From DEPARTMENT OF BIOSCIENCES AND NUTRITION
Karolinska Institutet, Stockholm, Sweden

INTERACTION AND REGULATION OF ASTHMA SUSCEPTIBILITY GENES

Christina Orsmark Pietras

Karolinska Institutet

Stockholm 2011

Published by Karolinska Institutet. Printed by Repro Print AB

ISBN 978-91-7457-308-4

ABSTRACT

Asthma is a disorder characterized by symptoms such as wheezing, shortness of breath, chest tightness or coughing. It is a chronic inflammation in the airways and the inflammation is usually accompanied by limitations in airflow as a result of hyper-secretion of mucus and broncho-constriction. Asthma commonly coincides with other allergic diseases such as allergic sensitization and rhinoconjunctivitis. The prevalence of asthma and allergy in children is highest in affluent countries with up to 20% in English speaking countries. Asthma and allergic disease are complex disorders and have long been known to be influenced by both heritable components and environmental factors.

The overall aim with this thesis was to investigate asthma susceptibility genes and their genetic role, biological dependency, as well as how they interact in a context-dependent manner, either with other genes (I) or with environmental factors (II). We studied the functional difference between splice variants of a previously identified asthma susceptibility gene showing unique expression patterns in asthmatic patients (III). We also aimed to define global gene expression patterns in asthmatic children that could reveal novel insight about characteristics of severe therapy-resistant asthma in children (IV).

In study I, we examined the biologically linked asthma susceptibility gene Tenascin C (*TNC*) and its genetic role in asthma and allergy. We also investigated the biological and genetic interactions between *TNC* and the previously genetically identified asthma susceptibility gene Neuropeptide S receptor 1 (*NPSR1*). In study II, we investigated the interactive effects of *NPSR1* and environmental exposures related to farming lifestyle, as well as the effect of lipopolysaccharide (LPS), a proxy for farm animal exposure, on NPSR1 expression. We provide data showing that *TNC* has an independent genetic role in certain allergic diseases. We show biological interplay by a dose-dependent upregulation of *TNC* expression upon NPS-NPSR1 activation, and we conclude that interaction occurs between *TNC* and *NPSR1*altering the outcome of asthma and allergy. Genetic variations in *NPSR1* are not only dependent on other genes, but can also modify the effect of the environment, on the development of allergic diseases. Farm animal contact and farm milk consumption, introduced early in a child's life, has been proven to show protective effects against development of allergic diseases. In study II, we demonstrate that the protective effect of farm animal contact can be further modified depending on genetic variations in *NPSR1*, especially if the contact is initiated later in life. We also identified increased NPSR1 expression upon LPS stimulation of human monocytes. From these two studies we can confirm that interactive effects, both biological and genetic, are important in the development of asthma and allergy. We could also see that the genetic dependency is most likely to occur when the main effect of the individual genes, or environmental factors, investigated are not that dominant.

In study III, we investigated the function of *NPSR1* in more detail. The *NPSR1* gene encodes two functional receptor variants (A and B) with distinct intracellular C-termini. Previous studies have illustrated different expression pattern, especially in asthmatic airways, between the two receptor variants. We could in study III demonstrate that, upon activation, both receptor variants A and B signals through the same pathways and induces expression of in principal identical set of genes. However, with few exceptions, variant A constantly induced stronger signaling effects than variant B. The effect was seen on both second messenger level and on down-stream gene expression. These findings suggest an isoform specific link to NPSR1s role in allergic airways.

Among children with asthma around 5% suffer from chronic symptoms and/or severe exacerbation despite extensive treatment. The causes of this severe, therapy resistant asthma in childhood are poorly understood. In study IV we aimed to investigate global differences in gene expression in white blood cells from patients with severe, therapy-resistant asthma (SA, n=20), patients with controlled but persistent asthma (CA, n=20) and a group of healthy controls (Ctrl, n=19). We identified 1378 genes to be significantly differentially expressed between any of the contrasts (CA-Ctrl, SA-CA, SA-Ctrl) demonstrating that there are differences in gene expression between groups of asthma. Functional annotation and enrichment analysis identified three significantly differentially expressed pathways; bitter taste transduction, (upregulated mostly in SA), natural killer cell mediated cytotoxicity (upregulated in CA) and N-glycan biosynthesis (downregulated in SA). The bitter taste receptor family (*TAS2Rs*) has recently been shown to play a protective role in asthmatic airways e.g. by dilation of airways upon stimulation with bitter substances. Our finding is the first to propose a role for *TAS2Rs* in asthma outside the airway system. In conclusion our data indicates a separation in gene expression patterns between children with severe, therapy resistent asthma and controlled asthma, and suggests pathways revealing novel insight about the characteristics of severe therapy-resistant asthma.

From the finding in this thesis we can conclude and confirm that there is always a complex interplay between several genes and environmental factors altering the outcome of allergic disease. It is important to investigate these genes in more detail to unravel the functional mode of action. We can also see that by investigating clear defined subgroups of asthma it might be possible to identify new therapeutic targets for asthma.

LIST OF PUBLICATIONS

I. **Christina Orsmark Pietras**, Erik Melén, Johanna Vendelin, Sara Bruce, Annika Laitinen, Lauri A. Laitinen, Roger Launer, Josef Riedler, Erika von Mutius, Gert Doekes, Magnus Wickman, Marianne van Hage, Göran Pershagen, Annika Scheynius, Fredrik Nyberg, Juha Kere and the PARSIFAL Genetics Study Group.
Biological and genetic interaction between Tenascin C and Neuropeptide S receptor 1 in allergic diseases.
Human Molecular Genetics, 2008; 17: 1673-1682.

II. Sara Bruce, Fredrik Nyberg, Erik Melén, Anna James, Ville Pulkkinen, **Christina Orsmark Pietras**, Anna Bergström, Barbro Dahlén, Magnus Wickman, Erika von Mutius, Gert Doekes, Roger Launer, Josef Riedler, Waltraud Eder, Marianne van Hage, Göran Pershagen, Annika Scheynius, Juha Kere.
The protective effect of farm animal exposure on childhood allergy is modified by *NPSR1* polymorphisms.
Journal of Medical Genetics 2009; 46: 159-167.

III. **Christina Orsmark Pietras**, Johanna Vendelin, Francesca Anedda, Sara Bruce, Mikael Adner, Lilli Sundman, Ville Pulkkinen, Harri Alenius, Mauro D'Amato, Cilla Söderhäll, Juha Kere.
The asthma candidate gene *NPSR1* mediates isoform specific downstream signaling.
Submitted

IV. **Christina Orsmark Pietras**, Jon Konradsen, Björn Nordlund, Cilla Söderhäll, Christophe Pedroletti, Gunilla Hedlin, Juha Kere, Erik Melén.
Genome wide transcriptome analysis suggests novel mechanisms in severe childhood asthma.
Manuscript.

ADDITIONAL PUBLICATIONS

Donner J, Haapakoski R, Ezer S, Melén E, Pirkola S, Gratacòs M, Zucchelli M, Anedda F, Johansson LE, Söderhäll C, **Orsmark-Pietras C**, Suvisaari J, Martín-Santos R, Torrens M, Silander K, Terwilliger JD, Wickman M, Pershagen G, Lönnqvist J, Peltonen L, Estivill X, D'Amato M, Kere J, Alenius H, Hovatta I.
Assessment of the neuropeptide S system in anxiety disorders.
Biol Psychiatry. 2010 Sep 1;68(5):474-83.

Hellquist A, Järvinen TM, Koskenmies S, Zucchelli M, **Orsmark-Pietras C**, Berglind L, Panelius J, Hasan T, Julkunen H, D'Amato M, Saarialho-Kere U, Kere J.
Evidence for genetic association and interaction between the TYK2 and IRF5 genes in systemic lupus erythematosus.
J Rheumatol. 2009 Aug;36(8):1631-8.

Orsmark C, Skoog T, Jeskanen L, Kere J, Saarialho-Kere U.
Expression of allograft inflammatory factor-1 in inflammatory skin disorders.
Acta Derm Venereol. 2007;87(3):223-7.

Skoog T, Ahokas K, **Orsmark C**, Jeskanen L, Isaka K, Saarialho-Kere U.
MMP-21 is expressed by macrophages and fibroblasts in vivo and in culture.
Exp Dermatol. 2006 Oct;15(10):775-83.

Tentler D, Johannesson T, Johansson M, Råstam M, Gillberg C, **Orsmark C**, Carlsson B, Wahlström J, Dahl N.
A candidate region for Asperger syndrome defined by two 17p breakpoints.
Eur J Hum Genet. 2003 Feb;11(2):189-95.

Tentler D, Brandberg G, Betancur C, Gillberg C, Annerén G, **Orsmark C**, Green ED, Carlsson B, Dahl N.
A balanced reciprocal translocation t(5;7)(q14;q32) associated with autistic disorder: molecular analysis of the chromosome 7 breakpoint.
Am J Med Genet. 2001 Dec 8;105(8):729-36.

CONTENTS

LIST OF ABBREVIATIONS

ACT	Asthma control test
ASM	Airway smooth muscle
Asn	Asparagine
A549	Human epithelial lung carcinoma cells
BH	Benjamini Hochberg
Ca	Calcium
CA	Controlled persistent asthma
cAMP	Cyclic adenosine monophosphate
cDNA	Complementary DNA
CD69	CD69 molecule
CGA	Glycoprotein hormones, alpha polypeptide
CI	Confidence interval
CNV	Copy number/neutral variation
cRNA	Complementary RNA
DE	Differentially expressed
DNA	Deoxyribonucleic acid
$DRS_{methacholine}$	Slope of the dose-respons curve for provocation with methacholine
ECM	Extracellular matrix
eQTL	Expressed quantitative trait loci
FDR	False discovery rate
FE_{NO}	Fraction of nitric oxide in exhaled air
FEV_1	Forced expiratory volume during 1 second
GPCR	G protein-coupled receptor
G-protein	Guanine nucleotide-binding protein
GWA	Genome wide association
HEK293	Human embryonic kidney cells
HeZ	Heterozygous
HGP	Human genome project
HoZ	Homozygous
HWE	Hardy Weinberg equilibrium
IgE	Immunoglobin E
IHC	Immunohistochemical staining
Ile	Isoleucine
LD	Linkage disequilibrium
LPS	Lipopolysaccharide
LRT	Likelihood ratio test
MAF	Minor allele frequency
MAPK	Mitogen-activated protein kinase
miRNA	MicroRNA
mRNA	Messenger RNA
ncRNA	Non coding RNA
NGS	Next generation sequencing
NK-T cells	Natural killer T cells

NPS	Neuropeptide S
NPSR1	Neuropeptide S receptor 1
NPSR1-A	NPSR1 variant A
NPSR1-B	NPSR1 variant B
NTS	Neurotensin
OR	Odds ratio
PA	Problematic severe asthma
PARSIFAL	Prevention of allergy, risk factors for sensitization in children releated to farming and anthroposophic lifestyle
PBMC	Peripheral-blood mononuclear cells
PLC	Phospholipase C
qRT-PCR	Quantitative real-time polymerase chain reaction
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
SA	Severe asthma
SES	Socioeconomic status
SH-SY5Y	Human neuroblastoma cells
SNP	Single nucleotide polymorphism
snRNA	Small nuclear RNA
TAS2R	Taste receptor, type 2
T_H1, 2	T helper cells one and two
TNC	Tenascin C
T_{regs}	Regulatory T cells
tRNA	Transfer RNA
WBC	White blood cells

1 POPULÄRVETENSKAPLIG SAMMANFATTNING

I varje cell i vår kropp finns förpackningar, kromosomer som består av DNA, som i sin tur består gener. Av varje gen finns det två kopior, av vilka hälften kommer från mamman och hälften från pappan och vi har en uppsättning om ca 20,500 olika gener i varje cell. I generna, och runtomkring dem, finns förändringar som gör oss unika, dels ärvda från många generationer tillbaka, men även nya.

Många av våra vanligaste sjukdomar, såsom bipolära psykoser, hjärt- och kärl problem, reumatism, diabetes och astma styrs både av miljöfaktorer och ärftliga komponenter (såsom gener). Dessa folksjukdomar är så kallade komplexa sjukdomar. I komplexa sjukdomar är det ett samspel mellan arv och miljö som utgör själva risken för att utveckla sjukdom eller inte. Astma är en kronisk luftvägs sjukdom som kännetecknas framförallt av svårigheter att andas. Astma förekommer ofta tillsammans med andra allergiska symptom, så som hösnuva och allergi.

De faktorer som tros påverka utfallet av astma och allergi är främst; (a) Miljöfaktorer - där den mångfaldiga bakteriefloran man utsätts för på ett jordbruk (kontakt med djur, opastöriserad mjölk, etc.) anses ha en skyddande effekt och rökning, tidiga luftvägs infektioner etc. medför en större risk. (b) Ärftliga faktorer - runt 100 gener har identifierats med möjlighet att öka risk för sjukdomen. Många ärftliga faktorer är kopplade till funktioner i immunförsvaret eller luftvägarna, men flera gener har i dagsläget en okänd koppling till astma. Vad vetenskapen kommit fram till de senaste åren är att det sannolikt är förändringar i ett flertal gener som i samspel med flera olika miljöfaktorer, orsakar astma och allergi.

I denna avhandling har vi utforskat hur förändringar i gener kan påverka risken att utveckla astma och allergi. Hur dessa förändringar i olika gener samarbetar med varandra och med miljön för att antingen öka eller minska risken för sjukdom, och hur förändringarna påverkar genes funktion. Vi har även undersökt om det finns skillnader i genernas uttrycksnivåer mellan astmapatienter som är svårt sjuka, där inte ens behandling hjälper, och de som är sjuka men svarar på sin medicinering.

I studie I har vi undersökt förändringar i gener som är kopplade till astma. Vi har studerat hur dessa förändringar på DNA nivå (exempelvis SNPs) i sig själva kan påverka risken att utveckla astma och andra allergiska sjukdomar. Även hur generna kan samarbeta med varandra, gener emellan, och med olika miljö faktorer för att antingen ge en skyddande effekt eller en ökad risk för sjukdom. Med våra forskningsresultat kan vi bekräfta att detta samarbete pågår i kroppen. Vi visar på nya kombinationer av gener (Tenascin C, *TNC* och Neuropeptide S receptor 1, *NPSR1*) vars förändringar både kan öka och minska risken för astma och allergiska sjukdomar.

Tidigare studier har påvisat att barn som bor på en bondgård, har regelbunden kontakt med djuren på gården, och dricker opastöriserad mjölk utvecklar ett skydd mot astma och allergi. Detta gäller speciellt om denna kontakt inleds redan under fosterstadiet, dvs. den blivande mamman vistas i dessa miljöer, eller under de första åren. Vi kan i studie II visa att, i kombination med vissa förändringar i genen NPSR1, har dessa

miljöfaktorer även en skyddande effekt när de kommer in senare under ett barns uppväxt.

I studie III studerade vi funktionen av genen *NPSR1* i mer detalj. NPSR1 är en mottagare, en s.k. receptor, som sitter på cellens yta, och tar emot signaler från omgivningen och genom receptorn förmedlas dessa signaler in i cellen. Detta startar en kedja av händelser som talar om vad cellen ska göra. Det finns två varianter av NSPR1, A och B, och det som skiljer dem åt är den biten av receptorn som befinner sig inne i cellen. A varianten är den vanligaste förekommande i kroppen av de två men i studier på patienter med astma har man sett att B varianten är vanligare på vissa celler i luftvägarna. I cellkulturer på laboratoriet har vi studerat om ”svaret” (aktivering av andra proteiner och gener inne i cellen), som cellerna ger när de olika receptorvarianterna signalerar in till cellen, skiljer sig åt. Vad vi har kunnat påvisa är att variant A ger ett starkare svar än B varianten.

Inom astma finns det olika svårighetsgrader av symptom. Vissa patienter uppvisar mildare symptom, andra svårare symptom, och ytterligare några är svårt sjuka och svarar inte på behandling. För den sistnämnda gruppen är det viktigt att undersöka vad det är som gör att de inte svarar på behandling. Det långsiktiga målet är att identifiera faktorer som skulle kunna leda till nya läkemedel för denna patientgrupp. I studie IV har vi analyserat skillnader i genernas uttryck i vita blodkroppar från de astmatiker som är svårt sjuka och inte svarar på behandling, och de som har något mildare symptom och svarar bra på behandling. Vad vi kunde påvisa var att det fanns skillnader i genernas uttryck mellan grupperna. Patienter med mildare symptom hade i regel ett högre uttryck av de gener som är kopplade till immunförsvaret. Den svårt sjuka gruppen visade i stället ett högre uttryck av en grupp gener som hör till gruppen av smakreceptorer som i munnen känner av bittra substanser. Denna grupp av receptorer har nyligen identifierats även i luftvägarna och tros ha en skyddande effekt mot astma. Vad dessa receptorer har för funktion i vita blodkroppar återstår att se, men dessa fynd kan ge oss ledtrådar om vad som orsakar terapiresistens hos svårt sjuka astmatiker.

Sammanfattningsvis kan vi i denna avhandling påvisa ytterligare bevis för att förändringar i gener samarbetar med varandra och med olika miljöfaktorer och därmed påverkas utfallet av astma och allergiska sjukdomar. Vi visar på funktionella skillnader mellan två varianter av en receptor som tidigare kopplats till astma. Genom att jämföra olika patientgrupper med astma, de som svarar på behandling mot de som inte svarar på behandling, kan vi identifiera nya nätverk av gener som skiljer dessa grupper åt. På sikt skulle dessa resultat möjligen kunna leda fram till nya läkemedel mot astma.

2 BACKGROUND

2.1 INHERITANCE

Many diseases, as well as characteristics of our body rely mainly or partially on our inheritance. Each individual has the same set of genes but there are some distinct changes in the genome that make us unique. These changes can be passed on from generation to generation, and are thus inherited. In some cases, variations in *one* gene are both necessary and sufficient for a character to be expressed. Such characteristics are called Mendelian, or monogenic. Cystic fibrosis is an example of a monogenic disease caused by alterations in the gene *CFTR*.[1] However, most of our traits are non-Mendelian and determined by a combination of alterations in *several* genes and also influenced by environmental factors. Such characters are called multifactorial or complex traits/disorders. Asthma is a typical complex disease where the features of the disease are influenced both by the genetic make-up composed of alterations in several genes, as well as by exposure to various environmental factors. Other common complex disorders are bipolar mental disorders, coronary artery disease, Crohn´s disease and diabetes. In this thesis I will discuss disease, inheritance, identification and consequences of inheritance in the context of complex disorders.

2.2 ASTHMA AND ALLERGIC DISEASE IN CHILDREN

2.2.1 Definition and prevalence

Allergic disease is a broadly defined term that includes several phenotypic characteristic such as allergic sensitization, asthma, wheeze, and rhinoconjunctivitis (hay fever). Some of these features commonly coincide. Allergic sensitization is a response in the immune system against, in normal cases, harmless substances such as certain food or other allergens, and usually defined by an increase in specific IgE antibodies (in plasma), or a positive skin prick test. Asthma is characterised by chronic inflammation in the airways. The inflammation is usually accompanied by limitations in airflow as a result of hyper-secretion of mucus and broncho-constriction causing symptoms like wheezing (a whistling sound during breathing), shortness of breath, chest tightness or coughing. Asthma occurs in different degrees of severity and as allergic (combined with allergic sensitization) or non-allergic asthma. Currently, asthma can roughly be classified clinically as mild, moderate, or severe based on the presentation of symptoms in relation to the level of medication.[2, 3] Asthma in children is closely related to wheeze and the severity of asthma is often defined by the number of reoccurring wheezing episodes.[4] Rhinoconjunctivitis is characterised by sneezing, nasal congestion and itching of the nose, eyes or throat and is in children usually associated with an allergic sensitization to pollen or pets. The cause of asthma and allergic disorders is currently unknown. However, it is very likely that genetic and environmental factors interact with each other to determine the outcome of asthma.[5]

The prevalence of asthma and allergic disease has increased in developed countries over the last century and asthma is now one of the most common chronic diseases of

childhood. The International Study of Asthma and Allergies in Childhood (ISAAC) demonstrates that the prevalence of asthma symptoms varies worldwide, with English speaking countries and Latin America having the highest prevalence, up to 20%. In Africa, the Indian Sub-continent and in the Eastern Mediterranean, the disease appears to be less often recognized but more severe. The trend seen is an increased prevalence in more affluent countries but a more severe phenotype in less affluent countries.[4] Among children with asthma, around 5% suffer from severe asthma, characterized by chronic symptoms and/or severe exacerbations despite drug treatment.[6] The prevalence of rhinoconjunctivitis in the same study was on an average 15% for current symptoms.[7]

The clinical diagnosis of asthma is usually based on patterns of symptoms, response to therapy, together with clinical testing of e.g. lung capacity.[8] Lately it has been discussed if clinically defined asthma is one disease or composed of several sub-diseases. Proposed classifications based on combined data on clinical features, physiology, immunology, pathology, genetics, environment, response to treatment and other factors might in the future lead to new, more specific definitions.[9, 10] Currently, all available treatments, e.g avoidance of allergens and irritants together with medications such as inhaled corticosteroids and bronchodilators, are non-curative and only provide symptom prevention or relief.

2.2.2 Factors influencing the development of asthma and allergic disease

Asthma is a complex disorder and thought to be influenced by several different factors.[8] These can either act as risk factors, or protect against developing the disease. Examples of such factors are: a) The local environment- living on a farm and consumption of farm-produced products is believed to have a protective effect[11], whereas smoking, pollution, and exposure to molds is negative for the outcome[12]; b) Genetics- early twin studies[13, 14] to modern genome wide associations studies (GWAs)[15] all point towards a high heritable factor in asthma; c) Infections- since the early 1970s and up to now, studies have shown a relationship between viral lower respiratory tract infections and asthma later in childhood[12, 16]; d) Gender, age, family size, obesity, ethnicity, and antibiotics during early infancy are also factors thought to influence the development of asthma.[12, 17, 18]

There has been an accelerated increase in the prevalence of asthma and allergies in the Western societies over the past decades. To explain this phenomenon the term “hygiene hypothesis” was coined in 1989.[19] The hygiene hypothesis states that due to the “Westernized” lifestyle, e.g. smaller family size, improved living standards and better personal hygiene, children are exposed to a “cleaner” environment, with fewer infections resulting in an increased risk of developing allergic diseases. Recent data still show an increase of asthma and allergy in the affluent world. [4] However, this may not be the consequence of a “cleaner” environment in the previous sense, but rather due to a deficiency in exposure to a diverse microbial environment in.[20-22] The main proposed cellular mechanism behind asthma and allergy is thought to be a skewed balance in the immune response between the two white blood cell groups T helper cells 1 and 2 (T_H1-

T_H2). Shifts in balance towards $T_H1$would lead to autoimmune disease whereas a shift towards T_H2 would lead to allergic disorders. A positive exposure to a diversity of non-pathogenic microorganisms, together with other factors, is now suggested to regulate the immune system and to maintain the balance between T_H1 and T_H2 cells.[23] Since the protective effects are no longer believed to be due to an "unhygienic" environment, but rather due to a diverse exposure to a beneficial bacterial flora the term "hygiene hypothesis" has been proposed to be replaced by the "microbial deprivation hypothesis".[20]

2.3 THE HUMAN GENOME

2.3.1 Composition of the genome

The human genome is made up of strings of deoxyribonucleic acid (DNA) molecules consisting of unique nucleotide bases, i.e. adenine (A), cytosine (C), guanine (G) and thymine (T). Continuous variation of the composition of these four bases is what makes up all our genes. DNA normally occurs as a double helix, comprising two complementary strands where a distinct stretch of DNA, used as a template to synthesize a functional complementary ribonucleic acid (RNA) molecule (which in most cases encodes for a protein), is called a gene. According to the latest estimate there are around 20,500 protein coding genes in the human genome. [24] The double helix DNA structure is rolled up on histones and twisted and twisted until it is packed into a chromosome. There are 23 chromosome set in the human body and most cells (except the sperm and egg cells, which only have one) carry two copies of each chromosome (44 and XX for females and 44 and XY for males). Each unique gene is used as a template to synthesize (RNA). RNA molecules can be divided into two broad classes, messenger RNA (mRNA); that serves as templates for protein production, and noncoding RNAs (ncRNA); which are RNA molecules that do not translate into proteins. Included in the group of ncRNA are among others, ribosomal RNA (rRNA), transfer RNA (tRNA), small nuclear RNA (snRNA), microRNA (miRNA). The ncRNA have important functions in building blocks in ribosomes, perform functions as carriers of amino acids during translation and regulating the expression of other genes.[25] Genes, usually, consists of both exons and introns which make up the template for RNA. The introns are however removed by splicing and it is only the exons that constitute the final mRNA (or ncRNA), which in turn can be translated into a protein (*Fig. 1*). The proteins constitute the cornerstones in building the body and provide the functional machinery that is utilized by all cells to perform their tasks.

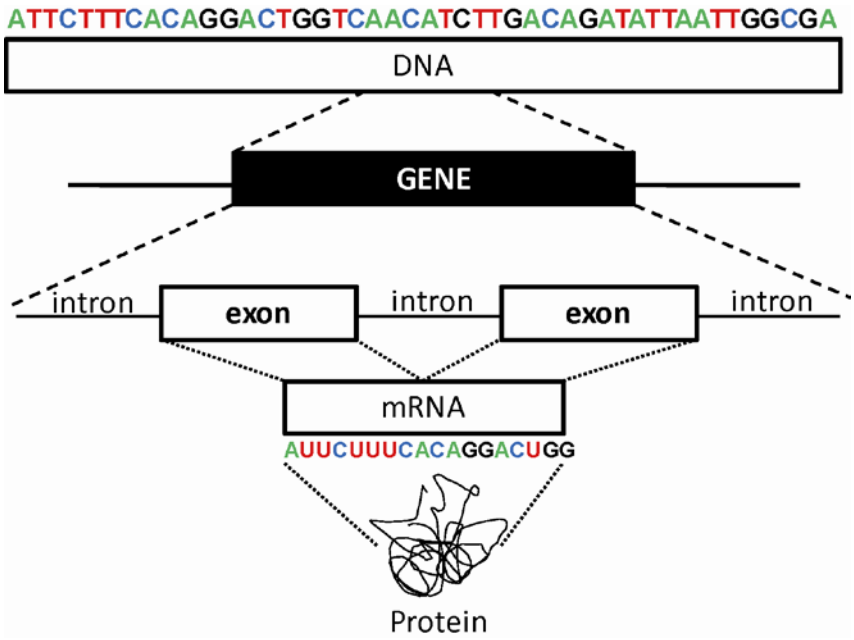

Figure 1*. DNA is made up by combinations of nucleotides with either an A, C, G or T base attached to it. A stretch of DNA makes up a gene which in turn is split up in exons and introns. The exons serve as the template for mRNA which codes for the proteins.*

2.3.2 Sequence variations in the genome

A unique chromosomal location defining the position of a gene or a DNA sequence is termed a locus (or several loci). An alternative version of a gene is called an allele. For one gene there are two alternative alleles in each cell, one on each chromosome, i.e. one inherited from the father and the other from the mother. An allele does not have to be the entire gene but can be designated to one specific nucleotide. The human genome contains numerous variations which can occur both within and among human populations. These variations are named single nucleotide polymorphisms (SNPs). Currently there are 12 million SNPs deposited in the GenBank® database (http://www.ncbi.nlm.nih.gov/genbank). A SNP can occur within an exon (coding SNP) or elsewhere in the genome (non-coding SNP).

A person is homozygous (HoZ) if it has the same base in a specific position (e.g. a “C” or a “T” on both alleles). A person is said to be heterozygous (HeZ) if the base in a specific position differ between the alleles (e.g. a “C” in one allele and a “T” on the other) (*Fig. 2*). Alternatively, the individual is said to have a “CC”, a “CT” or a “TT” genotype. If “C” is the common allele then “T” will consequently be the least frequent, rare allele (*Fig. 2*). The frequency of this rare or minor allele in the population is termed minor allele frequency (MAF). There are other variations in the human genome as well; copy number/neutral variations (CNVs), which is when a gene/part of a gene is deleted, duplicated, inverted or translocated, and epigenetic changes, which can be defined as “the structural variation of chromosomal regions so as to register, signal or enable altered activity states”.[26] The epigenetic changes do not affect the underlying DNA code but rather modifies how it is expressed.

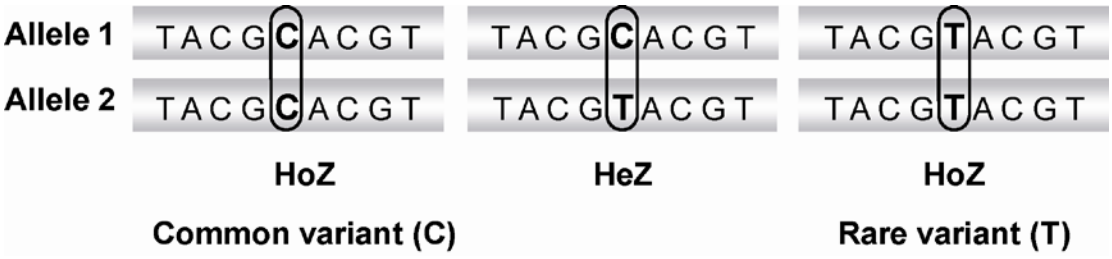

***Figure 2**. Illustrating three possible combinations of the two alleles. A person can either be homozygous (HoZ) or heterozygous (HeZ) for a genetic variation. The least frequent genotype (i.e. TT) will be the combination of the rare alleles (i.e. T).*

When a particular genotype present at one locus is *independent* of a genotype at a second locus, the genotypes are said to be in linkage *equilibrium*. When, on the other hand, a particular genotype present at one locus is *dependent* of a genotype at a second locus, the genotypes are said to be in linkage *disequilibrium* (LD). LD is a non-random association of genotypes. The genotypes within such a region form blocks called LD blocks. LD blocks are separated by hotspots of recombination and the genotypes in an LD block are most likely inherited together.

Although many of the variations discussed above do not lead to any visible phenotypic distinction between people, taken together all these variations result in a human-to-human genetic variation of approximately 0.5%.[27] As a result, each person carries an exclusive genome with a unique risk for each developing diseases, respond to the environment and respond to drug treatment. Of note, we are still 99.5% alike each other.

2.3.3 Finding disease genes: from past to presence

In 1953 the DNA double helix was discovered by Watson and Crick and the findings made way for new research in the field of genetics.[28] Since then, the field has simply exploded. In1990, the international Human Genome Project (HGP) was launched, with the goal to sequence the entire human genome. The project was completed and published in 2004[29], even though drafts of the human genome sequence, now performed as a race between the publically founded HGP and the private company Celera Genomics, were published already 2001, presenting roughly 90% of the total sequence.[30, 31] The drafts suggested that the genome contained 30,000-40,000 genes, not 100,000 as previously estimated.[32] However this number has been revised after completion of the project and is now estimated to ~20,500 distinct protein-coding genes.[24]

2.3.3.1 Linkage analysis

During the past years, linkage studies and candidate gene association studies have been the main approach to identify susceptibility genes for complex genetic disorders. Linkage studies are performed on families with affected individuals. Each family member is genotyped for genetic markers that are evenly spread throughout, a candidate gene region or the whole genome covering all chromosomes. If a region is

identified that contains a higher than expected number of shared alleles among the affected individuals, the region is said to be linked to the disease in question. The main advantage of this approach, when used at the whole genome level, is that it is hypothesis-free. The disadvantage is that the regions identified are quite large and requires further fine-mapping, i.e. positional cloning. Even when fine-mapping is performed, the identified region might harbor several susceptibility genes with small effects that together create the linkage peak thus making it difficult to elucidate which are the true candidate genes.

2.3.3.2 Candidate gene association studies

Candidate gene association studies (used in study I) mainly focus on previously identified candidate genes. They are usually performed on groups of unrelated "cases" and an appropriate unrelated control group ("controls"). A certain number of tagging SNPs (see *2.3.3.3*) or biologically interesting SNPs are genotyped in the gene in question and the frequency of each genetic variant is compared between cases and controls. A significantly altered frequency in cases vs. controls indicates that the genetic variant is related to disease susceptibility. The main advantage of association studies is that they are more easily powered, it is easier to obtain large numbers of unrelated individuals compared to collect large number of families. SNPs are used for case-control association studies both in small scale (candidate gene approach) and large scale (GWAs, see 2.3.3.4). Tests can be performed for either single SNP associations, which means that one SNP at the time is tested, or for a combination of SNPs in tight LD, a haplotype. Before any association analyses are performed, each SNP is tested for Hardy-Weinberg equilibrium (HWE), i.e. the allele frequencies received when genotyping the study groups are compared to the expected allele frequencies in a corresponding population. A deviation from HWE in random samples may indicate genotyping errors, leading to false conclusions.[33]

2.3.3.3 The HapMap project

In 2002, the international HapMap project was initiated to construct a genome-wide SNP database of common variations and to generate maps of stretches, or blocks, of DNA inherited together, i.e. SNPs in tight LD (http://www.hapmap.org).[34] This was initially done for 269 subjects in 4 different populations. Use of information from the HapMap project facilitates association studies, since it enables genotyping of only those SNPs that tag the blocks, i.e. tagging SNPs, and not every SNP in the human genome. If a tagging SNP is associated to disease, it is likely that the causative SNP/SNPs are in that haploblock (for an example of a haploblock/LD plot, structure, see *Fig. 5*). During the progress of the HapMap project, the development of genotyping arrays also took place (i.e. SNP chips). These arrays contain spots for today up to ~ 2 million SNPs and allow for a quick and cost effective way of extensively genotype many individuals (one per chip) at the same time without prior hypotheses.

2.3.3.4 Genome wide association studies (GWAs)

Based on the knowledge of tagging SNPs from HapMap, other genetic variants, and the development of SNP chips, large scale genome wide association (GWA) studies started to take place in 2005-6. Up to millions of SNPs can be tested for disease association in hundreds or thousands of individuals, which provides a great advantage over linkage analysis and candidate gene associations. The principle is the same as for regular case-control association analyses with one important exception. When association studies are performed with few SNPs in a defined candidate gene or region, a statistic significance of $p \leq 0.05$ is generally accepted. However, when many hundred thousands of SNPs are tested, as in GWAs, the number of tests performed will result in a large number of false positives if a $p \leq 0.05$ would be used. This multiple testing needs to be corrected for and in GWAs, an association is only regarded as significant if $p \leq 5 \times 10^{-8}$ (http://www.genome.gov/gwastudies). (See section 4.5.1 for more details on multiple testing)

2.3.3.5 Next generation sequencing and the 1000 genomes project

In parallel with the publication of an increasing amount of GWA studies, the sequencing technology has changed and vastly improved. What previously took years to finish is now completed within weeks with the Next Generations Sequencing (NGS) technology.[35] The major advantage over the array technology is that instead of only analyzing assigned spots, the whole genome can be investigated. The 1000 genomes project is a deep re-sequencing project, and an extension from the HapMap project aiming to provide detailed information on genetic variation using more than 1,000 genomes from populations all around the world (http://www.1000genomes.org). The goal is to identify 95% of all variantions with a greater frequency than 1%. This will hopefully provide information to evaluate the common disease/ many rare variants hypothesis (see section 2.3.3.6).

2.3.3.6 Common disease-common variants and the missing heritability

Much hope was put in the genetic information that would be gained from GWA studies. However, taking into account all these efforts to connect genetic variation to complex disorders, the risk effects (estimated as odds ratios, ORs) of individual genes have still turned out to be notably lower than the estimated total genetic risk. The gap between the expected proportion of genetic factors in disease and the actual findings from GWAs has been termed the "missing heritability".[36] The search for susceptibility genes in GWAs was founded on the "common disease-common variant" hypothesis; i.e. the genetic influences on susceptibility to common diseases are attributable to a limited number of common variants present in more than ~5% of the population.[32, 37] A possible explanation for the "missing heritability" has been that this hypothesis is not entirely correct, i.e. it is not only the common variants that contribute to common disease, but also the rare variants (less than ~0.5% frequent in the population). A new hypothesis was hence proposed; the "common disease-many rare variants" hypothesis.[38] It states that there are many large effect rare variants in the population, and each case of a common genetic disorder is due to the summation of the effects of a

few of these rare variants. Other possible explanation to the missing heritability has been attributed to the structural variants such as copy number variation, inversions, translocations and epigenetic effects, all of which are poorly detected by the genotyping arrays used and to the existence of gene-gene and gene-environment interactions.[36]

2.3.4 Factors influencing the genetic role in disease

Even if there is an evident strong genetic risk factor in a certain disease, there are many ways by which this risk can come across.

2.3.4.1 Direct and indirect variations in the DNA code

As discussed earlier, there are direct changes to the DNA, such as SNPs and copy number or copy neutral variations, that can affect a person's susceptibility for disease.[27] A SNP can lead to various functional changes, e.g. a new or abolished transcription factor binding site altering the expression of the gene, a splice variant of a gene resulting in distinct functions, structural changes of the protein hindering or enhancing the properties of a receptor, etc. These changes on the DNA level are more or less irreversible and usually affect all cells in the body, which is very convenient because the analysis can be performed on a sample from any tissue in the body, including blood. There are also variations that do not directly affect the DNA code, such as epigenetic changes. Epigenetic alterations include DNA methylation and histone modification and are structural adaptations of chromosomal regions that leave the DNA with a "memory".[26] In contrast to direct genetic changes to the DNA, epigenetic changes can be modulated both over time and by biological and environmental factors, and be cell type specific. Epigenetic alterations can affect the expression of genes, turning them on or off.[27]

2.3.4.2 Interactions; gene and environment

During the last years, it has become clear that neither genes, nor environment act alone to cause disease in the context of complex disorders. Genes collaborate with each other and with the environment in a complex pattern.[39] If a genetic factor functions through a complex mechanism that possibly involves several genes and/or environmental factors, the associated risk (or protective effect) might remain undetected if a particular gene is examined in isolation, as in regular gene-disease association studies (including GWAs). Therefore, it is important to take gene-gene (also termed epistasis) and gene-environment interactions in consideration. Interactions can be defined both functionally and statistically. A functional gene-gene interaction is caused by the molecular interaction that proteins and other genetic elements have with each other, either in the same pathway or in direct complexes.[40] A statistical interaction means that the outcome (or risk estimated) is different when a particular set of alleles from distinct loci are found in combination, than when they are considered apart. The difference is a departure from the expected effect if the two alleles were combined independently. The departure can be either from an additive model (adding up the effect from each allele)

or from a multiplicative model (multiplying the effect from each allele).[41] The same methods apply when gene-environment interactions are investigated.

2.3.4.2.1 Gene-gene interactions

Data supporting gene-gene interactions in complex diseases have emerged rather slowly. When searched for, interactions are commonly found, however they are not always easy to replicate. The reason might be that disease is due to accumulation and interaction of multiple variations acting together in the same network.[42] The main approach for gene-gene interaction studies performed to date has focused on genes with a reported biological or genetic association, in this case to asthma and allergy. A successively replicated network is the IL-4/IL-13 pathway, which is central for IgE regulation. Several studies have identified the outcome of asthma and allergy to be dependent on combinations of variations in *IL-4*, *IL-4Rα, IL13* or *STAT6*.[43-45] *IL-13* has also been implicated to interact with *GATA3*, a transcription activator of T_H2 cytokines, in childhood rhinitis [46] and *IL-4Rα* and *IL-9R* have been reported to modify the risk of childhood wheezing.[47] Other genes, such as *TLR2, IL-6, TGFβR2* and *FOXP3* (involved in development and function of regulatory T cells) have also been implicated to interact and modify the risk of sensitization against various allergens.[48]

2.3.4.2.2 Gene-environment interactions

Gene-environment interactions in asthma and allergy are more frequently reported, especially in relation to farming or farm-related exposures. The toll-like receptors (TLRs) consist of a family of innate immunity receptors with microbial molecules as ligands. The *TLR2* gene has been identified as a major determinant of asthma and allergy susceptibility in farmers' children, whereas the *TLR4* gene influences atopy in children heavily exposed to endotoxins.[49] The *CD14* gene, a pattern recognition receptor for microbial molecules, is dependent on exposure to animals and/or house dust endotoxins for its association to asthma, allergy or atopy.[50-52] It also modifies the protective effect of farm milk consumption on allergic diseases.[53] A study investigating SNPs in all *CD14*, *TLR2*, *TLR4* and *TLR9* genes identified asthma association to be dependent on country living.[54] The *NOD1* gene is an intracellular pattern-recognition receptor that initiates inflammation in response to bacteria and SNPs in *NOD1* modifies the protective effect of exposure to a farming environment.[55] Additional gene-environment interactions have also been investigated for exposures such as, e.g. air-pollutants[56] and tobacco smoke.[57, 58] All these studies have focused on environmental interactions with pre-defined asthma and allergy susceptibility genes. A recent large scale genome-wide gene-environment interaction study aimed to identify novel common gene variations, as well as replicate previously reported findings, however no significant interaction was identified.[59]

2.3.4.3 Gene expression

All direct and indirect variation in the DNA mentioned above can alter the RNA expression of a gene, which in turn can provide us with clues to the biological mechanisms behind genetic alterations. Gene expression can be upregulated,

downregulated or abolished. Either the expression of one, or a few genes are investigated using smaller scale methodologies, or a more large scale profile using expression microarrays (or RNA sequencing) is performed. The large scale approaches are valuable tools for identifying key regulatory networks or pathways important for a disease, and can also help to define sub-phenotypes within a clinically defined phenotype. Integrated with the genetic data on variations in the genome, it can also pinpoint individual changes in expression (expression quantitative trait loci, eQTL).[60] Due to e.g. indirect variations in the DNA and external and internal signals to the cell, gene expression is cell type specific. Therefore, when it comes to gene expression it is very important to consider which tissues or cells should be investigated. Gene expression can also be investigated *in vitro* by e.g. analyzing the downstream effects upon modification (overexpression or silencing) of one single gene in a specific cell type.

2.3.4.3.1 Gene expression in asthma

Variation in gene expression is an important mechanism in mediating susceptibility to disease and similar to candidate gene association studies, candidate gene expression studies on asthma have been performed. Focusing on differences between severe and mild asthma, some studies have reported both altered protein and mRNA expression for various inflammatory genes. Increased TNFα (a proinflammatory cytokine) levels have been found in bronchoalveolar lavage (BAL), biopsies from the airway lumen[61] and in peripheral-blood mononuclear cells (PBMCs).[62] Along with TNFα, IL33 (a promoter of T_H2 immunity) gene expression is increased in lung tissues from asthmatic patients.[63] IL-17 (a proinflammatory cytokine) is increased in serum from severe asthmatics compared to mild or moderate forms.[64] Investigations of global RNA expression in blood from asthmatic patients have so far been scarce. A transcriptional profile of T lymphocytes from asthmatic children revealed that children from a low socioeconomic status (SES) showed overexpression of genes regulating inflammatory processes compared to those from a high SES.[65] Profiling of subpopulations of PBMCs collected from asthmatic children during exacerbation vs. convalescence revealed upregulation of T_H2-associated functions in monocytes/ dendritic cells during the acute phase.[66]

2.3.4.4 Phenotypic heterogeneity

The phenotype is defined by the observed characteristics of an organism, influenced both by the genetic makeup and the environment. In complex disorders, such as asthma, the same genotype can result in several different phenotypes. In the same manner, different genotypes can give raise to the same phenotype. To be able to replicate genetic studies, the definition, measurements and validity of phenotyping need to be standardized.[67] This makes genetic studies, and especially large scale studies, where study groups from several collaborative groups are combined, troublesome. Even if guidelines are worked out[2, 3] and properly followed, they are often based on the phenotypic characteristics and will not take into account the existence of genetically different subgroups within the same phenotypic group. This has been shown by cluster analysis using 34 different variables on 726 subjects all defined as “severe asthmatics”, where not one but five distinct phenotypes were identified.[10] Even though large scale studies are important to identify over-all important genes, this supports that small scale

studies on well defined study groups also are needed to unravel subgroup-specific pathways. It also suggests that many of the replication problems of asthma susceptibility genes might be due to phenotypic heterogeneity.

2.4 GENETICS OF ASTHMA

It is now commonly accepted that asthma and asthma-related traits behave as typical complex disorders. The definition of a complex disorder is: "conditions in which various genetic hits that are individually mild may be capable of major phenotypic effects, when acting together and within a certain environmental context". As previously discussed there is good evidence that both multiple genes and environmental factors are part of the etiology of asthma.

2.4.1 Susceptibility genes in asthma

Over the past years, several susceptibility genes have been identified in asthma and allergy-related disorders, both biologically and genetically. The two main genetic approaches from start were genome-wide linkage studies and candidate-gene association studies, which have become outshined by the GWA studies. Until now, there are almost 100 well replicated susceptibility genes in asthma or related traits.[68] Most of the genes have been identified through hypothesis-driven studies where SNPs in genes in known pathways are tested against asthma and related phenotypes. Asthma susceptibility genes can be categorized into four main groups; genes associated with innate immunity and immunoregulation (e.g. *TLR*s, *NOD 1,2* and *HLA-DR-DQ-DP*), genes associated with adaptive immunity and T helper 2 cell (Th2) differentiation (e.g. *IL-4,12B, 13,5, IL4,5-RA, GATA 3, TBX21* and *STAT6*), genes associated with epithelial biology and mucosal immunity (e.g. *CCL5,11,24,26* and *FLG*), and genes associated with lung function/airway remodelling and disease severity (e.g. *ADRB2, GSTP1, NOS1, TNF, LTC4S, TNC* and *NOS1*). To these categories one can also add the genes identified through positional cloning (e.g. *ADAM33, DPP10, PHF11, HLA-G, CYPF1P2, IRAKM, COL29A1* and *NPSR1*).[69] The first GWAs on asthma was published in 2007, identifying *ORMDL3*[70] as a novel susceptibility gene. Since then a number of genes have been identified by GWA studies in relation to asthma, including *CTNNA3*[71], *PDE4D*[72], *TLE4* and *ChCHD9*[73], *DENND1B* and *CRB1*[74], *RAD50, SCG3* and *KIAA1271*[75], *HLA-DQ, IL33, IL18R1, SMAD3, GSDMA, IL2RB, RORA, GSDMB, IL13* and *SLC22A5*[15] (http://www.genome.gov/gwastudies). For many of these genes, the function is still uncertain and they might fall under any of the above mentioned categories. In the following sections I will discuss some of the asthma susceptibility genes examined in this thesis. These have either been primarily identified through genetic studies or through their biological properties.

2.4.1.1 Neuropeptide S receptor1

Neuropeptide S receptor 1 (*NPSR1* also *GPRA* or *GPR154*) is a G protein-coupled receptor (GPCR) (see section 2.4.1.3.1), first identified in 2004 as an asthma candidate gene through positional cloning.[76] The study showed both single SNP and haplotype

associations to asthma and total IgE in three separate populations. These findings have later been replicated in several studies and populations.[77-84] *NPSR1* is regarded as one of the robustly replicated susceptibility genes for asthma and asthma related traits. [68, 85] However, it has not, similar to many other previously identified susceptibility genes, reached appropriate significance level in any of the performed GWA studies. This might be due to poor SNP coverage on the microarrays used, but also due to phenotypic variation in the large study groups required for GWA studies. Another possible explanation might be gene-gene or gene-environment interactions. Previous studies have presented results suggesting that the environment might modify the risk effect of *NPSR1*.[77, 78] *NPSR1* has, apart from asthma, also been identified as a candidate gene for inflammatory bowel disease[86], sleep and circadian phenotypes[87] and anxiety[88].

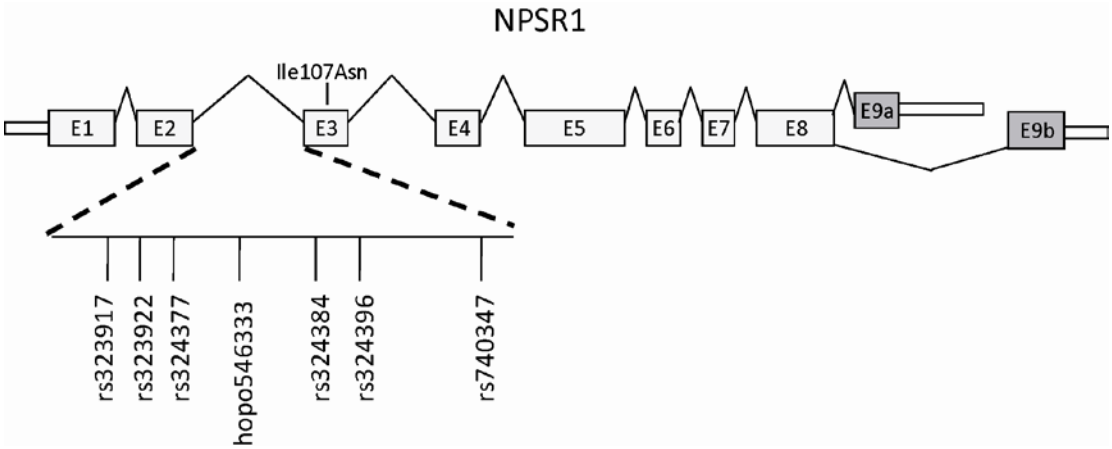

Figure 3*. A schematic picture of the NPSR1 gene. The gene is ~220kb and located at chromosome 7p. The SNPs used in study I and II are depicted and localised in intron two. Exon 3 harbours the only variant, Ile107Asn, that to date has proven to be functionally important. Alternative splicing in exon 9 result in two gene products, NPSR1-A and NPSR1-B. (Modified from study III)*

The *NPSR1* gene undergoes alternative splicing and several splice variants have been identified, out of which only two, NPSR1-A and NPSR1-B, are transported to the plasma membrane.[89] These full-length splice variants differ in their 3' ends with alternative terminal exons 9a and 9b (*Fig. 3*) encoding distinct carboxy-terminal peptide chains. What regulates this splicing event is not yet fully understood. The C-terminus is important for many stages during the lifespan of a GPCR and modifications can affect, e.g. transportation to the cell membrane, anchoring, downstream signaling[90] and also potentially dimerization.

Expression of NPSR1 has been identified in e.g. human bronchus, gastro-intestinal tract, skin, inflammatory cells[76, 89, 91, 92] and in murine brain[93]. Generally NPSR1-A and -B have a similar expression pattern but the A variant is more widely expressed than B.[91] However there are some discrepancies that might suggest important functionally distinct roles for NPSR1-A and -B. In the main, the A variant has a more prominent protein expression in smooth muscle whereas the B variant is dominant in the epithelial cells.[76, 89] The B variant have elevated protein expression in the bronchial smooth muscle layer and epithelial cells in asthmatic patients.[76] The A variant protein is expressed in the basal surface, whereas B is expressed in the apical surface of the colon epithelial cells and skin keratinocytes[89] and the A variant protein is uniquely expressed

in enteroendocrine cells in the gut[91]. Differences have also been detected in cell lines where e.g. a monocytic cell line stimulated with inflammatory mediators showed higher mRNA expression of variant B than A.[91] However, the expression pattern of NPSR1 is still under debate and there are results showing difficulties in detecting expression in some of the above mentioned tissues as well.[94]

The ligand for NPSR1, known as Neuropeptide S (NPS) is a 20-mer peptide. It activates signaling through NPSR1 by inducing both G_s and G_q pathways. It was first identified in brain and has been shown to regulate functions such as arousal, locomotion, food intake and anxiety.[93, 95] In general, expression of NPS follows the expression pattern of NPSR1.[89, 91]

Although there are many polymorphisms in NPSR1, only one, a non-synonymous variant Asn107Ile, has so far been described to be functionally important (*Fig.3*). It is situated in the first (out of three) extracellular loop. NPS stimulation of the 107^{Ile} variant results in increased second messenger response when investigating Ca^{2+} and cAMP accumulation, and MAPK phosphorylation, compared to 107^{Asn}.[96-98]

2.4.1.2 Tenascin C

Previous expression array studies from our group investigating NPS-NPSR1-A signaling identified tenascin C (*TNC)* as one of the differentially expressed target genes.[99] *TNC* is an extracellular matrix (ECM) protein functioning as an adhesion-modulating molecule. It has its main biological roles in cell communication and signal transduction. TNC belongs to a family of glycoproteins that displays highly restricted expression patterns but is upregulated in pathological states including inflammation, or in reparatory processes such as wound healing.[100, 101] Several previous studies have connected TNC expression to asthma and allergy both in mouse[102-106] and in human[107-112]. Genetic studies have so far been scarce. A few genome-wide linkage studies have linked the 9q33 region were *TNC* is situated to asthma or allergic disease and one candidate gene study on *TNC* has been reported, showing strong association to adult asthma for a coding SNP (Leu1677Ile). [113-117]

2.4.1.3 Taste receptor, type 2

The bitter taste receptor family (TAS2R) consists of more than 25 members. They belong to the GPCR family (see section 2.4.1.3.1) and are activated through binding of bitter compounds. Some TAS2Rs can respond to several bitter compounds, whereas others are activated only by a few. The expression pattern of TAS2Rs has been believed to be restricted to the oral cavity, however recent studies suggest gene expression and function of the TAS2Rs in the respiratory and gastrointestinal systems. In the respiratory system, TAS2Rs are expressed on the motile cilia emerging from human airway epithelial cells.[118] The motile cilia are important for propelling mucus and harmful material out of the lung and previous investigations have shown that viral infections and cigarette smoking, which acts as risk factors for asthma, causes a loss of airway cilia and might disrupt this defensive system.[119] Activation by bitter compounds causes an increase in the cilia intracellular Ca^{2+} concentration and increased ciliary beat

frequency.[118] Expression of TASR2s has also recently been identified in human airway smooth muscle (ASM). Stimulation with bitter tastants causes relaxation of isolated ASM and dilation of airways. Inhaled bitter tastants also causes decreased airway obstruction in a mouse model of asthma.[120] Overall, it seems like TAS2Rs might have a protective response effect in asthma. In the gastrointestinal system, TAS2Rs are expressed by enteroendocrine cells and are proposed to orchestrate an appropriate response to specific nutrients or harmful substances by the release of various peptides (reviewed by Sternini *et al*, 2008).[121]

2.4.1.3.1 G protein-coupled receptors

Both NPSR1 and TAS2Rs are G protein-coupled receptors. GPCRs are situated on the cell membrane and characterized by an extracellular N-terminus, a seven-transmembrane α-helix structure spanning the membrane and an intracellular C-terminus (*Fig. 4*). Upon extracellular stimulation, conformational changes of the receptor cause activation of a guanine nucleotide-binding (G) protein, after which the G-protein detaches from the receptor and modulates the activity of other intracellular proteins. NPSR1 couples to the G-proteins Gαq and Gαs.[93, 95] Gαq causes activation of the intracellular protein phospholipase C (PLC), which through a signal cascade causes release of intracellular Ca^{2+}. Gαs activates the intracellular cAMP-dependent pathway. TAS2Rs couples to the Gα protein, α-gustducin[122] which acts similar to Gαq and causes intracellular release of Ca^{2+}. However, even if this is regarded as the canonical TAS2R signal transduction cascade, there are indications of alternative signaling components as well.[123] GPCRs can also signal through G-protein independent pathways, such as the mitogen activated protein kinase (MAPK) pathway. [124]

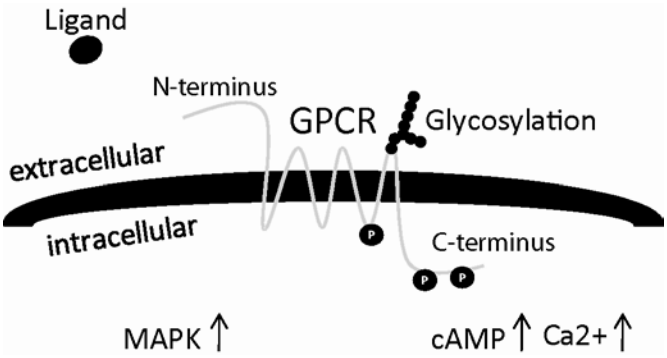

Figure 4*. A schematic picture of a G protein-coupled receptor. When a ligand attaches to the extracellular part of cell-membrane spanning GPCR (in grey) the receptor gets activated and the intracellular part induces a downstream signal cascade (e.g. cAMP, Ca^{2+} and MAPK). Many GPCRs are glycosylated which can modify the function of the receptor. P indicates phosphorylation sites.*

After activation, GPCRs are rapidly desensitized by phosphorylation and arrestin binding. It is mainly the serine (S) and threonine (T) residues within the C-termini and the third intracellular loop that are responsible for these events.[125] After desensitization, the receptor is internalized and targeted for degradation, redirected signaling through G-protein independent pathways (e.g. MAPK) or recycled back to the membrane.[124]

Phosphorylation site-directed mutations at the C-termini can severely impair both the ability to undergo phosphorylation and to recruit arrestin.[126, 127] Conformational changes improving C-terminal phosphorylation can also enhance arrestin binding and endocytosis.[128] The NPSR1 variants possesses distinct C-termini where the A variant contains five unique phosphorylation sites, and the B variant only two.

3 AIMS OF THESIS

The overall aim with this thesis was to investigate asthma susceptibility genes and their genetic role, biological function, as well as how they interact in a context-dependent manner, either with other genes or with environmental factors. We also aimed to define global gene expression patterns in asthma that could reveal novel insight about characteristics of severe therapy-resistant asthma in children.

The specific aims were:

I. To investigate the genetic role of Tenascin C in asthma and allergy and the biological and genetic interplay between the two asthma susceptibility genes Tenascin C and Neuropeptide S receptor 1.

II. To explore the interactive and biological effects between Neuropeptide S receptor 1 and environmental exposures related to farming lifestyle.

III. To examine downstream properties and functional differences between the two Neuropeptide S receptor 1 variants A and B.

IV. To identify differences in global patterns of gene expression between severe therapy-resistant asthma and controlled asthma in children.

4 MATERIALS AND METHODS

4.1 STUDY SUBJECTS AND MATERIALS

4.1.1 PARSIFAL (I, II)

The cross-sectional PARSIFAL study (Prevention of Allergy, Risk factors for Sensitization In children related to Farming and Anthroposophic Lifestyle) includes 14,893 school children 5 to 13 years from 5 Western European countries (Austria, Germany, the Netherlands, Sweden and Switzerland). The PARSIFAL study was originally designed to investigate the role of different lifestyles and environmental exposures in farm children, Steiner school children, and two corresponding rural and urban/suburban reference groups, respectively. The aim was to identify protective factors for development of asthma and allergic disorders. In Austria, Germany, the Netherlands and Switzerland, the recruitment of children were made at schools in rural areas known to have a high percentage of farmers. In Sweden, recruitment was made through the Farming Registry at the National Bureau of Statistics. Children with anthroposophic lifestyle were collected from Steiner schools and reference groups were enrolled from the corresponding geographical areas. The children's parents completed a detailed questionnaire on allergic diseases, infectious history and environmental exposures, and blood samples were obtained from the children after informed consent from the parents.[129] In the present studies (I, II) 3,113 children with available DNA and consent for genetic analysis (1,579 boys and 1,534 girls), were included.

The PARSIFAL material has been intensely investigated. The first report concluded that growing up on a farm, and leading an anthroposophic life style (to a lesser extent) protects from both sensitization and other allergic diseases in children.[129]

4.1.1.1 Group and outcome definitions

The *farm children* were defined as children currently living on a farm, and their reference group was recruited from children in the same area that did not meet the inclusion criteria for the farm children. The *Steiner school children* were recruited among children attending Steiner schools, whose families often act in accordance with an anthroposophic lifestyle. Their reference group was recruited from children attending other schools in similar suburban/rural areas.

Current rhinoconjunctivitis symptoms were defined as sneezing, runny nose, nasal block and itchy eyes in the child during the past 12 months without having a cold at the same time. *A doctor's diagnosis of asthma* was considered to be present for children reporting ever having been diagnosed with asthma, or with obstructive bronchitis more than once. *Current wheezing* was defined as at least one episode of wheezing during the past 12 months and *current atopic eczema* if the child had ever had an itchy rash intermittently for at least 6 months and, in addition, reported an itchy rash at any time during the past 12 months. *Atopic sensitization* was defined as at least one allergen-specific serum IgE test ≥0.35 kU/L or ≥3.5 kU/L against a

mixture of common allergens (Phadiatop®, e.g. pollens, cat, dust mite and mold), and/or a mixture of common food allergens (fx5®, e.g. milk, egg, fish and peanut), respectively. The term *Allergic symptoms* was defined by the combination of *doctor's diagnosis of asthma* and/or *current rhinoconjunctivitis.*

4.1.1.2 Investigated environmental exposures

The investigated environmental exposures (II) were; *ever regular farm animal contact, current regular farm animal contact, current regular visits to the barn, mother's farm animal contact during pregnancy, mother worked regularly in stable or barn during pregnancy ever consumption of farm milk. Current* refers to exposures that occurred during the past 12 months, and *regular* refers to exposures that occurred at least once a week. For *current regular farm animal contact* timing of the initiation of such regular exposure was further investigated; current regular contact with initiation in *first year of life, after first year but more than a year ago, past 12 months,* with *no current contact* as reference.

4.1.1.3 Definition of covariates

The following confounding or potentially confounding factors were included in the association and interaction analysis (study I, II): country-of origin (Austria/Germany/the Netherlands/Sweden/Switzerland), study group (farm/farm reference/Steiner school/Steiner reference), heredity (mother´s or father´s reported asthma and/or rhinoconjunctivitis) and gender.

4.1.2 Severe Asthma in Sweden (IV)

The severe asthma study in Sweden is a cross-sectional, observational study on children 6-18 years of age with problematic severe asthma and their reference group consisting of peers with controlled persistent asthma recruited from 14 clinics around Sweden.[130] The aim was to investigate differences between subtypes of asthma by exploring heritability, ethnology, patophysiology and clinical presentation. The main inclusion criteria to the study were a doctor diagnosis of asthma and daily high or low dose administration of inhaled corticosteroids (ICS). The patients were asked to fill out a questionnaire and were characterized by spirometry, methacholine provocation, measurement of fraction of nitric oxide in exhaled (FE_{NO}) and nasal air, and blood sampling for inflammatory biomarkers, atopy, DNA and RNA analyzes. Based on the clinical characterization the children were categorised to suffer from either problematic severe asthma (PA, n=54) or controlled persistent asthma (CA, n=39). Out of 93 children RNA was sampled from 57 patients. A group of healthy agematched controls were also recruited at Astrid Lindgren's Childrens Hospital, Stockholm. Informed consent was obtained from all participating children and their parents. The same team of two did all the sample collection and clinical characterization of the patients. [130]

Previous results from the study report that children with PA more often had parents with asthma, came from families with a lower socioeconomic status, were less physically active and had more comorbidity with rhinoconjunctivitis than children with

CA. Children with PA also exhibited lower FEV_1 values and increased bronchial hyper-responsiveness compared to children with CA.[130]

4.1.2.1 Outcome definitions

For gene expression analysis patients defined as *severe asthmatics* (SA) were recruited from the group of problematic severe asthma patients (PA), that were resistant to therapy and not exposed to pets in the household, smoking or other aggravating factors (n=20). *Controlled asthmatics* (CA) were recruited from the controlled, persistent asthma group (n=20), and *healthy controls* (ctrl) (n=19) were used as a reference group. The most severely affected patients from the SA (less than 19 points) and the least severely affected patients from the CA group (more than 20 and less than 24 points) were selected based on the self-reported asthma control using the Asthma Control Test® (ACT) score.[131, 132]

4.1.3 Cells and cell lines (I, II, III, IV)

Various cells and cell lines have been used to enable *in vitro* and *in vivo* studies. In study I and III, we primarily worked with NPSR1 overexpressing cells due to the relatively low endogenous expression of NPSR1 in most cell lines. The foremost cell line for NPSR1 overexpressing studies have been the human embryonic kidney cells (HEK293). The HEK293 cells have been either stably expressing NPSR1-A (I) or transiently transfected with NPSR1-A or -B (III). The vector used is under a pCMV promoter[89] and both NPSR1-A and -B constructs harbour the more potent 107^{Ile} isoform.[96] For replication and verification of results seen in the HEK293 cells NPSR1-A or -B transiently transfected human epithelial lung carcinoma cells (A549) and human neuroblastoma cells (SH-SY5Y) were used (III). In the experiments, NPSR1 overexpressing cells were stimulated with NPS, the ligand for NPSR1, for a pre-defined time and the results obtained by comparing the stimulated cells against the non-stimulated control cells (I, III).

Freshly drawn peripheral whole blood was used for study II and IV. In study II NPSR1-A and -B expression was investigated before and after lipopolysaccharide (LPS) stimulation (a proxy for farm animal exposure) of monocytes from healthy donors. In study IV, whole white blood cells from asthmatic patients and healthy controls were freshly obtained and immediately stored in *RNAlater*® until RNA extraction and later used for genome-wide RNA expression analysis.

4.2 GENETIC ANALYSIS

4.2.1 Genotyping (I)

DNA for all genotyping in the PARSIFAL samples was extracted from whole blood (Sweden, Switzerland, and the Netherlands), or from buffered white blood cells (Germany and Austria) using Qiagen kits. When performing genotyping, there is a choice ranging from a targeted small-scale approach up to a whole-genome large-scale

approach. If a low number of SNPs are to be investigated, that can be easily achieved by TaqMan® SNP genotyping assays where one or a few SNP are analyzed at the time. For whole-genome genotyping, the SNP genotyping arrays discussed earlier can be used. For an intermediate number of SNPs (e.g. >10) to be analyzed, we used MALDI-TOF mass spectrometry (matrix-assisted laser desorption/ionization-time of flight; Sequenom GmbH) (I, II). Using this technology, primer sequences for 25-30 SNPs can be pooled and analyzed at the same time in one sample. In study I, twelve SNPs in *TNC*, tagging the 50% of the most 3' 50 kb of the gene (*Fig. 5*) were successively genotyped. Seven SNPs tagging a haplotype block in *NPSR1* (*Fig. 3*) [76] had previously been genotyped in PARSIFAL [80] and were also used for downstream applications in these studies (I, II).

4.2.2 Association analysis (I)

Allelic association was analyzed in cases versus controls in the PARSIFAL material (I) using χ^2-test in Haploview 3.2[133]. Odds Ratios (OR) and 95% Confidence Intervals (CI) were calculated. Block-wise inheritance (haplotypes) was estimated using the LD measure D' in Haploview defined by Gabriel *et al*.[134] Haplotype association was tested in the statistical software 'R' (http://www.R-project.org) using the haplo.score algorithm implemented in the haplo.stats package. A permutation test was used to correct for multiple testing, and adjustment for the covariates; country-of origin, sampling group, and gender was performed on the haplotype analysis.

4.2.3 Interaction analysis (I, II)

Interaction analyses were performed in the PARSIFAL material both between genes (I), gene-gene interaction, and between gene and environmental exposures (II), gene-environmental interaction. This was done to investigate if the risk effect of one variable is dependent on the other. A multiple logistic regression model was used to test for interactions by adding an interaction term between the genotypes, or genotype and the environmental exposure investigated. *P*-values for departure from a multiplicative interaction model on the OR scale were obtained by likelihood-ratio tests between the models with and without interaction term. Permutation tests were used to estimate a global empirical *p*-value for interaction (see 4.5.1). The final regression models were adjusted for the confounders, or potential confounders; gender, study group, country-of-origin (I) and heredity (II).

4.3 RNA EXPRESSION ANALYSIS

When RNA expression is investigated, the extracted RNA molecules are used as templates for reverse transcription and the synthesizing of complementary DNA (cDNA) or RNA (cRNA). The cDNA or cRNA molecules are targeted by primers and/or hybridized to probes and the number of molecules measured by fluorescence intensity.

4.3.1 Expression arrays (III, IV)

To investigate global gene expression, we applied microarray technology. On the microarray, a number of DNA probes for each gene are spotted, labeled cDNA (Gene ST 1.0) or cRNA (HGU133plus2) are hybridized to the probes and fluorescence intensity measured. For transiently NPSR1-A or -B overexpressing HEK293 cells stimulated with NPS (III), the Affymetrix HGU133plus2 expression array was used. These arrays primarily tag the expression of the 3' end of the genes. When we investigated differential gene expression in white blood cells from severe asthmatics, controlled asthmatics and healthy controls (study IV), the Affymetrix Human Gene ST 1.0 expression array was used. These arrays are newer and give a more complete coverage of the whole transcriptome. Hybridizations and scannings were carried out using standard protocols for gene expression (www.affymetrix.com). Pre-processing and normalization of the data was performed using the statistical software 'R' by implementing relevant packages from Bioconductor (http://www.bioconductor.org) and KTH (http://www.biotech.kth.se/molbio/microarray). Six hybridizations were performed in study III, two technical replicates each of NPS-NPSR1-A, NPS-NPSR1-B and NPS-HEK293/empty vector. In study IV, 59 hybridizations were performed (20 SA, 20 CA and 19 Ctrls). After normalization and quality control, five samples were excluded as outliers, resulting in data for 17 SA, 19 CA and 18 Ctrls used in downstream analysis.

To assess differentially expressed genes, two different approaches were used. For the NPSR1-A or -B overexpressing cells (III), a t-test statistics, comparing the contrasts; NPS-NPSR1-A vs. NPS- HEK293/empty vector and NPS-NPSR1-B vs. HEK293/empty vector, were used with a B-value > 7 as cut-off for significantly differentially expressed genes. The B-statistics is the log-odds that the gene is differentially expressed and a B value of 7 corresponds to a probability of 99.99% that the gene is differentially expressed. Gene lists were generated for differentially expressed genes in NPSR1-A and -B separately. When we investigated differentially expressed genes in WBC from asthmatics (IV) an F-statistics as a global test (combines the t-statistics for all contrasts into an overall test of significance for that gene) was used to identify differentially expressed genes in any of the contrasts; MA-Ctrl, SA-MA, SA-Ctrl. The t-statistics were then used to reveal the contribution of each contrast to the global test and an adjusted p-value ≤ 0.05 was regarded as significant.

The enrichment analyses (IV) were assessed using the DAVID 6.7 database (http://david.abcc.ncifcrf.gov).[135, 136] Enrichment implies that there is a significantly larger amount of genes forming e.g a pathway in your data set than you would expect by chance. In this study, enrichment was considered significant if the adjusted p-value was ≤ 0.05 relative to the affymetrix Human Gene ST 1.0 background. Both enrichment for predefined pathways and enriched biological processes were investigated.

4.3.2 Quantitative real-time polymerase chain reaction (I, II, III)

To investigate differential gene expression in a more targeted approach, we used quantitative Real-Time Polymerase Chain Reaction (qRT-PCR). qRT-PCR was

performed on both NPS stimulated NPSR1 overexpressing cells (I, III) and LPS stimulated fresh monocytes (II). The genes targeted for qRT-PCR included both the expression of NPSR1 itself (II, III) and a selected group of NPS-NPSR1 downstream genes (I, III). When running qRT-PCR, you have an option of a primer-probe-primer combination or a primer-primer combination. The lack of a probe decreases the specificity but increases the sensitivity. Throughout study I, II and III, a primer-primer combination was used applying SYBR® Green chemistry and the PCR was run on a 7500 Fast Real-Time PCR system (Applied Biosystems). Relative quantification and calculation of the range of confidence was performed with the comparative ΔΔCT method.[137] To statistically verify differential gene expression, the paired Wilcoxon signed rank test (II) or Student's t-test (III) was used.

4.4 FUNCTIONAL ANALYSIS (I, II, III)

A selection of functional analyses has been performed to generate hypotheses for, strengthen, or verify the genetic findings. Immunohistochemistry (IHC) with antibodies targeting NPSR1-B and TNC was used (I) to identify the localization, and potential co-localization of the two proteins in bronchial biopsies from asthmatic patients and healthy controls. Fluorescence-based quantification, flow cytometry, with antibodies targeting NPSR1-A or NPSR1-B (II) or a common antibody targeting the N-terminal of both NPSR1 isoforms (III) was used to determine the NPSR1protein expression on the cell membrane w/wo NPS (III) or LPS (II) stimulation. Upon stimulation of GPCRs there is an accumulation of intracellular second messengers such as cAMP and Ca^{2+}. To measure cAMP accumulation after NPS stimulation of NPSR1-A or -B overexpressing cells, an enzyme immune assay was used. To measure the Ca^{2+} accumulation, cells were loaded with Fluo-4 NW and the increased fluorescence measured (III). Activation of transcription factors was investigated by a reporter assay for the three signaling pathways; cAMP/PKA, MAPK/JNK and MAPK/ERK, in which the relative luciferase activity was assayed. We also applied phosphorylation site-directed mutagenesis on the C-termini of NPSR1-A and -B to test if differences in signaling properties were due to dissimilar number of phosphorylation sites in the C-terminus. Serine (S) and threonine (T) residues were exchanged for a neutral alanine (A). These modified construct were then used for transient transfection and qRT-PCR as described in section 4.3.2.

4.5 STATISTICS

4.5.1 Multiple testing

Generally, the more tests that you run, the greater are the chances that false positives will be included as significant observations. A false positive (type I error) is when a result will appear as significant without any true effect. To control for type I errors, one can use various forms of multiple testing corrections.

This phenomenon was mentioned briefly in section 2.3.3.4 on GWA studies, where a huge number of tests are performed. However, it is also important to consider multiple testing when performing regular case-control association studies, as well as interaction

studies, since it will guide you on how true your positive findings are. The multiple testing problems in GWA studies are usually approached by a Bonferroni correction, where the significance level of 0.05 is divided by the number of tests performed. This is a very conservative approach and might disregard true positives (false negative, or type II error). In study I and II we instead used permutation tests to correct for multiple testing. A permutation test (also called a randomization test) treats the genotypes or haplotypes as fixed, while the disease association status (cases or controls) is randomized. This provides each individual with a new affection status. A number (e.g. >1000) of randomized association tests are executed and the proportion where a stronger association is found there than in the actual data provides the empirical *p*-value of the observation. This is done in a similar way for interactions.

In the genome-wide array expression analysis (study III, IV), a Benjamini Hochberg (BH) test was used for multiple testing correction, both to detect differentially expressed genes and to identify significantly enriched pathways when performing gene enrichment analysis. The BH method aims to control for false discovery rate (FDR) which is defined as the ratio between expected false positives to the total number of significant results. After correction, a new adjusted *p*-value can be provided.

4.5.2 Population structure

Population stratification is defined as a biased mixture in the genetic background of a population.[138] That is, if cases are collected from one geographic region and controls from another, the allele frequency of the investigated SNP might be different depending on the geographic region, rather than true difference between cases and controls. The "region" usually refers to different countries but this phenomenon can also occur within a country if little mixture between the people has occurred. The consequence of population stratification is identification of false positives. There are also other factors (e.g. gender, age) that can potentially bias the result in a similar way, i.e. the results are not a consequence of what you investigate but due to an underlying factor. All these factors can be included in models used for analyzing the data and thus attempted to be corrected for (I, II).

5 RESULTS AND DISCUSSION

5.1 ASSOCIATIONS (I)

An association study compares the frequency of alleles in controls and cases. A significant single SNP association takes place when one allele has an altered frequency in cases compared to controls. The corresponding allele is then referred to as a risk (or protective) allele. Haplotypes are investigated in a similar manner.

In study I we investigated *TNC* association to childhood allergy- and asthma-related traits by genotyping 12 SNPs in the PARSIFAL material. The whole ~100kb *TNC* gene splits up in two large LD blocks and the SNPs were picked to tag the most 3' LD block (~50kb) which previously had shown to be associated to asthma. [117] Figure 5 illustrates the LD block composed of the genotyped *TNC* SNPs and their relative position in the gene. Association was detected between several of the SNPs in *TNC* to *current rhinoconjunctivitis*, with the most significant associations to SNP rs3789873 (p=0.002). Weaker associations were seen for phenotypes such as *doctor's diagnosis of asthma* and *atopic sensitization*. Haplotype analysis revealed three blocks, as defined by Gabriel *et al*[134], with block 1 showing the strongest association to *current rhinoconjunctivitis* (p=0.0005). These results indicate that *TNC*, previously thought of mainly as a biomarker for asthma and inflammation[107-112], and with limited genetic studies in asthma and allergy related disorders[117], also is genetically associated to childhood rhinoconjunctivitis. However, rather than one causative SNP, there seemed to be many SNPs all contributing with smaller effects to the risk of disease. This might also be a consequence of the tight LD between the SNPs.

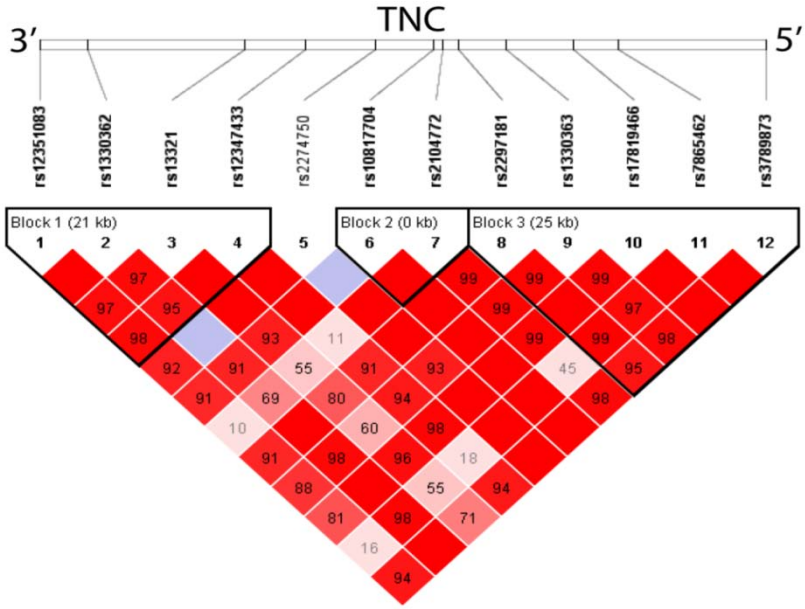

***Figure 5**. An LD plot assembled by the 12 TNC SNPs genotyped in the PARSIFAL study group. The LD plot, as defined by Gabriel et al [134], divides TNC in three haplotype blocks 1, 2 and 3.The numbers in each box correspond to the pair-wise linkage disequilibrium (LD) coefficient D' between SNPs, given in percentage. A D'=1 (indicated here by an empty square) indicates that no recombination has occurred and a D' of <1 indicates that recombination has occurred. (From study I).*

As for all genetic analyzes, it is always important to have clearly defined groups of healthy and affected idividuals. Since the PARSIFAL material is designed as a cross-sectional epidemiological study and not a case-control material in the proper sense, it might not be optimal for gene association (or gene interaction analysis, see section 5.2). The cases are much fewer than the controls and when investigating one outcome as cases there might be a mixture of the other outcomes in the controls. This lack of "clean controls" might lead to false negative results. For some outcomes, we attempted to remove those control subjects with closely related allergic phenotypes. However, due to loss of power there were no dramatic changes in the results. Nevertheless, the main risk of having mixed control groups is that we fail to identify true associations or interactions (false negatives), rather than identifying false positive association.

Many of the small genetic association studies that have been performed to date are now outshined by large-scale GWA studies. A massive number of patients and controls (sometimes more than 25, 000) [15] are recruited, often a fusion of already collected case-control materials. This is a very powerful tool to identify novel genes that are thought to have an overall large impact on the disease. However, if designed in a proper way, there are some advantages for the small scale candidate studies still. In large scale studies, it is challenging to have a universal classification of the phenotypes. Even though proper guidelines are utilized, a doctor´s diagnosis of asthma defined by one clinician might not be the same as a doctor´s diagnosis of asthma defined by another. Fusing these materials might thus lead to dilution of the genetic effects impacting each sub-phenotype. In small scale studies the phenotypic characteristics might be more homogenous enabling to identify genes with a specific role in a specific subtype of allergic diseases. Large scale studies also make it more cumbersome to pursue with investigating small genetic effects, which in combination with other small genetics effect or environmental factors might lead to an interactive increased risk effect.

5.2 INTERACTIONS (I, II)

Gene-gene or gene-environment interactions take place if the risk effect of one outcome (gene or environment) is dependent on the other outcome (gene or environment). That is, the risk (or protective) effect of the combination of two genes is significantly higher than simply adding, or multiplying those two risk effects up. In this thesis (study I and II) we have used a multiplicative model to test for interaction and investigated a departure from the odds ratio (OR) scale. A likelihood ratio test (LRT) was then used to compare the models with and without interaction terms and to test the null hypothesis of no interaction. The gained p-value indicates whether the effect (OR) of one genotype is altered by the effect of another genotype or environmental exposure.

As both *NPSR1* and *TNC* have been implicated as risk factors for asthma, and since *TNC* is regulated by NPS-NPSR1 activation[99], we aimed to explore the risk modification, or gene-gene interaction, between *TNC* and *NPSR1* in the PARSIFAL material (study I). We analyzed the 12 SNPs genotyped in *TNC* and 7 *NPSR1* SNPs previously genotyped and moderately associated with asthma and atopic sensitization in PARSIFAL.[80] The results showed that several significant interactions were taking

place between the *TNC* and *NPSR1* in subjects with *atopic sensitization* or *doctor's diagnosis of asthma.* The most significant interaction was seen between *TNC* rs3789873 and *NPSR1* rs324377 (p=0.002). However, there was no combination of SNPs that seemed to have a greater impact over the risk modification than the other combinations. The interaction analysis indicated that, depending on the *NPSR1* genotype, *TNC* variants can be associated with both an increased and a decreased risk of disease, and that several *TNC*NPSR1* SNP combinations can lead to similar risk modification. Interestingly, the interactive effects between the two genes were not seen in phenotypes were *TNC* by itself had a strong main effect (e.g. rhinoconjunctivitis). If a strong main effect is seen by one gene it might be less likely to be altered by another gene in that specific setting.

Previous studies on *NPSR1* have presented results suggesting that some of the conflicting association studies might be due to gene-environmental interactions.[77, 78] Since the PARSIFAL material has a wealth of information about environmental exposures related to farming lifestyle, we used the 7 *NPSR1* SNPs genotyped in PARSIFAL[80] to investigate potential gene-environmental interaction (study II). Farming is believed to be a protective environmental factor for the development of asthma and allergic disease, especially if exposure takes place during pregnancy or early life, involves contact with multiple animal species, and is combined with consumption of farm milk.[129, 139-145] In study II, we assessed whether the protective effects of farm related exposures was influenced by *NPSR1*. The protective effects of *regular farm contact, current regular visit to the barn, mother's farm animal contact during pregnancy, mother worked regularly in stable or barn during pregnancy* and *ever consumption of farm milk* were tested. When analyzing the combined phenotype *allergic symptoms* strong effect modification was seen for *current regular farm animal contact* especially by SNP rs323922 (p=0.001) and rs324377 (p=0.002). The protective effects of the other variables against development of allergic symptoms were not dependent on genotype. When investigating the effect of the timing of *current regular farm animal contact* there was a tendency towards a protective effect, regardless of genotype, if contact was initiated during the first year of life. However, if the contact was initiated after early infancy, the environmental effect differed considerably depending on the *NPSR1* background (*Fig. 6*).

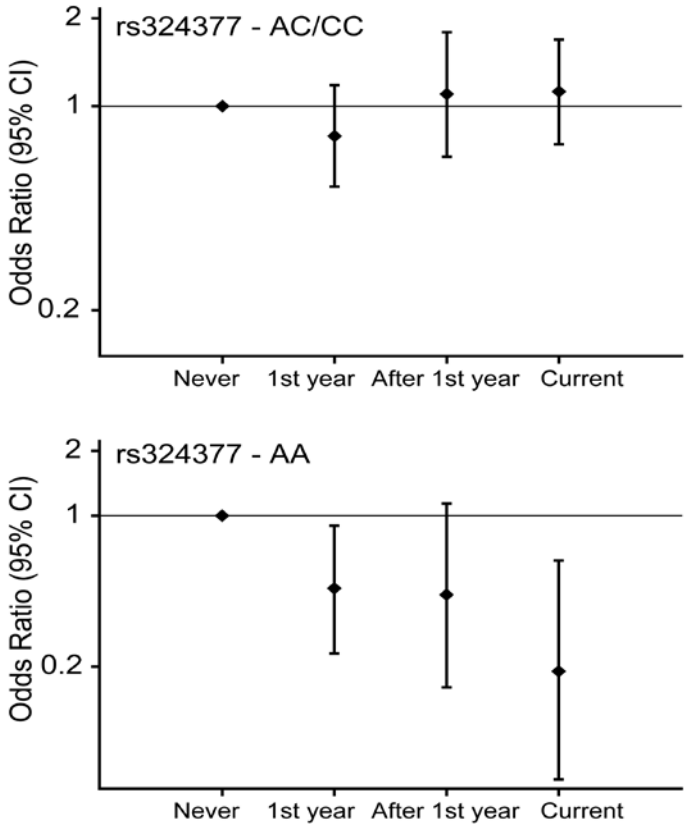

***Figure 6**. The effect of current regular farm animal contact on allergic symptoms is dependent on NPSR1 genotypes in the PARSIFAL children. Being homozygote for AA at rs324377 gives a protective effect against allergic symptoms, when current regular farm contact was initiated within last twelve months. Global p-value for interaction, p=0.003. All groups are compared to the reference group that has no current farm animal contact. (From study II).*

These results show that the effect of environment can differ depending on the genetic background. They also propose that when a strong protective effect from the environment already exists, as the documented protective effect of farm animal contact during pregnancy and early infancy, genotype interaction might have less effect. Later in life, when the environmental effect is not as dominant, the interplay with genetic factors becomes more important for the total outcome. Which is the same phenomenon as we observed in the *TNC* and *NPSR1* interaction analysis.

Both these studies highlight the importance of taking interactions into account when performing genetic or epidemiological studies. They also show that the results are not always straight forward to interpret. A risk genotype can be reversed to a non-risk genotype depending on the combination with another genotype in another gene. When interpreting the effect of the interaction (e.g. are the combinations of genotypes leading to risk or protective effect) the effect is also always mirrored to what you compare to, i.e. the baseline. The phenomenon of risk effects being reversed has been described as the flip-flop phenomenon [146] and discussed in the sense that many replication studies can replicate an association for a SNP, but in one population one allele is associated to

risk and in another population, the other allele. A possible explanation for this could be interactions both with other genes and environmental factors.

Even if interaction studies are very important in resolving the complex inheritance pattern, random interaction analysis will result in vast problems with multiple testing. In study I and II we evaluated already identified risk-modifying genes and genes with known or suspected biological or environmental interactions. This approach results in fewer tests and may be a more trustworthy finding since there is a relevant hypothesis behind the test. It is also a good approach to start with, even though the scope of novelty might be restricted. Even if the analyses are supported by a hypothesis, interaction analyses can still be troublesome. Due to the many different combinations investigated, relatively large materials are needed to have enough statistical power. For example, when a gene-gene interaction analysis is performed there will be 9 different genotype combinations (Table 1).

Table 1. *Illustrating the possible combinations for a gene-gene interaction*

	Gene 1		
Gene 2	Common HoZ	HeZ	Rare HoZ
Common HoZ	1	2	3
HeZ	4	5	6
Rare HoZ	7	8	9

HoZ; homozygote, HeZ; heterozygote

To be able to compare all possible combinations, all boxes must be filled with subjects carrying the specific combination of genotypes in gene 1 and gene 2. This might be easily achieved when common SNPs with high minor allele frequencies are used, but when rare alleles are investigated, box no 9 might not be easy to fill unless the number of study subjects is large. In study I and II, the power problem was addressed by combining phenotypes (e.g. *allergic symptoms* defined by the combination of *doctor's diagnosis of asthma* and/or *current rhinoconjunctivitis*) or genotypes, e.g. heterozygotes (HeZ) and rare homozygotes (HoZ) were in some cases combined to one outcome. Larger study groups to increase the power are necessary, but have to be approached with care. With larger study groups follows less homogenous phenotypes and instead of increased power we end up with less power to detect the specific interactive effects that might cause one specific subtype of asthma. This may be the reason why genome-wide interaction studies have not been very successful so far.[59]

Interactions do not only occur between genes or between gene and environment, but can also take place within a gene, intragenic interaction (Orsmark Pietras *et al*, unpublished data). *NPSR1* is a relatively large gene (~220kb), which splits up in several LD blocks. The size of the gene makes it likely that the blocks are inherited independently of each other. This opens up for the possibility that when risk SNPs located in different LD blocks are inherited together on the same allele, they might interact and give a joint risk effect. Unpublished data show that this phenomenon occurs in *NPSR1* (Orsmark Pietras *et al*, unpublished data). This might be a possible explanation to why *NPSR1* is very often identified as an asthma and allergy

susceptibility gene[76-84], but no causative SNP has been identified yet. Possibly, many SNPs with a moderate risk effect (that might not always show up as significant in association studies) act together to give a joint risk effect.

In study I and II we only investigated interactions between two factors (gene-gene or gene-environment). An issue that needs to be addressed is that in complex diseases there are probably multiple factors (many genes and many environmental factors) that interact. All these interactive effects might be one possible explanation for the "missing heritability" described earlier. It is possible that many small changes within the same gene or pathway (gene-gene, gene-environment, epigenetics) might act together and result in the same phenotypic characteristics e.g. asthma. These alterations do not necessarily have to be the same in each individual, hence confusing genetic analyses.

5.3 GENE EXPRESSION (I, II, III, IV)

To investigate the function of a gene, one can assess its RNA expression, i.e. identify where in the body the gene is expressed and how the expression of a gene alters upon various stimuli, changed conditions or disease states. This can be performed *in vivo* by investigating the expression in freshly sampled cells and tissue, or *in vitro* by manipulating cells in culture. There are several approaches to analyze RNA expression; either the expression of the gene itself is investigated, or the downstream effects that the gene of interest is causing. In this thesis, expression arrays (Affymetrix) have been used to detect global expression patterns (study III and IV) and qRT-PCR have been used for a more targeted gene-by-gene approach (study I, II, III).

In study I and III, NPSR1 overexpressing cells grown in culture were used to investigate the downstream effect of NPS stimulation. Global gene expression patterns downstream NPS stimulation of NPSR1 receptor variant A signaling has previously been assessed by microarray analyzes on NPSR1-A stably overexpressing HEK293 cells.[99] The study identified "cell proliferation", "morphogenesis" and "immune response" as the most enriched biological processes. One of the target genes, showing a $\geq$ 4-fold upregulation, was Tenascin C (*TNC*). In study I we investigated this regulation further and identified a dose-dependency between increased TNC expression and increased doses of NPS (0.1, 1 and 5µM). This implies a biological relationship between NPSR1 and TNC which lead us to further investigate the genetic interactions described earlier in this thesis (see 5.2). In study III, we investigated potential differences in downstream gene expression between the two NPSR1 variants A and B. HEK293 cells transiently overexpressing NPSR1-A or -B were stimulated with NPS and global gene expression assessed using microarray technology (*Fig. 7 a*).

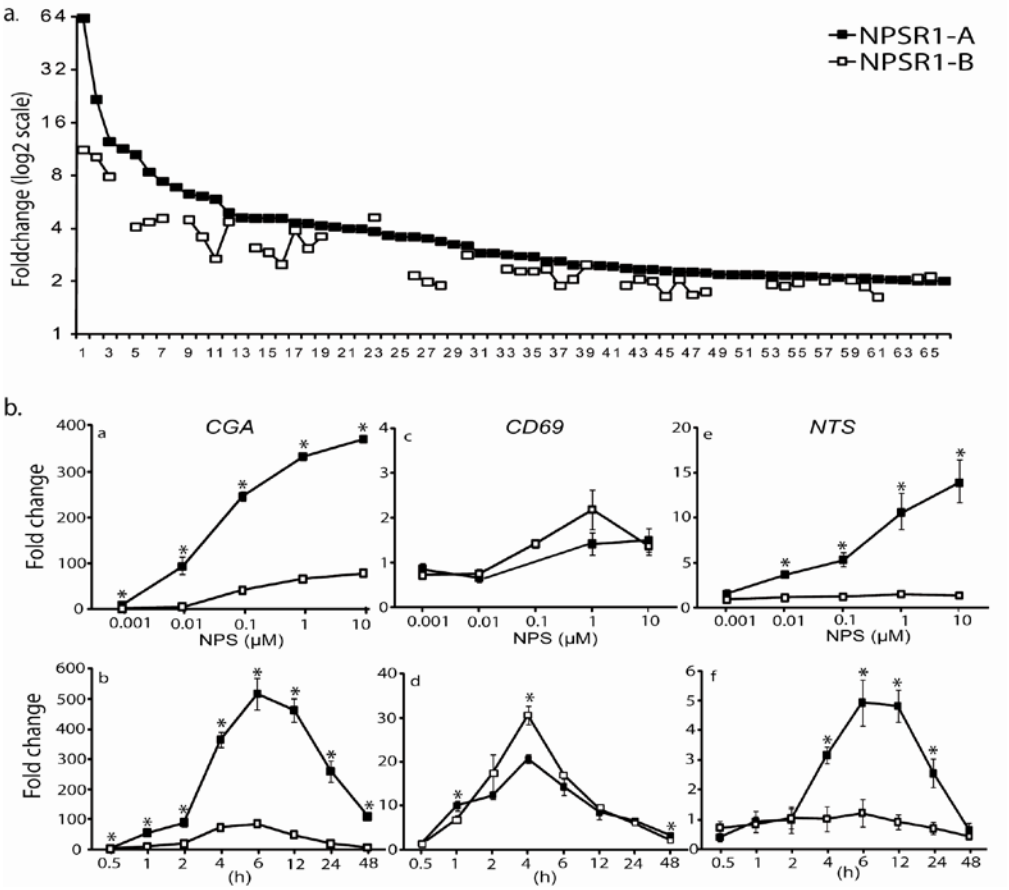

Figure 7*. Gene expression measured downstream NPS stimulation of transiently NPSR1-A or NPSR1-B overexpressing HEK293 cells. (a) NPSR1-A and -B differentially expressed genes as assessed using affymetrix expression-arrays revealed 66 genes induced by the -A isoform (cut off = ≥2 fold) of which 44 were induced by -B. (b) qRT-PCR verification showing relative expression of CGA, CD69 and NTS after NPS-NPSR1 concentration- and time-response experiments. Data shown as means relative to an NPS stimulated empty vector control ± SEM. * indicates at which concentrations or time points the difference between NPSR1-A and -B was significant (p ≤ 0.05).(From study III).*

The results demonstrated that both receptor variants in principal regulated the same set of genes, but that genes regulated by variant A constantly were induced to a higher degree. This was verified by qRT-PCR on 11 genes from the array experiment, where we further investigated both dose (1nM-10µM) (represented by *CGA*, *CD69* and *NTS* in *Fig. 7a*, *c*, *e*) and time (0.5-48h) (represented by *CGA*, *CD69* and *NTS* in *Fig. 7b*, *d*, *f*) dependency of NPS stimulation. The only notable exception was *CD69* (*Fig. 7c*, *d*), which was induced to a higher degree by receptor variant B. *CD69* is an early activator of regulatory T cells (T_{regs}), which are important in monitoring the balance between T_H1 cells and T_H2 cells. Recently, CD69 was reported to specifically control the pathogenesis of allergic airway inflammation.[147] The expression of the 11 genes from the array experiment was also investigated in two additional NPSR1-A or -B overexpressing cell lines, a lung cell line (A549) and a neuroblastoma cell line (SH-SY5Y). The results showed a similar trend, i.e. receptor variant A was a stronger inducer of gene expression, in these cells as well. Since these two isoforms have been shown to have explicit roles in asthma and allergy[76, 86, 89, 91, 92], these results might provide an isoform-specific link to pathogenetic processes in allergic airways (more

data on the differences between the two receptor variants are presented and discussed in section 5.4)

In study II, NPSR1 RNA expression was investigated, using qRT-PCR, in freshly obtained human monocytes cultured and stimulated for 6h and 24h with or without LPS, a potential proxy for farm animal exposure. The monocytes, derived from 10 healthy blood donors showed a modest trend towards upregulation of NPSR1-A and -B (however less for B) mRNA expression after 6h of LPS stimulation.

Since most genes have a distinct expression pattern in different cells, and the expression is regulated by a number of factors unique for that cell, it is difficult to decide what the proper cell types for investigations are. When assessing effects downstream NPS-NPSR1 signaling we have often used the HEK293 cell line (study I, III). It has the advantage of being a fast growing and easily transfected cell line. Previous studies on the NPS-NPSR1 system have also utilized HEK293 cells which makes the results easier to compare. However, one should be aware of that results obtained from one cell line might not be representative to other cells, or *in vivo*. This is why it is important to investigate several different cell lines (as in study III when we confirmed the results obtained by experiments in HEK293, in A549 and SH-SY5Y cells). In study II, we investigated how the expression of NPSR1 changed in monocytes upon LPS stimulation. Even if the choice of cell type is well motivated, as in study II when NPSR1 expression previously had been identified in monocytes [148], it is difficult to decide beforehand what the proper cells or tissues for investigation are. The assumed function of a gene might not always be the same as the true function. In addition, a gene can have different function in different settings. The solution, if there is one, might be to perform the experiments in many different cells and tissues and at different conditions.

In study IV, we investigated global gene expression in whole white blood cells (WBCs) obtained from healthy controls (Ctrl, n=18), children with persistent but controlled asthma (CA, n=19) and children with severe therapy resistant asthma (SA, n=17). The aim was to elucidate if there were differences in gene expression between subgroups of asthma. We identified 1378 genes that were differentially expressed (DE) in any of the comparisons CA-Ctrl, SA-CA and SA-Ctrl, with 355 genes exclusively DE in SA compared to both Ctrl and CA. In an attempt to control for the possibility that differences in expression not was due to differences in cell count between the groups (since WBCs were used), eosinophils and neutrophils were compared, and no significant difference in cell count was detected between SA and CA. We also investigated differences in other clinical characterization between the SA and CA groups used in this study (which were a selection of patients from the original study).[130] Apart from differences seen in asthma control score test (lower in SA) and dose of inhaled corticosteroid (higher in SA), which were the inclusion criteria for expression analysis, a significant differences was seen for methacholine responsiveness ($DRS_{methacholine}$, slope of the dose-respons curve for provocation with methacholine) (higher in SA, p=0.04). There were no significant differences seen for FEV_1 (forced expiratory volume during 1s), total WBC (white blood cells), FE_{NO} (fraction of nitric oxide in exhaled air) or total IgE.

To elucidate how homogenous the expression of these genes was in the pre-defined groups; Ctrl, CA and SA, we performed unsupervised hierarchical clustering and generated a heatmap of the 1378 DE genes (*Fig. 8*). The cluster analysis did not reveal three perfect column dendogram clusters, but rather illustrates that although extensively clinically characterized patients, there are some overlaps between the groups when it comes to gene expression. Previous cluster analysis with several different categories, such as phenotype and clinical testing parameters, show that we need to take a number of parameters into consideration in order to achieve the well-characterized subgroups of asthma that are needed to identify genetic risk factors.[10] The row clusters in the heatmap nicely show the pattern of up- or down-regulated genes in each subject.

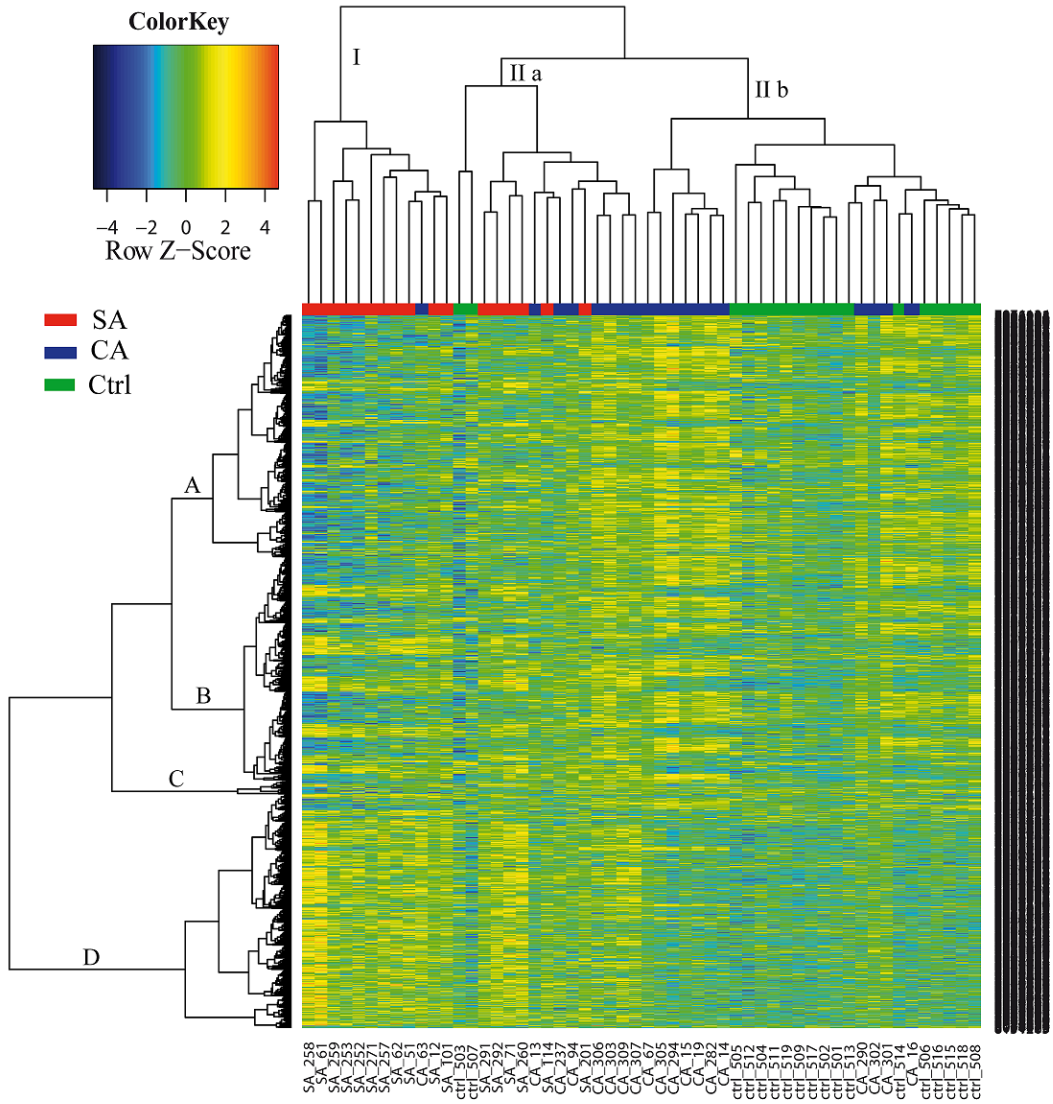

Figure 8. *Global gene expression patterns in asthma patients and healthy controls. Unsupervised hierarchical clustering and heat map illustrates each individual's expression pattern in all 1378 significant ($p \leq 0.05$) differentially expressed genes. Normalized gene expression is indicated by the row Z-score where yellow-red represents up- and blue down-regulated genes. SA, severe asthma; CA, controlled asthma; Ctrl, healthy controls. (From study IV).*

After identifying genes that could tell the groups apart, we performed gene enrichment analysis to investigate biologically relevant groups and pathways that were significantly enriched among these genes. Three pathways were identified; Bitter taste transduction (upregulated mostly in SA), Natural killer cell mediated cytotoxicity (mostly upregulated in CA) and N-Glycan biosynthesis (downregulated in SA).

The bitter taste receptor family (TAS2Rs) has recently been linked to a protective role in asthma pathogenesis. Expression of TAS2Rs has been identified on the motile cilia of human airway epithelial cells where stimulation with bitter compounds increases the ciliary beat frequency and enables a more efficient clearance of e.g. mucus out of the lung.[118] TAS2Rs have also been identified in human airway smooth muscle (ASM) where bitter tastants caused relaxation of isolated ASM and dilation of airways. Inhaled bitter tastants also caused decreased airway obstruction in a mouse model of asthma.[120] TAS2Rs have however not been connected to asthma and allergy outside the airway system. Here, we report upregulation of TAS2Rs in human white blood cells in patients suffering from asthma. The detailed role of taste receptors expressed in peripheral blood cells, however, warrants further investigation. Intriguingly, the upregulation of TAS2Rs was mostly seen in the severe therapy resistant asthmatics and this underlines the need for additional studies in order to better understand the mechanisms behind therapy-resistant asthma.

Natural killer (NK) cells are a type of cytotoxic lymphocytes that cause the target cell to die by apoptosis. Their role in asthma has been debated and studies have both shown a vast upregulation of NK T-cells in bronchoalveolar-lavage fluid from asthmatic patients [149], as well as no increase at all.[150] Murine models have supported the role of NK T-cells in asthma by showing that NK T-cell deficient mice do not develop, or showed impaired, allergen-induced airway hyperreactivity.[151, 152] Our results point towards a role for NK T-cells in asthma.

The N-glycanbiosynthesis pathway is responsible for the making of glycans that couples to proteins and lipids on the cell membrane. Glycosylation is one of the most common post-translational modifications, and almost half of all proteins are glycosylated.[153] Glycosylation patterns change extensively with T cell development and differentiation and recent data also states that these changes play a powerful role in regulating T cell responses.[154] It has been suggested that alterations in the distribution of glycoproteins at the cell surface contributes to many chronic human diseases, including autoimmunity.[155] Of note, NPSR1 and TAS2Rs both display N-glycosylation which is involved in trafficking of the receptor to the cell surface, and functioning of the receptor once attached in the cell membrane. For TAS2Rs, N-glycosylation has been reported to play a role in the maturation of the receptors, but is has been implicated that the N-glycan pattern does not have a major impact on the function of either NPSR1 or the TA2SRs.[156, 157]

A large number of the differentially expressed genes in SA were identified as ncRNA. ncRNA has a important role in transcriptional and post-transcriptional regulation.[25] We also identified 12 genes out of a list of 97 well-replicated asthma susceptibility genes in our list of 1378 DE genes, of which four were identified from GWA studies in asthma (*RORA, PDE4D, IL2RB* and *ORMDL3*). The majority were upregulated in CA, which

might be a consequence of that most association studies/GWAs are performed on a more controlled asthma phenotype.

Taken together, the data indicate a separation in gene expression patterns between children with severe, therapy resistant asthma and controlled asthma. It also reveals novel pathways characterizing the severe thearapy-resistent asthma phenotype.

In study IV, whole white blood cells were used for investigation of gene expression. As discussed above, the problem of knowing which cell type to use for investigation is applies here as well. Using whole white blood cells for our study was first of all motivated by the fact that the immune system is a large component in the etiology of asthma, secondly by the fact that withdrawal of blood is a less invasive procedure for the children compared to e.g. an airway biopsy. An obvious disadvantage is that we do not know which of the cell/cells in WBC that contributes to the effect seen. The observed effect might also be diluted by the fact that cell types showing differential expression for a certain gene are mixed with cell types not showing any differential expression. To obtain freshly separated cell populations for such a large number of study subjects (n=60) collected all over the country was however not feasible since RNA is a very instable molecule and will rapidly degrade if not immediately taken care of in a proper way.

When assessing direct changes in the DNA (as discussed in 2.3.2.1), cell type and timing is less important. When investigating indirect changes e.g. methylation studies and RNA expression, both the cell type and the timing matters. Following stimulation of cells with an agent (e.g. NPS or LPS), the gene expression appears different if you wait 2, 6 or 24 h after stimulation. In the same manner, a blood sample retrieved from a patient one day might have a totally different expression profile the next day. Taking this in to account, the chance of identifying differentially expressed genes acting in a common network, and expressed at the same time and in the same manner within a group of patients might be slim. When such genes nonetheless are identified, it is strongly indicative of them being important players.

The array technology, followed by next generation sequencing creates, enormous amounts of data. We need to learn how to sieve through the data flow in order to extract the maximal amount of information. Even when data from a relatively small study, as in study IV, are explored, there might be more information than we can process. As an example, the gene ontology or enrichment analysis commonly investigated in these types of studies can only identify already pre-defined pathways and categories. Even though identification of such pathways, as the upregulation of bitter taste receptors in severe asthmatics, might lead to novel and significant information, there might still be important data that we miss, due to the fact that we do not know what to search for.

5.4 BIOLOGICAL EFFECTS (I, II, III)

In the insatiable search for new disease candidate genes, supported by the technology explosion in genomics, it should be acknowledged that it is seldom the finding of the gene that cures the disease, but unraveling the function of it might do so. In this thesis

we have mainly performed functional studies on one of the asthma susceptibility genes, *NPSR1*.

The downstream signaling properties of the NPSR1 variant A stimulated with its ligand NPS have been previously investigated by our group [99], which led us in study III to investigate the signaling properties downstream NPSR1 variant B, and compare that to variant A. In contrast to our previous investigation of NPSR1-A signaling (utilizing a stable overexpressing HEK293 cell line) [99], we used transiently NPSR1-A and -B overexpressing HEK293 cells this time. The protein expression of both receptor variants on the plasma-membrane was determined by flow cytometry, using antibodies targeting the common N-terminal. The flow cytometry analysis revealed that the intensity of antibody-bound receptor on each cell was similar, both before and after NPS stimulation with equal number of positive cell counts, showing that the transfection efficiency between the two isoforms was similar, as well as the number of membrane bound receptors. When the downstream second messenger response of cAMP and Ca^{2+} was assayed, a similar pattern as for the expression analysis was seen. Taken together, NPSR1 variant A generated a greater response than variant B, which was also seen by downstream activation of various transcription factor complexes.

Previous studies investigating the functional properties of NPSR1 have mainly focused on differences in the NPSR1 variant A expressing either the amino acid 107^{Asn} or 107^{Ile} in the first extracellular loop of the receptor protein. Receptors carrying the 107^{Ile} variant has been proven to be more potent compared to those expressing the 107^{Asn} variant.[96-98, 158] Some studies have also assessed differences between variant A and B (expressing either 107^{Asn} or 107^{Ile}).[96-98] The constructs used in our studies both express 107^{Ile}, which is also the only present variant in rodents. The results from previous studies have varied from no differences in second messenger response between A and B [96, 97], to reported differences[98]. There have also been problems with achieving NPSR1-A and -B transient transfection systems to work, and in cases where stable systems have been compared to transient systems, the result has not been entirely similar. Our results (functional assays together with expression data), however, point towards differences in downstream effect between the two receptor isoforms, with NPSR1 variant A being a stronger inducer of downstream signaling than B.

A functioning phosphorylation of GPCRs is important for desensitization and rapid turnover, altering these sites has been demonstrated to affect the internal signaling properties of GPCRs.[125-128] Phosphorylation sites are situated in the third intracellular loop and in the C-terminus of GPCRs (see *Fig. 4).* Since the NPSR1 variants display distinct C-terminus, the number of phosphosrylation sites differ between the receptors, the A variant contains five whereas variant B only harbors two. In an attempt to understand what causes the different signaling properties between the receptor variants, we examined phosphorylation sites in the C-terminus. Residues in the C-terminal targeted for phosphorylation were exchanged to a neutral alanine, which diminished the phosphorylation sites. Expression of downstream target genes was investigated in NPS stimulated HEK293 cells overexpressing the altered NPSR1 constructs. However, alterations of phosphorylation sites could not explain the differences in signaling properties between the two receptor variants.

In study II, we investigated NPSR1-A and -B expression on the cell membrane of white blood cells with flow cytometry. CD14+ PBMCs from 10 blood donors were stimulated with or without LPS and specific antibodies for the two receptor variants were used for detection. Both isoforms were detected on the cell membrane in a similar intensity, but only the A variant responded to LPS treatment with a significant upregulation of the receptor protein on the cell membrane. In study I, we used IHC for detection of NPSR1-B and TNC protein in bronchial biopsies obtained from asthmatic patients and healthy controls. Both genes exhibited an upregulation in the asthmatic patients. NPSR1-B was expressed in the epithelial cells, whereas TNC was mainly observed in the sub-epithelial basement membrane layer.

Taken together, functional studies on candidate genes can help us to understand why a gene might be a genetically linked susceptibility gene for disease. As discussed in section 5.3, it is important to perform the experiments using different cell types and tissues to get a full picture, but all pieces of the puzzle are important.

6 CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Complex diseases are called “complex” for a reason. Most likely, many small changes in the genome, which differ between individuals, and environmental influences may give rise to the same phenotypic feature. Large scale *a priori* studies, such as GWA studies and NGS, are crucial to identify where to start looking. However small, more targeted studies with clearly defined phenotypes are needed to unravel pathways and networks leading to the specific disease state. The genes and changes along these pathways might be distinct in different groups of patients, however in the end they will make up the same phenotype. Once pathways and networks leading to a certain disease are defined, we can pinpoint target genes for therapy. These target genes might not be identified through genetic analysis, but the genetic analysis made it possible to identify where to look.

The results from one single study (association, interaction, expression or biological) are regarded as interesting. However, replication of a finding is what validates the results, and all these studies (study I, II, III, IV) would benefit from independent replications.

There is a discrepancy in the expression pattern between the two NPSR1 variants where e.g. A protein is uniquely expressed in the enteroendocrine cells in the gut. [91] The A variant is also the only variant upregulated in monocytes, both on RNA and protein level, by LPS stimulation. LPS also acts as a ligand for Toll-like receptors, which are expressed on enteroendocrine cells.[159] The B variant, on the other hand, is upregulated in asthmatic airways and the only gene that was stronger induced by NPS-NPSR1-B signaling was CD69, an early activator of T_{regs}. By creating a transgenic NPSR1-A or -B overexpressing mouse one could expand the environmental findings and investigate a possible connection between NPSR1 and microbial environment. It would also be interesting to assess differences between NPSR1 variant A and B in mouse models of asthma.

Our identification of upregulated bitter taste receptor (TAS2R) mRNA expression in white blood cells, primarily seen in severe therapy-resistant asthmatic childen, is a novel finding. Taken together with recent data which reported that TAS2Rs in the airways seems to act through a protective mechanism, where e.g. administration of bitter tastants caused relaxation of the airways in a mouse model of allergic inflammation, this network of TAS2Rs might provide totally new insight into the characteristics of asthma. However, the role of TAS2Rs in immune cells needs to be thoroughly investigated. First of all our finding needs to be replicated in an independent, however phenotypically closely related, study group. Secondly, by separating white blood cells in their subgroups (neutrophils, eosinophils, basophils, lymphocytes (B, T etc.) and monocytes), the exact expression pattern of TAS2Rs needs to be examined. Thirdly, the role that TAS2Rs play in the immune cells needs to be identified. The third point could be addressed *in vitro* by investigating the effect of stimulating cells with agonists for bitter taste receptors, by introducing antagonists or selective knock-down of the receptors in functional studies, or by assessing the

expression pattern of TAS2Rs after interventions that change the function of immunocompetent cells. The cause of severe therapy-resistant asthma is poorly understood and novel therapeutic approaches are required for this group of patients. The findings of the significantly upregulation of the TAS2R pathway in this specific group of patients is a new discovery that should be further exploited. It introduces the potential of an important endogenous protective system that may be a new therapeutic target in asthma.

Epigenetic studies are not within the scope of this thesis. It would, however, be interesting to investigate if the protective effect of prenatal and early-life farm exposure partially are due to epigenetic changes, and if these changes can be passed on to the next generation. The vision is that I may pass on to my children the protective effects that my mother introduced to me by living and working on a farm even though I am not currently exposed to farm life.

Clearly, more work is needed to fully understand the complex disease of asthma, but genetic studies hold the promise of providing the basis for discoveries aimed at treating asthmatic patients.

7 ACKNOWLEDGEMENTS

The work presented in this thesis has been carried out at the Department of Biosciences and Nutrition at Karolinska Institutet. I wish to express my sincere gratitude to everyone who in any way has contributed to this thesis.

To all children and families that participated in the PARSIFAL or Severe Asthma in Sweden study, without you none of this could have been done.

Juha, min inspirerande huvudhandledare. Du har kanske inte alltid varit den mest närvarande handledaren vilket många gånger lett till en viss frustration… Men när du väl tar dig an har du en förmåga att fokusera, uppmuntra och att alltid se saker från den ljusa sidan. Du har en positiv attityd som smittar av sig. Jag vill även tacka dig för ditt förtroende för mitt doktorandarbete. Jag har vid flertalet tillfällen önskat bättre vägledning inom mina projekt, men trots frånvaro av detta så har jag alltid känt stöd och uppmuntran från dig till att driva projekten på mitt sätt.

Ulpu, jag beklagar djupt din bortgång. Du var min handledare under min första tid som doktorand, då inom psoriasis-projektet. Jag minns dig som en professionell, engagerad och närvarande handledare.

Mina bihandledare, Sven-Erik, Tiina och Cilla.
Sven-Erik, på grund av projekt som inte gått som planerat så har samarbetena under mina doktorandår inte alltid varit så täta. Du har dock alltid varit väldigt välkomnande och jag har alltid känt ett stort engagemang när väl samarbete pågått. Tiina, tack för en fin introduktion under mitt första projekt om psoriasis. Cilla, du blev min bihandledare lagom efter halvtid men hann snabbt lära mig en av de viktigaste sakerna i livet, att hålla fokus! Jag är även glad för din höga ambition att både sätta dig in i mina gamla projekt så väl som nya, jag har alltid känt ett stort stöd från din sida.

Mina ofrivilliga bihandledare.
Erik, du är inte min bihandledare, varken officiellt eller vidare frivilligt. Men på något sätt har jag ändå lyckats nästlat mig in och fått mer hjälp från dig än från många andra. Tack för att du har tagit din tid! Kristian, jag måste nämna även dig här. Ingen har väl så outtröttligt som du lyssnat på mina långa haranger av osammanhängande resultat, torkat mina tårar vid jobbiga arbetssituationer, kommit med förlösande förslag när inget har fungerat, och delat min glädje när framgång har gått att skönja. Du är en beundransvärd människa!

I would also like to thank some of my former supervisors that I have encountered on the road to becoming a doctor. Dimitri Tentler, my supervisor when doing my master thesis in Uppsala. Little did I know about labwork when you, with some hesitation, lent me your lab bench. But when you, as a last lesson, taught me how to drink pure vodka from a plastic cup as a true Russian at you dissertation party, then I knew that you had turned me into a skilled laboratory person and I was accepted. No one has ever taught me that much on how to do proper labwork since. Jane Gitschier, who gave me a

position in her lab at University of California San Francisco (UCSF). The paper exercise was a mountain but you never gave up and it is a priceless experience to have done a “pre-doc” before starting the actual work as a doctoral student.

JKE gruppen, gamla som nya, tack för intressanta gruppmöten, roliga gruppresor, trevliga luncher och en stimulerande forskningsmiljö. Hong, Tiina, Lisa och Marco som delat kontor med mig. Hong, en dag ska jag lära mig mer än ett ord på kinesiska. Lisa (nu) och Marco (då), tänk att ha turen att dela rum med livs levande element. Tack för all värme! Tiina, tack för att du lyssnat på mina svamlande utläggningar om försök och resultat och försökt hjälpa. Myriam, Isabel, Kristiina, Cilla, Ingegerd, Virpi, Per, Helena, Daniel, Marie, Shintaro, Hasse, Gustaf, Elo, Natalie, Mauro, Dario, Francesca A, Francesca B, Lovisa (speciellt tack för att du läste kappan), Cissi, Katja, Pu. Tänk att ha en så stor grupp runt sig med så stor expertis inom så många områden, vilken ynnest. Alla ni på MAF och BEA (nu som då), tack för all hjälp med och omkring genotypning och microarray projekten. Det är alltid ett nöje att samarbeta med er!

Alla doktorander som fanns med när jag började, Heidi, Erik, Ulf, Sara, Anna, Hanna och Kristiina, men som disputerade en efter en. Nu är det bara du och jag kvar Kristiina, men inte så länge till! TMK (Heidi, Linda, Astrid, Sara, Hanna, Anna och Kristiina), den ypperliga ventilen när frustrationen över lab och labbande behöver vädras. Hoppas att vi kan hålla traditionen vid liv.

The asthma team in Helsinki, Ville, Johanna, Sini and Lilli. Many thanks for nice collaborations and discussions.

Sven-Erik’s grupp. Speciellt tack till Anna, Mikael och Ingrid som hjälpt till med artikel II och III.

PARSIFAL-G gruppen. Göran, Annika, Anna, Marianne, Magnus, Fredrik och Erik. Jag lärde mig otroligt mycket under stötandet och blötandet av de två interaktions artiklarna. Er expertis är inspirerande.

Gruppen på Astrid Lindgrens barnsjukhus. Gunilla, Christophe, Jon, Björn, Erik, Ann, Catharina. Vilket otroligt roligt och intressant samarbete vi har initierat. Det känns alltid så mycket mer motiverande att bedriva forskning när det faktiskt handlar om riktiga patienter.

Mina vänner, insamlade från gymnasietiden i Norrköping, universitets-studier i Uppsala och tillskansade via Kristians tid i Uppsala. Kommer ni ihåg att jag finns trots att jag aldrig hör av mig? Jag tror att vi alla är i en fas i livet där vi önskar att vi kunde träffas mycket mer än vad vi gör. Men ni ska veta att jag tycker så mycket om er och det är så kul när vi väl träffas!

Mina kära, kära systrar. Anna, Eva och Gunilla. Om det är någon jag skulle ringa mitt i natten (och jag inte vill väcka Kristian) så är det ni. Ni betyder oerhört mycket för mig och att veta att ni alltid finns där är en otrolig trygghet. Ni och era familjer är allt bra härliga!

Mina kära, kära föräldrar, Ingela och Sigward. Tack för att jag fick växa upp på en bondgård, varken astma eller allergi så långt ögat kan se ☺. Tack för att ni alltid har uppmuntrat mig att se världen och studera. Ni är alltid positiva och finns där när jag behöver stöd och råd, i smått som i stort. Tack också för att ni ställt upp och tagit hand om Thea när pusslet inte alltid gått att lösa.

Thea, och du där inne som ligger och sparkar i magen, mina älskade älskade barn. Det är ni som är livet…

Kristian, min älskade älskade man, det är du som håller mig under armarna och gör att jag kan leva livet…

8 REFERENCES

1. Kerem E, Corey M, Kerem BS, Rommens J, Markiewicz D, Levison H, et al. The relation between genotype and phenotype in cystic fibrosis--analysis of the most common mutation (delta F508). N Engl J Med 1990; 323:1517-22.
2. National Asthma Education and Prevention Program. Expert panel report 3: guidelines for the diagnosis and management of asthma. Bethesda (MD): National Heart, Lung, and Blood Institute. 2007:Publication no. 07-4051. Available at: http://www.nhlbi.nih.gov/guidelines/asthma.
3. Global Initiative for Asthma. Global Gtrategy for Asthma Management and Prevention (GINA). Bethesda (MD): National Institutes of Health; National Heart, Lung, and Blood Institute Updated 2008:Available at: http://www.ginasthma.org.
4. Lai C, Beasley R, Crane J, Foliaki S, Shah J, Weiland S. Global variation in the prevalence and severity of asthma symptoms: Phase Three of the International Study of Asthma and Allergies in Childhood (ISAAC). Thorax 2009.
5. Martinez FD. Genes, environments, development and asthma: a reappraisal. Eur Respir J 2007; 29:179-84.
6. Lang A, Carlsen KH, Haaland G, Devulapalli CS, Munthe-Kaas M, Mowinckel P, et al. Severe asthma in childhood: assessed in 10 year olds in a birth cohort study. Allergy 2008; 63:1054-60.
7. Ait-Khaled N, Pearce N, Anderson HR, Ellwood P, Montefort S, Shah J. Global map of the prevalence of symptoms of rhinoconjunctivitis in children: The International Study of Asthma and Allergies in Childhood (ISAAC) Phase Three. Allergy 2009; 64:123-48.
8. Lemanske RF, Jr., Busse WW. Asthma: clinical expression and molecular mechanisms. J Allergy Clin Immunol 2010; 125:S95-102.
9. Lotvall J, Akdis CA, Bacharier LB, Bjermer L, Casale TB, Custovic A, et al. Asthma endotypes: a new approach to classification of disease entities within the asthma syndrome. J Allergy Clin Immunol 2011; 127:355-60.
10. Moore WC, Meyers DA, Wenzel SE, Teague WG, Li H, Li X, et al. Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. Am J Respir Crit Care Med 2010; 181:315-23.
11. von Mutius E, Vercelli D. Farm living: effects on childhood asthma and allergy. Nat Rev Immunol 2010; 10:861-8.
12. Lowe L, Custovic A, Woodcock A. Childhood asthma. Curr Allergy Asthma Rep 2004; 4:159-65.
13. Sarafino EP, Goldfedder J. Genetic factors in the presence, severity, and triggers of asthma. Arch Dis Child 1995; 73:112-6.
14. Los H, Postmus PE, Boomsma DI. Asthma genetics and intermediate phenotypes: a review from twin studies. Twin Res 2001; 4:81-93.
15. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. N Engl J Med 2010; 363:1211-21.
16. Jackson DJ, Johnston SL. The role of viruses in acute exacerbations of asthma. J Allergy Clin Immunol 2010; 125:1178-87; quiz 88-9.
17. Melen E, Himes BE, Brehm JM, Boutaoui N, Klanderman BJ, Sylvia JS, et al. Analyses of shared genetic factors between asthma and obesity in children. J Allergy Clin Immunol 2010; 126:631-7 e1-8.

18. Foliaki S, Pearce N, Bjorksten B, Mallol J, Montefort S, von Mutius E. Antibiotic use in infancy and symptoms of asthma, rhinoconjunctivitis, and eczema in children 6 and 7 years old: International Study of Asthma and Allergies in Childhood Phase III. J Allergy Clin Immunol 2009; 124:982-9.
19. Strachan DP. Hay fever, hygiene, and household size. Bmj 1989; 299:1259-60.
20. Bjorksten B. The hygiene hypothesis: do we still believe in it? Nestle Nutr Workshop Ser Pediatr Program 2009; 64:11-8; discussion 8-22, 251-7.
21. Schaub B, Lauener R, von Mutius E. The many faces of the hygiene hypothesis. J Allergy Clin Immunol 2006; 117:969-77; quiz 78.
22. Ege MJ, Mayer M, Normand AC, Genuneit J, Cookson WO, Braun-Fahrlander C, et al. Exposure to environmental microorganisms and childhood asthma. N Engl J Med 2011; 364:701-9.
23. Vercelli D. Mechanisms of the hygiene hypothesis--molecular and otherwise. Curr Opin Immunol 2006; 18:733-7.
24. Clamp M, Fry B, Kamal M, Xie X, Cuff J, Lin MF, et al. Distinguishing protein-coding and noncoding genes in the human genome. Proc Natl Acad Sci U S A 2007; 104:19428-33.
25. Taft RJ, Pang KC, Mercer TR, Dinger M, Mattick JS. Non-coding RNAs: regulators of disease. J Pathol 2010; 220:126-39.
26. Bird A. Perceptions of epigenetics. Nature 2007; 447:396-8.
27. Baye TM, Martin LJ, Khurana Hershey GK. Application of genetic/genomic approaches to allergic disorders. J Allergy Clin Immunol 2010; 126:425-36; quiz 37-8.
28. Watson J, Crick F. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 1953; 171:737-8.
29. Finishing the euchromatic sequence of the human genome. Nature 2004; 431:931-45.
30. 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.
31. 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.
32. Lander ES. The new genomics: global views of biology. Science 1996; 274:536-9.
33. 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.
34. Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES. High-resolution haplotype structure in the human genome. Nat Genet 2001; 29:229-32.
35. Metzker ML. Sequencing technologies - the next generation. Nat Rev Genet 2010; 11:31-46.
36. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature 2009; 461:747-53.
37. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996; 273:1516-7.
38. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. PLoS Biol 2010; 8:e1000294.
39. Vercelli D. Gene-environment interactions in asthma and allergy: the end of the beginning? Curr Opin Allergy Clin Immunol 2010; 10:145-8.
40. Boone C, Bussey H, Andrews BJ. Exploring genetic interactions and networks with yeast. Nat Rev Genet 2007; 8:437-49.

41. Phillips PC. Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 2008; 9:855-67.
42. Moore JH, Williams SM. Epistasis and its implications for personal genetics. Am J Hum Genet 2009; 85:309-20.
43. 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.
44. Bottema RW, Nolte IM, Howard TD, Koppelman GH, Dubois AE, de Meer G, et al. Interleukin 13 and interleukin 4 receptor-alpha polymorphisms in rhinitis and asthma. Int Arch Allergy Immunol 2010; 153:259-67.
45. 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.
46. Huebner M, Kim DY, Ewart S, Karmaus W, Sadeghnejad A, Arshad SH. Patterns of GATA3 and IL13 gene polymorphisms associated with childhood rhinitis and atopy in a birth cohort. J Allergy Clin Immunol 2008; 121:408-14.
47. Melen E, Umerkajeff S, Nyberg F, Zucchelli M, Lindstedt A, Gullsten H, et al. Interaction between variants in the interleukin-4 receptor alpha and interleukin-9 receptor genes in childhood wheezing: evidence from a birth cohort study. Clin Exp Allergy 2006; 36:1391-8.
48. Bottema RW, Kerkhof M, Reijmerink NE, Thijs C, Smit HA, van Schayck CP, et al. Gene-gene interaction in regulatory T-cell function in atopy and asthma development in childhood. J Allergy Clin Immunol 2010; 126:338-46, 46 e1-10.
49. Eder W, Klimecki W, Yu L, von Mutius E, Riedler J, Braun-Fahrlander C, et al. Toll-like receptor 2 as a major gene for asthma in children of European farmers. J Allergy Clin Immunol 2004; 113:482-8.
50. 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.
51. 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.
52. Simpson A, John SL, Jury F, Niven R, Woodcock A, Ollier WE, et al. Endotoxin exposure, CD14, and allergic disease: an interaction between genes and the environment. Am J Respir Crit Care Med 2006; 174:386-92.
53. Bieli C, Eder W, Frei R, Braun-Fahrlander C, Klimecki W, Waser M, et al. A polymorphism in CD14 modifies the effect of farm milk consumption on allergic diseases and CD14 gene expression. J Allergy Clin Immunol 2007; 120:1308-15.
54. Smit LA, Siroux V, Bouzigon E, Oryszczyn MP, Lathrop M, Demenais F, et al. CD14 and toll-like receptor gene polymorphisms, country living, and asthma in adults. Am J Respir Crit Care Med 2009; 179:363-8.
55. Eder W, Klimecki W, Yu L, von Mutius E, Riedler J, Braun-Fahrlander C, et al. Association between exposure to farming, allergies and genetic variation in CARD4/NOD1. Allergy 2006; 61:1117-24.
56. Melen E, Nyberg F, Lindgren CM, Berglind N, Zucchelli M, Nordling E, et al. Interactions between glutathione S-transferase P1, tumor necrosis factor, and traffic-related air pollution for development of childhood allergic disease. Environ Health Perspect 2008; 116:1077-84.

57. 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.
58. Bottema RW, Reijmerink NE, Kerkhof M, Koppelman GH, Stelma FF, Gerritsen J, et al. Interleukin 13, CD14, pet and tobacco smoke influence atopy in three Dutch cohorts: the allergenic study. Eur Respir J 2008; 32:593-602.
59. Ege MJ, Strachan DP, Cookson WO, Moffatt MF, Gut I, Lathrop M, et al. Gene-environment interaction for childhood asthma and exposure to farming in Central Europe. J Allergy Clin Immunol 2011; 127:138-44, 44 e1-4.
60. Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M. Mapping complex disease traits with global gene expression. Nat Rev Genet 2009; 10:184-94.
61. Howarth PH, Babu KS, Arshad HS, Lau L, Buckley M, McConnell W, et al. Tumour necrosis factor (TNFalpha) as a novel therapeutic target in symptomatic corticosteroid dependent asthma. Thorax 2005; 60:1012-8.
62. Berry MA, Hargadon B, Shelley M, Parker D, Shaw DE, Green RH, et al. Evidence of a role of tumor necrosis factor alpha in refractory asthma. N Engl J Med 2006; 354:697-708.
63. Prefontaine D, Lajoie-Kadoch S, Foley S, Audusseau S, Olivenstein R, Halayko AJ, et al. Increased expression of IL-33 in severe asthma: evidence of expression by airway smooth muscle cells. J Immunol 2009; 183:5094-103.
64. Agache I, Ciobanu C, Agache C, Anghel M. Increased serum IL-17 is an independent risk factor for severe asthma. Respir Med 2010; 104:1131-7.
65. Chen E, Miller GE, Walker HA, Arevalo JM, Sung CY, Cole SW. Genome-wide transcriptional profiling linked to social class in asthma. Thorax 2009; 64:38-43.
66. Subrata LS, Bizzintino J, Mamessier E, Bosco A, McKenna KL, Wikstrom ME, et al. Interactions between innate antiviral and atopic immunoinflammatory pathways precipitate and sustain asthma exacerbations in children. J Immunol 2009; 183:2793-800.
67. Schulze TG, McMahon FJ. Defining the phenotype in human genetic studies: forward genetics and reverse phenotyping. Hum Hered 2004; 58:131-8.
68. Melén E, Kho AT, Sharma S, Gaedigk R, Leeder SJ, Mariani TJ, et al. Expression analysis of asthma candidate genes during human and murine lung development. Manuscript.
69. Vercelli D. Advances in asthma and allergy genetics in 2007. J Allergy Clin Immunol 2008; 122:267-71.
70. Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. Nature 2007; 448:470-3.
71. Kim SH, Cho BY, Park CS, Shin ES, Cho EY, Yang EM, et al. Alpha-T-catenin (CTNNA3) gene was identified as a risk variant for toluene diisocyanate-induced asthma by genome-wide association analysis. Clin Exp Allergy 2009; 39:203-12.
72. Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP, et al. Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. Am J Hum Genet 2009; 84:581-93.
73. Hancock DB, Romieu I, Shi M, Sienra-Monge JJ, Wu H, Chiu GY, et al. Genome-wide association study implicates chromosome 9q21.31 as a susceptibility locus for asthma in mexican children. PLoS Genet 2009; 5:e1000623.

74. Sleiman PM, Flory J, Imielinski M, Bradfield JP, Annaiah K, Willis-Owen SA, et al. Variants of DENND1B associated with asthma in children. N Engl J Med; 362:36-44.
75. Li X, Howard TD, Zheng SL, Haselkorn T, Peters SP, Meyers DA, et al. Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions. J Allergy Clin Immunol 2010; 125:328-35 e11.
76. 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.
77. Malerba G, Lindgren CM, Xumerle L, Kiviluoma P, Trabetti E, Laitinen T, et al. Chromosome 7p linkage and GPR154 gene association in Italian families with allergic asthma. Clin Exp Allergy 2007; 37:83-9.
78. Hersh CP, Raby BA, Soto-Quiros ME, Murphy AJ, Avila L, Lasky-Su J, et al. Comprehensive Testing of Positionally Cloned Asthma Genes in Two Populations. Am J Respir Crit Care Med 2007; 176:849-57.
79. Feng Y, Hong X, Wang L, Jiang S, Chen C, Wang B, et al. G protein-coupled receptor 154 gene polymorphism is associated with airway hyperresponsiveness to methacholine in a Chinese population. J. Allergy. Clin. Immunol. 2006; 117:612-7.
80. Melen E, Bruce S, Doekes G, Kabesch M, Laitinen T, Lauener R, et al. Haplotypes of G protein-coupled receptor 154 are associated with childhood allergy and asthma. Am. J. Respir. Crit. Care. Med. 2005; 171:1089-95.
81. 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.
82. Daley D, Lemire M, Akhabir L, Chan-Yeung M, He JQ, McDonald T, et al. Analyses of associations with asthma in four asthma population samples from Canada and Australia. Hum Genet 2009; 125:445-59.
83. Castro-Giner F, de Cid R, Gonzalez JR, Jarvis D, Heinrich J, Janson C, et al. Positionally cloned genes and age-specific effects in asthma and atopy: an international population-based cohort study (ECRHS). Thorax, 2010:65:124-31.
84. Vergara C, Jimenez S, Acevedo N, Martinez B, Mercado D, Gusmao L, et al. Association of G-protein-coupled receptor 154 with asthma and total IgE in a population of the Caribbean coast of Colombia. Clin Exp Allergy 2009; 39:1558-68.
85. Vercelli D. Discovering susceptibility genes for asthma and allergy. Nat Rev Immunol 2008; 8:169-82.
86. D'Amato M, Bruce S, Bresso F, Zucchelli M, Ezer S, Pulkkinen V, et al. Neuropeptide s receptor 1 gene polymorphism is associated with susceptibility to inflammatory bowel disease. Gastroenterology 2007; 133:808-17.
87. Gottlieb DJ, O'Connor GT, Wilk JB. Genome-wide association of sleep and circadian phenotypes. BMC Med Genet 2007; 8 Suppl 1:S9.
88. Donner J, Haapakoski R, Ezer S, Melen E, Pirkola S, Gratacos M, et al. Assessment of the neuropeptide S system in anxiety disorders. Biol Psychiatry; 68:474-83.
89. 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.
90. Ulloa-Aguirre A, Conn PM. Targeting of G protein-coupled receptors to the plasma membrane in health and disease. Front Biosci 2009; 14:973-94.
91. Sundman L, Saarialho-Kere U, Vendelin J, Lindfors K, Assadi G, Kaukinen K, et al. Neuropeptide S receptor 1 expression in the intestine and skin - putative

role in peptide hormone secretion. Neurogastroenterol Motil 2009:Epub ahead of print.
92. Pulkkinen V, Majuri ML, Wang G, Holopainen P, Obase Y, Vendelin J, et al. Neuropeptide S and G protein-coupled receptor 154 modulate macrophage immune responses. Hum Mol Genet 2006; 15:1667-79.
93. 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.
94. Allen IC, Pace AJ, Jania LA, Ledford JG, Latour AM, Snouwaert JN, et al. Expression and function of NPSR1/GPRA in the lung before and after induction of asthma-like disease. Am J Physiol Lung Cell Mol Physiol 2006.
95. Gupte J, Cutler G, Chen JL, Tian H. Elucidation of signaling properties of vasopressin receptor-related receptor 1 by using the chimeric receptor approach. Proc. Natl. Acad. Sci. U S A. 2004; 101:1508-13.
96. 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.
97. Bernier V, Stocco R, Bogusky MJ, Joyce JG, Parachoniak C, Grenier K, et al. Structure-function relationships in the neuropeptide S receptor: molecular consequences of the asthma-associated mutation N107I. J Biol Chem 2006; 281:24704-12.
98. Nepomuceno D, Sutton S, Yu J, Zhu J, Liu C, Lovenberg T, et al. Mutagenesis studies of neuropeptide S identify a suitable peptide tracer for neuropeptide S receptor binding studies and peptides selectively activating the I(107) variant of human neuropeptide S receptor. Eur J Pharmacol 2010; 635:27-33.
99. Vendelin J, Bruce S, Holopainen P, Pulkkinen V, Rytila P, Pirskanen A, et al. Downstream target genes of the Neuropeptide S--NPSR1 pathway. Hum. Mol. Genet. 2006; 15:2923-35.
100. Chiquet-Ehrismann R, Chiquet M. Tenascins: regulation and putative functions during pathological stress. J. Pathol. 2003; 200:488-99.
101. Jones FS, Jones PL. The tenascin family of ECM glycoproteins: structure, function, and regulation during embryonic development and tissue remodeling. Dev. Dyn. 2000; 218:235-59.
102. Nakahara H, Gabazza EC, Fujimoto H, Nishii Y, D'Alessandro-Gabazza CN, Bruno NE, et al. Deficiency of tenascin C attenuates allergen-induced bronchial asthma in the mouse. Eur. J. Immunol. 2006; 36:3334-45.
103. Roth-Kleiner M, Hirsch E, Schittny JC. Fetal lungs of tenascin-C-deficient mice grow well, but branch poorly in organ culture. Am. J. Respir. Cell. Mol. Biol. 2004; 30:360-6.
104. Cohen ED, Ihida-Stansbury K, Lu MM, Panettieri RA, Jones PL, Morrisey EE. Wnt signaling regulates smooth muscle precursor development in the mouse lung via a tenascin C/PDGFR pathway. J Clin Invest 2009; 119:2538-49.
105. Gueders MM, Hirst SJ, Quesada-Calvo F, Paulissen G, Hacha J, Gilles C, et al. Matrix metalloproteinase-19 deficiency promotes tenascin-C accumulation and allergen-induced airway inflammation. Am J Respir Cell Mol Biol 2010; 43:286-95.
106. Meuronen A, Karisola P, Leino M, Savinko T, Sirola K, Majuri ML, et al. Attenuated expression of tenascin-C in ovalbumin-challenged STAT4-/- mice. Respir Res 2011; 12:2.
107. Agarwal AR, Mih J, George SC. Expression of matrix proteins in an in vitro model of airway remodeling in asthma. Allergy. Asthma. Proc. 2003; 24:35-42.

108. Amin K, Ludviksdottir D, Janson C, Nettelbladt O, Bjornsson E, Roomans GM, et al. Inflammation and structural changes in the airways of patients with atopic and nonatopic asthma. BHR Group. Am. J. Respir. Crit. Care. Med. 2000; 162:2295-301.
109. Karjalainen EM, Lindqvist A, Laitinen LA, Kava T, Altraja A, Halme M, et al. Airway inflammation and basement membrane tenascin in newly diagnosed atopic and nonatopic asthma. Respir. Med. 2003; 97:1045-51.
110. Laitinen A, Altraja A, Kampe M, Linden M, Virtanen I, Laitinen LA. Tenascin is increased in airway basement membrane of asthmatics and decreased by an inhaled steroid. Am. J. Respir. Crit. Care. Med. 1997; 156:951-8.
111. Phipps S, Benyahia F, Ou TT, Barkans J, Robinson DS, Kay AB. Acute allergen-induced airway remodeling in atopic asthma. Am. J. Respir. Cell. Mol. Biol. 2004; 31:626-32.
112. Torrego A, Hew M, Oates T, Sukkar M, Chung KF. Expression and activation of TGF-{beta} isoforms in acute allergen-induced remodelling in asthma. Thorax 2007; 62:307-13.
113. Bu LM, Bradley M, Soderhall C, Wahlgren CF, Kockum I, Nordenskjold M. Genome-wide linkage analysis of allergic rhinoconjunctivitis in a Swedish population. Clin. Exp. Allergy. 2006; 36:204-10.
114. Dizier MH, Bouzigon E, Guilloud-Bataille M, Betard C, Bousquet J, Charpin D, et al. Genome screen in the French EGEA study: detection of linked regions shared or not shared by allergic rhinitis and asthma. Genes. Immun. 2005; 6:95-102.
115. Wjst M, Fischer G, Immervoll T, Jung M, Saar K, Rueschendorf F, et al. A genome-wide search for linkage to asthma. German Asthma Genetics Group. Genomics 1999; 58:1-8.
116. Yokouchi Y, Shibasaki M, Noguchi E, Nakayama J, Ohtsuki T, Kamioka M, et al. A genome-wide linkage analysis of orchard grass-sensitive childhood seasonal allergic rhinitis in Japanese families. Genes Immun 2002; 3:9-13.
117. Matsuda A, Hirota T, Akahoshi M, Shimizu M, Tamari M, Miyatake A, et al. Coding SNP in tenascin-C Fn-III-D domain associates with adult asthma. Hum. Mol. Genet. 2005; 14:2779-86.
118. Shah AS, Ben-Shahar Y, Moninger TO, Kline JN, Welsh MJ. Motile cilia of human airway epithelia are chemosensory. Science 2009; 325:1131-4.
119. Sisson JH, Papi A, Beckmann JD, Leise KL, Wisecarver J, Brodersen BW, et al. Smoke and viral infection cause cilia loss detectable by bronchoalveolar lavage cytology and dynein ELISA. Am J Respir Crit Care Med 1994; 149:205-13.
120. Deshpande DA, Wang WC, McIlmoyle EL, Robinett KS, Schillinger RM, An SS, et al. Bitter taste receptors on airway smooth muscle bronchodilate by localized calcium signaling and reverse obstruction. Nat Med 2010; 16:1299-304.
121. Sternini C, Anselmi L, Rozengurt E. Enteroendocrine cells: a site of 'taste' in gastrointestinal chemosensing. Curr Opin Endocrinol Diabetes Obes 2008; 15:73-8.
122. Wong GT, Gannon KS, Margolskee RF. Transduction of bitter and sweet taste by gustducin. Nature 1996; 381:796-800.
123. Margolskee RF. Molecular mechanisms of bitter and sweet taste transduction. J Biol Chem 2002; 277:1-4.
124. Lefkowitz RJ. Seven transmembrane receptors: something old, something new. Acta Physiol (Oxf) 2007; 190:9-19.

125. Marchese A, Paing MM, Temple BR, Trejo J. G protein-coupled receptor sorting to endosomes and lysosomes. Annu Rev Pharmacol Toxicol 2008; 48:601-29.
126. Miller WE, Houtz DA, Nelson CD, Kolattukudy PE, Lefkowitz RJ. G-protein-coupled receptor (GPCR) kinase phosphorylation and beta-arrestin recruitment regulate the constitutive signaling activity of the human cytomegalovirus US28 GPCR. J Biol Chem 2003; 278:21663-71.
127. Richardson RM, Pridgen BC, Haribabu B, Snyderman R. Regulation of the human chemokine receptor CCR1. Cross-regulation by CXCR1 and CXCR2. J Biol Chem 2000; 275:9201-8.
128. Suvorova ES, Gripentrog JM, Jesaitis AJ, Miettinen HM. Agonist-dependent phosphorylation of the formyl peptide receptor is regulated by the membrane proximal region of the cytoplasmic tail. Biochim Biophys Acta 2009; 1793:406-17.
129. Alfven T, Braun-Fahrlander 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; 61:414-21.
130. Konradsen JR, Nordlund B, Lidegran M, Pedroletti C, Gronlund H, Van Hage M, et al. Problematic severe asthma: A proposed approach to identifying children who are severely resistant to therapy. Pediatr Allergy Immunol 2010; 22:9-18.
131. Liu AH, Zeiger R, Sorkness C, Mahr T, Ostrom N, Burgess S, et al. Development and cross-sectional validation of the Childhood Asthma Control Test. J Allergy Clin Immunol 2007; 119:817-25.
132. Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, Marcus P, et al. Development of the asthma control test: a survey for assessing asthma control. J Allergy Clin Immunol 2004; 113:59-65.
133. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005; 21:263-5.
134. 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.
135. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 2003; 4:P3.
136. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009; 4:44-57.
137. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001; 25:402-8.
138. Cardon LR, Palmer LJ. Population stratification and spurious allelic association. Lancet 2003; 361:598-604.
139. Braun-Fahrlander C. Environmental exposure to endotoxin and other microbial products and the decreased risk of childhood atopy: evaluating developments since April 2002. Curr Opin Allergy Clin Immunol 2003; 3:325-9.
140. 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.
141. von Mutius E, Radon K. Living on a farm: impact on asthma induction and clinical course. Immunol Allergy Clin North Am 2008; 28:631-47, ix-x.

142. von Mutius E, Braun-Fahrlander C, Schierl R, Riedler J, Ehlermann S, Maisch S, et al. Exposure to endotoxin or other bacterial components might protect against the development of atopy. Clin Exp Allergy 2000; 30:1230-4.
143. Waser M, Michels KB, Bieli C, Floistrup H, Pershagen G, von Mutius E, et al. Inverse association of farm milk consumption with asthma and allergy in rural and suburban populations across Europe. Clin Exp Allergy 2007; 37:661-70.
144. Ege MJ, Frei R, Bieli C, Schram-Bijkerk D, Waser M, Benz MR, et al. Not all farming environments protect against the development of asthma and wheeze in children. J Allergy Clin Immunol 2007; 119:1140-7.
145. Ege MJ, Herzum I, Buchele G, Krauss-Etschmann S, Lauener RP, Roponen M, et al. Prenatal exposure to a farm environment modifies atopic sensitization at birth. J Allergy Clin Immunol 2008; 122:407-12, 12 e1-4.
146. Lin PI, Vance JM, Pericak-Vance MA, Martin ER. No gene is an island: the flip-flop phenomenon. Am. J. Hum. Genet. 2007; 80:531-8.
147. Miki-Hosokawa T, Hasegawa A, Iwamura C, Shinoda K, Tofukuji S, Watanabe Y, et al. CD69 Controls the Pathogenesis of Allergic Airway Inflammation. J Immunol 2009.
148. Pulkkinen V, Haataja R, Hannelius U, Helve O, Pitkanen OM, Karikoski R, et al. G protein-coupled receptor for asthma susceptibility associates with respiratory distress syndrome. Ann Med 2006; 38:357-66.
149. Akbari O, Faul JL, Hoyte EG, Berry GJ, Wahlstrom J, Kronenberg M, et al. CD4+ invariant T-cell-receptor+ natural killer T cells in bronchial asthma. N Engl J Med 2006; 354:1117-29.
150. Vijayanand P, Seumois G, Pickard C, Powell RM, Angco G, Sammut D, et al. Invariant natural killer T cells in asthma and chronic obstructive pulmonary disease. N Engl J Med 2007; 356:1410-22.
151. Akbari O, Stock P, Meyer E, Kronenberg M, Sidobre S, Nakayama T, et al. Essential role of NKT cells producing IL-4 and IL-13 in the development of allergen-induced airway hyperreactivity. Nat Med 2003; 9:582-8.
152. Lisbonne M, Diem S, de Castro Keller A, Lefort J, Araujo LM, Hachem P, et al. Cutting edge: invariant V alpha 14 NKT cells are required for allergen-induced airway inflammation and hyperreactivity in an experimental asthma model. J Immunol 2003; 171:1637-41.
153. Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim Biophys Acta 1999; 1473:4-8.
154. Daniels MA, Hogquist KA, Jameson SC. Sweet 'n' sour: the impact of differential glycosylation on T cell responses. Nat Immunol 2002; 3:903-10.
155. Dennis JW, Nabi IR, Demetriou M. Metabolism, cell surface organization, and disease. Cell 2009; 139:1229-41.
156. Clark SD, Tran HT, Zeng J, Reinscheid RK. Importance of extracellular loop one of the neuropeptide S receptor for biogenesis and function. Peptides 2010; 31:130-8.
157. Reichling C, Meyerhof W, Behrens M. Functions of human bitter taste receptors depend on N-glycosylation. J Neurochem 2008; 106:1138-48.
158. Okamura N, Habay SA, Zeng J, Chamberlin AR, Reinscheid RK. Synthesis and pharmacological in vitro and in vivo profile of 3-oxo-1,1-diphenyl-tetrahydro-oxazolo[3,4-a]pyrazine-7-carboxylic acid 4-fluoro-benzylamide (SHA 68), a selective antagonist of the neuropeptide S receptor. J Pharmacol Exp Ther 2008; 325:893-901.

159. Bogunovic M, Dave SH, Tilstra JS, Chang DT, Harpaz N, Xiong H, et al. Enteroendocrine cells express functional Toll-like receptors. Am J Physiol Gastrointest Liver Physiol 2007; 292:G1770-83.