmabesvär och 485 kontroller utan astmabesvär vid 4 års ålder, och 198 barn med astma och 192 kontroller utan astma vid 8 år. Frågeformulär fylldes i av föräldrarna när barnen var ca 2 månader, 1 2, 4 och 8 år. Barnen bjöds också in till läkarundersökning och blodprovstagning vid 4 och 8 års ålder.

Hos vuxna visade långvarig exponering för både trafik-NO_2 och uppvärmnings-SO_2 ett samband med IL-6 nivåer i blodet. Exempelvis var 30 års exponering för trafik-NO_2 förenad med 64,5% (95% CI 6,7 till 153,8%) ökning av IL-6 i serum per 28,8 $\mu g/m^3$ (motsvarande skillnaden mellan det 5:e och 95:e percentilvärdet för exponering). Positiva men icke-signifikanta samband sågs också mellan kortvarig exponering för trafikrelaterade luftföroreningar och inflammatoriska markörer (IL-6, TNF-α). Gen-miljö interaktioner observerades för flera *IL6*- och *TNF*-SNPar (SNP = single nucleotide polymorphism, dvs naturligt förekommande genetiska varianter) i förhållande till inflammationsmarkörer i blod. Exempelvis interagerade exponering för trafik-NO_2 sista året med *IL6*-174G/C på så sätt att luftföroreningseffekten på IL-6 nivåer ökade med varje ytterligare *IL6*-174C-allel, och exponering för uppvärmnings-SO_2 sista året var förenat med högre TNF-α nivåer hos *TNF*-308AA homozygoter men inte hos personer som var bärare av -308G. Även kortvariga exponeringar för luftföroreningar interagerade med *IL6* och *TNF* i förhållande till markörnivåer. Risken för hjärtinfarkt följde väl det mönster som sågs för effekten på blodmarkörer i varje genotypgrupp.

Hos barn sågs en samverkan mellan moderns rökning under graviditet och/eller spädbarnstid och 3 *TNF*-SNPar med avseende på astmabesvär i tidig ålder. Oddskvoten för att utveckla tidiga astmabesvär i relation till moderns rökning var 2,4 (95% CI 1,6-3,7) hos barn med genotypen *TNF*-857CC (homozygota för normalallelen C), medan ingen tobaksrelaterad risk sågs hos barn som bar på den sällsynta -857T-allelen. En antydd interaktion med moderns tidiga rökning sågs också för 3 *GSTP1*-SNPar i förhållande till tidiga men efter 2 år övergående astmabesvär. SNPar i generna *TNS1*,

ADAM19, *THSD4* och *ADCY2* som identifierats i genomwide-analyser av lungfunktion hos vuxna visade samband med lungfunktion också hos barn; *DAAM2* rs2395730 visade antydd interaktion med aktuell exponering för tobaksrök vid 8 års ålder.

Sammanfattningsvis visar resultaten att luftföroreningar modifierar nivåerna av inflammatoriska markörer i blod, och att denna effekt kan ändras av polymorfier i gener kopplade till inflammation, med påverkan både på blodmarkörer och hjärtinfarktriskrisk. Polymorfier i gener som kodar för inflammationsmarkörer (*TNF*) och antioxidativa enzymer (*GSTP1*) verkar också kunna modifiera effekten av tidig exponering för tobaksrök på astmabesvär i tidig barndom. Ett flertal genvarianterer av vikt för lungfunktion hos vuxna verkar även påverka lungfunktionen hos barn.

8 ACKNOWLEDGEMENTS

Completing this thesis work has been made with help and contribution of many people and I am happy to express my gratitude to them.

I would like to thank all the participants of our studies: senior Stockholm citizens from the SHEEP study, as well as children (now teenagers) and their parents from the BAMSE cohort. Thanks to your interest and participation these studies were realized.

Fredrik Nyberg, my main supervisor, thank you for sharing your deep knowledge and experience in Epidemiology, always giving me excellent guidance in this wide interdisciplinary research, for your patience and support, and for encouraging me to learn Swedish.

Erik Melén, my co-supervisor, thank you for your optimism and encouragement, your fascination with science and Genetic Epidemiology along with such important clinical work. Thank you for great explanations of most complicated things, and for seeing positive sides in every situation.

Göran Pershagen, my co-supervisor, thank you for giving me the opportunity to become a PhD student in your research group. Through all the years you have been like a captain on a ship – always calm, diplomatic and seeing the goal. Thank you for your invaluable guidance in research, your support and encouragement, and for your fantastic understanding and interest in Belarus.

Lars Alfredsson, my mentor, thank you for our discussions, for your help in identifying priorities and always positive attitude!

The SHEEP study team, most of all *Ulf de Faire* and *Karin Leander*, thank you for your expert knowledge in cardiovascular field and great discussions. It was a pleasure to work with you!

The BAMSE team, especially *Magnus Wickman*, *Inger Kull*, *Eva Hallner*, *André Lauber*, thank you for your joyful and positive attitude, keeping everyone interested and involved, both researchers and BAMSE participants.

Marike Boezen and *Salome Scholtens*, thank you so much for the wonderful time in Groningen and for your excellent guidance in genetic epidemiological research!

It was always a pleasure to collaborate with my co-authors. *Tom Bellander*, *Petter Ljungman* and *Mats Rosenlund*, thank you for your expertise in air pollution research, for your attention and help! *Juha Kere* and *Cecilia Lindgren*, thank you for making genetics interesting and exciting! *Jenny Hallberg* and *Anna Bergström*, thank you for your valuable advices regarding research of lung function in children!

Eva Undsén, thank you for your kindness and help!

Many thanks to the bright statistics experts, *Tomas Lind*, *Nicola Orsini* and *Gunnel Gråbergs*! Thank you for always finding time for me, and for your kind help with databases and programming!

Thanks to my colleagues for the inspiring team-spirit at the Unit of Environmental Epidemiology! *Helen*, my roommate, thank you for being my guide through all the stages of "PhD process", for your encouragement and humor, and for our exciting discussions about Swedish and Belarusian cultures! *Johanna*, thank you for your kindness and optimism, and for our talks about everything from science to human nature! *Lotta*, thank you for your friendship and for showing me beautiful places in Stockholm! *Anna Bessö*, when I just came to IMM, you were completing your PhD education, but you were always finding time for me and explaining me everything, thank you for your kind help! *Jenny*, thank you for your always positive attitude! *Misiu*, thank you for combining artistic ways and scientific seriousness and proving that research can be fun! *Jessica* and *Erica*, thank you for being the early birds to start working day, with so much positive energy to cheer us up! *Annika* and *Anna G.*, thank you for joining the "genetic research forces" of our Unit and making this research much more exciting! *Andrei*, thank you for bringing here the feeling of Belarus and for your enthusiasm! And many thanks to *Gösta*, *Anna S.*, *Olena, Maria-Pia,* and colleagues from other Units for the unforgettable time at IMM!

I would like to express my gratitude to many brilliant teachers whose student I had an honour to be through the years: *Galina Antonovna Prokopovich*, *Ekaterina Nikolayevna Kislenko*, *Rostislav Antonovich Lasskiy*, *Svetlana Petrovna Leschinskaya*, *Edward Bellinger*, *Ckees van Oijen*, and many others. I am very grateful to late *Liudmila Grigorievna Bortkevich*, my greatest teacher of Immunology, for her kindness and highest professionalism and enthusiasm in science. My dear teacher of Biology, *Tamara Yakovlevna Moroz*, thank you so much for believing in me, for giving me both knowledge and fascination with this wonderful science.

My dear friends, *Olya*, *Éva-Terézia*, *Slava*, *Asel*, *Solomiya*, *Tanya*, *Joanna*, *Luba*, *Cristina*, *Zoltán*, *Tove*, *Levent* and *Zhanna*, thank you for being in my life!

Dear *Andrey*, thank you so much for being my dear friend and teacher!

Dear *Liudmila Ivanovna*, thank you for your kind help and optimism!

Dear *Valentina Dmitriyevna*, I am so grateful for your kindness and wisdom!

Dear *Kenneth*, thank you so much for your care and support, and for always believing in me!

My dear brother *Siarhei*, thank you for your vast encyclopaedic knowledge and scientific curiosity, for guiding me in life. In many ways thanks to you I have chosen this path. My dear little nephews, *Yury*, *Stanislau* and *Ignat*, thank you for bringing so much joy!

Many thanks to my dear relatives for their support!
My dear grandmother *Alexandra*, thank you so much for your invincible sense of humor and joy of life!

I will always be grateful to my dear parents, *Lidziya* and *Dzmitry*.
My dearest mama, you have always been an inspiring example for me. Thank you so much for your love and care! *Матулечка, вялікі дзякуй!*

9 REFERENCES

1. Brunekreef B, Holgate ST. Air pollution and health. *Lancet* 2002;360(9341):1233-42.
2. Brook RD, Franklin B, Cascio W, Hong Y, Howard G, Lipsett M, et al. Air pollution and cardiovascular disease: a statement for healthcare professionals from the Expert Panel on Population and Prevention Science of the American Heart Association. *Circulation* 2004;109(21):2655-71.
3. WHO. *Air quality guidelines: Global update 2005 - Particulate matter, ozone, nitrogen dioxide and sulfur dioxide*. Copenhagen: World Health Organization, 2006.
4. Routledge HC, Ayres JG. Air pollution and the heart. *Occup Med (Lond)* 2005;55(6):439-47.
5. Brook RD. Cardiovascular effects of air pollution. *Clin Sci (Lond)* 2008;115(6):175-87.
6. Bhalla DK, Hirata F, Rishi AK, Gairola CG. Cigarette smoke, inflammation, and lung injury: a mechanistic perspective. *J Toxicol Environ Health B Crit Rev* 2009;12(1):45-64.
7. Prescott SL. Effects of early cigarette smoke exposure on early immune development and respiratory disease. *Paediatr Respir Rev* 2008;9(1):3-9; quiz 10.
8. Kvalvik LG, Skjaerven R, Haug K. Smoking during pregnancy from 1999 to 2004: a study from the Medical Birth Registry of Norway. *Acta Obstet Gynecol Scand* 2008;87(3):280-5.
9. Egebjerg Jensen K, Jensen A, Nohr B, Kruger Kjaer S. Do pregnant women still smoke? A study of smoking patterns among 261,029 primiparous women in Denmark 1997-2005. *Acta Obstet Gynecol Scand* 2008;87(7):760-7.
10. Lannero E, Wickman M, Pershagen G, Nordvall L. Maternal smoking during pregnancy increases the risk of recurrent wheezing during the first years of life (BAMSE). *Respir Res* 2006;7:3.
11. WHO. *WHO report on the global tobacco epidemic, 2009: Implementing smoke-free environments*. Geneva: World Health Organization, 2009.
12. Paul WE, editor. *Fundamental immunology*. 5th ed. Philadelphia: Lippincott Williams & Wilkins, 2003.
13. Gabay C. Interleukin-6 and chronic inflammation. *Arthritis Res Ther* 2006;8 Suppl 2:S3.
14. Libby P. Inflammation in atherosclerosis. *Nature* 2002;420(6917):868-74.
15. Willerson JT, Ridker PM. Inflammation as a cardiovascular risk factor. *Circulation* 2004;109(21 Suppl 1):II2-10.
16. Mills NL, Donaldson K, Hadoke PW, Boon NA, MacNee W, Cassee FR, et al. Adverse cardiovascular effects of air pollution. *Nat Clin Pract Cardiovasc Med* 2009;6(1):36-44.
17. Yudkin JS, Kumari M, Humphries SE, Mohamed-Ali V. Inflammation, obesity, stress and coronary heart disease: is interleukin-6 the link? *Atherosclerosis* 2000;148(2):209-14.
18. Ghio AJ, Kim C, Devlin RB. Concentrated ambient air particles induce mild pulmonary inflammation in healthy human volunteers. *Am J Respir Crit Care Med* 2000;162(3 Pt 1):981-8.
19. Ghio AJ, Hall A, Bassett MA, Cascio WE, Devlin RB. Exposure to concentrated ambient air particles alters hematologic indices in humans. *Inhal Toxicol* 2003;15(14):1465-78.
20. Mills NL, Tornqvist H, Gonzalez MC, Vink E, Robinson SD, Soderberg S, et al. Ischemic and thrombotic effects of dilute diesel-exhaust inhalation in men with coronary heart disease. *N Engl J Med* 2007;357(11):1075-82.
21. Ruckerl R, Greven S, Ljungman P, Aalto P, Antoniades C, Bellander T, et al. Air pollution and inflammation (interleukin-6, C-reactive protein, fibrinogen) in myocardial infarction survivors. *Environ Health Perspect* 2007;115(7):1072-80.
22. Seaton A, Soutar A, Crawford V, Elton R, McNerlan S, Cherrie J, et al. Particulate air pollution and the blood. *Thorax* 1999;54(11):1027-32.

23. Pietinalho A, Pelkonen A, Rytila P. Linkage between smoking and asthma. *Allergy* 2009;64(12):1722-7.
24. Noakes PS, Thomas R, Lane C, Mori TA, Barden AE, Devadason SG, et al. Association of maternal smoking with increased infant oxidative stress at 3 months of age. *Thorax* 2007;62(8):714-7.
25. Aycicek A, Ipek A. Maternal active or passive smoking causes oxidative stress in cord blood. *Eur J Pediatr* 2008;167(1):81-5.
26. Ward C, Lewis S, Coleman T. Prevalence of maternal smoking and environmental tobacco smoke exposure during pregnancy and impact on birth weight: retrospective study using Millennium Cohort. *BMC Public Health* 2007;7:81.
27. Noakes PS, Holt PG, Prescott SL. Maternal smoking in pregnancy alters neonatal cytokine responses. *Allergy* 2003;58(10):1053-8.
28. Carlsen KH, Carlsen KC. Respiratory effects of tobacco smoking on infants and young children. *Paediatr Respir Rev* 2008;9(1):11-9; quiz 19-20.
29. Dockery DW, Pope CA, 3rd, Xu X, Spengler JD, Ware JH, Fay ME, et al. An association between air pollution and mortality in six U.S. cities. *N Engl J Med* 1993;329(24):1753-9.
30. Pope CA, 3rd, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE, et al. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. *Am J Respir Crit Care Med* 1995;151(3 Pt 1):669-74.
31. Katsouyanni K, Touloumi G, Samoli E, Gryparis A, Le Tertre A, Monopolis Y, et al. Confounding and effect modification in the short-term effects of ambient particles on total mortality: results from 29 European cities within the APHEA2 project. *Epidemiology* 2001;12(5):521-31.
32. Rosenlund M, Berglind N, Pershagen G, Hallqvist J, Jonson T, Bellander T. Long-term exposure to urban air pollution and myocardial infarction. *Epidemiology* 2006;17(4):383-90.
33. Kunzli N, Jerrett M, Mack WJ, Beckerman B, LaBree L, Gilliland F, et al. Ambient air pollution and atherosclerosis in Los Angeles. *Environ Health Perspect* 2005;113(2):201-6.
34. Kunzli N, Perez L, von Klot S, Baldassarre D, Bauer M, Basagana X, et al. Investigating air pollution and atherosclerosis in humans: concepts and outlook. *Prog Cardiovasc Dis* 2011;53(5):334-43.
35. Engstrom G, Melander O, Hedblad B. Carotid intima-media thickness, systemic inflammation, and incidence of heart failure hospitalizations. *Arterioscler Thromb Vasc Biol* 2009;29(10):1691-5.
36. Bennet AM, Prince JA, Fei GZ, Lyrenas L, Huang Y, Wiman B, et al. Interleukin-6 serum levels and genotypes influence the risk for myocardial infarction. *Atherosclerosis* 2003;171(2):359-67.
37. Bennet AM, van Maarle MC, Hallqvist J, Morgenstern R, Frostegard J, Wiman B, et al. Association of TNF-alpha serum levels and TNFA promoter polymorphisms with risk of myocardial infarction. *Atherosclerosis* 2005.
38. Leander K, Wiman B, Hallqvist J, Falk G, De Faire U. The G-455A polymorphism of the fibrinogen Bbeta-gene relates to plasma fibrinogen in male cases, but does not interact with environmental factors in causing myocardial infarction in either men or women. *J Intern Med* 2002;252(4):332-41.
39. Leander K, Wiman B, Hallqvist J, Sten-Linder M, de Faire U. PAI-1 level and the PAI-1 4G/5G polymorphism in relation to risk of non-fatal myocardial infarction: results from the Stockholm Heart Epidemiology Program (SHEEP). *Thromb Haemost* 2003;89(6):1064-71.
40. Ridker PM, Hennekens CH, Buring JE, Rifai N. C-reactive protein and other markers of inflammation in the prediction of cardiovascular disease in women. *N Engl J Med* 2000;342(12):836-43.
41. Ridker PM, Rifai N, Stampfer MJ, Hennekens CH. Plasma concentration of interleukin-6 and the risk of future myocardial infarction among apparently healthy men. *Circulation* 2000;101(15):1767-72.
42. Ljungman P, Bellander T, Nyberg F, Lampa E, Jacquemin B, Kolz M, et al. DNA variants, plasma levels and variability of interleukin-6 in myocardial infarction survivors: results from the AIRGENE study. *Thromb Res* 2009;124(1):57-64.

43. Hilt B, Qvenild T, Holme J, Svendsen K, Ulvestad B. Increase in interleukin-6 and fibrinogen after exposure to dust in tunnel construction workers. *Occup Environ Med* 2002;59(1):9-12.
44. Sullivan JH, Hubbard R, Liu SL, Shepherd K, Trenga CA, Koenig JQ, et al. A community study of the effect of particulate matter on blood measures of inflammation and thrombosis in an elderly population. *Environ Health* 2007;6:3.
45. Ruckerl R, Ibald-Mulli A, Koenig W, Schneider A, Woelke G, Cyrys J, et al. Air pollution and markers of inflammation and coagulation in patients with coronary heart disease. *Am J Respir Crit Care Med* 2006;173(4):432-41.
46. Schwartz J. Air pollution and blood markers of cardiovascular risk. *Environ Health Perspect* 2001;109 Suppl 3:405-9.
47. Pekkanen J, Brunner EJ, Anderson HR, Tiittanen P, Atkinson RW. Daily concentrations of air pollution and plasma fibrinogen in London. *Occup Environ Med* 2000;57(12):818-22.
48. Stocks J, Dezateux C. The effect of parental smoking on lung function and development during infancy. *Respirology* 2003;8(3):266-85.
49. Alati R, Al Mamun A, O'Callaghan M, Najman JM, Williams GM. In utero and postnatal maternal smoking and asthma in adolescence. *Epidemiology* 2006;17(2):138-44.
50. Stern DA, Morgan WJ, Wright AL, Guerra S, Martinez FD. Poor airway function in early infancy and lung function by age 22 years: a non-selective longitudinal cohort study. *Lancet* 2007;370(9589):758-64.
51. Wilk JB, Djousse L, Arnett DK, Rich SS, Province MA, Hunt SC, et al. Evidence for major genes influencing pulmonary function in the NHLBI family heart study. *Genet Epidemiol* 2000;19(1):81-94.
52. Schunemann HJ, Dorn J, Grant BJ, Winkelstein W, Jr., Trevisan M. Pulmonary function is a long-term predictor of mortality in the general population: 29-year follow-up of the Buffalo Health Study. *Chest* 2000;118(3):656-64.
53. Hancock DB, Eijgelsheim M, Wilk JB, Gharib SA, Loehr LR, Marciante KD, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat Genet* 2010;42(1):45-52.
54. Ober C, Vercelli D. Gene-environment interactions in human disease: nuisance or opportunity? *Trends Genet* 2011;27(3):107-15.
55. de Maat MP, Kastelein JJ, Jukema JW, Zwinderman AH, Jansen H, Groenemeier B, et al. -455G/A polymorphism of the beta-fibrinogen gene is associated with the progression of coronary atherosclerosis in symptomatic men: proposed role for an acute-phase reaction pattern of fibrinogen. REGRESS group. *Arterioscler Thromb Vasc Biol* 1998;18(2):265-71.
56. Humphries SE, Luong LA, Ogg MS, Hawe E, Miller GJ. The interleukin-6 -174 G/C promoter polymorphism is associated with risk of coronary heart disease and systolic blood pressure in healthy men. *Eur Heart J* 2001;22(24):2243-52.
57. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010;466(7307):707-13.
58. Erdmann J, Linsel-Nitschke P, Schunkert H. Genetic causes of myocardial infarction: new insights from genome-wide association studies. *Dtsch Arztebl Int* 2010;107(40):694-9.
59. Ober C, Hoffjan S. Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun* 2006;7(2):95-100.
60. Holloway JW, Yang IA, Holgate ST. Genetics of allergic disease. *J Allergy Clin Immunol* 2010;125(2 Suppl 2):S81-94.
61. Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, Obeidat M, et al. Genome-wide association study identifies five loci associated with lung function. *Nat Genet* 2010;42(1):36-44.
62. Ljungman P, Bellander T, Schneider A, Breitner S, Forastiere F, Hampel R, et al. Modification of the interleukin-6 response to air pollution by interleukin-6 and fibrinogen polymorphisms. *Environ Health Perspect* 2009;117(9):1373-9.
63. Zanobetti A, Baccarelli A, Schwartz J. Gene-air pollution interaction and cardiovascular disease: a review. *Prog Cardiovasc Dis* 2011;53(5):344-52.

64. Bruunsgaard H, Christiansen L, Pedersen AN, Schroll M, Jorgensen T, Pedersen BK. The IL-6 -174G>C polymorphism is associated with cardiovascular diseases and mortality in 80-year-old humans. *Exp Gerontol* 2004;39(2):255-61.
65. Rosner SA, Ridker PM, Zee RY, Cook NR. Interaction between inflammation-related gene polymorphisms and cigarette smoking on the risk of myocardial infarction in the Physician's Health Study. *Hum Genet* 2005;118(2):287-94.
66. Polonikov AV, Ivanov VP, Solodilova MA. Genetic variation of genes for xenobiotic-metabolizing enzymes and risk of bronchial asthma: the importance of gene-gene and gene-environment interactions for disease susceptibility. *J Hum Genet* 2009;54(8):440-9.
67. Gilliland FD, Li YF, Dubeau L, Berhane K, Avol E, McConnell R, et al. Effects of glutathione S-transferase M1, maternal smoking during pregnancy, and environmental tobacco smoke on asthma and wheezing in children. *Am J Respir Crit Care Med* 2002;166(4):457-63.
68. Bouzigon E, Corda E, Aschard H, Dizier MH, Boland A, Bousquet J, et al. Effect of 17q21 variants and smoking exposure in early-onset asthma. *N Engl J Med* 2008;359(19):1985-94.
69. Kabesch M, Hoefler C, Carr D, Leupold W, Weiland SK, von Mutius E. Glutathione S transferase deficiency and passive smoking increase childhood asthma. *Thorax* 2004;59(7):569-73.
70. Li YF, Gauderman WJ, Conti DV, Lin PC, Avol E, Gilliland FD. Glutathione S-transferase P1, maternal smoking, and asthma in children: a haplotype-based analysis. *Environ Health Perspect* 2008;116(3):409-15.
71. Wenten M, Berhane K, Rappaport EB, Avol E, Tsai WW, Gauderman WJ, et al. TNF-308 modifies the effect of second-hand smoke on respiratory illness-related school absences. *Am J Respir Crit Care Med* 2005;172(12):1563-8.
72. Wu H, Romieu I, Sienra-Monge JJ, del Rio-Navarro BE, Anderson DM, Dunn EW, et al. Parental smoking modifies the relation between genetic variation in tumor necrosis factor-alpha (TNF) and childhood asthma. *Environ Health Perspect* 2007;115(4):616-22.
73. Lee YL, Lin YC, Lee YC, Wang JY, Hsiue TR, Guo YL. Glutathione S-transferase P1 gene polymorphism and air pollution as interactive risk factors for childhood asthma. *Clin Exp Allergy* 2004;34(11):1707-13.
74. Wang C, Salam MT, Islam T, Wenten M, Gauderman WJ, Gilliland FD. Effects of in utero and childhood tobacco smoke exposure and beta2-adrenergic receptor genotype on childhood asthma and wheezing. *Pediatrics* 2008;122(1):e107-14.
75. Li YF, Gauderman WJ, Avol E, Dubeau L, Gilliland FD. Associations of tumor necrosis factor G-308A with childhood asthma and wheezing. *Am J Respir Crit Care Med* 2006;173(9):970-6.
76. Melen E, Nyberg F, Lindgren CM, Berglind N, Zucchelli M, Nordling E, et al. Interactions between glutathione S-transferase P1, tumor necrosis factor, and traffic-related air pollution for development of childhood allergic disease. *Environ Health Perspect* 2008;116(8):1077-84.
77. Reijmerink NE, Kerkhof M, Koppelman GH, Gerritsen J, de Jongste JC, Smit HA, et al. Smoke exposure interacts with ADAM33 polymorphisms in the development of lung function and hyperresponsiveness. *Allergy* 2009;64(6):898-904.
78. Reuterwall C, Hallqvist J, Ahlbom A, De Faire U, Diderichsen F, Hogstedt C, et al. Higher relative, but lower absolute risks of myocardial infarction in women than in men: analysis of some major risk factors in the SHEEP study. The SHEEP Study Group. *J Intern Med* 1999;246(2):161-74.
79. Wickman M, Kull I, Pershagen G, Nordvall SL. The BAMSE project: presentation of a prospective longitudinal birth cohort study. *Pediatr Allergy Immunol* 2002;13 Suppl 15:11-3.
80. Bellander T, Berglind N, Gustavsson P, Jonson T, Nyberg F, Pershagen G, et al. Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. *Environ Health Perspect* 2001;109(6):633-9.

81. Levy D, Lumley T, Sheppard L, Kaufman J, Checkoway H. Referent selection in case-crossover analyses of acute health effects of air pollution. *Epidemiology* 2001;12(2):186-92.
82. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med* 2010;363(13):1211-21.
83. Vermylen C, De Vreker RA, Verstraete M. A rapid enzymatic method for assay of fibrinogen fibrin polymerization time (FPT test). *Clin Chim Acta* 1963;8:418-24.
84. Melen E, Kere J, Pershagen G, Svartengren M, Wickman M. Influence of male sex and parental allergic disease on childhood wheezing: role of interactions. *Clin Exp Allergy* 2004;34(6):839-44.
85. Hallberg J, Anderson M, Wickman M, Svartengren M. Factors in infancy and childhood related to reduced lung function in asthmatic children: a birth cohort study (BAMSE). *Pediatr Pulmonol* 2010;45(4):341-8.
86. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. *Eur Respir J* 2005;26(2):319-38.
87. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81(3):559-75.
88. Yanbaeva DG, Dentener MA, Creutzberg EC, Wesseling G, Wouters EF. Systemic effects of smoking. *Chest* 2007;131(5):1557-66.
89. Rodondi N, Marques-Vidal P, Butler J, Sutton-Tyrrell K, Cornuz J, Satterfield S, et al. Markers of atherosclerosis and inflammation for prediction of coronary heart disease in older adults. *Am J Epidemiol* 2010;171(5):540-9.
90. Basso F, Lowe GD, Rumley A, McMahon AD, Humphries SE. Interleukin-6 -174G>C polymorphism and risk of coronary heart disease in West of Scotland coronary prevention study (WOSCOPS). *Arterioscler Thromb Vasc Biol* 2002;22(4):599-604.
91. Jenny NS, Tracy RP, Ogg MS, Luong le A, Kuller LH, Arnold AM, et al. In the elderly, interleukin-6 plasma levels and the -174G>C polymorphism are associated with the development of cardiovascular disease. *Arterioscler Thromb Vasc Biol* 2002;22(12):2066-71.
92. Kelberman D, Hawe E, Luong LA, Mohamed-Ali V, Lundman P, Tornvall P, et al. Effect of Interleukin-6 promoter polymorphisms in survivors of myocardial infarction and matched controls in the North and South of Europe. The HIFMECH Study. *Thromb Haemost* 2004;92(5):1122-8.
93. Nauck M, Winkelmann BR, Hoffmann MM, Bohm BO, Wieland H, Marz W. The interleukin-6 G(-174)C promoter polymorphism in the LURIC cohort: no association with plasma interleukin-6, coronary artery disease, and myocardial infarction. *J Mol Med* 2002;80(8):507-13.
94. Padovani JC, Pazin-Filho A, Simoes MV, Marin-Neto JA, Zago MA, Franco RF. Gene polymorphisms in the TNF locus and the risk of myocardial infarction. *Thromb Res* 2000;100(4):263-9.
95. Keso T, Perola M, Laippala P, Ilveskoski E, Kunnas TA, Mikkelsson J, et al. Polymorphisms within the tumor necrosis factor locus and prevalence of coronary artery disease in middle-aged men. *Atherosclerosis* 2001;154(3):691-7.
96. Koch W, Kastrati A, Bottiger C, Mehilli J, von Beckerath N, Schomig A. Interleukin-10 and tumor necrosis factor gene polymorphisms and risk of coronary artery disease and myocardial infarction. *Atherosclerosis* 2001;159(1):137-44.
97. Zhang HF, Xie SL, Wang JF, Chen YX, Wang Y, Huang TC. Tumor necrosis factor-alpha G-308A gene polymorphism and coronary heart disease susceptibility: An updated meta-analysis. *Thromb Res* 2011.
98. Wilson AG, Symons JA, McDowell TL, McDevitt HO, Duff GW. Effects of a polymorphism in the human tumor necrosis factor alpha promoter on transcriptional activation. *Proc Natl Acad Sci U S A* 1997;94(7):3195-9.

99. Schuett H, Luchtefeld M, Grothusen C, Grote K, Schieffer B. How much is too much? Interleukin-6 and its signalling in atherosclerosis. *Thromb Haemost* 2009;102(2):215-22.
100. Ober C, Thompson EE. Rethinking genetic models of asthma: the role of environmental modifiers. *Curr Opin Immunol* 2005;17(6):670-8.
101. Spiteri MA, Bianco A, Strange RC, Fryer AA. Polymorphisms at the glutathione S-transferase, GSTP1 locus: a novel mechanism for susceptibility and development of atopic airway inflammation. *Allergy* 2000;55 Suppl 61:15-20.
102. Aynacioglu AS, Nacak M, Filiz A, Ekinci E, Roots I. Protective role of glutathione S-transferase P1 (GSTP1) Val105Val genotype in patients with bronchial asthma. *Br J Clin Pharmacol* 2004;57(2):213-7.
103. Nickel R, Haider A, Sengler C, Lau S, Niggemann B, Deichmann KA, et al. Association study of Glutathione S-transferase P1 (GSTP1) with asthma and bronchial hyper-responsiveness in two German pediatric populations. *Pediatr Allergy Immunol* 2005;16(6):539-41.
104. Dijkstra A, Postma DS, Noordhoek JA, Lodewijk ME, Kauffman HF, ten Hacken NH, et al. Expression of ADAMs ("a disintegrin and metalloprotease") in the human lung. *Virchows Arch* 2009;454(4):441-9.
105. Chen H, Herndon ME, Lawler J. The cell biology of thrombospondin-1. *Matrix Biol* 2000;19(7):597-614.
106. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010;467(7317):832-8.
107. Warburton D, Bellusci S, De Langhe S, Del Moral PM, Fleury V, Mailleux A, et al. Molecular mechanisms of early lung specification and branching morphogenesis. *Pediatr Res* 2005;57(5 Pt 2):26R-37R.
108. Watkins DN, Berman DM, Burkholder SG, Wang B, Beachy PA, Baylin SB. Hedgehog signalling within airway epithelial progenitors and in small-cell lung cancer. *Nature* 2003;422(6929):313-7.
109. Wilk JB, Chen TH, Gottlieb DJ, Walter RE, Nagle MW, Brandler BJ, et al. A genome-wide association study of pulmonary function measures in the Framingham Heart Study. *PLoS Genet* 2009;5(3):e1000429.
110. Pillai SG, Kong X, Edwards LD, Cho MH, Anderson WH, Coxson HO, et al. Loci identified by genome-wide association studies influence different disease-related phenotypes in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2010;182(12):1498-505.
111. Uhl GR, Liu QR, Drgon T, Johnson C, Walther D, Rose JE, et al. Molecular genetics of successful smoking cessation: convergent genome-wide association study results. *Arch Gen Psychiatry* 2008;65(6):683-93.
112. Cho MH, Boutaoui N, Klanderman BJ, Sylvia JS, Ziniti JP, Hersh CP, et al. Variants in FAM13A are associated with chronic obstructive pulmonary disease. *Nat Genet* 2010;42(3):200-2.
113. Bjerg A, Hedman L, Perzanowski M, Lundback B, Ronmark E. A strong synergism of low birth weight and prenatal smoking on asthma in schoolchildren. *Pediatrics* 2011;127(4):e905-12.
114. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature* 2009;461(7265):747-53.