Technische Universität München

Institut für Experimentelle Genetik

The role of the *STAT6* gene in the IL-4/IL-13 pathway: analyses of polymorphisms, functional studies and their effect on IgE regulation and the development of asthma

Michaela Schedel

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des Akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigten Dissertation

Vorsitzender: Univ.-Prof. Dr. rer. nat. Martin Hrabé de Angelis

Prüfer der Dissertation: 1. apl. Prof. Dr. rer. nat., Dr. rer. biol. hum. habil. Jerzy Adamski

2. Priv.-Doz. Dr. med. Michael Kabesch,
Ludwig-Maximilians-Universität München

3. Univ.-Prof. Dr. med. Dr. h.c. Hermann Wagner, Ph.D. (Melbourne)

Die Dissertation wurde am 06.08.2007 bei der Technischen Universität eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 06.11.2007 angenommen.

INDEX

1 Introduction

1.1 Clinical manifestation of atopic diseases

Atopic diseases are the most common chronic disorders during childhood. The clinical expression of atopic diseases includes asthma, atopic dermatitis and allergic rhinitis (hay fever) with an overall prevalence of approximately 25% (figure 1). Atopy is characterized by an allergic hypersensitivity generally accompanied by high levels of total and specific serum Immunglobulin E (IgE). This sensitization to ubiquitous allergens can be measured by skin-prick tests positivity against inhalant or food allergens.

The most common form of childhood asthma is the allergic or IgE-associated asthma affecting over 50% of all pediatric asthma patients in Germany. Asthma may be seen as a syndrome rather than a single disease entity. Hallmark characteristics include reversible obstruction of the lung in the presence of airway inflammation and bronchial hyperresponsiveness, mucus hypersecretion and smooth muscle hypertrophy [1,2]. Among its clinical symptoms are wheeze, chest tightness, shortness of breath and cough [2]. Yet, not all asthma symptoms are necessarily present in each patient or to the same degree of severity.

In population studies it was shown that elevated total serum IgE levels were strongly associated with asthma and bronchial hyperresponsiveness in children [3,4] but have also been reported in asthmatics independently of specific sensitization [5]. Thus, it was suggested that the presence of elevated serum IgE levels may be a good indicator for asthma and atopy in children at risk [1]. Regardless of these data, it is difficult to demonstrate the precise role of IgE in the context of asthma as the underling mechanisms are yet poorly understood. It is not clear if asthma and elevated IgE are two independent characteristics of allergic diseases or if increased levels of total serum IgE are an intermediate phenotype on the way to develop asthma.

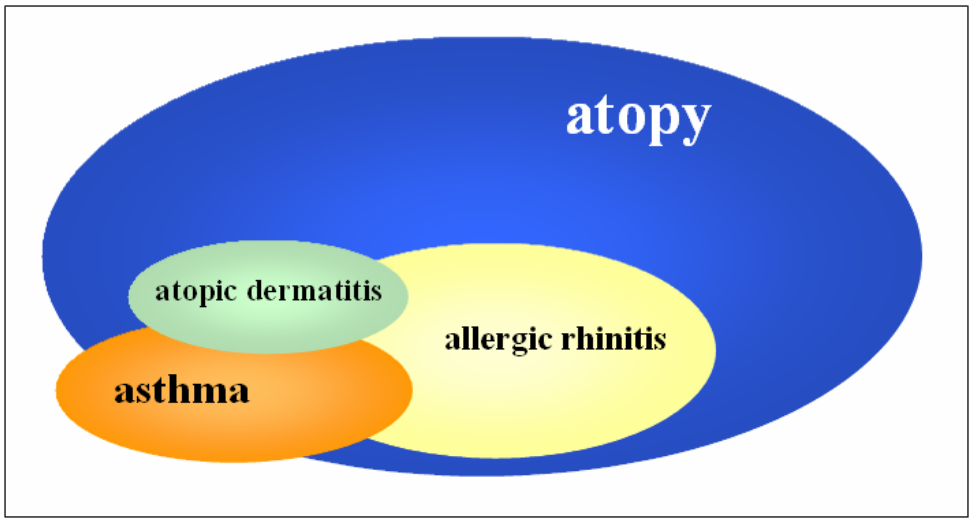

figure 1: The phenotypic heterogeneity of atopic diseases: approximately 40% of the general population are defined as being atopic by positive skin-prick-test. Asthma (10-20%), allergic rhinitis (20%) and atopic dermatitis (10-20%) considerably overlap. Modified from Feijen *et al.* [6].

1.2 Genes and environment contribute to the genesis of atopic diseases

In epidemiological studies it was shown that the prevalence for atopic diseases is yet higher in the Western world in comparison to most developing countries and has drastically increased in industrialized nations within the last decades. For example, the incidence of childhood asthma in the United States rose from 3.9% in the year 1980 up to 5.8% in 2003 [7]. Besides, a strong inconsistency in asthma rates between different populations was described. Recent trends in prevalence differ widely between populations as well as countries. While in some European and Asian countries the prevalence of asthma is still rising, other countries report a stagnation or even a slight decline in asthma prevalence rates [7].

As the clinical manifestation of allergic disorders is complex, the underlying pathophysiology remains to be unraveled. Further insight on the nature of allergic diseases has been provided by research efforts on the influence of environmental factors and genetic predisposition.

In a German study population of children with similar genetic background differences in the prevalence of atopic diseases have been described with a higher prevalence of current asthma (5.9%) and atopic sensitization (36.7%) in former West Germany (Munich) in contrast to 3.9% and 18.2% in the former eastern part (Dresden) of Germany [8]. Children from Dresden, in contrast to children from Munich, were exposed to much more microbial or infectious stimuli early in life (larger family size, more day care) putatively stimulating the immune system with signals conferring tolerance to allergens [9].

Several other population studies analyzed the relation between the prevalence of atopic diseases and the exposure to specific stimuli present in a farming environment. These data showed that children growing up on a farm had a significantly decreased risk of developing atopy and hay fever [10-12]. This effect was not detectable in children raised in the same rural area but not living on a farm. Thus, specific components exclusively present in the farming environment may prevent the development of asthma and atopic sensitization.

Besides environmental factors, genetic predisposition plays a crucial role as it has long been known that asthma clusters in families. In twin studies it was shown that the inheritance, attributable to genetic variation, was determined to be between 60 to 70 percent as monocygotic twins revealed higher asthma prevalence in comparison to dizygotic twins [6]. So far, a number of candidate genes for the development of atopic diseases have been identified by positional cloning or linkage studies in certain regions of the genome (e.g. 7p15-14 (*GPRA*) [13], 20p13 (*ADAM33*) [14] and 5q31 (*IL-4, IL-13, CD14*) [6]. An alternative approach was the hypothesis-based selection of a candidate gene. Hence, the respective gene is systematically screened for Single Nucleotide Polymorphisms (SNP) referring to a single base exchange which occurs randomly approximately every 1000bp. Detected SNPs are subsequently analyzed in population-based association studies to identify phenotype-specific alleles significantly occurring less frequent in a control population. Thus, a cluster of promising candidate genes for asthma and regulation of IgE levels have been proposed within the chromosomal region of

5q31-33, containing among others Interleukin 4 and Interleukin 13 (*IL-4* and *IL-13)*. The relation between atopic diseases and SNPs of *IL-4* and *IL-13* will be described in detail in sections 1.4.2.1 and 1.4.2.2.

However, the effect of genetic predisposition may only become detectable when certain environmental stimuli are present. Thus, it was hypothesized that both influences have to be taken into account when analyzing complex diseases such as asthma and allergies. For example, this has been demonstrated for the effect of genetic variations in the Toll-like receptor 2 (TLR2) on asthma. TLR2 is important for the response against bacterial compounds such as lipopolysaccharides (LPS) [15]. As mentioned before, children raised in a traditional farm environment showed a lower asthma prevalence, which may be due to increased exposure to microbial components. It was suggested that polymorphisms in *TLR2* may modulate the effect of microbial exposure [16]. Eder *et al*. described a significant association only in farming children carrying the polymorphic allele of a *TLR2* SNP (A-16934T). These children were less likely to be diagnosed with asthma and other atopic diseases in comparison to children expressing the wildtype allele. In contrast, no differences in the prevalence of atopic diseases depending on the genotype of A-16934T were observed among non-farming children. These results may suggest a gene-by-environment interaction between *TLR2* and immune responses against microbial components.

1.3 The T_H1/T_H2 paradigm regulates immune responses

In recent years it was proposed that the abnormal response of the airways to harmless allergens in atopic patients may largely be due to a skewed activation of naïve T-cells. Consequently, an imbalance of different types of T-helper cells (T_H1- versus T_H2-cells) was observed in asthmatics.

An immune response against pathogens like allergens or microbial components is mediated through the activation of naive $CD4^+$ T_H-cells (figure 2). Initially, naïve T-cells are functionally immature only capable to secrete the cytokine Interleukin-2 (IL-2) [17]. The activation and the differentiation of naive

$CD4^+$ T_H-cells require two independent signals. The first signal is induced through the interaction of antigen presenting cells (APC) with naive $CD4^+$ T_H-cells. The APCs, containing populations of dendritic cells (DC) or macrophages, present antigens through the Major Histocompatibility Complex (MHC) to the T-Cell Receptor (TCR) on the surface of $CD4^+$ T-cells. Once activated, naïve T-cells differentiate either into T_H1- or T_H2-cells which is determined by the release of a distinct cytokine pattern. The polarization towards T_H-subsets is clearly influenced by the activation of these two signals. However, the cytokine milieu plays the most critical role.

The T_H1 differentiation program, which protects against intracellular bacteria [18,19], is induced upon the release of IL-12 of macrophages and/or the contact with CD8 positive DCs. The activation of T_H1-cells leads to the expression of the transcription factor T-bet (T-box expressed in T-cells) in T-cells [20]. T-bet expression strongly correlates with the activation of the IFN-γ production in T_H1-cells [21]. Among other biological functions, IFN-γ is a key T_H1 factor in the activation of macrophages to up-regulate pro-inflammatory parameters such as IL-12, IL-15 and TNFα [22-24]. Hence, a positive feed-back loop on macrophages is initiated to boost the T_H1 differentiation and thus the response to microbial infections.

In contrast, the differentiation into T_H2-cells is induced through the cytokine IL-4 and/or the contact of naïve T-cells to CD8 negative DCs, followed by the expression of the transcription factor GATA3. Subsequently, T_H2-cells release both cytokines IL-4 and IL-13 which promote IgE production in human B-cells, growth and differentiation of human B-cells and monocytes [20]. T_H2-cells protect against nematodes but play also a crucial role in the allergic immune reactions [18,19].

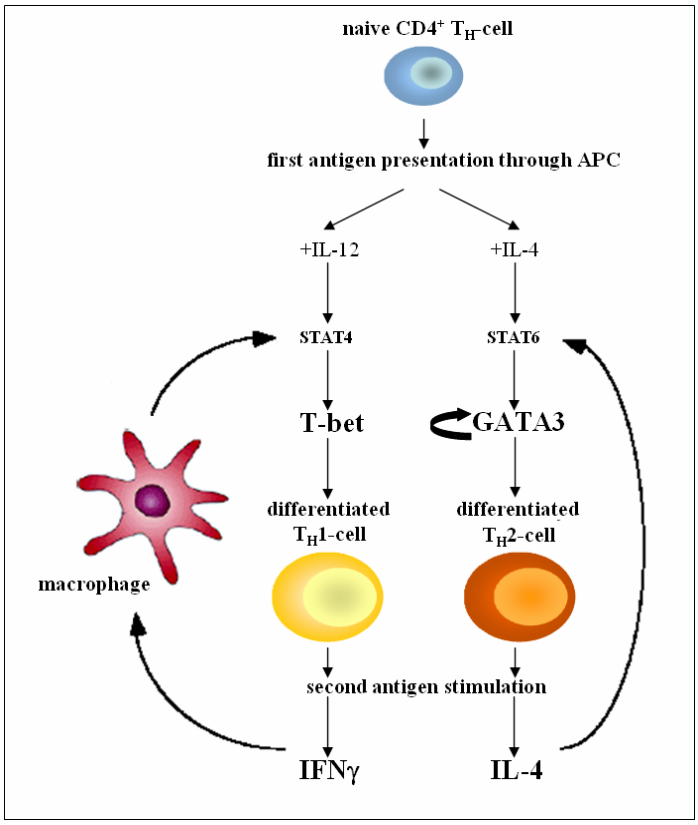

figure 2: Differentiation of naïve T-cells to T_H1- or T_H2-cells induced through the release of a distinct cytokine pattern. Modified from Rao *et al.* [25].

Based on *in vitro* and *in vivo* studies in mice as well as in humans, it was shown that the process of T_H1/T_H2 polarization is mutually regulated [26-28]. The production of IFN-γ does not only lead to the differentiation of T_H1-cells but also exerts an inhibitory function on the maturation of T_H2-cells. Likewise, the presence of IL-4 favors the T_H2-polarization but additionally inhibits the differentiation of T_H1-cells. The T_H-subset specific transcription factors T-bet and GATA3 also seem to regulate the process of T_H1/T_H2 polarization by suppressing the development of the opposite T_H-cell type [20]. It was initially assumed that this process was irreversible once the differentiation to either T_H1- or T_H2-cells was established [29,30]. However, recent data suggested that T-bet may be capable of forcing polarized T_H2-cells to produce IFN-γ comparable to T_H1-cells [31,32].

Several previous studies extended this T_H1/T_H2 paradigm by an additional T-cell subset, the T-regulatory cells (Treg). The Treg cells exert a suppressive function on both T_H1- and T_H2-cells to

maintain their balance [33-35]. An imbalance between both T-helper populations may induce the pathogenesis of human diseases. Thus, it was suggested that a strong shift towards T_H1-polarized cells may predispose to the development of autoimmune diseases, sarcoidosis and Crohn`s disease [36-38]. In contrast, a T_H2-dominated differentiation of naïve T-cells seem to be associated with allergic disorders including asthma and elevated total serum IgE levels [39,40].

1.4 The role of IgE in allergic diseases

An immediate immune response against common allergens is mediated through the recognition of allergens by specific IgE molecules on the surface of mast cells. These IgE molecules bound to mast cells by the IgE receptor FCεR1 trigger the release of vasoactive mediators and the transcription of cytokines (mast cell degranulation, figure 3). These mediators of the immediate-hypersensitive reaction induce mucus production, smooth muscle contraction and eventually inflammatory infiltration.

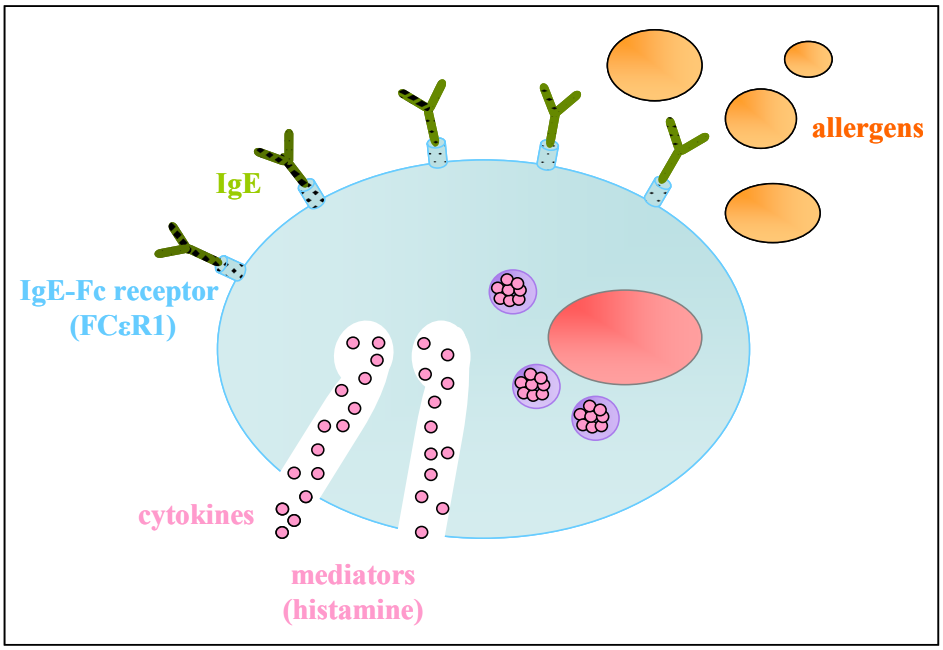

figure 3: IgE molecules bind to FCεR1 receptors located on mast cells. Antigens react with these bound IgE molecules which results in the secretion of multiple inflammatory mediators including histamine and cytokines.

However, IgE is not only important for the immediate allergic reaction but is also involved in the late-phase allergic reaction occurring many hours after the acute allergen contact. In asthmatics, IgE-mediated allergy may lead to a decrease in bronchial airflow as late as 4-8 hours after allergen challenge [41].

Humoral and cellular immune responses against allergens are activated through the binding of IgE to the low affinity receptor CD23 on B-cells. This interaction induces B-cells to present antigens efficiently and activates a specific T-cells response. Initially, IgM on naïve B-cells recognize antigens which are subsequently presented to the T-cell receptor via the MHC class II molecules (MHC II, figure 4). Thus, the IL-4/IL-13 pathway is activated to produce the cytokine IL-4 which binds to IL4-receptor α (IL-4Rα) on the B-cell (first signal). A second co-stimulatory signal is delivered through the interaction of CD40 and B7-1 on B-cells with CD40L and CD28 on T-cells. The concomitant presence of both signals induces class-switching to IgE in B-cells.

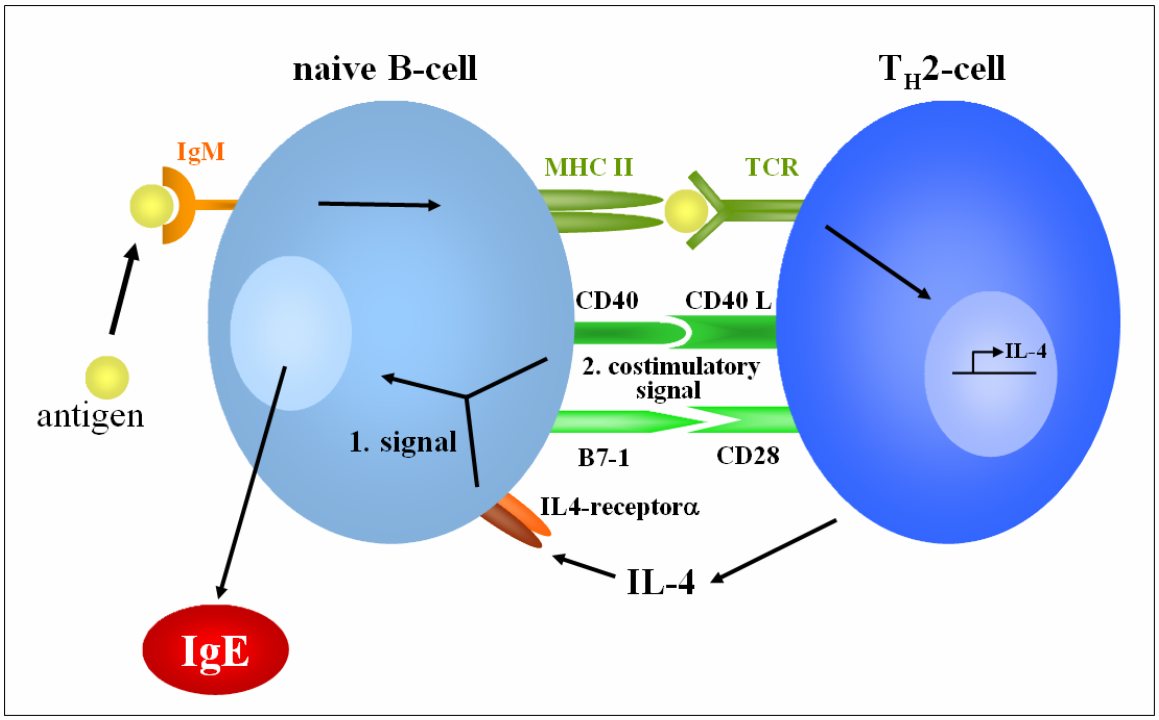

figure 4: Two signals are necessary to induce the IgE production: Antigens are recognized by IgM on the B-cell surface followed by their presentation of MHC class II molecules (MHC II). Antigens are presented to the T-cell receptor (TCR) which subsequently activates the first signal through the IL-4/IL-13 pathway. The binding of CD40 on the B-cell surface with its ligand CD40L on the T-cell triggers the expression of B7-1. The interaction of B7-1 and CD28 delivers a costimulatory signal that amplifies the cytokine synthesis and the T-cell proliferation. In response, IL-4 binds to the IL-4Rα, which in conjunction with the binding of CD40 triggers the IgE isotype switch, B-cell proliferation and expansion of the IgE-producing clone [41].

1.4.1 Activation of the IL-4/IL-13 pathway is crucial for allergic inflammation

1.4.1.1 IL-4/IL-13 pathway signals induce class-switching to IgE

IL-4 and IL-13 are the only cytokines which are capable to induce IgE production by B-cells. This is achieved by the activation of the IL-4/IL-13 pathway [42,43] (figure 5): IL-4 binds to the receptor composed of IL-4Rα chain and the common γ chain. IL-13 binds to its own specific receptor subunit IL-13Rα1 chain, to which IL-4 cannot bind, and additionally to the IL-4Rα chain (IL-4 Receptor α) [44]. After the binding of IL-4 and IL-13 to the receptors on the T-cell surface, the oligomerization of the receptors is induced. Hence, IL-4 activates the Janus tyrosine kinases Jak1 and Jak3. IL-13 transmits its signal through Jak1 and the Tyk2 kinase. The activated kinases initiate the phosphorylation of the intracellular molecule STAT6 (signal transducer and activator of transcription 6). Once

phosphorylated, STAT6 homodimers are formed which translocate to the nucleus. Hence, STAT6 binds to IL-4/IL-13 responsive regulatory gene regions. The recombination event is therefore activated leading to the transcription of germline epsilon messenger RNA (mRNA) for the class-switch from IgM to IgE in B-cells. Additionally, STAT6 induces the transcription factor GATA-3 which enables the differentiated T_H2-cell to maintain its characteristic features.

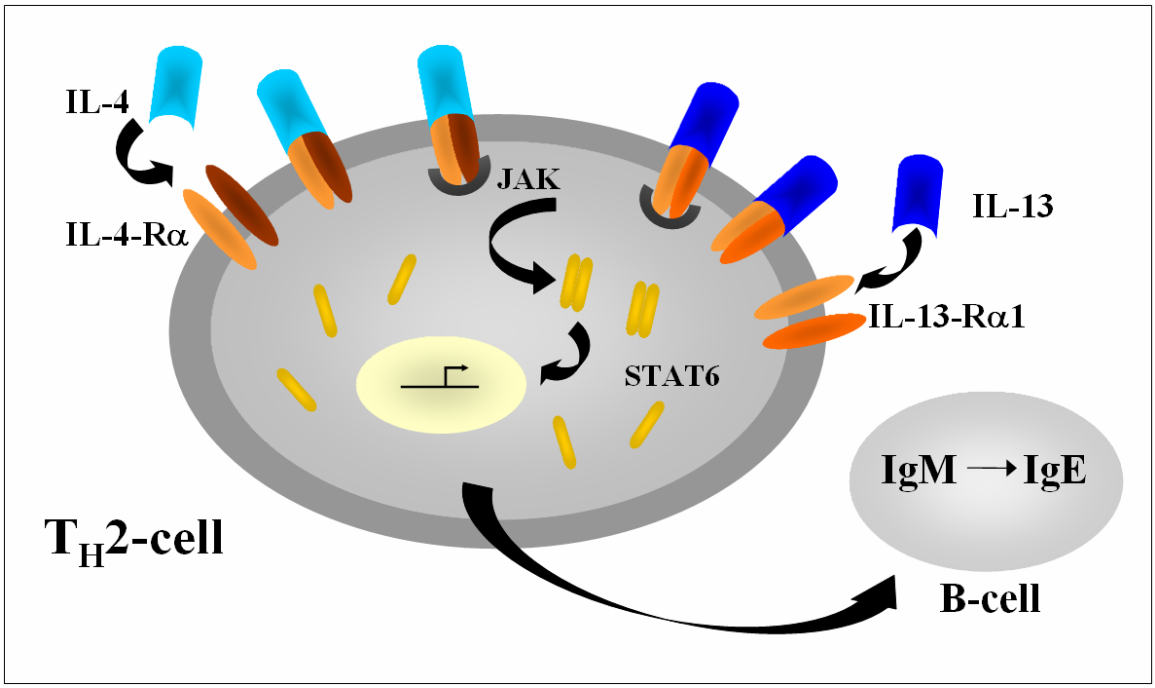

figure 5: IL-4 and IL-13, bind to a shared receptor on the surface of T_H2-cells inducing the activation of Janus tyrosine kinases (Jak). The phosphorylation of the intracellular molecule STAT6 through the activated Jaks leads to a homodimerisation of STAT6, which translocate to the nucleus and bind to specific regions of IL-4 - and IL-13 - inducible genes. Thus, crucial target genes for the synthesis of IgE are activated.

1.4.1.2 The critical role of IL-4 and IL-13 in the development of atopic diseases

Both cytokines, *IL-4* and *IL-13,* are located on chromosome 5q31 in close proximity to each other (25kb). Even though the proteins share only approximately 25% homology it was suggested that these genes emerged through a duplication event [45]. Both cytokines are primarily produced by T_H2-polarized cells but Natural Killer cells [46] and eosinophiles [47,48] are also an important source. IL-4 is additionally synthesized by B-cells [49].

IL-4 and IL-13 are both crucial for the IgE regulation but they exert further functions some of which are overlapping. Both cytokines are involved in T_H2 differentiation as well as the development and differentiation of B-cells and monocytes [50]. IL-4 and IL-13 can also inhibit the monocyte-derived production of proinflammatory cytokines [50].

The role of IL-4 in T_H2 differentiation was initially demonstrated in IL-4 deficient mice [51]. It was shown that IL-4 was necessary to induce a T_H2 response as the signal was strongly diminished after the disruption of the *IL-4* gene. Comparable effects were observed in $CD4^+$ T-cells from IL-13 deficient mice which had a significant impaired induction in T_H2 differentiation [52]. However, the strongest effect was seen in double knock-out mice (IL-4 and IL-13 negative) [53]. Only then, complete abrogation of eosinophil infiltration, IgE production, IL-5 secretion and differentiation to a T_H1-like phenotype was detectable. Thus, the expression of either IL-4 or IL-13 may be sufficient to induce a T_H2 response but their concomitant release seems to exert an additive effect on the immune response against allergens.

Furthermore, the crucial role of both cytokines in the pathogenesis of atopic diseases was demonstrated in IL-4 and IL-13 deficient mice: ovalbumin-sensitized IL-4 wildtype mice showed increased eosinophilic inflammation as well as the induction of bronchial hyperresponsiveness after additional provocation [54]. These effects were not detectable in IL-4 deficient mice, yet exposed to the same stimuli [54]. In addition, antigen challenge of IL-4 knock-out mice resulted in a substantial attenuation of IL-5 secretion [55]. To determine the specific time-point when IL-4 may affect immune responses, Coyle and coworkers administered neutralizing anti-IL-4 antibody either immediately before antigen challenge or during immunization [55]. The influence on eosinophilic inflammation was marginal when anti-IL-4 was administered before antigen challenge. Only when the anti-IL-4 antibody was given during immunization, a vigorous reduction in eosinophil infiltration, IL-5 secretion and IgE production was observed. From these data it was concluded that IL-4 is crucial for both, the local T_H2 response as well as the eosinophilic inflammation.

IL-13 deficient mice were unable to develop BHR even though airway inflammation and high levels of IL-4 and IL-5 were detectable [56]. However, the induction of BHR was restored after administration of recombinant IL-13 to these mice, indicating that IL-13 is also necessary for BHR. In contrast to IL-4 knock-out mice, only IL-13 deficient mice failed to generate goblet cells which are responsible for mucus overproduction, a characteristic feature of asthma [57]. Hence, IL-13 may be more relevant for mucus hypersecretion and subepithelial fibrosis.

1.4.1.3 The function of IL-4Receptor α in allergic inflammation

As both cytokines IL-4 and IL-13 mediate their signal through IL-4Rα, this receptor plays also a crucial role in allergic inflammation. IL-4Rα is expressed on hematopoetic as well as on non-hematopoetic cells [58]. It was shown that IL-4Rα activates the *IL-4* gene expression through a specific STAT6 DNA-binding target site within the *IL-4Rα* promoter [59]. This distinct region of *IL-4Rα* is also relevant for the IL-4 induced expression of CD23 and MHC class II [60].

In IL-4Rα knock-out mice, the IgE production was diminished and the differentiation to T_H2-cell was impaired [61]. Mice lacking the IL-4Rα were also not able to develop BHR, to induce mucus production and airway inflammation [62]. To further characterize the role of the *IL-4Rα*, murine radiation bone marrow chimera were used [63]. In a first step the bone marrow of *IL-4Rα* wildtype (+) or IL-4Rα knock out (-) mice was degraded after irradiation. Subsequently, bone marrow from *IL-4Rα* (+) or *IL-4Rα* (-) mice was transferred into recipient mice that expressed or lacked *IL-4Rα*. Hence, *IL-4Rα* deficient mice which obtained bone marrow from mice expressing IL-4Rα developed the expected goblet-cell metaplasia and allergic inflammation. In contrast, these responses were not detectable in *IL-4Rα* expressing mice which were transferred with *IL-4Rα* deficient bone marrow even in the presence of T_H2-cells and substantial eosinophilic inflammation. These data suggested that also goblet-cell

metaplasia and the mucus production after allergic inflammation strongly depend on the signal transduction through IL-4Rα.

1.4.1.4 STAT6 is crucial for the mediation of immune responses

In the downstream signaling of the IL-4/IL-13 pathway STAT6 is a major factor as IL-4 and IL-13 both signal through this intracellulare transcription factor. STAT6 is one of seven members of the STAT transcription factor family. STAT6 is ubiquitously expressed in all cell types but its expression may be augmented in B- and T-cells after cell stimulation [64]. STAT6 is classically activated through the IL-4/IL-13 pathway as described in detail in section 1.4.1.1. In addition, alternative pathways have been observed for *STAT6* induction. In a study of murine mast cells, STAT6 phosphorylation and activation after IL-15 stimulation was reported [65,66]. However, stimulation with IL-15 seemed to be less potent than IL-4 [65]. Patel *et al.* additionally observed a platelet-derived growth factor (PDGF)-induced activation of the *STAT6* gene in NIH3T3 mouse fibroblasts [67].

The importance of *STAT6* in T_H2 differentiation was predominantly shown in *in vivo* studies using *STAT6* deficient mice. Mice lacking *STAT6* (BALB/c strain) were not capable to differentiate naïve T-cells into T_H2-cells [68-70]. As the abrogation of *STAT6* induced a total defect in T_H2-cell generation, it was assumed that *STAT6* was not only sufficient but necessary for T_H2 differentiation. Further experiments with STAT6 fusion proteins supported the notion that STAT6 is of great relevance in T_H2 differentiation: when constitutively active forms of STAT6 were introduced into T_H1 differentiated cells, markedly reduced levels of IFN-γ accompanied by the expression of T_H2-type cytokines were detected [71,72].

As the IL-4/IL-13 pathway activates the class-switching process from IgM to IgE, it was analyzed if STAT6 additionally is a key regulator in IgE synthesis. It was shown that *STAT6* deficient mice were

not able to undergo class-switching to IgE although levels of IgG, IgA and IgM were present [68,69]. Accordingly, these mice did not contain IgE-producing B-cells [69].

Furthermore, *in vitro STAT6* deficient T-cells had a significantly impaired proliferative response to IL-4 stimulation. Thus, the relevance of STAT6 in IL-4–driven T-cell proliferation as well as differentiation was shown [72,73]. These observations were due to a defect to switch from G1 to the S phase of the cell cycle and an impaired ability to down-regulate the cell cycle–dependent kinase inhibitor $p27^{kip}$ after IL-4 stimulation [73].

However, STAT6 is not only relevant for T_H2 differentiation and IgE regulation. It is assumed that the JAK-STAT signaling pathway is involved in many of the physiologic events that are deregulated in asthma: analyses of pulmonary inflammation and bronchial hyperresponsiveness in mice were performed to identify the putative function of STAT6 in the development and perpetuation of allergic diseases. Following allergen provocation with ovalbumin, *STAT6* deficient mice (BALB/c strain) failed to develop BHR and airway eosinophilia [74,75]. Consistent with these observations, lower levels of T_H2 cytokines (IL-4, IL-5) but increased levels of the T_H1 cytokine IFN-γ were detectable. Moreover, the IgE production was completely abolished accompanied by significantly increased IgG2a response in the serum of these *STAT6* deficient mice.

Furthermore, it was suggested that STAT6 influences T_H2 recruitment and their effector function in allergic inflammation [76]. As T_H2-cell differentiation was completely abrogated in *STAT6* deficient mice, the effect of STAT6 on T_H2 recruitment had to be analyzed in *STAT6* wildtype mice. Therefore, T_H2-cells after allergen provocation of $STAT6^{+/+}$ mice were transferred either into $STAT6^{+/+}$ or $STAT6^{-/-}$ mice again followed by a challenge. The total number of cells of the bronchoalveolar lavage of *STAT6* deficient mice was significantly reduced in comparison to $STAT6^{+/+}$ mice. This observation was mainly due to the reduction of eosinophils to 1-2% in $STAT6^{-/-}$ mice. Other characteristics of

asthma were also impaired including mucus production, BHR and T_H2 chemokine production which control recruitment of T_H2-cells.

Several groups have investigated the expression and activation of STAT6 in asthmatic patients in comparison to healthy controls. Analyzing the immunochemistry of STAT6 revealed that patients with atopic and non-atopic asthma express significantly more STAT6 immunoactive cells than healthy controls [77]. This observation was surprising as the absolute number of cells effectively expressing the STAT6 protein was decreased in non-atopic versus atopic asthma patients. In contrast, Peripheral blood monocytes (PBMC) from asthmatic and allergic patients did not display significant differences in the level of STAT6 activity relative to healthy controls [78].

1.4.2 The influence of genetic predisposition on the development of asthma and atopy: an overview of IL-4/IL-13 pathway SNPs

Activation of the IL-4/IL-13 pathway is substantial for the allergic immune response. Therefore, it was hypothesized that genetic alterations within this pathway may modulate the development of asthma and IgE production.

1.4.2.1 The relation between IL-4 SNPs and allergic inflammation

Based on linkage studies, a susceptibility locus for allergic inflammation and IgE synthesis was attributed to a region on the short arm of chromosome 5 (5q31-33) where both genes *IL-4* and *IL-13* are located [79-81]. Resequencing of candidate genes and association studies followed to identify putative polymorphisms responsible for the observed linkage signal.

Rosenwasser and coworkers first described a polymorphism located in the promoter region of the *IL-4* gene (C-589T) to be associated with elevated IgE levels in atopic asthmatics [82]. Significant associations with this *IL-4* promoter polymorphism were observed in a number of studies but to other atopic phenotypes including asthma [83,84], atopic dermatitis [85] and bronchial hyperresponsiveness [86]. However, these results could not be replicated in other populations [87-89].

In 2003, a systematic screening for genetic variants in the complete *IL-4* gene was performed in 33 unrelated German subjects by our laboratory [90]. In total, 14 novel polymorphisms, all located in the intronic region of the gene, were identified and two previously reported SNPs were confirmed. In this Caucasian population the polymorphisms within the *IL-4* gene revealed two tightly linked clusters. Five SNPs were genotyped in a study of 1,120 German children (ISAAC II). The strongest association with asthma was observed with the promoter polymorphism C-589T. In accordance, two additional genotyped SNPs were significantly associated with asthma. This effect was due to their high linkage to the promoter polymorphism C-589T. A positive association between three SNPs (C-589T, C-33T, VE6523) and total IgE levels was described after stratification for positive skin-prick-test. Functional experiments have been performed to identify the putative regulatory relevance of the proximal *IL-4* promoter polymorphism C-589T [91]. These data showed alterations in the binding of a transcription factor and changes in the *IL-4* gene expression depending on the genotype of the SNP C-589T.

The relation between genetic alterations in the *IL-4* gene and the regulation of IgE was further investigated by Basehore and coworkers in three populations containing different ethnicities: Caucasians, African Americans and Hispanics [92]. In the Caucasian population nine SNPs were significantly associated with total serum IgE levels whereas five were also associated with asthma. One polymorphism (G3017T) located in intron 2 of the *IL-4* gene seemed to be the most consistent SNP over all three populations. Furthermore, associations of G3017T with total serum IgE levels in the African American as well as in the Hispanic population were described. This polymorphism was additionally associated with asthma in Hispanic subjects. Performing haplotype analyses, the

distribution and frequency of haplotype combinations was different across the three tested populations. Among the Caucasian and the African American population different haplotypes were associated with elevated total IgE levels. No significant relations between haplotypes and asthma have been observed.

1.4.2.2 IL-13 polymorphisms influence atopic diseases

As both cytokines, IL-4 and IL-13, are important in the activation of the IL-4/IL-13 pathway the putative role of genetic variants in the *IL-13* gene has also been investigated. In a mutation screening of the complete *IL-13* gene including all exons, introns and 1676bp of the promoter region, seven polymorphisms were identified [93]: two SNPs were located in the *IL-13* promoter, one in intron 3, one in exon 4 leading to a predicted amino acid change from arginine to glutamine (Arg130Gln) and three in the 3´UTR. The observed polymorphisms were genotyped in three different populations, all of Caucasian inheritance. The strongest effect was observed between the exonic SNP (G2044A) and elevated total serum IgE levels. For two genotyped *IL-13* promoter polymorphisms (A-1512C, C-1112T) a somewhat weaker but still statistically significant association to the same phenotype was reported. However, as all seven polymorphism within the *IL-13* gene are in high linkage disequilibrium with each other it is difficult to dissect which SNP is causal for the observed association signal.

In the *IL-13* gene, the polymorphisms G2044A in exon 4 and the promoter polymorphism C-1112T have been extensively investigated in further association studies to confirm their putative relevance on the development of atopic diseases. A significant association for G2044A with asthma [94,95], the regulation of total serum IgE [96-98] and atopic dermatitis [96] have been described among different ethnicities. The variation C-1112T in the promoter region of *IL-13* has been shown to be involved in the IgE regulation in another German population [88] as well as in the development of asthma [99,100], BHR and skin-test responsiveness in a Dutch population [100].

Functional studies were performed to better understand the regulatory influence of the genetic variants G2044A and C-1112T. As the polymorphism G2044A is located in exon 4 it was hypothesized that the alteration may modify the structure and the affinity of the IL-13 protein to its receptor. It was demonstrated that the affinity of the recombinant IL-13 protein containing either the wildtype or polymorphic allele of G2044A to IL-13Rα1 was almost unaffected [101]. In contrast, G2044A seemed to be relevant for the affinity of IL-13 to IL-13Rα2: the IL-13 protein expressing the polymorphic allele showed a lower affinity to IL-13Rα2, causing less clearance of IL-13. In accordance, IL-13 levels in the serum of asthmatic individuals carrying the polymorphic A allele were elevated in comparison to individuals with the wildtype allele. These data indicated that the exonic alteration in IL-13 may be important in the genesis of allergic diseases through the up-regulation of IL-13 levels. Recent data supported the functional relevance of the IL-13 SNP G2044A: IL-13 protein isolated from primary cells expressing the polymorphic A allele were significantly more active in inducing STAT6 phosphorylation and CD23 expression in monocytes and IgE switching in B-cells [102].

To further analyze the mechanism how the associated *IL-13* promoter polymorphism C-1112T may influence allergic diseases, this promoter SNP was investigated in T-cells by Cameron and coworkers [103]. In gene expression studies it was shown that the polymorphic T allele of the SNP enhanced the promoter activity of *IL-13* in primary human and murine $CD4^+$ T_H2-cells. This effect may be explained by observed changes in the binding properties of transcription factors (STAT6, Oct-1 and NFAT2) depending on the genotype. However, this outcome seemed to be strongly influenced by the nuclear environment as a reciprocal effect of the transcriptional activity was seen in non-polarized $CD4^+$ T-cells and in the Jurkat T-cell line accompanied by a modulated DNA/protein binding interaction. To support the hypothesis that SNP C-1112T may be involved in the regulation of atopic diseases, IL-13 secretion was compared between groups either expressing the wildtype C or the polymorphic T allele. A significant increase of IL-13 in the serum was observed in homozygous

carriers of the T allele [103]. Hence, a regulatory influence of the promoter polymorphism C-1112T on *IL-13* expression has now been well established.

1.4.2.3 The diverse picture of IL-4Rα SNPs and the genetic predisposition to allergic diseases

The central role of *IL-4Rα* in regulating the production of IgE encouraged numerous groups to investigate the influence of polymorphisms in the gene on receptor signaling and hence the effect on atopic diseases. In the *IL-4Rα* gene, 14 SNPs have so far been identified, all of which are located in the coding region [104-106]. Of these SNPs, eight polymorphisms are predicted to lead to an amino acid change implicating that the *IL-4Rα* is a rather polymorphic gene. However, association data for any of the reported SNPs are more or less contradictory. While some studies observed a relation between SNPs in the *IL4Rα* gene and the genetic predisposition to asthma [106-108], atopic disease [105,109] and IgE levels [110-112] others could not find an association with any atopic phenotype [113-116].

A study carried out by Ober and coworkers in three populations of different ethnicities gave evidence that different single SNPs in *IL-4Rα* were associated with atopy and asthma *per se* within the tested populations. Performing haplotype analyses, the association signal with asthma became stronger. Still, different haplotypes among these populations were responsible for the observed effects [106]. Comparable results were seen in two other populations [105,108]. In a study of German children, where the initial analysis did not show statistically significant association, only in haplotype analyses a positive association with elevated total serum IgE was demonstrated [117].

Also, functional studies were performed to investigate the influence of *IL-4Rα* polymorphisms on the development of atopic diseases. In mouse and human cell lines it was shown that the wildtype allele of I50V significantly upregulated the receptor activity of IL4Rα after IL-4 stimulation in contrast to the polymorphic allele [107]. Hence, an augmented STAT6 activity was detectable, subsequently leading to

increased cell proliferation. Additionally, I50 induced higher expression of CD23 and promoted IgE production [118].

The SNP Q576R was also investigated in functional experiments. 576R was associated with IgE regulation and induced the expression of higher CD23 levels after IL-4 stimulation [105]. It was argued that changes in the binding specificity of a tyrosine residue in close vicinity to Q576R may explain these observations. However, no effect on CD23 expression was observed by Risma *et al*. when the role of 576R was investigated [119]. Nevertheless, the combination of the polymorphic alleles of two SNPs (I75V-Q576R) induced the expression of IL-4Rα with enhanced sensitivity to IL-4.

1.4.2.4 Genetic variants in the STAT6 gene contribute to the development of allergic disorders

The relation between the development of atopic diseases and polymorphisms in crucial target genes (*IL-4*, *IL-13* and *IL4-Rα*) necessary for the activation of the IL-4/IL-13 pathway has yet been intensively analyzed in various population studies. Many studies have been performed to understand the biological function of STAT6 (described in detail in section 1.4.1.4) indicating its crucial role in the signal transduction of IL-4 and IL-13. However, only little data on the influence of genetic alterations within the *STAT6* gene and their effect on STAT6 function are yet available.

The *STAT6* gene itself maps to 12q13.3-q14.1. The whole gene comprises approximately 19kb (AF067572-AF067575) and contains 23 exons interrupted by 22 introns (figure 6). The first two exons do not contribute to the transcription, leaving a coding sequence of approximately 4kb [120].

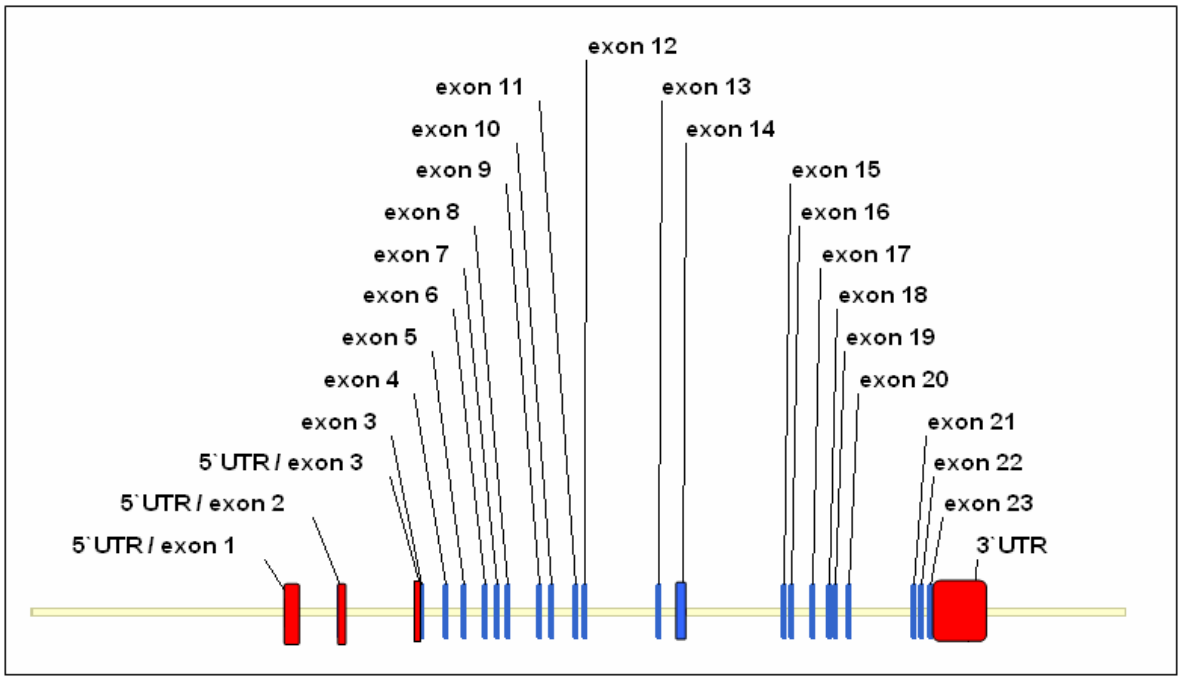

figure 6: The gene structure of *STAT6*. The gene encompasses about 19kb consisting of 23 exons (blue) interrupted by 22 intron. The first two exons do not contribute to the coding region. The untranslated regions (UTR) are marked in red.

The monomeric STAT6 protein consists of 847 amino acids with a size of 94kDa, respectively [121,122]. The domain organization of the STAT6 protein is separated into the N-terminal α-helix domain, the DNA binding domain, the src homology domain (SH2, allowing for interaction with phosphorylated tyrosines), and the transactivation domain at the C-terminus [123] (figure 7).

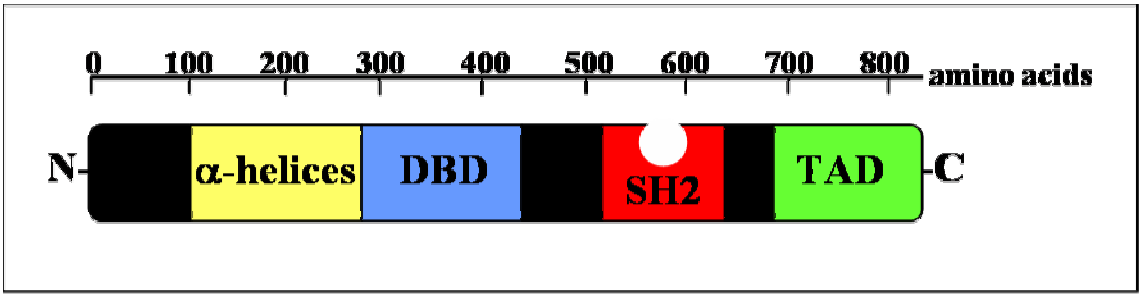

figure 7: The protein structure of STAT6. The protein consists of 847 amino acids. The domain organization is indicated by various colored regions: α-helices (yellow), DBD = DNA-binding domain (blue), SH2 = src homology 2 domain (red), and the TAD = transactivation domain (green)

Initial genetic studies of *STAT6* identified a GT dinucleotide repeat located in the first exon of *STAT6*. Its presence was investigated in three independent populations of Asian and Caucasian descent [124-126]. Among these populations of different ethnicity, deviating numbers of GT-repeats were described (Japanese: 13 - 16 [125], German and Swedish: 13 to 17 [124], Americans: 12 - 17 [126], British: 13 - 18 [126]). In the association study of Duetsch *et al.* 16 GT-repeats were significantly associated with an increase in the total number of eosionophils [124]. In contrast, subjects in the British study population with 16 GT-repeat seemed to protect against the development of asthma whereas carriers of 13 GT-repeats showed an inverse association with asthma, respectively [126]. Individuals with 13 GT-repeats furthermore had higher levels of total IgE and elevated *in vitro* transcriptional activity in three different cell lines [126]. Thus, it was assumed that the number of GT-repeat may be a marker for differences in the promoter activity of the *STAT6* gene.

In 2002, Duetsch and coworkers systematically screened the *STAT6* gene to identify genetic variants in addition to the previously identified GT repeat region [124]. Hence, the *STAT6* promoter and all 23 exons, including parts of neighboring introns, were analyzed to detect sequence alterations in 16 unrelated healthy controls and 16 asthmatics. In total, ten polymorphisms were identified. Two SNPs were located in the *STAT6* promoter, three polymorphisms were described in the intronic region (intron 16, intron 17, and intron 18) and five belonged to the 3`UTR in exon 23 of *STAT6*. An association study in a population of 108 Caucasian asthma sib-pair families (N = 474, 216 parents, 258 children) was performed with these ten identified SNPs. Six additional polymorphisms (three in intron 2, two in intron 8 and one in intron 12) from public databases (www.ncbi.nml.gov/SNP) were also included for genotyping. A significant association with single SNPs was described between total IgE levels and three *STAT6* polymorphisms located in intronic regions of the gene (intron 2, 17 and 18) and one SNP in the 3`UTR. Furthermore, for one polymorphism located in the 3`UTR a significant association with bronchial challenge was described. Testing the correlation of *STAT6* haplotypes with atopic diseases, no significant effect was observed. Potentially, this was due to the small sample size of the study population [124].

1.5 Objective of the present study

Atopic diseases like asthma and allergy are characterized through an IgE mediated immune response against ubiquitous allergens. The development of IgE is regulated through class-switching from IgM to IgE after activation of the IL-4/IL-13 cytokine pathway. The IL-4/IL-13 pathway is involved in many immunological aspects of the disease.

STAT6 is a bottle neck in the activation of the IL-4/IL-13 pathway. Hence, we thought to investigate if genetic variations at this point of the pathway influence the development of allergic diseases and alter immunological entities.

As previous small studies had already investigated the role of *STAT6* in IgE regulation and the development of asthma, we analyzed if this was true also in a well defined and large cross-sectional study population. Thus, the effect of *STAT6* variants was systematically tested in different populations and their role within the IL-4/IL-13 pathway was investigated.

After identifying the SNP putatively responsible for associations with IgE regulation, atopy and asthma development in the first part of the thesis, it was further analyzed how this polymorphism located in intron 2 may affect *STAT6* function. To address this question, the *STAT6* promoter was initially defined and potential interactions with intron 2 of *STAT6* were investigated. The *cis*-regulatory role of intron 2 was determined and the mechanism by which genetic variants may be involved in these processes was elucidated in more detail.

As this second part of the thesis revealed that the asthma and allergy associated polymorphism seem to act through the modulation of gene expression in an *in vitro* model, the *ex vivo STAT6* gene expression was analyzed in detail in the third part of the thesis on a population-based level.

2 Materials and Methods

2.1 Materials

2.1.1 Reagents and chemicals

100bp and 1kb Ladder (500µg/ml)	New England BioLabs, Ipswich, USA
30% polyacrylamid (19:1)	Biorad, Hercules, USA
γ[32P]-ATP	Hartmann Analytic, Braunschweig, Germany
β-mercaptoethanol	Sigma-Aldrich, Steinheim, Germany
ACK Lysis buffer	Bio Whittacker, Walkersville, USA
Agarose	Biozym, Hess, Germany
Ammoniumpersulfate	Sigma-Aldrich, Steinheim, Germany
Ampicillin	Sigma-Aldrich, Steinheim, Germany
anti-CD4-FITC antibody (4µl/5x10^5 cells)	Beckmann Coulter, Fullerton, USA)
Antipain hydrochloride	Sigma-Aldrich, Steinheim, Germany
Aprotinin	Sigma-Aldrich, Steinheim, Germany
BSA	New England BioLabs, Ipswich, USA
Benzamidin	Sigma-Aldrich, Steinheim, Germany
Betain	Sigma-Aldrich, Steinheim, Germany
Boric acid	Sigma-Aldrich, Steinheim, Germany
Bromphenol blue	Roth, Karlsruhe, Germany
Desoxyadenosintriphosphat dATP (100mM)	PeqLab, Erlangen, Germany MBI Fermentas, St. Leon-Rot, Germany
Desoxycytidintriphosphat dCTP (100mM)	PeqLab, Erlangen, Germany MBI Fermentas, St. Leon-Rot, Germany
Desoxyguanosintriphosphat dGTP (100mM)	PeqLab, Erlangen, Germany MBI Fermentas, St. Leon-Rot, Germany
Desoxythymidintriphosphat dTTP (100mM)	PeqLab, Erlangen, Germany MBI Fermentas, St. Leon-Rot, Germany
DTT	Sigma-Aldrich, Steinheim, Germany
Ethanol 100 %	Merck, Darmstadt, Germany
Ethidiumbromide (10mg/ml)	Biorad, Hercules, USA
Ethylene diamine tetraacetic acid EDTA (0.5 M)	Sigma-Aldrich, Steinheim, Germany
Fetal Calve Serum (FCS)	Gibco, Carlsbad, USA
Ficoll Paque™ Plus	Amersham, Piscataway, USA
Fluorescein Calibration Dye	Biorad, Hercules, USA
Glycerol	Merck, Darmstadt, Germany
Glycerol2-phosphate disodium salt hydrate	Sigma-Aldrich, Steinheim, Germany
H_2O bidest.	H. Kerndl GmbH, Weißenfeld,
Hepes solution	Sigma-Aldrich, Steinheim, Germany
Hydrogen chloride (HCl)	Sigma-Aldrich, Steinheim, Germany
Ionomycin	Sigma-Aldrich, Steinheim, Germany
Isopropanol 100 %	Merck, Darmstadt, Germany
Leupeptin	Sigma-Aldrich, Steinheim, Germany
LiChrosolv H_2O (HPLC)	Merck, Darmstadt, Germany
Magnesium chloride	Sigma-Aldrich, Steinheim, Germany
molecular H_2O	Eppendorf, Hamburg, Germany
N_{15}-Primer (200µM)	Metabion, Martinsried, Germany
NF-κB p50 antibody	Santa Cruz Biotechnology, Santa Cruz, USA
NF-κB p65 antibody	Santa Cruz Biotechnology, Santa Cruz, USA

Nonident P-40	Sigma-Aldrich, Steinheim, Germany
PBS	Sigma-Aldrich, Steinheim, Germany
Penicillin/Streptavidin (5000µg/ml)	Sigma-Aldrich, Steinheim, Germany
Pepstatin A	Sigma-Aldrich, Steinheim, Germany
Phenylmethanesulfonylfluoride	Sigma-Aldrich, Steinheim, Germany
Phorbol12-myristate 13-acetate (PMA)	Sigma-Aldrich, Steinheim, Germany
Poly (dl-dC)-Poly (dl-dC)	Amersham, Piscataway, USA
Polymer POP-6™	Applied Biosystems, Foster City, USA
Potassiumhydrogencarbonate ($KHCO_3$)	Merck, Darmstadt, Germany
Primer	Metabion, Martinsried, Germany
Recombinant human IL-4	R&D Systems, Minneapolis, USA
Rotiszint®Eco	Roth, Karlsruhe, Germany
RPMI 1640 Glutamax	Gibco, Carlsbad, USA
Sodiumchlorid (NH_4Cl)	Sigma-Aldrich, Steinheim, Germany
Sodiumchloride (NaCl)	Sigma-Aldrich, Steinheim, Germany
Sodium fluoride (NaF)	Sigma-Aldrich, Steinheim, Germany
Sodiumdodecylsulfate (SDS)	Serva, Heidelberg, Germany
Sodiumhydroxide (NaOH)	Merck, Darmstadt, Germany
Sodium orthovanadate (NaV)	Sigma-Aldrich, Steinheim, Germany
SP1 antibody	Santa Cruz Biotechnology, Santa Cruz, USA
SP2 antibody	Santa Cruz Biotechnology, Santa Cruz, USA
SP3 antibody	Santa Cruz Biotechnology, Santa Cruz, USA
SP4 antibody	Santa Cruz Biotechnology, Santa Cruz, USA
SpectroCLEAN Resin	Sequenom, San Diego, USA
SYBR®Green I Nucleic Acid Gel Stain	Cambrex Bio Science, Rockland, USA
T4 Polynucleotide Kinase	New England Biolabs, Ipswich, USA
TEMED	Sigma-Aldrich, Steinheim, Germany
Tris(hydroxymethyl)-aminomethan TRIS	Sigma-Aldrich, Steinheim, Germany
TRIS-HCl (1M)	Sigma-Aldrich, Steinheim, Germany
Trisbase	Sigma-Aldrich, Steinheim, Germany
Xylencynol	Merck, Darmstadt, Germany
YB1 antibody	Santa Cruz Biotechnology, Santa Cruz, USA

2.1.2 Enzymes

AccI	New England BioLabs, Ipswich, USA
AvaII	New England BioLabs, Ipswich, USA
BamHI	New England BioLabs, Ipswich, USA
BglII	New England BioLabs, Ipswich, USA
DpnI	Stratagene, La Jolla, USA
HotStar *Taq*®-Polymerase (5U/µl)	Qiagen, Hilden, Germany
Flexigene Protease	Qiagen, Hilden, Germany
FseI	New England BioLabs, Ipswich, USA
NEB *Taq* Polymerase (1U/µl)	New England BioLabs, Ipswich, USA
PfuUltra™ High Fidelity *Taq*	Stratagene, La Jolla, USA
Platinum® *Taq* High Fidelity	Invitrogen, Carlsbad, USA
Proteinase-K	Sigma-Aldrich. Steinheim, Germany
RNase free DNase	Qiagen, Hilden, Germany
RNase Out™	Invitrogen, Carlsbad, USA
RNase H	Invitrogen, Carlsbad, USA
Shrimp Alkaline Phosphatase (SAP)	Amersham, Piscataway, USA
ThermoScript RT	Invitrogen, Carlsbad, USA
Thermosequenase	Amersham, Piscataway, USA

2.1.3 Solutions and buffers

5xTBE buffer	54g Triszma base 27.5g boric acid 20ml 0.5M EDTA (pH 8.0) ad 1l H_2O bidest.
5M Betain	11.71g Betain 20ml dH_2O
0.5M Benzamidin	60.08mg ad 1ml dH_2O
0.5M β-Glycerophosphate	108.02mg ad 1ml dH_2O
buffer Y	200 mM TRIS-HCl (pH 8.55) 160 mM $(NH_4)_2SO_4$ 20 mM $MgCl_2$
DNA ladder	10µl 100bp or 1kb ladder 80µl 0.5x TBE-Buffer 10µl Loading Dye
10mM dNTPs	100µl each dATP(100mM), dGTP (100mM), dTTP (100mM), dCTP (100mM) 600µl dH_2O
0.5M DTT	77.13mg in 1ml 10mM NaAc
0.5M EDTA (pH 8.0)	86.14g EDTA ad 1l dH_2O
Ery-Lysis buffer (pH 7.4)	41.45g NH_4Cl (155mM) 5g $KHCO_3$ (10mM) 1ml EDTA (0.5M) ad 5l dH_2O
Ethidiumbromide (500µg/ml)	100µl Ethidiumbromide 1900µl dH_2O
Loading dye stock solution	0.25g bromphenol blue 0.25 g xylencynol 30% glycerol 70ml dH_2O
100mM Magnesium chloride	0.76g $MgCl_2$ ad 80ml dH_2O
10mM PMSF	1.74mg PMSF ad 1ml isopropanol
100mM Potassium chloride	0.6g KCl ad 80ml dH_2O
Proteinase-K (20mg/ml)	dilute 10mg in 500µl dH_2O
Proteinase-K buffer	5ml TRIS-HCl (1M. pH 7.5) 2ml EDTA (0.5M. pH 8.0) 3ml NaCl (5M) ad 1l H_2O bidest.
SDS 20 %	20g SDS ad 100ml dH_2O
0.5M Sodium fluoride	29.995mg NaF ad 1ml dH_2O
0.5M Sodium orthovanadate	91.96mg ad 1ml dH_2O
TE buffer (pH 8.0)	10mM Tris Cl (1M) 1mM EDTA (0.5M)

2.1.4 Reagent systems (Kits)

3-point calibrant	Sequinom, San Diego, USA
BCA Protein Assay Kit	Pierce, Rockford, USA
BigDye® Terminator v3.1 Cycle Sequencing Kit	Applied Biosystems, Foster City, USA
$CD4^+$ T-cell Isolation Kit	Miltenyi Biotec, Bergisch Gladbach, Germany
Dual-Luciferase® Reporter Assay	Promega, Madison, USA
Endo Free® Plasmid Maxi Kit	Qiagen, Hilden, Germany

FlexiGene DNA Kit	Qiagen, Hilden, Germany
HME MassEXTEND Nucleotide Tri-Mix-Pack	Sequinom, San Diego, USA
HotStar *Taq*® plusKit	Qiagen, Hilden, Germany
MassExtend Core Kit	Sequinom, San Diego, USA
MassExtend Nucleotide Mix Pack Tri 25	Sequinom, San Diego, USA
MassExtend Nucleotide Mix Pack Tri 26	Sequinom, San Diego, USA
MassExtend Nucleotide Mix Pack Tri28	Sequinom, San Diego, USA
QIAmp® RNA Blood Mini Kit	Qiagen, Hilden, Germany
QIAPrep® Spin Mini Prep Kit	Qiagen, Hilden, Germany
QIAquick® Gel Extraction Kits	Qiagen, Hilden, Germany
Quick Change® II Site-directed Mutagenesis Kit	Stratagene, La Jolla, USA
Quick Spin® Oligo Columns	Roche, Mannheim, Germany
PCR Master Mix™	Applied Biosystems, Foster City, USA
pGL3-Basic vector	Promega, Madison, USA
pGL3-Promoter vector	Promega, Madison, USA
Platinum® *Taq* High Fidelity Kit	Invitrogen, Carlsbad, USA
pRL-TK	Promega, Madison, USA
ThermoScript™ RT-PCR System	Invitrogen, Carlsbad, USA
Versagene RNA Purification Kit	Gentra, Minneapolis, USA

2.1.5 Consumables

Additionally to the standard laboratory equipment the following devices were used:

1.5 screw cap tubes	Sarstedt, Nümbrecht, Germany
15ml Falcon tubes 2059	Falcon, Bedford, USA
3MM Chromatography precut sheets	Whatman, Brentford, UK
25cm^2 flasks	Corning, Corning, USA
75cm^2 flask	Falcon, Bedford, USA
6-well plates	Falcon, Bedford, USA
96-wells Clusterplates	Qiagen, Hilden, Germany
96-well Low Profile Reaction Plate	PeqLab, Erlangen, Germany
96-well half-skirted plates	Applied Biosystems, Foster City, USA
384-wells Mikrotiter plates	ABgene, Epsom, UK
384-wells PCR plates	Biorad, Hercules, USA
384-well SpectroCHIP Bioarray	Sequenom, San Diego, USA
Cryogenitc vials	Corning, Corning, USA
Cyvette (0.2cm, 0.4cm)	Biorad, Hercules, USA
Freeze N´Squeeze™ Spin Columns	Biorad, Hercules, USA
Hyperfim™ MP	Amersham, Piscataway, USA
LS Coulms for Midi MACS or Quadro MACS	Miltenyi Biotec, Bergisch Gladbach, Germany
MultiScreen™ PCR plate	Millipore, Schwalbach, Germany
PCR foil	Eppendorf, Hamburg, Gemany
PCR foil Peel-it lite thermo sealing-foil	Eppendorf, Hamburg, Gemany
S-Monovette	Sarstedt Nümbrecht, Germany

2.1.6 Laboratory equipment

Additionally to the standard laboratory equipment the following devices were used:

96-Channel Pipettor	Beckman Coulter, Fullerton, USA
ABI PRISM® 3730 Genetic Analyzer™	Applied Biosystems, Foster City, USA

ABI PRISM® 7730 Sequence Detection System™	Applied Biosystems, Foster City, USA
Beta Counter LS 6000IC	Beckman Coulter, Fullerton, USA
Centrifuge 4K15	Sigma-Aldrich, Steinheim, Germany
Centrifuge 5810 R / 5417 R / 5415 D	Eppendorf, Hamburg, Germany
Centrifuge Rotanta 46RS	Hettich, Tuttlingen, Germany
Cuvette (10mm)	Hellma, Müllheim, Germany
Digitalcamera Kodak DC 290 Zoom	Kodak, Stuttgart, Germany
Dimple Plates (96-wells)	Sequenom, San Diego, USA
Electrophoresis Documentation and Analysis System	Kodak, Stuttgart; German
Eppendorf BioPhotometer	Eppendorf, Hamburg, Germany
Eppendorf Heat Sealer	Eppendorf, Hamburg, Germany
Film developer CP100	AGFA, Mortsel, Belgum
FLUOstar	BMD Labtech, Offenburg, Germany
Gel dryer 115/230	Biorad, Hercules, USA
Gel electrophoresis chambers	Biorad, Hercules, USA and PeqLab, Erlangen, Germany
Genepulser Xcell	Biorad, Hercules, USA
Hoefer apparatus (18cm x 24cm)	Hoefer, San Francisco, USA
Hyper cassette (8x10inch)	Amersham, Piscataway, USA
iCycler iQ™ Real Time PCR Detection System	Biorad, Hercules, USA
Incubator Hera Cell 240	Hereus, Hanau, Germany
Intesifying screen Kodak, Biomax MS (8x10inch)	Amersham, Piscataway, USA
LUMIstar	BMD Labtech, Offenburg, Germany
MassARRAY™ MALDI-TOF MS	Sequenom, San Diego, USA,
Micro Centrifuge II	NeoLab, Heidelberg, Germany
Microscope Axiovert 40C	Zeiss, Göttingen, Germany
Microwave	Bosch, Stuttgart, Germany
MWG Primus^{TM} Thermal Cycler	MWG, Biotech, Ebersberg, Germany
Nanoliterspotter Spectro Point Nanoliter Pipetting	Sequenom, San Diego, USA
Owl Separation Systems	Portsmouth, USA
Pipett workstation Genesis RSP 150 Workstation	Tecan, Crailsheim, Germany
Pipett workstation Multimek 96 Automated	Beckmann Coulter, Fullerton, USA
Pipett workstation Te-MO	Tecan, Crailsheim, Germany
Rotator Intelli Mixer RM-2M	NeoLab, Heidelberg, Germany
Rotator Roto-Shake Genie	Scientific Industries, Bohemia,
Scale Kern 440-33	Kern, Ballingen, Germany
Scientific Industries Vortex Genie 2	Scientific Industries, Bohemia,
Siliciumchip 384-well SpectroCHIP Bioarray System™	Sequenom, San Diego, USA, Sequenom, San Diego, USA
Thermocycler Eppendorf Mastercycler	Eppendorf, Hamburg, Germany
Thermocycler PCR PTC 225 Tetrad Peltier	MJ Research, Boston, USA
Thin layer chromatography plate	Whatman, Brentford, UK
Waterbath Köttermann	Köttermann GmbH, Uelze, Germany
Verifuge 20RS	Hereus, Hanau, Germany

2.1.7 Software

3100 Data Collection Software Version	Applied Biosystems, Weiterstadt. Germany
ABI PRISM 3730 Data Collection Software v3.0	Sequenom, San Diego, USA
AliBaba 2.1	http://darwin.nmsu.edu/~molb470/fall2003/Projects/solorz/aliBaba_2_1.htm
Adobe Photoshop	Adobe Systems, Edinburgh, UK
BioEdit Sequence Alignment Editor	Ibis Therapeutics, Carlsbad, USA http://www.mbio.ncsu.edu/BioEdit/bioedit.html
Bioworks	Beckman Coulter, Fullerton, USA

CHIP Bioinformatics Tools	http://snpper.chip.org/bio/
DNA Sequencing Analysis Software Version 3.6.1	Applied Biosystems, Weiterstadt, Germany
EndNote 5.0	ISI ResearchSoft, Berkeley, USA
Gemeni 3.5	Tecan, Crailsheim, Germany
Haploview	http://www.broad.mit.edu/mpg/haploview/
eShadow	http://eshadow.dcode.org
MassARRAY Assay Designer 3.0	Sequenom, San Diego, USA
MassARRAY Typer v3.4	Sequenom, San Diego, USA
MATinspector	http://www.genomatix.de/
Microsoft Office	Microsoft, Redmont, USA
National Center for Biotechnology Information	http://www.ncbi.nlm.nih.gov/
NEB cutter 2.0	http://tools.neb.com/NEBcutter2/index.php
NetPrimer	http://www.premierbiosoft.com/netprimer/netprlaunch/netprlaunch.html
Primer Express™	Applied Biosystems, Foster City, USA
Protein forecasting 1.0	http://www.aber.ac.uk/~phiwww/prof /index.html
SpectroDESIGNER	Sequenom, San Diego, USA
Spectro TYPER_RT 2.0.0.3	Sequenom, San Diego, USA, Germany
SAS/Genetics software 9.1.3.	SAS, Cary, USA
Vector NTI	Invitrogen, Carlsbad, USA
Vista Genome Browser	http://pipeline.lbl.gov/cgi-bin/gateway2

2.1.8 Cells and cell lines

Jurkat T-cell line	DSMZ, Braunschweig, Germany
E.coli DH10B E.coli XL10-Gold	Stratagene, La Jolla, USA

2.2 Methods

2.2.1 Study populations

2.2.1.1 Study populations for association analyses

Between 1995 and 1996 a cross sectional study was conducted to assess the prevalence of asthma and allergies in schoolchildren (9 to 11 years) in Munich (ISAAC II = International Study of Asthma and Allergies in Children), southern Germany, and Leipzig, eastern Germany [127]. Phenotypes of the children were determined for asthma, atopic dermatitis and hay fever by self administered questionnaires. Objective tests for atopy (skin prick test, total and specific serum IgE measurements) and lung function (basic lung function and provocation tests) were performed. Additionally, blood was drawn to obtain serum, plasma and DNA.

For the purpose of this genetic analysis, only children of German origin of whom both DNA and IgE data were available were included (N = 1,120, Munich n = 528, Leipzig n = 592).

2.2.1.1.1 Questionnaire

Standardized core questionnaires were given to the parents of the children through the schools including detailed questions on incidence and severity on atopic symptoms like wheezing, shortage of breath, atopic eczema and allergic rhinitis. Children whose parents reported a doctors` diagnosis of asthma or recurrent “asthmatic” or “spastic” bronchitis were classified as having asthma. Likewise, atopic dermatitis and hay fever were defined as the parents report of a doctors` diagnosis. Additional to the pulmonary case history, information was available on the affection of other diseases of the child, age and origin of the child and the parents, family history of atopic diseases, living conditions and the socioeconomic status [9,128].

2.2.1.1.2 Skin examination

Participating children were examined for visible flexural dermatitis according to a photographic protocol under detailed instructions. This check involved the inspection of skin around eyes, around the sides and front of the neck, in front of the elbows, behind the knees and in front of the ankles [9,128].

2.2.1.1.3 Skin-Prick-Test

Skin-Prick-Tests were performed with participating children for six common aeroallergens with highly standardized extracts and lancets including a positive and negative control (histamine and NaCl, respectively). In Munich, children were tested for *dermatophagoides pteronyssinus*, *dermatophagoides farinae*, *alternaria tenuis*, cat dander and mixed grass and tree pollen using extracts of ALK (Horsholm, Denmark). In contrast, in Leipzig extracts of *dermatophagoides pteronyssinus*, grass, birch and hazelnut pollen, cat and dog hair (Stallerkit, Stalleergènes, France) were used. A child was classified to be sensitized to specific allergens if a wheel size reaction of ≥ 3mm in comparison to the negative control was observed [128,129].

2.2.1.1.4 Pulmonary function and bronchial challenge

In Munich, bronchial hyperresponsiveness was assessed by inhaling hyperosmolar (4,5%) saline for periods of increasing duration of 0.5, 1, 2, 4 and 8 minutes whereas FEV1 (Forced Expiratory Volume) was measured 1min after each inhalation period and followed by the next challenge period after 3min. If the FEV1 fell by ≥15% from the baseline value, it was defined as a positive result. If the fall in FEV1 was lower, the duration of inhalation was doubled or the measurement was repeated under the same conditions [128].

In contrast, participating children in Leipzig were exposed to cold-air-challenge for provocation. Challenge was conducted for 4min with isocapnic hyperventilation (22 times the FEV1 per min) of dry, cold air (-15°C, measured at the mouthpiece). Bronchial hyperresponsiveness was defined as a fall of FEV1 as large or larger than the value of the 95th percentile of the reference population [129]. All children were asked to withhold bronchodilator medications before challenge.

2.2.1.1.5 Blood collection, DNA extraction and IgE measurement

DNA was isolated from EDTA blood using the methods described in detail in section 2.2.2. For further genetic analyses the DNA concentration was determined by photo-spectrometry (2.2.4) and pre-amplified to increase the amount of accessible DNA (2.2.7).

For the measurement of total and specific IgE levels venous blood was taken. Total serum IgE levels were assessed using the Imulite system (BPC Biermann, Bad Nauheim, Germany). Levels of specific serum IgE were measured in one central laboratory (Free University of Berlin, Germany) against a panel of aeroallergens (mixed grass pollen, birch pollen, mugwort pollen, D. *pteronyssinus*, cat dander, dog dander, *Clasdosporium herbarum*) and a panel of food allergens (egg white, milk protein, cod fish, wheat flour, peanut, soybean) both by Fluorescence Enzyme Immuno Assay (SX1 CAP, Pharmacia, Lund, Sweden) [9].

2.2.1.2 Study populations for gene expression analyses

2.2.1.2.1 Study population for the characterization of STAT6 isoforms

The population comprised 40 unrelated, randomly selected German volunteers (non-atopic and atopic adults). DNA was isolated from EDTA blood of each individual using the methods described in detail in section 2.2.2 DNA concentration was determined by photo-spectrometry (2.2.4).

The technique RFLP (Restriktion Fragment Length Polymorphism, 2.2.8.4) was applied to genotype the polymorphism C2892T in this population. Subsequently, ten donors each either carrying the homozygous wildtype (N = 5) or the polymorphic (N = 5) allele of SNP C2892T were selected for the initial *STAT6* gene expression analysis.

2.2.1.2.2 Study population for the gene expression analyses of STAT6 isoforms (overall STAT6, STAT6d, STAT6e)

In 2000, the cross-sectional PARSIFAL (Prevention of Allergy-Risk factors for Sensitization In children related to Farming and Anthroposophic Lifestyle) study was conducted to assess the role of environmental and lifestyle factors in the development of allergies and asthma. The goal of this study was to investigate if characteristics of farming and anthroposophic lifestyle may protect from atopic diseases. Therefore, 14,901 children between the age of 5 to 13 years were recruited in five European countries: Austria, Germany, Netherlands, Sweden and Switzerland. Children were recruited from families practicing traditional small scale farming, families with an anthroposophic lifestyle and respective reference groups [130]. A child who lived on a farm and whose family ran a farm was termed as a *farm child* [131] . Children with an *anthroposophic lifestyle* were recruited through Rudolph Steiner schools. The *reference children* did not fulfill these criteria but were selected as they lived in the neighborhood of farms (country side but no farm exposure) and Steiner school children.

The questionnaires on health outcomes and farm exposure were based on ISAAC II core questions and the ALEX (Allergy and Endotoxin) study protocol. In addition, information on the child`s farm activities and the mother`s farm exposure during pregnancy was collected. Questions related to the child`s contact with different farm animals, consumption of farm milk, regular stable or barn visits and helping with haying were also included [131].

For gene expression analyses, only children from the Swiss branch of the PARSIFAL study were included. DNA was accessible for 311 children and RNA was collected in 324 samples. In total, 239 individuals were analyzed in which both DNA as well as RNA samples were available in sufficient quantities.

2.2.2 DNA-Extraction

DNA was extracted from fresh EDTA whole blood using the standard protocol of the salting out procedure by Miller [132]. Through an initial centrifugation step (2000-3000g, 15–24°C, at least 15min) the EDTA blood was separated into different fractions of erythrocytes, DNA-containing layer of leukocytes and plasma. The plasma was then taken off and frozen at -20°C for further analyses. The pellet was resuspended in lysis buffer whereby the cloudy solution became clear after lysis of all cells was completed (approximately 10min). This step was followed by an additional centrifugation step (10min, 500g). The supernatant of the formed leukocyte pellet at the bottom of the tube was decanted and the pellet was again resuspended in lysis buffer. These steps were repeated in total three times. An incubation step at 37°C was performed overnight after the addition of 5ml Proteinase K buffer, 25µl Proteinase K (20mg/ml) and 250µl 20% SDS. Therefore, proteins were digested and DNA went into solution. On the following day, 1500µl NaCl (5M) were added to each sample, vortexed and centrifuged for 30min at 3200g. The supernatant containing the DNA was then transferred into a new tube and precipitated with 20ml 100% ethanol and washed with 70% ethanol. After 10min centrifugation at 3200g the supernatant was decanted and the pellet air-dried. The extracted DNA was solved in 150-200µl TE buffer and stored at 4°C for following applications. DNA concentration was determined using a photo-spectrometer (2.2.3, Eppendorf, Hamburg, Germany).

The general salting out method was only applicable for fresh blood, therefore blood that was stored at –20°C DNA had to be extracted using the FlexiGene-Kit (Qiagen, Hilden, Germany), also based on a salting out method, following the manufacturer`s protocol (Qiagen, Hilden, Germany).

2.2.3 Gel electrophoreses of nucleic acids

As a method for identification, cleaning and quality control of DNA, plasmids and PCR products, gel electrophoresis in corresponding agarose gels (1.5 – 3%) depending on the size of the fragments were performed in 0.5x TBE buffer. The method depends on the fact that nucleic acids under neutral pH conditions reveals a negative charge which leads to the migration in an electric field, whereas its pace is proportional to the logarithm of the molecular weight.

Adding the DNA incoperating reagent ethidiumbromide (0.5µg/ml), it is possible to visualize double stranded nucleic acids under UV light (302nm). The size of a fragment was assessed with a standard of a defined length (100bp or 1kb ladder, NEB, Ipswich, USA).

2.2.4 Measurement of nucleic acids using photo-spectrometry

The yield of DNA, plasmid or PCR product was determined through the measurement of extinction (OD260nm = DNA, RNA, oligonucleotide; OD280nm = protein) using an Eppendorf (Hamburg, Germany) photo-spectrometer. DNA-concentration is calculated applying the law of Lambert-Beer. Therefore, the DNA-concentration of 1µl of a solution with an OD_{260} = 1 measured in a cuvette with a thickness of 1cm results in 50ng/µl for DNA and 40ng/µl for RNA measurement. The proportion of OD_{260}/OD_{280} displays the degree of contamination indicated by an optimal range from 1.8 to 2.0. Values ≤ 1.6 refer to increased protein content or residues of phenol. In contrast, values ≥ 2.0 suggest single stranded DNA (hypochromic effect).

2.2.5 Polymerease-Chain-Reaction (PCR)

2.2.5.1 Background

The Polymerase-Chain-Reaction (PCR) is a method to amplify a specific nucleotide sequence of a defined length as a result of cycles of denaturation, polymerization and elongation.

A reaction of a cyclic amplification of a DNA fragment involves the use of specific forward and reverse primers (designed with VectorNTI, Invitrogen, Carlsbad, USA), heat-stable *Taq* polymerase, the four desoxynucleotidetriphosphates (dATP, dCTP, dGTP, dTTP) and template DNA. Each reaction was carried out under optimized conditions (e.g. buffer, length and temperature of cycle steps) for each of the amplified regions of interest.

In a first step, the template DNA is denaturated at 95°C to facilitate the binding of specific primers to the single stranded target sequence in the following annealing phase. The annealing temperature is dependant on the A/T and C/G content of the primers and is crucial for the specificity of the PCR. The annealing reaction is followed by the extension, also called elongation, starting at the free 3´OH of the primers catalyzed by the *Taq* polymerase leading to the assembly of the corresponding dNTPs. The newly synthesized DNA double strands serves as a matrix for the next cycle. The number of cycles of denaturation, annealing and extension are repeated until the exponential amplification reaches a plateau. Therefore, a theoretical maximum of 2^n double stranded DNA is contained after n cycles.

2.2.5.2 Experimental design

PCR reactions were carried out corresponding to the respective application with gene specific primer pairs under optimized conditions. table 1 and table 2 display a standard scheme of all general components and conditions of a PCR reaction which were adapted if necessary. To eliminate potential unspecific binding of primers, amplifications were performed under stringent PCR conditions. Therefore, different buffer systems, *Taq* polymerases, length and temperature of PCR steps were used.

components	amount	concentration
DNA template	3.3µl	20ng/µl
dNTP-Mix	1µl	10mM
PCR buffer	5µl	10x
forward primer	0.5µl	25µM
reverse primer	0.5µl	25µM
Taq-DNA-Polymerase	0.17µl	1U/µl
dH_2O	add to 50µl	

table 1: Standard components of a PCR reaction.

step	temperature	time
1) initial denaturation	95°C	2min
2) denaturation	95°C	20sec
annealing	62°C ± 7°C	20sec
extension	72°C	20sec
	→repeat step 2 40 times	
3) final extension	72°C	10min

table 2: Standard program of a PCR reaction.

2.2.6 Sequencing reaction and analyses

2.2.6.1 Background

The application of the sequencing reaction allows to determine the nucleotide sequence of a given DNA fragment either originated from a PCR product or plasmid DNA. DNA sequencing is applied using the chain termination method, developed by Frederick Sanger in 1975. Therefore, the respective DNA fragment is amplified only once following the principle of a standard PCR reaction because in addition to the regular dNTPs a marginal portion of fluorescence-marked didesoxynucleotides (ddNTP) is used [133]. The idea comprises that the enzymatic incoperation of the ddNTP induces the termination of an ongoing amplification of a DNA strand due to a missing 3`OH group of the

nucleotides. Furthermore, only one primer is added to the sequencing reaction leading to a termination after each of the elongation steps. As the selection process of the incoperation of the respective dNTP or ddNTP cannot be influenced, as many fragments have to be generated statistically representing each fragment at least once. The application of differently marked ddNTP allows for automatic detection of the terminated fragments of varying length.

The analysis of the respective DNA fragment was performed on a 48-capillary 3730 DNA Analyzer (Applied Biosystems, Fostercity, USA) through the electrophoretic separation of the nucleotides in capillaries (36cm) filled with Performance Optimized Polymer 7 (POP7, Applied Biosystems, Fostercity, USA). The analysis of each fragment, depending on the terminated, fluorescence-marked ddNTP, is performed through an argon laser activating the fluorophors which are recognized by a detector. The connected charge-couple device (CCD) camera converts the detected fluorescence to an electric signal. Due to the fact that smaller fragment migrate faster than larger fragments the ABI PRISM 3730 Data Collection Software version 3.0 is able to provide a chromatogram of all detected fragments of different lengths whereas the y-axis corresponds to the relative concentration of fluorescence and the time is displayed on the x-axis (figure 8).

2.2.6.2 Experimental design

Depending on the application either 25-50ng of PCR product or 500ng of plasmid DNA were used to determine the nucleotide sequence of the respective DNA fragment. The reaction was performed using a standard mastermix (table 3) of 5µl containing *Taq* polymerase, buffer, dNTPs, fluorescence-marked ddNTPs and specific primers (10µM) in a 96-well sequencing plate (PeqLab, Erlangen, Germany) using a standard Eppendorf PCR cycler (table 4, Hamburg, Germany). Application details are annotated in the respective section.

components	amount	concentration
DNA template: PCR-product plasmid DNA	0.5-2.5µl	25-50ng 500ng
Big Dye version 3.1	1.0µl	
sequencing buffer	1.0µl	5x
primer	0.5µl	10µM
HPLC-H_2O	add to 5µl	

table 3: Standard mastermix of a sequencing reaction.

step	temperature	time
denaturation	95°C	30sec
annealing	50°C	15sec
extension	60°C	4min
	→repeat 28 times	

table 4: Standard program of a sequencing reaction.

After completion of the sequencing reaction the products were precipitated with 100µl of 70% ethanol and incubated for 15min at room temperature followed by a centrifugation step (3000g, 30min). The supernatant was discarded and the pellet was resuspended in 150µl 70% ethanol as a final purification step. After centrifuging (3000g) for an additional 10min the supernatant was again discarded and the pellet was air-dried for 5min at room temperature. The purified reaction was stored at –20°C for further applications. The pellet was resuspended in 50µl HPLC-H_2O. 25µl were transferred to specific 96-well microtiter plates for sequencing analysis under described conditions.

2.2.6.3 Data analyses

The nucleotide sequences were analyzed in direct comparison to the respective reference sequence either using the software BioEdit Sequence Alignment Editor version 7.0.5.3 (Ibis Therapeutics, Carlsbad, USA) or the VectorNTI version 10.1.1 software (Invitrogen, Carlsbad, USA). As displayed in figure 8 only clearly analyzable sequences were used to perform subsequent experiments.

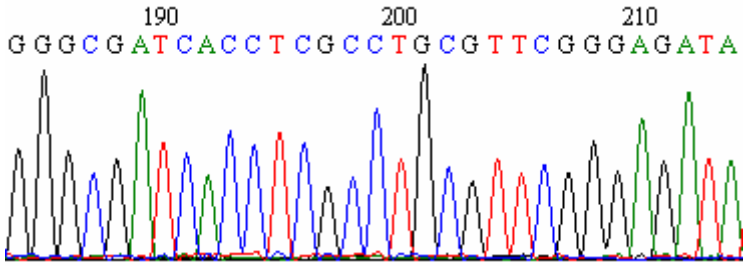

figure 8: Example of a clearly analyzable sequence.

2.2.7 Primer extension Preamplification (PEP)

2.2.7.1 Background

The aim of the Primer Extension Preamplification (PEP) PCR is to perform an unspecific amplification of the whole genome to increase the amount of available DNA of each individual. This is a necessary step to be able to investigate the relation between the genetic background of many genes, the development and manifestation of a disease. The PEP PCR was initially established by Zhang *et al.* [134] and is comparable to a regular PCR (2.2.5) only that random N15-Primers (random combination of 15 bases) instead of specific primers are used for the amplification. The reaction itself is performed under less stringent conditions. Due to the assortment of different unspecific primers, a random annealing to the complementary DNA of each sample occurs generating putatively overlapping products of different length. Hence, the amplification of the whole genome is achieved whereas the initially applied amount of DNA is increased by approximately 40-fold.

2.2.7.2 Experimental design

In the context of this work the PEP PCR was performed to increase the DNA amount of each individual from the study population of Munich and Leipzig (2.2.1.1). This was a necessary step as due to the study design no additional blood could be taken.

A PEP reaction was performed in 50µl (table 5) whereas 5ng of the double-stranded DNA was denaturated at 94°C, followed by an annealing of the random N15 primers under little stringent conditions beginning at 37°C but continuously increasing with each of the 90 cycles. As displayed in table 6 denaturation, annealing and extention (50 cycles) were then carried out to achieve an optimal amplification of the whole genome.

components	amount	concentration
DNA template	1.0µl	5ng/µl
dNTP-Mix	5.0µl	10mM
PCR buffer Y	5.0µl	10x
N15 primer	10µl	200µM
$MgCl_2$	5.0µl	25mM
Taq-DNA-Polymerase	0.17µl	1U/µl
dH_2O	23µl	

table 5: Standard components of a PEP-PCR reaction.

step	temperature	time
1)	94°C	3min
2)	94°C	1min
3)	37°C	2min
4)	37°C →repeat step 4 90 times	1sec + 0.2°C/sec (R=2.0°C/sec)
5)	55°C	4min
6)	15°C→repeat steps 2-5 49 times	
7)	72°C	5min

table 6: Standard program of a PEP-PCR reaction.

2.2.8 Genotyping by MALDI-TOF MS

2.2.8.1 Background

The basic principle of MALDI-TOF MS (Matrix Assisted Laser Desorption Ionisation-Time of Flight Mass Spectrometry) is based on laser pulse induced desportion and ionization of molecules which are embedded in a large excess of the respective sample on a crystalline matrix. The matrix is necessary to facilitate the evaporation of the sample by absorbing the laser energy, to prevent a photolytic damage of the fragments and to avoid their potential interaction. The method was first described by Karas *et. al.* [135].

The development of MALDI-TOF over the past few years has evolved to detect DNA fragments between 1,000 to 9,000Da (corresponding to 3-30 bases) with an accuracy of 0.1 – 0.001%. With the application of this high through-put process the analyses of Single Nucleotide Polymorphisms (SNP) in large study populations became more feasible [136].

The ionized DNA fragments are released from the matrix through the induced laser pulse and are accelerated in a constant filed of energy till the molecules reach the detector at the end of a linear tube. The time between the release of the molecules from the matrix and their identification at the detector reflects their molecular weight (figure 9). Since the ionized DNA-fragments are accelerated with a comparable potential, but differ in their terminal base at the 3`end (figure 10) depending on the genotype, deviations of the molecular weight and pace represent the genotype of each sample.

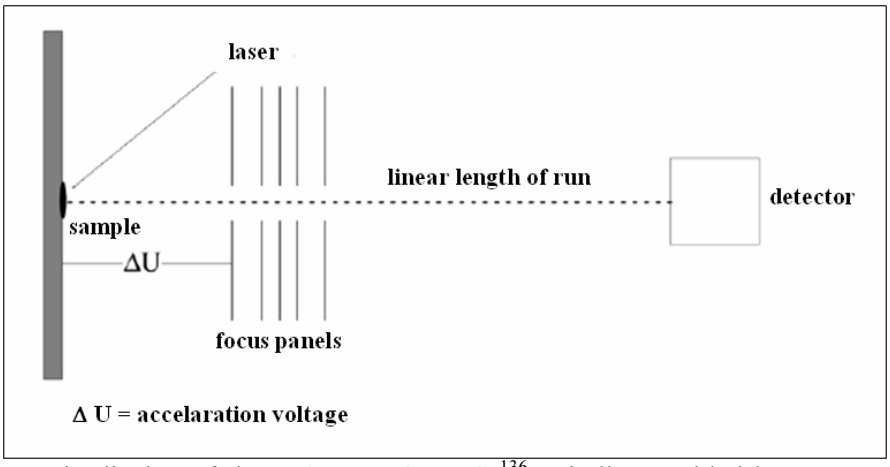

figure 9: Schematic display of the MALDI-TOF MS [136]. Kindly provided by Dr. H. Gohlke, GSF, Munich, Germany.

Genotyping with MALDI-TOF MS was performed using the Primer Extension procedure (hME = homogeneous Mass Extent™) using the technology commerlialized by Sequenom (Hamburg, Germany). In a first step a regular PCR product was generated to amplify approximately 100bp of the region of interest followed by a Base-Extension Reaction specific for each SNP which allowed to differentiate between PCR products of different masses corresponding to the allelic status of a polymorphism (described in detail in section 2.2.8.3.4) [136].

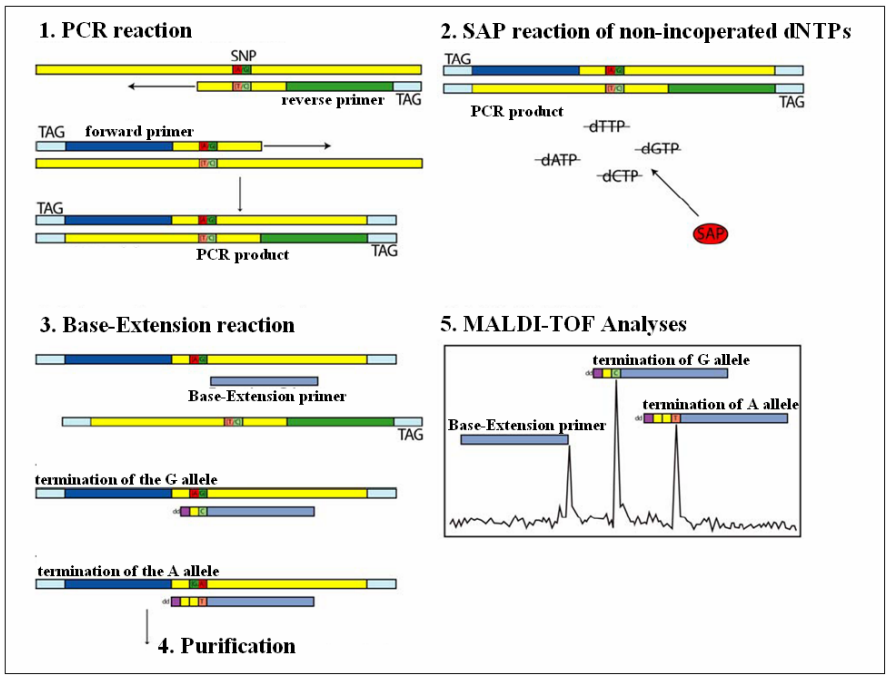

figure 10: Schematic display the hME reaction. Kindly provided by Dr. H. Gohlke, GSF, Munich, Germany.

2.2.8.2 Plate set-up for MALDI-TOF MS

The preamplified DNA of each sample of the study population of Munich and Leipzig (2.2.1.1) was pipetted on 96-well plates (Peqlab, Erlangen, Germany) in a defined scheme as displayed in figure 11 to control for pipetting errors. Additionally, four negative controls (dH_2O) were inserted at certain positions on each of the plates to identify putative contaminations.

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B			NC									
C												
D					NC							
E							NC					
F												
G											NC	
H												

figure 11: Layout of the pipetting scheme for one 96-well plate. NC = negative control. Kindly provided by the GSF, Munich, Germany.

Subsequently, the DNA was transferred and diluted to 1ng/µl onto 96-well "clusterplates" (Qiagen, Hilden, Germany). Four 96-well plates were combined to one 384-well plate (figure 12, ABgene, Surrey, UK) using the pipette workstation Genesis RSP 150 (Tecan, Crailsheim, Germany). The defined scheme of the negative controls allowed to monitor possible rotations of the respective 96-well plates after generating the 384-well plate as they should form a picture of four squares (figure 12). The DNA solution (5µl) was then air-dried. Thus, a final concentration of 5ng of each sample was used for genotyping.

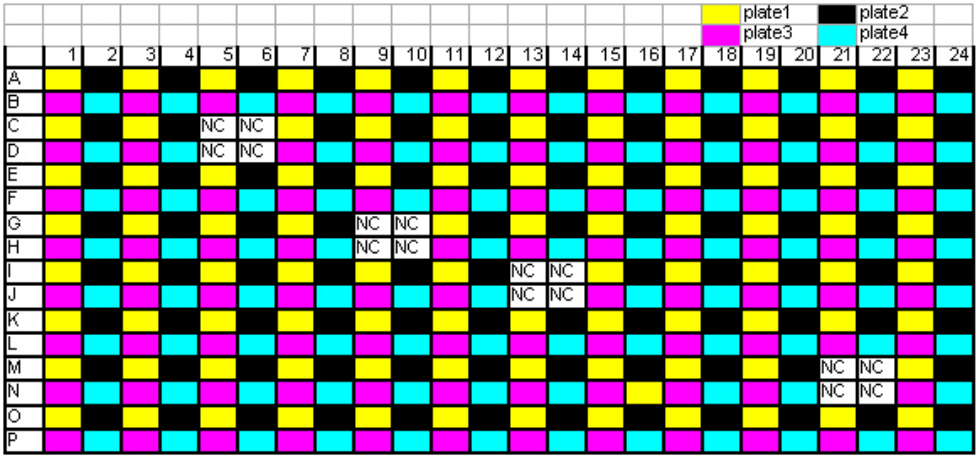

figure 12: Layout of the pipetting scheme for one 384-well plate consisting of four 96-well plates. NC = negative control. Kindly provided by the GSF, Munich, Germany.

2.2.8.3 Experimental design using the hME reaction

2.2.8.3.1 Primer design

To design primers for the standard PCR and the Base-Extension Reaction, the software *MassARRAY Assay Designer 3.0* (Sequenom, San Diego, USA) was applied to obtain the most specific primer binding only to the region of interest. Additionally, a known 10bp motif (= tag) was attached to each primer to perform simultaneous PCR reactions (multiplex) in one well due to different masses of the emerging PCR products. An artificially added sequence (= tag) to the SNP-specific primers was necessary for the deviation between the primers themselves and the Base-Extension PCR products. All standard PCR primers as well as the Base-Extension Reaction primers are displayed in table 7.

SNP	rs number	location	direction	PCR primer (5`3`)	Base-Extension primer (5`3`)
C-1839T	rs2598483	5`flanking	fwd	ACGTTGGATGTACTGTCTGTGGATGTCTGC	CACGTGCTTTGTATGTGCTCC
			rev	ACGTTGGATGTGTGCACATACGTGTTCACG	
G-1284A	rs3001428	5`flanking	fwd	ACGTTGGATGTGATTTCCAGAACCAGCTCC	CCACACACGTGCACTCATG
			rev	ACGTTGGATGTTCAAACCCTCTGCTGCCTG	
C2892T	rs324011	intron 2	fwd	ACGTTGGATGGATGCCCCTGGTTTAAGGTG	ATAGCCCTCCTAGGGAC
			rev	ACGTTGGATGCAGGGACCTCCCATAGATAG	
C9548T	rs3024974	intron 18	fwd	ACGTTGGATGTGACTGACCAAGGGTTGATG	GGGCTTAGTGCTTATCTG
			rev	ACGTTGGATGAAGGTGAGTGTGGTGGTATG	
T12888C	rs1059513	3`UTR	fwd	ACGTTGGATGAATTCCTGTTAGCCAGGTGG	ACGAAGAATCTCAGCCCT
			rev	ACGTTGGATGCGTTCACACAGCTATACACG	
T12949C	rs4559	3`UTR	fwd	ACGTTGGATGTGAACGTGTATGTACCTAGG	GCAACTAAGGTGCCAGCTATA
			rev	ACGTTGGATGTAGCATATGTCAGAGAGGCC	

table 7: General information of the six genotyped STAT6 polymorphisms: SNP name, database SNP accession number, location, and genotyping primers.

2.2.8.3.2 PCR reaction

To determine the genotype of the respective *STAT6* SNPs of each of the samples of the study population, 6µl of a PCR mastermix (table 8) were added to each of the wells of the 384-well plate containing 5ng of air-dried DNA. The amplification of each fragment (approximately 100bp) was carried out in Thermocyclers of MJ research (Waltham, USA) under standardized conditions (table 9). The efficiency of the PCR reaction and the size of the respective PCR products were controlled by applying 1µl of 24 samples on a 3% agarose gel whereas a negative control was tested to exclude potential contamination of the mastermix.

components	amount for single reaction	amount for 384-well plate (480x)	concentration
dNTP-Mix	0.6µl	288µl	2mM
$MgCl_2$	0.24µl	115.2µl	25mM
PCR buffer	0.6µl	288µl	10x
forward primer	0.1µl	48µl	10mM
reverse primer	0.1µl	48µl	10mM
HotStar Taq-DNA-Polymerase	0.02µl	9.6µl	5U/µl
dH_2O	4.34µl	2083.2µl	
total volume	**6µl**	**2784µl**	

table 8: Standard mastermix for the PCR reaction applying the hME method.

step	temperature	time
1) activation step	95°C	15min
2) denaturation	95°C	30sec
annealing	56°C	30sec
extension	72°C	1min
	→repeat step 2 45times	
3) final extension	72°C	10min

table 9: Standard protocol for the PCR reaction applying the hME method.

2.2.8.3.3 Shrimp Alkaline Phosphatase (SAP) reaction

To remove excessive dNTPs which were not incoperated, the catalytic enzyme Shrimp Alkaline Phosphatase (SAP, Amersham, Piscataway, USA) was added to the PCR products in order to prevent an unpredictable influence on the SNP-specific PCR reaction. A mastermix (table 10) was prepared and 2µl containing SAP, SAP buffer and HPLC-H_2O were added using a Multimek 96-channel autopipette (Beckman Coulter, Fullerton, USA). Incubation of the enzyme with the PCR product for 20 min at 37°C in a Thermocyclers (MJ research, Waltham, USA) was conducted followed by an inactivation step of the SAP enzyme at 85°C for 10 min.

components	amount for single reaction	amount for 384-well plate (480x)	concentration
hME buffer	0.17µl	81.6µl	10x
SAP (Shrimp Alkaline Phosphatase)	0.3µl	144µl	1U/µl
dH_2O	1.53µl	734.4µl	
total volume	**2µl**	**960µl**	

table 10: Standard mastermix for the SAP reaction applying the hME method.

2.2.8.3.4 Base-Extension-Reaction

With the help of the Base-Extension-Reaction, the genotype of each sample was identified. The designed primers (table 7) are generated to bind either 5` or 3` exactly one base in front of the SNP (depending whether it is a sense- or an anti-sense primer). The application of didesoxynucleotides (ddNTP) induces the termination of further amplification depending on the allelic status of a SNP which allows the identification of the genotype when comparing the molecular weights between groups (homozygous carriers of the wildtype allele, heterozygous and homozygous carriers of the polymorphic allele, figure 10).

For example, the genotype of a respective polymorphism is G/A. Therefore, the ddNTP G,T and C and the dNTP A are used for the Base-Extension-Reaction. Accordingly, the PCR reaction is stopped if the homozygous G allele is present. In contrast, the A as a dNTP and an additional ddNTP is incoperated if the individual is a homozygous carrier of the polymorphic A allele. In heterozygous individuals both conditions are present.

The Base-Extension Reaction was performed in 10µl reactions using the enzyme *Thermosequenase* (Amersham, Piscataway, USA) with a final concentration of 5,4µM of the extension primer and the mastermix displayed in table 11. For the Base-Extension-Reaction the denaturation was performed at 94°C for 2min, followed by 94°C for 5sec, 52°C for 5sec, and 72°C for 10sec for 55 cycles. All necessary pipetting steps were performed using the Multimek 96-channel autopipette (Beckman Coulter, Fullerton, USA). Thermocycling steps were carried out in a MJ research Thermocycler (Waltham, USA).

components	**amount for single reaction**	**amount for 384-well plate (480x)**
hME Extend-Mix	0.2µl	96µl
hME primer	0.054µl	25.92µl
Thermosequenase (MassExtend Enzyme)	0.018µl	8.64 µl
dH_2O	1.728µl	829.4µl
total volume	**2µl**	**960µl**

table 11: Standard mastermix for the Base-Extension-Reaction applying the hME method.

2.2.8.3.5 Purification of the Base-Extension Reactions

The final Base-Extension products were purified with SpectroCLEAN resin (Sequenom, San Diego, USA) to remove residues of salts from the reaction buffer and Mg^{2+}. 16µl of water were added to each Base-Extension Reaction conducted with a Multimek 96-channel autopipette (Beckman Coulter, Fullerton, USA), resulting in a total volume of 26µl. This solution was rotated for at least 10min and centrifuged at 800g for 3min in an Eppendorf Centrifuge 5810 (Hamburg, Germany).

2.2.8.3.6 MALDI-TOF Mass spectrometry

After performing the clean-up 8-12nl of each reaction solution were dispensed onto a 384-format SpectroCHIP (Sequenom, San Diego, USA) pre-spotted with a matrix of 3- hydroxypicolinic acid (3-HPA) by a SpectroPoint nanodispenser (Sequenom, San Diego, USA). Additional to the samples, a calibrant (Sequenom, San Diego, USA) was spotted containing a mixture of oligonucleotides of a know mass for the calibration of the MALDI-TOF system. The spotted SpectroCHIPs were then transferred onto a metallic sample carrier and data acquisition was conducted with the modified Bruker Biflex Matrix Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometer (Sequenom, San Diego, USA). Genotyping calls were made in real time with the MASSARRAY RT software (Sequenom, San Diego, USA).

2.2.8.4 Genotyping by Restriction Fragment Length Polymorphism (RFLP)

2.2.8.4.1 Background

The method Restriction Fragment Length Polymorphism (RFLP) is used for detecting polymorphisms. Purified DNA is usually first amplified by PCR (2.2.5). The DNA is then cut by suitable endonucleases (also called restriction enzymes) into fragments, which only cut at specific DNA recognition sequences. The genotype of each individual can therefore be identified, as the length of fragments on a agarose gel (2.2.4) varies depending on the DNA sequence. On that account, a restriction enzyme is selected which is predefined to cut only when a specific allele is present at a polymorphic locus. Additionally, the restriction enzyme used should possess a constitutive recognition site independent of the allelic status as a positive control recognition site. However, this method is not easily applicable for high-through-put genotyping of large study populations due to high costs, lack of automation potential and limitations in standardization.

2.2.8.4.2 Experimental design

RFLP was applied to genotype the polymorphism C2892T in the random sample of adults (N = 40, 2.2.1.2.1) and the PARSIFAL population (N = 311, 2.2.1.2.2) which were used for the *STAT6* gene expression study (described in detain in section 2.2.14). The restriction enzyme AvaII was selected as it cuts only when the wildtype allele of the SNP was present. Specific PCR primer were designed (VectorNTI, Invitrogen, Carlsbad, USA) whereby the forward primer (R_Int2fwd 5`-CCCACAGAACCAGAGCCAGGTCCAG-3`) introduced an artificial constitutive AvaII restriction site. In contrast, the reverse primer (R_Int2fwd 5`-CCTATCTCCTTGGGCAGCCAGGTGAC-3`) was designed to abolish an additional AvaII restriction site. The PCR was performed following the protocols displayed in table 12 and table 13.

components	amount	concentration
DNA template	3.3µl	20ng/µl
dNTP-Mix	1.0µl	10mM
PCR buffer (NEB)	5.0µl	10x
R_Int2fwd	0.3µl	25µM
R_Int2rev	0.3µl	25µM
Taq-DNA-Polymerase (NEB)	0.17µl	1U/µl
dH_2O	add to 50µl	

table 12: Components of the PCR reaction to genotype the polymorphism C2892T in the *STAT6* gene.

step	temperature	time
1) initial denaturation	95°C	2min
2) denaturation	95°C	30sec
annealing	58.9°C	40sec
extension	72°C	40sec
	→repeat step 2 43 times	
3) final extension	72°C	10min

table 13: Program of the PCR reaction to genotype the polymorphism C2892T in the *STAT6* gene.

13µl PCR product of each sample were supplemented with 1.5µl NEB buffer 4 and 0.5µl of AvaII and incubated for 3h at 37°C. The respective fragments were run on a 3% agarose gel at 100V for 1.5h. The expected sized of each fragment depending on the genotype are displayed in figure 13.

expected fragment size	genotype			
	uncut	homozygous wildtype	heterozygous	homozygous polymorphism
275bp				
256bp				
204bp				
52bp				
19bp				

figure 13: Expected fragment size depending on the genotype of C2892T after the digestion of the PCR product with AvaII. The "uncut" relates to 5µl of undigested PCR product.

2.2.9 Statistical analyses

Statistical analyses were performed to investigate putative associations between either single SNPs or haplotypes of the *STAT6* gene and the major atopic phenotypes asthma, atopic dermatitis, hay fever, atopy (indicated by skin-prick-test) and levels of total serum IgE.

2.2.9.1 Hardy-Weinberg equilibrium

The Hardy-Weinberg equilibrium (HWE) states that under common conditions, after one generation of random mating, the alleles for the next generation for any given individual are chosen randomly [137]. Therefore, the genotype frequencies at a single gene locus will become fixed at a particular equilibrium value. The HWE assumes that evolution will not arise in a population if the following criteria are met: mutation and natural selection is not occurring, the population is infinitely large, all members of a population breed, all mating is totally random, everyone produces the same number of offspring and there is no migration in or out of the population.

In the simplest case of a single locus with two alleles A and a with allele frequencies of *p* and *q:* p is defined as the frequency of the dominant allele and q as the frequency of the recessive allele for a trait. The HWE predicts that the genotypic frequencies for the AA homozygous to be p^2, the Aa heterozygous to be 2*pq* and the aa homozygous to be q^2:

$p^2 + 2pq + q^2 = 1$

These frequencies are called Hardy-Weinberg frequencies (or Hardy-Weinberg proportions). The Hardy–Weinberg principle may be applied in two ways: either a population is assumed to be in Hardy–Weinberg proportions, in which the genotype frequencies can be calculated, or if the genotype frequencies of all three genotypes are known, they can be tested for deviations that are statistically significant. The HWE is generally performed using Person`s chi-squared test by analyzing deviations between expected and observed allele frequencies.

If a significant ($p \leq 0.05$) deviation from the HWE is present, it can either be due to methodical difficulties with the genotyping procedure (frequent case) or due to biological reasons caused by evolution (rare case). Hence, the HWE was calculated using the SAS statistical software package (Version 8.2) for quality control of the genotyping method MALDI-TOF MS (2.2.8).

2.2.9.2 Statistical analyses for the association study

2.2.9.2.1 Odds Ratio and significance testing

The effect measurement for the chance to develop a disease among different groups is expressed through the Odds Ratio (OR). The Odds Ratio reflects in this study the ratio of the odds of an event (e.g. to develop a disease like asthma) occurring in one group (carriers of the rare allele) to the odds occurring in the second group (carriers of the frequent allele). An OR = 1 points out that the development of a disease is equally likely in both groups. An OR > 1 reflects that the risk of the manifestation of the respective disease is increased in carriers of the rare allele. In contrast, an OR < 1 indicates a protective effect when the polymorphic allele is present. Additionally, the 95% confidence interval (95% CI) is reported reflecting the reliability of the outcome data. The 95% CI states that the

calculated parameter will be i.e. 95 times within the given confidence interval if an experiment is repeated 100 times. P-values ≤ 0.05 indicate a significant effect.

2.2.9.2.2 Genetic models for statistical testing of qualitative traits

Applying different genetic models for statistical testing, associations between qualitative traits and the disease of interest e.g. asthma can be performed (table 14) [138]. The linear model is used to estimate the relationship between an in- or decrease of the probability to develop a disease with an accumulation of a defined allele. In contrast, the allelic model compares the frequency of alleles between patients and healthy controls. The dominant model (dominant on the rare allele) merges heterozygous and homozygous individuals for the rare allele to develop a disease in comparison to homozygous carriers of the wildtype allele. Whereas the recessive model compares homozygous carriers of the rare allele against the two other putative allelelic states taken together. This model may be difficult to apply depending on the allele frequency and the size of the population to receive enough statistical power. To analyze the association between a genotype and a phenotype for the allelic, dominant and the recessive model the χ^2-test is used. An Armitage-Trend-Test is required for the linear model.

For the association study of the *STAT6* gene all models were calculated. However, only the outcome data of the dominant model is displayed as this model gave the best fit to the data in independant study populations. Thus, the risk between *STAT6* single SNPs or haplotypes and the development of atopic phenotypes like asthma, atopic dermatitis, hay fever and atopy (measured by skin-prick-test) was calculated using the SAS statistical software package (Version 8.2).

model	group 1	group 2	group 3
linear	homozygous wildtype	heterozygous	homozygous SNP
allelic	allele 1	allele 2	
dominant	homozygous wildtype	heterozygous + homozygous SNP	
recessive	homozygous wildtype + heterozygous	homozygous SNP	

table 14: Different approaches of statistical models to calculate associations between a disease and a polymorphism.

2.2.9.2.3 Genetic analyses of quantitative traits

To analyze associations between SNPs and continuous variables (= quantitative traits) like total serum IgE levels the t-Test as statistical tool has to be applied. In general, the t-Test is used to compare means of continuous variables between groups also using different models (linear, allelic, dominant, recessive). To implement this tool, the analyzed data has to meet the criteria to be normally distributed across a study population. This feature is not accomplished for levels of IgE, respectively. Thus, IgE levels have to be log transformed leading to a normal distribution of the IgE values and the ability to perform the appropriate testing [138]. For better interpretation of the observed results, the data is presented in geometric means of total IgE. Additionally, Wilcoxson-testing is performed as a non-parametric-test which does meet the criteria of a normally distribution of IgE levels.

The associations between any genotyped *STAT6* SNP (single SNPs as well as hyplotypes) and geometric means of IgE levels were calculated with the dominant model applying the SAS statistical software package (Version 8.2). P-values ≤ 0.05 indicate a significant effect.

2.2.9.2.4 Haplotype analyses

A haplotype is a set of alleles of polymorphisms which are transmitted together to the next generation. In contrast to a genotype, these polymorphisms are located on a single chromatide and are not detectable on both strands of a chromosome.

Haplotype frequencies were estimated using the EM (expectation-maximization) algorithm. All haplotype analyses were implemented using the SAS/Genetics module [139]. To evaluate associations with traits of the respective phenotypes (asthma, hay fever and atopy measured by skin-prick-test), and the respective haplotypes of the *STAT6* gene, a haplotype trend regression model was applied, where the estimated probabilities of the haplotypes are modelled in a logistic regression as independent variables [139,140]. As this test is limited to binary variables, quantitative traits such as IgE levels had to be transferred into binary outsomes.

Based on all available German study subjects at the age of 9-11 years with existing IgE measurements, IgE percentiles were calculated (N = 4,400). Thus, for *STAT6* haplotype analyses the 50th (60.4 IU/ml), 66th (115 IU/ml) and 90th percentile (457 IU/ml) for total serum IgE was used as outcome variables.

2.2.9.2.5 Gene-by-gene interaction analyses

Additional to the haplotype analyses, the Cordell model for epistasis was used to test for statistically significant interactions in models with both additive and dominant effects and respective interaction terms [141]. Epistasis, the interaction between genes, is tested using a global test for interaction of the model with interaction terms versus the null model without interaction terms. This method identifies combinations of genes showing interaction, but does not identify individual haplotypes causing the interaction. As the power for tests of interaction is limited in comparison to test for main effects (genotypes), a significance level of 0.10 was used for the global test of interaction. All statistical analyses were carried out using the SAS statistical software package (Version 8.2), while the haplotype analyses were implemented using the SAS/Genetics module [139].

2.2.10 *In silico* analyses

Different application of *in silico* analyses for *STAT6* intron 2 were performed as these methods gave the possibility to generate first hypotheses for a putative functional role for this *STAT6* region containing the SNP C2892T.

2.2.10.1 Comparative genomics of humans, dog and mouse genome

Phylogenetic comparisons are used to gain information on an organism's evolutionary relationship. The search for functional DNA sequences that have been conserved between different organisms across a large distance in evolution is the classical approach of comparative genomics. Closely related species generally have a high overall sequence concordance, while the DNA of organisms distantly related usually show a high degree of dissimilarity. However, active conservation of coding regions and regulatory sequences is observed throughout evolution. Natural selection will tend to preserve genes and other functional regions, but will not exert strong preservative pressure on "junk" DNA without any functional relevance, allowing it to mutate more rapidly. Therefore, comparative genomic studies between evolutionarily distant species will identify regions of the human genome with a putative function role of a region of interest: not only exons or the promoter of a gene can be distinguished but also regulatory element within intronic domains can be identified.

To perform comparative genomics, the Java-based online available software VISTA Genome browser (http://pipeline.lbl.gov/cgi-bin/gateway2) was applied. This tool displays the respective similarity in percent between sequences of species which allows the comparison of smaller DNA-fragments as well as whole genes. Besides, conserved areas of the output file indicate also the location of the exons as well as the untranslated regions (UTRs) of the respective gene. Applying VISTA browser (2.2.10.1) sequence conservation across species can be used as a predictor for a putative functional influence of a genetic region of interest.

For the *STAT6* gene the region of intron 2 harboring the SNP C2892T was compared between the human (*Homo sapiens*), the dog (*Canis familiaris*) and the mouse genome (*Mus musculus*) to determine the conservation degree of this region across species using this approach.

2.2.10.2 Phylogenetic shadowing of the human sequence with different primate species

2.2.10.2.1 Background

Within the past years cross-species comparisons between distantly related genomes (human and mouse) have been extensively performed because approved tools were available for these applications. In contrast, primate sequence comparisons were very difficult to interpret due to the high degree of sequence similarity shared between these closely related species. Furthermore, no conclusive analysis was available. Recently, a novel method, phylogenetic shadowing, was pioneered for predicting functional elements in the human genome under selective pressure in multiple sequence alignments [142]. Previous experiments suggested that sequencing the genomes of as few as four to six primate species in addition to humans may be enough to identify much of the conserved functional primate specific DNA sequences in the human genome. With the means of this method information is retrieved that would inevitably be missed by only using more distant relatives.

The software eShadow (http://eshadow.dcode.org) enables to discriminate between regions of high conservation or detect hot spots for mutations. eShadow also considers at what time-point during evolution a mutation event has occurred which is subsequently included in the evalutation.

Using the statistical tool eShadow (http://eshadow.dcode.org), it was possible to predict the location of exons as well as transcriptional regulatory elements and the evaluation of protein domains. The eShadow software includes two distinct approaches for the identification of these functional elements: Hidden Markov Model Islands (HMMI) and Divergence Threshold (DT) scans of multiple sequence alignments [143].

2.2.10.2.2 Experimental design

For sequence comparisons of *STAT6* between humans and primates, DNA from six species containing chimpanzee, gorilla, orang-utan, old world (gibbon, green monkey) and new world (owl monkey) monkeys (figure 14) was acquired from the DNA respiratory of the European Collection of Cell Cultures (ECACC, Salisbury, U.K.). A standard PCR (2.2.5) was applied using the displayed conditions (table 1, table 2) with the specific primers 6391fwd (5´-CCCACAGAACCAGAGCCAGG-3`) and 6816rev (5´-CCGCAGGTGTTGGGGAAAG -3`) amplifying a 425bp region of the *STAT6* gene harboring the SNP C2892T. Sequencing reaction with the STAT6 primer 6391fwd was performed under standard conditions (table 3, table 4) with 25 – 50ng of purified PCR product (Gel Purification Kit, applying the manufacturer`s protocol, Qiagen) and analyzed through the ABI PRISM 3770 capillary DNA sequencer (2.2.6, Applied Biosystems, Foster City, USA). Sequence similarity was calculated using the eShadow software.

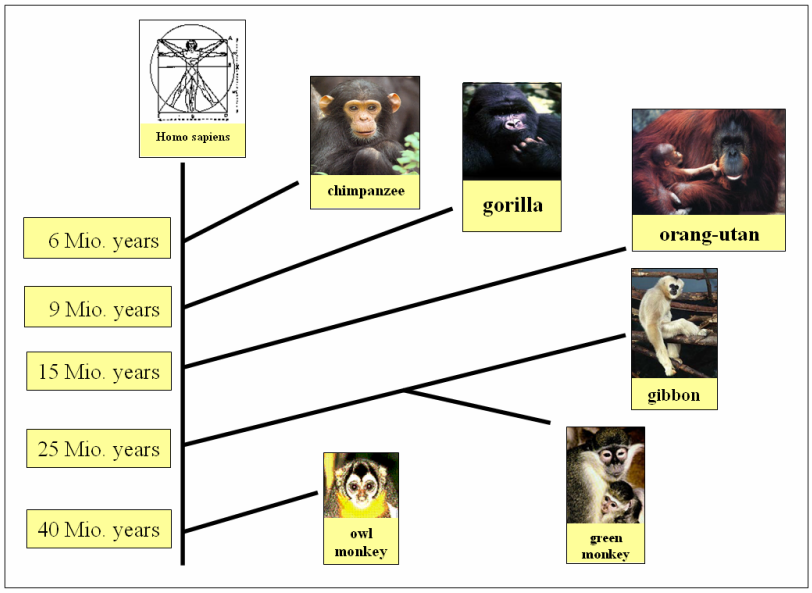

figure 14: Pedigree of Homo sapiens and different primate species used for phylogenetic shadowing. Kindly provided by PD. Dr. Michael Kabesch, University Childrens` Hospital, Munich, Germany.

2.2.10.3 In silico analyses of transcription factor binding

2.2.10.3.1 Background

Even if promoter and intronic regions (predicted to be non-coding) do not influence the amino acid sequence or structure of a protein, they may have profound effects. Putative regulatory elements, e.g. transcription factor binding sites, may be situated within these regions. So far, their functional role has not been investigated sufficiently. The binding of these transcription factors may thus be influenced through genetic variations in the corresponding DNA sequence and in turn affect gene expression. Transcription factors in general play an important role in the initiation of the RNA-polymerase. They are categorized in basal and tissue- or cell-specific transcription factors. Basal transcription factors are necessary for any transcriptional event and they either bind directly to the DNA or to other proteins of the initiation complex. Transcription factors reveal a very diverse structure possessing many different functions e.g. protein kinases or helicases. They are ubiquitous, meaning that they are present in almost all cells of an organism but play only a minor role in the specific gene regulation. In contrast, specific transcription factors activate distinct genes at certain time points in predefined cells and tissues. Once induced, they show a regulatory function either as an enhancer or a repressor after binding to specific DNA sequences.

The software MatInspector version 7.0 (www.genomatix.de) and AliBaba 2.1 (http://darwin.nmsu.edu/~molb470/fall2003/Projects/solorz/aliBaba_2_1.htm) predict putative binding sites for transcription factors within the area of interest by comparing the respective sequence with similarities of binding matrices from the database. Additional to these comparisons, an index is calculated representing the essential binding regions for the respective binding site. This index differentiates between the *matrix similarity* and the *core similarity*. A matrix is generally defined as a selective description of DNA patterns to bind transcription factors attributed to a database. Hence, the *matrix similarity* as the major outcome evaluates the matches of all necessary bases of the binding sequence of a transcription factor to the region of interest whereas a “good” sequence similarity

obtains values of >0.8. Each matrix contains a core sequence which is defined as the highest conserved positions of the matrix (usually 4bp) reflected by the *core similarity* with a maximum score of 1.0.

2.2.10.3.2 Experimental design

Applying MatInspector and AliBaba the concordance of the analyzed sequence with all available matrices of the database were calculated resulting in a qualitative output. Therefore, the program detected in a first step those matrices with a similarity of the core sequence (beyond a defined threshold). In a second step the matrices of the reference sequence to the sequence of interest were compared for putative concordances. Only those transcription factors were displayed which reached a defined threshold of the *matrix similarity*. Due to the fact that both DNA strands were analyzed with these programs, putative binding site were detected on the matrix strand as well as the complementary strand [144,145].

Both softwares were applied to identify putative transcription factor binding sites within intron 2 of the *STAT6* gene. Additional analyses were performed introducing the different allelic status of the polymorphism C2892T located within this region to detect potential changes in the binding properties of these transcription factors whereas the threshold of the *core similarity* was set to a minimum of 0.8.

2.2.11 Cell culture and isolation of different cell types

2.2.11.1 Cell culture of the Jurkat T-cell line

The Jurkat T-cell line (ACC 282, www.dsmz.de) is a human T-cell leukemia cell line which was established from the peripheral blood of a 14-year-old boy with acute lymphoblastic leukemia (ALL) at first relapse. Cells were cultured in RPMI 1640 Glutamax (Gibco, Carlsbad, USA) supplemented with 10% heat-inactivated FCS (Gibco, Carlsbad, USA) and penicillin/streptavidin (5µg/ml, Sigma-Aldrich, Steinheim, Germany) at 37°C with 5% CO_2. Jurkat T-cells grew in suspension as single cells or cell clumps and were split to 300,000 cells/ml every 60h.

2.2.11.2 Isolation of Peripheral Blood Monocytic Cells (PBMC)

The isolation of PBMCs from whole blood was performed within 24h, respectively. The total volume of retained blood was measured using a 10ml sterile pipette and transferred into 50ml conical tubes (Falcon, Bedford, USA). Whole blood was diluted and mixed of a ratio of 1:2 with 1x PBS (Gibco, Carlsbad, USA). The total amount of this solution (whole blood and 1x PBS) was divided into equal amounts of approximately 9ml in 15ml vial tubes (Sarstedt, Nümbrecht, Germany) which already contained 3ml Ficoll Paque™ Plus (Amersham, Piscataway, USA). The suspension was added very carefully and slowly as not to disturb the Ficoll layer. After a centrifugation step at 1400g (Verifuge 20RS, Heraeus, Hanau, Germany) for 30min at room temperature separate layers of serum, white blood cells, Ficoll and red blood cells are obtained. First, the serum was aliquoted into 2ml Cryogenic vials (Corning, Corning, USA) and frozen at –70°C for further analyses. Subsequently, the white blood cell layer was carefully removed using a 10ml sterile pipette without disturbing the layers of Ficoll and red blood cells. The white blood cells from each 15ml tube were transferred into a fresh 15ml tube and supplemented with RPMI to a total volume of 14ml. The supernatant was discarded after a centrifugation step at 1100g for 10min. Then, the pellet was resuspended in 10ml of RPMI. To count the total cell number 50µl were removed, supplemented with 50µl of ACK lysis buffer (Bio

Whittacker, Walkersville, USA) and incubated on ice for 7.5min. An additional 50µl of 0.4% trypan blue was added. The total number of white blood cells was determined by pipetting 10µl of cells to a counting chamber. Meanwhile, the 15ml tubes containing the rest of the cells were centrifuged at 1100g for 10min. Supernatant was discarded and the remaining pellet was either frozen at –80°C for RNA isolation (2.2.14.2) or a subsequent isolation of $CD4^+$ cells (2.2.11.3) was conducted.

2.2.11.3 Isolation of $CD4^+$ T-cells

After retaining PBMCs (2.2.11.2), an additional isolation step was performed to receive $CD4^+$ T-cells using the $CD4^+$ T-cell isolation kit (MACS, Miltenyi Biotec, Bergisch Gladbach, Germany). To prevent capping of antibodies on the cell surface and non-specific cell labeling, cells were kept on ice and only pre-cooled solutions were used. The PBMC pellet containing $1x10^5$ cells was resuspended in 90µl of 0,5% BSA-PBS/EDTA isolation buffer. 10µl of Biotin-Antibody Cocktail was added, mixed and incubated for 10min at 4°C. An additional incubation step for 15min at 4°C followed after adding 20µl Anti-Biotin MicroBeads. Cells were then washed with 1-2ml of isolation buffer. The supernatant was completely removed after a centrifugation step (300g, 10min, 4°C). In the meantime, the LC columns were prepared by rinsing with 2ml isolation buffer. Cells were subsequently resuspended in 500µl isolation buffer. The cell suspension was then applied onto the LD column. The unlabeled cells passing through were collected. The column was washed twice with 1ml isolation buffer by gradually adding the buffer once the column reservoir was empty. The total eluat was collected as it contained the unlabeled pre-enriched $CD4^+$ T-cell fraction. 2ml of the isolation buffer were pipetted onto the LD column. The fraction containing the magnetically labeled $CD4^+$ cells was immediately flushed out by firmly applying the plunger supplied with the column. The total cell number was determined and used for the preparation of nuclear extract as described in detail in 2.2.12.2. Additionally, a purity test (FACS) with an anti-CD4 antibody (4µl/$5x10^5$ cells, Beckmann Coulter, Fullerton, USA) was performed with 50µl of the $CD4^+$ T-cell as well as $CD4^-$ T-cells as a negative control.

2.2.12 Electrophoretic mobility shift assay

Electrophoretic Mobility Shift Assays (EMSA) were performed to verify the predicted binding of transcription factors to *STAT6* intron 2 depending on C2892T. Additionally, binding domains within the region harboring the polymorphism were investigated in detail.

2.2.12.1 Background

EMSA, also referred to as a gel shift assay or band shift assay, is a common technique used to study DNA/protein interactions. This procedure can determine whether a protein or mixture of proteins is capable of binding to a given DNA sequence. Thus, it can indicate if more than one protein molecule is involved in a binding complex. A mobility shift assay generally involves the electrophoretic separation of a DNA/protein mixture on a non-denaturated polyacrylamide gel whereas the speed at which different molecules move through the gel is dependant on their size and charge.

First, commercially available complementary single-stranded oligonucleotides representing the region of interest (approximately 20bp – 40bp) are annealed to double-strands. The purified double-stranded oligonucleotides (further referred to as "probes") are then end-labeled with radioactive ^{32}P and incubated with nuclear extract from cells containing a mixture of proteins. Formed DNA/protein complexes can be visualized on a polyacrylamide gel because of their increased size and slower migration in comparison to unbound DNA. These DNA/protein complexes are identified with specific antibodies against the respective protein. Thus, a "supershift" is achieved due to the increasing size of the DNA/protein/antibody complex. Competition experiments are preformed to verify the specificity of the observed complexes by pre-incubating the nuclear extract with a 100-fold molar excess of a competing unlabeled oligonucleotide prior to the addition of the labeled probe.

2.2.12.2 Preparation of nuclear extract

The Jurkat T-cell line was cultured under standard conditions for at least five but no more than 10 generations. 24h before the preparation of nuclear extract, Jurkat T-cells were set to 300,000 cells/ml. For each condition (unstimulated and stimulated) 14 x 10^6 cells were centrifuged at 300g for 10min (4°C) and resuspended in 7ml ice cold RPMI 1640+Glutamax (fetal calve serum free, Gibco, Carlsbad, USA) to receive 2 x 10^6 cells/ml. Cells were then transferred into 25cm^2 flasks (Corning, Corning, USA) and cultured for 3h either unstimulated or stimulated with PMA (50ng/ml)/Ionomycon (1µM). All further steps had to be performed on ice. After stimulation, cells were washed in 5ml ice cold 1x PBS followed by a centrifugation step at 300g for 10min (4°C). This washing step was repeated three times in total. After the last centrifugation the supernatant was decanted. The pellet was resuspended in 1ml ice cold 1x PBS and transferred into 1.5ml Eppendorf tubes. All liquid was aspirated from the pellet (14,000g, 2min, 4°C). Jurkat T-cells were then resuspended in 120µl of low-salt Buffer A (10mM HEPES, 3mM $MgCl_2$, 40mM KCl, 1mM DTT, 5% glycerol, 0.2% NP-40, 1mM PMSF, 10µg/ml aprotinin,10µg/ml leupeptin, 10µg/ml antipain, 10µg/ml pepstatin, 1mM benzamidine, 1mM sodium orthovanadate, 1mM sodium fluoride, 5mM β-glycerophosphate) to release the nuclei from the cells. The viability of the cells (5µl) was checked with 10µl of trypan blue after approximately 5min of incubation. The reactivity of trypan blue is based on the fact that its chromophores are negatively charged and do not interact with the cell unless the membrane is damaged. Therefore, all cells which excluded the dye were still viable. Centrifugation for 2min at 14,000g (4°C) was performed to separate the nuclei (sediment) from all cytoplasmic components (supernatant). By resuspending the pellet in an appropriate volume (approximately 60-100µl) of the high-salt Buffer C (20mM HEPES, 1.5mM $MgCl_2$, 420mM NaCl, 0.2mM EDTA, 1mM DTT, 25% glycerol, supplemented with protease and phosphatase inhibitors as in Buffer A), all proteins were released from the nuclei of the Jurkat T-cell (30min incubation on ice) followed by a centrifugation step (20min, 14,000g, 4°C). The nuclear extract of Jurkat T-cells was divided into 10µl aliquots, immediately frozen in liquid nitrogen and stored at –80°C for further analyses.

Nuclear extract of $CD4^+$ T-cells was obtain under equivalent conditions after performing all necessary isolation step for the respective cell type as described in detail in 2.2.11.3.

2.2.12.2.1 Quantification of nuclear extract

To quantify the amount of obtained protein, the concentration was determined using the BCA Protein Assay Kit (Pierce, Rockford, USA). A serial dilution for the necessary standard curve with BSA (2mg/ml) was pipetted in duplicates following the scheme presented in table 15.

dilution	amount of BSA	amount of H_2O	final concentration
1	1000µl of BSA (2.0mg/ml)	600µl	1.2mg/ml
2	500µl of BSA (1.2mg/ml)	500µl	0.6mg/ml
3	500µl of BSA (0.6mg/ml)	500µl	0.3mg/ml
4	500µl of BSA (0.3mg/ml)	500µl	0.15mg/ml

table 15: Dilution series to quantify protein content in nuclear extract using the BCA Protein Assay Kit (Pierce, Rockford, USA).

950µl of working reagent (= 50:1 reagent A:B) was added to 50µl of each dilution (1 – 4) whereas a 1:10 dilution of nuclear extract was prepared of each condition (5µl + 995µl working solution) followed by an incubation for 30min at 37°C. The quantification of the protein was determined using an Eppendorf photo-spectrometer.

2.2.12.3 Annealing and purification of the probes

2.2.12.3.1 Annealing of single-stranded oligonucleotides

Single-stranded oligonucleotides were obtained from Metabion (Martinsried, Germany) and diluted to 1µg/µl with freshly prepared TE (pH 8.0). All oligonucleotides used are presented in table 16.

probe name	probe sequence (5`3`)
NF-κB consensus site fwd	TCAGAGGGGACTTTCCGAGAGGCG
NF-κB consensus site rev	CGCCTCTCGGAAAGTCCCCTCTGA
SP consensus sequence_fwd	ATTCGATCGGGGCGGGGCGAGC
SP consensus sequence_rev	GCTCGCCCCGCCCCGATCGAAT
STAT6_C2892fwd	AGTGGTGGTGGGGACGGTCCCTAGGAGGGCT
STAT6_C2892rev	AGCCCTCCTAGGGACCGTCCCCACCACCACT
STAT6_2892Tfwd	AGTGGTGGTGGGGACAGTCCCTAGGAGGGCT
STAT6_2892Trev	AGCCCTCCTAGGGACTGTCCCCACCACCACT
block 1_fwd	CTGTTTGGTGGGGACAGTCCCTAGGAGGGCT
block 1_rev	AGCCCTCCTAGGGACTGTCCCCACCAAACAG
block 2_fwd	AGTGGGTTGTGGGACAGTCCCTAGGAGGGCT
block 2_rev	AGCCCTCCTAGGGACTGTCCCACAACCCACT
block 3_fwd	AGTGGTGGTGTTTCAAGTCCCTAGGAGGGCT
block 3_rev	AGCCCTCCTAGGGACTTGAAACACCACCACT
block 4_fwd	AGTGGTGGTGGGGACCTGAACTAGGAGGGCT
block 4_rev	AGCCCTCCTAGTTCAGGTCCCCACCACCACT
block 5_fwd	AGTGGTGGTGGGGACAGTCCAGCTTAGGGCT
block 5_rev	AGCCCTAAGCTGGACTGTCCCCACCACCACT
block 6_fwd	AGTGGTGGTGGGGACAGTCCCTAGGCTTTAG
block 6_rev	CTAAAGCCTAGGGACTGTCCCCACCACCACT
block 7_fwd	CTGTTTGGTGGGGACGGTCCCTAGGAGGGCT
block 7_rev	AGCCCTCCTAGGGACCGTCCCCACCAAACAG
block 8_fwd	AGTGGGTTGTGGGACGGTCCCTAGGAGGGCT
block 8_rev	AGCCCTCCTAGGGACCGTCCCACAACCCACT
block 9_fwd	AGTGGTGGTGTTTCAGGTCCCTAGGAGGGCT
block 9_rev	AGCCCTCCTAGGGACCTGAAACACCACCACT
block 10_fwd	AGTGGTGGTGGGGACTTGAACTAGGAGGGCT
block 10_rev	AGCCCTCCTAGTTCAAGTCCCCACCACCACT
block 11_fwd	AGTGGTGGTGGGGACGGTCCAGCTTAGGGCT
block 11_rev	AGCCCTAAGCTGGACCGTCCCCACCACCACT
block 12_fwd	AGTGGTGGTGGGGACGGTCCCTAGGCTTTAG
block 12_rev	CTAAAGCCTAGGGACCGTCCCCACCACCACT
S6_NF-κBfwd	CTATCTATGGGAGGTCCCTGGCTGCC
S6_NF-κBrev	GGGCAGCCAGGGACCTCCCATAGATAG
rs167769C_fwd	TTGGAGCAGAAACATCAGAGAAGGGACCT
rs167769C_rev	AGGTCCCTTCTCTGGACGTTTCTGCTCCAA
rs167769T_fwd	TTGGAGCAGAAACGTCAGAGAAGGGACCT
rs167769T_rev	AGGTCCCTTCTCTGGATGTTTCTGCTCCAA

table 16: Single-stranded oligonucleotides used for EMSA. Underlined sequences display the putative binding site for the respective transcription factor. For block 1-17 the underlined sequence indicates the mutated bp.

The annealing reaction was carried out in screw cap tubes (Sarstedt, Nümbrecht, Germany) containing 180µl of annealing buffer (TE pH 8.0, 100mM NaCl) and 10µl of the respective sense and anti-sense single-stranded oligonucleotide. After boiling the solutions for 5min to denature the oligonucleotides completely, the solutions were slowly cooled down to 37°C at room temperature. The double-stranded probes were stored at 4°C for the following purification.

2.2.12.3.2 Purification of double-stranded probes

To separate the double-stranded probes from residues of single-stranded oligonucleotides, a 15% polyacrylamidgel was prepared containing 30ml 30% polyacrylamid (19:1, Biorad, Hercules, USA), 12ml 5x TBE, 18ml of dH_2O, 30µl TEMED and 450µl of freshly diluted 10% APS. An additional aliquot of 30µl TEMED was mixed with 1000µl of the 15% gel to seal the cleaned glass plates (separated through spacer) to avoid leakage. After the seal was polymerized, the gel was pored and the 15well-comb inserted. The gel was air-dried for approximately 1h at room temperature followed by a pre-run in 0.6l of 0.5x TBE for at least 20min at 18mM (4°C).

To each annealing reaction (2.2.12.3) 40µl of 80% glycerol was added. Samples were vortexed and 120µl were loaded into each well of the gel. Additionally, 50µl of loading dye was applied to an extra well to control for the running pace of the probes, respectively. Approximately 3 – 4h later, the gel was carefully removed from the glass plates and each of the double-stranded probes was detected under UV light using a Thin Layer Chromatography plate (Whatman, Brentford, UK). The bands were then cut out of the gel, transferred into a 1.5ml Eppendorf tube and crushed with the back of a 1ml syringe (BD, Franklin Lakes, USA). To elute the probe from the gel, the Freeze`n Squeeze kit (Biorad, Hercules, USA) was applied. Therefore, the crushed gel was transferred onto the column and put at –20°C for 5min followed by a centrifugation step of the samples at 13,000g for 3min (room temperature). The flow-through was stored at 4°C. An additional 100µl of TE was applied to the column, centrifuged (13,000g, 3min, room temperature) and the appropriate samples were combined. To elute the entire probe, 200µl TE were added to each column and incubated for 2h at 37°C on a

shaker. After a last centrifugation step (10min, 13,000g, room temperature) the respective samples were combined. A precipitation step was performed using 0.3 volume NaAc (3M, pH 5.2) and 3 volumes of ice-cold 100% ethanol. The samples were stored at –20°C over night. Approximately 48h later, the probes were centrifuged at 4°C for 20-30min (13,000g) and the pellets were air-dried after a washing step with 1ml 70% ethanol. To resuspend the obtained pellet, 10µl of TE were added to each sample. After the probes went completely into solution the concentration was determined using an Eppendorf photo-spectrometer (2.2.4). For quality control 100ng were run on a 3% agarose gel. All samples were diluted to 50ng/µl whereas aliquots were stored at -20°C.

2.2.12.3.3 End-labeling of the probes with ^{32}P

The labeling reaction of the purified double-stranded probes (2.2.12.3.1) was carried out in a total volume of 25µl (table 17) using γ[^{32}P]-ATP (250µCi). The T4 Polynucleotide Kinase (New England Biolabs, Ipswich, USA) initiated the transfer of the phosphate group of the radioactive γ[^{32}P]-ATP to the 5` hydroxyl end of each probe. After an incubation step at 37°C for 15min this reaction was stopped through inactivating the reaction at 68°C for an additional 10min.

components	amount	concentration
double-stranded probe	2µl	(50ng/µl)
T4 Polynucleotide Kinase buffer	2.5µl	10x
T4 Polynucleotide Kinase	2µl	10U/µl
γ[^{32}P]-ATP	5µl	(250µCi)
dH_2O	13.5µl	

table 17: Mastermix for the labeling reaction of the double-stranded probes

Free radioactivity was removed by mini Quick Spin Oligo Columns (Roche, Mannheim, Germany). Therefore, the total volume of each reaction was increased to 50µl with dH_2O. To evenly resuspend the Sephadex matrix in the column buffer, the columns were vigorously inverted several times. The removal of excess buffer was achieved through a centrifugation step at 1000g for 1min at room temperature. The labeling reaction was immediately added to each column, centrifuged at 1000g for 4min (room temperature). 25µl aliquots were stored at –20°C. To determine the counts per minute (cpm) of the respective probes, 1µl was read using a beta-counter (Beta Counter LS 6000IC, Beckman Coulter, Fullerton, USA).

2.2.12.3.4 Binding reaction

In silico analyses suggested the binding of the transcription factor NF-κB at the respective *STAT6* site containing the SNP C2892T. First, different conditions for the binding reaction were tested to establish EMSA settings. Binding of NF-κB to a known NF-κB consensus binding sequence [146] with nuclear extract from Jurkat T-cells was analyzed. The different binding buffers applied for optimizing the binding conditions (20µl) are displayed in table 18. As buffer C, in which the nuclear extract was resuspended, already included NaCl and Glycerol, the additional amounts of both components were added up to a defined final concentration as displayed in table 18.

Additionally, a titration of the applied Poly(dI-dC)-Poly(dI-dC) (Amersham, Piscataway, USA) concentration was performed to obtain the most specific DNA/protein complex formation possible.

components	final concentration			
	buffer 1	**buffer 2**	**buffer 3**	**buffer 4**
nuclear extract	5µg	5µg	5µg	5µg
Tris HCl (pH 8.0)	10mM	10mM	10mM	
EDTA (0.5M)	1mM	1mM	0.5mM	0.1mM
β-mercaptoethanol		0,1mM		1.4mM
NaCl (1M)	100mM	80mM	50mM	100mM
Glycerol	10%	4%	4%	12.5%
Poly(dI-dC)-Poly(dI-dC)	0.25-5µg/reaction	0.125-5µg/reaction	1µg/reaction	1µg/reaction
DTT	10mM		0.5mM	
$MgCl_2$	0.5mM		1mM	
KCl				2mM
NP40				0.5%

table 18: Different binding buffers for the EMSA experiments were applied to identify the optimal binding conditions. For all further EMSAs binding buffer 2 with 1µg/reaction Poly(dI-dC)-Poly(dI-dC) was used.

Supershift and competition experiments were performed by either adding the corresponding antibody (4µg, Santa Cruz Biotechnology, Santa Cruz, USA) or unlabeled probes in a 100-fold molar excess prior to adding the labeled probe (30min incubation on ice) (table 19). After an additional incubation for 30min on ice, DNA/protein complexes were resolved on a 5% polyacrylamid gel (2.2.12.3.5).

probe	**antibodies (4µg each)**	**competition**
NF-κB consensus site	p50, p65	NF-κB consensus site, C2892, 2892T
SP consensus site	SP1, SP2, SP3, SP4	SP consensus site, C2892, 2892T
C2892	p50, p65, SP1, SP2, SP3, SP4	C2892, 2829T
2892T	p50, p65, SP1, SP2, SP3, SP4	2829T, C2892
S6_ NF-κB	p50, p65, SP1, SP2, SP3, SP4	NF-κB + SP consensus site, C2892, 2892T
rs167769C		rs167769C, rs167769T
rs167769T		rs167769T, rs167769C

table 19: For each probe supershift (4µg) and competition experiments (100-fold molar excess) were performed with the respective antibody and probe as indicated in the table.

2.2.12.3.5 PAGE (Polyacrylamid Gel Electrophoresis)

A 5% non-denaturated polyacrylamid gel (table 20) was prepared a day before performing the binding reaction as displayed in table 18 and described in detail in 2.2.12.3.4.

components	amount	concentration
dH_20	41ml	
TBE	6ml	5x
glycerol	3ml	80%
polyacrylamid	10ml	30% (19:1,filterd)
TEMED	30µl	10mg/ml
APS	450µl	10%

table 20: Components of a 5% non-denaturated polyacrylamid gel. APS is freshly diluted right before use.

The binding reaction of each EMSA sample was loaded into the pre-cleared wells of the gel. The gel was run for approximately 8h at 4°C in 0,5x TBE starting at 18mA (at the beginning about 180V, during course of electrophoresis voltage increased to >400V). After PAGE, the gel was transferred to a double layer of Whatman paper and dried on a gel dryer (Biorad, Hercules, USA) for 60 min at 80ºC. The dried gel was then exposed for an appropriate period of time to autoradiography film (Amersham Piscataway, USA) at -80ºC using a cassette with intensifying screens (Kodak, Biomax MS, Amersham, Piscataway, USA) on each side.

2.2.13 Luciferase gene expression studies

Luciferase gene expression experiments were performed to investigate the *STAT6* promoter activity in the human Jurkat T-cell in general and to assess if intron 2 may exert any function as a *cis*-regulatory element putatively depending on the genotype of the SNP C2892T.

2.2.13.1 Cloning of STAT6 gene expression vectors

2.2.13.1.1 Background

Genetic cloning in general is defined as the transfer of a DNA fragment of interest from one organism to a self-replicating genetic element such as a bacterial plasmid [147]. Cloning of any DNA sequence involves four steps: amplification, ligation, transformation and screening/selection. Initially, a region of interest needs to be amplified, commonly achieved by the means of PCR. Both, vector and insert, are then digested with the corresponding restriction enzyme to generate compatible ends for cloning. Subsequently, ligation is performed whereby the amplified fragment is inserted into the linearized vector. A re-circular complex (= recombinant DNA molecule) emerges which is then transformed into a donor cell (*E.coli*). Most commonly electroporation is applied for transformation, although a number of alternative techniques are available e.g. chemical sensitization or heat shock of cells. Finally, the transformed cells are cultured to increase the yield. It is necessary to identify the cell colonies that have been transformed with the vector containing the desired insert sequence. Most cloning vectors include selectable antibiotic resistance marker (e.g. ampicillin, canamycin), which allow only for cells with the vector to grow. However, this selection step does not assure that the DNA insert is present in the vector. Additionally, restriction assay and sequencing reaction are needed to control for the insertion, orientation and the correctness of the sequence of the fragment, respectively.

The use of reporter gene vector systems in general serves the purpose to analyze the expression and regulation of a gene of interest in eukaryotic cells. These types of genetic experiments can include the investigation of promoters and *cis*-regulatory elements.

In this study the pGL3-Basic vector (figure 15, Promega, Madison, USA) was used as the initial vector for all further generated plasmids. The pGL3-Basic vector lacks an eukaryotic promoter and enhancer sequence, which allowed maximum flexibility in cloning putative regulatory sequences. The vector contains a firefly luciferase reporter gene, an ampicillin gene conferring ampicillin resistance in *E.coli* and a multiple cloning region (MCR) upstream of the luciferase gene for eight restriction sites (KpnI, SacI, MluI, NheI, SmaI, XhoI, BglII, HindIII). Two additional restriction sites (SalI, BamHI), which carry compatible ends with two sites in the MCR (XhoI – SalI, BglII – BamHI), were located downstream of the luciferase gene to allow for an easy interchange of DNA inserts. Thus, positional effects of a putative genetic element may be tested.

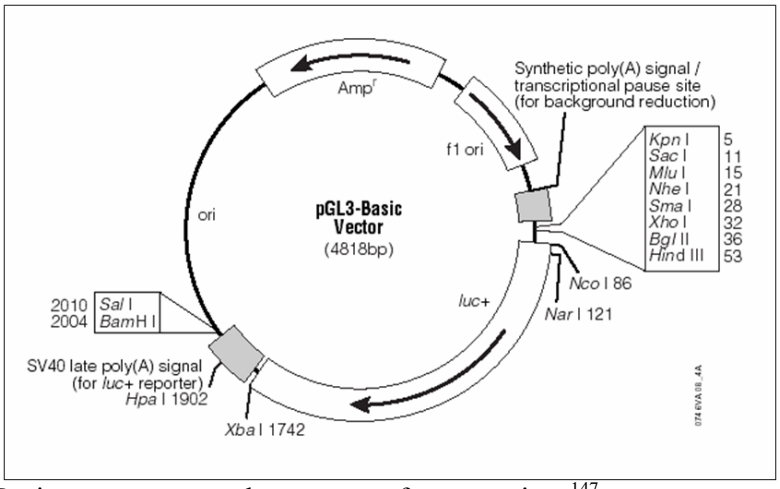

figure 15: pGL3-Basic vector map and sequence reference points [147].

2.2.13.1.2 STAT6 promoter gene expression vectors

Three pGL3-Basic vector constructs containing *STAT6* promoters of different sizes (pBP78 = 5643bp, pBP88 = 2588bp, pBP86 = 1081bp, figure 16) were initially available which were a kind gift from Bharvin Patel [120]. An additional *STAT6* promoter construct (pBP85) was generated in which approximately 200bp at the 5`end of pBP86 were deleted as a GT-repeat at the 5`end of *STAT6* was located within this region (figure 16). Previous findings already described a relation between the *STAT6* gene expression and another GT-repeat within the *STAT6* promoter [120]. The correct sequence of

all *STAT6* promoter constructs was verified by re-sequencing (2.2.6) using the primers indicated in table 21. Site-directed-mutagenesis (2.2.13.2, Stratagene , La Jolla, USA) was applied if necessary to delete missense mutations in the respective constructs (table 22).

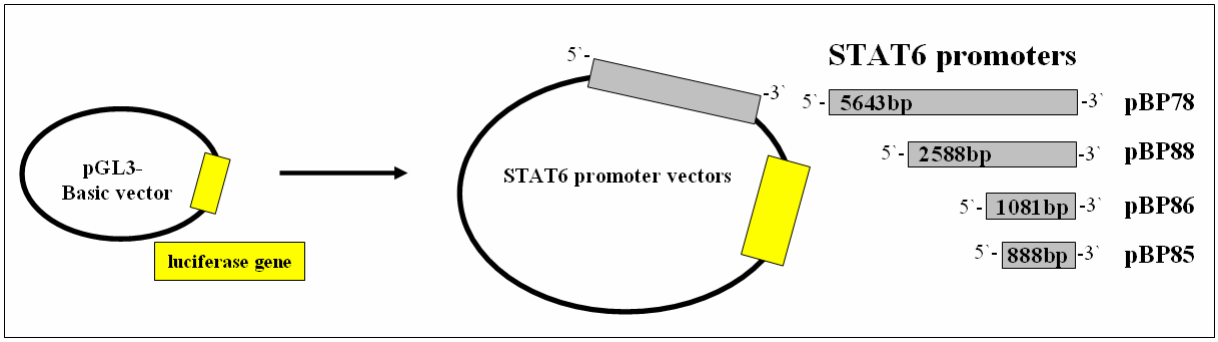

figure 16: The pGL3-Basic vector was used to generate four *STAT6* promoter vectors of different size (pBP78 = 5643bp, pBP88 = 2588bp, pBP86 = 1081bp and pBP85 = 888bp) to perform *STAT6* promoter studies.

primer name	primer sequence (5`3`)	*STAT6* promoter vector
RVprimer 3	CTAGCAAAATAGGCTGTCCC	pBP85, pBP86, pBP88, pBP78
pBP_660fwd	CCTCAGCAGCTTCCCACAATTC	pBP78
pBP_1164fwd	CCTGGGTTTAAGCAATTCTCGTG	pBP78
pBP_1819rev	CAGAAGGCATCAAGCGAAGTAAG	pBP78
pBP_1696fwd	TGGGCGACAGAGGGAGAC	pBP78
pBP_2265fwd	GGATCTTGCTATGTTGCCCAG	pBP78
pBP_2802fwd	CGATTTTGTGACTCACTATTGGG	pBP78
pBP_3132fwd	TCTCCTGACCTTGTGATCTGCC	pBP88, pBP78
pBP_3411fwd	CATTTGAGGGATTGACACCTGATTG	pBP88, pBP78
pBP_3805fwd	TGGACAAGCCAATGGACAGAG	pBP88, pBP78
pBP_4352fwd	GCACCAGGGGGCAGCATAG	pBP88, pBP78
pBP_5090rev	TAAAAAGCCGCCAGGGAGC	pBP86, pBP88, pBP78
pBP_4902fwd	CTGTGAGGAGAGAGCAAAGACC	pBP85, pBP86, pBP88, pBP78

table 21: Pimers used to verify the correctness of the sequence of the respective *STAT6* promoter pBP85, pBP86, pBP88, pBP78.

primer name	primer sequence (5`3`)	*STAT6* promoter vector
pBP_3717fwd	GCAGGCAGAGTGGGCACTCGGGAGCACATACAAAGC	pBP88, pBP78
pBP_5598fwd	TTTTTGGTGGTGGTGGTGGAAGGGGGGAGGTGCTAG	pBP85, pBP86, pBP88, pBP78

table 22: Pimers used for the respective *STAT6* promoter constructs to perform site-directed mutagenesis (2.2.13.2).

2.2.13.1.3 Cloning strategies of STAT6 intron 2

To study the effects of *STAT6* intron 2 on gene expression in a luciferase assay, a genomic region of 1652bp of the *STAT6* gene was amplified by the means of a standard PCR (table 23, STAT6_5164fwd 5`-CATCCTCTCCCTCATCTCTGG-3`, STAT6_6796rev 5`-CGACATGGATCCGCTGC-3`) following the protocol displayed in table 24. Using the restriction enzyme BamHI a fragment of 1287bp of this intronic region of the *STAT6* gene was cleaved and inserted into the pGL3-vector system. Due to compatible ends of the restriction enzymes BamHI and BglII it was possible to insert the *STAT6* region of interest upstream or downstream of the luciferase gene in the respective *STAT6* promoter constructs (pBP78, pBP86).

components	**amount**	**concentration**
High Fideltity PCR Buffer	1.5µl	10x
$MgSO_4$	0.6µl	50mM
STAT6_5164fwd	0.12µl	25µM
STAT6_6796rev	0.12µl	25µM
dNTP	0.3µl	10mM
template DNA	1µl	20ng/µl
dH_2O	11.3µl	
Platinum®*Taq* High Fidelity	0.06µl	5U/µl
total volume	**15µl**	

table 23: Components for the amplification of *STAT6* intron 2.

step	**temperature**	**time**
1) initial denaturation	94°C	2min
2) denaturation	94°C	30sec
annealing	54.4°C	45sec
extension	68°C	2min
	→repeat step 1 35 times	
3) final extension	68°C	7min

table 24: PCR conditions for the amplification of *STAT6* intron 2.

As the *STAT6* promoter vectors contained a constitutive restriction site for BamHI, intron 2 was first cloned into the compatible ends of pGL3-promoter vector (5`3` and 3`5`, figure 17, Promega, Madison, USA). To insert the polymorphic T allele of C2892T in intron 2, site-directed mutagenesis (Stratagene, La Jola, USA, 2.2.13.2) was performed with the following primer: 5`-CATGAGTGGTG GTGGGGACAGTCCCTAGGAGGGCTATC-3`. The correctness of the respective insert sequence was verified by re-sequencing (2.2.6) using the primers indicated in table 25.

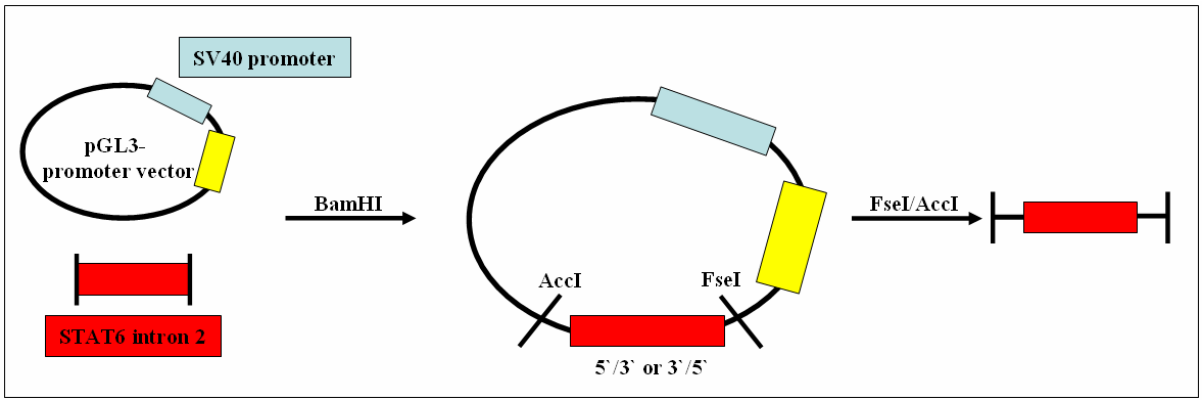

figure 17: *STAT6* intron 2 (1287bp) was cloned into the pGL3-promoter vector downstream of the luciferase gene using the BamHI restriction site. Intron 2 was subsequently isolated with a double digestion (FseI/AccI). Mutagenesis was performed to introduce the polymorphic allele of C2892T.

primer name	primer sequence (5`3`)	STAT6 intron 2: relative location to luciferase gene
pBP78_5610fwd	GTGGTGGAAGGGGGGAGG	upstream
GLprimer2	CTTTATGTTTTTGGCGTCTTCCA	upstream
STAT6_5945fwd	ACCACG CCCTGCTAAT TTTG	upstream and downstream
pGL3_2121fwd	GTTTCAGGTTCAGGGGGAGG	downstream
RVprimer4	GACGATAGTCATGCCCCGCG	downstream

table 25: Primers to verify the correctness of the sequence of *STAT6* intron 2.

After a double-digestion with FseI/AccI the purified *STAT6* intron 2 (either carrying the wildtype C or the polymorphic T allele) was then cloned into the respective *STAT6* promoter constructs (pBP78, pBP86), also digested with FseI/AccI, downstream of the luciferase gene (figure 18).

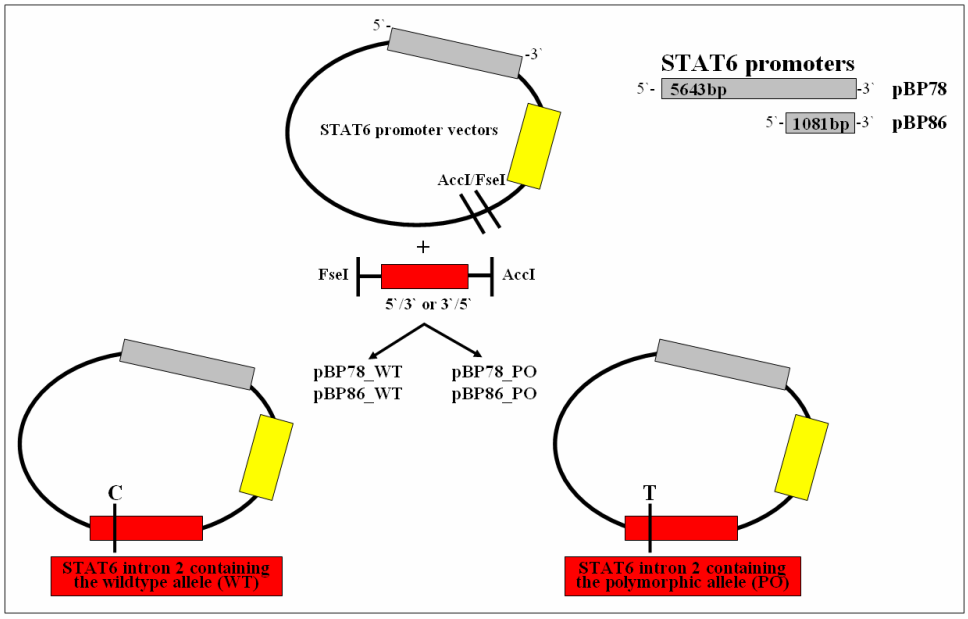

figure 18: *STAT6* intron 2 was double digestion (FseI/AccI) and cloned into the *STAT6* promoter constructs (pBP86 and pBp78) downstream of the luciferase gene in both orientations (5`3`and 3`5).

To investigate a putative *cis*-regulatory function of intron 2, this intronic region was additionally inserted in the *STAT6* promoter construct (pBP78) upstream of the luciferase gene in both orientations (5`3` and 3`5`, figure 19). Hence, intron 2 in the pGL3-promoter was digested with BamHI and the promoter vector pBP78 with BglII.

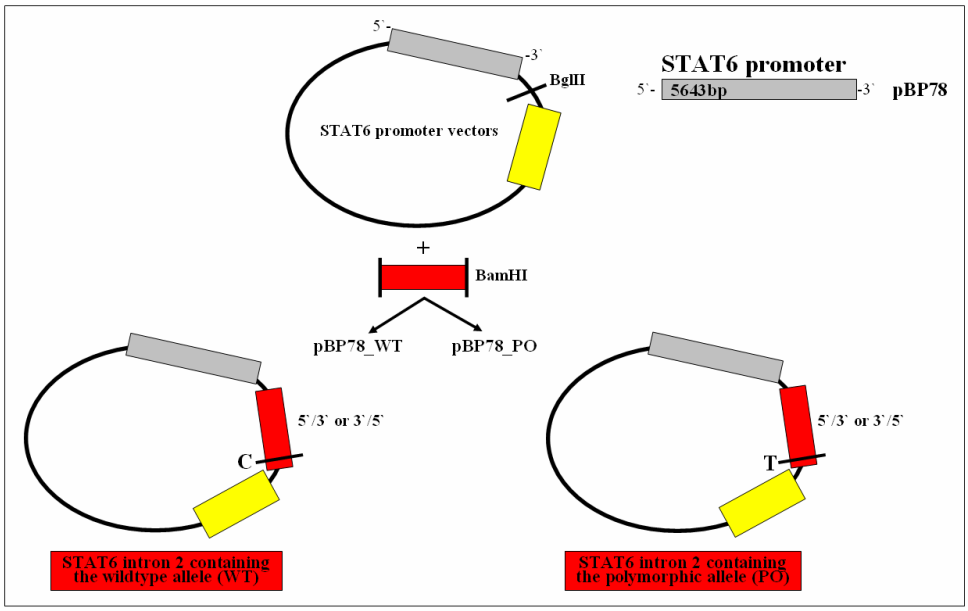

figure 19: *STAT6* intron 2 was additionally cloned into the *STAT6* pBP78 promoter constructs upstream of the luciferase gene to investigate a positional effect of a putative *cis*-regulatory function of intron 2.

2.2.13.2 Site-directed mutagenesis reaction

2.2.13.2.1 Background

Site-directed mutagenesis in general is used to introduce point mutations [148]. In the course of this study this method was used to either replace unwanted single point mutations in the sequence generated by the *Taq* polymerase in the process of cloning or to specifically introduce the polymorphic allele of a genetic variation under study (Quik Change® II Site-directed Mutagenesis Kit, Stratagene, La Jolla, USA).

The QuickChange® II site directed mutagenesis method is performed using *PfuUltra™* high-fidelity DNA polymerase for mutagenic primer-directed replication of both plasmid strands. The basic procedure utilizes a supercoil double-stranded DNA vector with an insert of interest and synthetic primers containing the desired mutation. The mutagenic primers are annealed to denaturated template DNA and extended during temperature cycling by the polymerase without primer displacement. Therefore, double-stranded DNA molecules are generated with one strand bearing the inserted mutations and containing nicks. These nicks are subsequently sealed by components in the enzyme blend. In a next step, the thermal cycling products are treated with the endonuclease DpnI. The DpnI endonuclease is specific for methylated and hemi-methylated DNA and is used to digest parental (bacterial) DNA template and to select for mutation-containing newly synthesized DNA. Hence, only the nicked vector DNA carrying the desired mutation is then transformed into ultra-competent cells, grown to increase the yield. The generated vectors putatively carrying the desired point mutation are re-sequenced to verify the insertion (2.2.6).

For the mutagenesis reaction specific primers are required. These primers have to anneal to the same strand of the template plasmid with a length of 25 – 45bp. To obtain optimal results, primers are designed with melting-temperatures above 75°C and the desired point mutation is located close to the middle of the respective primer (10 - 15bp of template-complementary sequence on both sides). A GC-content of more than 40% is required with a termination of one or more C or G at 3`end for ideal binding of the primers to the DNA strand. If multiple primers are applied for a simultaneous mutagenesis reaction, each primer has to be added in equimolar amounts.

2.2.13.2.2 Experimental design

The designed primers (Metabion, Martinsried, Germany) containing the point mutation were diluted to 1µg/µl in dH_2O whereas for each site-directed mutagenesis reaction the following pipetting scheme was used (table 26). The thermal cycling reaction was conducted comparable to a standard PCR reaction (denaturation, annealing and elongation). Though, the elongation had to be adjusted to the size of each plasmid (2kb/min, table 27). As all template vectors had a size larger than 5kb, this protocol was applied for the respective insertion of a point mutation.

template size	**<5kb**	**>5kb**
10x QuickChange Multi reaction buffer	2,5µl	2,5µl
Quick Solution	---	0,75µl
ds-DNA template	xµl (50ng)	xµl (100ng)
Mutagenesis primers: If using 1-3 primers, add 100ng of each primer If using 4-5 primers, add 50ng of each primer	xµl	xµl
dNTP mix	1µl	1µl
Qick change Multi enzyme blend	1µl	1µl
dH_2O	add to 25µl	add to 25µl

table 26: Standard mastermix for the site-directed mutagenesis reaction.

step	**temperature**	**time**
1) initial Denaturation	95°C	1min
2) Denaturation	95°C	1min
Annealing	55°C	1min
Elongation	65°C	2min/kb of plasmid length
	→repeat step 2 30times	

table 27: Standard program for the site-directed mutagenesis reaction.

Following the incoperation of the point mutation, the digestion of the parental methylated and hemi-methylated DNA (= non-mutated) was performed with 1µl DpnI restriction enzyme (10U/µl) (incubation at 37°C for 1h). To conduct the transformation, 45µl of XL10-Gold ultra-competent cells for each mutagenesis reaction was aliquoted in pre-chilled 15ml polypropylene tubes (Falcon 2059, Falcon, Bedford, USA) and 2µl of β-mercaptoethanol were added. After 10min incubation on ice, 3µl of DpnI treated DNA was applied to the pre-treated cells and incubated for an additional 30min on ice. Cells were then heat-pulsed for 30sec at 42°C and put on ice for 2min. Subsequently, 0.5ml of pre-heated NZY^+ media was added to each reaction. After 1h at 37°C on a shaker, each transformation reaction was plated in appropriate volumes on agarose plates containing ampicillin (50mg/ml). The obtained colonies were grown to increase the yield (Mini-prep Kit, Qiagen, Germany, following the manufacturer`s protocol) and the plasmid DNA was re-sequenced (2.2.6) to verify the insertion of the respective point mutation.

2.2.13.3 Transient transfection experiments using the Jurkat T-cell line

2.2.13.3.1 Background

In transfection experiments, foreign DNA is introduced into eukaryotic cells. Transfection typically involves opening transient gates in cells to allow for the entry of extra-cellular molecules. Cells that have been manipulated to accept foreign DNA are called "competent cells". Transient transfection is usually used to test how various genetic modifications affect the function of particular genes. There are various methods of introducing foreign DNA into a cell. Transfections can be achieved by treating the cells with calcium phosphate, electroporation, heat shock, and proprietary transfection reagents.

For most applications of transfection, it is sufficient that the transfected gene is only transiently expressed. Since the DNA introduced in the transfection process is usually not inserted into the nuclear genome, the foreign DNA is lost at the latest when the cells undergo mitosis. In contrast, if it is desired that the transfected gene actually remains in the genome of the cell and its daughter cells, a stable transfection must occur.

2.2.13.3.2 Experimental design

So far, the only *STAT6* promoter study was performed in NIH3T3 mouse fibroblast [120]. For *STAT6* transfection experiments in this study, a human T-cell line was selected as *STAT6* plays a crucial role in the IL-4/IL-13 pathway which is activated in T-cells. To perform reproducible and most efficient transient transfection experiments with the Jurkat T-cell line using the *STAT6* constructs, transfection settings were optimized for the following parameters: number of applied cells for each transfection, amount of co-transfected renilla vector, electroporation settings, length and time-point of application of different stimuli.

The Jurkat T-cells were grown in 50ml RPMI/Glutamax 1640 (Gibco, Carlsbad, USA) supplemented with 10% FCS and penicillin/streptavidin (5µg/ml, Sigma-Aldrich, Steinheim, Germany). Cells were set to 300.000 cells/ml 24h before each transfection experiment. On the day of the transfection, cells were counted and the appropriate number of cells for each transfection ($8x10^6$, $5x10^6$, $3x10^6$) was resuspended in 0.5ml serum free, ice-cold RPMI/Glutamax 1640. Cell numbers for transfection were varried to identify the optimal cell number with the highest transfection efficiency. After an incubation step of 5min on ice, the different amounts (300ng, 100ng, 50ng, 25ng, 12.5ng, 10ng, 0ng) of the co-transfected renilla vector pRL-TK (Promega, Madison, USA) were added to the cells and mixed by gently flicking the tube. pRL-TK is used to track transfection efficiency from the constituative thymidine kinase promoter. Multiple concentrations of co-transfected renilla vector pRL-TK were tested to assess the minimal amount necessary for each experiment. 4µg of reporter plasmid was applied followed by an additional incubation step for 5min on ice. This solution, containing cells as

well as both reporter vectors were transferred into ice-cold cuvettes (0.4cm, Biorad, Hercules, USA) for the electroporation. The transfection itself was performed using the Biorad square wave electroporation setting (Genepulser Xcell, Biorad, Hercules, USA) with one pulse for 50msec testing different voltages (240V, 220V, 200V) to reduce the number of dead cells during electroporation. After cells rested in the cuvettes for 15min on ice, cuvettes were rinsed with 500µl RPMI/Glutamax (10% FCS) and transferred into 6-well plates (Falcon, Bedford, USA) in a total volume of 4ml RPMI/Glutamax supplemented with 10% FCS and penicillin/streptavidin (5µg/ml). The transfected cells were incubated for 24h either unstimulated or stimulated at 37°C (5% CO_2) whereas the tested stimuli, concentrations, time-points and length of stimulation are displayed in table 28.

stimuli	**applied concentrations**	**time-point of stimulation**
PMA/ionomycin	50ng/ml, 1µM 25ng/ml, 0.5µM 12.5ng/ml, 0.25µM	directly after transfection for 24h
PMA/ionomycin	50ng/ml, 1µM 25ng/ml, 0.5µM 12.5ng/ml, 0.25µM	3h after transfection for 18h
IL-4	5ng/ml, 25ng/ml, 100ng/ml	directly after transfection for 24h
PMA/ionomycin/IL-4	50ng/ml, 1µM, 25ng/ml	directly after transfection for 24h

table 28: Induction of the respective constructs for transient transfection experiments using the Jurkat T-cell line were tested for different stimuli, concentrations, time-points and length of stimulation.

2.2.13.4 Quantification of transcriptional activity of the respective STAT6 constructs

2.2.13.4.1 Background

For all described experiments the Dual-Luciferase® Reporter Assay (Promega, Madison, USA) was used as it improves the experimental accuracy because of the simultaneous expression and measurement of two individual reporter enzymes within a single system [147]. The activity of the co-transfected "control" reporter pRL-TK provides an internal control that serves as a baseline response. Normalizing the activity of the reporter of interest to the activity of the internal control minimizes experimental variability caused by differences in cell viability, transfection efficiency, differences in pipetting volumes or cell lysis efficiency.

In the Dual-Luciferase® Reporter Assay, the activities of firefly (*Photinus pyralis* carried by the pGL3-Basic vector) and renilla (*Renilla reniformis* carried by the pRL-TK vector) luciferases is determined sequentially from a single sample. The firefly luciferase reporter is measured first by adding Luciferase Assay Reagent II (LAR II) to generate a stabilized luminescent signal. After quantifying the firefly luminescence, this reaction is quenched, and the renilla luciferase reaction is initiated by adding Stop & Glo® Reagent.

2.2.13.4.2 Experimental design

To measure the transcriptional activity, the transfected cells in RPMI/Glutamax (10% FCS) were transferred into 10ml tubes (Sarstedt, Nümbrecht, Germany) and each well was rinsed with 5ml PBS (Gibco, Carlsbad, USA). After a centrifugation step at 250g for 7min (Eppendorf, Hamburg, Germany), the supernatant was discarded and the pellets were resuspended in 160µl 1x Passive Lysis Buffer (PLB). During 30min incubation at room temperature the tubes were flicked every 5min for better lysis efficiency. Subsequently, 10µl of each cell lysate was transferred into 96-well plates (Greiner Bio-One, Frickenhausen, Germany). The Dual-Luciferase® Reporter Assay activity was measured using a plate-reading luminometer (LUMIstar, BMD Labtech, Offenburg, Germany)

equipped with two reagent injectors according to the following steps under constant shaking: 50µl of LAR II were injected and the firefly luciferase activity was measured after 2sec for a total of 10sec, then the 50µl of the Stop & Glo® Reagent was added and again after 2sec the renilla luciferase activity was detected for 10sec. Both activities were quantified as relative light units (RLU). The relative luciferase activity of each reporter vector of interest was calculated as the quotient of the firefly luciferase activity divided by the renilla luciferase activity.

As an additional control, the protein concentration of each sample was measured (BCA Protein Assay Kit) to be able to evaluate if comparable cell numbers were applied across experiments and to compare transfection efficiency between samples. A serial dilution for the necessary standard curve with BSA (10mg/ml) was pipetted in duplicates following the scheme presented in table 29.

dilution	amount of BSA (10mg/ml)	amount of PLB	final concentration
1	5µl	20µl	2mg/ml
2	4µl	21µl	1.6mg/ml
3	3µl	22µl	1.2mg/ml
4	2µl	23µl	0.8mg/ml
5	1µl	24µl	0.4mg/ml
6	10µl of dilution 9 (=2µl BSA (10mg/ml) + 48µl dH_2O)	15µl	0.16mg/ml
7	10µl of dilution 6	15µl	0.664mg/ml
8	0µl	25µl	2mg/ml

table 29: Dilution series to quantify the protein content after transient transfection experiments. PLB = passive lysis buffer

200µl of working reagent (= 50:1 reagent A:B) was added to 25µl of each dilution (1 – 8). For every sample a 1:2 dilution (10µl of lysate + 15µl of PLB) was prepared. The quantification of the protein was determined after 30min incubation at 37°C in a 96-well plate (Greiner Bio-One, Frickenhausen, Germany) using an FLUOstar Optima (BMD Labtech, Offenburg, Germany).

2.2.13.5 Transcriptional activity of STAT6 promoters and the cis-regulatory influence of STAT6 intron 2 depending on the genotype of C2892T

Jurkat T-cells ($8x10^6$) were transiently transfected in duplicates with 4µg of each *STAT6* reporter construct and co-transfected with 10ng pRL-TK using the Biorad square wave electroporation protocol with one pulse for 50msec at 200V (Biorad, Hercules, USA). Transfected cells were rested for 3h in 6-well plates, followed by stimulation with PMA/Ionomycin (12.5ng/ml, 0.25µM) and harvested after 21h with passive lysis buffer. Firefly and renilla luciferase activity was quantified as described in section 2.2.13.4. Transcriptional activity of each *STAT6* promoter (pBP78, pBP88, pBP86, pBP85, 2.2.13.1) was expressed in fold difference compared to the pGL3-Basic vector. Likewise, the *cis*-regulatory activity of intron 2 on the respective *STAT6* promoter (pBP78_WT, pBP86_WT) was expressed as fold difference compared to the pGL3-Basic vector containing only the particular *STAT6* promoter (pBP78, pBP86) without intron 2. To determine, if the allelic status of SNP C2892T influenced transcriptional activity of *STAT6* intron 2, fold difference of the relative transcriptional activity of the respective constructs carrying either the wildtype or the polymorphic allele was calculated and displayed in Relative Light Units (RLU).

2.2.14 Gene expression studies for the STAT6 gene performing Real-Time PCR

STAT6 mRNA levels were studied in two independent populations (N = 10, N = 347) to assess whether the gene expression *ex vivo* was influenced by the polymorphisms C2892T. Therefore, quantitative Real-time PCR was performed using the iCycer iQ™ Real Time PCR Detection System (Biorad, Hercules, USA,) and the ABI Prism 7700 Sequence Detection System™ (Perkin-Elmer-Applied-Biosystems, Foster City, USA).

2.2.14.1 Isolation of Peripheral Blood Monocytic Cells (PBMC)

Blood was drawn from ten donors each either carrying the homozygous wildtype (N = 5) or the polymorphic (N = 5) allele of SNP C2892T. The isolation of PBMCs from whole blood was performed as described in detail in section 2.2.11.2.

2.2.14.2 RNA isolation

RNA isolation of the PBMCs (2.2.11.2) was performed using the Versagene RNA Purification Kit (Gentra, Minneapolis, USA) for peripheral blood monocytic cells following the manufacturer`s protocol [149]. $5x10^7$ cells (frozen at –70°C) were vortexed for 20sec and resusupended in 600µl Lysis solution supplemented with 6µl Tris(2-caboxyethyl)phosphine (TCEP) reagent. Cells were vortexed at high speed to thoroughly homogenize the samples and lyse the cells (approximately for 2min). The entire lysate was gently pipetted up and down (20times) and transferred onto a Preclear Column. The Column was centrifuged for 1min at 400g. The flow-through, containing the lysate, was then loaded on a Purification Column in a clear tube and centrifuged for 1min at 14,000g. To wash the lysate, the Purification column was again transferred to a clear tube and 400µl of Wash Solution 1 was added. After a centrifugation step for 2min at 14,000g, the flow-trough was discarded followed by an addition centrifugation at 14,000g for 1min to completely remove the Wash Solution 1. In the next step, residues of DNA were degraded with a DNase treatment. Therefore, incubation with 50µl DNase

Solution was performed for 45min followed by two washing steps with 200µl DNase Wash Solution (centrifugation for 1min and 2min at 14,000g). The Purification column was transferred into a new tube and washed twice with 200µl Wash 2 Solution. To elute the RNA, 50µl of Elution Solution was added to the Purification Column in a new tube and centrifuged for 1min at 14,000g. The purified total RNA of each sample was then aliquoted on ice and immediately stored at –80°C. The concentration and purity was determined by photo-spectrometry (2.2.4).

2.2.14.3 Reverse Transcriptase for cDNA synthesis

The Reverse Transcriptase (RT) of total RNA isolated from PBMCs (2.2.11.2) was performed using the ThermoScript™ RT-PCR System from Invitrogen (Carlsbad, USA) to generate single stranded complementary cDNA out of mRNA [150]. Therefore, a denaturation reaction of the RNA was set up containing 1µg of RNA, 1µl Oligo(dT) primers (50µM) and 2µl dNTP Mix (10mM) adjusted with DEPC-treated water to a total volume of 12µl and incubated at 65°C for 5min. 8µl of a Mastermix with 5x cDNA Synthesis Buffer (4µl), 0.1M DTT (1µl), RNaseOUT™ (1µl, 40U/µl), DEPC-treated water (1µl) and ThermoScript RT (15U/µl, 1µl) was added to the solution on ice. The samples were transferred to a thermal cycler (Eppendorf, Hamburg, Germany) incubated for 60min at 55°C. To terminate the reaction, an incubation step was performed at 85°C for 5min. The remaining mRNA was degraded by 1µl of RNase H (37°C for 20min). The retained cDNA was stored at –20°C for further experiments.

2.2.14.4 Real-time PCR using the iCycler iQ™ Real-Time Detection System (Biorad)

2.2.14.4.1 Background

The Real-time PCR (RT-PCR), also called quantitative Real-time PCR (qRT-PCR), is used to simultaneously quantify and amplify a specific part of a given DNA molecule [151]. The procedure follows the general pattern of polymerase chain reaction (2.2.5), but the DNA is quantified after each round of amplification either due to the utilization of fluorescent dyes that intercalate with double-strand DNA (e.g. SYBR Green) or modified DNA oligonucleotide probes that fluoresce when hybridized with a complementary DNA.

The iClycer iQ™ Real Time PCR Detection System is based on a standard laboratory cycler supplied with an optical module which allows for real-time detection of PCR-products. Through the filter based optical design, discrimination between four fluorescent reporters (490nm-635nm) is possible and 96 samples can be scanned simultaneously.

The ability to monitor the real-time progress of the PCR is characterized by detection of fluorescence values during every cycle which represent the amount of product amplified. The more template present at the beginning of a reaction, the fewer number of cycles are needed to reach a point at which a reliable fluorescence signal is detected, which can be discriminated from the background. This point is defined as the C_T value (threshold cycle) which will always occur during the beginning of the exponential phase of amplification. Real-time PCR can therefore be used to measure the relative amount of mRNA present before amplification (indirectly through the usage of cDNA), enabling to quantify relative gene expression of a particular gene and comparing mRNA contents between different samples.

Using the iClycer iQ™ Real Time PCR Detection System, a melting curve is provided which is a powerful tool to confirm the specificity of the PCR product. The melting temperature (T_m) of each product is defined as the temperature at which the melting of double-stranded PCR occurs. The respective system records the total fluorescence generated by the SYBR Green binding to double-stranded DNA as temperature changes, and displays plots of fluorescence in real time as a function of temperature. Non-specific amplification can be observed as it may result in PCR products that melt at temperatures above or below the specific PCR product.

2.2.14.4.2 Experimental design

The *STAT6* gene expression was indirectly assessed through the measurement of cDNA (2.2.14.3) of donors either carrying the homozygous wildtype C (N = 5) or the polymorphic T (N = 5) allele of the SNP C2892T. For these ten donors the overall STAT6 mRNA levels were separately quantified using the forward primer STAT6_292fwd 5`-CCAGGATG-GCTCTCCACAG-3` and the reverse primer STAT6_933rev 5`-CATGGAGGAATCAGGGGC-3`. To specifically amplify the overall *STAT6* mRNA and to discriminate DNA contamination, the forward and the reverse primer were located in different exons: the forward primer was partially located in exon 16 and partially in exon 17 of the *STAT6* gene whereas the reverse primer was located in exon 19. Thus, a differentiation between the PCR product of interest and contamination through residues of DNA in the mRNA was detectable due to the sizes of emerging PCR products: a fragment size of 304bp was expected, whereas an addition PCR product of 957bp would have been observed with a contamination of DNA residues. Additionally, as a positive control, all primers were tested using genomic DNA as template. The melting curve of each PCR products was checked and PCR products were run on a 3% gel to verify the specific amplification of the respective fragment. Each fragment of interest underwent these control steps.

Real-time PCR experiments were performed with the HotStar*Taq*® DNA polymerase kit from Qiagen (Hilden, Germany) as the kit revealed the most reproducible results. This was partially due to the fact that the DNA polymerase used is provided in an inactive state with no polymerase activity at ambient temperatures for more specificity. Therefore, a heat activation step of 15min at 95°C had to be performed previous to the actual amplification reaction.

The standard mastermix with its components for each PCR reaction is shown in table 30. Deviations in the pipetting scheme are displayed in the respective tables for each fragment separately.

components	**amount**	**concentration**
PCR buffer	5.0µl	10x
dNTPs	1.0µl	10mM
forward primer	0.5µl	25pmol/µl
reverse primer	0.5µl	25pmol/µl
cDNA template (1:15 dilution)	6.0µl	1:15 dilution
SYBR Green	1.0µl	1000x
FITC	1.0µl	1750x
HotStar*Taq*® DNA polymerase (5U/µl)	0.17µl	5U/µl
dH_2O	add to 50µl	

table 30: Standard mastermix for the Real-time PCR using the iCycler IQ™ Detection System (Biorad, Hercules, USA).

PCR reaction was carried out in 96-well PCR plates (Biorad, Hercules, USA) in a total volume of 50µl. Each fragment (table 31) was optimized with the following conditions, whereas a mutual annealing temperature of 61.6°C for further experiments was used.

step	temperature	time
1) activation step	95°C	15min
2) denaturation	94°C	30sec
annealing	57°C + 10°C	30sec
extension	72°C	1min
	→repeat step 2 40 times	
3) final extension	72°C	10min
4) melting curve	96°C	1min
	55°C	30sec
	55°C + 0.5°C for 80 cycles	10sec

table 31: Standard program for the Real-time PCR using the iCycler IQ™ Detection System (Biorad, Hercules USA).

Using the primers STAT6_292fwd and STAT6_933rev for quantification of the overall *STAT6* mRNA, more peaks than expected were observed in the melting curve. When the PCR product was applied to a 3% agarose gel, bands of different sizes were detected in addition to the expected band of 304bp. Each band was eluted from the gel using the QIAquick® Gel Extraction Kit (Qiagen, Hilden, Germany) following the manufacturer`s protocol. The fragments were re-sequenced as described in section 2.2.6. More specific primers (as displayed in figure 20 and table 32) were designed to be able to quantify *STAT6* mRNA levels of different isoforms for each sample separately. Thus, overall STAT6 mRNA contains all isoforms except STATd and STAT6e. In contrast STAT6d and STAT6e account for the novel isoforms.

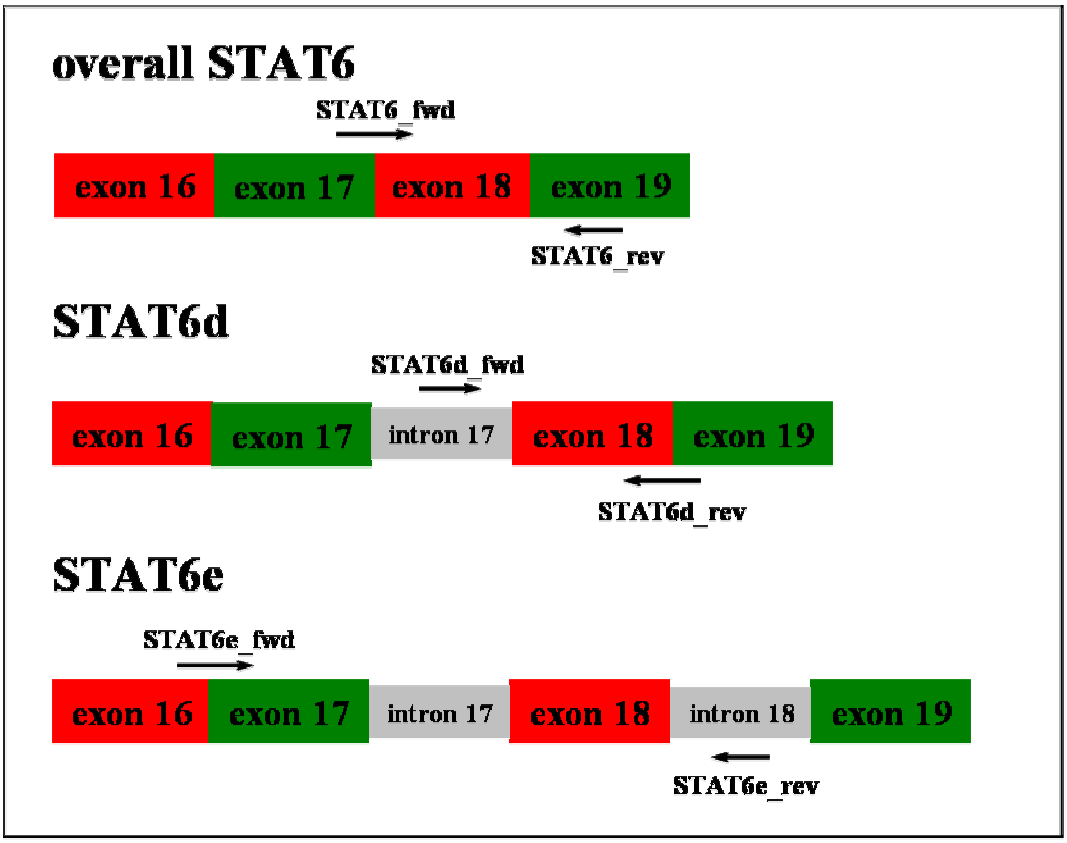

figure 20: Description of the observed *STAT6* splice variants: Overall STAT6 mRNA contains all isoforms except STATd and STAT6e. STAT6d and STAT6e are two novel identified splice variants. STAT6d consists of the displayed exons and intron 17 whereas STAT6e contains additionally to the exons intron 17 and 18. The localization of the respective primers for the Real-time PCR is indicated.

	primer name	primer sequence (5`3`)	localisation	fragment size	annotation
1	STAT6_fwd	AGCCACTCACAAGC-CTGAAC	exon 17-exon18	162bp	0.3µl of each primer
	STAT6_rev	CATGGAGGAATCAGGGGC	exon 19		
2	STAT6d_fwd	CTTCCTGTCACCCACTTCCTC	intron17	264bp	0.3µl of each primer
	STAT6d_rev	AGTGGTTGGTCC-CTTTCCAC	exon 18-exon19		
3	STAT6e_fwd	GGGGCCAGGATG-GCTCTC	exon 16-exon17	479bp	
	STAT6e_rev	TAAGCCCCTGACCTACCCACTG	intron 18		

table 32: Primer sequences, primer localization and size of the amplified fragment to quantify three *STAT6* isoforms (overall STAT6, STAT6d, STAT6e) using the iCycler IQ™ Detection System (Biorad, Hercules, USA). Annotations indicate deviations from the standard mastermix displayed in table 30.

2.2.14.4.3 Identification of the most robust housekeeping gene

A so called "housekeeping gene" was simultaneously analyzed which serves in general as an internal reference for normalization. This is a necessary step to control for variation in the amount of starting material and to exclude errors through misinterpretation of the expression profiles of the gene of interest. The ideal internal control should be expressed at a constant level in all individuals, should not be affected through stimulation and should be expressed at comparable levels as the gene under study (comparable C_T-values) [152]. Different housekeeping genes (*GAPDH, β-actin, 18SrRNA, PEX19*) were tested to identify the most specific and robust settings for the respective experimental design (table 33).

Performing Real-Time quantitative PCR with the iCycler Detection System (Biorad, Hercules, USA), the genes *GAPDH*, *β-actin* and *18SrRNA* showed comparable C_T-values among different samples ranging from 17.7 to 20.2 whereas *PEX19* reached the exponential phase not until the 30th cycle. Due to the so far unknown C_T-value of the *STAT6* gene, all four genes were further investigated.

Analyzing the melting curve of each PCR product indicated that each housekeeping gene showed only a single peak which supported the specificity of each fragment under study. After the PCR products were separated on a 3% agarose gel, the specific amplification of *18SrRNA* and *PEX19* was confirmed as only a single band with the expected size was detectable. In contrast, for the genes *GAPDH* and *β-actin* double bands were observed. In general, the primers were designed to allow for the identification of residues of genomic DNA because of the expected size of the emerging PCR product. Thus, it was excluded that the additional bands were due to DNA residues as these fragment would have been larger. Comparable results for these two genes were detectable applying different primer combinations and different PCR conditions. Therefore, *GAPDH* and *β-actin* were excluded for further analyses as these additional bands may emerge because of splice variants of the respective gene and may therefore influence the gene expression results.

Finally, 18SrRNA (primer combination 4) was used as the endogenous control for the quantification of the mRNA expression, as *STAT6* was expressed at equal levels indicated by comparable C_T-values.

	primer name	primer sequence (5`3`)	localization	fragment size	annotation
1	GAPDH_1455fwd	TGGGGAAGGTGAAGGTCG	exon 1	234bp	0.2M betain
	GAPDH_3783rev	GGAGACGACTAGGGGGT	exon 5		
2	GAPDH*	GAAGGTGAAGGTCGGAGTC	exon 1	226bp	0.2M betain
	GAPDH*	GAAATCCCATCACCATCTTC	exon 3		
3	β-actin_3144fwd	GTGGGCATGGGTCAGAAGG	exon 3	251bp	
	β-actin_3845rev	GGCTGGGGTGTTGAAGGTC	exon 4		
4	18SrRNAfwd*	AGTCCCTGCCCTTTGTACACA		69bp	1M betain
	18SrRNArev*	GATCCGAGGGCCTCACTAAAC			
5	18SrRNA_1446fwd	AGAGGGACAAGTGGCGTTCAG		222bp	0.5M betian
	18SrRNA_1666rev	GCTTATGACCCGCACTTACTGG			
6	PEX19fwd	TCCAAGGATGTGCTGTACC	exon 5	128bp	
	PEX19rev	TTTGCACATGACGCTGTGC	exon 6		

table 33: Primer sequences, primer localization and size of the amplified fragment for the housekeeping genes using the iCycler IQ™ Detection System (Biorad, Hercules, USA). Annotations indicate deviations from the standard mastermix displayed in table 30.

2.2.14.5 Real-time PCR using the TaqMan® ABI Prism 7700 (Perkin-Elmer- Applied-Biosystems)

In a colaborative project, the gene expression of *STAT6* isoforms (overall STAT6, STAT6d, STAT6e) was also quantified in a large study population (N = 324, 2.2.1.2.2) according to our specifications at the Department of Immunology of the Children`s Hospital Zurich (Switzerland), headed by PD Dr. Roger Lauener. These alternative methods of mRNA quantification will be described briefly in the following summary.

2.2.14.5.1 Background

Using *Taq*Man® chemistry, the ABI Prism 7700 Sequence Detection System™ (Perkin-Elmer-Applied-Biosystems, Foster City, USA) scaned 96 samples simultaneously detecting fluorescence between 500nm and 660nm. In contrast to the iCycler System, no melting curve is produced [153,154].

Unlike the iCycler System, the *Taq*Man® product from Applied-Biosystems is optimized for the use of a fluorogenic 5' exonuclease assay. Besides of the regular standard primers for the amplification of a fragment, a so called “probe” has to be applied which is designed to anneal specifically between the forward and reverse primer sites of the target sequence. In total, three primers have to bind to the region of interest to generate a signal which ensures a high specificity. The probe contains a fluorescent molecule (FAM) at the 5`end, called reporter, and at the 3`end a second fluorescent molecule (TAMRA), called quencher. The quencher emits an absorption spectrum equivalent to the emission spectrum of the reporter. The quencher therefore suppresses the signal of the reporter. If a template bearing the target sequence is present, the probe anneals to it. During PCR, the nuclease activity of the *Taq* polymerase cleaves the reporter dye from the probe. The reporter dye, now separated from the quencher, emits a fluorescent signal which is monitored at every cycle as additional reporter dye molecules accumulate. After the signal rises above background, the rate of increase is tracked during a number of linear cycles before the reaction reaches a plateau. These data are then used to calculate initial template levels.

2.2.14.5.2 Experimental design

RNA samples were collected from 324 individuals to analyze gene expression. Total RNA was isolated with the QIAmp® RNA Blood Mini Kit (Qiagen, Hilden, Germany) supplemented with RNase-free DNase (Qiagen, Hilden, Germany). Reverse transcription of RNA was performed with 300ng of total RNA in a final volume of 30µl with reagents form Applied Biosystems (Foster City, USA). Therefore, 50U MuLV Reverse Transcriptase, 6µl 25mM $MgCl_2$ solution, 3µl 10x RT-PCR buffer II, 3µl dNTP each 2.5mM, 15U RNase inhibitor, and 1.5µl of a 50µM random hexamers solution were used. The cDNA was generated in a MWG Primus^{TM} Thermal cycler with the following program: 10min at 25°C, 40min at 42°C, and 5min at 95°C.

Quantitative Real-Time PCR (*Taq*Man®) was performed using an ABI Prism 7700 Sequence Detection System™ (Applied Biosystems, Foster City, USA). The primers and probes for the three STAT6 isoforms and 18SrRNA (endogenous control) were designed with the primer design software Primer Express™ (Applied Biosystem, Foster City, USA) and purchased from Microsynth (Balgach, Switzerland, table 34).

	primer name	primer sequence (5`3`)	localisation	fragment size
1	STAT6_fwd	ATCCCAAGAAGCCCAAGGA	exon 17	81bp
	STAT6_rev	GGACATAACCCCTGCCATCC	exon 18	
	STAT6_probe	ACTACAAGC-CTGAACAGA	exon 17-exon 18	
2	STAT6d_fwd	CCAGAGCACTCCATGGCTGT	intron 17	162bp
	STAT6d_rev	GGTCATAAGAAGGCACCATGGT	exon 19	
	STAT6d_probe	TGGAAAG-GGACCAACCA	exon 18-exon 19	
3	STAT6e_fwd	TCACATGTACCTCCTTCCCTCC	intron 18	148bp
	STAT6e_rev	CACCTGGGGC-ACCATATCTG	exon 19-exon 20	
	STAT6e_probe	ATGAGCATGCAGCTTG	exon 19	

table 34: Primer sequences, primer localization and the size of the amplified fragment to quantify the *STAT6* isoforms (overall STAT6, STAT6d, STAT6e) using ABI Prism 7700 Sequence Detection System™ (Applied Biosystems, Foster City, USA).

Quantitative real-time PCR was performed with 3µl of the reverse transcription solution in a final volume of 25µl. The adequate amount of primers and probe was added, together with 12.5µl *Taq*Man Universal PCR Master Mix™ (Applied Biosystems, Foster City, USA). The gene expression values were normalized to the parallel measured endogenous control 18SrRNA. Data with the Comparative (ΔΔCt) method were analyzed according to the manufacturer's instructions (Applied Biosystem, Foster City, USA). Statistical analyses were performed with SAS 9.1.3 (The SAS Insitute, Cary, USA). Odds Ratios (OR) and 95% Confidence Intervals (95%CI) were calculated in a multiple logistic regression analysis.

3 Results

3.1 Association results for polymorphisms in the *STAT6* gene

STAT6 is a key regulator in the IL-4/IL-13 pathway as it is activated after binding of the cytokines IL-4 and IL-13 to their shared receptor IL-4Rα. This signal leads to the homoderimization of STAT6 molecules and their translocation into the nucleus, where STAT6 subsequently binds to crucial target genes and induces class-switching from IgM to IgE [42,43]. Therefore, it was hypothesized that STAT6 may influence the regulation of IgE and the development of asthma. To address this question an association study between *STAT6* polymorphisms and their putative influence on atopy-related phenotypes was conducted.

To perform the association analyses of the *STAT6* gene six polymorphisms in a large German population of 1,120 children (9 – 11 years) from Munich (N = 528) and Leipzig (N = 592) were genotyped. The analyzed *STAT6* SNPs were selected due to their location representing all SNP linkage clusters in the gene. Therefore, two SNPs located in the 5`UTR, one in intron 2, one in intron 18 and two in the 3`UTR were genotyped (figure 21). So far, no SNPs in exonic regions of the *STAT6* gene have been described in the public SNP databases (http://snpper.chip.org) in Caucasian populations with a minor allele frequency (MAF) ≥ 0.3.

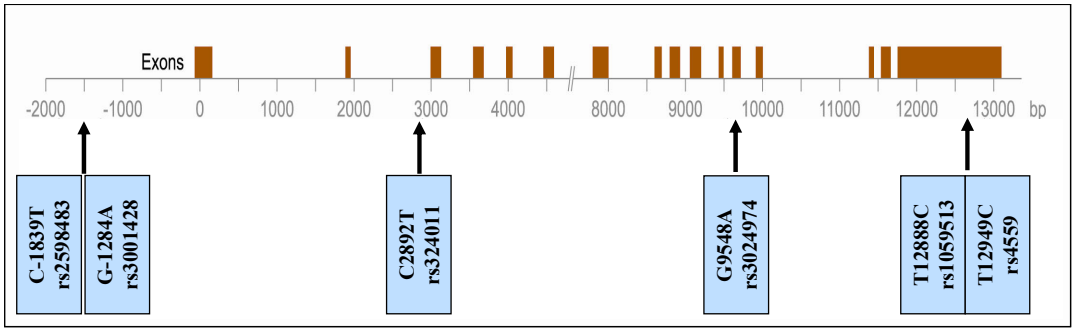

figure 21: Gene structure of *STAT6* indicating the location of all six genotyped SNPs with their relative position to the ATG and their respective rs number according to dbSNP (www.ncbi.nlm.nih.gov/SNP).

In our study population of 1,120 subjects, between 967 and 1,042 samples were genotyped successfully for the respective *STAT6* polymorphisms (figure 21, table 35) with the high-throughput method MALDI-TOF MS (see section 2.2.8). For all six polymorphisms a description of the SNPs, allele frequencies as well as genotyping details are given in table 35. The allele distribution observed in our Caucasian population was consistent with allele frequencies observed in previous studies [124]. As a quality control for the applied genotyping method, Hardy-Weinberg equilibrium was calculated for each of the analyzed SNPs. None of the genotyped polymorphisms showed a significant deviation.

SNP	rs number	location	alleles	minor allele frequency	successful genotyping call rate (%)
C-1839T	rs2598483	5`flanking	C/T	0.10	91.34
G-1284A	rs3001428	5`flanking	G/A	0.05	90.88
C2892T	rs324011	intron 2	C/T	0.37	89.23
G9548A	rs3024974	intron 18	G/A	0.10	92.27
T12888C	rs1059513	3`UTR	T/C	0.12	95.95
T12949C	rs4559	3`UTR	T/C	0.37	95.76

table 35: Description of genotyped *STAT6* polymorphisms with their relative position to the ATG, rs number according to dbSNP (www.ncbi.nlm.nih.gov/SNP), localisation in the gene, alleles, minor allele frequency and genotyping call rates.

3.1.1 *STAT6* polymorphisms are associated with total serum IgE levels

3.1.1.1 Single SNP analyses

Performing single SNP analyses with the six polymorphisms genotyped, one polymorphism in intron 2 (C2892T) and one in the 3`UTR (T12888C) showed significant associations with elevated total serum IgE levels in the pooled study population. For C2892T, geometric means for total serum IgE levels (IU/ml) were 45.1 in CC wildtype carriers versus 57.0 in CT and TT genotypes ($p = 0.014$, table 36). For T12888C, geometric means for total serum IgE levels (IU/ml) revealed values of 56.9 for wildtype

versus 41.5 in carriers of the C allele (p = 0.003), respectively. In the subset from Munich (N = 528) and from Leipzig (N = 592) both of our populations the same trends for association were observed separately as well as in the total study population (table 36).

SNP	genotype	geometric mean of total serum IgE (95% CI)		
		Munich	Leipzig	pooled sample
C-1839T (rs2598483)	**CC**	52.2 (44.9-61.4)	53.5 (46.9-61.0)	53.0 (47.9-58.7)
	CT+TT	43.4 (32.1-58.6)	53.1 (49.1-64.0)	47.4 (38.5-58.3)
	p-value	p=0.280	p=0.880	p=0.346
G-1284A (rs3001428)	**GG**	50.6 (437-58.5)	56.1 (49.1-64.0)	53.3 (48.3-58.8)
	GA+AA	53.3 (34.6-82.2)	39.9 (29.9-53.4)	45.1 (35.3-57.7)
	p-value	p=0.826	p=0.071	p=0.261
C2892T (rs324011)	**CC**	42.8 (34.9-53.2)	47.1 (39.3-56.5)	**45.1 (39.3-51.9)**
	CT+TT	56.5 (46.5-68.1)	57.6 (49.1-67.4)	**57.0 (50.5-64.4)**
	p-value	p=0.064	p=0.107	**p=0.014**
G9548A (rs3024974)	**GG**	48.8 (42.0-56-6)	52.0 (45.5-59.5)	50.4 (45.6-55.7)
	GA+GG	60.0 (42.6-84.5)	63.7 (49.1-82.6)	62.1 (50.5-76.4)
	p-value	p=0.261	p=0.172	p=0.073
T12888C (rs1059513)	**TT**	54.2 (46.6-63.1)	**59.5 (52.1-68.0)**	**56.9 (51.5-62.9)**
	TC+CC	44.8 (33.4-60.1)	**38.6 (30.8-48.4)**	**41.5 (34.5-49.8)**
	p-value	p=0.247	**p=0.002**	**p=0.003**
T12949C (rs4559)	**TT**	49.2 (39.7-61.0)	63.1 (51.7-77.1)	55.6 (48.1-64.4)
	TC+CC	53.5 (45.0-63.6)	50.1 (43.4-57.9)	57.7 (46.2-57.8)
	p-value	p=0.549	p=0.062	p=0.425

table 36: Association between the geometric mean of total IgE (95% confidence interval) and the genotyped *STAT6* polymorphisms in Munich (N = 528), Leipzig (N = 592) and the pooled population (N = 1,120) using the dominant model. Significant values are indicated in bold letters.

3.1.1.2 Haplotype analyses

Additionally, haplotype analyses were performed combining the six genotyped polymorphisms to assess whether particular haplotypes of the *STAT6* gene were associated with total serum IgE levels. For this analysis, 930 samples were available in which genotyping was successful in all of the *STAT6* polymorphisms. In these analyses, it was demonstrated that seven out of 32 possible haplotypes exceeded a frequency of 3% (figure 22). As continuous variables (such as serum IgE levels) cannot be applied for haplotype analyses these values were transformed into binary variables. IgE percentiles (50^{th}, 66^{th} and 90^{th}) instead of the widely used arbitrary cut-off of 100IU/ml were chosen as they may much better resemble the distribution of total IgE levels in our population. The 50^{th}, 66^{th} and 90^{th}

percentiles were calculated correlating with modestly, moderately and highly elevated serum IgE levels which derived from a representative cross-sectional German population of children 9 – 11 years of age (N = 4,400).

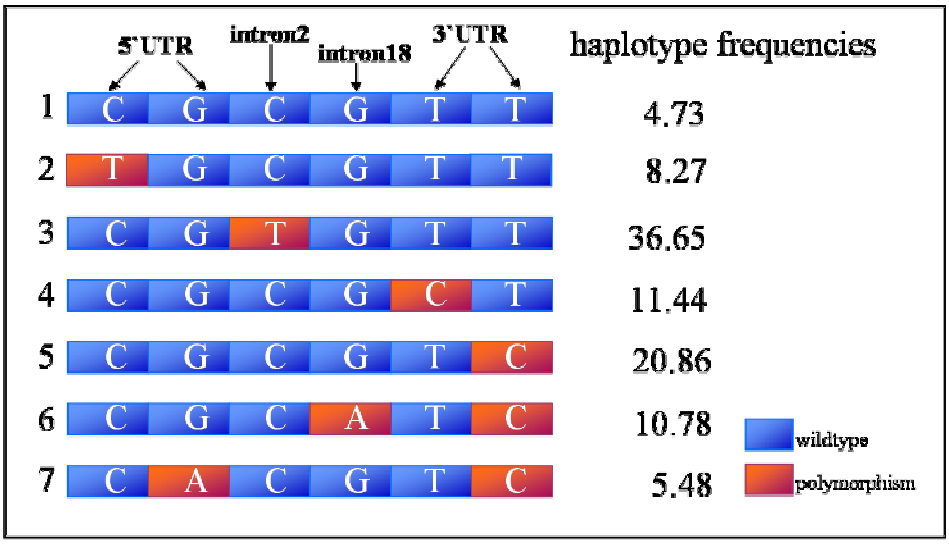

figure 22: *STAT6* haplotypes and haplotype frequencies (only haplotypes with a frequency > 3% were considered).

As shown in table 37, one high risk haplotype CGTGTT (haplotype 3) with increasing Odds Ratio (OR) of 1.58 (95% confidence interval CI 1.08 - 2.32, p = 0.020), 1.82 (95% CI 1.19 - 2.77, p = 0.006) and 3.92 (95% CI 1.93 - 7.96, p = 0.0002) for IgE percentiles 50th, 66th and 90th were observed.

In contrast, the haplotypic combination of CGCGCT (haplotype 4) revealed a borderline protective effect against the development of elevated total IgE levels with OR = 0.59 (95% CI 0.32 - 1.07, p = 0.084), 0.51 (95% CI 0.25 - 1.02, p = 0.056) and 0.27 (95% CI 0.07 - 1.10, p = 0.067) for the 50th, 66th and 90th IgE percentile. These haplotype effects were mainly driven by two single polymorphisms located in intron 2 (C2892T) and the 3`UTR (T12888C). Hence, these data indicated that specific SNP combinations seem to have an influence on IgE regulation in our German study population of children.

haplotype	Munich		
	50th percentile	66th percentile	90th percentile
1 (CGCGTT)	1.41 (0.37-5.5!) p=0.612	2.09 (0.50-8.68) p=0.310	0.82 (0.07-9.71) p=0.874
2 (TGCGTT)	**0.36 (0.14-0.94) p=0.036**	0.41 (0.14-1.22) p=0.109	0.43 (0.07-2.66) p=0.365
3 (CGTGTT)	1.33 (0.75-2.35) p=0.325	1.76 (0.94-3.26) p=0.077	**5.28 (1.96-14.2) p=0.001**
4 (CGCGCT)	0.91 (0.39-2.12) p=0.824	0.56 (0.21-1.49) p=0.243	0.40 (0.07-2.26) p=0.301
5 (CGCGTC)	0.90 (0.47-1.73) p=0.743	0.58 (0.27-0.12) p=0.152	0.22 (0.05-0.92) p=0.038
6 (CGCATC)	2.09 (0.88-4.96) p=0.094	2.14 (0.87-5.29) p=0.100	1.13 (0.25-4.97) p=0.877
7 (CACGTC)	1.10(0.30-3.97) p=0.887	1.09 (0.27-4.46) p=0.907	0.62 (0.05-7.18) p=0.702
haplotype	**Leipzig**		
	50th percentile	**66th percentile**	**90th percentile**
1 (CGCGTT)	2.21 (0.69-7.09) p=0.182	2.18 (0.65-7.30) p=0.206	not defined
2 (TGCGTT)	0.93 (0.34-2.57) p=0.895	0.85 (0.27-2.66) p=0.990	1.51 (0.23-10.11) p=0.670
3 (CGTGTT)	**1.80 (1.07-3.05) p=0.028**	**1.86 (1.05-3.31) p=0.034**	**2.95 (1.06-8.27) p=0.039**
4 (CGCGCT)	**0.38 (0.16-0.90) p=0.028**	0.46 (0.17-1.24) p=0.123	0.141 (0.01-1.57) p=0.112
5 (CGCGTC)	0.79 (0.41-1.51) p=0.473	0.83 (0.40-1.72) p=0.624	0.95 (0.25-3.55) p=0.940
6 (CGCATC)	1.18 (0.53-2.63) p=0.682	1.01 (0.41-2.45) p=0.990	1.92 (0.44-8.35) p=0.382
7 (CACGTC)	0.39 (0.12-1.20) p=101	0.21 (0.05-0.92) p=0.039	0.059 (0.001-3.32) p=0.166
haplotype	**pooled sample**		
	50th percentile	**66th percentile**	**90th percentile**
1 (CGCGTT)	1.86 (0.77-4.48) p=0.168	2.15 (0.85-5.40) p=0.104	0.19 (0.02-1.99) p=0.165
2 (TGCGTT)	0.54 (0.28-1.07) p=0.079	0.57 (0.26-1.25) p=0.160	0.77 (0.21-2.85) p=0.700
3 (CGTGTT)	**1.58 (1.08-2.32) p=0.020**	**1.82 (1.19-2.77) p=0.006**	**3.92 (1.93-7.96) p=0.0002**
4 (CGCGCT)	0.59 (0.32-1.07) p=0.084	0.51 (0.25-1.02) p=0.056	0.27 (0.07-1.10) p=0.067
5 (CGCGTC)	0.83 (0.52-1.32) p=0.427	0.69 (0.41-1.17) p=0.168	0.46 (0.17-1.20) p=0.110
6 (CGCATC)	1.55 (0.86-2.79) p=0.142	1.44 (0.77-2.72) p=0.256	1.42 (0.50-4.02) p=0.510
7 (CACGTC)	0.62 (0.26-1.44) p=0.264	0.46 (0.17-1.25) p=0.129	0.24 (0.03-1.9) p=0.177

table 37: Odds Ratio (95% confidence intervals) for IgE percentile (50th, 66th and 90th) of *STAT6* haplotypes in Munich (N = 528), Leipzig (N = 592) and the pooled population (N = 1,120). Significant values are indicated in bold letters.

3.1.2 Association between *STAT6* SNPs and other atopic phenotypes

In contrast to the described effects of *STAT6* SNPs on IgE regulation, no constant statistically significant association was observed between single SNPs or haplotypes in the *STAT6* gene and the prevalence of asthma, hay fever, specific IgE or atopy (measured by skin prick test reactivity ≥ 3mm) in our study population as presented in table 38 and table 39.

Therefore, it can be speculated that polymorphisms in the *STAT6* gene alone seem to be most relevant in IgE regulation. Less significant effects were observed for the manifestation of asthma and other atopic phenotypes when *STAT6* SNPs were analyzed *per se*.

SNP	Munich		
	asthma	hay fever	atopy
C-1839T (rs2598483)	1.19 (0.53-2.30) p=0.799	1.42 (0.67-3.03) p=0.336	1.33 (0.80-2.21) p=0.265
G-1284A (rs3001428)	1.99 (0.56-3.46) p=0.483	1.09 (0.37-3.23) p=0.874	1.24 (0.63-2.43) p=0.542
C2892T (rs324011)	1.13 (0.60-2.11) p=0.709	0.98 (0.50-1.93) p=0.950	1.18 (0.76-1.84) p=0.476
G9548A (rs3024974)	1.46 (0.71-2.99) p=0.299	1.33 (0.61-2.89) p=0.477	1.63 (0.96-2.77) p=0.067
T12888C (rs1059513)	0.58 (0.27-1.27) p=0.167	0.89 (0.41-1.91) p=0.762	**0.56 (0.33-0.96) p=0.034**
T12949C (rs4559)	1.05 (0.59-1.86) p=0.875	1.02 (0.54-1.92) p=0.948	1.04 (0.69-1.58) p=0.847
SNP	**Leipzig**		
	asthma	hay fever	atopy
C-1839T (rs2598483)	0.70 (0.24-2.05) p=0.510	0.45 (0.10-1.94) p=0.271	0.71 (0.41-1.24) p=0.227
G-1284A (rs3001428)	0.71 (0.21-2.40) p=0.581	0.94 (0.92-0.97) p=0.056	0.62 (0.32-1.17) p=0.138
C2892T (rs324011)	1.16 (0.56-2.43) p=0.688	1.47 (0.62-3.47) p=0.380	0.98 (0.66-1.45) p=0.918
G9548A (rs3024974)	1.01 (0.43-2.38) p=0.986	1.75 (0.74-4.14) p=0.198	1.31 (0.83-2.06) p=0.250
T12888C (rs1059513)	0.56 (0.21-1.49) p=0.240	0.45 (0.13-1.51) p=0.182	0.93 (5.59-1.47) p=0.757
T12949C (rs4559)	1.09 (0.54-2.20) p=0.812	0.60 (0.27-1.32) p=0.197	0.86 (0.59-1.27) p=0.461
SNP	**pooled sample**		
	asthma	hay fever	atopy
C-1839T (rs2598483)	0.96 (0.53-1.76) p=0.901	1.07 (0.56-2.04) p=0.851	0.99 (0.67-1.41) p=0.896
G-1284A (rs3001428)	1.02 (0.50-2.11) p=0.953	0.53 (0.19-1.48) p=0.216	0.848 (0.53-1.35) p=0.487
C2892T (rs324011)	1.16 (0.72-1.86) p=0.554	1.16 (0.68-1.97) p=0.587	1.06 (0.79-1.15) p=0.716
G9548A (rs3024974)	1.21 (0.70-2.10) p=0.492	1.14 (0.81-2.56) p=0.208	**1.45 (1.03-2.04) p=0.032**
T12888C (rs1059513)	0.58 (0.31-1.06) p=0.072	0.71 (0.37-1.35) p=0.291	0.74 (0.53-1.05) p=0.094
T12949C (rs4559)	1.04 (0.67-1.63) p=0.853	0.81 (0.50-1.33) p=0.408	0.95 (0.72-1.26) p=0.724

table 38: Odds Ratio (95% confidence intervals) for asthma, hay fever and atopy for single *STAT6* SNPs in Munich (N = 528), Leipzig (N = 592) and the pooled population (N = 1,120). Significant values are indicated in bold letters.

haplotype	Munich		
	asthma	hay fever	atopy
1 (CGCGTT)	0.58 (0.05-6.79) p=0.666	0.37 (0.02-6.98) p=0.503	2.09 (0.47-9.26) p=0.332
2 (TGCGTT)	0.87 (0.20-3.79) p=0.848	1.63 (0.40-6.72) p=0.500	1.54 (0.58-4.09) p=0.389
3 (CGTGTT)	1.25 (0.50-3.14) p=0.638	0.91 (0.33-2.51) p=0.853	1.30 (0.67-2.53) p=0.433
4 (CGCGCT)	0.40 (0.08-2.02) p=0.270	0.85 (0.18-3.94) p=0.838	0.35 (0.11-1.09) p=0.069
5 (CGCGTC)	0.64 (0.21-7.98) p=0.437	0.86 (0.27-2.82) p=0.809	0.47 (0.21-1.07) p=0.071
6 (CGCATC)	1.70 (0.46-6.23) p=0.426	1.47 (0.63-6.09) p=0.595	2.06 (0.80-5.30) p=0.136
7 (CACGTC)	2.43 (0.38-15.50) p=0.349	1.38 (0.16-12.28) p=0.77	1.43 (0.33-6.11) p=0.632
haplotype	**Leipzig**		
	asthma	hay fever	atopy
1 (CGCGTT)	2.38 (0.30-18.91) p=0.414	3.31 (0.41-26.84) p=0.262	1.79 (0.53-6.08) p=0.352
2 (TGCGTT)	not defined	0.08 (0.001-4.26) p=0.209	0.40 (0.11-1.40) p=0.151
3 (CGTGTT)	1.15 (0.38-3.46) p=0.802	2.02 (0.64-6.32) p=0.229	1.09 (0.61-1.95) p=0.765
4 (CGCGCT)	0.37 (0.04-3.09) p=0.357	0.24 (0.02-2.66) p=0.244	1.29 (0.52-3.19) p=0.581
5 (CGCGTC)	0.96 (0.24-3.80) p=0.954	0.60 (0.13-2.86) p=0.519	0.96 (0.47-1.96) p=0.902
6 (CGCATC)	1.38 (0.27-6.98) p=0.698	3.39 (0.76-15.21) p=0.110	1.59(0.67-3.75) p=0.294
7 (CACGTC)	0.86 (0.08-9.70) p=0.902	not defined	0.35 (0.09-1.39) p=0.136
haplotype	**pooled sample**		
	asthma	hay fever	atopy
1 (CGCGTT)	1.14 (0.23-5.60) p=0.870	1.23 (0.22-6.78) p=0.815	1.96 (0.76-5.04) p=0.163
2 (TGCGTT)	not defined	0.97 (0.26-3.59) p=0.959	0.84 (0.39-1.80) p=0.644
3 (CGTGTT)	1.19 (0.59-2.40) p=0.628	1.26 (0.60-2.67) p=0.544	1.18 (0.77-1.82) p=0.453
4 (CGCGCT)	0.40 (0.11-1.43) p=0.156	0.59 (0.15-2.02) p=0.373	0.74 (0.37-1.48) p=0.393
5 (CGCGTC)	0.76 (0.32-1.83) p=0.544	0.76 (0.30-1.95) p=0.572	0.69 (0.04-1.18) p=0.171
6 (CGCATC)	1.50 (0.55-4.14) p=0.432	2.06 (0.74-5.73) p=0.169	1.82 (0.96-3.44) p=0.065
7 (CACGTC)	1.48 (0.35-6.31) p=0.595	0.33 (0.04-2.55) p=0.285	0.67 (0.25-1.79) p=0.422

table 39: Odds ratios (95% confidence intervals) for asthma, hay fever and atopy for *STAT6* haplotypes in Munich (N = 528), Leipzig (N = 592) and the pooled population (N = 1,120).Significant values are indicated in bold letters.

3.2 Gene-by-gene interaction of polymorphisms in the IL-4/IL-13 pathway

As *STAT6* does not act alone but is an integrative part of the IL-4/IL-13 pathway it was hypothesized that genetic variations in *STAT6* may have more profound effects when analyzed in the context of other SNPs in the pathway genes *IL-4*, *IL-13* and *IL-4Rα.* Therefore, gene-by-gene interactions between the four molecules *IL-4*, *IL-13*, *IL-4Rα* and *STAT6* were studied using an extended haplotype analyses in the population of Munich and Leipzig (2.2.1.1).

Out of the originally 18 genotyped polymorphisms in *STAT6, IL-4*, *IL-13* and *IL-4Rα* of the IL-4/IL-13 pathway only those were selected (table 40) for the interaction analyses that revealed a MAF of ≥ 0.3 in both populations [90,93,117]. Additionally, the selection was based on positive associations in single gene studies and putative functional relevance of the respective SNPs.

Thus, the polymorphism located in the intronic region 2 of the *STAT6* gene was selected as *in silico* analyses gave evidence that the SNP C2892T was situated in a putative NF-κB site which additionally seemed to be highly conserved among primate species (3.3.2). This *STAT6* variant was analyzed in combination with the polymorphism C-589T in the proximal promoter region of *IL-4*, as it may alter the binding of a transcription factor (NFAT) [91]. Furthermore, the promoter polymorphism C-1112T in the *IL-13* gene was selected, as it changes binding of transcription factors and influences the expression of *IL-13* in T-cells [103]. In the *IL-4Rα* gene, SNP A148G was included in the analyses, as it leads to an amino acid change in the extra-cellular part of the receptor (I50V) which induces an increased downstream activation of *STAT6* and increased IgE production in B-cells [118]. All selected polymorphisms had previously been studied for their effect on atopic phenotypes in single gene analyses in our study populations [90,93,117].

Genotyping was successful in 92.6 - 98.7% of all samples and complete information on all four polymorphisms was available in 856 subjects. Assay description, allele frequencies as well as genotyping details are given in table 40.

According to its biological function, the effects of *STAT6* C2892T in combination with other SNPs in the IL-4/IL-13 pathway on two main outcomes were studied: the regulation of serum IgE levels and the development of childhood asthma.

gene (position from ORF)	rs number	alleles	minor allele frequency	call rate (%)
IL-4 (-589)	rs2243250	C/T	0.32	97.59
IL-13 (-1112)	rs1800925	C/T	0.46	92.58
IL-4Rα (148)	rs1805010	A/G	0.43	97.50
STAT6 (2892)	rs324011	C/T	0.37	98.68

table 40: Description of genotyped polymorphisms of the IL-4/IL-13 pathway: their relative position to the ATG, respective rs number according to dbSNP (www.ncbi.nlm.nih.gov/SNP), alleles, minor allele frequency and genotyping call rates.

3.2.1 IL-4/IL-13 pathway SNPs influence IgE regulation

First, the effects of single SNPs on total IgE levels were tested independently of the effects of polymorphisms in any of the other genes in the 856 samples available for gene-by-gene analysis. Thus, the isolated effects of single alleles of the respective SNP on the development of highly elevated serum IgE (90^{th} percentile) are shown in the upper panel of figure 23. As described in detail in section 3.1.1 *STAT6* revealed the maximum risk out of the four analyzed single SNPs with an OR of 1.73 (95% CI 1.21-2.47) for elevated total IgE levels.

In a second step, pairs of SNPs in the IL-4/IL-13 pathway were formed. The presence of both polymorphic alleles of *STAT6*_C2892T and *IL-13*_C-1112T resulted in a significant interaction term (OR = 2.92, 95% CI 0.91 - 9.30). The combination of the polymorphic alleles of *IL-13*_C-1112T/*IL-4*_C-589T and *IL-4Rα*_A148G/*IL-4*_C-589T reached statistical significance with an increasing OR of 6.69 (95% CI 1.60 - 28.01) and 8.71 (95% CI 2.47 - 30.70) (figure 23, middle panel).

Finally, the introduction of triplet carriers of the respective polymorphic alleles of *STAT6*_C2892T in combination with the genes *IL-4Rα*_A148G and *IL-4*_C-589T increased the risk for elevated IgE levels reaching an OR of 13.86 (95% CI 1.65-107.98) (figure 23, lower panel).

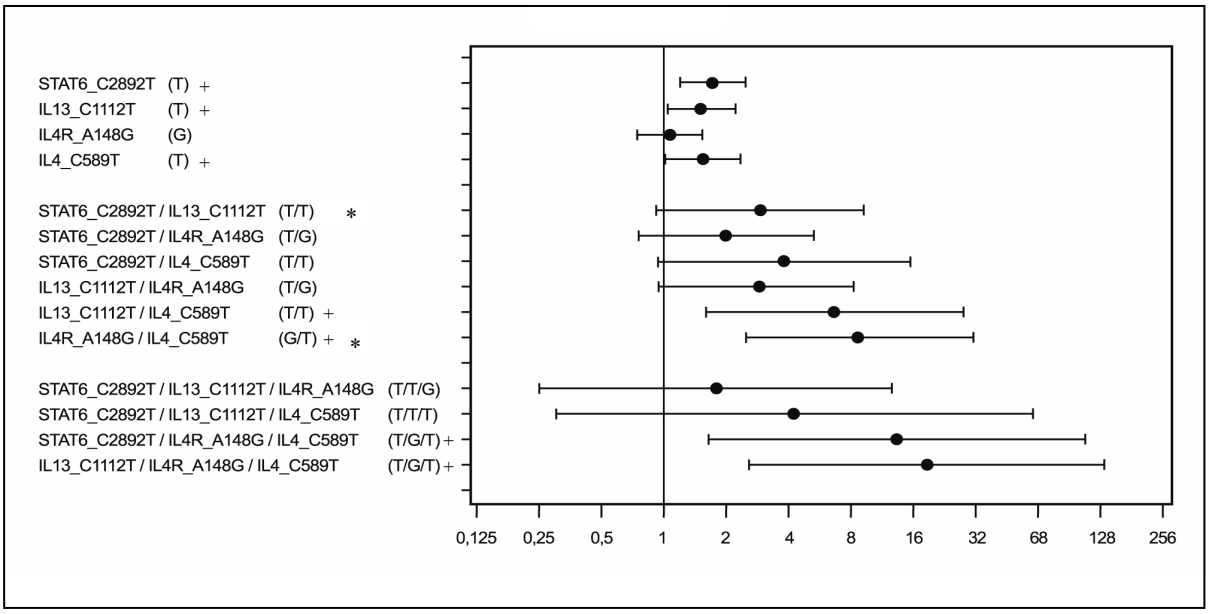

figure 23: Combination of the polymorphic alleles of IL-4/IL-13 pathway SNPs increase the risk to develop elevated IgE levels (90th percentile): in the upper panel, effects of single SNPs are shown. In the middle panel, haplotype pairs, and in the lower panel, haplotype triplets are presented. Significant associations (marked +) and haplotypes from genotype combinations showing statistically significant epistatic interactions (marked *).

3.2.2 IL-4/IL-13 pathway SNPs affect the risk to develop asthma

In a second analysis, genetic effects on asthma were studied according to the same stepwise procedure (figure 24). As described in section 3.1.2, the *STAT6* SNP C2892T does not seem to be sufficient to alter asthma risk by itself (figure 24, upper panel). In contrast, the OR for the development of asthma increased to a maximum of 4.55 (95% CI 1.59-12.97) in carriers of both polymorphic alleles of the SNPs C2892T in *STAT6* and C-1112T in *IL-13* for the interaction (figure 24, middle panel) and again to 26.81 (95% CI 3.05-236.62) when SNP *IL-4* C-589T was added to the model (figure 24, lower panel). Additionally, other triplet combination of *STAT6* C2892T with polymorphisms from the IL-4/IL-13 pathway reached borderline significance in relation to asthma development (figure 24, lower panel).

As indicated by increasingly large confidence intervals, the informative proportion of the study population decreased with the stepwise addition of SNPs (2.5 - 4.2% carry three risk alleles) and the analysis of four SNP interactions could not be performed reliably due to increasingly low numbers.

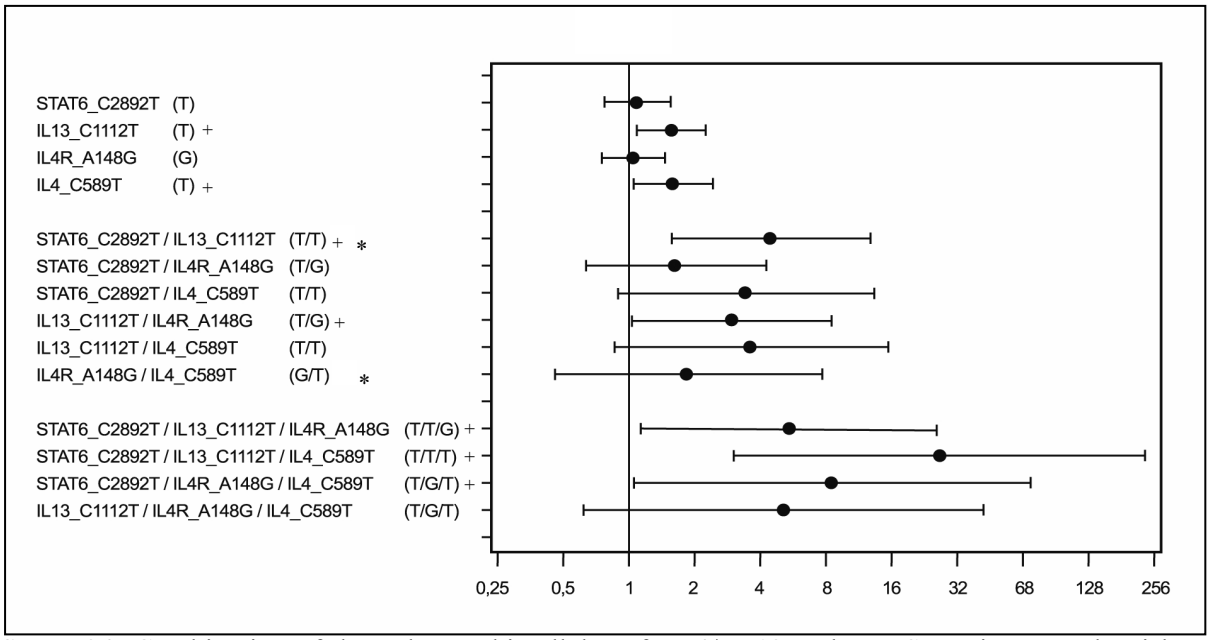

figure 24: Combination of the polymorphic alleles of IL-4/IL-13 pathway SNPs increase the risk to develop asthma: in the upper panel, effects of single SNPs are shown. In the middle panel, haplotype pairs, and in the lower panel, haplotype triplets are presented. Significant associations (marked +) and haplotypes from genotype combinations showing statistically significant epistatic interactions (marked *).

3.3 *In silico* analyses of intron 2 harboring C2892T

As SNP C2892T in intron 2 of the *STAT6* gene was associated with elevated total serum IgE levels and the development of asthma in single gene and gene-by-gene interaction analyses, further experiments were performed to determine the functional role of this intronic region in general and the effect of polymorphism C2892T on the *STAT6* gene.

3.3.1 Comparative genomics of the human, dog and mouse genome

So far, no information was available regarding whether this intronic region of the *STAT6* gene exerted any regulatory effect on STAT6 function. In a first step, the degree of conservation between the human, dog and mouse sequence of intron 2 was assessed. The approach of comparative genomics is based on the observation, that the sequence concordance across related species is especially high in regions of functional relevance, such as promoters, regulatory or coding regions, since these regions are actively conserved throughout the evolution. The data shown in figure 25 indicates that the polymorphism C2892T (high-lighted with a red error) is located within a region of almost 75% sequence concordance with the dog. In contrast, the mouse genome with a separation of approximately 75 million years from their last common ancestor revealed little sequence similarity (figure 25).

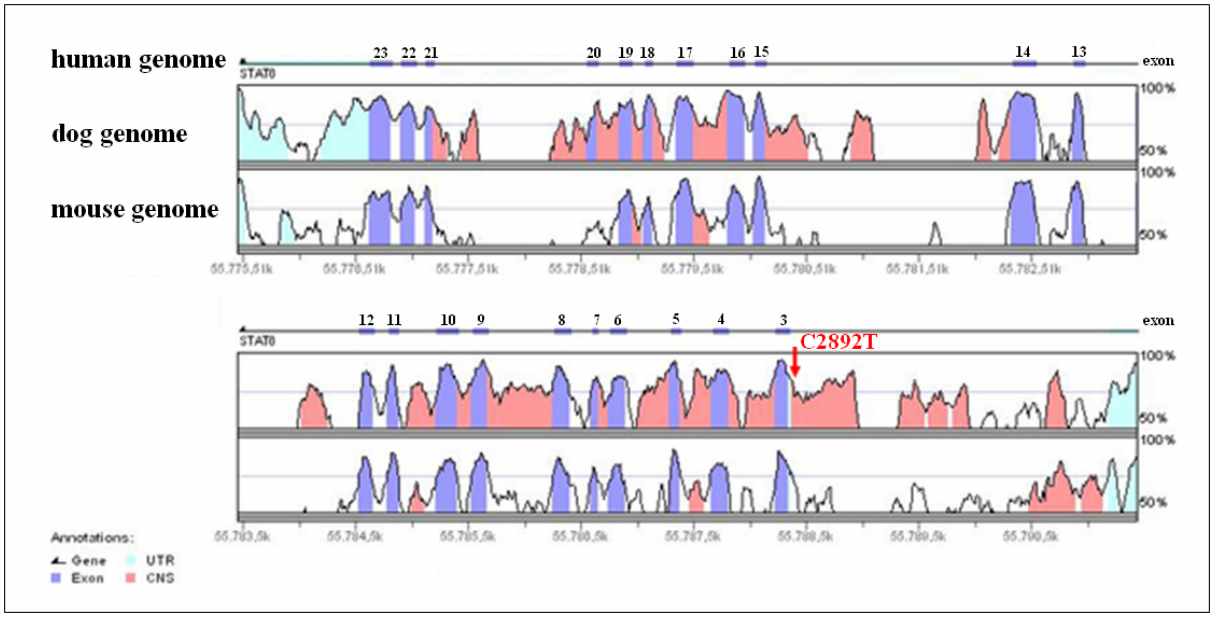

figure 25: Conservation of *STAT6* between the human, dog and mouse genome using Vista Browser (http://pipeline.lbl.gov/cgi-bin/gateway2). Exons are color-coded in purple, untranslated regions (UTR) in turquoise and conserved non-coding regions (CNS) in pink. SNP C2892T located in intron 2 of the *STAT6* gene is high-lighted with a red error.

3.3.2 Phylogenetic shadowing of the human sequence with different primate species

Comparative genomics with distant species will identify regions within the gene of interest with rather ancient biological function. However, changes that occurred more recently in primate species potentially representing the species specific gain of an evolutionary advantage will inevitably be missed. It was suggested that the comparison between the human genome and the respective sequence of a group of four to six primates may be enough to identify novel regulatory elements [142]. Especially intronic and intergenetic regions seem to have gained importance in highly evolved species. Therefore, this second approach was used to analyze the region of intron 2 of the *STAT6* gene (2.2.10.2) between humans and primate species comparing chimpanzee, gorilla, orang-utan, old world (gibbon, green monkey) and new world (owl monkey) monkeys with a shorter evolutionary dinstance from man (2.2.10.2, figure 14).

Using the newly developed eShadow analysis tool (2.2.10.2) a high degree of sequence similarity of 85% - 95% of this intronic region of *STAT6* harboring SNP C2892T was detected between humans and the other species analyzed (figure 26 and figure 27). This data further supported the hypothesis of a putative regulatory function of this intronic region of the *STAT6* gene.

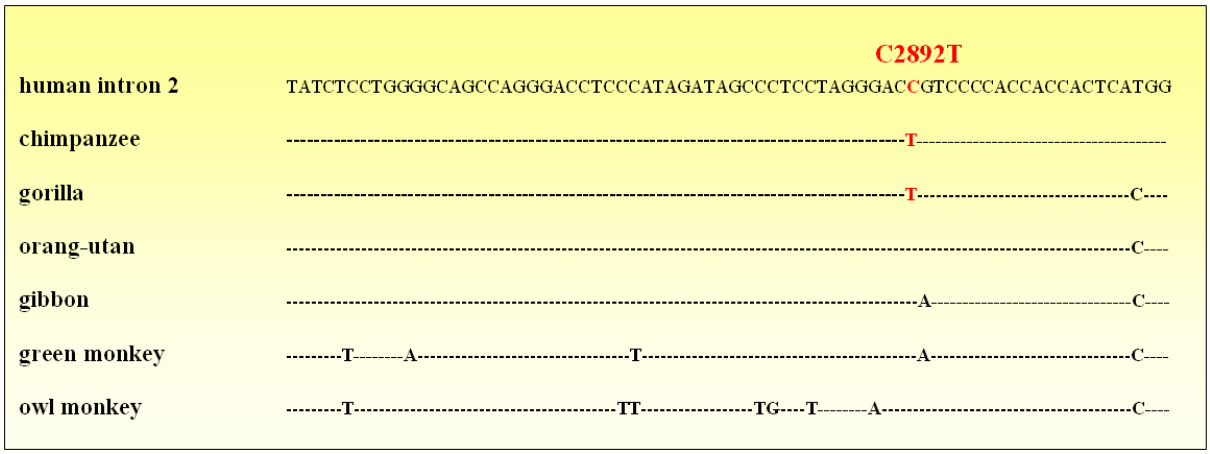

figure 26: Alignment of the human sequence of *STAT6* intron 2 containing the polymorphism C2892T to different primate species. Dashes indicate conserved positions.

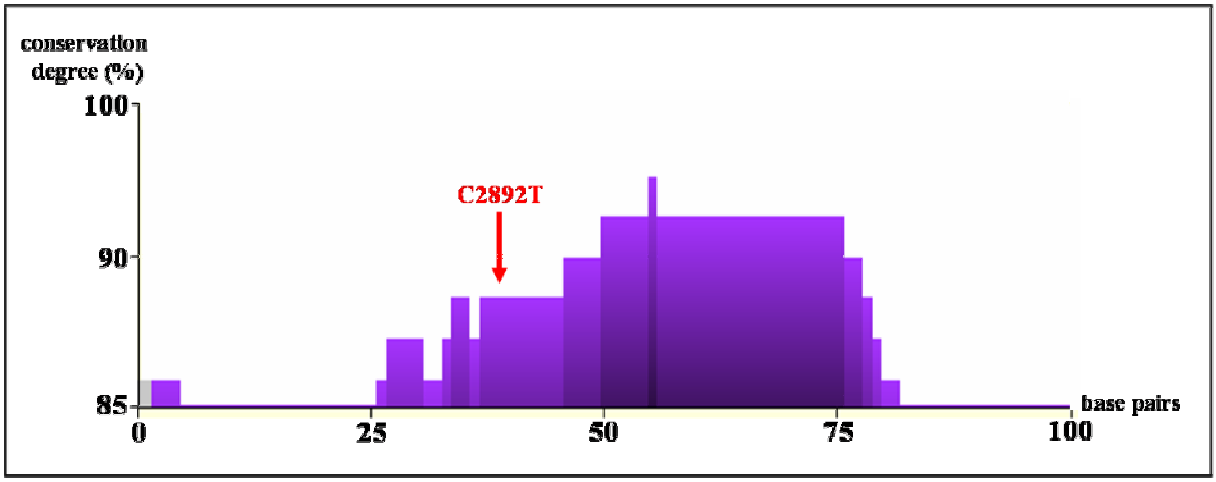

figure 27: Conservation of *STAT6* between the human and primate species containing chimpanzee, gorilla, orang-utan, old world (gibbon, green monkey) and new world (owl monkey) monkeys using eShadow (http://eshadow.dcode.org, 100bp window setting). SNP C2892T located in intron 2 of the *STAT6* gene is high-lighted with a red error.

3.3.3 *In silico* analysis for transcription factor binding to intron 2

As the region of intron 2 containing C2892T showed a high degree of conservation between primates, the functional relevance of this region on *STAT6* gene regulation was investigated in more detail. Thus, *in silico* analyses were performed to assess putative transcription factor binding sites in intron 2. Using AliBaba 2.1 a putative transcription factor binding site for NF-κB was predicted to occur only when the polymorphic T allele was present at position 2892. In contrast, Matinspector 7.0 predicted the binding of the transcription factor NF-κB to the complementary strand of *STAT6* intron 2 in the presence of the T and the C allele of SNP C2892T. However, the matrix similarity deviated depending on the genotype of the SNP: a score of 0.921 for the polymorphic T allele was detected whereas the wildtype C allele reached a matrix similarity of 0.888 indicating a stronger binding affinity of NF-κB to the polymorphic allele.

Additionally, a second putative binding site for NF-κB was suggested (applying both softwares) in close vicinity to the polymorphism within a distance of only 21bp. The binding of the transcription factor SP1 (Specificity protein), a member of the SP family (SP1-SP4), within this region was also postulated but its binding was indicated to be independent of the genotype of C2892T. In addition, a GATA binding site, which is a transcription factor relevant for the differentiation of T helper cells, may be located close to the SNP. All results are presented in figure 28.

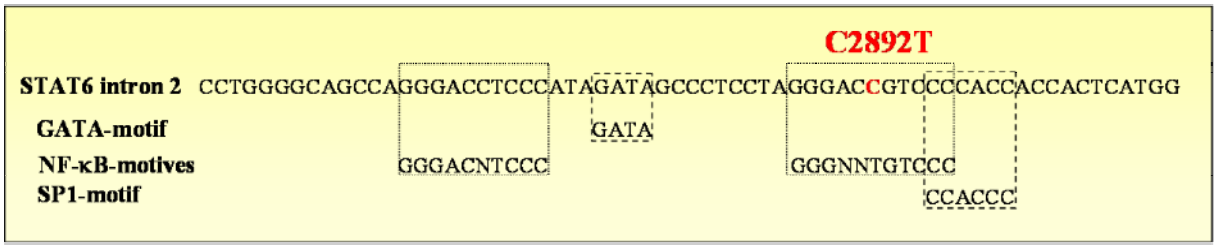

figure 28: Presentation of putative transcription factor binding sites in close vicinity to the polymorphism C2892T (marked in red) of *STAT6* intron 2 using the software programs Matinspector and AliBaba. The binding of the transcription factor NF-κB was postulated to bind only in the presence of the polymorphic allele of C2892T.

3.4 Transcription factor binding in intron 2 of the *STAT6* gene is influenced by the polymorphism C2892T

So far, our association studies demonstrated that the polymorphism C2892T in intron 2 of the *STAT6* gene plays a crucial role in mediating IgE regulation and asthma susceptibility in the IL-4/IL-13 pathway. *In silico* analyses supported the hypothesis of a putative functional role of intron 2 harboring SNP C2892T due to its high sequence concordance across species and the prediction of various regulatory elements within this region. *In silico* transcription factor binding analyses suggested that two transcription factors, NF-κB and SP1, may bind to the intron 2 region of interest in the *STAT6* gene. Their binding may be altered by the SNP C2892T. Thus, it was hypothesized that changes in the binding properties of these transcription factors may be associated with the observed effects on the population level. To test the relevance of this hypothesis, EMSA (Electrophoretic Mobility Shift Assay) experiments were performed with nuclear extract from human T-cells to examine putative DNA/protein binding of these transcription factors at the respective *STAT6* site of intron 2. A human T-cell line (Jurkat T-cell line) and primary T-cells (naïve $CD4^+$ T-cells) were chosen for these experiments as the activation of the IL-4/IL-13 pathway by *STAT6* takes place in T-cells.

In a first step it was the aim to prove that NF-κB and SP1 are capable to bind to the respective intron 2 region in T-cells *per se*. In a second step, the effect of SNP C2892T on these DNA/protein interactions was analyzed.

3.4.1 The role of NF-κB binding in the Jurkat T-cell line

The Jurkat T-cell line was selected for these experiments as STAT6 is an important T-cell transcription factor. Also, the IL-4/IL-13 pathway is activated in T-cells. Before studying NF-κB binding to the region of interest in intron 2 of the *STAT6* gene, it first had to be verified that NF-κB was present in the selected experimental system of the Jurkat T-cell line. Using a consensus-site for NF-κB [146], EMSA analyses with nuclear extract from Jurkat T-cells were tested for the presence of NF-κB and to optimize experimental conditions [103,155,156]. The best experimental conditions were achieved using the buffer conditions previously described by Agresti *et al*. [156]. Additionally, the Poly(dI-dC)-Poly(dI-dC) concentration was titrated ranging from 0.125µg up to 5.0µg/reaction: the most specific band pattern was observed with 1.0µg/reaction Poly(dI-dC)-Poly(dI-dC). Subsequently, the presence of NF-κB in Jurkat T-cells was established and was sufficiently proven.

For EMSA experiments with the NF-κB probe (26bp) nuclear extract from Jurkat T-cells either from unstimulated or stimulated (3h with PMA/Ionomycin) Jurkat T-cells was used. We observed three complexes, one of which was increased after stimulation (figure 29, lanes 1 and 2). These complexes were specific since they could be competed with a 100-fold molar excess of cold NF-κB probe (figure 29, lane 3). Competition experiments with the respective *STAT6* probes corresponding to the C or the T allele of C2892T provided first evidence that there may in fact be a difference in their binding affinity: 2892T was able to compete binding better than the wildtype C allele (figure 29, lanes 4 and 5). To identify specific NF-κB complexes supershift experiments were performed through pre-incubation with 4µg antibody of the NF-κB subunits p50 and p65 (Santa Cruz, California, USA). The p50 antibody induced the formation of a slower migrating complex whereas complex 2 could not be completely abrogated (figure 29, lane 6). In contrast, the p65 antibody led to a selective disruption of the NF-κB complex 2 (figure 29, lane 7). Hence, it was concluded that the NF-κB consensus sequence obtained a stronger affinity to the p65 subunit than to p50.

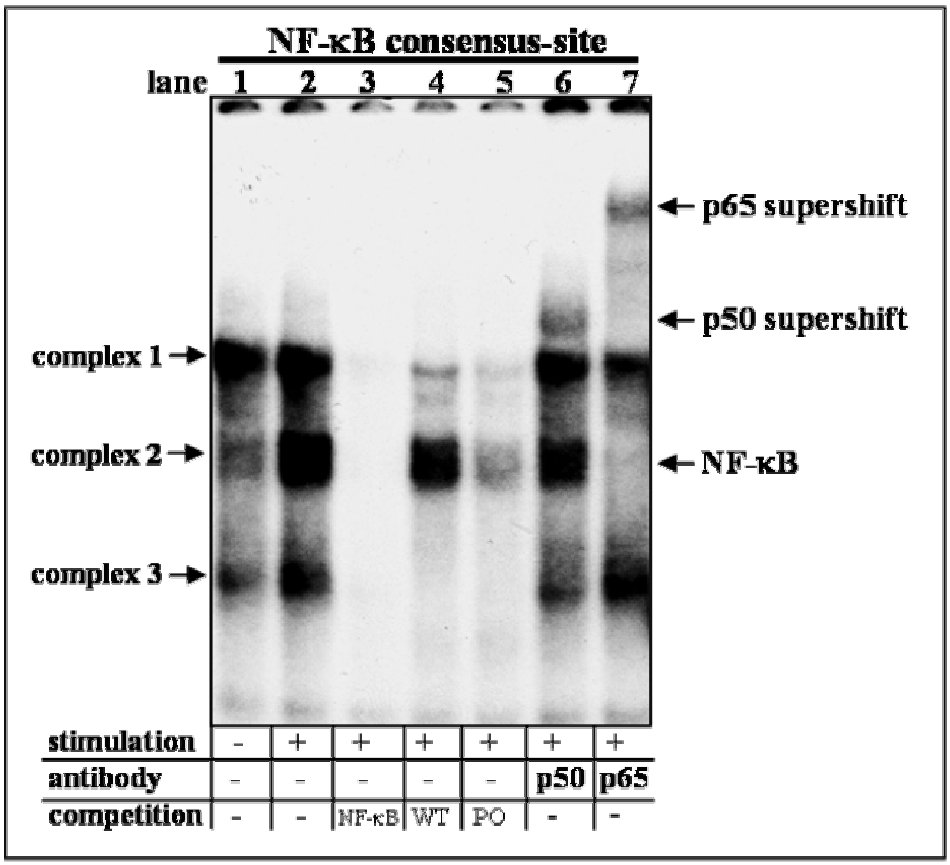

figure 29: EMSA analysis with a NF-κB consensus sequence using nuclear extract (5µg) from the Jurkat T-cell line cultered for 3h either left in medium or stimulated with PMA/Ionomycin (50ng/ml, 1µM). The competitors (100-fold molar excess) and supershift antibodies (4µg) used for each experiment are noted below the respective lanes.

3.4.2 The role of SP binding in Jurkat T-cells

As it was also postulated that the transcription factor SP1 may bind to the intronic *STAT6* region, this hypothesis was assessed in further experiments. Thus, it was tested if we would be capable to detect SP transcription factors in our EMSA settings using a 22bp SP consensus sequence [157] and nuclear extract from our Jurkat T-cell line (figure 30).

Incubation of the labeled probe with nuclear extract from the Jurkat T-cell line resulted in the formation of five specific complexes, all of which were disrupted by the addition of 100-fold molar excess of unlabeled SP consensus probe (figure 30, lane 2). Additionally, competition experiments with 100-fold molar excess of cold *STAT6* probes (31bp) either carrying the C or the T allele for SNP

C2892T were performed. Again it was indicated that the affinity of the two allelic variants was different (figure 30, lanes 3 and 4). In supershift experiments is was shown that both complexes 1 and 2 were a high-affinity site for SP1 as the antibody induced a strong supershift signal, the disruption of complex 1 and a much weaker signal of complex 2 (figure 30, lane 5). The remaining complex (figure 30, lane 5) was identified with a SP4 specific antibody because pre-incubation of antibodies SP1 and SP4 together completely abrogated this signal (figure 30, lane 11). Furthermore, the antibody SP3 led to a supershift detectable through a slower migrating complex and the loss of complexes 3, 4 and 5 (figure 30, lane 7, 10, 11). Using the antibodies SP1, SP3 and SP4 together in one reaction, all observed complexes were completely disrupted (figure 30, lane 13). Only the transcription factor SP2 did not show any binding to the SP consensus sequence using Jurkat T-cell nuclear extract (figure 30, lanes 6 and 9).

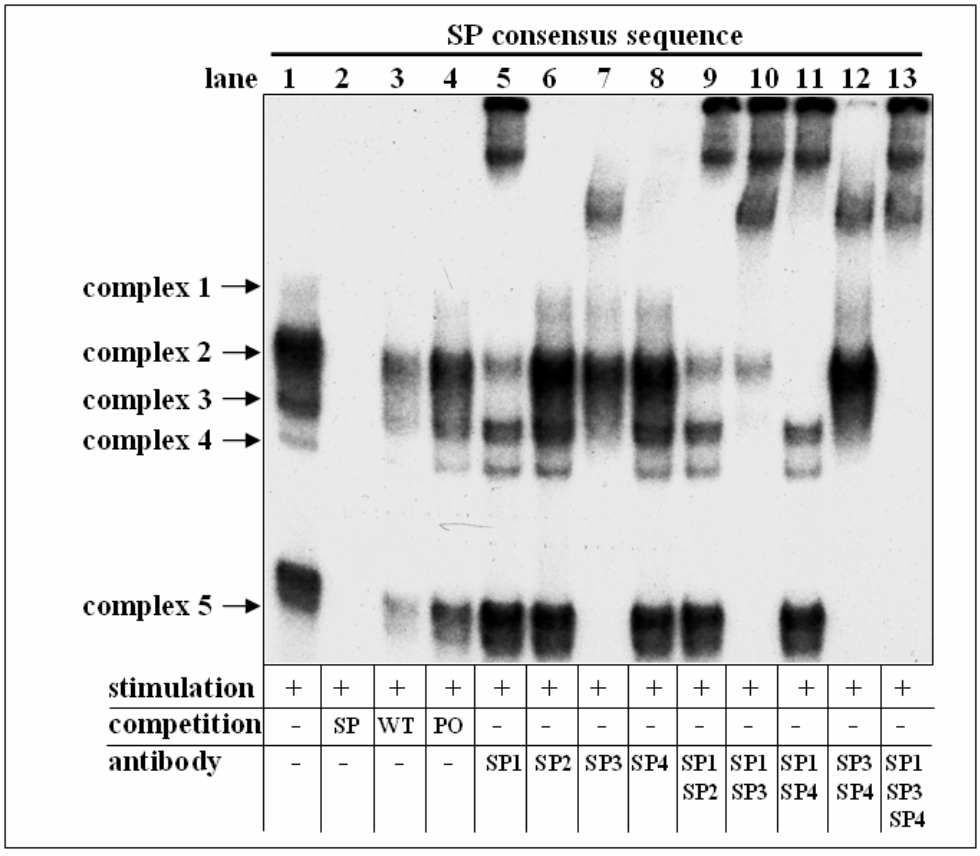

figure 30: EMSA analysis with a SP consensus sequence using nuclear extract (5µg) from the Jurkat T-cell line cultered for 3h either left in medium or stimulated with PMA/Ionomycin (50ng/ml, 1µM). The competitors (100-fold molar excess) and supershift antibodies (4µg) used for each experiment are noted below the respective lanes.

3.4.3 DNA/protein interaction (EMSA) of *STAT6* intron 2 depending on the allelic status of C2892T

3.4.3.1 General differences in DNA/protein interaction patterns depending on the allelic state of the probe harboring C2892T

To investigate differences in DNA/protein binding patterns between the region of intron 2 harboring SNP C2892T EMSA probes of 31bp containing either the C or T allele of SNP C2892T were generated. Binding patterns after incubation with unstimulated and stimulated nuclear extract from Jurkat T-cells were analyzed. Of note, in direct comparisons the general binding affinity of transcription factors to the 2892T probe (polymorphic, figure31, lanes 14-26, exposure time 42h) was much weaker than to C2892 probe (wildtype, figure31, lanes 1-13, exposure time 6h).

Incubation of the probe C2892 with either unstimulated or stimulated nuclear extract from the Jurkat T-cell line induced five complexes (figure 31, lanes 1 and 2). In contrast, when using the probe 2892T, an additional complex was observed with an increasing signal after stimulation. However, the affinity of any other complex was not influenced (figure 31, lanes 14 and 15). Competition experiments for both probes with a 100-fold molar excess of the respective unlabeled probe abrogated all observed complexes (figure 31, lanes 3 and 16). Further cross-competition experiments indicated that C2892 could not repress the additional complex 6 emerging with the incubation of 2892T (data not shown). Likewise, the probe containing the polymorphic T allele was not able to entirely complete the signal of the wildtype probe (data not shown).

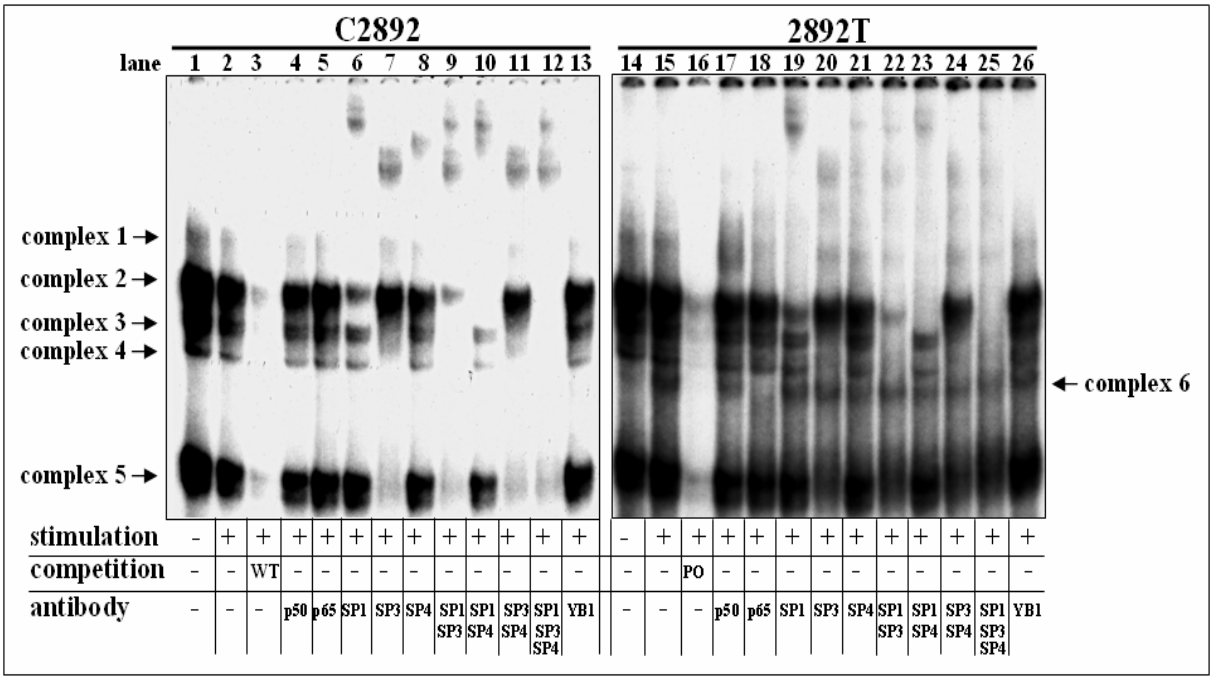

figure 31: EMSA analysis with a *STAT6* probe either carrying the wiltdtype (C2892, exposure time 6h) or the polymorphic (2892T, exposure time 42h) allele of the SNP C2892T using nuclear extract from the Jurkat T-cell line (5µg) cultered for 3h either left in medium or stimulated with PMA/Ionomycin (50ng/ml, 1µM). The competitors (100-fold molar excess) and supershift antibodies (4µg) used for each experiment are noted below the respective lanes.

3.4.3.2 EMSA analysis of STAT6 intron 2 reveal differences in NF-κB binding

Specific antibodies were pre-incubated with stimulated nuclear extract to identify the different transcription factors bound within this region of *STAT6* intron 2 (figure 31).

Since it was postulated that the transcription factor NF-κB may bind only to the polymorphic T allele, antibodies of both subunits p50 and p65 were used in supershift experiments. Indeed, the additional complex 6 which occurred selectively when the T allele was present contained NF-κB as the p50 antibody led to a weak supershift signal. The antibody p65 even induced the total abrogation of this complex (figure 31, lanes 17 and 18). No NF-κB binding at all was detectable with the probe corresponding to the C allele of *STAT6* intron 2 (figure 31, lanes 4 and 5).

3.4.3.3 EMSA analysis of STAT6 intron 2 reveal differnces in the transcription factor binding of members of the SP family

Members of the SP family were furthermore tested given that *in silico* analyses also postulated the binding of SP1. As specific SP1 antibodies induced the formation of a slower migrating supershift complex, the abrogation of complex 1 and the weakening of complex 2, it was concluded that both contained SP1 (figure 31, lane 6 and 19). Thus, the remaining complex (figure 31, lanes 6 and 19) was identified as SP4 due to the complete disruption of this signal when both antibodies SP1 and SP4 were applied together as shown in lanes 10 and 23 (figure 31). Complexes 3 and 4 contained SP3 because adding the specific antibody resulted in a slower migrating supershift signal (figure 31, lanes 7, 9, 11, 20, 22 and 24). Furthermore, through the triple incubation with antibodies for SP1, SP3 and SP4 together in one reaction all observed complexes consisting of members of the SP family were completely abolished (figure 31, lanes 13 and 25). In additional experiments it was shown that SP2 did not show any binding *per se* using the Jurkat T-cell nuclear. Specificity of these data was supported by an unrelated control antibody YB1 which did not alter the present band pattern (figure 31, lanes 13 and 26).

3.4.3.4 DNA/protein interaction of STAT6 intron 2 in a $CD4^+$ nuclear environment

To exclude that the observed effects with intron 2 were not due to cell line specific and putatively artificial conditions, further EMSA experiments were performed using nuclear extract from naïve $CD4^+$ T-cells: it was determined if intron 2 also exerts regulatory function in primary $CD4^+$ T-cells. Therefore, STAT6 DNA/protein interaction was compared between nuclear extract isolated from $CD4^+$ T-cells (figure 32, lanes 1-5) and the Jurkat T cell line (figure 32, lanes 6-12). Interestingly, using nuclear extract from $CD4^+$ T-cells one DNA/protein complex (complex 1, figure 32, lanes 1, 3 and 6) was absent in $CD4^+$ T-cells while it was present in Jurkat T-cells. In $CD4^+$ T-cells, complex 4 (figure 32, lanes 1, 3, 6 and 10) was only observed with probe C2892 while the same complex was consistently found in Jurkat T-cells. As in Jurkat T-cells, a strong NF-κB signal (figure 32, complex 6)

was detectable in $CD4^+$ T-cells when the polymorphic T allele was present (figure 32, lanes 3, 4 and 5) which was verified by specific antibodies (p50 and p65). Only a faint NF-κB binding was observed using C2892 as a probe (figure 32, lanes 1 and 2).

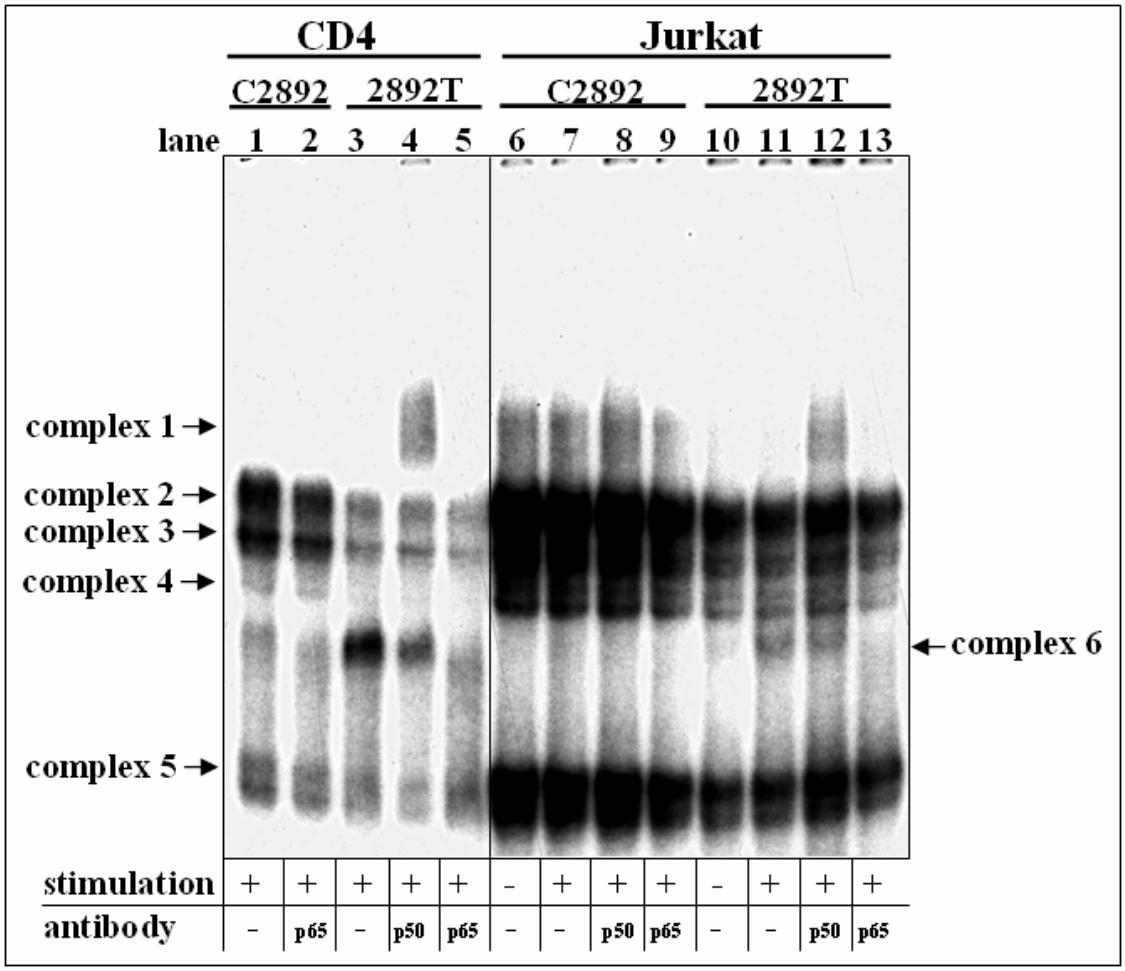

figure 32: EMSA analysis with a *STAT6* probe either carrying the wiltdtype (C2892) or the polymorphic (2892T) allele of the SNP C2892T using nuclear extract from freshly isolated $CD4^+$ T-cells or the Jurkat T-cell line (5µg) cultered for 3h either left in medium or stimulated with PMA/Ionomycin (50ng/ml, 1µM). The competitors (100-fold molar excess) and supershift antibodies (4µg) used for each experiment are noted below the respective lanes.

3.4.4 Extended binding analyses of the transcription factors NF-κB and members of the SP family

To further dissect how many basepairs aound SNP C2892T were necessary and sufficient for the binding of NF-κB and members of the SP family in the respective intron 2 region, we performed transversion experiments. Mutated EMSA probes were generated on wildtype or polymorphic background where one by one 5bp of the respective probe underwent transversion (table 16). Transversions *per se* refer to the substitution of a purine (adenosine, guanine) for a pyrimidine (cytosine, thymine) or vice versa to achieve maximal disruption of the DNA structure.

NF-κB binding was only impaired by the modification of the *STAT6* probe harboring the T allele represented by blocks 3 and 4 (figure 33). All other regions did not affect the binding of NF-κB, respectively. Additionally, block 3 abolished SP1, SP3 and SP4 transcription factor binding whereas in block 4 and 5 (figure 33) the binding affinity of the transcription factors appeared to be diminished.

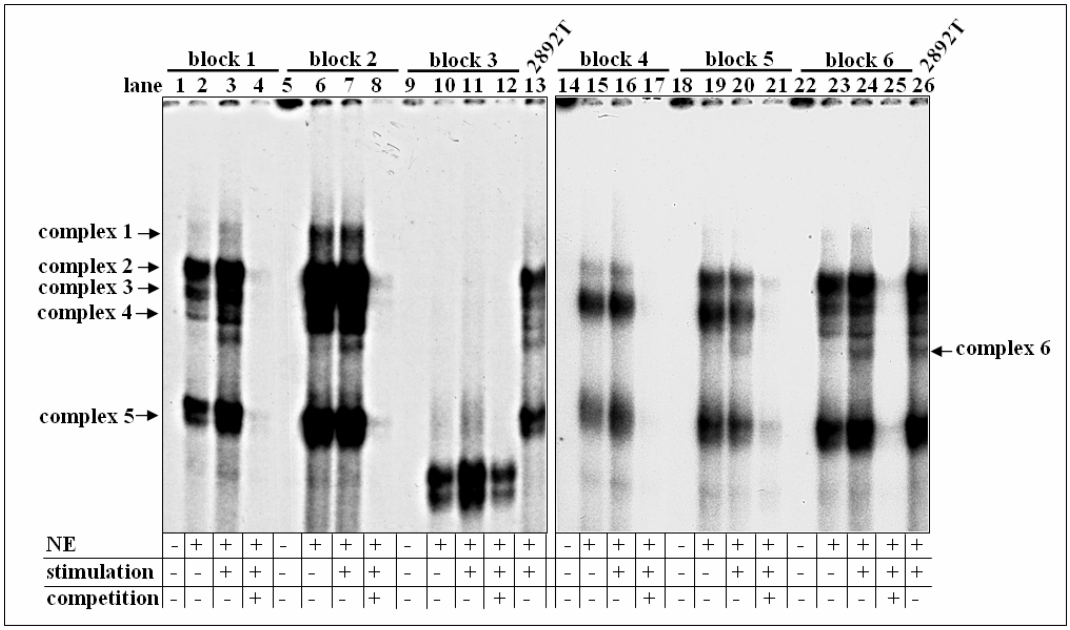

figure 33: EMSA analysis using probes with 5bp non-overlapping transversion on the background of the polymorphic *STAT6* probe. The experiments were performed with nuclear extract (NE) from the Jurkat T-cell line (5µg) cultered for 3h either left in medium or stimulated with PMA/Ionomycin (50ng/ml, 1µM). The competitors (100-fold molar excess) and supershift antibodies (4µg) used for each experiment are noted below the respective lanes.

Likewise six mutated probes with 5bp transversions on the wildtype background of C2892T were used to compare if the same domains were required for SP binding. Performing EMSA experiments with these mutated probes, it was demonstrated that the most relevant interaction of the SP family emerged within block 9, 10 and 11 (figure 34, corresponding to blocks 3-5 in figure 33). Block 12 (figure 34) only showed quantitative changes in the binding properties. Hence, the disruption of DNA/protein complexes affected an expanded region and was also much more severe in contrast to the probe containing the T allele of C2892T. All other adjacent blocks did not seem to influence the affinity of the transcription factors vigorously.

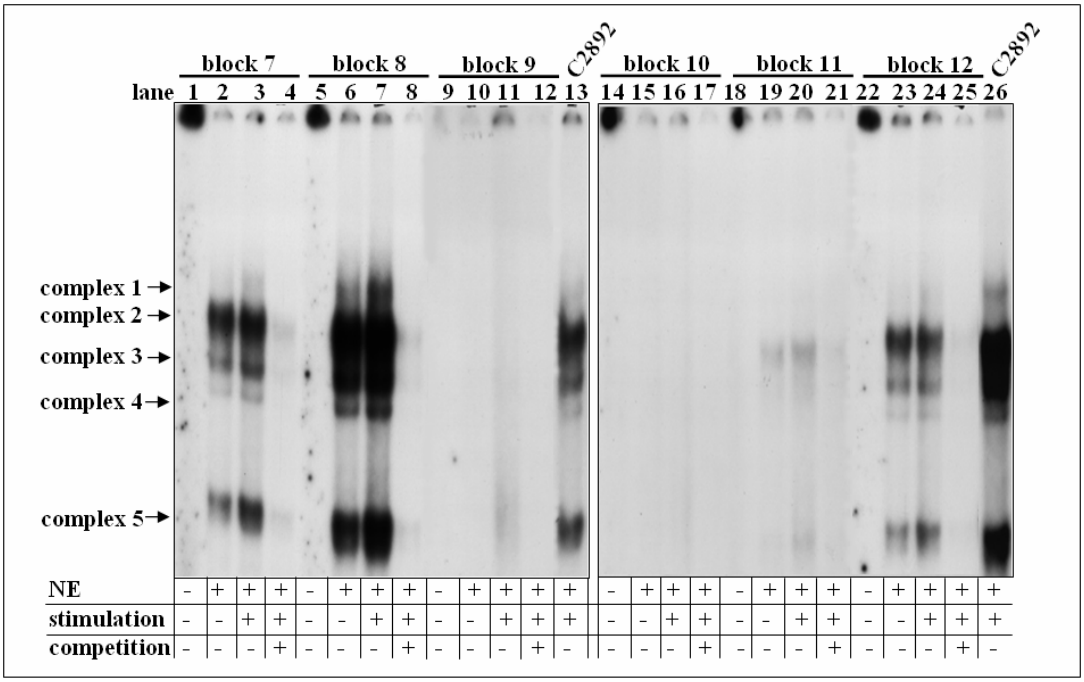

figure 34: EMSA analysis using probes with 5bp non-overlapping transversion on the background of the wildtype *STAT6* probe. The experiments were performed with nuclear extract (NE) from the Jurkat T-cell line (5µg) cultered for 3h either left in medium or stimulated with PMA/Ionomycin (50ng/ml, 1µM). The competitors (100-fold molar excess) and supershift antibodies (4µg) used for each experiment are noted below the respective lanes.

Thus, these experiments indicated that 5bp at each side of the polymorphic probe 2892T are necessary and sufficient for NF-κB binding. Partly overlapping but different regions seemed to be responsible for SP binding. This is schematically depicted in the following graphic overview (figure 35).

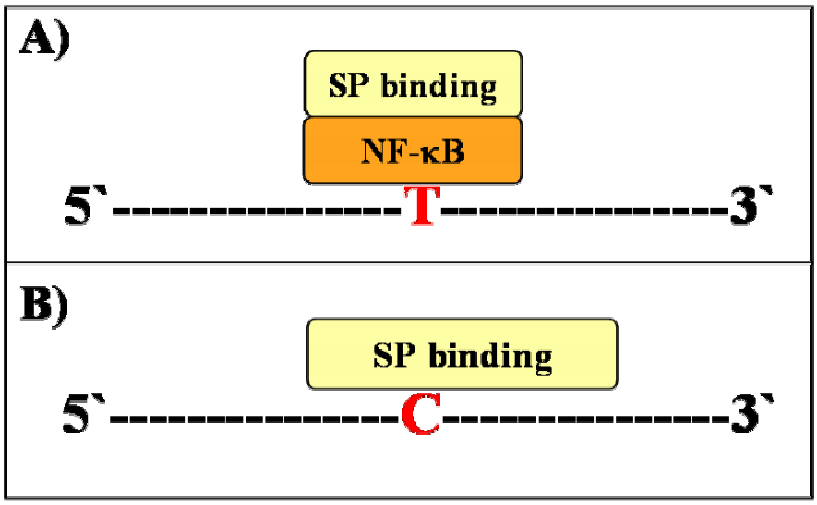

figure 35: Schematic differences of the relevant regions for the binding of transcription factors to the probe either representing the polymorphic T (A) or the wildtype C (B) allele of the SNP C2892T.

3.4.5 Analysis of a second putative NF-κB binding site in close vicinity to the polymorphism C2892T

A second putative binding site (S6_NF-κB) for NF-κB was predicted through *in silico* analyses (see section 3.3.3) in close vicinity to polymorphism C2892T within a distance of only 21bp. Thus, this putative NF-κB site was also analyzed by EMSA as it was speculated, that both NF-κB sites may interact with each other to influence S*TAT6* regulation.

EMSA experiments for this second potential NF-κB site (S6_NF-κB, figure 36) revealed the formation of five specific complexes (figure 36, lanes 2 and 3), all of which were abrogated by adding a 100-fold molar excess of the cold probe (figure 36, lane 4). However, the same amount of unlabeled NF-κB consensus sequence hardly diminished (figure 36, lane 5) the signal already indicating that the region of interest may not be capable of NF-κB binding. This first evidence was supported by

supershift experiments pre-incubating the nuclear extract with p50 and p65 antibody. These were neither able to disrupt any of the complexes nor did they induce a supershift signal (figure 36, lanes 6 and 7). Hence, it was concluded that this second region of *STAT6* intron 2 did not harbor a real high-affinity site for NF-κB, in contrast to the *in silico* prediction.

However, the observed complexes were virtually identical to the SP consensus sequence (figure 30, lanes 5-13). Adding a 100-fold molar excess of the cold SP consensus probe disrupted the formation of all complexes. Supershift experiments using specific antibodies for SP1, SP3 and SP4 confirmed this assumption as supershift patterns were completely analog (figure 36, lanes 8-13). Thus, the region predicted to harbor a NF-κB site turned out to be incapable to actually bind NF-κB, while the binding of members of the SP family was demonstrated by specific antibodies.

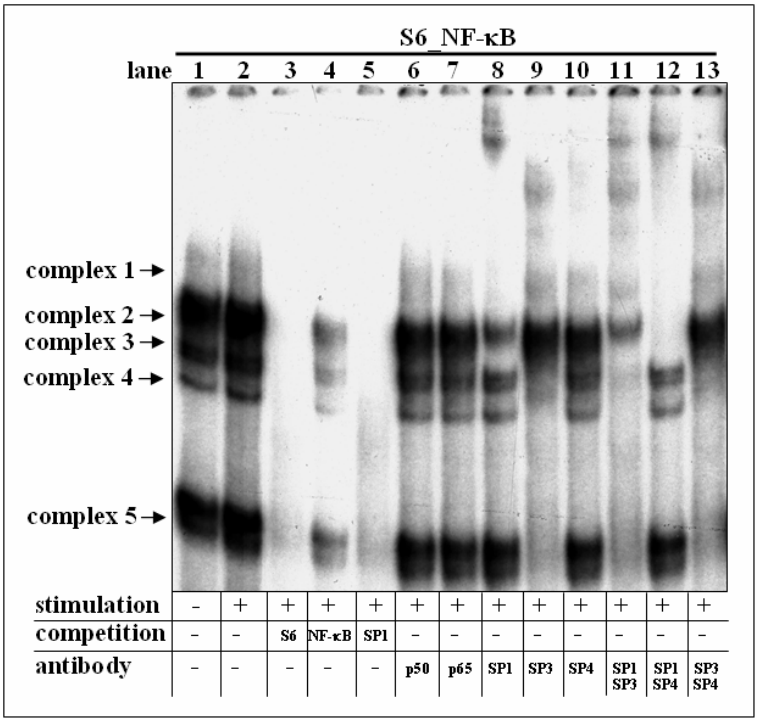

figure 36: EMSA analysis for a second putative NF-κB binding site in the *STAT6* gene located in close vicinity to the SNP C2892T (approximately only 21bp away). The experiments were performed with nuclear extract from the Jurkat T-cell line (5µg) cultered for 3h either left in medium or stimulated with PMA/Ionomycin (50ng/ml, 1µM) was used. The competitors (100-fold molar excess) and supershift antibodies (4µg) used for each experiment are noted below the respective lanes.

3.4.6 DNA/protein interaction analyses of a SNP in linkage disequilibrium with C2892T

So far all our functional experiments were focused on SNP C2892T. However, it may be argued that the observed association signals between C2892T and elevated IgE levels in our population are caused by other SNPs in linkage disequilibrium with C2892T which were not assed by genotyping. Applying the HapMap database, it became clear that C2892T was in high linkage disequilibrium (LD, $r^2 = 0.90$) with a polymorphism (rs167769) also located in *STAT6* intron 2 approximately 1600bp upstream of C2892T.

Thus, functional experiments concerning DNA/protein interaction analyses were performed to investigate if the association with elevated total serum IgE levels was due to rs167769 (figure 37). However, no qualitative difference in the binding pattern of transcription factors was observed for rs167769 comparing probes representing the wildtype (figure 37, lanes 3 and 4) or the polymorphic allele (figure 37, lanes 7 and 8).

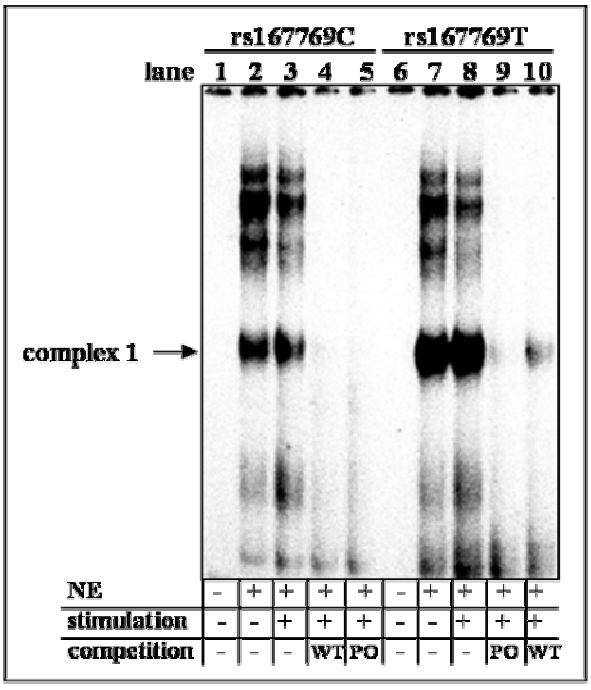

figure 37: EMSA analysis for rs167769 located in close vicinity to the SNP C2892T. The experiments were performed with nuclear extract from the Jurkat T-cell line (5µg) cultered for 3h either left in medium or stimulated with PMA/Ionomycin (50ng/ml, 1µM) was used. Competitors (100-fold molar excess) used for each experiment are noted below the respective lanes.

3.5 Transcriptional activity of *STAT6* intron 2 depending on C2892T

The previous EMSA experiments had shown that the genotype of SNP C2892T induced differences in the binding of NF-κB and members of the SP family in the region of *STAT6* intron 2. Thus, we hypothesized that changes in NF-κB and SP binding could affect *STAT6* gene expression. This hypothesis was tested in gene expression assays. As the polymorphism C2892T was not located in a promoter or coding region of the *STAT6* gene but in an intronic region, it was speculated that due to its position, intron 2 may act as a *cis*-regulatory element. Intially, the biological function of the *STAT6* promoter alone and in combination with intron 2 had to be investigated. Only then, the role of the polymorphism C2892T on gene expression could be studied.

3.5.1 Defining the role of the human *STAT6* promoter in the human Jurkat T-cell line

Up to date, the only published *STAT6* promoter studies were performed in NIH3T3 mouse fibroblasts. *STAT6* promoter constructs of different size were used to identify the minimal region which is necessary for maximal *STAT6* transcription. Previously, a minimal promoter with a size of approximately 1100bp was identified to induce the highest constitutive luciferase activity in comparison to the pGL3-Basic vector [120]. Due to the fact that no *STAT6* promoter studies had previously been performed in human cells, it first had to be determined if in human T-cells, comparable *STAT6* promoter activities could be observed. To address this question, transient transfection experiments were performed using *STAT6* promoters of different sizes (pBP78 = 5,6kb, pBP88 = 2,6kb and pBP86 = 1,1kbp) previously published [120]. These constructs were a kind gift from Bharvin Patel who had performed these mutual *STAT6* promoter studies. The respective promoter constructs were re-sequenced and as required, site-directed mutagenesis was conducted to obtain the accurate sequence of the respective *STAT6* promoter (2.2.13.2). A T-cell line was selected for these experiments as STAT6 plays a crucial role in the IL-4/IL-13 pathway which is activated in T-cells.

Transfection settings had to be optimized for our specific experimental conditions before the activity of the *STAT6* promoter in the Jurkat T-cell line could be analyzed properly. In a first step, the concentration of co-transfected renilla vector pRL-TK was titrated. In these experiments 10ng of pRL-TK per $8x10^6$ cells was determined to give optimal transfection results. Electroporation settings had to be optimized due to massive cell death in Jurkat T-cells with the settings suggested by the manufacturer. Numbers of cells for each transfection were adjusted to $8x10^6$ cells/transfection whereas the application of $5x10^6$ and $3x10^6$ cells/transfection led to a rigorously decreasing signal of detectable luciferase activity. Additionally, the electroporation settings using Jurkat T-cells were lowered from 240V to 200V leading to more reproducible results and less cell death.

Now, the appropriate stimuli for the activation of the *STAT6* promoter could be tested using different approaches. On the one hand PMA/Ionomycin was applied for stimulation. PMA/Ionomycin is known as a general T-cell activator without antigen specificity. We also tried to stimulate Jurkat T-cells with IL-4, as IL-4 would induce more specific T-cell stimulation through the activation of the IL-4/IL-13 pathway. As presented in figure 38, different concentrations of IL-4 were not able to activate the *STAT6* promoter pBP78 (5.6kb) in comparison to the unstimulated sample. This may be in line with previously published observations where a defect in the activation of the IL-4/IL-13 pathway was documented [158]. In contrast, after stimulation of 50ng/ml PMA and 1µM Ionomycin higher luciferase activity of the *STAT6* promoter was observed but also induced more cell death.

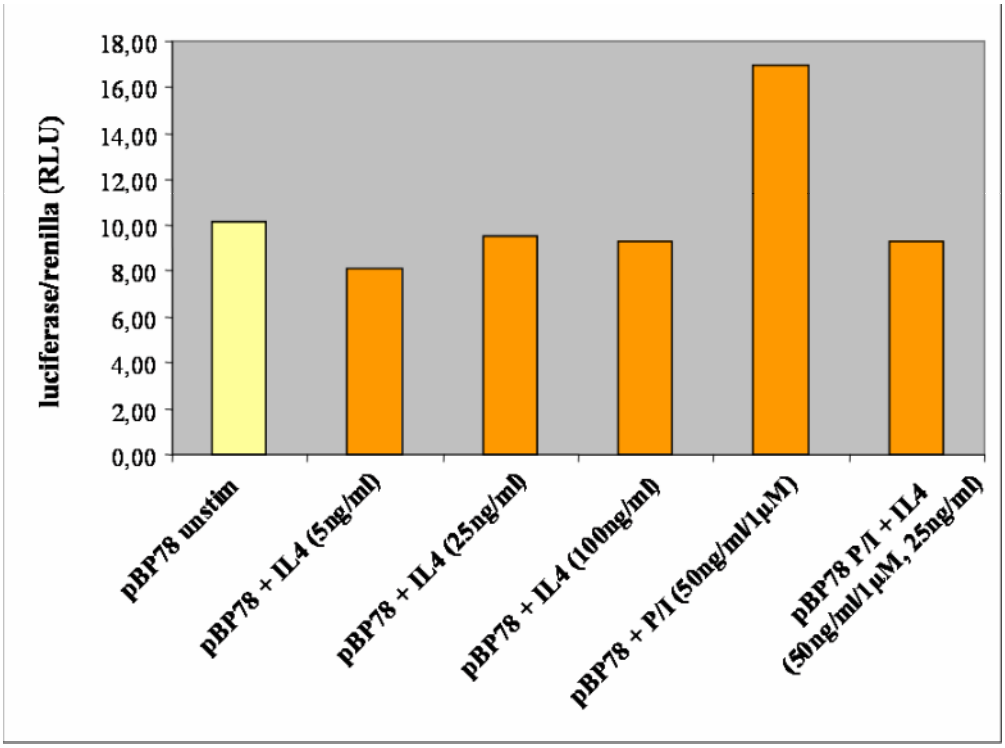

figure 38: Jurkat T-cells were transiently transfected with the *STAT6* promoter construct pBP78 (5.6kb). Cells were left in medium (yellow) or stimulated (orange) with the respective stimuli and harvested 24h after the transfection. Luciferase activity was normalized for transfection efficiency using the control plasmid pRL-TK expressed by renilla activity. The relative luciferase activity is presented in relative light units (RLU).

In further experiments with the pBP86 promoter (1.1kb) the concentration of PMA/Ionomycin and an additional time-point (3h after transfection = EPO) for the stimulation was tested. These data indicated that 12.5ng/ml PMA and 0.5µM Ionomycin 3h after transfection was sufficient for stimulating the *STAT6* promoter (figure 39).

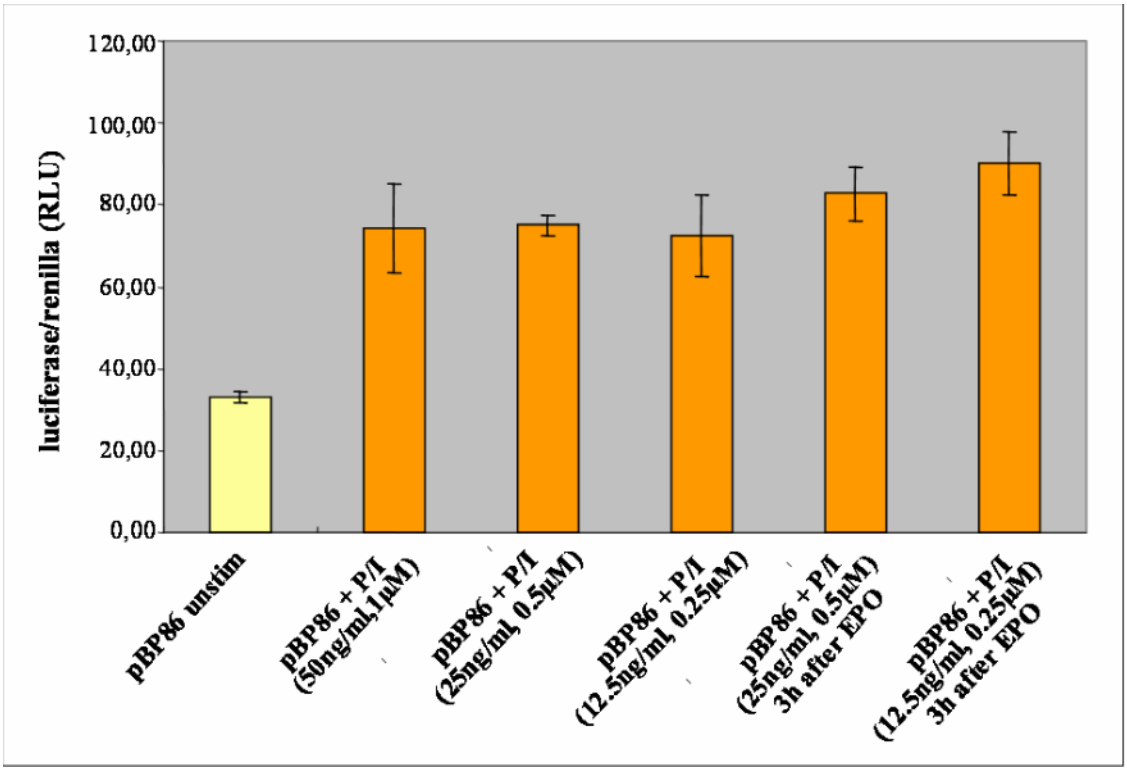

figure 39: Jurkat T-cells were transiently transfected with the *STAT6* promoter construct pBP86 (1.1kb). Cells were left in medium (yellow) or stimulated (orange) with the respective concentration of stimuli (PMA/Ionomycin). Cells were either stimulated right after or 3h after (= 3h after EPO) transfection. Cells were harvested 24h after the transfection. Luciferase activity was normalized for transfection efficiency using the control plasmid pRL-TK expressed by renilla activity. The relative luciferase activity is presented in relative light units (RLU).

Therefore, all further transfection experiments were performed as follows: $8x10^6$ cells were transfected with 4µg of the respective *STAT6* construct and co-transfected with 10ng of renilla pRL-TK vector using the square wave electroporation setting with one pulse for 50msec at 200V. Cells were stimulated with PMA/Ionomycin (12.5ng/ml, 0.5µM) 3h after transfection and harvested 21h after stimulation.

3.5.2 Characterization of the human *STAT6* promoter in the Jurkat T-cell line

As described before, the promoter pBP86 (1,1kb) of the *STAT6* gene was identified expressing the highest promoter activity in NIH3T3 mouse fibroblasts [120]. Using a human T-cell line (Jurkat T-cells) for *STAT6* promoter studies, data consistent with these previous reports were obtained. Also in human T-cells the pBP86 construct was the most active *STAT6* promoter with a 13-fold (unstimulated) and 15-fold (stimulated) up-regulation of relative luciferase activity (figure 40). In comparison, the largest promoter construct pBP78, with a size of approximately 5.6kb, led to a 4-fold increase in expression over the empty pGL3-Basic vector (figure 40).

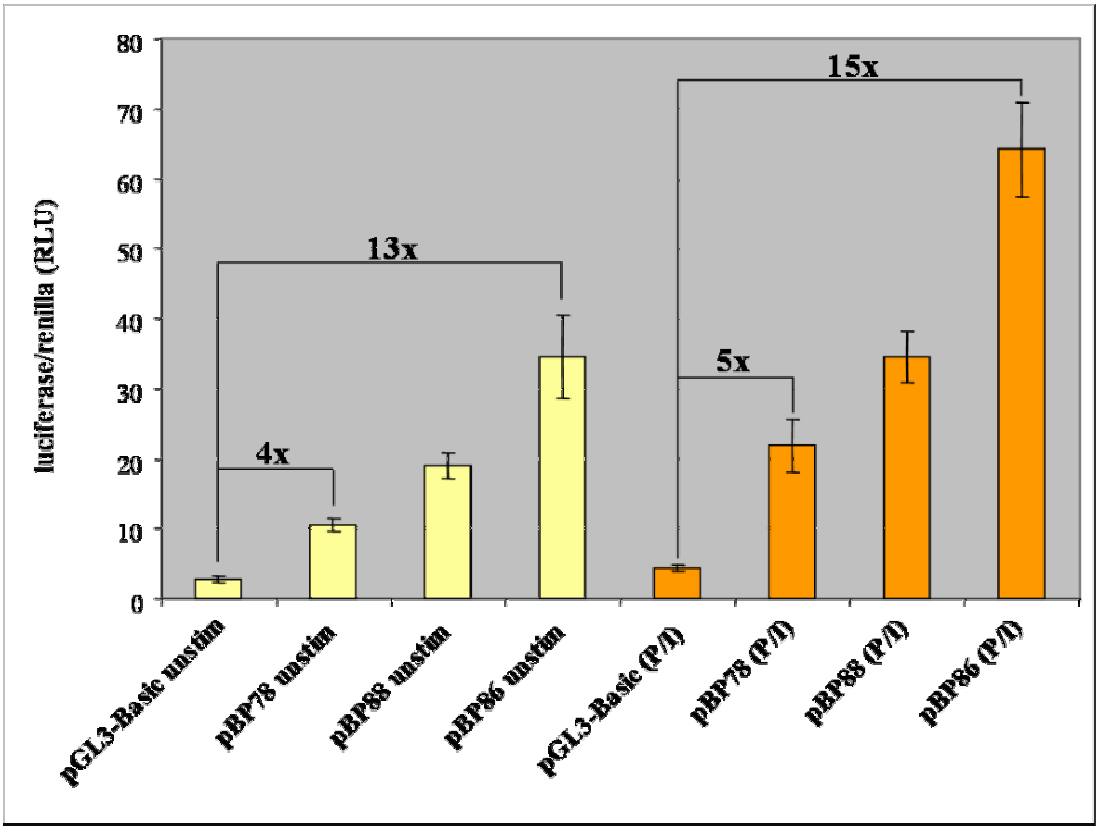

figure 40: To define the minimal *STAT6* promoter, Jurkat T-cells were transiently transfected either with the pGL3-Basic vector or the *STAT6* promoter constructs pBP78 (5.6kb), pBP88 (2.6kb) and pBP86 (1.1kb). Cells were left in medium (yellow) or stimulated (orange) with PMA/Ionomycin (0.25ng/ml, 12.5µM) 3h after transfection and harvested 21h after stimulation (N=10). Luciferase activity was normalized for transfection efficiency using the control plasmid pRL-TK expressed by renilla activity. The relative luciferase activity is presented in relative light units (RLU).

In addition to the previous vectors, a further *STAT6* promoter contruct was generated and compared to existing constructs (see section 2.2.13.1.2). This was necessary, as a previous study had suggested that GT-repeats may play a role in *STAT6* basal promoter activity [126]. We identified two GT-repeats in the *STAT6* promoter. The previously described GT-repeat was absent in pBP86 while a second, newly identified GT-repeat was still present in that vector. By deleting 200 bp at the 5' end of pBP86 a new vector (pBP85) was created now lacking all GT repeats in the *STAT6* promoter. Comparing the most active promoter pBP86 with the abbreviated pBP85 (lacking this GT-repeat) in our human gene expression system, no difference in their transcriptional activity was detectable (figure 41). Hence, it was concluded that this GT-repeat did not affect *STAT6* promoter activity.

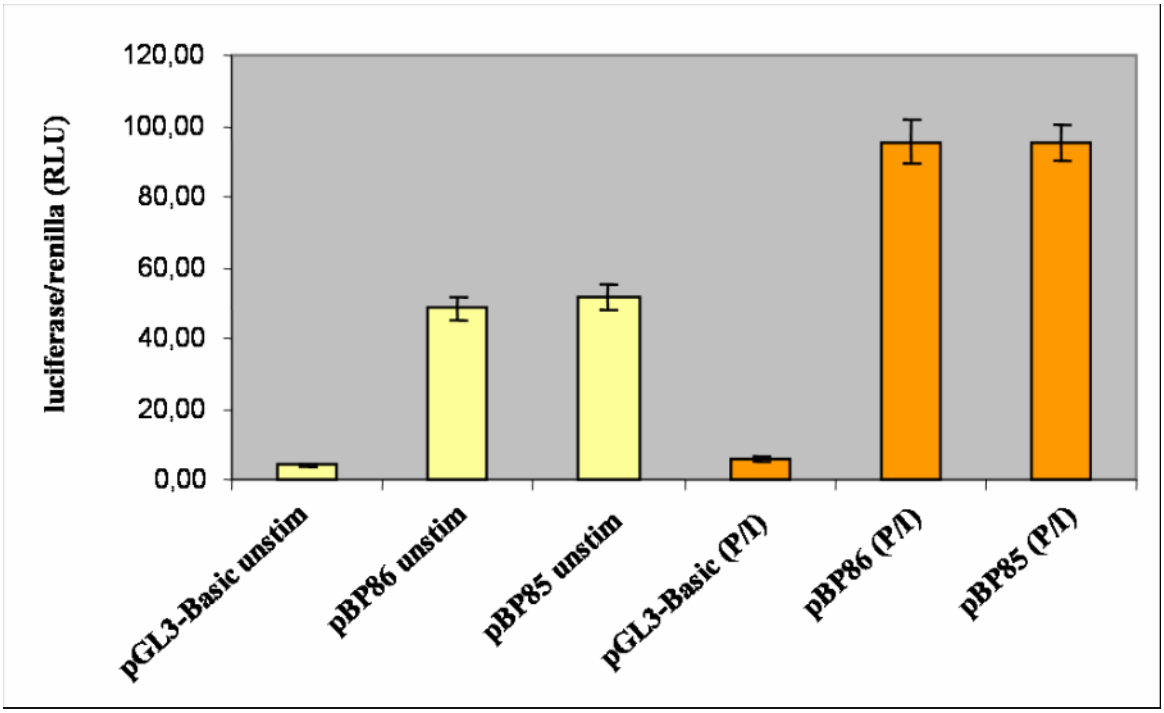

figure 41: To analyse the influence of a GT-repeat on the transcriptional activity of the *STAT6* promoter, Jurkat T-cells were transiently transfected either with the pGL3-Basic vector or the *STAT6* promoter constructs pBP86 and pBP85. Cells were left in medium (yellow) or stimulated (orange) with PMA/Ionomycin (0.25ng/ml, 12.5µM) 3h after transfection and harvested 21h after stimulation (N=10). Luciferase activity was normalized for transfection efficiency using the control plasmid pRL-TK expressed by renilla activity. The relative luciferase activity is presented in relative light units (RLU).

3.5.3 *STAT6* intron 2 acts as a silencing *cis*-regulatory element which is influenced by the allelic status of the polymorphism C2892T

While our previous experiments had established *STAT6* promoter function in a human cell line system, we next aimed to study the role of intron 2 in this system: intron 2 was located in a highly conserved region of the *STAT6* gene and DNA/protein interaction analyses supported its putative functional relevance as a *cis*-regulatory element.

Cis-regulatory elements may interact physically with promoter regions to exceed their regulatory function. This could e.g. be achieved by loop structures getting intronic regions in close proximity of certain promoter elements. For *STAT6* it was so far not known, if *cis*-regulatory mechanisms exist and which parts of the promoter may be involved. For intron 2, we hypothezised, according to our EMSA results with SNP C2892T, that this region must be involved in *STAT6* regulation and that the allelic state of C2892T may influence these suspected promoter-intron interactions.

Thus, transient transfection experiments were performed using the most active promoter pBP86 (1.1kbp) and the largest promoter pBP78 (5.6kb) both containing intron 2 downstream of the luciferase gene carrying either the wildtype C (corresponding to pBP86_WT) or the polymorphic T allele (corresponding to pBP86_PO) (see section 2.2.13.1). Transient transfection experiments were performed in Jurkat T-cells (unstimulated and stimulated with PMA/Ionomycin).

It was demonstrated that *STAT6* intron 2, expressing the C allele, (pBP86_WT) in combination with the pBP86 promoter (1.1kb) acted more as a silencing element indicated by the drop of approximately 20% of detectable luciferase activity in comparison to the empty pBP86 promoter construct. The outcome was not influenced by stimulation (figure 42). A similar effect was observed for the respective *STAT6* construct (pBP86_PO) containing the polymorphic T allele likewise with a decline of 30% (figure 42).

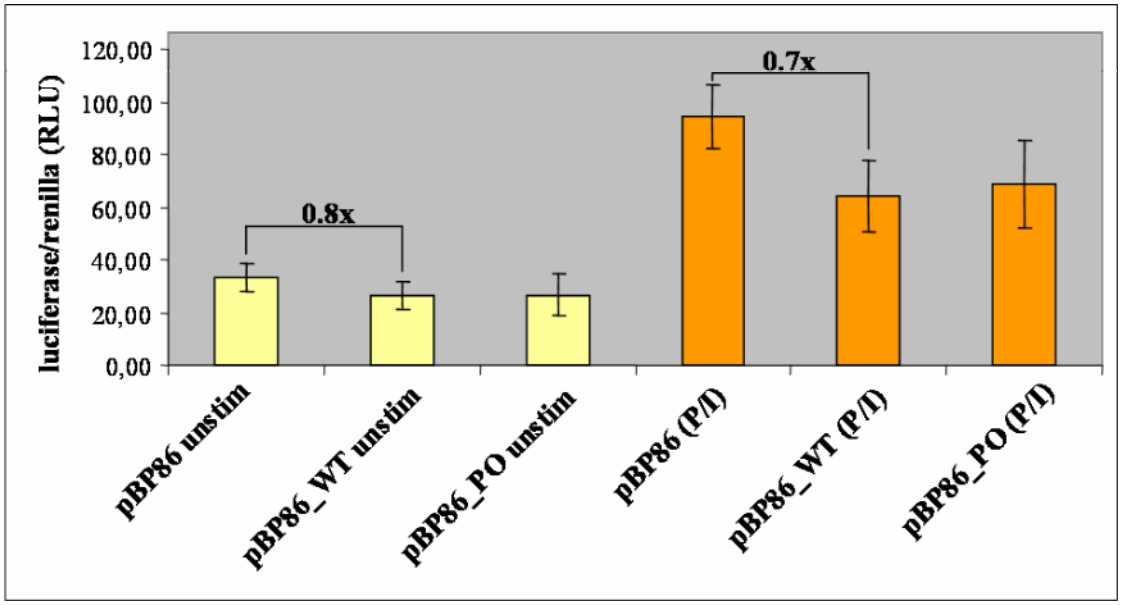

figure 42: The regulatory function of intron 2 depending on the polymorphism C2892T on the transcription activity of *STAT6* was investigated by transient transfection experiments. Therefore, Jurkat T-cells were transfected either with the empty pBP86 vector or the respective constructs carrying the wildtype (pBP86_WT) or the polymorphic allele (pBP86_PO) of the SNP C2892T. Cells were left in medium (yellow) or stimulated (orange) with PMA/Ionomycin (0.25ng/ml, 12.5µM) 3h after the transfection and harvested 21h after stimulation (N=10). Luciferase activity was normalized for transfection efficiency using the control plasmid pRL-TK expressed by renilla activity. The relative luciferase activity is presented in relative light units (RLU).

Comparable experiments were performed using the larger pBP78 promoter construct (5.6kb) as promoter background. Again, these *STAT6* promoter constructs additionally carrying *STAT6* intron 2 downstream of the luciferase gene with both allelic conditions were transfected (see section 2.2.13.1). Evaluating the relative luciferase activity of pBP78_WT, these constructs were likewise less active than the *STAT6* promoter pBP78 (figure 43). But interestingly, this repression of gene expression was drastically diminished when the T allele (pBP78_PO) at C2892T was introduced: the amount of detectable luciferase activity increased by 70% in direct comparison to the *STAT6* construct corresponding to the wildtype allele (pBP78_WT) of C2892T.

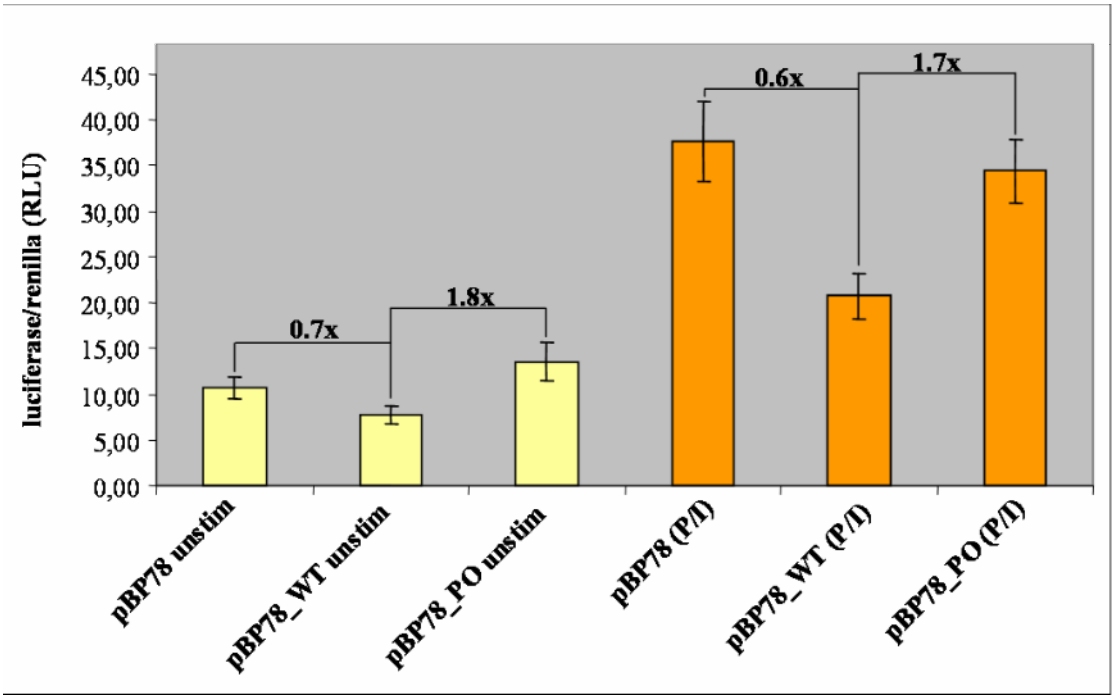

figure 43: The regulatory function of intron 2 depending on the polymorphism C2892T on the transcription activity of *STAT6* was investigated by transient transfection experiments. Therefore, Jurkat T-cells were transfected either with the empty pBP78 vector or the respective constructs carrying the wildtype (pBP78_WT) or the polymorphic allele (pBP78_PO) of the SNP C2892T. Cells were left in medium (yellow) or stimulated (orange) with PMA/Ionomycin (0.25ng/ml, 12.5µM) 3h after the transfection and harvested 21h after stimulation (N=10). Luciferase activity was normalized for transfection efficiency using the control plasmid pRL-TK expressed by renilla activity. The relative luciferase activity is presented in relative light units (RLU).

These data supported the hypothesis that *STAT6* intron 2 indeed possessed functional relevance. The intron itself acts as a silencing regulatory element. Additionally, it was shown that the repressing function depends on the allelic status of the analyzed polymorphism C2892T. This effect was more pronounced when the extended promoter contruct (pBP78) was used indicating that intron 2 interacts with a region of the *STAT6* promoter further upstream of the minimal promoter.

As a *cis*-regulatory element is defined to act independent of its orientation, further experiments were performed in which intron 2 was cloned 3`5` (reverse) into the respective site of the pBP78 vector downstream of the luciferase gene (see section 2.2.13.1). These experiments were now only performed with the pBP78 promoter which had been proven to be the most responsible contruct available in our experimental setting. When the orientation of the regulatory intron 2 element in the vector was changed, transcriptional activity of pBP78 was not significantly altered (figure 44). However, the allelic status of C2892T still influenced promoter activity: the T allele in *STAT6* intron 2 (pBP78_PO) increased transcription compared to the wildtype C allele (pBP78_WT) (figure 44).

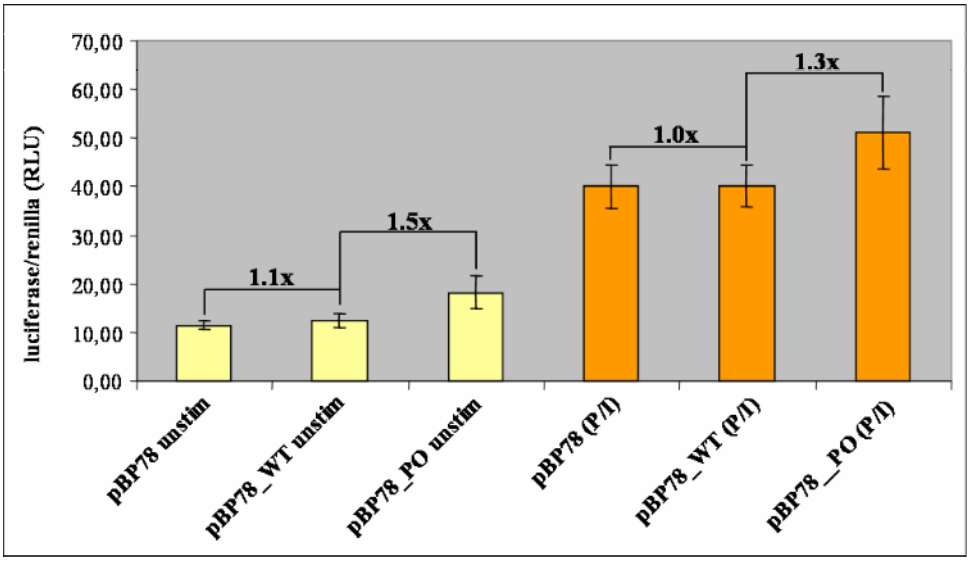

figure 44: To analyse whether the regulatory function of intron 2 with the polymorphism C2892T was independent on its orientation, transient transfection experiments were performed carrying intron 2 3`5` (reverse). Therefore, Jurkat T-cells were transfected either with the empty pBP78 vector or the respective constructs carrying the wildtype (pBP78_WT) or the polymorphic allele (pBP78_PO) of the SNP C2892T. Cells were left in medium (yellow) or stimulated (orange) with PMA/Ionomycin (0.25ng/ml, 12.5µM) 3h after the transfection and harvested 21h after stimulation (N=10). Luciferase activity was normalized for transfection efficiency using the control plasmid pRL-TK expressed by renilla activity. The relative luciferase activity is presented in relative light units (RLU).

Additional transfection experiments were conducted to assess if the intron 2 element may also have direct promoter activity in combination with the *STAT6* promoter construct pBP78. This hypothesis was tested with intron 2 being cloned in between the pBP78 promoter and the luciferase gene (see section 2.2.13.1). As presented in figure 45, intron 2 had no promoter activity and disrupted the promoter activity of pBP78 completely when it was situated directly upstream of the luciferase gene. Thus, the location of the intron 2 element inrespect to the *STAT6* promoter and the start of transcription appear to be important.

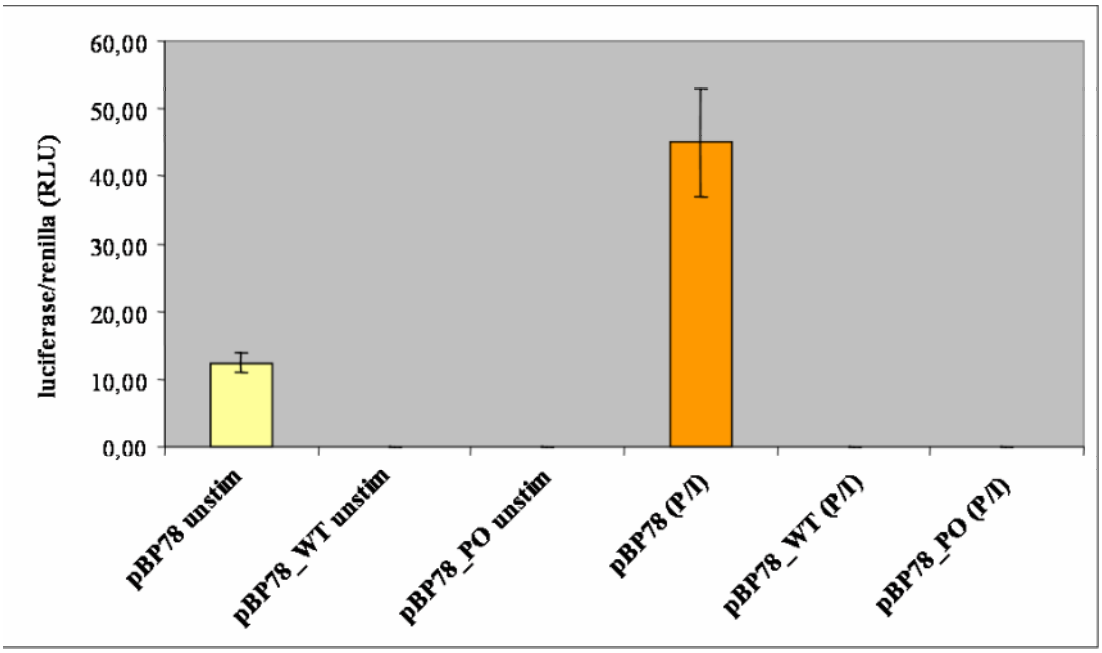

figure 45: To investigate whether the influence of intron 2 on *STAT6* gene expression might be affected by its positional vicinity to the STAT6 promoter, intron 2 with the polymorphism C2892T was cloned upstream of the luciferase gene. Therefore, Jurkat T-cells were transfected either with the empty pBP78 vector or the respective constructs carrying the wildtype (pBP78_WT) or the polymorphic allele (pBP78_PO) of the SNP C2892T. Cells were left in medium (yellow) or stimulated (orange) with PMA/ionomycin (0.25ng/ml, 12.5µM) 3h after the transfection and harvested 21h after stimulation (N=4). Luciferase activity was normalized for transfection efficiency using the control plasmid pRL-TK expressed by renilla activity. The relative luciferase activity is presented in relative light units (RLU).

3.6 *Ex vivo* gene expression of *STAT6* splice variants is modulated depending on the genotype status of the polymorphism C2892T

Based on our *in vitro* results, intron 2 of the *STAT6* gene seems to acts as a *cis*-regulatory element. It was demonstrated in the present study that intron 2 reduces *STAT6* gene expression, acting as a silencing element. Furthermore, the repressor function of intron 2 appeares to depend upon the SNP C2892T whereas the presence of the polymorphic T allele increases *STAT6* gene expression. This may be explained by the alternative binding of the transcription factor NF-κB since a novel NF-κB binding site was found only with the T allele of the polymorphism. While gene expression experiments so far were based on cell lines using *STAT6* expression constructs, in further experiments it was now examined if the polymorphism located in intron 2 may affect *STAT6* mRNA expression *ex vivo* on a population-based level.

3.6.1 Characterization of novel *STAT6* isoforms

In a first step, the impact of the SNP C2892T on the relative *STAT6* gene expression *ex vivo* was investigated. Ten individuals were defined from 40 randomly selected adult volunteers by determining the respective homozygote genotype for C2892T by the method Restriction Fragment Length Polymorphism (RFLP, see section 2.2.8.4). Then, the *STAT6* gene expression in PBMCs (Peripheral Blood Mononuclear Cell) was compared between these two groups consisting of five individuals each, either carrying wildtype C (CC) or polymorphic T (TT) allele homozygously. PCR of the samples was performed under standard conditions (see section 2.2.14.4) using the forward primer STAT6_292fwd which was partially located in exon 16 and partially in exon 17. The reverse primer STAT6_933rev was complementary to a region in exon 19 leading to an expected PCR fragment of 304bp when cDNA was used as template. The position of these primers allowed for the detection of putative residues of genomic DNA in the RNA identified through an emerging PCR fragment of 957bp (figure 46).

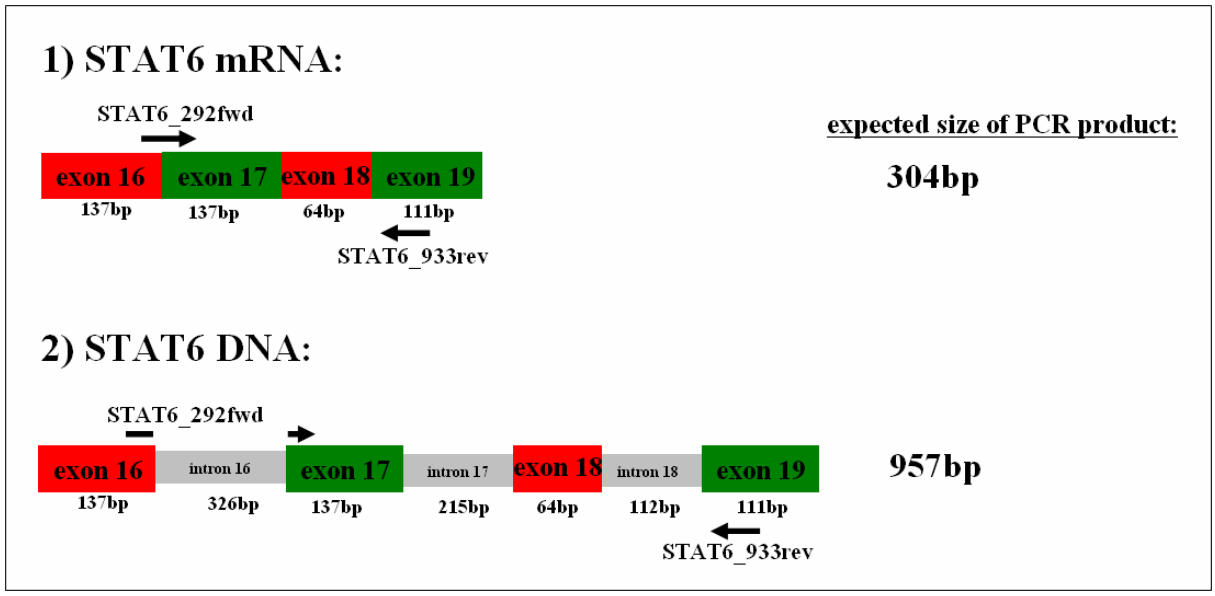

figure 46: Location and expected size of PCR fragments of the *STAT6* primers for quantitative real-time PCR: using the primers STAT6_292fwd and STAT6_933rev a PCR fragment of 304bp was expected using cDNA as a template. In contrast, a fragment of 957bp was expected using genomic DNA as template.

Surprisingly, three bands in addition (bands 2-4) to the expected fragment size of 304bp (band 1) were observed after the amplified fragments were separated on an agarose gel as displayed in figure 47. While band 1 appeared at the height of the targeted RNA fragment (304bp), it was speculated that band 2 (approximately 500bp) and 3 (approximately 600bp) may correspond to novel *STAT6* isoforms putatively containing additional intronic regions of the *STAT6* gene due to alternative splicing. The respective fragments (figure 47, bands 1 - 4) were eluted separately and re-sequenced (see section 2.2.6) to characterize their origin. Comparing the obtained sequences to the database, the hypothesis for band 1 - 3 was validated (figure 47). Band 1 was identified as the known *STAT6* (further referred to overall STAT6) splice variant [120]. The sequence of the fragment of band 2 contained not only the exons 16 to 19 but additionally the region of intron 17 leading to a fragment size of 519bp (further referred to as STAT6d, figure 49). The sequence of band 3 revealed a splice variant which included the respective exons as well as both intronic regions of intron 17 and 18 (expected fragment size of 631bp) and will further be referred to as STAT6e (figure 49). While performing these studies Tang *et al*. described the identification of a novel gene (*STAT6B*) homologous to members of the STAT family [159]. But we suggest that *STAT6B* represents a *STAT6* isoform similar to STAT6e. Band 4 was initially

assumed to represent residual DNA in the RNA due to inefficient DNase treatment. However, performing PCR with genomic DNA as template led to the detection of the expected PCR product of approximately 950bp, yet the observed band revealed a size close to 1100bp which could not be identified by sequencing. Therefore, the amplified region may represent an unspecific PCR product unrelated to *STAT6*.

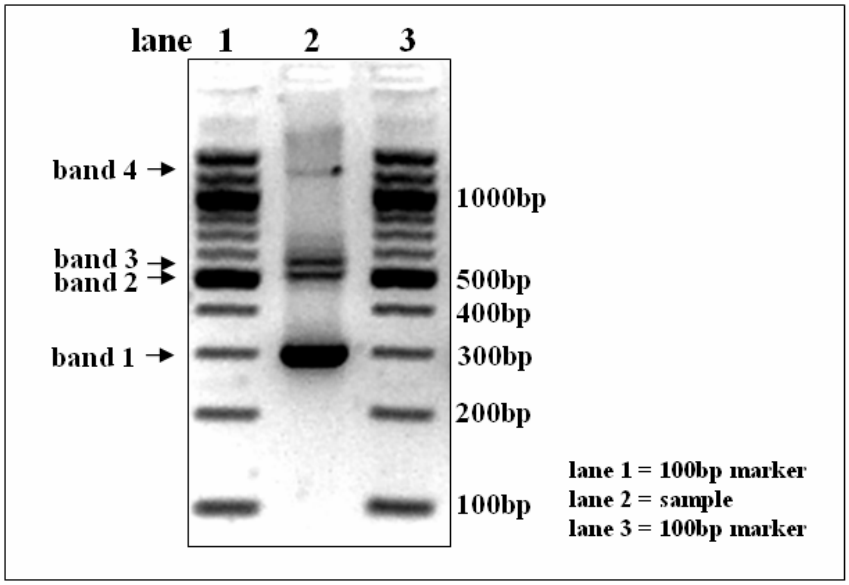

figure 47: PCR products using the primers STAT6_282fwd and STAT6_933rev separated on a 3% agarose gel. Band 1 represents STAT6, band 2 STAT6d and band 3 STAT6e verified by re-sequencing these PCR fragments. Band 4 is yet not identified but corresponds putatively to unspecific amplification unrelated to STAT6.

To further confirm the existence of these three *STAT6* splice variants, more specific primers (table 32) were applied to differentiate between the three isoforms in quantitative Real-Time PCR (figure 48, A and B). The amplification of the overall STAT6 (non-STAT6d, non-STAT6e) was obtained by using the forward primer partially located in exon 17 and exon 18 and the reverse primer in exon 19 (figure 49). Thus, the additional detection of STAT6d and STAT6e was excluded. The specificity for STAT6d was subsequently achieved with a forward primer binding to intron 17. Accordingly, the amplification of STAT6e was guaranteed with the reverse primer being located in intron 18 in the *STAT6* gene.

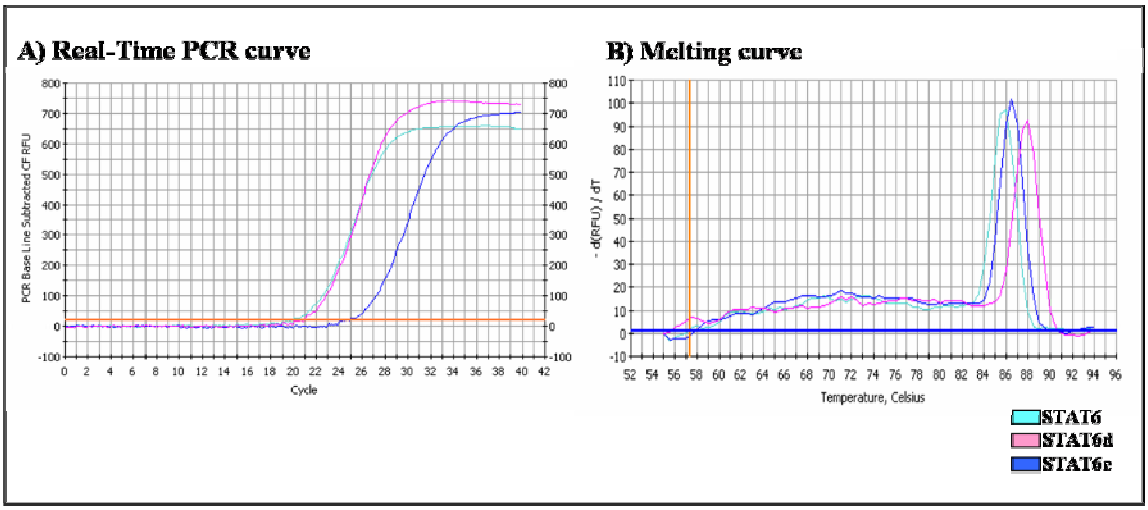

figure 48: A) Real-Time PCR curve of one sample using the iCycler IQ™ Detection System (Biorad, Hercules, USA) for the splice variants overall STAT6 and STAT6d + e. B) Melting curve of the same sample for the splice variants STAT6 and STAT6d + e.

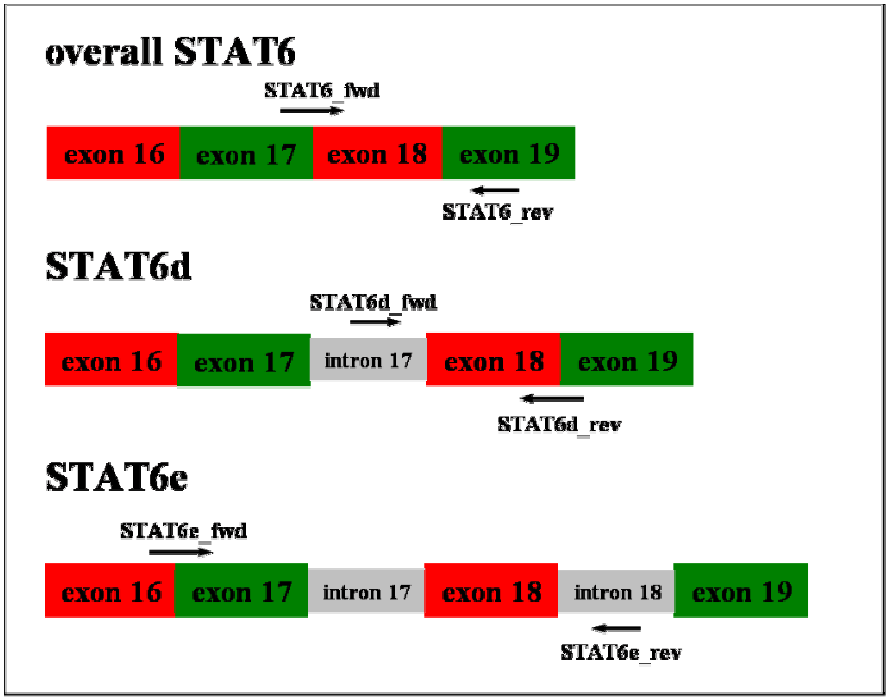

figure 49: Description of the observed STAT6 splice variants: Overall STAT6 mRNA contains all isoforms except STATd and STAT6e. STAT6d and STAT6e are two novel identified splice variants. STAT6d consists of the displayed exons as well as intron 17 whereas STAT6e contains intron 17 and 18 (marked in grey) in addition to the exons (red and green).

For the amplification of each *STAT6* isoform not only cDNA was used as template but also genomic DNA and total RNA as control templates to give supplementary evidence of the specificity of the respective primers as well as the PCR conditions.

3.6.2 The gene expression of specific *STAT6* splice variants is increased in carriers of the polymorphic allele of SNP C2892T

STAT6 is a key intracellular mediator in the IL-4/IL-13 pathway which influences IgE regulation and the development of asthma. It was assumed that elevated levels of *STAT6* mRNA may lead to increased STAT6 protein expression making STAT6 more readily available in the cell. Consequently, more *STAT6* may upregulate the downstream activation of the IL-4/IL-13 pathway. However, no study on *STAT6* gene expression has so far been performed. The characterization of different *STAT6* mRNA isoforms raised the question whether in addition to the known *STAT6* splice variant other isoforms may contribute to STAT6 protein regulation and function. Furthermore, it was hypothesized that the relative mRNA expression of different splice variants may be modulated by the genotype of the polymorphism C2892T. To test this hypothesis, participants of the PARSIFAL population of whom both genomic DNA and RNA (isolated from unstimulated PBMCs) were accessible, were used to assess if gene expression of different *STAT6* isoforms was influenced by the polymorphism C2892T *ex vivo*.

The gene expression of three *STAT6* (overall STAT6, STAT6d, STAT6e) isoforms was measured in 324 individuals (see section 2.2.1.2.2) applying Real-Time PCR (TaqMan®, Perkin-Elmer-Applied-Biosystems, see section 2.2.14.5) whereas the genotype of C2892T located in intron 2 of the *STAT6* gene was determined by RFLP (see section 2.2.8.4, N = 311). In total, for 239 individuals both RNA expression and SNP information was measured successfully. These 239 samples were now analyzed to compare *STAT6* variant expression levels depending on C2892T.

The relative expression of the overall STAT6 isoform normalized for the endogenous control 18SrRNA was slightly elevated in homozygous carriers of the C allele compared to TT individuals (GMR = 1.11, 95% CI 0.83 - 1.47, $p > 0.1$, figure 50), but this difference did not reach statistical significance. In contrast, there was a significant correlation between geometric mean ratios (GMR) of

gene expression of the novel isoform STAT6d (GMR = 1.37, 95% CI 1.06 - 1.78, 0.05 > p ≥ 0.01, figure 50) and homozygous carriers of the polymorphic T allele with almost 40% elevated STAT6d in comparisons to wildtype individuals. A similar effect was observed for STAT6e: the homozygous T allele led in average to 1.49 times higher STAT6e expression (GMR = 1.49, 95%CI 1.15 - 1.94, p < 0.01, figure 50). These results remained consistent after adjusting for confounder effects such as farming status, sex and age. Overall, TT individuals showed increased expression of all three *STAT6* isoforms even though significant differences were achieved only for STAT6d and STAT6e.

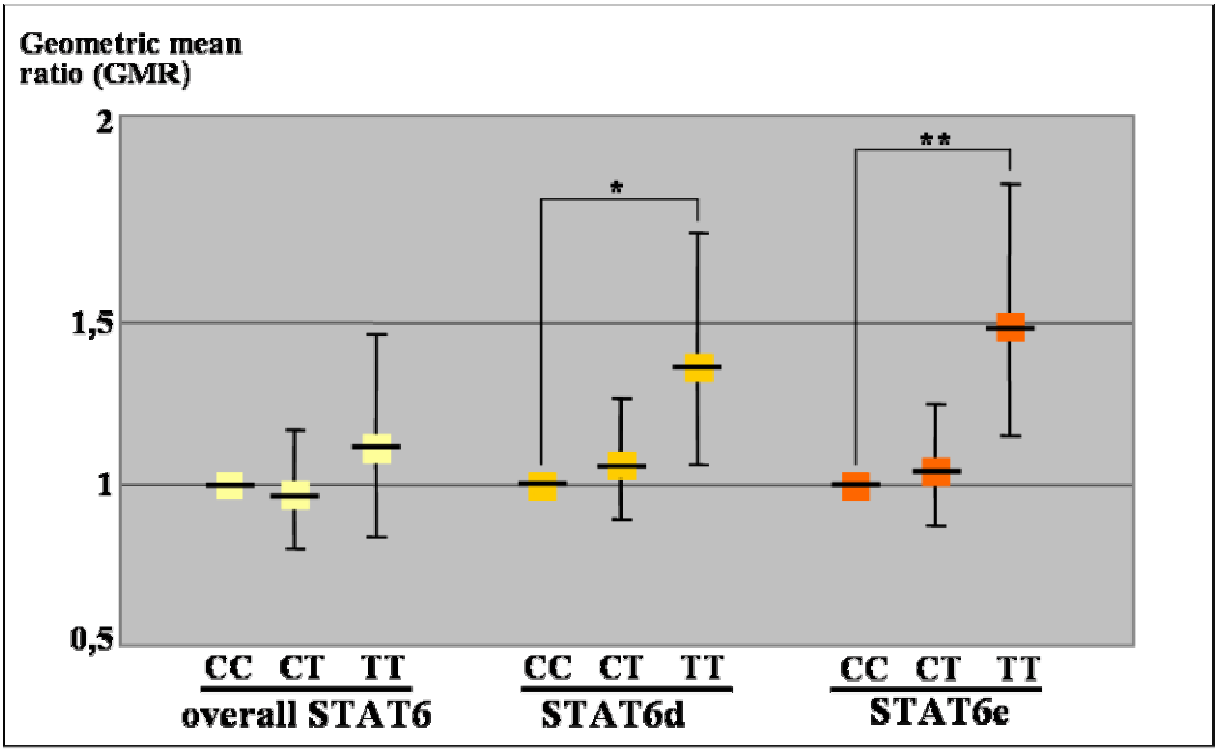

figure 50: Association between gene expression of three splice variants (overall STAT6, STAT6d, STAT6e) and SNP C2892T in intron 2 of the *STAT6* gene displayed in geometric mean ratios with the 95% CI. STAT6 gene expression was normalized for 18SrRNA (N = 239)
*0.05>p≥0.01, **p<0.01

3.6.3 Farming environment modulates gene expression of STAT6 isoforms

Because it was consistently observed that being exposed to farming environment in childhood was significantly protective against the development of allergic diseases in childhood [10,11], we hypothesized that these environmental stimuli present on farms also lead to alterations in *STAT6* gene expression. Previously, it was shown that farming children were exposed to higher levels of endotoxin and other bacterial components [160]. However, a relation to *STAT6* activation through these bacterial molecules has yet not been described. Nevertheless, various studies supported the hypothesis that LPS and other bacterial components affect the levels of IgE and may therefore interfere with the development of atopy [160,161]. The PARSIFAL study population (N = 324, 2.2.1.2.2), in which we had measured *STAT6* isoforms, enabled us to address this question. It allowed to investigate the relation of environmental influences of children growing up on a farm in comparison to reference children (non-farming children) and the putative effect of farm life on *STAT6* isoform (overall STAT6, STAT6d, STAT6e)expression levels.

For the overall *STAT6* isoform it was shown that children exposed to farming environment were significantly associated to express in average 0.85 times (15%) less overall STAT6 mRNA (GMR = 0.85, 95% CI 0.73 -0.99, $0.05 > p \geq 0.01$) in comparison to reference children (figure 51). Quite the opposite effect was observed for STAT6e as geometric means of mRNA expression were significantly increased by 1.26-fold (GMR = 1.26, 95% CI 1.09 - 1.45, $p < 0.01$, figure 51) among farm children in contrast to children not growing up on a farm. No significant association was observed for mRNA levels of the novel splice variant STAT6d and farming status (GMR = 1.04, 95% CI 0.9 - 1.21, $p > 0.1$, figure 51). Hence, in the population under study a significant gene-environment interaction between *STAT6* and the exposure to farming environment was described. However, the *STAT6* isoforms seemed to be inversely related as overall STAT6 mRNA levels were decreased whereas detectable STAT6e gene expression was elevated among farming children.

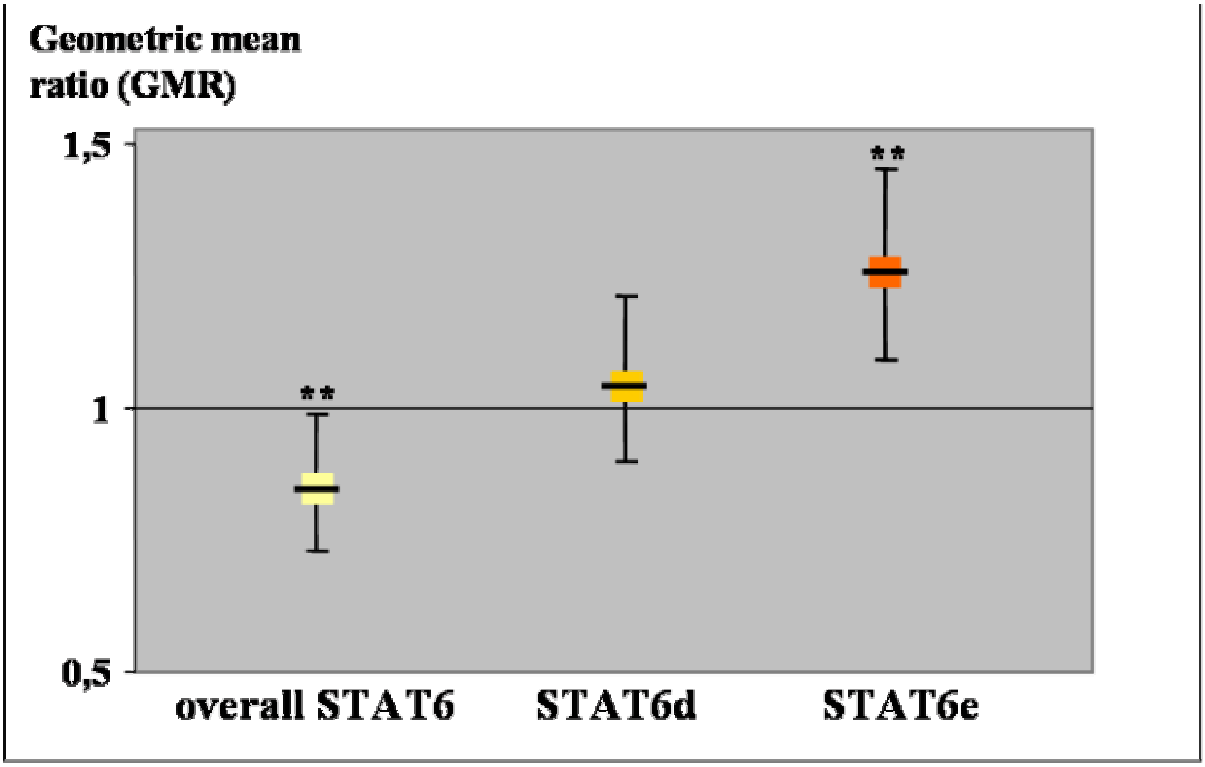

figure 51: Association between gene expression of three splice variants (overall STAT6, STAT6d, STAT6e) and farming status displayed in geometric mean ratios (GMR) with the 95% CI. The GMR of each STAT6 splice variant was compared between children that grew up on a farm in comparison to reference children (non-farming children). *STAT6* gene expression was normalized for 18SrRNA (N = 324), *0.05>p≥0.01, **p<0.01

4 Discussion

The intracellular transcription factor STAT6 is known to be a key mediator in the activation of the IL-4/IL-13 pathway, which is essential for the class-switch from IgM to IgE and affects the development of atopy and atopic diseases in many ways [45-48]. Thus, *STAT6* was identified as a promising target gene for the genetic regulation of immune responses and the development of atopic diseases. Genotyping six *STAT6* polymorphisms in 1,120 German school children (age 9 - 11 years), we observed a significant association between a high risk haplotype and total serum IgE levels. This effect was mainly driven by two polymorphisms located in intron 2 (C2892T) and the 3'UTR (T12888C) of the gene.

Furthermore, gene-by-gene interaction studies were performed to investigate if single SNP (Single Nucleotide Polymorphism) effects of *STAT6* are amplified when polymorphisms in other key genes of the IL-4/IL-13 pathway are concomitantly present. We observed a highly increased risk for elevated total serum IgE levels and the development of asthma when SNP C2892T in *STAT6* was present in combination with other polymorphisms in *IL-4*, *IL-13* and *IL-4Rα*.

Based on its significant associations with IgE and asthma, the role of SNP C2892T on *STAT6* regulation was analyzed in more detail. Thus, the highly conserved region of *STAT6* intron 2 was evaluated by *in vitro* gene expression experiments. Intron 2 was identified as a *cis*-acting silencing element leading to a reduction of *STAT6* promoter activity. However, this effect was strongly dependant upon the allelic status of the polymorphism C2892T as indicated by significantly altered *STAT6* promoter activity. These changes in transcriptional activity of *STAT6* seem to be caused by modulated patterns of transcription factor binding: the T allele at position C2892T induces the binding of NF-κB which is accompanied by the impairment of SP transcription factor affinity at the same region.

Following these *in vitro* results, *STAT6* mRNA expression was investigated on a population level (N = 324, complete genotype and expression data available in 239 individuals). Our work extended the understanding of *STAT6* gene regulation through the identification of one novel *STAT6* splice variant (STAT6d). The expression of two splice-variants (STAT6d and STAT6e), but not of the overall STAT6 cDNA, was significantly increased in carriers of the polymorphic T allele in comparisons to wildtype individuals. Additionally, expression of *STAT6* splice variants (overall STAT6 and STAT6e) seems to be modified by the exposure to farming environment.

Although *in vitro* and *ex vivo* effects of C2892T are intriguing, a second *STAT6* polymorphism, T12888C located in the 3`UTR of the *STAT6* gene also showed a strong association with elevated total serum IgE levels. Therefore, it may additionally contribute to *STAT6* regulation. Due to its location in the 3`UTR of *STAT6* a functional relevance for T12888C may be suggested in the context of mRNA stability rather than contributing to transcriptional activity. As shown previously, a polymorphism in the 3`UTR of the *TNFα* gene may modulate its mRNA stability, translational efficency and subsequently the transport of mRNA in the nucleus [162]. Based on our association signals, we would speculated that in the case of T12888C, mRNA stability may be enhanced in carriers of the T allele, thus raising the amount of *STAT6* mRNA for the translation into STAT6 molecules. Therefore, the effect of the SNP T12888C in the 3`UTR would have to be investigated not only on transcriptional but also on post-transcriptional aspects to better understand the complex mechanisms on *STAT6* regulation. These experiments have not yet been performed and are not part of this thesis. However, they are future challenges and should soon broaden our understanding of *STAT6* regulation.

In our association study of the *STAT6* gene in a large and well phenotyped Caucasian population a significant association between elevated total serum IgE levels and two *STAT6* polymorphisms located in intron 2 (C2892T) and the 3´UTR (T12888C) were shown (table 36). These results were consistent in our study population as the same trends for association were observed separately in both

populations from Munich and Leipzig as well as in the pooled study population. The relation of *STAT6* SNPs to increased levels of serum IgE levels confirmed the initial report of Duetsch *et al*. [124]. This study also demonstrated an association between SNP C2892T, T12888C and IgE regulation in a population of 108 Caucasian asthma sib-pair families (N=474, 216 parents, 258 children) from Germany and Sweden [124]. In addition, the association between polymorphism C2892T in intron 2 of the *STAT6* gene and elevated total serum IgE levels has been replicated in a third independent study population [163]: Weidinger and coworkers in cooperation with our laboratory reported a similar effect between increased levels of total serum IgE levels and the polymorphism C2892T in a cross-sectional cohort study of 1,407 German adults [163]. Thus, the association of SNP C2892T was found independently and repeatedly in three different studies comprising children as well as adults. These data indicate that C2892T may not only be relevant for childhood asthma but may also be involved in the development of atopic diseases in adulthood. It seems unlikely that these results have occurred spuriously due to a type I error, which may always be suspected in a single association study when multiple testing is performed.

Screening the whole gene, Duetsch *et al.* reported 13 polymorphisms in *STAT6* [124]. Out of these 13 polymorpisms we selected those six SNPs representing all SNP linkage disequilibrium (LD) clusters in the *STAT6* gene for our study (tagging SNPs). While these six SNPs capture all genotype information of the gene, genotyping is reduced to the necessary minimum. Using this approach no relevant associations are missed. At the same time, associations observed with these tagging SNPs may potentially be due to any of the polymorphisms represented by the tagging SNP. Thus, we analyzed extended LD of polymorphisms within the *STAT6* gene in LD with SNP C2892T and T12888C. The polymorphism C2892T is in high LD (r^2 = 0.90) with another polymorphism (rs167769) located in intron 2 of the *STAT6* gene approximately 1600bp upstream of C2892T. Due to the strong linkage disequilibrium between both SNPs, it cannot be dissected on a population genetic level which SNPs is responsible for the observed association with elevated IgE levels and asthma. Hence, we also tested if the genetic variant rs167769 may influence *STAT6* function. Performing

EMSA (Electrophoretic Mobility Shift Assay) experiments, no differences in the binding pattern of transcription factors were observed for the allelic variants of SNP rs167769 (figure 37). Thus, we concluded that polymorphism rs167769 may very unlikely be of functional relevance in *STAT6* regulation.

SNP T12888C located in the 3´UTR was not in linkage disequilibrium with any other *STAT6* polymorphism. As linkage disequilibrium decreases even within the *STAT6* gene, but the positive associations results are clearly focused on C2892T and T12888C, it can be assumed that no *STAT6* polymorphisms other than those two are responsible for the observed effects in our study population.

Nevertheless, as cases of extended linkage disequilibrium have been reported [90,164], it was excluded that polymorphisms outside the *STAT6* gene not enclosed in our initial LD analyses on which the SNP selection for genotyping was based may have influenced the observed associations. Thus, LD analyses were extended beyond the *STAT6* locus on chromosme 12 (in total 45kb were included) in this analysis using the HapMap database. This was possible as the HapMap database provides information of common sequence variations across the genome from a population of European descent, also including numerous SNPs within and around the *STAT6* gene. Hence, LD patterns in the region of interest around *STAT6* were characterized.

As *NAB2* (NGF1-A binding protein) is situated unusually close to *STAT6* (figure 52), with their coding regions separated only by 1,852bp [165], *NAB2* was considered to be a likely candidate which may have interfered with association signals from *STAT6*. *STAT6* and *NAB2* are transcribed in an anti-parallel direction with an overlap of 58bp at the 3` ends of their mRNAs. However, based on existing HapMap data, linkage disequilibrium between C2892T, T12888T and polymorphisms in the *NAB2* gene do not explain the associations with SNPs in this study. Therefore, this analysis suggests that no SNPs in the vicinity of the *STAT6* gene may be responsible for our observed association.

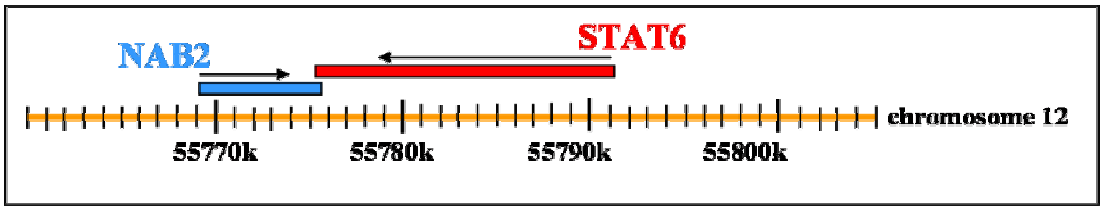

figure 52: Region on chromosome 12 which was included in extended LD analyses (in total 45kb) using the HapMap database(http://www.hapmap.org) containing *STAT6* and *NAB2*. Errors indicate the transcription direction of the respective gene.

In addition to single SNP analyses, associations to common haplotypes (MAF > 0.03) containing the six genotyped *STAT6* polymorphisms were investigated in our study population. While single SNP analysis only captures a snap shot of genetic information at a single base pair position haplotype analysis comprises the complete and transmitted genetic information of a gene locus. As genotype information was not available across different generations, haplotypes had to be estimated based on previously well established algorithms [166]. Haplotype Trend Regression models which result in odds ratio estimates for an association were applied to analyze the data. However, this analysis can only handle categorical but not continuous variables. Thus, IgE values had to be categorized. To achieve this, we used percentiles which were derived from all German children of the cross-sectional population of ISAAC II where IgE levels were available (N = 4,400) tested at the age of 9 - 11 years. These much better resemble the distribution of IgE levels in a population than the widely used arbitrary cut-off of 100 IU/ml.

Performing haplotype analyses, it was shown that two frequent *STAT6* haplotype combinations in the gene have profound effects on IgE regulation (table 37). While haplotype CGTGTT was strongly associated with elevated levels of total serum IgE, haplotype CGCGCT showed the opposite effect decreasing the risk for high IgE levels. Further analyses revealed that these haplotype effects were mainly driven by the presence or absence of the risk alleles at the position of two SNPs: C2892T and T12888C. The risk haplotype (CGTGTT) was also positively associated with elevated levels of total serum IgE in another independent study population of German adults (N = 1,407) [163]. It is not unexpected that haplotype associations are stronger than single SNP analyses, as observed in our study

like in many others. This may be due to the fact that the complete genetic information of a locus is analyzed and background may therefore be reduced.

In a next step, we performed gene-by-gene interaction analyses between major genes in the IL-4/IL-13 pathway. We hypothesized that genetic alterations in these genes may interact with each other and should not be investigated in isolation. Therefore, the single most relevant functional polymorphism of each of these genes was selected and included in this epigenetic analysis as presented in the results section. This was the first systematic interaction analysis of the IL-4/IL-13 pathway also including *STAT6* variants in the model. For the genes *IL-4*, *IL-13* and *IL-4Rα* those SNPs that were previously identified to be functional were chosen [103,107,167]. For *STAT6* we selected C2892T over T12888C as *in silico* analyses predicted C2892T to potentially influence gene expression while prediction for T12888C were less conclusive.

Thus, in extended haplotype analyses, we demonstrated that C2892T may indeed interact with other polymorphisms of the IL-4/IL-13 pathway (comprising *IL-4*, *IL-13* and *IL-4Rα*) and contributes to the genetic control of IgE levels and the development of asthma (figure 23 and figure 24). In contrast, those individuals that did not carry the polymorphism C2892T in the *STAT6* gene or any other analyzed SNP in the IL-4/IL-13 pathway were significantly protected from the development of asthma and less likely to have elevated levels of IgE [168]. Out of the four SNPs tested, *STAT6* C2892T and *IL-13* C-1112T are overrepresented in those extended haplotype combinations conferring the greatest risk to develop elevated IgE and/or asthma. This may indicate that these two polymorphisms are key in the deviation of the IL-4/IL-13 pathway while other SNPs in the IL-4/IL-13 pathway may facilitate a modifier function.

Howard *et al.* also performed gene-be-gene interaction analyses of the IL-4/IL-13 pathway in Dutch families (N = 233) including SNP C-1112T in *IL-13* and S478P in *IL-Rα* as they revealed the strongest association results in their single SNPs analyses [100,169]. A significant gene-by-gene interaction between the *IL-13* promoter SNP C-1112T and S478P in *IL4RA* with asthma susceptibility as well as increased levels of total serum IgE levels were observed. Yet, the interaction of these polymorphisms was stronger for the development of asthma: individuals with the common allele of S478P and the minor allele for C-1112T showed a 5fold increased risk for the development of asthma compared to individuals who did not carry the respective genotypes. The genetic interaction between other IL-4/IL-13 pathway SNPs was further addressed by Heinzmann and coworker in a population of British (N = 150 cases, N = 150 controls) and Japanese (N = 100 cases, N = 100 controls) subjects [94]. In total, they selected four polymorphisms including *IL-4* (C-589T), *IL-13* (Gln110Arg), *IL-4Rα* (Ile50Val, Arg551Gln) and *IL-13Rα1* (A1398G) for gene-by-gene interaction analyses. However, no significant effect between these polymorphisms and the development of asthma and atopy was observed, potentially due to the small sample size.

A combined analysis of SNPs in the IL-4/IL-13 pathway, in contrast to single SNPs, may prove valuable as a predictor of atopy and asthma in children. Even though the prevalence of children carrying SNPs in three or more genes of the pathway is low (2.5-4.2%), the extremely elevated risk to develop asthma (odds ratio = 26.81) in carriers may predict the disease in those individuals. One may argue that the population tested was relatively small for a study of multiple interaction terms. Yet, the effects were strong enough to reach statistical significance. These analyses may help to better understand the complex nature of gene-by-gene interactions in common diseases of complex genetic traits. As IgE switching is fundamental for human immunity and survival, the regulation of IgE is delicate and in part redundant, which may be a common observation in biologically important pathways. Thus, it can be deduced that a defined amount of genetic variance (maybe in combination with certain environmental factors) is necessary in these pathways to overwhelm the system and to

result in a profound, clinically measurable expression of a disease. Nevertheless, statistical interaction alone does not prove biological interaction but only hints to it. How valuable SNPs in the IL-4/IL-13 pathway are as disease predictors and if they could be used as screening tools for the development of asthma in childhood still remains to be tested and further replication studies including numerous SNPs of the IL-4/IL-13 pathway are necessary.

So far, functional studies have been performed for those polymorphisms in the genes *IL-4*, *IL-13* and *IL-4Rα* that we had included in the epigenetic analyses and that showed interaction with *STAT6* C2892T. For *IL-4* C-589T and *IL-13* C-1112T, both located in the proximal promoter region of the gene, transcription factor binding and gene expression was modulated by the presence of the polymorphic allele (*IL-4* -589T and *IL-13* -1112T) [91,103]. The polymorphism A148G in the *IL-4Rα* gene was shown to alter an amino acid in the extracellular part of the receptor (I50V). This modification was demonstrated to increase IL-4 mediated downstream activation of *STAT6* and an enhanced IgE production in B-cells [107]. However, the functional aspects of SNP C2892T in *STAT6* have not yet been investigated until now. In contrast to other polymorphisms in the IL-4/IL-13 pathway, where the putative functional role of the respective SNPs was suggested by their location in the promoter (*IL-4* and *IL-13*) or the change of an amino acid (*IL-4Rα*), the situation with the *STAT6* SNP C2892T seems more complex. This SNP is located in an intronic region and until recently, very little was known about the regulatory function of introns in the genome.

Surprisingly, with the completion of the Human Genome Project it was shown that the human genome contains more than 90% non-coding intronic and intergenetic DNA sequences but only 22,500 annotated genes; much less than other, more primitive species. These observations fueled a debate whether these intronic regions exert biological function and if intronic regions may play a much more prominent role in the progress and differentiation of evolution than previously thought. It would also be hard to understand, why these enormous stretches of sequence would be replicated and actively

kept in the genome if there would be no advantage in them. With more research focusing on understanding the role of introns in the past years it became evident that many functions are attributable to introns. Most of these newly identified roles seem to affect gene regulatory events, showing that non-coding regions can influence alternative splicing and may act as enhancer or silencer elements [170].

Yet, the identification of intronic sequences that direct expression of genes and define their regulatory function in humans remains a significant challenge as the majority of intronic sequences many not have any attributable function. One possibility to distinguish putative functional from irrelevant regions has evolved by the generation of databases allowing for interspecies comparisons of intronic (and exonic) sequences. This approach is mainly used to identify basic, functionally important regulatory regions of the genome (such as promoter, enhance and silencer regions), which are actively conserved throughout the evolution. However, when comparing distant species, novel improvements in gene regulation acquired recently during evolution may be missed [142]. Thus, it had been argued that comparing sequences within apes, primates and humans may better reflect recent advances in gene regulation. For our analyses of *STAT6* intron 2 we applied both approaches: while *STAT6* intron 2 was not strongly conserved between mice and man, intron 2 conservation was observed in dogs, apes, primates and humans. The C2892T SNP could be identified as soon as in primates, potentially indicating an evolutionary advantage of the SNP in highly developed species. Thus, it may be concluded that *STAT6* intron 2 has acquired functional relevance only recently.

Introns may affect gene function in different ways: e.g. by influencing splicing, transcription or 3D-structure of the gene. As in promoters, transcriptional regulation may be controlled by the binding of transcription factors also in intronic regions. Therefore, we performed prediction analysis of transcription factor binding in the intron 2 region harboring C2892T. These *in silico* analyses suggested that the T allele at position 2892 may create a novel binding site for the transcription factor NF-κB. Indeed, the binding of NF-κB was strongly dependant on the polymorphic T allele at position

2892 when performing Electrophoretic Mobility Shift Assays (EMSA, figure 31) as the SNP is located within the matrix of the NF-κB recognition site. In addition, the presence of the T allele reduced the binding of the SP (Specificity Proteins) transcription factor family, which was not predicted by *in silico* analysis, as the polymorphism is situated outside the core sequence of the SP binding site.

Our findings showed profound changes in the distinct binding of NF-κB to *STAT6* intron 2 in a human T-cell line (Jurkat T-cells) as well as in primary $CD4^+$ T-cells, induced by PMA/Ionomycin stimulation. The NF-κB complex binding to intron 2 contained both, p50 and p65 subunits. NF-κB has been shown to increase the gene expression of many cytokines, enzymes and adhesion molecules relevant in chronic inflammatory diseases. The production of IL-1β, TNF-α and IL-6 for example was increased in patients with asthma, rheumatoid arthritis and inflammatory bowel disease possibly by the activation of NF-κB [171]. NF-κB refers to a family consisting of five members: p50, p52, p65 (RelA), RelB and c-Rel [172] and was initially identified as a transcription factor of the expression of the kappa light-chain gene in murin B-lymphocytes [173]. These transcription factors exist in unstimulated cells in an inactive state bound by an inhibitor complex called IκB [174,175]. Only after stimulation, this inhibitor complex is degraded and thus releases the active NF-κB family members. Forming either homo- or heterodimers, they translocate to the nucleus to activate or repress gene transcription after binding to their respective binding site. Members of the NF-κB family have been found in numerous cell types and are known to control many genes crucial for immune and inflammatory responses [172] also relevant for the development of atopic diseases [176].

Previous studies have already described a functional synergism of STAT6 and NF-κB in gene regulation [177-180]. Within the IgE germline promoter recognition sites for STAT6 and NF-κB have been identified in close proximity to each other [177,178,180]. Only the concomitant binding of both transcription factors mediates the optimal activity of the IgE germline gene transcript in response to IL-4 stimulation [177]. Even though the binding sites of STAT6 and NF-κB in the IgE promoter region are not

overlapping, it was suggested that both transcription factors directly interact with each other to form a complex which promotes gene expression [178]. In addition, an interplay between STAT6 and NF-κB was also shown for the *eotaxin* gene which is a potent chemo-attractant for eosinophils, centrally involved in allergic inflammation [179]. As shown by these examples, the cooperation between STAT6 and NF-κB may increase the efficiency of specific gene transcription in settings relevant for the development of allergic diseases. Thus, we hypothesize that the presence of a NF-κB binding site within *STAT6* intron 2 may not be arbitrary but may represent a feedback mechanism by which NF-κB could directly influence the expression of its interaction partner STAT6.

Additionally to NF-κB, we observed the binding of members of the SP transcription factor family to the intronic region of *STAT6* containing C2892T in human T-cells. Similar SP binding patterns (SP1, SP3 and SP4) were detected for the C and the T allele, yet with different affinities (figure 31): a weaker binding affinity of SP proteins was present on the background of the polymorphic T allele.

The SP family consists of four members (SP1-SP4) whereas SP1, SP3 and SP4 are more closely related to each other than to SP2 [181]. Nevertheless, all SP members recognize the same consensus site through their highly conserved DNA-binding domain. SP1 and SP4 act as strong activating domains. In contrast, reports on the transcriptional properties of SP3 appear contradictory [182,183]. It was shown that SP3 can induce gene transcription but when SP1 and SP3 are concomitantly expressed at comparable levels, both transcription factors compete for the same binding site. It is assumed, that the ratio of SP1/SP3 transcription factors changes their ability to act as an activator or repressor [182,183]. That differences in SP binding may indirectly affect the development of allergies has previously been demonstrated for *CD14* [157]. CD14 is an important adapter molecule in the TLR system and a polymorphism (C-159T) in its promoter was associated with elevated soluble CD14 and decreased levels of total serum IgE [161]. The T allele at position -159 strongly impaired the affinity of SP1, SP2 and SP3 to the CD14 promoter which subsequently altered *CD14* gene expression.

At intron 2 position of the C2892T polymorphism in *STAT6* intron 2, SP binding seems to occur in unstimulated and stimulated cells. When the T allele is present NF-κB binding is induced and SP binding decreases. This effect may be due to a displacement of SP binding by NF-κB, as the matrix of these two transcription factors overlap at this locus. Thus, not only NF-κB *de novo* binding but also the alteration of SP binding could add to gene expression regulation of *STAT6*. Also the interaction of NF-κB and SP1 has previously been described for two promoters but not intronic regions: the transcriptional regulation of the FasL was dependant upon the concomitant binding of both transcription factors [184]. Furthermore, the histone acetylation of the monocyte chemoattractant protein 1 (MCP-1) gene was influenced through NF-κB and SP1 [185] suggesting a commonly found interaction between these two transcription factors.

As transcription factor binding may differ significantly between cell lines and primary cells as well as between different cell types we extended our EMSA experiments to primary $CD4^+$ lymphocytes extracted from peripheral blood in our experiments. While getting closer to the real life situation *ex vivo,* we still tried to benefit from the highly controllable experimental setting of EMSA experiments. Our data showed that the specific binding of NF-κB to 2892T was not only present in a T-cell line but was also found in the nuclear environment of primary $CD4^+$ T-cells (figure 32). Again, significant binding of NF-κB was only observed in the presence of the T allele.

Taken together, transcription factor binding analysis *in silico* and *in vitro* indicated that SNP C2892T changes binding properties of transcription factors NF-κB and members of the SP family, known to be associated with allergic inflammation. Next, we addressed the question if and how this may affect gene expression of *STAT6*. Thus, we generated luciferase constructs containing a *STAT6* specific promoter and added a large section of intron 2 containing SNP C2892T in different orientations downstream of the luciferase reporter insert to model the naturally occurring situation as closely as possible. This approach was selected specifically over an experiment in which an unspecific promoter

(e.g. SV40) would have been used to allow for direct interaction between the *STAT6* promoter and intron 2. We also decided to clone a large part of intron 2 in the reporter vector to not miss possible effects of the SNP C2892T occurring only in its natural intron 2 environment. Thus, all putative interaction sites in the respective *STAT6* promoter as well as in *STAT6* intron 2 were transferred into our experimental system.

To be able to perform these experiments we first had to characterize the *STAT6* promoter in human T-cells as all previously *STAT6* promoter studies had been performed in unstimulated NIH3T3 mouse fibroblasts [120]. The same human T-cell line (Jurkat T-cells) previously used for EMSA was selected for our gene expression experiments. In our experiments, Jurkat T-cells were found to constitutively express *STAT6* (data not shown) which is in contrast to previous reports which were unable to detect STAT6 expression in these cell lines [158]. We observed the highest luciferase induction in T-cells with the minimal *STAT6* promoter (pBP86, 1,1kb). These results were in accordance with the previously reported data in mouse fibroblasts [120].

Subsequently, we addressed the question if intron 2 in general may be capable to affect the gene expression of *STAT6* in a regulatory fashion. As indicated by our experiments gene expression was decreased in constructs containing intron 2 (figure 43 and figure 44). Furthermore, this negative effect on gene expression was independent of the intron 2 orientation within the construct, suggesting that intron 2 may act (a) in a *cis*-regulatory fashion and (b) as a silencing element.

In a next step, we now finally could test if the altered transcription factor binding of NF-κB and members of the SP family through C2892T had the ability to modify the repressor function of intron 2, hence altering transcriptional activity. In the presence of the polymorphic T allele of C2892T the activity of the largest *STAT6* promoter (pBP78, 5,6kb) was increased in comparison to the construct corresponding to the wildtpye C allele (figure 44). In contrast, the allelic status of C2892T did not

affect *STAT6* expression of the minimal promoter with the highest baseline activity (pBP86, 1,1kb). An influence of C2892T on *STAT6* activity was only detectable using the largest *STAT6* promoter construct (pBP78, 5,6kb). Thus, it is suggested that the region of intron 2 harboring C2892T may interact with distant *STAT6* promoter regions. At present, it is not fully understood how a *cis*-acting element mediates its regulatory function on its specific promoter. One possibility, suggested by different authors may be that the regulatory intronic region could form a loop to physically interact with its counter part in the promoter region. Other possibilities would be that NF-κB binding to 2892T in intron 2 may alter the 3dimensional structure of the *STAT6* region or that intronic NF-κB binding may influence the epigenetic signature of the *STAT6* locus. However, further studies are needed to investigate the modalities of interaction between the promoter and intron 2 in the *STAT6* gene.

Previous studies had also pointed out that the relative position of a promoter in relation to an intronic element may have a crucial influence on gene regulation. For the high affinity IgE receptor FcεRI it was for example shown that the fourth intron can serve as a repressor element as well as an enhancer depending on the relative position [186]: when intron 4 was cloned downstream of the luciferase gene a decrease in the FcεRI gene expression was demonstrated. Interestingly, when the same region was located upstream of its promoter, enhanced activity was obtained. When we measured *STAT6* gene expression with intron 2 located downstream of its promoter but 3` of the luciferase reporter, *STAT6* promoter activity was completely abrogated suggesting that *STAT6* intron 2 does not possess promoter activity *per se*.

Indisputably, the impact of C2892T on the transcriptional activity of *STAT6* observed in our experiments was fairly subtle. However, nothing else is to be expected from a regulatory SNP involved in a complex pathway and the development of a complex disease. The observed changes are similar to those previously described for other regulatory polymorphisms [103,157]: for *IL-13* C-1112T, which is strongly associated with IgE regulation and asthma in numerous studies on the population

level, molecular studies showed that gene expression varied approximately by 2.5fold in the presence of the polymorphic allele compared to the wildtype allele [103]. Similar observations were also made in functional studies of the *CD14* promoter polymorphism C-159T [157], which is also widely excepted as a candidate gene polymorphism for atopy [161].

As *in vitro* experiments indicated that transcriptional activity of *STAT6* may be affected through the polymorphism C2892T also *ex vivo STAT6* gene expression, assessed by the determination of mRNA levels, was studied on a population level (N = 324). Previously, four functional STAT6 splice variants had been described in humans (STAT6 wildtype, STAT6a, STAT6b, STAT6c): Historically, "STAT6 wildtype" was thought to be one STAT6 isoform, the sole gene product of STAT6. However, as we know now, the primers for STAT6 wildtype capture a mixture of STAT6 mRNAs when applied to human cDNA. "STAT6a" lacks parts of the 3`UTR [120], "STAT6b" is missing exon 4 and isoform "STAT6c" lacks 84bp of exon 16 [120,187]. It was shown that STAT6b and STAT6c were expressed in numerous human tissues at different levels and ratios [187]. The presence of the T allele of polymorphism C2892T increased overall STAT6 mRNA expression, but this trend was not statistically significant. This may have been due to the concomitant measurement of a mixture of further STAT6 mRNA isoforms (STAT6 wildtype, a, b and c) in our study population. Thus, a more profound SNP effect on a specific isoform could have been missed.

However, while performing these studies the identification of a new *STAT6* splice variant comprising intron 17 and 18 was reported [159]. Therefore, we investigated if polymorphism C2892T may affect the expression of this specific isoform (STAT6e). In addition, when studying this splice variant we detected the presence of another isoform (STAT6d) which comprises intron 17 but not intron 18. The expression levels of both novel STAT6 isoforms (STAT6d and STAT6e) were significantly elevated in homozygous carriers of the T allele of C2892T (figure 50).

However, one may argue that using Peripheral Blood Mononuclear Cells (PBMCs) is not sufficient to elucidate the influence of STAT6 (mainly acting on T- and B-cells) as various cell types are included. Yet, differences in mRNA expression for STAT6d and STAT6e, but not overall STAT6 (non-STAT6d and -STAT6e) were strong enough to reach statistical significance. Further analyses are needed to differentiate the role of C2892T on other splice variants of the *STAT6* gene specifically.

When analyzing *STAT6* gene expression in the PARSIFAL population which consists of children raised on farms and controls not exposed to farm life, we made a further interesting observation. *STAT6* isoform expression did not only depend on the C2892T genotype but also on the farming status of the individuals. The overall STAT6 mRNA levels (non-STAT6d and non-STAT6e) were significantly decreased while isoform STAT6e was significantly increased in farming children (figure 51). Thus, one may speculate that increased microbial exposure as observed on farms [130] may influence STAT6 regulation and consequently the IL-4/IL-13 pathway. This mechanism may contribute to the protective effect of farm life against the development of atopic diseases.

This hypothesis may be supported indirectly by recent studies indicating a relation between LPS exposure, levels of IgE [160,161] and *STAT6* isoform expression. First evidence on the induction of STAT6 by LPS stimulation was provided by Tang *et al.* [159]. Even though this work erroneously described the identification of a novel gene (*STAT6B*) homologous to members of the STAT family [159] we suggest that *STAT6B* represents the *STAT6* isoform STAT6e. In contrast to the known *STAT6* isoforms, *STAT6B* expression seemed to be sensitive to stimulation with LPS [159]. Based on these data, we hypothesized that the increased levels of STAT6e mRNA in our farming population could be due to specific stimulation of their immune system with components like LPS, present in abundance on farms.

It can also be speculated that the increased activation of the splice variants STAT6d and STAT6e in the presence of the polymorphic T allele may lead to a negative regulation of STAT6. A similar effect has already been described for the isoform STAT6c leading to a truncation of the SH2 domain (coded by exon 16) [187] which is necessary to bind phosphorylated tyrosine residues relevant for protein-protein interaction. It was shown previously that the expression of STAT6c suppresses the dimer formation of endogenous STAT6 and induces a decline in DNA-synthesis, tyrosine phosphorylation and the down-regulation of antigen presenting receptors (MHCII and FcγR1). Both splice variants STAT6d and STAT6e also originated from a region relevant for the functional SH2 domain in the *STAT6* gene. Whether these two splice variants also influence overall STAT6 mRNA levels by a similar effect is not yet known. Therefore, further investigations are needed to analyze if the presence of different *STAT6* mRNA isoforms may be efficient to modulate *STAT6* regulation or if these *STAT6* splice variants need to be translated into functionally active proteins to modulate STAT6 function.

As the analyzed *STAT6* splice variants were influenced by the polymorphism C2892T as well as certain environmental factors it may be suggested that the observations presented here may have implications of practical clinical relevance: the specific mRNA measurement of *STAT6* isoforms could be used as a diagnostic approach to predict the development of atopic disorders. The control of transcriptional activity of the *STAT6* gene may even be used as a new therapeutic approach. The applications of a respective phosphorylated *cis*-element decoy for the *STAT6* gene has already been tested in mouse models which diminished the expression and production of IL-4 [188]. Additionally, the proliferation of murine T_H2-cells and primary $CD4^+$ T-cells was abrogated. However, as the STAT6 transcription factor plays a central role in many other immune regulatory pathways caution is necessary before next steps are taken. At present it is difficult to imagine to transfer the gained knowledge on *STAT6* regulation into the design of therapeutics blocking STAT6 without causing unwanted side-effects on other cells and/or mechanisms.

Our results suggest that polymorphism C2892T, located within an intronic region of *STAT6*, influences the regulation of the *STAT6* gene and increases total serum IgE levels and the risk for asthma. As NF-κB and STAT6 are key players in many inflammatory mechanisms[178], it may be speculated that *de novo* NF-κB binding in this particular regulatory region of the *STAT6* intron 2 may amplify different inflammatory responses. While it clearly increases the risk for an allergic immune response this may be beneficial in some circumstances, such as the defense against parasite infections. Further studies, e.g. in populations at high risk for parasitic and infections, will need to address the question if this hypothesis is true and if the polymorphic allele of SNP C2892T is advantageous in these diseases and environments.

Additionally, STAT6 is not only involved in inflammation such as atopic diseases and helminthes infections but it also regulates a broad range of other cellular processes (as described in detail in section 1.4.1). Thus, STAT6 was shown to play a critical role in the proliferation of fibroblast and vascular smooth muscle cells [189,190]. In addition, the cytokine-stimulated proliferation of lymphocytes seems to be regulated by *STAT6* partially due to the control of cell cycle proteins like the cell cycle–dependent kinase inhibitor $p27^{kip}$ through *STAT6* dependent pathways [73]. It was shown that *STAT6* deficient T-cells had an impaired proliferative response to IL-4 stimulation because of modulated levels of $p27^{kip}$. Furthermore, the constituative expression of *STAT6* has been observed in prostate cancer tissue, Hogkin lymphomas, primary mediastinal large B-cell lymphomas, cutaneous T-cell lymphomas and adult T-cell leukemia/lymphomas [191]. The mechanisms leading to the constitutive expression of *STAT6* in some of these tumor tissues are still unknown but it could be speculated, based on our functional experiments that SNP C2892T may also influence *STAT6* expression in these circumstances. Which C2892T allele may be beneficial in the situation of tumor related *STAT6* expression, is purely speculative at this point but may be an interesting question for further studies.

Some *STAT6* SNPs have already been studied for there role in cancer treatment responses [192] as *STAT6* was shown to cause an increase in *μ-opioid receptor* gene expression [193]. Hence, *STAT6* polymorphisms (and SNPs in the genes *μ-opioid receptor, β-arrestin2 and diphosphate-glucuronysyltrans- ferease)* was studied in cancer patients achieving good analgesia by morphine (= controls) and those who did not tolerate morphine but required alternative opioids (= switchers) for pain control. In this study eight *STAT6* polymorphisms were analyzed. A significant association was observed for SNP rs167769 (C-1714T), the polymorphism we found to be in high linkage disequilibrium with C2892T but without functional implication by itself. The frequency of the polymorphic allele was significantly increased in the group of switchers compared to the control group. These data may indicate that not only *STAT6*, but even the very same SNP found to increase IgE expression and asthma, may influence opioid receptor tolerance, a mechanism completely unrelated to allergy.

This broad involvement of STAT6 in many mechanisms and pathways shows its importance but on the other hand, limits its use or application in any therapeutic approach. While our findings may suggest that *STAT6* polymorphisms, together with other SNPs from the IL-4/IL-13 pathway, may be used to identify individuals at high risk to develop asthma and allergy in a diagnostic approach in the near future, targeting STAT6 for therapeutic approaches seems to be a dangerous endeavour. Too many mechanisms and pathways depend on it and blocking or enhancing its function outside a very specific situation or cell may disturb many different and unexpected systems.

Indisputably, many further mechanisms may contribute to the development of such complex disorders as asthma and atopy. These mechanisms may most likely operate through gene-by-gene as well as gene-by-environment interactions. The work presented here indicates how crucial synergistic effects are in a regulatory pathway. C2892T should not be seen in isolation but needs to be seen in the right perspective within the IL-4/IL-13 pathway, together with other SNPs [103,107,119,167]. While we have so far

shown gene-by-gene interaction on the population level, this work has investigated *STAT6* SNP function in isolation. It is conceivable that SNPs within IL-4/IL-13 pathway genes may interact not only in a purely additive way, independent of each other but that also biological interactions may occur. However, before these comprehensive functional analyses may be attempted in the future, the functional properties of each of the associated SNPs need to be dissected one by one. The work presented here has set the ground for these kinds of studies in the IL-4/IL-13 pathway together with others [94,103]. Only then we will be able to fully understand the mechanisms by which IL-4/IL-13 pathway SNPs are capable to influence the development of asthma and allergy in such a strong way.

5 Summary

Atopic diseases like asthma and allergies are characterized by an IgE mediated immune response against common allergens. *STAT6* (Signal Transducer and Activator of Transcription 6) is a key mediator in the activation of the IL-4/IL-13 pathway inducing the class-switch to IgE. Thus, genetic variations within *STAT6* may influence T_H2 mediated diseases like asthma and atopy. Aim of the present work was to identify *STAT6* polymorphisms which may influence the predisposition for atopic diseases in children. Furthermore, the functional influences of putative causal *STAT6* SNPs relevant for associations with IgE regulation, atopy and asthma development were investigated.

The present work indicates that one *STAT6* haplotype significantly increased the risk for elevated IgE levels mainly driven by two genetic variations located in intron 2 (C2892T) and the 3`UTR (T12888C). Furthermore, we demonstrated that C2892T strongly contributes to the genetic control of IgE and the development of asthma in combination with polymorphisms of the IL-4/IL-13 pathway (*IL-4*, *IL-13* and *IL-4Rα*). *In vitro* experiments revealed that intron 2 reduced the *STAT6* expression, acting as a silencer regulatory element. However, this repressor function was strongly dependent on SNP C2892T: the existence of the polymorphic T allele increased *STAT6* gene expression significantly compared to the wildtype C allele. In addition, *ex vivo STAT6* expression was affected by the polymorphism C2892T. These effects seem to be due to altered DNA/protein interaction as the binding of the transcription factor NF-κB was only detectable in the presence of the T allele in a human T-cell line as well as in $CD4^+$ primary T-cells.

These data contribute to a better understanding of *STAT6* regulation *per se*. Additionally, the functional relevance of a genetic variation within *STAT6* intron 2 was demonstrated. These results give new insight into the complex and subtle mechanisms leading to the development of allergic diseases.

6 Zusammenfassung

Asthma und andere atopische Erkrankungen sind duch eine IgE-vermittelte Immunreaktion gegen Allergene gekennzeichnet, wobei die Zytokine IL-4 und IL-13 die Produktion von IgE maßgeblich steuern. Der intrazelluläre Transkriptionsfaktor STAT6 (Signal Transducer and Activator of Transcription 6) spielt bei der IL-4/IL-13 Signalübertragung eine entscheidende Rolle. Genetische Varianten im *STAT6* Gen könnten daher die Entstehung von Asthma und Atopie beeinflussen. Ziel dieser Studie war es, diese genetischen Varianten in *STAT6* zu identifizieren und ihre funktionelle Rolle bei der Entwicklung von Asthma und Atopie zu charakterisieren.

Unsere Assoziationsstudien zeigten, dass ein *STAT6* Haplotyp den Serum IgE Spiegel signifikant erhöht, wofür hauptsächlich zwei Polymorphismen verantwortlich waren (C2892T in Intron 2 und T12888C in der 3'UTR). Zusammen mit anderen Polymorphismen der IL-4/IL-13 Signalkaskade (*IL-4*, *IL-13, IL-4Rα*) beeinflusst der SNP C2892T auch die Entstehung von Asthma. In funktionellen Studien konnte Intron 2 des *STAT6* Gens als ein *cis*-regulatorisches Element identifiziert werden, wobei die Repressorfunktion von Intron 2 durch den SNP C2892T wesentlich beeinflusst wurde: die Expression des polymorphen T Alleles erhöhte die *STAT6* Expression signifikant im Vergleich zum Wildtypallel in Zelllinien, aber auch in einer *ex vivo* Genexpressionsstudie. Für diese Effekte scheinen Änderungen im Bindungsverhalten des Transkriptionsfaktors NF-κB im Intron 2 verantwortlich zu sein, welches ebenfalls vom Genotyp des SNP C2892T abhängig ist.

Die vorliegende Arbeit trägt somit zu einem besseren Verständnis der *STAT6* Regulation im Allgemeinen bei. Die funktionelle Relevanz einer für die IgE Regulation und das Auftreten von Asthma wichtigen genetischen Variante im *STAT6* Gen wurde gezeigt. Diese Ergebnisse liefern neue Einblicke in die komplexen und subtilen Mechanismen, die zur Entstehung allergischer Erkrankungen führen.

7 References

1. de Marco, R. et al. Prognostic factors of asthma severity: a 9-year international prospective cohort study. *J Allergy Clin Immunol* **117**, 1249-56 (2006).
2. Cookson, W. O. & Palmer, L. J. Investigating the asthma phenotype. *Clin Exp Allergy* **28 Suppl 1**, 88-9; discussion 108-10 (1998).
3. Burrows, B., Martinez, F. D., Halonen, M., Barbee, R. A. & Cline, M. G. Association of asthma with serum IgE levels and skin-test reactivity to allergens. *N Engl J Med* **320**, 271-7 (1989).
4. Peat, J. K. et al. Serum IgE levels, atopy, and asthma in young adults: results from a longitudinal cohort study. *Allergy* **51**, 804-10 (1996).
5. Beeh, K. M., Ksoll, M. & Buhl, R. Elevation of total serum immunoglobulin E is associated with asthma in nonallergic individuals. *Eur Respir J* **16**, 609-14 (2000).
6. Feijen, M., Gerritsen, J. & Postma, D. S. Genetics of allergic disease. *Br Med Bull* **56**, 894-907 (2000).
7. Eder, W., Ege, M. J. & von Mutius, E. The asthma epidemic. *N Engl J Med* **355**, 2226-35 (2006).
8. von Mutius, E. et al. Prevalence of asthma and atopy in two areas of West and East Germany. *Am J Respir Crit Care Med* **149**, 358-64 (1994).
9. Weiland, S. K. et al. Prevalence of respiratory and atopic disorders among children in the East and West of Germany five years after unification. *Eur Respir J*, 862-70 (1999).
10. Braun-Fahrlander, C. et al. Prevalence of hay fever and allergic sensitization in farmer's children and their peers living in the same rural community. SCARPOL team. Swiss Study on Childhood Allergy and Respiratory Symptoms with Respect to Air Pollution. *Clin Exp Allergy* **29**, 28-34 (1999).
11. Von Ehrenstein, O. S. et al. Reduced risk of hay fever and asthma among children of farmers. *Clin Exp Allergy* **30**, 187-93 (2000).
12. Riedler, J., Eder, W., Oberfeld, G. & Schreuer, M. Austrian children living on a farm have less hay fever, asthma and allergic sensitization. *Clin Exp Allergy* **30**, 194-200 (2000).
13. Laitinen, T. et al. Characterization of a common susceptibility locus for asthma-related traits. *Science* **304**, 300-4 (2004).
14. Van Eerdewegh, P. et al. Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. *Nature* (2002).
15. Akira, S., Takeda, K. & Kaisho, T. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nat Immunol* **2**, 675-80 (2001).
16. Eder, W. et al. Toll-like receptor 2 as a major gene for asthma in children of European farmers. *J Allergy Clin Immunol* **113**, 482-8 (2004).
17. Glimcher, L. H. & Murphy, K. M. Lineage commitment in the immune system: the T helper lymphocyte grows up. *Genes Dev* **14**, 1693-711 (2000).
18. Romagnani, S. Lymphokine production by human T cells in disease states. *Annu Rev Immunol* **12**, 227-57 (1994).
19. Abbas, A. K., Murphy, K. M. & Sher, A. Functional diversity of helper T lymphocytes. *Nature* **383**, 787-93 (1996).
20. Rengarajan, J., Szabo, S. J. & Glimcher, L. H. Transcriptional regulation of Th1/Th2 polarization. *Immunol Today* **21**, 479-83 (2000).
21. Szabo, S. J. et al. A novel transcription factor, T-bet, directs Th1 lineage commitment. *Cell* **100**, 655-69 (2000).
22. Yoshida, A., Koide, Y., Uchijima, M. & Yoshida, T. O. IFN-gamma induces IL-12 mRNA expression by a murine macrophage cell line, J774. *Biochem Biophys Res Commun* **198**, 857-61 (1994).

23. Doherty, T. M., Seder, R. A. & Sher, A. Induction and regulation of IL-15 expression in murine macrophages. *J Immunol* **156**, 735-41 (1996).
24. Hayes, M. P., Freeman, S. L. & Donnelly, R. P. IFN-gamma priming of monocytes enhances LPS-induced TNF production by augmenting both transcription and MRNA stability. *Cytokine* **7**, 427-35 (1995).
25. Rao, A. & Avni, O. Molecular aspects of T-cell differentiation. *Br Med Bull* **56**, 969-84 (2000).
26. Lack, G. et al. Nebulized IFN-gamma inhibits the development of secondary allergic responses in mice. *J Immunol* **157**, 1432-9 (1996).
27. Parronchi, P. et al. IL-4 and IFN (alpha and gamma) exert opposite regulatory effects on the development of cytolytic potential by Th1 or Th2 human T cell clones. *J Immunol* **149**, 2977-83 (1992).
28. Skapenko, A., Niedobitek, G. U., Kalden, J. R., Lipsky, P. E. & Schulze-Koops, H. Generation and regulation of human Th1-biased immune responses in vivo: a critical role for IL-4 and IL-10. *J Immunol* **172**, 6427-34 (2004).
29. Scott, P. IFN-gamma modulates the early development of Th1 and Th2 responses in a murine model of cutaneous leishmaniasis. *J Immunol* **147**, 3149-55 (1991).
30. Murphy, E. et al. Reversibility of T helper 1 and 2 populations is lost after long-term stimulation. *J Exp Med* **183**, 901-13 (1996).
31. Mullen, A. C. et al. Hlx is induced by and genetically interacts with T-bet to promote heritable T(H)1 gene induction. *Nat Immunol* **3**, 652-8 (2002).
32. Szabo, S. J. et al. Distinct effects of T-bet in TH1 lineage commitment and IFN-gamma production in CD4 and CD8 T cells. *Science* **295**, 338-42 (2002).
33. Romagnani, S. Regulatory T cells: which role in the pathogenesis and treatment of allergic disorders? *Allergy* **61**, 3-14 (2006).
34. Stock, P., DeKruyff, R. H. & Umetsu, D. T. Inhibition of the allergic response by regulatory T cells. *Curr Opin Allergy Clin Immunol* **6**, 12-6 (2006).
35. Akbari, O., Stock, P., DeKruyff, R. H. & Umetsu, D. T. Role of regulatory T cells in allergy and asthma. *Curr Opin Immunol* **15**, 627-33 (2003).
36. Romagnani, S. The Th1/Th2 paradigm. *Immunol Today* **18**, 263-6 (1997).
37. Moller, D. R. et al. Enhanced expression of IL-12 associated with Th1 cytokine profiles in active pulmonary sarcoidosis. *J Immunol* **156**, 4952-60 (1996).
38. Romagnani, P., Annunziato, F., Baccari, M. C. & Parronchi, P. T cells and cytokines in Crohn's disease. *Curr Opin Immunol* **9**, 793-9 (1997).
39. Robinson, D. S. et al. Predominant TH2-like bronchoalveolar T-lymphocyte population in atopic asthma. *N Engl J Med* **326**, 298-304 (1992).
40. Effros, R. M. & Nagaraj, H. Asthma: new developments concerning immune mechanisms, diagnosis and treatment. *Curr Opin Pulm Med* **13**, 37-43 (2007).
41. Oettgen, H. C. & Geha, R. S. IgE in asthma and atopy: cellular and molecular connections. *J Clin Invest* **104**, 829-35 (1999).
42. Jiang, H., Harris, M. B. & Rothman, P. IL-4/IL-13 signaling beyond JAK/STAT. *J Allergy Clin Immunol* **105**, 1063-70 (2000).
43. Corry, D. B. & Kheradmand, F. Induction and regulation of the IgE response. *Nature* **402**, B18-23 (1999).
44. Hilton, D. J. et al. Cloning and characterization of a binding subunit of the interleukin 13 receptor that is also a component of the interleukin 4 receptor. *Proc Natl Acad Sci U S A* **93**, 497-501 (1996).
45. Minty, A. et al. Interleukin-13 is a new human lymphokine regulating inflammatory and immune responses. *Nature* **362**, 248-50 (1993).
46. Akbari, O. et al. Essential role of NKT cells producing IL-4 and IL-13 in the development of allergen-induced airway hyperreactivity. *Nat Med* **9**, 582-8 (2003).
47. Chen, L. et al. IL-4 induces differentiation and expansion of Th2 cytokine-producing eosinophils. *J Immunol* **172**, 2059-66 (2004).

48. Schmid-Grendelmeier, P. et al. Eosinophils express functional IL-13 in eosinophilic inflammatory diseases. *J Immunol* **169**, 1021-7 (2002).
49. Johansson-Lindbom, B. & Borrebaeck, C. A. Germinal center B cells constitute a predominant physiological source of IL-4: implication for Th2 development in vivo. *J Immunol* **168**, 3165-72 (2002).
50. Dolganov, G. et al. Coexpression of the interleukin-13 and interleukin-4 genes correlates with their physical linkage in the cytokine gene cluster on human chromosome 5q23-31. *Blood* **87**, 3316-26 (1996).
51. Kopf, M. et al. Disruption of the murine IL-4 gene blocks Th2 cytokine responses. *Nature* **362**, 245-8 (1993).
52. McKenzie, G. J. et al. Impaired development of Th2 cells in IL-13-deficient mice. *Immunity* **9**, 423-32 (1998).
53. McKenzie, G. J., Fallon, P. G., Emson, C. L., Grencis, R. K. & McKenzie, A. N. Simultaneous disruption of interleukin (IL)-4 and IL-13 defines individual roles in T helper cell type 2-mediated responses. *J Exp Med* **189**, 1565-72 (1999).
54. Brusselle, G., Kips, J., Joos, G., Bluethmann, H. & Pauwels, R. Allergen-induced airway inflammation and bronchial responsiveness in wild-type and interleukin-4-deficient mice. *Am J Respir Cell Mol Biol* **12**, 254-9 (1995).
55. Coyle, A. J. et al. Interleukin-4 is required for the induction of lung Th2 mucosal immunity. *Am J Respir Cell Mol Biol* **13**, 54-9 (1995).
56. Walter, D. M. et al. Critical role for IL-13 in the development of allergen-induced airway hyperreactivity. *J Immunol* **167**, 4668-75 (2001).
57. Wills-Karp, M. et al. Interleukin-13: central mediator of allergic asthma. *Science* **282**, 2258-61 (1998).
58. Chatila, T. A. Interleukin-4 receptor signaling pathways in asthma pathogenesis. *Trends Mol Med* **10**, 493-9 (2004).
59. Kotanides, H. & Reich, N. C. Interleukin-4-induced STAT6 recognizes and activates a target site in the promoter of the interleukin-4 receptor gene. *J Biol Chem* **271**, 25555-61 (1996).
60. Ryan, J. J. et al. Growth and gene expression are predominantly controlled by distinct regions of the human IL-4 receptor. *Immunity* **4**, 123-32 (1996).
61. Noben-Trauth, N. et al. An interleukin 4 (IL-4)-independent pathway for CD4+ T cell IL-4 production is revealed in IL-4 receptor-deficient mice. *Proc Natl Acad Sci U S A* **94**, 10838-43 (1997).
62. Cohn, L. et al. Th2-induced airway mucus production is dependent on IL-4Ralpha, but not on eosinophils. *J Immunol* **162**, 6178-83 (1999).
63. Kelly-Welch, A. E. et al. Complex role of the IL-4 receptor alpha in a murine model of airway inflammation: expression of the IL-4 receptor alpha on nonlymphoid cells of bone marrow origin contributes to severity of inflammation. *J Immunol* **172**, 4545-55 (2004).
64. Arinobu, Y. et al. Augmentation of signal transducer and activation of transcription (STAT)6 and STAT3 expression in stimulated B and T cells. *Biochem Biophys Res Commun* **277**, 317-24 (2000).
65. Masuda, A., Matsuguchi, T., Yamaki, K., Hayakawa, T. & Yoshikai, Y. Interleukin-15 prevents mouse mast cell apoptosis through STAT6-mediated Bcl-xL expression. *J Biol Chem* **276**, 26107-13 (2001).
66. Bulanova, E. et al. Mast cells express novel functional IL-15 receptor alpha isoforms. *J Immunol* **170**, 5045-55 (2003).
67. Patel, B. K. et al. Stat6 and Jak1 are common elements in platelet-derived growth factor and interleukin-4 signal transduction pathways in NIH 3T3 fibroblasts. *J Biol Chem* **271**, 22175-82 (1996).
68. Kaplan, M. H., Schindler, U., Smiley, S. T. & Grusby, M. J. Stat6 is required for mediating responses to IL-4 and for development of Th2 cells. *Immunity* **4**, 313-9 (1996).
69. Shimoda, K. et al. Lack of IL-4-induced Th2 response and IgE class switching in mice with disrupted Stat6 gene. *Nature* **380**, 630-3 (1996).

70. Takeda, K. et al. Essential role of Stat6 in IL-4 signalling. *Nature* **380**, 627-30 (1996).
71. Kurata, H., Lee, H. J., O'Garra, A. & Arai, N. Ectopic expression of activated Stat6 induces the expression of Th2-specific cytokines and transcription factors in developing Th1 cells. *Immunity* **11**, 677-88 (1999).
72. Zhu, J., Guo, L., Watson, C. J., Hu-Li, J. & Paul, W. E. Stat6 is necessary and sufficient for IL-4's role in Th2 differentiation and cell expansion. *J Immunol* **166**, 7276-81 (2001).
73. Kaplan, M. H., Daniel, C., Schindler, U. & Grusby, M. J. Stat proteins control lymphocyte proliferation by regulating p27Kip1 expression. *Mol Cell Biol* **18**, 1996-2003 (1998).
74. Kuperman, D., Schofield, B., Wills-Karp, M. & Grusby, M. J. Signal transducer and activator of transcription factor 6 (Stat6)-deficient mice are protected from antigen-induced airway hyperresponsiveness and mucus production. *J Exp Med* **187**, 939-48 (1998).
75. Tomkinson, A. et al. The failure of STAT6-deficient mice to develop airway eosinophilia and airway hyperresponsiveness is overcome by interleukin-5. *Am J Respir Crit Care Med* **160**, 1283-91 (1999).
76. Mathew, A. et al. Signal transducer and activator of transcription 6 controls chemokine production and T helper cell type 2 cell trafficking in allergic pulmonary inflammation. *J Exp Med* **193**, 1087-96 (2001).
77. Christodoulopoulos, P. et al. TH2 cytokine-associated transcription factors in atopic and nonatopic asthma: evidence for differential signal transducer and activator of transcription 6 expression. *J Allergy Clin Immunol* **107**, 586-91 (2001).
78. Miller, R. L., Eppinger, T. M., McConnell, D., Cunningham-Rundles, C. & Rothman, P. Analysis of cytokine signaling in patients with extrinsic asthma and hyperimmunoglobulin E. *J Allergy Clin Immunol* **102**, 503-11 (1998).
79. Palmer, L. J. et al. Linkage of chromosome 5q and 11q gene markers to asthma-associated quantitative traits in Australian children. *Am J Respir Crit Care Med* **158**, 1825-30 (1998).
80. Martinez, F. D. et al. Linkage of circulating eosinophils to markers on chromosome 5q. *Am J Respir Crit Care Med* **158**, 1739-44 (1998).
81. Meyers, D. A. et al. Evidence for a locus regulating total serum IgE levels mapping to chromosome 5. *Genomics* **23**, 464-70 (1994).
82. Rosenwasser, L. J. Promoter polymorphism in the candidate genes, IL-4, IL-9, TGF-beta1, for atopy and asthma. *Int Arch Allergy Immunol* **118**, 268-70 (1999).
83. Noguchi, E. et al. Association of asthma and the interleukin-4 promoter gene in Japanese. *Clin Exp Allergy* **28**, 449-53 (1998).
84. Sandford, A. J. et al. Polymorphisms in the IL4, IL4RA, and FCERIB genes and asthma severity. *J Allergy Clin Immunol* **106**, 135-40. (2000).
85. Kawashima, T. et al. Linkage and association of an interleukin 4 gene polymorphism with atopic dermatitis in Japanese families. *J Med Genet* **35**, 502-4 (1998).
86. Doull, I. J. et al. Allelic association of gene markers on chromosomes 5q and 11q with atopy and bronchial hyperresponsiveness. *Am J Respir Crit Care Med* **153**, 1280-4 (1996).
87. Dizier, M. H. et al. Indication of linkage of serum IgE levels to the interleukin-4 gene and exclusion of the contribution of the (-590 C to T) interleukin-4 promoter polymorphism to IgE variation. *Genet Epidemiol* **16**, 84-94 (1999).
88. Liu, X. et al. Associations between total serum IgE levels and the 6 potentially functional variants within the genes IL4, IL13, and IL4RA in German children: the German Multicenter Atopy Study. *J Allergy Clin Immunol* **112**, 382-8 (2003).
89. Walley, A. J. & Cookson, W. O. Investigation of an interleukin-4 promoter polymorphism for associations with asthma and atopy. *J Med Genet* **33**, 689-92 (1996).
90. Kabesch, M. et al. A complete screening of the IL4 gene: novel polymorphisms and their association with asthma and IgE in childhood. *J Allergy Clin Immunol* **112**, 893-8 (2003).
91. Rosenwasser, L. J. et al. Promoter polymorphisms in the chromosome 5 gene cluster in asthma and atopy. *Clin Exp Allergy* **25 Suppl 2**, 74-8; discussion 95-6 (1995).

92. Basehore, M. J. et al. A comprehensive evaluation of IL4 variants in ethnically diverse populations: association of total serum IgE levels and asthma in white subjects. *J Allergy Clin Immunol* **114**, 80-7 (2004).
93. Graves, P. E. et al. A cluster of seven tightly linked polymorphisms in the IL-13 gene is associated with total serum IgE levels in three populations of white children. *J Allergy Clin Immunol* **105**, 506-13 (2000).
94. Heinzmann, A. et al. Genetic variants of IL-13 signalling and human asthma and atopy. *Hum Mol Genet* **9**, 549-59 (2000).
95. Moissidis, I. et al. Association of IL-13, RANTES, and leukotriene C4 synthase gene promoter polymorphisms with asthma and/or atopy in African Americans. *Genet Med* **7**, 406-10 (2005).
96. Liu, X. et al. An IL13 coding region variant is associated with a high total serum IgE level and atopic dermatitis in the German multicenter atopy study (MAS-90). *J Allergy Clin Immunol* **106**, 167-70 (2000).
97. Wang, M. et al. A common IL-13 Arg130Gln single nucleotide polymorphism among Chinese atopy patients with allergic rhinitis. *Hum Genet* **113**, 387-90 (2003).
98. Heinzmann, A. et al. Association study of the IL13 variant Arg110Gln in atopic diseases and juvenile idiopathic arthritis. *J Allergy Clin Immunol* **112**, 735-9 (2003).
99. van der Pouw Kraan, T. C. T. M. et al. An Il-13 promoter polymorphism associated with increased risk of allergic asthma. *Genes Immunol* **1**, 61-65 (1999).
100. Howard, T. D. et al. Identification and association of polymorphisms in the interleukin-13 gene with asthma and atopy in a Dutch population. *Am J Respir Cell Mol Biol* **25**, 377-84 (2001).
101. Arima, K. et al. Upregulation of IL-13 concentration in vivo by the IL13 variant associated with bronchial asthma. *J Allergy Clin Immunol* **109**, 980-7 (2002).
102. Vladich, F. D. et al. IL-13 R130Q, a common variant associated with allergy and asthma, enhances effector mechanisms essential for human allergic inflammation. *J Clin Invest* **115**, 747-54 (2005).
103. Cameron, L. et al. Th2 cell-selective enhancement of human IL13 transcription by IL13-1112C>T, a polymorphism associated with allergic inflammation. *J Immunol* **177**, 8633-42 (2006).
104. Deichmann, K., Bardutzky, J., Forster, J., Heinzmann, A. & Kuehr, J. Common polymorphisms in the coding part of the IL4-receptor gene. *Biochem Biophys Res Commun* **231**, 696-7 (1997).
105. Hershey, G. K., Friedrich, M. F., Esswein, L. A., Thomas, M. L. & Chatila, T. A. The association of atopy with a gain-of-function mutation in the alpha subunit of the interleukin-4 receptor [see comments]. *N Engl J Med* **337**, 1720-5 (1997).
106. Ober, C. et al. Variation in the interleukin 4-receptor alpha gene confers susceptibility to asthma and atopy in ethnically diverse populations. *Am J Hum Genet* **66**, 517-26 (2000).
107. Mitsuyasu, H. et al. Ile50Val variant of IL4R alpha upregulates IgE synthesis and associates with atopic asthma. *Nat Genet* **19**, 119-20 (1998).
108. Beghe, B. et al. Polymorphisms in the interleukin-4 and interleukin-4 receptor alpha chain genes confer susceptibility to asthma and atopy in a Caucasian population. *Clin Exp Allergy* **33**, 1111-7 (2003).
109. Adjers, K., Pessi, T., Karjalainen, J., Huhtala, H. & Hurme, M. Epistatic effect of IL1A and IL4RA genes on the risk of atopy. *J Allergy Clin Immunol* **113**, 445-7 (2004).
110. Kruse, S. et al. The polymorphisms S503P and Q576R in the interleukin-4 receptor alpha gene are associated with atopy and influence the signal transduction. *Immunology* **96**, 365-71 (1999).
111. Hackstein, H., Hofmann, H., Bohnert, A. & Bein, G. Definition of human interleukin-4 receptor alpha chain haplotypes and allelic association with atopy markers. *Hum Immunol* **60**, 1119-27 (1999).

112. Kauppi, P. et al. A second-generation association study of the 5q31 cytokine gene cluster and the interleukin-4 receptor in asthma. *Genomics* **77**, 35-42 (2001).
113. Haagerup, A. et al. No linkage and association of atopy to chromosome 16 including the interleukin-4 receptor gene. *Allergy* **56**, 775-9 (2001).
114. Tanaka, K. et al. Lack of association between atopic eczema and the genetic variants of interleukin-4 and the interleukin-4 receptor alpha chain gene: heterogeneity of genetic backgrounds on immunoglobulin E production in atopic eczema patients. *Clin Exp Allergy* **31**, 1522-7 (2001).
115. Noguchi, E. et al. Lack of association of atopy/asthma and the interleukin-4 receptor alpha gene in Japanese. *Clin Exp Allergy* **29**, 228-33 (1999).
116. Noguchi, E. et al. No association between atopy/asthma and the ILe50Val polymorphism of IL-4 receptor. *Am J Respir Crit Care Med* **160**, 342-5 (1999).
117. Woitsch, B. et al. A comprehensive analysis of interleukin-4 receptor polymorphisms and their association with atopy and IgE regulation in childhood. *Int Arch Allergy Immunol* **135**, 319-24 (2004).
118. Mitsuyasu, H. et al. Cutting edge: dominant effect of Ile50Val variant of the human IL-4 receptor alpha-chain in IgE synthesis. *J Immunol* **162**, 1227-31 (1999).
119. Risma, K. A. et al. V75R576 IL-4 receptor alpha is associated with allergic asthma and enhanced IL-4 receptor function. *J Immunol* **169**, 1604-10 (2002).
120. Patel, B. K., Keck, C. L., O'Leary, R. S., Popescu, N. C. & LaRochelle, W. J. Localization of the human stat6 gene to chromosome 12q13.3-q14.1, a region implicated in multiple solid tumors. *Genomics* **52**, 192-200 (1998).
121. Hou, J. et al. An interleukin-4-induced transcription factor: IL-4 Stat. *Science* **265**, 1701-6 (1994).
122. Quelle, F. W. et al. Cloning of murine Stat6 and human Stat6, Stat proteins that are tyrosine phosphorylated in responses to IL-4 and IL-3 but are not required for mitogenesis. *Mol Cell Biol* **15**, 3336-43 (1995).
123. Hebenstreit, D., Wirnsberger, G., Horejs-Hoeck, J. & Duschl, A. Signaling mechanisms, interaction partners, and target genes of STAT6. *Cytokine Growth Factor Rev* **17**, 173-88 (2006).
124. Duetsch, G. et al. STAT6 as an asthma candidate gene: polymorphism-screening, association and haplotype analysis in a Caucasian sib-pair study. *Hum Mol Genet* **11**, 613-21 (2002).
125. Tamura, K. et al. Novel dinucleotide repeat polymorphism in the first exon of the STAT-6 gene is associated with allergic diseases. *Clin Exp Allergy* **31**, 1509-14 (2001).
126. Gao, P. S. et al. Variation in dinucleotide (GT) repeat sequence in the first exon of the STAT6 gene is associated with atopic asthma and differentially regulates the promoter activity in vitro. *J Med Genet* **41**, 535-9 (2004).
127. Weiland, S. K. et al. Prevalence of respiratory and atopic disorders among children in the East and West of Germany five years after unification. *Eur Respir J* **14**, 862-70 (1999).
128. Weiland, S. K. et al. Phase II of the International Study of Asthma and Allergies in Childhood (ISAAC II): rationale and methods. *Eur Respir J* **24**, 406-12 (2004).
129. von Mutius, E., Weiland, S. K., Fritzsch, C., Duhme, H. & Keil, U. Increasing prevalence of hay fever and atopy among children in Leipzig, East Germany. *Lancet* **351**, 862-6 (1998).
130. Schram, D. et al. Bacterial and fungal components in house dust of farm children, Rudolf Steiner school children and reference children--the PARSIFAL Study. *Allergy* **60**, 611-8 (2005).
131. Ege, M. J. et al. Prenatal farm exposure is related to the expression of receptors of the innate immunity and to atopic sensitization in school-age children. *J Allergy Clin Immunol* **117**, 817-23 (2006).
132. Miller, S. A., Dykes, D. D. & Polesky, H. F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* **16**, 1215 (1988).
133. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-7 (1977).

134. Zhang, L. et al. Whole genome amplification from a single cell: implications for genetic analysis. *Proc Natl Acad Sci U S A* **89**, 5847-51 (1992).
135. Karas, M. & Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* **60**, 2299-301 (1988).
136. Griffin, T. J. & Smith, L. M. Single-nucleotide polymorphism analysis by MALDI-TOF mass spectrometry. *Trends Biotechnol* **18**, 77-84 (2000).
137. Crow, J. F. Hardy, Weinberg and language impediments. *Genetics* **152**, 821-5 (1999).
138. Sasieni, P. D. From genotypes to genes: doubling the sample size. *Biometrics* **53**, 1253-61 (1997).
139. SAS.
140. Zaykin, D. V. et al. Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. *Hum Hered* **53**, 79-91 (2002).
141. Cordell, H. J. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum Mol Genet* **11**, 2463-8 (2002).
142. Boffelli, D. et al. Phylogenetic shadowing of primate sequences to find functional regions of the human genome. *Science* **299**, 1391-4 (2003).
143. Ovcharenko, I., Boffelli, D. & Loots, G. G. eShadow: a tool for comparing closely related sequences. *Genome Res* **14**, 1191-8 (2004).
144. Quandt, K., Frech, K., Karas, H., Wingender, E. & Werner, T. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res* **23**, 4878-84 (1995).
145. Cartharius, K. et al. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* **21**, 2933-42 (2005).
146. Li-Weber, M., Giasi, M. & Krammer, P. H. Involvement of Jun and Rel proteins in up-regulation of interleukin-4 gene activity by the T cell accessory molecule CD28. *J Biol Chem* **273**, 32460-6 (1998).
147. Promega. pGL3 Luciferase Reporter Assay Manual. (2002).
148. Stratgene. QuikChange Multi-Site Directed Mutagenesis Manual. (2003).
149. Gentra. Versagene RNA Purifiaction Manual. (2004).
150. Invitrogen. ThermoScript Reverse Transcriptase Manual. (2003).
151. Biorad. iCycler iQ Real-Time PCR Detection System Manual.
152. Bustin, S. A. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* **25**, 169-93 (2000).
153. AppliedBiosystems. Real-Time PCR vs Traditional PCR.
154. AppliedBiosystems. Data Analysis on the ABI Prism 7700 Sequence Detection System: Setting, Baseline and Thresholds, Manual.
155. Kim, H., Kim, Y. N. & Kim, C. W. Oxidative stress attenuates Fas-mediated apoptosis in Jurkat T cell line through Bfl-1 induction. *Oncogene* **24**, 1252-61 (2005).
156. Agresti, A. & Vercelli, D. c-Rel is a selective activator of a novel IL-4/CD40 responsive element in the human Ig gamma4 germline promoter. *Mol Immunol* **38**, 849-59 (2002).
157. LeVan, T. D. et al. A common single nucleotide polymorphism in the CD14 promoter decreases the affinity of Sp protein binding and enhances transcriptional activity. *J Immunol* **167**, 5838-44 (2001).
158. Georas, S. N. et al. Stat6 inhibits human interleukin-4 promoter activity in T cells. *Blood* **92**, 4529-38 (1998).
159. Tang, X., Marciano, D. L., Leeman, S. E. & Amar, S. LPS induces the interaction of a transcription factor, LPS-induced TNF-alpha factor, and STAT6(B) with effects on multiple cytokines. *Proc Natl Acad Sci U S A* **102**, 5132-7 (2005).
160. von Mutius, E. et al. Exposure to endotoxin or other bacterial components might protect against the development of atopy. *Clin Exp Allergy* **30**, 1230-4 (2000).
161. Baldini, M. et al. A Polymorphism* in the 5' flanking region of the CD14 gene is associated with circulating soluble CD14 levels and with total serum immunoglobulin E. *Am J Respir Cell Mol Biol* **20**, 976-83 (1999).

162. Di Marco, S., Hel, Z., Lachance, C., Furneaux, H. & Radzioch, D. Polymorphism in the 3'-untranslated region of TNFalpha mRNA impairs binding of the post-transcriptional regulatory protein HuR to TNFalpha mRNA. *Nucleic Acids Res* **29**, 863-71 (2001).
163. Weidinger, S. et al. Association of a STAT 6 haplotype with elevated serum IgE levels in a population based cohort of white adults. *J Med Genet* **41**, 658-63 (2004).
164. Abecasis, G. R. et al. Extent and Distribution of Linkage Disequilibrium in Three Genomic Regions. *Am J Hum Genet* **68**, 191-197 (2001).
165. Svaren, J. et al. The Nab2 and Stat6 genes share a common transcription termination region. *Genomics* **41**, 33-9 (1997).
166. Excoffier, L. & Slatkin, M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* **12**, 921-7 (1995).
167. Rosenwasser, L. J. & Borish, L. Promoter polymorphisms predisposing to the development of asthma and atopy. *Clin Exp Allergy* **28 Suppl 5**, 13-5; discussion 26-8 (1998).
168. Schedel, M. et al. Polymorphismen in der IL-4/IL-13 Signalkaskade beeinflussen die IgE Regulation und Asthma bronchiale. *Allergologie* **28**, 307-314 (2005).
169. Howard, T. D. et al. Gene-gene interaction in asthma: IL4RA and IL13 in a Dutch population with asthma. *Am J Hum Genet* **70**, 230-6 (2002).
170. Pozzoli, U. & Sironi, M. Silencers regulate both constitutive and alternative splicing events in mammals. *Cell Mol Life Sci* **62**, 1579-604 (2005).
171. Barnes, P. J. & Karin, M. Nuclear factor-kappaB: a pivotal transcription factor in chronic inflammatory diseases. *N Engl J Med* **336**, 1066-71 (1997).
172. Bonizzi, G. & Karin, M. The two NF-kappaB activation pathways and their role in innate and adaptive immunity. *Trends Immunol* **25**, 280-8 (2004).
173. Kopp, E. B. & Ghosh, S. NF-kappa B and rel proteins in innate immunity. *Adv Immunol* **58**, 1-27 (1995).
174. Finco, T. S., Beg, A. A. & Baldwin, A. S., Jr. Inducible phosphorylation of I kappa B alpha is not sufficient for its dissociation from NF-kappa B and is inhibited by protease inhibitors. *Proc Natl Acad Sci U S A* **91**, 11884-8 (1994).
175. Verma, I. M., Stevenson, J. K., Schwarz, E. M., Van Antwerp, D. & Miyamoto, S. Rel/NF-kappa B/I kappa B family: intimate tales of association and dissociation. *Genes Dev* **9**, 2723-35 (1995).
176. Angelini, F. et al. Nuclear factor kappaB activity is increased in peripheral blood mononuclear cells of children affected by atopic and non-atopic eczema. *Int J Immunopathol Pharmacol* **20**, 59-67 (2007).
177. Stutz, A. M. & Woisetschlager, M. Functional synergism of STAT6 with either NF-kappa B or PU.1 to mediate IL-4-induced activation of IgE germline gene transcription. *J Immunol* **163**, 4383-91 (1999).
178. Shen, C. H. & Stavnezer, J. Interaction of stat6 and NF-kappaB: direct association and synergistic activation of interleukin-4-induced transcription. *Mol Cell Biol* **18**, 3395-404 (1998).
179. Matsukura, S. et al. Activation of eotaxin gene transcription by NF-kappa B and STAT6 in human airway epithelial cells. *J Immunol* **163**, 6876-83 (1999).
180. Messner, B., Stutz, A. M., Albrecht, B., Peiritsch, S. & Woisetschlager, M. Cooperation of binding sites for STAT6 and NF kappa B/rel in the IL-4-induced up-regulation of the human IgE germline promoter. *J Immunol* **159**, 3330-7 (1997).
181. Suske, G. The Sp-family of transcription factors. *Gene* **238**, 291-300 (1999).
182. Hata, Y., Duh, E., Zhang, K., Robinson, G. S. & Aiello, L. P. Transcription factors Sp1 and Sp3 alter vascular endothelial growth factor receptor expression through a novel recognition sequence. *J Biol Chem* **273**, 19294-303 (1998).
183. Prowse, D. M., Bolgan, L., Molnar, A. & Dotto, G. P. Involvement of the Sp3 transcription factor in induction of p21Cip1/WAF1 in keratinocyte differentiation. *J Biol Chem* **272**, 1308-14 (1997).

184. Yao, P. L., Lin, Y. C., Sawhney, P. & Richburg, J. H. Transcriptional regulation of FasL expression and participation of sTNF-alpha in response to sertoli cell injury. *J Biol Chem* **282**, 5420-31 (2007).
185. Guo, Z., Boekhoudt, G. H. & Boss, J. M. Role of the intronic enhancer in tumor necrosis factor-mediated induction of manganous superoxide dismutase. *J Biol Chem* **278**, 23570-8 (2003).
186. Takahashi, K., Nishiyama, C., Hasegawa, M., Akizawa, Y. & Ra, C. Regulation of the human high affinity IgE receptor beta-chain gene expression via an intronic element. *J Immunol* **171**, 2478-84 (2003).
187. Patel, B. K., Pierce, J. H. & LaRochelle, W. J. Regulation of interleukin 4-mediated signaling by naturally occurring dominant negative and attenuated forms of human Stat6. *Proc Natl Acad Sci U S A* **95**, 172-7 (1998).
188. Wang, L. H., Yang, X. Y., Kirken, R. A., Resau, J. H. & Farrar, W. L. Targeted disruption of stat6 DNA binding activity by an oligonucleotide decoy blocks IL-4-driven T(H)2 cell response. *Blood* **95**, 1249-57 (2000).
189. Kriebel, P., Patel, B. K., Nelson, S. A., Grusby, M. J. & LaRochelle, W. J. Consequences of Stat6 deletion on Sis/PDGF- and IL-4-induced proliferation and transcriptional activation in murine fibroblasts. *Oncogene* **18**, 7294-302 (1999).
190. Baetta, R. et al. Upregulation and activation of Stat6 precede vascular smooth muscle cell proliferation in carotid artery injury model. *Arterioscler Thromb Vasc Biol* **20**, 931-9 (2000).
191. Bruns, H. A. & Kaplan, M. H. The role of constitutively active Stat6 in leukemia and lymphoma. *Crit Rev Oncol Hematol* **57**, 245-53 (2006).
192. Ross, J. R. et al. Clinical response to morphine in cancer patients and genetic variation in candidate genes. *Pharmacogenomics J* **5**, 324-36 (2005).
193. Kraus, J. et al. Regulation of mu-opioid receptor gene transcription by interleukin-4 and influence of an allelic variation within a STAT6 transcription factor binding site. *J Biol Chem* **276**, 43901-8 (2001).

8 Table index

9 Figure index

10 Abbreviations

95% CI	95% Confidence Interval
ABI	Applied Biosystems
ADAM33	A Disintegrin and Metalloproteinase Domain 33
ALEX	Allergy and Endotoxin study population
ALL	acute lymphoblastic leukemia
APC	antigen presenting cell
APS	Ammonium persulfate
ATP	Adenosintriphosphate
BAL	bronchoalveolar lavage
BALB/c	an inbred albino strain of a lab mouse
BCA	bicinchoninic acid
BDB	DNA-binding domain
BHR	bronchial hyperresponsiveness
bp	basepair
BSA	Bovine serum albumin
CCD	charge-couple device
CD14	monocyte Differentiation antigen14
CD23 (FCεR2)	low affinity IgE receptor
CD28	cluster designation 28
CD4	cluster designation 4
CD40	cluster designation 40
CD40L	cluster designation 40 ligand
CD8	cluster designation 8
cDNA	coding DNA
CNS	conserved non-coding sequence
CpG	cytosine and guanine separated by a phosphate
cpm	counts per minute
C_T	threshold cycle
DC	dendritic cell
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DPP10	Dipetidyl peptidase 10
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen (German Resource Centre for Biological Material)
DT	Divergence Threshold
E.coli	Escherichia coli
e.g.	for example
ECACC	European Collection of Cell Cultures
EGR1	Early Growth Response 1
EM	expectation-maximization
EMSA	Electrophoretic Mobility Shift Assay
EPO	electroporation
ETS	environmental tobacco smoke
FACS	fluorescence activated cell sorting
FCεR1	high affinity IgE receptor
FCS	Fetal Calve Serum
FEV	Forced Expiratory Volume
fwd	forward
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GATA3	GATA-binding protein 3
GMR	geometric mean ratio
GPRA	G Protein-coupled Receptor for Asthma Susceptibility
GST	glutathione-S-transferease

GSTM1	GSTmu
GSTT1	GST-theta
HLA G	Histocompatibility antigen , class I
hME	homogeneous Mass Extent™
HMMI	Hidden Markov Model Islands
HPLC	High Performance Liquid Chromatography
HWE	Hardy-Weinberg equilibrium
IκB	inhibitor complex for NF-κB
I50V	isoleucine mutation to valine at amino acid 50
IFN-γ	Interferon gamma
IgA	Immunglobulin A
IgE	Immunglobulin E
IgG	Immunglobulin G
IgM	Immunglobulin M
IL	Interleukin
IL-12	Interleukin-12
IL-13	Interleukin-13
IL-15	Interleukin-15
IL-2	Interleukin-2
IL-4	Interleukin-4
IL-4Rα	Interleukin-4 Receptor α
IL-5	Interleukin-5
IL-9	Interleukin-9
ISAAC	International Study of Asthma and Allergies in Children
JAK	Janus Tyrosine Kinase
LAR	Luciferase Assay Reagent II
LD	Linkage Disequilibrium
LD	Low Density
LPS	lipopolysaccharide
MAF	Minor Allele Frequency
MALDI-TOF MS	Matrix Assisted Laser Desorption-Time of Flight Mass Spectrometry
MAPK	Mitogen-Activated Protein Kinase
MCR	multiple cloning region
MHC	Major Histocompatibility Complex
mRNA	messenger RNA
MZF-1	Myeloid Zinc Finger Protein 1
NAB2	EGR1 binding protein 2
NC	negative control
NEB	New England Biolabs
NF-κB	Nucleofactor κB
NFAT2	Nuclear Factor of Activated T-cells 2
NIH3T3	mouse fibroblast cell line
NZY$^+$	casein hydrolysate amine
Oct-1	Octamer binding protein 1
OD	optical density
OR	Odds Ratio
ORF	Open Reading Frame
p	p-value
P/I	PMA/Ionomycin
PAGE	Polyacrylamid Gel Electrophoresis
PARSIFAL	Prevention of Allergy-Risk factors for Sensitization In children related to Farming and Anthroposophic Lifestyle
PBMC	Peripheral blood monocytes
pBP	promoter Bharvin Patel
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PDGF	platelet-derived growth factor
PEP	Primer Extension Preamplification

PEX19	Peroxisome Biogenesis Factor 19
PLB	Passive Lysis Buffer
PMA	Phorbol12-myristate 13-acetate
PO	polymorphic allele
qRT-PCR	quantitative Real-time PCR
rev	reverse
RFLP	Restriction Fragment Length Polymorphism
RLU	Relative Light Units
RNA	Ribonucleic acid
RS virus	Respiratory Syncitial Virus
RT	Reverse Transcriptase
S6	STAT6
SAP	Shrimp Alkaline Phosphatase
SH2	src homology 2 domain
siRNA	small interfering RNA
SNP	Single Nucleotide Polymorphism
SP	Specificity Proteins
STAT1	Signal Transducer and Activator of Transcription 1
STAT2	Signal Transducer and Activator of Transcription 2
STAT4	Signal Transducer and Activator of Transcription 4
STAT6	Signal Transducer and Activator of Transcription 6
TAD	transactivation domain
T-bet	T-box expressed in T-cells
TCR	T-cell receptor
T_H cells	T-Helper cell
T_m	melting temperature
TLR	Toll-like receptor
TNFα	Tumor necrosis factor alpha
TRANSFAC	Transcription Factor Classification
Treg	T-regulatory cell
t-Test	statistical test: Student`s t distribution
USA	United States of America
UTR	Untranslated Region
VE	variable element
WT	wildtype allele
YB1	Y-box binding protein 1
YY1	Ying-Yang 1

11 Presentations at scientific meetings

- **Oral presentation: European Association of Asthma and Clinical Immunology,** 09.-13.06.07, Göteborg, Sweden
  Schedel M., Kabesch M. An enhanced NF-κB-depending STAT6 gene expression may explain the association between a STAT6 polymorphism and the risk for elevated total IgE levels
- **Poster presentation: 29. Jahrestagung der Gesellschaft für Kinderheilkunde**, 22.-24.3.07, Munich, Germany
  Schedel M., Kabesch M. Ein STAT6 Polymorphismus erhöht das Asthmarisiko und den IgE Spiegel durch NF-κB vermittelte Genexpressionserhöhung
- **Poster presentation: European Respiratory Society**, 02.09.-06.09.06, Munich, Germany
  Schedel M., Depner M., Schoen C., Weiland S. K., Vogelberg C., Niggemann B., Lau S., Fritzsch C., Illig T., Klopp N., Wahn U., von Mutius E., Nickel R., Kabesch M. The role of *ADAM33* polymorphisms in two German Populations
- **Oral presentation and Poster presentation in the Junior Poster Session: World Allergie Congress,** 26.06- 01.07.05, Munich, Germany
  Schedel M, Carr D, Klopp N, Woitsch B, Illig T, Stachel D, Schmid I, Fritzsch C, Weiland S, von Mutius E, Kabesch M. Polymorphisms in the *STAT6* gene influence total serum IgE levels in children
- **Poster presentation: World Allergie Congress,** 26.06- 01.07.05, Munich, Germany
  Schedel M, Schön C, Carr D, Wjist M, Illig T, Klopp N, Fritzsch C, Weiland S, von Mutius E, Kabesch M. *ADAM33* polymorphisms influence the development of asthma
- **Oral presentation: ERS 3rd Lung Science Conference**, 17.-19.03.05, Taormina, Italy
  Schedel M, Carr D, Klopp N, Woitsch B, Illig T, Stachel D, Schmid I, Fritzsch C, Weiland S, von Mutius E, Kabesch M. A genetic variation in a putative NF-κB enhancer region of the *STAT6* gene influences IgE levels
- **Poster presentation: NGFN Meeting**, 2.-4.06.04, Berlin, Germany
  Schedel M, Carr D, Klopp N, Woitsch B, Illig T, Stachel D, Schmid I, Fritzsch C, Weiland S, von Mutius E, Kabesch M. Polymorphisms in the *STAT6* gene influence total serum IgE levels in children
- **Oral presentation: American Thoracic Society**, 21.-26.05.04, Orlando, USA
  Schedel M, Carr D, Klopp N, Woitsch B, Illig T, Stachel D, Schmid I, Fritzsch C, Weiland S, von Mutius E, Kabesch M. A *STAT6* haplotype influences the regulation of serum IgE levels
- **Poster presentation: 26. Jahrestagung der Gesellschaft für Kinderheilkunde**, 25.-27.3.04, Hamburg; Germany
  Schedel M, Carr D, Klopp N, Woitsch B, Illig T, Stachel D, Schmid I, Fritzsch C, Weiland S, von Mutius E, Kabesch M. Polymorphismen im *STAT6*-Gen beeinflussen den Gesamt IgE-Spiegel bei Kindern

Poster Award:Junior Poster Session, World Allergy Congress 2005, Munich, Germany
Polymorphisms in the STAT6 gene influence total serum IgE levels in children

Abstract Award for Oral Presentation: European Association of Asthma and Clinical Immunology, 2007, Göteborg, Sweden
An enhanced NF-κB-depending STAT6 gene expression may explain the association between a STAT6 polymorphism and the risk for elevated total IgE levels

12 List of publications

- ***Schedel M**, *Pinto LA, Schaub B, Rosenstiel P, Cherkasov D, Cameron L, Klopp N, Illig T, Vogelberg C, Weiland SK, von Mutius E, *Lohoff M, *Kabesch M. *IRF1* gene variations influence IgE regulation and atopy, Am J Respir Crit Care Med. 2007 Dec 13 [Epub ahead of print]

- **Schedel M,** Depner M, Schoen C, Weiland SK, Vogelberg C, Niggemann B, Lau S, Fritzsch C, Illig T, Klopp N, Wahn U, von Mutius E, Nickel R, Kabesch M. The role of polymorphisms in *ADAM33*, a disintegrin and metalloprotease 33, in childhood asthma and lung function in two german populations, **Respir Res.** 2006 Jun 19;7:91

- Kabesch M, **Schedel M**, Carr D, Woitsch B, Fritzsch C, Weiland SK, von Mutius E. Interleukin 4/ Interleukin 13 pathway genetics strongly influence the serum IgE levels and childhood asthma. **J Allergy Clin Immunol**. 2006 Feb;117(2):269-74

- **Schedel M**, Carr D, Woitsch B, Schmid I, Fritzsch C, Weiland SK, von Mutius E, Kabesch M. Polymorphis-men in der IL-4/IL-13 Signalkaskade beeinflussen die IgE Regulation und Asthma bronchiale. **Allergologie** 2005; 28/8, 307-314

- Hysi P, Kabesch M, Moffatt MF, **Schedel M**, Carr D, Zhang Y, Boardman B, von Mutius E, Weiland SK, Leupold W, Fritzsch C, Klopp N, Musk AW, James A, Nunez G, Inohara N, Cookson WO. NOD1 variation, immunoglobulin E and asthma. **Hum Mol Genet.** 2005 Apr 1;14 (7):935-41. Epub 2005 Feb 17.

- Weidinger S, Klopp N, Wagenpfeil S, Rummler L**, Schedel M**, Kabesch M, Schafer T, Darsow U, Jakob T, Behrendt H, Wichmann HE, Ring J, Illig T. Association of a STAT 6 haplotype with elevated serum IgE levels in a population based cohort of white adults. **J Med Genet.** 2004 Sep;41(9):658-63

- **Schedel M**, Carr D, Woitsch B, Stachel D, Schmid I, Fritzsch C, Weiland SK, von Mutius E, Kabesch M. A signal transducer and activator of transcription 6 haplotype influences the regulation of serum IgE levels. **J Allergy Clin Immunol** 2004; 114:1100-5.

13 Acknowledgements

I would like to gratefully and sincerely thank my advisor, PD Dr. M. Kabesch, for his guidance, support and encouragement throughout my PhD studies. I have been amazingly fortunate to have a mentor who was always there to listen and to give advice. He inspired me to develop a critical scientific thinking with the freedom to explore my own ideas. The valuable technical and editorial comments fundamentally improved the contents of this dissertation.

Besides my advisor, my thanks go to the members of my committee Prof. Dr. J. Adamski, Prof. Dr. Dr. H. Wagner and Prof. Dr. M. Hrabe de Angelis for providing valuable comments and knowledge.

Carrying out the research would not have been possible without the help of Prof. Dr. E. von Mutius, PD. Dr. T. Illig, Dr. N. Klopp and PD. Dr. R. Lauener who participated in my project with great interest and enthusiasm.

The friendship of Ph.D. L. Cameron and Dr. med B. Schaub is greatly appreciated and has provided many interesting and good-spirited discussions relating to this research. They had strong confidence in me and offered substantial moral support and optimism to overcome challenging situations.

Special thanks go to my colleagues for helping me at any time and providing a good working atmosphere. Each member of the group supported the present work with various aspects. I would like to acknowledge M. Korman who implemented interesting discussions and assistance in computer problems. I appreciate the efforts of D. Carr and M. Depner for helping me to better understand statistical analyses. In particular, I am grateful to I. Dahmen, A. Pleiss, K. Suttner, L. A. Pinto and S. Zeilinger who dedicated their valuable time and practical assistance throughout this project.

13. ACKNOWLEDGEMENT

Many friends were an essential stabilizing force throughout the past years. I could not have finished this dissertation without their constant encouragement and support. I want to express particular gratitude to F. Böcking, S. Buschmann, I. Ruoss, A. Lenz and S. Gersting. Thank you for listening to my complaints and frustrations.

Finally, but most importantly, I would like to thank my family. Their unconditional patience and care helped me to stay focused on my goals in life. I deeply appreciate their belief in me.