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

GENETIC STUDIES ON CHILDHOOD ASTHMA AND ALLERGY – ROLE OF INTERACTIONS

Erik Melén

Karolinska Institutet

Stockholm 2006

Published and printed by
REPROPRINT AB
Stockholm 2006
www.reproprint.se
Gävlegatan 12 B SE-100 31 Stockholm, Sweden

ISBN 91-7140-686-7

Heredity deals the cards; environment plays the hand
CL Brewer

För att nå kärnan måste man knäcka skalet
Gammalt ordspråk

ABSTRACT

The occurrence of asthma and allergic diseases is influenced by inherited and environmental factors, and symptoms of asthma and allergy usually begin in early childhood. The overall aim with this thesis was to study the role of genetic factors for the development of childhood asthma and allergy, and to evaluate potential interaction between genetic and environmental factors.

Using the BAMSE birth cohort study, children with wheezing episodes up to the age of four were classified into the following groups: transient wheezing (n=266, 8%), persistent wheezing (n=319, 9%) and late-onset wheezing (n=195, 6%). Children with persistent and late onset wheezing had the highest occurrence of sensitisation to inhalant allergens (23% and 30%, respectively), whereas lower mean peak expiratory flow values were seen in children with transient and persistent wheezing (mean difference –8.9 and –8.5 l/min, respectively). Both maternal and paternal allergic disease were of importance for all wheezing outcomes in the children, but the influence of parental allergic disease on the risk of persistent wheezing seemed to be more pronounced in boys than in girls.

For the genetic analyses, around 500 children with asthma symptoms up to four years and 500 controls were selected from the BAMSE study. Single nucleotide polymorphisms (SNP) and their corresponding haplotypes in six candidate genes for asthma and allergy were analysed and their associations with various phenotypes were evaluated. Variations in the *IL9R* gene seemed to influence the susceptibility to both wheezing and sensitisation, predominantly in boys. No overall effect of the *IL4RA* SNPs was observed and only weak associations to wheezing and sensitisation were indicated when haplotypes were considered. Variants in the *ADRB2* gene showed no overall association to any of the outcomes, whereas the *TNF-α* -308 SNP seemed to affect the risk of sensitisation at the age of four. Ala114Val was the only SNP in the *GSTP1* gene that showed any association (particularly to asthma). For the *GPRA* association analyses, asthma and allergic sensitisation were used as major outcomes and the study was designed to evaluate the role of certain haplotypes on these study subjects both from BAMSE and a multinational European project (PARSIFAL). Both risk haplotypes (H5/H6) and non-risk haplotypes (H1/H3) could be identified, and these haplotypes seemed to predominantly influence the risk of sensitisation, but also asthma and allergic rhinoconjunctivitis.

Interaction analyses between the *IL9R* and *IL4RA* genes showed that the effect of *IL4RA* SNPs on wheezing up to the age of four was modified by SNPs in the *IL9R* gene. Combinations of the *IL4RA* Gln576Arg variant and an intron *IL9R* variant seemed to influence the risk of wheezing particularly, and both risk and non-risk combinations were observed.

Air pollution from road traffic in the study area was evaluated as nitrogen oxides (traffic-NO_x) and inhalable particulate matter (traffic- PM_{10}) using emission databases and dispersion modelling. Individual exposure levels during the first year of life were estimated through geocoding of the children's home addresses. Significant gene-environment interaction effects were suggested between SNPs in the *GSTP1* gene and exposure to traffic-NO_x during the first year of life with regard to allergic sensitisation at 4 years. Heterozygous *GSTP1* carriers seemed to have the most pronounced risk of disease and this pattern was seen for all *GSTP1* SNPs tested. Similar interaction was seen for exposure to traffic- PM_{10}.

In summary, we have shown that parental allergic disease is important for development of wheezing up to the age of four, but the hereditary influence seemed to be more pronounced in boys than in girls. Variants in several of the analyzed genes were associated with symptoms of asthma and allergic sensitisation. The association between these genetic variants and allergic diseases are likely to be influenced by other genetic variants, here exemplified by gene-gene interaction between *IL4RA* and *IL9R* variants, and environmental factors, here exemplified by gene-environment interaction between *GSTP1* variants and exposure to traffic-NO_x.

LIST OF PUBLICATIONS

I. **Melén E**, Kere J, Pershagen G, Svartengren S, Wickman M.
Influence of male sex and parental allergic disease on childhood wheezing: role of interactions.
Clin Exp Allergy 2004; 34: 839-844

II. **Melén E***, Gullstén H*, Zucchelli M, Lindstedt A, Nyberg F, Wickman M, Pershagen G, Kere J.
Sex specific protective effects of interleukin-9 receptor haplotypes on childhood wheezing and sensitisation.
J Med Genet 2004; 41:e123

III. **Melén E***, Bruce S*, Doekes G, Kabesch M, Laitinen T, Lauener R, Lindgren CM, Riedler J, Scheynius A, van Hage M, Kere J, Pershagen G, Wickman M, Nyberg F and the PARSIFAL Genetics Study Group.
Haplotypes of G Protein-coupled Receptor 154 are associated with childhood allergy and asthma.
Am J Respir Crit Care Med 2005;171:1089-95

IV. **Melén E**, Umerkajeff S, Nyberg F, Zucchelli M, Gullsten H, Lindstedt A, Wickman M, Pershagen G, Kere J.
Interaction between polymorphisms in the *IL4RA* and *IL9R* genes in childhood wheezing: Evidence from a prospective birth cohort.
Submitted

V. **Melén E**, Lindgren CM, Berglind N, Zucchelli M, Nordling E, Lindstedt A, Mäkelä V, Morgenstern R, Nyberg F, Kere J, Bellander T, Wickman M, Pershagen G.
Interaction between variants in asthma-susceptibility genes and long term exposure to air pollutants.
Manuscript

* Authors have contributed equally to the study.

CONTENTS

LIST OF ABBREVIATIONS

ADRB2	Beta-2-adrenergic receptor
Ala	Alanine
Arg	Arginine
BAMSE	Children, Allergy, Milieu, Stockholm, Epidemiological survey
CEPH	Centre d' Etude du Polymorphisme Humain
CI	Confidence interval
DNA	Deoxyribonucleic acid
EM	Expectation-maximization
ETS	Environmental tobacco smoke
FDR	False discovery rate
Gln	Glutamine
GPRA	G-Protein coupled receptor for asthma
GSTP1	Glutathione S-transferase P1
HWE	Hardy Weinberg equilibrium
IgE	Immunoglobulin E
IL	Interleukin
IL4RA	Interleukin-4 receptor alpha
IL9R	Interleukin-9 receptor
ISAAC	International Study of Asthma and Allergies in Childhood
Ile	Isoleucine
LD	Linkage disequilibrium
LTA	Lymphotoxin alpha
MALDI-TOF	Matrix-assisted laser desorption/ionisation-time of flight
mRNA	Messenger ribonucleic acid
NCBI	National Center for Biotechnology Information
NpS	Neuro Peptide S
NO_x	Nitrogen oxides
PARSIFAL	Prevention of Allergy – Risk factors for Sensitisation In children related to Farming and Anthroposophic Lifestyle
OR	Odds ratio
PCR	Polymerase chain reaction
PEF	Peak expiratory flow
PM_{10}	Particle matter (≤10 µm size)
RSV	Respiratory syncytial virus
SNP	Single nucleotide polymorphism
SO_2	Sulphur dioxide
TNF-α	Tumor necrosis factor alpha
UTR	Untranslated region
Val	Valine

BACKGROUND

CHILDHOOD WHEEZING, ASTHMA AND ATOPY

Definition of asthma

Asthmatic disease is associated with a number of intermediate phenotypes, such as early childhood wheezing induced by viral infections, asthmatic symptoms induced by exercise or inhalation of cold air, and allergic asthma. The common feature is a chronic inflammation of the airways that involves a number of inflammatory cells and cellular mechanisms. This inflammation is usually accompanied by airflow limitation as a result of mucus hypersecretion and bronchoconstriction, which may cause symptoms like wheeze, breathlessness, chest tightness or cough.[1] Asthma is diagnosed clinically on the basis of these symptoms and there is no gold standard definition, although many attempts have been made to define asthma in terms of its impact on lung function (e.g., airflow limitation, its reversibility and airway hyperresponsiveness). In children, recurrent wheezing is considered to be the major criterion for asthma and a certain number of episodes within the last 12 months do usually cover the asthma diagnosis according to definitions.[2, 3] In young children, the asthma definition has been proposed to be replaced by wheezing, which also may include symptoms over time; transient, late onset and persistent wheezing[4], or the association with allergic sensitisation; non-atopic and atopic wheezing.[5]

Allergic disease is considered to be a major public health issue. The prevalence of asthma symptoms in children (aged 13-14 years) shows large world-wide variations, with the highest 12-month prevalences reported from the UK, Australia and New Zealand (about 25-30%), and the lowest prevalences from Eastern European countries, India, and Ethiopia (about 2-10%).[6] The prevalence of childhood asthma using doctor's diagnosis or more stringent definitions than wheezing only is however around 6-8% in Europe with an estimated incidence rate of 0.9/100/year.[7-10] The increased prevalence of asthma symptoms during the past decades may have reached its peak according to recent studies in several European countries,[11-13] although there is evidence that the rise in prevalence actually has continued after 1988.[14]

Risk factors for early respiratory symptoms

The variations in prevalence of asthma symptoms and other allergic manifestations suggest that different risk factors may have an impact on susceptible individuals in different parts of the world. Well known factors that influence the risk for asthma in the westernized world include family history and family size, infant feeding, sex, atopy, environmental tobacco smoke and other pollutants, but also lower respiratory tract infections (Figure 1).[15] Almost 20 years ago, the so called "hygiene hypothesis" was introduced suggesting that the decrease in infectious burden during early life and changes in gut flora may have led to increased predisposition to allergy and asthma during childhood.[16, 17] Accordingly, presence of older siblings in the home and day care attendance during the first 6 months of life have been found to protect against the

development of asthma between 6-13 years of age.[18] Certain infections may account for this protection, especially viral infections other than lower respiratory tract infections (e.g., herpes virus).[19] Studies on the effects of specific environment with other exposure patterns than commonly seen today (e.g., farming lifestyle) have given support to the hygiene hypothesis with respect to the development of allergic diseases.[20, 21] From an immunologic point of view, reduced activity of T-regulatory cells, which may lead to reduced immune suppression, has been emphasized lately as a basis for the mechanisms behind the hypothesis, rather than lack of shifting of allergen-specific responses from the Th2 to the Th1 phenotype originally suggested.[22]

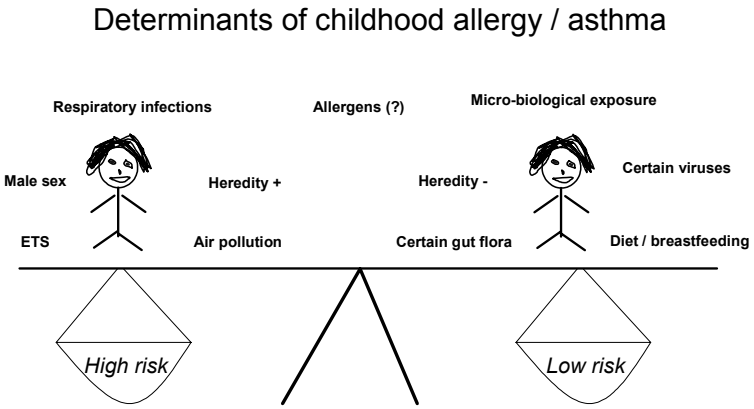

Figure 1. Genetic and environmental factors that may influence the risk of childhood allergy and asthma.

Effects of exposure to air pollutants

Traffic-related air pollutants are well known triggers for asthma exacerbations, but the results are not consistent regarding the role of air pollutants as causative agents in the initial disease development.[23, 24] Some epidemiological studies have also shown positive associations between air pollution and sensitisation to allergens in both adults and school children.[25-28] We have recently reported effects of exposure to air pollutants from the local traffic early in life and asthma-related outcomes up to the age of four in the BAMSE cohort (Nordling et al in manuscript). The results show that exposure to traffic-related air pollution during the first year of life (using NO_x and PM_{10} as indicators) was associated with an increased risk of persistent wheezing and sensitisation to inhalant allergens.

From childhood to adolescence

Considerable efforts have been made to identify childhood factors that predict asthma and atopy later in life and the cohort studies conducted over the world have contributed to a great extent in our understanding of the natural history of these conditions. One key finding from the Tucson Children's Respiratory Study is that events occurring already during the first year of life affects the total IgE levels, which in turn is a prognostic marker for later development of asthma.[4] The critical role of early life events is further underlined in that the majority of all cases with persistent asthma experience their first symptoms already in childhood.[29] Other theories suggest that important effects may start already *in utero*.[30, 31]

Factors that affect lung function in early infancy are of particular interest, as increased airway responsiveness already at 1 month of age has been shown to be associated with abnormal airway function and physician diagnosed asthma by 6 years of age.[32] Exposure to tobacco smoke is one such well known factor, and there are several studies that show negative effects of exposure already *in utero* on lung function later in life.[33-36] Wheeze in the early years is very common and may affect up to 30-40% of all children.[4] Up to 73% of the Dunedin cohort members in New Zealand reported any wheezing episode up to the age of 26 years.[37] More than 80% of children who wheeze during the first year of life and 60% of those who wheeze during the first two years of life do not wheeze after the age of 3 years.[4] However, these children have been shown to have lower levels of lung function already at birth, and also at the age of 5-7 years and even by 16 years of age.[4, 38, 39] Further, the earlier the age at onset of wheeze, the greater the risk of relapse later in life.[37] Although asthma severity seems to improve from childhood to adulthood, the majority of preteenage children with established asthma will also experience asthma symptoms around the age of 25-30.[40, 41] Complete remission of childhood asthma, usually defined as no symptoms, no medication and normal lung function, is estimated to around 10-20% after 25-30 years.[42, 43] Thus, early determinants for both wheezing susceptibility and lung function capacity seem to be of importance for the respiratory status later in life.

Atopy

Atopy can be defined as a personal and/or familial tendency to become sensitised and produce IgE antibodies in response to ordinary exposures to allergens, usually proteins.[44] Using this definition, atopy is referred to as an immunological event that leads to the production of IgE antibodies, which may (or may not) cause typical symptoms of asthma, rhinoconjunctivitis, or eczema in relation to exposure to the allergen in point. In pre-school children, up to 25% are sensitised (regardless of any symptoms) and the IgE antibody synthesis is typically directed towards food allergens in infancy followed by IgE antibody production against inhalant allergens later in childhood.[45, 46] The prevalence of sensitisation seems to peak in 20-30 year old individuals (around 35% prevalence) and declines at higher ages.[47] Atopy remains one of the strongest risk factors for the development of asthma, and may also have negative effects on lung function, independent of asthma, already at the age of 3 years.[37, 46, 48]

HUMAN GENETICS

The human genome

In 2001, the first analyses of the working draft human genome sequence were published as a joint effort by the publicly sponsored Human Genome Project and a private company, Celera Genomics.[49, 50] This happened almost 150 years after Gregor Mendel's discovery of the laws of heredity (1866) and almost 50 years since the determination of the DNA structure (1953).[51] The human genome contains approximately 30,000 protein-coding genes, and several other functional elements, such as non-protein-coding genes and DNA sequences related to chromosome dynamics.[49] In humans, the DNA is organized in 23 pairs of chromosomes (22 autosomes and one pair of sex chromosomes, XX or XY) found in the nucleus of the cell and is formed by large polymers with a linear backbone of sugar (deoxyribose) and nitrogen bases (adenine, A, cytosine, C, guanine, G, and thymine, T) attached to each sugar residue.[52] A nucleoside is a sugar with an attached base, and a nucleoside with a phosphate group attached is called a nucleotide, which is the basic repeat unit of the DNA strand and one important source of variation in the genome. The structure of the DNA strands is a double helix with two DNA strands bound together in an antiparallel way (see Figure 2).

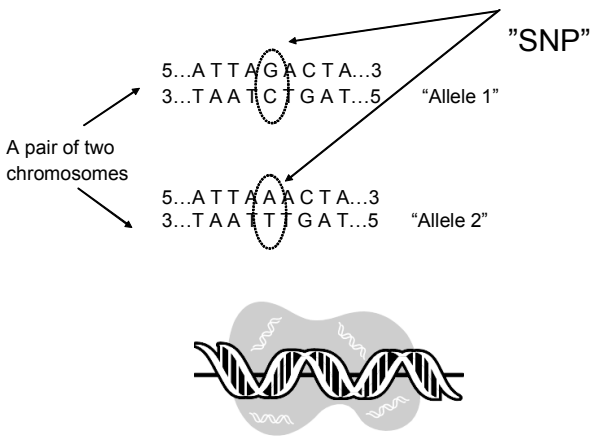

Figure 2. A schematic figure of the antiparallel DNA strands, a single nucleotide polymorphism (SNP) and the DNA double helix.

A gene is an ordered sequence of nucleotides located in a particular position (locus) on a particular chromosome that encodes a specific functional product (e.g., a protein or a RNA molecule). As we have two copies of each gene on the autosomal chromosomes deriving from the mother and father respectively, a gene may exist in two alternative forms, or alleles. An individual having two identical alleles at a particular locus is said to be homozygous, whereas an individual with two different alleles is said to be heterozygous. Polymorphism refers to the existence of two or more genetic variants at a population frequency of > 1% (< 1% are usually considered mutations), and single

nucleotide polymorphism to any variation at a single nucleotide (usually only two variants, e.g., presence of adenine, A or cytosine, C as the nitrogen base in the nucleotide). More than 9 million SNPs have been identified in the human genome and the average density is one SNP every 1,250 nucleotides, although large variations exist between genomic regions.[53, 54] Deletions of one or several nucleotides (or even the whole gene, e.g., *GSTM1* deletion) is also a form of polymorphisms and their relevance for medical genetic studies have gained attention lately.[55] According to the central dogma in genetics, messenger RNA (mRNA) is transcribed from the DNA in the nuclei of the cells, and further translated into proteins in the cytoplasmic ribosomes (Figure 3).[52] Three nucleotides in the mRNA molecule (called codons) encode each amino acid in the protein. Thus, a polymorphism in the DNA strand may lead to a different mRNA codon, which in turn could change the amino acid in point and thereby alter function of the protein (e.g., an enzyme or a receptor) that may be involved in the disease pathophysiology.

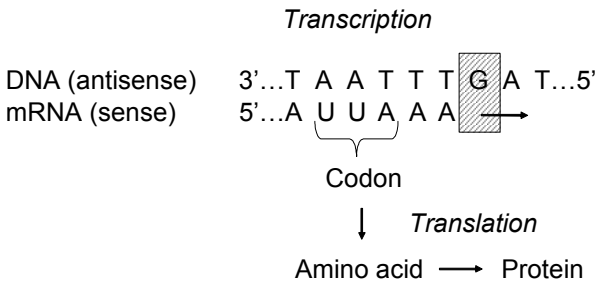

Figure 3. A schematic picture of DNA transcription and translation of mRNA into protein. Note that adenine (A) in the DNA strand is transcribed into uracile (U) in the mRNA strand.

Mode of inheritance

According to Mendel's laws of inheritance, a phenotype (e.g., disease or any other character) shows a dominant pattern of inheritance if it is manifest in the heterozygote (one copy of the disease related gene is enough) and recessive if manifest only in the homozygous carriers (two copies of the disease related gene are required).[52] Other types of genetic mechanisms include additive genetic effects: a mechanism of quantitative inheritance such that the combined effects of genetic alleles at two or more gene loci are equal to the sum of their individual effects. A multiplicative model assumes that the combined effect of two or more gene loci is equal to the product of their individual effects. Co-dominance refers to the situation in which two different

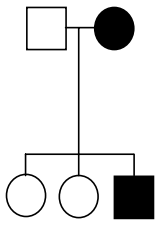

Figure 4. A small pedigree (only one family) with affected mother and son.

alleles for a trait are expressed unblended in the phenotype of heterozygous individuals. Thus, both alleles influence the phenotype, and neither dominant nor recessive inheritance pattern is seen. Type AB blood is an example of a co-dominant phenotype.

While most monogenic diseases follow Mendel's laws of inheritance (hence called Mendelian diseases), multifactorial or complex diseases do not (hence called non-Mendelian diseases). These diseases may still have a clear heritable component, but the occurrence depends on many genes on several chromosomes as well as environmental components. By studying human mutation rates, population genetics and the allelic spectra of some well-characterised monogenic disorders, it has been proposed that if the overall frequency of disease alleles is high, the frequency of some of the individual disease alleles will also be high, and vice versa.[56] This model has proven to accurately predict the allele diversity for monogenic disorders and lends support to the common disease / common variant hypothesis (CD/CV) for complex diseases. This hypothesis states that rather few alleles with relatively high frequencies are most likely to contribute to the genetic risk for common diseases, which is a fundamental assumption for the design of most association and linkage studies. This hypothesis could for instance be exemplified by the effect of the *PPARγ* variant Pro12Ala on diabetes type II (allele frequency of ~85%) and the *ApoE4* allele on Alzheimer's disease (allele frequency of ~15%) and fits well with the observed pattern of relative risks in extended families.[56-58] However, the hypothesis does not necessarily hold for all genes involved in complex diseases.[59]

The genes associated with complex diseases are usually referred to as susceptibility genes to distinguish them from causative genes. Asthma and allergic diseases belong to these complex diseases and no simple mode of inheritance that accounts for most asthma susceptible genes has been concluded despite considerable efforts.[60]

Linkage disequilibrium

Linkage disequilibrium (LD) can be described as non-independence of alleles, or association between alleles at different (but nearby) sites, and refers to the number of historical recombinations between the two sites (see review[61]). It is usually measured as r^2 or D' and both measures range from 0 (complete equilibrium) to 1 (complete disequilibrium). The correlation coefficient, r^2, is a direct statistical measure of the correlation between the sites, whereas D' is a pure indicator of historical recombination (values <1 indicates that recombination has occurred) and may be less intuitive to interpret, especially intermediate values. LD is a function of the recombination rate between the loci, population size and unexpected events such as mutations or selection mechanisms. As expected, LD is inversely related to the distance between the markers, but show also considerably variation in different genomic regions and in different populations.[62] Knowledge about historical recombination and consequent LD structure in the genome is vital for the design of genome wide linkage studies and indirect association studies in the search for new disease related genes.

Haplotype block structure

Several variations in a gene may cause amino acid changes in the ultimate protein product, and the joint effect of these amino acid changes can have a much larger influence on the function of the protein than any single amino acid change. Studying the effect of haplotypes rather than single SNPs is therefore an attractive approach in the search for disease related genes. Ideally, haplotype construction should be based on family data where you can identify alleles by descent and thereby construct the correct haplotypes. However, parental genotype information is usually lacking in case control studies, and the haplotypes have to be inferred by other methods. The key issue is to get a correct estimate of the haplotypes, that is, the combination of alleles at different markers along the same chromosome. There has been much attention in developing these inference methods in the past years.[63] Examples of different statistical approaches in different programs include Bayesian models (e.g., PHASE or Haplotyper), parsimonious models (e.g., the Clark algorithm), maximum likelihood models (e.g., HAPLO or SNPHAP), or maximum likelihood models combined with regression based approaches.[63-66] The expectation-maximization (EM) algorithm[67] is widely used and is for instance implemented in SNPHAP and several other programs. The algorithm relies on LD, HWE, relatively few missing data, normal homozygosity levels and sufficient sample size. However, even under very poor working conditions the expected error in haplotype frequency is about 5% for samples of size ≥100 and haplotypes of few SNPs length.[68]

The observed LD pattern across large genomic regions is the basis for the hypothesis of haplotype block structures in the human genome.[69, 70] In 2005, The International HapMap Consortium presented a haplotype map of the human genome containing more than 1 million SNPs genotyped in four populations, which will hopefully be of great help for the design and analyses of forthcoming genetic studies.[71] However, the use of these haplotype blocks in genetic studies on complex diseases has been questioned.[61] The major issues have been the large variation in block structure between genomic regions and populations, haplotype block breakdown when new SNPs are added to the analyses and the lack of data on the functional effects of the observed blocks. Moreover, there is no widely accepted definition of a haplotype block, although the definition by Gabriel et al using confidence bounds (95%) on D' is one of the most widely used.[70] Other definitions are based on frequency cut off levels of the observed

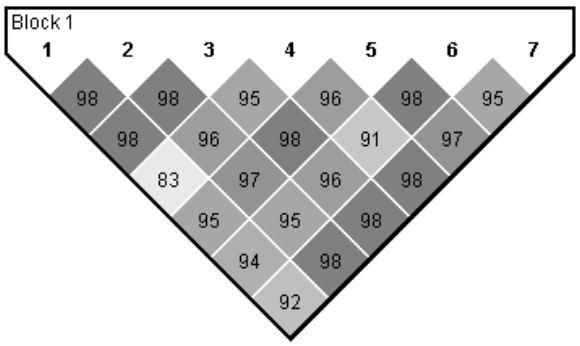

Figure 5. The 70-kb haplotype block comprised in the *GPR*A gene, as defined by Gabriel et al., using the BAMSE and PARSIFAL samples. The numbers in each box correspond to the pair-wise linkage disequilibrium coefficient, D' between respective SNPs (paper III).

haplotypes (e.g., frequency <0.01, “Four gamete rule”) or search for a "spine" of strong LD running from one marker to another (“Solid spine of LD”).[72, 73] The choice of method will consequently affect the pictured block structure, as well as the set of chosen SNPs within the region of interest.[74]

Indirect and direct genetic analyses

Depending on study design and prior knowledge in the area, genetic association analyses can either be direct or indirect. Direct analyses usually focus on variants that lead to amino acid changes in the protein, or affect RNA transcription or protein translation. One example is the *IL4RA* Q576R variant (Gln576Arg) which is associated with asthma and allergy and also show altered receptor activity dependent on the Q or R allele being present.[75] Indirect analyses rely on linkage disequilibrium between the tested variants and the causal variants, and constitute the basis of linkage studies with microsatellite markers and association studies based on haplotype tagging SNPs.[76] Genome-wide linkage studies using microsatellite markers have been the basis for the successful positional cloning of new susceptibility genes, but SNP-based linkage studies are now beginning to be reported, as the human genome sequence has been deciphered and most SNPs have been mapped.[77]

Association analyses at a single locus (e.g., a SNP) in a case control data set can be based either on the genotype frequencies or the allele frequencies. Usually, the aim is to estimate the genotype relative risk, that is, the relative risk of disease given a particular genotype compared to a reference genotype. For the genotype analyses in complex diseases, different modes of inheritance are usually tested (e.g., recessive, dominant, additive or multiplicative) if there is no á priori knowledge about inheritance mode for the specific gene or variant. Different genotype trend tests based on these inheritance modes have also been proposed in order to increase power and robustness.[78, 79] For variants with low population frequencies, too few rare homozygotes may be a problem in genotype analyses. On the other hand, information about the mode of inheritance for a specific gene may be of particular interest, and possible gene dosage effects may also give valuable information. Allele based association analysis is a commonly used alternative to genotype associations, whereby the frequency of the disease related allele is compared in cases and controls (usually by the chi2-test). When the Hardy Weinberg equilibrium holds (see “Success rates and Hardy Weinberg equilibrium” in the METHODS section) and a multiplicative mode of inheritance can be assumed, the allele based test statistic has shown to be a good approximation of the genotype relative risk.[80]

GENETICS OF ASTHMA AND ALLERGY

There is good evidence that both inherited and environmental factors influence the risk of developing allergic disease in childhood and these conditions belong to the group of complex diseases as discussed above. Studies on twins support the genetic impact of asthma in that the concordance rate of asthma is higher in monozygotic twins compared with dizygotic twins. The heritability of asthma has been suggested to be as high as 60-70%, even for asthma in pre-school children.[81, 82] The first linkage analyses for asthma susceptibility and IgE regulation genes were initiated in the early 90's and were initially promising in the search for a major gene.[83] However, the first candidate gene (*ADAM 33*) was not identified not until 2002.[84] Additional five genes (*DPP10*; chromosome 2q14, *PHF11*; chromosome 13q14, *GPRA*; chromosome 7p, *HLA-G*; chromosome 6p21 and *CYFIP2*; chromosome 5q33) have been identified by positional cloning to date, which has led to exciting new knowledge about the pathophysiology of asthma and allergy.[85-89] A recent review on asthma genetics identified 25 genes that have been associated with asthma or atopy in six or more populations and 54 genes in 2-5 populations.[90] The authors conclude that total number of genes that contribute to susceptibility to asthma or atopy may exceed 100. Segregation analyses have suggested that airway responsiveness, a typical feature of asthma, is genetically distinct from atopy.[91] Asthma and atopy are, however, clinically correlated to a high degree, and most candidate genes for asthma are also candidate genes for atopy. Whether this is simply due to the co-occurrence of these conditions or due to a common genetic pathophysiology of these traits is not fully investigated.

A short background is given below for each of the genes analysed within this thesis.

Interleukin-9 receptor (IL9R)

The *IL9R* gene is located on the pseudoautosomal region of X and Y chromosomes (Xq28 and Yq12[92, 93]) and had prior to our study been associated with asthma in two separate family-based data sets.[94, 95] One of the studies suggested a sex-specific genetic effect based on haplotype analyses of microsatellite markers, but associations between SNPs and asthma/allergy had not been reported.[95] The IL9 receptor belongs to the haematopoietin receptor superfamily[96] and is expressed on T cells, mast cells, macrophages, eosinophils and neutrophils.[97, 98] Signals from the IL9 receptor have been shown to be crucial for immunologic processes such as T cell development[99] and prevention of apoptosis.[100, 101] IL9 may also have a key role in the development of allergy, as IL9 can act directly on B lymphocytes (through IL9R) and regulate IgE synthesis.[102, 103]

Interleukin-4 receptor alpha (IL4RA)

A number of polymorphisms have been identified in the *IL4RA* gene (chromosome 16p12), several of which have been investigated in relation to asthma and atopy.[90] Although conflicting results exist regarding the effect of a particular *IL4RA* variant on asthma or allergy, *IL4RA* is considered to be an important gene for asthma susceptibility based on a number of positive findings across studies and populations.

IL4 is also one of the key cytokines in the Th2-type inflammatory response that is considered to be of particular importance for asthma and allergy.[104] Recently, both immunologic and genetic studies suggest that IL4RA may interact with other cytokines and receptors. Gene-gene interactions, the effect of one locus being altered by effects at another locus, have for instance been observed between *IL4RA* and *IL4*, as well as *IL4RA* and *IL13*, with respect to asthma (and to food allergy to some extent) [105-108], and between *IL4RA* and *IL10* with respect to RSV bronchiolitis.[109] Also, studies on gene-environment effects in the first year of life have also highlighted the *IL4RA* gene as a potential effect modifier of environmental stimuli.[110]

G-Protein coupled receptor for asthma (GPRA/GPR154)

GPR154 (alias *GPRA*) on chromosome 7p, was the fourth candidate gene for asthma-related traits identified through positional cloning.[88] The genetic evidence was supported by single nucleotide polymorphism (SNP) and haplotype associations to asthma and total IgE in three separate populations, a distinct distribution of protein isoforms between bronchial biopsies from healthy and asthmatic adults and increased expression of the *GPR154* gene in experimentally induced lung inflammation in mice. The natural GPRA agonist has recently been identified as NpS, Neuro Peptide S, through studies on murine brain tissues, and GPRA is therefore also known as NpS receptor (or PGR14, VRR1).[111, 112] GPRA and NpS are co-expressed in tissues relevant for asthma and allergy, e.g., the bronchial epithelium, and activation of GPRA with NpS results in inhibition of cell growth.[113] A specific polymorphism in the *GPRA* gene, Asn107Ile, has been associated with enhanced NpS-signalling effects.[112] These studies support the functional role GPRA might have in relation to asthma-related diseases. The association between *GPRA* variants and asthma / atopy has to date been replicated in two separate studies, including our study (paper III).[114] The Asn107Ile was not directly analyses in these studies, but from Laitinen et al, it can be concluded that it separates the observed non-risk haplotypes (H1, H3) from the others.[88] Variants in the *GPRA* gene seem, however, not to influence the risk of atopic dermatitis / eczema.[115, 116]

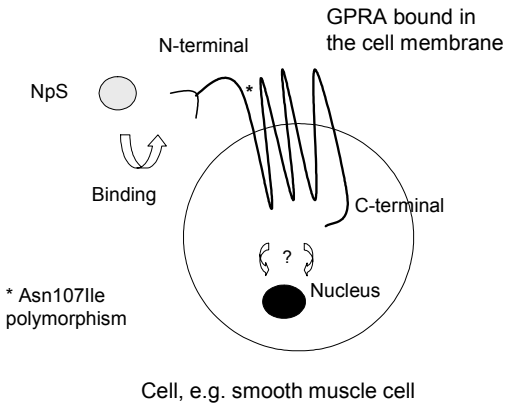

Figure 6. A schematic figure of the seven transmembrane receptor GPRA and its ligand NpS. The exact mechanism by which GPRA operate is still unknown.

ß2-adrenergic receptor (ADRB2)

The pathophysiology of asthma is partly due to a reduced function of the ß-adrenergic system, and ß-adrenergic receptor agonists (short acting or long acting) are one of the cornerstones in asthma treatment, both for children and adults. The *ADRB2* gene is located on human chromosome 5q31, where several studies have detected linkage with asthma or asthma-associated traits.[117] A number of coding SNPs in the *ADRB2* gene that may affect the receptor activity have also been identified. In a recent meta-analysis on the association between *ADRB2* polymnorphisms and asthma-related phenotypes, it was concluded that neither the Arg16Gly nor Gln27Glu polymorphisms were associated with mild asthma or bronchial hyperresponsiveness, but the Arg16Gly was rather strongly associated with nocturnal asthma and asthma severity.[118] The Arg16Gly polymorphism has also been shown to affect the long-term response to ADRB2 agonist use,[119, 120] but also the risk of having asthma in relation to smoking.[121]

Tumor necrosis factor alpha (TNF-α)

TNF-α is a well known cytokine with a wide range of pro-inflammatory effects and has been suggested to have an important role in the pathophysiology of several diseases including asthma. Although variants in the *TNF-α* gene (chromosome 6p21) have been associated with asthma and allergy in a number of studies, is still unclear to what extent these variants actually influence disease susceptibility, and the effects of *TNF-α* polymorphisms have been suggested to be dependent on polymorphisms in the *LTA*-gene.[90, 122, 123] The -308A variant in the *TNF-α* promoter region has been in particular focus because enhanced in vitro transcription and increased TNF levels in human white blood cells have been observed.[124, 125] Special attention has also been given to possible gene-environment effects between *TNF-α* variants and exposure to airway irritants, such as ozone and environmental tobacco smoke with respect to respiratory illness.[126, 127]

Glutathione S-transferase P1 (GSTP1)

Members of the glutathione S-transferase (GST) supergene family constitute an important intracellular protective system against electrophiles, oxidative stress and the formation of hazardous reactive oxygen species.[128, 129] Formation of these reactive oxygen species is a key component of airway inflammation and can be triggered by environmental stimuli such as air pollutants or viral infections. The GSTP1 enzyme is of particular interest in relation to the respiratory system, as it might provide more than 90% of the glutathione-S-transferase activity in the lung.[130] Variants in the *GSTP1* gene (chromosome 11q13) have been associated with asthma and allergy in several studies, regardless of any environmental exposure.[90] Recent studies support the presence of gene-environment interaction, or effect modification, between exposure to air pollutants and variants in the *GSTP1* gene (Ile105Val) with respect to childhood asthma and also to the variability of nasal allergic responses after challenge to diesel particles.[131, 132] Deficiency of two other GST genes, GSTM1 and GSTT1 has also been shown to influence the effect of passive smoking on the risk of childhood asthma and wheezing.[133, 134]

AIMS

The overall aim of this thesis was to study the role of hereditary and genetic factors in the development of childhood asthma and allergy, and to evaluate potential interaction effects between genetic and environmental factors.

The specific aims were:

I. To evaluate the influence of parental allergic disease on the development of wheezing and allergy

II. To investigate the influence of *IL9R* gene variants on childhood wheezing and allergy

III. To assess the impact of *GPRA* variants on childhood allergic disease, including allergic sensitisation, asthma and rhinoconjunctivitis in two different study populations

IV. To investigate the influence of *IL4RA* gene variants on childhood wheezing and allergy and to evaluate potential interaction effects between the *IL4RA* and *IL9R* variants

V. To assess interactions between exposure to ambient air pollution and variants in genes controlling the inflammatory response and antioxidative system that may affect the development of asthma and allergy in childhood

METHODS

STUDY POPULATIONS AND QUESTIONNAIRES

All papers (I-V) in the thesis are based on the BAMSE study. Besides, paper III also includes children from the European cross-sectional PARSIFAL study (see below).

BAMSE

The main purpose of the study has been to identify early risk factors for development of allergic disease in children, such as indoor climate, infant feeding and exposure to pets and tobacco smoke. Although studies on genetic factors and gene environment interaction effects were not covered in the very beginning of the BAMSE project, these issues have been addressed as the project has proceeded.

The BAMSE study is a prospective birth cohort study, conducted at the Department of Environmental and Occupational Health, Stockholm County Council and the Institute of Environmental Medicine, Karolinska Institutet. In the study area of four districts in Stockholm county (Stockholm city, Sundbyberg, Solna and Järfälla), 7,221 infants were born during the recruitment period, 1994-96. At the first visit to the Child Health Centres the families received information about the study from the attending nurse when the infant was approximately three weeks of age. However, 477 families could never be reached because the correct address was not registered. The actively excluded group included 699 families who planned to move within one year, 57 families with a seriously ill child and 331 with insufficient knowledge in the Swedish language. At the latter part of the enrolment period, another 169 children who had an older sibling already enrolled in the study were also excluded. Thirteen hundred and ninety-nine (1,399) families never answered the questionnaire or declined participation. Consequently, the final study cohort was made up of the 4,089 children (2,065 boys and 2,024 girls), which constitutes 75% of the 5,488 eligible children born during the recruitment period in the study area. Sixteen per cent of the children included have one or two parents born outside Scandinavia.

At the children's age of 2 months the parents answered a detailed questionnaire dealing with living conditions, socio-economic status and heredity for allergic diseases and environmental exposures at home. The questionnaire design and procedure have been described in detail elsewhere.[135, 136] New questionnaires were distributed to the parents when the children were 1, 2 and 4 years of age, now focusing on the child's health, especially symptoms of allergic diseases, and key environmental exposure factors, such as environmental tobacco smoke and pet contact. The response rate for the questionnaires was 96%, 92% and 91% respectively.

PARSIFAL

The cross-sectional PARSIFAL study (Prevention of Allergy – Risk factors for Sensitisation In children related to Farming and Anthroposophic Lifestyle) on the association between environmental and lifestyle factors and allergies and asthma was initiated in 2000.[21] A total of 14,893 children aged 5–13 years from five Western European countries were included. It was designed to investigate the role of different lifestyles and environmental exposures in farm children, Steiner school children (mainly from anthroposophic families) and two corresponding reference groups to identify protective factors for development of asthma and allergic disorders. In Austria, Germany, the Netherlands and Switzerland farm children were recruited from schools in rural areas known to have a high percentage of farmers; in Sweden through the Farming Registry at the National Bureau of Statistics. Children with anthroposophic lifestyle were recruited from classes in Steiner schools. The respective reference groups were recruited with similar methods from the same geographical areas. Parents completed a detailed questionnaire on allergic diseases, infectious history and environmental exposures, largely based on questions from ISAAC phase II[137], BAMSE, the ALEX study[138] and an earlier Swedish study focused on the anthroposophic lifestyle.[139]

AIR POLLUTION ASSESSMENT

A detailed description of the methods for assessing exposure to various air pollutants has been described elsewhere (Nordling et al in manuscript). Spatial distribution of air pollution from traffic in the study area was evaluated as source-specific nitrogen oxides (traffic-NO_x) and inhalable particulate matter (traffic-PM_{10}) using emission databases and dispersion modelling. Air pollution from residential heating was evaluated as sulphur dioxide (house heating- SO_2). For traffic-NO_x and heating-SO_2, emission databases were available for 1990 and 2000, whereas data bases on traffic-PM_{10} only were available for the year 2000. Monthly levels of traffic-NO_x and heating-SO_2 were calculated by interpolation between 1990 and 2000 assuming a linear change in air pollution levels between these years. For traffic-generated PM_{10}, the levels from the year 2000 were used for the whole study period. Estimated individual levels for the first 12 months of life for each child were then calculated through geocoding of the children's home addresses (standard GIS computer software in combination with a regional geographical address database). The geographical distribution of air pollution was assessed in three layers of different resolution, applied to regional/countryside area (500× 500 m), urban area (100× 100 m), and inner-city area (25× 25 m). Calibrations of the models were performed to minimize deviation when compared to available measured levels of total concentrations for the corresponding period.[140]

BLOOD SAMPLES

In the BAMSE study, 2,965 children attended a clinical investigation (73%) at the mean age of 4 years and venous blood samples were obtained from 2,614 children (64%). Serum IgE antibodies to inhalant and food allergens were analyzed with

Phadiatop® (a mixture of cat, dog, horse, birch, timothy, mugwort, *Dermatophagoides pteronyssinus* and *Cladosporium herbarum* allergens), and fx5® (a mixture of milk, egg white, soy bean, peanut, fish and wheat allergens, Phadia AB, Uppsala, Sweden). A positive result was defined as a concentration $\geq$0.35 kU_A/L. Additional analysis of specific IgE antibodies was made if the screening was positive. Children with history of wheezing were somewhat more likely to have a blood sample drawn compared with non-wheezing children, 74% vs 69% ($p<0.05$), whereas no difference was seen with regard to a number of basic characteristics (e.g., age, sex, parental allergic disease, maternal smoking, socio-economics and presence of furred pets at home).

In the PARSIFAL study, parental consent for blood sampling was obtained for 8,788 children, of these 4,854 were finally invited for blood sampling and 4,049 were able to give a blood sample. Serum IgE antibodies to inhalant and food allergens were analyzed with Phadiatop® and fx5® as described above.

LUNG FUNCTION TESTS

At the four year clinical investigation in BAMSE, lung function tests (peak expiratory flow, PEF, using the normal range Ferraris Peak Flow Meter®, Ferraris Medical Limited, UK) were performed on 2,926 (72%) children.[141] The best of the three PEF recordings were used for analysis and acceptable tests were obtained from 2,828 children. All tests were supervised by an experienced nurse.

OUTCOME DEFINITIONS

Wheezing

Since the BAMSE study is a prospective study with follow up on several occasions, our main outcomes have taken age of onset and duration of symptoms into account. Transient wheezing was defined as ≥ 3 episodes of wheezing between three months and two years of age, but no episode in the last 12 months at 4 years. Persistent wheezing was defined as ≥ 1 episode of wheezing between three months and two years of age and ≥ 1 episode in the last 12 months at 4 years, and late onset wheezing as no episode of wheezing between three months and two years of age, but ≥ 1 episode in the last 12 months at 4 years. Steroid treated wheezing at 4 years was defined as at least one episode of wheezing in the last 12 months and prescription of inhaled steroids.

Asthma

Asthma at the age of four was defined as a physician's diagnosis of asthma according to the question "Has your child ever been diagnosed with asthma by a physician?"
"Current asthma" at 4 years was defined as physician-diagnosed asthma and one or more episodes of wheezing in the last 12 months at 4 years. Allergic asthma (or atopic asthma) was derived from the combination a physician's diagnosis of asthma and sensitisation (see below).

In the PARSIFAL study, asthma was defined as a physician's diagnosis of asthma according to the question "Has your child ever been diagnosed with any of the following diseases by a physician?", which required any of the answers "*Once, or several times, for asthma AND/OR several times for spastic, obstructive or asthmatic bronchitis"*.

Allergic sensitisation

In both BAMSE and PARSIFAL, allergic sensitisation was defined as an IgE level ≥0.35 kU/L by either Phadiatop® or fx5®.

Allergic rhinoconjunctivitis

In PARSIFAL, allergic rhinoconjunctivitis was defined as a questionnaire response of having had rhinitis and conjunctivitis symptoms without concurrent cold within the last 12 months, in combination with sensitisation to inhalant allergens (Phadiatop®). The young age of the participants and the questionnaire data available in BAMSE did not allow for a comparable diagnosis.

SAMPLE SELECTION AND DNA EXTRACTION

In the BAMSE study, approximately 5 millilitres (ml) blood was drawn in an EDTA tube from most children who consent to give blood at the clinical investigation, but in some cases less than 5 ml was obtained. For the genetic analyses, 2,298 of these blood samples were available after exclusion of 69 samples because of too little blood, 81 samples due to lack of questionnaire data and 166 samples because parental consent to genetic analyses was not obtained. In a case-cohort sampling design, a sequential random sample of 709 children (357 girls, 352 boys) from the "genetics cohort" (n=2,298) was selected as a subcohort (Figure 7) until 542 children with no defined wheezing were included. These were used as random controls (282 girls, 260 boys), whereas children who fulfilled any wheezing criterion up to the age of four were identified as wheezing cases (n=167). In addition, all the other 375 children in the "genetics cohort" fulfilling the wheezing criteria were included, resulting in a total of 542 wheezing cases (214 girls, 328 boys). However, due to occasional failure in DNA extraction (n=29), the total number of DNA samples were reduced to 1,055 (530 wheezing cases and 525 controls). The randomly sampled subcohort of 709 children made it possible also to analyse different outcomes, such as sensitisation using non-sensitised children as comparison group. One hundred and ninety-five children (27.5%) in the subcohort were sensitised to either inhalant or food allergens.

The randomly selected subcohort appeared representative of the original cohort in that no significant differences were seen concerning a number of basic characteristics (e.g., ethnic background, sex, parental allergic diseases, smoking mothers and wheezing prevalence.) Children in the subcohort were however somewhat more often sensitised compared to the full cohort (27.5% vs. 24.0%, p=0.06) and a minimal difference was seen regarding age at investigation (4.31 vs. 4.29 years, p=0.01).

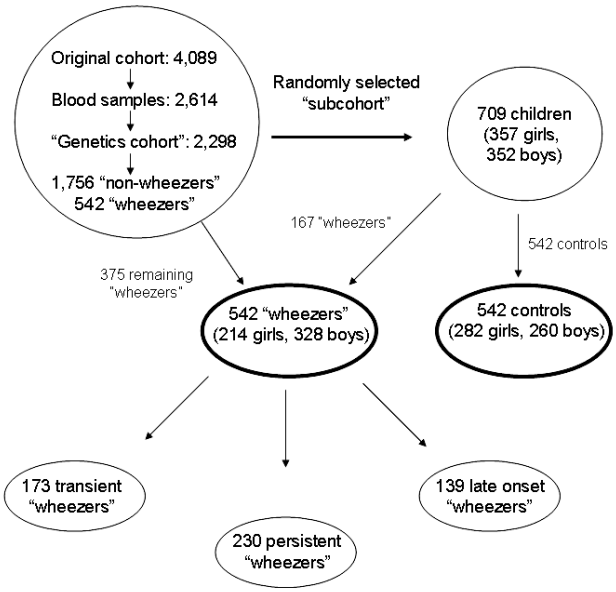

Figure 7. Selection of cases and controls using a case-cohort sampling design. A sequential random sample of 709 children from the "genetics cohort" (n=2,298) was selected as a subcohort. From this random subcohort, children with no defined wheezing were used as controls (n=542), whereas children who fulfilled any wheezing criterion up to the age of four were identified as wheezing cases (n=167). An additional 375 children were identified as wheezing cases from the overall "genetics cohort", resulting in 542 wheezing cases and 542 controls.

The blood was centrifuged whereby the blood cells were separated from the plasma. Both tubes were then frozen until later use. DNA was extracted from peripheral blood leukocytes using a standard non-enzymatic method or the PUREGENE KIT (Gentra Systems). The average DNA stock concentration in the samples was 237 ng/µl and the average total DNA yield 56.0 µg.

The PARSIFAL blood samples were collected in EDTA tubes and 3,113 children (1,579 boys, 1,534 girls) had complete questionnaire data, adequate DNA material and consent to genetic analyses. DNA was extracted from 1-5 ml whole blood (Sweden, Switzerland and the Netherlands) using the QIAGEN KIT or from buffered white blood cells (Germany and Austria) using a standard non-enzymatic method and the QIAGEN KIT. The average DNA stock concentration ranged from 62.5 ng/µl (Austria) to 386 ng/µl (Germany) and the average total DNA yield from 7.2 µg (Austria) to 33.2 µg (the Netherlands).

GENOTYPING PROCEDURE

The DNA samples were genotyped using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (SEQUENOM Inc., San Diego, California). PCR assays and associated extension reactions for each SNP were designed using the SpectroDESIGNER software (Sequenom Inc., San Diego, California) and primers were obtained from Metabion GmbH (Planegg-Martinsried, Germany). All amplification reactions were run in the same conditions in a total volume of 5 µl with 2.5 ng of genomic DNA, 1 pmol of each amplification primer, 0.2 mM of each dNTP, 2.5 mM MgCl2 and 0.2U of HotStarTaq DNA Polymerase (Qiagen). Reactions were heated at 95°C for 15 min, subjected to 45 cycles of amplification (20 s at 94°C, 30 s at 60°C, 30 s at 72°C) before a final extension of 7 min at 72°C. Extension reactions were conducted in a total volume of 9 µl using 5 pmol of allele-specific extension primer and the Mass EXTEND Reagents Kit before being cleaned using SpectroCLEANER (Sequenom Inc., San Diego, California) on a MULTIMEK 96 automated 96-channels robot (Beckman Coulter, Fullerton, California). Clean primer extension products were loaded onto a 384-elements chip with a nanoliter pipetting system (SpectroCHIP, SpectroJet, Sequenom) and analyzed by a MassARRAY mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany). The resulting mass spectra were analyzed for peak identification using the SpectroTYPER RT 2.0 software (Sequenom). For each SNP, two independent scorers confirmed all genotypes.

LABORATORY QUALITY ASSESSMENTS

Positive and negative controls

The DNA samples were arranged on a 96-well plate with two negative (H_20) and two positive controls (CEPH.1331-1) positioned uniquely on each sample plate. The controls were used to assess both genotyping quality (consistency of calls) and sample management (e.g., controls at the correct position).

Success rates and Hardy Weinberg equilibrium (HWE)

Each assay was controlled with regard to success rate, both in the optimizing step (35 unrelated individuals including three CEPH samples) and in the actual project. Assays with success rates < 85% were discarded or redesigned and then again optimized and evaluated.

The relationship between allele frequencies (exemplified by p as the frequency of the dominant allele and q as the frequency of the recessive allele for a trait controlled by a pair of alleles A and a in a given population) and the predicted genotype frequencies in a random sample was originally described by Hardy and Weinberg in 1908.[52] The equation reads $p^2 + 2pq + q^2 = 1$, where, p^2 is the predicted frequency of homozygous (AA) people in a population, $2pq$ is the predicted frequency of heterozygous (Aa) people, and q^2 is the predicted frequency of homozygous (aa) people.

If the observed genotypes in a given population differ from the predicted genotypes according to the equation above, deviation from the HWE is said to exist. Estimation of the HWE in each assay is a standard tool for genotyping control, as deviations from HWE may indicate genotyping errors or non-specific assay primers if the studied population meets basic criteria such as random mating and no inbreeding. Evolutionary selection, genetic drift, chance alone or genotype association with the disease of interest (if cases and controls are analysed together) may also lead to deviations from the HWE.[142] In our data, a cut off level of 0.01 (using an ordinary chi^2-test) was used to indicate deviation from the HWE.

Amelogenin test

This is a sex-specific assay that can be used to evaluate concordance between reported sex and genetically determined sex, in order to find errors in blood sampling, transportation and/or data handling. The human amelogenin gene is located on both the X- and Y-chromosomes as single copies in X and Y homologous regions.[143] Several PCR primer sets have been developed and the most commonly used PCR-based sex test is the one described by Sullivan et al.[144] The test has been shown to be very reliable in terms of accuracy and reproducibility.[145] However, there are reports of erroneous gender identification using the amelogenin test.[146] The assay was run in all BAMSE and PARSIFAL samples in the beginning of each project.

STATISTICS

Logistic regression and confounding control

In paper I, a multinomial logistic regression model was used to analyze associations between hereditary factors and wheezing phenotypes. The results were presented as Odds Ratios (OR) with 95% confidence intervals (CI) adjusted for potential confounders; sex, parental allergic disease (defined as doctor-diagnosed asthma and/or allergy (hay fever or allergy to pollen/pets) in mother, father or both), socio-economic status (based on the parents' profession), mother's age at enrolment, maternal smoking during pregnancy or at enrolment, and pet ownership (cat, dog or rodents in the home or at relatives at enrolment), which were identified by running several models with a number of covariates. Logistic regression models were also used in paper II, IV and V to test the effect of a particular haplotype (paper II), gene-gene interaction (paper IV) and gene-environment interaction (paper V). These models were also adjusted for a number of potential confounding factors (paper II and IV as listed above plus ethnicity) although no confounding effect was present in paper II and IV. In paper V, adjustments were made for municipality, socioeconomic status, heredity, maternal smoking during pregnancy and/or at enrolment, construction year of the residence, damp or mould in the home at birth and sex of the child. Additional factors were also evaluated in paper V (ethnicity, mother's age at enrolment, pet ownership, breastfeeding, number of siblings), but these variables had no confounding effect and were therefore left out of the model. For the continuous outcome variable PEF (peak expiratory flow), a linear regression was used to test differences in PEF-values between wheezing outcomes

(paper I; adjustment was made for sex, height, age and body weight). All regression analyses were performed with Stata, Statistical Software: releases 7.0 and 8.0 (College Station, Texas, USA).

Genetic association analyses

In paper II-V, allelic association analyses between the SNPs in each gene and the outcomes of interest were performed to get an estimate of the single locus associations. Block-wise inheritance of the SNPs in each gene was checked by estimating the relative linkage disequilibrium by the measure D' or r^2 using the Haploview program (based on the EM algorithm; paper III-V), which was also used to estimate the haplotype frequencies in cases and controls, respectively.[73] The Haploblocks program developed by M. Zucchelli and J. Kere (also based on the EM algorithm; unpublished data) was used in paper II for estimation of the relative LD between the SNPs and for estimation of the individual haplotypes in each gene in paper II-IV (the haplotypes were inferred for cases and controls together). For the haplotype association analyses, each individual contributed two haplotypes (one from each chromosome) to the analyses and the frequency and estimated counts of each haplotype were then assessed against all others using standard procedures for OR calculation. In paper III, haplotype association was also tested in the statistical software R[147], using the haplo.score algorithm, which generates both a global and haplotype-specific score statistics by comparing estimated haplotype frequencies in cases and controls.[65] Adjustment was made for a variable with one category for each combination of country-of-origin and sampling group, since both BAMSE and PARSIFAL individuals were included.

Gene-gene and gene-environment interactions

Interaction effects between different factors can be assessed by several methods and under different assumptions (e.g., on additive or multiplicative scales, and considering various effect measures) depending on study design and underlying hypothesis.[148, 149] The terms biological and statistical interaction are also used sometimes.[149] Biological interaction can be defined as participation of two (or more) risk factors in the same causal mechanism of disease (both factors are thus part of the same sufficient cause for the disease).[150] Statistical interaction is usually referred to as effect modification, or heterogeneity of effects, assessed through a statistical model with a defined scale for the outcome (e.g., risk or log-risk scale). In this thesis, it has been the intention to estimate *biologically relevant* interaction effects and for that purpose, both the additive and multiplicative models have been used.

In paper I, the interaction analyses were based on an additive scale as described by Rothman and Greenland.[150] Under this model, the expected OR for individuals jointly exposed to two factors A and B is the sum of the individual effects of A and B separately. The interaction effect (or the relative excess risk due to interaction) with CI 95% between parental allergic disease and male sex were estimated using the following categories: a) female sex and no parental disease (reference group); b) male sex and no parental disease; c) female sex and parental disease and d) male sex and parental

disease. The interaction effect was calculated as follows: $(OR_d - OR_a) - (OR_c - OR_a) - (OR_b - OR_a)$ and takes the value of zero if the joint effect of the two factors is only additive.

Estimation of gene-gene interaction, or epistasis, (paper IV) in the search of multiple genes affecting complex disorders is usually addressed by assessing departure from multiplicative effects on the OR scale using logistic regression (logit (Disease) = $\beta_0 + \beta_1 A + \beta_2 B + \beta_3 AB$), followed by LRTs (likelihood-ratio test) or Wald tests .[107, 148, 151-153] Under this model, the expected OR for individuals jointly exposed to two risk factors A and B is the product of the individual effects of A and B separately. A LRT between the models with and without interaction terms is performed to test the null hypothesis of no genotype-genotype interaction and the obtained p-value estimates a departure from a multiplicative interaction model on the OR scale and indicates whether the effect (OR) of one genotype is altered by effects of another genotype.
When the mode of inheritance is unknown, which is the case for asthma and allergy,[60] a log-additive coding of the genotypes has been suggested (each SNP coded 0 (wild type homozygous), 1 (heterozygous) and 2 (rare homozygous) if possible).[154]

Estimation of gene-environment interactions are also usually performed on a multiplicative interaction scale using a logistic regression framework.[110, 155, 156] Using this approach, both categorical and continuous variables can be included in the interaction analyses. In paper V, such tests for gene-environment interactions between genotypes of interest and exposure to NOx, PM_{10} and SO_2 with respect to asthma and allergy related outcomes were performed. The obtained p-value indicates in this case whether the effect (on the OR scale) of a certain exposure is altered by presence of a specific genotype.

Correction for multiple testing

In most research fields, a number of hypotheses are usually tested in a study rather than a single hypothesis. These tests may involve several phenotypes, multiple genetic or environmental factors or multiple test statistics. The more hypotheses being tested, the higher probability of obtaining false positive results by chance, given that the significance level is fixed (e.g., at 0.05). In genetic association studies, correction for multiple testing has been an issue for some time, as the number of markers tested is usually quite large, nowadays even up to 500,000 or more (e.g., by the recently developed 500 K SNP chips). Nevertheless, too many reports of genetic associations have been proven false positive and this issue has now become a major concern. The factors that particularly determine the magnitude of false positive findings are the prior probability of a true association (thus the biological plausibility), significance level and statistical power in the study.[157]

The traditional Bonferroni correction for multiple testing (to obtain a corrected p-value, the nominal (unadjusted) p-value for association is simply divided with the number of tests performed) has been shown to be too conservative, partly because markers are usually not independent of each other (see "Linkage disequilibrium" in METHODS) and phenotypes are not independent of each other (e.g., asthma and allergy). A less

conservative framework has been developed to control for the false discovery rate (FDR), which may be used for both correction of multiple tests of main effects as well as interactions.[110, 158, 159] The FDR is defined as the ratio of expected false positive associations to the total number of significant associations, and can be expressed either as the FDR itself or as a new corrected p-value for significance (paper V).

A permutation test is another way to test whether the association observed in the data is likely to occur by chance or not. The variables of interest (in our case the genotypes or haplotypes) are treated as fixed, while the phenotypes (case or control) are randomized, thus giving each individual a new affection status. The association tests are then preformed again (e.g., χ^2-tests). The proportion of 10-50,000 such iterations (randomised χ^2-tests), where a stronger association is found than in the actual data, provides the so called empirical p-value for association after correction for multiple tests (paper II-V). Similar approach can also be used for estimation of an overall interaction between for example markers in two genes (paper IV) or between several markers and an environmental exposure.

An alternative approach is to reduce the number of tests that are actually performed. This reduction can be achieved by looking at the haplotype level, thus combining the different alleles into haplotypes (preferably using a haplotype tagging strategy), and analyze the overall distribution of haplotypes in cases versus controls by for example the χ^2-test. Thereby, the association between several haplotypes (both risk and non-risk) and the outcome is tested simultaneously (paper II-IV). Other reduction techniques have been developed for screening purposes in family based studies, but these methods will not be further discussed here as they are not applicable to our case-control data.[160, 161]

RESULTS

CHARACTERIZATION OF WHEEZING PHENOTYPES AND INFLUENCE OF PARENTAL ALLERGIC DISEASE

Using the full BAMSE cohort, children with wheezing episodes up to the age of four were classified into the following groups: transient, persistent and late-onset wheezing (Figure 8). Children with persistent and late onset wheezing had the highest occurrence of sensitisation to inhalant allergens (23% and 30%, respectively). Significantly lower mean PEF-values were seen in children with transient and persistent wheezing (mean difference –8.9 and –8.5 l/min, respectively), whereas the PEF values in children with late onset wheezing did not significantly differ from those with no defined wheezing.

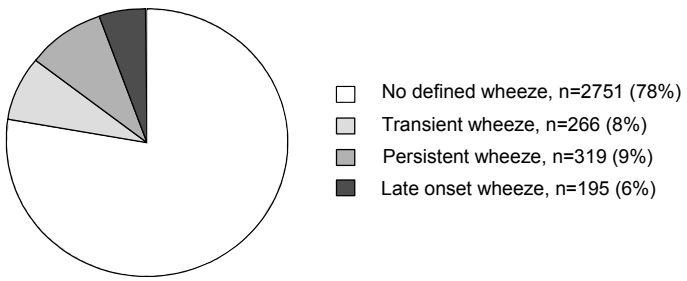

Figure 8. Prevalence of wheezing phenotypes up to the age of 4 years in the BAMSE cohort.

Both maternal and paternal allergic disease was of importance for all wheezing outcomes in the children, although less effect was seen on transient wheezing. Maternal allergic disease appeared slightly more associated with persistent wheezing than paternal disease ($p=0.03$), whereas no significant difference was seen in children with transient or late onset wheezing ($p>0.05$). The influence of parental allergic disease on the risk of persistent wheezing seemed to be more pronounced in boys than in girls; parental allergic disease (particularly maternal allergic disease) was only a minor risk factor among girls, but constituted a 3-fold risk increase among boys (Table 1). Significant interaction effects were also obtained using other outcomes, such as any wheezing at 4 years (persistent and late onset wheezing combined), steroid treated wheezing at 4 years or wheezing and concomitant sensitisation.

	No heredity	Heredity
	OR 95% CI	OR 95% CI
Girls	1.0 (Ref)	1.4 (1.0-2.1)
Boys	1.0 (0.7-1.5)	2.9 (2.1-4.0)
Interaction effect (additive model)		1.4 (1.1-1.7)
P-value interaction multiplicative model*		0.009

Table 1: Interaction analysis of sex and parental allergic disease in children with persistent wheezing. The "Interaction effect" was calculated from an additive model, whereas the p-value* for interaction was calculated using a multiplicative interaction model.

OVERALL GENOTYPING RESULTS; QUALITY ASPECTS

In each of the selected candidate genes (except for *GPRA*), a varying number of SNPs were excluded already in the optimizing step based on monomorphic status, assay success rate or deviation from HWE. For example, out of 29 potential SNPs in the *IL4RA* gene, 10 markers were found monomorphic, two had a success rate below 85% and one were out of HWE ($p<0.01$), whereas only one *GSTP1* marker out of seven was excluded due to a success rate below 85%. Overall, the success rate for each SNP was >90% in each project and the amelogenin assay indicated than sampling and genotyping errors were below 5% in the BAMSE samples and below 3% in the PARSIFAL samples. An internal control using an *IL9R* SNP (rs731477) genotyped repeatedly 3 times, gave however an estimated genotyping error of 1.2% (BAMSE only). As the projects proceeded, it became evident that some DNA samples failed repeatedly for several assays, and these samples were therefore excluded (e.g., 25 samples with complete genotyping failure in the *IL4RA* project (paper IV), 66 samples and 73 samples with >50% genotyping failure in the GPRA and air pollution study, respectively (paper III and V)). Allele and haplotype frequencies in all projects showed overall good agreement with previously published frequencies or data from NCBI or the HapMap project.[71] For the GPRA project, the frequencies could directly be compared between the BAMSE and PARSIFAL samples, which showed that both the allelic and haplotype frequencies were very concordant.

ASSOCIATION BETWEEN CANDIDATE GENES AND RESPIRATORY SYMPTOMS / ALLERGIC SENSITISATION

Table 2 summarizes the associations between SNPs / haplotypes in the selected candidate genes and various outcomes of interest in the BAMSE (and PARSIFAL) study (paper II-V). Only nominal p-values are presented in the table, and the issue of correction for multiple testing has been addressed somewhat differently in respective study (see the Statistics section in METHODS). Variations in the *IL9R* gene seemed to

Table 2. Significant associations between SNPs / haplotypes and various outcomes

Gene and SNP / haplotype	Location	Nominal p-values	Outcome
IL9R			
rs731476 [A/G]	Intron 1	0.003	Any wheeze
rs731478 [C/G]	Intron 1	0.004	Any wheeze
rs731478 [C/G]	Intron 1	0.043	Sensitisation
GAGC haplotype	Intron 1	0.0007	Any wheeze (boys)
GAGC haplotype	Intron 1	0.049	Sensitisation
IL4RA			
rs1805011 [A/C]	Exon 9 (E375A)	0.047	Transient wheeze
CGAA haplotype	5' UTR - Intron 3	0.028	Any wheeze
TAGG haplotype	5' UTR - Intron 3	0.037	Sensitisation
CGGGATA haplotype	5' UTR - Exon 9	0.007	Sensitisation
*GPRA**			
rs324384 [T/C]	Intron 2	0.019	Sensitisation
rs324384 [T/C]	Intron 2	0.044	Allergic asthma
rs324396 [C/T]	Intron 2	0.029	Sensitisation
rs324396 [C/T]	Intron 2	0.020	Asthma (+ allergic asthma)
H1 haplotype	Intron 2	0.038	Sensitisation
H3 haplotype	Intron 2	0.046	Allergic rhinoconjunctivitis†
H5 haplotype	Intron 2	0.019	Sensitisation
H5 haplotype	Intron 2	0.026	Asthma‡ (+ allergic asthma)
H6 haplotype	Intron 2	0.015	Sensitisation
ADRB2			
No associations			
TNFa			
rs1799724 [C/T]	-857 (promoter)	0.009	Transient wheeze
rs1800629 [G/A]	-308	0.031	Asthma
rs1800629 [G/A]	-308	0.033	Persistent wheeze
rs1800629 [G/A]	-308	0.0005	Sensitisation
rs1800610 [C/T]	Intron 1	0.029	Transient wheeze
TCGC haplotype	-1031 - intron 1	0.002	Sensitisation
TCAC haplotype	-1031 - intron 1	0.0008	Sensitisation
TCAC haplotype	-1031 - intron 1	0.036	Asthma
TCAC haplotype	-1031 - intron 1	0.038	Persistent wheeze
TTGT haplotype	-1031 - intron 1	0.017	Transient wheeze

Table 2, continuation.

GSTP1			
rs1799811 [C/T]	A114V	0.0004	Asthma
rs1799811 [C/T]	A114V	0.034	Persistent wheeze
AGTTTC	Intron 4 - exon 6	0.008	Asthma

* PARSIFAL and BAMSE analysed together

† PARSIFAL only

‡ Current asthma was used as the outcome in BAMSE

influence the susceptibility to both wheezing and sensitisation. When the analyses between *IL9R* SNPs and wheezing were run separately in boys and girls, both allelic and haplotype associations (GAGC haplotype) indicated an effect predominantly in boys. However, the interaction between sex and the GAGC haplotype was not statistically significant (p=0.10). No overall effect of the *IL4RA* SNPs was observed and only weak associations to wheezing and sensitisation could be proved when haplotypes were considered. Variants in the *ADRB2* gene showed no overall association to any of the outcomes, whereas the *TNF-α*-308 SNP seemed to affect the risk of sensitisation at the age of four. For some *TNF-α* variants, associations were also seen to wheezing and asthma, but none of these associations was significant after correction for multiple tests (10,000 permutations). Ala114Val was the only SNP in the *GSTP1* gene that showed any association (particularly to asthma).

For the *GPRA* association analyses, asthma and allergic sensitisation were used as major outcomes and the study was designed to evaluate the role of certain haplotypes on these outcomes using both BAMSE and PARSIFAL study subjects. Supported by phylogenetic analyses, both risk haplotypes (H5/H6) and non-risk haplotypes (H1/H3) could be identified, and these haplotypes seemed to predominantly influence the risk of sensitisation, but also asthma, allergic asthma and allergic rhinoconjunctivitis.

GENE-GENE INTERACTIONS

Interaction (or epistasis) analyses between variants in the *IL9R* and *IL4RA* genes showed that the effect of *IL4RA* SNPs on wheezing up to the age of four was modified by SNPs in the *IL9R* gene (global test for interaction p=0.009). Figure 9 shows how the effect of *IL4RA* Q576R genotypes (coded as RR/QR vs QQ, Q=Gln, R=Arg) on wheezing is dependent on the *IL9R* rs731476 genotype (p-value for interaction 0.005), and both increased risk (if rs731476 = GG) and decreased risk (if rs731476 = AA) could be observed. Interaction effects between the *IL4RA* and *IL9R* genes were also suggested on the haplotype level, more specifically *IL4RA* P503-R576 haplotype and *IL9R* GAGC haplotype. In the presence of the *IL9R* GAGC haplotype, *IL4RA* P503-R576 showed a protective effect on wheezing (OR= 0.58, 95% CI 0.35-0.96), whereas no effect was seen in the absence of *IL9R* GAGC (p-value for interaction= 0.013). No gene-gene interaction was observed with regard to sensitisation.

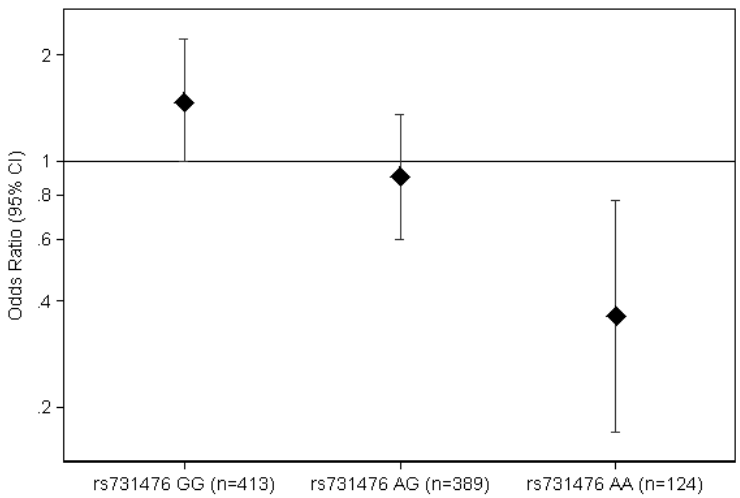

Figure 9. The effect of *IL4RA* Q576R genotypes (QR/RR versus QQ) on wheezing, in relation to *IL9R* SNP rs731476 genotypes.

GENE-ENVIRONMENT INTERACTIONS

The major finding in the study on interactions between air pollution and asthma-susceptibility genes, was that the risk of allergic sensitisation related to exposure to traffic-NO_x (an indicator of pollutants generated by motor vehicles) was increased in individuals heterozygous for the *GSTP1* SNPs tested, including the Ile105Val polymorphism (nominal p-values for interaction 0.0005-0.0085). The p-value cut off that corresponds to a nominal value of <0.05 after correction for multiple tests was estimated to 0.0085. An example is given in Figure 10, which shows the OR for sensitisation associated with exposure to traffic-NO_x (5-95 percentile) for different *GSTP1* genotypes (Ile105Val and Ala114Val exemplified). As expected, interaction effects were also seen for exposure to traffic-PM_{10} on sensitisation since these exposure variables were highly correlated in this data set (r=0.94). Weak evidence of interaction between SNPs in the *ADRB2* gene and exposure to house heating-SO_2 was observed using asthma and persistent wheezing as outcomes, although these interactions were not significant after corrections for multiple tests. No interaction was seen between exposure to air pollutants and *TNF-*α variants.

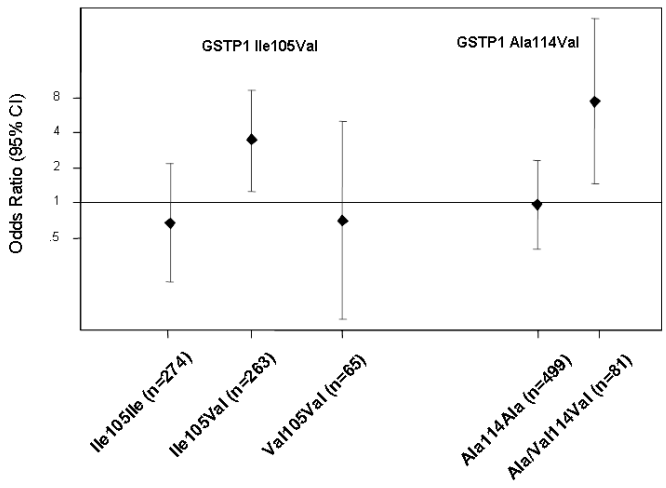

Figure 10. OR for sensitisation associated with exposure to NO_x for different *GSTP1* genotypes (Ile105Val and Ala114Val). The OR is expressed as a 90 percentile increase in NO_x levels (5-95 percentile) during the first year of life.

DISCUSSION

INFLUENCE OF PARENTAL ALLERGIC DISEASE

In paper I, we found that both maternal and paternal allergic disease (asthma and/or allergy) seemed to increase the risk of transient, persistent and late onset wheezing, and a dominating influence from maternal allergic disease was only seen in children with persistent wheezing. Several studies support the importance of maternal asthma for childhood asthma, especially in children < 5 years of age[162, 163], whereas other show that maternal and paternal asthma confer similar risk.[164] For children older than 5 years, there is accumulating evidence that paternal history of asthma has an equal or even stronger effect on asthma and airway hyper-responsiveness compared to maternal history.[163, 165] The maternal effects have been suggested to be partly regulated by the interaction between mother and foetus *in utero*[30, 31], which is further supported by associations between maternal allergic disease and risk of high total IgE at birth.[166] Linkage studies have attempted to genetically disentangle effects from the mother and father, and both maternal and paternal imprinting has been suggested for allergic sensitisation and asthma-related outcomes.[167-169] Genomic imprinting refers to the differential expression of genes depending on whether they come from the mother or father, and has also been reported to play a role in other complex diseases, such as diabetes.[170]

WHY ARE BOYS MORE SUSCEPTIBLE TO ALLERGIC DISEASES IN CHILDHOOD?

It is well known that boys are more susceptible to childhood asthma and allergy, whereas girls are more susceptible later in life.[171-175] Environmental factors have a major impact on the likelihood of developing asthma and allergy, and boys and girls are probably exposed differently in the environment in which they grow up. However, already at birth, twice as many boys compared to girls have been reported to have a high neonatal total IgE, which suggests that sex-specific events may occur already *in utero*.[176] Whether this is due to genetic, immunologic or hormonal factors is unclear. Boys have also been reported to have smaller airways relative to lung size than girls and may therefore be more sensitive to airway irritants such as infections or other airborne triggers.[177] In paper I, we found that the influence of parental allergic disease, particularly maternal, on the risk of wheezing was more pronounced in boys than in girls, which lends support to a sex-specific genetic influence. Similar results have been suggested from the PIAMA study, where boys have an increased risk of sensitisation at 4 years if the mother was allergic.[176] Interaction effects have also been observed in one-year-old children, where the lung function in male infants was affected more by maternal history of asthma than in female infants.[177] The interplay between mother, placenta and foetus during pregnancy has been shown to be sex-specific also from the mother's point of view; asthmatic mothers carrying a female foetus experience worse asthma compared to mothers carrying a male foetus.[178] An explanation may be that the female foetus up regulates maternal inflammatory pathways, whereas the male foetus does not.[179] As a result of this inflammation, reduced foetal growth is seen in female,

but not male foetuses. These data may seem somewhat contradictory to an increased susceptibility of asthma and allergy in boys, but there are indications that low birth weight may be associated with lower risk of subsequent asthma and allergy[180], and that factors responsible for foetal growth may also lead to the development of an immune system predisposed to the increased susceptibility of asthma and/or atopy.[181] Nevertheless, it highlights the importance of potential sex-specific *in utero* events that may influence the development of asthma and allergy.

Sex-specific genetic effects have been observed in a number of human traits, with regard to heritability, genome-wide linkages and gene expression.[182, 183] For asthma-related traits, there is evidence of linkage between specific autosomal loci and affected individuals of a particular sex and age of onset of wheezing[184], as well as lung function and eosinophilia.[183] It has even been suggested that by not taking sex-specific effects into account, the power to detect linkage may be substantially reduced.[185] Sex-specific genetic effects on asthma have been reported for a number of genes, for example *interleukin-1β* (affecting male asthma), *ADRB2* (bronchial hyperresponsiveness in females), *vitamin D receptor* (IgE levels in girls) and *COX2* (asthma in females).[186-189] A skewed male/female ratio could also imply that X-and Y-linked genes are involved in some way. Recent data suggest that up to 15% of the X-linked genes escape inactivation to some degree and this observation does not include the pseudoautosomal regions, which are known to escape inactivation.[190] In paper II, we report that variants in the X- and Y-chromosome gene *IL9R* were associated with wheezing and sensitisation, predominantly in boys. Although the major finding was that the GAGC haplotype had a protective effect primarily in boys, it indicates that the genetic influence of *IL9R* variants, protective or not, is larger in boys compared to girls. Given the strong LD in the region, and the fact that we were not able to study any functional effects of the selected SNPs, it is perhaps more relevant to interpret our results as a sex-specific association to the *IL9R* gene region rather than to any specific variants.

CANDIDATE GENES AND ASTHMA-RELATED PHENOTYPES

As presented in the results section, we have analysed the association between variants in six genes (*IL9R*, *IL4RA*, *GPRA*, *ADRB2*, *TNF-α* and *GSTP1*) and asthma-related phenotypes in childhood. *IL9R*, *GPRA* and *TNF-α* showed the most consistent findings in that several variants were associated with more than one outcome tested. However, *IL4RA*, *ADRB2* and *TNF-α* belong to the group of genes that have been replicated in > 10 study samples, and *GPRA* and *GSTP1* have been replicated in 6-10 samples.[90] We also consider the *IL9R* gene to be an asthma susceptibility gene, since it has been associated with asthma in two family-based data sets besides our study, and no negative study has to our knowledge been published.[94, 95]

Replication is considered the gold standard for genetic association studies, but there is yet not a single gene that has been replicated in all studies. Replication of original genetic findings often correlate only moderately well to subsequent research on the same association, with the strongest effect estimate typically occurring in the first study.[191, 192] For postitionally cloned genes, this fate is not surprising, as linkage studies rely on a strong genetic signal in order to identify a particular locus or gene.[193] A

number of factors will influence this signal (e.g., population characteristics, disease heterogeneity, environmental factors, other genetic factors and study power) and these factors will have to act jointly if a strong signal is to be detected in a linkage study. Upon replication attempts, these conditions will not be the same as in the original study; other environmental and genetic factors will have stronger or weaker influence. Given the complexity of asthma-related phenotypes, numerous combinations between these factors are possible, and the most likely scenario is that the effect of a cloned gene in a replication study is diluted by the effects of other factors, thereby giving a weaker association.

The effect size of a particular variant in a susceptibility gene for complex diseases across populations is usually around 1.2-2 on the odds ratio scale.[191] The first postitionally cloned asthma gene, ADAM33, was in a recent meta-analysis reported to have a maximum odds ratio of 1.46 for a common SNP using asthma as outcome[194], which is of the same magnitude as we found for *GPRA* H5 haplotype and allergic asthma, odds ratio 1.47 (paper III).

Traditionally, replication studies have focused on analyses of specific SNPs or haplotypes that have been identified in the original study. Due to reasons discussed above, all replication studies will not find associations to a given variant in a particular gene regardless whether the gene was identified by cloning or by other means (e.g., candidate gene approach). The *GPRA* gene for instance, has been considered replicated in two separate studies, whereas another study could not replicate the most significant SNP (rs323922) from the original study.[88, 195] If we only focus on rs323922 in the two positive replication studies, no association was found in these studies either (paper III,[114]). Still, we consider the two studies to have replicated the association between *GPRA* and asthma / allergy based on associations with other SNPs and several haplotypes, among them H6. This haplotype was not considered a risk haplotype in the original study, partly because of lower power to detect any association with H6 in the Finnish and Canadian populations. The phenomenon of different risk alleles in different populations is very common in genetic association studies, and it has been suggested that the gene, rather than any specific SNP or haplotype should be the unit of replication.[196] With the new large scale genotyping possibilities that enables examination of all relevant variations within a gene or region, for example by haplotype-tagging methods or chip-based SNP analyses, such multilocus approach as well as other similar multivariate methods is promising for future studies.[197-200]

ASTHMA DEFINITIONS AND THEIR RELEVANCE FOR GENETIC STUDIES

Despite a number of common characteristics for asthma-related diseases, primarily represented by airway inflammation and upper respiratory symptoms, these conditions constitute a heterogeneous group of phenotypes.[1] As discussed above, different environmental and genetic factors will probably influence each phenotype in a specific manner. The hypothesis commonly tested in genetic association studies, is whether a particular variant is associated with disease in the study sample of interest, e.g., cases of childhood asthma. A proper definition of cases and controls is the basis for all clinically oriented research and these definitions should be well considered. The choice of case

definition will naturally reflect the aim of the study; is the aim to demonstrate a genetic association between a particular variant and a very carefully defined phenotype, e.g., asthma patients aged X-Y years with sensitisation to ragweed and concomitant rhinoconjunctivitis, *or* are we interested in evaluation of the association between this variant and a broader definition, e.g., doctor's diagnosis of asthma in children? In general, the better cases and controls are characterized, the more distinct conclusions can hopefully be drawn in *that particular setting*, and comparison between studies with exact definitions is also facilitated. The broader definition will obviously represent a more heterogeneous phenotype and there is also a clear risk of disease misclassification, which will reduce the power to detect an association if there is one.[201] Still, this definition is important, as it hopefully reflects every day practice in the clinic and the question to be answered is *how important is this genetic variant from a population point of view*, given that cases and controls have been collected properly.

The definitions used in both the BAMSE and PARSIFAL studies are based on questionnaire data (except IgE and PEF measurements) and are quite broad and general. First of all, the study designs did not allow inclusion of bronchial hyperresponsiveness or examination by a physician in the disease criteria. Atopic status was however assessed and was also used for the phenotypes allergic asthma / rhinoconjunctivitis and atopy-associated wheezing. It should be noted that these definitions reflect co-occurrence of specific symptoms such as wheezing, and presence of IgE antibodies towards common allergens as signs of an immunological event, rather than a confirmed allergy-induced asthma or wheezing.[44] The BAMSE children are by study design only up to 4 years of age, and the majority of the children with these definitions reflect non-atopic wheezing (71%, paper II). Nevertheless, one of the primary aims with the BAMSE genetic study was precisely to investigate the genetic influence of such wheezing and the definitions are similar to those used for instance in the Tucson Children's Respiratory Study.[4] Follow-up of the Tucson study up to the age of 16 years showed that transient wheezers (defined at 6 years) do not have an increased risk of subsequent wheeze, whereas around 50% of the persistent and late onset wheezers were symptomatic at the age of 16.[38] Thus, the majority of the wheezing children in BAMSE will probably grow out of symptoms in a few years time. Although we hypothesized that some of the genetic markers analysed could be used to disentangle persistent from transient wheeze on a genetic basis, no such pattern was obvious. In study II and IV, all wheezing cases was instead analysed together in order to increase the study power and reduce the number of hypotheses tested.

Doctor's diagnosis of asthma presumably represents a more severe form of respiratory disease than the wheezing outcomes and this definition is quite commonly used in studies on asthma.[7, 8, 114, 202] Although the majority of preteenage children with established asthma will continue be asthmatic, childhood and adult asthma may differ in a number of parameters, including sex ratio, airflow obstruction, sensitisation and structural changes of the airways.[40, 42, 203] These differences are likely to be important when determinants for these phenotypes are to be identified, and there may be both common and unique risk factors for childhood and adulthood asthma.[204]

The BAMSE study has the advantage of being prospective and longitudinal, which enables observation of the natural course of asthma-related diseases in a rather large

population of children. Several phenotypes may also be studied, e.g., wheezing and sensitisation. The subcohort sampling design is particularly attractive when multiple phenotypes are of interest, as other outcomes in subcohort can be studied still using properly selected controls. For this thesis, we have focused on sensitisation as a complementary outcome, but the BAMSE study has also been used as a replication data set for genetic studies on e.g. eczema.[205] Having the opportunity to study several outcomes is naturally a major advantage in cohort studies that one should make use of. The number of phenotype subgroups related to asthma and allergy is however almost infinite; asthma symptoms +/- abnormal lung function, +/- specific IgE to inhalant or food allergens, +/- rhinoconjunctivitis, +/- eczema etc. Each of these subgroups may have specific risk factors associated with susceptibility and prognosis. Although is it tempting to try to find specific genetic factors that are associated to each of these subgroups, it is usually not feasible due to power limitations and the BAMSE study is no exception in this respect.

GENERAL ASPECTS ON INTERACTIONS

Although there is a wide spread knowledge that asthma and allergy results from multiple factors including gene-gene and gene-environment interactions, the vast majority of all studies on genetic factors do not take these matters into consideration. One of the main reasons for this situation is the central problem of how to model interactions, and how to take high-order interactions, that is, interaction between several factors, into consideration. How do we evaluate possible combinations between the suggested 100 candidate genes for asthma, and how do we further assess possible interactions with relevant environmental exposures? Traditional models for evaluating interaction, e.g., by stratification, calculating risks for "double-exposed", "single-exposed" and "none-exposed" in a regular 2x2 table and logistic frameworks with interaction terms may be sufficient to evaluate possible interaction between a limited number of factors, but these models are inadequate for analyses of multiple variables. Still, the traditional models may be applicable in targeted and hypothesis-driven analyses. In paper IV and V, such interaction effects were estimated between just two factors (paper IV, two genetic factors and paper V, one genetic and one environmental factor) using the logistic framework. Secondly, assessment of interactions demands very large data sets and power issues easily becomes a problem in studies with ordinary sample sizes.[206]

Interaction analyses are usually performed when there is a hypothesis that two or more factors have a joint action that is of particular interest, and that the joint effect may be larger than expected based on biological relevant mechanisms. What is usually tested in the multiplicative model is whether the effect of exposure to a factor A will be influenced by the presence or not of factor B, that is, if the effect of A is independent of factor B. A schematic figure of such interaction is presented in figure 11a-b.

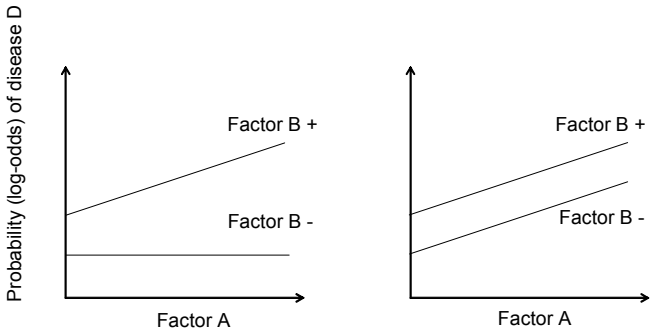

Figure 11a

Figure 11b

Figure 11a. Probability (expressed as log odds) of disease D as a function of exposure to factor A in two groups, exposed and unexposed to factor B. Interaction between factor A factor B is present on a multiplicative scale, as the effect of A is not similar in the groups with B+ and B- (the lines are not parallel). Figure 11b shows a situation when interaction between factor A factor B is not present on a multiplicative scale, as the effect of A is similar in the groups with B+ and B- (the lines are parallel). Both factor A and B could for example be genetic or environmental factors.

In paper I, an additive model was originally used to assess the interaction effect between male sex and parental allergic disease. As presented in Table 1, the same conclusion that interaction is present would be drawn if a multiplicative scale was used, which strengthen the findings. In other situations, interaction may be present on one scale, but not the other. Which scale is then the most relevant from a biological point of view? Do risk factors operate in an additive or multiplicative fashion? Obviously, there is no general answer and the different models can be applied in different settings, although word of caution has been given due to confounding and over-simplification issues regarding the additive interaction model.[207, 208]

Role of gene-gene interactions / epistasis

Gene-gene interaction is a rather broad definition that indicates that two or more DNA variations interact directly (DNA-DNA or DNA-mRNA) to change splicing, transcription or translation levels, or indirectly by the effect of their protein products in a way that is beyond what is expected by their separate effects.[209] Epistasis is another term that is used to describe how effects of a given gene are masked or enhanced by one or more other genes.[210] The endpoint of interest is usually disease risk, but may also be intermediate phenotypes such as receptor signalling or cytokine production. As discussed above, a widely used estimation of gene-gene interaction is to assess departure from multiplicative effects on the OR scale using logistic regression followed by LRTs or Wald tests.[107, 148, 151-153] Such approach was used in paper IV, when

interaction effects were estimated between 4 SNPs in the *IL4RA* gene and 4 SNPs in the *IL9R* gene. Based on significant SNP-SNP interactions and a significant overall test for interaction, it seems as if variants in the two genes have a joint effect on the susceptibility to develop wheezing in children. Interestingly, both increased and decreased risks for wheezing could be observed for the specific combinations (e.g., *IL4RA* Q576R and *IL9R* rs731476 SNPs). Given the substantial difference in allele frequencies observed between populations,[71] the phenomenon of epistasis, in this case *IL4RA* and *IL9R* polymorphisms, may in part explain discrepant results from different studies.[211] Evidence from research on hypertension suggests that lack of replication of single locus results often may be due to the fact that epistasis effects are more important than the main effect of a particular locus.[153, 212] In this case, there may not be any direct interaction between the *IL4RA* and *IL9R* genes, but the interaction could rather mirror the combined effect of the protein products encoded by the different genetic variants. As wheezing in this age group is primarily non-atopic, usually triggered by viral infections, the joint mechanism could indirectly be related to susceptibility to viral infections. Experimental evidence is however needed to prove such a hypothesis, although *IL4RA* variants have been shown affect susceptibility of viral infections in other settings and IL9 is for example a key cytokine upon RSV infection.[110, 213, 214]

Given the limitations with the traditional regression based interaction analyses, a number of new statistical models have recently been developed that allow for evaluation of high-order gene-gene interactions, such as Multifactor dimensionality reduction (MDR), Multivariate adaptive regression splines (MARS) and the Focused interaction testing framework (FITF).[159, 215, 216] Yet, no specific model is suitable for assessment of all respects of gene-gene interactions, e.g., categorical or continuous traits, linear or non-linear interactions, missing data, genetic heterogeneity and multiple test issues (see recent review by Thornton-Wells et al[209]). By combining two or more methods in a stepwise approach to genetic analyses, some of these limitations may however be overcome.

Role of gene-environment interactions / effect modification

Despite the estimated high heritability of asthma and allergy, environmental factors are of uttermost importance for the development of these diseases. It has been argued that the dramatic increase in asthma prevalence over the last decades can not be attributed to genetic factors, as our genetic set up has not changed so rapidly.[217] The explanation should rather be sought in changes of environmental exposure patterns, and how the environment affects genetically predisposed individuals.

Variants in many genes are likely to modify the physiologic and immunologic response to various environmental agents. CD14, which is one of the receptors for endotoxin, is one of the most studied genes in this context, and has been shown to modify the effect of environmental tobacco smoke, animal exposure and endotoxin exposure.[155, 156, 218] Gene environment interactions have not only been addressed in candidate gene studies; genome wide screens for asthma and bronchial hyperresponsiveness taking the influence of passive tobacco smoke have also been undertaken. Two separate studies showed linkage to chromosome 5q in families where the children were exposed to

tobacco smoke, thereby highlighting the importance of environmental exposure also in the search for new asthma genes.[219, 220]

Studying the effects of various air pollutants on respiratory health in relation to an individual's genetic setup is an attractive approach, as there is evidence from both epidemiological and experimental research in this area. The BAMSE study is well suited for studying the effects of long term exposure air pollutants, as there is excellent follow up rate, good phenotypic information and the possibility to assess individual exposure levels during the whole study period. The main finding was the gene-environment interaction effects between *GSTP1* variants and exposure to traffic-related NO_x and particles with regard to sensitisation, and similar results have been reported with regard to ragweed sensitisation and childhood asthma.[131, 132] However, it is surprising to find the strongest effect in heterozygous individuals, a pattern that could be seen for all SNPs tested and for several outcomes. These findings contradict the results by Gilliland et al, although study design (e.g., observational setting and long term effects vs. experimental setting and short term effects) and subjects differ somewhat.[131] In that study, 19 ragweed sensitised patients (age 20-34 years) were genotyped for the Ile105Val polymorphism and challenged intranasally with allergen alone and with allergen plus diesel exhaust particles. Patients homozygous for the Ile105 variant showed larger increase in nasal IgE responses to ragweed and histamine release compared with Ile105Val heterozygotes. No Val105 homozygous individuals were found among the patients. However, an increased risk of other diseases in relation to toxic exposures has been associated with Ile105Val heterozygosity, for example Parkinson's disease following exposure to pesticides and tobacco smoke.[221, 222] The highest risk of chemotherapy-induced leukaemia has also been observed in heterozygotes, which is further supported by experiments on carcinoma cell lines that show increased chemotherapeutic sensitivity in heterozygote subjects.[223, 224]

The antioxidative system, in this case represented by GSTP1, is a complex network and involves a number of important enzymes and proteins. The GSTP1 enzyme may exert its effect alone as a dimer (or multimer), but may also interact with other enzymes like the c-Jun N-terminal kinase (JNK) and the 1-cystein peroxiredoxin.[225-228] From an immunogenetic point of view, one possible explanation to these observations is that the enzymatic capacity of GSTP1 dimers may depend on the genotypic status. The interplay between GSTP1 and for example the c-Jun N-terminal kinase (JNK), which activates the transcription of number of genes including cytokines, growth factors, immunoglobulins, inflammatory enzymes,[229] may also be affected by the *GSTP1* genotype status. The GSTP1 enzymatic activity has been shown to differ between isoenzymes with isoleucine or valine 105; Val105 enzymes having for example a higher catalytic capacity for polycyclic aromatic hydrocarbons, which represent a widespread class of environmental pollutants, but a lower conjugation capacity for 1-chloro-2,4-dinitrobenzene.[230-232] However, it is unclear if these characteristics of the GSTP1 enzyme relate to the finding in our study. From this study, we conclude that variants in the *GSTP1* gene seem to modify the effect of long term exposure to ambient air pollution with respect to sensitisation to common allergens in children, but the mechanisms behind this observation need to be further addressed.

METHODOLOGICAL CONCERNS

Neither the BAMSE nor the PARSIFAL studies were originally designed for the primary aim to study genetic effects on allergic disease in childhood, which have influenced the possibilities to perform genetic analyses in these data sets. There is for instance no parental DNA available for family based analyses, although case-control studies are well suited for association studies given that subjects have been selected properly.[233] If cases and controls have different genetic backgrounds with inherent gene frequency differences, false positive associations may be observed due to the phenomenon called population stratification. Although population stratification may often be relatively unlikely to cause bias in real-life settings[234], it is difficult to completely rule out in creating false positive results in any case-control study. For the BAMSE cohort, this is unlikely a problem as both cases and controls were generated from the same study base with a rather homogenous population. Ethnicity was also included in the confounding model to rule out any potential effect (paper II, IV, V). Further, the subcohort was shown to be representative of the full original cohort. In paper III, children from five different countries with different environmental backgrounds were included, which could potentially give rise to a population stratification problem. However, the haplotypes frequencies were quite similar in the five countries and the estimated haplotype association p-values were adjusted for country of origin and study group, which should rule out any major residual effect. Genomic control by additional genotyping of several unrelated markers in both cases and controls is an alternative experimental approach to handle potential population stratification.[235]

Confounding from other factors was not considered to be a problem in the association studies, and all logistic models were controlled with a number of potential confounders. Regarding exposure to air pollutants, differential exposure misclassification is unlikely since these variables were assessed using residential address histories and dispersion modelling from emission databases. Geo-coding errors may however have influenced the assessed air pollution levels. In this model we estimated the outdoor air pollution levels from traffic. Yet, most children spend a large part of their time indoors, which makes the exposure assessment less precise. This error is also likely to be non-differential in relation to the outcomes studied and would tend to attenuate any true association. By using exposure during the first year of life only, we avoid possible reverse causality induced by avoidance behaviour due to the child's disease.

Some DNA samples in the BAMSE study were of poor quality, and a number samples failed repeatedly for several assays. The reason for repeated failure may be due to low DNA concentration (the mean stock concentration in the 73 samples excluded in paper V was 49.5 ng/µl compared to 237 ng/µl in all samples), DNA breakdown due to long term storage or errors in the initial sample management. Although the number of samples and accordingly the study power is reduced when samples are removed from the analyses, we believe that data quality and accuracy are improved by such action.

The BAMSE genetic study was primarily designed to study the genetic association to childhood wheezing and around 500 cases were included in the study. Using the case cohort sampling design, other outcome definitions could also be studied such as

sensitisation to either inhalant or food allergens (n=197). Initial power calculations showed that using 200 cases and 400 controls with an estimated allele frequency of 10% in the controls, and a relative risk of 2.0, the power would be 89% ($p<0.05$). The corresponding power for a relative risk of 1.5 and allele frequency of 20% in the controls was estimated to 74%. Thus, for estimation of the genetic effects from single markers, the study power was relatively good. The power to detect interactions, either gene-gene or gene-environment interactions, is clearly reduced as the number of cases and controls with each combination of variables will be limited. However, interaction based analyses can be even more powerful than single locus approaches when the main effects are weak.[211]

CONCLUSIONS

Based on the presented studies in this thesis, the following conclusions can be drawn:

- Parental allergic disease is important for the development of wheezing up to the age of four, but the hereditary influence seems to be more pronounced in boys than in girls

- The genetic analyses also provided evidence of sex specific effects in that variations in the *IL9R* gene were associated with wheezing and allergic sensitisation predominantly in boys

- Variants in a recently identified susceptibility gene for asthma (*GPRA*) showed association to several phenotypes, including sensitisation, asthma and allergic rhinoconjunctivitis

- Variants in the *IL4RA* gene alone may not exert any major influence on susceptibility to asthma-related diseases in childhood, but in combination with other genes, such as *IL9R, IL4RA* may be an important gene for disease susceptibility

- The risk of allergy following exposure to air pollutants early in life seems to be modified by variants in genes controlling the antioxidative system, such as *GSTP1*

SAMMANFATTNING PÅ SVENSKA

Astma- och allergisjukdomar är idag mycket vanliga i befolkningen och man räknar med att mer än 1/3 av alla barn någon gång drabbas av astma- eller allergibesvär. Såväl ärftliga som miljörelaterade faktorer påverkar risken att insjukna. Samverkan mellan dessa faktorer är sannolikt av stor betydelse. Astma och allergi är vanligare hos pojkar än flickor, men det är oklart om det beror på miljörelaterade, immunologiska eller genetiska faktorer. Den övergripande målsättningen med denna avhandling har varit att studera betydelsen av ärftlighet för utvecklingen av allergisjukdom hos små barn samt att studera eventuell samverkan mellan arv och miljörelaterad exponering. Ett flertal gener bidrar sannolikt till utvecklingen av astma och allergi och det har konstaterats att upp till 100 gener kan vara inblandade på olika sätt. Vi har valt att fokusera på dels gener som är av särskilt intresse för astma och allergi hos barn, dels gener som kan antas samverka med miljöfaktorer, i detta fall exponering för luftföroreningar.

Avhandlingen omfattar fem delstudier. Samtliga delar utgår från BAMSE-studien med ca 4000 barn i Stockholm, vilka har följts från födseln till 4 års ålder. För de genetiska analyserna har ca 1000 prover valts ut, varav ca 500 från barn med astmabesvär och 500 från barn utan astmabesvär. I delstudie III ingår dessutom prover från ca 3000 barn i åldern 5-13 år, vilka ingår i den europeiska tvärsnittsstudien PARSIFAL. PARSIFAL-studien har till syfte att studera arv och miljö i relation till olika livsstilar.

Studie I

Samtliga barn i BAMSE-studien ingick i denna delstudie. Målsättningen var att belysa hur ärftliga faktorer påverkar utveckling av astmabesvär hos pojkar respektive flickor. I studien gjordes en noggrann karaktärisering av barnens sjukdomsutveckling över tid. Tre typer av barnastma identifierades upp till 4 års ålder, (i) övergående astmabesvär (ii) kvarstående astmabesvär och (iii) sent debuterande astmabesvär. Vi fann att barn med kvarstående eller sent debuterande astmabesvär var allergiska (d.v.s. hade specifika IgE-antikroppar mot vanliga allergen) i störst utsträckning (upp till 30% av barnen). Trots att barn med övergående astmabesvär inte hade några symptom vid 4 års ålder såg vi hos dessa barn en lätt sänkt lungfunktion i samband med 4-årskontrollen. Även barn med kvarstående astmabesvär hade en lätt sänkt lungfunktion jämfört med barn utan astmabesvär. Astma och allergi hos föräldrarna ökade risken för alla typer av astmabesvär. Betydelsen av ärftlighet verkade dock vara större för pojkar än för flickor. Resultaten talar för att ärftlighet för astma och allergi påverkar pojkar och flickor olika.

Studie II

Interleukin-9 och dess receptor är involverad i flera steg av utvecklingen av astma och allergi. Genen för interleukin-9 receptorn (*IL9R*) återfinns på X- och Y-kromosomerna och har i tidigare studier kopplats till astma. Målsättningen med denna delstudie var att studera naturligt förekommande varianter (polymorfier) i genen för *IL9R* och allergisjukdomar hos barn upp till 4 år samt att studera om effekterna av dessa polymorfier skiljer sig mellan pojkar och flickor. Ett flertal polymorfier studerades initialt och slutligen valdes fyra ut för komplett analys. Flera av varianterna visade sig ha samband med astmabesvär och allergi. En specifik kombination av dessa polymorfier, s.k. haplotyp, hade en skyddande effekt mot utveckling av både

astmabesvär och allergi upp till 4 års ålder. Störst skydd mot astmabesvär sågs hos pojkar med haplotypen GAGC. Samma haplotyp var dock inte associerad med något skydd hos flickor. Resultaten tyder på att varianter i *IL9R*-genen påverkar pojkar i större utsträckning än flickor.

Studie III

I denna delstudie var målsättningen att studera sambandet mellan varianter i genen för *GPRA* och astma- och allergiutveckling i två olika studiematerial, BAMSE och PARSIFAL. Sammanlagt analyserades ca 4000 prover. *GPRA* har nyligen identifierats och utgör en stark kandidatgen för astma- och allergisjukdomar. Genen kodar för en G-proteinkopplad receptor som återfinns i bl.a. glatta muskelceller och epitelceller. Studier pågår för närvarande för att kartlägga dess funktion. I denna studie kunde vi framför allt se samband mellan olika haplotyper (och polymorfier) i *GPRA*-genen och allergi. Både ökad och minskad risk för allergi kunde kopplas till olika haplotyper. Samband kunde dessutom ses mellan *GPRA*-varianter och astma, allergisk astma och hösnuva. Sammantaget ger studien stöd för att *GPRA*-genen är involverad i utvecklingen av astma och allergi hos barn.

Studie IV

Ett flertal polymorfier i genen som kodar för interleukin-4 receptorn alpha (*IL4RA*) har kopplats ihop med astma och allergi. Immunologiska och genetiska studier har också visat att IL4RA samverkar med andra cytokiner och deras receptorer i relation till astma. I denna delstudie analyserade vi sambandet mellan polymorfier i *IL4RA*-genen och utveckling av astmasymptom och allergi upp till fyra års ålder. Vi kunde dock inte påvisa något tydligt sådant samband. Vidare studerades effekten av kombinationer av *IL4RA*- och *IL9R*-varianter i relation till astmasymptom och allergi. Vi kunde då påvisa samverkan (interaktion) mellan flera polymorfier i *IL4RA*- och *IL9R*-generna i relation till astmasymptom. Vi identifierade kombinationer som gav såväl ökad som minskad risk för astmasymptom. Resultaten stöder hypotesen att astma och allergi är multifaktoriella sjukdomar där flera gener samverkar i sjukdomsutvecklingen.

Studie V

Exponering för olika luftföroreningar (t.ex. kväveoxider och partiklar) kan ge upphov till inflammation i luftvägarna och öka risken för astma- och allergibesvär. Vissa individer kan vara extra känsliga för sådan exponering. Målsättningen med denna delstudie var att studera i vilken utsträckning samverkan mellan arv och exponering för trafikrelaterade luftföroreningar påverkar risken för astma- och allergibesvär. Ett antal gener valdes ut för analyser, vilka bl.a. kodar för inflammationsmolekyler (*TNF-α*), antioxidativa enzymer (*GSTP1*) samt luftvägsreceptorer (*ADRB2*). Data från individuell exponering för luftföroreningar upp till 4 års ålder finns tillgängliga för alla barn i BAMSE-studien. Varianter i *GSTP1*-genen visade sig ha betydelse för vilken effekt exponering för kvävedioxider och partiklar under första levnadsåret hade på risken att utveckla allergi. Barn med en särskild *GSTP1*-variant hade en flerfaldigt ökad risk för allergi kopplat till luftföroreningar jämfört med barn utan denna variant. Resultaten talar för att effekten av exponering för luftföroreningar för risken att utveckla astma- och allergibesvär påverkas av varianter i gener som involverade i kroppens immunförsvar.

SLUTSATSER

- Astma och allergi hos föräldrarna ökar risken för att utveckla astmabesvär upp till 4 års ålder, men betydelsen av ärftlighet verkar vara större för pojkar än för flickor

- Sambandet mellan varianter i *IL9R*-genen och astma/allergi är tydligast hos pojkar, vilket stödjer fyndet att ärftlighet kan påverka flickor och pojkar olika

- Varianter i den nyupptäckta ”astma-genen” *GPRA* påverkar risken för flera olika typer av astmarelaterade besvär hos barn, bl.a. allergi, astma och hösnuva

- Samband kunde inte påvisas mellan varianter i *IL4RA*-genen och astma/allergi, men i kombination med *IL9R*-varianter kan olika *IL4RA*-varianter medföra både ökad och minskad risk för astmabesvär

- Varianter i gener som är involverade i kroppens immunförsvar, t.ex. *GSTP1*, har sannolikt betydelse för vilken effekt exponering för luftföroreningar har på risken att utveckla astma- och allergibesvär hos barn

ACKNOWLEDGEMENTS

First of all, I would like to thank all children and families in the BAMSE and PARSIFAL studies. Your excellent participation has been the basis for all results and conclusions in this thesis.

A number of people have contributed to this work and I would like to express my sincere gratitude and appreciation to all of you.

Magnus Wickman, my fantastic supervisor who has been a true mentor and friend throughout the years. Thank you for all your support and encouragement in research matters, clinical work and other things that are important in life. It is a pleasure to do paediatric allergy research together with you.

Göran Pershagen, my co-supervisor and second mentor, thank you for your excellent guidance through my PhD-project and for sharing your deep knowledge in science in general and epidemiology in particular.

Juha Kere, my co-supervisor and third mentor, thank you for generously hosting me in your laboratory and for introducing me to the exciting field of genetics. You have truly inspired me to future research in this field.

The past and present directors at the Centre for Allergy research, Cfa, *Bengt Björkstén, Lars Gustafsson, Sven-Erik Dahlén* and *Gunilla Hedlin*, thank you for giving me the opportunity to complete my thesis within the interdisciplinary Cfa network and for fruitful and inspiring discussions on allergy research. Thank you *Gunilla Jakobsson Ekman* and *Marianne Heibert-Arnlind* for your commitment and support in the Cfa Phd-programme.

My close collaborators within the genetic projects who have contributed invaluably with their expertise knowledge: *Cecilia Lindgren* for inspiring talks, your deep knowledge in genetics and efforts in facilitating the laboratory workflow. *Fredrik Nyberg* for fruitful discussions on genetic epidemiology and your constructive feed-back. *Marco Zucchelli* for all co-operation and your crucial assistance in statistical genetics.

I am very thankful to my friends and colleagues in the BAMSE group; *Catarina Almqvist, Maria Böhme, Lotta Egmar, Gunnel Emenius, Jenny Hallberg* (also Cfa), *Inger Kull, Eva Lannerö* and *Eva Östblom* (also Cfa). Thank you for inspiring discussions on outcomes, phenotypes and all kinds of risk factors. Thank you for your warm support and encouragement, especially *Catarina, Lotta* and *Inger,* during the critical periods of my thesis work. I hope we will continue to collaborate with the BAMSE project for many years. Thank you *Eva Hallner* and *Stina Gustavsson* for invaluable assistance with project matters and blood sample management. Thank you *André Lauber* for excellent computer assistance.

I would also like to express my gratitude to many colleagues at the Department of Occupational and Environmental Health: *Magnus Svartengren* for generously hosting me at the Department and for inspiring scientific discussions. *Niklas Berglind* for excellent statistical support together with *Magnus Lundqvist*. *Tom Bellander* and *Emma Nordling* for sharing your excellent knowledge in research on air pollutants. *Petter Ljungman, Erik Lampa, Mats Rosenlund, Malin Schollin, Helena Svensson* for valuable discussions and kind support.

Thank you all friends and colleagues at the Clinical Research Centre and Department of Biosciences at Novum – you have really taught me what genetic research is in real life. *Harriet Gullstén Viitalahti* for all assistance in the initial part of the genetic project. *Linda Berglind,*

Astrid Lindstedt and *Ville Mäkelä* for all assistance with the genotyping and DNA sample management – you have really done the work! *Sara Bruce* for nice collaboration and excellent analyses in paper III. My room mates *Ulf Hannelius* and *Anna Hellquist* and the other PhD-students *Heidi Anthoni, Christina Orsmark* and *Hanna Pettersson* for inspiring conversations about genetic research. I would also like to thank *Myriam Peyrard* and *Hannele Anthoni* for assistance with the DNA samples, *Päivi Kiviluoma* for database management and *Päivi Onkamo* for initial study design.

All Cfa PhD-students *Charlotte Admyre, Jennie Axén, Lydia Bennedich Kahn, Helen Flöistrup, Nina Gunnarsson, Pär Gyllfors, Susanna Jernelöv, Agne Liedén, Birgitta Marklund, Cecilia Moberg, Mika Nokela, Petter Olsson, Pirjo Savlin, Tanja Sobko and Pernilla Wiebert* for fruitful interdisciplinary discussions, which has given me a glimpse of completely new aspects of allergy-related diseases. I do appreciate our collaboration.

Thank you *Anna Bergström, Marianne van Hage* and *Annika Scheynius* for nice collaboration in the PARSIFAL project, *Ralf Morgenstern* for valuable discussions on antioxidant systems and *Sarah Umerkajeff* for your analyses in paper IV. Thank you also *Gert Doekes, Michael Kabesch, Tarja Laitinen, Roger Lauener* and *Josef Riedler* for nice collaboration in paper III.

Anders Lindfors, for inspiring me to become a paediatrician and for great experiences during a number of allergy congresses and *Bo Magnusson* and the Astrid Lindgren Children's Hospital for giving me the opportunity to complete my thesis during my residency in paediatrics.

Eva Undsén and *Ingrid Wikblad* for always being kind and helpful.

Göran Karlsson, thank you for arranging an elective research period on cockroach allergens in Charlottesville, USA 10 years ago and for initiating my research in allergy. I am grateful to *Martin Chapman* and *Thomas Platts-Mills* who made the visit in Charlottesville a true pleasure and for inspiring me to do further research in allergy.

All my invaluable *friends* outside work and *relatives*, for your interest and support.

My father *Ingemar*, for highlighting the importance of a proper phenotype definition and for your true interest and encouragement in my work. My mother *Birgitta*, for highlighting the importance that I see patients once in a while and for encouraging other important things in life. My brother *Jonas* and my sister *Karin* and your families for all your love and friendship. *Agneta, Nils* and *Anders* for your love and support. Finally, my lovely wife and best friend *Sophia*, thank you for standing by my side and always being there for me.

Sources of research support: The Swedish Environmental Protection Agency, The Swedish Asthma and Allergy Association, The Swedish Foundation for Health Care Science and Allergy Research, The Swedish Society of Medicine, Konsul Th.C. Bergh's Foundation, The Swedish Heart and Lung Foundation and Stockholm County Council.

REFERENCES

1. NHLBI/WHO Workshop Report: Global Strategy for Asthma Management and Prevention. 2003.
2. Bergmann RL, Edenharter G, Bergmann KE, Forster J, Bauer CP, Wahn V, et al. Atopic dermatitis in early infancy predicts allergic airway disease at 5 years. Clin Exp Allergy 1998; 28:965-70.
3. Koopman LP, Brunekreef B, de Jongste JC, Neijens HJ. Definition of respiratory symptoms and disease in early childhood in large prospective birth cohort studies that predict the development of asthma. Pediatr Allergy Immunol 2001; 12:118-24.
4. Martinez FD, Wright AL, Taussig LM, Holberg CJ, Halonen M, Morgan WJ. Asthma and wheezing in the first six years of life. The Group Health Medical Associates. N Engl J Med 1995; 332:133-8.
5. Taussig LM, Wright AL, Holberg CJ, Halonen M, Morgan WJ, Martinez FD. Tucson Children's Respiratory Study: 1980 to present. J Allergy Clin Immunol 2003; 111:661-75.
6. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee. Lancet 1998; 351:1225-32.
7. Lau S, Illi S, Sommerfeld C, Niggemann B, Bergmann R, von Mutius E, et al. Early exposure to house-dust mite and cat allergens and development of childhood asthma: a cohort study. Multicentre Allergy Study Group. Lancet 2000; 356:1392-7.
8. Perzanowski MS, Ronmark E, Platts-Mills TA, Lundback B. Effect of cat and dog ownership on sensitization and development of asthma among preteenage children. Am J Respir Crit Care Med 2002; 166:696-702.
9. Kull I, Almqvist C, Lilja G, Pershagen G, Wickman M. Breast-feeding reduces the risk of asthma during the first 4 years of life. J Allergy Clin Immunol 2004; 114:755-60.
10. Ronmark E, Jonsson E, Platts-Mills T, Lundback B. Incidence and remission of asthma in schoolchildren: report from the obstructive lung disease in northern Sweden studies. Pediatrics 2001; 107:E37.
11. Anderson HR, Ruggles R, Strachan DP, Austin JB, Burr M, Jeffs D, et al. Trends in prevalence of symptoms of asthma, hay fever, and eczema in 12-14 year olds in the British Isles, 1995-2002: questionnaire survey. Bmj 2004; 328:1052-3.
12. Braun-Fahrlander C, Gassner M, Grize L, Takken-Sahli K, Neu U, Stricker T, et al. No further increase in asthma, hay fever and atopic sensitisation in adolescents living in Switzerland. Eur Respir J 2004; 23:407-13.
13. Zollner IK, Weiland SK, Piechotowski I, Gabrio T, von Mutius E, Link B, et al. No increase in the prevalence of asthma, allergies, and atopic sensitisation among children in Germany: 1992-2001. Thorax 2005; 60:545-8.
14. Burr ML, Wat D, Evans C, Dunstan FD, Doull IJ. Asthma prevalence in 1973, 1988 and 2003. Thorax 2006; In press.
15. Lowe L, Custovic A, Woodcock A. Childhood asthma. Curr Allergy Asthma Rep 2004; 4:159-65.
16. Strachan DP. Hay fever, hygiene, and household size. Bmj 1989; 299:1259-60.
17. Martinez FD, Holt PG. Role of microbial burden in aetiology of allergy and asthma. Lancet 1999; 354 Suppl 2:SII12-5.
18. Ball TM, Castro-Rodriguez JA, Griffith KA, Holberg CJ, Martinez FD, Wright AL. Siblings, day-care attendance, and the risk of asthma and wheezing during childhood. N Engl J Med 2000; 343:538-43.
19. Illi S, von Mutius E, Lau S, Bergmann R, Niggemann B, Sommerfeld C, et al. Early childhood infectious diseases and the development of asthma up to school age: a birth cohort study. Bmj 2001; 322:390-5.
20. Braun-Fahrlander C, Gassner M, Grize L, Neu U, Sennhauser FH, Varonier HS, et al. Prevalence of hay fever and allergic sensitization in farmer's children and their peers living in the same rural community. SCARPOL team. Swiss Study on Childhood Allergy and Respiratory Symptoms with Respect to Air Pollution. Clin Exp Allergy 1999; 29:28-34.
21. Alfvén T, Braun-Fahrländer C, Brunekreef B, von Mutius E, Riedler J, Scheynius A, et al. Allergic diseases and atopic sensitization in children related to farming and anthroposophic lifestyle – the PARSIFAL study. Allergy 2006; In press.

22. Romagnani S. The increased prevalence of allergy and the hygiene hypothesis: missing immune deviation, reduced immune suppression, or both? Immunology 2004; 112:352-63.
23. Nicolai T. Air pollution and respiratory disease in children: what is the clinically relevant impact? Pediatr Pulmonol Suppl 1999; 18:9-13.
24. Pope CA, 3rd. Epidemiology of fine particulate air pollution and human health: biologic mechanisms and who's at risk? Environ Health Perspect 2000; 108 Suppl 4:713-23.
25. Wyler C, Braun-Fahrlander C, Kunzli N, Schindler C, Ackermann-Liebrich U, Perruchoud AP, et al. Exposure to motor vehicle traffic and allergic sensitization. The Swiss Study on Air Pollution and Lung Diseases in Adults (SAPALDIA) Team. Epidemiology 2000; 11:450-6.
26. Kramer U, Koch T, Ranft U, Ring J, Behrendt H. Traffic-related air pollution is associated with atopy in children living in urban areas. Epidemiology 2000; 11:64-70.
27. Janssen NA, Brunekreef B, van Vliet P, Aarts F, Meliefste K, Harssema H, et al. The relationship between air pollution from heavy traffic and allergic sensitization, bronchial hyperresponsiveness, and respiratory symptoms in Dutch schoolchildren. Environ Health Perspect 2003; 111:1512-8.
28. Penard-Morand C, Charpin D, Raherison C, Kopferschmitt C, Caillaud D, Lavaud F, et al. Long-term exposure to background air pollution related to respiratory and allergic health in schoolchildren. Clin Exp Allergy 2005; 35:1279-87.
29. Lowe L, Custovic A, Woodcock A. Childhood asthma. Curr Allergy Asthma Rep 2003; 3:109-14.
30. Holt PG, Jones CA. The development of the immune system during pregnancy and early life. Allergy 2000; 55:688-97.
31. Bousquet J, Jacot W, Yssel H, Vignola AM, Humbert M. Epigenetic inheritance of fetal genes in allergic asthma. Allergy 2004; 59:138-47.
32. Palmer LJ, Rye PJ, Gibson NA, Burton PR, Landau LI, Lesouef PN. Airway responsiveness in early infancy predicts asthma, lung function, and respiratory symptoms by school age. Am J Respir Crit Care Med 2001; 163:37-42.
33. Tager IB, Ngo L, Hanrahan JP. Maternal smoking during pregnancy. Effects on lung function during the first 18 months of life. Am J Respir Crit Care Med 1995; 152:977-83.
34. Gilliland FD, Berhane K, Li YF, Rappaport EB, Peters JM. Effects of early onset asthma and in utero exposure to maternal smoking on childhood lung function. Am J Respir Crit Care Med 2003; 167:917-24.
35. Lau S, Illi S, Sommerfeld C, Niggemann B, Volkel K, Madloch C, et al. Transient early wheeze is not associated with impaired lung function in 7-yr-old children. Eur Respir J 2003; 21:834-41.
36. 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.
37. Sears MR, Greene JM, Willan AR, Wiecek EM, Taylor DR, Flannery EM, et al. A longitudinal, population-based, cohort study of childhood asthma followed to adulthood. N Engl J Med 2003; 349:1414-22.
38. Morgan WJ, Stern DA, Sherrill DL, Guerra S, Holberg CJ, Guilbert TW, et al. Outcome of Asthma and Wheezing in the First 6 Years of Life: Follow-up through Adolescence. Am J Respir Crit Care Med 2005; 172:1253-8.
39. Lowe LA, Simpson A, Woodcock A, Morris J, Murray CS, Custovic A. Wheeze phenotypes and lung function in preschool children. Am J Respir Crit Care Med 2005; 171:231-7.
40. Kjellman B, Gustafsson PM. Asthma from childhood to adulthood: asthma severity, allergies, sensitization, living conditions, gender influence and social consequences. Respir Med 2000; 94:454-65.
41. Roorda RJ, Gerritsen J, Van Aalderen WM, Schouten JP, Veltman JC, Weiss ST, et al. Risk factors for the persistence of respiratory symptoms in childhood asthma. Am Rev Respir Dis 1993; 148:1490-5.
42. Vonk JM, Postma DS, Boezen HM, Grol MH, Schouten JP, Koeter GH, et al. Childhood factors associated with asthma remission after 30 year follow up. Thorax 2004; 59:925-9.
43. Panhuysen CI, Vonk JM, Koeter GH, Schouten JP, van Altena R, Bleecker ER, et al. Adult patients may outgrow their asthma: a 25-year follow-up study. Am J Respir Crit Care Med 1997; 155:1267-72.
44. Johansson SG, Bieber T, Dahl R, Friedmann PS, Lanier BQ, Lockey RF, et al. Revised nomenclature for allergy for global use: Report of the Nomenclature Review Committee of the World Allergy Organization, October 2003. J Allergy Clin Immunol 2004; 113:832-6.

45. Wickman M, Ahlstedt S, Lilja G, van Hage Hamsten M. Quantification of IgE antibodies simplifies the classification of allergic diseases in 4-year-old children. A report from the prospective birth cohort study--BAMSE. Pediatr Allergy Immunol 2003; 14:441-7.
46. Illi S, von Mutius E, Lau S, Nickel R, Niggemann B, Sommerfeld C, et al. The pattern of atopic sensitization is associated with the development of asthma in childhood. J Allergy Clin Immunol 2001; 108:709-14.
47. Jarvis D, Luczynska C, Chinn S, Potts J, Sunyer J, Janson C, et al. Change in prevalence of IgE sensitization and mean total IgE with age and cohort. J Allergy Clin Immunol 2005; 116:675-82.
48. Lowe L, Murray CS, Custovic A, Simpson BM, Kissen PM, Woodcock A. Specific airway resistance in 3-year-old children: a prospective cohort study. Lancet 2002; 359:1904-8.
49. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature 2001; 409:860-921.
50. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science 2001; 291:1304-51.
51. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 1953; 171:737-8.
52. Strachan T, Read A. Human Molecular Genetics. Second ed. Oxford: BIOS Scientific Publishers Ltd; 1999.
53. Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 2001; 409:928-33.
54. Reich DE, Schaffner SF, Daly MJ, McVean G, Mullikin JC, Higgins JM, et al. Human genome sequence variation and the influence of gene history, mutation and recombination. Nat Genet 2002; 32:135-42.
55. McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, Barrett JC, et al. Common deletion polymorphisms in the human genome. Nat Genet 2006; 38:86-92.
56. Reich DE, Lander ES. On the allelic spectrum of human disease. Trends Genet 2001; 17:502-10.
57. Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl MC, Nemesh J, et al. The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. Nat Genet 2000; 26:76-80.
58. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 1993; 261:921-3.
59. Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? Am J Hum Genet 2001; 69:124-37.
60. Los H, Koppelman GH, Postma DS. The importance of genetic influences in asthma. Eur Respir J 1999; 14:1210-27.
61. Wall JD, Pritchard JK. Haplotype blocks and linkage disequilibrium in the human genome. Nat Rev Genet 2003; 4:587-97.
62. Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, et al. Linkage disequilibrium in the human genome. Nature 2001; 411:199-204.
63. Crawford DC, Nickerson DA. Definition and clinical importance of haplotypes. Annu Rev Med 2005; 56:303-20.
64. Zaykin DV, Westfall PH, Young SS, Karnoub MA, Wagner MJ, Ehm MG. Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. Hum Hered 2002; 53:79-91.
65. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002; 70:425-34.
66. Clayton D. http://www-gene.cimr.cam.ac.uk/clayton/software/snphap.txt.
67. Excoffier L, Slatkin M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. Mol Biol Evol 1995; 12:921-7.
68. Fallin D, Schork NJ. Accuracy of haplotype frequency estimation for biallelic loci, via the expectation-maximization algorithm for unphased diploid genotype data. Am J Hum Genet 2000; 67:947-59.
69. Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES. High-resolution haplotype structure in the human genome. Nat Genet 2001; 29:229-32.
70. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. Science 2002; 296:2225-9.

71. Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. Nature 2005; 437:1299-320.
72. Wang N, Akey JM, Zhang K, Chakraborty R, Jin L. Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. Am J Hum Genet 2002; 71:1227-34.
73. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005; 21:263-5.
74. Nothnagel M, Rohde K. The effect of single-nucleotide polymorphism marker selection on patterns of haplotype blocks and haplotype frequency estimates. Am J Hum Genet 2005; 77:988-98.
75. Hershey GK, Friedrich MF, Esswein LA, Thomas ML, Chatila TA. The association of atopy with a gain-of-function mutation in the alpha subunit of the interleukin-4 receptor. N Engl J Med 1997; 337:1720-5.
76. Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, et al. Haplotype tagging for the identification of common disease genes. Nat Genet 2001; 29:233-7.
77. Carlson CS, Eberle MA, Kruglyak L, Nickerson DA. Mapping complex disease loci in whole-genome association studies. Nature 2004; 429:446-52.
78. Freidlin B, Zheng G, Li Z, Gastwirth JL. Trend tests for case-control studies of genetic markers: power, sample size and robustness. Hum Hered 2002; 53:146-52.
79. Slager SL, Schaid DJ. Case-control studies of genetic markers: power and sample size approximations for Armitage's test for trend. Hum Hered 2001; 52:149-53.
80. Chiano MN, Clayton DG. Genotypic relative risks under ordered restriction. Genet Epidemiol 1998; 15:135-46.
81. Koeppen-Schomerus G, Stevenson J, Plomin R. Genes and environment in asthma: a study of 4 year old twins. Arch Dis Child 2001; 85:398-400.
82. Duffy DL, Martin NG, Battistutta D, Hopper JL, Mathews JD. Genetics of asthma and hay fever in Australian twins. Am Rev Respir Dis 1990; 142:1351-8.
83. Daniels SE, Bhattacharrya S, James A, Leaves NI, Young A, Hill MR, et al. A genome-wide search for quantitative trait loci underlying asthma. Nature 1996; 383:247-50.
84. Van Eerdewegh P, Little RD, Dupuis J, Del Mastro RG, Falls K, Simon J, et al. Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. Nature 2002; 418:426-30.
85. Allen M, Heinzmann A, Noguchi E, Abecasis G, Broxholme J, Ponting CP, et al. Positional cloning of a novel gene influencing asthma from chromosome 2q14. Nat Genet 2003; 35:258-63.
86. Zhang Y, Leaves NI, Anderson GG, Ponting CP, Broxholme J, Holt R, et al. Positional cloning of a quantitative trait locus on chromosome 13q14 that influences immunoglobulin E levels and asthma. Nat Genet 2003; 34:181-6.
87. Nicolae D, Cox NJ, Lester LA, Schneider D, Tan Z, Billstrand C, et al. Fine mapping and positional candidate studies identify HLA-G as an asthma susceptibility gene on chromosome 6p21. Am J Hum Genet 2005; 76:349-57.
88. Laitinen T, Polvi A, Rydman P, Vendelin J, Pulkkinen V, Salmikangas P, et al. Characterization of a common susceptibility locus for asthma-related traits. Science 2004; 304:300-4.
89. Noguchi E, Yokouchi Y, Zhang J, Shibuya K, Shibuya A, Bannai M, et al. Positional identification of an asthma susceptibility gene on human chromosome 5q33. Am J Respir Crit Care Med 2005; 172:183-8.
90. Ober C, Hoffjan S. Asthma genetics 2006: the long and winding road to gene discovery. Genes Immun 2006; In press.
91. Palmer LJ, Burton PR, Faux JA, James AL, Musk AW, Cookson WO. Independent inheritance of serum immunoglobulin E concentrations and airway responsiveness. Am J Respir Crit Care Med 2000; 161:1836-43.
92. Ciccodicola A, D'Esposito M, Esposito T, Gianfrancesco F, Migliaccio C, Miano MG, et al. Differentially regulated and evolved genes in the fully sequenced Xq/Yq pseudoautosomal region. Hum Mol Genet 2000; 9:395-401.
93. Vermeesch JR, Petit P, Kermouni A, Renauld JC, Van Den Berghe H, Marynen P. The IL-9 receptor gene, located in the Xq/Yq pseudoautosomal region, has an autosomal origin, escapes X inactivation and is expressed from the Y. Hum Mol Genet 1997; 6:1-8.
94. Holroyd KJ, Martinati LC, Trabetti E, Scherpbier T, Eleff SM, Boner AL, et al. Asthma and bronchial hyperresponsiveness linked to the XY long arm pseudoautosomal region. Genomics 1998; 52:233-5.

95. Kauppi P, Laitinen T, Ollikainen V, Mannila H, Laitinen LA, Kere J. The IL9R region contribution in asthma is supported by genetic association in an isolated population. Eur J Hum Genet 2000; 8:788-92.
96. Renauld JC, Druez C, Kermouni A, Houssiau F, Uyttenhove C, Van Roost E, et al. Expression cloning of the murine and human interleukin 9 receptor cDNAs. Proc Natl Acad Sci U S A 1992; 89:5690-4.
97. Abdelilah S, Latifa K, Esra N, Cameron L, Bouchaib L, Nicolaides N, et al. Functional expression of IL-9 receptor by human neutrophils from asthmatic donors: role in IL-8 release. J Immunol 2001; 166:2768-74.
98. Demoulin JB, Renauld JC. Interleukin 9 and its receptor: an overview of structure and function. Int Rev Immunol 1998; 16:345-64.
99. De Smedt M, Verhasselt B, Kerre T, Vanhecke D, Naessens E, Leclercq G, et al. Signals from the IL-9 receptor are critical for the early stages of human intrathymic T cell development. J Immunol 2000; 164:1761-7.
100. Bauer JH, Liu KD, You Y, Lai SY, Goldsmith MA. Heteromerization of the gamma chain with the interleukin-9 receptor alpha subunit leads to STAT activation and prevention of apoptosis. J Biol Chem 1998; 273:9255-60.
101. Demoulin JB, Uyttenhove C, Van Roost E, DeLestre B, Donckers D, Van Snick J, et al. A single tyrosine of the interleukin-9 (IL-9) receptor is required for STAT activation, antiapoptotic activity, and growth regulation by IL-9. Mol Cell Biol 1996; 16:4710-6.
102. Dugas B, Renauld JC, Pene J, Bonnefoy JY, Peti-Frere C, Braquet P, et al. Interleukin-9 potentiates the interleukin-4-induced immunoglobulin (IgG, IgM and IgE) production by normal human B lymphocytes. Eur J Immunol 1993; 23:1687-92.
103. Soussi-Gounni A, Kontolemos M, Hamid Q. Role of IL-9 in the pathophysiology of allergic diseases. J Allergy Clin Immunol 2001; 107:575-82.
104. Chatila TA. Interleukin-4 receptor signaling pathways in asthma pathogenesis. Trends Mol Med 2004; 10:493-9.
105. Howard TD, Koppelman GH, Xu J, Zheng SL, Postma DS, Meyers DA, et al. Gene-gene interaction in asthma: IL4RA and IL13 in a Dutch population with asthma. Am J Hum Genet 2002; 70:230-6.
106. Lee SG, Kim BS, Kim JH, Lee SY, Choi SO, Shim JY, et al. Gene-gene interaction between interleukin-4 and interleukin-4 receptor alpha in Korean children with asthma. Clin Exp Allergy 2004; 34:1202-8.
107. Liu X, Beaty TH, Deindl P, Huang SK, Lau S, Sommerfeld C, et al. Associations between specific serum IgE response and 6 variants within the genes IL4, IL13, and IL4RA in German children: The German Multicenter Atopy Study. J Allergy Clin Immunol 2004; 113:489-95.
108. Kabesch M, Schedel M, Carr D, Woitsch B, Fritzsch C, Weiland SK, et al. IL-4/IL-13 pathway genetics strongly influence serum IgE levels and childhood asthma. J Allergy Clin Immunol 2006; 117:269-74.
109. Hoebee B, Bont L, Rietveld E, van Oosten M, Hodemaekers HM, Nagelkerke NJ, et al. Influence of promoter variants of interleukin-10, interleukin-9, and tumor necrosis factor-alpha genes on respiratory syncytial virus bronchiolitis. J Infect Dis 2004; 189:239-47.
110. Hoffjan S, Nicolae D, Ostrovnaya I, Roberg K, Evans M, Mirel DB, et al. Gene-environment interaction effects on the development of immune responses in the 1st year of life. Am J Hum Genet 2005; 76:696-704.
111. Xu YL, Reinscheid RK, Huitron-Resendiz S, Clark SD, Wang Z, Lin SH, et al. Neuropeptide S: a neuropeptide promoting arousal and anxiolytic-like effects. Neuron 2004; 43:487-97.
112. Reinscheid RK, Xu YL, Okamura N, Zeng J, Chung S, Pai R, et al. Pharmacological characterization of human and murine neuropeptide s receptor variants. J Pharmacol Exp Ther 2005; 315:1338-45.
113. Vendelin J, Pulkkinen V, Rehn M, Pirskanen A, Raisanen-Sokolowski A, Laitinen A, et al. Characterization of GPRA, a novel G protein-coupled receptor related to asthma. Am J Respir Cell Mol Biol 2005; 33:262-70.
114. Kormann MS, Carr D, Klopp N, Illig T, Leupold W, Fritzsch C, et al. G-Protein-coupled Receptor Polymorphisms Are Associated with Asthma in a Large German Population. Am J Respir Crit Care Med 2005; 171:1358-62.
115. Soderhall C, Marenholz I, Nickel R, Gruber C, Kehrt R, Rohde K, et al. Lack of association of the G protein-coupled receptor for asthma susceptibility gene with atopic dermatitis. J Allergy Clin Immunol 2005; 116:220-1.

116. Veal CD, Reynolds NJ, Meggitt SJ, Allen MH, Lindgren CM, Kere J, et al. Absence of association between asthma and high serum immunoglobulin E associated GPRA haplotypes and adult atopic dermatitis. J Invest Dermatol 2005; 125:399-401.
117. Wills-Karp M, Ewart SL. Time to draw breath: asthma-susceptibility genes are identified. Nat Rev Genet 2004; 5:376-87.
118. Contopoulos-Ioannidis DG, Manoli EN, Ioannidis JP. Meta-analysis of the association of beta2-adrenergic receptor polymorphisms with asthma phenotypes. J Allergy Clin Immunol 2005; 115:963-72.
119. Israel E, Chinchilli VM, Ford JG, Boushey HA, Cherniack R, Craig TJ, et al. Use of regularly scheduled albuterol treatment in asthma: genotype-stratified, randomised, placebo-controlled cross-over trial. Lancet 2004; 364:1505-12.
120. Lee DK, Currie GP, Hall IP, Lima JJ, Lipworth BJ. The arginine-16 beta2-adrenoceptor polymorphism predisposes to bronchoprotective subsensitivity in patients treated with formoterol and salmeterol. Br J Clin Pharmacol 2004; 57:68-75.
121. Wang Z, Chen C, Niu T, Wu D, Yang J, Wang B, et al. Association of asthma with beta(2)-adrenergic receptor gene polymorphism and cigarette smoking. Am J Respir Crit Care Med 2001; 163:1404-9.
122. Randolph AG, Lange C, Silverman EK, Lazarus R, Weiss ST. Extended haplotype in the tumor necrosis factor gene cluster is associated with asthma and asthma-related phenotypes. Am J Respir Crit Care Med 2005; 172:687-92.
123. Wang TN, Chen WY, Wang TH, Chen CJ, Huang LY, Ko YC. Gene-gene synergistic effect on atopic asthma: tumour necrosis factor-alpha-308 and lymphotoxin-alpha-NcoI in Taiwan's children. Clin Exp Allergy 2004; 34:184-8.
124. Messer G, Spengler U, Jung MC, Honold G, Blomer K, Pape GR, et al. Polymorphic structure of the tumor necrosis factor (TNF) locus: an NcoI polymorphism in the first intron of the human TNF-beta gene correlates with a variant amino acid in position 26 and a reduced level of TNF-beta production. J Exp Med 1991; 173:209-19.
125. 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; 94:3195-9.
126. Yang IA, Holz O, Jorres RA, Magnussen H, Barton SJ, Rodriguez S, et al. Association of tumor necrosis factor-alpha polymorphisms and ozone-induced change in lung function. Am J Respir Crit Care Med 2005; 171:171-6.
127. 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; 172:1563-8.
128. Hayes JD, Strange RC. Potential contribution of the glutathione S-transferase supergene family to resistance to oxidative stress. Free Radic Res 1995; 22:193-207.
129. Strange RC, Spiteri MA, Ramachandran S, Fryer AA. Glutathione-S-transferase family of enzymes. Mutat Res 2001; 482:21-6.
130. Fryer AA, Hume R, Strange RC. The development of glutathione S-transferase and glutathione peroxidase activities in human lung. Biochim Biophys Acta 1986; 883:448-53.
131. Gilliland FD, Li YF, Saxon A, Diaz-Sanchez D. Effect of glutathione-S-transferase M1 and P1 genotypes on xenobiotic enhancement of allergic responses: randomised, placebo-controlled crossover study. Lancet 2004; 363:119-25.
132. 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; 34:1707-13.
133. 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; 166:457-63.
134. 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; 59:569-73.
135. Lannero E, Kull I, Wickman M, Pershagen G, Nordvall SL. Environmental risk factors for allergy and socioeconomic status in a birth cohort (BAMSE). Pediatr Allergy Immunol 2002; 13:182-7.
136. 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.

137. Asher MI, Keil U, Anderson HR, Beasley R, Crane J, Martinez F, et al. International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. Eur Respir J 1995; 8:483-91.
138. Riedler J, Braun-Fahrlander C, Eder W, Schreuer M, Waser M, Maisch S, et al. Exposure to farming in early life and development of asthma and allergy: a cross-sectional survey. Lancet 2001; 358:1129-33.
139. Alm JS, Swartz J, Lilja G, Scheynius A, Pershagen G. Atopy in children of families with an anthroposophic lifestyle. Lancet 1999; 353:1485-8.
140. 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; 109:633-9.
141. Hallberg J, Anderson M, Wickman M, Svartengren M. Sex Influences on Lung Function and Medication in Childhood Asthma. Acta Paediatr 2006; In press.
142. Hosking L, Lumsden S, Lewis K, Yeo A, McCarthy L, Bansal A, et al. Detection of genotyping errors by Hardy-Weinberg equilibrium testing. Eur J Hum Genet 2004; 12:395-9.
143. Nakahori Y, Takenaka O, Nakagome Y. A human X-Y homologous region encodes "amelogenin". Genomics 1991; 9:264-9.
144. Sullivan KM, Mannucci A, Kimpton CP, Gill P. A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. Biotechniques 1993; 15:636-8, 40-1.
145. Steinlechner M, Berger B, Niederstatter H, Parson W. Rare failures in the amelogenin sex test. Int J Legal Med 2002; 116:117-20.
146. Michael A, Brauner P. Erroneous gender identification by the amelogenin sex test. J Forensic Sci 2004; 49:258-9.
147. R Development Core Team. R: A language and environment for statistical computing.: R Foundation for Statistical Computing, Vienna, Austria, 2004.
148. Cordell HJ. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. Hum Mol Genet 2002; 11:2463-8.
149. Ahlbom A, Alfredsson L. Interaction: A word with two meanings creates confusion. Eur J Epidemiol 2005; 20:563-4.
150. Rothman K, Greenland S. Modern Epidemiology: Philadelphia: Lippincott Williams and Wilkins; 1998.
151. Cordell HJ, Clayton DG. A unified stepwise regression procedure for evaluating the relative effects of polymorphisms within a gene using case/control or family data: application to HLA in type 1 diabetes. Am J Hum Genet 2002; 70:124-41.
152. Millstein J CD, Gilliland FD, Gauderman WJ. A testing framework for identifying suseptibility genes in the presence of epistasis. Am J Hum Genet 2006; 78:15-27.
153. Moore JH, Williams SM. New strategies for identifying gene-gene interactions in hypertension. Ann Med 2002; 34:88-95.
154. Schaid DJ. General score tests for associations of genetic markers with disease using cases and their parents. Genet Epidemiol 1996; 13:423-49.
155. Choudhry S, Avila PC, Nazario S, Ung N, Kho J, Rodriguez-Santana JR, et al. CD14 tobacco gene-environment interaction modifies asthma severity and immunoglobulin E levels in Latinos with asthma. Am J Respir Crit Care Med 2005; 172:173-82.
156. Eder W, Klimecki W, Yu L, von Mutius E, Riedler J, Braun-Fahrlander C, et al. Opposite effects of CD 14/-260 on serum IgE levels in children raised in different environments. J Allergy Clin Immunol 2005; 116:601-7.
157. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 2004; 96:434-42.
158. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 1995; 57:289-300.
159. Millstein J, Conti DV, Gilliland FD, Gauderman WJ. A testing framework for identifying susceptibility genes in the presence of epistasis. Am J Hum Genet 2006; 78:15-27.
160. Lange C, DeMeo D, Silverman EK, Weiss ST, Laird NM. Using the noninformative families in family-based association tests: a powerful new testing strategy. Am J Hum Genet 2003; 73:801-11.
161. Lange C, Lyon H, DeMeo D, Raby B, Silverman EK, Weiss ST. A new powerful non-parametric two-stage approach for testing multiple phenotypes in family-based association studies. Hum Hered 2003; 56:10-7.

162. Jaakkola JJ, Nafstad P, Magnus P. Environmental tobacco smoke, parental atopy, and childhood asthma. Environ Health Perspect 2001; 109:579-82.
163. Litonjua AA, Carey VJ, Burge HA, Weiss ST, Gold DR. Parental history and the risk for childhood asthma. Does mother confer more risk than father? Am J Respir Crit Care Med 1998; 158:176-81.
164. London SJ, James Gauderman W, Avol E, Rappaport EB, Peters JM. Family history and the risk of early-onset persistent, early-onset transient, and late-onset asthma. Epidemiology 2001; 12:577-83.
165. Raby BA, Van Steen K, Celedon JC, Litonjua AA, Lange C, Weiss ST. Paternal history of asthma and airway responsiveness in children with asthma. Am J Respir Crit Care Med 2005; 172:552-8.
166. Bergmann RL, Schulz J, Gunther S, Dudenhausen JW, Bergmann KE, Bauer CP, et al. Determinants of cord-blood IgE concentrations in 6401 German neonates. Allergy 1995; 50:65-71.
167. Leaves NI, Bhattacharyya S, Wiltshire S, Cookson WO. A detailed genetic map of the chromosome 7 bronchial hyper-responsiveness locus. Eur J Hum Genet 2002; 10:177-82.
168. Lee YA, Wahn U, Kehrt R, Tarani L, Businco L, Gustafsson D, et al. A major susceptibility locus for atopic dermatitis maps to chromosome 3q21. Nat Genet 2000; 26:470-3.
169. Strauch K, Bogdanow M, Fimmers R, Baur MP, Wienker TF. Linkage analysis of asthma and atopy including models with genomic imprinting. Genet Epidemiol 2001; 21 Suppl 1:S204-9.
170. Huxtable SJ, Saker PJ, Haddad L, Walker M, Frayling TM, Levy JC, et al. Analysis of parent-offspring trios provides evidence for linkage and association between the insulin gene and type 2 diabetes mediated exclusively through paternally transmitted class III variable number tandem repeat alleles. Diabetes 2000; 49:126-30.
171. Anderson HR, Pottier AC, Strachan DP. Asthma from birth to age 23: incidence and relation to prior and concurrent atopic disease. Thorax 1992; 47:537-42.
172. Skobeloff EM, Spivey WH, St Clair SS, Schoffstall JM. The influence of age and sex on asthma admissions. JAMA 1992; 268:3437-40.
173. Weitzman M, Gortmaker S, Walker DK, Sobol A. Maternal smoking and childhood asthma. Pediatrics 1990; 85:505-11.
174. Johnson CC, Peterson EL, Ownby DR. Gender differences in total and allergen-specific immunoglobulin E (IgE) concentrations in a population-based cohort from birth to age four years. Am J Epidemiol 1998; 147:1145-52.
175. Sears MR, Burrows B, Flannery EM, Herbison GP, Holdaway MD. Atopy in childhood. I. Gender and allergen related risks for development of hay fever and asthma. Clin Exp Allergy 1993; 23:941-8.
176. Kerkhof M, Wijga A, Smit HA, de Jongste JC, Aalberse RC, Brunekreef B, et al. The effect of prenatal exposure on total IgE at birth and sensitization at twelve months and four years of age: The PIAMA study. Pediatr Allergy Immunol 2005; 16:10-8.
177. Young S, Sherrill DL, Arnott J, Diepeveen D, LeSouef PN, Landau LI. Parental factors affecting respiratory function during the first year of life. Pediatr Pulmonol 2000; 29:331-40.
178. Kwon HL, Belanger K, Holford TR, Bracken MB. Effect of fetal sex on airway lability in pregnant women with asthma. Am J Epidemiol 2006; 163:217-21.
179. Murphy VE, Gibson PG, Giles WB, Zakar T, Smith R, Bisits AM, et al. Maternal asthma is associated with reduced female fetal growth. Am J Respir Crit Care Med 2003; 168:1317-23.
180. Bernsen RM, de Jongste JC, Koes BW, Aardoom HA, van der Wouden JC. Perinatal characteristics and obstetric complications as risk factors for asthma, allergy and eczema at the age of 6 years. Clin Exp Allergy 2005; 35:1135-40.
181. Beasley R, Leadbitter P, Pearce N, Crane J. Is enhanced fetal growth a risk factor for the development of atopy or asthma? Int Arch Allergy Immunol 1999; 118:408-10.
182. Rinn JL, Snyder M. Sexual dimorphism in mammalian gene expression. Trends Genet 2005; 21:298-305.
183. Weiss LA, Pan L, Abney M, Ober C. The sex-specific genetic architecture of quantitative traits in humans. Nat Genet 2006; 38:218-22.
184. Srivastava A, Cotton C, Paterson AD. Sex- and age-of-onset-based locus heterogeneity in asthma. Genet Epidemiol 2001; 21 Suppl 1:S154-9.
185. Towne B, Siervogel RM, Blangero J. Effects of genotype-by-sex interaction on quantitative trait linkage analysis. Genet Epidemiol 1997; 14:1053-8.
186. Karjalainen J, Nieminen MM, Aromaa A, Klaukka T, Hurme M. The IL-1beta genotype carries asthma susceptibility only in men. J Allergy Clin Immunol 2002; 109:514-6.

187. Raby BA, Lazarus R, Silverman EK, Lake S, Lange C, Wjst M, et al. Association of vitamin D receptor gene polymorphisms with childhood and adult asthma. Am J Respir Crit Care Med 2004; 170:1057-65.
188. Szczeklik W, Sanak M, Szczeklik A. Functional effects and gender association of COX-2 gene polymorphism G-765C in bronchial asthma. J Allergy Clin Immunol 2004; 114:248-53.
189. Ulbrecht M, Hergeth MT, Wjst M, Heinrich J, Bickeboller H, Wichmann HE, et al. Association of beta(2)-adrenoreceptor variants with bronchial hyperresponsiveness. Am J Respir Crit Care Med 2000; 161:469-74.
190. Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 2005; 434:400-4.
191. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. Nat Genet 2001; 29:306-9.
192. Weiss ST, Raby BA. Asthma genetics 2003. Hum Mol Genet 2004; 13 Spec No 1:R83-9.
193. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 1995; 11:241-7.
194. Blakey J, Halapi E, Bjornsdottir US, Wheatley A, Kristinsson S, Upmanyu R, et al. Contribution of ADAM33 polymorphisms to the population risk of asthma. Thorax 2005; 60:274-6.
195. Shin HD, Park KS, Park CS. Lack of association of GPRA (G protein-coupled receptor for asthma susceptibility) haplotypes with high serum IgE or asthma in a Korean population. J Allergy Clin Immunol 2004; 114:1226-7.
196. Neale BM, Sham PC. The future of association studies: gene-based analysis and replication. Am J Hum Genet 2004; 75:353-62.
197. Van Steen K, McQueen MB, Herbert A, Raby B, Lyon H, Demeo DL, et al. Genomic screening and replication using the same data set in family-based association testing. Nat Genet 2005; 37:683-91.
198. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 2005; 6:95-108.
199. Matsuzaki H, Dong S, Loi H, Di X, Liu G, Hubbell E, et al. Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. Nat Methods 2004; 1:109-11.
200. Wang WY, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. Nat Rev Genet 2005; 6:109-18.
201. Pershagen G. Challenges in epidemiologic allergy research. Allergy 1997; 52:1045-9.
202. Celedon JC, Litonjua AA, Ryan L, Weiss ST, Gold DR. Lack of association between antibiotic use in the first year of life and asthma, allergic rhinitis, or eczema at age 5 years. Am J Respir Crit Care Med 2002; 166:72-5.
203. Jenkins HA, Cherniack R, Szefler SJ, Covar R, Gelfand EW, Spahn JD. A comparison of the clinical characteristics of children and adults with severe asthma. Chest 2003; 124:1318-24.
204. Mandhane PJ, Greene JM, Cowan JO, Taylor DR, Sears MR. Sex differences in factors associated with childhood- and adolescent-onset wheeze. Am J Respir Crit Care Med 2005; 172:45-54.
205. Ekelund E, Saaf A, Tengvall-Linder M, Link J, Melen E, Wickman M, et al. Elevated Expression and Genetic Association Links the SOCS3 Gene to Atopic Dermatitis. Swedish Medical Association, Annual Meeting 2005; Abstract.
206. Hunter DJ. Gene-environment interactions in human diseases. Nat Rev Genet 2005; 6:287-98.
207. Skrondal A. Interaction as departure from additivity in case-control studies: a cautionary note. Am J Epidemiol 2003; 158:251-8.
208. Darroch J. Biologic synergism and parallelism. Am J Epidemiol 1997; 145:661-8.
209. Thornton-Wells TA, Moore JH, Haines JL. Genetics, statistics and human disease: analytical retooling for complexity. Trends Genet 2004; 20:640-7.
210. Moore JH. A global view of epistasis. Nat Genet 2005; 37:13-4.
211. Marchini J, Donnelly P, Cardon LR. Genome-wide strategies for detecting multiple loci that influence complex diseases. Nat Genet 2005; 37:413-7.
212. Moore JH. The ubiquitous nature of epistasis in determining susceptibility to common human diseases. Hum Hered 2003; 56:73-82.
213. Hoebee B, Rietveld E, Bont L, Oosten M, Hodemaekers HM, Nagelkerke NJ, et al. Association of severe respiratory syncytial virus bronchiolitis with interleukin-4 and interleukin-4 receptor alpha polymorphisms. J Infect Dis 2003; 187:2-11.
214. McNamara PS, Flanagan BF, Baldwin LM, Newland P, Hart CA, Smyth RL. Interleukin 9 production in the lungs of infants with severe respiratory syncytial virus bronchiolitis. Lancet 2004; 363:1031-7.

215. Ritchie MD, Hahn LW, Roodi N, Bailey LR, Dupont WD, Parl FF, et al. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. Am J Hum Genet 2001; 69:138-47.
216. Cook NR, Zee RY, Ridker PM. Tree and spline based association analysis of gene-gene interaction models for ischemic stroke. Stat Med 2004; 23:1439-53.
217. Becker AB, Chan-Yeung M. Primary prevention of asthma. Curr Opin Pulm Med 2002; 8:16-24.
218. Zambelli-Weiner A, Ehrlich E, Stockton ML, Grant AV, Zhang S, Levett PN, et al. Evaluation of the CD14/-260 polymorphism and house dust endotoxin exposure in the Barbados Asthma Genetics Study. J Allergy Clin Immunol 2005; 115:1203-9.
219. Colilla S, Nicolae D, Pluzhnikov A, Blumenthal MN, Beaty TH, Bleecker ER, et al. Evidence for gene-environment interactions in a linkage study of asthma and smoking exposure. J Allergy Clin Immunol 2003; 111:840-6.
220. Meyers DA, Postma DS, Stine OC, Koppelman GH, Ampleford EJ, Jongepier H, et al. Genome screen for asthma and bronchial hyperresponsiveness: interactions with passive smoke exposure. J Allergy Clin Immunol 2005; 115:1169-75.
221. Kelada SN, Stapleton PL, Farin FM, Bammler TK, Eaton DL, Smith-Weller T, et al. Glutathione S-transferase M1, T1, and P1 polymorphisms and Parkinson's disease. Neurosci Lett 2003; 337:5-8.
222. Menegon A, Board PG, Blackburn AC, Mellick GD, Le Couteur DG. Parkinson's disease, pesticides, and glutathione transferase polymorphisms. Lancet 1998; 352:1344-6.
223. Allan JM, Wild CP, Rollinson S, Willett EV, Moorman AV, Dovey GJ, et al. Polymorphism in glutathione S-transferase P1 is associated with susceptibility to chemotherapy-induced leukemia. Proc Natl Acad Sci U S A 2001; 98:11592-7.
224. Kimura S, Imagawa Y, Satake K, Tsukuda M. The relationship of the human glutathione S-transferase P1 polymorphism and chemotherapeutic sensitivity in head and neck squamous carcinoma. Int J Mol Med 2004; 14:185-9.
225. Adler V, Pincus MR. Effector peptides from glutathione-S-transferase-pi affect the activation of jun by jun-N-terminal kinase. Ann Clin Lab Sci 2004; 34:35-46.
226. Adler V, Yin Z, Fuchs SY, Benezra M, Rosario L, Tew KD, et al. Regulation of JNK signaling by GSTp. Embo J 1999; 18:1321-34.
227. Manevich Y, Feinstein SI, Fisher AB. Activation of the antioxidant enzyme 1-CYS peroxiredoxin requires glutathionylation mediated by heterodimerization with pi GST. Proc Natl Acad Sci U S A 2004; 101:3780-5.
228. Ralat LA, Manevich Y, Fisher AB, Colman RF. Direct Evidence for the Formation of a Complex between 1-Cysteine Peroxiredoxin and Glutathione S-Transferase pi with Activity Changes in Both Enzymes. Biochemistry 2006; 45:360-72.
229. Bennett BL, Sasaki DT, Murray BW, O'Leary EC, Sakata ST, Xu W, et al. SP600125, an anthrapyrazolone inhibitor of Jun N-terminal kinase. Proc Natl Acad Sci U S A 2001; 98:13681-6.
230. Hu X, Singh SV. Differential catalytic efficiency and enantioselectivity of murine glutathione S-transferase isoenzymes in the glutathione conjugation of carcinogenic anti-diol epoxides of chrysene and benzo(g)chrysene. Arch Biochem Biophys 1997; 345:318-24.
231. Sundberg K, Johansson AS, Stenberg G, Widersten M, Seidel A, Mannervik B, et al. Differences in the catalytic efficiencies of allelic variants of glutathione transferase P1-1 towards carcinogenic diol epoxides of polycyclic aromatic hydrocarbons. Carcinogenesis 1998; 19:433-6.
232. Watson MA, Stewart RK, Smith GB, Massey TE, Bell DA. Human glutathione S-transferase P1 polymorphisms: relationship to lung tissue enzyme activity and population frequency distribution. Carcinogenesis 1998; 19:275-80.
233. Cardon LR, Bell JI. Association study designs for complex diseases. Nat Rev Genet 2001; 2:91-9.
234. Wacholder S, Rothman N, Caporaso N. Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. J Natl Cancer Inst 2000; 92:1151-8.
235. Devlin B, Roeder K. Genomic control for association studies. Biometrics 1999; 55:997-1004.