Genetic Variants in ST2 Previously Associated with Asthma are Associated with Asthma and sST2 Levels in a Brazilian Population Living in an Area Endemic for Schistosoma mansoni

Bergljót Rafnar Karlsdóttir

Thesis for BSc-degree in Medicine
University of Iceland
School of Health Sciences
Genetic Variants in ST2 Previously Associated with Asthma are Associated with Asthma and sST2 Levels in a Brazilian Population Living in an Area Endemic for Schistosoma mansoni

Bergljót Rafnar Karlsdóttir

Thesis for the degree of B.Sc. in Medicine
Supervisors: Dr. Kathleen C. Barnes, Dr. Candelaria Vergara, Dr. Li Gao, Dr. Rasika Mathias

1Department of Medicine, University of Iceland 2 Division of Allergy and Clinical Immunology, Department of Medicine, The Johns Hopkins University
**Abstract**

**Introduction.** Asthma is a common, heterogenous syndrome of the respiratory system estimated to affect as many as 300 million people worldwide. A strong overlap between immune response mechanisms responsible for both asthma and resistance to schistosomiasis, a helminth infection, in addition to genetic linkage and association data, supports the hypothesis that genetic determinants conferring resistance to schistosomiasis are also associated with greater risk of asthma. Genome-wide association studies (GWAS) have identified *IL1RL1*, also known as *ST2*, as one of the most consistently associated candidate genes for asthma. The aims of this study are to determine whether genetic variants of *ST2*, previously associated with asthma and serum ST2 levels in asthmatics, are co-associated with resistance to schistosomiasis and serum ST2 levels in participants exposed to schistosomiasis.

**Materials and methods.** 697 Brazilian DNA samples stored in -80°C were selected and prepared for genotyping. Four *ST2* SNPs associated with asthma and soluble serum ST2 (sST2) levels in an African-American population from Baltimore/Washington were genotyped using TaqMan ABI 7900. Hardy Weinberg testing and tests for Mendelian inconsistencies were performed using PLINK, and pairwise linkage disequilibrium estimations using Haploview. Tests for genetic association were performed using a generalized estimating equation (GEE) under a dominant model on the soluble adult worm antigen (SWAP)-specific IgE/IgG4 ratio (a measure of resistance to *Schistosoma mansoni* infection) and sST2 levels. Variables were log-transformed adjusting for age and gender as needed.

**Results.** Three *ST2* markers (rs1420101, rs12712135 and rs6543119) previously associated with asthma in African Americans are co-associated with changes in sST2 levels in a study population of Brazilians. One marker positively associated with asthma in African-Americans (rs76930359) is negatively associated with wheeze in the study population. The four genetic variants in *ST2* are not associated with prevalence or severity of infection by *S. mansoni*, IgE/IgG4 ratio, nor tIgE levels in this population.

**Conclusion.** The coded allele C in the SNP rs76930359 is negatively associated with wheeze and thus could be linked with protection against asthma in this population. The SNPs rs1420101, rs12712135 and rs6543119 are associated with changes in sST2 levels in both Brazilian and African American populations and may be an interesting area of study, for example, in the research on sST2 as a biomarker for disease. Further analyses are needed to better understand the role of variants in *ST2* in asthma, expression of sST2, and infection by *Schistosoma mansoni.*
## Table of contents

Abstract .................................................................................................................................................. i
Table of contents .................................................................................................................................... ii
List of figures .......................................................................................................................................... iv
List of tables and graphs ...................................................................................................................... v
List of abbreviations ........................................................................................................................... vi
Acknowledgements ............................................................................................................................ viii
1 Introduction ........................................................................................................................................ 1
  1.1 Atopy and allergy .......................................................................................................................... 1
      1.1.1 Definitions of atopy and allergy .............................................................................................. 1
  1.2 Asthma ........................................................................................................................................... 1
      1.2.1 Prevalence and definition of asthma ....................................................................................... 1
      1.2.2 Modifiers of risk for asthma .................................................................................................. 2
      1.2.3 Asthma and the immune system ............................................................................................. 2
  1.3 Race and the genetics of allergic disease ....................................................................................... 2
      1.3.1 Linkage Disequilibrium (LD) and analyses of genetic association ........................................ 3
  1.4 Genome-wide association studies (GWAS) of asthma ................................................................. 3
      1.4.1 Advantages of GWAS ........................................................................................................... 3
      1.4.2 Limitations to GWAS ............................................................................................................ 4
      1.4.3 Meta-analyses of GWAS – GABRIEL and EVE ................................................................. 4
  1.5 *IL-33* and *ST2* .......................................................................................................................... 5
      1.5.1 *IL-33* .................................................................................................................................... 5
      1.5.2 *ST2* ...................................................................................................................................... 7
      1.5.3 sST2 levels in disease .............................................................................................................. 8
  1.6 Schistosomiasis .............................................................................................................................. 9
      1.6.1 Prevalence of schistosomiasis ................................................................................................. 9
      1.6.2 *Schistosoma* species and their life cycles .............................................................................. 9
      1.6.3 Schistosomiasis and the immune system ............................................................................... 10
  1.7 Asthma and schistosomiasis ........................................................................................................... 10
      1.7.1 Interaction between asthma and infection with *S. mansoni* ................................................. 10
1.7.2 Co-associations in genetics of asthma and schistosomiasis .......................... 11

2 Aim of study ........................................................................................................... 11

3 Materials and methods .......................................................................................... 12

3.1 Previously existing data - location and study population .................................. 12

3.1.1 Demographic information and sample collection ........................................... 12

3.1.2 Evidence of asthma and schistosomiasis ....................................................... 12

3.1.3 Assessment of water contact ......................................................................... 13

3.2 Genotyping ......................................................................................................... 13

3.2.1 The Taqman method ..................................................................................... 13

3.2.2 Genotyping Quality Control .......................................................................... 14

3.3 Measurements of sST2, tIgE and the IgE/IgG4 ratio ......................................... 14

3.3.1 sST2 Levels .................................................................................................. 14

3.3.2 Total IgE measurements ............................................................................... 14

3.3.3 IgE/IgG4 ....................................................................................................... 14

3.4 Statistical Methods. .......................................................................................... 14

3.4.1 Data testing and analyses ............................................................................. 14

3.4.2 Population stratification ............................................................................... 15

4 Results .................................................................................................................... 16

4.1 Clinical characteristics of the analyzed subjects ................................................. 16

4.2 sST2 levels compared to other traits in the total population ............................ 18

4.3 Analyzed SNPs – a closer look .......................................................................... 19

4.4 GEE analyses of ST2 markers and other traits ................................................. 21

5 Discussion ............................................................................................................ 24

5.1 Clinical characteristics compared to other populations .................................... 24

5.2 Associations between sST2 levels and other traits ......................................... 26

5.3 Associations between genetic variations in ST2 and other traits .................... 26

5.4 Conclusions ....................................................................................................... 29

References ............................................................................................................... 30
List of figures

Figure 1  IL-33/ST2 signaling .................................................................................. 6
Figure 2  Human ST2 ............................................................................................... 7
Figure 3  Transmission cycle of schistosomes .......................................................... 9
Figure 4  Plots of the first principal components from analysis on Brazilians founders..... 15
List of tables and graphs

Table 1 Clinical characteristics of the analyzed subjects ............................................. 16
Table 2 Clinical characteristics of the analyzed subjects distributed by gender ............ 17
Table 3 Pairwise GEE analysis of sST2 and other traits in the total population............. 18
Table 4 HWE and MAFs of sST2 markers in a Brazilian population............................ 19
Table 5 GEE analysis of ST2 markers and wheeze in a Brazilian population.............. 21
Table 6 GEE analysis of ST2 markers and sST2 levels in a Brazilian population........... 22
Table 7 GEE analysis of ST2 markers and SWAP, total IgE and egg count in a Brazilian population................................................................. 23

Graph 1 Gene structure of ST2 and LD pattern of analyzed markers in a Brazilian population........................................................................................................... 20
List of abbreviations

AFR: African
ASN: Asian
CEU: CEPH (Utah residents with ancestry from northern and western Europe)
CHB: Han Chinese in Beijing, China
Chr.: Chromosome
CI: Confidence Interval
COPD: Chronic Obstructive Pulmonary Disease
DNA: Deoxyribonucleic Acid
ELISA: Enzyme-Linked Immunosorbent Assay
EUR: European
GEE: Generalized Estimating Equation
GWAS: Genome-Wide Association Study
HWE: Hardy Weinberg Equilibrium
IFNg: Interferon gamma
Ig: Immunoglobulin
IL-: Interleukin
IL1RACP: IL-1 Receptor Accessory Protein
IL1RL1: Interleukin-1 Receptor-Like 1
IRAK: IL-1R-Associated Kinase
JNK: JUN N-terminal Kinase
kb: kilobases
LD: Linkage disequilibrium
MAF: Minor allele frequency
MyD88: Myeloid Differentiation primary response protein 88
NF-κB: Nuclear Factor kappa-B
No.: Number
ORMDL3: Orosomucoid like 3
PC: Principal Component
PCA: Principal Component Analysis
PCR: Polymerase chain reaction
SNP: Single nucleotide polymorphism
ST2: suppressor of tumorigenicity 2
SWAP: Soluble adult worm antigen preparation
Th: T-helper
tIgE: total Immunoglobulin E
TIR: Toll-like/Interleukin-1 receptor
TNFα: Tumor Necrosis Factor alpha
TR: Transmembrane
TRAF-6: TNF Receptor-Associated Factor 6
USA: United States of America
YRI: Yoruba in Ibadan, Nigeria
Acknowledgements

I would first and foremost like to thank my supervisor, Kathleen Barnes, for allowing me the opportunity to work with her and her team at Johns Hopkins University. Thank you to everybody at the Division of Allergy and Clinical Immunology who helped me with this project, in particular; Monica Campbell for teaching me the Taqman method for genotyping, Nicholas Rafaels for helping with statistical analyses, Candelaria Vergara for helping with interpretation of data and Romina Ortiz for her assistance and friendship. I would also like to thank my parents, Þórunn Rafnar and Karl Ólafsson, for their love, help and support.
1 Introduction

1.1 Atopy and allergy

1.1.1 Definitions of atopy and allergy
Atopy refers to the immune system producing specific IgE (Immunoglobulin E) in response to otherwise harmless environmental substances known as allergens. This is reflected in either a positive skin prick test or the presence of specific IgE to one or more known allergens in serum. Elevated total IgE is also sometimes used as an equivalent to atopy.

Allergy is a detrimental immune response to allergens which results in allergic diseases such as asthma, atopic dermatitis, eczema, food allergy and allergic rhinitis. Atopy and allergic diseases are strongly associated with each other, however, not everybody with clinical manifestations of an allergic disease can be shown to be atopic nor do all atopic persons display signs of an allergic disease.1,2

1.2 Asthma

1.2.1 Prevalence and definition of asthma
Asthma is a common, heterogeneous syndrome of the respiratory system with a global prevalence ranging from 1% to 18% of the population in different countries.3-7 Asthma prevalence is increasing in most countries, especially among children.3 It is estimated to affect as many as 300 million people worldwide and is expected to increase to 400 million by the year 2025.3,4

Asthma is believed to account for 1 in every 250 deaths worldwide, resulting in 250,000 deaths each year.3,4 Many of the deaths are preventable, being due to suboptimal long-term medical care and delay in obtaining assistance during the final acute attack.8,9

Asthma is a chronic, obstructive lung disease and is characterized by the presence of chronic inflammation in the lower airways with variable and reversible airway obstruction and bronchial hyperresponsiveness. Clinical symptoms include recurrent episodes of coughing, wheezing, breathlessness, and chest tightness.2,3

A hallmark of other obstructive pulmonary diseases, such as emphysema and chronic bronchitis, is change in patient spirometry. These changes are not always present in asthma and as there are no known biomarkers or tests that can definitively diagnose asthma, diagnosis is based on the presence of clinical features and exclusion of diseases that might mimic the symptoms of asthma.3,9
1.2.2 Modifiers of risk for asthma

Asthma has many different phenotypes that depend on both genetics, with heritability estimates varying between 35% and 95%, and environmental factors as well as the interaction between these factors. Several associations have been introduced by epidemiologic studies between environmental exposures at critical times in development and the subsequent risk for asthma and other allergic diseases. Modifiers of risk that have been reported in the literature are for example race/ethnicity, sex, passive and active tobacco smoke exposure and many more. Further research is needed to clarify the importance of each and to determine the mechanisms through which each of these factors modifies risk for asthma and allergy.

1.2.3 Asthma and the immune system

Asthma is predominantly a disease of the Th2 (T-helper 2) immune response with production of inflammatory cytokines such as IL-4 (Interleukin 4), IL-5 and IL-13 and subsequent production of IgE. Recent studies have demonstrated a role played by the Th1 immune response as well as other cytokines, such as IL-25 and IL-33, in the pathogenesis of asthma.

1.3 Race and the genetics of allergic disease

The population of this study was ascertained from five communities (Buri, Camarao, Genipapo, Sempre Viva and Cobo) in the district of Conde, Bahia, Brazil. Brazilians represent a trihybrid population as the result of the admixture of West African, European, and Amerindian ancestral populations, with a substantial African component in Bahia. This is important to note because of racial differences in both expression of asthma and genetics. Phenotypic expression of the disease may vary between racial groups and can also be confounded by environmental differences. The frequency of genetic variation varies between races so an important disease susceptibility allele common to one race and uncommon in another may not be detected due to lack of statistical power and sample size required to identify the rarer variant. In addition, for some genetic variants, the most common form (allele) is the less frequent form in a different race.
1.3.1 Linkage Disequilibrium (LD) and analyses of genetic association

Another point to note is that the correlation between genetic variants in a gene may differ based on historical geographical ancestry. A SNP (single nucleotide polymorphism) can be strongly associated with one or more SNPs in the same gene or even across neighboring genes, termed linkage disequilibrium (LD). When SNPs are in full LD they are always inherited together and thus represent a single genetic signal. However, LD between SNPs can vary between populations. In order to locate genetic variants which influence disease, such as asthma, several different approaches have been explored. Linkage analyses of families with a high prevalence of disease could only identify large regions containing many co-segregated genes and candidate gene studies tended to be underpowered and hard to replicate. Thus association studies of asthma have, until recently, been largely unsuccessful.

1.4 Genome-wide association studies (GWAS) of asthma

The number of identified asthma susceptibility genes has increased rapidly over the last few years, especially with the application of the GWAS approach. In an asthma GWAS 300,000 to more than a million DNA (deoxyribonucleic acid) polymorphisms covering the genome are investigated for association with asthma in large samples of cases and control subjects.39

1.4.1 Advantages of GWAS

The greatest advantage of genome-wide approaches is that they can discover novel genes and pathways involved in disease pathogenesis. Other, older methods (the candidate gene studies, for example) focus primarily on polymorphisms in genes with a known function. In candidate gene studies of asthma bias lay heavily on genes with immunologic functions though genes with other functions have also been identified. GWASs do not require the studying of families40 and should, in theory, be able to detect associations between disease and common variation in the genome.
1.4.2 Limitations to GWAS

A limitation of GWAS approaches is the statistical burden that results from the large number of tests performed and the resulting requirement for very large sample sizes to achieve genome-wide statistical significance. The correction for multiple tests will recommend a typical $P$-value requirement of less than $10^{-7}$. Thus, real existing associations which do not reach this limit will not be recognized by a GWAS analysis.

Another limitation of the GWAS is that it does not detect less common risk variants, both because the genotyping platforms include mostly common variants and because of the reduced power to detect associations with SNPs with low (less than 1%) minor allele frequencies (MAFs). In fact, the GWASs to date, across many diseases including asthma, suggest that the risk alleles identified by this approach account for a very small proportion of the genetic risk thus indicating that rarer variants may have a larger effect on disease risk. For example, the seven most associated SNPs in the GABRIEL meta-analysis (discussed below) could accurately classify individuals with asthma with only 35% sensitivity. Therefore, although the variants show significant associations with asthma, they are in general poor predictors of disease status, reflecting their small effect size and common allele frequencies in the population.

Correlation found between SNPs in GWASs can also make it difficult to determine which SNPs should be investigated further when they are in high LD with each other. For example, *ORMDL3* (Orosomucoid like 3) was identified as a novel asthma susceptibility locus on chromosome 17p21 in the first GWAS of asthma, published in 2007. Association was observed between SNPs in multiple genes in this region on chromosome 17 leading to the article “Guilt by Association” which raised the question as to whether *ORMDL3* is the relevant gene. In this example, it will be necessary to identify all associated variants, common and rare, and test them separately to reach a final conclusion.

1.4.3 Meta-analyses of GWAS – GABRIEL and EVE

The need for large sample sizes to achieve statistical significance has encouraged collaborations and meta-analyses of GWASs where smaller studies are combined to increase power. Two meta-analyses of asthma GWAS have recently been completed, one by the GABRIEL Consortium of European Investigators including subjects of European ancestry
and one by the EVE Consortium of U.S. Investigators including racially and ethnically
diverse subjects from the USA and Mexico.\textsuperscript{44,45}

Both large meta-analyses of asthma GWASs produced similar results.\textsuperscript{2,41,44,45} SNPs in or near seven loci were associated with asthma in both studies, and SNPs in or near four of these loci had $P$-values at or near genome-wide levels of significance in both studies with contributions from ethnically diverse samples. Two of these loci were $IL1RL1$ (Interleukin 1 Receptor-Like 1), also known as and from now on referred to as $ST2$ (Suppressor of Tumorigenicity 2), and $IL-33$. Both can be considered robustly associated asthma susceptibility genes.

\section*{1.5 $IL-33$ and $ST2$}

$IL-33$ and $ST2$ are two of the most consistently associated candidate genes for asthma.\textsuperscript{41,46}

\subsection*{1.5.1 $IL-33$}

The $IL-33$ gene, located on chromosome 9, spans approximately 42.2 kilobases (kb), harboring 8 exons.\textsuperscript{39} The gene encodes a protein, also named $IL-33$, which is a member of the IL-1 cytokine family and was first discovered as a nuclear factor in endothelial cells of high endothelial venules in 2005.\textsuperscript{47}

$IL-33$ has two main functions; as an intracellular transcription factor, influencing gene transcription independent of its receptor, and as a cytokine, signaling through its receptor complex.

$IL-33$ works as a nuclear factor by binding directly to the chromatin\textsuperscript{48,49} as well as to the $NF-\kappa B$ (Nuclear Factor kappa-B) proteins p50 and p65.\textsuperscript{50} In doing so, $IL-33$ is capable of regulating the expression of pro-inflammatory genes, such as $IL-6$, $IL-8$, and the p65 subunit of $NF-\kappa B$\textsuperscript{51,52} and thus might be a direct regulator of the inflammatory response.

Cellular necrosis and tissue damage can lead to the release of the full-length functional $IL-33$ protein from leaking cells, promoting inflammation.\textsuperscript{53,54} In contrast, during apoptosis, $IL-33$ is cleaved by caspases and released in its biologically inactive form.\textsuperscript{53,55,56} (Figure 1)

This mechanism was probably retained through evolution to limit the release of bioactive full-length $IL-33$ during programmed cell death.\textsuperscript{54}
Because IL-33 is released on injury or damage, it has been described as an “alarmin,” translating damage into activation of the inflammatory response.

Once released, biologically active IL-33 activates a heterodimeric receptor complex containing an isoform of ST2 and IL1RacP (IL-1 Receptor Accessory Protein).\(^\text{47,57}\) (Figure 1) Binding of IL-33 to the ST2/IL1RacP receptor complex recruits signaling adaptor proteins. These proteins activate pathways that induce gene expression leading to, for example, cytokine and chemokine synthesis.\(^\text{54}\) A large number of cell types relevant to asthma pathogenesis have been shown to express ST2 and to be responsive to IL-33,\(^\text{58}\) including Th2 cells,\(^\text{59}\) mast cells,\(^\text{60,61}\) invariant natural killer T cells,\(^\text{62}\) eosinophilic and basophilic granulocytes\(^\text{63,64}\) and epithelial cells.\(^\text{65}\)

By activating these cells, IL-33 has been shown to mediate a wide range of responses, including Th17- mediated airway inflammation\(^\text{66}\) and neutrophil influx.\(^\text{67}\) However, the best described activity of IL-33 is the activation of innate and adaptive immune responses characterized by the production of IL-4, IL-5, and IL-13.\(^\text{39}\)

Through its effects on both the innate and adaptive immune responses, IL-33 can promote the pathogenesis of asthma and exacerbate various other allergic diseases as well as being host-protective against helminthic infection.\(^\text{54,68}\)
1.5.2  **ST2**

The ST2 gene, which is located on chromosome 2q12, spans approximately 40.5 kb, harboring 11 exons and a distal and a proximal promoter.\(^\text{39}\) It was first identified in oncogene- or serum-stimulated fibroblasts\(^\text{54,69,70}\) and encodes proteins with an extracellular region carrying three immunoglobulin-like domains, a transmembrane (TM) domain, and an intracellular region harboring a Toll-like/IL-1 receptor (TIR) domain.\(^\text{40}\) (Figure 2)

Three transcripts of the ST2 protein are expressed through alternative splicing; a short isoform encoding the soluble protein sST2,\(^\text{69}\) a long isoform encoding the full transmembrane receptor ST2L,\(^\text{71}\) and a less well-known variant that encodes a truncated protein with two immunoglobulin-like domains and a hydrophobic tail called ST2V.\(^\text{39,72}\)

Binding of ST2L with its ligand on the surface of basophils, eosinophils and mast cells promotes their activation,\(^\text{63}\) increased adhesion and survival\(^\text{73}\) and degranulation,\(^\text{74}\) respectively. Functionally, ST2L acts to transduce the IL-33 signal (discussed above) to the intracellular compartment while the soluble sST2 functions as a decoy receptor, capturing IL-33 and inhibiting its function.\(^\text{39}\) (Figure 1) sST2 corresponds to the extra-cellular domain of ST2L except for nine amino-acids in the C-terminal region.\(^\text{75}\) sST2 has a relatively ubiquitous tissue distribution showing the highest levels of the secreted form in the lung followed by the heart and the brain.\(^\text{76}\)

---

**Figure 2**  **Human ST2.** A schematic representation of the ST2 gene with its intron/exon structure and a depiction of the protein structure of sST2 and ST2L. Also depicted are the immunoglobulin (Ig) domains common to sST2 and ST2L and the transmembrane (TM) and the class-identifying Toll-like/Interleukin-1 receptor (TIR) domain unique to ST2L. (From Kakkar, R. and R.T. Lee 2009.)
1.5.3 sST2 levels in disease

sST2 levels in serum have been studied and used as a biomarker for disease severity and outcome, for example in sepsis,\textsuperscript{77} dengue,\textsuperscript{78} chronic, obstructive pulmonary disease (COPD),\textsuperscript{79} acute myocardial infarction\textsuperscript{80} and heart failure\textsuperscript{76,80} as well as atopic asthma.\textsuperscript{81} Because the elevated sST2 levels are not specific for a certain disease group they cannot be used as a definitive diagnostic test. However, sST2 has emerged as a powerful prognostic marker in many clinical settings.\textsuperscript{82}
1.6 Schistosomiasis

1.6.1 Prevalence of schistosomiasis
Schistosomiasis is an often neglected tropical disease endemic in 74 countries in regions of Africa, Asia, and South America.\(^8^3\) It is a major public health concern, and according to recent estimates approximately 200 million individuals are infected, with a further 700 million people living at risk of infection.\(^8^4,8^5\)

1.6.2 *Schistosoma* species and their life cycles
Schistosomiasis is caused mainly by the blood flukes *Schistosoma mansoni*, *S. haematobium*, and *S. japonicum*.\(^8^3\) *S. japonicum* (China and Southeast Asia) and *S. mansoni* (Africa, Arabia, and South America) dwell in peri-intestinal venula, and cause intestinal and hepatosplenic schistosomiasis. *S. haematobium* (Africa and Arabia) live in the perivesical plexus and causes urinary schistosomiasis of the bladder, ureters, and kidneys. Infection occurs where the human host comes into contact with water that harbors the intermediate snail host for *Schistosoma* species. Schistosomes have a life cycle that runs from the primary host to water, to the secondary host of a freshwater snail, back to water and then penetrating the skin through water contact the helminth returns to the human.\(^8^6\) (Figure 3).

![Figure 3](image_url) Transmission cycle of schistosomes. (A) Adult schistosomes in small blood vessels around the intestines or bladder. (B) Schistosome eggs leave the body with excreta. Left-to-right: *S. haematobium*, *S. mansoni*, and *S. japonicum*. (C) Miracidium hatches and swims freely in water. (D) Snail intermediate hosts are infected by miracidium. Left-to-right: *Oncomelania*, *Biomphalaria*, and *Biomphalaria*. (E) Cercariae hatch from snails 4 to 6 weeks later. (F) Cercariae infect humans through the skin. (G) Cercariae develop into schistosomula, which migrate with the bloodstream to the portal vein and mature into adult schistosomes, which mate and migrate to the destination. (From Gryseels B, Polman K, Clerinx J, et al. 2006)
The species of *Schistosoma* that causes most infections in Brazil is *S. mansoni* and will therefore be the main focus of this paper.

### 1.6.3 Schistosomiasis and the immune system

Soon after infection and before the egg-laying phase, the immune response is predominantly the Th1 inflammatory immune response, with high levels of TNFα (tumor necrosis factor alpha), IL-1, IL-6 and IFNγ (interferon gamma) characterizing the acute phase of schistosomiasis. The natural progress of the disease leads to liver and intestinal injury caused mainly by the immune response against *S. mansoni* eggs deposited in these sites. In the chronic phase of the disease, there is a predominance of the Th2 immune response to egg antigens, with low production of IFNγ and high levels of IL-4, IL-5, IL-13 and IL-10 as well as IgE. Resistance to reinfection with schistosomes has been associated with a Th2 immunity and consequent high level of IgE. High levels of IgE against SWAP (soluble adult worm antigen preparation) have been associated with resistance to reinfection, whereas high levels of IgG4 (Immunoglobulin G4) against adult worm and egg antigens have been associated with susceptibility to reinfection. The balance between IgE and IgG4 might determine resistance or susceptibility to *S. mansoni* infection, thus the IgE/IgG4 ratio has become a biomarker of *Schistosoma* resistance.

### 1.7 Asthma and schistosomiasis

#### 1.7.1 Interaction between asthma and infection with *S. mansoni*

Exposure to allergens in an individual with asthma and infection with *S. mansoni* each elicit a Th2-mediated immune response and stimulate production of IgE. In regions endemic for extracellular parasitic diseases, such as schistosomiasis, there is a lower prevalence of atopy overall and, among patients with asthma, less severe disease. This indicates that helminth infections protect against allergic disease. The reverse is also true: asthma and history of atopy appear to confer resistance against helminthic parasites such as *S. mansoni*. Additionally, persons with asthma who also have schistosomiasis have been shown to experience worsening in asthma severity after antihelminthic treatments.
1.7.2 Co-associations in genetics of asthma and schistosomiasis

There is extensive evidence for a genetic basis for schistosomiasis and of particular interest is how many of the associated loci overlap with asthma susceptibility.\textsuperscript{103} For example, the linkage to schistosomiasis in chromosome 5q31–q33 is in the same locus where some of the most compelling evidence for linkage to asthma and atopy have been reported\textsuperscript{104-108} and multiple polymorphisms in genes within the Th2, IgE-mediated pathway are co-associated with both traits.\textsuperscript{103}

2 Aim of study

Asthma and schistosomiasis both have a strong genetic component that often overlaps. One of the most commonly associated candidate genes for asthma is \textit{ST2}, encoding the receptor for IL-33. The aim of this study is to determine whether genetic variants in the \textit{ST2} gene, which have previously been positively associated with asthma and sST2 levels in an African-American population,\textsuperscript{109} are also associated with asthma, schistosomiasis, sST2 levels and other related factors in a study population from Brazil.
3 Materials and methods

3.1 Previously existing data - location and study population

This study was performed using samples and data acquired for previous studies of this Brazilian population.

Asthmatic subjects and their families were recruited in five communities (Buri, Camarao, Genipapo, Sempre Viva and Cobo of the district of Conde, Bahia, Brazil) between July 23 and September 2, 2004. Those who had an available DNA sample for genotyping and complete phenotype data were included in this analysis.

3.1.1 Demographic information and sample collection

Basic demographic information and data on environment exposure was collected as well as information on family relationships for all first-degree relatives. Willing participants were asked to give blood samples for storage and future genotyping and serum for measuring total IgE levels. All adults provided written consent (or verbal consent recorded by a witness), and children gave verbal assent and written consent was obtained from a parent or guardian. Children under the age of 6 were excluded.

Blood was collected through venipuncture and serum was separated for measuring total-IgE (tIgE) levels.

The research protocol was approved by the Institutional Review Boards of the Johns Hopkins University School of Medicine and of the Federal University of Bahia and was endorsed by the National Commission for Ethics in Human Research in Brazil.

3.1.2 Evidence of asthma and schistosomiasis

In this population, asthmatic individuals were identified using a modified ISAAC questionnaire.

Previous investigations in this region verified the endemicity of schistosomiasis. Regular, publicly administered mass-treatment campaigns against helminthic infection last occurred during 2001.
Two fecal samples were collected from subjects at an interval of 2–40 days. Stool samples were tested by using the Kato-Katz method\textsuperscript{113,114} to estimate the number of \textit{S. mansoni} eggs per gram of fecal matter from three subsamples taken from the two independent samples per individual. Then the arithmetic mean was calculated for each subject. This test has been shown to give reproducible results when 3–5 exams are performed.\textsuperscript{114}

### 3.1.3 Assessment of water contact.

A previously developed environmental exposures questionnaire\textsuperscript{115} was administered to quantitate the intensity of exposure to \textit{S. mansoni}.

### 3.2 Genotyping

DNA was extracted from whole blood using standard protocols from 697 participants from the Brazil study. Four \textit{ST2} SNPs were selected for genotyping based on prior association with asthma and sST2 levels in an African-American population from Baltimore/Washington.\textsuperscript{109} Of the four SNPs; rs12712135, rs1420101, rs6543119 and rs76930359, only one had prior published associations. SNP rs1420101 is significantly associated with asthma and eosinophil count in European and East-Asian populations.\textsuperscript{46}

### 3.2.1 The Taqman method

Genotyping of the polymorphisms was performed by using TaqMan probe-based, 5´nuclease allelic discrimination assay on the 7900HT Sequence Detection System (Applied Biosystems, Foster City, California, USA). TaqMan-validated assays and Master mix were manufactured by Applied Biosystems. PCR was conducted in a 5-mL volume by using a universal Master mix and four predesigned and validated TaqMan assays for the SNPs. The thermal cycling conditions were as follows: 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds/60°C for 1 minute and an extension step of 60°C for 5 minutes.
3.2.2 Genotyping Quality Control.
Non-template negative and genotyping-positive controls were included in each genotyping plate. Automatic calling was performed with a quality value of greater than 99%. 10% of the samples were genotyped in duplicate with 99% reproducibility.

3.3 Measurements of sST2, tIgE and the IgE/IgG4 ratio

3.3.1 sST2 Levels
sST2 levels were measured using the Presage ST2 Assay\textsuperscript{82} (Critical Diagnostics, California, USA) by ELISA (Enzyme-Linked Immunosorbent Assay) in serum samples using the protocol suggested by the manufacturer.

3.3.2 Total IgE measurements
Total-IgE levels were measured using chemiluminescence (ADVIA Centaur Bayer Corporation, Salvador, Brazil) and a ratio of intensity of response was obtained using a confirmed non-allergic reference sample.\textsuperscript{116}

3.3.3 IgE/IgG4
IgG4 specific for SWAP, soluble egg antigen, IL-4-inducing principle from \textit{S. mansoni} eggs and Sm22.6 was measured by using an indirect enzyme-linked immunosorbent assay as described elsewhere.\textsuperscript{83,117,118}

3.4 Statistical Methods.

3.4.1 Data testing and analyses
Hardy-Weinberg testing and tests for Mendelian inconsistencies were performed using PLINK, and pairwise linkage disequilibrium estimations using Haploview.
All analyses were performed using generalized estimating equations (GEE) to adjust for relatedness in the Brazil population and adjusting for the first 2 principal components (PCs).
In addition, before analysis, sST2, SWAP, and tIgE were log10-transformed and adjusted by age and gender. Egg count was log10-transformed and adjusted by age, gender and water contact index. Along with the first two PCs, wheeze was adjusted by egg count (yes/no) and gender.

### 3.4.2 Population stratification

Average population admixture and individual admixture estimates were determined using a panel of 237 SNPs\(^\text{119}\) A model-based clustering method was used by grouping data for the total sample in three ancestral populations (\(K = 3\)) to reflect the admixture history in South America with the software STRUCTURE (version 2.3.3). Assessment of stratification was done by using principal component analysis (PCA) using the smartpca program from the software package eigenstrat\(^\text{120}\) including the three putative ancestral populations available in HapMap (CEU: Utah residents with ancestry from northern and western Europe, CHB: Han Chinese in Beijing, China and YRI: Yoruba in Ibadan, Nigeria) as reference. (Figure 4) Principal components were determined for each individual in this population included in the current analysis.

![Plots of the first principal components (PC1, PC2) from analysis on Brazilians founders.](image)

Individuals are color plotted by asthma status (Cases/Controls) in each sample. Included are the putative parental populations available in HapMap (CEU, CHB and YRI).
4 Results

4.1 Clinical characteristics of the analyzed subjects

Table 1 shows the characteristics of the study population which consisted of 697 Brazilians. Of those, 196 were founders, defined as unrelated individuals; namely the parents in the oldest generation, and in-laws in subsequent generations. This group is older than the total population.

Of note, Table 1 shows that the reported prevalence of wheeze is 30% (212/697) in the total analyzed population and 31.6% (62/196) in founders. A higher percentage of males reported wheeze compared to females.

Mean sST2 levels are similar between all research groups with mean serum levels between 28 and 30 ng/mL. Mean total IgE levels are highest in the group with wheeze and lowest in founders.

*S. mansoni* infection is most prevalent and most severe in the wheeze group but least prevalent and least severe in the founders.

Table 1 Clinical characteristics of the analyzed subjects

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total</th>
<th>Founders</th>
<th>Wheeze(+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>697</td>
<td>196</td>
<td>212</td>
</tr>
<tr>
<td>Males; N (%)</td>
<td>307 (44.1%)</td>
<td>82 (41.8%)</td>
<td>100 (47.2%)</td>
</tr>
<tr>
<td>Age; Mean (SD)</td>
<td>27.3 (19.0)</td>
<td>42.6 (19.2)</td>
<td>27.4 (19.0)</td>
</tr>
<tr>
<td>Wheeze(+); N (%)</td>
<td>212 (30.4%)</td>
<td>62 (31.6%)</td>
<td>212 (100%)</td>
</tr>
<tr>
<td>sST2*[^]; Mean (95% CI)</td>
<td>28.4 (27.6-29.2)</td>
<td>30.1 (28.3-32.0)</td>
<td>28.2 (26.8-29.6)</td>
</tr>
<tr>
<td>Total IgE[^]^; Mean (95% CI)</td>
<td>2773 (2533-3036)</td>
<td>2635 (2211-3141)</td>
<td>3123 (2669-3654)</td>
</tr>
<tr>
<td>IgE/IgG4; Mean (95% CI)</td>
<td>1.11 (0.93-1.34)</td>
<td>0.74 (0.55-0.98)</td>
<td>0.92 (0.66-1.29)</td>
</tr>
<tr>
<td>S. mansoni(+); N (%)</td>
<td>228 (43.4%)</td>
<td>58 (39.2%)</td>
<td>74 (44.3%)</td>
</tr>
<tr>
<td>Egg count*[^]; Mean(95%CI)</td>
<td>54.9 (46.8-64.4)</td>
<td>44.7 (31.2-64.1)</td>
<td>56.6 (43.8-73.0)</td>
</tr>
</tbody>
</table>

*Geometric mean. ^[ng/mL]. No: Number, CI: Confidence Interval
Table 2 shows the clinical characteristics of the study population distributed by gender.

The main information Table 2 adds is that males have higher prevalence of wheeze and higher mean sST2 and tIgE levels compared to females in both total population and founder groups.

Males also have a higher prevalence and severity of *Schistosoma mansoni* infection and correspondingly lower IgE/IgG4 ratios.

For these reasons, further analyses were adjusted by gender and age as described in the Statistical methods section.

### Table 2 Clinical characteristics of the analyzed subjects distributed by gender

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total males</th>
<th>Total females</th>
<th>Founder males</th>
<th>Founder females</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>307</td>
<td>390</td>
<td>82</td>
<td>114</td>
</tr>
<tr>
<td>Age; Mean (SD)</td>
<td>26.1 (18.8)</td>
<td>28.2 (19.2)</td>
<td>42.2 (19.6)</td>
<td>42.9 (19.0)</td>
</tr>
<tr>
<td>Wheeze(+)N (%)</td>
<td>100 (32.6%)</td>
<td>112 (28.7%)</td>
<td>28 (34.2%)</td>
<td>34 (29.8%)</td>
</tr>
<tr>
<td>sST2*^; Mean (95% CI)</td>
<td>31.3 (30.0-32.7)</td>
<td>26.2 (25.3-27.1)</td>
<td>35.1 (32.0-38.5)</td>
<td>26.8 (24.8-28.9)</td>
</tr>
<tr>
<td>Total IgE*^; Mean (95% CI)</td>
<td>3578 (3196-4006)</td>
<td>2669 (1987-2589)</td>
<td>3568 (2856-4456)</td>
<td>2114 (1665-2718)</td>
</tr>
<tr>
<td>IgE/IgG4; Mean (95% CI)</td>
<td>0.87 (0.69-1.10)</td>
<td>1.35 (1.03-1.75)</td>
<td>0.54 (0.37-0.79)</td>
<td>0.93 (0.61-1.41)</td>
</tr>
<tr>
<td>S. mansoni(+)N (%)</td>
<td>115 (51.1%)</td>
<td>113 (37.5%)</td>
<td>30 (50.9%)</td>
<td>28 (31.5%)</td>
</tr>
<tr>
<td>Egg count*; Mean (95% CI)</td>
<td>64.9 (51.5-81.8)</td>
<td>46.3 (37.2-57.6)</td>
<td>62.6 (37.6-104)</td>
<td>31.2 (18.8-51.6)</td>
</tr>
</tbody>
</table>

*Geometric mean. ^[ng/mL]*
4.2 sST2 levels compared to other traits in the total population

Table 3 shows the pairwise GEE analysis between sST2 levels and other covariates in the total study population.

sST2 levels are observed to be positively associated with infection with *Schistosoma mansoni* and total IgE levels with a *P*-value of less than 0.05, but not significantly associated with wheeze or *Schistosoma* egg count.

sST2 levels are associated with the male gender and are negatively associated with the IgE/IgG4 ratio.

There is a trend of higher age – higher sST2 levels, however, the *P*-value is borderline significant.

### Table 3 Pairwise GEE analysis of sST2 and other traits in the total population.
Comparing unadjusted log(sST2) to all other covariates.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimation</th>
<th>Standard error</th>
<th>Log(P)</th>
<th><em>P</em>-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>-0.0966</td>
<td>0.0098</td>
<td>22.2301</td>
<td>5.89*10^{-23}</td>
</tr>
<tr>
<td>Age</td>
<td>5.00*10^{-4}</td>
<td>3.00*10^{-4}</td>
<td>1.3019</td>
<td>0.0499</td>
</tr>
<tr>
<td>Wheeze</td>
<td>0.0192</td>
<td>0.0121</td>
<td>0.943</td>
<td>0.114</td>
</tr>
<tr>
<td>Log(egg count)</td>
<td>0.0264</td>
<td>0.0162</td>
<td>0.9866</td>
<td>0.103</td>
</tr>
<tr>
<td><em>S. mansoni</em> infection (Yes/No)</td>
<td>0.0613</td>
<td>0.0108</td>
<td>7.8441</td>
<td>1.43*10^{-8}</td>
</tr>
<tr>
<td>Log((IgE/IgG4)+1)</td>
<td>-0.0423</td>
<td>0.0129</td>
<td>2.9812</td>
<td>0.001</td>
</tr>
<tr>
<td>Log(IgE)</td>
<td>0.0436</td>
<td>0.0082</td>
<td>6.9973</td>
<td>1.01*10^{-7}</td>
</tr>
</tbody>
</table>
4.3 Analyzed SNPs – a closer look

Table 4 summarizes the SNPs genotyped, their chromosomal positions, alleles and minor allele frequencies (MAF) in the total population.

The coded allele represents the allele analyzed as reference to conferring risk or protection of any of the traits.

All four markers; rs12712135, rs1420101, rs6543119 and rs76930359 were in Hardy Weinberg Equilibrium (HWE) so they were all included in the analysis.

Table 4 HWE and MAFs of sST2 markers in a Brazilian population.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Chr.</th>
<th>Position</th>
<th>Allele1</th>
<th>Allele2</th>
<th>Coded allele</th>
<th>MAF</th>
<th>HWE</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs12712135</td>
<td>2</td>
<td>102930948</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>0.437</td>
<td>0.878</td>
</tr>
<tr>
<td>rs1420101</td>
<td>2</td>
<td>102957716</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>0.308</td>
<td>0.463</td>
</tr>
<tr>
<td>rs6543119</td>
<td>2</td>
<td>102963072</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>0.299</td>
<td>0.858</td>
</tr>
<tr>
<td>rs76930359</td>
<td>2</td>
<td>102924684</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>0.077</td>
<td>0.265</td>
</tr>
</tbody>
</table>

Chr.: Chromosome
Graph 1 shows the location of the four genotyped SNPs in the ST2 gene and their LD pattern. SNP rs76930359 is in the promoter region. SNPs rs12712135, rs1420101 and rs6543119 are located in intronic regions; introns 1, 5 and 8 respectively. SNPs rs1420101 and rs6543119 are contained in a 5 kb haplotype block and are in high LD with a $R^2$ value of 0.93 in this population.
4.4 GEE analyses of ST2 markers and other traits

Table 5 shows associations between the genotyped SNPs and wheeze in a Brazilian population. Before analysis wheeze was adjusted by *Schistosoma mansoni* egg count (yes/no) and gender.

One of the four markers is significantly associated with wheeze in the Brazilian population.

SNP rs76930359 has a coded allele C that is negatively associated with wheeze in the study population.

No association was found between the other three markers and wheeze.

<table>
<thead>
<tr>
<th>Marker – coded allele</th>
<th>Estimation</th>
<th>Standard error</th>
<th>Log(P)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs12712135 – G</td>
<td>0,0353</td>
<td>0,0317</td>
<td>0,5776</td>
<td>0,264</td>
</tr>
<tr>
<td>rs1420101 – C</td>
<td>-0,0575</td>
<td>0,0352</td>
<td>0,9911</td>
<td>0,102</td>
</tr>
<tr>
<td>rs6543119 – A</td>
<td>-0,0582</td>
<td>0,0318</td>
<td>1,1753</td>
<td>0,0668</td>
</tr>
<tr>
<td>rs76930359 – C</td>
<td>-0,1021</td>
<td>0,044</td>
<td>1,6901</td>
<td>0,0204</td>
</tr>
</tbody>
</table>
Table 6 shows associations between genotyped SNPs and sST2 levels in a Brazilian population. Before analysis sST2 levels were log10-transformed and adjusted by age and gender.

Three markers are significantly associated with sST2 levels in the study population.

SNP rs12712135 has a coded allele G that is positively associated with sST2 levels in a Brazilian population.

SNP rs1420101 with a coded allele C and SNP rs6543119 with the coded allele A are both negatively associated with sST2 levels in the study population.

No association was found with SNP rs76930359.

<table>
<thead>
<tr>
<th>Marker – coded allele</th>
<th>Estimation</th>
<th>Standard error</th>
<th>Log(P)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs12712135 – G</td>
<td>0,0585</td>
<td>0,0122</td>
<td>5,8137</td>
<td>1,53*10^{-6}</td>
</tr>
<tr>
<td>rs1420101 – C</td>
<td>-0,0593</td>
<td>0,012</td>
<td>6,0758</td>
<td>8,40*10^{-7}</td>
</tr>
<tr>
<td>rs6543119 – A</td>
<td>-0,0604</td>
<td>0,0106</td>
<td>7,8668</td>
<td>1,35*10^{-8}</td>
</tr>
<tr>
<td>rs76930359 – C</td>
<td>0,0367</td>
<td>0,021</td>
<td>1,0902</td>
<td>0,0812</td>
</tr>
</tbody>
</table>
Table 7 shows associations between the genotyped SNPs and SWAP-specific IgE/IgG4, tIgE levels and *Schistosoma mansoni* egg count.

Before analysis, SWAP and tIgE were log10-transformed and adjusted by age and gender, egg count was log10-transformed and adjusted by age, gender and water contact index

None of the four SNPs genotyped in the Brazilian study population showed significant association with any of these traits.

**Table 7 GEE analysis of ST2 markers and SWAP, total IgE and egg count in a Brazilian population.**

<table>
<thead>
<tr>
<th>Marker</th>
<th>SWAP</th>
<th>tIgE</th>
<th>Egg count</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimation</td>
<td>Standard error</td>
<td>Log(f)</td>
</tr>
<tr>
<td>rs12712135</td>
<td>0.0349</td>
<td>0.0015</td>
<td>0.2436</td>
</tr>
<tr>
<td>rs1420101</td>
<td>-0.061</td>
<td>0.0739</td>
<td>0.3869</td>
</tr>
<tr>
<td>rs6543119</td>
<td>-0.015</td>
<td>0.064</td>
<td>0.0886</td>
</tr>
<tr>
<td>rs76930359</td>
<td>-0.014</td>
<td>0.0902</td>
<td>0.0577</td>
</tr>
</tbody>
</table>
5 Discussion

5.1 Clinical characteristics compared to other populations.

The Brazilian study population has a 30% prevalence of wheeze. (Table 1) This is quite high compared to the global prevalence of asthma; ranging from 1% to 18% of the population in different countries.3-7 One of the reasons behind this difference could be that this study population has a substantial African admixture.37 African ancestry is a known risk factor for asthma and high total IgE levels in African admixed populations.119 Environmental factors can also play a role in the differences in prevalence among different populations.

The mean total-IgE levels are also very high in all groups of the study population, especially in the wheeze group that displays mean tIgE levels of more than 3000 ng/mL. (Table 1) This population has a high prevalence of parasitic infections.98,99,101 Total IgE includes the quantification of IgE against parasites and this factor can account for the high total IgE levels. The measurements of specific IgE levels against relevant allergens can help us to disentangle the influence of parasites in the high values observed in this population.

The males have higher tIgE levels compared to the females. (Table 2) This corresponds to previous findings that at all ages, males have higher total IgE concentrations than females.1 The mean tIgE levels in all groups are much higher than the reference value of 4-274 ng/mL (1.5-114 kU/L with one kU/L being equal to 2.4 ng/mL121) established in 1981122 and even higher than the clinically relevant threshold for predicting allergy which is 480 ng/mL.122,123 Other studies have shown that after 14 years of age, total serum IgE levels over 800 ng/mL are considered abnormally elevated and strongly associated with atopic disorders, such as allergic rhinitis, extrinsic asthma, and atopic dermatitis.123,124 This corresponds to the high prevalence of wheeze in this Brazilian population. However, IgE levels can vary greatly depending on birthplace and background123,125,126 and also seem to have a genetic component.116 Therefore if diagnosis is based on tIgE it is important to study the reference values compatible with that patient’s particular population.

A higher percentage of males presents with wheeze than females (Tables 1 and 2) in the study population. In general, prior to the age of 14, asthma prevalence is nearly twice as high in boys as in girls127 but by adulthood the prevalence of asthma is greater in women.3
The reason for this gender-related difference is not clear although it is speculated this may be due to lung size, which is smaller in males than in females at birth but larger in adulthood. The founder group is, as expected, older than the total population and has both the lowest prevalence and severity of *Schistosoma* infection. (Table 1) This indicates that there is an age-related resistance to reinfection that has also been demonstrated in other studies.

The Presage sST2 assay showed mean sST2 levels to be relatively similar between the total population, founder- and wheeze groups around 28-30 ng/mL. (Table 1) When distributed by gender there was a higher level of sST2 in males than in females, (Table 2) corresponding with sST2 levels being significantly associated with the male gender. (Table 3) Therefore it is clear, both in this study and others (discussed below), that sST2 levels are higher in males compared to females. The reasons and possible consequences of this difference are unknown. Perhaps it has to do with hormonal differences or other gender-specific traits. In any case it would be interesting to study whether sST2 levels as a biomarker predict different outcomes between the sexes.

The mean sST2 levels are in general in the higher range compared to other populations; for example compared to an Austrian study which identified gender-specific reference intervals of 4–31 and 2–21 ng/mL for males and females, respectively, using samples obtained from blood donors in Austria. A study performed in the United States established combined reference intervals for both males (8.6–49.3 ng/mL) and females (7.2–33.5 ng/mL) that were significantly higher than the Austrian levels and more in keeping with the sST2 levels measured in this Brazilian study population. These variations in sST2 levels between studies may be accredited to genetics and/or demographic differences as well as variation in screening and exclusion criteria.

Contrary to studies showing inverse association between asthma and schistosomiasis, the wheeze group in this population shows both higher prevalence and severity of *S. mansoni* infection. (Table 1) An unknown factor is whether the population with wheeze would have an even higher prevalence of asthma and/or more severe symptoms if introduced to anti-helminthic treatments as has been shown in other studies. Asthma has also been shown to be linked to various environmental factors which are not included in this study and could thus confound the results.
The IgE/IgG4 ratio is lowest in the wheeze group, (Table 1) further indicating its role as a biomarker for resistance to *Schistosoma* spp. infection. The total IgE levels are, as expected, also highest in the wheeze group as asthma is closely associated with serum tIgE levels.

### 5.2 Associations between sST2 levels and other traits

As discussed in the Introduction, elevated sST2 levels have been linked to various diseases such as acute asthma, heart failure, acute myocardial infarction and more. In the pairwise GEE analysis comparing sST2 levels to all other covariates the sST2 levels are, however, not significantly associated with wheeze in this population. (Table 3)

In this study, wheeze was determined by the ISAAC questionnaire and used as an equivalent of asthma. Studies have various different definitions of asthma, with wheeze being one criterion. However, it is only one of multiple expressions of this heterogeneous disease which could explain the varying results between studies.

Both *Schistosoma mansoni* infection and mean tIgE levels are significantly and positively associated with sST2 levels in this study population. (Table 3) It is possible that the parasitic infection could mask the symptoms of wheeze in some of the non-wheeze controls in the population through dampening of the asthma immune response discussed in the Introduction.

Correspondingly, there is a significant negative association between sST2 levels and the SWAP-specific IgE/IgG4 ratio, previously described as a biomarker for *Schistosoma* resistance.

### 5.3 Associations between genetic variations in ST2 and other traits

All four genotyped SNPs; rs12712135, rs1420101, rs6543119 and rs76930359, had previously been positively associated with asthma and sST2 levels in a study population of African-Americans from the Baltimore/Washington area in the USA. These data are unpublished but are the basis for this study and the selection of SNPs for genotyping.
The GEE analyses showed that, in contrast to these prior association results, only one of the genotyped SNPs (rs76930359) was significantly associated with wheeze in this study population of Brazilians. (Table 5)

The other three; rs12712135, rs1420101, rs6543119, were significantly associated with sST2 levels. (Table 6) It is possible that we found more of a signal in the study of sST2 levels than in other components, such as wheeze, because the sST2 levels are more standardized and objective in terms of measurement.

None of the genotyped SNPs showed significant association with tIgE, *Schistosoma mansoni* egg count nor SWAP-specific IgE/IgG4. (Table 7)

The SNP rs76930359 was negatively associated with wheeze (Table 5) in the Brazilian study population as opposed to the prior positive association in the African American population. This indicates that the coded allele C in rs76930359 confers “protection” against wheeze in Brazilians but increases risk for asthma in a different population.

A possible explanation of the discrepancies between study results is the difference in allelic frequencies between study populations. Allele frequencies of three reference populations (AFR – Africans, ASN – Asians and EUR – Europeans) were therefore located on Ensembl Genome Browser in order to compare the allele frequencies with those of the Brazilian study population.

The negative association of rs76930359 with wheeze is not very robust with a *P*-value of 0.0204 (Table 5) however, the minor allele frequency is 0.077 (Table 4) which means only 7.7% of chromosomes in the study population display the coded allele. In the reference populations, the less frequent allele of the Brazilians, C, is the more frequent one with allele frequencies of 97%, 100% and 80% in African, Asian and European populations respectively.

Thus it is clear that a much larger study population is needed to detect a signal of this SNP in the Brazilian study population than in the reference populations. A solution to this problem would be simply to increase the number of subjects in the study population in order to increase the power of the study.
The three genotyped SNPs associated with sST2 levels in the Brazilian population did not all show the same association. SNP rs12712135 indicated a positive association with sST2 levels in this study population corresponding to prior findings in an African-American population. Conversely, both SNPs rs1420101 and rs6543119 were negatively associated with sST2 levels in the study population. They had very similar results (Table 6) that are likely the result of the same signal as they are in high LD with each other with an $R^2$ value of 0.93. (Graph 1)

The most likely explanation for the different association findings between populations is that the SNPs in and of themselves do not convey the effect but are rather in varying LD with causal and/or protective genetic variants for asthma in different populations. Distribution of allelic frequencies, as seen in the SNP rs76930359, can also vary between populations and thus increase or decrease the power needed to detect the causal variant. Various environmental factors could also be influencing phenotypic expression.

Additionally, as mentioned above, the traits studied, such as asthma/wheeze and sST2 levels, can be different between studies depending on definition, inclusion criteria, age of population and many other factors. These same reasons can explain why the SNP rs1420101, previously associated with asthma in European and East-Asian populations,\textsuperscript{46} showed no association with wheeze in the Brazilians.

The minor allele frequency of the coded allele C in the Brazilian study population is 30.8%. (Table 4) This is lower than the corresponding allele frequencies in all reference populations; AFR: 59%, ASN: 57% and EUR: 66%.\textsuperscript{133} In order to detect a signal from this marker in the study population, the sample size must be larger to reach significance than in studies of the reference populations. A logical direction of research now would be to study larger sample sizes of several different populations to determine and possibly replicate association between rs1420101 and asthma.
5.4 Conclusions

In conclusion, it is clear that heterogeneity exists between populations. Thus, even though this study did not replicate the findings of the previous study in African-Americans, it does not negate those results. However, further analyses are needed to better understand the role of variants in $ST2$ in asthma, expression of $ST2$ and infection by *Schistosoma mansoni*. Interesting areas of study in Brazilians lie in larger sample sizes due to varying allele frequencies between populations.

SNP rs76930359 and genetic variations in high LD with this SNP are of note in studying asthma resistance in this population. The other three SNPs may warrant future study, for example, in the research on sST2 as a biomarker for disease.

It is evident that the asthma-associated genetic variants discovered to date are only the tip of the iceberg. The common variants discovered by the GWAS approach (and meta-analyses of GWAS) do not seem to have a large effect on the risk of asthma and are poor predictors of disease status. This common disease may therefore be the result of an additive affect of many variants or of rarer variants that have a larger effect size and play a bigger part in the heritability of asthma. A future way of detecting these markers is through whole-genome or exome sequencing in order to discover causal variants.
References


69. Tominaga S. A putative protein of a growth specific cDNA from BALB/c-3T3 cells is highly similar to the extracellular portion of mouse interleukin 1 receptor. FEBS letters. Dec 4 1989;258(2):301-304.


100. Araujo MI, de Carvalho EM. Human schistosomiasis decreases immune responses to allergens and clinical manifestations of asthma. *Chemical immunology and allergy*. 2006;90:29-44.


125. Grigoreas C, Pappas D, Galatas ID, Kollias G, Papadimos S, Papadakis P. Serum total IgE levels in a representative sample of a Greek population. I. Correlation with age, sex, and skin reactivity to common aeroallergens. *Allergy.* Apr 1993;48(3):142-146.


